Tensor Analysis on Manifolds in Mathematical Physics with Applications to Relativistic Theories.

by **Valter Moretti**
www.science.unitn.it/∼moretti/home.html

Department of Mathematics,
Faculty of Science,
University of Trento
Italy

Contents

Acknowledgments.

The author is grateful to Dr. Riccardo Aramini who read these notes carefully correcting several misprints and mistakes. The author would like to thank Gian Pietro Peruzzi for corrections in some proofs.

Chapter 1

Basics on differential geometry: topological and differentiable manifolds.

1.1 Basics 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.

1.1.1. We recall 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 $X_1, \ldots, X_n \in \mathcal{T}$, then $\cap_{i=1,\ldots,n} X_i \in \mathcal{T}$.
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. $\mathbb{R}^n$.
1.1.2. 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 $U \subset K$.
1.1.2. If X is a topological space and $f : X \to \mathbb{R}$ is any function, the **support** of f, $\mathrm{supp} f$, is the closure of the set of the points $x \in X$ with $f(x) \neq 0$.
1.1.3. If $(X, \mathcal{T})$ and $(Y, \mathcal{U})$ are topological spaces, a mapping $f : X \to Y$ 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 $f : X \to Y$, 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 $h : X \to Y$, $U \subset X$ is either open or closed if and only if $h(U) \subset Y$ is such.

1.1.4. If $(X, \mathcal{T})$ is a topological space, a class $\mathcal{B} \subset \mathcal{T}$ is called **basis of the topology**, if each open set turns out to be union of elements of $\mathcal{B}$. A topological space which admits a *countable* basis of its topology is said to be **second countable**. If $(X, \mathcal{T})$ is second countable, from any basis $\mathcal{B}$ it is possible to extract a subbasis $\mathcal{B}' \subset \mathcal{B}$ which is countable. It is clear that second countability is a topological property.

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

1.1.6. 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 $\mathcal{A} \subset \mathcal{T}$, $\mathcal{T}_{\mathcal{A}} := \cap_{\mathcal{T} \subset C_{\mathcal{A}}} \mathcal{T}$ 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

$$(\cup_{i \in I'} B_i) \cap (\cup_{j \in I''} B_j) = \cup_{k \in K} B_k$$

for every choice of $I', I'' \subset I$ and a corresponding $K \subset I$.

1.1.7. 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 $f : X \to Y$ is continuous, X, Y being topological spaces, $f \restriction_A : A \to f(A)$ is continuous with respect to the induced topologies on A and $f(A)$ by X and Y respectively, for every subset $A \subset X$.

1.1.8. If $(X, \mathcal{T})$ is a topological space and $p \in X$, a **neighborhood** of p is an open set $U \subset X$ with $p \in U$. If X and Y are topological spaces and $x \in X$, $f : X \to Y$ 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 $f : X \to Y$ as above is continuous if and only if it is continuous in every point of X.

1.1.9. 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 $\mathcal{T}$.

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.

Finally, it turns out that if $f : X \to Y$ 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.

1.1.10. 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 $\gamma : [0, 1] \to X$ 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.

1.1.11. 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).

If $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ are topological spaces, the former is compact and $\phi : X \to Y$ 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.

1.1.12. 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.

1.1.13. A **semi metric space** is a set X endowed with a **semidistance**, that is $d : X \times X \to [0, +\infty)$, 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. It is very simple to show that a mapping $f : A \to M_2$, 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 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.

1.1.14. 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 $p : X \to [0, +\infty)$ 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.

1.1.1 The topology of $\mathbb{R}^n$.

A few words about the usual topology of $\mathbb{R}^n$ are in order. 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 $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ 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** (i.e. there is a ball $B_r(x) \subset \mathbb{R}^n$ with $r < \infty$ with $K \subset B_r(x)$).

Exercises 1.1.

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 $U \in \mathbb{R}^n$. 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_r 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 $f : M \to N$ 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 $F : X \to Y$ 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.)

1.1.2 Topological Manifolds.

Definition 1.1. (**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 $\phi : U_p \to V_p$

where $V_p \subset \mathbb{R}^n$ is an open set (equipped with the topology induced by $\mathbb{R}^n$).$\diamondsuit$

Remark 1.1.
(1) The homeomorphism ϕ 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 $\mathbb{R}^n$ were 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 $p_0 \not\in \mathbb{R}$. 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 $U \cap V = 0$ and $0 \in U$, $p_0 \in V$. Anyhow, 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 a 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 1.1.4) 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 $\mathbb{S}^2 \subset \mathbb{R}^3$. $\mathbb{S}^2 := \{(x, y, x) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$. $\mathbb{S}^2$ 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 Π_p be the plane tangent at $\mathbb{S}^2$ in p equipped with the topology induced by $\mathbb{R}^3$. With that topology Π_p is homeomorphic to $\mathbb{R}^2$ (the reader should prove it). Let ϕ be the orthogonal projection of $\mathbb{S}^2$ on Π_p. It is quite simply proved that ϕ is continuous with respect to the considered topologies and ϕ 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 Π_p (that is $\mathbb{R}^2$). The same procedure can be used to define local homeomorphisms referred to neighborhoods of each point of $\mathbb{S}^2$.

1.2 Differentiable Manifolds.

If $f : \mathbb{R}^n \to \mathbb{R}^n$ 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.

1.2.1 Local charts and atlas.

Definition 1.2. (**k-compatible local charts.**) Consider a topological manifold M with dimension n. A **local chart** or **local coordinate system** on M is pair (U, ϕ) where $U \subset M$ is open, $U \neq \emptyset$, and $\phi : p \mapsto (x^1(p), \ldots, x^n(p))$ is a homeomorphism from U to the open set $\phi(U) \subset \mathbb{R}^n$. Moreover:
(a) a local chart (U, ϕ) is called **global** chart if $U = M$;
(b) two local charts (U, ϕ), (V, ψ) are said to be C^k**-compatible**, $k \in (\mathbb{N} \setminus \{0\}) \cup \{\infty\}$, if either $U \cap V = \emptyset$ or, both $\phi \circ \psi^{-1} : \psi(U \cap V) \to \mathbb{R}^n$ and $\psi \circ \phi^{-1} : \phi(U \cap V) \to \mathbb{R}^n$ 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 1.3. (**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 :
(1) $\mathcal{A}$ covers M, i.e., $M = \cup_{i \in I} U_i$,
(2) the charts of $\mathcal{A}$ are pairwise C^k-compatible. ♢

Remark 1.2. 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 ∞ is an atlas of all orders.

1.2.2 Differentiable structures.

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

Definition 1.4. (**C^k-differentiable structure and differentiable manifold.**) 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 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, ϕ) is not C^k-compatible with some local chart of $\mathcal{M}$.
A topological manifold equipped with a differentiable structure of order $k \in (\mathbb{N} \setminus \{0\}) \cup \{\infty\}$ is said to be a **differentiable manifold** of order k. $\diamondsuit$

We leave to the reader the proof of the following proposition.

Proposition 1.1. *Referring to definition 1.4, if the local charts* (U, ϕ) *and* (V, ψ) *are separately* C^k *compatible with all the charts of a* C^k *atlas, then* (U, ϕ) *and* (V, ψ) *are* C^k *compatible.* $\diamondsuit$

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 are compatible with every chart of $\mathcal{A}$.

Comments 1.1. **(1)** $\mathbb{R}^n$ has a natural structure of C^∞-differentiable 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 $\vec{\cdot} : \mathbb{A}^n \times \mathbb{A}^n \to V$ 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^∞ 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 $f : \mathbb{A}^n \to \mathbb{R}^n$ 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^∞ atlas. Moreover, if $g : \mathbb{A}^n \to \mathbb{R}^n$ is another mapping defined as above with a different choice of O and the basis in V, $f \circ g^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ and $g \circ f^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ are C^∞ because they are linear non homogeneous transformations. Therefore, there is a C^∞ atlas containing all of the Cartesian coordinate systems defined by different choices of origin O and basis in V. The C^∞-differentiable structure generated by that atlas naturally makes the affine space a n-dimensional C^∞-differentiable manifold.

(3) The sphere $\mathbb{S}^2$ defined above gets a C^∞-differentiable structure as follows. Considering all of local homeomorphisms defined in remark (4) above, they turn out to be C^∞ compatible and define a C^∞ atlas on $\mathbb{S}^2$. That atlas generates a C^∞-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^∞-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^∞-differentiable structures. E.g., it happens for $n = 4$.

Important note. From now on "differential" and "differentiable" without further indication mean C^∞-differential and C^∞-differentiable respectively. Due to comment (4) above, we develop the theory in the C^∞ case only. However, several definitions and results may be generalized to the C^k case with $1 \leq k < \infty$

Exercises 1.2.
1.2.1. Show that the group $SO(3)$ is a three-dimensional differentiable manifold.

1.2.3 Differentiable functions and diffeomorphisms.

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

Definition 1.5. (**Differentiable functions and diffeomorphisms.**) Consider a continuous mapping $f : M \to N$, where M and N are differentiable manifolds with dimension m and n.
(1) f is said to be **differentiable** at $p \in M$ if the function:

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

is differentiable, for some local charts (V, ψ), (U, ϕ) on N and M respectively with $p \in U$, $f(p) \in V$ and $f(U) \subset V$.
(2) f is said to be **differentiable** if it is differentiable at every point of M.
The **real vector space** of all differentiable functions from M to N is indicated by $D(M|N)$ or $D(M)$ for $N = \mathbb{R}$.
If M and N are differentiable manifolds and $f \in D(M|N)$ is bijective and $f^{-1} \in D(N|M)$, f is called **diffeomorphism** from M to N. If there is a diffeomorphism from the differentiable manifold M to the differentiable manifold N, M and N are said to be **diffeomorphic**.$\diamondsuit$

Remark 1.3.
(1) It is clear that a differentiable function (at a point p) is continuous (in p) by definition.
(2) It is simply proved that the definition of function differentiable 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 addictive 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. 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 differentiable 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 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 C^k-differentiable manifolds ($k > 0$) 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 $f : M \to \mathbb{R}$ with $f : x \mapsto x$, whereas the differentiable structure of N is given by the assignment of the global chart $g : N \to \mathbb{R}$ with $g : x \mapsto x^3$. Notice that the differentiable structure of M differs from that of N because $f \circ g^{-1} : \mathbb{R} \to \mathbb{R}$ is not differentiable in $x = 0$. On the other hand M and N are diffeomorphic! Indeed a diffeomorphism is nothing but the map $\phi : M \to N$ 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 different 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 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 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 any 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, $f : U \to \mathbb{R}^n$ ($U \subset \mathbb{R}^n$) 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 a celebrated theorem shows that a 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.

1.3 Some Technical Lemmata. Differentiable 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 differentiable 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 1.1. *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 differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ 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* $y \not\in B_r(x)$. $\diamondsuit$

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 nonnegative, 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 nonnegative, it vanishes for $||x - y|| \geq r$ and takes the constant value 1 if $||x - y|| \leq r/2$ so that $G_r = B_{r/2}(x)$ $\Box$.

Lemma 1.2. *Let M be a differentiable manifold. For every $p \in M$ and every open neighborhood of p, U_p, there is a open neighborhoods of p, V_p and a mapping $h \in D(M)$ 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) $h(q) = 0$ *if* $x \not\in U_p$.
h is called **hat function** *centered on p with support contained in U_p.* $\diamondsuit$

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 $\overline{Z}^X \subset Y$ is closed in X and Y is open, it 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, $\overline{Z}^Y \supset \overline{Z}^X$ implies $\overline{Z}^Y = \overline{Z}^X$.)
Let us start with our construction. Consider a local chart (W, ϕ) with $p \in W$, and define $U_p = \phi^{-1}(B_r(x))$, $B_r(x)$ being an open ball with finite radius $r > 0$ centered on $x := \phi(p)$. Referring to lemma 1.1, define $V_p := \phi^{-1}(G_x)$ so that $\overline{V_p}^W = \phi^{-1}(\overline{G_x})$ because $\phi : W \to \phi(W)$ is a homeomorphism. For the same reason, as $\overline{G_x}$ is compact, $\overline{V_p}^W$ is compact as well and one has $\overline{V_p} := \overline{V_p}^M = \overline{V_p}^W = \phi^{-1}(\overline{G_x}) \subset \phi^{-1}(B_r(x)) = U_p$ and (1) holds true consequently. Finally define $h(q) := f(\phi(q))$ for $q \in W$ and $h(q) := 0$ if $q \notin W$. Notice that the support of f in $\phi(W)$ is compact and it coincides with that of h in W, that in turn coincides with that of h in M due to the initially mentioned result ($\overline{Z}^X = \overline{Z}^Y$), in particular $supp(h) \subset W$ is a compact set in M. By construction the function h automatically satisfies all requirements (2)-(4) barring the smoothness property that 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, ψ_q) such that $h \circ \psi_q^{-1}$ is $C^\infty(\psi_q(U_q))$. If $q \in W$ the local chart is (W, ϕ) itself. Then notice that $M \setminus \overline{W}$ is open and thus, for every $q \in M \setminus \overline{W}$, starting from local chart (U'_q, ψ'_q) with $q \in U'_q$ and sufficiently restricting the open neighborhood U'_q around q, we can construct a local chart $(U_q, \psi_q := \psi'_q \restriction_{U_q})$ with $q \in U_q \subset M \setminus \overline{W}$. Therefore $h \circ \psi_q^{-1}$ is $C^\infty(\psi_q(U_q))$ because $h \circ \psi_q^{-1} = 0$. It remains to consider the case $q \in \partial W := M \setminus (W \cup (M \setminus \overline{W}))$ only. It would be enough to prove that, if $q \in \partial W$, there is an open set $U_q \ni q$ with $U_q \cap supp(h) = \emptyset$. (In this way, restricting U_q if necessary, one has a local chart (U_q, ψ_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. If $q \in \partial W$ then $q \notin W$ and thus $q \notin supp(h) \subset W$. Since M is Hausdorff, for every $x \in supp(h)$ there are open neighborhoods O_x and $O_q^{(x)}$ of x and q respectively with $O_x \cap O_q^{(x)} = \emptyset$; Since $supp(h)$ is compact we can extract a finite covering of $supp(h)$ made of open sets O_{x_i}, $i = 1, 2, \ldots, N$ and consider the open set $U_q := \cap_{i=1,\ldots,N} O_q^{(x_i)}$. By construction $U_q \cap supp(h) \subset U_q \cap_{i=1,\ldots,N} O_q^{(x_i)} = \emptyset$ as wanted. $\Box$

Remark 1.4. Hausdorff property plays a central rôle 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 1.1, 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 $n = 1, 2, \ldots$. That sequence converges both to 0 and p_0 and trivially $\lim_{n\to+\infty} h(1/n) = h(0) = 1 \neq h(p_0) = 0$.

1.3.1 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 1.6. (**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.$\diamondsuit$

It is simply proved that a second-countable, Hausdorff, topological space X is paracompact if it is **locally compact**, i.e. every point $x \in X$ admits an open neighborhood U_p such that $\overline{U_p}$ is compact. As a consequence *every topological (or differentiable) manifold is paracompact* because it is Hausdorff, second countable and locally homeomorphic to $\mathbb{R}^n$ which, in turn, is locally compact.

Remark 1.5. It is possible to show (see Kobayashi and Nomizu: *Foundations of Differential Geometry.* Vol I, Interscience, New York, 1963) that, if X is a paracompact topological space which is also Hausdorff and locally homeomorphic to $\mathbb{R}^n$, X is second countable. *Therefore, a topological manifold can be equivalently defined as a paracompact topological space which is Hausdorff and locally homeomorphic to* $\mathbb{R}^n$.

1.3.2 Existence of a differentiable partition of unity.

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

Definition 1.7. (**Partition of Unity**.) Given a locally finite covering of a differentiable manifold M, $\mathcal{C} = \{U_i\}_{i\subset I}$, where every U_i is open, a **partition of unity** subordinate to $\mathcal{C}$ is a collection of functions $\{f_j\}_{j\in J} \subset D(M)$ such that:
(1) *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 $x \in M$.$\diamondsuit$

Remark 1.6.

(1) Notice that, for every $x \in M$, the sum in property (3) above is finite because of the locally 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 $f_i : M \to \mathbb{R}$ is *analytic* and $supp f_i \subset U_i$ 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 differentiable manifold made of open sets whose closures are compact. Using paracompactness one finds a subsequent locally finite covering which made of open sets whose closures are compact.

Theorem 1.1. (**Existence of a partition of unity**.) *Let M be a differentiable 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 $\mathcal{C}$.*

Proof. See Kobayashi and Nomizu: *Foundations of Differential Geometry.* Vol I, Interscience, New York, 1963. $\Box$

Chapter 2

Tensor Fields in Manifolds and Associated Geometric Structures.

2.1 Tangent and cotangent space in a point.

We introduce the *tangent space* by a direct construction. A **differentiable curve** or **differentiable path** $\gamma : \;\; (-\epsilon_\gamma, +\epsilon_\gamma) \;\; \to N$, $\epsilon_\gamma > 0$, where N is a differentiable manifold, is a mapping of $D(M|N)$, with $M = (-\epsilon_\gamma, +\epsilon_\gamma)$ equipped with the natural differentiable structure induced by $\mathbb{R}$. ϵ_γ depends on γ. If $p \in M$ is any point of a n-dimensional differentiable manifold, Q_p denotes the set of differentiable curves γ with $\gamma(0) = p$.
Then consider the relation on Q_p:

$$\gamma \sim \gamma' \quad \text{if and only if} \quad \frac{dx^i_\gamma}{dt}|_{t=0} = \frac{dx^i_{\gamma'}}{dt}|_{t=0} \,.$$

Above, we have singled out a local coordinate system $\phi : q \mapsto (x^1, \ldots x^n)$ defined in a neighborhood U of p, and $t \mapsto x^i_\gamma(t)$ denotes the i-th component of the mapping $\phi \circ \gamma$. Notice that the above relation is well defined, in the sense that it does not depend on the particular coordinate system about p used in the definition. Indeed if $\psi : q \mapsto (y^1, \ldots y^n)$ is another coordinate system defined in a neighborhood V of p, it holds

$$\frac{dx^i_\gamma}{dt}|_{t=0} = \frac{\partial x^i}{\partial y^j}|_{\psi\circ\gamma(0)} \frac{dy^j_\gamma}{dt}|_{t=0} \,.$$

The $n \times n$ matrices $J(q)$ and $J'(q)$ of coefficients, respectively,

$$\frac{\partial x^i}{\partial y^j}|_{\psi(q)} \,,$$

and

$$\frac{\partial y^k}{\partial x^l}|_{\phi(q)} \,,$$

defined in each point $q \in U \cap V$, are non-singular. This is because, deriving the identity:

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

one gets:

$$\frac{\partial x^i}{\partial y^j}|_{\psi(q)} \frac{\partial y^j}{\partial x^k}|_{\phi(q)} = \frac{\partial x^i}{\partial x^k}|_{\phi(q)} = \delta^i_k \ .$$

This is nothing but

$$J(q) J'(q) = I \ ,$$

and thus

$$det J(q) \ \ det J'(q) = 1 \ ,$$

which implies $det J'(q), det J(q) \neq 0$. Therefore the matrices $J(q)$ and $J'(q)$ are invertible and in particular: $J'(q) = J(q)^{-1}$. Using this result, one simply gets that the definition

$$\gamma \sim \gamma' \quad \text{if and only if} \quad \frac{dx^i_\gamma}{dt}|_{t=0} = \frac{dx^i_{\gamma'}}{dt}|_{t=0} \ ,$$

can equivalently be stated as

$$\gamma \sim \gamma' \quad \text{if and only if} \quad \frac{dy^j_\gamma}{dt}|_{t=0} = \frac{dy^j_{\gamma'}}{dt}|_{t=0} \ .$$

$\sim$ is well defined and is an *equivalence relation* as one can trivially prove. Thus the quotient space $T_pM := Q_p/ \sim$ is well defined too. If $\gamma \in Q_p$, the associated equivalence class $[\gamma] \in T_pM$ is called the **vector tangent to** γ **in** p.

