

# Geometric Methods in Mathematical Physics II: Tensor Analysis on Manifolds and General Relativity

by **Valter Moretti**

[www.science.unitn.it/~moretti/home.html](http://www.science.unitn.it/~moretti/home.html)

Department of Mathematics,  
University of Trento  
Italy

Lecture notes authored by Valter Moretti and freely downloadable from the web page  
<http://www.science.unitn.it/~moretti/dispense.html> and are licensed under a Creative  
Commons Attribuzione-Non commerciale-Non opere derivate 2.5 Italia License.

**None is authorized to sell them**

# Contents

<b>1</b>	<b>Introduction</b>	<b>7</b>
<b>2</b>	<b>Topological and smooth manifolds</b>	<b>8</b>
2.1	Some topology . . . . .	8
2.1.1	Basic concepts and results of general topology . . . . .	8
2.1.2	The topology of $\mathbb{R}^n$ . . . . .	11
2.1.3	Topological manifolds . . . . .	12
2.2	Differentiable manifolds . . . . .	14
2.2.1	Local charts and atlas . . . . .	14
2.2.2	Differentiable structures . . . . .	14
2.2.3	Differentiable functions and diffeomorphisms . . . . .	18
2.3	Technicalities: smooth partitions of unity . . . . .	20
2.3.1	Paracompactness . . . . .	22
2.3.2	Existence of a smooth partition of unity . . . . .	23
<b>3</b>	<b>Tensor Fields on Manifolds and Associated Geometric Structures</b>	<b>25</b>
3.1	Tangent and cotangent space at a point . . . . .	25
3.1.1	Tangent vectors as derivations . . . . .	25
3.1.2	More on the linear space of derivations . . . . .	29
3.1.3	Cotangent space . . . . .	32
3.2	Vector and tensor fields . . . . .	33
3.2.1	Tensor fields . . . . .	33
3.2.2	Lie brackets of vector fields . . . . .	36
3.2.3	Fields of $k$ -forms and exterior derivative . . . . .	37
3.3	Elementary bundles . . . . .	38
3.3.1	Tangent and cotangent bundle . . . . .	38
3.3.2	Fiber bundles . . . . .	42
<b>4</b>	<b>Differential mapping, submanifolds, Lie derivative</b>	<b>45</b>
4.1	Differential of mappings and its applications . . . . .	45
4.1.1	Push forward . . . . .	45

4.1.2	Pull back . . . . .	47
4.2	Rank of a smooth map and related notions . . . . .	47
4.2.1	Rank of a smooth map . . . . .	47
4.2.2	Immersions, submersions, submanifolds . . . . .	49
4.2.3	Theorem of regular values . . . . .	52
4.3	Flow of a vector field, Lie derivative, Frobenius Theorem . . . . .	53
4.3.1	Almost all about the flow of a vector field . . . . .	53
4.3.2	Lie derivative of vector fields . . . . .	56
4.3.3	Lie derivative of tensor fields . . . . .	58
4.3.4	Commuting vector fields and associated commuting flows . . . . .	60
4.3.5	Frobenius theorem . . . . .	63
<b>5</b>	<b>(Pseudo) Riemannian manifolds and related metric tools</b>	<b>66</b>
5.1	(Pseudo) Riemannian manifolds . . . . .	66
5.1.1	Types of manifolds equipped with metrics . . . . .	66
5.1.2	Length of curves . . . . .	67
5.1.3	Local and global flatness . . . . .	68
5.1.4	Existence of Riemannian and Lorentzian metrics . . . . .	69
5.2	Induced metrics and related notions . . . . .	71
5.2.1	Metric induction machinery . . . . .	71
5.2.2	Induced metric on submanifolds . . . . .	71
5.2.3	Isometries . . . . .	74
5.2.4	Killing fields . . . . .	75
5.2.5	Normal and co-normal bundle to an embedded submanifold . . . . .	77
5.3	Borel measure induced by a metric . . . . .	82
5.3.1	Invariant measure in coordinates . . . . .	82
5.3.2	Borel measure on the whole manifold . . . . .	83
<b>6</b>	<b>Affine connections and related geometric tools</b>	<b>86</b>
6.1	Affine connections and covariant derivatives . . . . .	86
6.1.1	The problem of the absence of an affine-space structure . . . . .	86
6.1.2	Affine connections and covariant derivative operator and vector . . . . .	88
6.1.3	Connection coefficients . . . . .	90
6.1.4	Covariant derivative tensor . . . . .	91
6.1.5	Transformation rule of the connection coefficients. . . . .	91
6.1.6	Assignment of a connection. . . . .	92
6.1.7	Covariant derivative of tensor fields. . . . .	93
6.2	The Levi-Civita connection on (pseudo) Riemannian manifolds . . . . .	94
6.2.1	Torsion tensor field . . . . .	95
6.2.2	The Levi-Civita connection. . . . .	96
6.2.3	Killing fields again . . . . .	98
6.2.4	The divergence theorem . . . . .	100

6.3	Parallel transport . . . . .	101
6.3.1	Covariant derivatives of a vector field along a curve . . . . .	102
6.3.2	Parallel transport of vectors and geodesics . . . . .	103
6.4	Affine and metric geodesics . . . . .	106
6.4.1	First order formulation of the geodesic equation . . . . .	106
6.4.2	Affine parameters and length coordinate . . . . .	108
6.5	Metric geodesics: the variational approach in coordinates . . . . .	109
6.5.1	Basic notions of elementary calculus of variations in $\mathbb{R}^n$ . . . . .	109
6.5.2	Euler-Poisson equations . . . . .	112
6.5.3	Geodesics of a metric from a variational point of view . . . . .	116
<b>7</b>	<b>The Exponential Map of Affine and Metric Connections</b>	<b>121</b>
7.1	Exponential map at a point . . . . .	121
7.1.1	The exponential map . . . . .	121
7.1.2	Normal neighborhoods, geodesic balls, injectivity radius . . . . .	125
7.1.3	Hopf-Rinow's theorem . . . . .	128
7.2	Normal coordinates and geodesically convex sets . . . . .	128
7.2.1	Normal coordinates around a point . . . . .	129
7.2.2	The Gauss lemma and local length minimizing/maximizing geodesics . . .	130
7.2.3	Riemannian normal coordinates adapted to a geodesic. . . . .	140
7.2.4	Convex normal neighborhoods and Whitehead's theorem . . . . .	143
7.2.5	Coverings of convex neighborhoods with a nice intersection property . . .	146
<b>8</b>	<b>General Relativity: a Geometric Presentation</b>	<b>151</b>
8.1	The spacetime of General Relativity . . . . .	151
8.1.1	Types of vectors and curves in a spacetime . . . . .	152
8.1.2	Time orientation of a spacetime . . . . .	153
8.1.3	Proper time and rest space associated to observers . . . . .	155
8.2	A physical justification of the mathematical formalism . . . . .	156
8.2.1	Einstein's equivalence principle in Physics . . . . .	157
8.2.2	Equivalence principle and normal coordinates . . . . .	159
8.3	The Strong Equivalence Principle and the extension of the formalism . . . . .	161
8.3.1	The Strong Equivalence Principle . . . . .	161
8.3.2	Conservation laws and Killing fields in General Relativity . . . . .	165
8.3.3	Energy extraction from rotating black holes . . . . .	168
8.4	Newtonian correspondence and some consequence . . . . .	169
8.4.1	Newtonian correspondence . . . . .	169
8.4.2	Gravitational redshift . . . . .	172
8.4.3	Gravitational time dilation . . . . .	174
8.5	Extended reference frames in General Relativity and related notions . . . . .	175
8.5.1	A general notion of extended reference frame . . . . .	175
8.5.2	Classes of stationary spacetimes and coordinates . . . . .	178

8.5.3	The problem of the spatial metric of an extend reference frame . . . . .	182
8.5.4	Einstein's synchronizability in stationary spacetimes . . . . .	190
8.5.5	The metric of a rotating platform and the Sagnac Effect . . . . .	192
8.6	Fermi-Walker's transport in Lorentzian manifolds. . . . .	194
8.6.1	The equations of Fermi-Walker's transport . . . . .	194
8.6.2	Fermi-Walker's transport and Lorentz group . . . . .	198
<b>9</b>	<b>Curvature</b>	<b>201</b>
9.1	Curvature tensors of affine and metric connections . . . . .	201
9.1.1	Curvature tensor and Riemann curvature tensor . . . . .	202
9.1.2	A geometric meaning of the curvature tensor . . . . .	206
9.1.3	Properties of curvature tensor and Bianchi identity . . . . .	208
9.1.4	Riemann tensor and Killing vector fields . . . . .	210
9.2	Local flatness and curvature tensor . . . . .	212
9.2.1	A local analytic version of Frobenius theorem. . . . .	212
9.2.2	Local flatness is equivalent to zero curvature tensor . . . . .	213
<b>10</b>	<b>Gravitation in General Relativity</b>	<b>216</b>
10.1	Geodesic deviation and local flatness of spacetime . . . . .	216
10.1.1	External gravitational interaction in General Relativity . . . . .	216
10.1.2	Geodesic deviation as description of external gravitational interaction . .	217
10.1.3	Geodesic deviation of causal geodesics and local flatness . . . . .	221
10.2	The Einstein equations . . . . .	222
10.2.1	Ricci tensor, contracted Bianchi identity and Einstein tensor . . . . .	222
10.2.2	The Einstein Equations . . . . .	224
10.2.3	The meaning of the cosmological constant . . . . .	229
10.2.4	Diffeomorphism invariance . . . . .	229
10.3	The birth of the cosmology as a science . . . . .	230
10.3.1	The FLRW cosmological model . . . . .	231
10.3.2	The observed expansion of the Universe . . . . .	234
10.3.3	Evidence of spatial flatness: $k = 0$ . . . . .	238
10.3.4	Big bang, cosmic microwave background, dark matter, and dark energy .	238
10.3.5	Open problems . . . . .	243

**Acknowledgments.** I am grateful to Carlo Cianfarani and Leonardo Goller for having pointed out several typos and mistakes of various nature.

# Chapter 1

## Introduction

The content of these lecture notes covers the second part<sup>1</sup> of the lectures of a graduate course in Modern Mathematical Physics at the University of Trento. The course has two versions, one is geometric and the other is analytic. These lecture notes only concern the geometric version of the course. The analytic version regarding applications to linear functional analysis to quantum and quantum relativistic theories is covered by my books [Morettia], [Morettib] and the chapter [KhMo15].

The idea of this second part is to present into a concise but rigorous fashion some of the most important notions of differential geometry and use them to formulate an introduction to the General Theory of Relativity.

---

<sup>1</sup>The first part apperas in [Mor20].

## Chapter 2

# Topological and smooth manifolds

This introductory chapter introduces the fundamental building block of these lectures, the notion of smooth manifold. An un introduction is necessary regarding some topological structures.

### 2.1 Some topology

We start with a very quick recap on general topology and next we pass to the notion of topological manifold.

#### 2.1.1 Basic concepts and results of general topology

Let us summarize several basic definitions and results of general topology. The proofs of the various statements can be found in every textbook of general topology or also in textbooks on geometry [Seri90].

**0. (Topological spaces)** We remind the reader that a **topological space** is a pair  $(X, \mathcal{T})$  where  $X$  is a set and  $\mathcal{T}$  is a class of subsets of  $X$ , called **topology**, which satisfies the following three properties.

- (i)  $X, \emptyset \in \mathcal{T}$ .
- (ii) If  $\{X_i\}_{i \in I} \subset \mathcal{T}$ , then  $\cup_{i \in I} X_i \in \mathcal{T}$  (also if  $I$  is *uncountable*).
- (iii) If  $X_1, \dots, X_n \in \mathcal{T}$ , then  $\cap_{i=1, \dots, n} X_i \in \mathcal{T}$ .

As an example, consider any set  $X$  endowed with the class  $\mathcal{P}(X)$ , i.e., the class of all the subsets of  $X$ . That is a very simple topology which can be defined on each set, e.g.  $\mathbb{R}^n$ .

**1. (Open and closed sets)** If  $(X, \mathcal{T})$  is a topological space, the elements of  $\mathcal{T}$  are said to be **open sets**. A subset  $K$  of  $X$  is said to be **closed** if  $X \setminus K$  is open. It is a trivial task to show that the (also uncountable) intersection closed sets is a closed set. The **closure**  $\overline{U}$  of a set  $U \subset X$  is the intersection of all the closed sets  $K \subset X$  with  $U \subset K$ .

**2. (Neighborhoods and relata)** If  $(X, \mathcal{T})$  is a topological space and  $p \in X$ , a (open) **neighborhood** of  $p$  is an open set  $U \subset X$  with  $p \in U$ .  $q \in X$  is an **accumulation point** of  $A \subset X$  if for every neighborhood  $U \ni q$ , it holds  $U \cap A \neq \emptyset$ . It turns out that  $C \subset X$  is closed if and only



if it contains its accumulation points. The union of the set of the accumulation points of  $A$  and  $A$  coincides with  $\overline{A}$ . The **interior**  $Int(A)$  of  $A \subset X$  is made of the points  $p \in A$  such that there is a neighborhood  $U \ni p$  with  $U \subset A$ . The **exterior**  $Ext(A)$  of  $A \subset X$  is made of the points  $p \in X$  such that there is a neighborhood  $U \ni p$  with  $U \subset X \setminus A$ . The **boundary** or **frontier**  $\partial A$  of a set  $A \subset X$  is made of the points  $x \in X$  which are accumulation points for both  $Int(A)$  and  $Ext(A)$ . It turns out that  $C \subset X$  is closed if and only if  $\partial C \subset C$ . Finally  $\overline{A} = A \cup \partial A$ .

**3. (Basis of a topology)** If  $(X, \mathcal{T})$  is a topological space, a family  $\mathcal{B} \subset \mathcal{T}$  is a topological **basis** if every  $A \in \mathcal{T}$  is the union of elements of  $\mathcal{B}$ .

**4. (Support of a function)** If  $X$  is a topological space and  $f : X \rightarrow \mathbb{R}$  is any function, the **support** of  $f$ ,  $\text{supp} f$ , is the closure of the set of the points  $x \in X$  with  $f(x) \neq 0$ .

**5. (Continuous functions)** If  $(X, \mathcal{T})$  and  $(Y, \mathcal{U})$  are topological spaces, a mapping  $f : X \rightarrow Y$  is said to be **continuous** if  $f^{-1}(T)$  is open for each  $T \in \mathcal{U}$ . The composition of continuous functions is a continuous function. An injective, surjective and continuous mapping  $f : X \rightarrow Y$ , whose inverse mapping is also continuous, is called **homeomorphism** from  $X$  to  $Y$ . If there is a homeomorphism from  $X$  to  $Y$  these topological spaces are said to be **homeomorphic**. There are properties of topological spaces and their subsets which are preserved under the action of homeomorphisms. These properties are called **topological properties**. As a simple example notice that if the topological spaces  $X$  and  $Y$  are homeomorphic under the homeomorphism  $h : X \rightarrow Y$ ,  $U \subset X$  is either open or closed if and only if  $h(U) \subset Y$  is such.

**6. (Open and closed functions)** If  $(X, \mathcal{T})$  and  $(Y, \mathcal{U})$  are topological spaces, a mapping  $f : X \rightarrow Y$  is said to be **open** if  $f(A) \in \mathcal{U}$  for each  $A \in \mathcal{T}$ . It is called **closed** if  $f(C)$  is a closed set in  $Y$  for every closed set  $C \subset X$ .

**7. (Second countability and Lindelöf's lemma)** A topological space which admits a *countable* basis of its topology is said to be **second countable**. **Lindelöf's lemma** proves that If  $(X, \mathcal{T})$  is second countable, from any covering of  $X$  made of open sets it is possible to extract a countable subcovering. It is clear that second countability is a topological property.

**8. (Generated topology)** It is a trivial task to show that, if  $\{\mathcal{T}_i\}_{i \in I}$  is a class of topologies on the set  $X$ ,  $\cap_{i \in I} \mathcal{T}_i$  is a topology on  $X$ , too.

If  $\mathcal{A}$  is a class of subsets of  $X \neq \emptyset$  and  $C_{\mathcal{A}}$  is the class of topologies  $\mathcal{T}$  on  $X$  with  $\mathcal{A} \subset \mathcal{T}$ ,  $\mathcal{T}_{\mathcal{A}} := \cap_{\mathcal{T} \in C_{\mathcal{A}}} \mathcal{T}$  is called the **topology generated by  $\mathcal{A}$** . Notice that  $C_{\mathcal{A}} \neq \emptyset$  because the set of parts of  $X$ ,  $\mathcal{P}(X)$ , is a topology and includes  $\mathcal{A}$ .

It is simply proved that if  $\mathcal{A} = \{B_i\}_{i \in I}$  is a class of subsets of  $X \neq \emptyset$ ,  $\mathcal{A}$  is a basis of the topology on  $X$  generated by  $\mathcal{A}$  itself if and only if  $B_i \cap B_j = \cup_{k \in K} B_k$  for every choice of  $i \in I, i' \in I'$  and a corresponding  $K \subset I$ .

**9. (Induced topology)** If  $A \subset X$ , where  $(X, \mathcal{T})$  is a topological space, the pair  $(A, \mathcal{T}_A)$  where,  $\mathcal{T}_A := \{U \cap A \mid U \in \mathcal{T}\}$ , defines a topology on  $A$  which is called the topology **induced** on  $A$  by  $X$ . The inclusion map, that is the map,  $i : A \hookrightarrow X$ , which sends every  $a$  viewed as an element of  $A$  into the same  $a$  viewed as an element of  $X$ , is continuous with respect to that topology. Moreover, if  $f : X \rightarrow Y$  is continuous,  $X, Y$  being topological spaces,  $f|_A : A \rightarrow f(A)$  is continuous with respect to the induced topologies on  $A$  and  $f(A)$  by  $X$  and  $Y$  respectively, for every subset  $A \subset X$ .

**10. ((finite) Product topology)** If  $(X_i, \mathcal{T}_i)$ , for  $i = 1, \dots, N$  (finite) are topological spaces,  $\times_{i=1}^N X_i$  can be endowed with the topology generated by all possible sets  $\times_{i=1}^N A_i$  with  $A_i \in \mathcal{T}_i$ . This topology is called the **product topology** and the products  $A_i \in \mathcal{T}_i$  form a basis of that topology. With this definition  $K_i \in \mathcal{T}_i$  is compact (see below) if every  $K_i \subset X_i$  are compact. Furthermore the surjective maps called **canonical projections**  $\pi_k : \times_{i=1}^N X_i \ni (x_1, \dots, x_N) \mapsto x_k \in X_k$  are open functions.

**11. (Continuity at a point)** If  $X$  and  $Y$  are topological spaces and  $x \in X$ ,  $f : X \rightarrow Y$  is said to be **continuous in**  $x$ , if for every neighborhood of  $f(x)$ ,  $V \subset Y$ , there is a neighborhood of  $x$ ,  $U \subset X$ , such that  $f(U) \subset V$ . It is simply proved that  $f : X \rightarrow Y$  as above is continuous if and only if it is continuous in every point of  $X$ .

**12. (Connected spaces)** A topological space  $(X, \mathcal{T})$  is said to be **connected** if there are no open sets  $A, B \neq \emptyset$  with  $A \cap B = \emptyset$  and  $A \cup B = X$ . A subset  $C \subset X$  is connected if it is a connected topological space when equipped with the topology induced by  $\mathcal{T}$ . Finally, it turns out that if  $f : X \rightarrow Y$  is continuous and the topological space  $X$  is connected, then  $f(Y)$  is a connected topological space when equipped with the topology induced by the topological space  $Y$ . In particular, connectedness is a topological property.

**13. (Connected components)** On a topological space  $(X, \mathcal{T})$ , the following equivalence relation can be defined:  $x \sim x'$  if and only if there is a connected subset of  $X$  including both  $x$  and  $x'$ . In this way  $X$  turns out to be decomposed as the disjoint union of the equivalence classes generated by  $\sim$ . Those maximal connected subsets are called the **connected components** of  $X$ . Each connected component is always closed and it is also open when the class of connected components is finite.

**14. (Path connection)** A topological space  $(X, \mathcal{T})$  is said to be **connected by paths** if, for each pair  $p, q \in X$  there is a continuous path  $\gamma : [0, 1] \rightarrow X$  such that  $\gamma(0) = p$ ,  $\gamma(1) = q$ . The definition can be extended to subset of  $X$  considered as topological spaces with respect to the induced topology. It turns out that a topological space connected by paths is connected. A connected topological space whose every point admits an open connected by path neighborhood is, in turn, connected by path.

Connectedness by paths is a topological property.

**15. (Compact sets 1)** If  $Y$  is any set in a topological space  $X$ , a **covering** of  $Y$  is a class  $\{X_i\}_{i \in I}$ ,  $X_i \subset X$  for all  $i \in I$ , such that  $Y \subset \cup_{i \in I} X_i$ . A topological space  $(X, \mathcal{T})$  is said to be **compact** if from each covering of  $X$  made of open sets,  $\{X_i\}_{i \in I}$ , it is possible to extract a covering  $\{X_j\}_{j \in J \subset I}$  of  $X$  with  $J$  finite. A subset  $K$  of a topological space  $X$  is said to be compact if it is compact as a topological space when endowed with the topology induced by  $X$  (this is equivalent to say that  $K \subset X$  is compact whenever every covering of  $K$  made of open sets of the topology of  $X$  admits a finite subcovering).

**16. (Compact sets 2)** If  $(X, \mathcal{T})$  and  $(Y, \mathcal{S})$  are topological spaces, the former is compact and  $\phi : X \rightarrow Y$  is continuous, then  $Y$  is compact. In particular compactness is a topological property.

Each closed subset of a compact set is compact. Similarly, if  $K$  is a compact set in a **Hausdorff topological space** (see below),  $K$  is closed. Each compact set  $K$  is **sequentially compact**, i.e., each sequence  $S = \{p_k\}_{k \in \mathbb{N}} \subset K$  admits some **accumulation point**  $s \in K$ , (i.e., each

neighborhood of  $s$  contains some element of  $S$  different from  $s$ ). If  $X$  is a topological **metric space** (see below), sequentially compactness and compactness are equivalent.

**17. (Hausdorff property)** A topological space  $(X, \mathcal{T})$  is said to be **Hausdorff** if each pair  $(p, q) \in X \times X$  admits a pair of neighborhoods  $U_p, U_q$  with  $p \in U_p, q \in U_q$  and  $U_p \cap U_q = \emptyset$ . If  $X$  is Hausdorff and  $x \in X$  is a limit of the sequence  $\{x_n\}_{n \in \mathbb{N}} \subset X$ , this limit is *unique*. Hausdorff property is a topological property.

**18. (Metric spaces)** A **semi metric space** is a set  $X$  endowed with a **semidistance**, that is  $d : X \times X \rightarrow [0, +\infty)$ , with  $d(x, y) = d(y, x)$  and  $d(x, y) + d(y, z) \geq d(x, z)$  for all  $x, y, z \in X$ . If  $d(x, y) = 0$  implies  $x = y$  the semidistance is called **distance** and the semi metric space is called **metric space**. Either in semi metric space or metric spaces, the **open metric balls** are defined as  $B_s(y) := \{z \in \mathbb{R}^n \mid d(z, y) < s\}$ .  $(X, d)$  admits a preferred topology called **metric topology** which is defined by saying that the open sets are the union of metric balls. Any metric topology is a Hausdorff topology. A set  $G \subset X$  is said to be **bounded** if  $B_s(y) \subset G$  for some  $y \in X$  and  $s \in [0, +\infty)$ . A sequence  $\{x_n\}_{n \in \mathbb{N}} \subset X$  **converges to**  $y \in X$  if  $\lim_{n \rightarrow +\infty} d(x_n, y) = 0$ . A sequence is **convergent** if it converges to some point of the metric space.

It is very simple to show that a mapping  $f : A \rightarrow M_2$ , where  $A \subset M_1$  and  $M_1, M_2$  are semimetric spaces endowed with the metric topology, is continuous with respect to the usual " $\epsilon - \delta$ " definition if and only if  $f$  is continuous with respect to the general definition of given above, considering  $A$  a topological space equipped with the metric topology induced by  $M_1$ .

**19. (Completeness of metric spaces)** A sequence  $\{x_n\}_{n \in \mathbb{N}} \subset X$  in a metric space  $(X, d)$  is said to be **Cauchy** if for every  $\epsilon > 0$  there is  $N_\epsilon > 0$  such that  $d(x_n, x_m) < \epsilon$  if  $n, m > N_\epsilon$ . Every convergent sequence is Cauchy. A metric space such that every Cauchy sequence is convergent is said to be **complete**. Completeness is **not** a topological property.

**20. (Normed spaces)** If  $X$  is a vector space with field  $\mathbb{K} = \mathbb{C}$  or  $\mathbb{R}$ , a semidistance and thus a topology can be induced by a **seminorm**. A semi norm on  $X$  is a mapping  $p : X \rightarrow [0, +\infty)$  such that  $p(av) = |a|p(v)$  for all  $a \in \mathbb{K}, v \in X$  and  $p(u + v) \leq p(u) + p(v)$  for all  $u, v \in X$ . If  $p$  is a seminorm on  $V$ ,  $d(u, v) := p(u - v)$  is the **semidistance induced by  $p$** . A seminorm  $p$  such that  $p(v) = 0$  implies  $v = 0$  is called **norm**. In this case the semidistance induced by  $p$  is a distance.

**21. (Completeness and topological equivalence of finite-dimensional normed spaces)** It turns out that if a normed space  $(X, p)$  (on  $\mathbb{R}$  or  $\mathbb{C}$ ) has finite dimension, then it is complete as a metric space with the distance  $d(x, y) := p(x - y)$  and all norms on  $X$  produce the same topology.

### 2.1.2 The topology of $\mathbb{R}^n$

A few words about the usual topology of the real vector space  $\mathbb{R}^n$  are in order [Seri90]. That topology, also called the **Euclidean topology**, is a metric topology induced by the usual distance

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

where  $x = (x_1, \dots, x_n)$  and  $y = (y_1, \dots, y_n)$  are points of  $\mathbb{R}^n$ . That distance can be induced by a norm  $\|x\| = \sqrt{\sum_{i=1}^n (x_i)^2}$ . As a consequence, an open set with respect to that topology is any set  $A \subset \mathbb{R}^n$  such that either  $A = \emptyset$  or each  $x \in A$  is contained in a open metric ball  $B_r(x) \subset A$  (if  $s > 0$ ,  $y \in \mathbb{R}^n$ ,  $B_s(y) := \{z \in \mathbb{R}^n \mid \|z - y\| < s\}$ ). The open balls with arbitrary center and radius are a basis of the Euclidean topology. A relevant property of the Euclidean topology of  $\mathbb{R}^n$  is that it admits a countable basis i.e., it is second countable. To prove that it is sufficient to consider the open balls with rational radius and center with rational coordinates. It turns out that any open set  $A$  of  $\mathbb{R}^n$  (with the Euclidean topology) is connected by paths if it is open and connected. It turns out that a set  $K$  of  $\mathbb{R}^n$  endowed with the Euclidean topology is compact if and only if  $K$  is closed and bounded. As a metric space  $\mathbb{R}^n$  is complete since it is finite dimensional and every norm on it produces the same topology.

### Exercises 2.1.

1. Show that  $\mathbb{R}^n$  endowed with the Euclidean topology is Hausdorff.
2. Show that the open balls in  $\mathbb{R}^n$  with rational radius and center with rational coordinates define a countable basis of the Euclidean topology.  
(*Hint.* Show that the considered class of open balls is countable because there is a one-to-one mapping from that class to  $\mathbb{Q}^n \times \mathbb{Q}$ . Then consider any open set  $U \subset \mathbb{R}^n$ . For each  $x \in U$  there is an open ball  $B_{r_x}(x) \subset U$ . Since  $\mathbb{Q}$  is dense in  $\mathbb{R}$ , one may change the center  $x$  to  $x'$  with rational coordinates and the radius  $r_x$  to  $r'_{x'}$  which is rational, in order to preserve  $x \in C_x := B_{r'_{x'}}(x')$ . Then show that  $\cup_x C_x = U$ .)
3. Consider the subset of  $\mathbb{R}^2$ ,  $C := \{(x, \sin \frac{1}{x}) \mid x \in (0, 1]\} \cup \{(x, y) \mid x = 0, y \in \mathbb{R}\}$ . Is  $C$  path connected? Is  $C$  connected?
4. Show that the disk  $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$  is homeomorphic to  $\mathbb{R}^2$ . Generalize the result to any open ball (with center and radius arbitrarily given) in  $\mathbb{R}^n$ . (*Hint.* Consider the mapping  $(x, y) \mapsto (x/(1 - \sqrt{x^2 + y^2}), y/(1 - \sqrt{x^2 + y^2}))$ . The generalization is straightforward).
5. Let  $f : M \rightarrow N$  be a continuous bijective mapping and  $M, N$  topological spaces, show that  $f$  is a homeomorphism if  $N$  is Hausdorff and  $M$  is compact.  
(*Hint.* Start by showing that a mapping  $F : X \rightarrow Y$  is continuous if and only if for every closed set  $K \subset Y$ ,  $F^{-1}(K)$  is closed. Then prove that  $f^{-1}$  is continuous using the properties of compact sets in Hausdorff spaces.)
6. Consider two topological spaces  $X, Y$  and their product  $X \times Y$  equipped with the product topology. Prove that, for every  $x \in X$  the map  $Y \ni y \mapsto (x, y) \in L_x := \{(x, z) \mid z \in Y\}$  is an homeomorphism if  $L_x$  is equipped with the topology induced from  $X \times Y$ .

### 2.1.3 Topological manifolds

**Definition 2.2.** (**Topological Manifold.**) A topological space  $(X, \mathcal{T})$  is called **topological manifold** of dimension  $n$  if  $X$  is Hausdorff, second countable and is **locally homeomorphic** to  $\mathbb{R}^n$ , that is, for every  $p \in X$  there is a neighborhood  $U_p \ni p$  and a homeomorphism  $\phi : U_p \rightarrow V_p$  where  $V_p \subset \mathbb{R}^n$  is an open set (equipped with the topology induced by  $\mathbb{R}^n$ ). ■

**Remark 2.3.**

- (1) The homeomorphism  $\phi$  may have co-domain given by  $\mathbb{R}^n$  itself.
- (2) We have assumed that  $n$  is fixed, anyway one may consider a Hausdorff connected topological space  $X$  with a countable basis and such that, for each  $x \in X$  there is a homeomorphism defined in a neighborhood of  $x$  which maps that neighborhood into  $\mathbb{R}^n$  where  $n$  may depend on the neighborhood and the point  $x$ . An important theorem due to Whitehead shows that, actually,  $n$  must be a constant if  $X$  is connected. This result is usually stated by saying that *the dimension of a topological manifold is a topological invariant*.
- (3) The Hausdorff requirement could seem redundant since  $X$  is locally homeomorphic to  $\mathbb{R}^n$  which is Hausdorff. The following example shows that this is not the case. Consider the set  $X := \mathbb{R} \cup \{p_0\}$  where  $p_0 \notin \mathbb{R}$ . Define a topology on  $X$ ,  $\mathcal{T}$ , given by all of the sets which are union of elements of  $\mathcal{E} \cup \mathcal{T}_{p_0}$ , where  $\mathcal{E}$  is the usual Euclidean topology of  $\mathbb{R}$  and  $U \in \mathcal{T}_{p_0}$  iff  $U = (V_0 \setminus \{0\}) \cup \{p_0\}$ ,  $V_0$  being any neighborhood of 0 in  $\mathcal{E}$ . The reader should show that  $\mathcal{T}$  is a topology. It is obvious that  $(X, \mathcal{T})$  is not Hausdorff since there are no open sets  $U, V \in \mathcal{T}$  with  $U \cap V = \emptyset$  and  $0 \in U$ ,  $p_0 \in V$ . Nevertheless, each point  $x \in X$  admits a neighborhood which is homeomorphic to  $\mathbb{R}$ :  $R = \{p_0\} \cup (\mathbb{R} \setminus \{0\})$  is homeomorphic to  $\mathbb{R}$  itself and is an open neighborhood of  $p_0$ . It is trivial to show that there are sequences in  $X$  which admit two different limits.
- (4) The simplest example of topological manifold is  $\mathbb{R}^n$  itself. An apparently less trivial example is an open ball (with finite radius) of  $\mathbb{R}^n$ . However it is possible to show (see exercise 2.1.4) that an open ball (with finite radius) of  $\mathbb{R}^n$  is homeomorphic to  $\mathbb{R}^n$  itself so this example is rather trivial anyway. One might wonder if there are natural mathematical objects which are topological manifolds with dimension  $n$  but are not  $\mathbb{R}^n$  itself or homeomorphic to  $\mathbb{R}^n$  itself. A simple example is a sphere  $\mathbb{S}^2 \subset \mathbb{R}^3$ .  $\mathbb{S}^2 := \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$ .  $\mathbb{S}^2$  is a topological space equipped with the topology induced by  $\mathbb{R}^3$  itself. It is obvious that  $\mathbb{S}^2$  is Hausdorff and has a countable basis (the reader should show it). Notice that  $\mathbb{S}^2$  is not homeomorphic to  $\mathbb{R}^2$  because  $\mathbb{S}^2$  is **compact** (being closed and bounded in  $\mathbb{R}^3$ ) and  $\mathbb{R}^2$  is not compact since it is not bounded.  $\mathbb{S}^2$  is a topological manifold of dimension 2 with local homeomorphisms defined as follows. Consider  $p \in \mathbb{S}^2$  and let  $\Pi_p$  be the plane tangent at  $\mathbb{S}^2$  in  $p$  equipped with the topology induced by  $\mathbb{R}^3$ . With that topology  $\Pi_p$  is homeomorphic to  $\mathbb{R}^2$  (the reader should prove it). Let  $\phi$  be the orthogonal projection of  $\mathbb{S}^2$  on  $\Pi_p$ . It is quite simply proved that  $\phi$  is continuous with respect to the considered topologies and  $\phi$  is bijective with continuous inverse when restricted to the open semi-sphere which contains  $p$  as the south pole. Such a restriction defines a homeomorphism from a neighborhood of  $p$  to an open disk of  $\Pi_p$  (that is  $\mathbb{R}^2$ ). The same procedure can be used to define local homeomorphisms referred to neighborhoods of each point of  $\mathbb{S}^2$ . ■

## 2.2 Differentiable manifolds

If  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  it is obvious the meaning of the statement " $f$  is differentiable". However, in mathematics and in physics there exist objects which look like  $\mathbb{R}^n$  but are not  $\mathbb{R}^n$  itself (e.g. the sphere  $\mathbb{S}^2$  considered above), and it is useful to consider real valued mappings  $f$  defined on these objects. What about the meaning of " $f$  is differentiable" in these cases? A simple example is given, in mechanics, by the configuration space of a material point which is constrained to belong to a circle  $\mathbb{S}^1$ .  $\mathbb{S}^1$  is a topological manifold. There are functions defined on  $\mathbb{S}^1$ , for instance the mechanical energy of the point, which are assumed to be "differentiable functions". What does it mean? An answer can be given by a suitable definition of a differentiable manifold. To that end we need some preliminary definitions.

### 2.2.1 Local charts and atlas

**Definition 2.4.** ( *$k$ -compatible local charts.*) Consider a topological manifold  $M$  with dimension  $n$ . A **local chart** or **local coordinate system** on  $M$  is pair  $(V, \phi)$  where  $V \subset M$  is open,  $V \neq \emptyset$ , and  $\phi : p \mapsto (x^1(p), \dots, x^n(p))$  is a homeomorphism from  $U$  to the open set  $\phi(U) \subset \mathbb{R}^n$ . Moreover:

- (a) local chart  $(V, \phi)$  is called **global** chart if  $V = M$ ;
- (b) two local charts  $(V, \phi), (U, \psi)$  are said to be  $C^k$ -**compatible**,  $k \in (\mathbb{N} \setminus \{0\}) \cup \{\infty\}$ , if either  $U \cap V = \emptyset$  or, both  $\phi \circ \psi^{-1} : \psi(U \cap V) \rightarrow \mathbb{R}^n$  and  $\psi \circ \phi^{-1} : \phi(U \cap V) \rightarrow \mathbb{R}^n$  are of class  $C^k$ . ■

The given definition allow us to define a *differentiable atlas* of order  $k \in (\mathbb{N} \setminus \{0\}) \cup \{\infty\}$ .

**Definition 2.5.** (**Atlas on a manifold.**) Consider a topological manifold  $M$  with dimension  $n$ . A **differentiable atlas** of **order**  $k \in (\mathbb{N} \setminus \{0\}) \cup \{\infty\}$  on  $M$  is a class of local charts  $\mathcal{A} = \{(U_i, \phi_i)\}_{i \in I}$  such that :

- (a)  $\mathcal{A}$  covers  $M$ , i.e.,  $M = \cup_{i \in I} U_i$ ,
- (b) the charts of  $\mathcal{A}$  are pairwise  $C^k$ -compatible. ■

**Remark 2.6.** An atlas of order  $k \in \mathbb{N} \setminus \{0\}$  is an atlas of order  $k-1$  too, provided  $k-1 \in \mathbb{N} \setminus \{0\}$ . An atlas of order  $\infty$  is an atlas of all orders. ■

### 2.2.2 Differentiable structures

Finally, we give the definition of *differentiable structure* and *differentiable manifold* of order  $k \in (\mathbb{N} \setminus \{0\}) \cup \{\infty\}$ .

**Definition 2.7.** ( *$C^k$ -differentiable structure.*) Consider a topological manifold  $M$  with dimension  $n$ , a **differentiable structure** of **order**  $k \in (\mathbb{N} \setminus \{0\}) \cup \{\infty\}$  on  $M$  is an atlas  $\mathcal{M}$  of

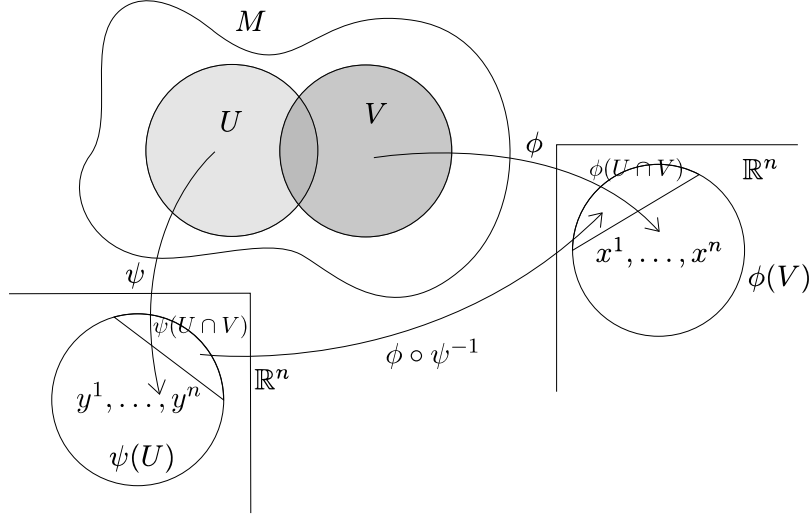


Figure 2.1: Two local charts on the differentiable manifold  $M$

order  $k$  which is maximal with respect to the  $C^k$ -compatibility requirement. In other words if  $(U, \phi) \notin \mathcal{M}$  is a local chart on  $M$ ,  $(U, \phi)$  is not  $C^k$ -compatible with some local chart of  $\mathcal{M}$ . ■

**Definition 2.8.** ( $C^k$ -differentiable manifold.) A topological manifold of dimension  $n$  equipped with a differentiable structure of order  $k \in (\mathbb{N} \setminus \{0\}) \cup \{\infty\}$  is said to be a **differentiable manifold** of dimension  $n$  and order  $k$ . A **smooth manifold** is differentiable manifolds of order  $C^\infty$ . ■

We henceforth denote the dimension of a manifold  $M$  with the symbol  $\dim(M)$ . We leave to the reader the proof of the following proposition.

**Proposition 2.9.** Referring to definition 2.7, if the local charts  $(U, \phi)$  and  $(V, \psi)$  are separately  $C^k$  compatible with all the charts of a  $C^k$  atlas, then  $(U, \phi)$  and  $(V, \psi)$  are  $C^k$  compatible.

This result implies that given a  $C^k$  atlas  $\mathcal{A}$  on a topological manifold  $M$ , there is exactly one  $C^k$ -differentiable structure  $\mathcal{M}_{\mathcal{A}}$  such that  $\mathcal{A} \subset \mathcal{M}_{\mathcal{A}}$ . This is the differentiable structure which is called **generated** by  $\mathcal{A}$ .  $\mathcal{M}_{\mathcal{A}}$  is nothing but the union of  $\mathcal{A}$  with the class of all of the local charts which are compatible with every chart of  $\mathcal{A}$ .

**Remark 2.10.** If  $M$  is a differentiable manifold with differentiable structure  $\mathcal{M}$  and  $U \subset M$  is open, it is possible to define on  $U$  a differentiable structure  $\mathcal{M}(U)$  **induced** by  $M$  of the same order and dimension. This happens when equipping  $U$  with the topology induced by  $M$  and, for

every local chart  $(V, \phi) \in \mathcal{M}$  with  $V \cap U \neq \emptyset$ , defining a corresponding local chart  $(U \cap V, \phi|_{U \cap V})$  on  $U$ . By definition,  $\mathcal{M}(U)$  is the differentiable structure induced by the atlas constructed in that way on  $U$ . ■

### Examples 2.11.

(1)  $\mathbb{R}^n$  has a natural structure of smooth manifold which is connected and path connected. The differentiable structure is that generated by the atlas containing the global chart given by the canonical coordinate system, i.e., the components of each vector with respect to the canonical basis.

(2) Consider a **real  $n$ -dimensional affine space**,  $\mathbb{A}^n$ . This is a triple  $(\mathbb{A}^n, V, \vec{\cdot})$  where  $\mathbb{A}^n$  is a set whose elements are called **points**,  $V$  is a real  $n$ -dimensional vector space and  $\vec{\cdot} : \mathbb{A}^n \times \mathbb{A}^n \rightarrow V$  is a mapping such that the two following requirements are fulfilled.

(i) For each pair  $P \in \mathbb{A}^n$ ,  $v \in V$  there is a *unique* point  $Q \in \mathbb{A}^n$  such that  $\overrightarrow{PQ} = v$ .

(ii)  $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$  for all  $P, Q, R \in \mathbb{A}^n$ .

$\overrightarrow{PQ}$  is called vector with **initial point**  $P$  and **final point**  $Q$ . An affine space equipped with a (pseudo) scalar product (defined on the vector space) is called **(pseudo) Euclidean space**.

Each affine space is a connected and path-connected topological manifold with a natural  $C^\infty$  differential structure. These structures are built up by considering the class of natural *global* coordinate systems, the **Cartesian coordinate systems**, obtained by fixing a point  $O \in \mathbb{A}^n$  and a vector basis for the vectors with initial point  $O$ . Varying  $P \in \mathbb{A}^n$ , the components of each vector  $\overrightarrow{OP}$  with respect to the chosen basis, define a bijective mapping  $f : \mathbb{A}^n \rightarrow \mathbb{R}^n$  and the Euclidean topology of  $\mathbb{R}^n$  induces a topology on  $\mathbb{A}^n$  by defining the open sets of  $\mathbb{A}^n$  as the sets  $B = f^{-1}(D)$  where  $D \subset \mathbb{R}^n$  is open. That topology does not depend on the choice of  $O$  and the basis in  $V$  and makes the affine space a topological  $n$ -dimensional manifold. Notice also that each mapping  $f$  defined above gives rise to a  $C^\infty$  atlas. Moreover, if  $g : \mathbb{A}^n \rightarrow \mathbb{R}^n$  is another mapping defined as above with a different choice of  $O$  and the basis in  $V$ ,  $f \circ g^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $g \circ f^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  are  $C^\infty$  because they are linear non homogeneous transformations. Therefore, there is a  $C^\infty$  atlas containing all of the Cartesian coordinate systems defined by different choices of origin  $O$  and basis in  $V$ . The  $C^\infty$ -differentiable structure generated by that atlas naturally makes the affine space a  $n$ -dimensional  $C^\infty$ -differentiable manifold.

(3) The sphere  $\mathbb{S}^2$  defined above gets a  $C^\infty$ -differentiable structure as follows. Considering all of local homeomorphisms defined in remark (4) above, they turn out to be  $C^\infty$  compatible and define a  $C^\infty$  atlas on  $\mathbb{S}^2$ . That atlas generates a  $C^\infty$ -differentiable structure on  $\mathbb{S}^n$ . (Actually it is possible to show that the obtained differentiable structure is the only one compatible with the natural differentiable structure of  $\mathbb{R}^3$ , when one requires that  $\mathbb{S}^2$  is an *embedded submanifold* of  $\mathbb{R}^3$ .)

(4) A classical theorem by Whitney shows that if a topological manifold admits a  $C^1$ -differentiable structure, then it admits a  $C^\infty$ -differentiable structure which is contained in the former. Moreover a topological  $n$ -dimensional manifold may admit none or several different and *not diffeomorphic* (see below)  $C^\infty$ -differentiable structures. E.g., it happens for  $n = 4$ . ■



Few words about the notion of *orientation* are in order.

**Definition 2.12.** Let  $M$  be a  $C^k$  manifold with  $k \geq 1$ .

- (a) An atlas  $\mathcal{A} = \{(U_i, \phi_i)\}_{i \in I}$  on  $M$  is called **oriented** if the determinant of the Jacobian matrix of  $\phi_i \circ \phi_j^{-1}$  is everywhere positive for all  $i, j \in I$ .
- (b)  $M$  is said to be **orientable** if it admits an oriented atlas.
- (c) A maximal oriented atlas on  $M$  (supposed to be orientable) is called **orientation** and an **oriented manifold** is an orientable manifold with the choice of a preferred orientation. ■

**Remark 2.13.**

(1) It is easy to prove that an orientable manifold  $M$  admits orientations as a consequence of the Zorn lemma. If  $M$  is also connected, then it admits exactly *two* orientations.

(2) An example of a manifold which is not orientable is the *Möbius strip* we shall introduce in (1) Examples 3.30. ■

To conclude, we introduce a useful technical notion. Given two differentiable manifolds  $M$  and  $N$  of the same order  $k$ , we can construct a third one of the order  $k$ , on their Cartesian product  $M \times N$ .

**Definition 2.14. (Product manifold.)** Let  $M$  and  $N$  be two differentiable manifolds of order  $k$ . The **product (differentiable) manifold**  $M \times N$  is the differentiable manifold of order  $k$  and dimension  $\dim(M) + \dim(N)$  constructed on the Cartesian product  $M \times N$ , equipped with the product topology and whose differentiable structure is induced by the following atlas.

$$\mathcal{A}(M \times N) := \{(U \times V, \phi \oplus \psi) \mid (U, \phi) \in \mathcal{A}(M), (V, \psi) \in \mathcal{A}(N)\}$$

where  $\mathcal{A}(M)$  and  $\mathcal{A}(N)$  are the differentiable structures on  $M$  and  $N$ , respectively, and

$$\phi \oplus \psi : M \times N \ni (p, q) \mapsto (\phi(p), \psi(q)) \in \mathbb{R}^{\dim(M) + \dim(N)}$$

■

We develop the theory in the  $C^\infty$  case only. However, several definitions and results may be generalized to the  $C^k$  case with  $1 \leq k < \infty$ .

**Exercises 2.15.**

1. Show that the group  $SO(3)$  is a three-dimensional smooth manifold.
2. Prove that the cylinder  $C := \{(x, y, z) \mid z \in \mathbb{R}, x^2 + y^2 = 1\}$  equipped with the natural  $C^\infty$ -differentiable structure constructed similarly to (3) Examples 2.11 can also be constructed as the product manifold of  $\mathbb{R}$  and a circle in  $\mathbb{R}^2$  equipped with the natural smooth differentiable structure constructed similarly to (3) Examples 2.11.

### 2.2.3 Differentiable functions and diffeomorphisms

Equipped with the given definitions, we can state the definition of a differentiable function.

**Definition 2.16. (Differentiable functions and curves.)** Consider a continuous mapping  $f : M \rightarrow N$ , where  $M$  and  $N$  are smooth manifolds with dimension  $m$  and  $n$ .

- (a)  $f$  is said to be  $C^\infty$ -**differentiable** or **smooth** at  $p \in M$  if the function:

$$\psi \circ f \circ \phi^{-1} : \phi(U) \rightarrow \mathbb{R}^n,$$

is differentiable, for some local charts  $(V, \psi)$ ,  $(U, \phi)$  on  $N$  and  $M$  respectively with  $p \in U$ ,  $f(p) \in V$  and  $f(U) \subset V$ .

- (b)  $f$  is said to be **smooth** if it is smooth at every point of  $M$ . In particular, if  $M = I \subset \mathbb{R}$  an open interval, then  $f$  above is said a **smooth curve** in  $N$ .

The class of all smooth functions from  $M$  to  $N$  is indicated by  $D(M|N)$  or  $D(M)$  for  $N = \mathbb{R}$ . Evidently,  $D(M)$  is a **real vector space** with the vector space structure induced by

$$(af + bg)(p) := af(p) + bg(p) \quad \text{for every } p \in M \text{ if } a, b \in \mathbb{R} \text{ and } f, g \in D(M).$$

■

**Definition 2.17. (Diffeomorphisms)** Let  $M$  and  $N$  be smooth manifolds.

- (a) if  $f \in D(M|N)$  is bijective and  $f^{-1} \in D(N|M)$ , then  $f$  is called (smooth) **diffeomorphism** from  $M$  to  $N$ .
- (b) If there is a diffeomorphism from the differentiable manifold  $M$  to the differentiable manifold  $N$ ,  $M$  and  $N$  are said to be **diffeomorphic** (through  $f$ ).
- (c) If  $M$  is orientable and  $N = M$ , a diffeomorphism  $f : M \rightarrow M$  is said to be **orientation-preserving** if, for every local chart  $(U, \phi)$ , the local chart  $(f(U), \phi \circ f^{-1})$  stays in the same orientation of the former.

■

**Remark 2.18.**

- (1) It is clear that a smooth function (at a point  $p$ ) is continuous (in  $p$ ) by definition.
- (2) It is simply proved that the definition of function smooth at a point  $p$  does not depend on the choice of the local charts used in (1) of the definition above.
- (3) Notice that  $D(M)$  is also a commutative **ring** with multiplicative and additive unit elements if endowed with the product rule  $f \cdot g : p \mapsto f(p)g(p)$  for all  $p \in M$  and sum rule  $f + g : p \mapsto f(p) + g(p)$  for all  $p \in M$ . The unit elements with respect to the product and sum are respectively the constant function 1 and the constant function 0. However  $D(M)$  is not a field, because there are elements  $f \in D(M)$  with  $f \neq 0$  without (multiplicative) inverse element.

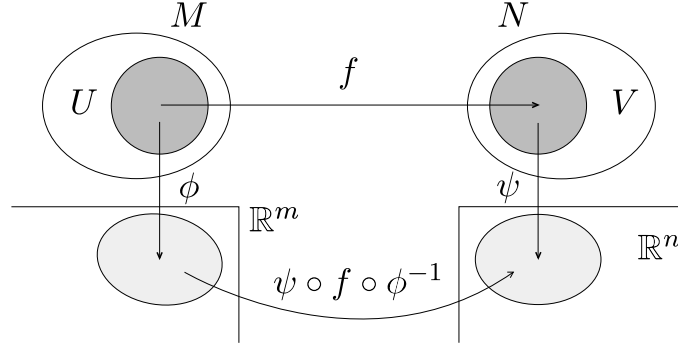


Figure 2.2: Coordinate representation of  $f : M \rightarrow N$

It is sufficient to consider  $f \in D(M)$  with  $f(p) = 0$  and  $f(q) \neq 0$  for some  $p, q \in M$ .

(4) Consider two smooth manifolds  $M$  and  $N$  such that they are defined on the same topological space but they can have different differentiable structures. Suppose also that they are diffeomorphic. Can we conclude that  $M = N$ ? In other words:

*Is it true that the smooth differentiable structure of  $M$  coincides with the differentiable structure of  $N$  whenever  $M$  and  $N$  are defined on the same topological space and are diffeomorphic?*

The following example shows that the answer can be *negative*. Consider  $M$  and  $N$  as one-dimensional smooth manifolds whose associated topological space is  $\mathbb{R}$  equipped with the usual Euclidean topology. The differentiable structure of  $M$  is defined as the differentiable structure generated by the atlas made of the global chart  $f : M \rightarrow \mathbb{R}$  with  $f : x \mapsto x$ , whereas the differentiable structure of  $N$  is given by the assignment of the global chart  $g : N \rightarrow \mathbb{R}$  with  $g : x \mapsto x^3$ . Notice that the smooth differentiable structure of  $M$  differs from that of  $N$  because  $f \circ g^{-1} : \mathbb{R} \rightarrow \mathbb{R}$  is not differentiable in  $x = 0$ . On the other hand  $M$  and  $N$  are diffeomorphic! Indeed a diffeomorphism is nothing but the map  $\phi : M \rightarrow N$  completely defined by requiring that  $g \circ \phi \circ f^{-1} : x \mapsto x$  for every  $x \in \mathbb{R}$ .

(5) A subsequent very intriguing question arises by the remark (4):

*Is there a topological manifold with dimension  $n$  which admits smooth differentiable structures which are not diffeomorphic, differently from the example given above?*

The answer is yes. More precisely, it is possible to show that  $1 \leq n < 4$  the answer is negative, but for some other values of  $n$ , in particular  $n = 4$ , there are topological manifolds which admit smooth differentiable structures that are not diffeomorphic. When the manifold is  $\mathbb{R}^n$  or a submanifold, with the usual topology and the usual differentiable structure, the remaining non-diffeomorphic smooth differentiable structures are said to be *exotic*. The first example was found by Whitney on the sphere  $\mathbb{S}^7$ . Later it was proved that the same space  $\mathbb{R}^4$  admits exotic structures. Finally, if  $n \geq 4$  once again, there are examples of topological manifolds which do not admit a smooth differentiable structure (also up to homeomorphisms).

It is intriguing to remark that 4 is the dimension of the spacetime.

(6) Similarly to differentiable manifolds, it is possible to define *analytic* manifolds. In that case all the involved functions used in changes of coordinate frames,  $f : U \rightarrow \mathbb{R}^n$  ( $U \subset \mathbb{R}^n$ ) must be analytic (i.e. that must admit Taylor expansion in a neighborhood of any point  $p \in U$ ). Analytic manifolds are convenient spaces when dealing with Lie groups. (Actually the celebrated *Gleason-Montgomery-Zippin theorem*, solving Hilbert's fifth problem, shows that a  $C^1$ -differentiable Lie groups is also an analytic Lie group.) It is simply proved that an affine space admits a natural analytic atlas and thus a natural analytic manifold structure obtained by restricting the natural differentiable structure. ■

## 2.3 Technicalities: smooth partitions of unity

In this section we present a few technical results which are very useful in several topics of differential geometry and tensor analysis. The first two lemmata concerns the existence of particular smooth functions which have compact support containing a fixed point of the manifold. These functions are very useful in several applications and basic constructions of differential geometry (see next chapter).

**Lemma 2.19.** *If  $x \in \mathbb{R}^n$  and  $x \in B_r(x) \subset \mathbb{R}^n$  where  $B_r(x)$  is any open ball centered in  $x$  with radius  $r > 0$ , there is a neighborhood  $G_x$  of  $x$  with  $\overline{G_x} \subset B_r(x)$  and a smooth function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  such that:*

- (1)  $0 \leq f(y) \leq 1$  for all  $y \in \mathbb{R}^n$ ,
- (2)  $f(y) = 1$  if  $y \in \overline{G_x}$ ,
- (3)  $f(y) = 0$  if  $y \notin B_r(x)$ .

**Proof.** Define

$$\alpha(t) := e^{\frac{1}{(t+r)(t+r/2)}}$$

for  $t \in [-r, -r/2]$  and  $\alpha(r) = 0$  outside  $[-r, -r/2]$ .  $\alpha \in C^\infty(\mathbb{R})$  by construction. Then define:

$$\beta(t) := \frac{\int_{-\infty}^t \alpha(s) ds}{\int_{-r}^{-r/2} \alpha(s) ds}.$$

This  $C^\infty(\mathbb{R})$  function is non-negative, vanishes for  $t \leq -r$  and takes the constant value 1 for  $t \geq -r/2$ . Finally define, for  $y \in \mathbb{R}^n$ :

$$f(y) := \beta(-||x - y||).$$

This function is  $C^\infty(\mathbb{R}^n)$  and non-negative, it vanishes for  $||x - y|| \geq r$  and takes the constant value 1 if  $||x - y|| \leq r/2$  so that  $G_r = B_{r/2}(x)$  □.

**Lemma 2.20.** *Let  $M$  be a smooth manifold. For every  $p \in M$  and every open neighborhood of  $p$ ,  $U_p$ , there is a open neighborhood of  $p$ ,  $V_p$  and a function  $h \in D(M)$  such that:*

- (1)  $\overline{V_p} \subset U_p$ ,
- (2)  $0 \leq h(q) \leq 1$  for all  $q \in M$ ,
- (3)  $h(q) = 1$  if  $q \in \overline{V_p}$ ,
- (4)  $\text{supp } h \subset U_p$  is compact.

$h$  is called **hat function** centered on  $p$  with support contained in  $U_p$ .

**Proof.** First of all we notice the following elementary result.

Let  $X$  be a topological Hausdorff space,  $Y \subset X$  is any open set, and  $Z \subset Y$  another open set with respect to the topology on  $Y$  induced by  $X$ . If the closure  $\overline{Z}^Y$  of  $Z$  referred to  $Y$  is compact then  $\overline{Z}^Y = \overline{Z}^X$  the closure of  $Z$  with respect to  $X$ . (Indeed, as  $\overline{Z}^Y$  is compact in  $Y$  it is compact in  $X$  and thus closed since  $X$  is Hausdorff. Therefore  $\overline{Z}^Y \supset \overline{Z}^X$  since the former is a closed set in  $X$  that contains  $Z$  and the latter is the smallest closed set in  $X$  including  $Z$ . On the other hand since  $\overline{Z}^X (\subset \overline{Z}^Y) \subset Y$  is closed in  $X$  and  $Y$  is open,  $\overline{Z}^X$  has to be closed in  $Y$  too because  $Y \setminus \overline{Z}^X = Y \cap (X \setminus \overline{Z}^X)$  is open it being the intersection of two open sets. As  $\overline{Z}^Y$  is the smallest closed set containing  $Z$  and  $\overline{Z}^X \supset Z$ , also the converse inclusion  $\overline{Z}^Y \subset \overline{Z}^X$  holds. Thus  $\overline{Z}^Y = \overline{Z}^X$ .)

Let us start with our construction. Obviously we can always assume that  $U_p$  is sufficiently small around the point  $p$ . In particular we can assume that there exists a local chart  $(W, \phi)$  with  $p \in W$  and  $U_p \subset W$ . In this case, define  $x := \phi(p)$  and fix  $r > 0$  such that  $U_p \supset \phi^{-1}(\overline{B_r(x)})$ ,  $B_r(x)$  being an open ball with finite radius  $r$  centered on  $x := \phi(p)$ . Finally, referring to lemma 2.19, define  $V_p := \phi^{-1}(G_x)$  so that  $\overline{V_p}^W = \phi^{-1}(\overline{G_x})$  because  $\phi : W \rightarrow \phi(W)$  is a homeomorphism. For the same reason, as  $\overline{G_x}$  is compact,  $\overline{V_p}^W$  is compact as well and one has  $\overline{V_p} := \overline{V_p}^M = \overline{V_p}^W = \phi^{-1}(\overline{G_x}) \subset \phi^{-1}(\overline{B_r(x)}) \subset U_p$  and (1) holds true consequently. Finally define  $h(q) := f(\phi(q))$  for  $q \in W$  and  $h(q) := 0$  if  $q \notin W$ . The support of  $f$  in  $\phi(W)$  is compact included in  $\overline{B_r(x)}$  and  $\phi^{-1}$  maps the support of  $f$  to the support of  $h$  in  $W$  which results to be compact as well. The support of  $h$  in  $W$  therefore coincides with the support of  $h$  in  $M$ . In particular  $\text{supp } h \subset \phi^{-1}(\overline{B_r(x)}) = \overline{\phi^{-1}(B_r(x))}^W = \overline{\phi^{-1}(B_r(x))}^M \subset U_p$ . By construction the function  $h$  also satisfies all requirements (2)-(4) barring the smoothness property we go to prove. To establish it, it is enough to prove that, for every  $q \in M$  there is a local chart  $(U_q, \psi_q)$  such that  $h \circ \psi_q^{-1}$  is  $C^\infty(\psi_q(U_q))$ . If  $q \in W$  the local chart is  $(W, \phi)$  itself. If  $q \notin W$ , there are two possibilities. In the first case, there is an open set  $U_q$  including  $q$  and with  $U_q \cap W = \emptyset$ . Taking  $U_q$  sufficiently small, we can assume that a local chart  $(U_q, \psi_q)$  exists. Therefore  $h \circ \psi_q^{-1} = 0$  constantly so that it is  $C^\infty(\psi_q(U_q))$ . It remains to consider the case  $q \in \partial W$ , i.e. every open set including  $q$  has non-empty intersection with  $W$ , although  $q \notin W$ . It would be enough to prove that, if  $q \in \partial W$ , there is an open set  $U_q \ni q$  with  $U_q \cap \text{supp } h = \emptyset$ . (In this way, restricting  $U_q$  if necessary, one has a local chart  $(U_q, \psi_q)$  with  $q \in U_q$  and  $h \circ \psi_q^{-1} = 0$  so that  $h \circ \psi_q^{-1}$  is smooth.) Let us prove the existence of such a  $U_q$ . Since  $q \notin W$  we have  $q \notin \text{supp } h \subset W$ . Since  $M$  is Hausdorff, for every  $x \in \text{supp } h$  there are open neighborhoods  $O_x$  and  $O_q^{(x)}$  of  $x$  and  $q$

respectively with  $O_x \cap O_q^{(x)} = \emptyset$ ; Since  $\text{supp } h$  is compact we can extract a finite covering of  $\text{supp } h$  made of open sets  $O_{x_i}$ ,  $i = 1, 2, \dots, N$  and consider the open set  $U_q := \cap_{i=1, \dots, N} O_q^{(x_i)}$ . By construction  $U_q \cap \text{supp } h \subset U_q \cap_{i=1, \dots, N} O_q^{(x_i)} = \emptyset$  as wanted.  $\square$

**Remark 2.21.** Hausdorff property plays a central role in proving the smoothness of hat functions defined in the whole manifold by the natural extension  $f(q) = 0$  outside the initial smaller domain  $W$ . Indeed, first of all it plays a crucial rôle in proving that the support of  $h$  in  $W$  coincides with the support of  $h$  in  $M$ . This is not a trivial result. Using the non-Hausdorff, second-countable, locally homeomorphic to  $\mathbb{R}$ , topological space  $M = \mathbb{R} \cup \{p_0\}$  defined in remark (3) after definition 2.2, one simply finds a counterexample. Define the hat function  $h$ , as said above, first in a neighborhood  $W$  of  $0 \in \mathbb{R}$  such that  $W$  is completely contained in the real axis and  $h$  has support compact in  $W$ . Then extend it on the whole  $M$  by stating that  $h$  vanishes outside  $W$ . The support of the extended function  $h$  in  $M$  differs from the support of  $h$  referred to the topology of  $W$ : Indeed the point  $p_0$  belongs to the former support but it does not belong to the latter. As an immediate consequence the extended function  $h$  is not continuous (and not differentiable) in  $M$  because it is not continuous in  $p_0$ . To see it, take the sequence of the reals  $1/n \in \mathbb{R}$  with  $n = 1, 2, \dots$ . That sequence converges both to 0 and  $p_0$  and trivially  $\lim_{n \rightarrow +\infty} h(1/n) = h(0) = 1 \neq h(p_0) = 0$ .

### 2.3.1 Paracompactness

Let us make contact with a very useful tool of differential geometry: the notion of paracompactness. Some preliminary definitions are necessary.

If  $(X, \mathcal{T})$  is a topological space and  $\mathcal{C} = \{U_i\}_{i \in I} \subset \mathcal{T}$  is a covering of  $X$ , the covering  $\mathcal{C}' = \{V_j\}_{j \in J} \subset \mathcal{T}$  is said to be a **refinement** of  $\mathcal{C}$  if every  $j \in J$  admits some  $i(j) \in I$  with  $V_j \subset U_{i(j)}$ .

A covering  $\{U_i\}_{i \in I}$  of  $X$  is said to be **locally finite** if each  $x \in X$  admits an open neighborhood  $G_x$  such that the subset  $I_x \subset I$  of the indices  $k \in I_x$  with  $G_x \cap U_k \neq \emptyset$  is finite.

**Definition 2.22.** (**Paracompactness.**) A topological space  $(X, \mathcal{T})$  is said to be **paracompact** if every covering of  $X$  made of open sets admits a locally finite refinement.  $\blacksquare$

It is possible to prove [KoNo96] that if a topological space  $X$  is

- (1) *second-countable*,
- (2) *Hausdorff*,
- (3) **locally compact**, i.e., for every  $x \in X$  there is  $U_p \ni p$  open such that  $\overline{U_p}$  is compact,

then it is paracompact.

As a consequence *every topological (or  $C^k$ -differentiable) manifold is paracompact* because it is

Hausdorff, second countable and locally homeomorphic to  $\mathbb{R}^n$  which, in turn, is locally compact.

**Remark 2.23.** It is possible to show (see [KoNo96]) that, if  $X$  is a paracompact topological space which is also Hausdorff and locally homeomorphic to  $\mathbb{R}^n$ , then  $X$  is second countable (i.e., it is a topological manifold) if and only if its connected components are countable.

Looking at our definition of topological manifold we see that a topological manifold could be equivalently defined as a topological space which is

- (1) paracompact,
- (2) Hausdorff,
- (3) locally homeomorphic to  $\mathbb{R}^n$ ,

if we exclude the case of a topological space whose connected components are uncountably many<sup>1</sup>. Some textbooks adopt this alternative definition. ■

### 2.3.2 Existence of a smooth partition of unity

The paracompactness of a smooth manifold has an important consequence, namely the existence of a smooth partition of unity.

**Definition 2.24.** (**Partition of Unity.**) Given a locally finite covering of a smooth manifold  $M$ ,  $\mathcal{C} = \{U_i\}_{i \in I}$ , where every  $U_i$  is open, a **partition of unity** subordinate to  $\mathcal{C}$  is a collection of functions  $\{f_i\}_{i \in I} \subset D(M)$  such that:

- (1)  $\text{supp} f_i \subset U_i$  for every  $i \in I$ ,
- (2)  $0 \leq f_i(x) \leq 1$  for every  $i \in I$  and every  $x \in M$ ,
- (3)  $\sum_{i \in I} f_i(x) = 1$  for every  $x \in M$ . ■

**Remark 2.25.**

(1) Notice that, for every  $x \in M$ , the sum in property (3) above is finite because of the local finiteness of the covering.

(2) It is worth stressing that there is no analogue for a partition of unity in the case of an *analytic* manifold  $M$ . This is because if  $f_i : M \rightarrow \mathbb{R}$  is *analytic* and  $\text{supp} f_i \subset U_i$  where  $U_i$  is sufficiently small (such that, more precisely,  $U_i$  is not a connected component of  $M$  and  $M \setminus U_i$  contains a nonempty open set),  $f_i$  must vanish everywhere in  $M$ . ■

Using sufficiently small coordinate neighborhoods it is possible to get a covering of a smooth manifold made of open sets whose closures are compact. Using paracompactness one finds a

---

<sup>1</sup>Case which is not very interesting for applications.

subsequent locally finite covering which made of open sets whose closures are compact.

**Theorem 2.26.** (**Existence of a smooth partition of unity.**) *Let  $M$  be a smooth manifold and  $C = \{U_i\}_{i \in I}$  a locally finite covering made of open sets such that  $\overline{U_i}$  is compact. There is a partition of unity subordinate to  $C$ .*

**Proof.** See [KoNo96]. □

**Remark 2.27.** Observe that we can always assume that  $C$  above is countable, because starting from a covering of open sets we can always extract a countable subcovering due to Lindelöf lemma since a manifold is second countable. This subcovering remains locally finite if the original one was locally finite and it is made of relatively compact sets if the original covering was made of such sets. ■

Paracompactness has other important technical implications, one is stated in the following result which is an immediate consequence of a result by Stone<sup>2</sup> re-adapted to Hausdorff spaces.

**Theorem 2.28.** *A Hausdorff topological space  $X$  is paracompact if and only if every covering  $\mathcal{C}$  of  $X$  made of open sets admits a **\*-refinement** of open sets. That is another covering  $\mathcal{C}^*$  of open sets such that, for every  $V \in \mathcal{C}$ ,*

$$\bigcup \{V' \in \mathcal{C}^* \mid V' \cap V \neq \emptyset\} \subset U_V$$

for some  $U_V \in \mathcal{C}$ .

---

<sup>2</sup>A. H. Stone, *Paracompactness and product spaces*, Bull. Amer. Math. Soc. vol. 54 (1948) pp. 977-982 and see also the review part in E. Michael *Yet Another Note on Paracompact Spaces*, Proceedings of the American Mathematical Society, Apr., 1959, Vol. 10, No. 2 (Apr., 1959), pp. 309-314.



## Chapter 3

# Tensor Fields on Manifolds and Associated Geometric Structures

This chapter is devoted to introduce and discuss the notion of smooth vector field on a smooth manifold and all related concepts. The chapter concludes with the introduction of the notion of fiber bundle as direct generalization of the tangent and cotangent space manifolds.

### 3.1 Tangent and cotangent space at a point

In this section we construct the notion of tangent and cotangent space at a point of a manifold. There are several equivalent approaches in the literature (at least three) we adopt here a quite abstract, though very powerful, approach where tangent vectors are viewed as differential operators [doC92, KoNo96, Seri90].

#### 3.1.1 Tangent vectors as derivations

To interpret the vectors tangent at a point of a smooth manifold in terms of differential operators we need two preliminary definitions.

**Definition 3.1.** (**Derivations**) Let  $M$  be a smooth manifold. A **derivation** in  $p \in M$  is a  $\mathbb{R}$ -linear map  $D_p : D(M) \rightarrow \mathbb{R}$ , such that, for each pair  $f, g \in D(M)$ :

$$D_p fg = f(p)D_p g + g(p)D_p f .$$

The  $\mathbb{R}$ -vector space of the derivations at  $p$  is indicated by  $\mathcal{D}_p M$ . ■

**Remark 3.2.** The above said linear structure in  $\mathcal{D}_p M$  is evidently defined as

$$(aD_p + bD_p')f := aD_p f + bD_p' f \quad \forall a, b \in \mathbb{R}, \forall f \in D(M)$$

if  $D_p, D'_p \in \mathcal{D}_p M$ , observing that  $D''_p := aD_p + bD'_p$  still satisfies Definition 3.1. ■

Derivations exist and, in fact, some of them can be built up as follows. Consider a local coordinate system around  $p$ ,  $(U, \phi)$ , with coordinates  $(x^1, \dots, x^n)$ . If  $f \in D(M)$  is arbitrary, operators

$$\frac{\partial}{\partial x^k} \Big|_p : f \mapsto \frac{\partial f \circ \phi^{-1}}{\partial x^k} \Big|_{\phi(p)}, \quad (3.1)$$

are derivations. The subspace of  $\mathcal{D}_p M$  spanned by those derivations has the same dimension as  $M$  and it is actually independent of the choice of the local chart around  $p$ .

**Proposition 3.3.** *Let  $M$  be a smooth manifold of dimension  $n$  and  $p \in M$ . The following facts are true.*

(a) *The subspace  $T_p M$  of  $\mathcal{D}_p M$  spanned by the  $n$  derivations*

$$\frac{\partial}{\partial x^k} \Big|_p,$$

*constructed out of a local chart  $\phi : U \ni q \mapsto (x^1(q), \dots, x^n(q)) \in \mathbb{R}^n$  around  $p$  as in (3.1), is independent of the choice of  $(U, \phi)$ .*

(b)  *$T_p M$  has dimension  $n$  since the  $n$  derivations  $\frac{\partial}{\partial x^k} \Big|_p$  are a basis of that subspace.*

(c) *If choosing another local chart around  $p$ ,  $\psi : V \ni q \mapsto (y^1(q), \dots, y^n(q)) \in \mathbb{R}^n$ , the identities hold*

$$\frac{\partial}{\partial y^k} \Big|_p = \frac{\partial x^r}{\partial y^k} \Big|_{\psi(p)} \frac{\partial}{\partial x^r} \Big|_p \quad k = 1, 2, \dots, n. \quad (3.2)$$

*(we adopt above the convention of summation over the repeated index  $r$ ) where the Jacobian matrix of elements  $\frac{\partial x^r}{\partial y^k} \Big|_{\psi(p)}$  is non-singular.*

**Proof.** Directly from (3.1), we have that

$$\frac{\partial}{\partial y^k} \Big|_p = \frac{\partial x^r}{\partial y^k} \Big|_{\psi(p)} \frac{\partial}{\partial x^r} \Big|_p.$$

The matrix  $J$  of coefficients  $\frac{\partial x^r}{\partial y^k} \Big|_{\psi(p)}$  is not singular because we can compose the maps  $\phi$  and  $\psi$  as follows

$$id_{\phi(U \cap V)} = \phi \circ \psi^{-1} \circ \psi \circ \phi^{-1} : \phi(U \cap V) \rightarrow \phi(U \cap V)$$

which in coordinates means

$$x^r = x^r(y^1(x^1, \dots, x^n), \dots, y^n(x^1, \dots, x^n))$$

so that, using the the rule of derivation of a composed function,

$$\delta_s^r = \frac{\partial x^r}{\partial x^s} = \frac{\partial x^r}{\partial y^k} \frac{\partial y^k}{\partial x^s},$$

where everything is computed on the image of  $p$  in the respective chart. This means in particular that the  $n \times n$  Jacobian matrix of elements  $\frac{\partial x^r}{\partial y^k}|_{\psi(p)}$  is invertible and thus it is not singular. As a consequence, the subspaces spanned by the all the derivations  $\frac{\partial}{\partial y^k}|_p$ ,  $k = 1, \dots, n$ , and the one spanned by all derivations  $\frac{\partial}{\partial x^k}|_p$ ,  $k = 1, \dots, n$ , coincide. To conclude the proof of all statements it is now sufficient to prove that the derivations  $\frac{\partial x^r}{\partial y^k}|_{\psi(p)}$ , for  $k = 1, \dots, n$  are linearly independent. Let us prove it. It is sufficient to use  $n$  functions  $f^{(j)} \in D(M)$ ,  $j = 1, \dots, n$ , such that  $f^{(j)} \circ \phi(q) = x^j(q)$  when  $q$  belongs to an open neighborhood of  $p$  contained in  $U$ . This implies the linear independence of the considered derivations. In fact, if:

$$c^k \frac{\partial}{\partial x^k}|_p = 0,$$

then

$$c^k \frac{\partial f^{(j)}}{\partial x^k}|_p = 0,$$

which is equivalent to  $c^k \delta_k^j = 0$  or :

$$c^j = 0 \quad \text{for all } j = 1, \dots, n.$$

The existence of the functions  $f^{(j)}$  can be straightforwardly proved by using Lemma 2.20. The mapping  $f^{(j)} : M \rightarrow \mathbb{R}$  defined as

$$f^{(j)}(q) = h(q)\phi^j(q) \text{ if } q \in U, \text{ where } \phi^j : q \mapsto x^j(q) \text{ for all } q \in U,$$

$$f^{(j)}(q) = 0 \text{ if } q \in M \setminus U,$$

turns out to be  $C^\infty$  on the whole manifold  $M$  and satisfies  $(f^{(j)} \circ \phi)(q) = x^j(q)$  in a neighborhood of  $p$  provided  $h$  is any hat function centered in  $p$  with support completely contained in  $U$ .  $\square$

**Remark 3.4.** All the above construction leading to Definition 3.5 below is valid also if the manifold has differentiability order  $k \geq 1$  but  $k \neq \infty$ .  $\blacksquare$

We are now in a position to state the fundamental definition.

**Definition 3.5. (Tangent space at a point.)** Let  $M$  be a smooth manifold of dimension  $n$  and  $p \in M$ . The  $n$  dimensional subspace  $T_p M$  of  $\mathcal{D}_p M$  spanned by the  $n$  derivations

$$\frac{\partial}{\partial x^k}|_p,$$

constructed out of a local chart  $\phi : U \ni q \mapsto (x^1(q), \dots, x^n(q)) \in \mathbb{R}^n$  around  $p$  as in (3.1) and independent of the choice of  $(U, \phi)$ , is called **tangent space** at  $p$ .  $\blacksquare$

**Remark 3.6.** With the given definition, it arises that any  $n$ -dimensional **Affine space**  $\mathbb{A}^n$  admits two different notions of vector applied to a point  $p$ . Indeed there are the vectors in the space of translations  $V$  used in the definition of  $\mathbb{A}^n$  itself. These vectors are also called **free**

**vectors.** On the other hand, considering  $\mathbb{A}^n$  as a smooth manifold as said in (2) Examples 2.11, one can think of those vectors as **applied vectors** to every point  $p$  of  $\mathbb{A}^n$ . What is the relation between these vectors and those of  $T_p\mathbb{A}^n$ ? Take a basis  $\{e_i\}_{i \in I}$  in the vector space  $V$  and a origin  $O \in \mathbb{A}^n$ , then define a Cartesian coordinate system centered on  $O$  associated with the given basis, that is the global coordinate system:

$$\phi : \mathbb{A}^n \rightarrow \mathbb{R}^n : p \mapsto (\langle \overrightarrow{Op}, e^{*1} \rangle, \dots, \langle \overrightarrow{Op}, e^{*n} \rangle) =: (x^1, \dots, x^n).$$

Now also consider the bases  $\frac{\partial}{\partial x^i}|_p$  of each  $T_p\mathbb{A}^n$  induced by these Cartesian coordinates. It results that there is a natural isomorphism  $\chi_p : T_p\mathbb{A}^n \rightarrow V$  which identifies each  $\frac{\partial}{\partial x^i}|_p$  with the corresponding  $e_i$

$$\chi_p : v^i \frac{\partial}{\partial x^i}|_p \mapsto v^i e_i.$$

Indeed the map defined above is linear, injective and surjective by construction. Moreover using different Cartesian coordinates  $y^1, \dots, y^n$  associated with a basis  $f_1, \dots, f_n$  in  $V$  and a new origin  $O' \in \mathbb{A}^n$ , one has

$$y^i = A^i{}_j x^j + C^i$$

where

$$e_k = A^j{}_k f_j \quad \text{and} \quad C^i = \langle \overrightarrow{O'O}, f^{*i} \rangle.$$

Thus, it is immediately proved by direct inspection that, if  $\chi'_p$  is the isomorphism

$$\chi'_p : u^i \frac{\partial}{\partial y^i}|_p \mapsto u^i f_i,$$

it holds  $\chi_p = \chi'_p$ . Indeed

$$\chi_p : v^i \frac{\partial}{\partial x^i}|_p \mapsto v^i e_i$$

can be re-written, if  $[B_i{}^k]$  is the inverse transposed matrix of  $[A^p{}_q]$

$$A^i{}_j u^j B_i{}^k \frac{\partial}{\partial y^k}|_p \mapsto A^i{}_j u^j B_i{}^k f_k.$$

But  $A^i{}_j B_i{}^k = \delta_j^k$  and thus

$$\chi_p : v^i \frac{\partial}{\partial x^i}|_p \mapsto v^i e_i$$

can equivalently be re-written

$$u^j \frac{\partial}{\partial y^j}|_p \mapsto u^j f_j,$$

that is  $\chi_p = \chi'_p$ . In other words the isomorphism  $\chi : T_p\mathbb{A}^n \rightarrow V$  does not depend on the considered Cartesian coordinate frame: it is a natural isomorphism. ■

We have a final important definition.

**Definition 3.7.** Let  $M$  be a smooth manifold,  $I \subset \mathbb{R}$  an open interval and  $\gamma : I \rightarrow M$  a smooth curve. The **tangent vector** to  $\gamma$  for  $t = t_0 \in I$  is

$$\gamma'(t_0) := \frac{dx_\gamma^i}{dt} \Big|_{t=t_0} \frac{\partial}{\partial x^i} \Big|_{\gamma(t_0)},$$

where  $\phi : U \ni p \mapsto (x^1(p), \dots, x^n(p))$  is a local chart defined in  $U \ni \gamma(t_0)$ , and  $x_\gamma^k(t) := \phi(\gamma(t))$  for  $t \in (\phi \circ \gamma)^{-1}(U)$ . ■

**Remark 3.8.** It is immediately proved that the definition above is actually independent of the chosen local chart  $(U, \phi)$  as a straightforward consequence of (3.2). ■

### 3.1.2 More on the linear space of derivations

Let us focus again on the full space of derivations to study its connection with the subspace  $T_p M$ . We have the following general properties, in particular items (2) and (3) below.

**Lemma 3.9.** Let  $M$  be a differential manifold. Take any  $p \in M$  and any  $D_p \in \mathcal{D}_p M$ .

(1) If  $h \in D(M)$  vanishes in a open neighborhood of  $p$ , then

$$D_p h = 0.$$

(2) For every  $f, g \in D(M)$ ,

$$D_p f = D_p g,$$

provided  $f(q) = g(q)$  in an open neighborhood of  $p$ .

(3) If  $h \in D(M)$  is constant in a neighborhood of  $p$ ,

$$D_p h = 0.$$

**Proof.** By linearity, (1) entails (2). Let us prove (1). Let  $h \in D(M)$  a function which vanishes in a small open neighborhood  $U$  of  $p$ . Shrinking  $U$  if necessary, by Lemma 1.2 we can find another neighborhood  $V$  of  $p$ , with  $\overline{V} \subset U$ , and a function  $g \in D(M)$  which vanishes outside  $U$  taking the constant value 1 in  $\overline{V}$ . As a consequence  $g' := 1 - g$  is a function in  $D(M)$  which vanishes in  $\overline{V}$  and take the constant value 1 outside  $U$ . If  $q \in U$  one has  $g'(q)h(q) = g'(q) \cdot 0 = 0 = h(q)$ , if  $q \notin U$  one has  $g'(q)h(q) = 1 \cdot h(q) = h(q)$  hence  $h(q) = g'(q)h(q)$  for every  $q \in M$ . As a consequence

$$D_p h = D_p g' h = g'(p) D_p h + h(p) D_p g' = 0 \cdot D_p h + 0 \cdot D_p g' = 0.$$

The proof of (3) is straightforward. It is sufficient to show that the thesis holds true if  $h$  is constant everywhere in  $M$ , then (2) implies the thesis in the general case. If  $h$  is constant, let  $g \in D(M)$  a hat function with  $g(p) = 1$ . By linearity, since  $h$  is constant

$$D_p(hg) = hD_pg.$$

On the other hand, since  $D_p$  is a derivation,

$$D_p(hg) = hD_pg + g(p)D_ph.$$

Comparing with the identity above, one gets

$$g(p)D_ph = 0.$$

Since  $g(p) = 1$ , one has  $D_ph = 0$ . □

Our final goal to describe the space of derivations is the proof of  $T_pM = \mathcal{D}_p(M)$ . This remarkable result will be established under the explicit hypothesis that the manifold is smooth ( $C^\infty$ ) as we always assume.

We remind the reader that an open set  $U \subset \mathbb{R}^n$  is said to be a open **starshaped neighborhood** of  $p \in \mathbb{R}^n$  if  $U$  is a open neighborhood of  $p$  and the closed  $\mathbb{R}^n$  segment  $\overline{pq}$  is completely contained in  $U$  whenever  $q \in U$ . Every open ball centered on a point  $p$  is an open starshaped neighborhood of  $p$ . Therefore these sets form a local basis of the Euclidean topology around every point  $p \in \mathbb{R}^n$ .

**Lemma 3.10.** (Flander's lemma.) *If  $f : B \rightarrow \mathbb{R}$  is  $C^\infty(B)$  where  $B \subset \mathbb{R}^n$  is an open starshaped neighborhood of  $p_0 = (x_0^1, \dots, x_0^n)$ , there are  $n$  smooth mappings  $g_i : B \rightarrow \mathbb{R}$  such that, if  $p = (x^1, \dots, x^n)$ ,*

$$f(p) = f(p_0) + \sum_{i=1}^n g_i(p)(x^i - x_0^i),$$

with

$$g_i(p_0) = \frac{\partial f}{\partial x^i} \Big|_{p_0}$$

for all  $i = 1, \dots, n$ .

**Proof.** let  $p = (x^1, \dots, x^n)$  belong to  $B$ . The points of the segment  $\overline{p_0p}$  are given by

$$y^i(t) = x_0^i + t(x^i - x_0^i)$$

for  $t \in [0, 1]$ . As a consequence, the following equation holds

$$f(p) = f(p_0) + \int_0^1 \frac{d}{dt} f(p_0 + t(p - p_0)) dt = f(p_0) + \sum_{i=1}^n \left( \int_0^1 \frac{\partial f}{\partial x^i} \Big|_{p_0 + t(p - p_0)} dt \right) (x^i - x_0^i).$$

If

$$g_i(p) := \int_0^1 \frac{\partial f}{\partial x^i} \big|_{p_0 + t(p-p_0)} dt ,$$

so that

$$g_i(p_0) = \int_0^1 \frac{\partial f}{\partial x^i} \big|_{p_0} dt = \frac{\partial f}{\partial x^i} \big|_{p_0} ,$$

the equation above can be re-written:

$$f(x) = f(p_0) + \sum_{i=1}^n g_i(p)(x^i - x_0^i) .$$

By construction the functions  $g_i$  are  $C^\infty(B)$  as a direct consequence of theorems concerning derivation under the symbol of integration (based on Lebesgue's dominated convergence theorem).  $\square$

**Theorem 3.11.** *Let  $M$  be a smooth manifold and  $p \in M$ . It holds*

$$\mathcal{D}_p(M) = T_p M$$

*since every derivation at  $p$  can uniquely be decomposed as a linear combination of derivations  $\{\frac{\partial}{\partial x^k} \big|_p\}_{k=1,\dots,n}$  for each local chart  $\phi : U \ni q \mapsto (x^1(q), \dots, x^n(q))$  with  $U \ni p$ .*

**Proof.** Let us prove that, if  $D_p \in \mathcal{D}_p M$  and considering the local chart  $(U, \phi)$  with coordinates  $(x^1, \dots, x^n)$  around  $p$ , then there are  $n$  reals  $c^1, \dots, c^n$  such that

$$D_p f = \sum_{k=1}^n c^k \frac{\partial f \circ \phi^{-1}}{\partial x^k} \big|_{\phi(p)} ,$$

for all  $f \in D(M)$ . As these reals do not depend on  $f$ , we can write

$$D_p = \sum_{k=1}^n c^k \frac{\partial}{\partial x^k} \big|_p$$

To prove it, we start from the expansion due to Lemma 3.9 and valid in a neighborhood  $U_p \subset U$  of  $\phi(p)$  (which is the image according to  $\phi^{-1}$  of a starshaped neighborhood of  $\phi(p)$  in  $\mathbb{R}^n$ ):

$$(f \circ \phi^{-1})(\phi(q)) = (f \circ \phi^{-1})(\phi(p)) + \sum_{i=1}^n g_i(\phi(q))(x^i - x_p^i) ,$$

where  $\phi(q) = (x^1, \dots, x^n)$  and  $\phi(p) = (x_p^1, \dots, x_p^n)$  and

$$g_i(\phi(p)) = \frac{\partial (f \circ \phi^{-1})}{\partial x^i} \big|_{\phi(p)} .$$

If  $h_1, h_2$  are hat functions centered on  $p$  (see Lemma 2.20) with supports contained in  $U_p$  define  $h := h_1 \cdot h_2$  and  $f' := h \cdot f$ . The multiplication of  $h$  and the right-hand side of the local expansion for  $f$  written above gives rise to an expansion valid on the whole manifold:

$$f'(q) = f(p)h(q) + \sum_{i=1}^n g'_i(q)r_i(q)$$

where the functions  $g'_i, r_i \in D(M)$  and

$$r_i(q) = h_2(q) \cdot (x^i - x_p^i) = (x^i - x_p^i) \quad \text{in a neighborhood of } p$$

while

$$g'_i(p) = h_1(p) \cdot \frac{\partial(f \circ \phi^{-1})}{\partial x^i} \Big|_{\phi(p)} = \frac{\partial(f \circ \phi^{-1})}{\partial x^i} \Big|_{\phi(p)}.$$

Moreover, Lemma 3.9 assures that  $D_p f' = D_p f$  since  $f = f'$  in a neighborhood of  $p$ . As a consequence

$$D_p f = D_p f' = D_p \left( f(p)h(q) + \sum_{i=1}^n g'_i(q)r_i(q) \right).$$

Since  $q \mapsto f(p)h(q)$  is constant in a neighborhood of  $p$ ,  $D_p f(p)h(q) = 0$  by Lemma 3.9. Moreover

$$D_p \left( \sum_{i=1}^n g'_i(q)r_i(q) \right) = \sum_{i=1}^n (g'_i(p)D_p r_i + r_i(p)D_p g'_i),$$

where  $r_i(p) = (x_p^i - x_p^i) = 0$ . Finally we have found

$$D_p f = \sum_{i=1}^n c^i g'_i(p) = \sum_{i=1}^n c^i \frac{\partial f \circ \phi^{-1}}{\partial x^i} \Big|_{\phi(p)},$$

where the coefficients

$$c^i = D_p r_i$$

do not depend on  $f$  by construction. This is the thesis and the proof ends.  $\square$

### 3.1.3 Cotangent space

As  $T_p M$  is a vector space, one can define its dual space. This space plays an important role in differential geometry.

**Definition 3.12.** (**Cotangent space at a point.**) Let  $M$  be a  $n$ -dimensional manifold. For each  $p \in M$ , the dual space  $T_p^* M$  is called the **cotangent space** at  $p$  and its elements are called **1-forms** in  $p$  or, equivalently, **covectors** in  $p$ . If  $(x^1, \dots, x^n)$  are coordinates around  $p$  inducing



the basis  $\{\frac{\partial}{\partial x^k}|_p\}_{k=1,\dots,n}$ , the associated dual basis in  $T_p^*M$  is denoted by  $\{dx^k|_p\}_{k=1,\dots,n}$ . ■

**Exercises 3.13.** Show that, changing local coordinates,

$$dx'^k|_p = \frac{\partial x'^k}{\partial x^i}|_p dx^i|_p,$$

and if  $\omega_p = \omega_{pi} dx^i|_p = \omega'_{pr} dx'^r|_p$ , then

$$\omega'_{pr} = \frac{\partial x^i}{\partial x'^r}|_p \omega_{pi}.$$

## 3.2 Vector and tensor fields

The introduced definitions allows one to introduce the tensor algebra  $\mathcal{A}_{\mathbb{R}}(T_p M)$  of the tensor spaces obtained by tensor products of spaces  $\mathbb{R}$ ,  $T_p M$  and  $T_p^* M$  [Mor20]. Using tensors defined on each point  $p \in M$  one may define *tensor fields*.

### 3.2.1 Tensor fields

**Definition 3.14. (Differentiable Tensor Fields.)** Let  $M$  be a  $n$ -dimensional smooth manifold.

(a) A **smooth tensor field**  $t$  is an assignment

$$M \ni p \mapsto t_p \in \mathcal{A}_{\mathbb{R}}(T_p M)$$

where

- (i)  $t_p$  are of the same kind independently of  $p \in M$ ;
  - (ii) varying  $p \in M$ , the components of  $t_p$  with respect to the canonical bases of  $\mathcal{A}_{\mathbb{R}}(T_p M)$  given by tensor products of bases  $\{\frac{\partial}{\partial x^k}|_p\}_{k=1,\dots,n} \subset T_p M$  and  $\{dx^k|_p\}_{k=1,\dots,n} \subset T_p^* M$  are smooth in every local chart  $U \ni p \mapsto (x^1(p), \dots, x^n(p)) \in \mathbb{R}^n$  on  $M$ .
- (b) A smooth **vector field** and a smooth **1-form** (equivalently called **covector field**) are assignments of tangent vectors and 1-forms respectively as stated in (a).