Definition 2.1. (**Tangent space.**) If M is a differentiable manifold and $p \in M$, the set $T_pM := Q_p/ \sim$ defined as above is called the **tangent space** at M in p. $\diamondsuit$

As next step we want to define a *vector space* structure on T_pM. If $\gamma \in [\eta]$,$\gamma' \in [\eta']$ with $[\eta], [\eta'] \in T_pM$ and $\alpha, \beta \in \mathbb{R}$, define $\alpha[\eta] + \beta[\eta']$ as the equivalence class of the differentiable curves $\gamma'' \in Q_p$ such that, in a local coordinate system about p,

$$\frac{dx^i_{\gamma''}}{dt}|_{t=0} := \alpha \frac{dx^i_\gamma}{dt}|_{t=0} + \beta \frac{dx^i_{\gamma'}}{dt}|_{t=0} \ ,$$

where the used curves are defined for $t \in] - \epsilon, +\epsilon[$ with $\epsilon = Min(\epsilon_\gamma, \epsilon_{\gamma'})$. Such a definition does *not* depend on both the used local coordinate system and the choice of elements $\gamma \in [\eta]$,$\gamma' \in [\eta']$, γ'' we leave the trivial proof to the reader. The proof of the following lemma is straightforward and is left to the reader.

Lemma 2.1. *Using the definition of linear combination of elements of* T_pM *given above,* T_pM *turns out to be a vector space on the field* $\mathbb{R}$. *In particular the null vector is the class* $\mathbf{0}_p \in T_pM$*,* *where* $\gamma_{0p} \in \mathbf{0}_p$ *if and only if, in local coordinates about* p, $x^i_\gamma(t) = x^i(p) + tO^i(t)$ *where every* $O^i(t) \to 0$ *as* $t \to 0$. $\diamondsuit$

To go on, fix a chart (U, ψ) about $p \in M$, consider a vector $V \in \mathbb{R}^n$. Take the differentiable curve Γ_V contained in $\psi(U) \subset \mathbb{R}^n$ (n is the dimension of the manifold M) which starts form $\psi(p)$ with initial vector V, $\Gamma_V : t \mapsto tV + \psi(p)$ with $t \in]-\delta, \delta[$ with $\delta > 0$ small sufficiently. Define a mapping $\Psi_p : \mathbb{R}^n \to T_pM$ by $\Psi_p : V \mapsto [\psi^{-1}(\Gamma_V)]$ for all $V \in \mathbb{R}^n$.

We have a preliminary lemma.

Lemma 2.2. *Referring to the given definitions,* $\Psi_p : \mathbb{R}^n \to T_pM$ *is a vector space isomorphism. As a consequence,* $dimT_pM = dim\mathbb{R}^n = n$. $\diamondsuit$

Proof. $\Psi_p : \mathbb{R}^n \to T_pM$ is injective since if $V \neq V'$, $\psi^{-1}(\Gamma_V) \not\sim \psi^{-1}(\Gamma_{V'})$ by construction. Moreover Ψ_p is surjective because if $[\gamma] \in T_pM$, $\psi^{-1}(\Gamma_V) \sim \gamma$ when $V = \frac{d}{dt}|_{t=0}\psi(\gamma(t))$. Finally it is a trivial task to show that Ψ_p is linear if T_pM is endowed with the vector space structure defined above. Indeed $\alpha\Psi_p(V) + \beta\Psi_p(W)$ is the class of equivalence that contains the curves η with (in the considered coordinates)

$$\frac{dx^i_\eta}{dt}|_{t=0} = \alpha V^i + \beta W^i \,.$$

One curve that satisfies this constraint is

$$(-\epsilon, \epsilon) \ni t \mapsto t(\alpha V + \beta W) + \psi(p) \,.$$

Thus, in particular

$$\alpha\Psi_p(V) + \beta\Psi_p(W) = [(-\epsilon, \epsilon) \ni t \mapsto \psi^{-1}(t(\alpha V + \beta W) + \psi(p))]$$

for some $\epsilon > 0$. But also, by definition of Ψ_p, it holds

$$[(-\epsilon, \epsilon) \ni t \mapsto t(\alpha V + \beta W) + \psi(p)] = \Psi_p(\alpha V + \beta W) \,.$$

As a consequence

$$\Psi_p(\alpha V + \beta W) = \alpha\Psi_p(V) + \beta\Psi_p(W)$$

and this concludes the proof. $\Box$

Definition 2.2. (**Basis induced by a chart.**) Let M be a differentiable manifold, $p \in M$, and take a chart (U, ψ) with $p \in U$. If $E_1, \ldots, E_n$ is the canonical basis of $\mathbb{R}^n$, $e_{pi} = \Psi_p E_i$, i=1,...,n, define a basis in T_pM which we call the **basis induced in T_pM by the chart** (U, ψ). $\diamondsuit$

Proposition 2.1. *Let M be a n-dimensional differentiable manifold. Take $p \in M$ and two local charts (U, ψ), (U', ψ') with $p \in U, U'$ and induced basis on T_pM, $\{e_{pi}\}_{i=1,\ldots,n}$ and $\{e'_{pj}\}_{j=1,\ldots,n}$ respectively. If $t_p = t^i e_{pi} = t'^j e'_{pj} \in T_pM$ then*

$$t'^j = \frac{\partial x'^j}{\partial x^k}|_{\psi(p)} t^k \ ,$$

or equivalently

$$e_{pk} = \frac{\partial x'^j}{\partial x^k}|_{\psi(p)} e'_{pj} \ ,$$

where $x'^j = (\psi' \circ \psi^{-1})^j(x^1, \ldots, x^n)$ in a neighborhood of $\psi(p)$. ♢

Proof. We want to show the thesis in the latter form, i.e.,

$$e_{pk} = \frac{\partial x'^j}{\partial x^k}|_{\psi(p)} e'_{pj} \ .$$

Each vector E_j of the canonical basis of the space $\mathbb{R}^n$ associated with the chart (U, ψ) can be viewed as the tangent vector of the differentiable curve $\Gamma_k : t \mapsto tE_k + \psi(p)$ in $\mathbb{R}^n$. Such a differentiable curve in $\mathbb{R}^n$ defines a differentiable curve in M, $\gamma_k : t \mapsto \psi^{-1}(\Gamma_k(t))$ which starts from p. In turn, in the set $\psi'(U') \subset \mathbb{R}^n$ this determines a curve $\Lambda_k : t \mapsto \psi'(\gamma_k(t))$. In coordinates, such a differentiable curve is given by

$$x'^j(t) = x'^j(x^1(t), \ldots, x^n(t)) = x'^j(x^1_p, \ldots, t + x^k_p, \ldots, x^n_p) \ ,$$

where x^k_p are the coordinates of p with respect to the chart (U, ψ). Taking the derivative at $t = 0$ we get the components of the representation of E_k with respect to the canonical basis $E'_1, \cdots, E'_n$ of $\mathbb{R}^n$ associated with the chart (U', ψ'). In other words, making use of the isomorphism Ψ_p defined above and the analogue Ψ'_p for the other chart (U', ψ'),

$$(({\Psi'_p}^{-1} \circ \Psi_p)E_k)^j = \frac{\partial x'^j}{\partial x^k}|_{\psi(p)} \ ,$$

or

$$({\Psi'_p}^{-1} \circ \Psi_p)E_k = \frac{\partial x'^j}{\partial x^k}|_{\psi(p)} E'_j \, .$$

As Ψ'_p is an isomorphism, that is equivalent to

$$\Psi_p E_k = \frac{\partial x'^j}{\partial x^k}|_{\psi(p)} \Psi'_p E'_j \ ,$$

but $e_{pr} = \Psi_p E_r$ and $e'_{pi} = \Psi_p E'_i$ and thus we have proved that

$$e_{pk} = \frac{\partial x'^j}{\partial x^k}|_{\psi(p)} e'_{pj} \ ,$$

which is the thesis.□

2.1.1 Vectors as derivations.

We want to show that there is a natural isomorphism between T_pM and $\hat{\mathcal{D}}_pM$, the latter being the space of the *derivations* generated by operators $\frac{\partial}{\partial x^k}|_p$. We need two preliminary definitions.

Definition 2.3. (**Derivations**) Let M be a differentiable manifold. A **derivation** in $p \in M$ is a $\mathbb{R}$-linear map $D_p : D(M) \to \mathbb{R}$, 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 in p is indicated by $\mathcal{D}_pM$.$\diamondsuit$

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

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

are derivations. Notice also that, changing coordinates about p and passing to (V, ψ) with coordinates $(y^1, \ldots, y^n)$ one gets:

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

Since the matrix J of coefficients $\frac{\partial x^r}{\partial y^k}|_{\psi(p)}$ is not singular as we shown previously, the vector space spanned by derivations $\frac{\partial}{\partial y^k}|_p$, for $k = 1, \ldots, n$, coincides with that spanned by derivations $\frac{\partial}{\partial x^k}|_p$ for $k = 1, \ldots, n$. In the following we shall indicate such a common subspace of $\mathcal{D}_p(M)$ by $\hat{\mathcal{D}}_pM$.
To go on, let us state and prove an important locality property of derivations.

Lemma 2.3. *Let M be a differential manifold. Take any $p \in M$ and any $D_p \in \mathcal{D}_pM$.*
(1) *If $h \in D(M)$ vanishes in a open neighborhood of p,*

$$D_p h = 0 \,.$$

(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,*

$$D_p h = 0 \,.$$

$\diamondsuit$

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

$$D_p(hg) = h D_p g\,.$$

On the other hand, since D_p is a derivation,

$$D_p(hg) = h D_p g + g(p) D_p h\,.$$

Comparing with the identity above, one gets

$$g(p) D_p h = 0\,.$$

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

As a final proposition we precise the interplay between $\mathcal{D}_p M$ and $T_p M$ proving that actually they are the same $\mathbb{R}$-vector space via a natural isomorphism.
A technical lemma is necessary. We remind the reader that a 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.

Lemma 2.4. (**Flander's lemma**.) *If $f : B \to \mathbb{R}$ is $C^\infty(B)$ where $B \subset \mathbb{R}^n$ is an open starshaped neighborhood of $p_0 = (x_0^1, \ldots, x_0^n)$, there are n differentiable mappings $g_i : B \to \mathbb{R}$ such that, if $p = (x^1, \ldots, x^n)$,*

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

with

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

for all $i = 1, \ldots, n$. ♢

Proof. let $p = (x^1, \ldots, x^n)$ belong to B. The points of $\overline{p_0 p}$ 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

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

If

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

so that

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

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 dominate convergence theorem). $\Box$

Proposition 2.2. *Let M be a differentiable manifold and $p \in M$. There is a natural $\mathbb{R}$-vector space isomorphism $F : T_pM \to \mathcal{D}_pM$ such that, if $\{e_{pi}\}_{i=1,\ldots,n}$ is the basis of T_pM induced by any local chart about p with coordinates $(x^1, \ldots, x^n)$, it holds:*

$$F : t^k e_{pk} \mapsto t^k \frac{\partial}{\partial x^k}|_p \, ,$$

for all $t_p = t^k e_{pk} \in T_pM$. In particular the set $\{\frac{\partial}{\partial x^k}|_p\}_{k=1,\ldots,n}$ is a basis of $\mathcal{D}_pM$.
As a consequence:

(a) *derivations in p are one-to-one represented by a vectors in T_pM and this representation does not depend on the choice of a particular local chart about p;*

(b) *every derivation in p can uniquely be decomposed as a linear combination of derivations $\{\frac{\partial}{\partial x^k}|_p\}_{k=1,\ldots,n}$ for each local chart about p referred to coordinates $(x^1, \ldots, x^n)$.* $\diamondsuit$

Proof. The mapping

$$F : t^k e_{pk} \mapsto t^k \frac{\partial}{\partial x^k}|_p$$

is a linear mapping from a n-dimensional vector space to the vector space generated by the derivations $\{\frac{\partial}{\partial x^k}|_p\}_{k=1,\ldots,n}$. As said above this latter space is denoted by $\hat{\mathcal{D}}_pM$. F is trivially surjective, then it defines a isomorphism if $\{\frac{\partial}{\partial x^k}|_p\}_{k=1,\ldots,n}$ is a basis of $\hat{\mathcal{D}}_pM$ or, it is the same, if the vectors $\hat{\mathfrak{D}}_p M$ are linearly independent. Let us prove that these vectors are, in fact, linearly independent. If (U, ϕ) is the considered local chart, with coordinates $(x^1, \ldots, x^n)$, it is sufficient to use n functions $f^{(j)} \in D(M)$, $j = 1, \ldots, n$ 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 $c^k \delta^j_k = 0$ or :

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

The existence of the functions $f^{(j)}$ can be straightforwardly proved by using lemma 1.2. The mapping $f^{(j)} : M \to \mathbb{R}$ defined as:
$f^{(j)}(q) = h(q)\phi^j(q)$ if $q \in U$, where $\phi^j : q \mapsto x^j(q)$ for all $q \in U$,
$f^{(j)}(q) = 0$ if $q \in M \setminus U$,
turns out to be C^∞ 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.
The isomorphism F does not depend on the used basis and thus it is natural. Indeed,

$$F : t^k e_{pk} \mapsto t^k \frac{\partial}{\partial x^k}|_p$$

can be re-written as:

$$F : (\frac{\partial x^k}{\partial x'^i} t'^i)(\frac{\partial x'^r}{\partial x^k} e'_{pr}) \mapsto (\frac{\partial x^k}{\partial x'^i} t'^i)(\frac{\partial x'^r}{\partial x^k} \frac{\partial}{\partial x'^r}|_p) \,.$$

Since

$$\frac{\partial x^k}{\partial x'^i} \frac{\partial x'^r}{\partial x^k} = \delta^r_i \,,$$

the identity above is noting but:

$$F : t'^i e'_{pi} \mapsto t'^i \frac{\partial}{\partial x'^i}|_p \,.$$

To conclude the proof it is sufficient to show that $\hat{\mathfrak{D}}_p M = \mathfrak{D}_p M$. In other words it is sufficient to show that, if $D_p \in \mathfrak{D}_p M$ and considering the local chart about p, (U, ϕ) with coordinates $(x^1, \ldots, x^n)$, there are n reals $c^1, \ldots, c^n$ such that

$$D_p f = \sum_{k=1}^{n} c^k \frac{\partial f \circ \phi^{-1}}{\partial x^k}|_p$$

for all $f \in D(M)$. To prove this fact we start from the expansion due to lemma 2.3 and valid in a neighborhood $U_p \subset U$ of $\phi(p)$:

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

where $\phi(q) = (x^1, \ldots, x^n)$ and $\phi(p) = (x^1_p, \ldots, x^n_p)$ and

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

If h_1, h_2 are hat functions centered on p (see lemma 1.2) 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

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

while

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

Moreover, by lemma 2.3, $D_p f' = D_p f$ 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 2.3. Moreover

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

where $r_i(p) = (x^i_p - x^i_p) = 0$. Finally we have found

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

where the coefficients

$$c^i = D_p r_i$$

do not depend on f by construction. This is the thesis and the proof ends. □

Remark 2.1. With the given definition, it arises that any n-dimensional **Affine space** $\mathbb{A}^n$ admits two different notions of vector. 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 differentiable manifold as said in the comment (2) after proposition 1.1, one can define vectors in every point p of $\mathbb{A}^n$, namely the vectors of T_pM. What is the relation between these two notions of vector? 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:

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

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 $\chi_p : T_p\mathbb{A}^n \to V$ which identifies each $\frac{\partial}{\partial x^i}|_p$ with the corresponding e_i[1].

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

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

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

where

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

Thus, it is immediately proved by direct inspection that, if χ'_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 $A^i{}_j B_i{}^k = \delta^k_j$ and thus

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

---

[1] This is equivalent to say the initial tangent vector at a differentiable curve $\gamma :]\epsilon, \epsilon[\to \mathbb{A}^n$ which start from p can be computed both as an element of V: $\dot\gamma|_p = \lim_{h\to 0} \frac{\overrightarrow{\gamma(0)\gamma(h)}}{h}$ or an element of $T_p\mathbb{A}^n$ using the general procedure for differentiable manifolds. The natural isomorphism is nothing but the identification of these two notions of tangent vector.

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 χ does not depend on the considered Cartesian coordinate frame, that is it is natural.

2.1.2 Cotangent space.

As T_pM is a vector space, one can define its dual space. This space plays an important rôle in differential geometry.

Definition 2.4. (**Cotangent space.**) Let M be a n-dimensional manifold. For each $p \in M$, the dual space T_p^*M is called the **cotangent space** on p and its elements are called **1-forms** in p or, equivalently, **covectors** in p. If $(x^1, \ldots, x^n)$ are coordinates about p inducing the basis $\{\frac{\partial}{\partial x^k}|_p\}_{k=1,\ldots,n}$, the associated dual basis in T_p^*M is denoted by $\{dx^k|_p\}_{k=1,\ldots,n}$. $\diamondsuit$

Exercises 2.1.
2.1.1. Let $\gamma : (-\epsilon, +\epsilon) \to M$ be a differentiable curve with $\gamma(0) = p$. Show that the tangent vector at γ in p is:

$$\dot{\gamma}|_p := \frac{dx^i_\gamma}{dt}|_{t=0} \frac{\partial}{\partial x^i}|_p \ ,$$

where $(x^1, \ldots, x^n)$ are local coordinates defined in the neighborhood of p, U, where γ is represented by $t \mapsto x^i_\gamma(t)$, $i = 1, \ldots, n$.
2.1.2. 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

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

2.2 Tensor fields.

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

Definition 2.5. (**Differentiable Tensor Fields.**) Let M be a n-dimensional manifold. A **differentiable tensor field** t is an assignment $p \mapsto t_p$ where the tensors $t_p \in \mathcal{A}_{\mathbb{R}}(T_pM)$ are of the same kind and have differentiable components with respect to all of the canonical bases of

$\mathcal{A}_{\mathbb{R}}(T_pM)$ given by tensor products of bases $\{\frac{\partial}{\partial x^k}|_p\}_{k=1,\ldots,n} \subset T_pM$ and $\{dx^k|_p\}_{k=1,\ldots,n} \subset T_p^*M$ induced by all of local coordinate systems on M.

In particular a differentiable **vector field** and a differentiable 1-**form** (equivalently called **covector field**) are assignments of tangent vectors and 1-forms respectively as stated above.

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 $(1,1)$. $\diamondsuit$

Remark 2.2.

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

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

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 $(x^1,\cdots,x^n)$ and $(y^1,\cdots,y^n)$ are the coordinates of the same point $p \in M$ in two different local charts,

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

Then, in local coordinates,

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

(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) If X is a differentiable vector field on a differentiable manifold, X defines a derivation at each point $p \in M$: if $f \in D(M)$,

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

where $x^1,\ldots,x^n$ are coordinates defined about p. More generally, every differentiable vector field X 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 differentiable 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 differentiable for all of $f \in D(M)$. Indeed, if X is a differentiable contravariant vector field and if $f \in D(M)$, one has that $X(f) : M \ni p \mapsto X_p(f)$ is a differentiable function too as it having a coordinate representation

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

in every local coordinate chart (U, ϕ) and all the involved function being differentiable. 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, ϕ) must be differentiable. 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 ϕ (the coordinate x^i) and h a hat function centered on q with support in U.
Similarly, the differentiability of a covariant vector field ω is equivalent to the differentiability of each function $p \mapsto \langle X_p, \omega_p \rangle$, for all differentiable vector fields X.
(5) If $f \in D(M)$, the **differential** of f **in** p, df_p is the 1-form defined by

$$df_p = \frac{\partial f}{\partial x^i}|_p dx^i|_p \ ,$$

in local coordinates about 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 differentiable vector field denoted by df and called the **differential** of f.
(6) The set of contravariant differentiable vector fields on any differentiable manifold M defines a 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 addictive 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 addictive unit elements" is well known and it is called **module**.

The following lemma is trivial but useful in applications.

Lemma 2.5. *Let p be a point in a differentiable manifold M. If t is any tensor in $\mathcal{A}_{\mathbb{R}}(T_pM)$, there is a differentiabile tensor field in M, Ξ such that $\Xi_p = t$.* $\Diamond$

Proof. Consider a local coordinate frame (U, ϕ) defined in an open neighborhood U of p. In U a tensor field Ξ' 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 definition2.4) a differentiable function $h : \phi(U) \to \mathbb{R}$ such that $h(\phi(p)) = 1$ and h vanishes outside a small neighborhood of $\phi(p)$ whose closure is completely contained in $\phi(U)$. Ξ defined as $(h \circ \phi)(r) \cdot \Xi'(r)$ if $r \in U$ and $\Xi(r) = 0$ outside U is a differentiable tensor fields on M such that $\Xi_p = t$.$\Box$

2.2.1 Lie brackets.

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

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

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 2.2, 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_pM (denoted by $[X,Y]_p$). On the other hand, varying the point p one gets $[X,Y]_p(f)$ is a differentiable function if $f \in D(M)$. This is because (see remark 4 above), as X and Y are differentiable 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 (differentiable) vector field on M.

Definition 2.6. (**Lie Bracket.**) Let X, Y be a pair of contravariant differentiable vector fields on the differentiable manifold M. The **Lie bracket** of X and Y, $[X,Y]$, is the contravariant differentiable vector field associated with the differential operator

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

for $f \in D(M)$.$\diamondsuit$

Exercises 2.2.
2.2.1. Show that in local coordinates

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

2.2.2. Prove that the Lie brackets define a **Lie algebra** in the real vector space of the contravariant differentiable vector fields on any differentiable manifold M. In other words $[\,,\,]$ enjoys the following properties, where X, Y, Z are contravariant differentiable vector fields,
antisymmetry, $[X,Y] = -[Y,X]$;
$\mathbb{R}$**-linearity**, $[\alpha X + \beta Y, Z] = \alpha[X,Z] + \beta[Y,Z]$ for all $\alpha, \beta \in \mathbb{R}$;

Jacobi identity, $[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]=0$ (0 being the null vector field);

Important note. From now *tensor (vector, covector) field* means *differentiable* tensor (vector, covector) field.

2.3 Tangent and cotangent space manifolds.

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

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

It is possible to endow TM with a structure of a differentiable 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 differentiable 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 $\phi : U \to \mathbb{R}^n$. It is straightforwardly proved that $\mathcal{B}$ is a basis of the topology of M. Then consider the class $T\mathcal{B}$ of subsets of TM, V, defined as follows. Take $(U,\phi) \in \mathcal{M}$ with $\phi : p \mapsto (x^1(p),\ldots,x^n(p))$, and an open nonempty set $B \subset \mathbb{R}^n$ and define

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

where $\hat{\phi}_p : \mathbb{R}^n \to T_pM$ is the linear isomorphism associated with ϕ: $\hat{\phi}_p : (v^1_p,\ldots,v^n_p) \mapsto v^i_p \frac{\partial}{\partial x^i}|_p$.
Let $\mathcal{T}_{T\mathcal{B}}$ denote the topology generated on TM by the class $T\mathcal{B}$ of all the sets V obtained by varying U and B as said above. $T\mathcal{B}$ itself is a basis of that topology. Moreover $\mathcal{T}_{T\mathcal{B}}$ is second-countable and Hausdorff by construction. Finally, it turns out that TM, equipped with the topology $\mathcal{T}_{T\mathcal{B}}$, is locally homeomorphic to $M \times \mathbb{R}^n$, that is it is locally homeomorphic to $\mathbb{R}^{2n}$. Indeed, if (U,ϕ) is a local chart of M with $\phi : p \mapsto (x^1(p),\ldots,x^n(p))$, we may define a local chart of TM, (TU,Φ), where

$$TU := \{(p,v) \mid p \in U \;\;,\;\; v \in T_pM\}$$

by defining

$$\Phi : (p,v) \mapsto (x^1(p),\ldots,x^n(p),v^1_p,\ldots,v^n_p)\;,$$

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

$$\bigcup TU = TM\;,$$

TM is locally homeomorphic to $\mathbb{R}^{2n}$.
The next step consists of defining a differentiable structure on TM. Consider two local charts on TM, (TU,Φ) and (TU',Φ') respectively induced by two local charts in M, (U,ϕ) and (U',ϕ'). As a consequence of the given definitions (TU,Φ) and (TU',Φ') are trivially compatible. Moreover,

the class of charts (TU, Φ) induced from all the charts (U, ϕ) 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 differentiable manifold with dimension $2n$.

An analogous procedure gives rise to a natural differentiable structure for

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

Definition 2.7. (**Tangent and Cotangent Space Manifolds**.) Let M be a differentiable manifold with dimension n and differentiable structure $\mathcal{M}$. If (U, ϕ) is any local chart of $\mathcal{M}$ with $\phi : p \mapsto (x^1(p), \ldots, x^n(p))$ define

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

and

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

where $B \subset \mathbb{R}^n$ is any open nonempty set and $\hat{\phi}_p : \mathbb{R}^n \to T_pM$ and ${}^*\hat{\phi}_p : \mathbb{R}^n \to T_p^*M$ are the linear isomorphisms naturally induced by ϕ. Finally define $\Phi : TU \to \phi(U) \times \mathbb{R}^n \subset \mathbb{R}^{2n}$ and ${}^*\Phi : T^*U \to \phi(U) \times \mathbb{R}^n \subset \mathbb{R}^{2n}$ such that

$$\Phi : (p, v) \mapsto (x^1(p), \ldots, x^n(p), v_p^1, \ldots, v_p^n) ,$$

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

$${}^*\Phi : (p, v) \mapsto (x^1(p), \ldots, x^n(p), \omega_{1p}, \ldots, \omega_{p_n}) ,$$

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

The **tangent space (manifold)** associated with M is the 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 nonempty sets of $\mathbb{R}^n$,

(2) the differentiable structure induced by the atlas

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

The **cotangent space (manifold)** 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

nonempty sets of $\mathbb{R}^n$,
(2) the differentiable structure induced by the atlas

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

♢

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.

Remark 2.3. 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 (C^∞) differentiable structure of M. Let the coordinates of the associated local chart on TM, (TU, Φ) be indicated by x^1, x^2, v^1, v^2 with $x^i \in \mathbb{R}$ associated with ϕ 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:

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

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 2.7 and concerning canonical projections, sections and lift of differentiable curves.

Definition 2.8. (**Canonical projections, sections, lifts**.) Let M be a differentiable manifold. The surjective differentiable mappings

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

and

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

are called **canonical projections onto** TM and T^*M respectively.
A **section** of TM (respectively T^*M) is a differentiable map $\sigma : M \to TM$ (respectively T^*M), such that $\Pi(\sigma(p)) = p$ (respectively $^*\Pi(\sigma(p)) = p$) for every $p \in M$.
If $\gamma : t \mapsto \gamma(t) \in M$, $t \in I$ interval of $\mathbb{R}$, is a differentiable curve, the **lift of** γ, Γ, is the differentiable curve in TM,

$$\Gamma : t \mapsto (\gamma(t), \dot\gamma(t)) \, .$$

♢

Chapter 3

Differential mapping and Submanifolds.

3.1 Push forward.

A useful tool in differential geometry is the differential of a differentiable function also called **push forward**.

Definition 3.1. (**Differential or push forward of a mapping**.) If $f : N \to M$ is a differentiable function from the differentiable manifold N to the differentiable manifold M, for every $p \in N$, the **differential of** f at p or **push forward of** f at p

$$df_p : T_pN \to T_{f(p)}M \ ,$$

is the linear mapping defined by

$$(df_pX_p)(g) := X_p(g \circ f)$$

for all differentiable vector fields X on N and differentiable functions $g \in D(M)$.$\diamondsuit$

Remark 3.1.
(1) Take two local charts (U, ϕ) in N and (V, ψ) in M about p and $f(p)$ respectively and use the notation $\phi : U \ni q \mapsto (x^1(q), \ldots, x^n(q))$ and $\psi : V \ni r \mapsto (y^1(r), \ldots, y^m(r))$. Then define $\tilde{f} := \psi \circ f \circ \phi^{-1} : \phi(U) \to \mathbb{R}^m$ and $\tilde{g} := g \circ \psi^{-1} : \psi(V) \to \mathbb{R}$. $\tilde{f}$ and $\tilde{g}$ "represent" f and g, respectively, in the fixed coordinate systems. By construction, it holds

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

That is, with obvious notation

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

In other words

$$((df_p X)g)^k = \left(\frac{\partial \tilde{f}^k}{\partial x^i} X^i\right) \frac{\partial \tilde{g}}{\partial y^k}|_f \, .$$

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

$$df_p : \; X^i(p)\frac{\partial}{\partial x^i}|_p \;\mapsto\; X^i(p)\frac{\partial(\psi \circ f \circ \phi^{-1})^k}{\partial x^i}|_{\phi(p)} \; \frac{\partial}{\partial y^k}|_{f(p)} \, .$$

That formula is more often written

$$df_p : \; X^i(p)\frac{\partial}{\partial x^i}|_p \;\mapsto\; X^i(p)\frac{\partial y^k}{\partial x^i}|_{(x^1(p),\ldots x^n(p))} \; \frac{\partial}{\partial y^k}|_{f(p)} \, ,$$

where it is understood that $\psi \circ f \circ \phi^{-1} : (x^1, \ldots, x^n) \mapsto (y^1(x^1, \ldots, x^n), \ldots, y^m(x^1, \ldots, x^n))$.
(2) With the meaning as in the definition above, often df is indicated by f_* and $g \circ f$ is denoted by f^*g.

3.2 Rank of a differentiable map: immersions and submersions.

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.
Notice that, if (U, ϕ) and (V, ψ) are local charts about p and $f(p)$ respectively, the rank of the Jacobian matrix of the function $\psi \circ f \circ \phi^{-1} : \phi(U) \to \mathbb{R}^m$ 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 3.2. If $f : N \to M$ is a differentiable function from the differentiable manifold N to the differentiable manifold N 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, ϕ) and (V, ψ) being a pair of local charts about 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 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.$\diamondsuit$

It is clear that if N is a differentiable manifold and $U \subset N$ is an open set, U is Hausdorff second countable and locally homeomorphic to $\mathbb{R}^n$. Thus we can endow U with a differentiable structure naturally induced by that of N itself, by restriction to U of the domains of the local charts on N. We have the following remarkable results.

Theorem 3.1. *Let $f : N \to M$ be a differentiable function with M and N differentiable manifolds with dimension m and n respectively and take $p \in N$.*
(1) *If $n \geq m$ and the rank of f at p is m, i.e. df_p is surjective, for any local chart (V, ψ) about $f(p)$ there is a local chart (U, ϕ) about p such that*

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

(2) *If $n \leq m$ and the rank of f at p is n, i.e. df_p is injective, for any local chart (U, ϕ) about p there is a local chart (V, ψ) about $f(p)$ such that*

$$\psi \circ f \circ \phi^{-1}(x^1, \ldots, x^n) = (x^1, \ldots, x^m, 0, \ldots, 0) \, ;$$

(3) *If $n = m$, the following statements are equivalent*
(a) *$df_p : T_pN \to T_{f(p)}N$ is a linear isomorphism,*
(b) *f defines a* **local diffeomorphism about** *p, i.e. there is an open neighborhood U of p and an open neighborhood V of $f(p)$ such that $f \restriction_U : \to V$ defined on the differentiable manifold U equipped with the natural differentiable structure induced by N and evaluated on the differentiable manifold V equipped with the natural differentiable structure induced by M. $\diamondsuit$*

Sketch of the proof. Working in local coordinates in N and M and passing to work with the Jacobian matrices of the involved functions (a) and (b) are direct consequences of Dini's implicit function theorem. Let us pass to consider (c). Suppose that $g := f \restriction_U$ is a diffeomorphism onto V. In that case $g^{-1} : V \to U$ is a diffeomorphism to and $g \circ f = id_U$. Working in local coordinates about 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 $J[f]_p \neq 0$ and, via the remark (1) above, this 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

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

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

3.3 Submanifolds.

Let us consider the definitions involved with the notion of submanifold.

Definition 3.3. If $f : N \to M$ is a differentiable function from the differentiable manifold N to the differentiable 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 is called **embedding** if
(i) it is injective and
(ii) $f : N \to f(N)$ is a homeomorphism when $f(N)$ is equipped with the topology induced by M.$\diamondsuit$

Definition 3.4. Let M, N be two differentiable manifolds and suppose that, regardless the respective differentiable structures, $N \subset M$. N is said to be a **differentiable submanifold** of M if the inclusion map $i : N \hookrightarrow M$ is differentiable and is an embedding. In that case, the differentiable structure of N is said to be **induced** by M.$\diamondsuit$

An equivalent definition can be given by using the statement in (b) of the following proposition.

Proposition 3.1. *Let M a differentiable manifolds with dimension m and let $N \subset M$.*
(a) *If N is a differentiable manifold with dimension n which is 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, ϕ), such that if we use the notation, $\phi : q \mapsto (x^1(q), \ldots, x^m(q))$, it holds

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

(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. That differentiable structure is obtained as follows. The maps $N \cap U_p \ni q \mapsto (x^1(q), \ldots, x^n(q))$ defines a local charts about 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 $\mathcal{N}$.* $\diamondsuit$

Sketch of proof. (a) if N is a submanifold of M, 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 3.1 (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 ϕ 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 about 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. $\Box$

Examples 3.1.

1. The map $\gamma : \mathbb{R} \ni t \mapsto (\sin t, \cos t) \subset \mathbb{R}^2$ is an immersion, since $d\gamma \neq 0$ (which is equivalent to say that $\dot{\gamma} \neq 0$) everywhere. Anyway that is not an embedding since γ is not injective.

2. However the set $C := \gamma(\mathbb{R})$ is a submanifold of $\mathbb{R}^2$ 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 3.1. 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 3.1, there is a local chart (U, ψ) of $\mathbb{R}^2$ about 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 (θ, 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 about 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 3.1 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 **2.2.3** with a differentiable structure and make it a one-dimensional differentiable 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 differentiable manifold with dimension 1 which is not diffeomorphic to $\mathbb{R}$ and cannot be considered a submanifold of $\mathbb{R}^2$.

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 a submanifold of $\mathbb{R}^2$. Actually, differently from above, here the problem concerns the smoothness of the inclusion map at $(0, 0)$ rather that 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 $\mathbb{R}^2$ about the point $(0, 0)$ such that the requirements of proposition 3.1 are fulfilled sue to the cusp in that point of the curve C. However, it is simply defined a differentiable structure on C which make it a one-dimensional differentiable manifold. It is sufficient to consider the differentiable structure generated by the global chart given by the inverse of the homeomorphism $f : \mathbb{R} \ni t \mapsto (|t|, t)$.

6. Let us consider once again the cylinder $C \subset \mathbb{E}^3$ defined in the example 4.1.2. C is a submanifold of $\mathbb{E}^3$ in the sense of the definition 3.3 since the construction of the differential structure made in the example 4.1.2 is that of proposition 3.1 starting from cylindrical coordinates $\theta, r' := r-1, z$.

3.3.1 Theorem of regular values.

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

Theorem 3.2. (**Theorem of regular values**.) *Let $f : N \to M$ be a differentiable function from the differentiable manifold N to the differentiable 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 submanifold of N.*$\diamondsuit$

Remark 3.2. A know theorem due to Sard, show that the *measure* of the set of singular values of any differentiable function $f : N \to M$ must vanish. This means that, if $S \subset M$ is the set of singular values of f, for every local chart (U, ϕ) 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 3.2. .

1. In *analytical mechanics*, consider a system of N material points with possible positions $P_k \in \mathbb{R}^3$, $k = 1, 2, \ldots, N$ and c constraints given by assuming $f_i(P_1, \ldots, P_N) = 0$ where the c functions $f_i : \mathbb{R}^{3N} \to \mathbb{R}$, $i = 1, \ldots, c$ are differentiable. If the constraints are *functionally independent*, i.e. the Jacobian matrix of elements $\frac{\partial f_i}{\partial x_k}$ has rank c everywhere, $x^1, x^2, \ldots, x^{3N}$ being the coordinates of $(P_1, \ldots, P_N) \in (\mathbb{R}^3)^N$, the configuration space is a submanifold of $\mathbb{R}^{3N}$ with dimension $3N - c$. This result is nothing but a trivial application of theorem 3.2.

2. Consider the same example 3.1.2 from another point of view. As a set the circumference $C = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1\}$ is $f^{-1}(0)$ with $f : \mathbb{R}^2 \to \mathbb{R}$ defined as $f(x, y) := x^2 + y^2 - 1$. The value 0 is a regular value of f because $df_p = 2xdx + 2ydy \neq 0$ if $f(x, y) = 0$ that is $(x, y) \in \mathbb{C}$. As a consequence of theorem 3.2 C can be equipped with the structure of submanifold of $\mathbb{R}^2$. This structure is that defined in the example 3.1.2.

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 differentiable. 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 $\mathbb{R}^3$.

Chapter 4

Riemannian and pseudo Riemannian manifolds.

Definition 4.1. (**(Pseudo) Riemannian Manifolds.**) A connected differentiable manifold M equipped with a symmetric $(0,2)$ differentiable tensor Φ field which defines a signature-constant (pseudo) scalar product $(\,|\,)_p$ in each space $T^*_pM \otimes T^*_pM$ is called **(pseudo) Riemannian manifold**. Φ is called **(pseudo) metric** of M.
In particular a $n-$*dimensional* pseudo Riemannian manifold is called **Lorentzian** if the signature of the pseudo scalar product is $(1, n-1)$ (i.e. the canonical form of the metric reads $(-1, +1, \cdots, +1)$.) $\diamondsuit$

Comments 4.1.
(1) It is possible to show that each differentiable manifold can be endowed with a metric.
(2) Assume that $\gamma : [a,b] \to M$ is a differentiable curve on a (pseudo) Riemannian manifold, i.e., $\gamma \in C^\infty([a,b])$ where $\gamma \in C^\infty([a,b])$ means $\gamma \in C^\infty((a,b))$ and furthermore, the limits of derivatives of every order toward a^+ and b^- exists and are finite. It is possible to define the (pseudo) length of γ as

$$L(\gamma) = \int_a^b \sqrt{|(\dot\gamma(t)|\dot\gamma(t))|}dt\,.$$

Above and from now on $(\dot\gamma(t)|\dot\gamma(t))$ indicates $(\dot\gamma(t)|\dot\gamma(t))_{\gamma(t)}$.
(3) A (pseudo) Riemannian manifold M is path-connected and the path between to points $p, q \in M$ can be chosen as differentiable curves. Then, if the manifold is Riemannian (not pseudo), define

$$d(p,q) := \inf \left\{ \int_a^b \sqrt{|(\dot\gamma(t)|\dot\gamma(t))|}dt \;\middle|\; \gamma : [a,b] \to M\,, \gamma \in C^\infty([a,b])\,,\; \gamma(a) = p\,, \gamma(b) = q \right\}\,.$$

$d(p,q)$ is a distance on M, and M turns out to be *metric space* and the associated metric topology coincides with the topology initially given on M.

4.1 Local and global flatness.

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

Definition 4.2. (**Flatness.**) A n-dimensional (pseudo) Riemannian manifold M is said to be **locally flat** if, for every $p \in M$, there is a local chart (U, ϕ) with $p \in U$, which is **canonical**, i.e.,

$$(g_q)_{ij} = diag(-1, \ldots, -1, +1, \ldots, +1)$$

for each $q \in U$, where

$$\Phi(q) = (g_q)_{ij} dx^i|_q \otimes dx^j|_q$$

is the (pseudo) metric represented in the local coordinates $(x^1, \ldots x^n)$ defined by ϕ. (In other words all the bases $\{\frac{\partial}{\partial x^k}|_q\}_{k=1,\ldots,n}$, $q \in U$, are (pseudo) orthonormal bases with respect to the pseudo metric tensor.)
A (pseudo) Riemannian manifold is said to be **globally flat** if there is a global chart which is canonical.$\diamondsuit$

In other words, a (pseudo) Riemannian manifold is locally flat if admits an atlas made of canonical local charts. If that atlas can be reduced to a single chart, the manifold is globally flat.

Examples 4.1.
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. To show it, 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**. 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 $\chi_p : V \to T_pM$ defined in the remark after proposition 2.2 above. The (pseudo) scalar product $(|)$ on V can be exported in each $T_p\mathbb{A}^n$ by defining $(u|v)_p := (\chi_p^{-1}u|\chi_p^{-1}u)$ for all $u, v \in T_p\mathbb{A}^n$. By this way the bases $\{\frac{\partial}{\partial x^i}|_p\}_{i=1,\ldots,n}$ 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 differentiable 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, (θ, z) with $\theta \in]0, \pi[$, $z \in \mathbb{R}$ obtained by restriction of usual cylindric coordinates in $\mathbb{E}^3$ (r, θ, 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 $u, v \in T_pC$ 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 (θ, 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 $\mathbb{M}^4$. Hence it is equipped with a pseudo-metric $\Phi = g_{ij} dx^i \otimes dx^j$ 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*. That is the spacetime of *Special Relativity Theory*.