The vector space of smooth (i.e., smooth) vector fields on  $M$  with linear structure

$$(aX + bY)_p := aX_p + bY_p \quad \text{for all } a, b \in \mathbb{R} \text{ and all vector fields } X, Y$$

is denoted by  $\mathfrak{X}(M)$ . ■

For tensor fields the same terminology referred to tensors is used. For instance, a tensor field  $t$  which is represented in local coordinates by  $t^i_j(p) \frac{\partial}{\partial x^i}|_p \otimes dx^j|_p$  is said to be of order  $(1, 1)$ .

The tensor fields of a given type form a vector space with respect to the point-by-point linear combination.

**Remark 3.15.**

(1) It is clear that to assign on a smooth manifold  $M$  a smooth tensor field  $T$  (of any kind and order) it is *necessary and sufficient* to assign a set of smooth functions

$$(x^1, \dots, x^n) \mapsto T^{i_1 \dots i_m}_{j_1 \dots j_k}(x^1, \dots, x^n)$$

in every local coordinate patch (of the whole differentiable structure of  $M$  or, more simply, of an atlas of  $M$ ) such that they satisfy the usual rule of transformation of components of tensors: if  $(x^1, \dots, x^n)$  and  $(y^1, \dots, y^n)$  are the coordinates of the same point  $p \in M$  in two different local charts,

$$T^{i_1 \dots i_m}_{j_1 \dots j_k}(x^1, \dots, x^n) = \frac{\partial x^{i_1}}{\partial y^{k_1}}|_p \cdots \frac{\partial x^{i_m}}{\partial y^{k_m}}|_p \frac{\partial y^{l_1}}{\partial x^{j_1}}|_p \cdots \frac{\partial y^{l_m}}{\partial x^{j_m}}|_p T'^{k_1 \dots k_m}_{l_1 \dots l_k}(y^1, \dots, y^n).$$

Then, in local coordinates,

$$T(p) = T^{i_1 \dots i_m}_{j_1 \dots j_k} \frac{\partial}{\partial x^{i_1}}|_p \otimes \cdots \otimes \frac{\partial}{\partial x^{i_m}}|_p \otimes dx^{j_1}|_p \otimes \cdots \otimes dx^{j_k}|_p.$$

(2) It is obvious that the differentiability requirement of the components of a tensor field can be checked using the bases induced by a single atlas of local charts. It is not necessary to consider all the charts of the differentiable structure of the manifold.

(3) Every  $X \in \mathfrak{X}(M)$  defines a derivation at each point  $p \in M$ . Indeed, if  $f \in D(M)$ ,

$$X_p(f) := X^i(p) \frac{\partial f}{\partial x^i}|_p,$$

where  $x^1, \dots, x^n$  are coordinates defined around  $p$ . More generally, every  $X \in \mathfrak{X}(M)$  defines a linear mapping from  $D(M)$  to  $D(M)$  given by

$$f \mapsto X(f) \quad \text{for every } f \in D(M),$$

where  $X(f) \in D(M)$  is defined as

$$X(f)(p) := X_p(f) \quad \text{for every } p \in M.$$

(4) For (contravariant) vector fields  $X$  on a smooth manifold  $M$ , a requirement equivalent to the differentiability is the following: the function  $X(f) : p \mapsto X_p(f)$  (where we used  $X_p$  as a derivation) is smooth for all of  $f \in D(M)$ . Indeed, if  $X$  is a smooth contravariant vector field and if  $f \in D(M)$ , one has that  $X(f) : M \ni p \mapsto X_p(f)$  is a smooth function too as it having a coordinate representation

$$X(f) \circ \phi^{-1} : \phi(U) \ni (x^1, \dots, x^n) \mapsto X^i(x^1, \dots, x^n) \frac{\partial f}{\partial x^i}|_{(x^1, \dots, x^n)}$$

in every local coordinate chart  $(U, \phi)$  and all the involved function being smooth. Conversely, if  $p \mapsto X_p(f)$  defines a function in  $D(M)$ ,  $X(f)$ , for every  $f \in D(M)$ , the components of  $p \mapsto X_p$  in every local chart  $(U, \phi)$  must be smooth. This is because, in a neighborhood of  $q \in U$

$$X^i(q) := X_q(f^{(i)})$$

where the function  $f^{(i)} \in D(M)$  vanishes outside  $U$  and is defined as  $r \mapsto x^i(r) \cdot h(r)$  in  $U$ , where  $x^i$  is the  $i$ -th component of  $\phi$  (the coordinate  $x^i$ ) and  $h$  a hat function centered on  $q$  with support in  $U$ .

Similarly, the differentiability of a covariant vector field  $\omega$  is equivalent to the differentiability of each function  $p \mapsto \langle X_p, \omega_p \rangle$ , for all smooth vector fields  $X$ .

(5) If  $f \in D(M)$ , the **differential** of  $f$  at  $p$ ,  $df_p$  is the 1-form defined by

$$df_p = \frac{\partial f}{\partial x^i} \Big|_p dx^i \Big|_p, \quad (3.3)$$

in local coordinates around  $p$ . The definition does not depend on the chosen coordinates. As a consequence of remark (1) above, varying the point  $p \in M$ ,  $p \mapsto df_p$  defines a covariant smooth vector field denoted by  $df$  and called the **differential** of  $f$ .

Notice that

$$X_p(f) = \langle X_p, df_p \rangle, \quad (3.4)$$

for every smooth vector field  $X$  and  $f \in D(M)$  at each point  $p \in M$ .

(6) As we know,  $\mathfrak{X}(M)$  is vector space with field given by  $\mathbb{R}$ . Notice that if  $\mathbb{R}$  is replaced by  $D(M)$ , the obtained algebraic structure is not a vector space because  $D(M)$  is a commutative ring with multiplicative and additive unit elements but fails to be a field as remarked above. However, the incoming algebraic structure given by a "vector space with the field replaced by a commutative ring with multiplicative and additive unit elements" is well known and it is called **module**.

The following lemma is trivial but useful in applications.

**Lemma 3.16.** *Let  $p$  be a point in a smooth manifold  $M$ . If  $t$  is any tensor in  $\mathcal{A}_{\mathbb{R}}(T_p M)$ , there is a differentiable tensor field in  $M$ ,  $\Xi$  such that  $\Xi_p = t$ .*

**Proof.** Consider a local coordinate frame  $(U, \phi)$  defined in an open neighborhood  $U$  of  $p$ . In  $U$  a tensor field  $\Xi'$  which have constant components with respect the bases associated with the considered coordinates. We can fix these components such that  $\Xi'_p = t$ . One can find (see remark 2 after definition 3.12) a smooth function  $h : \phi(U) \rightarrow \mathbb{R}$  such that  $h(\phi(p)) = 1$  and  $h$  vanishes outside a small neighborhood of  $\phi(p)$  whose closure is completely contained in  $\phi(U)$ .  $\Xi$  defined as  $(h \circ \phi)(r) \cdot \Xi'(r)$  if  $r \in U$  and  $\Xi(r) = 0$  outside  $U$  is a smooth tensor fields on  $M$  such that  $\Xi_p = t$ .  $\square$

**Remark 3.17.** *From now tensor (vector, covector, etc.) field means smooth tensor (vector, covector, etc.) field.* ■

### 3.2.2 Lie brackets of vector fields

Contravariant smooth vector fields can be seen as differential operators (derivations at each point of the manifold) acting on smooth scalar fields. It is possible to obtain such an operator by an appropriate composition of two vector fields. To this end, consider the application  $[X, Y] : D(M) \rightarrow D(M)$ , where  $X, Y \in \mathfrak{X}(M)$ , defined as follows

$$[X, Y](f) := X(Y(f)) - Y(X(f)) ,$$

for  $f \in D(M)$ . It is clear that  $[X, Y]$  is linear. Actually it turns out to be a derivation too. Indeed, a direct computation shows that it holds

$$X(Y(fg)) = fX(Y(g)) + gY(X(f)) + (X(f))(Y(g)) + (X(g))(Y(f))$$

and

$$Y(X(fg)) = fY(X(g)) + gX(Y(f)) + (Y(f))(X(g)) + (Y(g))(X(f)) ,$$

so that

$$[X, Y](fg) = f[X, Y](g) + g[X, Y](f) .$$

Using proposition 3.11, this fact shows that, for each point  $p$  of  $M$ ,  $[X, Y]_p : f \mapsto ([X, Y](f))(p)$  is a derivation. Hence it is represented by a contravariant vector of  $T_p M$  (denoted by  $[X, Y]_p$ ). On the other hand, varying the point  $p$  one gets  $[X, Y]_p(f)$  is a smooth function if  $f \in D(M)$ . This is because (see remark 4 above), as  $X$  and  $Y$  are smooth vector fields,  $X(f)$  and  $Y(f)$  are in  $D(M)$  if  $f \in D(M)$  and thus  $X(Y(f))$  and  $Y(X(f))$  are in  $D(M)$  too. Thus, using remark 4 above, one gets that, as we said,  $M \ni p \mapsto [X, Y]_p$  is a (smooth) vector field on  $M$ .

**Definition 3.18. (Lie Bracket.)** Let  $X, Y \in \mathfrak{X}(M)$  for the smooth manifold  $M$ . The **Lie bracket** of  $X$  and  $Y$ ,  $[X, Y]$ , is the contravariant smooth vector field associated with the differential operator

$$[X, Y](f) := X(Y(f)) - Y(X(f)) ,$$

for  $f \in D(M)$ . ■

By direct computation, in coordinates one easily sees that

$$[X, Y]_p = \left( X^i(p) \frac{\partial Y^j}{\partial x^i} \Big|_p - Y^i(p) \frac{\partial X^j}{\partial x^i} \Big|_p \right) \frac{\partial}{\partial x^j} \Big|_p , \quad (3.5)$$

where  $\phi : U \ni p \mapsto (x^1(p), \dots, x^n(p)) \in \mathbb{R}^n$  is any fixed local chart.

**Exercises 3.19.** Prove that the Lie brackets define a **Lie algebra** on the vector space  $\mathfrak{X}(M)$  for every smooth manifold  $M$ . In other words

$$[\cdot, \cdot] : \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow \mathfrak{X}(M)$$

enjoys the following properties, where  $X, Y, Z$  are contravariant smooth vector fields,

- (a) **antisymmetry**,  $[X, Y] = -[Y, X]$ ;
- (b)  **$\mathbb{R}$ -linearity**,  $[aX + \beta Y, Z] = b[X, Z] + \beta[Y, Z]$  for all  $a, b \in \mathbb{R}$ ;
- (c) **Jacobi identity**,  $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$  (0 being the null vector field).

### 3.2.3 Fields of $k$ -forms and exterior derivative

**Definition 3.20.** Let  $M$  be a smooth manifold and  $\Lambda^k(T_p^*M)$  is the space of  $k$ -forms on  $T_p^*M$  for  $k = 0, 1, \dots, \dim(M)$  [Mor20]. A smooth assignment

$$\omega : M \ni p \mapsto \omega_p \in \Lambda^k(T_p^*M)$$

is called a **field of  $k$ -form** or simply  **$k$ -form** on  $M$ , where as usual smoothness concerns the components of the said tensor field. The vector space of  $k$ -forms on  $M$  will be denoted by  $\Omega^k(M)$ , for  $k = 0, 1, \dots, \dim(M)$ , where  $\Omega^0(M) := D(M)$ . ■

The local expression of a  $k$ -form can be given in terms of the *exterior product* [Mor20] of the elements of the basis of each  $T_p^*M$  associated with every local chart  $U \ni p \mapsto (x^1, \dots, x^n) \in \mathbb{R}^n$

$$\omega_p = \sum_{1 \leq i_1 < \dots < i_k \leq n} (\omega_p)_{i_1 \dots i_k} dx^{i_1}|_p \wedge \dots \wedge dx^{i_k}|_p. \quad (3.6)$$

A very useful machinery to deal with  $k$ -forms is the so-called **exterior derivative**. That is a linear map  $d_k : \Omega^k(M) \rightarrow \Omega^{k+1}(M)$  where we assume  $\Omega^k(M) := \{0\}$  if  $k > n := \dim(M)$ . In local coordinates, if (3.6) is true,

$$(d_k \omega)_p := \sum_{1 \leq i_1 < \dots < i_k \leq n} \left( \sum_{j=1}^n \frac{\partial \omega_{i_1 \dots i_k}}{\partial x^j} dx^j|_p \right) \wedge dx^{i_1}|_p \wedge \dots \wedge dx^{i_k}|_p. \quad (3.7)$$

**Remark 3.21.** As is usual dealing with the exterior derivation, we use the simplified notation  $d$  in place of  $d_k$ . ■

It is not difficult to prove the following properties of the exterior derivative.

- (1)  $d : \Omega^k(M) \rightarrow \Omega^{k+1}(M)$  is linear;
- (2)  $d\omega = df$  if  $\omega = f \in \Omega^0(M) = D(M)$ , where  $df$  is the differnetial of  $f$ ;
- (3)  $dd\omega = 0$  if  $\omega \in \Omega^k(M)$ ;
- (4)  $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta$  if  $\omega \in \Omega^k(M)$  and  $\eta \in \Omega^h(M)$ ;

A  $k$ -form  $\omega$  is said to be **closed** if  $d\omega = 0$  and it is **exact** if  $\omega = d\eta$  for a  $(k-1)$ -form  $\eta$ . Evidently exact forms are exact but the converse is false. However the celebrated *Poincaré lemma* is valid [KoNo96]. To state that result, we recall the following technical definition.

**Definition 3.22.** Let  $M$  a smooth manifold. An open set  $U \subset M$  is said to be **smoothly contractible** (with respect to a point  $p \in U$ ) if, for some  $p \in U$ , a jointly smooth map exists  $c^{(p)} : (-\epsilon, 1 + \epsilon) \times U \rightarrow U$ , for some  $\epsilon$ , such that

$$(i) \quad c^{(p)}(0, q) = q \text{ for all } q \in U;$$

$$(ii) \quad c^{(p)}(1, q) = p. \quad \blacksquare$$

**Theorem 3.23.** (**Poincaré lemma.**) *Let  $U$  be a open set of a smooth manifold  $M$  that is smoothly contractible (with respect to some  $p \in U$ ) and  $\omega \in \Omega^k(M)$ . If  $d\omega|_U = 0$  – where  $U$  is viewed as a smooth manifold with the structure induced from  $M$  – then  $\omega|_U = d\eta$  for some  $\eta \in \Omega^{k-1}(U)$ .*

For instance, the image  $U := \phi(B)$  of an open ball  $B \subset \mathbb{R}^n$  through a local chart  $\phi$  on  $M$  is contractible. Therefore there is a local basis around every fixed point of a smooth manifold  $M$  which is made of contractible open sets. As a consequence, the Poincaré lemma is always valid in a sufficiently small neighborhood of every  $p \in M$ .

**Remark 3.24.** If  $k = 1$ , the hypotheses on  $U$  can be relaxed to the requirement that it is simply connected as is well known in elementary analysis.  $\blacksquare$

## 3.3 Elementary bundles

### 3.3.1 Tangent and cotangent bundle

If  $M$  is a smooth manifold and with dimension  $n$ , we can consider the set

$$TM := \{(p, v) \mid p \in M, v \in T_p M\}.$$

It is possible to endow  $TM$  with a structure of a smooth manifold with dimension  $2n$ . That structure is naturally induced by the analogous structure of  $M$ .

First of all let us define a suitable second-countable Hausdorff topology on  $TM$ . If  $M$  is a  $n$ -dimensional smooth manifold with differentiable structure  $\mathcal{M}$ , consider the class  $\mathcal{B}$  of all (open) sets  $U \subset M$  such that  $(U, \phi) \in \mathcal{M}$  for some  $\phi : U \rightarrow \mathbb{R}^n$ . It is straightforwardly proved that  $\mathcal{B}$  is a basis of the topology of  $M$ . Then consider the class  $T\mathcal{B}$  of all subsets  $V$  of  $TM$  defined as follows.

$$(a) \quad \text{take } (U, \phi) \in \mathcal{M} \text{ with } \phi : p \mapsto (x^1(p), \dots, x^n(p));$$

$$(b) \quad \text{take an open nonempty set } B \subset \mathbb{R}^n;$$

(b) define

$$V := \{(p, v) \in TM \mid p \in U, v \in \hat{\phi}_p B\},$$

where  $\hat{\phi}_p : \mathbb{R}^n \rightarrow T_p M$  is the linear isomorphism associated to  $\phi$ :

$$\hat{\phi}_p : (v_p^1, \dots, v_p^n) \mapsto v_p^i \frac{\partial}{\partial x^i} \Big|_p \quad (3.8)$$

Let  $\mathcal{T}_{TB}$  us finally denote the class of all sets which are unions of the above sets  $V$  adding also the empty set to the class.

It is easy to prove that  $\mathcal{T}_{TB}$  is a topology and the sets  $V$  (varying the chart, the point  $p$  and the set  $B$ ) is a basis of that topology.  $\mathcal{T}_{TB}$  is also second-countable and Hausdorff:

- (i) the Hausdorff property immediately arises from the analogous property of the topologies on  $M$  and on  $\mathbb{R}^n$ ;
- (ii) second countability is evident if observing that  $\mathcal{T}_{TB}$  admits a basis made of a countable class of sets  $V$  constructing by
  - (a) choosing a countable basis of elements  $U$  (such that  $(U, \phi) \in \mathcal{M}$  for some  $\phi$ ) for the topology of  $M$  (it exists since  $M$  is second countable),
  - (b) choosing the points  $p \in U$  of rational (with sign)  $\phi$ -coordinates ,
  - (c) choosing the sets  $B \subset \mathbb{R}^n$  as open balls of rational radius.

Finally, it turns out that  $TM$ , equipped with the topology  $\mathcal{T}_{TB}$ , is locally homeomorphic to  $\mathbb{R}^n \times \mathbb{R}^n$ . Indeed, if  $(U, \phi)$  is a local chart of  $M$  with  $\phi : U \ni p \mapsto (x^1(p), \dots, x^n(p)) \in \mathbb{R}^n$ , we may define a local chart of  $TM$ ,  $(TU, \Phi)$ , where

$$TU := \{(p, v) \mid p \in U, v \in T_p M\}$$

by defining

$$T\phi : (p, v) \mapsto (x^1(p), \dots, x^n(p), v_p^1, \dots, v_p^n),$$

where  $v = v_p^i \frac{\partial}{\partial x^i} \Big|_p$ . Notice that  $T\phi$  is injective and  $T\phi(TU) = \phi(U) \times \mathbb{R}^n \subset \mathbb{R}^{2n}$ . As a consequence of the definition of the topology  $\mathcal{T}_{TB}$  on  $TM$ , every  $T\phi$  defines a local homeomorphism from  $TM$  to  $\mathbb{R}^{2n}$ . As the union of domains of every  $T\phi$  is  $TM$  itself

$$\bigcup TU = TM,$$

$TM$  is locally homeomorphic to  $\mathbb{R}^{2n}$ .

The next step consists of defining a smooth differentiable structure on  $TM$ . Consider two local charts on  $TM$ ,  $(TU, T\phi)$  and  $(TU', T\phi')$  respectively induced by two local charts  $(U, \phi)$  and  $(U', \phi')$  of the differentiable structure of  $M$ . As a consequence of the given definitions,  $(TU, T\phi)$  and  $(TU', T\phi')$  are trivially compatible. Moreover, the class of charts  $(TU, T\phi)$  induced from all the charts  $(U, \phi)$  of the differentiable structure of  $M$  defines an atlas  $\mathcal{A}(TM)$  on  $TM$  (in

particular because, as said above,  $\bigcup TU = TM$ ). The differentiable structure  $\mathcal{M}_{\mathcal{A}(TM)}$  induced by  $\mathcal{A}(TM)$  makes  $TM$  a smooth manifold with dimension  $2n$ .

An analogous procedure gives rise to a natural smooth differentiable structure for

$$T^*M := \{(p, \omega) \mid p \in M, \omega_p \in T_p^*M\}.$$

**Definition 3.25. (Tangent and Cotangent Bundles or Spaces.)** Let  $M$  be a smooth manifold with dimension  $n$  and differentiable structure  $\mathcal{M}$ . If  $(U, \phi)$  is any local chart of  $\mathcal{M}$  with  $\phi : p \mapsto (x^1(p), \dots, x^n(p))$  define

$$TU := \{(p, v) \mid p \in U, v \in T_pM\}, \quad T^*U := \{(p, \omega) \mid p \in U, \omega \in T_p^*M\}$$

and

$$V := \{(p, v) \mid p \in U, v \in \hat{\phi}_p B\}, \quad {}^*V := \{(p, \omega) \mid p \in U, \omega \in {}^*\hat{\phi}_p B\},$$

where  $B \subset \mathbb{R}^n$  are open nonempty sets and  $\hat{\phi}_p : \mathbb{R}^n \rightarrow T_pM$  and  ${}^*\hat{\phi}_p : \mathbb{R}^n \rightarrow T_p^*M$  are the linear isomorphisms naturally induced by  $\phi$  as in (3.8). Finally define  $T\phi : TU \rightarrow \phi(U) \times \mathbb{R}^n \subset \mathbb{R}^{2n}$  and  $T^*\phi : T^*U \rightarrow \phi(U) \times \mathbb{R}^n \subset \mathbb{R}^{2n}$  such that

$$T\phi : (p, v) \mapsto (x^1(p), \dots, x^n(p), v_p^1, \dots, v_p^n),$$

where  $v = v_p^i \frac{\partial}{\partial x^i} \big|_p$  and

$$T^*\phi : (p, \omega) \mapsto (x^1(p), \dots, x^n(p), \omega_{1p}, \dots, \omega_{np}),$$

where  $\omega = \omega_{ip} dx^i \big|_p$ .

(a) The **tangent bundle** associated with  $M$  is the smooth manifold obtained by equipping

$$TM := \{(p, v) \mid p \in M, v \in T_pM\}$$

with:

- (1) the topology generated by the sets  $V$  above varying  $(U, \phi) \in \mathcal{M}$  and  $B$  in the class of open non-empty sets of  $\mathbb{R}^n$ ,
- (2) the differentiable structure induced by the atlas

$$\mathcal{A}(TM) := \{(U, T\phi) \mid (U, \phi) \in \mathcal{M}\}.$$

The local charts  $(TU, T\phi)$  of  $\mathcal{A}(TM)$  are said **natural local charts** on  $TM$  or also local chart adapted to the fiber-bundle structure of  $TM$ .

(b) The **cotangent bundle** associated with  $M$  is the manifold obtained by equipping

$$T^*M := \{(p, \omega) \mid p \in M, \omega \in T_p^*M\}$$



with:

- (1) the topology generated by the sets  ${}^*V$  above varying  $(U, \phi) \in \mathcal{M}$  and  $B$  in the class of open non-empty sets of  $\mathbb{R}^n$ ,
- (2) the differentiable structure induced by the atlas

$${}^*\mathcal{A}(TM) := \{(U, T^*\phi) \mid (U, \phi) \in \mathcal{M}\}.$$

The local charts  $(TU, T^*\phi)$  of  ${}^*\mathcal{A}(TM)$  are said **natural local charts** on  $T^*M$  or also local chart adapted to the fiber-bundle structure of  $T^*M$ .

The tangent bundle and cotangent bundle are also called **tangent space** and **cotangent space** respectively. ■

From now on we denote the tangent space, including its differentiable structure, by the same symbol used for the “pure set”  $TM$ . Similarly, the cotangent space, including its differentiable structure, will be indicated by  $T^*M$ .

For future reference, it is interesting to write down explicitly the relation between the coordinates of two local charts  $T\phi : TU \ni (p, v) \mapsto (x^1, \dots, x^n, v^1, \dots, v^n) \in \mathbb{R}^{2n}$  and  $\Psi : TV \ni (p, v) \mapsto (x'^1, \dots, x'^n, v'^1, \dots, v'^n) \in \mathbb{R}^{2n}$  on  $TM$  induced from two local charts, respectively,  $(U, \phi)$  and  $(V, \psi)$  on  $M$ , when  $U \cap V \neq \emptyset$ . By direct inspection one finds

$$x'^a = x'^a(x^1, \dots, x^n) \tag{3.9}$$

$$v'^a = \sum_{b=1}^n \frac{\partial x'^a}{\partial x^b} \big|_{\phi(p)} v^b. \tag{3.10}$$

We leave to the reader to write down the analog for  $T^*M$ .

**Remark 3.26.** It should be clear that the atlas  $\mathcal{A}(TM)$  (and the corresponding one for  $T^*M$ ) is not maximal and thus the differential structure on  $TM$  ( $T^*M$ ) is larger than the definitory atlas.

For instance suppose that  $\dim(M) = 2$ , and let  $(U, M)$  be a local chart of the smooth differentiable structure of  $M$ . Let the coordinates of the associated local chart on  $TM$ ,  $(TU, T\phi)$  be indicated by  $x^1, x^2, v^1, v^2$  with  $x^i \in \mathbb{R}$  associated with  $\phi$  and  $v^i \in \mathbb{R}$  components in the associated bases in  $T_{\phi^{-1}(x^1, x^2)}M$ . One can define new local coordinates on  $TU$ :

$$y^1 := x^1 + v^1, \quad y^2 := x^1 - v^1, \quad y^3 := x^2 + v^2, \quad y^4 := x^2 - v^2.$$

The corresponding local chart is admissible for the differential structure of  $TM$  but, in general, it does not belong to the atlas  $\mathcal{A}(TM)$  naturally induced by the differentiable structure of  $M$ .

There are some definitions related with definition 3.25 and concerning canonical projections, sections and lift of differentiable curves.

**Definition 3.27.** (**Canonical projections, sections, lifts.**) Let  $M$  be a smooth manifold. The surjective open smooth mappings

$$\Pi : TM \rightarrow M \quad \text{such that} \quad \Pi(p, v) \mapsto p ,$$

and

$$*\Pi : T^*M \rightarrow M \quad \text{such that} \quad *\Pi(p, \omega) \mapsto p ,$$

are called **canonical projections onto**  $TM$  and  $T^*M$  respectively.

A **section** of  $TM$  (respectively  $T^*M$ ) is a smooth map  $\sigma : M \rightarrow TM$  (respectively  $T^*M$ ), such that  $\Pi(\sigma(p)) = p$  (respectively  $*\Pi(\sigma(p)) = p$ ) for every  $p \in M$ .

If  $\gamma : I \ni t \mapsto \gamma(t) \in M$ ,  $I$  being an interval of  $\mathbb{R}$ , is a smooth curve, the **lift of**  $\gamma$ ,  $\Gamma$ , is the smooth curve in  $TM$ ,

$$\Gamma : I \ni t \mapsto (\gamma(t), \gamma'(t)) \in TM .$$

■

### 3.3.2 Fiber bundles

The structure of thangent bundle can be generalized to the general notion of *fiber bundle*. This ia a manifold  $E$  such that it can be locally viewed as the product manifold  $F \times M$  where  $F$  and  $M$  are canonical manifolds.

**Definition 3.28.** (**Fiber bundle.**) We say that we have a **Fiber bundle**  $E$  with **basis**  $M$ , **canonical fiber**  $F$ , and **canonical projection**  $\pi : E \rightarrow M$ , – where  $E, M, F$  are smooth manifolds and  $\pi$  a smooth function – if the following condition is satisfied.

Every point  $p \in M$  admits an open neighborhood  $U_p \subset M$  such that the open set  $\pi^{-1}(U_p)$  is diffeomorphic to  $F \times U_p$  through a diffeomorphism  $\phi : F \times U_p \rightarrow \pi^{-1}(U)$  which satisfies

$$\pi(\phi(x, y)) = y , \quad \forall (x, y) \in F \times U_p .$$

Above  $\pi^{-1}(U_p)$  is equipped with the natural differentiable structure induced by  $E$  and  $U_p$  is equipped with the natural differentiable structure induced by  $M$ .

A **local section** of  $E$  is a smooth map  $f : V \rightarrow E$ , where  $V \subset M$  is open, such that  $\pi(f(p)) = p$  for every  $p \in V$ . If  $V = M$  the local section is called **global**. The space of sections is denoted by  $\Gamma(E)$ . ■

**Remark 3.29.**

(1) Due to the requirement  $\pi(\phi(x, y)) = y$  for all  $(x, y) \in F \times U_p$  the map  $\pi : E \rightarrow M$  is necessarily surjective and it is also a *submersion* according to Definition 4.6 (we shall introduce in the next chapter) since it coincides through a diffeomorphism with the canonical projection

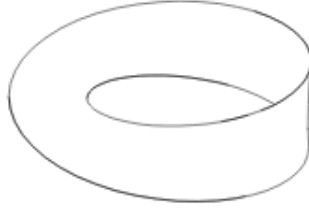


Figure 3.1: The Möbius strip

$F \times U_p \ni (x, y) \mapsto y \in U_p$  which is a submersion as it follows working in local coordinates of the natural atlas of the product manifold.

(2) A fiber bundle  $E$  as in Definition 3.28 is said to be **trivializable** if the open neighborhood  $U_p$  coincides with the whole basis  $M$ . It turns out that if  $M$  is diffeomorphic to  $\mathbb{R}$ , then  $E$  is always trivializable. The spacetime of the classical physics, when the basis is the time axis, is trivializable, but not in a canonical way, every reference frame defines a different trivialization [Mor20b].

(3) It should be evident that  $TM$  is a fiber bundle with basis  $M$ , canonical fiber  $\mathbb{R}^{\dim(M)}$ , and canonical projection  $\Pi$ . A smooth manifold  $M$  is said to be **parallelizable** if  $TM$  admits  $m := \dim(M)$  linearly independent sections. In other words, there must exist  $m$  smooth vector fields such that they define a basis of  $T_p M$  for every  $p \in M$ . This, in turn, is equivalent to saying that  $TM$  is trivializable. Every affine space is parallelizable. the spheres  $S^1$ ,  $S^3$ ,  $S^7$  are parallelizable, whereas all remaining sphere,  $S^2$  in particular, are not parallelizable. ■

### Examples 3.30.

(1) The natural issue is whether or not there exist fiber bundles which are not globally diffeomorphic to  $F \times M$ . The famous so-called *Möbius strip* is a well known counterexample. It is *locally* constructed as the product  $U \times S^1$ , where  $U = (-1, 1)$  and  $S^1$  is the unit circle  $S^1 := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ , but it is not diffeomorphic to it, more weakly,  $U$  is its canonical fiber and  $S^1$  is its basis.

(2) It is not guaranteed that  $\Gamma(E)$  is non-trivial (this is the case for the fiber bundle over  $S^1$  with fiber  $F := \mathbb{R} \setminus \{0\}$  obtained by taking the Möbius bundle and removing the zero section), sometimes only local sections exists. However, in the special case of  $TM$  and  $T^*M$  global sections exist. Since  $TM$  is a fiber bundle the smooth vector fields on  $M$  are sections of  $TM$  and *vice versa*:

$$\mathfrak{X}(M) = \Gamma(TM) .$$

An analogous fact is true for  $T^*M$ :

$$\Omega^1(M) = \Gamma(T^*M) .$$

■

## Chapter 4

# Differential mapping, submanifolds, Lie derivative

This chapter discusses the interplay of the notion of smooth transformation and that of smooth manifold. This interplay passes through and link together fundamental mathematical concepts as the various notions of submanifold, that of flow of a smooth vector field and Lie derivative.

### 4.1 Differential of mappings and its applications

A fundamental mathematical tool in differential geometry is the *differential* of a smooth function. It gives rise to a variety of interesting and pervasive mathematical structures.

#### 4.1.1 Push forward

**Definition 4.1.** (**Differential of a mapping or push forward.**) If  $f : N \rightarrow M$  is a smooth function from the smooth manifold  $N$  to the smooth manifold  $M$ , the **differential of  $f$**  at  $p$  or **push forward of  $f$**  at  $p$  is the linear mapping

$$df_p : T_p N \ni X_p \mapsto df X_p \in T_{f(p)} M ,$$

defined by

$$df X_p(g) := X_p(g \circ f) ,$$

for all vectors  $X_p \in T_p N$  and all smooth functions  $g \in D(M)$ . ■

**Remark 4.2.**

(1) Notice that the definition above is an extension of (3.3) when  $M = \mathbb{R}$  as it is even more evident from the discussion below.

(2) With the meaning of  $df$  as in the definition above,  $df$  is often indicated by  $f_*$ . ■

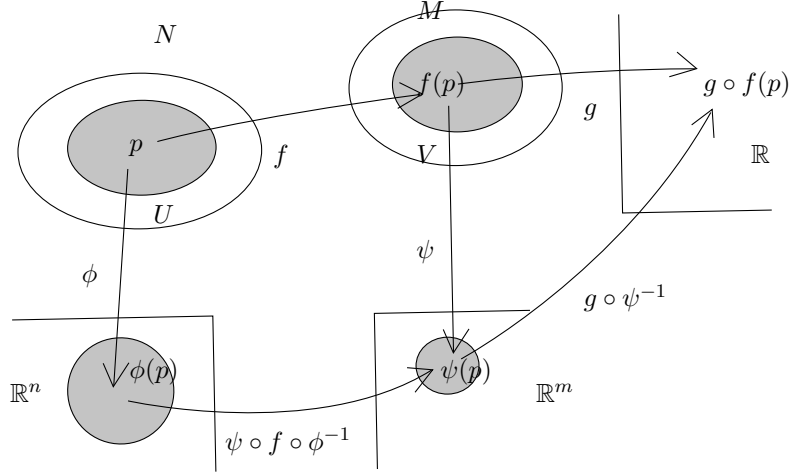


Figure 4.1: Diagrams to compute the differential

Let us study the explicit form of the differential map in components. Take two local charts  $(U, \phi)$  in  $N$  and  $(V, \psi)$  in  $M$  around  $p$  and  $f(p)$  respectively and use the notation  $\phi : U \ni q \mapsto (x^1(q), \dots, x^n(q))$  and  $\psi : V \ni r \mapsto (y^1(r), \dots, y^m(r))$ . Then define

$$\tilde{f} := \psi \circ f \circ \phi^{-1} : \phi(U) \rightarrow \mathbb{R}^m \quad \text{and} \quad \tilde{g} := g \circ \psi^{-1} : \psi(V) \rightarrow \mathbb{R}.$$

$\tilde{f}$  and  $\tilde{g}$  represent  $f$  and  $g$ , respectively, in the fixed coordinate systems. By construction, it holds

$$X_p(g \circ f) = X_p^i \frac{\partial}{\partial x^i} (g \circ f \circ \phi^{-1}) = X_p^i \frac{\partial}{\partial x^i} (g \circ \psi^{-1} \circ \psi \circ f \circ \phi^{-1}).$$

That is, with obvious notation

$$X_p(g \circ f) = X_p^i \frac{\partial}{\partial x^i} (\tilde{g} \circ \tilde{f}) = X_p^i \frac{\partial \tilde{g}}{\partial y^k} \Big|_f \frac{\partial \tilde{f}^k}{\partial x^i} = \left( \frac{\partial \tilde{f}^k}{\partial x^i} X_p^i \right) \frac{\partial \tilde{g}}{\partial y^k} \Big|_f.$$

In other words

$$((df X_p)g)^k = \left( \frac{\partial \tilde{f}^k}{\partial x^i} X_p^i \right) \frac{\partial \tilde{g}}{\partial y^k} \Big|_f.$$

This means that, with the said notations, the following very useful coordinate form of  $df_p$  can be given

$$df_p : X_p^i \frac{\partial}{\partial x^i} \Big|_p \mapsto X_p^i \frac{\partial (\psi \circ f \circ \phi^{-1})^k}{\partial x^i} \Big|_{\phi(p)} \frac{\partial}{\partial y^k} \Big|_{f(p)}. \quad (4.1)$$

That formula is more often written

$$df_p : X_p^i \frac{\partial}{\partial x^i} \Big|_p \mapsto X_p^i \frac{\partial y^k}{\partial x^i} \Big|_{(x^1(p), \dots, x^n(p))} \frac{\partial}{\partial y^k} \Big|_{f(p)}, \quad (4.2)$$

where it is understood that  $\psi \circ f \circ \phi^{-1} : (x^1, \dots, x^n) \mapsto (y^1(x^1, \dots, x^n), \dots, y^m(x^1, \dots, x^n))$ .

### 4.1.2 Pull back

The dual application of the push forward, which transforms covariant vectors, is called *pull back* since it proceeds in the opposite direction.

**Definition 4.3.** (**Pull back.**) If  $f : N \rightarrow M$  is a smooth function from the smooth manifold  $N$  to the smooth manifold  $M$ , the **pull back** at  $q = f(p)$

$$f_q^* : T_q^* M \ni \omega_q \mapsto f^* \omega_q \in T_p^* N ,$$

defined by

$$\langle X_p, f^* \omega_q \rangle = \langle df X_p, \omega_q \rangle$$

for all vector  $X_p \in T_p M$ . ■

A straightforward computations strictly analogous to the previous one proves that, in coordinates,

$$f_q^* : (\omega_q)_i dy^i|_q \mapsto (\omega_q)_k \frac{\partial y^k}{\partial x^i}|_{(x^1(p), \dots, x^n(p))} dx^i|_p , \quad (4.3)$$

where it is understood that  $\psi \circ f \circ \phi^{-1} : (x^1, \dots, x^n) \mapsto (y^1(x^1, \dots, x^n), \dots, y^m(x^1, \dots, x^n))$ .

## 4.2 Rank of a smooth map and related notions

The notion of differential allows one to define the *rank* of a map and associated definitions useful in distinguishing among the various types of submanifolds of a given manifold.

### 4.2.1 Rank of a smooth map

Consider a smooth map between two smooth manifolds  $f : N \rightarrow M$ . If  $(U, \phi)$  and  $(V, \psi)$  are local charts around  $p$  and  $f(p)$  respectively, the rank of the Jacobian matrix of the function  $\psi \circ f \circ \phi^{-1} : \phi(U) \rightarrow \mathbb{R}^m$  computed in  $\phi(p)$  does not depend on the choice of those charts. This is because any change of charts transforms the Jacobian matrix into a new matrix obtained by means of left or right composition with nonsingular square matrices and this does not affect the range.

**Definition 4.4.** If  $f : N \rightarrow M$  is a smooth function from the smooth manifold  $N$  to the smooth manifold  $M$  and  $p \in N$ .

- (a) The **rank** of  $f$  at  $p$  is the rank of  $df_p$  (that is the rank of the Jacobian matrix of the function  $\psi \circ f \circ \phi^{-1}$  computed in  $\phi(p) \in \mathbb{R}^n$ ,  $(U, \phi)$  and  $(V, \psi)$  being a pair of local charts around  $p$  and  $f(p)$  respectively);
- (b)  $p$  is called a **critical point** or **singular point** of  $f$  if the rank of  $f$  at  $p$  is not maximal. Otherwise  $p$  is called **regular point** of  $f$ ;

- (c) If  $p$  is a critical point of  $f$ ,  $f(p)$  is called **critical value** or **singular value** of  $f$ . A **regular value** of  $f$ ,  $q$  is a point of  $M$  such that every point in  $f^{-1}(q)$  is a regular point of  $f$ . ■

We have the following remarkable results concerning regular points.

**Theorem 4.5.** *Let  $f : N \rightarrow M$  be a smooth function with  $M$  and  $N$  smooth manifolds with dimension  $m$  and  $n$  respectively and take  $p \in N$ .*

- (1) *If  $n \geq m$  and  $p$  is a regular point, i.e.  $df_p$  is surjective, then  $f$  looks like the canonical projection of  $\mathbb{R}^n$  onto  $\mathbb{R}^m$  around  $p$ .*

*In other words, for any local chart  $(V, \psi)$  around  $f(p)$  there is a local chart  $(U, \phi)$  around  $p$  such that*

$$\psi \circ f \circ \phi^{-1}(x^1, \dots, x^m, \dots, x^n) = (x^1, \dots, x^m) \quad \text{if } (x^1, \dots, x^n) \in \phi(U).$$

- (2) *If  $n \leq m$  and  $p$  is a regular point, i.e.  $df_p$  is injective, then  $f$  looks like the canonical injection of  $\mathbb{R}^n$  into  $\mathbb{R}^m$  around  $p$ .*

*In other words, for any local chart  $(U, \phi)$  around  $p$  there is a local chart  $(V, \psi)$  around  $f(p)$  such that*

$$\psi \circ f \circ \phi^{-1}(x^1, \dots, x^n) = (x^1, \dots, x^n, 0, \dots, 0) \quad \text{if } (x^1, \dots, x^n) \in \phi(U).$$

- (3) *If  $n = m$ , then the following statements are equivalent:*

(a)  *$p$  is a regular point, i.e.,  $df_p : T_p N \rightarrow T_{f(p)} M$  is a linear isomorphism;*

(b)  *$f$  defines a **local diffeomorphism around  $p$** . In other words, there is an open neighborhood  $U$  of  $p$  and an open neighborhood  $V$  of  $f(p)$  such that  $f|_U : U \rightarrow V$  is a diffeomorphism from the smooth manifold  $U$  equipped with the natural differentiable structure induced<sup>1</sup> by  $N$  to the smooth manifold  $V$  equipped with the similar natural differentiable structure induced by  $M$ .*

**Sketch of proof.** Working in local coordinates in  $N$  and  $M$  and passing to work with the Jacobian matrices of the involved functions (1) and (2) are direct consequences of Dini's implicit function theorem. Let us pass to consider (3). Suppose that  $g := f|_U$  is a diffeomorphism onto  $V$ . In that case  $g^{-1} : V \rightarrow U$  is a diffeomorphism to and  $g \circ f = id_U$ . Working in local coordinates around  $p$  and  $f(p)$  and computing the Jacobian matrix of  $g \circ f$  in  $p$  one gets  $J[g]_{f(p)} J[f]_p = I$ . This means that both  $\det J[g]_{f(p)}$  and  $\det J[f]_p$  cannot vanish. In particular  $\det J[f]_p \neq 0$  is equivalent to the fact that  $df_p$  is a linear isomorphism. Conversely, assume that  $df_p$  is a linear isomorphism. In that case both (1) and (2) above hold and there is a pair of open neighborhoods  $U \ni p$  and  $V \ni f(p)$  equipped with coordinates such that

$$\psi \circ f \circ \phi^{-1}(x^1, \dots, x^m) = (x^1, \dots, x^m),$$

which means that  $\psi \circ f \circ \phi^{-1}(x^1, \dots, x^m) : \phi(U) \rightarrow \psi(V)$  is the (restriction of) identity map on  $\mathbb{R}^m$ . This fact immediately implies that  $f|_U$  is a diffeomorphism onto  $V$ . □

---

<sup>1</sup>According to Remark 2.10.



### 4.2.2 Immersions, submersions, submanifolds

Let us pass to introduce the notion of submanifold.

**Definition 4.6.** If  $f : N \rightarrow M$  is a smooth function from the smooth manifold  $N$  to the smooth manifold  $M$  then,

- (a)  $f$  is called **submersion** if  $df_p$  is surjective for every  $p \in N$ ;
- (b)  $f$  is called **immersion** if  $df_p$  is injective for every  $p \in N$ ;
- (c) An immersion  $f : N \rightarrow M$  is called **embedding** if
  - (i) it is injective and
  - (ii)  $f : N \rightarrow f(N)$  is a homeomorphism when  $f(N)$  is equipped with the topology induced by  $M$ . ■

**Definition 4.7.** Let  $f : N \rightarrow M$  a smooth map between two smooth manifolds.

- (a) If  $f$  is an injective immersion, then  $N$  is called **immersed smooth submanifold** of  $M$  through  $f$ .
- (a) If  $f$  is an embedding, then  $N$  is called **embedded smooth submanifold** of  $M$  (or simply **smooth submanifold** of  $M$ ) through  $f$ .

In the absence of further specification, if  $N \subset M$  as sets, the injective map  $f : N \rightarrow M$  is assumed to be the **inclusion map**  $i : N \hookrightarrow M$ , defined as  $i : N \ni x \mapsto x \in M$  (and however  $N$  must be autonomously equipped with a smooth manifold structure). ■

An equivalent definition of embedded submanifold for  $N \subset M$  can be given by using the statement in (b) of the following proposition.

**Proposition 4.8.** Let  $M$  be a smooth manifold with dimension  $m$  and let  $N \subset M$ .

- (a) If  $N$  is a smooth manifold with dimension  $n$  which is an embedded submanifold of  $M$ , the following pair of conditions are satisfied:
  - (i) the topology of  $N$  is that induced by  $M$ ,
  - (ii) for every  $p \in N$  (and thus  $p \in M$ ) there is an open (in  $M$ ) neighborhood of  $p$ ,  $U_p$  and a local chart of  $M$ ,  $(U_p, \phi)$ , such that if we use the notation,  $\phi : q \mapsto (x^1(q), \dots, x^m(q))$ , it holds

$$\phi(N \cap U_p) = \{(x^1, \dots, x^m) \in \phi(U) \mid x^{n+1} = 0, \dots, x^m = 0\}.$$

- (b) If (i) and (ii) hold for some fixed  $n \leq m$ ,  $N$  can be equipped with a differentiable structure  $\mathcal{N}$  so that it results to be a submanifold with dimension  $n$  of  $M$  with embedding given by the inclusion map. That differentiable structure is obtained as follows. The maps  $N \cap U_p \ni q \mapsto (x^1(q), \dots, x^n(q))$  define a local chart around every point  $p \in N$  with domain  $V_p = N \cap U_p$ . The set of these charts is an atlas whose generated differentiable structure is  $\mathcal{N}$ .

**Sketch of proof.** (a) if  $N \subset M$  is a smooth submanifold of  $M$  (the embedding being the inclusion map), the topology of  $N$  must be that induced by  $M$  because the inclusion map is a homeomorphism from the topological manifold  $N$  to the subset  $N \subset M$  equipped with the topology induced by  $M$ . Using Theorem 4.5 (items (2) and (3)) where  $f$  is replaced by the inclusion map one straightforwardly proves the validity of (ii).

(b) Under the given hypotheses equip  $N$  with the topology induced by  $M$ . As a consequence  $N$  turns out to be Hausdorff and second countable. By direct computation, it result that, if the conditions (i),(ii), are satisfied, the local charts with domains  $V_p$  defined in (b), varying  $p \in N$ , are: (1) local homeomorphisms from  $N$  to  $\mathbb{R}^n$  (this is because the maps  $\phi$  are local homeomorphisms from  $M$  to  $\mathbb{R}^m$ ), (2) pairwise compatible (this is because these charts are restrictions of pairwise compatible charts). Since there is such a chart around every point of  $N$ , the set of the considered charts is an atlas of  $N$ . Using such an atlas it is simply proved by direct inspection that the inclusion map  $i : N \hookrightarrow M$  is an embedding.  $\square$

**Remark 4.9.** If  $f : N \rightarrow M$  is an *injective* immersion, according to (2) in Theorem 4.5, for every  $p \in N$  we can always find two local charts  $(U, \phi)$  around  $p$  in  $N$ , with  $\phi(q) = (x^1(q), \dots, x^n(q))$ , and  $(V, \psi)$  around  $f(p)$  in  $M$ , with  $\psi(r) = (y^1(r), \dots, y^m(r))$ , such that

- (a)  $f(U)$  coincides with the set  $y^{n+1} = \dots = y^m = 0$  in  $\psi(V)$  and
- (b)  $(x^1(r), \dots, x^n(r)) = (y^1(f(r)), \dots, y^n(f(r)))$  for every  $r \in U$

The fundamental difference with the corresponding result for (embedded) submanifolds is that now  $f(N) \cap V \subsetneq f(U)$  in general. Instead for embeddings we can always arrange the neighborhoods in order that  $f(N) \cap V = f(U)$  as a consequence of Proposition 4.8.  $\blacksquare$

#### Examples 4.10.

1. The map  $\gamma : \mathbb{R} \ni t \mapsto (\sin t, \cos t) \in \mathbb{R}^2$  is an immersion, since  $d\gamma \neq 0$  (which is equivalent to say that  $\gamma' \neq 0$ ) everywhere. Anyway that is not an embedding since  $\gamma$  is not injective. So  $\mathbb{R}$  is not an immersed manifold nor and embedded manifold through  $\gamma$ .
2. Changing perspective and looking at the only image  $C = \gamma(\mathbb{R})$  this set turns out to be an embedded submanifold of  $\mathbb{R}^2$  (through the inclusion map) if  $C$  is equipped with the topology induced by  $\mathbb{R}^2$  and the differentiable structure is that built up by using (b) of proposition 4.8. In fact, take  $p \in C$  and notice that there is some  $t \in \mathbb{R}$  with  $\gamma(t) = p$  and  $d\gamma_p \neq 0$ . Using (2) of theorem 4.5, there is a local chart  $(U, \psi)$  of  $\mathbb{R}^2$  around  $p$  referred to coordinates  $(x^1, x^2)$ , such that the portion of  $C$  which has intersection with  $U$  is represented by  $(x^1, 0)$ ,  $x^1 \in (a, b)$ . For

instance, such coordinates are polar coordinates  $(\theta, r)$ ,  $\theta \in (-\pi, \pi)$ ,  $r \in (0, +\infty)$ , centered in  $(0, 0) \in \mathbb{R}^2$  with polar axis (i.e.,  $\theta = 0$ ) passing through  $p$ . These coordinates define a local chart around  $p$  on  $C$  in the set  $U \cap C$  with coordinate  $x^1$ . All the charts obtained by varying  $p$  are pairwise compatible and thus they give rise to a differentiable structure on  $C$ . By proposition 4.8 that structure makes  $C$  a submanifold of  $\mathbb{R}^2$ . On the other hand, the inclusion map, which is always injective, is an immersion because it is locally represented by the trivial immersion  $x^1 \mapsto (x^1, 0)$ . As the topology on  $C$  is that induced by  $\mathbb{R}$ , the inclusion map is a homeomorphism. So the inclusion map  $i : C \hookrightarrow \mathbb{R}^2$  is an embedding and this shows once again that  $C$  is a submanifold of  $\mathbb{R}^2$  using the definition itself.

**3.** Consider the set in  $\mathbb{R}^2$ ,  $C := \{(x, y) \in \mathbb{R}^2 \mid x^2 = y^2\}$ . It is *not* possible to give a differentiable structure to  $C$  in order to have a one-dimensional submanifold of  $\mathbb{R}^2$ . This is because  $C$  equipped with the topology induced by  $\mathbb{R}^2$  is not locally homeomorphic to  $\mathbb{R}$  due to the point  $(0, 0)$ .

**4.** Is it possible to endow  $C$  defined in **3** with a differentiable structure and make it a one-dimensional smooth manifold? The answer is yes.  $C$  is connected but is the union of the disjoint sets  $C_1 := \{(x, y) \in \mathbb{R}^2 \mid y = x\}$ ,  $C_2 := \{(x, y) \in \mathbb{R}^2 \mid y = -x, x > 0\}$  and  $C_3 := \{(x, y) \in \mathbb{R}^2 \mid y = -x, x < 0\}$ .  $C_1$  is homeomorphic to  $\mathbb{R}$  defining the topology on  $C_1$  by saying that the open sets of  $C_1$  are all the sets  $f_1(I)$  where  $I$  is an open set of  $\mathbb{R}$  and  $f_1 : \mathbb{R} \ni x \mapsto (x, x)$ . By the same way,  $C_2$  turns out to be homeomorphic to  $\mathbb{R}$  by defining its topology as above by using  $f_2 : \mathbb{R} \ni z \mapsto (e^z, -e^z)$ .  $C_3$  enjoys the same property by defining  $f_3 : \mathbb{R} \ni z \mapsto (-e^z, e^z)$ . The maps  $f_1^{-1}, f_2^{-1}, f_3^{-1}$  also define a global coordinate system on  $C_1, C_2, C_3$  respectively and separately, each function defines a local chart on  $C$ . The differentiable structure generated by the atlas defined by those functions makes  $C$  a smooth manifold with dimension 1 and cannot be considered a submanifold of  $\mathbb{R}^2$ . It is clear that  $C$  equipped with the said differentiable structure is however an immersed submanifold of  $\mathbb{R}^2$  through the inclusion map.

**5.** Consider the set in  $\mathbb{R}^2$ ,  $C = \{(x, y) \in \mathbb{R}^2 \mid y = |x|\}$ . This set cannot be equipped with a suitable differentiable structure which makes it an embedded submanifold of  $\mathbb{R}^2$  (through the inclusion map). Actually, differently from above, here the problem concerns the smoothness of the inclusion map at  $(0, 0)$  rather than the topology of  $C$ . In fact,  $C$  is naturally homeomorphic to  $\mathbb{R}$  when equipped with the topology induced by  $\mathbb{R}^2$ . Nevertheless there is no way to find a local chart in  $\mathbb{R}^2$  around the point  $(0, 0)$  such that the requirements of proposition 4.8 are fulfilled due to the corner in that point of the curve  $C$ . However, it is simply defined a differentiable structure on  $C$  which make it a one-dimensional smooth manifold. It is sufficient to consider the differentiable structure generated by the global chart given by the inverse of the homeomorphism  $f : \mathbb{R} \ni t \mapsto (|t|, t)$ . The inclusion map does not permit to view  $C$  as an immersed submanifold of  $\mathbb{R}^2$  due to smoothness problems with the corner.

**6.** Let us consider once again the cylinder  $C \subset \mathbb{E}^3$  defined in the example 5.8.2.  $C$  is an embedded submanifold of  $\mathbb{E}^3$  (referring as usual to the inclusion map). Since the construction of the differential structure made in the example 5.8.2 is that of proposition 4.8 starting from cylindrical coordinates  $\theta, r' := r - 1, z$ .

We finally quote a fundamental theorem<sup>2</sup> due to Whitney concerning the fact that every smooth manifold can be always seen as an embedded submanifold of a  $\mathbb{R}^n$  space with  $n$  sufficiently large.

**Theorem 4.11.** (Whitney embedding theorem) *If  $M$  is a smooth  $n$ -dimensional manifold (Hausdorff and second countable), then there exists a smooth embedding  $f : M \rightarrow \mathbb{R}^{2n}$ .*

**Remark 4.12.** From now on, *smooth submanifold* means *embedded smooth manifold*. ■.

### 4.2.3 Theorem of regular values

To conclude this concise discussion about submanifolds, we state (without proof) a very important theorem with various application in mathematical physics.

**Theorem 4.13.** (Theorem of regular values.) *Let  $f : N \rightarrow M$  be a smooth function from the smooth manifold  $N$  to the smooth manifold  $M$  with  $\dim M < \dim N$ .*

*If  $y \in M$  is a regular value of  $f$ ,  $P := f^{-1}(\{y\}) \subset N$  is a smooth submanifold of  $N$ .*

**Remark 4.14.** A know theorem due to Sard, proves that the *measure* of the set of singular values of any smooth function  $f : N \rightarrow M$  must vanish. This means that, if  $S \subset M$  is the set of singular values of  $f$ , for every local chart  $(U, \phi)$  in  $M$ , the set  $\phi(S \cap U) \subset \mathbb{R}^m$  has vanishing Lebesgue measure in  $\mathbb{R}^m$  where  $m = \dim M$ . ■

### Examples 4.15.

1. In *analytical mechanics*, consider a system of  $N$  material points with possible positions  $P_k \in \mathbb{R}^3$ ,  $k = 1, 2, \dots, N$  and  $c$  constraints given by assuming  $f_i(P_1, \dots, P_N) = 0$  where the  $c$  functions  $f_i : \mathbb{R}^{3N} \rightarrow \mathbb{R}$ ,  $i = 1, \dots, c$  are smooth. If the constraints are *functionally independent*, i.e. the Jacobian matrix of elements  $\frac{\partial f_i}{\partial x_k}$  has rank  $c$  everywhere,  $x^1, x^2, \dots, x^{3N}$  being the coordinates of  $(P_1, \dots, P_N) \in (\mathbb{R}^3)^N$ , the configuration space is a submanifold of  $\mathbb{R}^{3N}$  with dimension  $3N - c$ . This result is nothing but a trivial application of theorem 4.13.
2. Consider (2) in Exercises 2.15 from another point of view. As a set the circumference  $C = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$  is  $f^{-1}(0)$  with  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  defined as  $f(x, y) := x^2 + y^2 - 1$ . The value 0 is a regular value of  $f$  because  $df_p = 2x dx + 2y dy \neq 0$  if  $f(x, y) = 0$  that is  $(x, y) \in C$ . As a consequence of Theorem 4.13,  $C$  can be equipped with the structure of submanifold of  $\mathbb{R}^2$ . This structure is that defined in (2) Example 4.10.
3. Consider a function  $z = g(x, y)$  in  $\mathbb{R}^3$  where  $(x, y) \in U$ ,  $U$  being any open set and suppose that the function  $g$  is smooth. Since  $dz_p \neq 0$ , the function  $f(x, y, z) := z - g(x, y)$  satisfies  $df_p = dz|_p + \frac{\partial g}{\partial x}|_p dx|_p + \frac{\partial g}{\partial y}|_p dy|_p \neq 0$  for every point  $p \in U \times \mathbb{R}$ . In particular this fact happens for the points such that  $f(p) = 0$ . As a consequence the map  $z = g(x, y)$  defines a two-dimensional submanifold embedded in  $\mathbb{R}^3$ .

---

<sup>2</sup>M. Adachi, *Embeddings and Immersions*, translated by Hudson, Kiki, AMS (1993).

### 4.3 Flow of a vector field, Lie derivative, Frobenius Theorem

This section is devoted to present some further applications of the Lie commutator of vector fields. First of all, it has the interpretation of *Lie derivative*, which is one of the two notions of derivative of vector fields we consider in this book. Another use of the Lie parenthesis concerns the theory of integrable distributions and the celebrated *Frobenius Theorem*. To introduce the notion of Lie derivative we need a preparatory discussion about the concept of *flow of a vector field*.

#### 4.3.1 Almost all about the flow of a vector field

We start by introducing the notion of *local flow* of  $X \in \mathfrak{X}(M)$  for a smooth manifold  $M$ . Let us consider the problem of finding a smooth curve

$$J_p \ni t \mapsto \gamma_p(t), \quad (4.4)$$

where  $J_p \ni 0$  is any open interval of  $\mathbb{R}$ , that solves the following Cauchy problem

$$\gamma'_p = X_{\gamma_p(t)} \quad \text{satisfying} \quad \gamma_p(0) = p \in M. \quad (4.5)$$

It is known (see, e.g. [Lee03]) that, for every  $p \in M$ , there exists and is unique a solution that is not the restriction of a solution of the same problem defined on a larger domain. This solution is called the **maximal solution** of the Cauchy problem (4.5). It will be henceforth denoted by

$$I_p \ni t \mapsto \gamma_p(t). \quad (4.6)$$

It turns out that every other solution of the same Cauchy problem is a restriction of it. The following elementary but crucial result is valid.

**Proposition 4.16.** *Referring to the maximal solutions (4.6) of the Cauchy problem (4.5),*

$$A := \cup_{p \in M} I_p \times \{p\}$$

*is an open set in  $\mathbb{R} \times M$  including  $\{0\} \times M$  and the map  $\Phi^{(X)} : A \rightarrow M$  defined as*

$$\Phi_t^{(X)}(p) := \gamma_p(t) \quad (4.7)$$

*called **(local) flow** of  $X \in \mathfrak{X}(M)$ , is jointly smooth in its variables.*

**Proof.** See [Lee03]. □

**Definition 4.17.** A smooth vector field  $X$  on the smooth manifold  $M$  is said to be **complete** if all the maximal solutions of (4.5) are **complete**, i.e., defined in all  $(-\infty, +\infty)$ . ■

We can now state and prove a proposition regarding the basic local properties of  $\Phi^{(X)}$ .

**Proposition 4.18.** *Let  $M$  be a smooth manifold and  $X \in \mathfrak{X}(M)$ . The following facts are valid.*

(1) *For every  $p \in M$  there are an open neighborhood  $U_p \ni p$  and an open interval  $L_p \ni 0$  such that,*

- (a) *if  $t \in L_p$ , then  $\Phi_t^{(X)}(U_p)$  is open and  $\Phi_t^{(X)}|_{U_p} : U_p \rightarrow \Phi_t^{(X)}(U_p)$  is a diffeomorphism;*
- (b) *if  $q \in U_p$  and  $s, t, s+t \in L_p$ , then*

$$\Phi_0^{(X)}(q) = q \quad \text{and} \quad \Phi_t^{(X)} \circ \Phi_s^{(X)}(q) = \Phi_s^{(X)} \circ \Phi_t^{(X)}(q) = \Phi_{t+s}^{(X)}(q). \quad (4.8)$$

(2) *If  $X$  is complete, then  $\{\Phi_t^{(X)}\}_{t \in \mathbb{R}}$  is **one-parameter group of diffeomorphisms** of  $M$ :*

$$\Phi_0^{(X)} = id_M \quad \text{and} \quad \Phi_t^{(X)} \circ \Phi_s^{(X)} = \Phi_{t+s}^{(X)} \quad \forall s, t \in \mathbb{R} \quad (4.9)$$

*In particular,*

$$\Phi_{-t}^{(X)} = (\Phi_t^{(X)})^{-1} \quad \forall t \in \mathbb{R}. \quad (4.10)$$

**Proof.** We start the proof by observing that, from the uniqueness theorem for the maximal solutions of the Cauchy problem (4.5) with a generic initial time  $t_0$  in place of 0, it easily follows that

$$\Phi_0^{(X)}(q) = q \quad \text{and} \quad \Phi_t^{(X)} \circ \Phi_s^{(X)}(q) = \Phi_{t+s}^{(X)}(q) \quad (4.11)$$

for every fixed  $p \in M$  and, regarding the latter identity, every corresponding pair  $(s, t)$  such that  $(s, q) \in A$  and  $(s+t, q) \in A$ .

(2) If  $X$  is complete then  $A = \mathbb{R} \times M$  and thus these conditions are always satisfied for all  $t, s \in \mathbb{R}$  and all  $q \in M$ , proving the last assertion in the thesis.

(1) Since  $A$  is open in view of the Proposition 4.16, and  $\Phi^{(X)}$  is continuous and defined around the points  $(0, p)$ , the preimage of an open neighborhood of  $(0, p) \in A$  is an open set which we can always choose of the form  $(-\delta, \delta) \times U_p$ , where  $U_p$  is a suitably small open neighborhood of  $p$  and  $\delta > 0$  is small enough. Defining  $L_p := (-\delta/2, \delta/2)$ , we have that  $t, s \in L_p$  and  $q \in U_p$  entail  $(t, q), (s, q), (t+s, q) \in (-\delta, \delta) \times U_p \subset A$  and (4.11) are satisfied in view of the initial observation. More strongly we also have

$$\Phi_t^{(X)} \circ \Phi_s^{(X)}(q) = \Phi_{t+s}^{(X)}(q) = \Phi_{s+t}^{(X)}(q) = \Phi_s^{(X)} \circ \Phi_t^{(X)}(q),$$

completing the proof of (4.8). Notice that the second identity in (4.11) implies that

$$\Phi_{-t}^{(X)} \circ \Phi_t^{(X)}(q) = q \quad (4.12)$$

is valid for  $q \in U$  and  $t \in I$ . Since both composed maps are differentiable, it immediately implies that  $d\Phi_t^{(X)}$  is bijective on  $T_p M$  so that (Theorem 4.5)

$$\Phi_t^{(X)}|_{U_p} : U_p \rightarrow \Phi_t^{(X)}(U_p)$$

is a local diffeomorphism around each point of its domain. In particular, it is an open map. Therefore the set  $\Phi_t^{(X)}(U_p)$  must be open because  $U_p$  is open. Since  $\Phi_t^{(X)}|_{U_p}$  is also injective for (4.12), it defines a (global) diffeomorphism onto its image.  $\square$

The following useful proposition extends a well known fact from  $\mathbb{R}^n$  to smooth manifolds.

**Proposition 4.19.** *Let  $M$  be a smooth manifold and  $X \in \mathfrak{X}(M)$ . Assume that for a maximal solution  $\gamma_p : I_p \rightarrow M$  of (4.5) there is a compact set  $K \subset M$  such that  $\gamma(t) \in K$  for all  $I_p \cap [0, +\infty)$ , then  $\sup I_p = +\infty$ . Similarly, if  $\gamma(t) \in K$  for all  $I_p \cap (-\infty, 0]$ , then  $\inf I_p = -\infty$ .*

**Proof.** See [Lee03].  $\square$

**Remark 4.20.** As a consequence of Proposition 4.19, if  $X$  has compact support – this is the case when  $M$  itself is compact in particular – then  $X$  is complete and the flow of  $X$  is global as in (2) of Proposition 4.18.  $\blacksquare$

A finer result that completes Proposition 4.18 is the following one.

**Proposition 4.21.** *Let  $M$  be a smooth manifold and  $X \in \mathfrak{X}(M)$ . Referring to (4.6) and (4.5) define*

$$D_t := \{p \in M \mid I_p \ni t\} \quad \forall t \in \mathbb{R}.$$

$\Phi^{(X)}$  and the sets  $D_t$  enjoy the following properties.

- (1)  $D_t$  is open (possibly  $D_t = \emptyset$ ).
- (2)  $\bigcup_{t>0} D_t = M$  and  $\bigcup_{t<0} D_t = M$ .
- (3)  $\Phi_t^{(X)} : D_t \rightarrow D_{-t}$  is a diffeomorphism with inverse  $\Phi_{-t}^{(X)}$  for every  $t \in \mathbb{R}$ .
- (4) If  $s, t \in \mathbb{R}$  and  $p \in M$  are such that  $\Phi_s^{(X)} \circ \Phi_t^{(X)}(p)$  is well defined, then  $p \in D_{t+s}$  and
$$\Phi_s^{(X)} \circ \Phi_t^{(X)}(p) = \Phi_{s+t}^{(X)}(p).$$
- (5) If  $s, t \in \mathbb{R}$  have the same sign, then  $\Phi_s^{(X)} \circ \Phi_t^{(X)}$  is well defined on  $D_{t+s}$  and the identity above is valid for every  $p \in D_{s+t}$ .

**Proof.** See [War83].  $\square$ .

**Remark 4.22.** Evidently, if either  $0 \leq t < s$  or if  $s < t \leq 0$ , then  $D_s \subset D_t$ .  $\square$

### 4.3.2 Lie derivative of vector fields

We aim now to define the notion of *derivative of a vector field  $X$  with respect to another vector field  $Y$*  on a generic manifold  $M$ . The major obstruction one faces in trying to extend to manifolds the standard definition in  $\mathbb{R}^n$ ,

$$\nabla_Y X = \lim_{h \rightarrow 0} \frac{1}{h} (X_{p+hY_p} - X_p) ,$$

is the absence of a common affine structure independent of the chosen coordinate system: here  $X_{p+hY_p} \in T_{p+hY_p}M$ , whatever meaning we give to  $p+hY_p$ , whereas  $X_p \in T_pM$  and the difference of vectors of two different tangent spaces is not defined. That difference can be defined in the  $\mathbb{R}^n$  of a local coordinate system, but the arising definition would depend on the choice of that local chart. There are at least two ideas in differential geometry to circumvent the problem. One exploits the notion of *flow of  $Y$*  defined above and its push forward and it leads to the notion of *Lie derivative*. The other one relies upon the notion of *affine connection* and will be discussed in the next chapter.

**Definition 4.23.** If  $X, Y \in \mathfrak{X}(M)$  for the smooth manifold  $M$  and  $p \in M$ , the **Lie derivative** at  $p \in M$  of  $X$  with respect to  $Y$  is the vector of  $T_pM$

$$(\mathcal{L}_Y X)_p := \lim_{t \rightarrow 0} \frac{1}{t} \left( d\Phi_{-t}^{(Y)} X_{\Phi_t^{(Y)}(p)} - X_p \right) \quad (4.13)$$

■

**Remark 4.24.** We stress that

$$d\Phi_{-t}^{(Y)} X_q \in T_{\Phi_{-t}^{(Y)}(q)}M$$

by definition of differential, so that

$$d\Phi_{-t}^{(Y)} X_{\Phi_t^{(Y)}(p)} \in T_{\Phi_{-t}^{(Y)} \circ \Phi_t^{(Y)}(p)}M = T_pM$$

where we have exploited (4.8) assuming that  $t$  is sufficiently close to  $p$ . As a consequence, (4.13) is well defined as the difference of two vectors in  $T_pM$  appears in the right-hand side circumventing the problem illustrated above the definition. This apparently cumbersome procedure defines an intrinsic notion of derivative of a vector field with respect to another vector field in a fixed point of a manifold. This notion of derivative is physically useful and it enters for instance the theorems relating dynamically conserved quantities and symmetries of a physical system through the various versions of Noether theorem. ■

Let us explicitly compute the above Lie derivative. Since

$$d\Phi_{-t}^{(Y)} X_{\Phi_0^{(Y)}(p)} = X_p$$



due to the definition of the differential and the former in (4.8), another way to write (4.13) is

$$(\mathcal{L}_Y X)_p = \left. \frac{d}{dt} \right|_{t=0} d\Phi_{-t}^{(Y)} X_{\Phi_t^{(Y)}(p)}$$

In other words, in components of a local chart around  $p$  with coordinates  $x^1, \dots, x^n$

$$(\mathcal{L}_Y X)_p^k = \left. \frac{d}{dt} \right|_{t=0} \left( \left. \frac{\partial x_{-t}^k}{\partial x^j} \right|_{(x_t^1, \dots, x_t^n)} X^j(x_t^1, \dots, x_t^n) \right),$$

where  $(x_s^1, \dots, x_s^n)$  are the coordinates of  $\Phi_s^{(Y)}(p)$  when  $p$  has coordinates  $(x^1, \dots, x^n)$  so that, in particular,  $(x_0^1, \dots, x_0^n) = (x^1, \dots, x^n)$ . Computing the derivative – using Schwarz' theorem in the first addend – we have

$$(\mathcal{L}_Y X)_p^k = \left[ \left. \frac{\partial}{\partial x^j} \frac{\partial x_{-t}^k}{\partial t} \right|_{t=0} + \left. \frac{\partial x_{-t}^l}{\partial t} \right|_{t=0} \left. \frac{\partial}{\partial x^l} \frac{\partial x_{-t}^k}{\partial x^j} \right] X^j(x^1, \dots, x^n) + \left. \frac{\partial x^k}{\partial x^j} \frac{\partial X^l}{\partial x^l} \frac{\partial x_t^l}{\partial t} \right|_{t=0},$$

where all functions are evaluated at  $(x^1, \dots, x^n)$ . Since

$$\frac{\partial x^a}{\partial x^b} = \delta_b^a \quad \text{and} \quad \left. \frac{\partial x_t^k}{\partial t} \right|_{t=0} = - \left. \frac{\partial x_{-t}^k}{\partial t} \right|_{t=0} = Y^k,$$

the final result is

$$(\mathcal{L}_Y X)_p^k = Y_p^j \left. \frac{\partial X^k}{\partial x^j} \right|_{\phi(p)} - \left. \frac{\partial Y^k}{\partial x^j} \right|_{\phi(p)} X_p^j. \quad (4.14)$$

Comparing with (3.5), we conclude that the Lie derivative of vector fields and the Lie bracket are the same mathematical object:

$$\mathcal{L}_Y X = [Y, X] = -[X, Y] = -\mathcal{L}_X Y. \quad (4.15)$$

In particular, as already suggested by the notation above, the Lie derivative  $(\mathcal{L}_Y X)_p$  defines a smooth vector field  $\mathcal{L}_Y X \in \mathfrak{X}(M)$  out of  $X, Y \in \mathfrak{X}(M)$  when varying the point  $p \in M$ .

An immediate and useful consequence of (4.15), of the antisymmetry of the Lie bracket, and the Jacobi identity (see Exercise 3.19) is the identity

$$\mathcal{L}_{[X, Y]} = \mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X. \quad (4.16)$$

**Remark 4.25.** If comparing the notion of Lie derivative  $\mathcal{L}_Y X$  written in coordinates as in (4.14) with the notion of directional derivative  $\nabla_Y X$  in  $\mathbb{R}^n$ , we see that the difference just amounts to the last term subtracted in the right-hand side of (4.14). As this term contains the derivative of the components of  $Y$ , we conclude that

$$(\mathcal{L}_{fY} X)_p = f(p)(\mathcal{L}_Y X)_p - X_p(f)Y_p \neq f(p)(\mathcal{L}_Y X)_p$$

in general, differently from

$$(\nabla_{fY}X)_p = f(p)(\nabla_YX)_p$$

where  $f \in D(M)$ . It is possible to repair this feature (which however accounts for the different definition of this type of derivative with respect the standard one in  $\mathbb{R}^n$ ) introducing another notion of derivative of vector fields on manifolds, the covariant one, as we shall see in the next chapter. ■

### 4.3.3 Lie derivative of tensor fields

It is possible to extend the action of the Lie derivative  $\mathcal{L}_Y$  on tensor fields on a smooth manifold  $M$ . This is done by imposing the following requirements.

- (a) If  $T$  is a smooth tensor field on  $M$ , then  $\mathcal{L}_YT$  is a smooth tensor field of the same type as  $T$  – thus it belongs to the same vector space of  $T$  – and the map  $T \mapsto \mathcal{L}_YT$  is linear;
- (b)  $\mathcal{L}_Yf := Y(f)$  for every  $f \in D(M)$ ;
- (c)  $\mathcal{L}_Y$  acts as a derivation with respect to the point-by-point contraction:

$$\mathcal{L}_Y\langle X, \omega \rangle = \langle \mathcal{L}_YX, \omega \rangle + \langle X, \mathcal{L}_Y\omega \rangle$$

for every smooth vector field  $X$  and covariant vector field  $\omega$ ;

- (d)  $\mathcal{L}_Y$  acts as a derivation with respect to the point-by-point tensor product:

$$\mathcal{L}_Y(T \otimes T') = (\mathcal{L}_YT) \otimes T' + T \otimes (\mathcal{L}_YT')$$

where  $T$  and  $T'$  are smooth tensor fields.

Let us illustrate, for instance, how to compute the Lie derivative of a covariant vector field  $\omega$ . Fixing a local chart  $U \ni p \mapsto (x^1, \dots, x^n) \in \mathbb{R}^n$  on  $M$  we have

$$\omega_p = \left\langle \frac{\partial}{\partial x^k} \Big|_p, \omega_p \right\rangle dx^k \Big|_p = (\omega_p)_k dx^k \Big|_p.$$

Now fix  $p \in U$  and take a hat function  $h$  centered on  $p$  and supported in  $U$  and define

$$\frac{\partial}{\partial x^k} \Big|_q := h(q) \frac{\partial}{\partial x^k} \Big|_q.$$

assuming that the left-hand side vanishes outside  $U$ . The left-hand side is a smooth vector field on  $M$  whose components coincide with those of  $\frac{\partial}{\partial x^k} \Big|_q$  in a neighborhood of  $p$ . Taking advantage of (b) and (c), we have in a sufficiently small neighborhood of  $p$ ,

$$Y^j \frac{\partial \omega_k}{\partial x^j} = \mathcal{L}_Y \left\langle \frac{\partial}{\partial x^k}, \omega \right\rangle = \left\langle \mathcal{L}_Y \frac{\partial}{\partial x^k}, \omega \right\rangle + \left\langle \frac{\partial}{\partial x^k}, \mathcal{L}_Y \omega \right\rangle.$$

Since from (4.15),

$$\left( \mathcal{L}_Y \frac{\partial}{\partial x^k} \right)_q = - \frac{\partial Y^h}{\partial x^k} \Big|_{\phi(q)} \frac{\partial}{\partial x^h} \Big|_q$$

we finally find for  $q = p$ ,

$$Y_p^j \frac{\partial \omega_k}{\partial x^j} \Big|_{\phi(p)} + \frac{\partial Y^h}{\partial x^k} \Big|_{\phi(p)} (\omega_p)_h = \left\langle \frac{\partial}{\partial x^k} \Big|_p, (\mathcal{L}_Y \omega)_p \right\rangle .$$

$(\mathcal{L}_Y \omega)_p$  is a covariant vector due to (a) and therefore the right-hand side is the  $k$ -th component in the considered reference frame. In summary

$$(\mathcal{L}_Y \omega)_p = \left[ Y_p^j \frac{\partial \omega_k}{\partial x^j} \Big|_{\phi(p)} + \frac{\partial Y^j}{\partial x^k} \Big|_{\phi(p)} (\omega_p)_j \right] dx^k \Big|_p .$$

At this point it is easy, using (d), to prove the general form of the action of  $\mathcal{L}_Y$  on a tensor field of order  $(r, s)$  (both sides are evaluated at a point  $p$  in the domain of the used local chart)

$$\begin{aligned} (\mathcal{L}_Y T)^{i_1 \dots i_r}_{j_1 \dots j_s} &= Y^k \frac{\partial T^{i_1 \dots i_r}_{j_1 \dots j_s}}{\partial x^k} - T^{k \dots i_r}_{j_1 \dots j_s} \frac{\partial Y^{i_1}}{\partial x^k} + \dots - T^{i_1 \dots i_{r-1} i_k}_{j_1 \dots j_s} \frac{\partial Y^{i_n}}{\partial x^k} \\ &\quad + T^{i_1 \dots i_r}_{k \dots j_s} \frac{\partial Y^k}{\partial x^{j_1}} + \dots + T^{i_1 \dots i_k}_{j_1 \dots j_{s-1} k} \frac{\partial Y^k}{\partial x^{j_s}} . \end{aligned} \quad (4.17)$$

The case of a mixed tensor is analogous.

#### Exercises 4.26.

1. Prove that, with the definition above, if  $f \in D(M)$ ,

$$(\mathcal{L}_Y f)_p := \lim_{t \rightarrow 0} \frac{1}{t} \left( f(\Phi_t^{(Y)}(p)) - f(p) \right) . \quad (4.18)$$

2. Prove that, with the definition above, if  $\omega$  is a covariant vector field,

$$(\mathcal{L}_Y \omega)_p := \lim_{t \rightarrow 0} \frac{1}{t} \left( \Phi_t^{(Y)*} \omega_{\Phi_t^{(Y)}(p)} - \omega_p \right) . \quad (4.19)$$

3. Extend (4.13), (4.18), (4.19) to the case of a generic tensor field  $S$  of type  $(r, s)$ ,

$$(\mathcal{L}_Y S)_p := \lim_{t \rightarrow 0} \frac{1}{t} \left( \underbrace{d\Phi_{-t}^{(Y)} \otimes \dots \otimes d\Phi_{-t}^{(Y)}}_{r \text{ times}} \otimes \underbrace{\Phi_t^{(Y)*} \otimes \dots \otimes \Phi_t^{(Y)*}}_{s \text{ times}} S_{\Phi_t^{(Y)}(p)} - S_p \right) . \quad (4.20)$$

where  $\otimes$  is the point-by-point tensor product of operators.

4. Proves that (4.16) extends to the action of the Lie derivative on every tensor field.

*(Hint. Prove the thesis according to the following steps. (1) Use the axioms (1)-(d) observing that the identity is trivial when the Lie derivative acts on  $f \in D(M)$ , then (2) it extends it to 1-forms using (c) and finally (3) use (d) to extend the results to all tensor fields).*

5. Prove that, as a consequence of (4.20), if  $Y \in \mathfrak{X}(M)$ , then

$$\omega \in \Omega^k(M) \quad \text{entails} \quad \mathcal{L}_Y \omega \in \Omega^k(M) .$$

5. Prove the **Cartan magic formula**, also known as **Cartan homotopy formula**,

$$\mathcal{L}_X \omega = i_X d\omega + di_X \omega , \quad (4.21)$$

where

- (a)  $\omega \in \Omega^k(M)$ ,
- (b)  $d$  is the *exterior derivative* introduced in Section 3.2.3,
- (c)  $i_X : \Omega^k(M) \rightarrow \Omega^{k-1}(M)$  is the point-by-point **interior product** (see [Mor20] where it is denoted by  $X \lrcorner$ ) for a given  $X \in \mathfrak{X}(M)$ , which can be equivalently defined as

$$(i_{X_p} \omega_p)(X_{p1}, \dots, X_{pk-1}) := \omega(X_p, X_{p1}, \dots, X_{pk-1})$$

for every  $\omega \in \Omega^k(M)$  and  $X, X_1, \dots, X_{k-1} \in \mathfrak{X}(M)$ .

#### 4.3.4 Commuting vector fields and associated commuting flows

Let us consider  $X, Y \in \mathfrak{X}(M)$  for a smooth manifold  $M$ , our aim is to investigate the interplay of the commutativity of the vectors at  $p$ :

$$[X, Y]_p = 0$$

and the commutativity of the corresponding flows:

$$(\Phi_t^{(X)} \circ \Phi_u^{(Y)})(p) = (\Phi_t^{(Y)} \circ \Phi_u^{(X)})(p) .$$

We have the following two important results.

**Proposition 4.27.** *Let  $M$  be a smooth manifold and  $X, Y \in \mathfrak{X}(M)$ . Suppose that, for a fixed  $p \in M$  and for a pair of open intervals  $I, I' \subset \mathbb{R}$  containing 0, it holds*

$$(\Phi_t^{(X)} \circ \Phi_u^{(Y)})(p) = (\Phi_t^{(Y)} \circ \Phi_u^{(X)})(p) \quad \text{if } (t, u) \in I \times I' .$$

Then

$$[X, Y]_p = 0 ,$$

that is tantamount to saying

$$(\mathcal{L}_Y X)_p = (\mathcal{L}_X Y)_p = 0 .$$

**Proof.** If we compute the  $u$ -partial derivative at  $u = 0$  of both sides in a local chart around  $p$ , and next the  $t$ -partial derivative at  $t = 0$ , we get

$$\sum_{k=1}^n \frac{\partial X^i}{\partial x^k} Y_p^k = \sum_{k=1}^n \frac{\partial^2 (\Phi_u^{(Y)})^i}{\partial u \partial x^k} \Big|_{u=0} X_p^k .$$

Schwarz' theorem implies that we can interchange the derivatives in the right-hand side producing

$$\sum_{k=1}^n \frac{\partial X^i}{\partial x^k} Y_p^k = \sum_{k=1}^n \frac{\partial Y^i}{\partial x^k} X_p^k ,$$

that mens

$$[X, Y]_p = 0 .$$

The last statement is an obvious consequence of (4.15).  $\square$ .

The converse fact also holds also if it is a bit more difficult to prove.

**Theorem 4.28.** *Let  $M$  be a smooth manifold and  $X, Y \in \mathfrak{X}(M)$  such that*

$$[X, Y] = 0 .$$

*Suppose that, for  $p \in M$  and for a pair of open intervals  $I, I' \subset \mathbb{R}$  containing 0, the function  $(\Phi_u^{(Y)} \circ \Phi_t^{(X)})(p)$  is well defined when  $(t, u) \in I \times I'$ , then also  $(\Phi_t^{(X)} \circ \Phi_u^{(Y)})(p)$  is well defined for  $(t, u) \in I \times I'$  and it holds*

$$(\Phi_u^{(Y)} \circ \Phi_t^{(X)})(p) = (\Phi_t^{(X)} \circ \Phi_u^{(Y)})(p) .$$

**Proof.** For a fixed  $t \in I$ , let us consider the function taking values in  $TM$  (notice that  $p \in M$  is fixed)

$$I' \ni u \mapsto \psi_{t,p}(u) := \frac{\partial}{\partial t} (\Phi_u^{(Y)} \circ \Phi_t^{(X)})(p) - X_{(\Phi_u^{(Y)} \circ \Phi_t^{(X)})(p)} \in T_{(\Phi_u^{(Y)} \circ \Phi_t^{(X)})(p)} M .$$

We want to prove that this function is constant. Computing the  $u$ -derivative at  $u = u_0$  (without writing this specification explicitly) and taking Schwarz' theorem into account, in a local coordinate chart defined around  $(\Phi_{u_0}^{(Y)} \circ \Phi_t^{(X)})(p) \in M$  we have

$$\begin{aligned} \frac{d\psi_{t,p}}{du} &= \frac{\partial^2}{\partial t \partial u} (\Phi_u^{(Y)} \circ \Phi_t^{(X)})(p) - \frac{\partial}{\partial u} X_{(\Phi_u^{(Y)} \circ \Phi_t^{(X)})(p)} \\ &= \frac{\partial}{\partial t} Y_{\Phi_t^{(X)}(p)}^i \frac{\partial}{\partial x^i} - \frac{\partial X^i}{\partial x^k} Y_{\Phi_t^{(X)}(p)}^k \frac{\partial}{\partial x^i} \end{aligned}$$

$$= \frac{\partial Y^i}{\partial x^k} X_{\Phi_t^{(X)}(p)}^k \frac{\partial}{\partial x^i} - \frac{\partial X^i}{\partial x^k} Y_{\Phi_t^{(X)}(p)}^k \frac{\partial}{\partial x^i} = [X, Y]_{\Phi_t^{(X)}(p)}^i \frac{\partial}{\partial x^i} = 0.$$

We conclude that

$$\psi_{t,p}(u) = \psi_{t,p}(0) = \frac{\partial}{\partial t} \Phi_t^{(X)}(p) - X_{\phi_t^{(\mathbf{x})}(p)} = \frac{d}{dt} \Phi_t^{(X)}(p) - X_{\phi_t^{(\mathbf{x})}(p)} = 0.$$

According to the definition of  $\psi_{t,p}$ , we have established that

$$\frac{\partial}{\partial t} (\Phi_u^{(Y)} \circ \Phi_t^{(X)})(p) = X_{(\phi_u^{(Y)} \circ \phi_t^{(X)})(p)}.$$

The found identity implies that, keeping  $u \in I'$  fixed

$$I \ni t \mapsto \gamma_{u,p}(t) := (\Phi_u^{(Y)} \circ \Phi_t^{(X)})(p)$$

satisfies the differential equation

$$\frac{d}{dt} \gamma_{u,p}(t) = X_{\gamma_{u,p}(t)} \quad \text{with initial condition} \quad \gamma_{u,p}(0) = \Phi_u^{(Y)}(p).$$

The uniqueness theorem implies

$$\gamma_{u,p}(t) = \Phi_t^{(X)}(\Phi_u^{(Y)}(p)).$$

Comparing it with the definition of  $\gamma_{u,p}$ , we conclude that

$$(\Phi_u^{(Y)} \circ \Phi_t^{(X)})(p) = (\Phi_t^{(X)} \circ \Phi_u^{(Y)})(p)$$

ending the proof. □

**Remark 4.29.** Many physical systems are described on differential manifolds and their temporal evolution is represented by the curves tangent to a certain vector field  $Z$  called the **dynamical vector**. This is the case both in Lagrangian and in Hamiltonian mechanics where, respectively,  $M$  is the *spacetime of kinetic states* or the *spacetime of phases*. Locally, or also globally if the field is complete, the evolution is therefore represented by the flow  $\Phi_t^{(H)}$ .

Another vector field  $X$  defines a **continuous dynamical symmetry** of the system if, by definition,  $[X, Z] = 0$  everywhere. As a consequence of the proved theorem, this implies that where both sides are defined

$$\Phi_t^{(Z)} \circ \Phi_u^{(X)}(p) = \Phi_u^{(X)} \circ \Phi_t^{(Z)}(p).$$

The identity above says that, if transforming at each instant the evolution (where  $I_p$  is a suitable open interval containing 0)

$$I_p \ni t \mapsto \Phi_t^{(Z)}(p) \in M$$

of the initial state  $p \in M$  with the transformation  $\Phi_u^{(X)}$ , the resulting curve

$$I_p \ni t \mapsto \Phi_u^{(X)} \circ \Phi_t^{(Z)}(p) \in M$$

is still a possible evolution (with different initial state), since it equals

$$I_p \ni t \mapsto \Phi_t^{(Z)} \circ \Phi_u^{(X)}(p) \in M.$$

This is the most elementary notion of dynamical symmetry. We stress that it only relies only on the equations of the dynamics and not on more sophisticated notions as the Lagrangian or Hamiltonian function. ■

#### 4.3.5 Frobenius theorem

Lie bracket are of relevance also dealing with the problem of the *integrability of smooth distributions* as we briefly discuss here. Let us start by introducing the notion of *smooth distribution*.

**Definition 4.30.** (**Smooth distribution.**) If  $M$  is a smooth manifold of dimension  $n$ , a **smooth distribution of rank  $k \leq n$**  in  $M$  is an assignment of  $k$ -dimensional subspaces

$$W : M \ni p \mapsto W_p \subset T_p M$$

which is smooth in the following sense. For every fixed  $p \in M$ , there are  $k$  vector fields  $X_{(1)}, \dots, X_{(k)} \in \mathfrak{X}(M)$  such that, in an open set  $U \ni p$ , they define a basis of  $W_q$  when evaluated at  $q \in U$ . ■

The simplest case of a smooth distribution defined in the domain  $U$  of a local chart  $\phi : U \ni q \mapsto (x^1, \dots, x^n) \in \mathbb{R}^n$  on  $M$  viewed as a smooth manifold in its own right, is provided by

$$W_q := \text{Span} \left( \frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^k} \right), \quad q \in U.$$

It is evident that, in this case,  $U$  is foliated by the  $k$ -dimensional embedded submanifolds  $\Sigma_{x^{k+1}, \dots, x^n}$  at constant  $x^{k+1}, \dots, x^n$ .  $U$  is the union of these disjoint surfaces so that, every  $p \in U$  is contained in one of them. The interesting fact is that

$$T_q \Sigma_{x^{k+1}, \dots, x^n} = W_q \quad \text{if } q \in \Sigma_{x^{k+1}, \dots, x^n},$$

so that these surfaces are also everywhere tangent to the smooth distribution.

We would like to extend these properties to more general smooth distributions. In general, it is impossible to find *embedded* submanifolds tangent to a smooth distribution as above, at least when considering maximal integral manifolds (see Example 4.34 and Remark 4.35). However if we stick to *immersed* submanifolds, a crucial result due to Frobenius provides necessary and sufficient conditions.

To go on, observe that in the example above  $[\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}] = 0$  and, more generally, if two smooth vector fields  $X, Y$  are tangent to an embedded (or simply immersed) submanifold  $\Sigma$  at  $p \in \Sigma$ , by direct inspection exploiting a suitable coordinate system around  $p$  and adapted to  $\Sigma$  (as in (b) of 4.5) one sees that  $[X, Y]_p \in T_p \Sigma$ . This condition which seems relevant (and it is) deserves a definition.

**Definition 4.31.** (**Involutive smooth distribution.**) A smooth distribution  $W$  of rank  $k \leq \dim(M)$  in the smooth manifold  $M$  is said to be **involutive** if

$$[X, Y]_p \in W_p \quad \forall p \in M,$$

if  $X, Y \in \mathfrak{X}(M)$  are such that

$$X_p, Y_p \in W_p \quad \forall p \in M.$$

■

To state the theorem we need the crucial definition.

**Definition 4.32.** (**Integrable smooth distribution.**) A smooth distribution  $W$  of rank  $k \leq \dim(M)$  in the smooth manifold  $M$  is said to be **integrable** if, for every  $p \in M$ , there exist an immersed submanifold  $N^{(p)} \subset M$  such that

$$di_q N^{(p)} = W_q \quad \text{for every } q \in N^{(p)},$$

where  $i : N \rightarrow M$  is the inclusion map.  $N^{(p)}$  is called **integral manifold through  $p$** . ■

We are in a position to state the theorem by Frobenius in both local and global version.

**Theorem 4.33.** (**Frobenius Theorem.**) *Let  $M$  be a smooth manifold of dimension  $n$  and  $W$  a smooth distribution of rank  $k \leq n$  therein. The following facts are valid.*

- (a)  *$W$  is integrable if and only if it is involutive.*
- (b) *If  $W$  is involutive and  $p \in M$ , then there is a local chart  $\phi : U \ni q \mapsto (x^1, \dots, x^n)$  with  $\phi(U) = (-\delta, \delta)^n$  for some  $\delta > 0$  and  $\phi(p) = (0, \dots, 0)$ , such that*

(1) *the slices*

$$\Sigma_{x^{k+1}, \dots, x^n} := \{\phi(x^1, \dots, x^n) \mid x^{k+1}, \dots, x^n \text{ constant}\}$$

*with the smooth structure induced from  $M$  are integral manifolds so that, in particular,*

$$\text{Span} \left( \frac{\partial}{\partial x^1} \Big|_q, \dots, \frac{\partial}{\partial x^k} \Big|_q \right) = W_q \quad \text{if } q \in U,$$

- (2) *if  $N^{(p)} \subset U$  is a connected integral manifold through  $p$ , then  $N^{(p)}$  coincides (as smooth manifold) with one of those slices.*



**Proof** See [War83].  $\square$

**Examples 4.34.** Let  $\mathbb{T}^2$  be the torus obtained by identifying 0 and 1 along both the  $x$  axis and the  $y$  axis of  $\mathbb{R}^2$ . Consider the smooth vector fields in  $\mathbb{T}^2 \times \mathbb{R}_z$  giving rise to a smooth distribution of rank 2.

$$X_{(1)} := \sqrt{2} \frac{\partial}{\partial x} + \frac{\partial}{\partial y}, \quad X_{(2)} = \frac{\partial}{\partial z}$$

(the former smoothly extended on the identification lines). Evidently they satisfy the hypotheses of Frobenius theorem since they commute. However, their maximal integral manifolds are only immersed. Notice in fact that the maximal integral lines of  $X$  are dense in  $\mathbb{T}^2$ , this fact prevent for having the topology induced from  $\mathbb{T}^2 \times \mathbb{R}_z$  on the integral manifolds which therefore cannot be embedded submanifolds. For every  $p \in \mathbb{T}^2 \times \mathbb{R}_z$  we can however cut the maximal integral manifold around  $p$ , in the unique local slice passing through  $p$ , to obtain an integral manifold through  $p$  which is also an embedded submanifold.