4.2 Existence of Riemannian metrics.

It is possible to show that any differentiable manifold can be equipped with a Riemannian metric. This result is a straightforward consequence of the existence of a partition of unity (see Section 1). Thus, in particular, it cannot be extended to the analytic case.

Theorem 4.1. *If M is a differentiable manifold, it is possible to define a Riemannian metric Φ on M.*

Proof. Consider a covering of M, $\{U_i\}_{i \in I}$, made of coordinate domains whose closures are compact. Then, using paracompactness, extract a locally finite subcovering $\mathcal{C} = \{V_j\}_{j \in J}$. By construction each V_j admits local coordinates $\phi_j : V_j \to \mathbb{R}^n$. For every $j \in J$ define, in the bases associated with the coordinates, a component-constant Riemannian metric g_j. If $\{h_j\}_{j \in J}$ is a partition of unity subordinate to $\mathcal{C}$ (see theorem 1.1), $\Phi := \sum_{j \in J} h_j g_j$ is well-defined, differentiable (in a neighborhood of any point the sum encompasses only a finite number of differentiable functions) and defines a strictly positive scalar product on each point of M.$\Box$

4.3 Induced metric and isometries.

Let M be a (pseudo) Riemannian manifold with (pseudo) metric tensor Φ. If N is another manifold equipped with an immersion $i_p : N \to M$ it is possible to induce to N a covariant symmetric differentiable tensor field Φ_N associated with Φ. If Φ_N is nondegenerate, it defines a (pseudo) metric called the (pseudo) metric on N induced by M. The procedure is straightforward. Indeed $di_p : T_pN \to T_pM$ is injective and, as a consequence, any $v \in T_pN$ can be seen as a vector in a subspace of T_pM, that subspace being di_pT_pN. In turn we can define the bilinear symmetric form in $T_pN \times T_pN$:

$$\Phi_{Np}(v|u) := \Phi(di_pv|di_pu)$$

Varying $p \in N$ and assuming that $u = U(p)$, $v = V(p)$ where U and V are differentiable vector fields in N, one sees that the map $p \mapsto \Phi_{Np}(V(p)|U(p))$ must be differentiable because it is composition of differentiable functions. We conclude that $p \mapsto \Phi_{Np}$ define a covariant symmetric differentiable tensor field on N. The procedure works, in particular, whenever $N \subset M$ is a submanifold since the inclusion $i : N \hookrightarrow M$ is an embedding.

Definition 4.3. *Let M be a (pseudo) Riemannian manifold with (pseudo) metric tensor Φ and $i : N \to M$ an embedding. The covariant symmetric differentiable tensor field on N, Φ_N, defined by*

$$\Phi_{Np}(v|u) := \Phi(di_pv|di_pu) \quad \textit{for all } p \in N \textit{ and } u, v \in T_pN$$

is called the **metric induced on N by M**.
If $N \subset M$ is connected, $i : N \hookrightarrow M$ is the canonical inclusion and Φ_N is not degenerate, and thus (N, Φ_N) is a (pseudo) Riemannian manifold, it is called (pseudo) Riemannian submanifold of M.

Particular attention deserves the case where $i : N \to M$ is a diffeomorphism and N is equipped with a metric Φ' as well.

Definition 4.4. *Let (M, Φ) and (N, Φ') be two (pseudo) Riemannian manifolds. A diffeomorphism $\phi : N \to M$ is called* **isometry***, and (M, Φ) and (N, Φ') are said to be* **isometric***, if it results $\Phi_N = \Phi'$, or equivalently $\Phi'_M = \Phi$.*

Remark 4.1.
(1) We stress that, in general, Φ_N is not a (pseudo) metric on N because there are no guarantee for it being nondegenerate. Nevertheless, if Φ is a proper metric, i.e. it is positive defined, Φ_N is necessarily positive defined by construction. In that case, when $N \subset M$, (N, Φ_N) is a Riemannian submanifold of M if and only if N is connected.
(2) What is the coordinate form of Φ_N? Fix $p \in N$, a local chart in N, (U, ϕ) with $p \in U$ and another local chart in M, (V, ψ) with $p \in V$ once again. Use the notation $\phi : q \mapsto (y^1(q), \ldots, y^n(q))$ and $\psi : r \mapsto (x^1(r), \ldots, x^m(r))$. The inclusion map $i : N \hookrightarrow M$ admits the coordinate representation in a neighborhood of p

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

Finally, in the considered coordinate frames one has $\Phi = g_{ij}dx^i \otimes dx^j$ and $\Phi_N = g_{(N)kl}dy^k \otimes dy^l$. With the given notation, if $u \in T_pN$, using the expression of df_p given in the remark (1) after definition 3.1 with $f = i$, one sees that, in our coordinate frames

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

As a consequence, using the definition of Φ_N in definition 4.3, one finds

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

Thus

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

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 g_{ij} and the tensor g_{Nkl} evaluated at the same point p with coordinates $(y^1, \ldots, y^n)$ in N and $(x^1(y^1, \ldots, y^n), \ldots, x^m(y^1, \ldots, y^n))$ in M reads

$$g_{(N)kl}(p) \ = \ \frac{\partial x^i}{\partial y^k}|_{(y^1,\ldots y^n)} \ \frac{\partial x^j}{\partial y^l}|_{(y^1,\ldots y^n)} \ g_{ij}(p) \ .$$

Examples 4.2.
1. Let us consider the submanifold given by the cylinder $C \subset \mathbb{E}^3$ defined in the example 4.1.2. 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

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

where x^1, x^2, x^3 are local coordinates in $\mathbb{E}^3$ defined about a point $q \in C$ and y^1, y^2 are analogous coordinates on C defined about 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 about 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

$$\Phi = du^1 \otimes du^1 + du^2 \otimes du^2 + du^3 \otimes du^3 \ ,$$

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

$$g_{rr} = \frac{\partial x^i}{\partial r} \frac{\partial x^j}{\partial r} \delta_{ij} = 1$$

$$g_{\theta\theta} = \frac{\partial x^i}{\partial \theta}\frac{\partial x^j}{\partial \theta}\delta_{ij} = r^2$$

$$g_{\theta\theta} = \frac{\partial x^i}{\partial z}\frac{\partial x^j}{\partial z}\delta_{ij} = 1$$

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

$$\Phi = dr \otimes dr + r^2 d\theta \otimes d\theta + dz \otimes dz$$

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

$$\Phi_C = \frac{\partial x^i}{\partial y^k}\frac{\partial x^j}{\partial y^l}\ g_{ij} dy^j \otimes dy^l = r|^2_C d\theta \otimes d\theta + dz \otimes dz = d\theta \otimes d\theta + dz \otimes dz\ .$$

That is

$$\Phi_C = d\theta \otimes d\theta + dz \otimes dz\ .$$

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 about 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 $\mathbb{M}^4$, that is the affine four-dimensional space $\mathbb{A}^4$ equipped with the scalar product (defined in the vector space of 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

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

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 Σ is actually a submanifold of $\mathbb{M}^4$ with dimension 3. A global coordinate system on Σ 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 Σ? It can be obtained, in components, by the relation

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

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. In Theory of Relativity such submanifolds with degenerate

induced metric are called “null submanifolds” or “light-like manifolds”.
3. In view of **1** in Examples 4.1 and of the introduced tools, we can re-state the definition of locally flat (pseudo)Riemannian manifold as follws.

Definition 4.5. (**Flatness.**) A n-dimensional (pseudo) Riemannian manifold M is said to be **locally flat** if it admits a covering $\{U_i\}_{i\in\mathscr{I}}$ made of open subsets, such that every (pseudo) Riemannian submanifold $(U_i, g\restriction_{U_i})$ is isometric to a (pseudo) Euclidean space with signature (m,p), $\mathbb{E}^n$.$\diamondsuit$

Chapter 5

Covariant Derivative. Levi-Civita's Connection.

5.1 Affine connections and covariant derivatives.

Consider a differentiable manifold M. Suppose for simplicity that $M = \mathbb{A}^n$, the n-dimensional affine space. The global coordinate systems obtained by fixing an origin $O \in \mathbb{A}^n$, a basis $\{e_i\}_{i=1,\ldots,n}$ in V, the vector space of $\mathbb{A}^n$ and posing:

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

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 $(x^1, \ldots, x^n)$ and $(y^1, \ldots, y^n)$ are related by non-homogeneous linear transformations determined by real constants ${A^i}_j, B^i$,

$$y^i = {A^i}_j x^j + B^i\ ,$$

where the matrix of coefficients ${A^i}_j$ is non-singular.
Let $(x^1, \ldots, x^n)$ be a system of Cartesian coordinates on $\mathbb{A}^n$. Each vector field X can be decomposed as $X_p = X^i_p \frac{\partial}{\partial x^i}|_p$. Changing coordinate system but remaining in the class of Cartesian coordinate systems, components of vectors transform as

$$X'^i = {A^i}_j X^j\ ,$$

if the primed coordinates are related with the initial ones by:

$$x'^i = {A^i}_j x^j + B^i\ .$$

If Y is another differentiable 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:

$$(\nabla_X Y)_p := X^j_p \frac{\partial Y^i_p}{\partial x^j} \frac{\partial}{\partial x^i}|_p\ ,$$

or, using the index notation and omitting the index p,

$$(\nabla_X Y)^i = X^j \frac{\partial Y^i}{\partial x^j} \, .$$

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 pass to another Cartesian coordinate system we trivially get:

$$(\nabla_X Y)'^i_p = A^i{}_j (\nabla_X Y)^j_p \, ,$$

since the coefficients $A^i{}_j$ do *not* depend on p and the action of derivatives on these coefficients do not produce added terms in the transformation rule above. Hence, the given definition 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 in p to the value of Y in a point $q = p + hX_p$, where the notation means nothing but that $\overrightarrow{pq} = h\chi_p Y_p$, $\chi_p : T_p\mathbb{A}^n \to V$ being the natural isomorphism between $T_p\mathbb{A}^n$ and the vector space V of the affine structure of $\mathbb{A}^n$ (see the remark after proposition 2.2). This difference has to be divided by h and the limit $h \to 0$ 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 $\chi_p : T_p\mathbb{A}^n \to V$ and $\chi_q : T_q\mathbb{A}^n \to V$. As a consequence $\mathcal{A}[q,p] := \chi_p^{-1} \circ \chi_q : T_q\mathbb{A}^n \to T_p\mathbb{A}^n$ is a well-defined vector space isomorphism. The very definition of $(\nabla_X Y)_p$ can be given as

$$(\nabla_X Y)_p := \lim_{h \to 0} \frac{\mathcal{A}[p + hX_p, p] Y_{p+hX_p} - Y_p}{h} \, .$$

Passing in 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 $(\nabla_X Y)_p$. Without such a structure, that is in a generic manifold, it is not so simple to define the notion of derivative of a vector field in a point. Remaining in the 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 the constant coefficients A^i_j have to be replaced by $\frac{\partial x'^i}{\partial x^j}|_p$ which depend on p. What is the form of $\nabla_X Y$ in generic coordinate systems? And what about the definition of $\nabla_X Y$ in general differentiable manifolds which are *not* affine spaces? We shall see that the answer to these questions enjoy an interesting interplay.

5.1.1 Affine connections.

The key-idea to give a general answer to the second question is to generalize the properties of the operator ∇_X above.

Definition 5.1. (**Affine Connection and Covariant Derivative.**) Let M be a differentiable manifold. An **affine connection** or **covariant derivative** ∇, is a map

$$\nabla : (X, Y) \mapsto \nabla_X Y \ ,$$

where $X, Y, \nabla_X Y$ are differentiable contravariant vector fields on 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 differentiable functions f, g and differentiable vector fields X, Y, Z;
(2) $(\nabla_Y fX)_p = Y_p(f)X_p + f(p)(\nabla_Y X)_p$ for all differentiable vector field X, Y and differentiable functions f;
(3) $(\nabla_X(\alpha Y + \beta Z))_p = \alpha(\nabla_X Y)_p + \beta(\nabla_X Z)_p$ for all $\alpha, \beta \in \mathbb{R}$ and differentiable vector fields X, Y, Z.
The contravariant vector field $\nabla_Y X$ is called the **covariant derivative of X with respect to** Y (and the affine connection ∇). $\diamondsuit$

Remark 5.1.
(1) It is very important that the three relations written in the definition are understood pointwisely also if, very often, they are written without referring to any 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.
In particular this means that, it make sense to define the derivative in p, $\nabla_{X_p} Y$, where $X_p \in T_pM$ is a simple vector and not a vector field. Indeed one can always extend X_p to a vector field in M using lemma 2.5, 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 by 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 that vanishes outside a coordinate patch (U, ϕ) with coordinates $x^1, \ldots, x^n$, 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 in p and h vanishes in a neighborhood of p, $X_p(h) = 0$ (cf lemma 2.3). 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 ∇ through the class of isomorphisms $\mathcal{A}[q, p]$. In fact,

$$(\nabla_X Y)_p := \lim_{h\to 0} \frac{\mathcal{A}[p + hX_p, p]Y_{p+hX_p} - Y_p}{h}$$

satisfies all the requirements above. The point is that, 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. Exercise 5.1.1 below provided an appropriate answer.

5.1.2 Connection coefficients.

Let us come back to the general definition 5.1 In components referred to any local coordinate system $(x^1, \ldots, x^n)$ 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 5.1. 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 the index $_*$ below. It holds (using notations introduced in the remark (2) above):

$$(\nabla_X Y)_p = \nabla_{X^i(p)\frac{\partial}{\partial x^i}|_p}\left(Y^j_*\frac{\partial}{\partial x^j}_* + Z\right)$$

where the vector field 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 remark (3) above. Hence

$$(\nabla_X Y)_p = \nabla_{X^i(p)\frac{\partial}{\partial x^i}|_p} Y^j_*\frac{\partial}{\partial x^j}_* = X^i(p)Y^j_*(p)\nabla_{\frac{\partial}{\partial x^i}|_p}\frac{\partial}{\partial x^j}_* + X^i(p)\frac{\partial Y^j}{\partial x^i}_*|_p\frac{\partial}{\partial x^j}_*|_p\,.$$

In other words, in our hypotheses:

$$(\nabla_X Y)_p = X^i(p)Y^j(p)\nabla_{\frac{\partial}{\partial x^i}|_p}\frac{\partial}{\partial x^j}_* + X^i(p)\frac{\partial Y^j}{\partial x^i}|_p\frac{\partial}{\partial x^j}|_p\,.$$

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 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

$$\nabla_{\frac{\partial}{\partial x^i}|_p}\frac{\partial}{\partial x^j} = \langle\nabla_{\frac{\partial}{\partial x^i}|_p}\frac{\partial}{\partial x^j},\, dx^k|_p\rangle\frac{\partial}{\partial x^k}|_p \ := \Gamma^k_{ij}(p)\frac{\partial}{\partial x^k}|_p\,.$$

The coefficients $\Gamma^k_{ij} = \Gamma^k_{ij}(p)$ are differentiable 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 can be written down as:

$$(\nabla_X Y)^i = X^j\left(\frac{\partial Y^i}{\partial x^j} + \Gamma^i_{jk}Y^k\right)\,.$$

Fix a differentiable contravariant vector field X and $p \in M$. The linear map $Y_p \mapsto (\nabla_{Y_p}X)_p$ (taking the remark (2) above into account) and lemma 2.5 define a tensor, $(\nabla X)_p$ of class $(1,1)$ in $T^*_pM \otimes T_pM$ such that the (only possible) contraction of Y_p and $(\nabla X)_p$ is $(\nabla_Y X)_p$. Varying

$p \in M$, $p \mapsto (\nabla X)_p$ define a smooth $(1,1)$ tensor field ∇X because in local coordinates its components are differentiable because they are given by coefficients

$$\frac{\partial X^i}{\partial x^j} + \Gamma^i_{jk} X^k =: \nabla_j X^i =: {X^i}_{,j} \,.$$

∇X is called **covariant derivative of** X (with respect to the affine connection ∇).
In components we have

$$(\nabla_Y X)^i = Y^j {X^i}_{,j} \,.$$

5.1.3 Transformation rule of the connection coefficients.

Now we are interested in the transformation rule of the connection coefficients under change of coordinates. We pass from local coordinates $(x^1, \ldots, x^n)$ to local coordinates $(x'^1, \ldots, x'^n)$ and the connection coefficients change form Γ^k_{ij} to ${\Gamma'}^h_{pq}$.

$$\Gamma^k_{ij} = \langle \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}\,, dx^k \rangle = \langle \nabla_{\frac{\partial x'^p}{\partial x^i}\frac{\partial}{\partial x'^p}} (\frac{\partial x'^q}{\partial x^j}\frac{\partial}{\partial x'^q}), \frac{\partial x^k}{\partial x'^h} dx'^h \rangle = \frac{\partial x^k}{\partial x'^h}\frac{\partial x'^p}{\partial x^i} \langle \nabla_{\frac{\partial}{\partial x'^p}} (\frac{\partial x'^q}{\partial x^j}\frac{\partial}{\partial x'^q}), dx'^h \rangle \,.$$

Expanding the last term we get

$$\frac{\partial x^k}{\partial x'^h}\frac{\partial x'^p}{\partial x^i} \nabla_{\frac{\partial}{\partial x'^p}} (\frac{\partial x'^q}{\partial x^j}) \; \langle \frac{\partial}{\partial x'^q}\,, dx'^h \rangle + \frac{\partial x^k}{\partial x'^h}\frac{\partial x'^p}{\partial x^i}\frac{\partial x'^q}{\partial x^j} \langle \nabla_{\frac{\partial}{\partial x'^p}} \frac{\partial}{\partial x'^q}\,, dx'^h \rangle \,,$$

which can be re-written as

$$\frac{\partial x^k}{\partial x'^h}\frac{\partial x'^p}{\partial x^i}\frac{\partial^2 x'^h}{\partial x'^p \partial x^j} + \frac{\partial x^k}{\partial x'^h}\frac{\partial x'^p}{\partial x^i}\frac{\partial x'^q}{\partial x^j} {\Gamma'}^h_{pq}$$

or

$$\Gamma^k_{ij} = \frac{\partial x^k}{\partial x'^h}\frac{\partial^2 x'^h}{\partial x^i \partial x^j} + \frac{\partial x^k}{\partial x'^h}\frac{\partial x'^p}{\partial x^i}\frac{\partial x'^q}{\partial x^j} {\Gamma'}^h_{pq} \,.$$

The obtained result shows that the connection coefficients do not define a tensor because of the non-homogeneous former term in the right-hand side above.

Remark 5.2.
If ∇ is the affine connection naturally associated with the affine structure of an affine space $\mathbb{A}^n$, it is clear that $\Gamma^k_{il} = 0$ in every Cartesian coordinate system. As a consequence, in a generic coordinate system

$$\Gamma^k_{ij} = \frac{\partial x^k}{\partial x'^h}\frac{\partial^2 x'^h}{\partial x^i \partial x^j}$$

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)?". The answer is

$$(\nabla_X Y)^i = X^j (\frac{\partial Y^i}{\partial x^j} + \Gamma^i_{jk} Y^k) \,,$$

where the coefficients Γ^i_{jk} are defined as

$$\Gamma^k_{ij} = \frac{\partial x^k}{\partial x'^h}\frac{\partial^2 x'^h}{\partial x^i \partial x^j} ,$$

the primed coordinates being Cartesian coordinates.

5.1.4 Torsion tensor

By Schwarz' theorem, the inhomogeneous term in

$$\Gamma^k_{ij} = \frac{\partial x^k}{\partial x'^h}\frac{\partial^2 x'^h}{\partial x^i \partial x^j} + \frac{\partial x^k}{\partial x'^h}\frac{\partial x'^p}{\partial x^i}\frac{\partial x'^q}{\partial x^j}\Gamma'^h{}_{pq} ,$$

drops out when considering the transformation rules of coefficients:

$$T^i_{jk} := \Gamma^i_{jk} - \Gamma^i_{kj} .$$

Hence, these coefficients define a tensor field which, in local coordinates, is represented by:

$$T(\nabla) = (\Gamma^i_{jk} - \Gamma^i_{kj})\frac{\partial}{\partial x^i} \otimes dx^j \otimes dk^k .$$

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

$$((\nabla_X Y)_p - (\nabla_Y X)_p - [X,Y]_p)^k = T(\nabla)_p{}^k{}_{ij} X^i_p Y^j_p ,$$

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 can be defined as a bilinear mapping which associates pairs of differentiable vector fields X, Y with a differentiable vector field $T(\nabla)(X,Y)$ along the rule

$$T(\nabla)_p(X_p, Y_p) = \nabla_{X_p} Y - \nabla_{Y_p} X - [X,Y]_p .$$

It is worthwhile stressing that this identity shows that *the difference of* $\nabla_{X_p} Y - \nabla_{Y_p} X$ *and* $[X,Y]_p$ *does not depend on the shape of the fields* X *and* Y*, but on the values attained by them at the point* p *only.* 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 the second definition of torsion tensor field we have the following useful result.

Proposition 5.1. *Let* ∇ *be an affine connection on a differentiable manifold* M*. If* ∇ *is* torsion free*, i.e., the torsion tensor* $T(\nabla)$ *field vanishes on* M*,*

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

for every pair of contravariant differentiable vector fields X, Y*.* $\diamondsuit$

5.1.5 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.

Proposition 5.2. *The assignment of an affine connection on a differentiable manifold M is completely equivalent to the assignment of coefficients $\Gamma^k_{ij}(p)$ in each local coordinate system, which differentiably depend on the point p and transform as*

$$\Gamma^k_{ij}(p) = \frac{\partial x^k}{\partial x'^h}|_p \frac{\partial^2 x'^h}{\partial x^i \partial x^j}|_p + \frac{\partial x^k}{\partial x'^h}|_p \frac{\partial x'^p}{\partial x^i}|_p \frac{\partial x'^q}{\partial x^j}|_p \Gamma'^h_{pq}(p) \,, \tag{5.1}$$

under change of local coordinates. More precisely:
(a) *if an affine connection ∇ is given, coefficients Γ^i_{ij} associated with ∇ which satisfy (5.1) are defined by*

$$\Gamma^k_{ij}(p) := \langle \nabla_{\frac{\partial}{\partial x^i}|_p} \frac{\partial}{\partial x^j}|_p \,, dx^k|_p \rangle \,,$$

(b) *if coefficients $\Gamma^k_{ij}(p)$ are assigned for every point $p \in M$ and every coordinate system of an atlas of M, such that (5.1) hold, an affine connection associated with this assignment is given by*

$$(\nabla_X Y)^i_p = X^j_p \left(\frac{\partial Y^i}{\partial x^j}|_p + \Gamma^i_{jk}(p) Y^k_p\right) .$$

in every coordinate patch of the atlas, for all vector fields X, Y and every point $p \in M$;
(c) *if ∇ and ∇' are two affine connections on M such that the coefficients $\Gamma^k_{ij}(p)$ and $\Gamma'^k_{ij}(p)$ respectively associated to the connections as in (a) coincide for every point $p \in M$ and every coordinate system about p in a given atlas on M, then $\nabla = \nabla'$.* $\diamondsuit$

Note. Shortly, after we have introduced the notion of geodesic segment and parallel transport, we shall come back to the geometrical meaning of the covariant derivative.

5.2 Covariant derivative of tensor fields.

If M is a differentiable manifold equipped with an affine connection ∇, it is possible to extend the action of the covariant derivatives to all differentiable tensor fields. In other words if X is a differentiable vector field and u is a differentiable tensor field, it is possible to define a new tensor field $\nabla_X u$, *of the same order as* u, uniquely. It is done by assuming the following further requirements on the action of ∇_X, *which are supposed to hold true point-wisely, but we omit* $_p$ *everywhere for the shake of notational simplicity*;
(4) $\nabla_X(\alpha u + \beta v) = \alpha \nabla_X u + \beta \nabla_X v$ for all $\alpha, \beta \in \mathbb{R}$, differentiable tensor fields u, v and differentiable vector fields X.
(5) $\nabla_X f := X(f)$ for all differentiable vector fields X and differentiable functions f.
(6) $\nabla_X(t \otimes u) := (\nabla_X t) \otimes u + t \otimes \nabla_X u$ for all differentiable tensor fields u, t and vector fields X.

(7) $\nabla_X\langle Y,\eta\rangle = \langle\nabla_X Y,\eta\rangle + \langle Y,\nabla_X\eta\rangle$ for all differentiable vector fields X, Y and differentiable covariant vector fields η.

In particular, the action of ∇_X on covariant vector fields turns out to be defined by the requirements above as follows.

$$\nabla_X\,\eta = \langle\frac{\partial}{\partial x^k},\nabla_X\eta\rangle\,dx^k = \Big(\nabla_X\langle\frac{\partial}{\partial x^k},\eta\rangle\Big)\,dx^k - \langle\nabla_X\frac{\partial}{\partial x^k},\eta\rangle\,dx^k\,,$$

where

$$\nabla_X\langle\frac{\partial}{\partial x^k},\eta\rangle = \nabla_X\eta_k = X(\eta_k) = X^i\frac{\partial\eta_k}{\partial x^i}\,,$$

and

$$\langle\nabla_X\frac{\partial}{\partial x^k},\eta\rangle = X^i\eta_r\langle\nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^k},dx^r\rangle = X^i\eta_r\Gamma^r_{ik}\,.$$

Putting all together we have:

$$(\nabla_X\,\eta)_k dx^k = X^i(\frac{\partial\eta_k}{\partial x^i} - \Gamma^r_{ik}\eta_r)\,dx^k\,,$$

which is equivalent to:

$$(\nabla\eta)_{ki} = \eta_{k,i} := \frac{\partial\eta_k}{\partial x^i} - \Gamma^r_{ik}\eta_r\,,$$

where we have introduced the **covariant derivative** of the covariant vector field η, $\nabla\eta$, as the unique tensor field of tensors in $T^*_pM\otimes T^*_pM$ such that the contraction of X_p and $(\nabla\eta)_p$ (with respect to the space corresponding to the index i) is $(\nabla_{X_p}\eta)_p$.

It is simply proved that, given an affine connection ∇, there is exactly one map which transforms tensor fields to tensor fields (preserving their order) and satisfies all the requirements above. Uniqueness is straightforward. Indeed, as shown above, the action on covariant tensor fields is uniquely fixed, the action on scalar fields in defined in (5), finally (6) determines the action on generic tensor fields. 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, ∇t, which has to be interpreted as the unique tensor field of tensors in $T^*_pM\otimes S_pM$ (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 $(\nabla_{X_p}t)_p$:

$$\begin{aligned}(\nabla t)^{i_1\ldots i_l}{}_{j_1\ldots j_k r}(p) := {}& \frac{\partial t^{i_1\ldots i_l}{}_{j_1\ldots j_k}}{\partial x^r}|_p + \Gamma^{i_1}_{sr}(p)t^{s\ldots i_l}{}_{j_1\ldots j_k}(p) + \ldots + \Gamma^{i_l}_{sr}(p)t^{i_1\ldots s}{}_{j_1\ldots j_k}(p)\\ & - \Gamma^s_{rj_1}(p)t^{i_1\ldots i_l}{}_{s\ldots j_k}(p) - \ldots - \Gamma^s_{rj_k}(p)t^{i_1\ldots i_l}{}_{j_1\ldots s}(p)\,. \end{aligned} \tag{5.2}$$

The reader can easily check the validity of requirements (4)-(7) for the map $\nabla : (X,t)\mapsto\nabla_X t$ defined in that way.

The following notation is used in the literature:

$$t^{i_1\ldots i_l}{}_{j_1\ldots j_k,r} := (\nabla t)^{i_1\ldots i_l}{}_{j_1\ldots j_k r} = \nabla_r t^{i_1\ldots i_l}{}_{j_1\ldots j_k}\,.$$

5.3 Levi-Civita's connection.

Let us show that, if M is (pseudo) Riemannian, there is a preferred affine connection completely determined by the metric. This is **Levi-Civita's affine connection**.

Theorem 5.1. *Let M be a (pseudo) Riemannian manifold with metric locally represented by $\Phi = g_{ij} dx^i \otimes dx^j$. There is exactly one affine connection ∇ such that :*
(1) *it is* **metric***, i.e.,* $\nabla \Phi = 0$
(2) *it is* **torsion free***, i.e.,* $T(\nabla) = 0$.
That is the **Levi-Civita connection** *which is defined by the connection coefficients, called* **Christoffel's coefficients***,:*

$$\Gamma^i_{jk} = \{_j{}^i{}_k\} := \frac{1}{2} g^{is} \left(\frac{\partial g_{ks}}{\partial x^j} + \frac{\partial g_{sj}}{\partial x^k} - \frac{\partial g_{jk}}{\partial x^s} \right) .$$

♢

Proof. Assume that a connection with the required properties exits. Expanding (1) and rearranging the result, we have:

$$-\frac{\partial g_{ij}}{\partial x^k} = -\Gamma^s_{ki} g_{sj} - \Gamma^s_{kj} g_{is} ,$$

twice cyclically permuting indices and changing the overall sign we get also:

$$\frac{\partial g_{ki}}{\partial x^j} = \Gamma^s_{jk} g_{si} + \Gamma^s_{ji} g_{ks} ,$$

and

$$\frac{\partial g_{jk}}{\partial x^i} = \Gamma^s_{ij} g_{sk} + \Gamma^s_{ik} g_{js} .$$

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:

$$\frac{\partial g_{ki}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k} = 2\Gamma^s_{ij} g_{sk} .$$

Contracting both sides with $\frac{1}{2} g^{kr}$ and using $g_{sk} g^{kr} = \delta^r_s$ we get:

$$\Gamma^r_{ij} = \frac{1}{2} g^{rk} (\frac{\partial g_{ki}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k}) = \frac{1}{2} g^{rk} (\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}) = \{_i{}^r{}_j\} .$$

We have proved that, if a connection satisfying (1) and (2) exists, its connection coefficients have the form of has the form of Christoffel's coefficients. This fact also implies that, if such a connection exists, it must be unique. The coefficients

$$\{_j{}^i{}_k\}(p) := \frac{1}{2} g^{is}(p) (\frac{\partial g_{ks}}{\partial x^j}|_p + \frac{\partial g_{sj}}{\partial x^k}|_p - \frac{\partial g_{jk}}{\partial x^s}|_p)$$

define an affine connection because they transform as:

$$\{_i{}^k{}_j\}(p) = \frac{\partial x^k}{\partial x'^h}|_p \frac{\partial^2 x'^h}{\partial x^i \partial x^j}|_p + \frac{\partial x^k}{\partial x'^h}|_p \frac{\partial x'^p}{\partial x^i}|_p \frac{\partial x'^q}{\partial x^j}|_p \ \{_p{}^h{}_q\}'(p) ,$$

as one can directly verify. This concludes the proof. □

Remark 5.3.
(1) The 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,

$$g_{ki}\nabla_l t^{ij}{}_r = \nabla_l \left(g_{ki} t^{ij}{}_r\right) .$$

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 Levi-Civita's connection. In particular, this implies that the connection ∇ 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, θ, ϕ 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 δ_{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 $\frac{\partial}{\partial \phi} = \sqrt{g_{\phi\phi}}\mathbf{e}_\phi$.

Exercises 5.1.
1. Show that, if ∇_k are affine connections on a manifold M, then $\nabla = \sum_k f_k \nabla_k$ is an affine connection on M if the smooth functions $f_k : M \to \mathbb{R}$ satisfy $f_k \geq 0$ and $f_k(q) = 0$ in a neighborhood of p, barring a finite number of indices k depending on $p \in M$, and $\sum_k f_k(p) = 1$ for every $p \in M$ (i.e. $\sum_k f_k \nabla_k$ is a 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 4.1, there is a Riemannian metric Φ defined on M. As a consequence M admits the Levi-Civita connection associated with Φ. (2) Let ω, η be a pair of co-vector fields defined in M and X a vector field in M. Suppose that they are somewhere nonvanishing and $\omega \neq \eta$ (these fields exist due to lemma 2.5 and using Φ to pass to co-vector fields from vector fields). Let Ξ be the tensor field with $\Xi_p := X_p \otimes \omega_p \otimes \eta_p$ for every $p \in M$. If $\Gamma^i{}_{jk}$ are the Levi-Civita connection coefficients associated with Φ in any coordinate patch in M, define $\Gamma'^i{}_{jk} := \Gamma^i{}_{jk} + \Xi^i{}_{jk}$ in the same coordinate patch. By construction these coefficients transforms as connection coefficients under a change of coordinate frame. As a consequence of proposition 5.2 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

$$\Gamma^i_{ij}(p) = \frac{\partial \ln \sqrt{|g|}}{\partial x^j}|_p \,.$$

where $g(p) = det[g_{ij}(p)]$ in the considered coordinates.

Solution. Notice that the sign of g is fixed it depending on the signature of the metric. It holds

$$\frac{\partial \ln \sqrt{|g|}}{\partial x^j} = \frac{1}{2g}\frac{\partial g}{\partial x^j} \,.$$

Using the formula for expanding derivatives of determinants and expanding the relevant determinants in the expansion by rows, one sees that

$$\frac{\partial g}{\partial x^j} = \sum_k (-1)^{1+k} cof_{1k} \frac{\partial g_{1k}}{\partial x^j} + \sum_k (-1)^{2+k} cof_{2k} \frac{\partial g_{2k}}{\partial x^j} + \ldots + \sum_k (-1)^{n+k} cof_{nk} \frac{\partial g_{nk}}{\partial x^j} \,.$$

That is

$$\frac{\partial g}{\partial x^j} = \sum_{i,k} (-1)^{i+k} cof_{ik} \frac{\partial g_{ik}}{\partial x^j} \,,$$

On the other hand, Cramer's formula for the inverse matrix of $[g_{ik}]$, $[g^{pq}]$, says that

$$g^{ik} = \frac{(-1)^{i+k}}{g} cof_{ik}$$

and so,

$$\frac{\partial g}{\partial x^j} = g g^{ik} \frac{\partial g_{ik}}{\partial x^j} \,,$$

hence

$$\frac{1}{2g}\frac{\partial g}{\partial x^j} = \frac{1}{2} g^{ik} \frac{\partial g_{ik}}{\partial x^j}$$

But direct inspection proves that

$$\Gamma^i_{ij}(p) = \frac{1}{2} g^{ik} \frac{\partial g_{ik}}{\partial x^j} \, .$$

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 $divX := \nabla_i X^i$ with respect to the Levi-Civita connection can be computed by using:

$$(divV)(p) = \frac{1}{\sqrt{|g(p)|}} \frac{\partial \sqrt{|g|} V^i}{\partial x^i} |_p \, .$$

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 the remark (2) above).

7. Execute the exercise 3.1.3 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 defined by:

$$\Delta f := g^{ij} \nabla_j \nabla_i f \, ,$$

where ∇ is the Levi-Civita connection. Show that, in coordinates:

$$(\Delta f)(p) = \frac{1}{\sqrt{|g(p)|}} \left(\frac{\partial}{\partial x^i} \sqrt{|g|} g^{ij} \frac{\partial}{\partial x^j} \right) |_p f \, .$$

9. Consider cylindrical coordinates in $\mathbb{R}^3$, (r, θ, 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, θ, ϕ). Show that:

$$\Delta f = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2\theta} \frac{\partial^2 f}{\partial \phi^2} \, .$$

5.4 Geodesics: parallel transport approach.

Take a manifold M equipped with an affine connection ∇. It is possible to generalize the concept of straight line by introducing the concept of *geodesic*. Consider a smooth regular curve $\gamma : (a,b) \to M$ (we remark that the definition of a curve used here includes a *preferred choice* for the parameter t and it is assumed that $\dot\gamma(t) \neq 0$ for $t \in [a,b]$ so that the map $\gamma : (a,b) \to \gamma((a,b))$ is bijective) and we also assume that $\gamma((a,b))$ is an embedded submanifold of M. Next fix a smooth vector field V defined on γ, that is a smooth assignation $t \mapsto V(t) \in T_{\gamma(t)}M$. It is possible to extend V to a smooth vector field $\widetilde{V}$ defined in a neighborhood N of $\gamma((a,b))$. (Indeed, in view of the fact that $\gamma((a,b))$ is an embedded submanifold of dimension 1, in a open neighborhood U in M of any point $p \in \gamma((a,b))$, γ is described, in a local chart, as $x^k = 0$ for $k > 1$ and x^1 runs along the considered portion of γ. $V(q)$ for $q \in \gamma((a,b)) \cap U$ is defined exactly on the line $x^k = 0$ for $k > 1$ and takes the form $V = V^k(x^1)\frac{\partial}{\partial x^k}|_{\gamma(q)}$ thereon. Thus one can extend V assuming that, in coordinates, it verifies $\widetilde{V}^k(x^1,\ldots,x^n) := V^k(x^1)$ in that local chart. Using a partition of the unity subordinated to an open covering of $\gamma((a,b))$ made of such domains of local charts one easily obtains the smooth extension $\widetilde{V}$ along the whole curve γ.) Hence $\widetilde{V} \restriction_{\gamma((a,b))} = V$. Then we may consider the field $\nabla_{\dot\gamma(t)}V(t) = (\nabla_V \widetilde{V}) \restriction_{\gamma((a,b))}$. It is a trivial task to show that the obtained restriction defines a vector field on $\gamma((a,b))$ which does not depend on the extension $\widetilde{V}$ of V and thus the used notation is appropriate. As a matter of fact, in local coordinates, we have

$$(\nabla_{\dot\gamma(t)}\widetilde{V}(t))^i \restriction_{\gamma(t)} = \dot\gamma^k(t)\frac{\partial \widetilde{V}^i}{\partial x^k} + \Gamma^i_{jk}(\gamma(t))\frac{dx^j}{dt}V^k$$

so that:

$$(\nabla_{\dot\gamma(t)}V(t))^i = \frac{dV^i}{dt} + \Gamma^i_{jk}(\gamma(t))\frac{dx^j}{dt}V^k\,, \tag{5.3}$$

where the independence from the extension is apparent. More generally, even if $\gamma((a,b))$ is not an embedded submanifold, but it is a smooth curve without self-intersections and if $(a,b) \ni t \mapsto V(t) \in T_{\gamma(t)}M$ is smoothly defined, the right hand-side of (5.3) does not depend on the chosen local chart, as one can prove by direct inspection using the rule of transformation of the connection coefficients when changing coordinates. This fact provides an intrinsic definition of $\nabla_{\dot\gamma(t)}V(t)$ as a vector defined at every point $\gamma(t)$. In particular, if focusing on the special case $V := \dot\gamma(t)$, we have:

$$(\nabla_{\dot\gamma(t)}\dot\gamma(t))^i = \frac{d^2x^i}{dt^2} + \Gamma^i_{jk}(\gamma(t))\frac{dx^j}{dt}\frac{dx^k}{dt}\,, \tag{5.4}$$

where γ is given by $n = dimM$ smooth functions $x^i = x^i(t)$. If ∇ is Levi-Civita's connection in $\mathbb{R}^n$ or in an affine space referred to a metric which is everywhere constant and diagonal in Cartesian coordinates, in Cartesian coordinate system it holds respectively:

$$(\nabla_{\dot\gamma(t)}V(t))^i = \frac{dV^i}{dt}\,, \quad (\nabla_{\dot\gamma(t)}\dot\gamma(t))^i = \frac{d^2x^i}{dt^2}.$$

As a consequence, straight lines are the unique solutions of $\nabla_{\dot\gamma(t)}\dot\gamma^i(t) \equiv 0$ in those spaces. More precisely, if $\gamma = \gamma(t)$ is a solution of the equation above, in whatever (generally local) Cartesian coordinate system, the expression for the curve γ, parametrized by the parameter $t \in (c, d)$, has the form $x^i(t) = a^i t + b^i$ for $2n$ constants $a^1, \ldots, a^n, b^1, \ldots, b^n$.

5.4.1 Parallel transport and geodesics.

In general manifolds we have the following definition which, in a sense, extends the concept of straight line. Notice that we can omit the hypothesis that V is extendible in a neighbourhood of γ since the equation of geodesic transport does not require it to be stated.