**Remark 4.35.** The result in the example above is general. According to (b) in Theorem 4.33 we can always select a local slice passing through  $p$ . Considered alone, this slice is an embedded submanifold (see Proposition 4.8) tangent to the distribution.

## Chapter 5

# (Pseudo) Riemannian manifolds and related metric tools

This chapter is devoted to introduce some basic notions of the theory of (pseudo) Riemannian manifolds. The overall goal – which drives the choice of the treated subjects – is exploiting those notions in discussing relativistic theories. We use in the text the term *pseudo Riemannian manifold*. These manifolds are also known in the literature as *semi Riemannian manifolds* [O’Ne83, BEE96, ?].

### 5.1 (Pseudo) Riemannian manifolds

We start by discussing the notion of (pseudo) Riemannian manifold and its basic features.

#### 5.1.1 Types of manifolds equipped with metrics

**Definition 5.1.** ((pseudo) Riemannian Manifolds.) Let  $M$  be a connected smooth manifold of dimension  $n$  equipped with a symmetric  $(0, 2)$  smooth tensor field  $\mathbf{g}$  which defines a non-degenerate quadratic form  $\mathbf{g}_p(, ) : T_p M \times T_p M \rightarrow \mathbb{R}$  with constant signature  $(r, s)$ .

- (a)  $(M, \mathbf{g})$  is called **Riemannian manifold** if the signature of  $\mathbf{g}$  is  $(n, 0)$  – the canonical form of the metric therefore reading  $(1, +1, \dots, +1)$  – and thus  $\mathbf{g}_p$  is a scalar product. In this case  $\mathbf{g}$  is called **metric tensor** of  $M$ .
- (b)  $(M, \mathbf{g})$  is called **pseudo Riemannian manifold** if the signature of  $\mathbf{g}$  is  $(r, s)$  with  $rs \neq 0$  – the canonical form of the metric therefore reading  $\underbrace{(-1, \dots, -1)}_{r \text{ times}}, \underbrace{(+1, \dots, +1)}_{s \text{ times}}$  – and thus  $\mathbf{g}_p$  is a pseudo scalar product. In this case  $\mathbf{g}$  is called **pseudo metric tensor** of  $M$ .
- (c)  $(M, \mathbf{g})$  is called **Lorentzian manifold** if the signature of  $\mathbf{g}$  is  $(1, n - 1)$  – the canonical form of the metric therefore reading  $(-1, +1, \dots, +1)$  – and thus  $\mathbf{g}_p$  is a Lorentzian pseudo scalar product the canonical form of the metric reads. In this case  $\mathbf{g}$  is called **Lorentzian**

pseudo metric tensor of  $M$ . ■

**Remark 5.2.** *In the rest of the paper we shall often omit the term pseudo and Lorentzian when the meaning of the used terms is clear from the context. The (pseudo, Lorentzian) metric tensor will be often called simply the “metric”.*

### 5.1.2 Length of curves

We start by extending the notion of *smooth curve* in  $M$  (Definition 2.16).

**Definition 5.3.** (**Piecewise smooth curve.**) Let  $M$  be a smooth manifold. A **piecewise smooth curve** in  $M$  is a *continuous* map  $\gamma : I \rightarrow M$  defined on the interval  $I$ , possibly including one or both endpoints, such that

- (a)  $\gamma$  is smooth in  $I$  except for a finite number of points (possibly including one or both endpoints);
- (b) the limits of the derivatives of every order towards those singular points exist and are finite, but they can be different depending on the side of the considered point. ■

**Proposition 5.4.** *A connected smooth manifold  $M$  is connected by piecewise smooth curves. In other words, if  $p, q \in M$  there is a (continuous!) piecewise smooth curve  $\gamma : [a, b] \rightarrow M$  with  $\gamma(a) = p$  and  $\gamma(b) = q$ .*

**Proof.** Let  $A \subset M$  the set of points which are connected to a given  $p \in M$  with some piecewise smooth path. It is easy to prove that  $A$  is open, because if  $q \in A$  and  $\gamma$  connects  $p$  and  $q$ , dealing with a local coordinate chart around  $q$ , every point  $q'$  in an open coordinate ball centered on  $q$  can be connected to  $p$  by adding a smooth path from  $q$  to  $q'$ . With this procedure one has a piecewise smooth path from  $p$  to  $q'$ . With an analogous argument, the set  $B \subset M$  which are not connected by piecewise smooth paths to  $p$  is proved to be open as well. In summary,  $M = A \cup B$  where  $A$  and  $B$  are open disjoint sets. Since  $M$  is connected, it must be either  $M = A$  and  $B = \emptyset$  or  $M = B$  and  $A = \emptyset$ . The second possibility is not allowed as  $p \in A$ , hence  $M = A$ .  $\square$

**Definition 5.5.** (**Length of a curve.**) Let  $\gamma : I \ni t \mapsto \gamma(t) \in M$  be a piecewise smooth curve in a (pseudo) Riemannian manifold  $(M, \mathbf{g})$ , where  $I \subset \mathbb{R}$  is a bounded interval. The **length** of  $\gamma$  is, using the notation  $\mathbf{g}(\gamma'(t), \gamma'(t)) := \mathbf{g}_{\gamma(t)}(\gamma'(t), \gamma'(t))$ ,

$$L_{\mathbf{g}}(\gamma) = \int_I \sqrt{|\mathbf{g}(\gamma'(t), \gamma'(t))|} dt, \quad (5.1)$$

where the tangent vector  $\gamma'(t)$  in the integrand is defined as the limit towards its discontinuity points where  $\gamma$  is not smooth. ■

**Remark 5.6.**

(1) It is easy to prove that the  $L_{\mathbf{g}}(\gamma)$  is re-parametrization invariant: if  $u = u(t)$  is a smooth function with  $\frac{du}{dt} > 0$  on  $I$ , then  $L_{\mathbf{g}}(\gamma_1) = L_{\mathbf{g}}(\gamma)$  where  $\gamma_1(u) = \gamma(t(u))$  defined on  $u(I)$ .

(2) Let  $(M, \mathbf{g})$  be explicitly *Riemannian*.

$$d_{\mathbf{g}}(p, q) := \inf \{ L_{\mathbf{g}}(\gamma) \mid \gamma : [a, b] \rightarrow M, \gamma \text{ piecewise smooth}, \gamma(a) = p, \gamma(b) = q \} \quad (5.2)$$

is a *distance* function in  $M$  so that  $(M, d_{\mathbf{g}})$  is a *metric space*. The proof is quite elementary. What it is not trivial is the fact that [KoNo96] the associated metric topology coincides with the topology initially given on  $M$ . ■

### 5.1.3 Local and global flatness

A physically relevant property of a (pseudo) Riemannian manifold concerns its *flatness*.

**Definition 5.7.** (**Flatness.**) A  $n$ -dimensional (pseudo) Riemannian manifold  $(M, \mathbf{g})$  with generic signature  $(r, s)$  is said to be **locally flat** if, for every  $p \in M$ , there is a local chart  $U \ni q \mapsto (x^1(q), \dots, x^n(q)) \in \mathbb{R}^n$  with  $p \in U$ , where  $\mathbf{g}$  takes its **canonical form** in components:

$$\mathbf{g}_q = (g_q)_{ij} dx^i|_q \otimes dx^j|_q \quad \text{where} \quad (g_q)_{ij} = \text{diag}(\underbrace{-1, \dots, -1}_{r \text{ times}}, \underbrace{+1, \dots, +1}_{s \text{ times}}), \quad q \in U.$$

In other words all the bases  $\{\frac{\partial}{\partial x^k}|_q\}_{k=1, \dots, n}$ ,  $q \in U$ , are **canonical (pseudo) orthonormal bases** with respect to the pseudo metric tensor.

$(M, \mathbf{g})$  is said to be **globally flat** if there is a global chart as above. ■

In other words, a (pseudo) Riemannian manifold is locally flat if admits an atlas made of **canonical local charts** which are, by definition, charts where the components of the metric take the  $\text{diag}(\underbrace{-1, \dots, -1}_{r \text{ times}}, \underbrace{+1, \dots, +1}_{s \text{ times}})$  constantly. If that atlas can be reduced to a single chart, the manifold is globally flat.

### Examples 5.8.

1. Any  $n$ -dimensional **(pseudo) Euclidean space**  $\mathbb{E}^n$  with signature  $(m, p)$ , i.e, a  $n$ -dimensional affine space  $\mathbb{A}^n$  whose vector space  $V$  is equipped with a (pseudo) scalar product  $(|)$  with signature  $(m, p)$  is a (pseudo) Riemannian manifold which is globally flat. As special cases we have

- (a)  $\mathbb{R}^n$  equipped with the standard metric,
- (b) the **Minkowski spacetime** of dimension  $n$ , denote by  $\mathbb{M}^n$  which correspond to the Lorentzian choice  $(m, p) = (1, n - 1)$ .

To illustrate how the (pseudo) Riemannian manifold structure is constructed, first of all we notice that the presence of a (pseudo) scalar product in  $V$  singles out a class of Cartesian coordinates systems called **(pseudo) orthonormal Cartesian coordinates systems** called **Minkowskian coordinates** in  $\mathbb{M}^n$ . These are the Cartesian coordinate systems built up by starting from any origin  $O \in \mathbb{A}^n$  and any (pseudo) orthonormal basis in  $V$ . Then consider the isomorphism  $\chi_p : V \rightarrow T_p M$  defined in the remark after proposition 3.11 above. The (pseudo) scalar product  $(\cdot | \cdot)$  on  $V$  can be exported in each  $T_p \mathbb{A}^n$  by defining  $\mathbf{g}_p(u, v) := (\chi_p^{-1} u | \chi_p^{-1} v)$  for all  $u, v \in T_p \mathbb{A}^n$ . By this way the bases  $\{\frac{\partial}{\partial x^i}|_p\}_{i=1, \dots, n}$  associated with (pseudo) orthonormal Cartesian coordinates turn out to be (pseudo) orthonormal. Hence the (pseudo) Euclidean space  $\mathbb{E}^n$ , i.e.,  $\mathbb{A}^n$  equipped with a (pseudo) scalar product as above, is a globally flat (pseudo) Riemannian manifold.

**2.** Consider the cylinder  $C$  in  $\mathbb{E}^3$ . Referring to an orthonormal Cartesian coordinate system  $x, y, z$  in  $\mathbb{E}^3$ , we further assume that  $C$  is the set corresponding to triples or reals  $\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = 1\}$ . That set is a smooth manifold when equipped with the natural differentiable structure induced by  $\mathbb{E}^3$  as follows. First of all define the topology on  $C$  as the topology induced by that of  $\mathbb{E}^3$ .  $C$  turns out to be a topological manifold of dimension 2. Let us pass to equip  $C$  with a suitable differential structure induced by that of  $\mathbb{E}^3$ . If  $p \in C$ , consider a local coordinate system on  $C$ ,  $(\theta, z)$  with  $\theta \in ]0, \pi[$ ,  $z \in \mathbb{R}$  obtained by restriction of usual cylindric coordinates in  $\mathbb{E}^3$   $(r, \theta, z)$  to the set  $r = 1$ . This coordinate system has to be chosen (by rotating the origin of the angular coordinate) in such a way that  $p \equiv (r = 1, \theta = \pi/2, z = z_p)$ . There is such a coordinate system on  $C$  for any fixed point  $p \in C$ . Notice that it is not possible to extend one of these coordinate frame to cover the whole manifold  $C$  (why?). Nevertheless the class of these coordinate system gives rise to an atlas of  $C$  and, in turn, it provided a differentiable structure for  $C$ . As we shall see shortly in the general case, but this is clear from a synthetic geometrical point of view, each vector tangent at  $C$  in a point  $p$  can be seen as a vector in  $\mathbb{E}^3$  and thus the scalar product of vectors  $u, v \in T_p C$  makes sense. By consequence there is a natural metric on  $C$  induced by the metric on  $\mathbb{E}^3$ . The Riemannian manifold  $C$  endowed with that metric is locally flat because in coordinates  $(\theta, z)$ , the metric is diagonal everywhere with unique eigenvalue 1. It is possible to show that there is no global canonical coordinates on  $C$ . The cylinder is locally flat but not globally flat.

**3.** In *Einstein's General Theory of Relativity*, the *spacetime* is a four-dimensional Lorentzian manifold  $M^4$ . Hence it is equipped with a pseudo-metric  $\mathbf{g} = g_{ij} dx^i \otimes dx^j$  with hyperbolic signature  $(1, 3)$ , i.e. the canonical form of the metric reads  $(-1, +1, +1, +1)$  (this holds true if one uses units to measure length such that the speed of the light is  $c = 1$ ). The points of the manifolds are called **events**. If the spacetime is *globally* flat and it is an affine four dimensional space, it is called *Minkowski Spacetime*,  $\mathbb{M}^4$  defined in (1) above. That is the spacetime of *Special Relativity Theory* (see [Mor20]).

#### 5.1.4 Existence of Riemannian and Lorentzian metrics

Every smooth manifold can be equipped with a *Riemannian* metric. The result is not generally valid if considering different signatures. This result is a straightforward consequence of the

existence of a smooth partition of unity (see Section 2.3.2). Thus, in particular, it cannot be extended to the analytic case.

**Theorem 5.9.** *If  $M$  is a (connected) smooth manifold, it is possible to define a Riemannian metric  $\mathbf{g}$  on  $M$ .*

**Proof.** Consider a covering of  $M$ ,  $\{U_i\}_{i \in I}$ , made of (open by definition) coordinate domains whose closures are compact. Then, using paracompactness, refine the covering to a locally finite covering  $\mathcal{C} = \{V_j\}_{j \in J}$ . The closures of the sets  $V_j$  are compact since they are closed sets included in compact sets  $\overline{U_{i(j)}}$ . By construction each  $V_j$  admits local coordinates  $\phi_j : V_j \rightarrow \mathbb{R}^n$ . For every  $j \in J$  define, in the bases associated with the coordinates, a Riemannian metric  $\mathbf{g}_j$  whose components are constants in the considered chart, e.g.,  $(\mathbf{g}_j)_{hk} = \delta_{hk}$ . If  $\{h_j\}_{j \in J}$  is a partition of unity subordinate to  $\mathcal{C}$  (see Theorem 2.26),  $\mathbf{g}_p := \sum_{j \in J} h_j(p)(\mathbf{g}_j)_p$  is a well-defined symmetric form at each point which is also smooth when varying  $p$  (the sum is finite in a neighborhood of  $p$  so we can compute the derivatives at every order). To conclude observe that  $\mathbf{g}_p(v, v) = \sum_{j \in J} h_j(p)(\mathbf{g}_j)_p(v, v) \geq 0$  because the right hand side is the sum of non-negative reals. Eventually,  $\mathbf{g}_p(v, v) = 0$  implies  $v = 0$  because, as  $\sum_{j \in J} h_j(p) = 1$ , and  $h_j(p) \geq 0$ , there must be some  $j_0 \in J$  with  $h_{j_0}(p) > 0$ . As a consequence,  $0 = \mathbf{g}_p(v, v) := \sum_{j \in J} h_j(p)(\mathbf{g}_j)_p(v, v)$  implies  $(\mathbf{g}_{j_0})_p(v, v) = 0$  which, in turn, yields  $v = 0$ .  $\square$

A sufficient condition for the existence of a Lorentzian metric on a smooth manifolds is the following one.

**Theorem 5.10.** *Let  $M$  a (connected) smooth manifold, it is possible to define a Riemannian metric  $\mathbf{g}$  on  $M$ . If there is  $V \in \mathfrak{X}(M)$  with  $V_p \neq 0$  for every  $p \in M$ , then it is possible to define a Lorentzian metric on  $M$ .*

**Proof.** Fix a Riemannian metric  $\mathbf{g}^{(E)}$  of  $M$  according to Theorem 5.9 and define the smooth unit vector field  $M \ni p \mapsto E_p := (\mathbf{g}_p^{(E)}(V_p, V_p))^{-1/2} V_p \in T_p M$ . Denoting by  $T_p M \ni v_p \mapsto v_p^\flat := \mathbf{g}^{(E)}(v_p, \cdot) \in T_p^* M$  the standard isomorphism between  $T_p^* M$  and  $T_p M$  induced by the metric, the  $(0, 2)$  smooth tensor field defined point-by point by

$$\mathbf{g}_p := \mathbf{g}_p^{(E)} - 2E_p^\flat \otimes E_p^\flat$$

is immediately proved to be a Lorentzian metric just completing  $E_p$  to an orthonormal basis of  $\mathbf{g}_p^{(E)}$  and observing that in this basis  $\mathbf{g}_p$  is represented by the matrix  $\text{diag}(-1, +1, \dots, +1)$ .  $\square$

**Remark 5.11.** Notice that  $E$  above is a non-vanishing timelike vector. The condition in the hypothesis of the stated theorem actually assures the existence of a Lorentzian metric admitting a time orientation as we shall discuss later. *Vice versa*, if a Lorentzian metric admits a time orientation then a non vanishing smooth vector necessarily exists since it defines the time orientation.  $\square$

## 5.2 Induced metrics and related notions

This section is devoted to describe how the metric of a (pseudo) Riemannian manifold  $M$  induces (possibly degenerated) metrics on other smooth manifolds which are related to  $M$  through a smooth immersion.

### 5.2.1 Metric induction machinery

Let  $M$  be a (pseudo) Riemannian manifold with metric tensor  $\mathbf{g}$ . If  $N$  is another smooth manifold – possibly with the same dimension as  $M$  – and we are given a smooth (not necessarily injective) immersion  $i : N \rightarrow M$ . Then it is possible to induce on  $N$  a covariant symmetric smooth tensor field  $\mathbf{g}^{(N)}$  associated to  $\mathbf{g}$ . Indeed, taking advantage of the differential  $di_p : T_p N \rightarrow T_p M$ , we can define the bilinear symmetric form in  $\mathbf{g}^{(N)} : T_p N \times T_p N \rightarrow \mathbb{R}$ ,

$$\mathbf{g}^{(N)}(x, y) := \mathbf{g}(di_p x, di_p y) .$$

Varying  $p \in N$  and assuming that  $x = X_p, y = Y_p$  where  $U$  and  $V$  are smooth vector fields in  $N$ , one sees that the map  $N \ni p \mapsto \mathbf{g}^{(N)}(X_p, Y_p)$  must be differentiable because it is composition of differentiable functions. We conclude that  $N \ni p \mapsto \mathbf{g}_p^{(N)}$  define a covariant symmetric smooth tensor field on  $N$ .

**Remark 5.12.**

(1) If  $N$  is connected and  $(M, \mathbf{g})$  is properly Riemannian, then  $(N, \mathbf{g}^{(N)})$  is necessarily a Riemannian manifold as well. In fact,  $\mathbf{g}^{(N)}(u, v) = \mathbf{g}(di_p u, di_p v) \geq 0$  and  $0 = \mathbf{g}^{(N)}(u, u) = \mathbf{g}(di_p u, di_p u)$  implies  $di_p u = 0$ , so that  $u = 0$  since  $di_p$  is injective.

(2) Evidently, if  $i_q^* : T_q^* N \rightarrow T_p^* M$  with  $q = i(p)$  is the pull back operator, then

$$\mathbf{g}_p^{(N)} = i^* \otimes i^* \mathbf{g}_q .$$

This formula is more often written

$$\mathbf{g}_p^{(N)} = i^* \mathbf{g}_q .$$

interpreting  $i_q^*$  as the pull back of covariant  $(0, 2)$  tensors. ■

### 5.2.2 Induced metric on submanifolds

We specialize the previous definition to the case where the immersion is a smooth embedding  $i : N \hookrightarrow M$  provided by the canonical inclusion.

**Definition 5.13.** Let  $M$  be a (pseudo) Riemannian manifold with metric tensor  $\mathbf{g}$  and  $i : N \rightarrow M$  an embedding. The covariant symmetric differentiable tensor field on  $N$ ,  $\mathbf{g}^{(N)}$ , defined by

$$\mathbf{g}^{(N)}(x, y) := \mathbf{g}(di_p x, di_p y) \quad \text{for all } p \in N \text{ and } x, y \in T_p N$$

is called the **metric induced on  $N$  from  $M$** , even if it is not a (pseudo) metric in general (!). If  $N \subset M$  is connected,  $i : N \hookrightarrow M$  is the canonical inclusion and  $\mathbf{g}^{(N)}$  is not degenerate with constant signature, then the (pseudo) Riemannian manifold  $(N, \mathbf{g}^{(N)})$  is called **(pseudo) Riemannian submanifold** of  $M$ . ■

**Remark 5.14.** We stress that, in general,  $\mathbf{g}^{(N)}$  is not a (pseudo) metric on  $N$  because there are no guarantee for it being nowhere non-degenerate and also its signature could change. Nevertheless, according to (1) in Remark 5.12, if  $(M, \mathbf{g})$  is properly *Riemannian*, then  $\mathbf{g}^{(N)}$  is necessarily positive definite by construction. In that case, when  $N \subset M$ ,  $(N, \mathbf{g}^{(N)})$  is a Riemannian submanifold of  $M$ , assuming that  $N$  is connected. ■

What is the coordinate form of  $\mathbf{g}^{(N)}$ ? Let us assume in general that  $i : N \hookrightarrow M$  is the inclusion map which furthermore defines a smooth embedding. Fix  $p \in N$ , a local chart in  $N$ ,  $(U, \phi)$  with  $p \in U$  and another local chart in  $M$ ,  $(V, \psi)$  with  $p \in V$  once again. Use the notation  $\phi : q \mapsto (y^1(q), \dots, y^n(q))$  and  $\psi : r \mapsto (x^1(r), \dots, x^m(r))$ . We have therein

$$\tilde{i} := \psi \circ i \circ \phi^{-1} : (y^1, \dots, y^n) \mapsto (x^1(y^1, \dots, y^n), \dots, x^m(y^1, \dots, y^n))$$

and

$$\mathbf{g} = g_{ij} dx^i \otimes dx^j \quad \text{and} \quad \mathbf{g}^{(N)} = g_{(N)kl} dy^k \otimes dy^l.$$

With the given notation, if  $u \in T_p N$ , then

$$(di_p u)^i = \frac{\partial x^i}{\partial y^k} |_{\phi(p)} u^k.$$

As a consequence, with  $\mathbf{g}^{(N)}$  as in Definition 5.13, one finds

$$g_{kl}^{(N)} u^k v^l = \mathbf{g}^{(N)}(u, v) = g_{ij} \frac{\partial x^i}{\partial y^k} u^k \frac{\partial x^j}{\partial y^l} v^l = \left( \frac{\partial x^i}{\partial y^k} \frac{\partial x^j}{\partial y^l} g_{ij} \right) u^k v^l.$$

Thus

$$\left( g_{kl}^{(N)} - \frac{\partial x^i}{\partial y^k} \frac{\partial x^j}{\partial y^l} g_{ij} \right) u^k v^l = 0.$$

Since the values of the coefficients  $u^r$  and  $v^s$  are arbitrary, each term in the matrix of the coefficients inside the parentheses must vanish. We have found that the relation between the tensor  $\mathbf{g}$  and the tensor  $\mathbf{g}^{(N)}$  evaluated at the same point  $p$  with coordinates  $(y^1, \dots, y^n) = \phi(q)$  in  $N$  and  $(x^1(y^1, \dots, y^n), \dots, x^m(y^1, \dots, y^n))$  in  $M$  reads

$$(g_p^{(N)})_{kl} = \frac{\partial x^i}{\partial y^k} |_{\phi(p)} \frac{\partial x^j}{\partial y^l} |_{\phi(p)} (g_p)_{ij}. \quad (5.3)$$

**Examples 5.15.**

1. Let us consider the submanifold given by the cylinder  $C \subset \mathbb{E}^3$  defined in the example 5.8.2.



It is possible to induce a metric on  $C$  from the natural metric of  $\mathbb{E}^3$ . To this end, referring to the formulae above, the metric on the cylinder reads

$$g_{kl}^{(C)} = \frac{\partial x^i}{\partial y^k} \frac{\partial x^j}{\partial y^l} g_{ij}.$$

where  $x^1, x^2, x^3$  are local coordinates in  $\mathbb{E}^3$  defined around a point  $q \in C$  and  $y^1, y^2$  are analogous coordinates on  $C$  defined around the same point  $q$ . We are free to take cylindrical coordinates adapted to the cylinder itself, that is  $x^1 = \theta, x^2 = r, x^3 = z$  with  $\theta = (-\pi, \pi)$ ,  $r \in (0, +\infty)$ ,  $z \in \mathbb{R}$ . Then the coordinates  $y^1, y^2$  can be chosen as  $y^1 = \theta$  and  $y^2 = z$  with the same domain. These coordinates cover the cylinder without the line passing for the limit points at  $\theta = \pi \equiv -\pi$ . However there is such a coordinate system around every point of  $C$ , it is sufficient to rotate (around the axis  $z = u^3$ ) the orthonormal Cartesian frame  $u^1, u^2, u^3$  used to define the initially given cylindrical coordinates. In global orthonormal coordinates  $u^1, u^2, u^3$ , the metric of  $\mathbb{E}^3$  reads

$$\mathbf{g} = du^1 \otimes du^1 + du^2 \otimes du^2 + du^3 \otimes du^3,$$

that is  $\mathbf{g} = \delta_{ij} du^i \otimes du^j$ . As  $u^1 = r \cos \theta$ ,  $u^2 = r \sin \theta$ ,  $u^3 = z$ , the metric  $\mathbf{g}$  in local cylindrical coordinates of  $\mathbb{E}^3$  has components

$$\begin{aligned} g_{rr} &= \frac{\partial x^i}{\partial r} \frac{\partial x^j}{\partial r} \delta_{ij} = 1 \\ g_{\theta\theta} &= \frac{\partial x^i}{\partial \theta} \frac{\partial x^j}{\partial \theta} \delta_{ij} = r^2 \\ g_{zz} &= \frac{\partial x^i}{\partial z} \frac{\partial x^j}{\partial z} \delta_{ij} = 1 \end{aligned}$$

All the mixed components vanish. Thus, in local coordinates  $x^1 = \theta, x^2 = r, x^3 = z$  the metric of  $\mathbb{E}^3$  takes the form

$$\mathbf{g} = dr \otimes dr + r^2 d\theta \otimes d\theta + dz \otimes dz$$

The induced metric on  $C$ , in coordinates  $y^1 = \theta$  and  $y^2 = z$  has the form

$$\mathbf{g}^{(C)} = \frac{\partial x^i}{\partial y^k} \frac{\partial x^j}{\partial y^l} g_{ij} dy^k \otimes dy^l = r^2 d\theta \otimes d\theta + dz \otimes dz = d\theta \otimes d\theta + dz \otimes dz.$$

That is

$$\mathbf{g}^{(C)} = d\theta \otimes d\theta + dz \otimes dz.$$

In other words, the local coordinate system  $y^1, y^2$  is canonical with respect to the metric on  $C$  induced by that of  $\mathbb{E}^3$ . Since there is such a coordinate system around every point of  $C$ , we conclude that  $C$  is a locally flat Riemannian manifold.  $C$  is not globally flat because there is no global coordinate frame which is canonical and cover the whole manifold.

**2.** Let us illustrate a case where the induced metric is degenerate. Consider Minkowski spacetime  $\mathbb{M}^4$  (see (1) Examples 5.8 and [Mor20]), that is the affine four-dimensional space  $\mathbb{A}^4$  equipped

with the scalar product – defined in the vector space of translations  $V$  associated with  $\mathbb{A}^4$  and thus induced on the manifold – with signature  $(1, 3)$ . In other words,  $\mathbb{M}^4$  admits a (actually an infinite class) Cartesian coordinate system with coordinates  $x^0, x^1, x^2, x^3$  where the metric reads

$$\mathbf{g} = g_{ij} dx^i \otimes dx^j = -dx^0 \otimes dx^0 + \sum_{i=1}^3 dx^i \otimes dx^i .$$

Now consider the submanifold

$$\Sigma = \{p \in \mathbb{M}^4 \mid (x^0(p), x^1(p), x^2(p), x^3(p)) = (u, u, v, w) , \ u, v, w \in \mathbb{R}\}$$

We leave to the reader the proof of the fact that  $\Sigma$  is actually a submanifold of  $\mathbb{M}^4$  with dimension 3. A global coordinate system on  $\Sigma$  is given by coordinates  $(y^1, y^2, y^3) = (u, v, w) \in \mathbb{R}^3$  defined above. What is the induced metric on  $\Sigma$ ? It can be obtained, in components, by the relation

$$g^{(\Sigma)} = g_{pq}^{(\Sigma)} dy^p \otimes dy^q = g_{ij} \frac{\partial x^i}{\partial y^p} \frac{\partial x^j}{\partial y^q} dy^p \otimes dy^q .$$

Using  $x^0 = y^1, x^1 = y^1, x^2 = y^2, x^3 = y^3$ , one finds  $g_{(\Sigma)33} = 1$ ,  $g_{(\Sigma)3k} = g_{(\Sigma)k3} = 0$  for  $k = 1, 2$  and finally,  $g_{(\Sigma)11} = g_{(\Sigma)22} = 0$  while  $g_{(\Sigma)12} = g_{(\Sigma)21} = 1$ . By direct inspection one finds that the determinant of the matrix of coefficients  $g_{(\Sigma)pq}$  vanishes and thus the induced metric is degenerate, that is it is not a metric. Such submanifolds with degenerate induced metric are called “null submanifolds” or “light-like manifolds”. ■

### 5.2.3 Isometries

Another interesting case is when the manifolds connected by the immersion  $i : N \rightarrow M$  are *both* (pseudo) Riemannian and  $i$  is a diffeomorphism or a smooth embedding. In that case we have two metrics on  $N$  to be compared.

**Definition 5.16.** (**Isometry.**) *Let  $(M, \mathbf{g})$  and  $(N, \mathbf{g}')$  be two (pseudo) Riemannian manifolds. A diffeomorphism  $\phi : N \rightarrow M$  is called **isometry**, and  $(M, \mathbf{g})$  and  $(N, \mathbf{g}')$  are said to be **isometric**, if it results  $\mathbf{g}^{(N)} = \mathbf{g}'$  or, equivalently,  $\mathbf{g}'^{(M)} = \mathbf{g}$ .*

With the help of this new notion, we can equivalently state the definition of locally flat (pseudo) Riemannian manifold as follows.

**Definition 5.17.** (**Flatness.**) *A  $n$ -dimensional (pseudo) Riemannian manifold  $(M, \mathbf{g})$  is said to be **locally flat** if it admits a covering  $\{U_i\}_{i \in I}$  made of open subsets, such that every (pseudo) Riemannian submanifold  $(U_i, \mathbf{g}^{(U_i)})$  is isometric to an open set of a (pseudo) Euclidean space with constant signature  $(m, p)$  constructed out an affine space  $\mathbb{A}^{m+p}$ .* ■

A weaker version of isometry is an *isometric embedding*, where it is not required that the two manifolds have the same dimension.

**Definition 5.18.** (**Isometric embedding.**) Let  $(M, \mathbf{g})$  and  $(N, \mathbf{g}')$  be two (pseudo) Riemannian manifolds. A smooth embedding  $i : N \rightarrow M$  is called **isometric embedding** if  $\mathbf{g}^{(N)} = \mathbf{g}'$  is valid in  $N$ . ■

A natural question is whether or not a Riemannian/Lorentzian  $n$ -dimensional manifold can be isometrically embedded in a  $\mathbb{R}^d$  space viewed as a Riemannian manifold with the standard flat metric or, respectively, in a Minkowski spacetime  $\mathbb{M}^d$  (see (1) Examples 5.8) with sufficiently large  $d$ . The answer to the former case was provided by Nash with his famous isometric embedding theorem exploiting another achievement by Nash<sup>1</sup>: a generalization of the implicit function theorem valid for a certain types of Fréchet spaces, nowadays known as the *Nash-Moser theorem*.

**Theorem 5.19.** (**Nash's Riemannian isometric embedding theorem.**) *If  $(M, \mathbf{g})$  is a Riemannian compact manifold, then there is a smooth isometric embedding  $f : M \rightarrow \mathbb{R}^d$  for a sufficiently large  $d$  with respect to  $\dim(M)$ .*

In 2011 an analogous theorem was obtained by O. Müller and M. Sánchez for Lorentzian manifolds<sup>2</sup> which are *globally hyperbolic* [O'Ne83, BEE96, Min19].

**Theorem 5.20.** (**Müller-Sánchez' Lorentzian isometric embedding theorem.**) *If  $(M, \mathbf{g})$  is a globally hyperbolic Lorentzian manifold, then there is a smooth isometric embedding  $f : M \rightarrow \mathbb{M}^d$  for a sufficiently large  $d$  with respect to  $\dim(M)$ , where  $\mathbb{M}^d$  is the  $d$ -dimensional Minkowski spacetime ((1) Examples 5.8).*

Actually the quoted theorem proves more facts about that embedding, but to properly describe them we should add some further material.

#### 5.2.4 Killing fields

Let us pass to discuss another facet of the notion of isometry. Consider a smooth vector field  $K$  on the (pseudo) Riemann manifold  $(M, g)$ . Its flow  $\Phi^{(K)}$  has an action on the metric through the pull back, producing a new metric  $(\mathbf{g}_t^{(K)})_p$  for every  $p \in M$ :

$$(\mathbf{g}_t^{(K)})_p = \Phi_t^{(K)*} \otimes \Phi_t^{(K)*} \mathbf{g}_{\Phi_t^{(K)}(p)} . \quad (5.4)$$

If  $(\mathbf{g}_t^{(K)})_p = \mathbf{g}_p$ , i.e.,

$$\Phi_t^{(K)*} \otimes \Phi_t^{(K)*} \mathbf{g}_{\Phi_t^{(K)}(p)} = \mathbf{g}_p , \quad (5.5)$$

---

<sup>1</sup>J. Nash, *The imbedding problem for Riemannian manifolds*, Annals of Mathematics, 63 (1): 20–63, (1956)

<sup>2</sup>O. Müller and M. Sánchez, *Lorentzian manifolds isometrically embeddable in  $\mathbb{L}^N$* , Trans. Amer. Math. Soc. 363 (2011), 5367–5379

for every  $p \in M$  and for all  $t \in \mathbb{R}$  where the left-hand side is defined for the corresponding  $p$ , then  $K$  is called *Killing (vector) field* of  $(M, \mathbf{g})$ . In practice,  $\Phi^{(K)}$  defines a smooth local one-parameter group of isometries of  $(M, \mathbf{g})$ . If  $K$  is complete for instance when  $M$  is compact, then  $\{\Phi_t^{(K)}\}_{t \in \mathbb{R}}$  is a properly smooth local one-parameter group of isometries. Taking the derivative at  $t = 0$  of both sides of (5.5) we have the infinitesimal version of it in terms of Lie derivative of tensor fields:

$$\mathcal{L}_K \mathbf{g}_p = 0 \quad \text{for all } p \in M. \quad (5.6)$$

It is not difficult, taking advantage of the group structure of  $\Phi^{(K)}$ , to prove that (5.5) is actually *equivalent* to (5.6). All that leads to the following definition.

**Definition 5.21.** Let  $(M, \mathbf{g})$  be a (pseudo) Riemannian manifold. A smooth vector field  $K \in \mathfrak{X}(M)$  is said to be a **Killing field** for the metric  $\mathbf{g}$  if it satisfies the so-called **Killing equation** (5.6). ■

According to (4.17), the Killing equation takes the following form in coordinates (where we do not explicitly indicate the point  $p \in M$ )

$$K^c \frac{\partial g_{ab}}{\partial x^c} + g_{ac} \frac{\partial K^c}{\partial x^b} + g_{cb} \frac{\partial K^c}{\partial x^a} = 0. \quad (5.7)$$

After having introduced the notion of *covariant derivative* associated to the *Levi-Civita connection*, we shall restate this equation into a more popular version (especially for physicists).

Making use of (4) in Exercises 4.26, we have the following relevant fact.

**Proposition 5.22.** *The Killing vector fields of a (pseudo) Riemannian manifold  $(M, \mathbf{g})$  form a Lie algebra with respect to the Lie bracket. In other words, if  $K, H$  are Killing fields of  $(M, \mathbf{g})$ , then*

- (a)  $aK + bH$  is a Killing field of  $(M, \mathbf{g})$  for every  $a, b \in \mathbb{R}$ ,
- (b)  $[K, H]$  is a Killing field of  $(M, \mathbf{g})$ .

**Proof.** (a) is a trivial consequence of the Killing equation and of  $\mathcal{L}_{aK+bH} = a\mathcal{L}_K + b\mathcal{L}_H$ . Regarding (b), observe that from (4) in Exercises 4.26,

$$\mathcal{L}_{[K, H]} \mathbf{g} = \mathcal{L}_K \mathcal{L}_H \mathbf{g} - \mathcal{L}_H \mathcal{L}_K \mathbf{g} = 0$$

ending the proof. □

**Remark 5.23.** It is possible to prove (Proposition 9.12) that the dimension of the vector space of Killing fields cannot exceed  $\frac{n(n+1)}{2}$  on a smooth (pseudo) Riemannian manifold with dimension  $n$ . When that dimension is reached, as for Minkowski spacetime or *deSitter spacetime*, the manifold is called **maximally symmetric** and it manifests remarkable properties. ■

**Exercises 5.24.**

1. Let  $(M, \mathbf{g})$  and  $(M', \mathbf{g}')$  be two (pseudo) Riemannian manifolds. If  $f : M \rightarrow M'$  is an isometry, prove that  $L_{\mathbf{g}'}(f \circ \gamma) = L_{\mathbf{g}}(\gamma)$  for every piecewise smooth curve  $I \ni t \mapsto \gamma(t) \in M$ .

2. Let  $N \subset M$  be a (connected) embedded submanifold of the (pseudo) Riemannian manifold  $M$  and  $I \ni t \mapsto \gamma(t) \in N$  a piecewise smooth curve. Prove that  $L_{\mathbf{g}^{(N)}}(\gamma) = L_{\mathbf{g}}(\gamma)$  where this identity holds also when  $\mathbf{g}^{(N)}$  is somewhere degenerated and changes the signature provided the definition (5.1) is extended to this case.

3. Prove that (5.6) implies (5.5).

(Hint. Compute the  $t$  derivative of  $\Phi_t^{(K)*} \otimes \Phi_t^{(K)*} \mathbf{g}_{\Phi_t^{(K)}(p)}$  observing that  $\Phi_{t+h}^{(K)}(p) = \Phi_h^{(K)} \Phi_t^{(K)}(p)$  so that  $\Phi_{t+h}^{(K)*} \otimes \Phi_{t+h}^{(K)*} \mathbf{g}_{\Phi_{t+h}^{(K)}(p)} = \Phi_h^{(K)*} \otimes \Phi_h^{(K)*} \mathbf{g}_{\Phi_h^{(K)}(q)}$  where  $q = \Phi_t^{(K)}(p)$ .)

4. Consider a (pseudo) Euclidean space  $\mathbb{E}^n$ , i.e., an affine space equipped with a (pseudo) scalar product in the space of translations  $V$ . Prove that each vector  $K \in V$  defines a complete Killing field (when viewed as vector field with constant components in every Cartesian coordinate system). In other words, translations are *global* smooth one-parameter groups of isometries in (pseudo) Euclidean spaces. This result applies in particular to every Euclidean space and to Minkowski spacetime ((1) Examples 5.8).

## 5.2.5 Normal and co-normal bundle to an embedded submanifold

Let us consider an embedded submanifold  $S \subset M$  of a (pseudo) Riemannian manifold  $(M, \mathbf{g})$ . Let us denote by  $s$  and  $m$  the dimensions of  $S$  and  $M$  respectively. If  $p \in S$  the tangent space  $T_p S$  coincides (through the differential of the canonical inclusion map) with a subspace of  $T_p M$  of dimension  $s$ , therefore we can complete  $T_p S$  as  $T_p M = T_p S \oplus V_p$ , where  $V_p \subset T_p M$  is a subspace with dimension  $m - s$ . We prove now that  $V_p$  can be chosen orthogonal to  $T_p S$  with respect to  $\mathbf{g}_p$  when  $\mathbf{g}_p^{(S)}$  is non-degenerate.

**Proposition 5.25.** *Let  $S \subset M$  be an embedded  $s$ -dimensional smooth submanifold of a (pseudo) Riemannian manifold  $(M, \mathbf{g})$  of dimension  $m$ . If  $\mathbf{g}^{(S)}$  is non-degenerate at  $p \in S$ , it is possible to decompose*

$$T_p M = T_p S \oplus N_p S$$

where  $\oplus$  denotes the direct decomposition of vector spaces and

(a)  $N_p S$  is subspace of  $T_p M$  with dimension  $m - s$ ,

(b)  $\mathbf{g}_p(n, t) = 0$  if  $n \in N_p S$  and  $t \in T_p S$ .

When defined,  $N_p S$  is called the **normal space** to  $S$  at  $p$ .

**Proof.** According to Proposition 4.8, we can find a local chart  $(U, \phi)$  of  $M$  around  $p \in M$  with coordinates  $\phi : U \ni q \mapsto (x^1(q), \dots, x^s(q), y^1(q), \dots, y^{m-s}(q)) \in \mathbb{R}^n$ , such that  $(y^1, \dots, y^{m-s})$  are the coordinates of a chart  $(V, \psi)$  on  $S$  with  $V := U \cap S$  where

$$\phi(V) = \{(x^1, \dots, x^s, 0, \dots, 0) \mid (x^1, \dots, x^s, 0, \dots, 0) \in \psi(V)\}.$$

With this choice,  $T_p S$  is the span in  $T_p M$  of the vectors  $\frac{\partial}{\partial x^k}|_p$  for  $k = 1, \dots, s$ . Consider the covectors  $dy^1|_p, \dots, dy^{m-s}|_p \in T_p^* M$ . By definition, they are linearly independent and they satisfy

$$\left\langle \frac{\partial}{\partial x^k}|_p, dy^r|_p \right\rangle = 0.$$

If we define the  $m - s$  linearly independent vectors  $N_{pr} := \mathbf{g}(dy^r|_p, \cdot) \in T_p M$ , the identities above read

$$\mathbf{g}_p \left( \frac{\partial}{\partial x^k}|_p, N_{pr} \right) = 0, \quad k = 1, \dots, s, \quad r = 1, \dots, m - s.$$

To conclude, it is sufficient to prove that

$$\left\{ \frac{\partial}{\partial x^k}|_p \right\}_{k=1, \dots, s} \cup \{N_{pr}\}_{r=1, \dots, m-s}$$

is a set of  $m$  linearly independent vectors when  $\mathbf{g}_p^{(S)}$  is non-degenerate. To this end consider a zero linear combination

$$0 = \sum_{k=1}^s c^k \frac{\partial}{\partial x^k}|_p + \sum_{r=1}^{n-s} d^r N_{pr}. \quad (5.8)$$

We want to prove that this is only possible if  $c^k = d^r = 0$  for all values of  $k$  and  $r$ . If  $\mathbf{g}_p^{(S)}$  is non-degenerated, we can pass from the basis of the  $\frac{\partial}{\partial x^k}|_p$  of  $T_p S$  to a canonical basis  $e_1, \dots, e_s$  such that  $\mathbf{g}^{(S)}(e_i, e_j) = \pm \delta_{ij}$  and the identity above reads

$$0 = \sum_{k=1}^s c'^k e_k + \sum_{r=1}^{n-s} d^r N_{pr}, \quad (5.9)$$

where  $c^k = \sum_{l=1}^s A^k_l c'^l$  for some nonsingular matrix  $[A^k_l]_{k,l=1, \dots, s}$ . Taking the scalar product of both sides of (5.9) with  $e_l$ , since  $\mathbf{g}_p(N_{pr}, e_l) = 0$  by construction, we have

$$c'^l = 0, \quad l = 1, 2, \dots, s \quad \text{so that} \quad c^k = 0 \quad k = 1, 2, \dots, s.$$

Inserting the result in (5.9) we have,

$$0 = \sum_{r=1}^{n-s} d^r N_{pr},$$

which implies  $d^r = 0$  for  $r = 1, 2, \dots, m - s$ , since the  $N_{pr}$  are linearly independent. All coefficients of the combination (5.8) vanish and thus the set

$$\left\{ \frac{\partial}{\partial x^k}|_p \right\}_{k=1, \dots, s} \cup \{N_{pr}\}_{r=1, \dots, m-s}$$

is made of linearly independent vectors concluding the proof.  $\square$

If  $\mathbf{g}^{(S)}$  is degenerate, we cannot decompose  $T_p S$  as before, however we can characterize  $T_p S$  using the cotangent space as follows.

**Proposition 5.26.** *Let  $S \subset M$  be an embedded  $s$ -dimensional smooth submanifold of a (pseudo) Riemannian manifold  $(M, \mathbf{g})$  of dimension  $m$ . For  $p \in S$ , consider the subspace of  $T_p^* M$*

$$N_p^* S := \{\omega \in T_p^* M \mid \omega(v) = 0 \text{ if } v \in T_p S\}.$$

*The following facts are true.*

(a)  $\dim(N_p^* S) = m - s,$

(b) *If  $\mathbf{g}_p^{(S)}$  is nondegenerate, then  $N_p S$  is isomorphic to  $N_p^* S$  through the canonical isomorphism  $T_p M \ni v \mapsto v^\flat := \mathbf{g}(v, \cdot) \in T_p^* M$ .*

$N_p^* S$  is called the **co-normal space** to  $S$  at  $p$ .

**Proof.** Using the same coordinate system as in the the proof of Proposition 5.25, we can expand  $\omega \in T_p^* M$  as

$$\omega = \sum_{k=1}^s c_k dx^k|_p + \sum_{r=1}^{m-s} d_r dy^r|_p.$$

Imposing the condition that defines  $N_p^* S$  is equivalent to requiring that the action of  $\omega$  is zero when acting on every vector  $\frac{\partial}{\partial x^k}|_p$  for  $k = 1, \dots, s$ . This means that  $c_k = 0$  for  $k = 1, 2, \dots, s$ . In other words  $N_p^* S$  is exactly made of all possible covectors of the form

$$\omega = \sum_{r=1}^{m-s} d_r dy^r|_p \quad d_r \in \mathbb{R} \quad r = 1, \dots, m - s.$$

This result proves (a). Regarding (b), as we have seen in the proof of Proposition 5.25, the isomorphism  $T_p M \ni v \mapsto \mathbf{g}(v, \cdot) \in T_p^* M$  transforms the found basis  $\{dy^r|_p\}_{r=1, \dots, m-s}$  of  $N_p^* S$  to a basis of  $N_p S$  when  $\mathbf{g}_p^{(S)}$  is non degenerate (otherwise  $N_p S$  is not defined).  $\square$

**Definition 5.27.** If  $S$  is a smooth embedded submanifold of the (pseudo) Riemannian manifold  $(M, \mathbf{g})$  then

$$NS := \bigcup_{p \in S} N_p S \quad \text{and} \quad N^* S := \bigcup_{p \in S} N_p^* S$$

(the former defined when  $\mathbf{g}_p^{(S)}$  is non degenerate for every  $p \in S$ ), equipped with the natural fiber bundle structure induced by  $TM$  and  $T^* M$  respectively, are said the **normal bundle** and the **co-normal bundle** to  $S$ .  $\blacksquare$

**Examples 5.28.**

1. If  $(M, \mathbf{g})$  is a Lorentzian manifold the following classification of vectors and covectors exists. If  $X_p \in T_p M \setminus \{0\}$ ,

- (i) If  $X_p$  is **timelike** if  $\mathbf{g}_p(X_p, X_p) < 0$ ,
- (ii) if  $X_p$  is **spacelike** if  $\mathbf{g}_p(X_p, X_p) > 0$ ,
- (iii) if  $X_p$  is **lightlike** if  $\mathbf{g}_p(X_p, X_p) = 0$ .

Covectors  $n_p \in T_p^*M$  are classified similarly using the contravariant representation  $n^\sharp \in T_pM$  where  $\mathbf{g}_p(n_p^\sharp, \cdot) = n_p$ .

**2.** Let us consider an embedded submanifold  $S$  of co-dimension 1 (i.e., of dimension  $\dim(M) - 1$ ) in a Lorentzian manifold  $(M, \mathbf{g})$ . In this case,  $\dim(N_p^*S) = 1$  and this leads to the following popular classification of  $S$  at  $p$ .

Take  $n_p \in N_p^*S \setminus \{0\}$  so that all elements of  $N_p^*S$  are of the form  $an_p$  for  $a \in \mathbb{R}$ . (See [Mor20] for the standard terminology on vectors in Lorentzian manifolds adopted below.)

- (i) If  $n_p$  is **timelike**, then  $S_p$  is said to be **spacelike** at  $p$ ,
- (ii) if  $n_p$  is **spacelike**, then  $S_p$  is said to be **timelike** at  $p$ ,
- (iii) if  $n_p$  is **lightlike**, then  $S_p$  is said to be **lightlike** at  $p$ .

$S$  is said to be **spacelike**, **timelike**, *lightlike* if it is respectively **spacelike**, **timelike**, *lightlike* at each  $p \in S$ .

Let us assume that  $m := \dim(M)$  and focus on the three cases.

- (i) The tangent space of  $S$  at  $p$  is made of  $m - 1$  orthogonal spacelike vectors. This is evident if referring to a canonical orthonormal basis  $\{e_k\}_{k=1, \dots, m}$  of  $T_pM$  with  $e^{*1} = an$  for some  $a \neq 0$ . The vectors annihilated by  $n$  are the span of  $e_2, \dots, e_m$  which are spacelike. Hence  $g_p^{(S)}$  is a proper scalar product and  $(S, \mathbf{g}^{(S)})$  (if  $S$  is connected) is a proper Riemannian manifold of co-dimension 1.
- (ii) Assuming  $m \geq 3$ , the tangent space of  $S$  at  $p$  is spanned by a set of  $m - 2$  spacelike vectors and one timelike vector mutually orthogonal. This is evident if referring to a canonical orthonormal basis  $\{e_k\}_{k=1, \dots, m}$  of  $T_pM$  with  $e^{*2} = an$  for some  $a \neq 0$ . The vectors annihilated by  $n$  are the span of  $e_1$  and  $e_3, \dots, e_m$ . The first one is timelike and the remaining  $m - 2$  are spacelike. Hence  $g_p^{(S)}$  is a Lorentzian scalar product and  $(S, \mathbf{g}^{(S)})$  (if  $S$  is connected) is a Lorentzian manifold of co-dimension 1.
- (iii) Assuming  $m \geq 3$ , the tangent space of  $S$  at  $p$  is spanned by a set of  $m - 2$  spacelike vectors and one timelike vector mutually orthogonal. This is evident if referring to a canonical orthonormal basis  $\{e_k\}_{k=1, \dots, m}$  of  $T_pM$  with  $n = c(e_1^* + e_2^*)$  for some  $c \neq 0$ . The vectors annihilated by  $n$  are the span of  $e_1 - e_2$  and  $e_3, \dots, e_m$ . The first one is lightlike and the remaining  $m - 2$  are spacelike. Hence  $g_p^{(S)}$  is degenerated since  $e_1 - e_2 \neq 0$ , but  $\mathbf{g}^{(S)}(e_1 - e_2, u) = 0$  for every  $u \in T_pS$ .



**3.** In the case (iii) above it is interesting to study the contravariant form  $n^\sharp$  of the co-normal vector  $n$ . Using the said basis where  $n = c(e_1^* + e_2^*)$ , we have  $n^\sharp = c(e_1 - e_2)$ . We stress that  $n^\sharp \in T_p S$  as a consequence, instead of being a vector “normal” to  $T_p S$  as it happens when  $\mathbf{g}^{(S)}$  is non-degenerate. The point is that, with Lorentzian metric, lightlike vectors are normal to themselves. ■

**Exercises 5.29.** Consider a lightlike embedded submanifold  $\Sigma$  in the Lorentzian manifold  $(M, \mathbf{g})$  where  $\dim(M) = 4$  (but the result extends to a generic  $\dim(M) \geq 2$  with the same proof). Proves that, if  $p \in \Sigma$ , then there is an open neighborhood  $U$  of  $p$  in  $M$  equipped with local coordinates  $(u, v, r, s)$ , another open neighborhood  $S \subset \Sigma$ , of  $p$  such that  $v, r, s$  are coordinates on  $S$  – corresponding to  $u = 0$  – and furthermore  $\frac{\partial}{\partial v}$  is lightlike and  $\frac{\partial}{\partial r}, \frac{\partial}{\partial s}$  are spacelike.

**Solution.** There is a local coordinate system  $(u, x, y, z)$  in  $M$  with domain an open neighborhood of  $p$  such that a neighborhood  $S \subset \Sigma$  of  $p$  is represented by  $u = 0$ . Since  $\Sigma$  is lightlike,  $g(du^\sharp, du^\sharp) = 0$ . The vectors  $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$  are therefore tangent to  $\Sigma$  in  $S$  and  $x, y, z$  are coordinates in  $S$  (viewed as an embedded submanifold). Now observe that

$$0 = \mathbf{g}(du^\sharp, du^\sharp) = \langle du^\sharp, du \rangle ,$$

so that  $du_q^\sharp \in T_q S$  as well. This smooth vector field can be integrated in  $S$  since the conditions of Frobenius theorem are trivially satisfied. This means that we can change coordinates  $x, y, z$  in  $S$ , passing to a new local coordinate system  $u, v, r, s$  around  $p$  such that  $\frac{\partial}{\partial v} = du^\sharp$ . Let us study the nature of the remaining coordinates  $r, s$ . By construction,  $\frac{\partial}{\partial v}$  is lightlike. Therefore for every  $q \in S$  we can arrange an orthonormal basis of  $T_q M$  where, for some constant  $k \neq 0$ ,

$$\frac{\partial}{\partial v} \equiv k(1, 0, 0, 1)^t .$$

Just in view of the definition of dual basis, we have that

$$\left\langle \frac{\partial}{\partial r}, du \right\rangle = 0 ,$$

which means

$$\mathbf{g} \left( \frac{\partial}{\partial r}, \frac{\partial}{\partial v} \right) = 0 .$$

Using the said basis and assuming

$$\frac{\partial}{\partial r} \equiv (a, b, c, d)^t ,$$

the orthogonality condition implies

$$\partial_r \equiv (a, b, c, a)^t .$$

Hence,

$$\mathbf{g} \left( \frac{\partial}{\partial r}, \frac{\partial}{\partial r} \right) = b^2 + c^2 \geq 0 .$$

However, if  $b = c = 0$ , we would have that  $\frac{\partial}{\partial r}$  is linearly dependent from  $\frac{\partial}{\partial v}$  which is not possible by construction. We conclude that

$$\mathbf{g}\left(\frac{\partial}{\partial r}, \frac{\partial}{\partial r}\right) = b^2 + c^2 > 0$$

Therefore  $\frac{\partial}{\partial r}$  is spacelike. The same argument proves that  $\frac{\partial}{\partial r}$  is spacelike as well.

### 5.3 Borel measure induced by a metric

This section is devoted to prove that every (pseudo) Riemannian manifold has a canonical Borel measure which admits Radon-Nikodym derivative with respect to the Lebesgue measure (restricted to the Borel  $\sigma$ -algebra) associated to the coordinates of every local chart. This smooth density is constructed out of the components of the metric in those coordinates.

#### 5.3.1 Invariant measure in coordinates

If  $(M, \mathbf{g})$  is a (pseudo) Riemannian manifold, for every local chart  $\phi : U \ni p \mapsto (x^1(p), \dots, x^n(p))$  we can define a preferred positive Borel measure [Coh80, Rud87]:

$$\mu_{\phi}^{(\mathbf{g})}(E) := \int_{\phi(U)} 1_{\phi(E)}(\phi(p)) \sqrt{|\det g_{\phi}(p)|} dx^1 \cdots dx^n . \quad (5.10)$$

Above  $E$  is every Borel set in  $U$  and  $g_{\phi}(p)$  is the matrix of coefficients of  $\mathbf{g}$  in the coordinates of  $(U, \phi)$  evaluated at  $p \in U$  and  $dx^1 \cdots dx^n$  denotes the standard Lebesgue measure of  $\mathbb{R}^n$ , finally  $1_A(x) := 1$  if  $x \in A$  and  $1_A(x) := 0$  otherwise. The crucial property of  $\mu_{\phi}$  is its invariance under change of coordinates.

**Lemma 5.30.** *If  $\phi : U \ni p \mapsto (x^1(p), \dots, x^n(p))$  and  $\psi : V \ni p \mapsto (y^1(p), \dots, y^n(p))$  are local charts on the (pseudo) Riemannian manifold  $(M, \mathbf{g})$  and  $E \subset U \cap V$  is a Borel set, then*

$$\mu_{\phi}^{(\mathbf{g})}(E) = \mu_{\psi}^{(\mathbf{g})}(E) . \quad (5.11)$$

**Proof.** In coordinates

$$g_{\phi}(p) = J|_{\phi(p)} g_{\psi}(p) J|_{\phi(p)}^t ,$$

where

$$J_p = \left[ \frac{\partial y^k}{\partial x^h} |_{\phi(p)} \right]_{h,k=1,\dots,n} .$$

Hence

$$\sqrt{|\det g_{\phi}(p)|} = |\det J_{\phi(p)}| \sqrt{|\det g_{\psi}(p)|} ,$$

so that

$$\begin{aligned}\mu_{\phi}^{(\mathbf{g})}(E) &= \int_{\phi(E)} \sqrt{|\det g_{\phi}(p)|} dx^1 \cdots dx^n = \int_{\phi(E)} |\det J_{\phi(p)}| \sqrt{|\det g_{\psi}(p)|} dx^1 \cdots dx^n \\ &= \int_{\psi(E)} \sqrt{|\det g_{\psi}(p)|} dy^1 \cdots dy^n = \mu_{\psi}^{(\mathbf{g})}(E),\end{aligned}$$

which is the thesis.  $\square$

### 5.3.2 Borel measure on the whole manifold

To go on, observe that, as a consequence of Fubini-Tonelli theorem applied to the counting measure over  $\mathbb{N}$ , we have that, if  $\{s_{ij}\}_{i,j \in \mathbb{N}}$  are *non-negative* reals, then

$$\sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} s_{ij} = \sum_{j \in \mathbb{N}} \sum_{i \in \mathbb{N}} s_{ij} = \sum_{(i,j) \in \mathbb{N} \times \mathbb{N}} s_{ij}$$

if one of the three series converges. If one of the three series diverges (necessarily to  $+\infty$ ) the remaining two do the same.

Furthermore, observe that since a domain  $U$  of a local chart on  $M$  is open, the Borel sets in  $U$  are exactly the intersections  $E \cap U$  where  $E$  is any Borel in  $M$ .

**Remark 5.31.** We recall that a positive measure  $\mu : \mathcal{B}(X) \rightarrow [0, +\infty]$ , where  $\mathcal{B}(X)$  is the Borel  $\sigma$ -algebra on the topological Hausdorff locally-compact space  $X$  is said to be **regular** if both the following properties are true.

- (a) (**External regularity**)  $\mu(A) = \inf\{\mu(B) \mid B \supset A, B \text{ open set in } X\}$  if  $A \in \mathcal{B}(X)$ .
- (b) (**Internal regularity**)  $\mu(A) = \sup\{\mu(K) \mid K \subset A, K \text{ compact in } X\}$  if  $A \in \mathcal{B}(X)$ .

Every topological manifold is Hausdorff and locally-compact so that this definition applies and it applies to smooth manifolds *a fortiori*.  $\blacksquare$

**Theorem 5.32.** *Let  $(M, g)$  be a (pseudo) Riemannian manifold. There exist a unique positive Borel measure  $\mu^{(\mathbf{g})}$  on  $M$  such that*

$$\mu^{(\mathbf{g})}(E) = \mu_{\phi}^{(\mathbf{g})}(E), \tag{5.12}$$

*for every local chart  $U \ni p \mapsto (x^1(p), \dots, x^n(p))$  and for every Borel set  $E \subset U$ , where  $\mu_{\phi}^{(\mathbf{g})}$  is defined as in (5.10).*

*The Borel measure  $\mu^{(\mathbf{g})}$  is regular.*

**Proof.**

*Existence of  $\mu^{(\mathbf{g})}$ .* For every point  $p \in M$ , take a local chart  $(V, \phi)$  with  $V \ni p$ , next shrink the domain  $V$  around  $p$  to another open set  $U$  such that  $\overline{U}$  is compact and  $\overline{U} \subset V$ . Finally, exploiting paracompactness, refine the final covering to a locally finite covering  $\{(U_i, \phi_i)\}_{i \in I}$  which can be always assumed to be countable since  $M$  is second countable (see Remark 2.27). Notice that  $U_i$  and its compact closure are by construction covered by a coordinate patch. To conclude the construction, taking advantage of Theorem 2.26, define a partition of unity  $\{\chi_i\}_{i \in I}$  subordinate to  $\{(U_i, \phi_i)\}_{i \in I}$  and

$$\mu^{(\mathbf{g})}(E) := \sum_{i \in I} \int_M 1_E(p) \chi_i(p) d\mu_{\phi_i}^{(\mathbf{g})}(p) \in [0, +\infty], \quad (5.13)$$

for every Borel set  $E \subset M$ . Notice that  $\chi_i$  is supported in  $U_i$  so that the integration on  $M$  is actually restricted to  $U_i$  and the right-hand side is well-defined. Using the remark about the Fubini-Tonelli theorem for the counting measure, it is easy to prove that  $\mu^{(\mathbf{g})}$  is  $\sigma$ -additive on the Borel algebra of  $M$  and, it being a non-negative function, it defines a positive Borel measure on  $M$ . Let us pass to prove (5.12). If  $E \subset U$ , where  $(U, \phi)$  is a coordinate patch, we have

$$1_E = 1_E 1_U$$

so that, using Lemma 5.30,

$$\begin{aligned} \mu^{(\mathbf{g})}(E) &:= \sum_{i \in I} \int_M 1_E(p) 1_U(p) \chi_i(p) d\mu_{\phi_i}^{(\mathbf{g})}(p) = \sum_{i \in I} \int_U 1_E(p) \chi_i(p) d\mu_{\phi_i}^{(\mathbf{g})}(p) = \sum_{i \in I} \int_U 1_E(p) \chi_i(p) d\mu_{\phi}^{(\mathbf{g})}(p) \\ &= \int_U \chi_E(p) \sum_{i \in I} \chi_i(p) d\mu_{\phi}^{(\mathbf{g})}(p) = \int_U \chi_E(p) d\mu_{\phi}^{(\mathbf{g})}(p) = \mu_{\phi}^{(\mathbf{g})}(E), \end{aligned}$$

where we have interchanged the sum with the integral using the *monotone convergence theorem*, and eventually we have used the fact that  $\sum_{i \in I} \chi_i(p) = 1$  for a partition of unity.

*Uniqueness of  $\mu^{(\mathbf{g})}$ .* Referring to the same locally finite countable open covering  $\{U_i\}_{i \in I}$  and the associated partition of the unity  $\{\chi_i\}_{i \in I}$  used in the existence part of the proof, every positive Borel measure  $\nu$  on  $M$  satisfies

$$\nu(E) = \sum_{i \in I} \int_M 1_E(p) \chi_i(p) d\nu(p)$$

for every Borel set  $E \subset M$ , because  $1_E(p) = \sum_{i \in I} \int_M 1_E(p) \chi_i(p)$  and taking advantage of the monotone convergence theorem. If furthermore  $\nu(F) = \mu_{\phi}^{(\mathbf{g})}(F)$  when  $F \subset U$  is Borel, we also have  $\nu(F) = \mu_{\phi_i}^{(\mathbf{g})}(F)$  when  $F \subset U_i$  in particular. Hence,

$$\nu(E) = \sum_{i \in I} \int_M 1_E(p) \chi_i(p) d\nu(p) = \sum_{i \in I} \int_{U_i} 1_E(p) \chi_i(p) d\nu(p)$$

$$= \sum_{i \in I} \int_{U_i} 1_E(p) \chi_i(p) d\mu_{\phi_i}^{(\mathbf{g})}(p) = \sum_{i \in I} \int_M 1_E(p) \chi_i(p) d\mu_{\phi_i}^{(\mathbf{g})}(p) = \mu^{(\mathbf{g})}(E)$$

due to (5.13).

*Regularity of  $\mu^{(\mathbf{g})}$ .* To conclude,  $\mu^{(\mathbf{g})}$  is regular because it is a positive Borel measure on a Hausdorff locally-compact space with a countable topological basis and such that every compact set has finite measure (Proposition 7.2.3 in [Coh80]). Indeed  $M$  has a countable topological basis by definition. Furthermore, if  $K \subset M$  is compact, we can cover it with a finite number  $N$  of open sets  $U_i$  such that  $\overline{U_i}$  is compact and  $\overline{U_i} \subset V_i$  for a suitable local chart  $(V_i, \phi_i)$  (for instance  $U_i$  can be the image of a coordinate ball with finite radius). Next observe that  $\mu^{(\mathbf{g})}(\overline{U_i}) = \mu_{\phi_i}^{(\mathbf{g})}(\overline{U_i}) < +\infty$  because (a) compact sets have finite Lebesgue measure in  $\mathbb{R}^n$  and furthermore (b) the factor  $\sqrt{|\det g_{\phi_i}(p)|}$  in (5.10) is continuous and thus bounded for  $p \in \overline{U_i}$ . Finally  $\mu^{(\mathbf{g})}(K) \leq \sum_{i=1}^N \mu^{(\mathbf{g})}(\overline{U_i}) < +\infty$ .  $\square$

**Definition 5.33.** (**Measure induced from the metric.**) The measure  $\mu^{(\mathbf{g})}$  constructed in Theorem 5.32 is called **measure induced** from (or **associated to**) the metric  $\mathbf{g}$  on  $M$ .  $\blacksquare$

**Remark 5.34.**

(1) It is evident that the measure  $\mu^{(\mathbf{g})}$  coincides with the standard Lebesgue measure in every orthonormal Cartesian coordinate chart of an Euclidean space. However the same result is valid also in Minkowski spacetime ((1) Examples 5.8) referring to Minkowskian coordinates (i.e. pseudo orthonormal Cartesian coordinates see [Mor20]).

(2) If  $S$  is an embedded submanifold of a (pseudo) Riemannian manifold  $(M, \mathbf{g})$ , a measure  $\mu^{(\mathbf{g}^{(S)})}$  can be analogously defined for the metric induced to  $S$  from  $(M, \mathbf{g})$ . We know that this metric may be degenerate. This happens in particular on (everywhere) lightlike submanifolds of co-dimension 1 (see Example 5.28). In this case,  $\det \mathbf{g}_{\phi}^{(S)}(p) = 0$ . As a consequence, the induced measure is the trivial one.  $\blacksquare$

**Exercises 5.35.** Let  $(M, \mathbf{g})$  and  $(M', \mathbf{g}')$  be a pair of (pseudo) Riemannian manifolds. If  $f : M \rightarrow M'$  is an isometry, prove that  $\mu^{(\mathbf{g}')} (f(E)) = \mu^{(\mathbf{g})}(E)$ .

(Hint. Prove that  $\mu^{(\mathbf{g}')} \circ f$  coincides with  $\mu^{(\mathbf{g})}$  on the domains of a suitable atlas associated with a partition of unity, then exploit (5.10).)

## Chapter 6

# Affine connections and related geometric tools

The notion of affine connection has a number of applications in physics from Classical Mechanics to General Relativity. This chapter presents a general overview on the general notion of covariant derivative of tensor fields on a smooth manifold also for non-metric manifolds, concluding with the notion of affine (and metric) geodesic.

### 6.1 Affine connections and covariant derivatives

Our goal is to give some precise meaning to the derivative  $\nabla_X Y$  of a smooth vector field  $Y$  with respect to another smooth vector field  $X$  on a generic smooth manifold  $M$ .

#### 6.1.1 The problem of the absence of an affine-space structure

Suppose for the moment that  $M = \mathbb{A}^n$ , a  $n$ -dimensional affine space. The global coordinate systems obtained by fixing an origin  $O \in \mathbb{A}^n$ , a basis  $\{e_i\}_{i=1,\dots,n}$  in  $V$ , the vector space of the translations in  $\mathbb{A}^n$  and defining

$$\phi : \mathbb{A}^n \rightarrow \mathbb{R}^n : p \mapsto (\langle \overrightarrow{Op}, e^{*1} \rangle, \dots, \langle \overrightarrow{Op}, e^{*n} \rangle)$$

are called *Cartesian coordinate systems*. These are not (pseudo) orthonormal Cartesian coordinates because there is no given metric. As is well known, different Cartesian coordinate systems  $\phi : \mathbb{A}^n \ni p \mapsto (x^1(p), \dots, x^n(p))$  and  $\psi : \mathbb{A}^n \ni p \mapsto (x'^1(p), \dots, x'^n(p))$  are related by non-homogeneous linear transformations determined by real constants  $A^i_j, B^i$ ,

$$x'^i = A^i_j x^j + B^i,$$

where the matrix of coefficients  $A^i_j$  is non-singular. As a consequence if  $X = X^i \frac{\partial}{\partial x^i}$  is a vector field decomposed with respect to the first Cartesian coordinate system, its components transform

as

$$X'^i = A^i_j X^j,$$

when passing to the second one. If  $Y$  is another vector field, we may try to define the *derivative of  $X$  with respect to  $Y$* , as the contravariant vector which is represented in a Cartesian coordinate system by

$$(\nabla_X Y)_p := X_p^j \frac{\partial Y^i}{\partial x^j} \Big|_p \frac{\partial}{\partial x^i} \Big|_p. \quad (6.1)$$

The question is: “The form of  $(\nabla_X Y)^i$  is preserved under change of coordinates?” If we give the definition using an initial Cartesian coordinate system and we next pass to another Cartesian coordinate system we trivially get:

$$(\nabla_X Y)^i_p = A^i_j (\nabla_X Y)^j_p, \quad (6.2)$$

since the coefficients  $A^i_j$  do *not* depend on  $p$ . Hence, definition (6.1) does *not* depend on the used particular Cartesian coordinate system and gives rise to a  $(1,0)$  tensor which, *in Cartesian coordinates*, has components given by the usual  $\mathbb{R}^n$  directional derivatives of the vector field  $Y$  with respect to  $X$ . The given definition can be re-written into a more intrinsic form which makes clear a very important point. Roughly speaking, to compute the derivative in  $p$  of a vector field  $Y$  with respect to  $X$ , one has to subtract the value of  $Y$  at  $p$  to the value of  $Y$  at a point  $q = p + hX_p$ , where the notation means nothing but that  $\vec{pq} = h\chi_p Y_p$ ,  $\chi_p : T_p \mathbb{A}^n \rightarrow V$  being the natural isomorphism between  $T_p \mathbb{A}^n$  and the vector space  $V$  of the affine structure of  $\mathbb{A}^n$  (see Remark 3.6). This difference has to be divided by  $h$  and the limit  $h \rightarrow 0$  defines the wanted derivatives. It is clear that, as it stands, that procedure makes no sense. Indeed  $Y_q$  and  $Y_p$  belong to different tangent spaces and thus the difference  $Y_q - Y_p$  is not defined. However the affine structure gives a meaning to that difference. In fact, one can use the natural isomorphisms  $\chi_p : T_p \mathbb{A}^n \rightarrow V$  and  $\chi_q : T_q \mathbb{A}^n \rightarrow V$ . As a consequence  $\mathcal{A}[q, p] := \chi_p^{-1} \circ \chi_q : T_q \mathbb{A}^n \rightarrow T_p \mathbb{A}^n$  is a well-defined vector space isomorphism. The very definition of  $(\nabla_X Y)_p$  can be given as

$$(\nabla_X Y)_p := \lim_{h \rightarrow 0} \frac{\mathcal{A}[p + hX_p, p] Y_{p+hX_p} - Y_p}{h}. \quad (6.3)$$

Passing to Cartesian coordinates, it is simply proved that the definition above coincides with that given at the beginning. On the other hand, it is obvious that the affine structure plays a central rôle in the definition of  $(\nabla_X Y)_p$ . In a generic manifold, in the absence of the affine structure, it is not so simple to define the notion of derivative of a vector field with respect to another vector field. Sticking to an affine space  $\mathbb{A}^n$ , but using arbitrary coordinate systems, one can check by direct inspection that the components of the tensor  $\nabla_X Y$  are *not* the  $\mathbb{R}^n$  usual directional derivatives of the vector field  $Y$  with respect to  $X$ . This is because when passing from Cartesian coordinates to generic coordinates, the constant coefficients  $A^i_j$  have to be replaced by  $\frac{\partial x'^i}{\partial x^j} \Big|_p$  which depend on  $p$  and (6.1) cannot be achieved.

In summary, two issues pop out here.

- (a) What is the form of  $\nabla_X Y$  in a generic coordinate system of an affine space?

(b) What about the definition of  $\nabla_X Y$  in a general smooth manifold?

The rest of the section is devoted to answer both questions.

### 6.1.2 Affine connections and covariant derivative operator and vector

The key-idea to give a general answer to the second question is to axiomatically adopt in a generic manifold  $M$  the properties of the operator  $\nabla_X$  grasped when  $M = \mathbb{A}^n$ .

**Definition 6.1.** (**Affine Connection and Covariant Derivative.**) Let  $M$  be a smooth manifold. An **affine connection** or **covariant derivative operator**  $\nabla$ , is a map

$$\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \ni (X, Y) \mapsto \nabla_X Y \in \mathfrak{X}(M),$$

which obeys the following requirements for every point  $p \in M$ :

- (1)  $(\nabla_{fY+gZ} X)_p = f(p)(\nabla_Y X)_p + g(p)(\nabla_Z X)_p$ , for all  $f, g \in D(M)$  and  $X, Y, Z \in \mathfrak{X}(M)$ ;
- (2)  $(\nabla_Y fX)_p = Y_p(f)X_p + f(p)(\nabla_Y X)_p$  for all  $X, Y \in \mathfrak{X}(M)$  and  $f \in D(M)$ ;
- (3)  $(\nabla_X(aY + bZ))_p = a(\nabla_X Y)_p + b(\nabla_X Z)_p$  for all  $a, b \in \mathbb{R}$  and  $X, Y, Z \in \mathfrak{X}(M)$ .

The contravariant vector field  $\nabla_Y X$  is called the **covariant derivative vector of  $X$  with respect to  $Y$**  (and the affine connection  $\nabla$ ). ■

**Remark 6.2.**

(1) It is very important that the three relations written in the definition are understood point-wise also if, very often, they are written without referring to a point of the manifold. For instance, (1) could be re-written  $\nabla_{fY+gZ} X = f\nabla_Y X + g\nabla_Z X$ .

(2) The identity (1) implies that, if  $X_p = X'_p$  then

$$(\nabla_X Z)_p = (\nabla_{X'} Z)_p,$$

in other words:  $(\nabla_X Z)_p$  depends on the value  $X_p$  attained at  $p$  by  $X$ , but not on the other values of  $X$ . This property defines a severe distinction between  $\nabla_X Z$  and the analogous Lie derivative  $\mathcal{L}_X Z$ .

In particular this means that, it make sense to define the derivative in  $p$ ,  $\nabla_{X_p} Y$ , where  $X_p \in T_p M$  is a simple vector and not a vector field. Indeed one can always extend  $X_p$  to a vector field in  $M$  using lemma 3.16, and the derivative does not depend on the choice of such an extension.

As a consequence, the following alternative notations are also used for  $(\nabla_X Z)_p$ :

$$(\nabla_X Z)_p = (\nabla_{X_p} Z)_p = \nabla_{X_p} Z.$$

To show that  $(\nabla_X Z)_p = (\nabla_{X'} Z)_p$  if  $X_p = X'_p$ , by linearity, it is sufficient to show that  $(\nabla_X Z)_p = 0$  if  $X_p = 0$ . Let us prove this fact. First suppose that  $X$  vanishes in a neighborhood  $U_p$  of  $p$ .



Let  $h \in D(M)$  be a function such that  $h(p) = 0$  and  $h(q) = 1$  in  $M \setminus U_p$  (such a function can be constructed as the difference of the constant function 1 and a suitable hat function centered in  $p$ ). By definition  $X = hX$  and thus  $(\nabla_X Z)_p = h(p)(\nabla_X Z)_p = 0(\nabla_X Z)_p = 0$ .

Then suppose that  $X_p = 0$  but, in general  $X_q \neq 0$  if  $q \neq p$ . If  $g$  is a hat function centered on  $p$  compactly supported in the domain of the local chart  $(U, \phi)$  with coordinates  $x^1, \dots, x^n$ , we can write

$$X = gX^i g \frac{\partial}{\partial x^i} + X'.$$

By construction,  $X'$  vanishes in a neighborhood of  $p$  where

$$X = gX^i g \frac{\partial}{\partial x^i},$$

since  $g = 1$  thereon. Putting all together and using the condition (1), one gets, where every scalar function  $gX^i$  is well-defined on the whole manifold,

$$(\nabla_X Z)_p = g(p)X^i(p)(\nabla_{g \frac{\partial}{\partial x^i}} Z)_p + (\nabla_{X'} Z)_p.$$

The first term in the right-hand side vanishes because  $X^i(p) = 0$  by hypotheses, the second vanishes too because  $X'$  vanishes in a neighborhood of  $p$ . Hence  $(\nabla_X Z)_p = 0$  if  $X_p = 0$ .

**(3)** The requirement (2) entails that, if  $Y = Y'$  in a neighborhood of  $p$  then

$$(\nabla_X Y)_p = (\nabla_X Y')_p.$$

In other words:  $(\nabla_X Y)_p$  depends on the behaviour of  $Y$  in a (arbitrarily small) neighborhood of  $p$ .

To show it, it is sufficient to prove that  $(\nabla_X Y)_p = 0$  if  $Y$  vanishes in a neighborhood  $U$  of  $p$ . To prove it, notice that, under the given hypotheses:  $Y = hY$  where  $h \in D(M)$  is a function which vanishes in a neighborhood of  $p$ ,  $V \subset U$  and takes the constant value 1 outside  $U$ . As a consequence

$$(\nabla_X Y)_p = (\nabla_X hY)_p = h(p)(\nabla_X Y)_p + X_p(h)Y_p = 0 + X_p(h)Y_p.$$

Since  $X_p$  is a derivation at  $p$  and  $h$  vanishes in a neighborhood of  $p$ ,  $X_p(h) = 0$  (cf Lemma 3.9). This proves that  $(\nabla_X Y)_p = 0$ .

**(4)** It is clear that the affine structure of  $\mathbb{A}^n$  provided automatically an affine connection  $\nabla$  simply defining it as the standard derivative in every fixed Cartesian coordinate system. (We know that this definition does not depend on the choice of the Cartesian global chart.) The converse is not true: an affine connection does *not* determine any affine structure on a manifold.

**(5)** An important question concerns the existence of an affine connection for a given differentiable manifold. It is possible to successfully tackle that issue after the formalism is developed further.

(1) and (4) in Exercises 6.14 below provide an appropriate answer.

### 6.1.3 Connection coefficients

Let us come back to the general Definition 6.1, in components referred to a local coordinate system  $\phi : U \ni q \mapsto (x^1(q), \dots, x^n(q)) \in \mathbb{R}^n$  defined in a neighborhood  $U$  of  $p \in M$ , we can compute  $(\nabla_X Y)_p$ . To this end we decompose  $X$  and  $Y$  along the local bases made of vectors  $\partial/\partial x^i|_q$  defined for  $q \in U$ . Actually these vectors and the components  $Y^p$  are not defined in the whole manifold as required if one wants to use Definition 6.1. Nevertheless one can define these fields on the whole manifold by multiplying them with suitable hat functions which equal 1 constantly in a neighborhood of  $p$  and vanishes outside the domain of the considered coordinate map. The fields so obtained will be indicated with a prime '. It holds (using notations introduced in (2) Remark 6.2 above):

$$(\nabla_X Y)_p = \nabla_{X^i(p) \frac{\partial}{\partial x^i}|_p} \left( Y^{j'} \frac{\partial}{\partial x^j} \right)' + Z$$

where the vector field  $Z$  vanishes in a neighborhood of  $p$  since, there,  $Y^{j'} \frac{\partial}{\partial x^j} = Y$ . As a consequence, the field  $Z$  does not give contribution to the computation of the covariant derivative in  $p$  by (3) in Remark 6.2. Hence

$$(\nabla_X Y)_p = \nabla_{X^i(p) \frac{\partial}{\partial x^i}|_p} Y^{j'} \frac{\partial}{\partial x^j} = X^i(p) Y^{j'}(p) \nabla_{\frac{\partial}{\partial x^i}|_p} \frac{\partial}{\partial x^j} + X^i(p) \frac{\partial Y^{j'}}{\partial x^i} \Big|_p \frac{\partial}{\partial x^j} \Big|_p.$$

In other words, in our hypotheses:

$$(\nabla_X Y)_p = X^i(p) Y^j(p) \nabla_{\frac{\partial}{\partial x^i}|_p} \frac{\partial}{\partial x^j} + X^i(p) \frac{\partial Y^j}{\partial x^i} \Big|_p \frac{\partial}{\partial x^j} \Big|_p.$$

Notice that, if  $i, j$  are fixed, the coefficients  $\nabla_{\frac{\partial}{\partial x^i}|_p} \frac{\partial}{\partial x^j}$  define a  $(1, 0)$  tensor field in  $p$  which is the derivative of  $\frac{\partial}{\partial x^j}$  with respect to  $\frac{\partial}{\partial x^i}|_p$ . This derivative does not depend on the used extension of the field  $\frac{\partial}{\partial x^j}$  since  $\frac{\partial}{\partial x^j} = \frac{\partial}{\partial x^j}$  in a neighborhood of  $p$ . For this reason we shall henceforth write  $\nabla_{\frac{\partial}{\partial x^i}|_p} \frac{\partial}{\partial x^j}$  instead of  $\nabla_{\frac{\partial}{\partial x^i}|_p} \frac{\partial}{\partial x^j}$ '. It holds

$$\nabla_{\frac{\partial}{\partial x^i}|_p} \frac{\partial}{\partial x^j} = \left\langle \nabla_{\frac{\partial}{\partial x^i}|_p} \frac{\partial}{\partial x^j}, dx^k \Big|_p \right\rangle \frac{\partial}{\partial x^k} \Big|_p := \Gamma_{ij}^k(p) \frac{\partial}{\partial x^k} \Big|_p.$$

The coefficients  $\Gamma_{ij}^k = \Gamma_{ij}^k(p)$  are smooth functions of the considered coordinates and are called **connection coefficients**.

Using these coefficients and the above expansion, in components, the covariant derivative of  $Y$  with respect to  $X$  at  $p$  can be written down as:

$$(\nabla_{X_p} Y)_p^i = X_p^j \left( \frac{\partial Y^i}{\partial x^j} \Big|_{\phi(p)} + \Gamma_{jk}^i(p) Y_p^k \right).$$

### 6.1.4 Covariant derivative tensor

Fix  $X \in \mathfrak{X}(M)$  and  $p \in M$ . The linear map  $Y_p \mapsto (\nabla_{Y_p} X)_p$  and Lemma 3.16 defines a tensor,  $(\nabla X)_p$  of type  $(1, 1)$  in  $T_p^* M \otimes T_p M$  such that the (only possible) contraction of  $Y_p$  and  $(\nabla X)_p$  is  $(\nabla_{Y_p} X)_p$ . Varying  $p \in M$ ,  $p \mapsto (\nabla X)_p$  defines a smooth  $(1, 1)$  tensor field  $\nabla X$  because in local coordinates its components are differentiable they being

$$\frac{\partial X^i}{\partial x^j} + \Gamma_{jk}^i X^k.$$

$\nabla X$  is called **covariant derivative tensor of  $X$**  (with respect to the affine connection  $\nabla$ ).

**Notation 6.3.** The following notation is often used in textbooks

$$\nabla_j X^i := X^i_{;j} := \frac{\partial X^i}{\partial x^j} + \Gamma_{jk}^i X^k. \quad (6.4)$$

Notice the different positions of the corresponding indices passing from the first to the second notation. ■

Using the introduced notation, the relation between the covariant derivative tensor and the covariant derivative vector is stated in components as

$$(\nabla_Y X)^i = Y^j X^i_{;j}.$$

### 6.1.5 Transformation rule of the connection coefficients.

We are now interested in the transformation rule of the connection coefficients under change of coordinates. We pass from local coordinates  $\phi : U \ni q \mapsto (x^1, \dots, x^n) \in \mathbb{R}^n$  to local coordinates  $\psi : V \ni q \mapsto (x'^1, \dots, x'^n) \in \mathbb{R}^n$  such that  $p \in U \cap V$  and the connection coefficients evaluated at  $p$  change from  $\Gamma_{ij}^k(p)$  to  $\Gamma'_{pq}{}^h(p)$ . We do not write  $p$  explicitly in the following for shortness.

$$\Gamma_{ij}^k = \langle \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}, dx^k \rangle = \langle \nabla_{\frac{\partial x'^p}{\partial x^i} \frac{\partial}{\partial x'^p}} \left( \frac{\partial x'^q}{\partial x^j} \frac{\partial}{\partial x'^q} \right), \frac{\partial x^k}{\partial x'^h} dx'^h \rangle = \frac{\partial x^k}{\partial x'^h} \frac{\partial x'^p}{\partial x^i} \langle \nabla_{\frac{\partial}{\partial x'^p}} \left( \frac{\partial x'^q}{\partial x^j} \frac{\partial}{\partial x'^q} \right), dx'^h \rangle.$$

Expanding the last term we get

$$\frac{\partial x^k}{\partial x'^h} \frac{\partial x'^p}{\partial x^i} \nabla_{\frac{\partial}{\partial x'^p}} \left( \frac{\partial x'^q}{\partial x^j} \right) \langle \frac{\partial}{\partial x'^q}, dx'^h \rangle + \frac{\partial x^k}{\partial x'^h} \frac{\partial x'^p}{\partial x^i} \frac{\partial x'^q}{\partial x^j} \langle \nabla_{\frac{\partial}{\partial x'^p}} \frac{\partial}{\partial x'^q}, dx'^h \rangle,$$

which can be re-written as

$$\frac{\partial x^k}{\partial x'^h} \frac{\partial x'^p}{\partial x^i} \frac{\partial^2 x'^h}{\partial x'^p \partial x^j} + \frac{\partial x^k}{\partial x'^h} \frac{\partial x'^p}{\partial x^i} \frac{\partial x'^q}{\partial x^j} \Gamma'_{pq}{}^h.$$

In summary,

$$\Gamma_{ij}^k(p) = \frac{\partial x^k}{\partial x'^h} \Big|_{\psi(p)} \frac{\partial^2 x'^h}{\partial x^i \partial x^j} \Big|_{\phi(p)} + \frac{\partial x^k}{\partial x'^h} \Big|_{\psi(p)} \frac{\partial x'^p}{\partial x^i} \Big|_{\phi(p)} \frac{\partial x'^q}{\partial x^j} \Big|_{\phi(p)} \Gamma'_{pq}{}^h(p).$$

The obtained result shows that the connection coefficients do *not* define a tensor field because of the presence of a non-homogeneous first term in the right-hand side above.

**Remark 6.4.**

If  $\nabla$  is the affine connection naturally associated with the affine structure of an affine space  $\mathbb{A}^n$ , it is clear that  $\Gamma_{il}^k = 0$  in every Cartesian coordinate system. As a consequence, in a generic coordinate system, it must hold

$$\Gamma_{ij}^k = \frac{\partial x^k}{\partial x'^h} \frac{\partial^2 x'^h}{\partial x^i \partial x^j}$$

where the primed coordinates are Cartesian coordinates and the left-hand side does not depend on the choice of these Cartesian coordinates. This result gives the answer of the question "What is the form of  $\nabla_X Y$  in generic coordinate systems (of an affine space)?" raised at the end of Section 6.1.1. The answer is

$$(\nabla_X Y)^i = X^j \left( \frac{\partial Y^i}{\partial x^j} + \Gamma_{jk}^i Y^k \right),$$

where the coefficients  $\Gamma_{jk}^i$  are defined as

$$\Gamma_{ij}^k = \frac{\partial x^k}{\partial x'^h} \frac{\partial^2 x'^h}{\partial x^i \partial x^j},$$

the primed coordinates being Cartesian coordinates.

### 6.1.6 Assignment of a connection.

All the procedure used to define an affine connection can be reversed obtaining the following result. We leave the straightforward proof of the proposition below to the reader.

**Notation 6.5.** Now and henceforth occasionally we write  $|_p$  in place of  $|\phi(p)$  or  $|\psi(p)$  for the sake of shortness. ■

**Proposition 6.6.** *The assignment of an affine connection on a differentiable manifold  $M$  is completely equivalent to the assignment of coefficients  $\Gamma_{ij}^k(p)$  in each local coordinate system, which smoothly depend on the point  $p$  and transform as*

$$\Gamma_{ij}^k(p) = \frac{\partial x^k}{\partial x'^h} |_p \frac{\partial^2 x'^h}{\partial x^i \partial x^j} |_p + \frac{\partial x^k}{\partial x'^h} |_p \frac{\partial x'^p}{\partial x^i} |_p \frac{\partial x'^q}{\partial x^j} |_p \Gamma_{pq}^h(p), \quad (6.5)$$

under change of local coordinates. More precisely,

- (a) if an affine connection  $\nabla$  is given, coefficients  $\Gamma_{ij}^k$  associated with  $\nabla$  which satisfy (6.5) are defined by

$$\Gamma_{ij}^k(p) := \left\langle \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} |_p, dx^k |_p \right\rangle,$$

- (b) if coefficients  $\Gamma_{ij}^k(p)$  are assigned for every point  $p \in M$  and every coordinate system of an atlas of  $M$ , such that (6.5) hold, an affine connection associated with this assignment is given by

$$(\nabla_X Y)_p^i = X_p^j \left( \frac{\partial Y^i}{\partial x^j} \Big|_p + \Gamma_{jk}^i(p) Y_p^k \right).$$

in every coordinate patch of the atlas, for all vector fields  $X, Y$  and every point  $p \in M$ ;

- (c) if  $\nabla$  and  $\nabla'$  are two affine connections on  $M$  such that the coefficients  $\Gamma_{ij}^k(p)$  and  $\Gamma'_{ij}{}^k(p)$  respectively associated to the connections as in (a) coincide for every point  $p \in M$  and every coordinate system around  $p$  in a given atlas on  $M$ , then  $\nabla = \nabla'$ . ■

### 6.1.7 Covariant derivative of tensor fields.

If  $M$  is a smooth manifold equipped with an affine connection  $\nabla$ , it is possible to extend the action of the covariant derivatives to all smooth tensor fields. In other words, if  $X \in \mathfrak{X}(M)$  and  $u$  is a smooth tensor field, it is possible to define a new tensor field  $\nabla_X u$  interpreted as the **covariant derivative of  $u$  with respect to  $X$** . It is done by assuming the following further requirements on the action of  $\nabla_X$  in addition to the four requirements in Definition 6.1, which are supposed to hold point-by-point (we omit  $p$  everywhere for the shake of notational simplicity).

- (4)  $\nabla_X u$  is of the same type as  $u$ ,
- (5)  $\nabla_X (au + bv) = a\nabla_X u + b\nabla_X v$  for all  $a, b \in \mathbb{R}$ , smooth tensor fields  $u, v$  of the same kind .
- (6)  $\nabla_X f := X(f)$  for all  $f \in D(M)$ .
- (7)  $\nabla_X (t \otimes u) := (\nabla_X t) \otimes u + t \otimes \nabla_X u$  for all smooth tensor fields  $u, t$ ,
- (8)  $\nabla_X \langle Y, \eta \rangle = \langle \nabla_X Y, \eta \rangle + \langle Y, \nabla_X \eta \rangle$  for all  $Y \in \mathfrak{X}(M)$  and  $\eta \in \Omega^1(M)$ .