Definition 5.2. Let M be a differentiable manifold equipped with an affine connection ∇. Let $\gamma : (a, b) \to M$ be a smooth curve in M with $\dot\gamma(t) \neq 0$ if $t \in (a, b)$ without self intersections. A C^1 vector field T defined on $\gamma((a, b))$, i.e. a C^1 map $(a, b) \ni t \mapsto T(t) \in T_{\gamma(t)}M$, is said to be **transported along γ parallely to** $\dot\gamma$ (and with respect to ∇) if

$$\nabla_{\dot\gamma(t)}T(\gamma(t)) \equiv 0 \quad \text{for all } t \in (a, b) ,$$

where the left-hand side has to be interpreted as in (5.3) referring to any coordinate patch about any point on γ.
γ is a **geodesics** if it transports its tangent vector parallely to itself. In other words γ satisfies the **geodesic equation**:

$$\nabla_{\dot\gamma(t)}\dot\gamma(t) \equiv 0 \quad \text{for all } t \in (a, b).$$

$\diamondsuit$

In the (semi) Riemannian case we have an important result which, in particular holds true for *Levi-Civita connections.*

Proposition 5.3. *If a differentiable manifold M admits both a (pseudo)metric Φ and an affine connection ∇ such that $\nabla\Phi \equiv 0$ (i.e.,* **the connection is metric***),* **the parallel transport preserves the scalar product***. In other words, if X, Y are vector fields defined in a neighborhood of a differentiable curve $\gamma = \gamma(t)$, and both X, Y are parallely transported along γ, it turns out that $t \mapsto (X(\gamma(t))|Y(\gamma(t)))$ is constant.* $\diamondsuit$

Proof. The connection is metric and thus in any local coordinate patch around any point on γ:

$$\frac{d}{dt}(X(\gamma(t))|Y(\gamma(t))) = \frac{d}{dt}\left[g_{ij}(\gamma(t))X^i(\gamma(t))Y^j(\gamma(t))\right]$$

$$= \dot\gamma^k \frac{\partial g_{ij}}{\partial x^k} X^i(\gamma)Y^j(\gamma) + g_{ij}(\gamma)\dot\gamma^k \frac{dX^i(\gamma(t))}{dt} Y^j(\gamma) + g_{ij}(\gamma)\dot\gamma^k X^i(\gamma)\frac{dY^j(\gamma(t))}{dt} .$$

Now the condition $g_{ij,k} = 0$ can be re-written:

$$\frac{\partial g_{ij}}{\partial x^k} = \Gamma^s_{ki} g_{sj} + \Gamma^s_{kj} g_{is}$$

that inserted above produces:

$$\frac{d}{dt}(X(\gamma(t))|Y(\gamma(t))) = (\nabla_{\dot\gamma}X|Y) + (\nabla_{\dot\gamma}Y|X) = 0\,.$$

□

Remark 5.4.
(1) Let M be a differentiable manifold equipped with an affine connection ∇. From known theorems of ordinary differential equations, if $p \in M$ and $v \in T_pM$, there is only one geodesic segment $\gamma = \gamma(t)$ which starts from $\gamma(0) = p$ with initial tangent vector $\dot\gamma(0) = v$ and defined in a neighborhood of 0. This is because the geodesic equation is a second-order equation written in *normal form* in any coordinate system about p and the remaining term in the geodesic equation is smooth. The correct background where one can profitably study the properties of the geodesic equation is TM where the geodesic equation reduces to a first-order equation. Then the theorem of global existence and uniqueness assures that, for a fixed $p \in M$ and $v \in T_pM$ there is only one maximally extended geodesic with those initial data.
A straightforward consequence of the (local) uniqueness theorem is that the tangent vector of a *non constant* geodesic $\gamma : (a, b) \to M$ cannot vanish in any point.
(2) If one changes the parameter of a non constant geodesic $t \mapsto \gamma(t)$, $t \in (a, b)$ to $u = u(t)$ where that mapping is smooth and $du/dt \neq 0$ for all $t \in (a, b)$, the new differentiable curve $\gamma' : u \mapsto \gamma(t(u))$ does not satisfy the geodesic equation in general. Anyway, working in local coordinates and using (5.4), and the geodesic equation for γ, one finds immediately

$$\nabla_{\dot\gamma'(u)}\dot\gamma'(u) = \frac{d^2t}{du^2}\dot\gamma(t)\,.$$

Since $\dot\gamma(t) \neq 0$, we see that γ' satisfies the geodesic equation too, *if and only if* $u = kt + k'$ for some constants $k \in \mathbb{R} \setminus \{0\}$ and $k' \in \mathbb{R}$. These transformations of the parameter of geodesics which preserve the geodesic equations are called **affine transformations** (of the parameter).
(3) If $\gamma : (a, b) \to M$ is fixed, the parallel transport condition

$$\nabla_{\dot\gamma(t)}V(\gamma(t)) \equiv 0 \quad \text{for all } t \in (a, b)\,.$$

can be used as a differential equation. Expanding the left-hand side in local coordinates $(x^1, \ldots, x^n)$ one finds a first-order differential equation for the components of V referred to the bases of elements $\frac{\partial}{\partial x^k}|_{\gamma(t)}$:

$$\frac{dV^i}{dt} = -\Gamma^i_{jk}(\gamma(t))\dot\gamma^j V^k(\gamma(t))\,.$$

As the equation is linear, in normal form with smooth known functions in the right-hand side, the initial vector $V(\gamma(a))$ determines $V = V(\gamma(t))$ uniquely along the *whole* curve, more precisely, along the portion of the curve contained in the domain of the used local coordinates. Using a covering of open sets equipped with coordinates, $V = V(\gamma(t))$ turns out to be fixed along the

whole curve γ. In a certain sense, one may view the solution $t \mapsto V(t)$ as the "transport" and "evolution" of the initial condition $V(\gamma(a))$ along γ itself.
The global existence and uniqueness theorem has an important consequence. If $\gamma : [a, b] \to M$ is any differentiable curve and $u, v \in (a, b)$ with $u < v$, the notion of parallel transport along γ produces an vector space isomorphism $\mathcal{P}_\gamma[\gamma(u), \gamma(v)] : T_{\gamma(u)} \to T_{\gamma(v)}$ which associates $V \in T_{\gamma(u)}$ with that vector in $T_{\gamma(u)}$ which is obtained by parallely transporting V in $T_{\gamma(u)}$.
If ∇ is metric, proposition 5.3 implies that $\mathcal{P}_\gamma[\gamma(u), \gamma(v)]$ also preserves the scalar product, in other words, it is an isometric isomorphism.
(4) Consider a Riemannian manifold M. Let $\gamma = \gamma(t)$ be a non constant geodesic segment with $t \in [a, b)$ with respect to the Levi-Civita connection. The **length ascissa** or **length parameter**

$$s(t) := \int_a^t \sqrt{(\dot\gamma(t')|\dot\gamma(t'))}\, dt' ,$$

defines a linear function $s = kt + k'$ with $k \neq 0$ and thus s can be used to re-parametrize the geodesic. Indeed $(\dot\gamma(t')|\dot\gamma(t'))$ is constant by proposition 5.3 and $(\dot\gamma(t')|\dot\gamma(t')) \neq 0$ because $\dot\gamma(t') \neq 0$.
(5) If the manifold M is equipped with an affine connection M, it is possible to show that each point of $p \in M$ admits a neighborhood U such that, if $q \in U$, there is a unique geodesic segment γ completely contained in U from p to q. These neighborhoods are called **geodesically convex**.

Examples 5.1. 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 $\Phi = g_{ij} dx^i \otimes dx^j$ 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 the spacetime is flat and it is an affine four dimensional space, it is called *Minkowski spacetime*. That is the spacetime of *Special Relativity Theory*.
If $V \in T_pM$, $V \neq 0$, for some event $p \in M$, V is called *timelike*, *lightlike* (or *null*), *spacelike* if, respectively $(V|V) < 0$, $(V|V) = 0$, $(V|V) > 0$. A curve $\gamma : \mathbb{R} \to M$ is defined similarly referring to its tangent vector $\dot\gamma$ provided $\dot\gamma$ preserves the sign of $(\dot\gamma|\dot\gamma)$ along the curve itself. The evolution of a particle is represented by a *world line*, i.e., a timelike differentiable curve $\gamma : u \mapsto \gamma(u)$ and the length parameter (length ascissa) along the curve

$$t(u) := \int_a^u \sqrt{|(\dot\gamma(u')|\dot\gamma(u'))|}\, du' ,$$

(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 welding the tangent space $T_{\gamma(t)}M$ is naturally decomposed as $T_{\gamma(t)}M = L(\dot\gamma(t)) \oplus \Sigma_{\gamma(t)}$, where $L(\dot\gamma(t))$ is the linear space spanned by $\dot\gamma(t)$ and $\Sigma_{\gamma(t)}$ is the orthogonal space to $L(\dot\gamma(t))$. It is simple to prove that the metric $\Phi_{\gamma(t)}$ 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 γ has a timelike, lightlike, spacelike initial tangent vector, any other tangent vector along γ 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.

5.4.2 Back on the meaning of the covariant derivative.

The notion of parallel transport respect to an affine connection enable us to give a more geometrical meaning of the notion of covariant derivative. As remarked in Section 5.1, if M is a differentiable manifold and we aim to compute the derivative of a vector field X with respect to another vector field Y in a point $p \in M$, we should compute something like the following limit

$$\lim_{h\to 0} \frac{X(p+hY) - X(p)}{h} \, .$$

Unfortunately, there are two problems involved in the formula above:

(1) What does it mean $p+hY$? In general, we have not an affine structure on M and we cannot move points thorough M under the action of vectors as in affine spaces.
(N.B. The reader should pay attention on the fact that affine connections and affine structures are different objects!).
(2) $X(p) \in T_pM$ but $X(p+hY) \in T_{p+hY}M$. If something like $p+hY$ makes sense, we expect that $p+hY \neq p$ because derivatives in p should investigate the behaviour of the function $q \mapsto X(q)$ in a "infinitesimal" neighborhood of p. So the difference $X(p+hY) - X(p)$ does not make sense because the vectors belong to different vector spaces!

As we have seen in Section 5.1, if M is an affine space $\mathbb{A}^n$ the candidate definition above can be improved into

$$(\nabla_Y X)_p := \lim_{h\to 0} \frac{\mathcal{A}[p+hY_p, p]X_{p+hY_p} - X_p}{h} \, .$$

(see Section **3.1** for notation) which turns out to coincide with the definition given via the affine connection naturally associated with the affine structure of $\mathbb{A}^n$. Is it possible to extend such a (equivalent) definition of derivative in the case of a manifold M equipped with an affine connection ∇? The answer is yes. Fix p and $Y(p)$ and consider the unique geodesic segment $[0,\epsilon) \ni h \mapsto \gamma(h)$ starting from p with initial vector $Y(p)$. Consider the point $\gamma(h)$. Formally we can view that point as "$p+hY$". Using that interpretation $X(p+hY)$ has to be interpreted as $X(\gamma(h))$ and the problem (1) becomes harmless. That is not the whole story because

$$X(\gamma(h)) - X(p)$$

does not make sense anyway since the vectors belong to different vector spaces.
As we are equipped with geodesics, we can move the vectors along them using the notion of

parallel transport. In practice, to improve our idea we may say that

$$X(p + hY)$$

must actually be understood as

$$\mathcal{P}_\gamma^{-1}[p, \gamma(h)]X(\gamma(h)) \,,$$

where

$$\mathcal{P}_\alpha[\alpha(u), \alpha(v)] : T_{\alpha(u)} \to T_{\alpha(v)}$$

is the vector-space isomorphism, introduced in the remark (3) after proposition 5.3, induced by the parallel transport along a differentiable curve $\alpha : [a, b] \to M$ for $u < v$ and $u, v \in [a, b]$. Within this interpretation

$$X(p + hY) - X(p) = \mathcal{P}_\gamma^{-1}[p, \gamma(h)]X(\gamma(h)) - X(p)$$

makes sense because both $\mathcal{P}_\gamma^{-1}[p, \gamma(h)]X(\gamma(h))$ and $X(p)$ belong to the same vector space $T_p(M)$. Notice that, in general

$$\mathcal{P}_\gamma^{-1}[p, \gamma(h)]X(\gamma(h)) \neq X(p) \,.$$

Summarizing, if M is equipped with an affine connection ∇, the derivative of X with respect to Y in p can be define as

$$D_Y^\nabla X|_p := \lim_{h\to 0} \frac{\mathcal{P}_\gamma^{-1}[p, \gamma(h)]X(\gamma(h)) - X(p)}{h} \,.$$

Let us show that the notion of derivative defined above is nothing but the covariant derivative $\nabla_Y X$ referred to the affine connection ∇. To this end, take a local coordinate system about p. From the equation of parallel transport, if $\mathcal{P}^{-1} := \mathcal{P}_\gamma[p, \gamma(h)]$ we have

$$X^i(\gamma(h)) - \left(\mathcal{P}^{-1}X(\gamma(h))\right)^i + h\, Y^j(\gamma(h))\, \Gamma^i_{jk}(\gamma(h)) \, \left(\mathcal{P}^{-1}X(\gamma(h))\right)^k = hA^i(h) \,,$$

where $A^i(h) \to 0$ as $h \to 0^+$. That identity can equivalently be written

$$\left(\mathcal{P}^{-1}X(\gamma(h))\right)^i = X^i(\gamma(h)) + h\, Y^j(p)\, \Gamma^i_{jk}(p) \, \left(\mathcal{P}^{-1}X(\gamma(h))\right)^k + hO^i(h) \,,$$

where we have dropped some infinitesimal functions which are now embodied in O^i with $O^i(h) \to 0$ as $h \to 0^+$. Using that expansion in the definition of $D_Y^\nabla X|_p$ we get:

$$\left(D_Y^\nabla X|_p\right)^i := \lim_{h\to 0} \frac{X^i(\gamma(h)) - X^i(p) + h\, Y^j(p)\, \Gamma^i_{jk}(p) \, \left(\mathcal{P}^{-1}X(\gamma(h))\right)^k - hO(h)}{h} \,.$$

Equivalently:

$$\left(D_Y^\nabla X|_p\right)^i := \lim_{h\to 0} \frac{X^i(\gamma_{p,Y}(h)) - X^i(p)}{h} + \lim_{h\to 0} Y^j(p)\Gamma^i_{jk}(p) \, \left(\mathcal{P}_\gamma^{-1}[p, \gamma(h)]X(\gamma(h))\right)^k \,,$$

and thus

$$\left(D^{\nabla}_{Y} X|_p\right)^i = Y^k(p)\frac{\partial X^i}{\partial x^k}|_p + Y^j(p)\Gamma^i_{jk}(p)X^k(p) = (\nabla_Y X)^i(p)\,.$$

Let us summarize our results into a Proposition.

Proposition 5.4. *Let M be a differentiable manifold equipped with an affine connection ∇. If X and Y are differentiable contravariant vector fields in M and $p \in M$,*

$$(\nabla_Y X)(p) = \lim_{h\to 0} \frac{\mathcal{P}^{-1}_{\gamma}[p, \gamma(h)]X(\gamma(h)) - X(p)}{h}\,,$$

where, $\gamma : [0, \epsilon) \to M$ is the unique geodesic segment referred to ∇ starting from p with initial tangent vector $Y(p)$ and

$$\mathcal{P}_\alpha[\alpha(u), \alpha(v)] : T_{\alpha(u)} \to T_{\alpha(v)}$$

is the vector-space isomorphism induced by the ∇ parallel transport along a differentiable curve $\alpha : [a, b] \to M$ for $u < v$ and $u, v \in [a, b]$. $\diamondsuit$

5.5 Geodesics: variational approach.

There is another approach to determine geodesics with respect to Levi-Civita's connection in a Riemannian manifold. Indeed, geodesics satisfy a *variational principle* because, roughly speaking, they stationarize the length functional of curves.

5.5.1 Basic notions of elementary variation calculus in $\mathbb{R}^n$.

Let us remind some basic notion of elementary variation calculus in $\mathbb{R}^n$. Fix an open nonempty set $\Omega \subset \mathbb{R}^n$, a closed interval $I = [a, b] \subset \mathbb{R}$ with $a < b$ and take a nonempty set

$$G \subset \{\gamma : I \to \Omega \mid \gamma \in C^{2k}(I)\}$$

for some fixed integer $0 < k < +\infty$ ($\gamma \in C^l([a, b])$ means that $\gamma \in C^l((a, b))$ and the limits toward either a^+ and b^- of derivatives of γ exist and are finite up to the order l).
A **variation** V of $\gamma \in G$, if exists, is a map $V : [0, 1] \times I \to U$ such that, if V_s denotes the function $t \mapsto V(s, t)$:
(1) $V \in C^{2k}([0, 1] \times I)$ (i.e., $V \in C^{2k}((0, 1) \times (a, b))$ and the limits awards the points of the boundary of $(0, 1) \times (a, b)$ all the derivatives of order up to $2k$ 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 γ 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 defined appropriately.

Lemma 5.1. *Let $\Omega \subset (\mathbb{R}^n)^k$ be an open nonempty set, $I = [a,b]$ with $a < b$. Fix $(p, P_1, \ldots, P_{k-1})$ and $(q, Q_1, \ldots, Q_{k-1})$ in Ω. Let D denote the space of elements of $\{\gamma : I \to \mathbb{R}^n \mid \gamma \in C^{2k}(I)\}$ such that:*
(1) $\left(\gamma(t), \frac{d^1\gamma}{dt^1}, \ldots, \frac{d^{k-1}\gamma}{dt^{k-1}}\right) \in \Omega$ *for all* $t \in [a,b]$,
(2) $\left(\gamma(a), \frac{d^1\gamma}{dt^1}|_a, \ldots, \frac{d^{k-1}\gamma}{dt^{k-1}}|_a\right) = (p, P_1, \ldots, P_{k-1})$ *and* $\left(\gamma(b), \frac{d^1\gamma}{dt^1}|_b, \ldots, \frac{d^{k-1}\gamma}{dt^{k-1}}|_b\right) = (q, Q_1, \ldots, Q_{k-1})$.
Within 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, $\eta : [a,b] \to \mathbb{R}^n$ 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 $r = 1, \ldots, k-1$. In particular, the result holds for every $c < C$, if $C > 0$ is sufficiently small. As a consequence endowing D with the norm topology induced by the norm

$$||\gamma||_k := \max\left\{\sup_I ||\gamma||, \sup_I \left|\left|\frac{d\gamma}{dt}\right|\right|, \ldots, \sup_I \left|\left|\frac{d^k\gamma}{dt^k}\right|\right|\right\} \, ,$$

it turns out that every punctured metric ball $B^*_r(\gamma) = \{\gamma' \in D \setminus \{\gamma\} \mid ||\gamma - \gamma'|| < r\}$ is not empty for every $r > 0$.
$\diamondsuit$

Proof. The only nontrivial fact we have to show is that there is some $C > 0$ such that

$$\left(\gamma(t) \pm sc\eta(t), \frac{d^1}{dt^1}(\gamma(t) \pm sc\eta(t)), \ldots, \frac{d^{k-1}}{dt^{k-1}}(\gamma(t) \pm sc\eta(t))\right) \in \Omega$$

for every $s \in [0,1]$ and every $t \in I$ provided $0 < c < C$. From now on for a generic curve $\tau : I \to \mathbb{R}^n$,

$$\tilde{\tau}(t) := \left(\tau(t), \frac{d^1\tau(t)}{dt^1}, \ldots, \frac{d^{k-1}\tau(t)}{dt^{k-1}}\right) \, .$$

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 Ω. Then, using the compactness of $\tilde{\gamma}([a,b])$ we can extract a finite subcovering. If Ω' 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 $|| \; ||$ 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 Ω. 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 η 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 $s \to 0^+$ 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 uniformly on $I = I' \cup I''$. Since the proof can be given for each component of the curve, we get that $||(\tilde{\gamma}(t) + sc\tilde{\eta}(t)) - \tilde{\gamma}(t)|| \to 0$ uniformly in $t \in I$ as $sc \to 0^+$. 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 $t \in I$.
Decreasing C if necessary, by a similar proof we get 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 proved as follows. Let $\gamma' \in D \setminus \{\gamma\}$. Such a curve exists with the form $\gamma'(t) = sc\eta(t)$, with $0 < c < C$ due to the previous part of the proposition. By construction $||\gamma - \gamma'|| < r_0$ 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 $r \in (0, r_0)$. □

Exercises 5.2. .

1. In the same hypotheses of lemma 5.1, 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.
(*Hint*. Note that the proof is obvious.)