In particular, the action of  $\nabla_X$  on covariant vector fields turns out to be defined by the requirements above as follows.

$$\nabla_X \eta = \left\langle \frac{\partial}{\partial x^k}, \nabla_X \eta \right\rangle dx^k = \left( \nabla_X \left\langle \frac{\partial}{\partial x^k}, \eta \right\rangle \right) dx^k - \left\langle \nabla_X \frac{\partial}{\partial x^k}, \eta \right\rangle dx^k,$$

where

$$\nabla_X \left\langle \frac{\partial}{\partial x^k}, \eta \right\rangle = \nabla_X \eta_k = X(\eta_k) = X^i \frac{\partial \eta_k}{\partial x^i},$$

and

$$\left\langle \nabla_X \frac{\partial}{\partial x^k}, \eta \right\rangle = X^i \eta_r \left\langle \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^k}, dx^r \right\rangle = X^i \eta_r \Gamma_{ik}^r.$$

Putting all together we have:

$$(\nabla_X \eta)_k dx^k = X^i \left( \frac{\partial \eta_k}{\partial x^i} - \Gamma_{ik}^r \eta_r \right) dx^k.$$

We can define

$$(\nabla\eta)_{ki} = \eta_{k,i} := \frac{\partial\eta_k}{\partial x^i} - \Gamma_{ik}^r \eta_r ,$$

as **covariant derivative (tensor)** of the covariant vector field  $\eta$ .  $\nabla\eta$  it is the unique tensor field of type  $(0, 2)$  such that the contraction of  $X_p$  and  $(\nabla\eta)_p$  is  $(\nabla_{X_p}\eta)_p$ .

It is simply proved that, given an affine connection  $\nabla$ , there is exactly one map which transforms smooth tensor fields to smooth tensor fields (preserving their type) and satisfies all the requirements above.

- (a) Uniqueness is straightforward. Indeed, as shown above, the action on covariant tensor fields is uniquely fixed, the action on scalar fields is defined in (6), finally (7) determines the action on generic tensor fields.
- (a) The proof of existence is constructive: in components the uniquely-determined action of the connection on tensor fields is the following. First of all introduce the **covariant derivative** of the tensor field  $t$ ,  $\nabla t$ , which has to be interpreted as the unique tensor field of tensors in  $T_p^*M \otimes S_pM$  ( $S_pM$  being the space of the tensors in  $p$  which contains  $t_p$ ) such that the contraction of  $X_p$  and  $(\nabla t)_p$  (with respect to the space corresponding to the index  $r$ ) is  $(\nabla_{X_p}t)_p$ :

$$\begin{aligned} (\nabla t)^{i_1 \dots i_l}_{j_1 \dots j_k r}(p) &:= \frac{\partial t^{i_1 \dots i_l}_{j_1 \dots j_k}}{\partial x^r} \Big|_p + \Gamma_{sr}^{i_1}(p) t^{s \dots i_l}_{j_1 \dots j_k}(p) + \dots + \Gamma_{sr}^{i_l}(p) t^{i_1 \dots s}_{j_1 \dots j_k}(p) \\ &\quad - \Gamma_{rj_1}^s(p) t^{i_1 \dots i_l}_{s \dots j_k}(p) - \dots - \Gamma_{rj_k}^s(p) t^{i_1 \dots i_l}_{j_1 \dots s}(p) . \end{aligned} \quad (6.6)$$

The reader can easily check the validity of requirements (4)-(8) for the map

$$\nabla : (X, t) \mapsto \nabla_X t$$

defined in that way.

**Notation 6.7.** The following notation is used in the literature,

$$t^{i_1 \dots i_l}_{j_1 \dots j_k, r} := (\nabla t)^{i_1 \dots i_l}_{j_1 \dots j_k r} = \nabla_r t^{i_1 \dots i_l}_{j_1 \dots j_k} .$$

Again, pay attention to the different position of the index  $r$ . ■

## 6.2 The Levi-Civita connection on (pseudo) Riemannian manifolds

We pass to discuss an important affine connection which is always present in (pseudo) Riemannian manifolds. This connection enjoys a general property related to the notion of *torsion*. So we start by introducing that concept.

### 6.2.1 Torsion tensor field

By Schwarz' theorem, the inhomogeneous term in

$$\Gamma_{ij}^k = \frac{\partial x^k}{\partial x'^h} \frac{\partial^2 x'^h}{\partial x^i \partial x^j} + \frac{\partial x^k}{\partial x'^h} \frac{\partial x'^p}{\partial x^i} \frac{\partial x'^q}{\partial x^j} \Gamma_{pq}^h,$$

drops out when considering the transformation rules of coefficients:

$$T_{jk}^i := \Gamma_{jk}^i - \Gamma_{kj}^i.$$

Hence, these coefficients define the components of a tensor field which, in local coordinates, is represented by:

$$T = (\Gamma_{jk}^i - \Gamma_{kj}^i) \frac{\partial}{\partial x^i} \otimes dx^j \otimes dk^k.$$

This tensor field is symmetric in the covariant indices and is called **torsion tensor field of the connection**. It is straightforwardly proved that for any pair of differentiable vector fields  $X$  and  $Y$

$$((\nabla_X Y)_p - (\nabla_Y X)_p - [X, Y]_p)^k = T_p^k{}_{ij} X_p^i Y_p^j,$$

for every point  $p \in M$ . That identity provided an intrinsic definition of torsion tensor field associated with an affine connection. In other words, the torsion tensor at  $p$  can be defined as a bilinear mapping which associates pairs of smooth vector fields  $X, Y$  to a smooth vector field  $T_p(\nabla)(X_p, Y_p)$  according with the rule

$$T_p(X_p, Y_p) = \nabla_{X_p} Y - \nabla_{Y_p} X - [X, Y]_p. \quad (6.7)$$

**Remark 6.8.** It is worthwhile stressing that the difference of  $\nabla_{X_p} Y - \nabla_{Y_p} X$  and  $[X, Y]_p$  does not depend on the behaviour of the fields  $X$  and  $Y$  around  $p$ , but only on the values attained by them exactly at the point  $p$ . Notice that this fact is false if considering separately  $\nabla_{X_p} Y - \nabla_{Y_p} X$  and  $[X, Y]_p$ . ■

There is a nice interplay between the absence of torsion of an affine connection and Lie brackets. In fact, using (6.7), we end up with the following useful result.

**Proposition 6.9.** *Let  $\nabla$  be an affine connection on a differentiable manifold  $M$ . If  $\nabla$  is torsion free, i.e., the torsion tensor field  $T$  vanishes on  $M$ ,*

$$[X, Y] = \nabla_X Y - \nabla_Y X, \quad (6.8)$$

for every  $X, Y \in \mathfrak{X}(M)$ .

### 6.2.2 The Levi-Civita connection.

Let us show that, if  $M$  is (pseudo) Riemannian, there is a preferred affine connection which is torsion free and completely determined by the metric. This is the celebrated *Levi-Civita affine connection*.

**Theorem 6.10.** *Let  $(M, \mathbf{g})$  be a (pseudo) Riemannian manifold with metric locally represented by*

$$\mathbf{g} = g_{ij} dx^i \otimes dx^j .$$

*There is exactly one affine connection  $\nabla$  such that*

(1) *it is **metric**, i.e.,  $\nabla \mathbf{g} = 0$ ,*

(2) *it is **torsion free**, i.e.,  $T = 0$ .*

*That is the **Levi-Civita connection** which is defined by the connection coefficients, called **Christoffel's coefficients**,:*

$$\Gamma_{jk}^i = \{^i_{jk}\} := \frac{1}{2} g^{is} \left( \frac{\partial g_{ks}}{\partial x^j} + \frac{\partial g_{sj}}{\partial x^k} - \frac{\partial g_{jk}}{\partial x^s} \right) . \quad (6.9)$$

**Proof.** Assume that a connection with the required properties exists. Expanding (1) and rearranging the result, we have:

$$-\frac{\partial g_{ij}}{\partial x^k} = -\Gamma_{ki}^s g_{sj} - \Gamma_{kj}^s g_{is} ,$$

twice cyclically permuting indices and changing the overall sign we get also:

$$\frac{\partial g_{ki}}{\partial x^j} = \Gamma_{jk}^s g_{si} + \Gamma_{ji}^s g_{ks} ,$$

and

$$\frac{\partial g_{jk}}{\partial x^i} = \Gamma_{ij}^s g_{sk} + \Gamma_{ik}^s g_{js} .$$

Summing side-by-side the obtained results, taking the symmetry of the lower indices of connection coefficients, i.e. (2), into account as well as the symmetry of the (pseudo) metric tensor, it results:

$$\frac{\partial g_{ki}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k} = 2\Gamma_{ij}^s g_{sk} .$$

Contracting both sides with  $\frac{1}{2}g^{kr}$  and using  $g_{sk}g^{kr} = \delta_s^r$  we get:

$$\Gamma_{ij}^r = \frac{1}{2} g^{rk} \left( \frac{\partial g_{ki}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k} \right) = \frac{1}{2} g^{rk} \left( \frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} \right) = \{^r_{ij}\} .$$



We have proved that, if a connection satisfying (1) and (2) exists, its connection coefficients have the form of Christoffel's coefficients. This fact also implies that, if such a connection exists, it must be unique. The coefficients

$$\{^i_{jk}\}(p) := \frac{1}{2}g^{is}(p) \left( \frac{\partial g_{ks}}{\partial x^j}|_p + \frac{\partial g_{sj}}{\partial x^k}|_p - \frac{\partial g_{jk}}{\partial x^s}|_p \right),$$

define an affine connection because they transform as:

$$\{^k_{ij}\}(p) = \frac{\partial x^k}{\partial x'^h}|_p \frac{\partial^2 x'^h}{\partial x^i \partial x^j}|_p + \frac{\partial x^k}{\partial x'^h}|_p \frac{\partial x'^p}{\partial x^i}|_p \frac{\partial x'^q}{\partial x^j}|_p \{^h_{pq}\}'(p),$$

as one can directly verify with a lengthy computation. This concludes the proof.  $\square$

**Exercises 6.11.** Prove that  $\nabla$  is metric if and only if

$$X_p \mathbf{g}(Y, Z) = \mathbf{g}(\nabla_{X_p} Y, Z_p) + \mathbf{g}(Y_p, \nabla_{X_p} Z) \quad \text{for every } p \in M \text{ and } X, Y, Z \in \mathfrak{X}(M). \quad (6.10)$$

(*Hint.* Prove the thesis in coordinates and use the arbitrariness of  $X, Y, Z$ .)

**Remark 6.12.**

(1) The practical meaning of the requirement (1) is the following. One expects that, in the simplest case, the operation of computing the covariant derivative commutes with the procedure of raising and lowering indices. That is, for instance,

$$g_{ki} \nabla_l t^{ij}{}_r = \nabla_l (g_{ki} t^{ij}{}_r) .$$

The requirement (1) is, in fact, equivalent to the commutativity of the procedure of raising and lowering indices and that of taking the covariant derivative as it can trivially be proved noticing that, in components, requirement (1) read:

$$\nabla_l g_{ij} = 0 .$$

(2) This remark is very important for applications. Consider a (pseudo) Euclidean space  $\mathbb{E}^n$ . In any (pseudo) orthonormal Cartesian coordinate system (and more generally in any Cartesian coordinate system) the affine connection naturally associated with the affine structure has vanishing connection coefficients. As a consequence, that connection is torsion free. In the same coordinates, the metric takes constant components and thus the covariant derivative of the metric vanishes too. Those results prove that the affine connection naturally associated with the affine structure is the Levi-Civita connection. In particular, this implies that the connection  $\nabla$  used in elementary analysis is nothing but the Levi-Civita connection associated to the metric of  $\mathbb{R}^n$ . The exercises below show how such a result can be profitably used in several applications.

(3) A point must be stressed in application of the formalism: using non-Cartesian coordinates in  $\mathbb{R}^n$  or  $\mathbb{E}^n$ , as for instance polar spherical coordinates  $r, \theta, \phi$  in  $\mathbb{R}^3$ , one usually introduces a local basis of  $T_p \mathbb{R}^3$ ,  $p \equiv (r, \theta, \phi)$  made of *normalized-to-1* vectors  $\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi$  tangent to the

curves obtained by varying the corresponding coordinate. These vectors do not coincide with the vector of the natural basis  $\frac{\partial}{\partial r}|_p, \frac{\partial}{\partial \theta}|_p, \frac{\partial}{\partial \phi}|_p$  because of the different normalization. In fact, if  $g = \delta_{ij} dx^i \otimes dx^j$  is the standard metric of  $\mathbb{R}^3$  where  $x^1, x^2, x^3$  are usual orthonormal Cartesian coordinates, the same metric has coefficients different from  $\delta_{ij}$  in polar coordinates. By construction  $g_{rr} = g(\frac{\partial}{\partial r}, \frac{\partial}{\partial r}) = 1$ , but  $g_{\theta\theta} = g(\frac{\partial}{\partial \theta}, \frac{\partial}{\partial \theta}) \neq 1$  and  $g_{\phi\phi} = g(\frac{\partial}{\partial \phi}, \frac{\partial}{\partial \phi}) \neq 1$ . So  $\frac{\partial}{\partial r} = \mathbf{e}_r$  but  $\frac{\partial}{\partial \theta} = \sqrt{g_{\theta\theta}} \mathbf{e}_\theta$  and  $\frac{\partial}{\partial \phi} = \sqrt{g_{\phi\phi}} \mathbf{e}_\phi$ .

### 6.2.3 Killing fields again

The Levi-Civita connection  $\nabla$  permits us to recast the Killing equation (5.7) for the Killing field  $K$  into a more popular form.

$$\nabla_a K_b + \nabla_b K_a = 0. \quad (6.11)$$

Let us prove it in a proposition.

**Proposition 6.13.** *A smooth vector field  $K \in \mathfrak{X}(M)$  satisfies the Killing equation (5.7)*

$$\mathcal{L}_X \mathbf{g} = 0$$

*with respect to the metric  $\mathbf{g}$  of a (pseudo) Riemannian manifold if and only if it satisfies (6.11).*

**Proof.** If  $\mathcal{L}_X \mathbf{g} = 0$  is equivalent to  $(\mathcal{L}_X \mathbf{g})(Y, Z) = 0$  for every  $X, Y, Z \in \mathfrak{X}(M)$ .

$$\begin{aligned} 0 &= (\mathcal{L}_X \mathbf{g})(Y, Z) = \mathcal{L}_X \mathbf{g}(Y, Z) - \mathbf{g}(\mathcal{L}_X Y, Z) - \mathbf{g}(Y, \mathcal{L}_X Z) \\ &= X \mathbf{g}(Y, Z) - \mathbf{g}([X, Y], Z) - \mathbf{g}(Y, [X, Z]) \\ &= \mathbf{g}(\nabla_X Y, Z) + \mathbf{g}(Y, \nabla_X Z) - \mathbf{g}([X, Y], Z) - \mathbf{g}(Y, [X, Z]) \\ &= \mathbf{g}(\nabla_Y X, Z) + \mathbf{g}(Y, \nabla_Z X), \end{aligned}$$

In the first line we used the fact that the Lie derivative satisfies the Leibnitz rule with respect to contractions. In the third line we exploited the fact that the Levi-Civita connection is metric (see (6.10)) and in the fourth we used that it is also torsion free (see (6.8)). Finally

$$0 = \mathbf{g}(\nabla_Y X, Z) + \mathbf{g}(Y, \nabla_Z X) = Y^a Z^b (\nabla_a X_b) + Y^a Z^b (\nabla_b X_a)$$

is equivalent to (6.11) for the arbitrariness of  $Y$  and  $Z$ .  $\square$

#### Exercises 6.14.

1. Show that, if  $\nabla_k$  are affine connections on a manifold  $M$ , then  $\nabla = \sum_k f_k \nabla_k$  is an affine connection on  $M$  if  $\{f_k\}_{k \in K}$  is a smooth partition of unity. (At each point  $\sum_k f_k \nabla_k$  is a finite convex linear combination of connections).

2. Show that a differentiable manifold  $M$  (1) always admits an affine connection, (2) it is possible to fix that affine connection in order that it does not coincide with any Levi-Civita connection for whatever metric defined in  $M$ .

**Solution.** (1) By theorem 5.9, there is a Riemannian metric  $\mathbf{g}$  defined on  $M$ . As a consequence  $M$  admits the Levi-Civita connection associated with  $\mathbf{g}$ . (2) Let  $\omega, \eta$  be a pair of co-vector fields defined in  $M$  and  $X$  a vector field in  $M$ . Suppose that they are somewhere non-vanishing and  $\omega \neq \eta$  (these fields exist due to Lemma 3.16 and using  $\mathbf{g}$  to pass to co-vector fields from vector fields). Let  $\Xi$  be the tensor field with  $\Xi_p := X_p \otimes \omega_p \otimes \eta_p$  for every  $p \in M$ . If  $\Gamma_{jk}^i$  are the Levi-Civita connection coefficients associated with  $\mathbf{g}$  in any local chart on  $M$ , define  $\Gamma_{jk}^i := \Gamma_{jk}^i + \Xi_{jk}^i$  in the same coordinate patch. By construction these coefficients transforms as connection coefficients under a change of coordinate frame. As a consequence of proposition 6.6 they define a new affine connection in  $M$ . By construction the found affine connection is not torsion free and thus it cannot be a Levi-Civita connection.

**3.** Show that the coefficients of the Levi-Civita connection on a manifold  $M$  with dimension  $n$  satisfy

$$\Gamma_{ij}^i(p) = \frac{\partial \ln \sqrt{|\det g_\phi(p)|}}{\partial x^j} \Big|_{\phi(p)}.$$

where  $g_\phi(p) = [g_{ij}(p)]$  in every local chart  $\phi : U \ni q \mapsto (x^1(q), \dots, x^n(q)) \in \mathbb{R}^n$ .

**Solution.** Notice that the sign of  $\det g_\phi$  is fixed it depending on the signature of the metric. It holds

$$\frac{\partial \ln \sqrt{|\det g|}}{\partial x^j} = \frac{1}{2 \det g} \frac{\partial \det g}{\partial x^j}.$$

Using the formula for expanding derivatives of determinants and expanding the relevant determinants in the expansion by rows, one sees that

$$\frac{\partial \det g}{\partial x^j} = \sum_k (-1)^{1+k} \text{cof}_{1k} \frac{\partial g_{1k}}{\partial x^j} + \sum_k (-1)^{2+k} \text{cof}_{2k} \frac{\partial g_{2k}}{\partial x^j} + \dots + \sum_k (-1)^{n+k} \text{cof}_{nk} \frac{\partial g_{nk}}{\partial x^j}.$$

That is

$$\frac{\partial \det g}{\partial x^j} = \sum_{i,k} (-1)^{i+k} \text{cof}_{ik} \frac{\partial g_{ik}}{\partial x^j},$$

On the other hand, Cramer's formula for the inverse matrix of  $[g_{ik}]$ ,  $[g^{pq}]$ , says that

$$g^{ik} = \frac{(-1)^{i+k}}{g} \text{cof}_{ik}$$

and so,

$$\frac{\partial \det g}{\partial x^j} = (\det g) g^{ik} \frac{\partial g_{ik}}{\partial x^j},$$

hence

$$\frac{1}{2 \det g} \frac{\partial \det g}{\partial x^j} = \frac{1}{2} g^{ik} \frac{\partial g_{ik}}{\partial x^j}.$$

But direct inspection proves that

$$\Gamma_{ij}^i(p) = \frac{1}{2} g^{ik} \frac{\partial g_{ik}}{\partial x^j}.$$

Putting all together one gets the thesis.)

4. Prove, without using the existence of a Riemannian metric for any differentiable manifold, that every differentiable manifold admits an affine connection.

(*Hint.* Use a proof similar to that as for the existence of a Riemannian metric: Consider an atlas and define the trivial connection (i.e, the usual derivative in components) in each coordinate patch. Then, making use of a suitable partition of unity, glue all the connections together paying attention to the fact that a convex linear combinations of connections is a connection.)

5. Show that the **divergence** of a vector field  $\text{div} X := \nabla_i X^i$  with respect to the Levi-Civita connection can be computed in coordinates as

$$(\text{div} V)(p) = \frac{1}{\sqrt{|\det g_\phi(p)|}} \frac{\partial \sqrt{|\det g_\phi|} V^i}{\partial x^i} \Big|_{\phi(p)} .$$

6. Use the formula above to compute the divergence of a vector field  $V$  represented in polar spherical coordinates in  $\mathbb{R}^3$ , using the components of  $V$  either in the natural basis  $\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \phi}$  and in the normalized one  $\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi$  (see (2) in Remark Rem16.).

7. Solve exercise (6) for a vector field in  $\mathbb{R}^2$  in polar coordinates and a vector field in  $\mathbb{R}^3$  is cylindrical coordinates.

8. The **Laplace-Beltrami** operator (also called **Laplacian**) on differentiable functions is locally defined by:

$$\Delta f := g^{ij} \nabla_j \nabla_i f ,$$

where  $\nabla$  is the Levi-Civita connection. Show that, in coordinates:

$$(\Delta f)(p) = \frac{1}{\sqrt{|\det g_\phi(p)|}} \left( \frac{\partial}{\partial x^i} \sqrt{|\det g_\phi|} g^{ij} \frac{\partial}{\partial x^j} \right) \Big|_{\phi(p)} f .$$

9. Consider cylindrical coordinates in  $\mathbb{R}^3$ ,  $(r, \theta, z)$ . Show that:

$$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r} \frac{\partial f}{\partial r} + \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2} .$$

10. Consider spherical polar coordinates in  $\mathbb{R}^3$ ,  $(r, \theta, \phi)$ . Show that:

$$\Delta f = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2} .$$

#### 6.2.4 The divergence theorem

The Levi-Civita covariant derivative permits to extend the validity of the classic divergence theorem. Actually there are many versions of that theorem which is a subcase of Stokes-Poincaré theorem for  $k$ -forms [KoNo96].

**Theorem 6.15.** (Divergence theorem in covariant form.) *Let  $(M, \mathbf{g})$  be a (pseudo) Riemannian manifold and  $\Omega \subset M$  an open relatively compact set whose boundary  $\partial\Omega$  is piecewise*

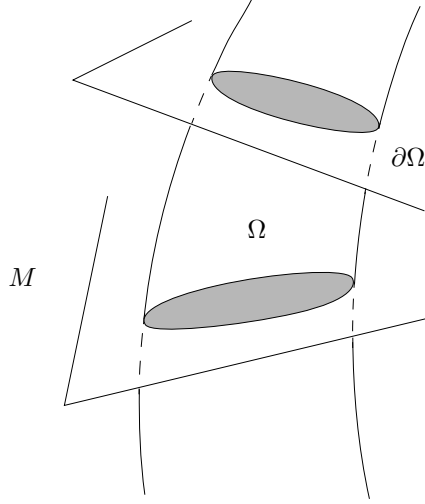


Figure 6.1: An example of domain for the divergence theorem

smooth and it is made of the union of a finite number of co-dimension 1 submanifolds – each with non degenerate induced metric if  $\mathbf{g}$  is semi Riemannian – which meet in a finite number co-dimension 2 submanifolds.

If  $X$  is a smooth field defined on  $\bar{\Omega}$ , and  $\partial\Omega$  is orientable in the sense that we can define a continuous and piecewise smooth normal co-vector  $n$  outward oriented, then

$$\int_{\Omega} \nabla \cdot X d\mu(\mathbf{g}) = \int_{\partial\Omega} \langle X, n \rangle d\mu(\mathbf{g}^{(\partial\Omega)}) . \quad (6.12)$$

Above  $n$  is normalized to 1 on the co-dimension 1 submanifolds and  $\nabla \cdot X$  is the divergence computed with respect to the Levi-Civita connection, in local coordinates

$$\nabla \cdot X = \nabla_a X^a .$$

This theorem has a big impact to the formulation of conservation laws in relativistic theories.

### 6.3 Parallel transport

Let us focus on a manifold  $M$  equipped with an affine connection  $\nabla$ . It is possible to generalize the concept of *straight line* or *affine segment* by introducing the concept of *geodesic*. To do it, we have to introduce some mathematical technology regarding derivatives along a given smooth curve.

### 6.3.1 Covariant derivatives of a vector field along a curve

Let us start with a very natural definition.

**Definition 6.16.** If  $M$  is a smooth manifold with an affine connection  $\nabla$  and  $\gamma : (a, b) \rightarrow M$  a smooth curve, a **smooth vector field  $X$  on  $\gamma$**  is a map

$$(a, b) \ni t \mapsto X(t) \in T_{\gamma(t)}M ,$$

whose components are smooth functions in every local chart around  $\gamma(t)$  for every  $t \in (a, b)$ .  $\mathfrak{X}(\gamma)$  denotes the space of smooth vector field  $X$  on  $\gamma$ . ■

**Remark 6.17.** Notice that we have not requested than  $\gamma$  is injective, so that we may have  $\gamma(t_1) = \gamma(t_2)$  but  $X(t_1) \neq X(t_2)$ . ■

A special case is when  $X(t) = Y|_{\gamma(t)}$  for some  $Y \in \mathfrak{X}(M)$ . In this case we can define the derivative of  $X$  respect to  $\gamma'$  at  $t = t_0$  trivially as

$$\nabla_{\gamma'(t_0)}X := \nabla_{\gamma'(t)}Y ,$$

where the right-hand side is the standard covariant derivative of a vector field with respect to a vector at a point of  $M$ . In a local chart  $(U, \phi)$  around  $\gamma(t_0)$ , where  $\phi(\gamma(t)) = (x^1(t), \dots, x^n(t))$ ,

$$\begin{aligned} (\nabla_{\gamma'(t_0)}X)^b &= \frac{dx^a}{dt}\bigg|_{t_0} \frac{\partial Y^b}{\partial x^a}\bigg|_{\phi(\gamma(t_0))} + \Gamma_{ac}^b(\gamma(t_0)) \frac{dx^a}{dt}\bigg|_{t_0} Y^c(\gamma(t_0)) \\ &= \frac{dY^b(\gamma(t))}{dt}\bigg|_{t=t_0} + \Gamma_{ac}^b(\gamma(t_0)) \frac{dx^a}{dt}\bigg|_{t_0} Y^c(\gamma(t_0)) \\ &= \frac{dX^b(t)}{dt}\bigg|_{t=t_0} + \Gamma_{ac}^b(\gamma(t_0)) \frac{dx^a}{dt}\bigg|_{t_0} X^c(t_0) . \end{aligned}$$

In the last line only the restriction of  $Y$  to  $\gamma$  is used. The final result can be used to give the wanted definition of the covariant derivative of a smooth vector field on  $\gamma$  defined as in Definition 6.16.

**Proposition 6.18.** Let  $M$  be a smooth manifold with an affine connection  $\nabla$ ,  $\gamma : (a, b) \rightarrow M$  a smooth curve. Then there is a unique map associating a smooth field  $X$  on  $\gamma$  to another smooth field  $X'$  on  $\gamma$  satisfying the following requirements.

- (a)  $(aX + bY)'(t) := aX'(t) + bY'(t)$  if  $a, b \in \mathbb{R}$ ,  $X, Y \in \mathfrak{X}(\gamma)$ , and  $t \in (a, b)$ ;
- (b)  $(fX)'(t) := \frac{df}{dt}X(t) + f(t)X'(t)$  if  $f \in D((a, b))$ ,  $X \in \mathfrak{X}(\gamma)$ , and  $t \in (a, b)$ ;
- (c)  $X'(t) = \nabla_{\gamma(t)}Y$  if  $X(t) = Y|_{\gamma(t)}$  for some  $Y \in \mathfrak{X}(M)$ .

If  $t_0 \in (a, b)$ ,  $X'(t_0)$  is called **(covariant) derivative of  $X$  on  $\gamma$  at  $t_0$**  and it is denoted by  $\nabla_{\gamma'(t_0)} X$ .

In a local chart  $(U, \phi)$  around  $\gamma(t_0)$  where  $\phi(\gamma(t)) = (x^1(t), \dots, x^n(t))$ , it holds

$$(\nabla_{\gamma'(t_0)} X)^b = \frac{dX^a(t)}{dt} \Big|_{t=t_0} + \Gamma_{ac}^b(\gamma(t_0)) \frac{dx^a}{dt} \Big|_{t_0} X^c(t_0). \quad (6.13)$$

for every  $X \in \mathfrak{X}(\gamma)$ .

**Proof.** Using (a), (b), and (c) and dealing with a local chart  $(U, \phi)$ , one finds (6.13). On the other hand, by direct inspection one immediately sees that the right-hand side of (6.13) does not depend on the local chart.  $\square$

### 6.3.2 Parallel transport of vectors and geodesics

The introduced results and definitions permit us to introduce two crucial notions with fundamental impact in the formulation of General Relativity.

**Definition 6.19.** (**Parallel transport and geodesic curves.**) Let  $M$  be a smooth manifold equipped with an affine connection  $\nabla$  and consider a smooth curve  $\gamma : (a, b) \rightarrow M$ .

- (a) A vector field  $X \in \mathfrak{X}(\gamma)$  is said to be **parallelly transported along  $\gamma$**  (and with respect to  $\nabla$ ) if

$$\nabla_{\gamma'(t)} X = 0 \quad \text{for all } t \in (a, b),$$

- (b)  $\gamma$  is a **geodesic curve** if it transports its tangent vector parallelly to itself:

$$\nabla_{\gamma'(t)} \gamma'(t) \equiv 0 \quad \text{for all } t \in (a, b).$$

This equation is called **geodesic equation**.  $\blacksquare$

**Remark 6.20.** If  $M = \mathbb{A}^n$  equipped with the natural affine connection whose connection coefficients vanish in every Cartesian coordinate system, the geodesic equation reads in one of these global charts

$$\frac{d^2 x^a}{dt^2} = 0.$$

We conclude that  $I \ni t \mapsto \gamma(t) \in \mathbb{A}^n$  is a geodesic curve if and only if it is (a restriction of) an affine straight line

$$\gamma(t) = p + tv \quad t \in \mathbb{R},$$

for some  $p \in \mathbb{A}^n$  and some  $v \in V$ . This result is obviously valid for the Levi-Civita connection of every (pseudo) Euclidean space.  $\blacksquare$

In the (pseudo) Riemannian case we have an important result which, in particular holds true for *Levi-Civita connections*.

**Proposition 6.21.** *Let  $M$  be a smooth manifold equipped with*

- (a) *a (pseudo) metric  $\mathbf{g}$ ,*
- (b) *an affine connection  $\nabla$ .*

*If the connection is metric, i.e.,  $\nabla \mathbf{g} = 0$  in  $M$ , then the parallel transport preserves the scalar product: if  $X, Y \in \mathfrak{X}(\gamma)$  are parallelly transported along the smooth curve  $\gamma : (a, b) \ni t \rightarrow M$ , then*

$$\mathbf{g}_{\gamma(t_1)}(X(t_1), Y(t_1)) = \mathbf{g}_{\gamma(t_2)}(X(t_2), Y(t_2))$$

*for every  $t_1, t_2 \in (a, b)$*

**Proof.** Assume that there is a single local chart containing  $\gamma(t_1)$  and  $\gamma(t_2)$ . In local coordinates

$$\begin{aligned} \frac{d}{dt} \mathbf{g}(X(t), Y(t)) &= \frac{d}{dt} [g_{ij}(\gamma(t)) X^i(t) Y^j(t)] \\ &= \gamma'^k \frac{\partial g_{ij}}{\partial x^k} X^i(t) Y^j(t) + g_{ij}(t) \frac{dX^i(t)}{dt} Y^j(t) + g_{ij}(\gamma) X^i(t) \frac{dY^j(t)}{dt}. \end{aligned} \quad (6.14)$$

Now the condition  $\nabla \mathbf{g} = 0$  can be re-written:

$$\frac{\partial g_{ij}}{\partial x^k} = \Gamma_{ki}^s g_{sj} + \Gamma_{kj}^s g_{is}$$

that, taking (6.13) into account, inserted in (6.14) produces:

$$\frac{d}{dt} \mathbf{g}(X(t), Y(t)) = \mathbf{g}(\nabla_{\gamma'} X, Y) + \mathbf{g}(\nabla_{\gamma'} Y, X) = 0,$$

since  $\nabla_{\gamma'} X = \nabla_{\gamma'} Y = 0$  by hypothesis.

If a local chart including the whole  $\gamma(a, b)$  does not exist, the thesis is however valid in some interval  $[t_0, t_1]$  with  $t_0, t_1 \in (a, b)$ . Let  $S$  be the set of the reals  $u \in (t_0, b)$  such that  $\mathbf{g}_{\gamma(t)}(X(t), Y(t))$  is constant in  $[t_0, u]$  and let  $T := \sup S \leq b$ . Assume  $T < b$ . There is a local chart containing  $\gamma(T)$  and thus, using the argument above in that local chart, since  $\mathbf{g}_{\gamma(t)}(X(t), Y(t))$  is constant in a left neighborhood of  $T$ , we conclude that  $\mathbf{g}_{\gamma(t)}(X(t), Y(t))$  takes the same constant value also in a right neighborhood of  $T$ . This is impossible for definition of  $T$ , so that  $T = b$  which means that  $\mathbf{g}_{\gamma(t)}(X(t), Y(t))$  is constant in  $[t_0, b)$ . An analogous procedure applies to the left endpoint  $a$  proving that  $\mathbf{g}_{\gamma(t)}(X(t), Y(t))$  is constant in  $(a, t_1]$  concluding the proof.  $\square$

**Remark 6.22.** If  $\gamma : (a, b) \rightarrow M$  is a fixed smooth curve, the parallel transport condition

$$\nabla_{\gamma'(t)} V(t) \equiv 0 \quad \text{for all } t \in (a, b).$$



can be used as a differential equation. Expanding the left-hand side in a local chart  $\phi : U \ni p \mapsto (x^1, \dots, x^n) \in \mathbb{R}^n$  one finds a first-order differential equation for the components of  $V$  referred to the bases of elements  $\frac{\partial}{\partial x^k}|_{\gamma(t)}$ :

$$\frac{dV^i}{dt} = -\Gamma_{jk}^i(\gamma(t))\gamma'^j(t)V^k(t).$$

As the equation is linear, in normal form with smooth known functions in the right-hand side, the initial vector  $V(\gamma(0))$ , assuming  $0 \in (a, b)$  and  $\gamma(0) \in U$ , determines  $V = V(\gamma(t))$  uniquely the maximal sub interval  $(a', b') \subset (a, b)$  such that  $0 \in (a', b')$  and  $\gamma(a', b') \subset U$ . Using an argument similar to the one exploited in the proof of Proposition 6.21, one finally proves that  $V = V(t)$  is uniquely defined in the whole interval  $(a, b)$ .

In a certain sense, one may view the solution  $(a, b) \ni t \mapsto V(t)$  as the “transport” and “evolution” of the initial condition  $V(0)$ , where we assume  $0 \in (a, b)$ , along  $\gamma$ .

If  $u, v \in (a, b)$  with  $u < v$ , the notion of parallel transport along  $\gamma$  produces an vector space isomorphism  $\mathcal{P}_\gamma[u, v] : T_{\gamma(u)} \rightarrow T_{\gamma(v)}$  which associates  $V \in T_{\gamma(u)}$  with that vector in  $T_{\gamma(v)}$  obtained by parallelly transporting  $V$  in  $T_{\gamma(u)}$ .

If  $\nabla$  is metric, Proposition 6.21 implies that  $\mathcal{P}_\gamma[u, v]$  also preserves the scalar product: it is an isometric isomorphism. ■

### Exercises 6.23.

1. Prove that, if  $\nabla$  is an affine connection on  $M$  with torsion tensor  $T$ , then the functions  $\Gamma_{bc}^{a'} := \Gamma_{bc}^a - T_{bc}^a$  define another (torsion-free) connection  $\nabla'$  and that  $\nabla$  and  $\nabla'$  share the same geodesics.

2. Prove that if  $\gamma : I \rightarrow M$  is a geodesic curve with respect to the Levi-Civita connection and  $K$  is a Killing field, then the *conservation equation* holds

$$\frac{d}{dt} \mathbf{g}(K_{\gamma(t)}, \gamma'(t)) = 0.$$

**Solution.** We work in local coordinates. Expanding the left-hand side,

$$\frac{d}{dt} K_a \gamma'^a = \gamma'^b \nabla_b (K_a \gamma'^a) = \gamma'^b \gamma'^a \nabla_b K_a + \gamma'^b K_a \nabla_b \gamma'^a = \gamma'^b \gamma'^a \nabla_b K_a + 0.$$

Finally

$$\gamma'^b \gamma'^a \nabla_b K_a = \frac{1}{2} \gamma'^b \gamma'^a \nabla_b K_a + \frac{1}{2} \gamma'^a \gamma'^b \nabla_a K_b = \frac{1}{2} \gamma'^b \gamma'^a (\nabla_b K_a + \nabla_a K_b) = 0.$$

3. Let  $I \subset \mathbb{R}$  be a bounded interval. Prove that a smooth curve  $I \ni u \mapsto \gamma(u)$  satisfying  $\nabla_{\gamma'} \gamma' = f(u) \gamma'$  and  $\gamma' \neq 0$ , for some known smooth function  $f : I \rightarrow \mathbb{R}$ , can be always re-parametrized to a geodesic segment changing the parameter from  $u$  to  $t = t(u)$  where the function is smooth and  $dt/du > 0$ .

**Solution.** Changing parametrization from  $u$  to  $t = t(u)$ , one sees that the geodesic equation is satisfied with respect to the parameter  $t$  if and only if  $\frac{d}{du}(\frac{dt}{du})^{-1} = f(u)$ . As a cosequence, for  $t_0 \in \mathbb{R}$  and  $u_0 \in I$  fixed arbitrarily,

$$t(u) = t_0 + \int_{u_0}^u \frac{d\eta}{C + \int_{\eta_0}^{\eta} f(\xi) d\xi}$$

where we can always choose  $C \in \mathbb{R}$  such that  $C + \int_{\eta_0}^{\eta} f(\xi) d\xi > 0$  if  $\xi \in I$ , so that  $\frac{dt}{du} > 0$  and the re-parametrization is permitted.

## 6.4 Affine and metric geodesics

This last section is devoted to discuss some general technical property of affine and metric geodesics.

### 6.4.1 First order formulation of the geodesic equation

Let  $M$  be a smooth manifold equipped with an affine connection  $\nabla$ . The geodesic equation  $\nabla_{\gamma'} \gamma'(t) = 0$  for a smooth curve  $\gamma : I \rightarrow M$  can be seen as an equation for a curve in  $TM$  written in natural local coordinates  $(x^1, \dots, x^n, v^1, \dots, v^n)$  adapted to the fiber bundle structure of  $TM$  as

$$\frac{dv^a}{dt} = -\Gamma_{bc}^a(x^1(t), \dots, x^n(t))v^b(t)v^c(t), \quad (6.15)$$

$$\frac{dx^a}{dt} = v^a(t). \quad (6.16)$$

If we change the local chart, remaining in the atlas of natural local charts on  $TM$ , we have the analogous equations

$$\frac{d\bar{v}^d}{dt} = -\bar{\Gamma}_{hk}^d(\bar{x}^1(t), \dots, \bar{x}^n(t))\bar{v}^h(t)\bar{v}^k(t), \quad (6.17)$$

$$\frac{d\bar{x}^d}{dt} = \bar{v}^d(t), \quad (6.18)$$

where  $\Gamma_{bc}^a$  and  $\bar{\Gamma}_{hk}^d$  are connected by the relations (6.5). The crucial fact is the following one. If, in every natural local chart, we define the components at each  $(p, v) \in TM$

$$\Gamma^a|_{p,v} := v^a, \quad \dot{\Gamma}^a|_{p,v} := -\Gamma_{bc}^a(p)v^b v^c,$$

it results that these components define a smooth vector field of  $T(TM)$ , locally represented as

$$\Gamma_{p,v} := \Gamma^a|_{p,v} \frac{\partial}{\partial x^a} |_p + \dot{\Gamma}^a|_{p,v} \frac{\partial}{\partial v^a} |_p.$$

In other words, when changing local natural coordinates, relations (6.5) imply

$$\bar{\Gamma}_{(p,v)}^a = \frac{\partial \bar{x}^a}{\partial x^b} \Gamma_{(p,v)}^a, \quad \dot{\bar{\Gamma}}_{(p,v)}^a = \frac{\partial^2 \bar{x}^a}{\partial x^b \partial x^c} v^c \Gamma_{(p,v)}^a + \frac{\partial \bar{x}^a}{\partial x^b} \dot{\Gamma}_{(p,v)}^a,$$

which are the correct transformation rules of components of a vector of  $T_{(p,v)}TM$  arising from (3.9) and (3.10). Since all the involved functions are smooth, according to (1) in Remark 3.15, this is sufficient to define a smooth vector field on  $TM$  viewed as a smooth manifold in its own right. In summary the geodesic equation can be rephrased as

$$\Lambda'(t) = \Gamma(\Lambda(t)), \quad (6.19)$$

for an unknown smooth curve  $I \ni t \mapsto \Lambda(t) \in TM$ .

With this information, using Proposition 4.16, we have the following result.

**Theorem 6.24.** *Let  $M$  be a smooth manifold equipped with an affine connection  $\nabla$ . The following facts are true.*

(a) *For every  $p \in M$  and  $v \in T_p M$  there exists a unique maximal solution of the geodesic equation  $\gamma_{p,v} : I_{p,v} \rightarrow M$ ,  $-I_{p,v} \subset \mathbb{R}$  being the maximal open interval of definition – such that*

$$(i) \quad \gamma_{p,v}(0) = p,$$

$$(ii) \quad \gamma'_{p,v}(0) = v.$$

(b)  *$\bigcup_{(p,v) \in TM} I_{p,v} \times \{(p,v)\}$  is an open set of  $\mathbb{R} \times TM$  which includes the trivial section*

$$\{(p, 0) \mid p \in M\} \equiv M.$$

**Remark 6.25.**

- (1) A geodesic curve is always a restriction of maximal solution of the geodesic equation.
- (2) A straightforward consequence of the uniqueness part of the theorem above is that the tangent vector of a *non constant* geodesic segment  $\gamma : I \rightarrow M$  cannot vanish in any point. ■

**Definition 6.26.** If the smooth manifold  $M$  is equipped with the affine connection  $\nabla$ ,

- (a) a **geodesic** is a maximal solution of the geodesic equation, i.e., a maximal geodesic curve;
- (b) a **geodesic segment**  $\eta := \gamma|_J$  is the restriction of a geodesic  $\gamma : I \rightarrow M$  to an interval  $J$  possibly including one or both its endpoints. ■

**Definition 6.27.** A manifold  $M$  equipped with an affine connection  $\nabla$  is said to be **geodesically complete**, or simply **complete**, if the maximal geodesics referred to  $\nabla$  are complete (i.e., the domain of every maximal geodesic is the whole real axis). ■

### 6.4.2 Affine parameters and length coordinate

If one changes the parameter of a non constant geodesic segment  $I \ni t \mapsto \gamma(t)$  to  $u = u(t)$  where that mapping is smooth and  $du/dt \neq 0$  for all  $t \in I$ , the new differentiable curve  $\gamma_1 : J \ni u \mapsto \gamma(t(u))$  does not satisfy the geodesic equation in general. However, one easily finds

$$\nabla_{\gamma'_1(u)} \gamma'_1(u) = - \left( \frac{du}{dt} \right)^3 \frac{d^2 t}{du^2} \gamma'(t).$$

Indeed,  $\gamma'_1(u) = \frac{dt}{du} \gamma'(t)$ , where  $t = t(u)$  is the inverse function of the re-parametrization, and thus

$$\nabla_{\gamma'_1(u)} \gamma'_1(u) = \nabla_{\frac{dt}{du} \gamma'(t)} \frac{dt}{du} \gamma'(t) = \frac{dt}{du} \nabla_{\gamma(t)} \frac{dt}{du} \gamma'(t) = \left( \frac{dt}{du} \right)^2 \nabla_{\gamma(t)} \gamma'(t) + \frac{dt}{du} \left( \nabla_{\gamma(t)} \frac{dt}{du} \right) \gamma'(t).$$

That is

$$\nabla_{\gamma'_1(u)} \gamma'_1(u) = 0 + \frac{dt}{du} \left( \frac{d}{dt} \frac{dt}{du} \right) \gamma'(t) = \frac{dt}{du} \left( \frac{d}{dt} \left( \frac{du}{dt} \right)^{-1} \right) \gamma'(t) = - \frac{dt}{du} \frac{\frac{d^2 u}{dt^2}}{\left( \frac{du}{dt} \right)^2} \gamma'(t) = - \left( \frac{du}{dt} \right)^3 \frac{d^2 t}{du^2} \gamma'.$$

Since both  $\gamma'(t) \neq 0$  and  $\frac{du}{dt} \neq 0$ , we see that  $\gamma_1$  satisfies the geodesic equation too, *if and only if*  $\frac{d^2 t}{du^2} = 0$ . This is equivalent to say that  $t = ku + k'$  for some constants  $k, k' \in \mathbb{R}$  with  $k = \frac{dt}{du} \neq 0$ . Equivalently  $u = cu + c'$  for some constants  $c, c' \in \mathbb{R}$  with  $c \neq 0$ .

These transformations of the parameter of geodesics which preserve the geodesic equations are called **affine transformations** (of the parameter) and the new parameters connected to the initial one are said **affine parameters** of the geodesic.

Let us pass to treat a specific case after a general definition.

**Definition 6.28.** If  $\gamma : I \rightarrow M$  is a piecewise smooth curve in the (pseudo) Riemannian manifold  $(M, \mathbf{g})$  with  $\gamma' \neq 0$  everywhere, the function

$$s(t) := \int_{t_0}^t \sqrt{|\mathbf{g}(\gamma'(t'), \gamma'(t'))|} dt'$$

is called **length coordinate**, referred to the fixed point  $t_0 \in I$ , or **length parameter** on  $\gamma$ . ■

Notice that  $s = s(t)$  is a strictly increasing  $C^1(I)$  function which is also piecewise smooth. It is more strongly smooth if  $\gamma$  is smooth. That function admits an inverse  $t = t(s)$  with the same regularity as  $s$  which can be used to re-parametrize  $\gamma$ .

If  $\gamma : I \rightarrow M$  is a geodesic segment with respect to the Levi-Civita connection, due to Proposition 6.21, we know that  $\mathbf{g}(\gamma'(t'), \gamma'(t'))$  is constant along the curve. If  $\mathbf{g}(\gamma'(t'), \gamma'(t')) \neq 0$  – which is always true for non-constant geodesics in Riemannian manifolds or for spacelike or timelike geodesics in Lorentzian manifolds (see below) – the length coordinate referred to  $t_0 \in I$ ,

$$s(t) := \int_{t_0}^t \sqrt{|\mathbf{g}(\gamma'(t'), \gamma'(t'))|} dt',$$

therefore defines a linear function  $s = kt + k'$  with  $k = \sqrt{\mathbf{g}(\gamma', \gamma')} \neq 0$  and thus  $s$  is always an affine parameter of the considered geodesic.

**Examples 6.29.** As we said, in *Einstein's General Theory of Relativity*, the *spacetime* is a four-dimensional Lorentzian manifold  $M$ . Hence it is equipped with a pseudo-metric  $\mathbf{g} = g_{ij}dx^i \otimes dx^j$  with hyperbolic canonic form  $(-1, +1, +1, +1)$  (this holds true if one uses units to measure length such that the speed of the light is  $c = 1$ ). The points of the manifolds are called **events**. If  $V \in T_p M$ ,  $V \neq 0$ , for some event  $p \in M$ ,  $V$  is called *timelike*, *lightlike* (or *null*), *spacelike* if, respectively  $\mathbf{g}(V, V) < 0$ ,  $\mathbf{g}(V, V) = 0$ ,  $\mathbf{g}(V, V) > 0$ . A smooth curve  $\gamma : \mathbb{R} \rightarrow M$  is classified similarly referring to its tangent vector  $\gamma'$  provided  $\gamma'$  preserves the sign of  $\mathbf{g}(\gamma', \gamma')$  along the curve itself. The evolution of a particle is represented by a *world line*, i.e., a timelike differentiable curve  $\gamma : I \ni u \mapsto \gamma(u)$  and the length parameter (length coordinate) along the curve

$$t(u) := \int_a^u \sqrt{|\mathbf{g}(\gamma'(u'), \gamma'(u'))|} du',$$

(notice the absolute value) represents the *proper time* of the particle, i.e., the time measured by a clock which co-moves with the particle. If  $\gamma(t)$  is an event reached by a world line, the tangent space  $T_{\gamma(t)}M$  is naturally decomposed as  $T_{\gamma(t)}M = L(\gamma'(t)) \oplus \Sigma_{\gamma(t)}$ , where  $L(\gamma'(t))$  is the linear space spanned by  $\gamma'(t)$  and  $\Sigma_{\gamma(t)}$  is the orthogonal space to  $L(\gamma'(t))$ . It is simple to prove that the metric  $\mathbf{g}_{\gamma(t)}$  induces a Riemannian (i.e., positive) metric in  $\Sigma_{\gamma(t)}$ .  $\Sigma_{\gamma(t)}$  represents the *local rest space* of the particle at time  $t$ .

Lightlike curves describe the evolution of particles with vanishing mass. It is not possible to define proper time and local rest space in that case.

As a consequence of the remark (3) above, if a geodesic  $\gamma$  has a timelike, lightlike, spacelike initial tangent vector, any other tangent vector along  $\gamma$  is respectively timelike, lightlike, spacelike. Therefore it always make sense to define timelike, lightlike, spacelike geodesics. Timelike geodesics represent the evolutions of points due to the gravitational interaction only. That interaction is represented by the metric of the spacetime. ■

## 6.5 Metric geodesics: the variational approach in coordinates

There is a second approach to define the geodesics with respect to the Levi-Civita connection in a (pseudo) Riemannian manifold. Indeed, geodesics satisfy a *variational principle* because, roughly speaking, they stationarize the length functional of curves and also another functional called *energy* functional.

### 6.5.1 Basic notions of elementary calculus of variations in $\mathbb{R}^n$

Let us recall some basic notion of elementary calculus of variations in  $\mathbb{R}^n$ . Fix an open non-empty set  $\Omega \subset \mathbb{R}^n$ , a closed interval  $I = [a, b] \subset \mathbb{R}$  with  $a < b$  and take a non-empty set

$$G \subset \{\gamma : I \rightarrow \Omega \mid \gamma \in C^k(I)\}$$

for some fixed integer  $0 < k < +\infty$ . Here and henceforth  $\gamma \in C^l([a, b])$ , for  $l = 1, 2, \dots, \infty$ , means that  $\gamma \in C^l((a, b))$  and the limits towards either  $a^+$  and  $b^-$  of all derivatives of  $\gamma$  exist and are finite up to the order  $l$ , and smooth means  $l = \infty$ .

A **variation**  $V$  of  $\gamma \in G$ , if exists, is a map  $V : [0, 1] \times I \rightarrow U$  such that, if  $V_s$  denotes the function  $t \mapsto V(s, t)$ :

- (1)  $V \in C^k([0, 1] \times I)$  (i.e.,  $V \in C^k((0, 1) \times (a, b))$  and the limits towards the points of the boundary of  $(0, 1) \times (a, b)$  of all the derivatives of order up to  $k$  exist and are finite),
- (2)  $V_s \in G$  for all  $s \in [0, 1]$ ,
- (3)  $V_0 = \gamma$  and  $V_s \neq \gamma$  for some  $s \in (0, 1]$ .

It is obvious that there is no guarantee that any  $\gamma$  of any  $G$  admits variations because both condition (2) and the latter part of (3) are not trivially fulfilled in the general case. The following lemma gives a proof of existence provided the domain  $G$  is suitably defined.

**Lemma 6.30.** *Let  $\Omega \subset (\mathbb{R}^n)^k$  be an open non-empty set,  $I = [a, b]$  with  $a < b$ . Fix  $(p, P_1, \dots, P_{k-1})$  and  $(q, Q_1, \dots, Q_{k-1})$  in  $\Omega$ . Let  $D$  denote the space of the elements of*

$$\{\gamma : I \rightarrow \mathbb{R}^n \mid \gamma \in C^k(I)\}$$

*such that:*

- (1)  $\left(\gamma(t), \frac{d^1\gamma}{dt^1}, \dots, \frac{d^{k-1}\gamma}{dt^{k-1}}\right) \in \Omega$  for all  $t \in [a, b]$ ,
- (2)  $\left(\gamma(a), \frac{d^1\gamma}{dt^1}|_a, \dots, \frac{d^{k-1}\gamma}{dt^{k-1}}|_a\right) = (p, P_1, \dots, P_{k-1})$  and  $\left(\gamma(b), \frac{d^1\gamma}{dt^1}|_b, \dots, \frac{d^{k-1}\gamma}{dt^{k-1}}|_b\right) = (q, Q_1, \dots, Q_{k-1})$ .

*With the given definitions and hypotheses, every  $\gamma \in D$  admits variations of the form*

$$V_{\pm}(s, t) = \gamma(t) \pm sc\eta(t),$$

*where  $c > 0$  is a constant,  $\eta : [a, b] \rightarrow \mathbb{R}^n$  is  $C^k$  with*

$$\eta(a) = \eta(b) = 0,$$

*and*

$$\frac{d^r\eta}{dt^r}|_a = \frac{d^r\eta}{dt^r}|_b = 0$$

*for  $r = 1, \dots, k-1$ . In particular, the result holds for every  $c < C$ , if  $C > 0$  is sufficiently small. As a consequence, endowing  $D$  with the topology induced by the norm*

$$\|\gamma\|_k := \max \left\{ \sup_I \|\gamma\|, \sup_I \left\| \frac{d\gamma}{dt} \right\|, \dots, \sup_I \left\| \frac{d^k\gamma}{dt^k} \right\| \right\},$$

every punctured metric ball  $B_r^*(\gamma) = \{\gamma_1 \in D \setminus \{\gamma\} \mid \|\gamma - \gamma_1\| < r\}$  is not empty for every  $r > 0$ .

**Proof.** The only non-trivial fact we have to show is that there is some  $C > 0$  such that

$$\left( \gamma(t) \pm sc\eta(t), \frac{d^1}{dt^1}(\gamma(t) \pm sc\eta(t)), \dots, \frac{d^{k-1}}{dt^{k-1}}(\gamma(t) \pm sc\eta(t)) \right) \in \Omega$$

for every  $s \in [0, 1]$  and every  $t \in I$  provided  $0 < c < C$ . From now on, for a generic curve, we define  $\tau : I \rightarrow \mathbb{R}^n$ ,

$$\tilde{\tau}(t) := \left( \tau(t), \frac{d^1 \tau(t)}{dt^1}, \dots, \frac{d^{k-1} \tau(t)}{dt^{k-1}} \right).$$

We can suppose that  $\overline{\Omega}$  is compact. (If not we can take a covering of  $\tilde{\gamma}([a, b])$  made of open balls of  $(\mathbb{R}^n)^k = \mathbb{R}^{nk}$  whose closures are contained in  $\Omega$ . Then, using the compactness of  $\tilde{\gamma}([a, b])$  we can extract a finite subcovering. If  $\Omega'$  is the union of the elements of the subcovering,  $\Omega' \subset \Omega$  is open,  $\overline{\Omega'} \subset \Omega$  and  $\overline{\Omega'}$  is compact and we may re-define  $\Omega := \Omega'$ .)  $\partial\Omega$  is compact because it is closed and contained in a compact set. If  $\|\cdot\|$  denotes the norm in  $\mathbb{R}^{nk}$ , the map  $(x, y) \mapsto \|x - y\|$  for  $x \in \tilde{\gamma}$ ,  $y \in \partial\Omega$  is continuous and defined on a compact set. Define  $m = \min_{(x, y) \in \tilde{\gamma} \times \partial\Omega} \|x - y\|$ . Obviously  $m > 0$  as  $\tilde{\gamma}$  is internal to  $\Omega$ . Clearly, if  $t \mapsto \tilde{\eta}(t)$  satisfies  $\|\tilde{\gamma}(t) - \tilde{\eta}(t)\| < m$  for all  $t \in [a, b]$ , it must hold  $\tilde{\eta}(I) \subset \Omega$ . Then fix  $\eta$  as in the hypotheses of the Lemma and consider a generic  $\mathbb{R}^{nk}$ -component  $t \mapsto \tilde{\gamma}^i(t) + sc\tilde{\eta}^i(t)$  (the case with  $-$  is analogous). The set  $I' = \{t \in I \mid \tilde{\eta}^i(t) \geq 0\}$  is compact because it is closed and contained in a compact set. The  $s$ -parametrized sequence of continuous functions,  $\{\tilde{\gamma}^i + sc\tilde{\eta}^i\}_{s \in [0, 1]}$ , monotonically converges to the continuous function  $\tilde{\gamma}^i$  on  $I'$  as  $s \rightarrow 0^+$  and thus converges therein uniformly by Fubini's theorem. With the same procedure we can prove that the convergence is uniform on  $I'' = \{t \in I \mid \tilde{\eta}^i(t) \leq 0\}$  and hence it is uniform on  $I = I' \cup I''$ . Since the proof can be given for each component of the curve, we get that  $\|(\tilde{\gamma}(t) + sc\tilde{\eta}(t)) - \tilde{\gamma}(t)\| \rightarrow 0$  uniformly in  $t \in I$  as  $sc \rightarrow 0^+$ . In particular  $\|(\tilde{\gamma}(t) + sc\tilde{\eta}(t)) - \tilde{\gamma}(t)\| < m$  for all  $t \in [a, b]$ , if  $sc < \delta$ . Define  $C := \delta/2$ . If  $0 < c < C$ ,  $sc < \delta$  for  $s \in [0, 1]$  and  $\|(\tilde{\gamma}(t) + sc\tilde{\eta}(t)) - \tilde{\gamma}(t)\| < m$  uniformly in  $t$  and thus  $\tilde{\gamma}(t) + sc\tilde{\eta}(t) \in \Omega$  for all  $s \in [0, 1]$  and  $t \in I$ .

Taking  $C$  smaller if necessary, by means of a similar procedure we prove that,  $\tilde{\gamma}(t) - sc\tilde{\eta}(t) \in D$  for all  $s \in [0, 1]$  and  $t \in I$ , if  $0 < c < C$ .

The last statement can be established as follows. Let  $\gamma_1 \in D \setminus \{\gamma\}$ . Such a curve exists with the form  $\gamma_1(t) = sc\eta(t)$ , with  $0 < c < C$  due to the previous part of the proposition. By construction  $\|\gamma - \gamma_1\| < r_0$  for some  $r_0 > 0$ . Since  $0 < sc < C$  for every  $s \in (0, 1)$ , the curve  $sc\eta$  belongs to  $B_{sr_0}^*(\gamma)$ . In other words  $B_r^*(\gamma)$  is not empty for every  $r \in (0, r_0)$ .  $\square$

### Exercises 6.31.

1. With the same hypotheses of Lemma 6.30, drop the condition  $\gamma(a) = p$  (or  $\gamma(b) = q$ , or both conditions or other similar conditions for derivatives) in the definition of  $D$  and prove the existence of variations  $V_{\pm}$  in this case, too.

We recall that, if  $G \subset \mathbb{R}^n$  and  $F : G \rightarrow \mathbb{R}$  is any sufficiently regular function,  $x_0 \in \text{Int}(G)$  is said to be a *stationary point* of  $F$  if  $dF|_{x_0} = 0$ . Such a condition can be re-written as

$$\frac{dF(x_0 + su)}{ds} \Big|_{s=0} = 0 ,$$

for all  $u \in \mathbb{R}^n$ . In particular, if  $F$  attains a local extreme value in  $x_0$  (i.e. there is a open neighborhood of  $x_0$ ,  $U_0 \subset G$ , such that either  $F(x_0) > F(x)$  for all  $x \in U_0 \setminus \{x_0\}$  or  $F(x_0) < F(x)$  for all  $x \in U_0 \setminus \{x_0\}$ ),  $x_0$  turns out to be a stationary point of  $F$ .

The definition of stationary point can be generalized as follows. Consider a functional on  $G \subset \{\gamma : I \rightarrow U \mid \gamma \in C^k(I)\}$ , i.e. a mapping  $F : G \rightarrow \mathbb{R}$ . We say that  $\gamma_0$  is a **stationary point of  $F$** , if the **variation of  $F$** ,

$$\delta_V F|_{\gamma_0} := \frac{dF[V_s]}{ds} \Big|_{s=0}$$

exists and vanishes for all variations of  $\gamma_0$ ,  $V$ .

**Remark 6.32.** There are different definition of  $\delta_V F$  related to the so-called Fréchet and Gateaux notions of derivatives of functionals. Here we adopt a third definition useful in our context. ■

### 6.5.2 Euler-Poisson equations

For suitable spaces  $G$  and functionals  $F : G \rightarrow \mathbb{R}$ , defining an appropriate topology on  $G$  itself, it is possible to show that if  $F$  attains a *local extremum* in  $\gamma_0 \in G$ , then  $\gamma_0$  must be a stationary point of  $F$ . We state a precise result after specializing the functional  $F$ .

We henceforth work with domains  $G$  of the form  $D$  defined in Lemma 6.30 and we focus attention on this kind of functionals,

$$F[\gamma] := \int_I \mathcal{F} \left( t, \gamma(t), \frac{d\gamma}{dt}, \dots, \frac{d^k \gamma}{dt^k} \right) dt , \quad (6.20)$$

where  $k$  is the same as that used in the Definition of  $D$  and  $\mathcal{F} \in C^k(I \times \Omega \times A)$ ,  $A \subset \mathbb{R}^n$  being an open set where the vectors  $d^k \gamma / dt^k$  take their values. Making use of Lemma 6.30, we can prove a second important lemma.

**Lemma 6.33.** *If  $F : D \rightarrow \mathbb{R}$  is the functional in (6.20) with  $D$  defined in Lemma 3.1,  $\delta_V F|_{\gamma_0}$  exists for every  $\gamma_0 \in D$  and every variation of  $\gamma_0$ ,  $V$  and*

$$\delta_V F|_{\gamma_0} = \sum_{i=1}^n \int_I \frac{\partial V^i}{\partial s} \Big|_{s=0} \left[ \frac{\partial \mathcal{F}}{\partial \gamma^i} + \sum_{r=1}^k (-1)^r \frac{d^r}{dt^r} \left( \frac{\partial \mathcal{F}}{\partial \frac{d^r \gamma^i}{dt^r}} \right) \right] \Big|_{\gamma_0} dt .$$



**Proof.** From known properties of Lebesgue's measure based on Lebesgue's dominate convergence theorem (notice that  $[0, 1] \times I$  is compact and all the considered functions are continuous therein), we can pass the  $s$ -derivative operator under the sign of integration obtaining

$$\delta_V F|_{\gamma_0} = \sum_{i=1}^n \int_a^b \left( \frac{\partial V^i}{\partial s} \Big|_{s=0} \frac{\partial \mathcal{F}}{\partial \gamma^i} + \sum_{r=1}^k \frac{\partial^{r+1} V^i}{\partial t^r \partial s} \Big|_{s=0} \frac{\partial \mathcal{F}}{\partial \frac{d^r \gamma^i}{dt^r}} \right) dt.$$

We have interchanged the derivative in  $s$  and  $r$  derivatives in  $t$  in the first factor after the second summation symbol, it being possible by Schwarz' theorem in our hypotheses. The following identity holds

$$\int_I \frac{\partial^{r+1} V^i}{\partial t^r \partial s} \frac{\partial \mathcal{F}}{\partial \frac{d^r \gamma^i}{dt^r}} dt = \int_I (-1)^r \frac{\partial V^i}{\partial s} \frac{d^r}{dt^r} \left( \frac{\partial \mathcal{F}}{\partial \frac{d^r \gamma^i}{dt^r}} \right) dt.$$

This can be obtained by using integration by parts and dropping boundary terms in  $a$  and  $b$  which vanish because they contains factors

$$\frac{\partial^{l+1} V^i}{\partial t^l \partial s} \Big|_{t=a \text{ or } b}$$

with  $l = 0, 1, \dots, k-1$ . These factors must vanish in view of the boundary conditions on curves in  $D$ :

$$\gamma(a) = p \quad \text{and} \quad \gamma(b) = q,$$

$$\frac{d^r t \gamma}{d^r t} \Big|_a = P_r$$

and

$$\frac{d^r \gamma}{d^r t} \Big|_b = Q_r$$

for  $r = 1, \dots, k-1$  which imply that the variations of any  $\gamma_0 \in D$  with their  $t$ -derivatives in  $a$  and  $b$  up to the order  $k-1$  vanish in  $a$  and  $b$  whatever  $s \in [0, 1]$ . Then the formula in thesis follows trivially.  $\square$

A third and last lemma is in order.

**Lemma 6.34.** *Suppose that  $f : [a, b] \rightarrow \mathbb{R}^n$ , with components  $f^i : [a, b] \rightarrow \mathbb{R}$ ,  $i = 1, \dots, n$ , is continuous. If*

$$\int_a^b \sum_{i=1}^n h^i(x) f^i(x) dx = 0$$

*for every  $C^\infty$  function  $h : \mathbb{R} \rightarrow \mathbb{R}^n$  whose components  $h^i$  have supports contained in  $(a, b)$ , then  $f(x) = 0$  for all  $x \in [a, b]$ .*

**Proof.** If  $x_0 \in (a, b)$  is such that  $f(x_0) > 0$  (the case  $< 0$  is analogous), there is an integer  $j \in \{1, \dots, n\}$  and an open neighborhood of  $x_0$ ,  $U \subset (a, b)$ , where  $f^j(x) > 0$ . Exploiting the

elementary mathematical technology introduced in Section 2.3, take a function  $g \in C^\infty(\mathbb{R})$  with  $\text{supp } g \subset U$ ,  $g(x) \geq 0$  therein and  $g(x_0) = 1$ , so that, in particular,  $f^j(x_0)g(x_0) > 0$ . Shrinking  $U$ , one finds another open neighborhood of  $x_0$ ,  $U'$  whose closure is compact and  $\overline{U'} \subset U$  and  $g(x)f^j(x) > 0$  on  $\overline{U'}$ . As a consequence  $\min_{\overline{U'}} g \cdot f^j = m > 0$ .

Below,  $\chi_A$  denotes the characteristic function of a set  $A$  and  $h : (a, b) \rightarrow \mathbb{R}^n$  is defined as  $h^j = g$  and  $h^i = 0$  if  $i \neq j$ . We have

$$0 = \int_a^b \sum_{i=1}^n h^i(x) f^i(x) dx = \int_U g(x) f^j(x) dx = \int_a^b \chi_U(x) g(x) f^j(x) dx$$

because the integrand vanish outside  $U$ . On the other hand, as  $\overline{U'} \subset U$  and  $g(x)f(x) \geq 0$  in  $U$ ,

$$\chi_U(x)g(x)f^j(x) \geq \chi_{\overline{U'}}(x)g(x)f^j(x)$$

and thus

$$0 = \int_a^b \sum_{i=1}^n h^i(x) f^i(x) dx \geq \int_{\overline{U'}} g(x) f^j(x) dx \geq m \int_{\overline{U'}} dx > 0$$

because  $m > 0$  and  $\int_{\overline{U'}} dx \geq \int_{U'} dx > 0$  since non-empty open sets have strictly positive Lebesgue measure.

The found result is not possible. So  $f(x) = 0$  in  $(a, b)$  and, by continuity,  $f(a) = f(b) = 0$ .  $\square$

We conclude the general theory with two theorems.

**Theorem 6.35.** *Let  $\Omega \subset (\mathbb{R}^n)^k$  be an open non-empty connected set,  $I = [a, b]$  with  $a < b$ . Fix  $(p, P_1, \dots, P_{k-1})$  and  $(q, Q_1, \dots, Q_{k-1})$  in  $\Omega$ . Let  $D$  denote the space of elements of the set  $\{\gamma : I \rightarrow \mathbb{R}^n \mid \gamma \in C^{2k}(I)\}$  such that*

- (1)  $\left(\gamma(t), \frac{d^1\gamma}{dt}, \dots, \frac{d^{k-1}\gamma}{dt^{k-1}}\right) \in \Omega$  for all  $t \in [a, b]$ ,
- (2)  $\left(\gamma(a), \frac{d^1\gamma}{dt}|_a, \dots, \frac{d^{k-1}\gamma}{dt^{k-1}}|_a\right) = (p, P_1, \dots, P_{k-1})$  and  $\left(\gamma(b), \frac{d^1\gamma}{dt}|_b, \dots, \frac{d^{k-1}\gamma}{dt^{k-1}}|_b\right) = (q, Q_1, \dots, Q_{k-1})$ .

Finally define

$$F[\gamma] := \int_I \mathcal{F} \left( t, \gamma(t), \frac{d\gamma}{dt}, \dots, \frac{d^k\gamma}{dt^k} \right) dt$$

where  $\mathcal{F} \in C^k(I \times \Omega \times A)$  for some open non-empty set  $A \subset \mathbb{R}^n$ .

With these hypotheses,  $\gamma \in D$  is a stationary point of  $F$  if and only if it satisfies the **Euler-Poisson equations** for  $i = 1, \dots, n$ :

$$\frac{\partial \mathcal{F}}{\partial \gamma^i} + \sum_{r=1}^k (-1)^r \frac{d^r}{dt^r} \left( \frac{\partial \mathcal{F}}{\partial \frac{d^r \gamma^i}{dt^r}} \right) = 0.$$

**Proof.** It is clear that if  $\gamma \in D$  fulfills Euler-Poisson equations,  $\gamma$  is an extremal point of  $F$  because of Lemma 6.33.

If  $\gamma \in D$  is a stationary point, Lemma 6.33 entails that

$$\sum_{i=1}^n \int_I \frac{\partial V^i}{\partial s} \Big|_{s=0} \left[ \frac{\partial \mathcal{F}}{\partial \gamma^i} + \sum_{r=1}^k (-1)^r \frac{d^r}{dt^r} \left( \frac{\partial \mathcal{F}}{\partial \frac{d^r \gamma^i}{dt^r}} \right) \right] \Big|_{\gamma_0} dt = 0$$

for all variations  $V$ . We want to prove that these identities valid for every variation  $V$  of  $\gamma$  imply that  $\gamma$  satisfies E-P equations. The proof is based on lemma 6.34 with

$$f^i = \left[ \frac{\partial \mathcal{F}}{\partial \gamma^i} + \sum_{r=1}^k (-1)^r \frac{d^r}{dt^r} \left( \frac{\partial \mathcal{F}}{\partial \frac{d^r \gamma^i}{dt^r}} \right) \right] \Big|_{\gamma_0}$$

and

$$h^i = \frac{\partial V^i}{\partial s} \Big|_{s=0}.$$

Indeed, the functions  $h^i$  defined as above range in the space of  $C^\infty(\mathbb{R})$  functions with support in  $(a, b)$  as a consequence of Lemma 6.30 when using variations  $V^i(s, t) = \gamma_0^i(t) + cs\eta^i(t)$  with  $\eta^i \in C^\infty(\mathbb{R})$  supported in  $(a, b)$ . In this case  $h^i = c\eta^i$ . The condition

$$\sum_{i=1}^n \int_I \frac{\partial V^i}{\partial s} \Big|_{s=0} \left[ \frac{\partial \mathcal{F}}{\partial \gamma^i} + \sum_{r=1}^k (-1)^r \frac{d^r}{dt^r} \left( \frac{\partial \mathcal{F}}{\partial \frac{d^r \gamma^i}{dt^r}} \right) \right] \Big|_{\gamma_0} dt = 0$$

becomes

$$c \int_a^b \sum_{i=1}^n h^i(x) f^i(x) dx = 0$$

for every choice of functions  $h_i \in C^\infty((a, b))$ ,  $i = 1, \dots, n$  and for a corresponding constant  $c > 0$ . Eventually, Lemma 6.30 implies the thesis.  $\square$

**Remark 6.36.** Notice that, for  $k = 1$ , Euler-Poisson equations reduce to the well-known **Euler-Lagrange** equations interpreting  $\mathcal{F}$  as the Lagrangian of a mechanical system.  $\blacksquare$

**Theorem 6.37.** *With the same hypotheses of theorem 6.35, endow  $D$  with the topology induced by the norm*

$$\|\gamma\|_k := \max \left\{ \sup_I \|\gamma\|, \sup_I \left\| \frac{d\gamma}{dt} \right\|, \dots, \sup_I \left\| \frac{d^k \gamma}{dt^k} \right\| \right\}.$$

*If the functional  $F : D \rightarrow \mathbb{R}$  attains an extremal value at  $\gamma_0 \in D$ ,  $\gamma_0$  turns out to be a stationary point of  $F$  and it satisfies Euler-Poisson's equations.*

**Proof.** Suppose  $\gamma_0$  is a local maximum of  $F$  (the other case is similar). In that case there is an open norm ball  $B \subset D$  centered in  $\gamma_0$ , such that, if  $\gamma \in B \setminus \{\gamma_0\}$ ,  $F(\gamma) < F(\gamma_0)$ . In particular if  $V_\pm = \gamma \pm s\eta$ ,

$$\frac{F(\gamma_0 \pm cs\eta) - F(\gamma_0)}{s} < 0$$

for every choice of  $\eta \in C^\infty(\mathbb{R})$  whose components are compactly supported in  $(a, b)$  and  $s \in [0, 1]$ .  $c > 0$  is a sufficiently small constant. The limit as  $s \rightarrow 0^+$  exists by Lemma 6.33. Hence

$$\delta_{V_\pm} F|_{\gamma_0} \leq 0.$$

Making explicit the left-hand side by lemma 6.33 one finds

$$\pm \sum_{i=1}^n \int_I \eta^i \left[ \frac{\partial \mathcal{F}}{\partial \gamma^i} + \sum_{r=1}^k (-1)^r \frac{d^r}{dt^r} \left( \frac{\partial \mathcal{F}}{\partial \frac{d^r \gamma^i}{dt^r}} \right) \right] \Big|_{\gamma_0} dt \leq 0,$$

and thus

$$\sum_{i=1}^n \int_I \eta^i \left[ \frac{\partial \mathcal{F}}{\partial \gamma^i} + \sum_{r=1}^k (-1)^r \frac{d^r}{dt^r} \left( \frac{\partial \mathcal{F}}{\partial \frac{d^r \gamma^i}{dt^r}} \right) \right] \Big|_{\gamma_0} dt = 0.$$

Exploiting Lemma 6.34 as in proof of Theorem 6.35, we conclude that  $\gamma_0$  satisfies Euler-Poisson's equations. As a consequence of Theorem 6.35,  $\gamma_0$  is a stationary point of  $F$ .  $\square$

### 6.5.3 Geodesics of a metric from a variational point of view

Let us pass to consider geodesics in Riemannian and Lorentzian manifolds. We state and prove a first theorem valid for properly Riemannian metrics and involves the length functional of a smooth curve.

**Theorem 6.38.** *Let  $(M, \mathbf{g})$  be a Riemannian manifold. Take  $p, q \in M$  such that there is a common local chart  $(U, \phi)$ ,  $\phi(r) = (x^1(r), \dots, x^n(r))$ , with  $p, q \in U$ , assuming  $U$  connected. Fix  $[a, b] \subset \mathbb{R}$ ,  $a < b$  and consider the **length functional***

$$L[\gamma] = \int_a^b \sqrt{g_{ij}(\gamma(t)) \frac{dx^i(\gamma(t))}{dt} \frac{dx^j(\gamma(t))}{dt}} dt,$$

*defined on the space  $S$  of smooth curves  $\gamma : [a, b] \rightarrow U$  ( $U$  being identified to the open set  $\phi(U) \subset \mathbb{R}^n$ ) with  $\gamma(a) = p$ ,  $\gamma(b) = q$  and everywhere non-vanishing tangent vector  $\gamma'$ .*

- (a) *If  $\gamma_0 \in S$  is a stationary point of  $L$ , there is a smooth bijection with smooth inverse,  $u : [0, L[\gamma_0]] \rightarrow [a, b]$ , such that  $\gamma \circ u$  is a geodesic with respect to the Levi-Civita connection connecting  $p$  to  $q$ .*
- (b) *If  $\gamma_0 \in S$  is a geodesic (connecting  $p$  to  $q$ ),  $\gamma_0$  is a stationary point of  $L$ .*

**Proof.** First of all, notice that the domain  $S$  of  $L$  is not empty ( $M$  is connected and thus path connected by definition) and  $S$  belongs to the class of domains  $D$  used in Theorem 6.35: now  $\Omega = \phi(U) \times (\mathbb{R}^n \setminus \{0\})$ .  $L$  itself is a specialization of the general functional  $F$  and the associated function  $\mathcal{F}$  is  $C^\infty$  (indeed the function  $x \mapsto \sqrt{x}$  is  $C^\infty$  in the domain  $\mathbb{R} \setminus \{0\}$ ).

(a) By Theorem 6.35, if  $\gamma_0 \in S$  is a stationary point of  $F$ ,  $\gamma_0$  satisfies in  $[a, b]$ :

$$\frac{d}{dt} \left[ \frac{g_{ki} \frac{dx^i}{dt}}{\sqrt{g_{rs} \frac{dx^r}{dt} \frac{dx^s}{dt}}} \right] - \frac{\frac{1}{2} \frac{\partial g_{ij}}{\partial x^k} \frac{dx^i}{dt} \frac{dx^j}{dt}}{\sqrt{g_{rs} \frac{dx^r}{dt} \frac{dx^s}{dt}}} = 0, \quad (6.21)$$

where  $x^i(t) := x^i(\gamma_0(t))$  and the metric  $g_{lm}$  is evaluated on  $\gamma_0(t)$ .

Since  $\gamma_0'(t) \neq 0$  and the metric is positive,  $g_{rs}(\gamma_0(t)) \frac{dx^r}{dt} \frac{dx^s}{dt} \neq 0$  in  $[a, b]$  and the function

$$s(t) := \int_a^t \sqrt{g_{rs}(\gamma_0(t)) \frac{dx^r}{dt} \frac{dx^s}{dt}} dt$$

takes values in  $[0, L[\gamma_0]]$  and, by trivial application of the fundamental theorem of calculus, is differentiable, injective with inverse differentiable. Let us indicate by  $u : [0, L[\gamma_0]] \rightarrow [a, b]$  the inverse function of  $s$ . By (6.21), the curve  $s \mapsto \gamma(u(s))$  satisfies the equations

$$\frac{d}{ds} \left[ g_{ki} \frac{dx^i}{ds} \right] - \frac{1}{2} \frac{\partial g_{ij}}{\partial x^k} \frac{dx^i}{ds} \frac{dx^j}{ds} = 0.$$

Expanding the derivative we get

$$\frac{d^2 x^i}{ds^2} g_{ki} + \frac{\partial g_{ki}}{\partial x^j} \frac{dx^i}{ds} \frac{dx^j}{ds} - \frac{1}{2} \frac{\partial g_{ij}}{\partial x^k} \frac{dx^i}{ds} \frac{dx^j}{ds} = 0.$$

These equations can be re-written as

$$\frac{d^2 x^i}{ds^2} g_{ki} + \frac{1}{2} \left[ \frac{\partial g_{ki}}{\partial x^j} \frac{dx^i}{ds} \frac{dx^j}{ds} + \frac{\partial g_{kj}}{\partial x^i} \frac{dx^j}{ds} \frac{dx^i}{ds} - \frac{\partial g_{ij}}{\partial x^k} \frac{dx^i}{ds} \frac{dx^j}{ds} \right] = 0.$$

Contracting with  $g^{rk}$  these equations become

$$\frac{d^2 x^r}{ds^2} + \frac{1}{2} g^{rk} \left[ \frac{\partial g_{ki}}{\partial x^j} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} \right] \frac{dx^i}{ds} \frac{dx^j}{ds} = 0,$$

which can be re-written as the geodesic equations with respect to the Levi-Civita connection:

$$\frac{d^2 x^r}{ds^2} + \{^r_{ij}\} \frac{dx^i}{ds} \frac{dx^j}{ds} = 0.$$

(b) A curve from  $p$  to  $q$ ,  $t \mapsto \gamma(t)$ , can be re-parametrized by its length parameter:  $s = s(t)$ ,  $s \in [0, L[\gamma]]$  where  $s(t) \in [0, L(\gamma_0)]$  is the length of the curve  $\gamma_0$  evaluated from  $p$  to  $\gamma(t)$ . In that case it holds

$$\int_0^s \sqrt{g_{rl}(\gamma_0(t(s)))} \frac{dx^r}{ds} \frac{dx^l}{ds} ds = s$$

and thus

$$\sqrt{g_{rl}(\gamma_0(t(s)))} \frac{dx^r}{ds} \frac{dx^l}{ds} = 1.$$

Now suppose that  $t \mapsto \gamma_0(t)$  is a geodesic. Thus  $t \in [a, b]$  is an affine parameter: there are  $c, d \in \mathbb{R}$  with  $c > 0$  such that  $t = cs + d$ . As a consequence

$$\sqrt{g_{rl}(\gamma_0(t))} \frac{dx^r}{dt} \frac{dx^l}{dt} = \frac{1}{c} \sqrt{g_{rl}(\gamma_0(t(s)))} \frac{dx^r}{ds} \frac{dx^l}{ds} \quad (6.22)$$

and thus

$$\sqrt{g_{rl}(\gamma_0(t))} \frac{dx^r}{dt} \frac{dx^l}{dt} = \frac{1}{c}. \quad (6.23)$$

Following the proof of (a) with reversed order, one finds that

$$\frac{d^2 x^r}{dt^2} + \{i \ r \ j\} \frac{dx^i}{dt} \frac{dx^j}{dt} = 0$$

yields

$$\frac{d}{dt} \left[ g_{ki} \frac{dx^i}{dt} \right] - \frac{1}{2} \frac{\partial g_{ij}}{\partial x^k} \frac{dx^i}{dt} \frac{dx^j}{dt} = 0,$$

or, since  $c > 0$ ,

$$c \frac{d}{dt} \left[ g_{ki} \frac{dx^i}{dt} \right] - c \frac{1}{2} \frac{\partial g_{ij}}{\partial x^k} \frac{dx^i}{dt} \frac{dx^j}{dt} = 0.$$

Noticing that  $c$  is constant and taking advantage of (6.23), these equations are equivalent to Euler-Poisson equations

$$\frac{d}{dt} \left[ \frac{g_{ki} \frac{dx^i}{dt}}{\sqrt{g_{rs} \frac{dx^r}{dt} \frac{dx^s}{dt}}} \right] - \frac{\frac{1}{2} \frac{\partial g_{ij}}{\partial x^k} \frac{dx^i}{dt} \frac{dx^j}{dt}}{\sqrt{g_{rs} \frac{dx^r}{dt} \frac{dx^s}{dt}}} = 0,$$

and this concludes the proof by Theorem 6.35.  $\square$

We can generalize the theorem to the case of a Lorentzian manifold.

**Theorem 6.39.** *Let  $(M, \mathbf{g})$  a Lorentzian manifold. Take  $p, q \in M$  such that there is a common local chart  $(U, \phi)$ ,  $\phi(r) = (x^1(r), \dots, x^n(r))$ , with  $p, q \in U$ . Fix  $[a, b] \subset \mathbb{R}$ ,  $a < b$  and consider the **timelike length functional**:*

$$L_T[\gamma] = \int_a^b \sqrt{\left| g_{ij}(\gamma(t)) \frac{dx^i(\gamma(t))}{dt} \frac{dx^j(\gamma(t))}{dt} \right|} dt,$$

*defined on the space  $S_T$  of smooth curves  $\gamma : [a, b] \rightarrow U$  ( $U$  being identified to the open set  $\phi(U) \subset \mathbb{R}^n$ ) with  $\gamma(a) = p$ ,  $\gamma(b) = q$  and  $\gamma$  is timelike, i.e.  $\mathbf{g}(\gamma', \dot{\gamma}) < 0$  everywhere.*

*Suppose that  $p$  and  $q$  are such that  $S_T \neq \emptyset$ .*

(a) If  $\gamma_0 \in S_T$  is a stationary point of  $L_T$ , there is a differentiable bijection with inverse differentiable,  $u : [0, L_T[\gamma_0]] \rightarrow [a, b]$ , such that  $\gamma \circ u$  is a timelike geodesic with respect to the Levi-Civita connection connecting  $p$  to  $q$ .

(a) If  $\gamma_0 \in S_T$  is a timelike geodesic (connecting  $p$  to  $q$ ),  $\gamma_0$  is a stationary point of  $L_T$ .

**Proof.** The proof is the same as the one of Theorem 6.38 with the specification that  $S_T$ , if non-empty, is a domain of the form  $D$  used in Theorem 6.35. In particular the set  $\Omega \subset \mathbb{R}^{2n}$  used in the definition of  $D$  is now the open set:

$$\{(x^1, \dots, x^n, v^1, \dots, v^n) \in \mathbb{R}^{2n} \mid (x^1, \dots, x^n) \in \phi(U) , (g_{\phi^{-1}(x^1, \dots, x^n)})_{ij} v^i v^j < 0\}$$

where  $g_{ij}$  represent the metric in the coordinates associated with  $\phi$ .  $\square$

**Theorem 6.40.** Let  $(M, \mathbf{g})$  be a Lorentzian manifold. Take  $p, q \in M$  such that there is a common local chart  $(U, \phi)$ ,  $\phi(r) = (x^1(r), \dots, x^n(r))$ , with  $p, q \in U$ . Fix  $[a, b] \subset \mathbb{R}$ ,  $a < b$  and consider the **spacelike length functional**:

$$L_S[\gamma] = \int_a^b \sqrt{g_{ij}(\gamma(t)) \frac{dx^i(\gamma(t))}{dt} \frac{dx^j(\gamma(t))}{dt}} dt ,$$

defined on the space  $S_S$  of smooth curves  $\gamma : [a, b] \rightarrow U$  ( $U$  being identified to the open set  $\phi(U) \subset \mathbb{R}^n$ ) with  $\gamma(a) = p$ ,  $\gamma(b) = q$  and  $\gamma$  is spacelike, i.e.  $\mathbf{g}(\gamma', \dot{\gamma}) > 0$  everywhere.

Suppose that  $p$  and  $q$  are such that  $S_S \neq \emptyset$ .

(a) If  $\gamma_0 \in S_S$  is a stationary point of  $L_S$ , there is a differentiable bijection with inverse differentiable,  $u : [0, L_S[\gamma_0]] \rightarrow [a, b]$ , such that  $\gamma \circ u$  is a spacelike geodesic with respect to the Levi-Civita connection connecting  $p$  to  $q$ .

(b) If  $\gamma_0 \in S_S$  is a spacelike geodesic (connecting  $p$  to  $q$ ),  $\gamma_0$  is a stationary point of  $L_S$ .

**Proof.** Once again the proof is the same as the one of Theorem 6.38 with the specification that  $S_S$ , if non-empty, is a domain of the form  $D$  used in Theorem 6.35. In particular the set  $\Omega \subset \mathbb{R}^{2n}$  used in the definition of  $D$  is now the open set:

$$\{(x^1, \dots, x^n, v^1, \dots, v^n) \in \mathbb{R}^{2n} \mid (x^1, \dots, x^n) \in \phi(U) , (g_{\phi^{-1}(x^1, \dots, x^n)})_{ij} v^i v^j > 0\}$$

where  $g_{ij}$  represent the metric in the coordinates associated with  $\phi$ .  $\square$

As a final result, we prove the following theorem which is valid both for Riemannian and Lorentzian manifolds. More precisely we consider a generic (pseudo) Riemannian case.

**Theorem 6.41.** Let  $(M, \mathbf{g})$  be a (pseudo) Riemannian manifold. Take  $p, q \in M$  such that there is a common local chart  $(U, \phi)$ ,  $\phi(r) = (x^1(r), \dots, x^n(r))$ , with  $p, q \in U$ , assuming  $U$  connected. Fix  $[a, b] \subset \mathbb{R}$ ,  $a < b$  and consider the **energy functional**:

$$E[\gamma] = \int_a^b g_{ij}(\gamma(s)) \frac{dx^i(\gamma(s))}{ds} \frac{dx^j(\gamma(s))}{ds} ds ,$$

defined on the space  $S$  of smooth curves  $\gamma : [a, b] \rightarrow U$  ( $U$  being identified to the open set  $\phi(U) \subset \mathbb{R}^n$ ) with  $\gamma(a) = p$ ,  $\gamma(b) = q$  and everywhere non-vanishing tangent vector  $\gamma'$ .

The curve  $\gamma_0 \in S$  is a stationary point of  $L$  if and only if it is a geodesic with respect to the metric  $\mathbf{g}$ , and the parameter  $s$  is an affine parameter.

**Proof.** The equations determining the stationary points read

$$\frac{d}{ds} \left[ g_{ki} \frac{dx^i}{ds} \right] - \frac{1}{2} \frac{\partial g_{ij}}{\partial x^k} \frac{dx^i}{ds} \frac{dx^j}{ds} = 0.$$

Expanding the derivative we get

$$\frac{d^2 x^i}{ds^2} g_{ki} + \frac{\partial g_{ki}}{\partial x^j} \frac{dx^i}{ds} \frac{dx^j}{ds} - \frac{1}{2} \frac{\partial g_{ij}}{\partial x^k} \frac{dx^i}{ds} \frac{dx^j}{ds} = 0.$$

Dealing with as in the proof of Theorem 6.38, these equations can be proved to be equivalent to the geodesic equations with respect to the Levi-Civita connection:

$$\frac{d^2 x^r}{ds^2} + \{i \ r \ j\} \frac{dx^i}{ds} \frac{dx^j}{ds} = 0.$$

□

**Exercises 6.42.** Show that the sets  $\Omega$  used in the proof of Theorem 6.39 and Theorem 6.40 are open in  $\mathbb{R}^{2n}$ .

(*Hint.* Prove that, in both cases  $\Omega = f^{-1}(E)$  where  $f$  is some continuous function on some appropriate space and  $E$  is some open set in that space.)

**Remark 6.43.**

(1) Working in  $TM$ , the four theorems proved above can be generalized by dropping the hypotheses of the existence of a common local chart  $(U, \phi)$  containing the differentiable curves.

(2) There is no guarantee of having a geodesic joining any pair of points in a (pseudo) Riemannian manifold. For instance consider the Euclidean space  $\mathbb{E}^2$  (see example 5.8.1), and take  $p, q \in \mathbb{E}^2$  with  $p \neq q$ . As everybody knows, there is exactly a geodesic segment  $\gamma$  joining  $p$  and  $q$ . If  $r \in \gamma$  and  $r \neq p$ ,  $r \neq q$ , the space  $M \setminus \{r\}$  is anyway a Riemannian manifold globally flat. However, in  $M$  there is no geodesic segment joining  $p$  and  $q$ .

As a general result, it is possible to show that in a (pseudo) Riemannian manifold, if two points are sufficiently close to each other there is at least one geodesic segments joining the points.

(3) There is no guarantee for having a *unique* geodesic connecting a pair of points in a (pseudo) Riemannian manifold if one geodesic at least exists. For instance, on a 2-sphere  $S^2$  with the metric induced by  $\mathbb{E}^3$ , there are infinite many geodesic segments connecting the north pole with the south pole. ■



## Chapter 7

# The Exponential Map of Affine and Metric Connections

We introduce in this chapter an important geometric tool called *exponential map* which, in its most elementary version, allows one to identify locally the manifold with the tangent space at  $p$ . It moreover provides a coordinate system such that the connection coefficients vanishes exactly at  $p$  whenever the connection is torsion-free. This mathematical tool is of central relevance in General Relativity since it permits to state a mathematically rigorous geometric version of *Einstein's equivalence principle* we shall introduce in the next chapter

### 7.1 Exponential map at a point

Let us start with discussing the exponential map at a point and some associated notions.

#### 7.1.1 The exponential map

Consider a smooth manifold  $M$  equipped with an affine connection  $\nabla$ . We have the first important definition.

**Definition 7.1.** (**Exponential map.**) Let  $M$  be a smooth manifold equipped with the affine connection  $\nabla$  and let denote by  $\gamma_{p,v} : I_{p,v} \rightarrow M$  the unique (maximal) geodesic with initial conditions  $\gamma_{p,v}(0) = p$ ,  $\gamma'_{p,v}(0) = v$ .

(1) If  $D_p E := \{v \in T_p M \mid I_{p,v} \ni 1\}$ , the map

$$\exp_p : D_p E \ni v \mapsto \gamma_{p,v}(1) \in M \quad (7.1)$$

is called **exponential map** at  $p$ .

(1) If  $DE := \{(p, v) \in TM \mid v \in D_p E\}$ , the map

$$\exp : DE \ni (p, v) \mapsto \gamma_{p,v}(1) \in M \quad (7.2)$$

is called **exponential map** on  $M$ . ■

We pass to prove that  $DE, D_pE \neq \emptyset$  which are also open sets. Furthermore  $\exp$  and  $\exp_p$  are smooth on the respective domains. Later, we will prove that, if further restrictions are imposed to the domain  $D_pE$ , then  $\exp_p$  becomes a local diffeomorphism from  $T_pM$  to  $M$ , so that points around  $p$  are smoothly one-to-one with vectors in  $T_pM$  around the zero vector.

We recall that a set  $B$  in an affine space is said to be **star-shaped** with **center**  $q \in B$ , if the unique segment joining  $q$  to every  $r \in B \setminus \{q\}$  is completely contained in  $B$ . An open ball in  $\mathbb{R}^n$  is a star-shaped set with respect to every point inside it.

**Theorem 7.2.** *Let  $M$  be a smooth manifold equipped with the affine connection  $\nabla$ .*

- (a)  *$DE$  is an open nonempty set of  $TM$  and  $\exp : DE \rightarrow M$  is smooth;*
- (b)  *$D_pE$  is an open nonempty set of  $T_pM$  for every  $p \in M$  and  $\exp : D_pE \rightarrow M$  is smooth.*

More precisely,

- (c)  *$DE$  is an open neighborhood of the trivial section  $M \times \{0\} \subset TM$ ;*
- (d)  *$D_pE$  is an open neighborhood of  $0 \in T_pM$  for every  $p \in M$ .*

Finally, for every  $p \in M$ ,

- (e)  *$v \in D_pE$  implies  $\lambda v \in D_pM$  so that  $D_pM$  is a star-shaped set with center  $0 \in T_pM$ ;*
- (f) *if  $v \in D_pE$ , then*

$$\exp_p(\lambda v) = \gamma_{p,v}(\lambda), \quad \text{for } \lambda \in [0, 1], \quad (7.3)$$

where  $I \ni \lambda \mapsto \gamma_{p,v}(\lambda)$  be the unique (maximal) geodesic with  $\gamma_{p,v}(0) = p$  and  $\gamma'_{p,v}(0) = v$ .

**Proof.** First of all consider the domain flow  $\Phi^{(\Gamma)}$  of the smooth vector field  $\Gamma$  on  $TM$  defined in Section 6.4, whose integral curves projected onto  $M$  are the (maximal) geodesics of  $M$ . This domain an open subset of  $TM \times \mathbb{R}$  according to Proposition 4.16. If we intersect this set with  $TM \times \{1\}$  we still have an open set<sup>1</sup> in  $TM$  and the restriction of the flow is still smooth thereon. Hence its composition with the canonical projection  $\Pi : TM \rightarrow M$  is still smooth (since  $\Pi$  is smooth). Since  $\Pi(\Phi^{(\Gamma)}) = \exp$ , we have established that  $DE$  is an open set of  $TM$  and  $\exp$  is thereon smooth. When intersecting  $DE$  with  $T_pM$  we still obtain an open set and the restriction  $\exp_p$  of  $\exp$  thereon remains smooth. We have so far established (a) and (b).

Let us prove (c) and (d). If  $\phi : U \ni q \mapsto (x^1, \dots, x^n)$  is a local chart on  $M$ , consider the associated natural coordinate patch on  $TM$  with coordinates

$$T\phi : TU \ni (q, v) \mapsto (x^1, \dots, x^n, v^1, \dots, v^n) \in \phi(U) \times \mathbb{R}^n,$$

---

<sup>1</sup>Make use of (6) in Exercises 2.1, noticing that we are dealing with spaces which (at least locally) are topological products.

and recast the system of differential equations defining geodesic curves  $I \ni t \mapsto \gamma(t)$  using that local chart as we did in Section 6.4

$$\frac{dv^a}{dt} = -\Gamma_{bc}^a(x^1(t), \dots, x^n(t))v^b(t)v^c(t), \quad (7.4)$$

$$\frac{dx^a}{dt} = v^a(t). \quad (7.5)$$

Above  $I \ni t \mapsto (x^1(t), \dots, x^n(t)) = \phi(\gamma(t))$  and  $\gamma'(t) = v^k(t) \frac{\partial}{\partial x^k} |_{\gamma(t)}$ .

If  $p \in U$ ,  $v \in T_p M$ , and  $t$  belongs to some open interval  $I_{p,v} \ni 0$  depending on  $p$  and  $v_p$ , let us indicate by

$$\gamma = \gamma(p, v, t)$$

the unique geodesic with initial conditions  $\gamma(0) = p$  and  $\gamma'(0) = v$ . In local coordinates, it determines a maximal solution in  $TU$  of (7.4)-(7.5) with the said initial conditions. Theorem 6.24 and the topology of  $TM$ , imply in particular that, if we fix  $q \in U$ , then there exists an open set of the form

$$V_r \times B_\delta(0) \times (-\epsilon, \epsilon) \subset \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$$

where

- (1)  $V_r \subset \phi(U)$  is an open ball of radius  $r > 0$  centered on  $\phi(q)$ ,
- (2)  $B_\delta(0)$  is an open ball of radius  $\delta > 0$  centered on the origin,
- (3)  $\epsilon > 0$ ,

such that

$$(T\phi)^{-1}(V_r \times B_\delta(0)) \times (-\epsilon, \epsilon) \ni (p, v, t) \mapsto \gamma(p, v, t)$$

is well defined and smooth. We want to prove that it is possible increase  $\epsilon$  and decrease  $\delta$  in order that the domain of the arising geodesic segments includes  $t = 1$ .

The uniqueness theorem entails that, for every  $\lambda > 0$ , the two geodesic segments

$$(-\epsilon, \epsilon) \ni t \mapsto \gamma(p, \lambda v, t) \quad \text{and} \quad (-\epsilon/\lambda, \epsilon/\lambda) \ni t \mapsto \gamma(p, v, \lambda t)$$

coincide, since both pass through  $p$  for  $t = 0$  and both have the same initial tangent vector  $\lambda v$  at  $t = 0$ . Hence

$$\gamma(p, \lambda v, t) = \gamma(p, v, \lambda t). \quad (7.6)$$

We can now fix  $\lambda > 0$  sufficiently small in order to have  $\epsilon' := \epsilon/\lambda > 1$ , concluding that, if  $(p, u) \in (T\phi)^{-1}(V_r \times B_{\delta'}(0))$  with  $\delta' = \lambda\delta$ , the map:

$$(T\phi)^{-1}(V_r \times B_{\delta'}(0)) \ni (p, v) \mapsto \gamma(p, v, t)$$

is well-defined for  $t \in (-\epsilon', \epsilon') \supset (-1, 1)$ . Obviously the function

$$(T\phi)^{-1}(V_r \times B_{\delta'}(0)) \ni (p, u_p) \mapsto \exp_p(u_p) = \gamma(p, u_p, 1) \quad (7.7)$$

is well defined and jointly smooth.

We have in particular established that  $DE$  includes open sets of the form  $(T\phi)^{-1}(V_{r_q} \times B_{\delta'_q}(0))$ , where  $V_{r_q}$  is an open set containing  $q \in M$ . In particular  $(T\phi)^{-1}(V_{r_q} \times B_{\delta'_q}(0)) \ni (q, 0) \in TM$ . Since all that is valid for every  $q \in M$ , we can assert that the open set  $DE$  is an open neighborhood of the trivial section  $M \times \{0\} \subset TM$  and evidently  $D_pM$  is an open neighborhood of  $0 \in T_pM$ .

We pass to (e) and (f). If  $v \in D_pE$ , then the domain of  $\gamma(p, v, \cdot)$ , which is of the form  $(-a, b)$ , for  $a, b > 0$  possibly infinite, includes 1. As a consequence, the domain of  $\gamma(p, \lambda v, \cdot)$ , which is of the form  $(-a/\lambda, b/\lambda)$  still includes 1 if  $|\lambda| \leq 1$  so that  $\lambda v \in D_pE$ . We can therefore apply (7.6) with  $t = 1$  finding

$$\exp_p(\lambda u) = \gamma(p, \lambda u, 1) = \gamma(p, u, \lambda),$$

which is the thesis.  $\square$

A special case of the theorem above is stated in the following corollary.

**Proposition 7.3.** *A manifold  $M$  equipped with an affine connection is geodesically complete if and only if  $DE = TM$  and in that case  $\exp : TM \rightarrow M$  is everywhere smooth.*

**Proof.** If  $M$  is geodesically complete, then  $DE = TM$  evidently. *Vice versa*, if  $DE = TM$  we conclude that the manifold is complete in view of the identity  $\gamma_{p,v}(\lambda) = \gamma_{p,\lambda v}(1)$ , since the right hand side is defined for every  $(p, \lambda v) \in TM$ .  $\square$

**Remark 7.4.** Theorem 7.15 below establishes some other features of Riemannian complete manifolds.  $\blacksquare$

We are in a position to state and prove the most important basic result concerning the local smooth identification of  $M$  and  $T_pM$ .

**Theorem 7.5.** *Let  $M$  be a smooth manifold equipped with the affine connection  $\nabla$  and  $p \in M$ . There is a open set  $B \subset D_pE$  – which is an open neighborhood of  $0 \in T_pM$  – and an open set  $N_p \subset M$  with  $p \in N_p$  such that*

$$\exp_p|_B : B \rightarrow N_p$$

*is a diffeomorphism (where  $B$  is endowed with the natural smooth structure induced by  $T_pM$  viewed as an affine space).*

**Proof.** According to (3) in Theorem 4.5, it is sufficient to prove that  $d\exp_p|_{v=0} : T_0T_pM \rightarrow T_pM$  is bijective. Now observe that, if  $f : N \rightarrow M$  is smooth and  $\eta : (-\alpha, \alpha) \rightarrow N$  is smooth and satisfies  $\eta'(0) = u$ , then

$$df_{\eta(0)}u = \frac{d}{d\lambda}|_{\lambda=0}f(\eta(\lambda)).$$

Applying this result to the case  $N = T_p M$ ,  $M = M$ ,  $f = \exp_p$ ,  $\eta(\lambda) = \lambda u$ , we have

$$d \exp_p |_0 u = \frac{d}{d\lambda} |_{\lambda=0} \exp_p(\lambda u) = \frac{d}{d\lambda} |_{\lambda=0} \gamma(p, u, \lambda) = u ,$$

where we have used (7.3)

$$\exp_p(\lambda u) = \gamma(p, u, \lambda) .$$

Since  $d \exp_p |_0 = T_0 T_p M \equiv T_p M \ni u \mapsto u \in T_p M$  the proof is over.  $\square$

### 7.1.2 Normal neighborhoods, geodesic balls, injectivity radius

Another natural mathematical notion we introduce is that of *normal neighborhood* of a point. It arises if compining (e) and (f) of Theorem 7.2 with Theorem 7.5.

**Definition 7.6.** (**Normal neighborhoods of a point**) Let  $M$  be a smooth manifold equipped with the affine connection  $\nabla$  and  $p \in M$ . A **normal neighborhood** of  $p$  is an open set of the form  $\exp_p(B)$ , such that

- (a)  $B \subset T_p M$  is an open star-shaped set with center  $0 \in T_p M$ ;
- (b)  $\exp_p |_B$  is a diffeomorphism.

The point  $p$  is called the **center** of the normal neighborhood. ■

**Remark 7.7.**

(1) As a trivial consequence of the definition, if  $U = \exp_p(B)$  is a normal neighborhood of  $p$  and  $B' \subset B$  is another open star-shaped neighborhood with center  $0 \in T_p M$ , then also  $U' = \exp_p(B')$  is a normal neighborhood of  $p$ .

(2) If  $p \in M$  and we fix a (positive) scalar product in  $T_p M$  (even if there is no metric on  $M$ , but there is a connection), it is evident that the open balls centered on the origin of  $T_p M$  of radius sufficiently small give rise to normal neighborhoods of with center  $p$  as a trivial consequence of the fact that all metric topologies in finite dimensional vector spaces are equivalent. ■

(2) in Remark 7.7 implies the following pair of technically interestig fact.

**Proposition 7.8.** *The class of normal neighborhoods of the points of a smooth manifold equipped with an affine connection is a basis of the topology.*

*Proof.* The open balls centered on the origin of  $T_p M$  of radius sufficiently small give rise, under the action of the local diffeomorphism  $\exp_p$ , to normal neighborhoods of with center  $p$ . Since it is valid for every  $p$  the assertion is evident. □

**Proposition 7.9.** *A connected smooth manifold  $M$  equipped with an affine connection  $\nabla$  is connected by piecewise geodesic curves. In other words, if  $p, q \in M$  there is a continuous curve*

$\gamma : [a, b] \rightarrow M$  with  $\gamma(a) = p$  and  $\gamma(b) = q$  and  $[a, b] = [a_1, a_2] \cup [a_2, a_3] \cup \cdots \cup [a_{n-1}, a_n]$ , where  $a_1 = a$ ,  $a_n = b$  and every  $\gamma|_{[a_k, a_{k+1}]}$  is a geodesic segment.

**Proof.** Let  $A \subset M$  the set of points which are connected to a given  $p \in M$  with some piecewise geodesic path. It is easy to prove that  $A$  is open, because if  $q \in A$  and  $\gamma$  connects  $p$  and  $q$ , dealing with a normal neighborhood  $N_q$  centered on  $q$ , every point  $q' \in N_q$  can be connected to  $p$  by adding geodesic segment  $q$  to  $q'$ . With this procedure one has a piecewise geodesic from  $p$  to  $q'$ . With an analogous argument, the set  $B \subset M$  which are not connected by piecewise geodesics to  $p$  is proved to be open as well. In summary,  $M = A \cup B$  where  $A$  and  $B$  are open disjoint sets. Since  $M$  is connected, it must be either  $M = A$  and  $B = \emptyset$  or  $M = B$  and  $A = \emptyset$ . The second possibility is not allowed as  $p \in A$ , hence  $M = A$ .  $\square$

To go on, let us specialize the considered manifold to a Riemannian one.

**Definition 7.10.** (**Geodesical ball**) If  $(M, \mathbf{g})$  is a Riemannian manifold

$$\exp_p(B_r(0)) \subset M$$

is called **geodesic ball** of **radius**  $r$  and **center**  $p$  if

$$B_r(0) := \left\{ v \in T_p M \mid \sqrt{\mathbf{g}_p(v, v)} < r \right\}, \quad (7.8)$$

where  $r > 0$ . ■

If  $(M, \mathbf{g})$  is Riemannian, it should be evident, according to (2) in Remark 7.7, that a geodesic ball centered on  $p$  with sufficiently small radius is a normal neighborhood of  $p$ . This fact is embodied in the following definition.

**Definition 7.11.** (**Injectivity radius**) If  $(M, \mathbf{g})$  is Riemannian and  $\nabla$  is the Levi-Civita connection.

(a) The **injectivity radius at**  $p$  is

$$Inj_p(M) := \sup \left\{ r > 0 \mid \exp_p(B_r(0)) \text{ is a normal neighborhood of } p \right\}.$$

where definition (7.8)

$$B_r(0) := \left\{ v \in T_p M \mid \sqrt{\mathbf{g}_p(v, v)} < r \right\}$$

has been adopted.

(b) The **injectivity radius of**  $M$  is

$$Inj(M) := \inf_{p \in M} Inj_p(M).$$

**Remark 7.12.**

(1) If  $M$  is compact and equipped with a smooth (properly) Riemannian metric, then its injectivity radius is strictly positive (see (4) in Exercises 7.38). This property is also shared with other types of manifolds called *of bounded geometry* which occupy and intermediate place between compact Riemannian manifolds and generic Riemannian manifolds.

(2) It turns out that if a Riemannian manifold has strictly positive injectivity radius then it is complete and thus  $\exp$  is smooth and everywhere defined on  $TM$  (see (5) in Exercises 7.38). ■

Normal neighborhoods are star-shaped with respect to geodesics in the sense of the following important technical theorem.

**Theorem 7.13.** *Let  $M$  be a smooth manifold equipped with the affine connection  $\nabla$  and  $p \in M$ . If  $U = \exp_p(B) \ni p$  is a normal neighborhood of  $p$  and  $q \in U \setminus \{p\}$ , then*

- (a) *there is a geodesic segment  $\gamma : [0, 1] \rightarrow U$  with  $\gamma(0) = p$  and  $\gamma(1) = q$ ;*
- (b) *the geodesic segment in (a) is unique and  $\gamma'(0) = (\exp_p|_B)^{-1}(q)$ .*

**Proof.** (a) By definition  $U = \exp_p(B)$  for some star-shaped open neighborhood  $B$  of center  $0 \in T_pM$ . If  $q \in U \setminus \{p\}$ , then  $(\exp_p|_B)^{-1}(q) =: v \in B$  and the segment  $[0, 1] \ni t \mapsto tv$  stays in  $B$  (it being star-shaped) and thus its image through  $\exp_p$  is completely included in  $U$  by construction.  $[0, 1] \ni t \mapsto \exp_p(tv) \in U$  is nothing but a geodesic segment joining  $p$  ( $t = 0$ ) and  $q$  ( $t = 1$ ) by Theorem 7.2.

(b) If  $\eta : I \rightarrow U$  is a geodesic segment such that  $\eta(0) = p$  and  $\eta(1) = q$ , since  $1 \in I$ , by definition,  $u := \eta'(0) \in D_pE$  so that (7.3) is valid, we also have

$$\eta(t) = \exp_p(tu) \quad \text{for } t \in [0, 1].$$

Comparing with the geodesic segments  $[0, 1] \ni t \mapsto \exp_p(tv)$ , with  $v \in B$ , constructed in (a) we have in particular  $\exp_p(1v) = \exp_p(1u) = q$ . This implies  $v = u$  – as well as the uniqueness of the said geodesic – just by applying  $(\exp_p|_B)^{-1}$  to both sides by noticing that also  $u \in B$ .

(We prove<sup>2</sup> that  $u \in B$ . As  $B \subset T_pM$  is star-shaped with center 0, if  $u \notin B$ , then there is  $t_0 \in [0, 1]$  such that  $tu \in B$  for  $t < t_0$  and  $tu \notin B$  for  $t > t_0$ , and  $t_0u \notin B$  because  $B$  is open and  $t_0u \in \partial B$ . Then consider a sequence  $B \ni t_nu \rightarrow t_0u$ . Since  $\eta : [0, 1] \rightarrow U$  is continuous,  $\eta(t_n) \rightarrow \eta(t_0) \in U$ . Since  $\eta(t_n), \eta(t_0) \in U = \exp_p|_B(B)$  and  $\exp_p|_B \rightarrow U$  is a diffeomorphism,

$$(\exp_p|_B)^{-1}(\eta(t_n)) \rightarrow (\exp_p|_B)^{-1}(\eta(t_0)) \in B.$$

As  $t_nu \in B$ , we find a contradiction

$$B \not\ni t_0u \leftarrow t_nu = (\exp_p|_B)^{-1}(\exp_p(t_nu)) = (\exp_p|_B)^{-1}(\eta(t_n)) \rightarrow (\exp_p|_B)^{-1}(\eta(t_0)) \in B.$$

---

<sup>2</sup>I add this proof because it is a bit more subtle than I expected and I cannot understand the proof of this point in O'Neill's book [O'Ne83].

□

**Remark 7.14.** We stress that the uniqueness property discussed in the Proposition above is strictly related to the requirement that the complete segment stays in the normal neighborhood. The result could be false if dropping this condition. Think of the unit-radius sphere  $S^2$  embedded in  $\mathbb{R}^3$  with the induced metric. The south hemisphere is a normal neighborhood of the south pole. If  $Q$  stays in the south hemisphere, then there is a unique geodesic joining it with the south pole and remaining in the south hemisphere. However there is another geodesic connecting the two points which passes through the north pole! The injectivity radius of  $S^2$  is, in fact,  $\pi$ . ■

### 7.1.3 Hopf-Rinow's theorem

To conclude, we state without proof [KoNo96, doC92] a fundamental result of Riemannian manifolds theory known as the *Hopf-Rinow theorem* (there are related but much weaker results in Lorentzian geometry [BEE96, Min19]).

**Theorem 7.15.** (**Hopf-Rinow's theorem.**) *Let  $(M, \mathbf{g})$  a Riemannian manifold. The following facts are equivalent.*

- (a) *The manifold is geodesically complete.*
- (b)  *$M$  equipped with the distance  $d_g$  induced by the metric (5.2) is a complete metric space.*
- (c) *Closed metrically bounded sets are compact.*

*If the previous properties are valid, then for every pair  $p, q \in M$  there exist a geodesic segment  $\gamma : [t_1, t_2] \rightarrow M$  such that  $\gamma(t_1) = p$  and  $\gamma(t_2) = q$ . Furthermore*

$$L_{\mathbf{g}}(\gamma) = d_{\mathbf{g}}(p, q) .$$

**Remark 7.16.** The converse of (c) is trivial in every smooth Riemannian manifold: a compact set  $K$  is closed because  $M$  is Hausdorff. Furthermore since  $K$  is compact, it admits a finite covering made of open metric balls of finite radius and therefore there is a larger ball of finite radius including all those balls and  $K$  as well. ■

## 7.2 Normal coordinates and geodesically convex sets

This section is devoted to introduce a type of local charts which has many applications in pure and applied mathematics especially to physics. We only present the basic construction. A more detailed discussion can be found in [O'Ne83, KoNo96].



### 7.2.1 Normal coordinates around a point

Consider a smooth manifold  $M$  equipped with an affine connection  $\nabla$ . If  $N_p$  is a normal neighborhood of  $p \in M$ , we can naturally define a local chart simply choosing a basis of  $T_p M$ .

**Definition 7.17.** (Normal coordinates around a point.) Let  $M$  be a smooth manifold endowed with an affine connection  $\nabla$  and  $p \in M$ . A local chart  $(U, \phi)$  is called a **system of normal coordinates centered on  $p$**  if

- (a)  $U$  is a normal neighborhood of  $p$ ;
- (a)  $\phi : U \ni q \mapsto (x^1, \dots, x^n) \in \mathbb{R}^n$  such that  $q = \exp_p(x^k e_k)$ , where  $\{e_k\}_{k=1, \dots, n} \in T_p M$  is a fixed vector basis. ■

It is clear that the coordinates are well defined since  $\exp_p|_B : B \rightarrow U$ , where  $B = \phi(U)$  is the star-shaped open set centered on  $0 \in TM$  that defines the normal neighborhood  $U$  of  $p$ , is a diffeomorphism by hypothesis. A more explicit expression for  $\phi$  is

$$\phi : U \ni q \mapsto (\langle (\exp_p|_B)^{-1}(q), e^{*1} \rangle, \dots, \langle (\exp_p|_B)^{-1}(q), e^{*n} \rangle) \in B.$$

Normal coordinates enjoy very nice properties.

**Proposition 7.18.** Let  $M$  be a smooth manifold equipped with a smooth affine connection  $\nabla$  which is torsion-free, the Levi-Civita connection in particular if  $(M, \mathbf{g})$  is (pseudo) Riemannian. The following facts hold for every system of normal coordinates  $(U, \phi)$  centered on  $p \in M$ .

- (a) If  $\Gamma_{bc}^a$  are the connection coefficients of  $\nabla$  in the said coordinates, then

$$\Gamma_{bc}^a(p) = 0 \quad \text{for } a, b, c = 1, 2, \dots, \dim(M),$$

- (b) If  $\nabla$  is the Levi-Civita connection of  $(M, \mathbf{g})$  and  $g_{ab}$  are the components of the metric in the said coordinates, then

$$\frac{\partial g_{ab}}{\partial x^c} \Big|_{\phi(p)} = 0 \quad \text{for } a, b, c = 1, 2, \dots, \dim(M).$$

- (c) If  $\gamma : [0, 1] \ni t \mapsto U$  is the unique geodesic segment joining  $p$  at  $t = 0$  with  $q$  at  $t = 1$ , then

$$\phi(\gamma(t)) = (tx^1(q), \dots, tx^n(q)) \quad \text{for every } t \in [0, 1] \quad (7.9)$$

**Proof.** (c) Since geodesics starting from  $p$  are described by  $\gamma : I \ni t \mapsto \exp_p(tx^k e_k)$ , where  $x^k e_k \in B$ , in the considered coordinates, such a geodesic has equation  $x^i(t) = tx^i$ , where  $x^1, \dots, x^n$  are the normal coordinates of the point reached by the said geodesic when  $t = 1$ . This identity proves (c).

(a) In view of the coordinate expression found above of the geodesics emanated from  $p$ , we have  $\frac{d^2 x^i}{dt^2} = 0$ . On the other hand, it must also hold

$$\frac{d^2 x^i}{dt^2} + \Gamma_{jk}^i(\gamma(t)) \frac{dx^j}{dt} \frac{dx^k}{dt} = 0 .$$

As a consequence, for every  $t \in [0, 1]$ :

$$\Gamma_{jk}^i(\gamma(t)) \frac{dx^j}{dt} \frac{dx^k}{dt} = 0 .$$

In particular, for  $t = 0$ ,

$$\Gamma_{jk}^i(p) x^j x^k = 0 , \quad \text{for all } (x^1, \dots, x^n) \text{ in a neighborhood of } 0 \in \mathbb{R}^n .$$

Evidently the statement above is therefore valid for all  $(x^1, \dots, x^n) \in \mathbb{R}$ . If the connection is torsion-free, i.e.  $\Gamma_{jk}^i = \Gamma_{kj}^i$ , the identities above entail that, using  $x^i = u^i + z^i$ ,

$$\Gamma_{jk}^i(p) u^j z^k = 0 \quad \text{for all } u^i, z^i \in \mathbb{R},$$

and thus:

$$\Gamma_{jk}^i(p) = 0 .$$

The proof of (b) is an immediate consequence (a) and of the identity  $\nabla \mathbf{g} = 0$ , which, in coordinates reads:

$$\frac{\partial g_{ki}}{\partial x^j} = \Gamma_{jk}^s g_{si} + \Gamma_{ji}^s g_{ks} . \quad (7.10)$$

□

### 7.2.2 The Gauss lemma and local length minimizing/maximizing geodesics

With the help of the notion of normal neighborhoods and normal coordinates it is possible to prove [KoNo96, O'Ne83, doC92, Pos01, Min19] some remarkable facts regarding local length minimizing/maximizing properties of geodesics. The crucial technical tool is a theorem known as the *Gauss Lemma*.

To state the theorem, observe that in a (pseudo) Riemannian manifold  $(M, \mathbf{g})$  every tangent space  $T_p M$  viewed as a manifold in its own right admits a *flat metric*  $\mathbf{G}$  induced by  $\mathbf{g}_p$ . This is the metric that is constant in every Cartesian coordinate system on  $T_p M$  and takes the components of  $\mathbf{g}_p$  constantly. As a consequence, a normal neighborhood  $N_p = \exp_p(B)$  centered on  $p \in M$  admits two natural metrics: one is the restriction to  $N_p$  of  $\mathbf{g}$ , the other is the restriction to  $N_p$  of  $\mathbf{G}$ , when taking the identification  $N_p = \exp_p(B)$  into account. By construction  $\mathbf{g}_p(u, v) = \mathbf{G}_p(u, v)$ , but

$$\mathbf{g}_q(u, v) \neq \mathbf{G}_q(u, v)$$

in general if  $q \in N_p \setminus \{p\}$  and  $u, v \in T_q M$ . A remarkable result is that, however, the geodesic segments exiting from  $p$  are the same for  $\mathbf{g}$  and  $\mathbf{G}$  as consequence of the discussion around

definition 7.17. This fact is evident if observing that in normal coordinates centered on  $p$ , the geodesic  $\gamma : [0, 1] \ni t \mapsto \exp(tu)$  has the form of a  $\mathbb{R}^n$  segment, i.e., a  $\mathbf{G}$ -geodesic,

$$[0, 1] \ni t \mapsto (tu^1, \dots, tu^n) \in \mathbb{R}^n,$$

and conversely all such segments are  $\mathbf{g}$ -geodesics. Relying upon this fact, the following important result completes the discussion about the interplay of  $\mathbf{g}$  and  $\mathbf{G}$  establishing that actually  $\mathbf{g}_q(u, v) = \mathbf{G}_q(u, v)$  is valid when one of the vectors  $u, v$  are suitably chosen.

**Theorem 7.19.** (**Gauss Lemma.**) *Let  $(M, \mathbf{g})$  be a (pseudo) Riemannian manifold and  $N_p = \exp_p(B)$ , with  $B \subset T_p M$ , a normal neighborhood centered on  $p \in M$ . Consider a geodesic segment  $\gamma : [0, 1] \rightarrow \gamma(t) \in B$  exiting from  $p$  and remaining in  $B$ , then*

$$\mathbf{g}_{\gamma(t)}(\gamma'(t), v) = \mathbf{G}_{\gamma(t)}(\gamma'(t), v) \quad \forall t \in [0, 1], \forall v \in T_{\gamma(t)} M, \quad (7.11)$$

where  $\mathbf{G}$  is the flat metric on  $N_p$  induced by  $\mathbf{g}_p$  and the identification  $N_p = \exp_p(B)$ ,  $B \subset T_p M$ .

**Proof.** If  $(M, \mathbf{g})$  is pseudo Riemannian and (7.11) is valid for geodesic segments such that the initial tangent vector  $w$  satisfies either  $\mathbf{g}(w, w) < 0$  or  $\mathbf{g}(w, w) > 0$ , then it is also valid for the case  $\mathbf{g}(w, w) = 0$  by the joint  $(t, w)$  continuity, when representing the relevant geodesics as  $\gamma : [0, 1] \rightarrow \exp(tw) \in B$ . Therefore we shall only consider the case  $\mathbf{g}(w, w) \neq 0$  in the rest of the proof. Furthermore the thesis is trivial when  $\gamma$  is the constant geodesic because  $\gamma'(t) = 0$ , so that we shall consider only non-constant geodesics.

Let us define an initial system of normal coordinates  $\psi : B \ni w \mapsto (x^1(w), \dots, x^n(w)) \in \mathbb{R}^n$  centered on  $p \equiv (0, \dots, 0)$  with  $g_{ij}(p) = \pm \delta_{ij}$ . In a conical neighborhood of a fixed geodesic  $\gamma : [0, 1] \rightarrow \exp(tu) \in B$ , we can now define a new coordinate system as follows. First observe that a (non lightlike and non-constant) geodesics exiting from  $p$  can be written in the said normal coordinates

$$[0, c] \ni r \mapsto (rn^1, \dots, rn^n)$$

where  $g_{ij}(0, \dots, 0)n^i n^j = \pm 1$  and  $c > 0$ . The sign depends on the nature of  $\mathbf{g}$  and the type of the initial vector  $u$ . Notice that we passed from the parametrization  $t$ , i.e.,  $\gamma : [0, 1] \rightarrow \exp(tu) \in B$  to the new affine parametrization  $r$ , i.e.,  $\gamma : [0, c_u] \rightarrow \exp(rn_u) \in B$  where we use the same symbol  $\gamma$  for the sake of simplicity. Evidently  $n_u$  is, up to sign, the unit vector parallel to  $u$  and  $c_u = \sqrt{|\mathbf{g}_p(u, u)|}$ .

We can next complete the coordinate  $r$  with  $n - 1$  coordinates  $\omega^1, \dots, \omega^{n-1}$  defined on the surface  $g_{ij}(0, \dots, 0)n^i n^j = \pm 1$  in  $\mathbb{R}^n$  identified with  $T_p M$ . The sign is again determined by the nature of  $\mathbf{g}$  and the type of the initial vector  $u$  of the initially fixed geodesic covered by our coordinate system: all that is done in some conical neighborhood  $0 < r < r_0$ ,  $(\omega^1, \dots, \omega^{n-1}) \in A$  — where  $A \subset \mathbb{R}^{n-1}$  is open — of the coordinate representation of  $u$ . In summary, the relation between normal coordinates and the new coordinate system reads

$$x^k = rn^k(\omega^2, \dots, \omega^n), \quad (7.12)$$

for some smooth functions  $n^k = n^k(\omega^2, \dots, \omega^n)$  on  $A$ . In the rest of the proof we shall use the notation

$$(y^1, \dots, y^n) = (r, \omega^2, \dots, \omega^n)$$

so that the metric takes the expression

$$\mathbf{g} = h_{ij} dy^i \otimes dy^j .$$

Within this new coordinate system the (portions of) non-constant geodesic segments exiting from  $p$  have the form

$$(0, c] \ni \lambda \mapsto (\lambda, \omega^2, \dots, \omega^{n-1})$$

for some constants  $c > 0$  and constant values  $\omega^1, \dots, \omega^n \in A$ . From the geodesical equation

$$\frac{d^2 y^k}{d\lambda^2} = \Gamma_{11}^k(\lambda, \omega^2, \dots, \omega^{n-1}) \frac{dy^1}{d\lambda} \frac{dy^1}{d\lambda} .$$

we have

$$0 = \Gamma_{11}^k(\lambda, \omega^2, \dots, \omega^{n-1}) .$$

Expanding the right-hand side according to (6.9), we find

$$0 = 2 \frac{\partial h_{11}}{\partial y^k} - \frac{\partial h_{1k}}{\partial y^1}$$

where both derivatives are computed along the geodesic. Notice that

$$h_{11} = \mathbf{g} \left( \frac{\partial}{\partial r}, \frac{\partial}{\partial r} \right) = \frac{\partial x^k}{\partial r} \frac{\partial x^h}{\partial r} g_{hk} = n^h n^k g_{hk} = \pm 1 .$$

That identity is valid all along the considered geodesic, i.e., varying  $y^1 = r = \lambda$ , since  $n^k$  are the components of its tangent vectors which is parallelly transported. On the other hand, this identity must be valid also changing the initial vector of the geodesic, i.e., changing the values  $\omega^2, \dots, \omega^{n-1}$  which are nothing but the coordinates  $y^2, \dots, y^n$ . In summary  $h_{11}$  is constant in the coordinates  $y^1, \dots, y^n$  so that

$$0 = \frac{\partial h_{1k}}{\partial y^1}$$

along each geodesic exiting from  $p$ . In particular

$$h_{1k}(\lambda, \omega^2, \dots, \omega^n) = h_{1k}(\lambda', \omega^2, \dots, \omega^n), \quad \lambda, \lambda' > 0 . \quad (7.13)$$

It is not difficult to prove from (7.12) that

$$h_{1k}(q) = \frac{\partial x^a}{\partial r} \Big|_q \frac{\partial x^b}{\partial \omega^k} \Big|_q g_{ab}(q) \rightarrow \pm \delta_{1k} \quad \text{if } q \rightarrow p$$

where  $p$  is the center of the initial normal coordinates and the sign may be negative only if the metric is Lorentzian. In summary, for the geodesic  $\gamma : [0, 1] \ni t \mapsto \exp(tu) \in B$  (reparametrization of  $[0, c_u] \ni \lambda \mapsto \exp(\lambda n_u) \in B$ ),

$$\mathbf{g}_{\gamma(t)}(\gamma'(t), v) = h_{hk}(y^1(t), \dots, y^n(t))u^h v^k = \pm u^1 v^1, \quad t > 0.$$

All the reasoning can be repeated as it stands by keeping the coordinate system in  $B$  and using the metric  $\mathbf{G}$  in place of  $\mathbf{g}$ , taking advantage of the fact that the coordinate representation of the  $\mathbf{G}$ -geodesics exiting from  $p$  is identical to the coordinate representation of the  $\mathbf{g}$ -geodesics exiting from  $p$ . We therefore have

$$\mathbf{G}_{\gamma(t)}(\gamma'(t), v) = H_{hk}(y^1(t), \dots, y^n(t))u^h v^k = \pm u^1 v^1, \quad t > 0.$$

Comparing the two obtained identities, we have

$$\mathbf{g}_{\gamma(t)}(\gamma'(t), v) = \mathbf{G}_{\gamma(t)}(\gamma'(t), v), \quad \forall v \in T_{\gamma(t)}M,$$

for  $t > 0$ . Continuity implies that the identity is valid also for  $t = 0$ .  $\square$

**Remark 7.20.** The Gauss lemma is very often stated as follows<sup>3</sup>. Take  $q \in N_p \setminus \{p\}$  and define  $u := (\exp|_B)^{-1}(q)$ . Viewing  $u, v \in B$  as vectors in  $T_u(T_p M)$ , so that both

$$(d\exp_p)_u u \in T_q M \quad \text{and} \quad (d\exp_p)_u v \in T_q M,$$

we have

$$\mathbf{g}_{\exp_p u}((d\exp_p)_u u, (d\exp_p)_u v) = \mathbf{g}_p(u, v). \quad (7.14)$$

■

We are now in a position to state and prove our first result about local minimization of length of curves by geodesical segments.

**Proposition 7.21.** *Consider a Riemannian manifold  $(M, \mathbf{g})$  and a normal neighborhood  $N_p = \exp_p(B)$  of  $p \in M$ . If  $q \in N_p \setminus \{p\}$  and  $\gamma : [0, 1] \rightarrow N_p$  is the unique geodesic segment joining the centre  $p$  with  $q$  remaining in  $N_p$ , then  $L_{\mathbf{g}}(\gamma)$  minimize the length of the piecewise smooth curves joining  $p$  and  $q$  remaining in  $N_p$ .*

**Proof.** Some preparatory tools are necessary specializing to the Riemannian case the coordinate system already exploited in the proof of the Gauss lemma. Consider a normal coordinate system centered on  $p$   $\phi : N_p \rightarrow (x^1, \dots, x^n) \in \mathbb{R}^n$  centered on  $p$  where  $N_p = \exp_p(B)$ . Assume that  $(g_p)_{ab} = \delta_{ab}$  are the components of the metric at  $p$  in the said coordinate system and define the **radial function**

$$r(q) := (\exp_p|_B)^{-1}(q) = \sqrt{\sum_{k=1}^n x^k(q)^2}, \quad q \in N_p.$$

---

<sup>3</sup>I find this formulation a little too obscure.

This function is smooth on  $N_p$  but the origin of the coordinates  $0 = \phi(p)$  where  $r$  is however continuous and the derivatives  $\frac{\partial r}{\partial x^k}$  are bounded in a neighborhood of 0. If  $\gamma : [0, 1] \rightarrow N_p$  is the unique geodesic from  $p$  to  $q$ , as a consequence of the preservation of the length of the tangent vector along a metric geodesic we find

$$L_{\mathbf{g}}(\gamma) = \int_0^1 \sqrt{|\mathbf{g}(\gamma'(t), \gamma'(t))|} dt = \int_0^1 \sqrt{|\mathbf{g}(\gamma'(0), \gamma'(0))|} dt = \int_0^1 r(q) dt = r(q),$$

where we have used the expression  $\gamma(t)^k = tx^k(q)$  (7.9) in normal coordinates. The function  $r$  satisfies a non-trivial fact stated in the following lemma which is consequence of the Gauss lemma.

**Lemma 7.22.** *If  $r$  is defined as above and referring to the used normal coordinates centered on  $p$ ,*

$$g_q^{hk} \frac{\partial r}{\partial x^k} |_{\phi(q)} = \frac{x^h(q)}{r(q)}. \quad (7.15)$$

**Proof.** First of all observe that, if using normal coordinates  $\phi : N_p \ni r \mapsto (x^1(r), \dots, x^n(r)) \in B$  centered on  $p$  (and we also assume that  $(g_p)_{ab} = \delta_{ab}$  as said for simplicity), the exponential map acts as the identity map in coordinates. Complete the coordinate  $r := \sqrt{\sum_{i=1}^n (x^i)^2}$  with coordinates  $\omega_2, \dots, \omega_n$  in a neighborhood of the fixed vector  $w \in T_p M$  corresponding to the fixed  $q \in M$  through  $\exp_p|_B$  in order to have a local chart in  $T_p M$  around  $w$ . The added coordinates are assumed to be coordinates on the embedded manifolds  $S_r := \{u \in T_p M \mid \sum_{i=1}^n (x^i)^2 = r^2\}$ . In particular,

$$\left\langle \frac{\partial}{\partial r}, dr \right\rangle = 1, \quad \left\langle \frac{\partial}{\partial \omega^k}, dr \right\rangle = 0 \quad k = 2, \dots, n. \quad (7.16)$$

Observe that

$$\frac{\partial}{\partial r} = \frac{\partial x^i}{\partial r} \frac{\partial}{\partial x^i} = \frac{x^k}{r} \frac{\partial}{\partial x^k}$$

satisfies

$$g_p \left( \frac{\partial}{\partial \omega^i}, \frac{\partial}{\partial r} \right) = 0, \quad i = 2, \dots, n. \quad (7.17)$$

All this structure in  $T_p M$  can be exported in  $N_p$  around  $q$  through the  $\exp_p$  its pushforward and the pullback of the inverse map. Preserving the names of the coordinates, (7.16) still holds in  $N_p$  around  $q$  by construction.

$$\left\langle \frac{\partial}{\partial r} |_q, dr |_q \right\rangle = 1, \quad \left\langle \frac{\partial}{\partial \omega^k} |_q, dr |_q \right\rangle = 0 \quad k = 2, \dots, n \quad (7.18)$$

Regarding (7.17), Theorem 7.19 immediately implies that

$$g_q \left( \frac{\partial}{\partial \omega^i} |_q, \frac{\partial}{\partial r} |_q \right) = 0, \quad i = 2, \dots, n. \quad (7.19)$$

Notice that now  $g_q$  appears in place of  $g_p$ . We can decompose the functional

$$g_q \left( \cdot, \frac{\partial}{\partial r} \Big|_q \right) \in T_q^* M$$

along the basis  $dr|_q, d\omega^k|_q$ . Taking (7.18) and (7.19) into account, the only possibility is

$$g_q \left( \cdot, \frac{\partial}{\partial r} \Big|_q \right) = c dr|_q ,$$

for some  $c \in \mathbb{R}$ . On the other hand,

$$1 = g_q \left( \frac{\partial}{\partial r} \Big|_q, \frac{\partial}{\partial r} \Big|_q \right) = c \left\langle \frac{\partial}{\partial r} \Big|_q, dr|_q \right\rangle = c1$$

where, in the right hand side, we used the fact that

$$g_p \left( \frac{\partial}{\partial r}, \frac{\partial}{\partial r} \right) = 1$$

in  $T_p M$ , as one proves by direct inspection, and next observe that (see also (7.21) below)  $\frac{\partial}{\partial r}|_q$  is the parallel transport of the vector  $\frac{\partial}{\partial r}|_p$  from  $p$  to  $q$  along the unique geodesic segment joining  $p$  and  $q$  in  $N_p$ , hence its length is preserved as due to Proposition 6.21. In summary  $c = 1$ , so that

$$g_q \left( \cdot, \frac{\partial}{\partial r} \Big|_q \right) = dr|_q .$$

This identity, in normal coordinates reads

$$(g_q)_{kh} \frac{x^h(q)}{r(q)} = \frac{\partial r}{\partial x^k} \Big|_{\phi(q)}$$

and this is an equivalent way to write (7.15).  $\square$

To go on, define the smooth vector field on  $N_p \setminus \{p\}$

$$R_q := \frac{x^k(q)}{r(q)} \frac{\partial}{\partial x^k} \Big|_q . \quad (7.20)$$

Finally observe that, if  $\gamma : [0, 1] \rightarrow N_p$  is the unique geodesic joining  $p$  and  $q$ , again from (7.9)

$$\mathbf{g}(R_q, R_q) = \frac{x^k(q)x^h(q)g_{kh}(q)}{r(q)^2} = \frac{\mathbf{g}(\gamma'(1), \gamma'(1))}{r(q)^2} = \frac{\mathbf{g}(\gamma'(0), \gamma'(0))}{r(q)^2} = 1 , \quad (7.21)$$

where we have exploited Proposition 6.21. Consider a smooth curve  $\alpha : [0, 1] \rightarrow N_p$  with  $\alpha(0) = p$  and  $\alpha(1) = q$ . Exploiting (7.21), we can orthogonally decompose

$$\alpha'(t) = \mathbf{g}(\alpha'(t), R_{\alpha(t)})R_{\alpha(t)} + N(t) ,$$

for  $t > 1$ , where  $N(t)$  is a smooth vector field defined on  $\alpha$  which is everywhere ortogonal to  $R_{\alpha(t)}$ . Hence

$$\sqrt{|\mathbf{g}(\alpha'(t), \alpha'(t))|} = \sqrt{\mathbf{g}(\alpha'(t), R_{\alpha(t)})^2 + \mathbf{g}(N(t), N(t))} \geq |\mathbf{g}(\alpha'(t), R_{\alpha(t)})| \geq \mathbf{g}(\alpha'(t), R_{\alpha(t)}) .$$

Taking advantage of (7.15),

$$\mathbf{g}(\alpha'(t), R_{\alpha(t)}) = g_{hk} \alpha'^h \frac{\partial r}{\partial x^k} = \frac{\partial r}{\partial x^k} |_{\phi(\alpha(t))} \alpha'^k(t) = \frac{dr(\alpha(t))}{dt} .$$

In summary,

$$L_{\mathbf{g}}(\alpha) = \int_0^1 \sqrt{|\mathbf{g}(\alpha'(t), \alpha'(t))|} dt \geq \int_0^1 |\mathbf{g}(\alpha'(t), R_{\alpha(t)})| dt \geq \int_0^1 \frac{dr(\alpha(t))}{dt} dt = r(q) = L_{\mathbf{g}}(\gamma) .$$

Concluding the proof for smooth curves. The piecewise smooth case can be worked out decomposing the various integrals into a finite sum of smooth pieces.  $\square$

As a consequence,  $L_{\mathbf{g}}(\gamma)$  coincides with  $d_{\mathbf{g}}^{(N_p)}(p, q)$  by definition (5.2) where  $(N_p)$  indicates that the distance is referred to  $N_p$  as the manifold where one has to define the used curves. It is possible to replace  $d_{\mathbf{g}}^{(N_p)}(p, q)$  with  $d_{\mathbf{g}}(p, q)$  if choosing  $N_p$  as a *geodesic ball* as we are going to prove. We remind that geodesical balls of sufficiently small radius are always normal neighborhoods of their centers. In this case, the geodesic segment joining the center of the ball with a point in the ball is also the unique distance-minimizing geodesic segment in the whole manifold.

**Proposition 7.23.** *Consider a Riemannian manifold  $(M, \mathbf{g})$  and a normal neighborhood  $N_p^{(r)} = \exp_p(B_r)$  of  $p \in M$  made of a geodesic ball of radius  $r > 0$ . If  $q \in N_p^{(r)} \setminus \{p\}$  and  $\gamma : [0, 1] \rightarrow N_p^{(r)}$  is the unique geodesic segment joining the centre  $p$  with  $q$  remaining in  $N_p^{(r)}$ , then the following holds.*

(a)  $L_{\mathbf{g}}(\gamma)$  minimizes the length of the piecewise smooth curves joining  $p$  and  $q$  in the whole  $M$ . In particular

$$d_{\mathbf{g}}(p, q) = L(\gamma) . \tag{7.22}$$

(b) The said  $\gamma$  is the unique minimizing geodesic segment joining  $p$  and  $q$  with domain  $[0, 1]$  in the whole manifold  $M$ . In other words, a geodesic segment  $\beta : [0, 1] \rightarrow M$  joining  $p$  and  $q$  (a priori not completely included in  $N_p^{(r)}$ ) satisfies either  $L(\beta) > L(\gamma)$  or  $\beta = \gamma$ .

**Proof.** (a) Consider a piecewise smooth curve  $\alpha$  joining  $p$  and  $q$  that is not completely contained in  $N_p^{(r)}$ . Let  $r_q < r$  be length of  $\gamma$ . Next consider another geodesical ball  $N_p^{(r')}$  centered on  $p$  of radius  $r' > r_q$  and  $r' < r$ . Let  $a$  be the point where  $\alpha$  meets for the first time the boundary of  $N_p^{(r')}$ . Using Proposition 7.21, the unique geodesic segment joining  $p$  and  $a$  in  $N_p^{(r')}$  (parametrized in  $[0, 1]$ ) has lenght  $r'$ . Therefore the portion of  $\alpha$  form  $p$  to  $a$  has already



length  $\geq r' > r_q = L(\gamma)$  so that  $L(\alpha) > L(\gamma)$  and the thesis follows. The proof of (b) is easy. If  $L(\gamma) = L(\beta)$ , then every point on  $\beta([0, 1])$  belongs to  $N_p^{(r)}$  because the length ascissa on  $\beta$  is nothing but the radial coordinate of the corresponding point referred to the center of the ball  $p$ . In this case  $\beta$  is a geodesic segment joining  $p$  and  $q$  and completely remaining in a normal neighborhood of  $p$ . The uniqueness property of these geodesics implies that  $\gamma = \beta$ .  $\square$

The existence of a minimizing geodesics in Riemannian manifolds is *globally* true in *complete manifolds* as established by Theorem 7.15, in the sense that every pair of points, also arbitrarily far, are joined by a geodesic segment whose length coincides with the distance of the points. Notice that this geodesic segment could be not unique if the one of the two points is not a normal neighborhood of the other (think of the North and South poles in 2-dimensional sphere equipped with the standard Riemannian metric). However if a piecewise smooth curve minimizes the length of the curves joining two points arbitrarily far to each other, it must be a geodesic segment up to reparametrization as established below.

**Proposition 7.24.** *A piecewise smooth curve segment  $\gamma : [a, b] \rightarrow M$  with  $\gamma(a) = p$  and  $\gamma(b) = q$  and  $p, q \in M$  for a Riemannian manifold  $(M, \mathbf{g})$  can be reparametrized to a geodesic segment defined in  $[0, 1]$  if it minimized the distance of  $p$  and  $q$ , i.e., if  $L(\gamma) = d_{\mathbf{g}}(p, q)$ .*

**Proof.** If  $a', b' \in [a, b]$  with  $a' < b'$ , it must hold  $L(\gamma|_{[a', b']}) = d_{\mathbf{g}}(\gamma(a'), \gamma(b'))$  otherwise it is easy to prove that  $L(\gamma) > d_{\mathbf{g}}(p, q)$ . Choosing  $a'$  and  $b'$  sufficiently close to each other, it turns out that  $\gamma|_{[a', b']}$  is included in a normal neighborhood of  $\gamma(a')$  and thus it can be reparametrized to a geodesic segment due to Proposition 7.21. Since  $\gamma([a, b])$  is compact, we can cover  $[a, b]$  with a finite covering of such intervals  $[a', b']$ . The uniqueness theorem of the differential equation defining geodesics easily implies that  $\gamma$  can be globally reparametrized to a geodesic segment.  $\square$

Let us pass to Lorentzian manifolds where we state similar local results [O'Ne83, BEE96, Min19].

**Proposition 7.25.** *Consider a Lorentzian manifold  $(M, \mathbf{g})$  and a normal neighborhood  $N_p$  of  $p \in M$ . If  $q \in N_p \setminus \{p\}$  assume that the unique geodesic segment  $\gamma : [0, 1] \rightarrow N_p$  joining the centre  $p$  with  $q$  completely contained in  $N_p$  is timelike. Then  $L(\gamma)$  maximize the length of the smooth timelike curves  $\alpha : I \rightarrow N_p$  joining  $p$  and  $q$ .*

*The property is valid also referring to piecewise smooth timelike curves  $I \ni t \mapsto \alpha(t) \in N_p$  such that<sup>4</sup>  $\mathbf{g}(\alpha'(t_i^-), \alpha'(t_i^+)) < 0$  with obvious notations, for every non-smoothness point  $t_i \in I$ .*

**Proof.** As in the Riemannian case we prove the thesis for smooth curves, the piecewise smooth case being a trivial generalization after a remark we shall state at the end of the proof.

---

<sup>4</sup>If  $(M, \mathbf{g})$  is time-oriented according to Def. 8.10, these curves are all future-directed or past-directed as stated in (2) Remark 8.11.

Some preparatory tools are necessary. Consider a normal coordinate system centered on  $p$   $\phi : N_p \rightarrow (x^1, \dots, x^n) \in \mathbb{R}^n$  centered on  $p$  where  $N_p = \exp_p(B)$ . Assume that  $(g_p)_{ab} = \eta_{ab} \equiv \text{diag}(-1, 1, \dots, 1)$  are the components of the metric at  $p$  in the said coordinate system and define the **time radial function**

$$r(q) := (\exp_p|_B)^{-1}(q) = \sqrt{x^1(q)^2 - \sum_{k=2}^n x^k(q)^2}, \quad q \in V_p,$$

in the open cone  $V_p$  around the  $x^1$  axis defined by the points with coordinates  $x^k$  such that the radicand is positive and known as the light cone at  $p$ .

This function is smooth on  $N_p$  but the boundary  $\partial V_p$  where  $r$  is however continuous and the derivatives  $\frac{\partial r}{\partial x^k}$  are bounded in a neighborhood of 0. If  $\gamma : [0, 1] \rightarrow N_p$  is the unique timelike geodesic from  $p$  to  $q \in V_p$ , as a consequence of the preservation of the length of the tangent vector along a metric geodesic we find

$$L_{\mathbf{g}}(\gamma) = \int_0^1 \sqrt{|\mathbf{g}(\gamma'(t), \gamma'(t))|} dt = \int_0^1 \sqrt{|\mathbf{g}(\gamma'(0), \gamma'(0))|} dt = \int_0^1 r(q) dt = r(q),$$

where we have used the expression  $\gamma(t)^k = tx^k(q)$  (7.9) in normal coordinates. We henceforth assume to arrange the coordinates in order that the  $\frac{\partial}{\partial x^1}$ -component of  $\gamma'(0)$  is positive. The function  $r$  satisfies Lemma 7.22 as well, with a minus sign.

**Lemma 7.26.** *If  $r$  is defined as above and referring to the used normal coordinates centered on  $p$ ,*

$$g_q^{hk} \frac{\partial r}{\partial x^k} |_{\phi(q)} = -\frac{x^h(q)}{r(q)}. \quad (7.23)$$

**Proof.** The proof is identical to that of Lemma 7.22, just paying attention to the different definition of the function  $r$ , taking the signs into account.  $\square$

To go on, define the smooth timelike vector field on  $N_p \setminus \{p\}$

$$R_q := \frac{x^k(q)}{r(q)} \frac{\partial}{\partial x^k} |_q. \quad (7.24)$$

Finally observe that, if  $\gamma : [0, 1] \rightarrow N_p$  is the unique timelike geodesic joining  $p$  and  $q$ , again from (7.9)

$$\mathbf{g}(R_q, R_q) = \frac{x^k(q)x^h(q)g_{kh}(q)}{r(q)^2} = \frac{\mathbf{g}(\gamma'(1), \gamma'(1))}{r(q)^2} = \frac{\mathbf{g}(\gamma'(0), \gamma'(0))}{r(q)^2} = -1, \quad (7.25)$$

where we have exploited Proposition 6.21. Consider a smooth timelike curve  $\alpha : [0, 1] \rightarrow N_p$  with  $\alpha(0) = p$  and  $\alpha(1) = q$ . Using (7.25), we can orthogonally decompose

$$\alpha'(t) = \mathbf{g}(\alpha'(t), R_{\alpha(t)})R_{\alpha(t)} + N(t),$$

for  $t > 1$ , where  $N(t)$  is a smooth vector field defined on  $\alpha$  which is everywhere ortogonal to  $R_{\alpha(t)}$ . Hence

$$\sqrt{|\mathbf{g}(\alpha'(t), \alpha'(t))|} = \sqrt{\mathbf{g}(\alpha'(t), R_{\alpha(t)})^2 - \mathbf{g}(N(t), N(t))} \leq |\mathbf{g}(\alpha'(t), R_{\alpha(t)})| = -\mathbf{g}(\alpha'(t), R_{\alpha(t)}) ,$$

where in the last passage we have used the fact that both  $\alpha'(t)$ ,  $R_{\alpha(t)}$  are timelike so that their scalar product is negative since it is at  $p$  with our choice of the coordinates and it cannot change later by continuity. Exploiting (7.15),

$$\mathbf{g}(\alpha'(t), R_{\alpha(t)}) = g_{hk} \alpha'^h \frac{\partial r}{\partial x^k} = -\frac{\partial r}{\partial x^x} |_{\phi(\alpha(t))} \alpha'^k(t) = -\frac{dr(\alpha(t))}{dt} .$$

In summary,

$$L_{\mathbf{g}}(\alpha) = \int_0^1 \sqrt{|\mathbf{g}(\alpha'(t), \alpha'(t))|} dt \leq \int_0^1 |\mathbf{g}(\alpha'(t), R_{\alpha(t)})| dt = \int_0^1 \frac{dr(\alpha(t))}{dt} dt = r(q) = L_{\mathbf{g}}(\gamma) .$$

Concluding the proof for the smooth case. The piecewise smooth case with the further property  $\mathbf{g}(\alpha'(t_i^-), \alpha'(t_i^+)) < 0$  can be proved analogously by decomposing the integral above into a sum of integrals over the smoothness subintervals. The only crucial fact is to assure  $|\mathbf{g}(\alpha'(t), R_{\alpha(t)})| = -\mathbf{g}(\alpha'(t), R_{\alpha(t)})$  on every subsegment. The first subinterval satisfies this requirement by construction, the remaining ones do as well because, using an inductive proof,  $\mathbf{g}(\alpha'(t_{i+1}^+), R_{\alpha(t_i)}) < 0$  arises from  $\mathbf{g}(\alpha'(t_i^-), R_{\alpha(t_i)}) < 0$  and  $\mathbf{g}(\alpha'(t_i^-), \alpha'(t_i^+)) < 0$  by direct inspection (see Proposition 8.7).  $\square$

We have a final result which established that, analogously to the Riemannian case, the timelike geodesic segments are the unique type of smooth curves that may maximize the length. Notice that there is no guarantee that these maximal curves exist. They exist under suitable hypotheses on the Lorentzian manifold [BEE96, Min19].

**Proposition 7.27.** *Consider, in the Lorentzian manifold  $(M, \mathbf{g})$ , a timelike piecewise smooth curve*

$$\gamma : [a, b] \rightarrow M$$

*such that  $\gamma(a) = p$  and  $\gamma(b) = q$ . If  $\gamma$  maximize the length of all the piecewise smooth curves joining  $p$  and  $q$   $I \ni t \mapsto \alpha(t) \in N_p$  such that  $\mathbf{g}(\alpha'(t_i^-), \alpha'(t_i^+)) < 0$  for every non-smoothness point  $t_i \in I$ , then  $\gamma$  can be re-parametrized to the unique geodesic segment defined on  $[0, 1]$  that joins  $p$  and  $q$ .*

**Proof.** If  $a', b' \in [a, b]$  with  $a' < b'$ , then  $L(\gamma|_{[a', b']})$  maximize the length of the timelike piecewise smooth cirves joining  $\gamma(a')$  and  $\gamma(b')$ , otherwise it is easy to prove that  $L(\gamma)$  has not the extremal property assumed in the hypotheses. Choosing  $a'$  and  $b'$  sufficiently close to each other, it turns out that  $\gamma|_{[a', b']}$  is included in a normal neighborhood of  $\gamma(a')$  and thus it can be reparametrized to a geodesic segment due two Proposition 7.25. Since  $\gamma([a, b])$  is compact, we can cover  $[a, b]$

with a finite covering of such intervals  $[a', b']$ . The uniqueness theorem of the differential equation defining geodesics easily implies that  $\gamma$  can be globally reparametrized to a timelike geodesic segment.  $\square$

### 7.2.3 Riemannian normal coordinates adapted to a geodesic.

Let us pass to consider a more complicated construction concerning normal coordinates *around a given geodesic*. In this case, we focus attention on the case where  $\nabla$  is the Levi-Civita connection.

Let  $(M, \mathbf{g})$  be a (pseudo) Riemannian manifold. Let  $\gamma : I \rightarrow M$  be a non-constant geodesic. In case  $\mathbf{g}$  is pseudo Riemannian, we assume that  $\mathbf{g}(\gamma', \gamma') \neq 0$ . The local chart around  $\gamma(t_0)$  for some  $t_0 \in I$  and adapted to  $\gamma$  is constructed with the following steps.

- (1) Fix a basis  $\{e_i\}_{i=2,\dots,n}$  of the subspace  $N_{\gamma(t_0)}\gamma$  of  $T_{\gamma(t_0)}M$  normal to  $\gamma'(t_0)$ .
- (2) Parallely transport that basis along  $\gamma$ . Due to Proposition 6.21, the transported vectors define a basis  $\{e_i(t)\}_{i=2,\dots,n}$  for the subspace  $N_{\gamma(t)}\gamma$  of  $T_{\gamma(t)}M$  normal to  $\gamma'(t)$ .
- (3) Consider the map

$$E : U \ni (t, v^2, \dots, v^n) \mapsto \exp_{\gamma(t)} \left( \sum_{i=2}^n v^i e_i(t) \right), \quad (7.26)$$

where  $U \subset I \times \mathbb{R}^{n-1}$  is an open set. We know that  $U$  exists as a consequence of Theorem 7.2 using also the fact that the maps  $I \ni t \mapsto (\gamma(t), e_i(t)) \in TM$  are continuous.

**Remark 7.28.** The geometric meaning of (7.26) should be clear: the map associates the coordinates  $(t, v^2, \dots, v^n)$  to the point in  $M$  reached by the geodesic which starts from  $\gamma(t)$  with initial vector  $\sum_{i=2}^n v^i e_i(t)$  normal to  $\gamma'(t)$ , when the affine parameter of that new geodesic takes the value 1.  $\blacksquare$

We have the following technically remarkable result.

**Proposition 7.29.** *Let  $(M, \mathbf{g})$  be a (pseudo) Riemannian manifold and  $\gamma : I \rightarrow M$  a geodesic with  $\mathbf{g}(\gamma', \gamma') \neq 0$ , define the map  $E$  as in (7.26) and fix  $\tau \in I$ . The following facts are true.*

- (a) *There is  $\epsilon > 0$  and an open set  $D \subset \mathbb{R}^{n-1}$  including  $0 \in \mathbb{R}^{n-1}$  such that  $E|_{(\tau-\epsilon, \tau+\epsilon) \times D}$  defines a diffeomorphism onto its image  $V \subset M$  and thus if  $\psi := E|_{(\tau-\epsilon, \tau+\epsilon) \times D}^{-1}$ ,*

$$\psi : V \ni q \mapsto (y^1, \dots, y^n) \in (\tau - \epsilon, \tau + \epsilon) \times D$$

*is a local chart around a portion of  $\gamma$  containing  $\gamma(\tau)$ .*

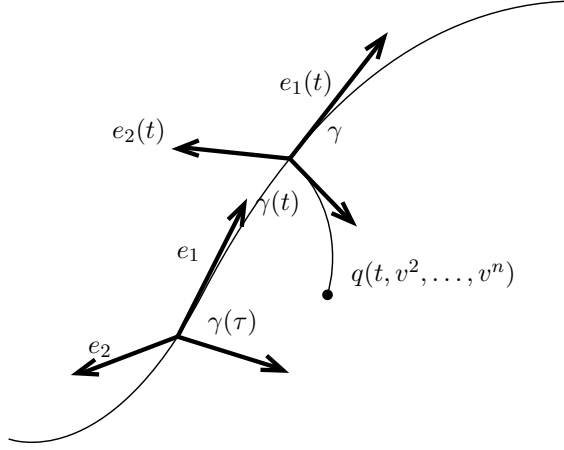


Figure 7.1: Normal coordinates around  $\gamma$

(b) The connection coefficients  $\Gamma_{bc}^a$  and the components of the metric  $g_{ab}$  in the said coordinate system satisfy

$$\Gamma_{bc}^a(\gamma(t)) = 0, \quad \frac{\partial g_{ab}}{\partial y^c}|_{\gamma(t)} = 0 \quad \text{if } a, b, c = 1, \dots, n \text{ and } t \in (\tau - \epsilon, \tau + \epsilon). \quad (7.27)$$

**Proof.** Consider an auxiliary system of normal coordinates  $x^1, \dots, x^n$  centered on the initial point  $p = \gamma(\tau)$  associated to the initial basis  $\{\gamma'(\tau)\} \cup \{e_i(\tau)\}_{i=2, \dots, n}$  of  $T_{\gamma(\tau)}M$ . In that case, exactly at  $t = \tau$ ,

$$\frac{\partial}{\partial x^1}|_p = \gamma'(\tau) \quad \text{and} \quad \frac{\partial}{\partial x^i}|_p = e_i(\tau) \quad \text{for } i = 2, \dots, n.$$

Using those coordinates to represent the image of the map  $E$ :  $E^k = E^k(t, v^1, \dots, v^n)$ , we find

$$\frac{\partial}{\partial t}\bigg|_{\gamma(\tau)} \left( \exp_{\gamma(t)} \left( \sum_{i=2}^n v^i e_i(t) \right) \right)^k = \frac{\partial}{\partial t}\bigg|_{t=\tau} \left( \exp_{\gamma(t)}(0) \right)^k + \sum_{i=2}^n v^i g^k(\tau, v^2, \dots, v^n) \bigg|_{(v^2, \dots, v^n) = (0, \dots, 0)},$$

where the functions  $g^k$  are smooth. Since  $\exp_{\gamma(t)}(0) = \gamma(t)$ , we conclude that:

$$\frac{\partial}{\partial t}\bigg|_{\gamma(\tau)} \left( \exp_{\gamma(t)} \left( \sum_{i=2}^n v^i e_i(t) \right) \right)^k = \frac{\partial}{\partial t}\bigg|_{t=\tau} \gamma^k(t) + 0 = \gamma'^k(\tau) = \left( \frac{\partial}{\partial x^1}\bigg|_{\gamma(\tau)} \right)^k = \delta_1^k,$$

For  $j = 2, \dots, n$ , it also arises (there is no sum over  $j$  in the second exponential)

$$\frac{\partial}{\partial v^j}\bigg|_{\gamma(\tau)} \left( \exp_{\gamma(t)} \left( \sum_{i=2}^n v^i e_i(t) \right) \right)^k = \frac{\partial}{\partial v^j}\bigg|_{(v^2, \dots, v^n) = (0, \dots, 0)} \left( \exp_{\gamma(\tau)}(v^j e_j(\tau) + 0) \right)^k = (e_j(\tau))^k = \delta_j^k.$$

Therefore the *rank* of the map (7.26) at  $\gamma(\tau)$  is  $n$  since using the said normal coordinates we have found that

$$\frac{\partial E^k}{\partial t}|_{\gamma(\tau)} = \delta_1^k, \quad \frac{\partial E^k}{\partial v^j}|_{\gamma(\tau)} = \delta_j^k \quad \text{for } k = 1, \dots, n \text{ and } j = 2, \dots, n$$

are the components of a  $n \times n$  matrix with determinant 1. The map  $E$  defines a local diffeomorphism around the coordinate representation of  $\gamma(\tau)$  from an open set we can always assume to be of the form  $(\tau - \epsilon, \tau + \epsilon) \times D$  to  $U \subset M$  and thus there is a local chart  $\psi : U \ni q \mapsto \psi(U) = (\tau - \epsilon, \tau + \epsilon) \times D$ . This concludes the proof of (a).

The proof of (b) is analogous to that of Proposition 7.18. Let us indicate by  $y^1, \dots, y^n$  the local coordinates defined above where  $y^1 = 1$  and  $y^k = v^k$  for  $k > 1$ .

- (i) In coordinates  $y^1, \dots, y^n$ , the geodesics starting from  $\gamma(t)$  with initial tangent vector  $\sum_{i=2}^n v^i \frac{\partial}{\partial y^i}|_{\alpha(t)}$  have equation  $y^1(\lambda) = t$  (constant!) and  $y^j(\lambda) = \lambda v^j$  for  $j = 2, \dots, n$ . Inserting in

$$\frac{d^2 y^k}{d\lambda^2} + \Gamma_{ij}^k \frac{dy^i}{d\lambda} \frac{dy^j}{d\lambda} = 0$$

and considering the result for  $\lambda = 0$  we find  $\Gamma_{ij}^k(\gamma(t))v^i v^j = 0$  if  $i, j = 2, \dots, n$ , which imply  $\Gamma_{ij}^k(\gamma(t)) = 0$  if  $i, j = 2, \dots, n$  for the arbitrariness of  $v^k$ .

- (ii) The vectors  $\frac{\partial}{\partial y^j}$  with  $j = 2, \dots, n$  satisfy the equation of parallel transport with respect to  $\gamma$ , that is with respect to  $\frac{\partial}{\partial y^1}$ . In coordinates  $y^1, \dots, y^n$ , this equation reads

$$\frac{d}{dt} \delta_j^k + \delta_1^i \Gamma_{ir}^k(\gamma(t)) \delta_j^r = 0,$$

this entails  $\Gamma_{1j}^k(\gamma(t)) = \Gamma_{j1}^k(\gamma(t)) = 0$  for  $j = 2, \dots, n$ .

- (iii) Since  $\gamma$  is a geodesic, in coordinates  $y^1, \dots, y^n$ , the curve

$$y^1 = t, \quad y^j = 0 \quad \text{if } j = 2, \dots, n$$

satisfies the geodesic equation.

$$\frac{d^2}{dt^2} \delta_1^i + \Gamma_{jk}^i(\gamma(t)) \delta_1^j \delta_1^k = 0.$$

Hence  $\Gamma_{11}^k(\gamma(t)) = 0$ .

We have established that all coefficients  $\Gamma_{jk}^i$  vanish along the considered portion of  $\gamma$ . To conclude, observe that, since

$$\frac{\partial g_{ki}}{\partial y^j} = \Gamma_{jk}^s g_{si} + \Gamma_{ji}^s g_{ks},$$

then also the derivatives of the components of the metric vanishes along  $\gamma$ .  $\square$

**Definition 7.30.** Let  $(M, \mathbf{g})$  be a (pseudo) Riemannian manifold and  $\gamma : I \rightarrow M$  a geodesic with  $\mathbf{g}(\gamma', \gamma') \neq 0$ . The local chart defined as in (7.26) around a portion of  $\gamma$  is called **(Riemannian) normal coordinate system around  $\gamma$** , or **adapted to  $\gamma$** .  $\blacksquare$

#### 7.2.4 Convex normal neighborhoods and Whitehead's theorem

Consider a manifold  $M$  with an affine connection  $\nabla$  (thus in particular a Riemannian or a Lorentzian manifold). It is possible to make stronger the notion of normal neighborhood as follows [O'Ne83, KoNo96, Pos01].

**Definition 7.31.** (**Convex normal neighborhoods.**) Let  $M$  be a smooth manifolds endowed with an affine connection  $\nabla$ . A **convex normal neighborhood** in  $M$  is an open set  $U$  that is a normal neighborhood of *every*  $p \in U$ .  $\blacksquare$

The existence of normal convex neighborhoods in (pseudo) Riemannian manifolds and the fact that they form a basis of the topology is due to Whitehead. The following result is more generally valid [KoNo96, Pos01] for smooth manifolds equipped with affine connections<sup>5</sup>.

**Theorem 7.32.** (**Whitehead's Theorem**) *Let  $M$  be a smooth manifold equipped with an affine connection  $\nabla$  and  $p \in M$ . Every sufficiently small normal neighborhood  $C$  of  $p$  is also convex normal and the class of convex normal neighborhoods is a basis of the topology of  $M$ . More precisely, if  $(U, \phi)$  is a system of normal coordinates centered on  $p$ , then there is a number  $\delta > 0$  such that, if  $B_r(0) \subset T_p M$  is an open coordinate ball centered on  $0 \in T_p M$  with radius  $r < \delta$ , then  $C := \exp_p(B_r(0))$  is a convex normal neighborhood of  $p$ .*

**Proof.** Fix a system of normal coordinates  $(U, \phi)$  centered on  $p \in M$ . We have a first lemma where we use notations as Definition 3.25.

**Lemma 7.33.** *There is a neighborhood  $W$  of  $(p, 0) \in DE \subset TM$  of the form*

$$W = (T\phi)^{-1}(\phi(V) \times B),$$

*where  $V \subset U$  is a smaller neighborhood  $V \ni p$  and  $B \subset \mathbb{R}^n$  is an open ball centered on the origin of  $\mathbb{R}^n$  such that*

$$[0, 1] \times W \ni (t, q, v) \mapsto \exp_q(tv) \in U,$$

*so that the both the projection onto  $M$  of the domain and the image of that function are simultaneously covered by the coordinate system  $\phi$ .*

---

<sup>5</sup>The proof in [O'Ne83] seems to me affected by a gap here closed with Lemma 7.33.

**Proof of Lemma 7.33.** Consider the map

$$[0, 1] \times DM \ni (t, q, v) \mapsto F(t, q, v) := \exp_q(tv) \in M .$$

This map is continuous as a consequence of Theorem 7.2 and  $F(t, p, 0) = p$  for every  $t \in [0, 1]$ . Using the fact that  $[0, 1]$  is compact, it is easy to prove that we can fix an open neighborhood  $W$  of  $(p, 0) \in DE \subset TM$  such that  $F(t, q, v) \in U$  if  $(t, q, v) \in [0, 1] \times W$ . Indeed, due to continuity, for every  $t \in [0, 1]$  there is an open neighborhood  $W_t$  of  $(t, p, 0)$  such that  $F(W_t) \subset U$ . We can extract a finite subcovering  $\{W_{t_i}\}_{i=1,2,\dots,N}$  of the compact  $[0, 1] \times \{p\} \times \{0\}$ . We can define  $W := \bigcap_{i=1,\dots,N} W_{t_i}$  which is open because finite intersection of open sets.

We can finally assume, shrinking  $W$  around  $(p, 0)$  if necessary, that  $W \subset TU$  and, taking the topology of  $TM$  into account and using the coordinates of  $U$ ,  $T\phi(W) = \phi(V) \times B$  where  $V \subset U$  open,  $V \ni p$  and  $B \subset \mathbb{R}^n$  is an open ball centered on the origin of  $\mathbb{R}^n$ .  $\square$

Next we pass to study the map

$$E : W \ni (p, v) \mapsto (p, \exp_p v) \in U \times U ,$$

which is well defined in view of the previous lemma by  $E$ .

**Lemma 7.34.** *It holds that the differential*

$$dE|_{(p,0)} : T_{(p,0)}(TM) \rightarrow T_p U \times T_p U$$

*is bijective. Therefore, according to (3) in Theorem 4.5, there is an open neighborhood  $W' \subset W$  of  $(p, 0) \in TM$  which is diffeomorphically mapped onto an open neighborhood of  $(p, p) \in U \times U$ .*

**Proof of Lemma 7.34.** We can compute the rank of  $E$  at  $(p, 0)$  using the local chart  $(U, \phi)$ . In those coordinates, the map  $E$  is represented by

$$\begin{aligned} (x^1, \dots, x^n, v^1, \dots, v^n) &\mapsto (E^1, \dots, E^n, \tilde{E}^1, \dots, \tilde{E}^n) \\ &:= (x^1, \dots, x^n, \exp_p(v^1, \dots, v^n)^1, \dots, \exp_p(v^1, \dots, v^n)^n) \end{aligned}$$

where  $(p, 0) \equiv (0, \dots, 0, 0, \dots, 0)$ . Hence, for  $h, k = 1, \dots, n$ ,

$$\begin{aligned} \frac{\partial E^k}{\partial x^h}|_{(p,0)} &= \delta_h^k , & \frac{\partial E^k}{\partial v^h}|_{(p,0)} &= 0 , \\ \frac{\partial \tilde{E}^k}{\partial x^h}|_{(p,0)} &= f_h^k(p) , & \frac{\partial \tilde{E}^k}{\partial v^h}|_{(p,0)} &= \delta_h^k , \end{aligned}$$

where we have used the fact that in normal coordinates centered on  $p$ ,

$$(\exp_p(v))^k = v^k .$$



The  $2n \times 2n$  matrix with the components as above define an injective map as one immediately proves by direct inspection, concluding the proof.  $\square$

At this point we can assume that  $E(W') = U' \times U'$  where  $U' \subset U$  is a smaller neighborhood of  $p$ . By definition of the map  $E$ , it must be  $\Pi(W') = U'$ . In other words, for every  $q \in U'$

$$\exp_q(B_q) = U' \quad \text{where } B_q \text{ is an open neighborhood of } 0 \in T_q M.$$

$B_q \subset T_q M$  is open because it is the intersection of  $E^{-1}(U' \times U') \subset TM$  and the fiber  $T_q M$  and  $E^{-1}(U' \times U')$  is open<sup>6</sup>.

To conclude it is sufficient to prove that  $B_q$  is star-shaped for every  $q \in U'$  since, for the very definition of normal neighborhood, it means that  $U'$  is a normal neighborhood of every  $q \in U'$ . Star-shaped means that  $E(q, \lambda v) \in U'$  if  $t \in [0, 1]$ . In turn, this is the same as requiring that every geodesic segment  $[0, 1] \ni t \mapsto \exp_q(tv)$  is included in  $U'$  for  $v \in B_q$ . This fact is established in the following final lemma.

**Lemma 7.35.** *Choose  $U' := \exp_p(B_r(0))$  where  $B_\delta(0) \subset T_p M$  is a coordinate ball of radius  $r > 0$  centered on  $0 \in T_p M$ . There is  $\delta > 0$  such that, if  $r < \delta$ , then  $\exp_q(tv) \subset U'$  for every  $q \in U'$  and  $t \in [0, 1]$ .*

**Proof of Lemma 7.35.** We can perform all computations in normal coordinates of  $(U, \phi)$  due to Lemma 7.33, since  $[0, 1] \times W \ni (t, q, v) \mapsto \exp_q(tv) \in U$  and we are still working inside that neighborhood  $W \subset TM$  of  $(p, 0)$ , because  $W \supset W' = E(U' \times U')$  in our construction. Now we are free to choose  $U'$  as in the hypothesis provided  $r > 0$  is sufficiently small since the normal neighborhoods form a basis of the topology.

In normal coordinates centred on  $p$ ,  $U'$  is therefore represented by the ball

$$B_r(0) := \{x \in \mathbb{R}^3 \mid \|x\| < r\}$$

for some  $r > 0$  and where  $x = (x^1, \dots, x^n)$ , the center of the chart coinciding with  $\phi(p)$ .

To go on, assuming that  $\nabla$  is torsion-free, consider the quadratic form defined by the coefficients  $G_{ab}(q)$  at each point  $q \in U$ , where  $q$  has coordinates  $(x^1, \dots, x^n)$ ,

$$G_{ab}(q) := \delta_{ab} + \sum_c \Gamma_{ab}^c(q) x^c.$$

(Till the end of the proof we will not take advantage of Einstein's summation convention on repeated indices.) If the coefficients  $\Gamma_{ab}^c$  are not symmetric, we can replace  $\nabla$  with the torsion-free connection whose coefficients are  $\frac{1}{2}(\Gamma_{ab}^c + \Gamma_{ba}^c)$  since it has the same geodesics as  $\nabla$ . For  $q = p$ , the quadratic form is strictly positive and therefore, since the term added to  $\delta_{ab}$  is smooth and vanishes at the origin, it remains strictly positive provided  $\|x\| < \delta$  for some  $\delta > 0$ . Let us correspondingly fix  $0 < r < \delta$ .

---

<sup>6</sup>Make use of (6) in Exercises 2.1 locally.

Consider a geodesic segment  $\gamma$  starting at  $q \in U'$  for  $t = 0$  and reaching  $q' \in U' \setminus \{q'\}$  at  $t = 1$  (i.e.,  $\exp_q(v) = q'$ ). Suppose that it is not completely contained in  $U'$ . Therefore, if denoting  $x(t) := \phi(\gamma(t))$ , so that  $\|x(0)\| < r$  and  $\|x(1)\| < r$ , the smooth function  $\|x(t)\|^2$  must reach a maximum value in  $t_0 \in (0, 1)$  with  $\|x(t_0)\| > r$ . In particular

$$\frac{d^2\|x(t)\|^2}{dt^2}\bigg|_{t_0} \leq 0. \quad (7.28)$$

Let us prove that this is not possible and thus  $\gamma([0, 1]) \subset U'$  concluding the proof. In fact, from (7.28) and the geodesic equation,

$$\begin{aligned} \frac{d^2\|x(t)\|^2}{dt^2}\bigg|_{t_0} &= 2 \sum_i \frac{dx^i}{dt}\bigg|_{t_0} \frac{dx^i}{dt}\bigg|_{t_0} + 2 \sum_i x^i(t_0) \frac{d^2x^i}{dt^2}\bigg|_{t_0} \\ &= 2 \sum_i \frac{dx^i}{dt}\bigg|_{t_0} \frac{dx^i}{dt}\bigg|_{t_0} - 2 \sum_{i,c,d} x^i(t_0) \Gamma_{cd}^i(x(t_0)) \frac{dx^c}{dt}\bigg|_{t_0} \frac{dx^d}{dt}\bigg|_{t_0} = \sum_{cd} G_{cd}(x(t_0)) \frac{dx^c}{dt}\bigg|_{t_0} \frac{dx^d}{dt}\bigg|_{t_0} > 0, \end{aligned}$$

where we have also used the fact that the tangent vector of the geodesic cannot vanish since the geodesic is non-constant. We have found a contradiction with (7.28).  $\square$

The proof of the theorem ends defining  $C := U'$ .  $\square$

**Remark 7.36.**

(1) If  $C$  is a convex normal neighborhood in  $M$ , the following facts are valid directly from the definition.

- (a) If  $B \subset T_q M$ ,  $q \in C$ , is an open star-shaped neighborhood with center the origin, then  $D := \exp_q(B)$  is automatically a normal neighborhood of  $q$ ;
- (b) there is a unique geodesic segment parametrized in  $[0, 1]$  joining  $q, r \in C$  with  $q \neq r$  and completely included in  $C$ ;
- (c) If  $C' \subset C$  is an open set such that every geodesic segment in  $C$  joining some given  $q \in C'$  to every other  $r \in C'$  is completely included in  $C'$ , then  $C'$  is a normal neighborhood of  $q$ .

(2) if  $(M, g)$  is Riemannian, Theorem 7.32 immediately implies that every sufficiently small geodesic ball is convex normal.

(3) Propositions 7.21 and 7.25 are *a fortiori* valid for the length of every geodesic segment, respectively timelike geodesic segment, joining two distinct points in an normal convex neighborhood and completely included in it.  $\blacksquare$

### 7.2.5 Coverings of convex neighborhoods with a nice intersection property

To conclude we state and prove a useful property of the basis of convex normal neighborhoods. If we consider a pair of convex normal neighborhoods  $C$  and  $C'$  such that  $C \cap C' \neq \emptyset$  and

$p, q \in C \cap C'$ , it is not necessarily true that the unique geodesic segment (parametrized in  $[0, 1]$ ) joining  $p$  and  $q$  in  $C$  coincides with the analogous unique geodesic segment joining  $p$  and  $q$  in  $C'$ . However that is evidently true when  $C \cap C'$  is convex normal as well since, in that case, the unique geodesic segment joining  $p$  and  $q$  in  $C \cap C'$  is also included in  $C$  and in  $C'$ .

A natural issue therefore arises. If  $M$  is a smooth manifold equipped with an affine connection  $\nabla$ , is there an open covering of convex normal neighborhoods  $\mathcal{C}$  such that, if  $C, C' \in \mathcal{C}$ , then  $C \cap C'$  is empty or convex normal?

We actually have a bit stronger result proved in the following proposition<sup>7</sup>.

**Proposition 7.37.** *Let  $M$  be a smooth manifold equipped with an affine connection  $\nabla$  and  $\mathcal{A}$  a covering of  $M$  made of open sets. Then there exists a covering  $\mathcal{C}$  of  $M$  made of convex normal neighborhoods (with respect to  $\nabla$ ) such that,*

- (a) *if  $C \in \mathcal{C}$ , then  $C \subset U_C \in \mathcal{A}$ , i.e.,  $\mathcal{C}$  is a refinement of  $\mathcal{A}$ ,*
- (b) *if  $C, C' \in \mathcal{C}$  and  $C \cap C' \neq \emptyset$ , then  $C \cap C'$  is convex normal.*

**Proof.** Consider the covering  $\mathcal{C}_0$  made of all convex normal neighborhoods that are subsets of the elements of  $\mathcal{A}$ .  $\mathcal{C}_0$  is well defined as the convex normal neighborhoods form a basis of the topology. Exploiting theorem 2.28, using the fact that  $M$  is Hausdorff and paracompact, consider a refinement  $\mathcal{A}_0$  of  $\mathcal{C}_0$  satisfying, for every  $V \in \mathcal{A}_0$ ,

$$\bigcup \{V' \in \mathcal{A}_0 \mid V' \cap V \neq \emptyset\} \subset C_V$$

for some  $C_V \in \mathcal{C}_0$ . The proof concludes by defining  $\mathcal{C}$  as the family of convex normal neighborhoods subsets of the elements of  $\mathcal{A}_0$  so that (a) is in particular true by construction. To prove (b), we start by observing that, if  $C, C' \in \mathcal{A}_0$ , then  $C \subset V$  and  $C' \subset V'$  for some  $V, V' \in \mathcal{A}_0$ ; if furthermore  $C \cap C' \neq \emptyset$ , we conclude that  $V \cap V' \neq \emptyset$  and thus  $C \cup C' \subset V \cup V' \subset C_V$ . Property (b) is now easy using convex normality of  $C_V$ . First the intersection  $C \cap C'$  is open. Next, if  $p, q \in C \cap C'$ , then the unique geodesic segment  $\gamma : [0, 1] \rightarrow C_V$  joining  $p$  and  $q$  is also completely included in  $C \cap C'$  since it must simultaneously stay in  $C$  and  $C'$ , they being convex normal as well. This means that  $C \cap C' = \exp_p(B)$  for some star-shaped open set  $B \subset T_p(C \cap C')$  centered on the origin of  $T_p M$ .  $\square$ .

### Exercises 7.38.

1. Consider a Riemannian manifold  $(M, g)$  and a geodesic ball  $D_r(p)$  centered on  $p$  with radius  $r > 0$  whose closure stays in a normal neighborhood  $N_p$  of  $p$  (this is the case in particular if  $r < \text{Inj}_p(M)$ ). Prove that  $r$  coincides to length of every geodesic segment joining the center to the boundary and completely included in the geodesic ball. In particular, using Proposition 7.21,

$$D_r(p) = \{q \in M \mid d_{\mathbf{g}}^{(N_p)}(p, q) < r\}. \quad (7.29)$$

---

<sup>7</sup>This proposition is stated in [O'Ne83] as a lemma, but the poof is only sketched concerning a non-trivial use of the paracompactness property.

**Solution.** The boundary of the geodesic ball is one-to-one with the boundary of the corresponding metric ball  $B_r(0) \subset T_p M$  since  $\exp_p$  is a diffeomorphism in an open set including the closure of the ball. A geodesic segment completely contained geodesic ball and joining the center  $p$  with a point  $q$  of the boundary, just in view of the definition of exponential map has the form  $[0, 1] \ni t \mapsto \exp_p(tv) = \gamma(t)$ , where  $q = \exp(v)$  and  $\mathbf{g}(v, v) = r^2$  since  $v \in \partial B_r(0)$ . Hence

$$L(\gamma) = \int_0^1 \sqrt{\mathbf{g}(\gamma'(u), \gamma'(u))} du = \int_0^1 \sqrt{\mathbf{g}(v, v)} du = \int_0^1 r du = r .$$

**2.** Consider a Riemannian or Lorentzian manifold  $(M, \mathbf{g})$  and a convex normal neighborhood  $U$ . Prove that the map called **Synge's world function** in General Relativity

$$U \times U \ni (x, y) \mapsto \sigma(x, y) := \mathbf{g}_x((\exp_x|_{B_x})^{-1}y, (\exp_x|_{B_x})^{-1}y)$$

is well defined, smooth, and

$$\sigma(x, y) = \mathbf{g}_x((\exp_y|_{B_y})^{-1}x, (\exp_y|_{B_y})^{-1}x) .$$

Here, if  $z \in U$ ,  $B_z$  is the starshaped neighborhood centered on  $0 \in T_z M$  such that  $\exp_z(B_z) = U$  according to the definition of convex normal neighborhood. Prove that in the Riemannian case

$$\sigma(x, y) = d_{\mathbf{g}}^{(U)}(x, y)^2$$

**3.** Consider a (pseudo) Riemannian manifold  $(M, \mathbf{g})$  and an open neighborhood  $A \subset M \times M$  of the diagonal  $\Delta := \{(x, x) \in M \times M \mid x \in M\}$ . Prove that there is an open set  $A_0 \subset A$  with  $A_0 \supset \Delta$  such that the Synge world function  $\sigma : A_0 \rightarrow \mathbb{R}$  is well defined and smooth thereon.

**Solution.** Since  $A$  is open,  $(x, x) \in A$ , and the products of elements of a topological basis of  $M$  form a topological basis of  $M \times M$ , we can always pick out a neighborhood  $O_x \subset M$  of every  $x$  such that  $O_x \times O_x \subset A$ . Therefore  $A_1 := \bigcup_{x \in M} O_x \times O_x \subset A$  is another open neighborhood of  $\Delta$  and  $\mathcal{A} := \{O_x\}_{x \in M}$  is an open covering of  $M$ . Applying proposition 7.37, we can construct a covering  $\mathcal{C}$  of  $M$  made of convex normal neighborhoods which are subsets of elements of  $\mathcal{A}$  and such that if  $C, C' \in \mathcal{C}$  and  $C \cap C' \neq \emptyset$ , then  $C \cap C'$  is a convex normal neighborhood. Finally, define  $A_0 := \bigcup_{C \in \mathcal{C}} C \times C$ . With that definition,  $A_0 \subset A$  and  $A_0 \supset \Delta$  as wanted. To conclude, suppose that  $(p, q) \in A_0$ . By definition, there must be a convex normal neighborhood  $C \in \mathcal{C}$  with  $C \ni p, q$ . According to exercise 2,

$$\sigma(p, q) = \mathbf{g}_p((\exp_q|_{B_q})^{-1}p, (\exp_q|_{B_q})^{-1}p) ,$$

is a smooth function in  $C \times C$ . If there is another  $C' \in \mathcal{C}$  such that  $p, q \in C'$ , the analogous value  $\sigma(p, q)$  coincides with the one obtained dealing with  $C$ . That is because that value is nothing  $g_q(v, v)$  of the initial tangent vector  $v$  to the unique geodesic segment in  $C$  (resp.  $C'$ ) from  $q$  to  $p$  (parametrized in  $[0, 1]$ ) and such geodesic segment joining the points in  $C$  (resp.  $C'$ ) coincides with the analogous geodesic segment joining the points in  $C \cap C'$  which is convex normal as well.

4. With  $\sigma(p, q)$  defined as in exercise 2 and referring to a coordinate system, prove that

$$\frac{\partial}{\partial x_{p'}^k} \big|_{p'=p} \sigma(p', q) = -2(g_p)_{kl} v^l,$$

where  $v = v^k \frac{\partial}{\partial x^k} \big|_p$  is the initial tangent vector in  $T_p M$  of the geodesic segment  $\gamma : [0, 1] \rightarrow U$  with  $\gamma(0) = p$  and  $\gamma(1) = q$ . Similarly

$$\frac{\partial}{\partial x_{q'}^k} \big|_{q'=q} \sigma(p, q') = 2(g_q)_{kl} u^l,$$

where  $u = u^k \frac{\partial}{\partial x^k} \big|_q$  is the tangent vector at  $q$  of the geodesic segment  $\gamma : [0, 1] \rightarrow U$  with  $\gamma(0) = p$  and  $\gamma(1) = q$ .

Finally prove that

$$\frac{\partial^2 \sigma(p, q)}{\partial x_p^k \partial x_q^h} \big|_{p=q=r} = -2(g_r)_{kh}.$$

(*Hint.* Start with the second question using Lemma 7.22 or Lemma 7.26 as is the case, observing that  $\sigma(p, q) = r^2(q)$ .)

5. Taking Theorem 7.32 into account, prove that a compact Riemannian manifold has a strictly positive injectivity radius.

**Solution.** If  $p \in M$  there is  $R_p > 0$  such that a geodesic ball  $\Gamma_{r_p}(p)$  of radius  $r_p < R_p$  centered on  $p$  is a convex normal neighborhood of  $p$ . For every  $p \in M$ , fix such a ball of radius  $0 < \rho_i := (r_p/2) - \epsilon_i$  for some small  $\epsilon_i$ . Using compactness extract a finite covering  $\{\Gamma_{\rho_i}(p_i)\}_{i=1, \dots, N}$  from the said covering. Every  $p \in M$  necessarily belongs to some  $\Gamma_{\rho_i}(p_i)$ . Since  $\Gamma_{\rho_i}(p_i) \subset \Gamma_{r_{p_i}}(p_i)$  and the latter is normal convex, a geodesic ball of radius  $r_i/2$  centered on  $p$  define a normal neighborhood of  $p$  from trivial properties of metric balls in  $\Gamma_{r_{p_i}}(p_i)$ , if taking (7.29) into account. Therefore  $\text{Inj}(M) \geq \min\{r_i/2 \mid i = 1, 2, \dots, N\}$ .