We recall the reader that, if $G \subset \mathbb{R}^n$ and $F : G \to \mathbb{R}$ is any sufficiently regular function, $x_0 \in Int(G)$ is said to be a *stationary point* of F if $dF|_{x_0} = 0$. Such a condition can be re-written as

$$\frac{dF(x_0 + su)}{ds}|_{s=0} = 0 \,,$$

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 $x \in U \setminus \{x_0\}$), 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 $G \subset \{\gamma : I \to U \mid \gamma \in C^{2k}(I)\}$, i.e. a mapping $F : G \to \mathbb{R}$. We say that γ_0 **stationary point of** F, if for all variations of γ_0, V, the **variation of** F,

$$\delta_V F|_{\gamma_0} := \frac{dF[V_s]}{ds}|_{s=0}$$

exists and vanishes.

Remark 5.5. 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.

5.5.2 Euler-Poisson equations.

For suitable spaces G and functionals $F : G \to \mathbb{R}$, defining an appropriate topology on G itself, it is possible to show that if F attains a *local extremum* in $\gamma_0 \subset G$, then γ_0 must be a stationary point of F. We state a precise result after the specialization of the functional F.
From now on we work on domains G of the form D defined in lemma 5.1 and we focus attention on functionals with the form

$$F[\gamma] := \int_I \mathcal{F}\left(t, \gamma(t), \frac{d\gamma}{dt}, \cdots, \frac{d^k\gamma}{dt^k}\right) dt\,, \tag{5.5}$$

where k is the same used in the definition of D and $\mathcal{F} \in C^k(\Omega)$. Making use of lemma 5.1 we can prove a second important Lemma.

Lemma 5.2. *If $F : D \to \mathbb{R}$ is the functional in (5.5) with D defined in Lemma 3.1, $\delta_V F|_{\gamma_0}$ exists for every $\gamma_0 \in D$ and every variation of γ_0, V and*

$$\delta_V F|_{\gamma_0} = \sum_{i=1}^{n} \int_I \left.\frac{\partial V^i}{\partial s}\right|_{s=0} \left.\left[\frac{\partial \mathcal{F}}{\partial \gamma^i} + \sum_{r=1}^{k} (-1)^r \frac{d^r}{dt^r}\left(\frac{\partial \mathcal{F}}{\partial \frac{d^r\gamma^i}{dt^r}}\right)\right]\right|_{\gamma_0} dt$$

♢

Proof. From known properties of Lebesgue's measure based on Lebesgue's dominate convergence theorem (notice that $[0,1] \times I$ is compact an all the considered functions are continuous therein), we can pass the s-derivative operator under the sign of integration obtaining

$$\delta_V F|_{\gamma_0} = \sum_{i=1}^{n} \int_a^b \left(\left.\frac{\partial V^i}{\partial s}\right|_{s=0} \frac{\partial \mathcal{F}}{\partial \gamma^i} + \sum_{r=1}^{k} \left.\frac{\partial^{r+1} V^i}{\partial t^r \partial s}\right|_{s=0} \frac{\partial \mathcal{F}}{\partial \frac{d^r\gamma^i}{dt^r}} \right) dt\,.$$

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 vanishes because they contains factors

$$\frac{\partial^{l+1} V^i}{\partial^l t \partial s}|_{t=a \text{ or } b}$$

with $l = 0, 1, \ldots, k-1$. These factors must vanish because the conditions on curves in D:

$$\gamma(a) = p \quad \text{and} \quad \gamma(b) = q\,,$$

$$\frac{d^r t\gamma}{d^r t}|_a = P_r$$

and

$$\frac{d^r \gamma}{d^r t}|_b = Q_r$$

for $r = 1, \ldots, k-1$ imply that the variations of any $\gamma_0 \in D$ with their t-derivatives in a and b up to the order $k-1$ have to vanish in a and b whatever $s \in [0,1]$. Then the formula in thesis follows trivially. □

A third and last lemma is in order.

Lemma 5.3. *Suppose that $f : [a,b] \to \mathbb{R}^n$, with components $f^i : [a,b] \to \mathbb{R}$, $i = 1, \ldots, n$, is continuous. If*

$$\int_a^b \sum_{i=1}^n h^i(x) f^i(x) dx = 0$$

for every C^∞ function $h : \mathbb{R} \to \mathbb{R}^n$ whose components h^i have supports contained in in (a,b), it has to hold $f(x) = 0$ 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 $j \in \{1, \ldots, n\}$ and an open neighborhood of x_0, $U \subset (a,b)$, where $f^j(x) > 0$. Using the remark (3) after definition 2.4, take a function $g \in C^\infty(\mathbb{R})$ with $\mathrm{supp}\, g \subset U$, $g(x) \geq 0$ 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', such that $\overline{U'} \subset U$ and $g(x) f^j(x) > 0$ on $\overline{U'}$. As a consequence $\min_{\overline{U'}} g \cdot f^j = m > 0$.
Below χ_A denotes the characteristic function of a set A and $h : (a,b) \to \mathbb{R}^n$ is defined as $h^j = g$ and $h^i = 0$ if $i \neq j$. Finally 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 $\int_{\overline{U'}} dx \geq \int_{U'} dx > 0$ because nonempty open sets have strictly positive Lebesgue measure.
The found result is not possible. So $f(x) = 0$ in (a,b) and, by continuity, $f(a) = f(b) = 0$. □

We conclude the general theory with two theorems.

Theorem 5.2. *Let* $\Omega \subset (\mathbb{R}^n)^k$ *be an open nonempty set,* $I = [a, b]$ *with* $a < b$. *Fix* $(p, P_1, \ldots, P_{k-1})$ *and* $(q, Q_1, \ldots, Q_{k-1})$ *in* Ω. *Let* D *denote the space of elements of* $\{\gamma : I \to \mathbb{R}^n \mid \gamma \in C^{2k}(I)\}$ *such that:*
(1) $\left(\gamma(t), \frac{d^1\gamma}{d^1t}, \ldots, \frac{d^{k-1}\gamma}{d^{k-1}t}\right) \in \Omega$ *for all* $t \in [a, b]$,
(2) $\left(\gamma(a), \frac{d^1\gamma}{d^1t}|_a, \ldots, \frac{d^{k-1}\gamma}{d^{k-1}t}|_a\right) = (p, P_1, \ldots, P_{k-1})$ *and* $\left(\gamma(b), \frac{d^1\gamma}{d^1t}|_b, \ldots, \frac{d^{k-1}\gamma}{d^{k-1}t}|_b\right) = (q, Q_1, \ldots, Q_{k-1})$.
Finally define

$$F[\gamma] := \int_I \mathcal{F}\left(t, \gamma(t), \frac{d\gamma}{dt}, \cdots, \frac{d^k\gamma}{dt^k}\right) dt$$

where $\mathcal{F} \in C^k(\Omega)$.
Under these hypotheses $\gamma \in D$ *is a stationary point of* F *if and only if it satisfies the* **Euler-Poisson equations** *for* $i = 1, \ldots, n$*:*

$$\frac{\partial \mathcal{F}}{\partial \gamma^i} + \sum_{r=1}^{k} (-1)^r \frac{d^r}{dt^r}\left(\frac{\partial \mathcal{F}}{\partial \frac{d^r\gamma^i}{dt^r}}\right) = 0\,.$$

◊

Proof. It is clear that if $\gamma \in D$ fulfills Euler-Poisson equations, γ is an extremal point of F because of lemma 5.2.
By lemma 5.2 once again, if $\gamma \in D$ is a stationary point, it must satisfy

$$\sum_{i=1}^{n} \int_I \left.\frac{\partial V^i}{\partial s}\right|_{s=0} \left.\left[\frac{\partial \mathcal{F}}{\partial \gamma^i} + \sum_{r=1}^{k} (-1)^r \frac{d^r}{dt^r}\left(\frac{\partial \mathcal{F}}{\partial \frac{d^r\gamma^i}{dt^r}}\right)\right]\right|_{\gamma_0} dt = 0$$

for all variations V. We want to prove that these identities valid for every variation V of γ entail that γ satisfies E-P equations. The proof is based on lemma 5.3 with

$$f^i = \left.\left[\frac{\partial \mathcal{F}}{\partial \gamma^i} + \sum_{r=1}^{k} (-1)^r \frac{d^r}{dt^r}\left(\frac{\partial \mathcal{F}}{\partial \frac{d^r\gamma^i}{dt^r}}\right)\right]\right|_{\gamma_0}$$

and

$$h^i = \left.\frac{\partial V^i}{\partial s}\right|_{s=0}\,.$$

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 5.1 if one uses 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

$$\sum_{i=1}^{n} \int_I \left.\frac{\partial V^i}{\partial s}\right|_{s=0} \left.\left[\frac{\partial \mathcal{F}}{\partial \gamma^i} + \sum_{r=1}^{k} (-1)^r \frac{d^r}{dt^r}\left(\frac{\partial \mathcal{F}}{\partial \frac{d^r\gamma^i}{dt^r}}\right)\right]\right|_{\gamma_0} dt = 0$$

becomes

$$c\int_a^b \sum_{i=1}^n h^i(x)f^i(x)dx = 0$$

for every choice of functions $h_i \in C^\infty((a,b))$, $i = 1, \ldots, n$ and for a corresponding constant $c > 0$ which does not affect the use of the lemma 5.1. Then, lemma 5.1 implies the thesis. □

Remark 5.6. Notice that, for $k = 1$, Euler-Poisson equations reduce to the well-known **Euler-Lagrange** equations $\mathcal{F}$ being the Lagrangian of a mechanical system.

Theorem 5.3. *With the same hypotheses of theorem 5.2, endow D with the norm topology induced by the norm*

$$||\gamma||_k := \max\left\{\sup_I ||\gamma||\,, \sup_I \left|\left|\frac{d\gamma}{dt}\right|\right|, \ldots, \sup_I \left|\left|\frac{d^k\gamma}{dt^k}\right|\right|\right\} .$$

If the functional $F : D \to \mathbb{R}$ attains an extremal value at $\gamma_0 \in D$, γ_0 turns out to be a stationary point of F and it satisfies Euler-Poisson's equations. ♢

Proof. Suppose that γ_0 defines a local maximum of F (the other case is similar). In that case there is an open norm ball $B \subset D$ centered in γ_0, such that, if $\gamma \in B \setminus \{\gamma_0\}$, $F(\gamma) < F(\gamma_0)$. In particular if $V_\pm = \gamma \pm sc\eta$,

$$\frac{F(\gamma_0 \pm cs\eta) - F(\gamma_0)}{s} < 0$$

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 $s \to 0^+$ exists by lemma 5.2. Hence

$$\delta_{V_\pm} F|_{\gamma_0} \leq 0 .$$

Making explicit the left-hand side by lemma 5.2 one finds

$$\pm\sum_{i=1}^n \int_I \eta^i \left[\frac{\partial\mathcal{F}}{\partial\gamma^i} + \sum_{r=1}^k (-1)^r \frac{d^r}{dt^r}\left(\frac{\partial\mathcal{F}}{\partial\frac{d^r\gamma^i}{dt^r}}\right)\right]\Bigg|_{\gamma_0} dt \leq 0\,,$$

and thus

$$\sum_{i=1}^n \int_I \eta^i \left[\frac{\partial\mathcal{F}}{\partial\gamma^i} + \sum_{r=1}^k (-1)^r \frac{d^r}{dt^r}\left(\frac{\partial\mathcal{F}}{\partial\frac{d^r\gamma^i}{dt^r}}\right)\right]\Bigg|_{\gamma_0} dt = 0\,.$$

Using lemma 5.3 as in proof of theorem 5.2 we conclude that γ_0 satisfies Euler-Poisson's equations. As a consequence of theorem 5.2, γ_0 is a stationary point of F. □

5.5.3 Geodesics from variational point of view.

We can pass to consider geodesics in Riemannian and Lorentzian manifolds. Let us state and prove a first theorem which is valid for properly Riemannian metrics and involves the length of a differentiable curve (see the comment (2) after definition 4.1).

Theorem 5.4. *Let M be a Riemannian manifold with metric locally denoted by g_{ij}. Take $p, q \in M$ such that there is a common local chart (U, ϕ), $\phi(r) = (x^1(r), \ldots, x^n(r))$, with $p, q \in U$. Fix $[a, b] \subset \mathbb{R}$, $a < b$ and consider the* **curve-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 (differentiable) curves $\gamma : [a, b] \to U$ (U being identified to the open set $\phi(U) \subset \mathbb{R}^n$) with $\gamma(a) = p$, $\gamma(b) = q$ and everywhere nonvanishing tangent vector $\dot{\gamma}$.
(a) *If $\gamma_0 \in S$ is a stationary point of L, there is a differentiable bijection with inverse differentiable, $u : [0, L[\gamma_0]] \to [a, b]$, 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), γ_0 is a stationary point of L.*$\diamondsuit$

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 5.2: 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^∞ (indeed the function $x \mapsto \sqrt{x}$ is C^∞ in the domain $\mathbb{R} \setminus \{0\}$).
(a) By theorem 5.2, if $\gamma_0 \in S$ is a stationary point of F, γ_0 satisfies in $[a, b]$:

$$\frac{d}{dt}\left[\frac{g_{ki}\frac{dx^i}{dt}}{\sqrt{g_{rs}\frac{dx^r}{dt}\frac{dx^s}{dt}}}\right] - \frac{\frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\frac{dx^i}{dt}\frac{dx^j}{dt}}{\sqrt{g_{rs}\frac{dx^r}{dt}\frac{dx^s}{dt}}} = 0\,, \tag{5.6}$$

where $x^i(t) := x^i(\gamma_0(t))$ and the metric g_{lm} is evaluated on $\gamma_0(t)$.
Since $\dot{\gamma}_0(t) \neq 0$ 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

$$s(t) := \int_a^s \sqrt{g_{rs}(\gamma_0(t))\frac{dx^r}{dt}\frac{dx^s}{dt}}dt$$

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 $u : [0, L[\gamma_0]] \to [a, b]$ the inverse function of s. By (5.6), 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

$$\frac{d^2x^i}{ds^2}g_{ki} + \frac{\partial g_{ki}}{\partial x^j}\frac{dx^i}{ds}\frac{dx^j}{ds} - \frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\frac{dx^i}{dt}\frac{dx^j}{dt} = 0\,.$$

These equations can be re-written as

$$\frac{d^2x^i}{ds^2}g_{ki} + \frac{1}{2}\left[\frac{\partial g_{ki}}{\partial x^j}\frac{dx^i}{ds}\frac{dx^j}{ds} + \frac{\partial g_{kj}}{\partial x^i}\frac{dx^j}{ds}\frac{dx^i}{ds} - \frac{\partial g_{ij}}{\partial x^k}\frac{dx^i}{ds}\frac{dx^j}{ds}\right] = 0\,.$$

Contracting with g^{rk} these equations become

$$\frac{d^2x^r}{ds^2} + \frac{1}{2}g^{rk}\left[\frac{\partial g_{ki}}{\partial x^j} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}\right]\frac{dx^i}{ds}\frac{dx^j}{ds} = 0\,,$$

which can be re-written as the geodesic equations with respect to Levi-Civita's connection:

$$\frac{d^2x^r}{ds^2} + \{{}_i{}^r{}_j\}\frac{dx^i}{ds}\frac{dx^j}{ds} = 0\,.$$

(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 γ_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

$$\sqrt{g_{rl}(\gamma_0(t(s)))\frac{dx^r}{ds}\frac{dx^l}{ds}} = 1\,.$$

Then suppose that $t \mapsto \gamma_0(t)$ is a geodesic. Thus $t \in [a, b]$ is an affine parameter. By the remark (4) after definition 5.2, there are $c, d \in \mathbb{R}$ with $c > 0$ such that $t = cs + d$. As a consequence

$$\sqrt{g_{rl}(\gamma_0(t))\frac{dx^r}{dt}\frac{dx^l}{dt}} = \frac{1}{c}\sqrt{g_{rl}(\gamma_0(t(s)))\frac{dx^r}{ds}\frac{dx^l}{ds}} \tag{5.7}$$

and thus

$$\sqrt{g_{rl}(\gamma_0(t))\frac{dx^r}{dt}\frac{dx^l}{dt}} = \frac{1}{c}\,. \tag{5.8}$$

Following the proof of (a) by a reversed order one proves that

$$\frac{d^2x^r}{dt^2} + \{{}_i{}^r{}_j\}\frac{dx^i}{dt}\frac{dx^j}{dt} = 0\,.$$

implies

$$\frac{d}{dt}\left[g_{ki}\frac{dx^i}{dt}\right] - \frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\frac{dx^i}{dt}\frac{dx^j}{dt} = 0\,,$$

or, since $c > 0$,

$$c\frac{d}{dt}\left[g_{ki}\frac{dx^i}{dt}\right] - c\frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\frac{dx^i}{dt}\frac{dx^j}{dt} = 0\,,$$

Using the fact that c is constant and (5.8), 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 5.2. □

We can generalize the theorem to the case of a Lorentzian manifold.

Theorem 5.5. *Let M be a Lorentzian manifold with metric locally denoted by g_{ij}. Take $p, q \in M$ such that there is a common local chart (U, ϕ), $\phi(r) = (x^1(r), \ldots, x^n(r))$, with $p, q \in U$. Fix $[a, b] \subset \mathbb{R}$, $a < b$ and consider the* **timelike-curve-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 (differentiable) curves $\gamma : [a, b] \to U$ (U being identified to the open set $\phi(U) \subset \mathbb{R}^n$) with $\gamma(a) = p$, $\gamma(b) = q$ and γ is **timelike***, i.e. $(\dot\gamma|\dot\gamma) < 0$ 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, $u : [0, L_T[\gamma_0]] \to [a, b]$, such that $\gamma \circ u$ is a timelike geodesic with respect to the Levi-Civita connection connecting p to q.*
(b) *If $\gamma_0 \in S_T$ is a timelike geodesic (connecting p to q), γ_0 is a stationary point of L_T.*♢

Proof. The proof is the same of theorem 5.4 with the specification that S_T, if nonempty, is a domain of the form D used in theorem 5.2. In particular the set $\Omega \subset \mathbb{R}^{2n}$ used in the definition of D is now the open set:

$$\{(x^1,\ldots,x^n,v^1,\ldots,v^n)\in\mathbb{R}^{2n}\;|\;(x^1,\ldots,x^n)\in\phi(U)\;,\;(g_{\phi^{-1}(x^1,\ldots,x^n)})_{ij}v^iv^j<0\}$$

where g_{ij} represent the metric in the coordinates associated with ϕ. □

Theorem 5.6. *Let M be a Lorentzian manifold with metric locally denoted by g_{ij}. Take $p, q \in M$ such that there is a common local chart (U, ϕ), $\phi(r) = (x^1(r), \ldots, x^n(r))$, with $p, q \in U$. Fix $[a, b] \subset \mathbb{R}$, $a < b$ and consider the* **spacelike-curve-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 (differentiable) curves $\gamma : [a, b] \to U$ (U being identified to the open set $\phi(U) \subset \mathbb{R}^n$) with $\gamma(a) = p$, $\gamma(b) = q$ and γ is **spacelike***, i.e. $(\dot\gamma|\dot\gamma) > 0$ 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, $u : [0, L_S[\gamma_0]] \to [a, b]$, 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), γ_0 is a stationary point of L_S.*$\diamondsuit$

Proof. Once again the proof is the same of theorem 5.4 with the specification that S_S, if nonempty, is a domain of the form D used in theorem 5.2. In particular the set $\Omega \subset \mathbb{R}^{2n}$ used in the definition of D is now the open set:

$$\{(x^1, \ldots, x^n, v^1, \ldots, v^n) \in \mathbb{R}^{2n} \mid (x^1, \ldots, x^n) \in \phi(U) \ , \ (g_{\phi^{-1}(x^1,\ldots,x^n)})_{ij} v^i v^j > 0\}$$

where g_{ij} represent the metric in the coordinates associated with ϕ. $\Box$

Exercises 5.3. .

1. Show that the sets Ω used in the proof of theorem 5.5. and theorem 5.6 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 5.7.
(1) Working in TM, the three theorems proved above can be generalized by dropping the hypotheses of the existence of a common local chart (U, ϕ) containing the differentiable curves.
(2) It is worth stressing that there is no guarantee for having a geodesic joining any pair of points in a (pseudo) Riemannian manifold. For instance consider the Euclidean space $\mathbb{E}^2$ (see example 4.1.1), and take $p, q \in \mathbb{E}^2$ with $p \neq q$. As everybody knows there is exactly a geodesic segment γ 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) It is worth stressing that 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^4 with the metric induced by $\mathbb{E}^3$, there are infinite many geodesic segments connecting the north pole with the south pole.
(4) It is possible to show that, in Riemannian manifolds, geodesics locally minimize the curve-length functional ("locally" means here that the endpoints are sufficiently close to each other). Conversely, in Lorentzian manifolds, timelike geodesics (see example 5.1) locally maximize the curve-length functional.