6. Prove that  $\text{Inj}(M) > 0$  for a Riemannian manifold  $(M, g)$ , then the manifold is geodesically complete,  $DE = TM$ , and thus in particular  $\exp : TM \rightarrow M$  is smooth. Conclude that compact Riemannian manifolds have the features above,

(*Hint.* Assume that there is a maximal geodesic  $\gamma : (a, b) \rightarrow M$  with  $b < +\infty$ . We can always suppose that the affine parameter is the arch length, so that  $\gamma$  has finite length if evaluated from every fixed  $r_0 \in (a, b)$  towards  $b$ . Consider a normal neighborhood of each element of a sequence of points  $\gamma(s_k)$  with  $s_k \rightarrow b$  which are geodesic balls. Prove that the radius of those geodesic balls must tend to 0 as  $n \rightarrow +\infty$ , implying that  $\text{Inj}(M) = 0$  against the hypothesis. The same argument is valid if  $a > -\infty$ .)

7. Consider a smooth vector field  $X$  on the complete Riemannian manifold  $(M, g)$ . Prove that, if

$$\sup_{p \in M} g_p(X, X) < +\infty,$$

then  $X$  is complete as well.

**Solution.** Let  $\gamma : (\alpha, \omega) \rightarrow M$  be an integral curve of  $X$  such that  $\omega < +\infty$  (the other case is similar). Take  $t_0 \in (\alpha, \omega)$  and consider the continuous map  $[t_0, \omega) \ni t \mapsto d_{\mathbf{g}}(\gamma(t_0), \gamma(t)) =: f(t)$ . If this function is bounded we have a contradiction since Theorem 7.15 proves that the closed metric balls of finite radius are compact and Proposition 4.19 implies  $\omega = +\infty$ . Since  $f$  is continuous and therefore bounded on each interval  $[t_0, u]$  with  $t_0 < u < \omega$ , it must exist a sequence  $t_n \rightarrow \omega$  such that  $d_{\mathbf{g}}(\gamma(t_0), \gamma(t_n)) \rightarrow +\infty$ . If  $s = s(t)$  is the length coordinate on  $\gamma$  evaluated from some point on the curve, it holds for  $t_n \rightarrow \omega$

$$\left| \frac{ds}{dt} \Big|_{u_n} \right| = \frac{|s(t_n) - s(t_0)|}{t_n - t_0} \geq \frac{d_{\mathbf{g}}(\gamma(t_0), \gamma(t_n))}{t_n - t_0} \rightarrow +\infty,$$

where we exploited the Lagrange theorem and the points  $u_n$  stay in  $[t_0, t_n]$ . This is impossible because due to the very definition of length coordinate,

$$\left( \frac{ds}{dt} \right)^2 = \mathbf{g}_{\gamma(t)}(X(\gamma(t)), X(\gamma(t)))$$

is bounded.

**8.** Extend to piecewise smooth curves the statements of Propositions 7.21 and 7.25.

(*Hint.* Use the same sketch of proof in every smoothness subinterval and then exploit the additivity property of the integral.)

**9.** Consider a Lorentzian manifold  $(M, g)$  and a Killing vector  $K$ . Prove that if  $K$  is light-like, i.e.,  $\mathbf{g}(K, K) = 0$  then it is parallelly transported with respect to itself:  $\nabla_K K = 0$ .

(*Hint.* Use the Killing equation in terms of covariant derivative and contract both sides with  $K$ .)

**10.** Prove Proposition 6.13 using the properties of normal coordinates. Prove more generally that

$$\nabla_a K_b + \nabla_b K_a = \mathcal{L}_K g_{ab} \quad (7.30)$$

where  $\nabla$  is the Levi Civita derivative associated to  $\mathbf{g}$ .

**Solution.** Fix a system of normal coordinates centered on  $p$  and observe that, exactly at  $p$ ,

$$\mathcal{L}_K \mathbf{g} = \frac{\partial g_{ab} K^b}{\partial x^a} + \frac{\partial g_{ba} K^a}{\partial x^b}$$

since  $\frac{\partial g_{ab}}{\partial x^c} \Big|_p = 0$  due to (b) in Proposition 7.29. In turn, for the same reason, that identity can be re-written

$$\mathcal{L}_K \mathbf{g} = \nabla_a K_b + \nabla_b K_a$$

exactly at  $p$ . This identity is now independent from the coordinate system and the point.

## Chapter 8

# General Relativity: a Geometric Presentation

We have eventually accumulated a number of mathematical tools which permit us to formalize the ideas that stay at the basis of the General Theory of Relativity.

**Remark 8.1.** We henceforth use the notions, definitions and notations introduced and discussed in [Mor20] (see also Example 6.29). ■

### 8.1 The spacetime of General Relativity

We start with a mathematical presentation of the notion of spacetime and its physical interpretation which are nothing but a generalization of the analogs in Special Relativity.

In General Relativity, exactly as in Special Relativity, the spacetime  $(M^4, \mathbf{g})$  is supposed to be a smooth<sup>1</sup> manifold equipped with a smooth Lorentzian metric  $\mathbf{g}$ . We however generally consider  $n$ -dimensional spacetimes  $(M^n, \mathbf{g})$ .

**Definition 8.2.** A  *$n$ -dimensional spacetime* is a  $n$ -dimensional Lorentzian manifold  $(M, \mathbf{g})$  (thus connected, Hausdorff, second countable and the signature of  $\mathbf{g}$  is  $(-1, +1, \dots, +1)$ ). ■

It should be clear the reason why we assume that the spacetime is connected: no communication would exist between different connected components, at least with the presently known physical laws.

**Remark 8.3.** There exist many hypotheses about the possible discontinuous nature – or in any case not described by a 4 dimensional Lorentzian manifold – of spacetime at very small scales: the Planck scales  $\sim 10^{-33}$  cm and  $\sim 10^{-43}$  s. Some quantum theory of spacetime should

---

<sup>1</sup>Actually, everything could be formulated with much less regularity [HaEl75], but we are not interested here in discussing these issues and we adopt the most comfortable regularity hypotheses.

pop out there. There are different proposals as the *string theory* in its different variants, the *em* loop quantum gravity and various approaches based on *non-commutative geometry*. However recent experimental observations made with the “Fermi Gamma-ray” space telescope, concerning the so-called  $\gamma$  -bursts, have lowered the threshold for the existence of some quantum gravity phenomena, such as violation of Lorentz symmetry, well below the Planck scale<sup>2</sup>. ■

### 8.1.1 Types of vectors and curves in a spacetime

As a general informal principle, the geometric structure and the physical interpretation of what exists around an event  $p \in M^n$  are the same as for the corresponding objects in Minkowski spacetime  $\mathbb{M}^n$ . We formalize later this idea in a specific principle of General Relativity. For the moment we make the following assumptions.

**Definition 8.4.** If  $(M^n, \mathbf{g})$  is a spacetime, a vector  $X_p \in T_p M^n$  is called

- (a) **spacelike** if  $\mathbf{g}(X_p, X_p) > 0$ ,
- (b) **timelike** if  $\mathbf{g}(X_p, X_p) < 0$ ,
- (c) **lightlike** if  $\mathbf{g}(X_p, X_p) = 0$  but  $X_p \neq 0$ ,
- (d) **causal** if it is timelike or lightlike.

The **open light cone** at  $p \in M^n$  is

$$V_p := \{X_p \in T_p M^n \mid \mathbf{g}(X_p, X_p) < 0\} \subset T_p M^n. \quad (8.1)$$

With that definition, the **light cone** at  $p$  is  $\partial V_p$  and **closed light cone** at  $p$  is  $\overline{V_p}$ . ■

Exactly as in Minkowski spacetime,  $V_p$  is the disjoint union of two halves whose closures meet at  $0 \in T_p M$ . Fixing a (pseudo) orthonormal basis  $\{e_k\}_{k=1,\dots,n} \subset T_p M^n$  with  $e_1$  timelike and the remaining vectors spacelike, and decomposing  $T_p M^n \ni v = v^k e_k$ , these two halves are

$$V_p^{(>)} := \left\{ v \in T_p M \mid \sum_{\alpha=2}^n (v^\alpha)^2 < (v^1)^2, \quad v^1 > 0 \right\}$$

and

$$V_p^{(<)} := \left\{ v \in T_p M \mid \sum_{\alpha=2}^n (v^\alpha)^2 < (v^1)^2, \quad v^1 < 0 \right\}.$$

Proposition 8.2 in [Mor20] establishes in particular that this decomposition of  $V_p$  does not depend on the used basis. Timelike vectors stay in those halves, whereas lightlike vectors belong to

---

<sup>2</sup>A. A. Abdo et al. *em* A limit on the variation of the speed of light arising from quantum gravity effects. Nature 462, 331-334 (19 November 2009).



the light cone at  $p$ , and causal vectors, fill the closed light cone at  $p$  which also contain  $0 \in T_p M^n$ .

**Definition 8.5.** If  $(M^n, \mathbf{g})$  is a spacetime, a piecewise smooth curve  $I \ni u \mapsto \gamma(u) \in M^n$  is called

- (a) **spacelike** if  $\gamma'(u)$  is spacelike for all  $u \in I$ ,
- (b) **timelike** if  $\gamma'(u)$  is timelike for all  $u \in I$ ,
- (c) **lightlike** if  $\gamma'(u)$  is light like (hence  $\gamma'(u) \neq 0$ ) for all  $u \in I$ .

The conditions above must be valid for both the tangent vectors at every singular point of the domain  $I$ .

- (d) **Causal curves** are those piecewise smooth curves which are piecewise timelike or lightlike indifferently. ■

**Remark 8.6.** The above classification is invariant under re-parametrization  $u = u(s)$  where  $u'(s) \neq 0$  everywhere. ■

### 8.1.2 Time orientation of a spacetime

To avoid physical problems with the chronological order of events, it is also assumed that there exist a *time orientation* of the spacetime, that is a continuous choice of one of the two halves  $V_p^{(>)}$  and  $V_p^{(<)}$  of every every  $V_p$ . That continuous choice is made by assigning (if possible) a smooth (continuous would be sufficient) timelike vector  $T$ . As a matter of fact,  $T_p$  chooses the half  $V_p^+$  of  $V_p$  defining the local future.

**Proposition 8.7.** Assume that two smooth timelike vector fields  $T$  and  $T'$  exist on the spacetime  $(M^n, g)$ . Then the following facts hold.

- (a)  $\mathbf{g}(T_p, T'_p) < 0$  for every  $p \in M^n$  or  $\mathbf{g}(T_p, T'_p) > 0$  for every  $p \in M^n$ .
- (b)  $T$  and  $T'$  choose the same half at  $p \in M^n$  if and only if  $\mathbf{g}(T_p, T'_p) < 0$  for every  $p \in M^n$ .

**Proof.** Suppose that  $\mathbf{g}(T_p, T'_p) < 0$  and  $\mathbf{g}(T_q, T'_q) > 0$ . Since  $M^n$  is connected then (Proposition 5.4) there is a continuous curve  $[a, b] \ni u \mapsto \gamma(u) \in M^n$  with  $\gamma(a) = p$  and  $\gamma(b) = q$ . As the function  $[a, b] \ni u \mapsto \mathbf{g}(T_{\gamma(u)}, T'_{\gamma(u)})$  is continuous, it must vanish for some  $t_0 \in (a, b)$ , but this is not possible because either  $\mathbf{g}(T_{\gamma(t_0)}, T'_{\gamma(t_0)}) > 0$  or  $\mathbf{g}(T_{\gamma(t_0)}, T'_{\gamma(t_0)}) < 0$  since both vectors are timelike. (b) can be proved as we did in Proposition 8.2 in [Mor20] or by direct inspection. □

All that leads to the following definition.

**Definition 8.8.** (**Time orientation.**) Consider a spacetime  $(M^n, \mathbf{g})$ .

- (a)  $(M^n, \mathbf{g})$  is called **time orientable** if it admits a smooth timelike vector field  $T$ .

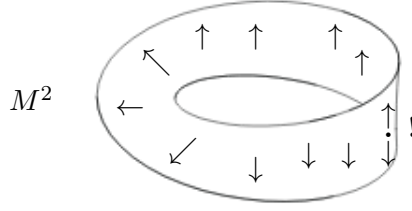


Figure 8.1: The Minkowski - Möbius strip spacetime cannot be time oriented

- (b) Two such vector fields  $T, T'$  are said to **define the same time-orientation** if

$$\mathbf{g}(T_p, T'_p) < 0$$

for every  $p \in M^n$  (i.e., they determine the same half of  $V_p$  at every point  $p \in M^n$ ).

- (c) An **orientation** on  $(M^n, \mathbf{g})$  is a class of equivalence of the above equivalence relation.
- (d) An **oriented spacetime** is a spacetime with a choice of a preferred time orientation.
- (e) If  $(M^n, \mathbf{g})$  is oriented,  $V_p^+ \subset V_p$  is the **future open cone** defined at  $p \in TM^n$  by the orientation and  $\overline{V_p^+}$  and  $\partial V_p^+$  are defined accordingly. ■

**Remark 8.9.**

- (1) It is not difficult to prove that, since  $M^n$  is assumed to be connected, if  $M^n$  is orientable, then it admits *exactly two* possible time orientations as it happened to  $\mathbb{M}^4$ .
- (2) There are spacetimes  $(M^n, \mathbf{g})$  which do not admit a time orientation. The basic example is the Möbius strip ((1) Examples 3.30) obtained by a two-dimensional Minkowski spacetime with spatial Minkowskian axis along the basis  $S^1$  and temporal Minkowskian axis along the fiber  $\mathbb{R}$ . ■

The given definition reflects into a finer classification of vectors and curves.

**Definition 8.10.** In a time-oriented spacetime  $(M^n, \mathbf{g})$ ,

- (a) a non-vanishing causal vector  $X_p \in \overline{V_p}$  is said to be **future-directed** (or **future-oriented**) if  $X_p \in \overline{V_p^+}$ ;
- (b) A piecewise smooth causal curve  $\gamma : I \rightarrow M^n$  is **future-directed** (or **future-oriented**) if  $\gamma'(u) \in \overline{V_p^+}$  for very  $t \in I$ , where that condition must be valid in particular for both the tangent vectors at every singular point of the domain  $I$ .

- (c) A (piecewise) smooth future-directed causal curve  $I \ni u \mapsto \gamma(u) \in M^n$ , where  $I$  is an interval possibly containing one or both its endpoints, is called **worldline**. ■

**Remark 8.11.**

(1) The above classification of curves is invariant under re-parametrization  $u = u(s)$  provided  $u'(s) > 0$  everywhere.

(2) In a time oriented spacetime  $(M^n, \mathbf{g})$ , a piecewise smooth timelike curve  $I \ni t \mapsto \gamma(t) \in M^n$  is future-directed or past-directed if and only if  $\mathbf{g}(\gamma'(t_i^-), \gamma'(t_i^+)) < 0$  with obvious notations, for every non-smoothness point  $t_i \in I$ . The proof is trivial by exploiting Proposition 8.7. ■

The notions of **past-directed** (or **past-oriented**) vector and causal curve are defined similarly. As in Special Relativity, physically speaking, worldlines describe the histories of physical objects (material points) evolving in the universe.

**Remark 8.12.** *From now on a spacetime will be always supposed to be time oriented.* ■

### 8.1.3 Proper time and rest space associated to observers

We now focus our attention to those worldlines that are timelike and future oriented. We shall call these worldlines **observers** from now on, since we want to equip them with some geometric tools corresponding to physical instruments co-moving with them and useful to explore the spacetime in an “infinitesimal” region around each event reached by the worldline.

According to the general principle that Special Relativity is assumed to be valid in a neighborhood of an event (and this principle will be clearly stated in the next section in terms of the so-called Strong Equivalence Principle) we can give the following definitions when dealing with a timelike worldline  $I \ni u \mapsto \gamma(u) \in M^n$ .

- (a) The affine parameter – the length coordinate up to the universal constant  $c > 0$  with the physical meaning of the *speed of light* –

$$\tau(u) := \frac{1}{c} \int_{u_0}^u \sqrt{|\mathbf{g}(\gamma'(\xi), \gamma'(\xi))|} d\xi, \quad (8.2)$$

is called **proper time** of the particle whose history is  $\gamma$ . It describes the temporal coordinate measured with an *ideal clock* at rest with the particle whose history is  $\gamma$ . When parametrizing a timelike worldline with the proper time, the tangent vector  $\gamma'(\tau)$  will be denoted by  $\dot{\gamma}(\tau)$  and that vector will be called the  **$n$ -velocity** of the material point. It holds

$$\mathbf{g}(\dot{\gamma}(\tau), \dot{\gamma}(\tau)) = -c^2. \quad (8.3)$$

- (b) The tangent space  $T_{\gamma(\tau)}M^n$  is *orthogonally* decomposed as

$$T_{\gamma(u)}M^n = d\Sigma_{\gamma(\tau)} \oplus \text{Span}(\dot{\gamma}(\tau)), \quad (8.4)$$

where  $d\Sigma_{\gamma(\tau)}$  is a  $n - 1$  subspace of  $T_{\gamma(\tau)}M^n$  equipped with a (positive) scalar product induced from the Lorentzian metric  $\mathbf{g}$ .  $d\Sigma_{\gamma(\tau)}$  is the **rest space** of  $\gamma$  at proper time  $\tau$ . It represents the “infinitesimal” *rest space* of a reference frame transported by the material point whose history is  $\gamma$ .

Exactly as in Special Relativity (see Section 8.2.2 of [Mor20]), if  $\sigma = \sigma(\tau)$  is another causal wordline crossing  $\gamma(\tau_0)$ , we can define the **velocity** of  $\sigma$  at  $\tau_0$  (proper time of  $\gamma$ ) **with respect to**  $\gamma$  as

$$\mathbf{v}_{\sigma,\tau,\gamma} := \frac{\delta X}{\delta \tau} \in d\Sigma_{\gamma(\tau)}, \quad (8.5)$$

where

$$\sigma' = \delta t \dot{\gamma}(\tau) + \delta X \quad \text{according to (8.4)}.$$

Exactly as in Proposition 8.16 of [Mor20] (using the same proof) the following statement is valid.

**Proposition 8.13.** *In a spacetime  $(M^n, \mathbf{g})$ , with the given definition (8.5) of the velocity  $\mathbf{v}_{\sigma,\tau,\gamma}$  of  $\sigma$  at  $\tau_0$  with respect to  $\gamma$ ,*

$$\|\mathbf{v}_{\sigma,\tau,\gamma}\| := \sqrt{\mathbf{g}(\mathbf{v}_{\sigma,\tau,\gamma}, \mathbf{v}_{\sigma,\tau,\gamma})} \leq c$$

*where the value  $c$  is reached if and only if  $\sigma'$  is lightlike at the considered event.*

As a consequence, also in General Relativity, lightlike worldlines describe *histories of material points travelling at the speed of light for every observer*. Furthermore, the speed of light if measured as in (8.5) is again the maximal admitted speed.

**Remark 8.14.** It is interesting to note that, in convex normal neighborhoods in  $M^n$ , the geodesics of  $(M^n, \mathbf{g})$  maximize the proper time measured between two events they join as a consequence of Proposition 7.25 and (3) in Remark 7.36. This result appears as a straightforward physical generalization of (c) in Proposition 8.16 in [Mor20], even if the proof of the latter is much more elementary than the former. ■

## 8.2 A physical justification of the mathematical formalism

Even if the above picture is a nice generalization of the structure of the Special Theory of Relativity and a mathematician may be happy just for that, the natural question arising in theoretical physicist’s mind should be now “Why should we relax the rigid but physically well motivated affine space structure of Minkowski spacetime into the one of smooth (though Lorentzian) manifold? What do we *physically* gain from this weaker assumption?”

The first answer to this crucial question is discussed in the next section and it is based on a deep physical principle discovered by Einstein.

### 8.2.1 Einstein's equivalence principle in Physics

The so-called *Equivalence Principle* relies upon Einstein's observation that the *gravitational mass* and the *inertial mass* of every known physical body experimentally coincide.

- (1) The **gravitational mass** is the constant  $M$  taking place in *Newton's universal gravitation formula*,

$$\vec{F} = -G \frac{MM'}{\|P - Q\|^3} (P - Q) ,$$

where  $\vec{F}$  is the gravitational force acting on the body with mass  $M$  and due to the body of mass  $M'$ .

- (2) The **inertial mass**  $m$  is instead the other constant, always proper of a given body, that enters Newton's second law of classical dynamics

$$\vec{F}(t, P, \vec{v}) = m\vec{a} .$$

Above  $\vec{a}$  is the acceleration of the considered body, here viewed as a material point  $P$ , with velocity  $\vec{v}$ , everything referred to an inertial reference frame  $\mathcal{I}$ , when the point is subjected to the force  $\vec{F}$ .

Newton postulated

$$M = m .$$

This coincidence of values has been experimentally checked with a very high precision with several experiments. Celebrated Eötvös' experiment in 1908 confirmed that identity with a sensibility of  $10^{-9}$  exploiting a torsional pendulum. Later, Dicke and co-workers experimentally confirmed the identity in 1964 with a sensibility of  $10^{-12}$ . In 2017 the satellite MICROSCOPE confirmed the identity with a sensibility of  $10^{-15}$ . If assuming the coincidence of gravitational and inertial masses, a crucial physical fact first observed by Einstein pops out, embodied in his famous *Equivalence Principle*.

**Equivalence Principle.** *It is always possible to locally cancel the dynamical effect of a gravitational field by means of a suitable choice of the reference frame where one describes the motion of a material point. Vice versa, it is always possible to locally create the dynamical effect of a gravitational field by means of a suitable choice of the reference frame where one describes the motion of a material point.*

**Remark 8.15.** *Locally above means in suitable small spatial regions of space and for suitably short intervals of time.* ■

Let us enter into the details of this idea sticking to framework of classical physics. Let us consider a classical gravitational field  $\vec{g} = \vec{g}(t, P)$  described in the rest coordinates of an inertial reference frame  $\mathcal{I}$ . The vector field  $\vec{g}(t, P)$  is therefore the *gravitational acceleration* vector

field for a material point  $P$  in the rest space if  $\mathcal{S}$  at time  $t$ . Let us change the reference frame passing to the *non-inertial* one  $\mathcal{S}'$  that is free falling in that gravitational field. This reference frame is practically constructed first of all staying at rest with a particle  $O$  in motion in  $\mathcal{S}$  with acceleration

$$\vec{a}_O|_{\mathcal{S}}(t) = \vec{g}(t, O(t)) .$$

Secondly, equipping  $O$  with a triple of Cartesian axes centered on  $O(t)$  which are assumed *not to rotate* in  $\mathcal{S}$ . If the material point  $P$ , with inertial mass  $m$  and gravitational mass  $M$ , is initially placed at  $O$  with some initial velocity, its motion in  $\mathcal{S}$  is described by Newton's equation

$$m\vec{a}_P|_{\mathcal{S}} = M\vec{g}(t, P(t)) .$$

Since the relative motion of  $\mathcal{S}$  and  $\mathcal{S}'$  is translational, we also have

$$\vec{a}_P|_{\mathcal{S}}(t) = \vec{a}_P|_{\mathcal{S}'}(t) + \vec{a}_O|_{\mathcal{S}}(t) .$$

We conclude that the equation of motion of  $P$  in  $\mathcal{S}'$  can be re-written

$$m\vec{a}_P|_{\mathcal{S}'}(t) = -m\vec{a}_O|_{\mathcal{S}}(t) + M\vec{g}(t, P(t)) ,$$

that is, taking advantage of  $\vec{a}_O|_{\mathcal{S}}(t) = \vec{g}(t, O(t))$ ,

$$m\vec{a}_P|_{\mathcal{S}'}(t) = -m\vec{g}(t, O(t)) + M\vec{g}(t, P(t)) .$$

Eventually, since  $m = M$ , we obtain

$$m\vec{a}_P|_{\mathcal{S}'}(t) = m(\vec{g}(t, P(t)) - \vec{g}(t, O(t))) .$$

We see that, if  $P$  is close to the origin  $O$  of  $\mathcal{S}'$  – and this is the case for sufficiently small times since  $P$  was initially placed at  $O$  – then

$$\vec{g}(t, P(t)) - \vec{g}(t, O(t)) \sim 0$$

so that  $\vec{a}_P|_{\mathcal{S}'}(t)$  is arbitrarily small and the motion of  $P$  in  $\mathcal{S}'$  turns out to be similar to a straight inertial motion: as if the point were not subjected to the gravitational field.

Conversely, in the absence of gravitational fields, we can however simulate the dynamical effect of a gravitational field by describing the motion of a given point  $P$  in a non-inertial reference frame  $\mathcal{S}'$ , which is suitably accelerated with respect to an inertial reference frame  $\mathcal{S}$ . Indeed, with the same definition of  $O$  at rest with  $\mathcal{S}'$  and the non-rotating Cartesian axes centered on  $O$ , suppose that the acceleration  $\vec{a}_O|_{\mathcal{S}}$  is given. The equation of motion of the material point  $P$  in absence of forces is, in the inertial reference frame  $\mathcal{S}$ ,

$$m\vec{a}_P|_{\mathcal{S}} = \vec{0} ,$$

Therefore, in the reference frame  $\mathcal{S}'$ ,

$$m\vec{a}_P|_{\mathcal{S}'}(t) = -m\vec{a}_O|_{\mathcal{S}} .$$

In other words, where we write  $M$  in place of  $m$  in the right-hand side since  $M = m$ ,

$$m\vec{a}_P|_{\mathcal{S}'}(t) = M\vec{g}. \quad (8.6)$$

Here a gravitational field  $\vec{g} = -\vec{a}_O|_{\mathcal{S}}$  seems to take place.

**Remark 8.16.** Replacing the gravitational field with another interaction, e.g., the electromagnetic one (assuming that the material point  $P$  carries an electrical charge) we would not be able to achieve the same results. *This feature is proper of the gravitational interaction.* ■

In the next section, we prove that if we adopt the geometric description of the spacetime of the General Relativity as we presented in Section 8.1 and if we further assume the further postulate:

**Geodesic Postulate.** *The worldlines of the material points classically described as free-falling bodies in a gravitational field are represented in  $(M^4, \mathbf{g})$  by causal future-directed geodesics of the Levi-Civita connection;*

then the first half of the equivalence principle becomes a well-known geometric fact of (pseudo) Riemannian geometry.

**Remark 8.17.**

(1) Another relevant fact connected with the above discussion is that, in a given gravitational field, the classical motion of a free falling body depends on its initial velocity but not on its mass (just because the two types of masses cancel each other in the equation of motion). Also this fact is compatible with Geodesic Postulate, since geodesics are completely determined by their initial point, their initial vector and no information is necessary concerning the mass of the material point. After all, we did not have yet introduced the notion of mass in the general relativistic description!

(2) As soon as we assume the Geodesic Postulate, we are committed to accept the idea that the properties of the classical gravitational field are now embodied in the metric  $\mathbf{g}$ . From this viewpoint, passing from Special to General Relativity,  $\mathbf{g}$  must have further properties than the metrical ones. Furthermore, we are also forced to assume that the gravitational interaction is *not* described by a 4-force, differently from the electromagnetic interaction for instance.

(3) In Special Relativity, causal future oriented geodesics are nothing but causal future oriented affine segments since the connection coefficients vanish in Minkowskian coordinates. The Geodesic Postulate in Special Relativity therefore selects the histories of material points that are not subjected to forces. ■

### 8.2.2 Equivalence principle and normal coordinates

Let us assume that causal geodesics are the histories of classically free-falling objects. We stick to the case of a fourdimensional spacetime  $M^4$ , but everything we say below does not depend on that choice.

Consider a timelike geodesic segment  $\gamma : I \ni t \mapsto M^4$  parametrized by its *proper time*  $t$ , so that  $\mathbf{g}(\dot{\gamma}, \dot{\gamma}) = -1$  where we are assuming that  $c = 1$ . Consider a normal coordinate system around  $\gamma$  including  $\gamma(t_0)$  in its domain for a fixed  $t_0 \in I$  according to Definition 7.30. Let us denote by  $(t, x^1, x^2, x^3)$  the coordinates of the normal chart.

The property  $\Gamma_{jk}^i(\gamma(t)) = 0$  (Proposition 7.29) produces a very clear formalization of part of Einstein's Equivalence Principle, which is automatically embodied in the mathematical formalism of the General Relativity when assuming the Geodesic Postulate.

If we study the histories causal geodesics  $\alpha = \alpha(t)$  of free-falling bodies leaving  $\gamma$  at some  $t_i$  using the coordinates  $(t, x^1, x^2, x^3)$ , we discover that they are represented as straight non-accelerated motions for sufficiently small times around  $t_1$ . In other words, *locally, the gravitational effect has been suppressed by a suitable choice of the reference frame.*

Without lack of generality, we may assume  $t_1 = 0$  as well as  $\gamma(0) \equiv (0, 0, 0, 0)$  in the used coordinates  $(y^1, y^2, y^3, y^4) = (t, x^1, x^2, x^3)$ . The equations of  $\alpha$  have the form

$$\frac{d^2 y^i}{d\lambda^2} = -\Gamma_{jk}^i(\alpha) \frac{dy^j}{d\lambda} \frac{dy^k}{d\lambda}, \quad i, j, k = 1, 2, 3, 4,$$

where  $\lambda$  is an affine parameter of the geodesic  $\alpha$ . We arrange  $\lambda$  such that  $\lambda = 0$  defines the event  $\gamma(0)$ , where  $\alpha$  departs from  $\gamma$ . Taylor expansion around  $\lambda = 0$  yields:

$$y^i(\lambda) = 0 + \lambda \frac{dy^j}{d\lambda} \Big|_{\lambda=0} + \frac{\lambda^2}{2} \frac{d^2 y^j}{d\lambda^2} \Big|_{\lambda=0} + \lambda^2 O^j(\lambda),$$

where  $O^j(\lambda) \rightarrow 0$  as  $\lambda \rightarrow 0$ . However, in the considered coordinates, in view of Proposition 7.29, we have  $\Gamma_{jk}^i(\alpha(0)) = 0$ , so that:

$$\frac{d^2 y^j}{d\lambda^2} \Big|_{\lambda=0} = -\Gamma_{jk}^i(\alpha(0)) \frac{dy^j}{d\lambda} \Big|_{\lambda=0} \frac{dy^k}{d\lambda} \Big|_{\lambda=0} = 0,$$

and thus

$$y^i(\lambda) = \lambda \frac{dy^j}{d\lambda} \Big|_{\lambda=0} + \lambda^2 O^j(\lambda). \quad (8.7)$$

In particular, for the first coordinate  $y^1 = t$ , we find

$$t(\lambda) = \lambda \frac{dt}{d\lambda} \Big|_{\lambda=0} + \lambda^2 O^1(\lambda). \quad (8.8)$$

Since the geodesics  $\gamma$  is timelike and  $\alpha$  is causal, it must be

$$\frac{dt}{d\lambda} \Big|_{\lambda=0} \neq 0.$$

This fact implies that, around  $t = 0$ , we can use  $t$  as a parameter for the smooth curve  $\alpha$ . Starting with the parameter  $t$  from scratch, we can write

$$x^\alpha(t) = t \frac{dx^\alpha}{dt} \Big|_{t=0} + t^2 O^\alpha(t), \quad \alpha = 1, 2, 3. \quad (8.9)$$



The velocity of  $\alpha$  respect to  $\gamma$  at  $t = 0$  has components

$$v^\alpha = \frac{dx^\alpha}{dt}|_{t=0}, \quad \alpha = 1, 2, 3,$$

(and this is also in agreement with the definition (8.5) since  $\frac{dt}{dt} = 1$ ) so that,

$$x^\alpha(t) = tv^\alpha + t^2 O^\alpha(t), \quad \alpha = 1, 2, 3. \quad (8.10)$$

where  $O^\alpha(t) \rightarrow 0$  as  $t \rightarrow 0$ .

Up to terms of third order in  $t$  – and this is a precise interpretation of the word “locally” in the formulation of the Equivalence Principle – the motion of  $\alpha$  is given by a constant-velocity motion. The acceleration (classically due to the gravitational field) has been suppressed by an appropriate choice of the coordinate system.

In this sense, the the first part of the equivalence principle is automatically encapsulated in the assumption that  $M^4$  is a Lorentzian manifold where causal geodesics describe the evolution of free-falling bodies. The possibility to locally simulate the existence of a gravitational field with the choice of the reference frame is a much more delicate issue since, up to now, actually we do not know what the gravitational field is described in General Relativity.

### 8.3 The Strong Equivalence Principle and the extension of the formalism

Once we have justified the geometric description of spacetime in General Relativity, in particular assuming the Geodesic Postulate, we can proceed with the construction of the formalism including further important physical notions like mass, energy, momentum for pointwise and extended physical objects.

#### 8.3.1 The Strong Equivalence Principle

The idea is to export as much as possible from Special Relativity to General Relativity exploiting the fact that in a neighborhood of an event the geometry of Special Relativity and the one of General Relativity are very similar when dealing with normal coordinates.

We start from the mathematical observation that, if dealing with normal coordinates centered on an event  $p$  (and defining those coordinates by fixing a pseudo orthonormal basis at the origin of  $T_p M^n$ ) all definitions and law of Special Relativity written in Minkowskian coordinates which at most include *first order derivatives* at  $p$  can be rewritten as they stand in normal coordinates at  $p$ . In particular because  $\Gamma_{ab}^a(p) = 0$  (Proposition 7.18) so that we cannot distinguish between ordinary derivatives  $\frac{\partial}{\partial x^k}|_p$  (used in Minkowskian coordinates of Special Relativity) and covariant derivatives  $\nabla_k|_p$ .

From the physical side, these coordinates can be viewed as the normal coordinates around a timelike geodesic passing through the said event. With this interpretation, the principle we

are about stating asserts that, in those free falling laboratories, the laws of Special Relativity valid in inertial reference frames, i.e., Minkowskian coordinates, are locally valid also in General Relativity, as if the gravitational interaction were not present, provided these laws are local and derivatives appear up to the first order. In this sense, inertial frames of the Special Relativity and free falling frames of General Relativity are “equivalent” for the formulation of all physical laws of a certain type. Once accepted the principle, the exported laws written in normal coordinates using the formalism of tensors and the covariant derivative have universal validity in every coordinate system around the given event of the spacetime of General Relativity.

**Strong Equivalence Principle.** *A physical law, including definitions and principles, that holds in Special Relativity and that can be stated in terms of identities of tensors and first-order derivatives of tensors at a given event of the spacetime is also valid in General Relativity, provided the derivatives are replaced by the corresponding covariant derivatives.*

The identification of the proper time  $\tau$  (measured with a clock at rest with an observer) with the length parameter and the identification of  $d\Sigma_{\gamma(\tau)}$  with the rest space of the observer represented by a timelike world line  $\gamma$ , as we did in Section 8.1.3 are immediate consequence of the Strong Equivalence Principle starting from the corresponding identifications in Special Relativity. However we spend some words about the physics involved in these notions in the spirit of the said principle. These definitions of local nature are proper of inertial frames of Special Relativity so that they can be exported to free falling observers frames provided by timelike geodesics and normal coordinates around them. Let now consider two worldlines  $\gamma = \gamma(u)$  and  $\gamma_1 = \gamma_1(u)$  crossing at  $p \in M^n$  with the same  $n$ -velocity and assume that  $\gamma_1$  is a timelike geodesic. Adopting the definitions in Section 8.1.3 we are actually assuming that  $\frac{d\tau}{du}$  and the rest space  $d\Sigma_{\gamma_1(u)}$  of  $\gamma_1$  are the same as for  $\gamma$  exactly at  $p$ . In other words *acceleration does not matter for ideal clocks and ideal rulers*. If one thinks that acceleration must matter, then the approach can be reversed assuming that, for a generic observer  $\gamma$  different from a timelike geodesic, the notion of proper time and rest space are at each instant *by definition* the ones of a free falling observer instantaneously at rest with  $\gamma$ .

We next pass to export to General Relativity the notion of  $n$ -momentum and mass of a material point and their basic properties exploiting the Strong Equivalence Principle. This is possible because all the involved laws and definitions can be stated referring to an event and at most first-order derivatives enter the game (see in particular remark 8.31 in [Mor20] concerning (d)).

If  $\gamma = \gamma(s)$  is a future-oriented causal world-line describing the history of a particle then

- (a) there is a smooth vector field  $P = P(s)$  parallel to  $\gamma'(s)$  which is called  **$n$ -momentum** and it identifies with the  $n$ -momentum of Special Relativity in every normal coordinate system centered on every event reached by the worldline;
- (b) if the worldline is a geodesic  $\nabla_{\gamma'}\gamma' = 0$ , it is assumed that  $P$  is the tangent vector referred

to an *affine parametrization* so that it is parallely transported along the worldline

$$\nabla_{\gamma'} P = 0 \quad \text{or, equivalently,} \quad \nabla_P P = 0. \quad (8.11)$$

Requirement (8.11) is more weakly valid at a possibly isolated event  $e$  reached by the worldline if  $\nabla_{\gamma'} \gamma' = 0$  is valid at that event;

- (c) the **mass** of the particle is defined as the function  $m(s) \geq 0$  such that

$$\mathbf{g}(P(s), P(s)) = -m(s)^2 c^2, \quad (8.12)$$

and it identifies with the mass of Special Relativity in every normal coordinate system centered on every event reached by the worldline;

- (d) Consider  $N_{in}$  material points evolving along worldlines till a common event  $e$  giving rise there to  $N_{out}$  material points still evolving along new worldlines. If the worldlines are geodesics or more weakly, if all (ingoing and outgoing) worldlines satisfy  $\nabla_{\gamma'} \gamma' = 0$  exactly at<sup>3</sup>  $e$ , then the sum of the  $n$ -momenta entering  $e$  is equal to the sum of the  $n$ -momenta exiting  $e$ :

$$\sum_{i=1}^{N_{in}} P_{(i)e}^{(in)} = \sum_{i=1}^{N_{out}} P_{(i)e}^{(out)}. \quad (8.13)$$

This identity makes sense because both sides are vectors in the same tangent space  $T_e M^n$ .

**Remark 8.18.**

- (1) Considering a timelike worldline  $\gamma(s)$  and passing to the *proper time* parametrization, its  $n$ -momentum can be re-written in terms of the mass and the  $n$ -velocity of the particle as

$$P(\tau) = m(\tau) \dot{\gamma}(\tau). \quad (8.14)$$

This representation is not allowed if the worldline is lightlike.

- (2) Notice that the mass may depend on the value of the parameter  $s$ , but it is constant if the point evolves along a geodesic.

- (3) Since  $P$  is parallel to  $\gamma'$ , we conclude that  $m(s) = 0$  if and only if  $\gamma'(s)$  is lightlike.

- (4) As in Special Relativity [Mor20], we can define the four momentum of a *photon* (a particle of light)  $P$  in a semiclassical (non-quantum) view. Referring to a pseudo orthonormal basis on an event of an observer represented by a worldline  $\gamma$  whose timelike unit vector  $e_0$  is parallel to  $\gamma'$ ,

$$P^0 = \hbar \frac{\omega}{c}, \quad P^\alpha = \hbar k^\alpha, \quad \alpha = 1, 2, 3,$$

where  $\omega$  is the pulsation of the light wave associated with the photon measured with respect to the proper time of the observer and  $k^\alpha$  the components of the wave vector measured in

---

<sup>3</sup>In this case, for each incoming worldline  $\gamma'$  is the tangent vector computed as right derivative at the final endpoint of its interval of definition, for the outgoing worldlines  $\gamma'$  is the tangent vector computed as left derivative at the initial endpoint of the interval of definition.

the rest space of the considered observer. Furthermore,  $\hbar = \frac{h}{2\pi}$  with  $h$  the Planck constant ( $6.626 \times 10^{-27} \text{ erg sec}$ ). Since

$$\sum_{\alpha=1}^3 (k^\alpha)^2 = \left(\frac{\omega}{c}\right)^2$$

from the electromagnetic theory, we have in this case  $P_a P^a = 0$ , so that *photons must be massless*. ■

A further step is to export some laws from Special to General Relativity concerning extended continuous systems.

As we know, in Special Relativity, the content of energy and momentum of a continuous system are encapsulated in a  $(2, 0)$  symmetric tensor field  $T_{ab}$  called the *stress energy tensor* of the system (see Section (85 of [Mor20])). That tensor field, if the system is isolated, satisfies a local law of the form, valid at each event  $p \in \mathbb{M}^4$ ,

$$\nabla_a T^{ab} = 0. \quad (8.15)$$

On the ground of the Strong Equivalence Principle we are committed to assume first of all that such a tensor field does exist also in General Relativity for extended systems. Sometime its specific form can be easily exported from Special Relativity to General Relativity since we can once more take advantage of the Strong Equivalence Principle. That is the case for the *ideal fluid* (Section 8.5.5 of [Mor20])

$$T^{ab} = \mu_0 V^a V^b + \rho \left( g^{ab} + \frac{V^a V^b}{c^2} \right), \quad (8.16)$$

where  $V$  is the field of 4-velocities of the fluid,  $\rho$  the pressure, and  $\mu_0$  the density of mass measured at rest with a particle of fluid. All those quantities are evaluated at an event  $p$ . The same result is valid for the stress energy tensor of the electromagnetic fields

$$T^{ab} = F^{ac} F^b{}_c - \frac{1}{4} g^{ab} F_{cd} F^{cd}. \quad (8.17)$$

where the electromagnetic tensor  $F_{ab}$  is defined as

$$[F^{ab}]_{a,b=0,1,2,3} = \left[ \begin{array}{c|ccc} 0 & E_x/c & E_y/c & E_z/c \\ \hline -E_x/c & 0 & B_z/c & -B_y/c \\ -E_y/c & -B_z/c & 0 & B_x/c \\ -E_z/c & B_y/c & -B_x/c & 0 \end{array} \right]. \quad (8.18)$$

Above,  $E^\alpha$  and  $B^\beta$  are the components of the electric and magnetic field at  $p$  in normal coordinates centered on  $p$  (with pseudo orthonormal axes).

Also the *Maxwell equations* can be directly exported to General Relativity from Special Relativity if written in terms of the electromagnetic tensor  $F^{ab}$

$$\nabla_a F^{ab} = -J^b, \quad \epsilon^{abcd} \nabla_b F_{cd} = 0,$$

where  $J$  is four-current density  $J^a = \rho_0 V^a$ , where  $\rho_0$  is the charge density computed at rest with the particles of a continuous charged body.

**Remark 8.19.** We stress that one should make use of the Strong Equivalence Principle *cum grano salis*. After all, that is mathematics whereas the last word is always of physics. It may happen that passing from flat to curved spacetime further terms show up in the formulation of physical laws. The Strong Equivalence Principle seems to work because it is usually applied to spacetimes which are very close to Minkowski spacetime, to assert that this regime is universally valid also in the presence of robust gravitational phenomena (i.e., when the metric is intrinsically different from Minkowski one) is a pure gamble. ■

### 8.3.2 Conservation laws and Killing fields in General Relativity

Coming back to the stress energy tensor, since (8.15) respects the hypotheses of the Strong Equivalence Principle, we should assume that it is also valid in General Relativity when the physical system does not interact with any other external system in turn described with its own stress energy tensor. That is exactly what happens in Special Relativity (Section 8.5.4 of [Mor20]). The system cannot be completely isolated however: gravitational interaction cannot be removed since it must present in some sense in the metric  $\mathbf{g}$  as soon as the spacetime is not flat.

A potential problem shows up with the interpretation of identity (8.15). In Special Relativity (8.15) is responsible for the conservation law of energy and momentum of the extended continuous system as discussed in Sections 8.5.3 and 8.5.4 of [Mor20]. The conserved quantities in  $\mathbb{M}^4$  are integrals of components  $T^{0a}$  over the 3-space at rest with an inertial system, where  $x^0$  is the temporal Minkowskian coordinate orthogonal to these 3-surfaces, and (8.15) is equivalent to the fact that those quantities are constant in time. In this sense (8.15) is interpreted as the law of conservation of momentum and energy in local form.

In General Relativity no extended inertial frames exist – just because the manifold is not flat so that no Minkowskian coordinates are at disposal – and the interpretation of (8.15) as a local version of a conservation law is much more delicate and, generally speaking, *definitely false*. In spite of this difficulty, (8.15) is commonly called the *conservation law of the stress energy tensor*.

From the physical side, we actually must expect that, in general, (8.15) does not correspond to a conservation law, though we can assume that it is valid using the Strong Equivalence Principle. This is because, as already stressed, it is impossible to completely isolate the system: we cannot remove the interaction with the gravitation field which is now described (in a way we still have to clarify) by the metric, at least when the spacetime is not (locally) flat. On the other hand, the gravitational interaction is embodied in the general formalism in a way that it does not permit to define a stress energy tensor of the gravitational field (though several very interesting attempts exist in suitable classes of spacetimes [LaLi80]).

To discuss how conservation laws are formalized in General Relativity and how they are related with identity (8.15), we start by treating the elementary case of a current  $J$ , a  $(1,0)$

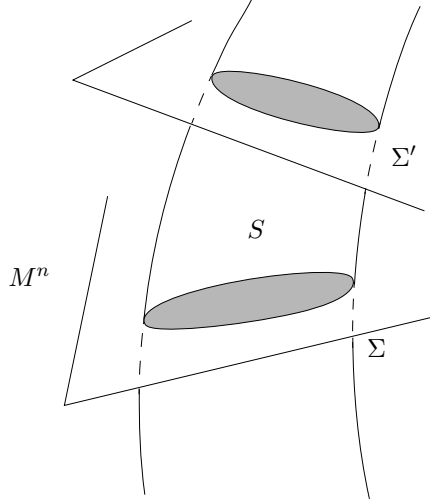


Figure 8.2: Conservation of  $Q$

smooth vector field satisfying a local conservation rule

$$\nabla_a J^a = 0 \quad (8.19)$$

everywhere in  $M^n$ . As an example,  $J$  can be the electric charge 4-current in  $M^4$ . Taking advantage of the divergence theorem (Theorem 6.15), if the support of  $J$  is confined in a tube  $S$  with timelike lateral surface (see figure 7.28) which intersect the relevant *spatial sections*  $\Sigma$  in compact sets, we can define

$$Q_\Sigma := \int_\Sigma \langle J, n \rangle d\mu(\mathbf{g}^{(\Sigma)}), \quad (8.20)$$

on a spacelike embedded submanifold  $\Sigma$  of co-dimension 1 and where  $n$  is future-oriented. This quantity, in view of the divergence theorem, remains constant if computed over another similar spacelike embedded submanifold  $\Sigma'$  when (8.19) is valid. In summary we find a *global conservation law*:

$$Q_\Sigma = Q_{\Sigma'}. \quad (8.21)$$

The idea is to interpret  $\Sigma$  and  $\Sigma'$  as the extended rest spaces of some reference frame. Or also of *two different* reference frames, proving that not only  $Q_\Sigma$  is conserved but that it is also independent of the reference frame (as it happens for the electric charge in particular).

We shall discuss this interpretation of  $\Sigma$  in Section 8.5.1, when introducing the notion of (extended) reference frame in General Relativity and for the moment we are just content with this opportunity.

The hypothesis that the support of  $J$  is confined in a spatially bounded<sup>4</sup> tube  $S$  can be relaxed assuming a suitable rapid decay of  $J$  on  $\Sigma$ .

<sup>4</sup>However also the strict requirement of compactness  $S \cap \Sigma$  can be required when  $J$  is a quantity associated to a solution of a hyperbolic equation in  $M^n$  just by imposing compactly supported initial conditions on spatial sections (Cauchy surfaces).

Coming back to (8.15) we prove that, *in the presence of a Killing field  $K$* , a conserved quantity can be defined which depends on both the stress energy tensor and the Killing field. This quantity is simply defined out of the current

$$J_K^a := K_b T^{ba} . \quad (8.22)$$

Indeed, if  $T^{ab}$  satisfies (8.15), in view of the Killing equation (6.11) and  $T^{ab} = T^{ba}$

$$\nabla_a J_K^a = \nabla_a (K_b T^{ba}) = (\nabla_a K_b) T^{ba} + K^a \nabla_a T^{ba} = (\nabla_a K_b) T^{ba} ,$$

but also the apparently surviving term vanishes, as

$$(\nabla_a K_b) T^{ba} = \frac{1}{2} (\nabla_a K_b) T^{ba} + \frac{1}{2} (\nabla_b K_a) T^{ab} = \frac{1}{2} (\nabla_a K_b) T^{ba} + \frac{1}{2} (\nabla_b K_a) T^{ba} = \frac{1}{2} (\nabla_a K_b + \nabla_b K_a) T^{ba} = 0 .$$

If we define

$$Q_\Sigma^{(K)} := \int_\Sigma \langle J_K, n \rangle d\mu^{(\mathbf{g}^{(\Sigma)})} , \quad (8.23)$$

then (8.21) is valid so that we can associate the pair made of a stress energy tensor and a Killing vector with a conserved quantity.

Usually, the conserved quantity  $Q_\Sigma^{(K)}$  is interpreted as a notion of *energy* when  $K$  is timelike, an energy *strictly referred to the notion of time appearing in the flow generated by  $K$  also known as Killing time*.

In Minkowski spacetime, every vector tangent to a coordinate of a Minkowskian reference system is a Killing vector. It is easy to see that using those vectors, the definition above of conserved quantities is in agreement with the standard one as discussed in Sections 8.5.3 and 8.5.4 of [Mor20].

**Remark 8.20.**

(1) The found result is once more a manifestation of the deep relation between dynamically conserved quantities and symmetries as it happens with Noether's theorem. In fact,  $K$  is a Killing vector, so that it reflects the existence of a symmetry of the metric expressed by the Lie version of Killing equation (5.7),

$$\mathcal{L}_K \mathbf{g} = 0 ,$$

which, in turn, means that the metric is invariant under the action of the flow generated by  $K$ . The difference with the standard formulation of Noether's theorem is that here the symmetry is of the background and not of the physical system associated to  $T^{ab}$ .

(2) The general interpretation of  $Q^{(K)}$  depends of the nature of the Lie group of continuous symmetries of  $(M, \mathbf{g})$  whose the Lie algebra of Killing vector represent the Lie algebra. For instance if that group includes a subgroup isomorphic to  $SO(3)$ , it is natural to interpret the three conserved quantities associated to the 3 Killing vectors generators of  $SO(3)$  as components of the *angular momentum* of the system, and so on. This abstract approach leads in Minkowski spacetime to the standard physical interpretation of the various conserved quantities. ■

As a special case of the discussion above, we can consider the case of a particle in a spacetime with a Killing vector  $K$  and suppose that the point is isolated from external interactions (barring the gravitational one). In this case, from (8.11) and the Killing equation, we have that, if we define,

$$Q(s) := \mathbf{g}(K_{\gamma(s)}, \gamma'(s)) , \quad (8.24)$$

then  $Q(s)$  is constant along the worldline

$$\frac{dQ}{ds} = 0 , \quad (8.25)$$

because

$$\frac{d}{ds} = \mathbf{g}(\nabla_{\gamma'} K_{\gamma(s)}, \gamma'(s)) + \mathbf{g}(K_{\gamma(s)}, \nabla_{\gamma'} \gamma'(s)) = \mathbf{g}(\nabla_{\gamma'} K_{\gamma(s)}, \gamma'(s)) ,$$

but the apparently surviving terms also vanishes since, from (6.11),

$$\mathbf{g}(\nabla_{\gamma'} K_{\gamma(s)}, \gamma'(s)) = \gamma'^a \gamma'^b \nabla_a K_b = \frac{1}{2}(\gamma'^a \gamma'^b \nabla_a K_b + \gamma'^b \gamma'^a \nabla_b K_a) = \frac{1}{2} \gamma'^a \gamma'^b (\nabla_a K_b + \nabla_b K_a) = 0 .$$

This is not the whole story because we can consider the case where many particles are involved establishing again a conservation law. If  $N_{in}$  material points evolving along geodesics in a neighborhood of  $e$  where collide and give rise there to  $N_{out}$  material points still evolving along geodesics at least close to  $e$ , then the sum of the quantities  $Q$  of the particles entering  $e$  is equal to the sum of the quantities  $Q$  of the particles exiting  $e$ :

$$\sum_{i=1}^{N_{in}} Q_{(i)}^{(in)}(e) = \sum_{i=1}^{N_{out}} Q_{(i)}^{(out)}(e) . \quad (8.26)$$

The proof of the identity above is a trivial consequence of (8.13) just taking the scalar product with  $K_e$  on both sides.

### 8.3.3 Energy extraction from rotating black holes

There exists an interesting theoretical phenomenon discovered by Penrose known as the *energy extraction process* from a rotating black hole. We do not want to deal here with the notion of rotating black hole (also known as Kerr's black hole). The only important piece of information is that the rotating black hole is a spacetime  $(M^4, \mathbf{g})$  which admits a Killing vector field  $K$  with special properties. In a spatial section of the spacetime, far from the so-called black hole region surrounded by the *event horizon*, the spacetime becomes flat and the metric tends to become the Minkowskian one. Similarly, far from the black hole region,  $K$  tends to become a standard Killing vector tangent to the temporal coordinate of a Minkowskian coordinate system<sup>5</sup>. Going along the opposite direction, before crossing the event horizon another spatial region exists,

---

<sup>5</sup>This Minkowskian reference frame is completely determined by the black hole and by  $K$ : it is the asymptotic reference frame where the spatial black hole region is viewed to rotate around an axis.



called the *ergosphere* which surrounds the black hole region, therein  $K$  becomes spacelike. This region exist just in view of the rotation of the black hole.

Consider a particle that starting from the Minkowskian region with an energy  $E$  with respect to the asymptotic Minkowskian reference frame, falls into the ergosphere, there it breaks into two particles and one of the two particles comes back, still along a geodesic, to the Minkowskian region. We can define  $E = -Q(s) = \mathbf{g}(K, P)$  according to (8.24) using the Killing vector  $K$ . The sign is chosen in order to have a positive energy  $E$  where both  $P$  and  $K$  are timelike and future directed as it happens in the asymptotic Minkowskian region. There  $E$  becomes the standard energy of the particle. The conservation law (8.26) of  $Q$  applied to the event where the initial particle breaks leads to the identity

$$-K_b P_{(1)}^b = -K_b P^b + K_b P_{(2)}^b \quad \text{i.e.} \quad E_1 = E - E_2 .$$

Furthermore, the values of  $E, E_1, E_2$  are also constant along the respective worldlines since they are geodesics and we can also apply (8.25). If  $K$  and  $P, P_{(1)}, P_{(2)}$  are causal and future-oriented, the energies  $E, E_1, E_2$  are all positive so that, in particular  $0 < E_2, E_1 \leq E$ . However, the identity above is still valid when  $K$  is not timelike, in this case  $-K_b P^b$  is conserved but it has not the meaning of energy and its sign can be arbitrary. As said the initial particle breaks just inside the ergosphere of a Kerr black hole where  $K$  is spacelike. Suppose also that part 2 remains inside the ergosphere whereas, as said, part 1 comes out and reaches the initial asymptotic Minkowskian travelling along a geodesic. In this case  $E_1 \geq 0$ , because the momentum of the particle future oriented as  $K$  is. However it is now permitted that  $E_2 < 0$ , because  $K$  is spacelike in the ergosphere even if  $P_{(2)}$  is still timelike and future directed therein. As  $E = E_1 + E_2$  it must be

$$E_1 > E > 0 .$$

$E$  and  $E_1$  are the energies the particles have also when they, respectively, leave and reach the asymptotic Minkowskian reference frame. As a matter of fact the asymptotic Minkowskian observer who launched the initial particle and receives the final one extracts energy from the black hole, more precisely from the ergosphere.

## 8.4 Newtonian correspondence and some consequence

Up to now, we have assumed that  $\mathbf{g}$  is somehow related with the gravitational interaction, but we did not yet investigate this fact. This is the goal of the next section, where shall grasp some pieces of information. We shall see that, when the spacetime is close to Minkowski spacetime and the used coordinates are well approximated by Minkowskian coordinates, then  $g_{00}$  turns out to be related to the classical gravitational potential. Later, we pass to discuss two immediate entangled phenomenological consequences of this result.

### 8.4.1 Newtonian correspondence

Let us consider a spacetime  $(M^4, \mathbf{g})$  of General Relativity and assume that there exist a chart  $\phi : U \ni p \mapsto (x^0, x^1, x^2, x^3) \in \mathbb{R}^4$  where we define  $t := x^0/c$  and the components of the metric

$\mathbf{g} = g_{ab}dx^a \otimes dx^b$  are very close to those of Special Relativity in Minkowskian coordinates. We stress that this is generally false in a generic spacetime over an extended region, but it is evidently valid in our universe in a spatial region which includes the Solar system, where also Newtonian physics works quite well. In this background, we study the equation of motion of the free-falling bodies, i.e., of the geodesics, for massive material points in order to compare this motion with the one predicted by the classical Newtonian theory of gravitation. In this intermediate regime, we expect to find some correspondence between classical and general relativistic notions.

Here is the list of our assumptions and approximations.

- (1) We assume that, with the following decomposition,

$$g_{ab} = \eta_{ab} + h_{ab}, \quad (8.27)$$

where  $\eta_{ab}$  is the standard Minkowski metric

$$[\eta_{ab}]_{a,b=0,1,2,3} = \text{diag}(-1, 1, 1, 1), \quad (8.28)$$

at each event of the chart the estimate is true

$$|h_{ab}| \ll 1. \quad (8.29)$$

As a consequence, we shall consider negligible the contribution of such terms in some passages below.

- (2) As an approximation of different nature, we also suppose that there is a timelike Killing field defined in  $U$  coinciding with  $\frac{\partial}{\partial t}$ . The Killing equation  $\mathcal{L}_K \mathbf{g} = 0$ , writing the components of  $\mathbf{g}$  as functions of the coordinates  $ct, x^1, x^2, x^3$ , immediately becomes

$$\frac{\partial g_{ab}}{\partial t} = 0, \quad (8.30)$$

so that we are actually supposing that the metric does not depend on the (Killing) time, as it happens in every Minkowskian reference frame in Special Relativity. This hypothesis could be relaxed making more precise the dependence on time of the metric components, but we stick to the basic case.

- (3) We shall suppose that the velocity of the material points, roughly defined as  $v^\alpha = \frac{dx^\alpha}{d\tau}$  ( $\alpha = 1, 2, 3$ ), has very small magnitude with respect to the speed of light  $c$ :

$$\left| \frac{dx^\alpha}{d\tau} \right| \ll c. \quad (8.31)$$

Consequently, we shall consider negligible some addends in the following formulas when they are multiplied with inverse powers of  $c$ .

This picture is a way to (at least locally) embody the Newtonian universe in the spacetime of General Relativity.

Let us pass to exploit these approximations in the geodesic equation for a timelike geodesic  $\gamma$  parametrized with its proper time  $\tau$ . First of all we get rid of  $\tau$  in favour of  $t$  observing that

$$-c^2 = \mathbf{g}(\dot{\gamma}, \dot{\gamma}),$$

is explicitly written

$$-1 = (-1 + h_{00}) \left( \frac{dt}{d\tau} \right)^2 + \frac{2}{c} \sum_{\alpha=1}^3 h_{0\alpha} \frac{dt}{d\tau} \frac{dx^\alpha}{d\tau} + \frac{1}{c^2} \sum_{\alpha,\beta=1}^3 (\delta_{\alpha\beta} + h_{\alpha\beta}) \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau}.$$

Dropping all terms in the right-hand side with factors  $1/c$  and  $1/c^2$  and also neglecting  $h_{00}$  with respect to 1 in view of (8.29), we conclude that

$$1 = \frac{dt}{d\tau}$$

is a good approximation in our context so that we can assume  $t = \tau$ . With this approximation, the geodesic equation reads (recall that  $x^0 = ct$ )

$$\frac{d^2 x^a}{dt^2} = -\Gamma_{bc}^a \frac{dx^b}{dt} \frac{dx^c}{dt} = -c^2 \left[ \Gamma_{00}^a \left( \frac{dt}{dt} \right)^2 + \frac{2}{c} \Gamma_{b0}^a \frac{dx^b}{dt} \frac{dt}{dt} + \frac{1}{c^2} \sum_{\alpha,\beta=1}^3 \Gamma_{\alpha\beta}^a \frac{dx^\alpha}{dt} \frac{dx^\beta}{dt} \right].$$

Dropping the terms of order  $1/c$  and  $1/c^2$  due to (8.31), we find

$$\frac{d^2 x^a}{dt^2} = -c^2 \Gamma_{00}^a.$$

Let us pass to expand the expression of  $\Gamma_{00}^a$ . Dropping all  $x^0$ -derivatives in view of (8.30), we find

$$\Gamma_{00}^a = \frac{g^{ab}}{2} \left( \frac{\partial g_{0b}}{\partial x^0} + \frac{\partial g_{b0}}{\partial x^0} - \frac{\partial g_{00}}{\partial x^b} \right) = -\frac{g^{ab}}{2} \frac{\partial g_{00}}{\partial x^b} = -\frac{g^{ab}}{2} \frac{\partial h_{00}}{\partial x^b},$$

so that the initial geodesic equation reduces to

$$g_{ab} \frac{d^2 x^b}{dt^2} = \frac{c^2}{2} \frac{\partial h_{00}}{\partial x^a}.$$

Namely

$$\sum_{\beta=1}^3 (\eta_{a\beta} + h_{a\beta}) \frac{d^2 x^\beta}{dt^2} = \frac{c^2}{2} \frac{\partial h_{00}}{\partial x^a}.$$

where we have omitted the contribution of  $b = 0$  since  $\frac{d^2 x^0}{dt^2} = 0$ . For  $a = \alpha = 1, 2, 3$ , we find

$$\frac{d^2 x^\alpha}{dt^2} + \sum_{\beta=1}^3 h_{\alpha\beta} \frac{d^2 x^\beta}{dt^2} = \frac{c^2}{2} \frac{\partial h_{00}}{\partial x^\alpha}.$$

As a first approximation, assuming all components  $\frac{d^2 x^\gamma}{d\tau^2}$  of similar magnitude, we can drop the second addend in the left-hand side in view of (8.29). The final equation is

$$\frac{d^2 x^\alpha}{dt^2} = \frac{c^2}{2} \frac{\partial h_{00}}{\partial x^\alpha}, \quad \alpha = 1, 2, 3. \quad (8.32)$$

That is just *Newton's law of the motion* of a material point free falling in a *gravitational static potential*<sup>6</sup>

$$\varphi(\vec{x}) = \frac{c^2}{2} h_{00}(\vec{x}), \quad (8.33)$$

where  $\vec{x} = (x^1, x^2, x^3)$ . Actually  $\varphi$  is defined up to an additive constant. It is however reasonable to keep the identification (8.33) as it stands, without adding additive constants, if we assume that far from the source of the gravitational field where  $\varphi$  vanishes also the metric becomes Minkowskian.

**Remark 8.21.** Notice that we have also found that, within our approximations,

$$g_{00}(t, \vec{x}) = -1 + \frac{2}{c^2} \varphi(\vec{x}). \quad (8.34)$$

Within this geometric interpretation of gravitational Newtonian mechanics, the motion of Earth around the Sun is due to the geometry of the spacetime of General Relativity. The timelike geodesics are here proved to be strongly different from the ones of Minkowski spacetime, where they are standard  $\mathbb{R}^4$  segments in Minkowskian coordinates. To appreciate better this difference from an intrinsic point of view, we should introduce some further geometric notions concerning the curvature of a (pseudo) Riemannian manifold as we shall do later. ■

#### 8.4.2 Gravitational redshift

Equation (8.34) has an important consequence which can be experimentally tested and, actually, it has been tested successfully several times. Not only: nowadays this consequence is also part of the standard GPS technology.

Let us consider a region of spacetime where a timelike Killing vector  $K$  is defined and suppose that, as in the previous section, we are dealing with a coordinate system  $t, x^1, x^2, x^3$  where

$$K^a = \frac{\partial}{\partial t}.$$

The metric is stationary in this system of coordinate and all physical stationary or periodic phenomena are naturally referred to that the Killing time. For instance, equilibrium thermodynamics is referred to the Killing time and relevant notion of thermodynamic energy is associated to that notion of time.

Consider a pair of observers which evolve in time with worldlines  $\gamma_1$  and  $\gamma_2$  that are integral

---

<sup>6</sup> $u(\vec{x}) := -\varphi(\vec{x})$  is the *gravitational potential energy* as a consequence.

curves of  $K$  placed at different spatial positions  $\vec{x}_1 = (x_1^1, x_2^2, x_3^3)$  and  $\vec{x}_2 = (x_2^1, x_2^2, x_2^3)$ . These observers are “at rest” in an extended reference frame where all gravitational phenomena are stationary. The local time they measure is however referred to ideal clocks they carry and thus it is the *proper time*  $\tau$  rather than the *Killing time*  $t$ . The difference of the two notions of time is completely described by

$$d\tau_i = \sqrt{|g_{00}(\vec{x}_i)|} dt = \sqrt{1 - \frac{2}{c^2} \varphi(\vec{x}_i)} dt \quad i = 1, 2. \quad (8.35)$$

where, in the last identity, we assumed that both the previous Newtonian approximation can be exploited and the arbitrary additive constant of  $\varphi$  was chosen as discussed.

Suppose that, at rest with the observer whose worldline is  $\gamma_1$ , there is an electromagnetic source emitting waves with frequency  $f_1 = \frac{c}{T_1}$ . The period  $T_1$  is referred to the proper time  $\tau_1$ . However the wave propagates reaching  $\gamma_2$  and defines a common periodic process with respect to the Killing time: when writing Maxwell’s equations in coordinates  $t, x^1, x^2, x^3$  no explicit dependence on time  $t$  exists in any place and the only explicit dependence on  $t$  is in the source. The solution of those equations must therefore have the same  $t$ -period of the source. This period is

$$\int_0^T \frac{1}{\sqrt{|g_{00}(\vec{x}_1)|}} d\tau = \frac{T_1}{\sqrt{|g_{00}(\vec{x}_1)|}}$$

due to (8.35). In summary,

$$\frac{T_1}{\sqrt{|g_{00}(\vec{x}_1)|}} = \frac{T_2}{\sqrt{|g_{00}(\vec{x}_2)|}},$$

so that

$$f_2 = f_1 \frac{\sqrt{|g_{00}(\vec{x}_1)|}}{\sqrt{|g_{00}(\vec{x}_2)|}}.$$

Within the Newtonian approximation,

$$f_2 = f_1 \sqrt{\frac{1 - \frac{2}{c^2} \varphi(\vec{x}_1)}{1 - \frac{2}{c^2} \varphi(\vec{x}_2)}}. \quad (8.36)$$

If  $\vec{x}_2$  is farther from the source of the gravitational force than the other observer, the denominator tends to 1 whereas the denominator remains finite and

$$f_2 < f_1.$$

This phenomenon is known as the *gravitational redshift*: the frequency detected in a weaker gravitational field of a source placed in a stronger gravitational field shifts towards the red direction in the spectrum. This phenomenon is not only well known, but modern GPS technology *take it into account* in the communications between Earth surface and satellites.

The existence of the gravitational redshift was first confirmed directly by the celebrated experiment by *Pound and Rebka* <sup>7</sup> in 1959. The result confirmed that the predictions of General Relativity with an accuracy of 10%. In 1980 another test<sup>8</sup>, which exploited a *maser*, strongly increased the accuracy of the measurement to about 0.01 %.

### 8.4.3 Gravitational time dilation

A related phenomenon (well illustrated in the 2014 movie *Interstellar*) is that of *time dilation* due to the gravitational field. Consider two twin brothers who initially stay at  $\vec{x}_1$  and next one of them quickly moves to  $\vec{x}_2$  where  $|\varphi(\vec{x}_2)| \gg |\varphi(\vec{x}_1)|$  and he spends there a long period of proper time

$$\Delta\tau_2 = \int_s^{s'} \sqrt{1 - \frac{2}{c^2}\varphi(\vec{x}_2)} dt = \sqrt{1 - \frac{2}{c^2}\varphi(\vec{x}_2)} \Delta t ,$$

much longer than the time spent to travel to and from  $\vec{x}_2$ . When he goes back to  $\vec{x}_1$ , he discovers that the age of his brother is increased of (disregarding the time of the trips)

$$\Delta\tau_1 = \int_s^{s'} \sqrt{1 - \frac{2}{c^2}\varphi(\vec{x}_i)} dt = \sqrt{1 - \frac{2}{c^2}\varphi(\vec{x}_1)} \Delta t$$

so that

$$\Delta\tau_1 = \sqrt{\frac{1 - \frac{2}{c^2}\varphi(\vec{x}_1)}{1 - \frac{2}{c^2}\varphi(\vec{x}_2)}} \Delta\tau_2 > \Delta\tau_2 .$$

**Remark 8.22.** The qualitative result would not change if taking the time necessary to travel into account, but the formula would not be so easy as it should include the details of the trip as velocities etc. ■

The existence of the gravitational time dilation was first established by the celebrated experiment by *Hafele and Keating* in 1971 using four cesium-beam atomic clocks aboard commercial airliners. The experiment also confirmed the time dilation predicted by Special Relativity. In 2010, Chou and collaborators<sup>9</sup> performed tests in which both gravitational and Special relativistic effects were tested. It was possible to confirm the gravitational time dilation phenomenon from a difference in elevation between two clocks of only 33 cm.

<sup>7</sup>Pound, R. V.; Rebka Jr. G. A. (November 1, 1959). *Gravitational Red-Shift in Nuclear Resonance*. Physical Review Letters. 3 (9): 439–441.

<sup>8</sup>Vessot, R. F. C.; Levine, M. W.; Mattison, E. M.; Blomberg, E. L.; Hoffman, T. E.; Nystrom, G. U.; Farrel, R.; Decher, P. B.; Eby, C. R.; Baugher, J. W.; Watts, D. L.; Teuber, F. D.; Wills (December 29, 1980). *Test of Relativistic Gravitation with a Space-Borne Hydrogen Maser*. Physical Review Letters. 45 (26): 2081–2084

<sup>9</sup>Chou, C. W.; Hume, D. B.; Rosenband, T.; Wineland, D. J. (2010). *Optical Clocks and Relativity*. Science. 329 (5999): 1630–1633.

## 8.5 Extended reference frames in General Relativity and related notions

This section is devoted to introduce the notion of *extended reference frame* in General Relativity and to discuss some related concepts. It is clear that the idea of reference frame in General Relativity has a more delicate status than in Special Relativity in view of the absence of inertial reference frames. Actually, inertial reference frames still exist locally in the sense of normal coordinate systems around a timelike geodesic. However this notion is a bit fuzzy also in view of the mixing of gravitational interaction and inertia arising from the equivalence principle. Furthermore it does not provide a canonical notion of rest space, but just an approximated rest space described in the subspace orthogonal to the geodesic in the tangent space of any event crossed by it. We are instead interested in a finitely extended structure obtained by collecting in same way a number of observers described by worldlines.

### 8.5.1 A general notion of extended reference frame

Let us consider a spacetime  $(M^n, \mathbf{g})$  which may be also a sub region of a larger spacetime. An extended reference frame is a mathematical machinery used to one-to-one associate a space position and a time location to every event of  $M$ . From a very general point of view we can proceed as follow.

- (a) *Spatial position of an event.* A very general procedure is to identify the spatial positions with the maximal integral curves of a future-oriented timelike smooth vector field  $T$ . Since  $T$  vanishes nowhere, these curves cannot intersect and there is one of them passing through any event of the spacetime. The space position of  $e \in M^n$  is exactly the unique worldline  $\gamma_e$  integral curve of  $T$  crossing  $e$ .
- (b) *Time location of an event.* A very general procedure is to identify the time locations with the values of a surjective smooth function  $t : M^n \rightarrow I$ , for an interval  $I \subset \mathbb{R}$ , satisfying natural requirements. First of all, *simultaneous events* should give rise to a suitable notion of *rest space* of the reference frame at a given time. To this end, we assume that  $dt$  is everywhere timelike and consider the family of the  $n - 1$ -surfaces

$$\{\Sigma_s\}_{s \in I} \quad \text{where} \quad \Sigma_s := \{p \in M^n \mid t(p) = s \in I\} \quad (8.37)$$

Since  $dt$  cannot vanish, Theorem 4.13 implies that every  $\Sigma_s$  is an embedded submanifold of co-dimension 1. By construction it is also spacelike. Since  $t$  is defined everywhere on  $M^n$ , the union of the surfaces  $\Sigma_s$  is the whole  $M^n$ . Observe that these surfaces are also pairwise disjoint by definition and they are one-to-one with the values of  $t$ . The time location of  $e \in M^n$  is the unique surface of the family that  $\Sigma_{t(e)} \ni e$ .

There is still a pair of issues to fix in this framework.

- (i) Suppose that both  $T$  and  $t$  are given as above on  $(M^n, \mathbf{g})$ . Let  $\Sigma_{t_0}$  be one of the  $n - 1$ -spacelike surfaces associated to  $t$ . For every  $p \in \Sigma_{t_0}$  there is exactly one integral curve  $\gamma_p$  of  $T$

passing through  $p$ . However it is not necessarily true the converse: it is not obvious that every  $\Sigma_{t_0}$  meets every integral curve of  $T$ . It seems instead natural to impose it to have a “number of space positions” constant in time. In this case, all the integral curves of  $T$  are one-to-one defined with the points of every fixed surface  $\Sigma_t$ .

(ii) Consider every  $p \in \Sigma_{t_0}$  and the unique integral line  $\gamma_p$  of  $T$  propagated from  $p$  and arrange the origin of the Killing time in order that  $\gamma_p(t_0) = p$  for every  $p \in \Sigma_{t_0}$ . With this choice,  $\Sigma_{t_0}$  evolves according to the Killing time, defining other surfaces

$$\Sigma_{t,t_0} := \{q \in M^n \mid q = \gamma_p(t), q \in \Sigma_{t_0}\}.$$

To avoid the existence of two different notions of rest space in the same reference frame, we impose as our last condition that

$$\Sigma_{t,t_0} = \Sigma_t \quad \forall t \in I.$$

This requirement will be actually stated as a condition valid for  $T$  and  $t$ :

$$\langle T, dt \rangle = 1 \quad \text{everywhere in } M^n. \quad (8.38)$$

Indeed, under this hypothesis

$$\frac{dt(\gamma(s))}{ds} = \langle \gamma'(s), dt \rangle = \langle T, dt \rangle = 1$$

that implies that the lenght of an interval of time measured along the integral curves of  $T$  in terms of the Killing time coincides with the lenght of the corresponding interval of time measured in terms of the time function  $t$ . In other words  $\gamma_p(t) \in \Sigma_t$  for every  $t \in I$  if we have arranged the origin of the Killing time in order that  $\gamma(t_0) \in \Sigma_{t_0}$ .

We are in a position to state our definition.

**Definition 8.23.** (**Extended reference frame.**) In a spacetime  $(M^n, \mathbf{g})$ , an **extended refence frame** (or **system**) over a region  $N \subset M^n$  is defined by  $N$  and a pair  $(T, t)$ , where

- (a)  $T$  is a smooth future-oriented timelike vector field defined on  $N$ ,
- (b)  $t : N \rightarrow I$  – where  $I \subset \mathbb{R}$  is an open interval – is a surjective smooth map with  $dt$  everywhere timelike.

The following requirements are also assumed,

- (1) every surface  $\Sigma_s$  defined in (8.37) meets all maximal integral line of  $T$ ,
- (2) requirement (8.38) is valid and we consequently assume that the origin of the parameter of every maximal integral curve of  $T$  is arranged to satisfy  $\gamma(t) \in \Sigma_t$  if  $t \in I$ .



$T$  is called **time vector** of the reference frame,  $t$  is called **time coordinate** of the reference frame, and a submanifold  $\Sigma_t$  is called **rest space** at time  $t$  of the reference frame. ■

Given a reference frame  $(N, T, t)$ , we can define a local chart **adapted** to the reference frame, equivalently called **co-moving** with it. It is a local chart  $\phi : U \ni p \mapsto (x^0(p), \dots, x^{n-1}(p)) \in \mathbb{R}^n$  such that, occasionally enumerating the coordinates from 0 to  $n - 1$ ,

- (1)  $U \subset N$ ,
- (2)  $\frac{\partial}{\partial x^0} = \frac{1}{c}T$  on  $U$ ,
- (2)  $x^0 = ct$  restricted to  $U$ , so that the remaining coordinates  $x^1, \dots, x^{n-1}$  define local coordinates on  $\Sigma_t$  for every fixed value of  $x^0 = t$ .

With this definition,

$$g_{00} = \mathbf{g} \left( \frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^0} \right) < 0,$$

furthermore,

$$g^{00} = \mathbf{g}(dt^\sharp, dt^\sharp) < 0,$$

and the Riemannian metric  $\mathbf{g}^{(\Sigma_t)}$  has components

$$g_{\alpha\beta} \quad \text{with } \alpha, \beta = 1, \dots, n-1, \text{ if } \quad \mathbf{g} = g_{ij}dx^i \otimes dx^j.$$

It is not difficult to see that if  $\phi : U \ni p \mapsto (x^0(p), \dots, x^{n-1}(p)) \in \mathbb{R}^n$  and  $\psi : V \ni p \mapsto (y^0(p), \dots, y^{n-1}(p)) \in \mathbb{R}^n$  are two local charts adapted to  $(N, T, t)$  and  $U \cap V \neq \emptyset$  then  $\psi \circ \phi^{-1} : \phi(U \cap V) \rightarrow \psi(U \cap V)$  has the coordinate expression

$$y^0 = x^0, \quad y^\alpha = y^\alpha(x^1, \dots, x^{n-1}), \quad \alpha = 1, \dots, n-1. \quad (8.39)$$

Let us prove that extend reference frames exist.

**Proposition 8.24.** *Let  $(M^n, \mathbf{g})$  be a spacetime. If  $p \in M^n$  and  $S_p \in T_p M$  is timelike and future directed, there is an extend reference frame  $(N, T, t)$  with  $p \in N$  and  $T_p = S_p$ .*

**Proof.** Extend  $S_p$  to a smooth vector field  $S$  which is therefore necessarily timelike and future-oriented around  $p$ . Fix a pseudo orthonormal basis  $e_0, e_2, \dots, e_{n-1}$  of  $T_p M$  with  $e_1$  parallel to  $S_p$ . Using  $\exp_p$ , we can extend the span of the spacelike  $n-1$  vectors  $e_1, \dots, e_{n-1}$  to an embedded  $n-1$ -submanifold  $\Sigma$  passing through  $p$ . Notice that the co-normal vector to  $\Sigma$  at  $p$  is  $\mathbf{g}(e_0, \cdot)$  which is timelike by construction so that, in a neighborhood  $N$  of  $p$ ,  $\Sigma$  remains spacelike and we restrict all the discussion to that neighborhood. As  $\Sigma$  is an embedded submanifold, restricting  $N$  if necessary, we can define a function  $t' : N \rightarrow \mathbb{R}$  which vanishes exactly on that surface and that  $dt$  is timelike in its domain (it is sufficient to use the coordinate  $t' = x^0$  where  $x^0, \dots, x^{n-1}$  are normal coordinates centered on  $p$ ). By construction  $f := \langle S, dt' \rangle$  cannot vanish because both vectors in the bracket are timelike. We can always assume that the sign of  $f$  is positive (possibly

changing the sign of the initial function  $t'$ ). The vector field  $T_q := f(q)^{-1}f(p)S_q$  satisfies (8.38) with respect to  $t(q) := f(p)^{-1}t'(q)$  and  $T_p = S_p$ . Finally, re-defining  $N := \{\Phi_s^{(T)}(\Sigma) \mid s \in I\}$  restricting the interval  $I \ni s$  and  $\Sigma$  if necessary, also the condition (1) turns out to be satisfied in view of the properties of the local flow  $\Phi^{(T)}$ . ■

**Remark 8.25.**

(1) It is clear that the illustrated definition of reference frame is an extension of the notion of Minkowskian reference frame  $\mathcal{F}$  in Special Relativity (see [Mor20]) when  $N = \mathbb{M}^4$ . In that case  $T = \mathcal{F} = \frac{\partial}{\partial t}$  and  $t = x^0/c$ , where  $x^0, x^1, x^2, x^3$  are a Minkowskian coordinate system co-moving with  $\mathcal{F}$ . Similarly, Minkowskian coordinates co-moving with an inertial reference frame [Mor20] in Special Relativity are a particular case of co-moving coordinates with an extended reference frame.

(2) Notice that  $\langle T, dt \rangle = 1$  implies that the contravariant form  $dt^\sharp$  of  $dt$  is *past-directed* since  $T$  is future-directed and  $\mathbf{g}(T, dt^\sharp) = \langle T, dt \rangle = 1 > 0$ .

(3) Given an extended reference frame we can decompose

$$T = ldt^\sharp + S.$$

The smooth function  $l$  is called **lapse function** and the smooth vector field  $S$  is called **shift vector field**. This decomposition, more often written in terms of components,

$$T^a = N(dt)^a + N^a,$$

plays an important role in the so called *ADM formalism* [Wal84] to tackle the problem of solving Einstein's equations of gravitation we shall introduce later.

(4) The rest space  $\Sigma_t$  is the standard candidate where to define the conserved quantities as in (8.20) and (8.23). ■

### 8.5.2 Classes of stationary spacetimes and coordinates

Related with the definition of extended reference frame is that of *stationary*, *static* and *ultrastatic* spacetime. Here, the timelike vector field used to define a reference frame is further assumed to be Killing. Furthermore in two cases it is supposed that the spacetime is foliated by spacelike surfaces as in the definition of reference frame. However requirements (1) and (2) are not assumed.

**Definition 8.26.** Let  $(M^n, \mathbf{g})$  be a spacetime.

- (1)  $(M^n, \mathbf{g})$  is called **stationary** if it admits a timelike Killing vector field  $K$ .
- (2)  $(M^n, \mathbf{g})$  **static** if it is stationary and all the maximal integral curves of the preferred timelike Killing vector field  $K$  meet  $n - 1$ -dimensional embedded submanifold  $\Sigma$  and are normal to it.

(3)  $(M^n, \mathbf{g})$  **ultrastatic** if it is static and the preferred timelike Killing vector field  $K$  satisfies

$$\mathbf{g}(K, K) = -1$$

everywhere in  $M^n$ . ■

**Remark 8.27.**

(1) Suppose that  $(M^n, \mathbf{g})$  is static with preferred timelike Killing vector  $K$  and preferred spatial section  $\Sigma$ . Evidently  $\Sigma$  is necessarily spacelike since its co-normal vector is timelike (its contravariant representation is parallel to  $K$ ).

If  $K$  is complete, we can define a family of  $n - 1$ -dimensional embedded submanifolds

$$\Sigma_t := \Phi_t^{(K)}(\Sigma), \quad t \in \mathbb{R}.$$

As each  $\Phi_t^{(K)}$  is an isometry, these surfaces are spacelike and orthogonal to  $K$ . By construction, their union is  $M^n$  since  $\Sigma$  meets each integral curve of  $K$  and they are also pairwise disjoint if every maximal integral curve of  $K$  meets  $\Sigma$  exactly once: if  $\gamma(t), \gamma(t') \in \Sigma$  then  $t = t'$ . In this case we have an extended reference frame  $(M^n, K, t)$  covering the whole manifold, where  $t$  is the parameter of the integral curves of  $K$ . More precisely,  $(M^n, \mathbf{g})$  turns out to be isometrically diffeomorphic to  $\mathbb{R} \times \Sigma$ , where this latter manifold is equipped with a metric of the form

$$\mathbf{g}(K, K)dt \otimes dt + \mathbf{h},$$

where  $\mathbf{h}$  is the metric induced on  $\Sigma$  from  $\mathbf{g}$  and  $\mathbf{g}(K, K)$  generally depends on the position in  $\Sigma$  but not on  $t$ . The isometry is the map  $\mathbb{R} \times \Sigma \ni (t, p) \mapsto \Phi_t^{(K)}(p)$ . (See (1) and (2) Exercises 8.31.)

If  $K$  is not complete and we cannot control the generic domain, this construction is valid only locally according with the domain of the flow  $\Phi^{(K)}$  in the following sense as the reader can prove easily taking advantage of Propositions 4.18 and 4.21. Every maximal integral curve  $\gamma$  of  $K$  meet  $\Sigma$  at a corresponding point  $p$ . Hence, there is an interval  $[0, \omega_p)$  with the property that, if  $t_1 \in [0, \omega_p)$ , an open neighborhood  $S_p \subset \Sigma$  of  $p$  exists such that the spacelike pairwise disjoint  $n - 1$ -dimensional embedded submanifolds

$$S_{pt} := \Phi_t^{(K)}(S_p) \quad t \in [0, t_1]$$

are well defined. The analogue is valid for an interval  $(\alpha_p, 0]$ . In this discussion  $\omega_p \in (0, +\infty]$  and  $\alpha_p \in [-\infty, +0)$

(2) It is possible to consider a weaker condition of static spacetime. We say that a spacetime  $(M^n, \mathbf{g})$  is **locally static** if there is a Killing vector  $K$  such that, for every  $p \in M^n$ , there is an embedded spacelike  $n - 1$ -dimensional submanifold  $\Sigma_p$  such that  $p \in \Sigma_p$  and  $K$  is normal to  $\Sigma$ . An example of spacetime that is locally static but not static is the open region of the 2-dimensional Minkowski spacetime defined by requiring  $t \in \mathbb{R}$  and  $|x| < 2 + \sin ct$ , where  $t, x$  are standard Minkoskian coordinates and  $K = \frac{\partial}{\partial t}$ . In this case, there is no common spacelike

section that meets all complete integral lines of  $K$ , though every section at constant  $t$  satisfies the requirement for a locally static spacetime. Some texts use the local requirement as the definition of static spacetime. ■

Let us examine the definition from the point of view of physics.

- (a) Stationary means that, evolving along the integral curves of  $K$ , the metric properties of the spacetime (including the gravitational ones) are seen as invariant.
- (b) A static spacetime has further properties. First of all, we can decompose all structures into time (along  $K$ ) and space (each surface  $\Sigma_t$ ) and, as above, the metric properties decomposed in this way are invariant when the (Killing) time varies. However, we can also *revert the sign of the Killing time*, around a given instant of time, redefined  $t = 0$ , and the metric properties of the spacetime remain fixed. This inversion appears in spacetime as a reflection with respect to every fixed rest space  $\Sigma_t$  (or a local version  $S_{pt}$ , when  $K$  is not complete). In some sense, for instance in black hole theory, a stationary non-static spacetime may involve a spatial *rotation* and the orientation of the rotation breaks the time reversal.
- (c) An ultrastatic spacetime has the further nice property that the Killing time and the proper time of the worldlines tangent to the Killing vector coincide.

**Remark 8.28.** We stress that, physically speaking, no equilibrium states, for instance in thermodynamics, are possible in the absence of a timelike Killing vector since sooner or later some metric or gravitational phenomena change the equilibrium. ■

In some cases it is convenient to adapt an extended reference frame to a static structure using the Killing vector  $K$  as the time vector and the surfaces  $\Sigma_t$  ( $S_{pt}$ , when  $K$  is not complete) as rest spaces. Furthermore the definitions above permit a corresponding definition of local charts.

**Definition 8.29.** Consider a spacetime  $(M^n, \mathbf{g})$  and a local chart  $\phi : U \ni p \mapsto (x^0, \dots, x^{n-1}) \in \mathbb{R}^n$ .

- (1)  $\phi$  is **stationary** if the spacetime is stationary with preferred timelike Killing vector  $K$ ,  $\frac{\partial}{\partial x^0} = K$  in  $U$ .
- (2)  $\phi$  is **static** if the spacetime is static with preferred timelike Killing vector  $K$  and both  $\frac{\partial}{\partial x^0} = K$  and  $\frac{\partial}{\partial x^\alpha}$  are normal to  $K$  if  $\alpha = 1, \dots, n-1$  in  $U$ .
- (3)  $\phi$  is **ultrastatic** if the spacetime is ultrastatic and  $\phi$  is static. ■

**Remark 8.30.**

- (1) Observe that a static and ultrastatic chart,  $x^0$  coincides to the parameter of the integral

curves of  $K$  and labels the surfaces  $\Sigma_{x^0}$ .

(2) Furthermore, again in the static and ultrastatic case, for every fixed  $x^0$  the coordinates  $x^1, \dots, x^{n-1}$  define a local chart of the spatial section  $\Sigma_{x^0}$  (or on a local version  $S_{px^0}$  when  $K$  is not complete) according to Remark 8.27.

(3) According to the given definitions, in a stationary chart,

$$\frac{\partial g_{ab}}{\partial x^0} = 0 .$$

In a static chart, it also holds

$$g_{0b} = g_{b0} = 0 \quad b = 2, \dots, n .$$

In a ultrastatic chart, we have in addition to the previous facts,

$$g_{00} = -1 .$$

### Exercises 8.31.

1. Consider a static spacetime  $(M^n, \mathbf{g})$  with preferred timelike Killing vector  $K$  and preferred orthogonal spatial section  $\Sigma$ . Assume that  $K$  is complete and each maximal integral curve of  $\gamma$  meets  $\Sigma$  exactly once (i.e., if  $\gamma(t), \gamma(t') \in \Sigma$  then  $t = t'$ ). Prove that  $(M^n, \mathbf{g})$  is isometrically diffeomorphic to  $\mathbb{R} \times \Sigma$ , where this manifold is equipped with the metric

$$\mathbf{g}(K, K)dt \otimes dt + \mathbf{h} ,$$

where  $\mathbf{h}$  is the metric induced on  $\Sigma$  from  $\mathbf{g}$  (thus it does not depend on  $t$ ), and  $\mathbf{g}(K, K)$  generally depends on the position in  $\Sigma$  but not on  $t$ . The isometry is the map

$$f : \mathbb{R} \times \Sigma \ni (t, p) \mapsto \Phi_t^{(K)}(p) \in M^n .$$

**Solution.** The map  $\mathbb{R} \times \Sigma \ni (t, p) \mapsto \Phi_t^{(K)}(p)$  is evidently surjective since every  $q \in M^n$  admits an integral line  $\gamma_q$  of  $K$  such that  $\gamma_q(t_1) = q$  and, with the said hypotheses,  $\gamma_q(t_0) = p \in \Sigma$ . Hence  $f(t_0, p) = q$ . The map is also injective. Indeed, if  $\Phi_t^{(K)}(p) = \Phi_{t'}^{(K)}(p')$ , then  $\gamma(t - t') = \Phi_{t-t'}^{(K)}(p) = p' \in \Sigma$  and thus  $\gamma(0) = \Phi_0^{(K)}(p) = p \in \Sigma$  implies  $t - t' = 0$  so that  $p = p'$ . The map  $f$  is smooth by construction. To prove that it is a diffeomorphism it is sufficient to prove that  $df_{(t,p)}$  has rank  $n$ , i.e., it is injective. For  $(s_{t_0}, v_{p_0}) \in T_{(t_0, p_0)}\mathbb{R} \times \Sigma = T_{t_0}\mathbb{R} \times T_{p_0}\Sigma$ , we have

$$df(s_{t_0}, v_{p_0}) = s_{t_0}K_{f(t_0, p_0)} + d\Phi_{t_0}^{(K)}v_{p_0} ,$$

where the pushforward  $d\Phi_{t_0}^{(K)}$  is the restriction to  $\Sigma \ni p$  of  $d(\Phi_{t_0}^{(K)}) : T_{p_0}M^n \rightarrow T_pM^n$ . As a consequence,  $d\Phi_{t_0}^{(K)}v_{p_0}$  is tangent to  $\Sigma_{t_0} := \Phi_{t_0}^{(K)}(\Sigma)$  by construction, whereas  $K_{f(t_0, p_0)}$  is normal to that surface. Hence  $s_{t_0}K_{f(t_0, p_0)} + d\Phi_{t_0}^{(K)}v_{p_0} = 0$  means that  $s_{t_0}K_{f(t_0, p_0)} = 0$  and  $d\Phi_{t_0}^{(K)}v_{p_0} = 0$ . Since  $K \neq 0$  everywhere as it is timelike, we conclude that  $s_{t_0} = 0$ . Similarly  $v_{p_0} = 0$  because  $d\Phi_{t_0}^{(K)} : T_pM^n \rightarrow T_pM^n$  is injective  $\Phi_{t_0}^{(K)} : M^n \rightarrow M^n$  being a diffeomorphism. We conclude

that  $f$  is a diffeomorphism. The fact that  $f$  is also an isometry (it preserves the scalar product) when the metric on the domain of  $f$  is  $\mathbf{g}(K, K)dt \otimes dt + \mathbf{h}$  is an immediate consequence of the above decomposition of  $df$  and of the fact that  $\Phi_{t_0}^{(K)} : M^m \rightarrow M^n$  is an isometry and it remains an isometry when restricting the domain to  $\Sigma$  and the co-domain to its image  $\Sigma_{t_0}$ , also observing that  $K$  is normal to every  $\Sigma_t$  by construction.

2. Assume all the hypotheses of the previous exercise, completeness of  $K$  in particular, except the one which requires that each maximal integral curve of  $\gamma$  meets  $\Sigma$  exactly once. Find a counterexample to the fact that  $(M^n, \mathbf{g})$  is isometrically diffeomorphic to  $\mathbb{R} \times \Sigma$  with the said metric.

(*Hint.* Consider the two-dimensional Minkowski spacetime constructed on  $\mathbb{R}^2$  with  $ct = x^0 = y$  and  $x^1 = x$ . Then construct a time oriented spacetime over a cylinder by identifying  $(x, y)$  and  $(x, y + T)$  for some  $T > 0$  fixed. Use  $K := \frac{\partial}{\partial y}$  smoothly extended at  $y = T$ . Notice that this  $K$  is complete.)

### 8.5.3 The problem of the spatial metric of an extended reference frame

We come here to a still controversial and still partially open problem regarding what physically meaningful metric one should define on the rest spaces  $\Sigma_t$  of an extended reference frame when the time vector field  $T$  is *not* orthogonal to  $\Sigma_t$ . This problem was tackled in various versions by Langevin, Ehrenfest, Born, Landau and Lifshitz, and Cattaneo. The prototype of this long standing issue is in particular the version of the problem stated for the metric of the *rotating platform* we shall briefly discuss in the next section<sup>10</sup>.

The problem arises when realizing that, if  $T$  is not orthogonal to  $\Sigma_t$ , then the “infinitesimal” rest space (see (8.4))  $d\Sigma_{\gamma(t)} \subset T_{\gamma(t)}M$  of an observer whose worldline is an integral curve of  $T$  is *not* tangent to  $\Sigma_t$  at  $\gamma(t)$ . Generally speaking, we expect that the “infinitesimal” rest space  $d\Sigma_{\gamma(t)}$  of an observer whose tangent vector is  $T$  should be an “infinitesimal portion” of the extended rest space of the reference frame constructed out  $T$ . With a more precise mathematical language we would expect that

$$d\Sigma_{\gamma(t)} = T_{\gamma(t)}\Sigma_t .$$

This is true if and only if  $T$  is normal to  $\Sigma_t$ .

For the same reason, it does not seem physically meaningful assuming that the metric on  $\Sigma_t$  coincides with the metric  $\mathbf{g}^{(\Sigma_t)}$  induced from  $M^n$ , since this induced metric does not produce the physically meaningful scalar product in  $d\Sigma_{\gamma(t)}$ . If  $T$  is instead normal to  $\Sigma_t$ , no problem arises, as the scalar product  $\mathbf{g}^{(\Sigma_t)}$  is just the one already present in the subspaces  $d\Sigma_{\gamma(t)}$ .

Any definition of a suitable metric  $\mathbf{h}_t$  in  $\Sigma_t$  should be based on some established physical fact. In Landau and Lifshitz’ approach, this physical fact is the experimental result stated in

<sup>10</sup>See, in particular, Rizzi, G.; Ruggiero, M.L. (2002). *Space geometry of rotating platforms: an operational approach*. Found. Phys. 32 (10): 1525–1556 and Rizzi, G.; Ruggiero, M. L. (2004). *Relativity in Rotating Frames*. Dordrecht: Kluwer.

the following postulate they *implicitly* assume.

**Back and Forth Journey Postulate.** *The speed of light is constantly  $c$  if measured in vacuum along an ideal ruler travelled by the light back and forth, independently of the state of motion of the reference frame where the ruler is at rest.*

**Remark 8.32.**

(1) It is worth remarking that this assumption is different from *Einstein's synchronization procedure* of distant clocks at rest to each other with a known distance  $L$  between them. In that case, they are synchronized just sending a particle of light from one clock at time  $\tau_0$  to the other clock requiring that it says  $\tau_0 + \frac{L}{c}$  on receiving the particle. Referring to the Back and Forth Journey Postulate there is nothing to synchronize instead, since with a closed path only one clock is exploited.

(2) The new principle may appear physically difficult to test in the general case if the spacetime is not stationary, since the metric properties of the a real ruler may change in time. A better version of the principle could concern “infinitesimally” short rulers in order that the time necessary for the light to run them can be taken as small as possible in comparison with the typical scale of time used by the metric to change<sup>11</sup>. Another point of view is that the postulate can be adopted only for stationary spacetimes and stationary coordinates in a stationary chart. In the rest of the section we shall apply the postulate for infinitesimally small rulers without any hypothesis about stationarity. ■

We are going to prove that the above postulate completely fixes the metric  $\mathbf{h}_t$  on every rest space of an extended reference frame. Furthermore, if  $T$  is normal to  $\Sigma_t$ , then  $\mathbf{h}_t$  coincides with the one induced from  $\mathbf{g}$ .

Let us consider an extended reference frame  $(N, T, t)$  in a spacetime  $(M^n, \mathbf{g})$  and fix an event  $p \in N$ . As a consequence of our definitions,  $\Sigma_{t(p)}$  is the rest space containing  $p$ . Another point  $q \in \Sigma_{t(p)}$  infinitesimally close to  $p$  in the same rest space  $\Sigma_{t(p)} = \Sigma_{t(q)}$  can be heuristically defined by fixing a very short vector  $\delta X_p \in T_p \Sigma_{t(p)}$ : in a chart around  $p$ , where  $p$  corresponds to the origin, the components of  $q$  are those of  $\delta X_p$ . That is our mathematical description of the “infinitesimally” short ruler said above.

Let next evolve these two points along the integral lines of  $T$ , respectively,  $\gamma_p$  and  $\gamma_q$ . We also suppose that a particle of light is launched from  $\gamma_q$  to  $\gamma_p$ , reaching the latter exactly at  $p$ ,

---

<sup>11</sup>I do not like very much this viewpoint since it seems to suggest that ideal rulers cannot exist, because their length necessarily changes when the metric evolves in time. It is clear that if adopting that radical point of view, then nothing may make sense in General Relativity, since the geometry, including its evolution in time described by the evolution of the metric, is just a description of the properties of the ideal rulers (and ideal clocks)! A crucial observation is that the structure of the ideal rulers (and clocks) is not described in General Relativity where their existence is only postulated. From the physical point of view a ruler is “ideally rigid” as a consequence of *other interactions*, different from the gravitational one. A safer point of view is that, under a certain scale, ideal rulers exist just in view of the predominant action of these other interactions. This interpretation however seems to imply that General Relativity is not a fundamental theory and that, in its standard presentation at least, it is a strictly macroscopic theory.

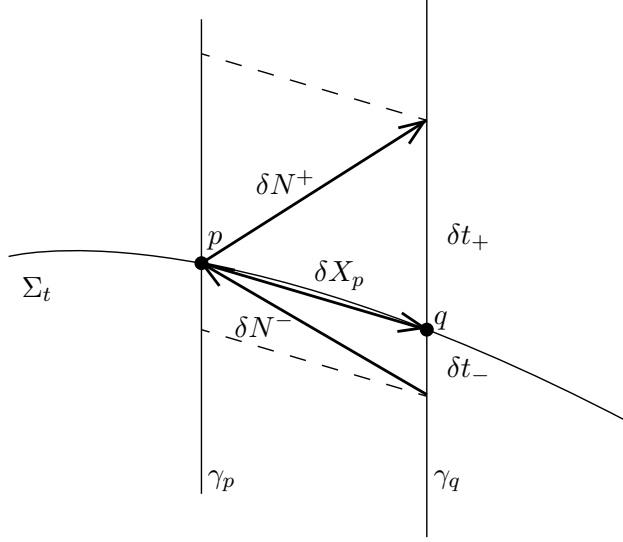


Figure 8.3: Determination of the physical length of  $\delta X_p$ .

where it is bounced back to  $\gamma_q$ . This is a back and forth journey along a ruler, so we have to impose that the velocity of that particle of light is  $c$ . This constraints defines the physical length of  $\delta X_p$  in terms of a metric on  $\Sigma_{t(p)}$  as we are going to prove.

As said, we arrange the starting time (defined by the time function) of the particle of light such that it reaches  $\gamma_p$  exactly at  $p$ . The lightlike geodesic describing the worldline of the particle of light in the first part of its trip has a tangent vector (supposed to be applied on  $p$ )

$$\delta N_p^{(-)} = \delta t_- T_p - \delta X_p ,$$

where  $\delta t_- > 0$ . The tangent vector for the second part of the trip is instead

$$\delta N_p^{(+)} = \delta t_+ T_p + \delta X_p ,$$

where again  $\delta t_+ > 0$ . (See Fig. 8.32)

The requirement that  $\delta N^{(\pm)}$  is lightlike reads

$$\mathbf{g}(\delta t_{\pm} T_p \pm \delta X_p, \delta t_{\pm} T_p \pm \delta X_p) = 0 ,$$

that is

$$(\delta t_{\pm})^2 \mathbf{g}(T_p, T_p) \pm 2\delta t_{\pm} \mathbf{g}(T_p, \delta X_p) + \mathbf{g}(\delta X_p, \delta X_p) = 0 . \quad (8.40)$$

We can solve these equations for the unknowns  $\delta t_{\pm} \geq 0$  taking the constraints  $-\mathbf{g}(T_p, T_p) > 0$  and  $\mathbf{g}(\delta X_p, \delta X_p) \geq 0$  into account:

$$\delta t_{\pm} = \frac{\pm \mathbf{g}(T_p, \delta X_p)}{-\mathbf{g}(T_p, T_p)} + \frac{\sqrt{\mathbf{g}(T_p, \delta X_p)^2 - \mathbf{g}(T_p, T_p) \mathbf{g}(\delta X_p, \delta X_p)}}{-\mathbf{g}(\delta X_p, \delta X_p)} . \quad (8.41)$$



The total amount of reference frame time spent by the particle of light in its complete trip is

$$\delta t = \delta t_+ + \delta t_- = 2 \frac{\sqrt{\mathbf{g}(T_p, \delta X_p)^2 - \mathbf{g}(T_p, T_p)\mathbf{g}(\delta X_p, \delta X_p)}}{-\mathbf{g}(T_p, T_p)}. \quad (8.42)$$

To obtain the value  $\delta\tau$  corresponding to  $\delta t$  in terms of the *proper time* measured along  $\gamma_p$ , it is necessary to multiply this result with  $\frac{\sqrt{-\mathbf{g}(T_p, T_p)}}{c}$ , obtaining

$$\delta\tau = \frac{2}{c} \sqrt{\mathbf{g}(\delta X_p, \delta X_p) + \frac{\mathbf{g}(T_p, \delta X_p)\mathbf{g}(T_p, \delta X_p)}{-\mathbf{g}(T_p, T_p)}}. \quad (8.43)$$

At this juncture, we are in a position to use the postulate, requiring that the back and forth speed of light computed along this ruler takes the value  $c$ , so that the total length of the ruler  $2L$  must satisfy

$$L = \frac{c\delta\tau}{2} = \sqrt{\mathbf{g}(\delta X_p, \delta X_p) + \frac{\mathbf{g}(T_p, \delta X_p)\mathbf{g}(T_p, \delta X_p)}{-\mathbf{g}(T_p, T_p)}}, \quad (8.44)$$

where  $L$  coincides with the physical length of  $\delta X_p$ . If a physically meaningful scalar product

$$\mathbf{h}_p : T_p\Sigma_{t(p)} \times T_p\Sigma_{t(p)} \rightarrow \mathbb{R}$$

exists in agreement with the Back and Forth Journey Postulate, we should have  $L^2 = \mathbf{h}(\delta X_p, \delta X_p)$ . Hence

$$\mathbf{h}(\delta X_p, \delta X_p) = \mathbf{g}(\delta X_p, \delta X_p) - \frac{\mathbf{g}(T_p, \delta X_p)\mathbf{g}(T_p, \delta X_p)}{\mathbf{g}(T_p, T_p)}$$

Since a scalar product is completely determined by its norm as a consequence of the *polarization identity*, if our metric exists, it has necessarily the form (where we relax the heuristic requirement of dealing with “short” vectors)

$$\mathbf{h}(X_p, Y_p) := \mathbf{g}(X_p, Y_p) - \frac{\mathbf{g}(T_p, X_p)\mathbf{g}(T_p, Y_p)}{\mathbf{g}(T_p, T_p)} \quad \text{for all } X_p, Y_p \in T_p\Sigma_{t(p)}. \quad (8.45)$$

The above symmetric bilinear form  $\mathbf{h}$  is actually defined in the full  $T_pM^n$ , but we are interested to it only when  $X_p, Y_p \in T_p\Sigma_{t(p)}$ . We have to prove that, with this choice of its domain, it is positive definite so that it deserves the status of a scalar product.

**Proposition 8.33.** *Let  $\mathbf{g} : V \times V \rightarrow \mathbb{R}$  be a Lorentzian metric over the real vector space  $V$  of dimension  $n$ . If  $T \in V$  is a timelike vector and  $\Sigma \subset V$  is a  $n - 1$ - dimensional subspace spanned by spacelike vectors, then the symmetric bilinear form*

$$\mathbf{h}(X, Y) := \mathbf{g}(X, Y) - \frac{\mathbf{g}(T, X)\mathbf{g}(T, Y)}{\mathbf{g}(T, T)} \quad \text{for all } X, Y \in \Sigma \quad (8.46)$$

*is a (positive) scalar product in  $\Sigma$ . Moreover*

$$\mathbf{h}(X, Y) := \mathbf{g}(PX, PY) \quad \text{for all } X, Y \in \Sigma. \quad (8.47)$$

where  $P : V \rightarrow V$  is the orthogonal projector onto

$$\text{Span}(T)^\perp := \{Y \in V \mid \mathbf{g}(Y, T) = 0\}.$$

(In other words,  $P$  is the linear map associating  $Y \in V$  with the second addend of the decomposition  $Y = Y_T + Y_T^\perp$  referred to the direct decomposition  $V = \text{Span}(T) \oplus \text{Span}(T)^\perp$ .)

**Proof.** Since

$$Q = \frac{\mathbf{g}(T, \cdot)}{\mathbf{g}(T, T)} T$$

is the orthogonal projector onto  $\text{Span}(T)$ , it must be

$$P = I - Q = I - \frac{\mathbf{g}(T, \cdot)}{\mathbf{g}(T, T)} T.$$

Inserting this expression in the right-hand side of (8.47), we obtain (8.46) with elementary computations. We also have that

$$\mathbf{h}(X, Y) = \mathbf{g}(PX, PY) \geq 0 \quad \text{for all } X, Y \in V.$$

To prove it, it is sufficient to select an orthonormal basis in  $\text{Span}(T)^\perp$  and to add a non-vanishing unit vector of  $\text{Span}(T)$  producing a pseudo orthonormal basis of  $V$ .  $\mathbf{g}$  takes its canonical form with signature  $(-1, 1, \dots, 1)$  in that basis so that  $\mathbf{g}(PX, PY) \geq 0$  trivially because  $\mathbf{g}(P\cdot, P\cdot)$  is the (positive) scalar product induced by  $\mathbf{g}$  on  $\text{Span}(T)^\perp$ . Finally  $\mathbf{h}(X, X) = 0$  implies  $X = 0$  for  $X \in \Sigma$  concluding the proof. Indeed,  $\mathbf{h}(X, X) = 0$  yields  $\mathbf{g}(PX, PX) = 0$  which means  $PX = 0$  because  $\mathbf{g}(P\cdot, P\cdot)$  is positive.  $PX = 0$  is equivalent to  $QX = X$ , that is  $X \in \text{Span}(T)$ , so that  $X \in \Sigma \cap \text{span}(T) = \{0\}$ .  $\square$

**Definition 8.34.** If  $(N, T, t)$  is an extended reference frame in the spacetime  $(M^n, \mathbf{g})$ , the **physical metric** on every rest space  $\Sigma_t$  is

$$\mathbf{h}^{(t)}(X_p, Y_p) := \mathbf{g}(X_p, Y_p) - \frac{\mathbf{g}(T_p, X_p)\mathbf{g}(T_p, Y_p)}{\mathbf{g}(T_p, T_p)} \quad \text{for all } X_p, Y_p \in T_p \Sigma_t \quad (8.48)$$

where  $p \in \Sigma_t$ . ■

**Remark 8.35.**

(1) The metric (8.48) was also introduced by Cattaneo<sup>12</sup> following a different mathematical procedure. We prefer here to insist on the physical meaning of that metric.

(2) In coordinates co-moving with the reference frame  $x^0 = ct, x^1, \dots, x^{n-1}$  we find

$$\mathbf{h}^{(t)} = \sum_{\alpha, \beta=1}^{n-1} h_{\alpha\beta}^{(t)} dx^\alpha \otimes dx^\beta \quad \text{where} \quad h_{\alpha\beta}^{(t)} = g_{\alpha\beta} - \frac{g_{0\alpha}g_{0\beta}}{g_{00}} \quad (8.49)$$

---

<sup>12</sup>See Rizzi, G.; Ruggiero, M. L. (2004). Relativity in Rotating Frames. Dordrecht: Kluwer.

and this is the form of the metric found by Landau and Lifshitz [LaLi80]. It is possible to prove (see (1) in Exercises 8.37) that the matrix of coefficients  $h_{\alpha\beta}^{(t)}$  is the inverse of the matrix of coefficients  $g^{\alpha\beta}$ , where  $g^{ab}$  are as usual the components of the metric in each cotangent space induced by  $\mathbf{g}$ . ■

It is now interesting to compute the *one-way* speed of light using the found metric. Suppose the worldline of a particle of light passing through  $p$  at time  $t$  has tangent vector  $N_p \in T_p M^n$ . We can uniquely decompose it as

$$N_p = \delta t T_p + \delta X_p \quad \text{where } \delta t > 0 \text{ and } \delta X \in T_p \Sigma_{t(p)}. \quad (8.50)$$

Imposing  $\mathbf{g}(N_p, N_p) = 0$ , the value of  $\delta t$  is the same as  $\delta t_+$  in (8.41). The speed of light at  $p$  is the ration of the length of  $\delta X_p$  referred to the metric  $\mathbf{h}^{(t)}$  and the proper time interval (measured along an integral curve of  $T$ ) corresponding to  $\delta t$ , i.e.,

$$\delta\tau = \frac{1}{c} \sqrt{-\mathbf{g}(T_p, T_p)} \delta t$$

With some lengthy but elementary computation we find<sup>13</sup>

$$\left\| \frac{\delta X_p}{\delta\tau} \right\|_{\mathbf{h}} = \frac{c}{1 + \frac{\mathbf{g}(T_p, \delta X_p)}{\sqrt{-\mathbf{g}(T_p, T_p)} \|\delta X_p\|_{\mathbf{h}(t)}}}$$

It is clear that the result does not depend on  $\delta X_p$  but only on the unit vector  $n_p \in T_p \Sigma_{t(p)}$  defined by it, always using the metric  $\mathbf{h}^{(t)}$  to normalize  $\delta X_p$ . The final formula is

$$c(t, n_p) = \frac{c}{1 + \frac{\mathbf{g}(T_p, n_p)}{\sqrt{-\mathbf{g}(T_p, T_p)}}}, \quad (8.51)$$

That is the *one-way* speed of light measured at  $p \in N$  with respect to the extended reference frame  $(N, T, t)$ . It is evident that  $c(t, n_p) = c$  independently from the direction  $n_p \in T_p \Sigma_t$  if and only if  $T_p$  is normal to  $\Sigma_t$ .

**Remark 8.36.**

(1) The *one-way* speed of light depends on which synchronization procedure one uses to synchronize the ideal clocks placed at different “infinitesimally close” position: one at  $p$  and the other at “ $q = p + \delta X_p$ ”, when approximating an “infinitesimal” neighborhood in  $\Sigma_{t(p)}$  of  $p$  with  $T_p \Sigma_{t(p)}$  as in Fig. 8.32. We stress that these clocks measure the *proper time*  $\tau$  and not the time  $t$  of the reference frame. The speed of light (8.51) is obtained assuming that both clocks say  $\tau = 0$

<sup>13</sup>Landau and Lifshitz use a different definition of velocity defined in a reference frame  $(N, T, t)$  when translating the approach of [LaLi80] to our framework. That definition, though defined into a complicated fashion in my view, is nothing but the one measured at rest with an observer whose worldline is and integral curve of  $T$  according to (8.5), completely disregarding the existence of  $t$  and  $\Sigma_t$ . In that case the speed of light is evidently always  $c$ .

in the rest space  $\Sigma_{t(p)}$  when the measure of the speed of a particle of light starts. This means that we are actually using a different synchronization procedure of ideal clocks than Einstein's one which imposes a different one-way speed of light (8.51). This different procedure is however compatible with the postulate about the constant value of  $c$  measured back and forth along an ideal ruler. Einstein's procedure is possible only when choosing the time vector  $T$  and the time function  $t$  such that  $T$  and  $\Sigma_t$  are orthogonal.

If  $T$  and  $\Sigma_t$  are orthogonal, then (8.41) implies that the time function  $t$  is such that

$$\delta t_+ = \delta t_- .$$

To better illustrate this identity consider two ideal clocks at rest with observers whose worldlines are tangent to  $T$  and suppose that those observers continuously exchange particles of light, as in the discussion around (8.41). A time function  $t$  permits to synchronize *à la* Einstein these two clocks in the events  $p$  and  $q$ , which are simultaneous according to  $t$ , if the time  $\delta t_-$  necessary to reach  $p$  from the worldline  $\gamma_q$ , measured from  $q$  along the curve, is equal to the one  $\delta t_+$  necessary to go back from  $p$  to  $\gamma_q$ , again measured from  $q$ . In this situation, if both observers set their ideal clocks at  $\tau = 0$  at the ( $t$ -simultaneous) events  $p$  and  $q$ , then the speed of light from  $p$  to  $q$  (and *viceversa*) turns out to be  $c$  directly from (8.51).

(2) In general, any synchronization imposed at  $p, q \in \Sigma_t$  of the proper time of ideal clocks transported by  $\gamma_p$  and  $\gamma_q$  gets lost in the future of those events due to the gravitational redshift which distinguishes  $t$  and  $\tau$ , in other words,  $\frac{d\tau}{dt} = \frac{\sqrt{-g_{00}}}{c}$  depends on the considered worldline tangent to  $T$ . Even if Einstein's synchronization is possible and it was imposed on  $\Sigma_t$ , without re-synchronizing the clocks (setting again  $\tau = 0$  for both clocks in another future rest space  $\Sigma_{t'}$ ) further measurements of the one-way speed of light between  $\gamma_p$  and  $\gamma_q$  at  $t'$  produce values different from  $c$ . All that is very different from what happens in Special Relativity, when the reference frame is inertial: there the clocks remain synchronized for ever after the first synchronization procedure and the one-way speed of light constantly takes the value  $c$ . The reason is that in Minkowskian coordinates  $g_{00}$  is constant in space and time. In *ultrastatic* spacetimes, for reference systems with  $T$  given by the preferred timelike Killing vector, Einstein's synchronization is similarly permanent.

(3) The natural measure induced by the metric  $\mathbf{h}^{(t)}$  (8.45) on  $\Sigma_t$  in co-moving coordinate is

$$d\nu^{(t)} := \sqrt{\det h^{(t)}} dx^1 \dots dx^n$$

where  $h^{(t)}$  is the matrix of the components  $h_{\alpha\beta}^{(t)}$  appearing in (8.49). It is possible to prove that (see (2) in Exercise 8.37)

$$d\nu^{(t)} = \frac{d\mu(\mathbf{g}^{(\Sigma_t)})}{\sqrt{\mathbf{g}(T, T)\mathbf{g}(dt^\sharp, dt^\sharp)}} \quad (8.52)$$

where the left-hand side is the geometric measure on  $\Sigma_t$  associated to the induced metric  $\mathbf{g}^{(\Sigma_t)}$ . Notice that, if  $T$  is normal to  $\Sigma_t$ , the two measures coincide as we have that

$$\sqrt{\mathbf{g}(T, T)\mathbf{g}(dt^\sharp, dt^\sharp)} = \sqrt{g_{00} g^{00}} = \sqrt{g_{00}(g_{00})^{-1}} = 1 ,$$

where we have used co-moving coordinates where, by hypothesis,  $g_{0\alpha} = g_{\alpha 0} = 0$ , so that  $g^{00} = (g_{00})^{-1}$ .

The geometric measure  $\mu^{(\mathbf{g}(\Sigma_t))}$  is the one entering the conservation laws via divergence theorem (see Section 8.3.2). This fact implies that, when writing the conserved quantities  $Q_{\Sigma_t}$  (8.20) over  $\Sigma_t$  using the physical measure  $\nu^{(t)}$ , a further factor must appear in the integrand

$$Q_{\Sigma_t} := \int_{\Sigma_t} \langle J, n \rangle \sqrt{\mathbf{g}(T, T) \mathbf{g}(dt^\sharp, dt^\sharp)} d\nu^{(t)}.$$

■

### Exercises 8.37.

1. Consider a coordinate system  $x^0, x^1, \dots, x^{n-1}$  comoving with an extended reference frame in a spacetime  $(M^n, \mathbf{g})$ . Let  $g_{ab}$  denote the components of the spacetime metric and  $h_{\alpha\beta}^{(t)}$ ,  $\alpha, \beta = 1, \dots, n-1$  denote the components of the physical metric on every rest space at  $x^0$  constant. Prove that

$$\sum_{\beta=1}^{n-1} g^{\alpha\beta} h_{\beta\gamma}^{(t)} = \delta_\gamma^\alpha, \quad \alpha, \gamma = 1, \dots, n-1, \quad (8.53)$$

where  $g^{ab}$  are the usual components of the metric in each cotangent space induced by  $\mathbf{g}$ .

**Solution.** The identity  $g^{ab} g_{bc} = \delta_c^a$  expands to following equations when specialising the values of the free indices  $a$  and  $c$ .

$$\sum_{\beta=1}^{n-1} g^{\alpha\beta} g_{\beta\gamma} + g^{\alpha 0} g_{\beta 0} = \delta_\gamma^\alpha, \quad \sum_{\beta=1}^{n-1} g^{\alpha\beta} g_{\beta 0} + g^{\alpha 0} g_{00} = 0, \quad \sum_{\beta=1}^{n-1} g^{0\beta} g_{\beta 0} + g^{00} g_{00} = 1.$$

The expression of  $g^{\alpha 0}$  obtained from the second equation inserted in the first one produces the wanted identity when taking (8.49) into account.

2. Prove (8.52):

$$d\nu^{(t)} = \frac{d\mu^{(\mathbf{g}(\Sigma_t))}}{\sqrt{\mathbf{g}(T, T) \mathbf{g}(dt^\sharp, dt^\sharp)}}.$$

**Solution.** Using the same notation as in the previous exercise, it is sufficient to prove that

$$\sqrt{\det h} = \frac{\sqrt{\det g^{(\Sigma_t)}}}{\sqrt{g_{00} g^{00}}},$$

where

$$g^{(\Sigma_t)} := [g_{\alpha\beta}]_{\alpha, \beta=1, \dots, n-1}, \quad h^{(t)} := [h_{\alpha\beta}^{(t)}]_{\alpha, \beta=1, \dots, n-1}.$$

In fact the value of  $\sqrt{g_{00} g^{00}}$  does not depend on the used co-moving chart and once the identity above is satisfied in all co-moving local chart, (8.52) arises by a standard use of a partition of

unity on  $\Sigma_t$ . From the Cramer rule to compute the inverse of a matrix, since  $g_{00}$  is an element of the inverse matrix of the matrix of coefficients  $g^{ab}$ , we have

$$g_{00} = \frac{\det \tilde{g}}{\det g^{-1}}$$

where  $\tilde{g}$  is the matrix of elements  $g^{\alpha\beta}$  with  $\alpha, \beta = 1, \dots, n-1$ . Taking (8.53) into account, this identity becomes

$$g_{00} = \frac{\det(h^{(t)})^{-1}}{\det g^{-1}}$$

so that

$$\det g = g_{00} \det h^{(t)}.$$

With the same argument, we have that

$$g^{00} = \frac{\det g^{(\Sigma_t)}}{\det g}.$$

Inserting the expression of  $\det g$  obtained from this identity in the penultimate equation, we find (8.52) when taking the square root of both sides.

#### 8.5.4 Einstein's synchronizability in stationary spacetimes

Consider an extended reference frame  $(N, T, t)$  in the spacetime  $(M^n, \mathbf{g})$  where  $T$  is *not* orthogonal to the rest spaces  $\Sigma_t$ . A natural question arises: if, keeping  $T$ , it is possible to re-define  $t$  in order to fulfill that requirement so that the physical metric  $\mathbf{h}^{(t)}$  and the geometric one  $\mathbf{g}^{(\Sigma_t)}$  coincides so that, in particular, the speed of light is constantly  $c$  for all times and spatial directions in accordance with Einstein's synchronization procedure (see (1) in Remark 8.36).

In general the answer is negative. However if  $(M^n, \mathbf{g})$  is *stationary* with preferred timelike Killing vector  $K = T$ , the problem can be definitely discussed.

We tackle this problem only locally in a given *co-moving stationary coordinate system*

$$\phi : U \ni p \mapsto (x^0 = ct, x^1, \dots, x^{n-1}) \in \mathbb{R}^n \quad \text{with } \phi(U) = I \times V$$

where  $I \subset \mathbb{R}$  is an interval and  $V \subset \mathbb{R}^{n-1}$  an open connected set. We stress that the components of the metric  $g_{ab}$  in this coordinate system do not depend on  $t = x^0/c$  and we can therefore identify  $V \subset \mathbb{R}^{n-1}$  with a *common rest space* independent of  $t$  where perform our analysis. This space is also called the **quotient space** since it is one-to-one with the equivalence classes of the events belonging to the integral curves of  $T$ . This space of curves actually always exists and can be endowed with the structure of a smooth manifold. The important fact is that, when  $T$  is a Killing vector, we can also give the quotient space a physically meaningful metric as we are doing.

In this framework, the problem is whether or not there is a smooth map  $f = f(x^1, \dots, x^{n-1})$  such that the new time function<sup>14</sup>

$$t'(x^1, \dots, x^{n-1}) := t(x^1, \dots, x^{n-1}) + f(x^1, \dots, x^{n-1}) \quad (8.54)$$

satisfies, for some smooth function  $k : V \rightarrow \mathbb{R} \setminus \{0\}$ ,

$$dt' = kT^\flat. \quad (8.55)$$

In fact, that condition is equivalent to

$$dt'^\sharp = kT$$

and we know that  $dt'^\sharp$  is (up to rescaling) the unique vector normal to  $\Sigma_{t'}$  due to Proposition 5.26. In components (8.55) reads

$$\delta_a^0 + \sum_{\alpha=1}^{n-1} \frac{\partial f}{\partial x^\alpha} \delta_\alpha^a = kg_{a0}.$$

Hence,

$$k = (g_{00})^{-1} \quad \text{and} \quad \frac{\partial f}{\partial x^\alpha} = \frac{g_{\alpha 0}}{g_{00}}, \quad \alpha = 1, \dots, n-1.$$

Using the notation  $\vec{x} := (x^1, \dots, x^{n-1})$  and  $\vec{g} := (g_{01}, \dots, g_{0n-1})$  and where  $\cdot$  denotes the standard scalar product in  $\mathbb{R}^{n-1}$ , from elementary analysis, we conclude that a smooth function  $f$  satisfying (8.55) when inserted in (8.54) exists if and only if

$$\int_0^1 \frac{\vec{g}(\vec{\eta}(u))}{g_{00}(\vec{\eta}(u))} \cdot \frac{d\vec{\eta}}{du} du = 0 \quad (8.56)$$

for every smooth curve  $[0, 1] \ni u \mapsto \vec{\eta}(u) \in V$  such that  $\vec{\eta}(0) = \vec{\eta}(1)$ . In this case, the function  $f$  is defined as

$$f(\vec{x}) = \int_0^1 \frac{\vec{g}(\vec{\eta}(u))}{g_{00}(\vec{\eta}(u))} \cdot \frac{d\vec{\eta}}{du} du \quad (8.57)$$

where  $[0, 1] \ni u \mapsto \vec{\eta}(u) \in V$  is any smooth curve joining a fixed point  $\vec{\eta}(0) = \vec{x}_0 \in V$  to  $\vec{\eta}(1) = \vec{x}$ .

A necessary condition, which is also sufficient when  $V$  is *simply connected*, is the *irrotationality condition*

$$\frac{\partial}{\partial x^\alpha} \left( \frac{g_{\beta 0}}{g_{00}} \right) = \frac{\partial}{\partial x^\beta} \left( \frac{g_{\alpha 0}}{g_{00}} \right) \quad \text{everywhere in } V \text{ for } \alpha, \beta = 1, \dots, n-1. \quad (8.58)$$

We conclude that these facts are equivalents for a stationary reference frame (in the domain of a co-moving local chart),

(1) condition (8.56) holds;

---

<sup>14</sup>Observe that  $t'$  is a good time function since  $\langle dt', T \rangle = \langle dt, T \rangle + c \sum_{\alpha=1}^{n-1} \frac{\partial f}{\partial x^\alpha} \langle dx^\alpha, \frac{\partial}{\partial x^0} \rangle = \langle dt, T \rangle = 1$ .

- (2) it is possible to change the time function in order to obtain rest spaces orthogonal to  $T$  (so that the region of spacetime included in the domain of the coordinates turns out to be *static* with respect to the Killing vector  $T$ );
- (3) it is possible, changing the time function, to synchronize the proper time ideal clocks at rest with the reference frame in accordance with Einstein's synchronization procedure.

A necessary condition for the validity of (1)-(3), which is also sufficient if the spatial domain of the coordinates is simply connected is (8.58).

### 8.5.5 The metric of a rotating platform and the Sagnac Effect

As an example which had a crucial impact on the development of this subject though it was born in Special Relativity (by Langevin, Ehrenfest, Born and other authors<sup>15</sup>), we quickly examine the issue concerning the *reference frame of the rotating platform*. In Minkowski spacetime  $\mathbb{M}^4$ , consider an inertial reference frame and a co-moving Minkowskian system of coordinates  $(t, x, y, z)$ . Let us first re-write the metric in the local chart of (spatial) cylindric coordinates

$$\mathbf{g} = -cdt \otimes dt + dr \otimes dr + r^2 d\phi \otimes d\phi + dz \otimes dz, \quad (8.59)$$

where  $t \in \mathbb{R}$ ,  $r \in (0, +\infty)$ ,  $\phi \in (-\pi, \pi)$ , and  $z \in \mathbb{R}$ . Finally we pass to describe the metric within another local system of coordinates related to the cylindrical ones by the following transformations, for a constant  $\omega > 0$ ,

$$t' = t, \quad r' = r, \quad \phi' = \phi + \omega t, \quad z' = z$$

These coordinates are co-moving with a new non-inertial reference frame, called the reference frame of the **rotating platform** defined in the region  $r < c/\omega$  in the domain of the cylindrical coordinates. The time function of this new extended reference frame is  $t' = t$  and the time vector is

$$T = \frac{\partial}{\partial t'} = \frac{\partial}{\partial t} - \omega \frac{\partial}{\partial \phi}.$$

It is tangent to the worldlines of the physical particles forming the rotating platform.

Evidently  $T$  is a Killing vector since the coefficients of the metric in (8.59) do not depend on  $t$  and  $\phi$ . It is easy to see that  $\mathbf{g}(T, T) < 0$ ,  $\mathbf{g}(dt^\sharp, dt^\sharp) < 0$  provided if  $r < c/\omega$ , and that  $\langle T, dt \rangle = 1$ . By direct inspection one sees that all integral curves of  $T$  in the domain  $r < c/\omega$  meet every 3-surface at constant  $t$ .

In the new coordinates, which are by construction stationary and co-moving with the reference frame of the rotating platform, the metric reads

$$\mathbf{g} = (-c^2 + \omega^2 r'^2) dt' \otimes dt' + dr' \otimes dr' + r'^2 d\phi' \otimes d\phi' + dz' \otimes dz' - r'^2 \omega (d\phi' \otimes dt' + dt' \otimes d\phi'). \quad (8.60)$$

---

<sup>15</sup>See Rizzi, G.; Ruggiero, M. L. (2004). *Relativity in Rotating Frames*. Dordrecht: Kluwer.



It is clear that the coordinates are stationary but they do not satisfy condition (8.58) because

$$\frac{\partial}{\partial r'} \frac{g_{0\phi}}{g_{00}} - \frac{\partial}{\partial \phi'} \frac{g_{0r}}{g_{00}} = \frac{\partial}{\partial r'} \frac{r'^2 \omega}{\omega^2 r'^2 - c^2} \neq 0,$$

and thus there is no chance to choose a new time function  $t$  in order that the conditions (1)-(3) are satisfied. In particular the physical metric  $\mathbf{h}$  on the rest space of the reference frame does not coincide with the one induced by Minkowski metric (the standard Euclidean three-dimensional metric). Indeed we have

$$\mathbf{h} = dr' \otimes dr' + dz' \otimes dz' + \frac{r'^2}{1 - \frac{\omega^2 r'^2}{c^2}} d\phi' \otimes d\phi'. \quad (8.61)$$

Along the radial direction the metric coincides with the Euclidean one, however along the circles everything changes. Using this metric one easily sees that the length of a circle of radius  $R$  is

$$L_R := \frac{2\pi R}{\sqrt{1 - \frac{\omega^2 R^2}{c^2}}} \quad (8.62)$$

This fact alone implies that the geometry of the rest spaces is not Euclidean (so that there is no way to obtain the metric in constant diagonal form by a suitable change of coordinates).

The instantaneous speed of light has the usual value  $c$  along the radial direction, instead along the circles of radius  $r'$ , we find

$$c_{\pm} = \frac{c}{1 \pm \frac{r'\omega}{c}},$$

according to the direction, where we have used (8.51). The analogous velocity measured with respect to the time  $t'$  is

$$c'_{\pm} = \sqrt{-g_{00}} \frac{1}{1 \pm \frac{r'\omega}{c}} = c \frac{\sqrt{1 - \frac{r'^2 \omega^2}{c^2}}}{1 \pm \frac{r'\omega}{c}}.$$

This means that when a particle of light travels along a complete circle of radius  $R$ , it takes a period of time

$$\Delta t_{\pm} = \frac{L_R}{c'_{\pm}} = \frac{2\pi R}{c} \left( 1 \pm \frac{R\omega}{c} \right).$$

Then a difference of the two interval of total time exists which depends on the direction along the circle of radius  $R$  the particle of light takes. For instance, we can assume to use a system of mirrors or a *fiber-optic cable* fixed to a circle of the platform. This difference can be measured in experiments and it is the well known **Sagnac effect** which was in the past the cause of a discussion about the validity of Special Relativity.

**Remark 8.38.** Actually in experiments one measures corresponding intervals  $\Delta\tau_{\pm}$  of *proper time* of a clock fixed to the platform in a point of the circuit. These intervals are however equal to  $\Delta t_{\pm}$  up to a common factor

$$\frac{1}{c} \sqrt{-g_{00}} = \sqrt{1 - \frac{R^2 \omega^2}{c^2}},$$

so that the  $\Delta\tau_+ - \Delta\tau_- \neq 0$ . ■

## 8.6 Fermi-Walker's transport in Lorentzian manifolds.

To conclude this chapter we discuss the technical notion of *Fermi-Walker's transport* or *non-rotating transport* along a worldline which has some applications in Special and General Relativity.

### 8.6.1 The equations of Fermi-Walker's transport

If  $(M^n, \mathbf{g})$  is a spacetime, consider a timelike smooth curve  $\gamma : (a, b) \rightarrow M^n$  parametrized by its proper time  $t$  and assume that there is a smooth vector field  $X$  defined on  $\gamma$  according to Definition 6.16. For the moment we also suppose that  $X(t) \in \Sigma_{\gamma(t)}$ ,  $\Sigma_{\gamma(t)}$  as usual denoting the rest space of the observer associated to  $\gamma$ , i.e., subspace of  $T_{\gamma(t)}M^n$  made of the vectors  $u$  with  $\mathbf{g}(u, \dot{\gamma}(t)) = 0$ .

**Remark 8.39.** For instance  $X$  could be the *spin* of a particle whose world line is  $\gamma$  itself when  $n = 4$ . ■

We want to formalize the idea of a vector  $X$  which both

- (a) does not rotate,
- (b) preserves metrical properties in  $\Sigma_{\gamma(t)}$

during its evolution along the worldline. Let us examine conditions (a) and (b) separately.

- (a) As  $T_{\gamma(t)}M^n$  is orthogonally decomposed as  $\text{Span}(\dot{\gamma}(t)) \oplus d\Sigma_{\gamma(t)}$ , the only possible infinitesimal deformations of  $X_{\gamma(t)}$  during an infinitesimal interval of time  $t$  must take place in the linear space  $\text{Span}(\dot{\gamma}(t))$  spanned by  $\dot{\gamma}$ . If  $X(t)$  does not satisfy  $X_{\gamma(t)} \in \Sigma_{\gamma(t)}$ , a direct generalization of the said condition is that the orthogonal projection of  $X_{\gamma(t)}$  onto  $d\Sigma_{\gamma(t)}$  involves deformations along  $\dot{\gamma}$  only during its evolution.
- (b) The condition about the preservation of metrical structures means that  $\mathbf{g}(X(t), X(t))$  is constant.

**Remark 8.40.**

- (1) Notice that  $\dot{\gamma}$  naturally satisfies both requirements.
- (2) The non-rotating and metric preserving conditions can be generalized to a generic set of vectors  $\{X_{(a)}(t)\}_{a \in A}$ . Here, condition (a) is formulated exactly as above for each vector separately, while condition (b) property means that the scalar products  $\mathbf{g}(X_{(a)}(t), X_{(b)}(t))$ , with  $a, b \in A$ , are preserved during evolution along the line for  $t \in (a, b)$ . ■

Our goal is to find a set of equations which are equivalent to the pair of conditions (a) and (b) also relaxing the condition  $X(t) \in d\Sigma_{\gamma(t)}$ . In this case, condition (a) is imposed only on the orthogonal projection

$$X(t) + \mathbf{g}(X(t), \dot{\gamma}(t)) \dot{\gamma}(t) \in d\Sigma_{\gamma(t)}$$

of  $X(t)$  onto  $\Sigma_{\gamma(t)}$ . Condition (a) therefore reads

$$\nabla_{\dot{\gamma}} [X(t) + \mathbf{g}(X(t), \dot{\gamma}(t)) \dot{\gamma}(t)] = \alpha(t) \dot{\gamma}(t), \quad (8.63)$$

for some suitable smooth scalar function  $\alpha$  we want to determine. Expanding (8.63). we find

$$\begin{aligned} \nabla_{\dot{\gamma}} X(t) + \mathbf{g}(\nabla_{\dot{\gamma}} X(t), \dot{\gamma}(t)) \dot{\gamma}(t) + \mathbf{g}(X(t), \nabla_{\dot{\gamma}} \dot{\gamma}(t)) \dot{\gamma}(t) + \mathbf{g}(X(t), \dot{\gamma}(t)) \nabla_{\dot{\gamma}} \dot{\gamma}(t) \\ = \alpha(t) \dot{\gamma}(t). \end{aligned} \quad (8.64)$$

Taking the scalar product with  $\dot{\gamma}(t)$  and using  $\mathbf{g}(\dot{\gamma}(t), \dot{\gamma}(t)) = -1$  we obtain

$$\mathbf{g}(\nabla_{\dot{\gamma}} X(t), \dot{\gamma}(t)) - \mathbf{g}(\nabla_{\dot{\gamma}} X(t), \dot{\gamma}(t)) - \mathbf{g}(X(t), \nabla_{\dot{\gamma}} \dot{\gamma}(t)) = -\alpha(t) \quad (8.65)$$

so that

$$\mathbf{g}(X(t), \nabla_{\dot{\gamma}} \dot{\gamma}(t)) = \alpha(t).$$

That identity, inserted in the right-hand side of (8.63), produces a definite formulation of condition (a),

$$\nabla_{\dot{\gamma}} [X(t) + \mathbf{g}(X(t), \dot{\gamma}(t)) \dot{\gamma}(t)] = \mathbf{g}(X(t), \nabla_{\dot{\gamma}} \dot{\gamma}(t)) \dot{\gamma}(t),$$

namely

$$\nabla_{\dot{\gamma}} X(t) + \nabla_{\dot{\gamma}} [\mathbf{g}(X(t), \dot{\gamma}(t)) \dot{\gamma}(t)] - \mathbf{g}(X(t), \nabla_{\dot{\gamma}} \dot{\gamma}(t)) \dot{\gamma}(t) = 0.$$

In summary, our final equation corresponding to the requirment (a) is

$$\nabla_{\dot{\gamma}} X(t) + \mathbf{g}(X(t), \dot{\gamma}(t)) \nabla_{\dot{\gamma}} \dot{\gamma}(t) + \mathbf{g}(\nabla_{\dot{\gamma}} X(t), \dot{\gamma}(t)) \dot{\gamma}(t) = 0. \quad (8.66)$$

Let us finally impose the metric preserving property (b). To this end we modify the previous equation as

$$\nabla_{\dot{\gamma}} X(t) + \mathbf{g}(X(t), \dot{\gamma}(t)) \nabla_{\dot{\gamma}} \dot{\gamma}(t) - \mathbf{g}(X(t), \nabla_{\dot{\gamma}} \dot{\gamma}(t)) \dot{\gamma}(t) + \left[ \frac{d}{dt} \mathbf{g}(X(t), \dot{\gamma}(t)) \right] \dot{\gamma}(t) = 0. \quad (8.67)$$

Now observe that, if both  $\dot{\gamma}$  and  $X$  satisfy the metric preserving property, we are committed to also assume that

$$\frac{d}{dt} \mathbf{g}(X(t), \dot{\gamma}(t)) = 0. \quad (8.68)$$

As a consequence, (8.67) specializes to

$$\nabla_{\dot{\gamma}} X(t) + \mathbf{g}(X(t), \dot{\gamma}(t)) \nabla_{\dot{\gamma}} \dot{\gamma}(t) - \mathbf{g}(X(t), \nabla_{\dot{\gamma}} \dot{\gamma}(t)) \dot{\gamma}(t) = 0. \quad (8.69)$$

We have found that if  $X$  satisfies both the non-rotating condition (a) and the metric preserving condition (b), then it satisfies (8.69). Let us discuss if that equation implies conditions (a) and (b). If the vectors of a family  $\{X_{(a)}(t)\}_{a \in A}$  satisfy (8.69), their scalar products along  $\gamma$  are constant as shown below so that they agree with the condition (b). Moreover, observing that  $\dot{\gamma}$  itself fulfils (8.69), we also have that (8.68) holds true for every  $X_{(a)}(t)$  and therefore they also satisfy the equivalent condition (8.66) which is the mathematical statement of condition (a). We conclude that (8.69) is equivalent to the couple of requirements (a) and (b). (8.69) is the wanted equation.

**Definition 8.41.** (**Fermi-Walker's Transport of a vector along a curve.**) Let  $(M^n, \mathbf{g})$  be a Lorentzian manifold and  $\gamma : (a, b) \rightarrow M^n$  a smooth timelike curve where  $t$  is the length parameter (i.e., the proper time). A smooth vector field  $X$  on  $\gamma$  is said to be **Fermi-Walker transported** along  $\gamma$  if

$$D_{\dot{\gamma}}^{(F)} X(t) = 0$$

for all  $t \in (a, b)$ , where we have introduced the **Fermi-Walker derivative** along  $\gamma$

$$D_{\dot{\gamma}}^{(F)} X(t) := \nabla_{\dot{\gamma}} X(t) + \mathbf{g}(X(t), \dot{\gamma}(t)) \nabla_{\dot{\gamma}} \dot{\gamma}(t) - (X(t), \nabla_{\dot{\gamma}} \dot{\gamma}(t)) \dot{\gamma}(t). \quad (8.70)$$

■

**Proposition 8.42.** *The notion of Fermi-Walker's transport along a curve  $\gamma$  as defined in Definition 8.41 enjoys the following properties.*

- (1) *It is **metric preserving**, i.e., if  $(a, b) \ni t \mapsto X(t)$  and  $(a, b) \ni t \mapsto X'(t)$  are Fermi-Walker transported along  $\gamma$ , then*

$$(a, b) \ni t \mapsto \mathbf{g}(X(t), X'(t))$$

*is constant.*

- (2)  *$(a, b) \ni t \mapsto \dot{\gamma}(t)$  is Fermi-Walker transported along  $\gamma$ .*

- (3) *If  $\gamma$  is a geodesic with respect to the Levi-Civita connection, then the notions of parallel transport and Fermi-Walker's transport along  $\gamma$  coincide.*

**Proof.** (1) Using the fact that the connection is metric one has:

$$\frac{d}{dt} \mathbf{g}(X(t), X'(t)) = \mathbf{g}(\nabla_{\dot{\gamma}} X(t), X'(t)) + \mathbf{g}(X(t), \nabla_{\dot{\gamma}} X'(t)). \quad (8.71)$$

Making use of the equation of Fermi-Walker's transport,

$$\nabla_{\dot{\gamma}} U(t) = -\mathbf{g}(U(t), \dot{\gamma}(t)) \nabla_{\dot{\gamma}} \dot{\gamma}(t) + \mathbf{g}(U(t), \nabla_{\dot{\gamma}} \dot{\gamma}(t)) \dot{\gamma}(t),$$

for both  $X$  and  $X'$  in place of  $U$ , the terms in the right-hand side of (8.71) cancel out each other. The proof of (2) is direct by noticing that

$$\mathbf{g}(\dot{\gamma}(t), \dot{\gamma}(t)) = -1$$

and

$$\mathbf{g}(\dot{\gamma}(t), \nabla_{\dot{\gamma}} \dot{\gamma}(t)) = \frac{1}{2} \frac{d}{dt} \mathbf{g}(\dot{\gamma}(t), \dot{\gamma}(t)) = -\frac{1}{2} \frac{d}{dt} 1 = 0.$$

The proof of (3) is trivial noticing that if  $\gamma$  is a geodesic, then  $\nabla_{\dot{\gamma}} \dot{\gamma}(t) = 0$  and (8.71) specializes to the equation of the parallel transport

$$\nabla_{\dot{\gamma}} U(t) = 0.$$

□

**Remark 8.43.**

(1) If  $\gamma : (a, b) \rightarrow M$  is fixed, the Fermi-Walker's transport condition

$$\nabla_{\dot{\gamma}} V(t) = \mathbf{g}(V(t), \nabla_{\dot{\gamma}} \dot{\gamma}(t)) \dot{\gamma}(t) - \mathbf{g}(V(t), \dot{\gamma}(t)) \nabla_{\dot{\gamma}} \dot{\gamma}(t)$$

can be used as a differential equation. Expanding both sides in local coordinates  $(x^1, \dots, x^n)$  one finds a first-order differential equation for the components of  $V$  referred to the bases of elements  $\frac{\partial}{\partial x^k}|_{\gamma(t)}$ . As the equation is in normal form and linear where everything known is smooth, the initial vector  $V(t_0)$ , where  $t_0 \in (a, b)$ , determines  $V$  uniquely along the whole curve (see the analogous comment for the case of the parallel transport). In a certain sense, one may view the solution  $(a, b)t \mapsto V(t)$  as the “transport” and “evolution” of the initial condition  $V(t_0)$  along  $\gamma$  itself.

The global existence and uniqueness theorem has an important consequence. If  $\gamma : [a, b] \rightarrow M^n$  is fixed and  $u, v \in (a, b)$  with  $u \neq v$ , the notion of parallel transport along  $\gamma$  produces an vector space isomorphism  $\mathcal{F}_\gamma[u, v] : T_{\gamma(u)}M^n \rightarrow T_{\gamma(v)}M^n$  which associates  $V \in T_{\gamma(u)}M^n$  with that vector in  $T_{\gamma(v)}M^n$  which is obtained by Fermi-Walker transporting  $V$  in  $T_{\gamma(u)}M^n$ . Notice that  $\mathcal{F}_\gamma[u, v]$  also preserves the scalar product by property (1) of proposition 8.42, i.e., it is an isometric isomorphism.

(2) The equation of Fermi-Walker's transport of a vector  $X$  in a  $n$ -dimensional spacetime  $(M^n, \mathbf{g})$  can be re-written

$$\frac{dX^a(t)}{dt} = X^b(t)[A_b(t)V^a(t) - A^a(t)V_b(t)], \quad (8.72)$$

where we have indicated the  $n$ -velocity by  $V(t) = \dot{\gamma}(t)$  and we have introduced the  $n$ -acceleration  $A(t) := \nabla_{\dot{\gamma}} \dot{\gamma}(t)$  of a worldline  $\gamma$  parametrized by the proper time  $t$ . Notice that  $\mathbf{g}(A(t), V(t)) = 0$  for all  $t$  and thus if  $A \neq 0$ , it turns out to be *spacelike* because  $V$  is timelike by definition. ■

**Examples 8.44.** The famous *Bargmann–Michel–Telegdi* (BMT) equation describes the *spin precession* of an electron in an external electromagnetic field  $F$  in Special Relativity. It reads (for  $c = 1$ )

$$\frac{dS^a}{d\tau} = \frac{e}{m} V^a V_b F^{bc} S_c + 2\mu(F^{ad} - V^a V_c F^{cd}) S_d. \quad (8.73)$$

Above the derivative is computed along the worldline  $\gamma$  of the electron whose 4-velocity is  $V$  and  $S$  is the *polarization vector*, describing the spin vector of the electron. It satisfies  $\mathbf{g}(V, S) = 0$  so that it is always spacelike and stays in the rest space of the electron, and there it is normalized to 1. Since this property is invariant changing the reference frame, we have  $\mathbf{g}(S, S) = 1$ . The coefficients  $e$ ,  $m$ , and  $\mu$  respectively denote the *charge*, the *mass*, and the *magnetic moment* of the electron.

The written equation can be rephrased in terms of the Fermi-Walker derivative. To this end, observe that the 4-velocity satisfies the equation of motion according to the *Lorentz force* (see [Mor20])

$$\frac{dV^a}{d\tau} = \frac{e}{m} F^{ab} V_b .$$

Inserting this identity in the right-hand side of (8.73), we find

$$\frac{dS^a}{d\tau} = V^a A^c S_c + 2\mu(F^{ad} - V^a V_c F^{cd}) S_d ,$$

where we have used notations as in (8.72) so that  $A$  denotes the 4-acceleration of the electron. Now observe that

$$A^a V^c S_c = 0$$

by definition of  $S$ . Therefore the equation above can be re-written as

$$\frac{dS^a}{d\tau} = S^c [A_c V^a - A^a V_c] + 2\mu(F^{ad} - V^a V_c F^{cd}) S_d .$$

In other words, as expected since it deals with the phenomenon of precession of the spin, the BMT equation says *how the spin fails to be transported without rotating along the worldline of the electron*. This failure is due to the presence of an external electromagnetic field:

$$D_\gamma^{(F)} S^a = 2\mu(F^{ad} - V^a V_c F^{cd}) S_d . \quad (8.74)$$

### 8.6.2 Fermi-Walker's transport and Lorentz group

The non-rotating property of Fermi-Walker's transport can be viewed from another point of view when  $M^n$  is the four-dimensional Minkowski spacetime  $\mathbb{M}^4$ . Consider the *orthochronous proper Lorentz group*  $SO(1, 3)^\uparrow$  [Mor20] represented by real  $4 \times 4$  matrices  $\Lambda : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ , with  $\Lambda = [\Lambda_j^i]$ ,  $i, j = 0, 1, 2, 3$ . Here the coordinate  $x^0$  represents the *time coordinate* and the remaining three coordinates are the space coordinates in Minkowski spacetime. These sets of coordinates are the natural coordinates of inertial frames and  $SO(1, 3)^\uparrow$  (with the addition of spacetime translations) describes the possible coordinate transformation between inertial reference frames. It is known that every  $\Lambda \in SO(1, 3)^\uparrow$  can uniquely be decomposed as

$$\Lambda = \Omega P ,$$

where  $\Omega, P \in SO(1, 3)^\uparrow$  are respectively a rotation of  $SO(3)$  of the spatial coordinates which does not affect the time coordinate, and a *pure Lorentz transformation*. In this sense, every pure

Lorentz transformation does not contain rotations and represents the coordinate transformation between a pair of pseudo-orthonormal reference frames in Minkowski spacetime [Mor20] which do not involve rotations in their reciprocal position.

Every pure Lorentz transformation can uniquely be represented as

$$P = e^{\sum_{i=1}^3 A_i K_i} ,$$

where  $(A_1, A_2, A_3) \in \mathbb{R}^3$  and  $K_1, K_2, K_3$  are matrices in the Lie algebra of  $SO(1, 3)$ ,  $so(1, 3)$ , called *boosts*. The elements of the boosts  $K_a = [K_{(a)}^i_j]$  are

$$K_{(a)}^0_j = K_{(a)}^i_0 = \delta_{ai} \quad \text{and} \quad K_{(a)}^i_j = 0 \quad \text{in all remaining cases.}$$

We have the expansion in the metric topology of  $\mathbb{R}^{16}$

$$P = e^{h \sum_{i=1}^3 A_i K_i} = \sum_{n=0}^{\infty} \frac{h^n}{n!} \left( \sum_{i=1}^3 A_i K_i \right)^n ,$$

and thus

$$P = I + h \sum_{i=1}^3 A_i K_i + hO(h) ,$$

where  $O(h) \rightarrow 0$  as  $h \rightarrow 0$ . The matrices of the form

$$I + h \sum_{i=1}^3 A_i K_i .$$

with  $h \in \mathbb{R}$  and  $(A_1, A_2, A_3) \in \mathbb{R}^3$  (notice that  $h$  can be re-absorbed in the coefficients  $A_i$ ) are called *infinitesimal pure Lorentz transformations*.

Now, consider a differentiable timelike curve  $\gamma : [0, \epsilon) \rightarrow \mathbb{M}^4$ , starting from  $p$  in a four dimensional Lorentzian manifold  $\mathbb{M}^4$ , and fix a pseudo-orthonormal basis in  $T_p \mathbb{M}^4$ ,  $e_0, e_1, e_2, e_3$  with  $e_1 = \dot{\gamma}(0)$ . We are assuming that the parameter  $t$  of the curve is the proper time. Consider the evolutions of  $e_i$ ,  $t \mapsto e_i(t)$ , obtained by using Fermi-Walker's transport along  $\gamma$ . We want to investigate the following issue.

*What is the Lorentz transformation which relates the basis  $\{e_i(t)\}_{i=0,\dots,3}$  with the basis of Fermi-Walker transported elements  $\{e_i(t+h)\}_{i=0,\dots,3}$  in the limit  $h \rightarrow 0$ ?*

In fact, we want to show that the considered transformation is an infinitesimal pure Lorentz transformation and, in this sense, it does not involves rotations.

To compare the basis  $\{e_i(t)\}_{i=0,\dots,3}$  with the basis  $\{e_i(t+h)\}_{i=0,\dots,3}$  we have to transport, by means of parallel transport, the latter basis in  $\gamma(t)$ . In other words, we intend to find the Lorentz transformation between  $\{e_i(t)\}_{i=0,\dots,3}$  and  $\{\mathcal{P}_\alpha^{-1}[\gamma(t), \gamma(t+h)]e_i(t+h)\}_{i=0,\dots,3}$ ,  $\alpha$  being the geodesic joining  $\gamma(t)$  and  $\gamma(t+h)$  for  $h$  small sufficiently. We define

$$e'_i(t+h) := \mathcal{P}_\alpha^{-1}[\gamma(t), \gamma(t+h)]e_i(t+h) .$$

**Remark 8.45.** The notation  $\mathcal{P}_\alpha^{-1}[\gamma(t), \gamma(t+h)]$  stresses that the parallel transport is performed along  $\alpha$ , but the joinend points  $\gamma(t), \gamma(t+h)$  (also) belong to the other curve  $\gamma$ . ■

We have

$$e'_i(t+h) - e_i(t) = h\nabla_{\dot{\gamma}} e_i(t) + hO(h) .$$

Using the equation of Fermi-Walker's transport we get

$$e'_i(t+h) - e_i(t) = h\mathbf{g}(e_i(t), A(t))e_0(t) - h\mathbf{g}(e_i(t), e_0(t))A(t) + hO(h) , \quad (8.75)$$

where  $A(t) = \nabla_{\dot{\gamma}} \dot{\gamma}(t)$  is the 4-acceleration of the worldline  $\gamma$  itself and  $O(h) \rightarrow 0$  as  $h \rightarrow 0$ . Notice that  $\mathbf{g}(A(t), e_0(t)) = 0$  by the (2) in Remarks 8.43 and thus

$$A(t) = \sum_{i=1}^3 A_i(t)e_i(t) , \quad (8.76)$$

for some triple of functions  $A_1, A_2, A_3$ . If  $\eta_{ab} = \text{diag}(-1, 1, 1, 1)$  and taking (8.76) and the pseudo orthonormality of the basis  $\{e_i(t)\}_{i=0,\dots,3}$  into account, (8.75) can be re-written

$$e'_i(t+h) = e_i(t) + h(A_i(t)e_0(t) - \eta_{i0}A(t)) + hO(h) . \quad (8.77)$$

If we expand  $e'_i(t+h)$  in components referred to the basis  $\{e_i(t)\}_{i=0,\dots,3}$ , (8.77) becomes

$$(e'_i(t+h))^j = \delta_i^j + h(A_i(t)\delta_0^j(t) - \eta_{i0}A_j(t)) + hO^j(h) , \quad (8.78)$$

where one should remind that  $A_0 = 0$ . As  $(e_i(t))^j = \delta_i^j$ , (8.78) can be re-written

$$e'_i(t+h) = I + h \left( \sum_{j=1}^3 A_j K_j \right) e_i(t) + hO(h) . \quad (8.79)$$

We have found that the infinitesimal transformation which connect the two bases is, in fact, an infinitesimal pure Lorentz transformation. Notice that this transformation depends on the 4-acceleration  $A$  and reduces to the identity (up to terms  $hO(h)$ ) if  $A = 0$ , i.e., if the curve is a timelike geodesic.



## Chapter 9

# Curvature

This chapter introduces the basic notion of *curvature tensor* for affine and metric connections, on the other hand it discusses how these notions play a central role in the physical development of General Relativity which we shall see in the next chapter.

### 9.1 Curvature tensors of affine and metric connections

According to Definition 5.7 a (pseudo) Riemannian manifold  $(M, \mathbf{g})$  is *locally flat* if it admits an atlas made of local *canonical charts*  $(U, \psi)$ . By definition, the coordinate representation of the metric  $\mathbf{g}$  metric in those coordinates has the constant form  $g_{ij} = \pm \delta_{ij}$ , with the sign determined by the signature of the metric.

In general the mentioned atlas does not contain *global canonical charts*, so that the manifold is not *globally flat*. This is the case, for instance of a 2-dimensional cylinder  $C \subset \mathbb{R}^3$ ,  $x^2 + y^2 = 1$ , equipped with the metric induced by the standard metric of  $\mathbb{R}^2$ : it is locally flat but not globally flat differently from  $\mathbb{R}^2$  itself.

Since the metric tensor is constant in canonical coordinates defined in a neighborhood  $U$  of any  $p \in M$ , the Levi-Civita connection is represented by trivial connection coefficients in those coordinates:  $\Gamma_{ij}^k = 0$ . As a consequence, in those coordinates,

$$\nabla_i \nabla_j Z^k = \frac{\partial^2 Z^k}{\partial x^i \partial x^j} = \frac{\partial^2 Z^k}{\partial x^j \partial x^i} = \nabla_j \nabla_i Z^k ,$$

for every  $Z \in \mathfrak{X}(M)$ . In other words, the covariant derivatives of vector fields commute

$$\nabla_i \nabla_j Z^k = \nabla_j \nabla_i Z^k .$$

Due to the intrinsic nature of covariant derivatives, that identity holds in any coordinate system. We have therefore established that *the local flatness condition of  $(M, \mathbf{g})$  implies commutativity of (Levi-Civita) covariant derivatives on vector fields on  $M$* . This fact completely characterizes locally flat (pseudo) Riemannian manifolds because the converse implication also holds, as we

shall prove in this chapter. Therefore, a (pseudo) Riemannian manifold can be considered “curved” whenever commutativity of (Levi-Civita) covariant derivatives fails. This property can be used to extend the notion of curved manifold to manifolds which are equipped with non-metric connections  $\nabla$  as we shall see shortly.

Some quantitative notion is necessary to measure the failure of commutativity of the covariant derivatives on vector fields. This notion is provided by the *curvature tensor* (field)  $R$  associated to  $\nabla$ . In this way, commutativity of the covariant derivatives in  $M$  is equivalent to the requirement that  $R = 0$  everywhere in  $M$ .

In the special case of the the Levi-Civita connection, it is possible to prove a stronger remarkable result: the condition  $R = 0$  everywhere is equivalent to the local flatness of the manifold, in the sense of the existence of an atlas made of charts where the metric takes its constant canonical diagonal form.

### 9.1.1 Curvature tensor and Riemann curvature tensor

To introduce the curvature tensor, let us discuss in more detail the commutativity property of covariant derivative acting on tensor fields using an intrinsic approach.

**Lemma 9.1.** *Let  $M$  be a smooth manifold endowed with a torsion-free affine connection  $\nabla$ . The following facts are equivalent.*

(a) *Covariant derivatives of contravariant vector fields commute*

$$\nabla_i \nabla_j Z^k = \nabla_j \nabla_i Z^k \quad (9.1)$$

*in every chart on  $M$  and all coordinate indices  $i, j, k$ , for every  $Z \in \mathfrak{X}(M)$ .*

(b) *The following identity holds everywhere on  $M$*

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z = 0, \quad (9.2)$$

*for all  $X, Y, Z \in \mathfrak{X}(M)$ .*

**Proof.** If  $X, Y \in \mathfrak{X}(M)$ , (9.1) entails

$$X^i Y^j \nabla_i \nabla_j Z^k = X^i Y^j \nabla_j \nabla_i Z^k,$$

which can be re-written,

$$X^i \nabla_i (Y^j \nabla_j Z^k) - X^i (\nabla_i Y^j) \nabla_j Z^k = Y^j \nabla_j (X^i \nabla_i Z^k) - Y^j (\nabla_j X^i) \nabla_i Z^k,$$

or

$$X^i \nabla_i (Y^j \nabla_j Z^k) - Y^j \nabla_j (X^i \nabla_i Z^k) - (X^i (\nabla_i Y^j) \nabla_j Z^k - Y^j (\nabla_j X^i) \nabla_i Z^k) = 0,$$

and finally

$$X^i \nabla_i (Y^j \nabla_j Z^k) - Y^j \nabla_j (X^i \nabla_i Z^k) - (X^i (\nabla_i Y^j) - Y^i (\nabla_i X^j)) \nabla_j Z^k = 0 .$$

Proposition 6.9, which relies on the hypothesis that  $\nabla$  is torsion-free, permits to recast the found identity into the intrinsic form

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z = 0 , \quad \forall X, Y, Z \in \mathfrak{X}(M) .$$

In turn, (9.2) implies (9.1) when specialized to  $X = \frac{\partial}{\partial x^i}$  and  $Y = \frac{\partial}{\partial x^j}$  (smoothly extended outside the domain of the coordinate patch) also using  $[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] = 0$ .  $\square$

**Remark 9.2.** In case  $\nabla$  has a torsion  $T$ , we can reformulate condition (9.2) in order to remain equivalent to condition (9.1) making explicit use of the torsion tensor  $T$ . In that case (9.3) below would be replaced by

$$R_p(X_p, Y_p, Z_p) = (\nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X,Y]} Z)_p + T(X_p, Y_p) Z_p .$$

However, we are not interested in this extension so that we leave all details to the interested reader.  $\blacksquare$

We are in a position to introduce a tensor field which is the quantitative measure of the failure of condition (9.2).

**Proposition 9.3.** *Let  $M$  be a smooth manifold equipped with an affine torsion-free connection  $\nabla$ . The following facts are true.*

- (a) *There is a unique smooth tensor field  $R$  such that  $R_p \in T_p^* M \otimes T_p^* M \otimes T_p^* M \otimes T_p M$  if  $p \in M$  and*

$$R_p(X_p, Y_p, Z_p) = (\nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X,Y]} Z)_p , \quad (9.3)$$

*if  $X, Y, Z \in \mathfrak{X}(M)$ .*

- (b) *In local coordinates, we have*

$$R_p(X_p, Y_p, Z_p)^l = (R_p)_{ijk}{}^l X_p^i Y_p^j Z_p^k , \quad (9.4)$$

*where*

$$R_{ijk}{}^l = \frac{\partial \Gamma_{ik}^l}{\partial x^j} - \frac{\partial \Gamma_{jk}^l}{\partial x^i} + \Gamma_{ik}^r \Gamma_{jr}^l - \Gamma_{jk}^r \Gamma_{ir}^l , \quad (9.5)$$

*and*

$$(R_p)_{ijk}{}^l := \left\langle R_p \left( \frac{\partial}{\partial x^i} |_p, \frac{\partial}{\partial x^j} |_p, \frac{\partial}{\partial x^k} |_p \right), dx_p^l \right\rangle .$$

**Proof.** By direct inspection one finds that, if  $R_{ijk}{}^l$  is defined as in the right-hand side of (9.5), then

$$(\nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X,Y]} Z)_p^l = R_{ijk}{}^l X_p^i Y_p^j Z_p^k, \quad (9.6)$$

because, in particular, all derivatives of the fields  $X, Y, Z$  cancel each other in view of Schwarz' theorem on crossed derivatives. Such an identity proves, in particular, that the map

$$\mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) \ni (X, Y, Z) \mapsto (\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z)_p,$$

depends only on the values attained at  $p$  by  $X, Y, Z$  and not on the values these vector fields take around this point. Since the map is multi  $\mathbb{R}$ -linear, it defines a tensor

$$R_p \in T_p^* M \otimes T_p^* M \otimes T_p^* M \otimes T_p M.$$

By construction, the components of  $R_p$  are given in (9.5) and (9.3) holds. Finally, by construction,  $R_p$  is smooth when varying  $p \in M$ .  $\square$

**Definition 9.4.** (**Curvature tensor and Riemann curvature tensor.**) *Let  $M$  be a smooth manifold equipped with an affine torsion-free connection  $\nabla$ .*

- (a) *The smooth tensor field  $R$  defined in Proposition 9.3 is called **curvature tensor (field) of  $\nabla$** .*
- (b) *If  $\nabla$  is the Levi-Civita connection obtained by a metric  $\mathbf{g}$ ,  $R$  is called **Riemann curvature tensor (field) of  $\mathbf{g}$** .*  $\blacksquare$

**Notation 9.5.** The following standard notation is exploited.

- (1)  $R(X, Y, Z)$  indicates the vector field that restricts to the vector  $R_p(X_p, Y_p, Z_p)$  at every  $p \in M$ .
- (2)  $R(X, Y) : \mathfrak{X}(M) \rightarrow \mathfrak{X}(M)$  is the differential operator

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

so that  $R(X, Y)Z = R(X, Y, Z)$ .  $\blacksquare$

**Exercises 9.6.**

- 1. Prove that

$$\nabla_i \nabla_j \omega_k - \nabla_j \nabla_i \omega_k = R_{ijk}{}^l \omega_l. \quad (9.7)$$

- 2. Prove that, in the general case, **Ricci's identity** holds:

$$\nabla_i \nabla_j \Xi^{i_1 \dots i_p}_{j_1 \dots j_q} - \nabla_j \nabla_i \Xi^{i_1 \dots i_p}_{j_1 \dots j_q} = - \sum_{u=1}^p R_{ijs}{}^u \Xi^{i_1 \dots s \dots i_p}_{j_1 \dots j_q} + \sum_{u=1}^p R_{ijj_u}{}^s \Xi^{i_1 \dots i_p}_{j_1 \dots s \dots j_q}.$$

To go on, we extend the notion of flatness to affine torsion-free and generally non-metric connections.

**Definition 9.7.** (Locally flat affine connection.) Let  $M$  be a smooth manifold equipped with an affine torsion-free connection  $\nabla$ .  $M$  and  $\nabla$  are said **locally flat** if there is an atlas of  $M$  where the connection coefficients vanish everywhere. ■

**Remark 9.8.** The Levi-Civita connection of a locally flat (pseudo) Riemannian manifold is evidently locally flat. ■.

To conclude, we state a proposition that summarize all that we found adding some further minor result. This statement will be completed later into a more general proposition.

**Proposition 9.9.** *Let  $M$  be a smooth manifold equipped with a torsion-free affine connection  $\nabla$ . The following facts are valid.*

(1) *The following statements are equivalent.*

(a) *Covariant derivatives of smooth tensor fields  $\Xi$  commute, i.e.,*

$$\nabla_i \nabla_j \Xi^A = \nabla_j \nabla_i \Xi^A ,$$

*in every local chart;*

(b) *covariant derivatives of smooth contravariant vector fields  $X$  commute;*

(c) *covariant derivatives of smooth covariant vector fields  $\omega$  commute;*

(d) *the curvature tensor associated with  $\nabla$  vanishes everywhere in  $M$ .*

(2) *If  $M$  is locally flat, then the equivalent facts (a)-(d) hold.*

(3) *If  $(M, \mathbf{g})$  is a locally flat (pseudo) Riemannian manifold and  $\nabla$  is the Levi-Civita connection, then the equivalent facts (a)-(d) hold.*

**Proof.** It is clear that (a)  $\Rightarrow$  (b) and also (b)  $\Rightarrow$  (d) in view of Lemma 9.1 and (9.6). Furthermore, (d)  $\Rightarrow$  (c). Next (c)  $\Rightarrow$  (d) in view of (1) in Exercises 9.6. Finally, (d)  $\Rightarrow$  (a) due to (2) in Exercises 9.6. The next statement in the thesis is consequence of (9.5): in the coordinates of the atlas  $R_{jkl}^i$  must vanish and, since they define a tensor, they vanish in every coordinate system, i.e.,  $R = 0$  in  $M$ . If  $\nabla$  is the Levi-Civita connection, local flatness implies there is an atlas where the coefficients of the metric are constant and thus the Levi-Civita connection coefficients vanish and one reduces to the previously considered case. □

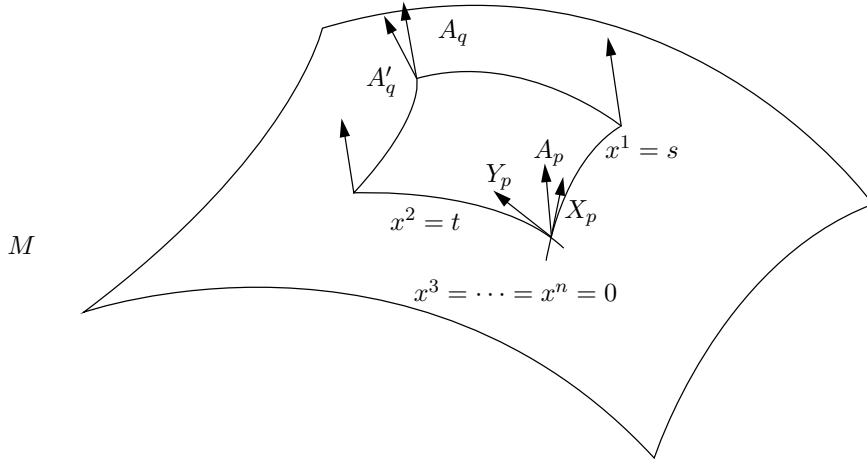


Figure 9.1: Difference of parallelly transported vectors

### 9.1.2 A geometric meaning of the curvature tensor

We want to spend some words about an interesting geometric meaning of the curvature tensor.

Consider a smooth  $M$  equipped with a torsion-free connection  $\nabla$  and a pair of linearly independent smooth vector fields  $X, Y$  such that  $[X, Y] = 0$ . In particular, according to Theorem 4.33, we can assume that  $X|_U = \frac{\partial}{\partial x}$  and  $Y|_U = \frac{\partial}{\partial y}$  in a local chart  $\phi : U \ni p \mapsto (x, y, x^3, \dots, x^n) \in \mathbb{R}^n$ . Starting from  $p \equiv (0, \dots, 0)$  we can reach the point  $q \equiv (h, k, 0, \dots, 0)$ , where  $h, k > 0$ , along two different swapped paths described in coordinates as follows.

- (1)  $\gamma_1 : [0, h+k] \rightarrow U$  made of a segment from  $(0, \dots, 0)$  to  $(h, 0, \dots, 0)$  followed by a segment from  $(h, 0, \dots, 0)$  to  $(h, k, \dots, 0)$ .
- (2)  $\gamma_2 : [0, h+k] \rightarrow U$  made of a segment from  $(0, \dots, 0)$  to  $(0, k, \dots, 0)$  followed by a segment from  $(0, k, \dots, 0)$  to  $(h, k, \dots, 0)$ .

A vector field  $A_p \in T_p M$  can be parallelly transported to  $q$  either along  $\gamma_1$ , giving rise to the vector  $A_q \in T_q M$ , or along  $\gamma_2$ , giving rise to the vector  $A'_q \in T_q M$ . If  $(M, \mathbf{g})$  is locally flat and  $\phi$  is a chart where the connection coefficients vanish, it is easy to see that  $A_q = A'_q$  by direct computation.

If  $M$  is not locally flat, in other words it is curved, one finds  $A_q \neq A'_q$  in general (See Fig 9.1.2). In the Riemannian case, think for instance of a sphere  $S^2$  embedded in  $\mathbb{R}^3$  with the induced metric. Intuitively speaking, the *curvature* is responsible for the failure of  $A_q = A'_q$ . This failure can be adopted as a definition of “curved manifold” differently from our approach where we used the failure of the commutativity of the covariant derivatives.

What we want to show now is the failure to have  $A_q = A'_q$  is however quantitatively measured by the Riemann tensor  $R_p$  in the limit  $h, k \rightarrow 0$ , so that in the (pseudo) Riemannian

case the two ideas agree.

To go on with computations, we henceforth write  $(h, k)$  in place of  $(h, k, 0, \dots, 0)$  and we directly indicate the points with their coordinates.

From the equation of parallel transport,

$$\frac{dA^a}{dt} = -\Gamma_{bc}^a(\gamma(t))A^b\gamma'^c(t),$$

the vector  $A(h, 0)$  obtained by parallel transport of  $A(0, 0) = A_p$  along the segment joining  $(0, 0)$  to  $(h, 0)$  satisfies

$$A^a(h, 0) = A^a(0, 0) - \int_0^h \Gamma_{1b}^a(x, 0)A^b(x, 0)dx.$$

Transporting this vector to  $q = (h, k)$  along the segment joining  $(h, 0)$  and  $(h, k)$  produces

$$\begin{aligned} A^a(h, k) &= A^a(0, 0) - \int_0^h \Gamma_{1b}^a(x, 0)A^b(x, 0)dx - \int_0^k \Gamma_{2b}^a(h, y)A^b(0, 0)dy \\ &\quad + \int_0^k \Gamma_{2c}^a(h, y) \int_0^h \Gamma_{1b}^c(x, 0)A^b(x, 0)dx dy. \end{aligned}$$

This is the vector  $A_q$  produced by using  $\gamma_1$ . The vector  $A'_q$  computed by using  $\gamma_2$  has instead the form

$$\begin{aligned} A'^a(h, k) &= A^a(0, 0) - \int_0^k \Gamma_{2b}^a(0, y)A^b(0, y)dy - \int_0^h \Gamma_{1b}^a(x, k)A^b(0, 0)dx \\ &\quad + \int_0^h \Gamma_{1c}^a(x, k) \int_0^k \Gamma_{2b}^c(0, y)A^b(0, y)dy dx. \end{aligned}$$

In summary,

$$\begin{aligned} A'^a(h, k) - A^a(h, k) &= \int_0^h \Gamma_{1c}^a(x, k) \int_0^k \Gamma_{2b}^c(0, y)A^b(0, y)dy dx - \int_0^k \Gamma_{2c}^a(h, y) \int_0^h \Gamma_{1b}^c(x, 0)A^b(x, 0)dx dy \\ &\quad + \int_0^h \Gamma_{1b}^a(x, 0)A^b(x, 0)dx - \int_0^k \Gamma_{2b}^a(0, y)A^b(0, y)dy \\ &\quad + A^b(0, 0) \left[ \int_0^k \Gamma_{2b}^a(h, y)dy - \int_0^h \Gamma_{1b}^a(x, k)dx \right]. \end{aligned}$$

We want to expand this difference in  $(h, k)$  up to the second order using the Taylor expansion around  $(0, 0) \equiv p$ . A direct computation proves that all first and second order derivatives vanishes barring

$$\frac{\partial^2}{\partial h \partial k} \Big|_{(0,0)} [A'^a(h, k) - A^a(h, k)] = A^b(0, 0) \left[ \frac{\partial \Gamma_{2b}^a}{\partial x^1} - \frac{\partial \Gamma_{1b}^a}{\partial x^2} + \Gamma_{1c}^a \Gamma_{2b}^c - \Gamma_{2c}^a \Gamma_{1b}^c \right] (0, 0).$$

From (9.5), remembering that  $X = \frac{\partial}{\partial x^1}$  and  $Y = \frac{\partial}{\partial x^2}$ , we conclude that

$$A'_q - A_q = \frac{hk}{2} R_p(X_p, Y_p) A_p + O_3(h, k),$$

where  $|O_3(h, k)| \leq C \|(h, k)\|^3$  in a neighborhood of  $(0, 0)$  for some constant  $C \geq 0$ .

We have obtained that the curvature tensor  $R$  is a measure of the failure of the parallel transport to be “commutative” when joining a pair of “infinitesimally” close points by using paths tangent to a pair of commuting smooth vector fields.

### 9.1.3 Properties of curvature tensor and Bianchi identity

The curvature tensor enjoys a set of useful properties which we are going to summarize in the proposition below. In the (pseudo) Riemannian case, these properties are very crucial in physics because they play a central role in General Relativity as we shall see shortly.

**Proposition 9.10.** *The curvature tensor  $R$  associated to an affine torsion-free connection  $\nabla$  on a smooth manifold  $M$  enjoys the following properties, where  $X, Y, Z, W$  are arbitrary in  $\mathfrak{X}(M)$ .*

- (1)  $R(X, Y)Z = -R(Y, X)Z$  or equivalently  $R_{ijk}{}^l = -R_{jik}{}^l$ .
- (2)  $R(X, Y, Z) + R(Y, Z, X) + R(Z, X, Y) = 0$  or equivalently  $R_{ijk}{}^l + R_{jki}{}^l + R_{kij}{}^l = 0$ .
- (3) If  $\nabla$  is the Levi-Civita connection, then

$$\mathbf{g}(R(X, Y)Z, W) = -\mathbf{g}(Z, R(X, Y)W) \quad \text{or equivalently} \quad R_{ijkl} = -R_{ijlk}$$

where  $R_{ijkl} := R_{ijk}{}^r g_{rl}$ .

- (4) If  $\nabla$  is the Levi-Civita connection, then the **Bianchi identity** holds

$$\nabla_Y R(Z, W) + \nabla_Z R(W, Y) + \nabla_W R(Y, Z) = 0 \quad \text{i.e.} \quad \nabla_k R_{ijp}{}^a + \nabla_i R_{jkp}{}^a + \nabla_j R_{kip}{}^a = 0.$$

- (5) if  $\nabla$  is the Levi-Civita connection, then

$$R_{ijkl} = R_{klij}.$$

**Proof.** (1) is an immediate consequence of the definition of the curvature tensor given in Proposition 9.3.

To prove (2), we start from the identity,

$$\nabla_{[i} \nabla_{j} \omega_{k]} := \frac{1}{3!} (\nabla_i \nabla_j \omega_k + \nabla_j \nabla_k \omega_i + \nabla_k \nabla_i \omega_j - \nabla_j \nabla_i \omega_k - \nabla_i \nabla_k \omega_j - \nabla_k \nabla_j \omega_i) = 0$$

which can be checked by direct inspection and using  $\Gamma_{pq}^r = \Gamma_{qp}^r$ . Next one directly finds by (9.5),  $\nabla_i \nabla_j \omega_k - \nabla_j \nabla_i \omega_k = R_{ijk}{}^l \omega_l$  (see Exercise 9.6.1) and thus  $\nabla_{[i} \nabla_{j} \omega_{k]} - \nabla_{[j} \nabla_{i} \omega_{k]} = R_{[ijk]}{}^l \omega_l$ .



Hence  $R_{[ijk]}^l \omega_l = 0$ . Since  $\omega$  is arbitrary  $R_{[ijk]}^l = 0$  holds. Using (1), it immediately leads to  $R_{ijk}^l + R_{jki}^l + R_{kij}^l = 0$ , i.e. (2).

(3) is nothing but the specialization of the identity (see Exercise 9.6.2)

$$\nabla_i \nabla_j \Xi^{i_1 \dots i_p}_{j_1 \dots j_q} - \nabla_j \nabla_i \Xi^{i_1 \dots i_p}_{j_1 \dots j_q} = - \sum_{u=1}^p R_{ijs}{}^{iu} \Xi^{i_1 \dots s \dots i_p}_{j_1 \dots j_q} + \sum_{u=1}^p R_{ijju}{}^s \Xi^{i_1 \dots i_p}_{j_1 \dots s \dots j_q}$$

to the case  $\Xi = \mathbf{g}$  and using  $\nabla_i g_{j_1 j_2} = 0$ .

(4) can be proved as follows. Start from

$$X^a{}_{,ij} - X^a{}_{,ji} = R_{ijp}{}^a X^p$$

and take another covariant derivative obtaining

$$X^a{}_{,ijk} - X^a{}_{,jik} - R_{ijp}{}^a X^p_{,k} = R_{ijp}{}^a{}_{,k} X^p$$

Permuting the indices  $ijk$  and summing the results, one has

$$\begin{aligned} & (X^a{}_{,ijk} - X^a{}_{,jik} - R_{ijp}{}^a X^p_{,k}) + (X^a{}_{,jki} - X^a{}_{,ikj} - R_{jkp}{}^a X^p_{,i}) \\ & + (X^a{}_{,kij} - X^a{}_{,kji} - R_{kip}{}^a X^p_{,j}) \\ & = R_{ijp}{}^a{}_{,k} X^p + R_{jkp}{}^a{}_{,i} X^p + R_{kip}{}^a{}_{,j} X^p. \end{aligned}$$

This identity can be re-arranged as

$$\begin{aligned} & (X^a{}_{,ijk} - X^a{}_{,ikj}) + (X^a{}_{,jki} - X^a{}_{,jik}) + (X^a{}_{,kij} - X^a{}_{,kji}) \\ & - (R_{ijp}{}^a X^p_{,k} + R_{jkp}{}^a X^p_{,i} + R_{kip}{}^a X^p_{,j}) \\ & = R_{ijp}{}^a{}_{,k} X^p + R_{jkp}{}^a{}_{,i} X^p + R_{kip}{}^a{}_{,j} X^p. \end{aligned}$$

Exploiting Ricci's identity (Exercise 9.6.2), we find

$$\begin{aligned} & (-R_{jkp}{}^a X^p_{,i} + R_{jki}{}^p X^a_{,p}) + (-R_{kip}{}^a X^p_{,j} + R_{kij}{}^p X^a_{,p}) + (-R_{ijp}{}^a X^p_{,k} + R_{ijk}{}^p X^a_{,p}) \\ & - (R_{ijp}{}^a X^p_{,k} + R_{jkp}{}^a X^p_{,i} + R_{kip}{}^a X^p_{,j}) = R_{ijp}{}^a{}_{,k} X^p + R_{jkp}{}^a{}_{,i} X^p + R_{kip}{}^a{}_{,j} X^p. \end{aligned}$$

After some trivial cancellations we end up with

$$(R_{jki}{}^p + R_{kij}{}^p + R_{ijk}{}^p) X^a_{,p} = R_{ijp}{}^a{}_{,k} X^p + R_{jkp}{}^a{}_{,i} X^p + R_{kip}{}^a{}_{,j} X^p.$$

Property (2) written in components eventually proves that the left-hand side vanishes, so that

$$(R_{ijp}{}^a{}_{,k} + R_{jkp}{}^a{}_{,i} + R_{kip}{}^a{}_{,j}) X^p = 0$$

for every vector field  $X$ . Since that field is arbitrary, we get

$$R_{ijp}{}^a{}_{,k} + R_{jkp}{}^a{}_{,i} + R_{kip}{}^a{}_{,j} = 0,$$

that is

$$\nabla_k R_{ijp}{}^a + \nabla_i R_{jkp}{}^a + \nabla_j R_{kip}{}^a = 0$$

as wanted proving (4). Contracting the indices with those of the vector fields  $Y^i, Z^j, W^k$  one gets also

$$\nabla_Y R(Z, W) + \nabla_Z R(W, Y) + \nabla_W R(Y, Z) = 0.$$

Property (5) is a immediate consequence of (1),(2), and (3)  $\square$

**Exercises 9.11.** Let  $R_{ijkl}$  be the Riemann tensor of a (pseudo) Riemannian manifold with dimension  $n$ . Prove that, at every point  $p \in M$ ,  $R_{ijkl}$  has  $n^2(n^2-1)/12$  independent components. (*Hint.* Use properties (1) and (2) and (3) above.)

#### 9.1.4 Riemann tensor and Killing vector fields

There is a nice interplay between Killing vector fields  $K$  and the Riemann tensor of a (pseudo) Riemannian manifold  $(M, \mathbf{g})$  we are going to illustrate.

Let us start from the definition of Riemann tensor which gives

$$\nabla_a \nabla_b X_c - \nabla_b \nabla_a X_c = R_{abc}{}^d X_d$$

according to (9.7). The Killing equation permits us to recast that identity to

$$\nabla_a \nabla_b X_c + \nabla_b \nabla_c X_a = R_{abc}{}^d X_d.$$

Permuting cyclically the indices we also have

$$\nabla_b \nabla_c X_a + \nabla_c \nabla_a X_b = R_{cab}{}^d X_d.$$

and

$$\nabla_c \nabla_b X_a + \nabla_c \nabla_a X_b = R_{abc}{}^d X_d.$$

Adding together the first two identities and subtracting the third one we have

$$2\nabla_b \nabla_c X_a = (R_{abc}{}^d + R_{bca}{}^d - R_{cab}{}^d) X_d = -R_{cab}{}^d X_d.$$

due to (1) in Proposition 9.10. In summary, a Killing vector field in a (pseudo) Riemannian manifold satisfies the identity

$$2\nabla_b \nabla_c X_a = -R_{cab}{}^d X_d. \quad (9.8)$$

Taking Proposition 5.22 into account, this identity implies the following important result.

**Proposition 9.12.** *Let  $(M, \mathbf{g})$  be a (pseudo) Riemannian manifold. The following facts are valid*

- (a) *If a smooth Killing vector field vanishes at  $p \in M$  and  $(\nabla_a X_b)_p = 0$  also holds, then  $X = 0$  everywhere in  $M$ .*

(b) The vector space of smooth Killing vector fields has dimension  $n(n+1)/2$  at most, where  $n = \dim(M)$

**Proof.** Define the tensor field  $D$  such that, in components,  $D_{ab} := \nabla_a X_b$ . Let  $\gamma$  be such a path connecting  $p$  and any fixed  $q \in U$  where  $U \ni p$  is the domain of a local chart. Along that curve, we can define the following system of differential equations in coordinates for the pair  $(X, D)$ ,

$$\begin{aligned}\frac{dX_b(\gamma(t))}{dt} &= \gamma^a(t) D_{ab}(\gamma(t)) , \\ \frac{dD_{bc}(\gamma(t))}{dt} &= -\frac{1}{2} R_{bca}{}^d(\gamma(t)) X_d(\gamma(t)) \gamma^a(t) .\end{aligned}$$

That system is written in normal form with right-hand side smooth and polynomial. As a consequence there is a unique solution along the whole curve for every choice of initial conditions. With the hypotheses of (a), this solution is the zero solution so that  $X(q) = 0$ . The same argument proves that the subset of  $M$  where  $X = 0$  is open. Evidently, it is also open the set where  $X \neq 0$ . Since  $M$  is connected and is the disjoint union of those open sets, one of them must be empty. The former contains at least  $p$ . Hence  $M$  coincides to the set of the points where  $X$  vanishes and it proves (a). The proof of (b) follows observing that as a consequence of (a), the vector space of Killing vector fields has exactly the same dimension of the space of the initial conditions of the above system.  $X_p$  has  $n$  components, whereas  $(D_{ab})_p$  has  $(n^2 - n)/2$  independent coefficients due to the Killing equation  $D_{ab} + D_{ba} = 0$ . In summary,  $n + n(n-1)/2 = n(n+1)/2$ .  $\square$

**Corollary 9.13.** *Let  $(M, g)$  be a (pseudo) Riemannian manifold. If a Killing vector field  $X$  on  $M$  vanishes in a set that admits an accumulation point, then it vanishes everywhere in  $M$ .*

**Proof.** Let  $p$  be the accumulation point and consider a local chart around  $p$ . By continuity,  $X_p = 0$ . Furthermore, the incremental ratios used to compute the partial derivatives  $\frac{\partial X^a}{\partial x^b}|_p$  admit sequences of zero elements when approaching  $p$ . Since the limits exist and are the partial derivatives, these derivatives must vanish at  $p$ . In summary,  $X$  satisfies the hypotheses of (a) in the previous theorem and thus it vanishes everywhere in  $M$ .  $\square$

**Remark 9.14.** There are several cases of (pseudo) Riemannian manifolds with the maximal number of Killing vector fields, these spaces are said **maximally symmetric spaces**. For instance Minkowski space and Euclidean spaces are always maximally symmetric, but also de Sitter spacetime in General Relativity. These spaces have the nice property that the Riemann tensor takes the form [KoNo96]

$$R_{abcd} = \frac{S}{n(n-1)} (g_{ac}g_{bd} - g_{bc}g_{ad})$$

where

$$S := g^{ac} R_{adc}{}^d$$

is the *curvature scalar* which, in the considered case, is also constant in  $M$ . ■

## 9.2 Local flatness and curvature tensor

This section is devoted to establish the fundamental result concerning the interplay between curvature tensor and local flatness. We know from Proposition 9.9 that the curvature tensor must vanish if the manifold is locally flat (either in the metric or in the affine case). We aim to show that the converse implication is also true: the curvature tensor vanishes everywhere *if and only* if  $M$  is locally flat. We shall prove this remarkable result both for (pseudo) Riemannian manifolds and for smooth manifolds endowed with a torsion-free affine connection.

### 9.2.1 A local analytic version of Frobenius theorem.

A preliminary technical lemma is necessary. This is an analytic version [Shi72]<sup>1</sup> in  $\mathbb{R}^n$  of the well-known *Frobenius Theorem* discussed in Theorem 4.33) and locally equivalent to it.

**Theorem 9.15.** *Let  $A \subset \mathbb{R}^n$  be an open set and let  $F_{ij} : A \times \mathbb{R}^m \rightarrow \mathbb{R}$  be a set of  $C^1$  mapping,  $i = 1, \dots, n$ ,  $j = 1, \dots, m$ . Consider the system of differential equations*

$$\frac{\partial X_j}{\partial x^i} = F_{ij}(x^1, \dots, x^n, X_1, \dots, X_m), \quad (9.9)$$

*with initial conditions at a fixed  $p \in A$ ,*

$$X_j(p) = X_j^{(0)} \quad j = 1, \dots, m, \quad (9.10)$$

*where  $X_j = X_j(x^1, \dots, x^n)$  are unknown real-valued  $C^1$  functions and  $X_j^{(0)} \in \mathbb{R}$  given constants. Then there exists an open set  $U \subset A$  with  $p \in U$  such that a solution  $\{X_j\}_{j=1, \dots, m}$  of (9.9)-(9.10) exists and is unique in  $U$  if the following **Frobenius conditions** hold in  $A \times \mathbb{R}^m$*

$$\begin{aligned} & \frac{\partial F_{ij}(x^1, \dots, x^n, Y_1, \dots, Y_m)}{\partial x^k} + \sum_{r=1}^m \frac{\partial F_{ij}(x^1, \dots, x^n, Y_1, \dots, Y_m)}{\partial Y^r} F_{kr}(x^1, \dots, x^n, Y_1, \dots, Y_m) \\ &= \frac{\partial F_{kj}(x^1, \dots, x^n, Y_1, \dots, Y_m)}{\partial x^i} + \sum_{r=1}^m \frac{\partial F_{kj}(x^1, \dots, x^n, Y_1, \dots, Y_m)}{\partial Y^r} F_{ir}(x^1, \dots, x^n, Y_1, \dots, Y_m). \end{aligned}$$

**Remark 9.16.**

(1) Frobenius conditions are nothing but the statement of Schwarz' theorem referred to the solution  $\{X_j\}_{j=1, \dots, m}$ , if it is  $C^2$ :

$$\frac{\partial^2 X_j}{\partial x^i \partial x^k} = \frac{\partial^2 X_j}{\partial x^k \partial x^i},$$

---

<sup>1</sup>See also Remark 1.61 in [War83] and the nice discussion presented in H.A. Hakopian and M.G. Tonoyan, *Partial differential analogs of ordinary differential equations and systems*. New York J. Math.10 (2004) 89–116.

written in terms of the functions  $F_{ij}$ , making use of the differential equation (9.9) itself.

(2) If the functions  $F_{ij}$  are  $C^\infty$ , then the solution  $X_j$  is also  $C^\infty$ , as it arises by differentiating both sides of (9.9) an arbitrary number of times. ■

### 9.2.2 Local flatness is equivalent to zero curvature tensor

We can state and prove the crucial theorem relating local flatness and the fact that the curvature tensor vanishes everywhere.

**Theorem 9.17.** *Let  $M$  be a smooth manifold equipped with a torsion-free affine connection  $\nabla$ . The following facts are equivalent.*

- (a)  $\nabla$  is locally flat.
- (b) The curvature tensor  $R$  vanishes everywhere in  $M$ .

*If  $(M, \mathbf{g})$  is (pseudo) Riemannian and  $\nabla$  is the Levi-Civita connection, then the above conditions are equivalent to*

- (c)  $(M, \mathbf{g})$  is locally flat.

**Proof.** It is evident that (c) is equivalent to (a) if  $\nabla$  is the Levi Civita-connection. Indeed, (c) implies (a) evidently. The converse is also true. In fact, assuming (a), the very definition of Christoffel coefficients (6.9) and (7.10) imply that if the connection coefficients vanish in a chart, then the components of the metric are constant there. Next, with a linear transformation, we can simultaneously put the components of the metric in canonical form.

Let us pass to prove that (a) and (b) are equivalent. By Proposition 9.9, we know that (a) implies (b). We have to show that (b) implies (a): if  $R = 0$  everywhere, then there is an open neighborhood of each  $p \in M$  where the connection coefficients vanish. To this end, fix  $p \in M$  and take vector basis  $e_1, \dots, e_n$  of  $T_p M$ . The proof consists of two steps.

- (A) We prove that there are  $n$  smooth vector fields  $X_{(1)}, \dots, X_{(n)}$  defined in a sufficiently small neighborhood  $U$  of  $p$  such that

- (i)  $(X_{(a)})_p = e_a$ ,
- (ii)  $\nabla X_{(a)} = 0$  for  $a = 1, \dots, n$ ,
- (iii)  $\{(X_{(a)})_q\}_{a=1, \dots, n}$  is also a basis of  $T_q M$  if  $q \in U$ .

- (B) We prove that there is a coordinate system  $y^1, \dots, y^n$  defined in  $U$ , such that

$$(X_{(a)})_q = \frac{\partial}{\partial y^a} \Big|_q,$$

for every  $q \in U$  and  $i = 1, \dots, n$ .

(A) and (B) imply the thesis:

$$\Gamma_{jk}^i = \left\langle \nabla_{\frac{\partial}{\partial y^j}} \frac{\partial}{\partial y^k}, dy^i \right\rangle = \left\langle \nabla_{\frac{\partial}{\partial y^j}} X_{(k)}, dy^i \right\rangle = \langle 0, dy^i \rangle = 0 \quad \text{everywhere in } U.$$

*Proof of (A).* The condition  $\nabla X = 0$  (we omit the index  $(a)$  for the sake of simplicity), using a local coordinate system around  $p$  reads

$$\frac{\partial X^i}{\partial x^r} = -\Gamma_{rj}^i X^j.$$

Theorem 9.15 assures that a solution locally exist (with fixed initial condition) if, in a neighborhood of  $p$ ,

$$-\frac{\partial \Gamma_{rj}^i}{\partial x^s} X^j + \Gamma_{rj}^i \Gamma_{sq}^j X^q$$

equals

$$-\frac{\partial \Gamma_{sj}^i}{\partial x^r} X^j + \Gamma_{sj}^i \Gamma_{rq}^j X^q$$

for all the values of  $i, r, s$ . Using the absence of torsion ( $\Gamma_{kl}^i = \Gamma_{lk}^i$ ) and (9.5), the given condition can be rearranged into

$$R_{srj}{}^i X^j = 0,$$

which holds because  $R = 0$  in  $M$ . Using the found result, in a sufficiently small neighborhood  $U$  of  $p$  we can define the vector fields  $X_{(1)}, \dots, X_{(n)}$  such that (i)  $(X_{(a)})_p = e_a$ , (ii)  $\nabla X_{(a)} = 0$ . We finally prove that (iii) these vectors are everywhere linearly independent. Suppose that for some  $c^1, \dots, c^n$  and for some  $q \in U$

$$\left( \sum_{a=1}^n c^a X_{(a)} \right)_q = 0.$$

We want to prove that  $c^1 = \dots = c^n = 0$ . Define the vector field  $S := \sum_{a=1}^n c^a X_{(a)}$  on  $U$  using the found vector fields  $X_{(a)}$  and the said constants. The equation  $\nabla S = 0$ , with initial condition  $S_q = 0$ , just in view of the validity of Frobenius conditions, admits the *unique* solution  $S = 0$  in a neighborhood of  $q \in U$ , so that  $S$  must vanish therein. This fact proves the subset  $O \subset U$  where  $S = 0$  is open. On the other hand, the subset  $O'$  of  $U$  where  $S \neq 0$  is evidently open since  $S$  is continuous. We therefore have the disjoint decomposition  $U = O \cup O'$  made of open subsets. We can always assume that  $U$  is connected and this fact implies that  $U = O$ , since by hypothesis  $O$  includes at least a point. In particular,  $\sum_{a=1}^n c^a (X_{(a)})_p = 0$  which implies  $c^1 = \dots = c^n = 0$  as wanted, since the vectors  $(X_{(a)})_p$  form a basis.

*Proof of (B).* Fix coordinates  $x^1, \dots, x^n$  on  $U$  and, for every  $q \in U$ , consider the dual basis  $\omega_q^{(1)}, \dots, \omega_q^{(n)} \in T_q^* M$  of the basis  $(X_{(1)})_q, \dots, (X_{(n)})_q \in T_q M$ . The covariant vector fields  $\omega^{(a)}$  are smooth<sup>2</sup> and satisfy  $\nabla \omega^{(a)} = 0$ . Indeed, for every vector field  $Z$  one has:

$$0 = Z(\delta_a^b) = \nabla_Z \langle X_{(a)}, \omega^{(b)} \rangle = \langle \nabla_Z X_{(a)}, \omega^{(b)} \rangle + \langle X_{(a)}, \nabla_Z \omega^{(b)} \rangle = 0 + \langle X_{(a)}, \nabla_Z \omega^{(b)} \rangle,$$

---

<sup>2</sup>In a coordinate system defined in  $U$ , we have  $\omega_j^{(a)} = (M^{-1})_j^{(a)}$ , where  $M$  is the matrix of smooth coefficients  $X_{(a)}^k$ . Hence also the inverse matrix has smooth coefficients.

and thus  $\nabla\omega^{(b)} = 0$  because the vectors  $X_{(a)}$  form a basis. We seek for smooth functions  $y^a = y^a(x^1, \dots, x^n)$ , where  $a = 1, \dots, n$ , defined on  $U$  (or in a smaller open neighborhood of  $p$  contained in  $U$ ) such that

$$\frac{\partial y^a}{\partial x^i} = \omega_i^{(a)} \quad \text{for } i = 1, \dots, n. \quad (9.11)$$

Once again Theorem 9.15 assures that these functions exists provided

$$\frac{\partial \omega_i^{(a)}}{\partial x^r} = \frac{\partial \omega_r^{(a)}}{\partial x^i}$$

for  $a, i, r = 1, \dots, n$  in a neighborhood of  $p$ . Exploiting the absence of torsion of the connection, the condition above can be re-written in the equivalent form

$$\nabla_r \omega_i^{(a)} = \nabla_i \omega_r^{(a)},$$

which holds true because  $\nabla\omega^{(a)} = 0$ . Notice that the found set of smooth functions  $y^a = y^a(x^1, \dots, x^n)$ ,  $a = 1, \dots, n$  satisfy

$$\det \left[ \frac{\partial y^a}{\partial x^i} \right] \neq 0.$$

This is because, from (9.11),  $\det \left[ \frac{\partial y^a}{\partial x^i} \right] = 0$  would imply that the forms  $\omega^{(a)}$  are not linearly independent and that is not possible because they form a basis. We have proved that the functions  $y^a = y^a(x^1, \dots, x^n)$ ,  $a = 1, \dots, n$  define a local coordinate system around  $p$ . To conclude, we notice that

$$\langle X_{(a)}, \omega^{(b)} \rangle = \delta_a^b$$

can be re-written in view of (9.11)

$$X_{(a)}^i \frac{\partial y^b}{\partial x^j} \left\langle \frac{\partial}{\partial x^i}, dx^j \right\rangle = \delta_a^b,$$

that is

$$X_{(a)}^i \frac{\partial y^b}{\partial x^i} = \delta_a^b.$$

Therefore:

$$X_{(a)}^i = \frac{\partial x^i}{\partial y^a},$$

or, equivalently,

$$(X_{(a)})_q = \frac{\partial}{\partial y^a} \Big|_q,$$

is valid in a neighborhood of  $p$ . □

## Chapter 10

# Gravitation in General Relativity

This chapter has a twofold goal. On the one hand it clarifies what it is meant for *external* gravitational field in General Relativity. This is done by taking advantage of the mathematical tools presented in the previous chapter regarding the notion of curvature of a manifold. The presence of an external gravitational field is encoded in the appearance of the so-called *geodesic deviation* of free falling bodies which, in turn, is equivalent to a non-vanishing Riemann curvature tensor. On the other hand, the chapter discusses how the matter *generates* its own gravitational field according to the very famous *Einstein equations*. We conclude with a short discussion about the *relativistic cosmology* arising from the Einstein Equations, pointing out also some problems with the recent astronomical observations.

### 10.1 Geodesic deviation and local flatness of spacetime

In this section we introduce the notion of *geodesic deviation*. Afterwards we analyze the interplay of local flatness and geodesic deviation measured for *causal* geodesics (more precisely *timelike*) starting from the remark that, from a physical viewpoint, the geodesic deviation can be measured for causal geodesic, observing the motion of (infinitesimal) falling bodies, but it can hardly be evaluated on spacelike geodesics. We establish in particular that a generic spacetime is (locally) flat if and only if there is no geodesic deviation for *timelike geodesics*.

#### 10.1.1 External gravitational interaction in General Relativity

The issue we want to investigate now is the physical meaning in General Relativity of a statement such “in the region  $\Omega$  of a spacetime  $(M^n, \mathbf{g})$ , the material points are subjected to an *external gravitational interaction*.” (For the moment we disregard the gravitational interaction generated by themselves.)

The Newtonian interpretation of the statement is based on the notion of *gravitational force* present in  $\Omega$ . The relativistic extension of the notion of force is that of *n-force*, so it may seem natural to exploit it. Here an obstruction of physical nature pops up. The Equivalence



Principle, which states that the gravitational field can be cancelled locally, does not allow the use of a  $n$ -force: a  $n$ -force cannot be cancelled by the choice of the reference frame it being a tensor (if it vanishes in a coordinate system it must vanish in all of them). If we make use of the notion of  $n$ -force we cannot encompass the Equivalence Principle in the formalism, exactly as it happens in Newtonian physics.

A meaningful idea is looking for some phenomenon that cannot be cancelled by any suitable choice of the reference frame and trying to embody it in the formalism to describe the presence of an external gravitational interaction. What we cannot cancel with the choice of the reference frame is the *relative acceleration* between *two or more* free falling bodies, except in the (classical) case of a uniform static gravitational field, where this relative acceleration does not exist. However, in this case, sooner or later relative accelerations appear, unless considering the unphysical idealization of classical physics of a static uniform gravitational field infinitely extended in space and time. On the basis of these remarks, we shall keep the relative acceleration as an operational definition of the presence of external gravitational interaction acting on the bodies.

On the other hand, we already know from Section 8.4.1 that the failure of the metric to be flat (Minkowskian), seems to be appropriate to describe the relativistic notion of external gravitational interaction. We would like to include also this fact in the formalism. Indeed, we shall see that the use of the relative acceleration is also in agreement with the idea that the local flatness of the spacetime physically corresponds to the absence of external gravitational interaction. This agreement passes through the properties of the Riemann tensor.

### 10.1.2 Geodesic deviation as description of external gravitational interaction

We need some preliminary definition to formalize the ideas discussed above. We shall adopt here the most mathematically comfortable hypotheses for our physical goals, a technically different formalization of the same notion is discussed and used in [O'Ne83].

**Definition 10.1.** Let  $(M^n, \mathbf{g})$  be a spacetime and consider a family of geodesics

$$\gamma_s : I \ni t \mapsto \gamma_s(t) \in M^n \quad \text{where} \quad s \in J$$

with  $I, J \subset \mathbb{R}$  open non-empty intervals,  $t$  a (common) affine parameter, and assume that the map

$$\gamma : J \times I \ni (s, t) \mapsto \gamma_s(t) \in M^n$$

is an immersion. The map  $\gamma$  will be called a **smooth congruence of geodesics**. ■

If we define

$$T_{\gamma_s(t)} := d\gamma \frac{\partial}{\partial t} \Big|_{(s,t)}, \quad S_{\gamma_s(t)} := d\gamma \frac{\partial}{\partial s} \Big|_{(s,t)}, \quad (10.1)$$

we have that  $T$  and  $S$  are everywhere linearly independent. Furthermore, since  $(s, t) \in J \times I$  are coordinates on  $J \times I$  viewed as smooth 2-dimensional manifold, Theorem 4.5 implies that there is a neighborhood  $U$  of  $(s_0, t_0) \in J \times I$  and a local chart in a neighborhood  $V$  of  $\gamma_{s_0}(t_0) \in M^n$

with coordinates  $x^1, \dots, x^n$ , such that the points  $\gamma_s(t) \in V$  have coordinates  $(s, t, 0, \dots, 0)$  if  $(s, t) \in U$ . In this local chart

$$T_{\gamma_s(t)} = \frac{\partial}{\partial t} \Big|_{\gamma(s,t)}, \quad S_{\gamma_s(t)} = \frac{\partial}{\partial s} \Big|_{\gamma(s,t)}, \quad (10.2)$$

where now  $s, t, x^3, \dots, x^n$  are coordinates on  $V$ . Therefore  $T$  and  $S$  are in that way locally extended to smooth vector fields in  $M^n$ . With this extension, evidently,

$$[T, S]_{\gamma_s(t)} = 0, \quad \text{for every } (s, t) \in I \times J, \quad (10.3)$$

since that commutator can be viewed as the commutator of the vector fields tangent to the first two coordinates of the said chart around  $p$  in  $M$ . The coordinate  $t$  can be chosen as the length parameter for spacelike geodesics, or the proper time for timelike geodesics.

At least when  $S$  is spacelike,  $\nabla_T S$  defines the *relative velocity*, referred to the parameter  $t$ , between infinitesimally close geodesics (say  $\gamma_s$  and  $\gamma_{s+\delta s}$ ). Similarly,  $\nabla_T(\nabla_T S)$  defines the *relative acceleration*, referred to the parameter  $t$ , between infinitesimally close geodesics.

Smooth congruences of geodesics exist as stated in the following lemma.

**Lemma 10.2.** *Let  $(M^n, \mathbf{g})$  be a spacetime,  $p \in M^n$  and  $T_p, S_p \in T_p M^n$  linearly independent vectors. Then there exists a smooth congruence of geodesics  $\gamma : J \times I \rightarrow M^n$  with  $I, J \ni 0$ , such that*

$$(a) \quad p = \gamma_0(0),$$

$$(b) \quad T_{\gamma_0(0)} = T_p \text{ and } S_{\gamma_0(0)} = S_p$$

**Proof.** Fix normal coordinates  $\phi : N \ni q \mapsto (x^1, \dots, x^n) \in A := \phi(N) \subset \mathbb{R}^n$  around  $p$  with  $\frac{\partial}{\partial x^1} \Big|_p = T_p$  and  $\frac{\partial}{\partial x^2} \Big|_p = S_p$ . Finally consider the map

$$(-a, a)^n \mapsto f(x^1, \dots, x^n) = \exp_{\phi(x^1, \dots, x^n)} \left( x^1 \frac{\partial}{\partial x^1} \Big|_{\phi(x^1, \dots, x^n)} \right) \in M^n$$

where  $a > 0$  is sufficiently small such that the function above is well defined. By direct inspection, one sees that

$$\frac{\partial(\phi \circ f)^a}{\partial x^b} \Big|_0 = \delta_b^a.$$

Thus  $f : (-\delta, \delta)^n \rightarrow M^n$  defines a local diffeomorphism if  $\delta > 0$  is sufficiently small. Hence the 2-dimensional embedded submanifold defined in normal coordinates by  $x^3 = x^4 = \dots = x^n = 0$  and  $x^1 = t, x^2 = s$  arbitrarily varying but remaining in  $(-\delta, \delta)^n$  is mapped by  $f$  to an embedded 2-dimensional manifold  $\Sigma$  of  $M^n$  (*a fortiori* an immersion) passing through  $p$  for  $t = s = 0$  and defined by

$$\gamma_s(t) = \exp_{\phi(0, s, 0, \dots, 0)} \left( t \frac{\partial}{\partial x^1} \Big|_{\phi(0, s, 0, \dots, 0)} \right).$$

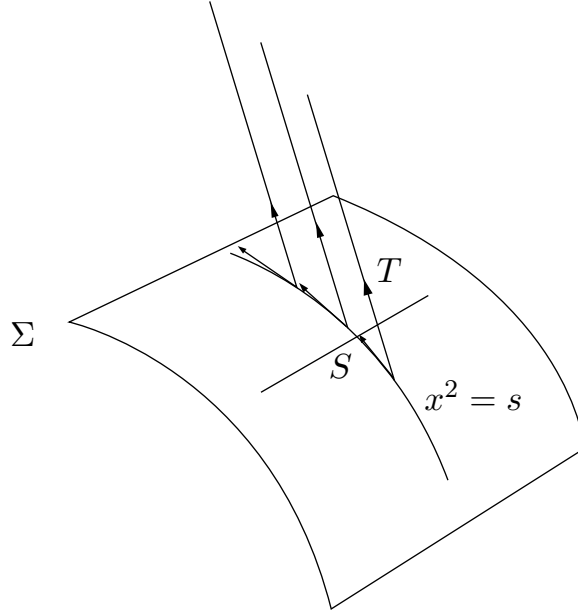


Figure 10.1: Construction of  $\gamma$

By construction this is a smooth congruence of geodesics and  $T_{\gamma_0(0)} = T_p$  and  $S_{\gamma_0(0)} = S_p$ .  $\square$

We are in a position to state a general definition that considers all possible cases including spacelike geodesics.

**Definition 10.3.** If  $\gamma : J \times I \rightarrow M^n$  is a smooth congruence of geodesics in the spacetime  $(M^n, \mathbf{g})$ , the smooth field on  $\gamma$

$$\nabla_T \nabla_T S$$

is said **geodesic deviation** of  $\gamma$ , where  $T$  and  $S$  are defined in (10.2).  $\blacksquare$

To this respect we have the following celebrated result.

**Proposition 10.4.** If  $\gamma : J \times I \rightarrow M^n$  is a smooth congruence of geodesics in the spacetime  $(M^n, \mathbf{g})$ , the **equation of geodesic deviation** is valid at every  $p \in \gamma(J \times I)$ ,

$$\nabla_T \nabla_T S = R(S, T)T \quad \text{equivalently} \quad (\nabla_T \nabla_T S)^k = R_{ijl}{}^k S^i T^j T^l. \quad (10.4)$$

**Remark 10.5.**

(1) Though  $S$  and  $T$ , as vector fields in  $M^n$  outside  $\gamma(J \times I)$ , were defined with the help of an

arbitrary coordinate system around every point  $p \in \gamma_s(t)$ , the equation above does not depend on this arbitrary choice as (1) the right-hand side is evaluated exactly on  $\gamma_s(t)$  and (2) the left-hand side can be completely interpreted using Definition 6.16 and Proposition 6.18, when first  $S$  and next  $\nabla_T S$  are viewed as smooth vector fields defined exactly on  $\gamma_s$ .

(2) In a (pseudo) Riemannian manifold, every vector field  $S = S(t)$  defined along a geodesic segment  $\gamma = \gamma(t)$ , with tangent vector  $\gamma'(t) := T(t)$ , satisfying (10.4) is called a **Jacobi field** of that geodesic. ■

**Proof of Proposition 10.4.** From (10.3), we have

$$(\nabla_T \nabla_T S)^k = T^i \nabla_i (T^j \nabla_j S^k) = T^i \nabla_i (S^j \nabla_j T^k),$$

so that

$$(\nabla_T \nabla_T S)^k = (T^i \nabla_i S^j) \nabla_j T^k + S^j T^i \nabla_i \nabla_j T^k.$$

Using (9.4) in the last term, we have

$$(\nabla_T \nabla_T S)^k = (T^i \nabla_i S^j) \nabla_j T^k + S^j T^i \nabla_j \nabla_i T^k - R_{ijl}{}^k S^j T^i T^l.$$

Exploiting again (10.3) in the first term on the right-hand side, we find

$$(\nabla_T \nabla_T S)^k = (S^i \nabla_i T^j) \nabla_j T^k + S^j T^i \nabla_j \nabla_i T^k - R_{ijl}{}^k S^j T^i T^l.$$

Swapping the summed indices  $i$  e  $j$  in the first term on the right-hand side, we eventually obtain

$$(\nabla_T \nabla_T S)^k = (S^j \nabla_j T^i) \nabla_i T^k + S^j T^i \nabla_j \nabla_i T^k - R_{ijl}{}^k S^j T^i T^l.$$

Observe that the last addend can be transformed to

$$-R_{ijl}{}^k S^j T^i T^l = R_{jil}{}^k S^j T^i T^l = R_{ijl}{}^k S^i T^j T^l$$

where we have taken (i) of Proposition 9.10 into account. The overall result can be rearranged to

$$(\nabla_T \nabla_T S)^k = S^j \nabla_j (T^i \nabla_i T^k) + R_{ijl}{}^k S^i T^j T^l.$$

The geodesic equation now implies that  $T^i \nabla_i T^k = 0$  producing (10.4). □

The found result permits us to conclude our mathematical discussion with this paramount result (which will be however improved in the next section from the physical point of view), where one sees that in a region of the spacetime there is no external gravitational field (i.e., there is geodesic deviation) if and only if that region is locally flat.

**Theorem 10.6.** *Let  $\Omega \subset M^n$  be an open subset of a spacetime  $(M^n, \mathbf{g})$ . The region  $\Omega$ , viewed as spacetime in its own right, is locally flat if and only if the geodesic deviation  $\nabla_T \nabla_T S$  vanishes*

for every smooth congruence of geodesics taking values in  $\Omega$ .

**Proof.** If  $R = 0$  in  $\Omega$ , then no geodesic deviation exists trivially. Let us prove the converse fact. As established in Lemma 10.2, for every  $p \in \Omega$  and  $S_p, T_p \in T_p M^n$  linearly independent, there is a smooth congruence of geodesics taking values in a neighborhood of  $p$  such that the vector fields  $S$  and  $T$  evaluated at  $p$  coincide with  $S_p$  and  $T_p$  respectively. In the hypothesis that  $\nabla_T \nabla_T S = 0$  at  $p$ , the geodesic equation furnishes  $R_p(S_p, T_p)T_p = 0$ . Let us therefore use first  $T_p = U_p + V_p$  and next  $T_p = U_p - V_p$ . Subtracting side-by-side the obtained results, taking bi-linearity into account, one finds:

$$R_p(S_p, U_p)V_p + R_p(S_p, V_p)U_p = 0, \quad (10.5)$$

which is valid for every  $S_p, U_p, V_p \in T_p M$ . Identity (2) in Proposition 9.10 can be specialized here as:

$$R_p(S_p, U_p)V_p + R_p(U_p, V_p)S_p + R_p(V_p, S_p)U_p = 0. \quad (10.6)$$

Summing side-by-side (10.5) and (10.6), taking (1) in Proposition 9.10 into account, it arises  $2R_p(S_p, U_p)V_p + R_p(U_p, V_p)S_p = 0$ , which can be recast to  $2R_p(S_p, U_p)V_p - R_p(U_p, S_p)V_p = 0$ , where we employed Eq.(10.5) (with different names of the vectors). Using (1) in Proposition 9.10 again, we can restate the obtained result as:  $2R_p(S_p, U_p)V_p + R_p(S_p, U_p)V_p = 0$ . In other words  $R_p(S_p, U_p)V_p = 0$  for all vectors  $S_p, U_p, V_p \in T_p M$ , so that  $R_p = 0$  as wanted.  $\square$

### 10.1.3 Geodesic deviation of causal geodesics and local flatness

Even if the result established in Theorem 10.6 is mathematically interesting and it is also in the spirit of our initial issue, from experimentalist's viewpoint, we must admit that the geodesic deviation  $\nabla_T(\nabla_T S)$  can hardly be measured along spacelike geodesics, excluding particular cases of spacetimes as a static one, and referring to a very special choice of the field  $S$ . For this reason it is natural to try to prove Theorem 10.6 referring only to causal geodesics with  $S$  spacelike. More precisely we shall consider timelike geodesics. in this case,  $\nabla_T(\nabla_T S)$  has the direct meaning of *relative acceleration* of two infinitesimally close massive free falling bodies.

**Theorem 10.7.** *Let  $\Omega \subset M^n$  be an open subset of a spacetime  $(M^n, \mathbf{g})$ . The region  $\Omega$ , viewed as spacetime in its own right, is locally flat if and only if the geodesic deviation  $\nabla_T \nabla_T S$  vanishes for every smooth congruence of timelike geodesics taking values in  $\Omega$  with  $S$  spacelike.*

**Proof.** Evidently, if  $\Omega$  is locally flat, then  $\nabla_T \nabla_T S = 0$  for every type of smooth congruence of geodesics. Let us therefore prove that if the geodesic deviation is zero for all smooth congruences of timelike geodesics with  $S$  spacelike in  $\Omega$ , then  $R = 0$  therein. If  $p \in \Omega$ , for fixed  $S_p, T_p \in T_p M$ , respectively spacelike and timelike, there is a smooth congruence of geodesics  $\gamma : J \times I \rightarrow \Omega$  such that (a) and (b) of Lemma 10.2 are true. By continuity, shrinking  $I$  and  $J$  around 0 if necessary, all vectors  $S$  and  $T$  of  $\gamma$  are respectively timelike and spacelike. Our hypothesis therefore entails that  $0 = \nabla_T \nabla_T S|_p = R_p(S_p, T_p)T_p$  is valid for every choice of  $T_p \in T_p M$  timelike and  $S_p \in T_p M$

spacelike where, in particular,  $\mathbf{g}(S_p, T_p) = 0$ . To extend this result to all possible arguments of  $R$ , we start by noticing that  $R_p(S_p, T_p)T_p = 0$  is still valid if dropping the requirements  $S_p$  spacelike and  $\mathbf{g}(T_p, S_p) = 0$ . Indeed, if  $S_p \in T_p M$  is generic, we can decompose it as  $S_p = S'_p + cT_p$ , where  $c \in \mathbb{R}$  and  $S'_p$  is spacelike with  $\mathbf{g}(T_p, S'_p) = 0$ , for some timelike vector  $T_p$ . Then

$$R_p(S_p, T_p)T_p = R_p(S'_p, T_p)T_p + cR_p(T_p, T_p)T_p = 0 + cR_p(T_p, T_p)T_p = 0$$

where we have used (2) in Proposition 9.10. So, we can start with the hypothesis that

$$R_p(S_p, T_p)T_p = 0 \quad \text{for all } T_p, S_p \in T_p M \text{ with } T_p \text{ timelike.}$$

Let us show that this last constraint can be dropped, too. Fix  $S_p \in T_p M$  arbitrarily and consider the bi-linear map  $T_p M \ni T_p \mapsto F_{S_p}(T_p) := R_p(S_p, T_p)T_p$ . If we restrict  $F_{S_p}$  to one of the two *open* halves  $V_p^{(+)}$  of the light-cone at  $p$ , e.g. that containing the future-directed timelike vectors, we find  $F_{S_p} \upharpoonright_{V_p^{(+)}} = 0$  in view of the discussion above. Since every component of  $F_{S_p}$  is analytic (it being a polynomial in the components of  $T_p$ ) and defined on the connected open domain  $T_p M$ , it must vanish everywhere on  $T_p M$ . Summarizing, we have obtained that  $R_p(S_p, T_p)T_p = 0$  for every vectors  $T_p, S_p \in T_p M$ . At this point Theorem 10.6 implies that  $R_p = 0$  and thus the region  $\Omega$  is locally flat with respect to the metric  $\mathbf{g}$  according to Theorem 9.17.  $\square$

**Remark 10.8.** A similar result can be proved referring to congruences of light-like geodesics<sup>1</sup>.

## 10.2 The Einstein equations

This section is devoted to present the famous Einstein Equations describing how the matter generates the gravitational field in General Relativity. The gravitational interaction, roughly speaking, is now viewed as the curvature of the spacetime. However “curvature” is a quite vague notion. We can think that it is the Riemann tensor as it seems from the discussion about the geodesic deviation. However, we already know that in General Relativity, the gravitational phenomenology has various facets. For instance, it is related to metric notions, differently from what happens in classical physics.

On the other hand, the Riemann tensor is a quite complex object and it is not obvious that the equations that relate the presence of matter to the curvature – *viewing the matter as the source of the curvature* – involves the Riemann tensor and not some other tensor constructed out of it.

### 10.2.1 Ricci tensor, contracted Bianchi identity and Einstein tensor

In a (pseudo) Riemannian manifold and referring to the Levi-Civita (metric torsion-free) connection, there are several tensors which are obtained from Riemann tensor and they turn out

---

<sup>1</sup>See V. Moretti and R. Di Criscienzo, *How can we determine if a spacetime is flat?* Frontiers in Physics 1:12 (2013), DOI:10.3389/fphy.2013.00012.

to be useful in physics. Let us focus on the properties of the curvature tensor established in Proposition 9.10. By properties (1) and (3) the contraction of Riemann tensor over its first two or last two indices vanishes. Conversely, the contraction over the second and fourth (or equivalently, the first and the third using (2),(1) and (3)) indices gives rise to the only non-trivial second order tensor obtained by one contraction from the Riemann tensor. This tensor is known as the **Ricci tensor**:

$$Ric_{ij} := R_{ij} := R_{ikj}{}^k = R_{ki}{}^k{}_j. \quad (10.7)$$

By property (5) one has the symmetry of *Ric*:

$$Ric_{ij} = Ric_{ji}.$$

The contraction of *Ric* produces the so-called **curvature scalar**

$$S := R := R_k^k. \quad (10.8)$$

Another relevant tensor is the so-called **Einstein's tensor** which plays a crucial role in General Relativity as we shall see shortly. It is defined as

$$G_{ij} := Ric_{ij} - \frac{1}{2}g_{ij}S. \quad (10.9)$$

This tensor field satisfies a crucial identity as a consequence of Bianchi identity.

**Proposition 10.9.** *In a spacetime  $(M^n, \mathbf{g})$  the Einstein tensor satisfies the equations*

$$\nabla_a G^{ab} = 0,$$

*known as **contracted Bianchi identity**.*

**Proof.** Starting from Bianchi's identity

$$\nabla_k R_{ijp}{}^a + \nabla_i R_{jkp}{}^a + \nabla_j R_{kip}{}^a = 0$$

and contracting *k* and *a* one gets

$$\nabla_k R_{ijp}{}^k + \nabla_i R_{jkp}{}^k + \nabla_j R_{kip}{}^k = 0.$$

This identity can be rewritten as:

$$\nabla_k R_{ijp}{}^k + \nabla_i R_{jp} - \nabla_j R_{ip} = 0,$$

contracting *i* and *p* (after having raised the index *p*) it arises

$$\nabla_k R_j{}^k + \nabla_i R_j{}^i - \nabla_j R = 0.$$

Multiplying by 1/2 and changing the name of  $k$ :

$$\frac{1}{2}\nabla_i Ric_j^i + \frac{1}{2}\nabla_i Ric_j^i - \frac{1}{2}\nabla^i g_{ij}S = 0 .$$

Those are the equations written above, since they can be rearranged into:

$$\nabla^i \left( Ric_{ij} - \frac{1}{2}g_{ij}S \right) = 0$$

concluding the proof.  $\square$

**Remark 10.10.** As we said above, in a (pseudo)Riemannian manifold  $M$ , Ricci's tensor and the curvature scalar are the only non-vanishing tensors which can be obtained from Riemann tensor using contractions. If  $\dim M =: n \geq 3$ , using  $Ric$  and  $S$  it is possible to built up a tensor field of order 4 which *satisfies properties (1),(2) and (3) in proposition 9.10 and produces the same tensors as  $R_{ijkl}$  under contractions*. That tensor is

$$D_{ijkl} := \frac{2}{n-2}g_{i[k}Ric_{l]j} - g_{j[k}Ric_{l]i} - \frac{2}{(n-1)(n-2)}Sg_{i[k}g_{l]j} .$$

Above  $[ab]$  indicates anti-symmetrization with respect to  $a$  and  $b$ . As a consequence

$$C_{ijkl} := R_{ijkl} - D_{ijkl}$$

satisfies properties (1), (2) and (3) too and every contraction with respect to a pair of indices vanishes. The tensor  $C$ , defined in (pseudo) Riemannian manifolds, is called **Weyl's tensor** or **conformal tensor**. It behaves in a very simple manner under *conformal transformations*.

## 10.2.2 The Einstein Equations

We come to the most interesting and intriguing issue: writing down the relativistic equations that correspond to the Newtonian equations

$$\Delta\varphi = -4\pi G \mu , \tag{10.10}$$

where  $\varphi$  is the gravitational potential and  $\mu$  the density of mass. This equation says how the matter generates the gravitational field in classical physics. The only clue to find the corresponding equation in relativistic physics is perhaps the relation (8.34), found in Section 8.4.1, which relates  $\varphi$  and  $g_{00}$  in a semiclassical scenario

$$g_{00}(t, \vec{x}) = -1 + \frac{2}{c^2}\varphi(\vec{x}) ,$$

It approximatively holds under suitable hypotheses which include small curvatures and small velocities with respect to  $c$  and referring to a coordinate system where the metric is stationary and the components of the metric are close to the Minkowskian one.



Einstein spent 10 years on this formidable problem also eventually discovering his famous equations. In the final part of his research work, around 1915 and later, he received some technical help by various mathematicians like Grossmann and Levi-Civita.

*A posteriori*, we are about describing how to obtain those equations from physical principles, with the help of some mathematical result. We divide the procedure into several steps to emphasize the used physical and mathematical hypotheses.

- (1) Since (10.10) relates the gravitational field to the density of mass, the equations we are searching for should connect the density of mass to the curvature of the spacetime. It is possible to see, with concrete examples (we shall discuss this point later with more details), that the component  $T^{00}$  of the stress energy tensor, which coincide to the density of mass-energy, dominates over all remaining components in semi-classical scenarios. Einstein's idea was that in fully relativistic regimes the whole  $T^{ab}$  should replace  $\mu$  in the equation that corresponds to (10.10).
- (2) Since  $T^{ab}$  is a tensor field, the simplest equation should equates  $T^{ab}$  to some geometric tensor  $H_{ab}$  constructed out of the Riemann tensor which is also symmetric as  $T_{ab}$ ,

$$H_{ab} = kT_{ab} . \quad (10.11)$$

where  $k$  is some constant. Notice that, as  $H_{ab}$  is a  $(0, 2)$  tensor, the Riemann tensor is automatically ruled out.

**Remark 10.11.** This exclusion is actually in agreement with the physical evidence that the gravitational field (here the curvature of the manifold) propagates *outside* the source. An equation like (10.11) permits to have  $R_{abc}{}^d \neq 0$  outside the support of  $T_{ab}$ . ■

- (3) Another hypothesis by Einstein was that the tensor  $H_{ab}$  should be constructed with the derivatives of  $g_{ab}$  *up to the second order*, as the minimal extension of (10.10) and (8.34).
- (4) The final requirement arose from a property of  $T_{ab}$  due to the Strong Equivalence Principle. Indeed, *independently of the specific metric  $\mathbf{g}$  of the spacetime*, the stress energy tensor must satisfy

$$\nabla_a T^{ab} = 0 .$$

We stress that this is an *on shell* equation: it is valid when the matter whose stress-energy tensor is  $T^{ab}$  satisfies the laws of motion for the specific metric of the spacetime viewed as a given external gravitational interaction. Using this identity in (10.11), we must conclude that

$$\nabla_a H^{ab} = 0 . \quad (10.12)$$

We stress that (10.12) must hold for *every* metric  $\mathbf{g}$  automatically, since  $\nabla_a T^{ab} = 0$  is valid for every metric as stated above.

We know from the previous section that the Einstein tensor (10.9)

$$G_{ab} := Ric_{ab} - \frac{1}{2}g_{ab}S$$

is symmetric and satisfies the said requirements, including the fact that it is constructed with the derivatives, up to the second order of the metric, in components. The candidate equation is therefore

$$kT_{ab} = Ric_{ab} - \frac{1}{2}g_{ab}S. \quad (10.13)$$

It is worth stressing that in 1971 Lovelock established<sup>2</sup> an important result later generalized in various directions:

**Theorem 10.12.** (**Lovelock Theorem.**) *In a smooth Lorentzian manifold  $(M, g)$  with  $\dim(M) = 4$ , the most general smooth natural tensor field  $H_{ab}$  locally constructed out of the metric and its derivatives up to the second in coordinates and which satisfies (10.12) has the form*

$$H_{ab} = p G_{ab} + q g_{ab} \quad \text{for arbitrary constants } p, q \in \mathbb{R}. \quad (10.14)$$

As Einstein did, we initially set  $q = 0$  in a spacetime of dimension 4, so that our equation is (10.13) above. We want to fix the constant  $k$  by imposing that this equation produces (10.10) in some classical limit. First of all, observe that, taking the trace of both sides we have,

$$S - 2S = kT,$$

namely

$$S = -kT$$

where

$$T := T_a^a. \quad (10.15)$$

We can therefore re-write our equations as

$$Ric_{ab} = k \left( T_{ab} - \frac{1}{2}Tg_{ab} \right). \quad (10.16)$$

At this juncture, we assume to deal with a semi-classical stationary regime where, in a suitable coordinate system  $(x^0 = ct, x^1, x^2, x^3) = (x^0, \vec{x})$ , the metric appears to be close to the Minkowskian one, as we did in Section 8.4.1. So that

$$g_{ab} = \eta_{ab} + h_{ab}, \quad g^{ab} = \eta^{ab} - h^{ab} \quad (10.17)$$

---

<sup>2</sup>D. Lovelock, *The Einstein Tensor and Its Generalizations*. Journal of Mathematical Physics. 12 (3): 498–502 (1971) and see also A. Navarro and J. Navarro *Lovelock's theorem revisited*. Journal of Geometry and Physics 61 (2011) 1950–1956 for a modern view.

where  $h_{ab}$  and  $h^{ab}$  do not depend on  $x^0$ ,  $|h_{ab}| \ll 1$  and we use the metric  $\eta$  to raise and lower indices, so that

$$h^{ab} := \eta^{ac} \eta^{bd} h_{cd} ,$$

in particular. Dropping systematically all terms containing products of  $h_{ab}$  (or  $h^{ab}$ ), we obtain the second identity in (10.17). Regarding  $T$ , we assume the simplest case of a gas of non-interacting particles

$$T^{ab} = \mu_0 V^a V^b \quad (10.18)$$

where  $\mu_0$  is the density of matter measured at rest with the particles whose 4-velocity is  $V$ . We assume that  $\mu$  is as the same order of magnitude as  $h$  and its derivatives, so that we also neglect product of  $\mu$  and  $h$ . Expanding  $V$  as in Special Relativity and assuming the components of the spatial velocities very small with respect to  $c$  or better that the source is at rest in the said coordinate frame,

$$V^a = c \delta_0^a ,$$

and, in our first-order approximation,

$$T = -\mu_0 c^2 ,$$

and (10.16) produces (neglecting second order terms some of them being products of  $\mu$  and  $h_{ab}$ )

$$Ric_{00} = \frac{k}{2} \mu_0 c^2 . \quad (10.19)$$

Finally,

$$Ric_{00} = R_{0j0}{}^j = \frac{\partial \Gamma_{00}^j}{\partial x^j} - \frac{\partial \Gamma_{j0}^j}{\partial x^0} + \Gamma_{jk}^j \Gamma_{00}^k - \Gamma_{0k}^j \Gamma_{j0}^k .$$

The second addend is zero in view of our stationarity hypothesis. The last two terms are of the second order in the perturbation  $h$  so that we can neglect them. The surviving term, dropping all terms containing  $x^0$  derivatives in its explicit formula, yields

$$Ric_{00} = -\frac{1}{2} \sum_{\alpha=1}^3 \frac{\partial^2 h_{00}}{\partial (x^\alpha)^2} = -\frac{1}{2} \Delta_{\vec{x}} h_{00} .$$

Inserting in (10.19),

$$-\Delta_{\vec{x}} h_{00} = k \mu_0 c^2 ,$$

which, compared with

$$-1 + h_{00}(t, \vec{x}) = -1 + \frac{2}{c^2} \varphi(\vec{x}) ,$$

furnishes

$$-\Delta_{\vec{x}} \varphi = \frac{k c^4}{2} \mu_0 .$$

Comparing with (10.10), we find

$$k = \frac{8\pi G}{c^4}.$$

In summary, **Einstein equations** read

$$\frac{8\pi G}{c^4} T_{ab} = Ric_{ab} - \frac{1}{2} g_{ab} S. \quad (10.20)$$

If the components  $T_{ab}$  are assigned at every point of  $M^4$ , the components  $g_{ab}$  of the metric are the unknown in these differential equations which are non-linear so that its solution is very difficult in general.

**Remark 10.13.**

(1) The *vacuum* Einstein equations

$$Ric_{ab} - \frac{1}{2} g_{ab} S = 0, \quad (10.21)$$

can be rephrased to (contracting  $a$  and  $b$  thus observing that  $S = 0$ )

$$Ric_{ab} = 0. \quad (10.22)$$

These equations still permit the presence of curvature, since they do not require  $R = 0$ .

(2) In principle, the degrees of freedom of the gravitational field are the independent components of the metric tensor. Apparently they are 10 in four dimensions, taking the symmetry  $g_{ab} = g_{ba}$  into account. However, we can fix four degrees of freedom choosing our system of coordinates. Finally, the constraint  $\nabla_a G^{ab} = 0$  imposes further four conditions. Apparently, the remaining degrees of freedom are  $10 - 4 - 4 = 2$  and this turns out to be true at least in the linearized theory, whereas when dealing with the full theory the situation is more complicated and the *Cauchy problem* would deserve a detailed discussion. However the rough argument is in agreement with the idea that the gravitational interaction, in a linearized quantum version, is transported by quantum particles called *gravitons*. From general principles, gravitons must have 2 degree of freedom. ■

Non trivial consequences of Einstein equations received several experimental confirmations since 2016. We quote only three of them.

- (1) The the *perihelion precession of Mercury's orbit* already in 1916 (predicted by Einstein in 1916 with incomplete equations!).
- (2) The direct detection in 2016 by LIGO of *gravitational waves* and all the next direct detections also due to VIRGO. They in particular arise from the linearized Einstein equations for the perturbation  $h_{ab}$  (10.17) after a suitable choice of the reference frame (more precisely of the *gauge*):

$$-\frac{1}{c^2} \frac{\partial^2 \tilde{h}_{ab}}{\partial t^2} + \Delta_{\vec{x}} \tilde{h}_{ab} = -16\pi G T_{ab} \quad (10.23)$$

where  $\tilde{h}_{ab} = h_{ab} - \frac{1}{2}\eta^{cd}h_{cd}\eta_{ab}$ . Equation (10.23) is the standard *D'Alembert Equation with source* describing wave propagation in linear media. In the said approximation and viewing the background as Minkowski space, the speed of these waves is still the speed of light  $c$ .

- (3) The direct observation of a *black hole* in 2017 which is a very peculiar solution of the Einstein Equations: the event horizon of the black hole at the center of M87 was directly imaged at the wavelength of radio waves by the EHT.

### 10.2.3 The meaning of the cosmological constant

The Einstein Equations in the form (10.20) permit models of universe that are not stationary. For that reason, Einstein introduced a further term in an improved version of his equations, *a posteriori* the only permitted by Lovelock's theorem,

$$\frac{8\pi G}{c^4}T_{ab} = Ric_{ab} - \frac{1}{2}g_{ab}S + \Lambda g_{ab} . \quad (10.24)$$

The idea was to obtain in that way a stationary cosmological model by suitably tuning the **cosmological constant**  $\Lambda$ . On the ground of the observational data, it was later evident the true physical nature of that constant: it should assume values different from Einstein's ones. The correct value is such that it *implies* a non-stationary universe according to modern observations where the universe is observed to be expanding in space and this expansion has an acceleration. Astronomical observations suggest that  $\Lambda > 0$  and that it has a very small value in natural units:

$$\Lambda^{-1/2} \sim 10^9 \text{ light years}. \quad (10.25)$$

This value is of the same order of magnitude as the size of the observable Universe. As a consequence, we expect that the effects of the cosmological constant in the Einstein Equations can be disregarded unless we deal with cosmological length scales. This observation is in agreement with the procedure we followed to determine the value  $k = \frac{8\pi G}{c^4}$  where we completely neglected the possible presence of  $\Lambda$  in the Einstein Equations.

We can move the term  $\Lambda g_{ab}$  to the left-hand side of the Einstein Equations by viewing it as a part of the cosmological stress energy tensor. As we shall see shortly, this is the case in the standard cosmological model we describe in the next section. However this interpretation faces a strong unresolved problem called *the cosmological constant problem*: quantum field theory, interpreting the energy density associated to  $\Lambda$  as the *quantum vacuum energy*, suggests that the value of  $\Lambda$  should be around  $10^{120}$  times the value above.

### 10.2.4 Diffeomorphism invariance

The picture is now complete. Within the framework of General Relativity, the physical objects are described in the spacetime  $M^4$  by tensor fields  $S$  associated to every material physical system (including radiation) and the gravitational interaction/metric is described by the metric  $\mathbf{g}$  accompanied by the associated curvature tensors. We know that this picture is just an approximation which is valid at macroscopic scales, since quantum phenomena cannot be encompassed

by this framework because other notions and mathematical structures must be added. Actually it is by no means clear if a complete theory may be constructed in that way. However let us stick to this macroscopic scenario to discuss a crucial point related to long standing foundational issues with General Relativity actually not yet completely solved, since several details are still open. The question is whether or not the mere *smooth manifold structure* of  $M^4$ , before assigning  $S$  and  $\mathbf{g}$ , possesses any physical meaning. The commonly shared opinion about this issue is that the smooth manifold structure of  $M^4$  has no physical significance. This apparently philosophical discussion has actually a deep impact on concrete problems related to the Einstein Equations, since it gives rise to the idea that there are *gauge transformations* in General Relativity similar to the ones present in electromagnetism when describing the fields in terms of the (relativistic) potentials  $A$  instead of the (relativistic) strength field  $F$ .

Consider a Generally Relativistic scenario  $(M^4, S, \mathbf{g})$  and a diffeomorphism  $\phi : M^4 \rightarrow M^4$ . We can use  $\phi$  to transform  $S$  and  $\mathbf{g}$  according to the general pullback map between tensor fields of definite type constructed out the pushforward  $d\phi^{-1} := d(\phi^{-1})$  of vector fields and the pullback  $\phi^*$  of 1-forms, and still generically denoted by  $\phi^*S$ , where

$$(\phi^*S)_p := \underbrace{d\phi^{-1} \otimes \cdots \otimes d\phi^{-1}}_{r \text{ times}} \otimes \underbrace{\phi^* \otimes \cdots \otimes \phi^*}_{s \text{ times}} S_{\phi(p)} .$$

We therefore find an apparently different scenario  $(M^4, \phi^*S, \phi^*\mathbf{g})$ . The assumption of **diffeomorphism invariance** of General Relativity is that  $(M^4, S, \mathbf{g})$  and  $(M^4, \phi^*S, \phi^*\mathbf{g})$  describe the same physics for every diffeomorphism  $\phi : M^4 \rightarrow M^4$ . This property can be fruitfully exploited, for instance, when looking for solutions of the Einstein Equations. However, the most important consequence is that only gauge-invariant quantities can represent physical observables. For instance, the values of  $R_{abcd}R^{abcd}$  at a certain event of  $M^4$  have no physical meaning in general. They acquire a meaning if the events are intrinsically characterized. For instance, the values of  $R_{abcd}R^{abcd}$  at the events where  $S = 0$ . This prescription defines a gauge-invariant observable.

Roughly speaking, the diffeomorphism invariance can be viewed as the possibility to arbitrarily fix a sistem of coordinate. This possibility, if  $n = 4$ , permits us to choose 4 of the independent 10 components of the metric, or to impose some conditions among the metric coefficients without lack of physical information. This opportunity is in particular exploited when dealing with gravitational waves and also when studying the problem of solving the Einstein Equations as a Cauchy problem with given initial data, for instance in the *ADM formalism*.

### 10.3 The birth of the cosmology as a science

This section is devoted to focus on some applications of the General Relativity, in particular arising from the Einstein Equations. General Relativity has received many experimental confirmations especially regarding the description of the behaviour of bodies and light in an external gravitational field, now viewed as curvature of the spacetime. In particular, the celebrated *deflection of light by the Sun* and the *gravitational redshift of light*. Here we are instead interested

in consequences of Einstein Equations which concerns the *production* of the gravitational interaction from the matter. In this context one of the most relevant merit of Einstein Equations is that they permitted the birth of cosmology as a science. Newtonian cosmology was inconsistent for many reasons (e.g. the Olbers paradox), whereas a cosmology based on Einstein's equations gave rise to amazing results.

### 10.3.1 The FLRW cosmological model

The relativistic model of cosmology relies upon the Einstein Equations and a pair of further assumptions.

- (a) At cosmological scales, the clusters of galaxies viewed as a gas has a stress-energy tensor  $T_{ab}$  with the form of a ideal fluid (8.16), we re-write here with a different notation,

$$T_{ab} = \rho' V_a V_b + p' \left( g_{ab} + \frac{V_a V_b}{c^2} \right). \quad (10.26)$$

where as we know,  $\rho'$  is the density of mass,  $p'$  the pressure, and  $V$  the 4-velocity of the galaxies.

- (b) A **Cosmological Principle** (whose a mathematical formulation was due to the eminent mathematician H. Weyl before the explicit construction of the model by physicists) is valid for the general form of the large-scale spacetime: the metric and the spatial distribution of matter in the universe turns out to be homogeneous and isotropic in an extended reference frame co-moving with the cluster of galaxies.

The arising model is called **FLRW model**, after A. Friedmann, G. Lemaitre, H.P. Robertson, and A.G. Walker. It is an exact solution of Einstein Equations (10.24) with the added parameter  $\Lambda$ . Let us illustrate the mathematical structure of this model which actually defines a natural cosmic extended reference frame according to Definition 8.23.

1. The spacetime is postulated to be a 4-dimensional Lorentzian manifold diffeomorphic to  $(\alpha, \omega) \times \Sigma^{(k)}$  and the metric reads

$$\mathbf{g}_{FLRW} = -c^2 dt \otimes dt + a(t)^2 \mathbf{h}^{(k)}$$

where  $t \in (\alpha, \omega)$  and the metric on the said spatial section  $\Sigma_t$  (each diffeomorphic to  $\Sigma^{(k)}$ ) is

$$\mathbf{h}^{(k)} = \frac{1}{1 - kr^2} dr \otimes dr + r^2 (d\theta \otimes d\theta + \sin^2 \theta d\phi \otimes d\phi)$$

with  $x^0 = ct$  and  $x^1 = r, x^2 = \theta, x^3 = \phi$  formal spherical polar coordinates which cover (almost) completely the spatial sections. Finally the **scale parameter**  $a$  is fixed as  $a(t_0) = 1$  where  $t_0$  is “now”.

(1.1) The parameter  $k$  can only constantly take one of the values  $-1, 0, +1$  corresponding to the three *maximally symmetric* 3-dimensional Riemannian manifolds (see Remark 9.14) describing  $\Sigma_t$ .

- (a) The 3D *Hyperbolic space* also known as the *Lobachevskij space* for  $k = -1$ ,
- (b) the flat Euclidean 3D space for  $k = 0$ ,
- (c) the 3-sphere for  $k = 1$ .

More generally, we can also replace these spaces for quotients space with respect the discrete subgroup of isometries of the said Riemannian manifolds (e.g., a 3-torus in palce of the Eucliden space). The choice of the above spatial metric is due to the requirements of **homogeneity** and **isotropy**. We can focus on the metric  $\mathbf{h}^{(k)}$  since the factor  $a(t)$  is constant at every fix time and it does not affect the observations below which are pertinent for every fixed  $t$ . Thinking of  $(\Sigma_t, \mathbf{h}^{(k)})$  as an orientable manifold,  $\mathbf{h}^{(k)}$  has the following properties (and it is possible to prove that they charachterize completely it up to quotients with respect to discrete subgroups of isometries).

- (i) For every  $p, q \in \Sigma_t$ , there exist a orientation-preserving (Definition 2.17) isometry  $f : \Sigma_t \rightarrow \Sigma_t$  such that  $f(p) = f(q)$ . This property is the mathematical description of the physical notion of *homogeneity* of the metrical structures.
  - (ii) For every  $p \in \Sigma_t$ , the subgroup of the isometries that leave fixed  $p$  includes a subgroup which is isomorphic to  $O(3)^3$ . The isometries representing  $SO(3)$  are orientation preserving, whereas the one representing  $-I \in O(3)$  changes the orientation of the manifold. This property is the mathematical description of the physical notion of *isotropy* of the metrical structures.
- (1.2) Physically speaking, the integral lines of the future-directed vector field  $\frac{\partial}{\partial t}$  are the world-lines of the clusters of galaxies (it is therefore parallel to  $V$ ) and  $t$  is the proper time measured at rest with the clusters. The spacelike surfaces  $\Sigma_t$  crossed by the worldlines of the clusters of galaxies are the rest-space at time  $t$  co-moving with these clusters. (Notice that the time vector is orthogonal to the spatial sections, so the speed of light is  $c$ .)
- (1.3) On surfaces  $\Sigma_t$ , the large-scale density of energy  $\rho c^2$  (by definition measured at rest with the galaxies) and the pressure  $p$  is assumed to be homogeneous and isotropic as the metric structures, in other words,  $\rho$  and  $p$  are constant on every  $\Sigma_t$ , though they depend on  $t$ .

**2.** The evolution equations of  $\mathbf{g}_{FLRW}$  and the pressure  $p$  and the density of mass  $\rho$  are obtained from (10.24) with the source (10.26) and using the said metric. Of the 10 equations obtained

---

<sup>3</sup> $O(3)$  is a smooth manifold it being a Lie group. Here we are assuming that the map  $f_p : O(3) \times \Sigma_t \rightarrow \Sigma_t$  representing the said action of  $O(3)$  on  $\Sigma_t$  is smooth as well and that  $O(3) \ni R \mapsto f_p(R, \cdot) \in D(\Sigma_t)$  is injective, it preserves the group structure, and every element in its image is an isometry.



from (10.24), only 2 are independent and they are known as the **Friedmann equations**:

$$\left(\frac{\dot{a}(t)}{a(t)}\right)^2 = -\frac{kc^2}{a(t)^2} + \frac{8\pi G}{3} \left(\rho'(t) + \frac{\Lambda c^2}{8\pi G}\right), \quad (10.27)$$

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3} \left(\rho'(t) + \frac{\Lambda c^2}{8\pi G}\right) - \frac{4\pi G}{c^2} \left(p'(t) - \frac{\Lambda c^4}{8\pi G}\right). \quad (10.28)$$

These equations are obtained by the Einstein ones as follows. A straightforward computation proves that

$$R_{00} = -3\frac{\ddot{a}}{a}, \quad R_{\alpha\beta} = (a\ddot{a} + 2\dot{a}^2 + 2k)h_{\alpha\beta}, \quad R_{0\alpha} = R_{\alpha 0} = 0, \quad \alpha, \beta = 1, 2, 3, \quad (10.29)$$

and also

$$R = 6 \left( \frac{\dot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{k}{a^2} \right). \quad (10.30)$$

Using  $V_a = -c\delta_{a0}$  in (10.26), (10.27) is nothing but 00 component of the Einstein equations whereas (10.28) is obtained from the  $\alpha\beta$  components and taking the former equation into account. The form of the two Friedmann equations suggests to interpret

$$\rho(t) := \rho'(t) + \frac{\Lambda c^2}{8\pi G}, \quad p(t) := p'(t) - c^2 \frac{\Lambda c^2}{8\pi G} \quad (10.31)$$

as effective density and pressure with a contribution of the cosmological constant. In other words, the addend containing the cosmological constant in the Einstein Equations (10.24) is moved from the right-hand side to the left-hand side and it gives rise to a new source of the gravitational field,

$$T_{\Lambda}^{ab} = -\frac{\Lambda c^4}{8\pi G} g^{ab}. \quad (10.32)$$

This “modern” interpretation will play a crucial role later.

We stress that the solutions in  $a$  also fix their maximal interval  $(\alpha, \omega)$  of definition, which may result bounded from below or from above.

**3.** The Friedmann equations are accompanied with an **equation of state**, also known as **constitutive equation**, connecting the (effective) pressure  $p$  and the (effective) density of mass  $\rho$  as in (10.31).

(3.1) The simplest case, that permits to solve completely the equations above in the functions of time  $a$ ,  $\rho$ ,  $p$ , is

$$p = w\rho c^2, \quad (10.33)$$

where  $w$  is a constant.

(3.2) Realistic models consider **mixtures** of several non-interacting fluids. Each component has an equation of state as above with its own constant  $w$ . The density and the pressure are respectively the sum of the partial densities and pressures of the components.

(3.3) The equation of conservation  $\nabla_a T^{ab} = 0$  produces in particular the identity

$$\dot{\rho} = -3\frac{\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right). \quad (10.34)$$

This equation must be also a consequence of the Friedmann equations, in fact it also arises from the first Friedmann equation used in the second one.

(3.4) Each component of the fluid satisfies (10.34) separately when thinking the various parts independent.

4. The constant  $k$ , the scale factor  $a = a(t) > 0$ , the functions  $p = p(t)$ ,  $\rho = \rho(t)$  are finally determined from the two Friedmann equations and (10.33), and “initial conditions” at  $t = t_0$  grasped from the present astronomical observations.

Also (10.34) is useful in determining the solutions. Actually the evolution of  $\rho$  and  $p$  easily arise from the one of  $a$  just using (10.34) and the constitutive equation. Indeed, taking advantage of (10.33), equation (10.34) becomes

$$\frac{\dot{\rho}}{\rho} = -3(1+w)\frac{\dot{a}}{a},$$

which can be integrated giving rise to

$$\rho(t) = \rho(t_0)a(t)^{-3(1+w)}, \quad p(t) = w\rho(t_0)c^2a(t)^{-3(1+w)}. \quad (10.35)$$

where  $a(t_0) = 1$  and  $t_0$  is “now”.

From the experimental side, we can say that, by combining the observation data from *WMAP* and *Planck* with theoretical results, astrophysicists now agree that the universe is almost completely homogeneous and isotropic and thus well described by a FLRW spacetime. Homogeneity shows up when averaging over a very large scale greater than<sup>4</sup>  $400Mpc$ , the diameter of the observed universe being  $28.5Gpc$ .

### 10.3.2 The observed expansion of the Universe

According to experimental observations, the large-scale rest space is “expanding” and this expansion can be easily described in the FLRW model. In other words  $a : (\alpha, \omega) \ni t \mapsto (0, +\infty)$  is an *increasing* function of  $t$ . The net effect is that galaxies move away from each other isotropically. This fact is strongly corroborated by experimental observations by Hubble already in 1922 who discovered that the galaxies move away from Earth in all directions by observing the *red-shift* in the spectra of the emitted light, and that their speed increases proportionally to their distance from Earth as encoded in *Hubble law* we are going to discuss. The ratio

$$H(t) := \frac{\dot{a}(t)}{a(t)} \quad (10.36)$$

---

<sup>4</sup> $1pc \sim 31 \times 10^{15}m$ .

is called **Hubble parameter**. Its present value

$$H_0 := H(t_0) > 0 \quad (10.37)$$

– where  $t_0$  as usual is “now” in the cosmological context – is (improperly) called **Hubble constant**, even if it is not a constant! From (10.36), we have that (we write explicitly  $a(t_0)$  even if  $a(t_0) = 1$ )

$$\dot{a}(t_0) = H_0 a(t_0) .$$

If a galaxy has distance  $d(t)$  from us at time  $t$ , due to the uniform expansion of the scale parameter  $a$ , we have that

$$\frac{d(t)}{d(t_0)} = \frac{a(t)}{a(t_0)} .$$

As a consequence,

$$\left. \frac{\dot{d}(t)}{d(t_0)} \right|_{t=t_0} = \frac{\dot{a}(t_0)}{a(t_0)} = H_0$$

and we find the famous **Hubble law**

$$\dot{d}(t_0) = d(t_0) H_0 > 0 , \quad (10.38)$$

which says that the speed of a galaxy measured on  $\Sigma_{t_0}$  is proportional to its distance from Earth (or from any other observer evolving with an integral line of  $\frac{\partial}{\partial t}$ ) measured on the present rest space of the FLRW reference frame. The estimate value of  $H_0$ , is

$$H_0 = 2.2 \times 10^{-18} s^{-1} .$$

Notice that this notion of speed permits in principle values larger than  $c$ .

The value of  $\dot{d}(t_0)$  is very difficult to be directly measured because we actually are only able to observe signals reaching us from the past, in view of the finite propagation speed of the light! It can be indirectly measured obtained by measuring a *red shift phenomenon* associated to the expansion itself in view of the following argument.

First of all we introduce a vector field which almost is a Killing vector. Let us introduce the **conformal time**

$$\eta(t) := \int_{t_0}^t \frac{du}{a(u)} \quad (10.39)$$

Replacing the coordinate  $t$  for  $\eta$ , keeping the spatial coordinates, and writing  $a$  as a function of  $\eta$  by inverting the function above, the FLRW metric reads

$$\mathbf{g}_{FLRW} = a(\eta)^2 \left( -c^2 d\eta \otimes d\eta + \mathbf{h}^{(k)} \right) .$$

As a consequence, if  $K := \frac{\partial}{\partial \eta}$ ,

$$\mathcal{L}_K \mathbf{g}_{FLRW} = \mathcal{L}_K a(\eta)^2 \left( -c^2 d\eta \otimes d\eta + \mathbf{h}^{(k)} \right) = K(a(\eta)^2) \left( -c^2 d\eta \otimes d\eta + \mathbf{h}^{(k)} \right) + \frac{1}{a(\eta)^2} 0 .$$

Hence

$$\mathcal{L}_K \mathbf{g}_{FLRW} = K (a(\eta)^2) \frac{1}{a(\eta)^2} \mathbf{g}_{FLRW} = \frac{2}{a(\eta)} \frac{da}{d\eta} \mathbf{g}_{FLRW} = 2\dot{a}(t) \mathbf{g}_{FLRW}$$

where we used (10.39) in the last passage. Taking (7.30) into account, we have found that

$$\nabla_a K_b + \nabla_b K_a = 2\dot{a}(t) \mathbf{g}_{FLRW} . \quad (10.40)$$

This equation, in the general context of a (pseudo) Riemannian manifold, defines a **conformal Killing vector field**, when the right-hand side is a generic smooth scalar function multiplied with the metric. Hence the conformal time  $\eta$  defines a conformal Killing vector  $K = \frac{\partial}{\partial \eta}$ . A conformal Killing vector can be used to define conserved quantities for massless particles or extended systems whose stress energy tensor has zero trace (as the electromagnetic field). Let us illustrate the first case in our context. We assume that a particle of light, whose geodesic is parametrized with the affine parameter  $s$ , is emitted from a far galaxy at  $s = s_e$  and it is received here on Earth at  $s = s_0$  (“now”) and that  $P$  is the four momentum of the said photon ((4) in Remark see 8.18).  $P$  is paralelly transported with respect to itself along the geodesic and more precisely it is the tangent vector referred to an affine parametrization of the geodesic we can identify with  $s$  itself. Using also this fact, we have that the quantity  $K_a P^a$  is constant along the geodesic because

$$\begin{aligned} \frac{d}{ds} K_a P^a &= P^c \nabla_c (K_a P^a) = (\nabla_c K_a) P^c P^a + K_a P^c (\nabla_c P^a) \\ &= \frac{1}{2} (\nabla_a K_c + \nabla_c K_a) P^c P^a + 0 = 2\dot{a} P_c P^c + 0 = 0 . \end{aligned}$$

On the other hand, using the defintion of  $K = \frac{\partial}{\partial \eta} = a \frac{\partial}{\partial t}$  and (4) in Remark see 8.18,

$$K_a P^a = a(t(s)) \hbar \omega(\gamma(s)) = \text{constant} .$$

If  $f := \frac{\omega}{2\pi}$  is the frequence of the light, we conclude that the emitted frequency  $f(\gamma(s_e))$  and the received frequency  $f(\gamma(s_0))$  are therefore related by<sup>5</sup>

$$a(t_0) f(\gamma(s_0)) = a(t_e) f(\gamma(s_e)) ,$$

where  $t_0 > t_e := t(s_e)$ . So that what we can measure is the **red-shift parameter**

$$z := \frac{f(\gamma(s_e)) - f(\gamma(s_0))}{f(\gamma(s_e))} = 1 - \frac{a(t_e)}{a(t_0)} > 0 . \quad (10.41)$$

We remark that  $z > 0$  arises from the fact that  $a(t_0) > a(t_e)$  *since the Universe expands while the light travels from  $\gamma(s_e)$  to  $\gamma(t_0)$* . Positivity of  $z$ , measured along all directions around Earth,

---

<sup>5</sup>A different procedure similar to the one used for computing the gravitational red-shift that uses the fact that the vacuum Maxwell equations do not explicitly depend on  $\eta$ , leads to the same result.

is a qualitative proof of the isotropic expansion.

**Remark 10.14.** We can measure the red-shift because we know both  $f(\gamma(0))$  and  $f(\gamma(1))$  even if the former may seem inaccessible since it is the emission frequency in a galaxy far from us. Actually we possess this information as soon as we assume that the far galaxies are made of the same type of matter (chemical elements) as the stars in our galaxy. Each chemical element emits a very precise spectrum of frequencies

$$(f_1, f_n, \dots)$$

called the *emission spectrum* of that element which is nothing but a physical *signature* of it. Observing the spectral decomposition of the light emitted by a far galaxy we find the same spectra of each chemical element present in the stars of our galaxy multiplied with a common factor:

$$(f'_1, f'_n, \dots) = (rf_1, rf_n, \dots)$$

That common factor  $r < 1$  is just the ratio  $a(t(1))/a(t(0))$ . Obviously life is not so easy and there are many other local phenomena (gravitational red shift, kinematical Doppler phenomena etc) that may further modify the observed spectra, however the fundamental argument is the one outlined above. ■

To go on, assuming that  $t(1) - t(0)$  is “small”, we can approximate the right-hand side of (10.41) by the approximating denominator with Taylor expansion

$$a(t_e) = a(t_0) + (t_e - t_0)\dot{a}(t_0) + O(2) \simeq a(t_0) [1 - (t_e - t_0)H(t(0))] ,$$

so that, from (10.41),

$$z \simeq (t_0 - t_e)H_0 .$$

If the distance between Earth and the emitter galaxy is not too large, we can estimate the time  $t_0 - t_e$  interval as simply the distance measured now ( $t = t_0$ ) divided by the speed of light:

$$z \simeq \frac{d(t_0)}{c}H(t_0) . \tag{10.42}$$

From (10.38) we have an estimate of the speed of the galaxies (measured with the distance in  $\Sigma_t$ ) as a function of the red-shift parameter:

$$\dot{d}(t_0) \simeq cz . \tag{10.43}$$

**Remark 10.15.** The observed expansion of the Universe proves that at large scale the gravitational interaction is *repulsive* instead that attractive as we know very well. This fact is a completely new feature of the gravitational interaction due to the relativistic model of the gravitational interaction. ■

### 10.3.3 Evidence of spatial flatness: $k = 0$

Theoretically speaking, it is useful to introduce the **density parameter**

$$\Omega(t) := \frac{\rho(t)}{\rho_c(t)} \quad \text{where} \quad \rho_c(t) := \frac{3H(t)^2}{8\pi G} \quad (10.44)$$

is the **critical density**. The measured value of  $\rho_c(t_0) = 8.5 \times 10^{-27} \text{Kg/m}^3$ . The theoretical relevance of this quantity is due to the fact that the first Friedmann equation can be re-arranged to

$$\Omega(t) - 1 = \frac{k}{H(t)^2 a(t)^2}.$$

Hence we can conclude that the following possibilities exist, where we stress that  $\rho$  includes the contribution of the cosmological constant,

- (a)  $\rho(t) < \rho_c(t)$  is equivalent to  $\Omega(t) < 1$  which means  $k = -1$ : *open hyperbolic universe*.
- (b)  $\rho(t) = \rho_c(t)$  is equivalent to  $\Omega(t) = 1$  which means  $k = 0$ : *open flat universe*.
- (a)  $\rho(t) > \rho_c(t)$  is equivalent to  $\Omega(t) > 1$  which means  $k = +1$ : *closed spherical universe*.

Notice that  $k$  cannot change in time, so the if one of the above possibilities is valid at a certain time  $t$ , it must be always valid. From the observational side, the spatial geometry on the rest spaces  $\Sigma_{t_0}$  appears to be very close to the flat case  $k = 0$ , which is the same as saying that  $\rho(t_0) = \rho_c(t_0)$ .

**Remark 10.16.** “Open” refers above to the simplest (simply-connected) manifold  $\Sigma_t$  with  $k = -1, 0$ . However the shape of the spatial section of the Universe can be still “closed” (compact) also if  $k = -1, 0$  just by taking the quotient with respect to a discrete subgroup of isometries. ■

### 10.3.4 Big bang, cosmic microwave background, dark matter, and dark energy

From the late 1990s there is a standard cosmological model that includes both the observational information and the mathematical theory based on the FLRW geometric model and the Friedmann equations discussed above, it is known as the  **$\Lambda$ CDM model**. It was later extended by adding the *cosmological inflation*. It is defined by fixing all the free parameters previously discussed (the constants  $w$  of the mixture) and the initial conditions at some time, typically the present one. This model embodies several remarkable features of our Universe, some of them directly observed. We immediately quote a triple of interesting features of the  $\Lambda$ CDM model.

- (a) In the far past, around 13.8 billion years ago, measured with the proper time of the galaxies, all worldlines join at a single event where  $a = 0$ , thus of infinite curvature, called the **Big Bang**. It happened at the *finite* initial value  $\alpha$  of the interval of the argument of  $a : (\alpha, \omega) \ni t \mapsto (0, +\infty)$ , whereas  $\omega = +\infty$ .

- (b) There is a relic from the Big Bang called **cosmic background radiation**. One component is the **cosmic microwave background** which is a large-scale homogeneous and isotropic thermal (black body) radiation nowadays of the value of around  $3K$ . This is due to redshifted photons that have freely streamed from a cosmological epoch when the Universe was transparent for the first time to radiation. Its discovery in 1965 (by chance<sup>6</sup>) together with a number of detailed observations of its properties are considered nowadays one of the major confirmations of the Big Bang hypothesis.
- (c) From the big bang on,  $a(t)$  continued to increase with very different regimes: accelerating (inflation) in the first instants after the Big Bang, and later still decelerating and accelerating. Nowadays we are in an accelerating regime. All that is in agreement with the contemporary observations about the expansion of the Universe.

The model also explain the large-scale structure in the distribution of galaxies, the observed abundances of hydrogen (including deuterium), helium and lithium, and – as already stressed – the accelerating expansion of the universe observed in the light from distant galaxies and supernovae.

As a matter of fact, almost all said above<sup>7</sup> relies in particular on assuming that the source of the Friedmann equations consists of a mixture of ideal fluids with 4 *components*, whose features are inferred by direct or indirect experimental data (in particular by the space observatory *Planck*)

- (1) **Dark energy**: the energy density of this component *now* should amount to the 69.1% of the total energy density of the universe. It has a quite anomalous equation of state with  $w = -1$  involving a *negative* pressure:

$$p = -\rho c^2 .$$

In fact, it is described by the term  $\Lambda g_{ab}$  in (10.24) transferred in the left hand side and interpreted as a part of  $T_{ab}$ . This dominant term is responsible in the solution of Friedmann's equations to a positive acceleration  $\ddot{a}$  of the scale function  $a$ . This acceleration is an experimental fact: the observed expansion of the universe is accelerating!  $\Lambda > 0$  is tuned in the  $\Lambda$ CDM model just to account for this observed accelerated expansion.

- (2) **Cold dark matter** (CDM): the energy density of this component *now* should amount to the 25.9% of the total energy density of the universe. It is postulated in order to account for gravitational effects observed in very large-scale structures (as anomalous rotational curves of galaxies, anomalous gravitational lensing of light by galaxy clusters etc.) that cannot be explained by the observed quantity of standard matter. Its equation of state assumes *zero* pressure because  $w = 0$

$$p = 0 .$$

---

<sup>6</sup>By A. Penzias and R. Woodrow Wilson who measured the temperature to be around  $3K$ . Later, R. Dicke, P. J. E. Peebles, P. G. Roll and D. T. Wilkinson interpreted this radiation as a signature of the Big Bang.

<sup>7</sup>We omit to give details about the inflation mechanism which involves notions of quantum field theory.

- (3) **Ordinary matter:** the energy density of this component *now* should amount to the 4.1% of the energy density of the universe. It has the equation of state involving *zero* pressure for stars and intergalactic gas

$$p = 0 .$$

- (4) **Radiation:** Another independent part of the mixture has equation of state arising from  $w = \frac{1}{3}$ ,

$$p = \frac{1}{3}\rho c^2 .$$

This is a very negligible fraction made of massless particles, i.e., photons<sup>8</sup>, whose total energy density is *now* of  $\sim 3 \times 10^{-4}$  the energy density of ordinary matter. The value  $w = \frac{1}{3}$  for the radiation component is obtained by observing that the stress energy tensor of the EM field (8.17) satisfies  $T_a^a = 0$ . When assuming the same constraint of the corresponding stress-energy tensor (10.26) we find  $T_a^a = -\rho c^2 + 3p$ .

**Remark 10.17.** Ordinary matter and radiation are the only parts of the mixture actually directly detected in astronomical observations. ■

To see some consequences of these hypotheses, let us consider the first Friedmann equation. It can be specialized to

$$\left(\frac{\dot{a}(t)}{a(t)}\right)^2 = -\frac{kc^2}{a(t)^2} + \frac{8\pi G}{3}(\rho_m(t) + \rho_r(t) + \rho_\Lambda(t))$$

where  $\rho_m$  is the density of energy due to ordinary matter and dark matter,  $\rho_r$  is the energy density due to radiation and  $\rho_\Lambda$  is the density of energy due to dark energy. Taking the first equation in (10.35) into account

$$\left(\frac{\dot{a}(t)}{a(t)}\right)^2 = \frac{8\pi G}{3} \left[ -\frac{3kc^2}{8\pi G} a(t)^{-2} + \rho_m(t_0)a(t)^{-3} + \rho_r(t_0)a(t)^{-4} + \rho_\Lambda(t_0) \right] .$$

Now we neglect the term with the constant  $k$  which we know to be very small, obtaining

$$\left(\frac{\dot{a}(t)}{a(t)}\right)^2 = \frac{8\pi G}{3} [\rho_m(t_0)a(t)^{-3} + \rho_r(t_0)a(t)^{-4} + \rho_\Lambda(t_0)] ,$$

Making use of (10.37) and (10.44), we can recast this equation to

$$\left(\frac{\dot{a}(t)}{a(t)}\right)^2 = H_0 [\Omega_m(t_0)a(t)^{-3} + \Omega_r(t_0)a(t)^{-4} + \Omega_\Lambda(t_0)] ,$$

i.e.,

$$\frac{da}{dt} = a(t) \sqrt{H_0 [\Omega_m(t_0)a(t)^{-3} + \Omega_r(t_0)a(t)^{-4} + \Omega_\Lambda(t_0)]} . \quad (10.45)$$

---

<sup>8</sup>Neutrinos could be added but their treatment needs particular care in view of their oscillating masses.



All constants above are directly or indirectly known

$$H_0 \sim 2.2 \times 10^{-18} s^{-1}, \quad \Omega_m(t_0) \sim 0.308, \quad \Omega_r(t_0) \sim 10^{-4}, \quad \Omega_\Lambda(t_0) \sim 0.691. \quad (10.46)$$

Neglecting also the radiation term as  $\Omega_r(t_0) \sim 10^{-4}$ , equation (10.47) becomes

$$\frac{da}{dt} = a(t) \sqrt{H_0 [\Omega_m(t_0) a(t)^{-3} + \Omega_\Lambda(t_0)]}. \quad (10.47)$$

Since the right-hand side of this equation in normal form is smooth for  $a > 0$ , it has a unique maximal solution in that region for the initial condition  $a(t_0) = 1$ . It is

$$a(t) = \left( \frac{\Omega_m(t_0)}{\Omega_\Lambda(t_0)} \right)^{1/3} \sinh^{2/3} \left( \frac{t - t_{bb}}{t_\Lambda} \right) \quad \text{for } t > t_{bb}, \quad (10.48)$$

where

$$t_\Lambda := \frac{2}{3H_0\Omega_\Lambda(t_0)^{1/2}}$$

and we have set  $t_{bb} \in \mathbb{R}$  such that

$$1 = \left( \frac{\Omega_m(t_0)}{\Omega_\Lambda(t_0)} \right)^{1/3} \sinh^{2/3} \left( \frac{t_0 - t_{bb}}{t_\Lambda} \right).$$

It is evident from the shape of the function  $y = \sinh x$ , for  $x \geq 0$ , that the above condition is always satisfied for a unique  $t_{bb} \in \mathbb{R}$  and that it also satisfies  $t_{bb} < t_0$ . With the data (10.46), it turns out that

$$t_0 - t_{bb} \simeq 13.8 \times 10^9 \text{ years}.$$

Notice that  $a(t_{bb}) = 0$ : that is the Big Bang. In other words, the Big Bang was around  $13.8 \times 10^9$  years ago. Studying the second derivative  $\ddot{a}(t)$  of the specific solution (10.48), one easily sees that it becomes positive for  $t > t_p$  such that

$$a(t_p) = \left( \frac{\Omega_m(t_0)}{2\Omega_\Lambda(t_0)} \right)^{1/3}. \quad (10.49)$$

It turns out that  $a(t_p) \sim 0.6$ . This means that, in this model, the expansion is now accelerating after a phase of deceleration from the Big Bang. Notice that the responsibility for this acceleration is of  $\Lambda$  according to the previous remark. Another way to see it is the following one: setting  $\Lambda \rightarrow 0^+$  we have that  $a(t_p) \rightarrow +\infty$ .

The model above is valid for sufficiently large times when the contribution of  $\Lambda$  and  $\Omega_m$  are dominant with respect to the radiation. For very small values of  $a(t)$ , we cannot neglect the radiation term which becomes dominant with respect to the matter term and the  $\Lambda$  term in (10.45) as it increases as  $a^{-4}$  whereas the one of the matter increases as  $a^{-3}$  and  $\Lambda$  furnishes a constant term. Hence the dynamics is expected to be more complicated than the one described by (10.48) for small  $a$ . One may think that the presence of an initial the Big Bang is therefore

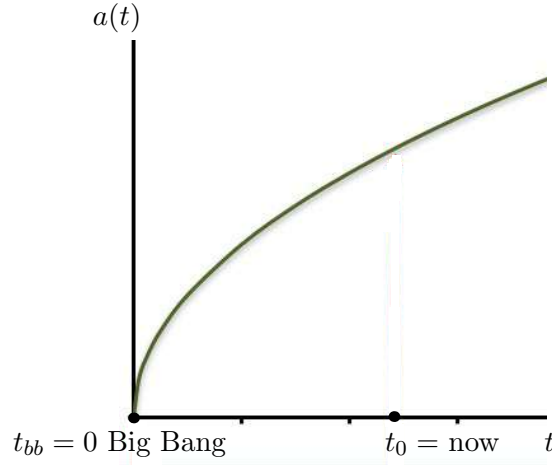


Figure 10.2: Evolution of  $a(t)$  with the origin of  $t$  re-arranged in order to have the Big Bang at  $t_{bb} = 0$ .

disputable. Actually, the Big Bang is a quite common feature of the evolution of  $a$  if  $\rho$  and  $p$  (defined in (10.31)) were positive before some time  $t_p < t_0$ . In fact, let us focus on the evolution of  $a$  for a universe containing an effective fluid with  $\rho > 0$  and  $p \geq 0$  (i.e., matter, radiation, and dark matter are dominant with respect to the dark energy before  $t_p$ ). With these hypotheses, from (10.28),

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}(\rho + 3p(t)) < 0 \Rightarrow \ddot{a}(t) < 0.$$

If we also assume that the universe has been expanding  $\dot{a}(t) > 0$ , then the curve  $a = a(t) \geq 0$  traced back in time necessarily reaches a singularity  $a(t_{bb}) = 0$  at some  $t_{bb} < t_p$  (see Fig. 10.17).

An overall *caveat* is however that it is not physically meaningful to assume that the components of the fluid are really independent. In particular, radiation and matter must have strongly interacted in the past, so that the previous rough model has to be refined into a more sophisticated scenario where all interactions between the various components of the cosmic fluid and their different quantum nature are taken into account. When the size of the Universe was very small, fundamental quantum phenomena of matter played a crucial role. Furthermore, it seems that very close to the Big Bang, a first accelerating expansion took place called *inflation* (see the next paragraph). This first accelerating expansion needs a more sophisticated model than the one outlined above and all known proposals to explain it are essentially of quantum nature. That period and, *a fortiori*, the instants before the cosmic inflation are still matter of discussion since no reliable theory of quantum gravity exists for the moment.

### 10.3.5 Open problems

An overall problem with the  $\Lambda$ CDM model above is evident. The largest part of the source of gravitational interaction made of dark energy and dark matter is *postulated to exist* just in view of the observations of the effects they would imply on ordinary matter when assuming the validity of Einstein's equations. No direct observation exist of dark energy and dark matter because, up to now, it seems that this sort of matter does not interact in any way with standard matter excluding the gravitational interaction. A number of proposals exist to explain the nature of the dark energy and the dark mass<sup>9</sup>. But also other proposals assume that the Einstein equations are incomplete or wrong<sup>10</sup> at large scales and the dark where the model matter and energy simply do not exist. There are also other problems with the apparently wrong predictions of the model at relatively small scales (sub-galaxy scale), which however could be related with the weird properties of dark matter.

Another important problem we only mention is the **cosmological horizon problem**. The cosmic microwave background appears to possess the same temperature in all directions and it is affected by very small perturbations. This experimental fact raises a serious problem. If we trace back the origin of that radiation assuming an expansion of the spatial section of the universe from the Big Bang to now without acceleration, then we discover that the radiation that reached us from different directions was produced in causally separated regions. How is it possible that these regions had the same temperature without causal interactions? **Cosmological inflation** is a proposal of solution to that problem. Within this theory, immediately after the Big Bang, a first isolated rapid exponential expansion of space, with  $\ddot{a} > 0$ , took place in the very early universe: from  $10^{-36}$  seconds after the Big Bang to some time between  $10^{-33}$  and  $10^{-32}$  seconds after the singularity. During the inflation period the parameter  $a$  expanded by a fantastic factor of at least  $10^{26}$ . The universe increased in size from a small and causally connected region in near equilibrium. Inflation then expanded the universe very rapidly, isolating nearby regions of spacetime by growing them beyond the limits of causal contact, freezing the uniformity at large distances. After that inflationary period, the universe continued to expand at a slower rate decelerating. As said above, the second acceleration of the expansion due to dark energy started after  $9 \times 10^9$  years and it is still present. The nature of the first accelerating expansion called inflation is not completely clear and there are several proposals to explain it especially from quantum field theory.

---

<sup>9</sup>J.P. Ostriker and S. Mitton *Heart of Darkness: Unraveling the mysteries of the invisible universe*. Princeton, Princeton University Press (2013).

<sup>10</sup>A. Maeder. *An Alternative to the  $\Lambda$ CDM Model: The Case of Scale Invariance*. The Astrophysical Journal 834 (2): 194 (2017).

# Bibliography

- [BEE96] J. K. Beem, P. E. Ehrlich, and K. L. Easley, *Global Lorentzian geometry*, 2nd ed., Marcel Dekker, Inc., New York, (1996).
- [doC92] M.P.do Carmo, *Riemannian Geometry*, Birkhäuser, Boston 1992.
- [Coh80] D. L. Cohn, *Measure Theory*, Birkhäuser, Boston 1980.
- [HaEl75] S. W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Space-Time*, Cambridge University Press (1975)
- [KhMo15] I.Khavkine and V. Moretti, *Algebraic QFT in Curved Spacetime and quasifree Hadamard states: an introduction* Chapter 5 of *Advances in Algebraic Quantum Field Theory* Springer 2015 (Eds R. Brunetti, C. Dappiaggi, K. Fredenhagen, and J.Yngvason)
- [KoNo96] K. Kobayashi and S. Nomizu: *Foundations of Differential Geometry*. Vol I, (1996)
- [Min19] E. Minguzzi: *Lorentzian causality theory*. Living Reviews in Relativity (2019) 22:3 <https://doi.org/10.1007/s41114-019-0019-x>
- [LaLi80] L.D. Landau and E.M. Lifshitz, *The Classical Theory of Fields*, 4th Revised English Edition, Butterworth-Heinemann (1980)
- [Lee03] J. M. Lee, *Introduction to Smooth Manifolds*, Springer, New York, 2003
- [Mor20] V. Moretti, *Multi-Linear Algebra, Tensors, Spinors with applications to Special Relativity*, Lecture Notes of the University of Trento, version 2020 <http://www.science.unitn.it/%7Emoretti/tensori.pdf>
- [Mor20b] V. Moretti, *Meccanica Analitica: Meccanica Classica, Meccanica Lagrangiana e Hamiltoniana e Teoria della Stabilità*, Springer-Verlag, Milano, (2020)
- [Morettia] V. Moretti, *Spectral Theory and Quantum Mechanics: Mathematical Foundations of Quantum Theories, Symmetries and Introduction to the Algebraic Formulation*, 2nd Edition, Springer, 2017

- [Morettib] V. Moretti, *Fundamental Mathematical Structures of Quantum Theory Spectral Theory: Foundational Issues, Symmetries, Algebraic Formulation*, Springer 2019
- [O'Ne83] B. O'Neill *Semi-Riemannian Geometry With Applications to Relativity*. Academic Press (1983)
- [Pos01] M.M. Postnikov, *Geometry VI Riemannian Geometry* Encyclopaedia of Mathematical Sciences Geometry. Springer 2001
- [Shi72] G. E. Shilov *Mathematical Analysis, Functions of Several Variables*; Parts 1-2, Nauka, Moscow, 1972 (Russian). French translation: *Analyse Mathématique. Fonctions de Plusieurs Variables Réelles; Première et deuxième parties*, Mir, Moscou, 1975
- [Rud87] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, 1987.
- [Seri90] E. Sernesi, *Geometria I and II*, Bollati Boringhieri, Torino, 1990.
- [Wal84] R.M. Wald, *General Relativity*, Chicago University Press, Chicago, 1984.
- [War83] F.W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, Springer 1983 Interscience, New York, 1963.