Chapter 6

Some advanced geometric tools.

6.1 Exponential map and its applications in General Relativity.

We go to introduce an important tool, a special coordinate system about any point of a manifold endowed with a connection, called *exponential map*. It allows one to identify locally the manifold with the tangent space at p. Moreover it provides a coordinate system such that the connection coefficients vanishes exactly at p whenever the connection is torsion-free. This fact is of central relevance in general relativity since it entails a mathematically rigorous statement of *Einstein's equivalence principle*.

6.1.1 The exponential map and normal coordinates about a point.

Consider a differentiable manifold M equipped with a smooth affine connection ∇ (however C^2 would be enough for what follows). If (U, ϕ) is a local chart on M with coordinates $x^1, \ldots, x^n$, consider the associated natural coordinate patch on TM with coordinates $(x^1, \ldots, x^n, \dot{x}^1, \ldots, \dot{x}^n)$. In other words, a tangent vector of T_pM with $U \ni p$ can be written as:

$$\dot{x}^i \frac{\partial}{\partial x^i}|_p \,.$$

The Cauchy problem associated with the geodesic equation, referred to ∇ and *viewed as an equation in TM*, reads in coordinates:

$$\left\{ \begin{array}{cl} \dfrac{d\dot{x}^i}{dt} = & -\Gamma(x(t))^i_{jk}\dfrac{d\dot{x}^j}{dt}\dfrac{d\dot{x}^k}{dt} \,, \\ \dfrac{dx^i}{dt} = & \dot{x}^i(t) \,, \\ x^i(0) = x^i_p & \dot{x}^i(0) = \dot{x}^i_p \,, \end{array} \right. \qquad i = 1, \ldots n \,. \tag{6.1}$$

In particular $(x^1_p, \ldots, x^n_p)$ are the coordinates of $p \in U$, the starting point of the geodesic with initial vector $\dot{x}_p = \dot{x}^i_p \frac{\partial}{\partial x^i}|_p$. Let us indicate the unique *maximal* (in U) solution of (6.1) by

$\gamma = \gamma(p, v_p, t)$, where $p \in U$, $v_p \in T_pM$ and t belongs to some open interval $I \ni 0$ generally depending on p and v_p. In the following, we represent (p, v_p) in terms of the corresponding coordinates $(x^1_p, \ldots, x^n_p, \dot{x}^1, \ldots, \dot{x}^n) \in \phi(U) \times \mathbb{R}^n \subset \mathbb{R}^n \times \mathbb{R}^n$.
As is well known from the general theory of differential equations, if one considers maximal solutions varying the initial conditions $p \in U$, $v_p \in T_pM$, and taking t in the corresponding interval, the obtained domain $\Omega \subset \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ of γ is *open* in $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$. As a consequence, if we fix $r \in U$, there is a set of the form $V_r \times B_\delta(0) \times (-\epsilon, \epsilon)$ – with $\epsilon > 0$, $B_\delta(0)$ being the open ball in $\mathbb{R}^n$ with radius $\delta > 0$ centered in 0 and $V_r \subset U$ being open neighborhood of r – such that

$$V_r \times B_\delta(0) \times (-\epsilon, \epsilon) \ni (p, v_p, t) \mapsto \gamma(p, v_p, t)$$

is well defined. Now (6.1) and the uniqueness theorem entail that, for every $\lambda > 0$, if the geodesic $t \mapsto \gamma(p, v_p, t)$ is defined for $t \in (-\epsilon, \epsilon)$, then the geodesic $t \mapsto \gamma(p, \lambda v_p, t)$ is defined for $t \in (-\epsilon/\lambda, \epsilon/\lambda)$, and

$$\gamma(p, \lambda v_p, t) = \gamma(p, v_p, \lambda t) \,. \tag{6.2}$$

(Indeed the right-hand side, considered as a function of $t \in (-\epsilon/\lambda, \epsilon/\lambda)$, satisfies the geodesic equation with tangent initial vector at p λv_p.) Therefore, we can fix $\lambda > 0$ sufficiently small in order to have $\epsilon' := \epsilon/\lambda > 1$. We conclude that, if $(p, u_p) \in V_r \times B_{\delta'}(0)$ with $\delta' = \lambda\delta$, the map:

$$V_r \times B_{\delta'}(0) \ni (p, u_p) \mapsto \gamma(p, u_p, t)$$

is well-defined for $t \in (-\epsilon', \epsilon') \supset (-1, 1)$. Therefore, the function

$$V_r \times B_{\delta'}(0) \ni (p, u_p) \mapsto \exp_p(u_p) := \gamma(p, u_p, 1) \tag{6.3}$$

is well defined. Notices that, standard theorems on regular dependence of solutions of differential equations from initial conditions entail that $(p, u_p) \mapsto \exp_p(u_p)$ is smooth as well.
For each $p \in V_r$, $B_{\delta'}(0)$ can be identified with an open *starshaped* neighborhood of the origin[1] of T_pM.

Definition 6.1. Let M be a differentiable manifold equipped with a smooth affine connection.
(a) The map (6.3) defined on a sufficiently small open set $E \subset TM$, (represented in coordinates by a product $V_r \times B_{\delta'}(0)$) is called **exponential map** on E.
(b) If $p \in M$, the restriction to the exponential map to $\{p\} \times U_0$, U_0 being a suitable starshaped neighborhood of 0 of the origin of T_pM, is called **exponential map** centered in p. $\diamondsuit$

Remark 6.1. By construction (use (6.2) with $t = 1$), the map

$$[0, 1] \mapsto \exp_p(\lambda u_p) \in M, \tag{6.4}$$

---

[1] An open *starshaped* neighborhood of the origin 0 of a topological vector space V is an open neighborhood U of 0 such that, if $v \in U$, the segment joining 0 to v belongs to U too.

is the unique geodesic segment starting from p, with initial tangent vector u_p and with affine parameter $\lambda \in [0,1]$.

We can now state the main result about the exponential map.

Theorem 6.1. *Let M be a differentiable manifold equipped with a smooth affine connection ∇ and consider the exponential map centered in a point $p \in M$. The following facts hold.*
(a) *In a sufficiently small starshaped neighborhood U_0 of the origin of T_pM, the exponential map defines a diffeomorphism onto the open set $\exp_p(U_0)$.*
(b) *If ∇ is torsion free and $\{e_{pi}\}_{i=1,\ldots,n} \subset T_pM$ is a basis, consider the local coordinate system defined on $\exp_p(U_0)$ which associates $q \in \exp_p(U_0)$ with the components of $\exp_p^{-1}(q)$ with respect to that basis, i.e.:*

$$q \mapsto \left(\langle \exp_p^{-1}(q), e_p^{*1}\rangle, \cdots, \langle \exp_p^{-1}(q), e_p^{*1}\rangle\right) .$$

In that coordinate patch the connection coefficients of ∇ vanishes at p.
(c) *If ∇ is Levi-Civita connection associated with a smooth metric g on M, and $\{e_{pi}\}_{i=1,\ldots,n} \subset T_pM$ is a basis, consider the local coordinate system defined on $\exp_p(U_0)$:*

$$q \mapsto \left(\langle \exp_p^{-1}(q), e_p^{*1}\rangle, \cdots, \langle \exp_p^{-1}(q), e_p^{*1}\rangle\right) .$$

In that coordinate patch the derivative of the components of the metric with respect to the coordinates, vanish at p. $\diamondsuit$

Proof. By (3) in theorem 3.1, to prove (a) it is sufficient showing that $d\exp_p|_p$ is non singular. Fix a local coordinate system about p with coordinates $x^1, \ldots, x^n$ and denote with $(\exp_p(v))^i$ the component x^i of any $\exp_p(v)$. Using the remark 6.1 one gets, *where we do not use the convention of summation over repeated indices*:

$$\frac{\partial}{\partial v^j}\bigg|_{v=0} \left(\exp\left(\sum_{i=1}^n v^i e_{pi}\right)\right)^k = \frac{\partial}{\partial v^j}\bigg|_{v=0} \left(\exp\left(v^j e_{pj}\right)\right)^k = \frac{\partial}{\partial \lambda}\bigg|_{\lambda=0} \gamma^k(p, e_j, \lambda) = e^k_{pj} ,$$

where e^k_{pj} is the k-th component of e_{pj} with respect to the basis $\{\partial/\partial x^i|_p\}_{i=1,\ldots,n}$. The matrix of these $\mathbb{R}^n$ vectors is nonsingular since $\{e_{pj}\}_{j=1,\ldots,n}$ is a basis too. The proof of (a) is concluded. Due to remark 6.1, in coordinates $y^1, \ldots, y^n$ defined on $\exp_p(U_0)$ which associates $q \in \exp_p(U_0)$ with the components of $\exp_p^{-1}(q)$ with respect to the basis $\{e_{pi}\}_{i=1,\ldots,n}$, a geodesic starting from p with initial vector $v^i e_{pi}$ has equation: $y^i(\lambda) = \lambda v^i$. Therefore $\frac{d^2 y^i}{d\lambda^2} = 0$. On the other hand it must also hold:

$$\frac{d^2 y^i}{d\lambda^2} + \Gamma^i_{jk}(y(\lambda))\frac{dy^j}{d\lambda}\frac{dy^k}{d\lambda} = 0 .$$

As a consequence, for every $\lambda[0,1]$:

$$\Gamma^i_{jk}(y(\lambda))\frac{dy^j}{d\lambda}\frac{dy^k}{d\lambda} = 0 .$$

In particular, if $\lambda = 0$:

$$\Gamma^i_{jk}(p)v^j v^k = 0\,, \quad \text{for all } v^j \in \mathbb{R}\,.$$

If the connection is torsion-free, i.e. $\Gamma^i_{jk} = \Gamma^i_{kj}$, the identities above entail, using $v^i = u^i + z^i$ for all $u^i, z^i \in \mathbb{R}$, $\Gamma^i_{jk}(p)u^j z^k = 0$ for all $u^i, z^i \in \mathbb{R}$, and thus:

$$\Gamma^i_{jk}(p) = 0\,.$$

The proof of (c) is an immediate consequence (b) and of the identity $\nabla g = 0$, which, in coordinates reads:

$$\frac{\partial g_{ki}}{\partial y^j} = \Gamma^s_{jk} g_{si} + \Gamma^s_{ji} g_{ks}\,.$$

□

Definition 6.2. Let M be a differentiable manifold equipped with a smooth affine connection ∇ and $p \in M$. Consider a starshaped neighborhood of 0, $U_0 \subset T_pM$ where $\exp_p$ defines a local diffeomorphism. If $\{e_{pi}\}_{i=1,\ldots,n} \subset T_pM$ is a basis, the local coordinate system defined on $\exp_p(U_0)$ which associates $q \in \exp_p(U_0)$ with the components of $\exp_p^{-1}(q)$ with respect to that basis, i.e.:

$$q \mapsto \left(\langle \exp_p^{-1}(q), e_p^{*1}\rangle, \cdots, \langle \exp_p^{-1}(q), e_p^{*1}\rangle\right)$$

is called **(Riemannian) normal coordinate system** centered in p. ♢

6.1.2 Riemannian normal coordinates adapted to a differentiable curve.

Let us pass to consider a more complicated construction concerning normal coordinates about a given differentiable curve. If M is a differentiable manifold endowed with a smooth Riemannian or Lorentzian metric g, let $\alpha : (a,b) \to M$ be a differentiable curve with $\dot\alpha(t) \neq 0$ for every $t \in (a,b)$. In case g is Lorentzian, we assume that α is timelike, i.e. $g(\dot\alpha(t), \dot\alpha(t)) < 0$ for every $t \in (a,b)$.

Take $t_0 \in (a,b)$, and consider a basis for the subspace $N_{\alpha(t_0)}\alpha$ of $T_{\alpha(t_0)}$ normal to $\dot\alpha(t_0)$, $\{e_{\alpha(t_0)i}\}_{i=2,\ldots,n}$. Then transport that basis along α using the parallel transport procedure. Due to proposition 5.3, the transported vectors define a basis $\{e_{\alpha(t)i}\}_{i=2,\ldots,n}$ for the subspace $N_{\alpha(t)}\alpha$ of $T_{\alpha(t)}M$ normal to $\dot\alpha(t)$. Finally consider the map:

$$\mathbb{R}^n \ni (t, v^2, \ldots, v^n) \mapsto \exp_{\alpha(t)}\left(\sum_{i=2}^{n} v^i e_{\alpha(t)i}\right) \tag{6.5}$$

The geometric meaning of (6.5) should be clear: the map associates $(t, v^2, \ldots, v^n)$ with the point in M reached by the geodesic starting from $\alpha(t)$ with initial vector $\sum_{i=2}^n v^i e_{\alpha(t)i}$, when the affine parameter of the geodesic takes the value 1.

Let us discuss on the definition domain of the map (6.5). Fix a local system of coordinates $x^1, \ldots, x^n$ in an open set $U \ni \alpha(t_0)$. The function exp on TM is defined in sufficiently small open neighborhoods $E \subset TM$ such that, in coordinates $x^1, \ldots, x^n, \dot{x}^1, \ldots, \dot{x}^n$, has the form $V \times B_\delta(0) \subset \mathbb{R}^n \times \mathbb{R}^n$, where V corresponds to an open neighborhood of $\alpha(t_0)$ and $B_\delta(0)$ is an open ball with radius $\delta > 0$ centered in the origin of $\mathbb{R}^n$. In the considered coordinates (6.5) reads

$$(t, v^2, \ldots, v^n) \mapsto \exp_{(x^1(t),\ldots,x^n(t))}\left(\sum_{k=1}^{n} \dot{x}^k(t, v^2, \ldots, v^n) \left.\frac{\partial}{\partial x^k}\right|_{(x^1(t),\ldots,x^n(t))}\right) , \tag{6.6}$$

where α is parametrized as $x^k = x^k(t)$ and

$$\dot{x}^k(t, v^2, \ldots, v^n) := \sum_{i=2}^{n} v^i \left\langle e_{(x^1(t),\ldots,x^n(t))i}, dx^k|_{(x^1(t),\ldots,x^n(t))}\right\rangle .$$

Since all involved functions are continuous, it is simply proved that the right hand side of (6.6) is well defined for $(t, v^2, \ldots, v^n) \in (t_0 - \epsilon, t_0 + \epsilon) \times D$, $D \subset \mathbb{R}^{n-1}$ being some open ball centered in the origin of $\mathbb{R}^{n-1}$.

Theorem 6.2. *Let M be a differentiable manifold endowed with a smooth Riemannian or Lorentzian metric g, and $\alpha : (a, b) \to M$ a differentiable curve with $\dot{\alpha}(t) \neq 0$ for every $t \in (a, b)$ and of time-like type if g is Lorentzian. Take $t_0 \in (a, b)$, consider a basis for the subspace $N_{\alpha(t_0)}\alpha$ of $T_{\alpha(t_0)}$ normal to $\dot{\alpha}(t_0)$, $\{e_{\alpha(t_0)i}\}_{i=2,\ldots,n}$ and transport that basis along α to $\{e_{\alpha(t)i}\}_{i=2,\ldots,n}$ for every $t \in (a, b)$. Finally, consider the map*

$$(t_0 - \epsilon, t_0 + \epsilon) \times D \ni (t, v^2, \ldots, v^n) \mapsto \exp_{\alpha(t)}\left(\sum_{i=2}^{n} v^i e_{\alpha(t)i}\right) \tag{6.7}$$

for some $\epsilon > 0$ and $D \subset \mathbb{R}^{n-1}$ being some open ball centered in the origin of $\mathbb{R}^{n-1}$. The following facts hold.
(a) *Shrinking if necessary the set of definition, the map (6.7) define a local diffeomorphism.*
(b) *In the local coordinate patch about α associated with coordinates*

$$(y^1, y^2, \ldots, y^n) := (t, v^2, \ldots, v^n) ,$$

the connection coefficients of Levi-Civita connection satisfy, for $i = 1, \ldots, n$,

$$\Gamma^i_{jk}(\alpha(t)) = 0 , \quad \textit{if } t \in (t_0 - \epsilon, t_0 + \epsilon) \textit{ and } (j, k) \neq (0, 0). \tag{6.8}$$

(c) *If α is a geodesic, in the local coordinate patch about α associated with coordinates*

$$(y^1, y^2, \ldots, y^n) := (t, v^2, \ldots, v^n) ,$$

the connection coefficients of Levi-Civita connection satisfy,

$$\Gamma^i_{jk}(\alpha(t)) = 0 , \quad \textit{if } t \in (t_0 - \epsilon, t_0 + \epsilon) \textit{ and } i, j, k = 1, \ldots, n. \tag{6.9}$$

Moreover, the derivatives of the metric coefficients with respect to the said coordinates, vanishes along α for $t \in (t_0 - \epsilon, t_0 + \epsilon)$. $\diamondsuit$

Proof. Consider normal coordinates $x^1, \ldots, x^n$ centered in $p = \alpha(t_0)$ associated to the basis $\{\dot\alpha(t_0)\} \cup \{e_{\alpha(t_0)i}\}_{i=2,\ldots,n}$. In that case $\partial/\partial x^1|_p = \dot\alpha(t_0)$ and $\partial/\partial x^i|_p = e_{\alpha(t_0)i}$ for $i = 2, \ldots, n$.
A straightforward computation proves that:

$$\frac{\partial}{\partial y^1}\Big|_{\alpha(t_0)} \left(\exp_{\alpha(t)}\left(\sum_{i=2}^{n} y^i e_{\alpha(t)i}\right)\right)^k = \frac{\partial}{\partial t}\Big|_{t=t_0} \left(\exp_{\alpha(t)}(0)\right)^k + \sum_{i=2}^{n} y^i g^k(t, y^2, \ldots, y^n)\Bigg|_{(y^2,\ldots,y^n)=(0,\ldots,0)},$$

where the functions g^k are smooth. Since $\exp_{\alpha(t)}(0) = \alpha(t)$, we conclude that:

$$\frac{\partial}{\partial y^1}\Big|_{\alpha(t_0)} \left(\exp_{\alpha(t)}\left(\sum_{i=2}^{n} y^i e_{\alpha(t)i}\right)\right)^k = \frac{\partial}{\partial t}\Big|_{t=t_0} \alpha^k(t) + 0 = \dot\alpha^k(t_0) = \left(\frac{\partial}{\partial y^1}\Big|_{\alpha(t_0)}\right)^k = \delta^k_1\,,$$

For $j = 2, \ldots, n$, using remark 6.1 one achieves (below there is no summation over repeated indices):

$$\frac{\partial}{\partial y^j}\Big|_{\alpha(t_0)} \left(\exp_{\alpha(t)}\left(\sum_{i=2}^{n} y^i e_{\alpha(t)i}\right)\right)^k = \frac{\partial}{\partial y^j}\Big|_{(y^2,\ldots,y^n)=(0,\ldots,0)} \left(\exp_{\alpha(t_0)}\left(y^j e_{\alpha(t_0)j}\right)\right)^k = (e_{\alpha(t_0)j})^k = \delta^k_j.$$

Therefore the range of the map (6.7) is n in $\alpha(t_0)$ and thus it defines a local diffeomorphism about that point. This concludes the proof of (a). The proof of (b) follows, dealing with as in the proof of the analogous statement in theorem 6.1, using the following facts. (i) The geodesics starting from $\alpha(t)$ with initial tangent vector $\sum_{i=2}^n v^i \frac{\partial}{\partial y^k}|_{\alpha(t)}$ have equation $y^1(\lambda) = t$ (constant!) and $y^j(\lambda) = \lambda v^j$ for $j = 2, \ldots, n$, this implies that $\Gamma^k_{ij}(\alpha(t)) = 0$ if $i, j = 2, \ldots, n$. (ii) The vectors $\partial/\partial y^j$ with $j = 2, \ldots, n$ satisfy the equation of parallel transport with respect to α, that is with respect to $\partial/\partial y^1$:

$$\frac{d}{dt}\delta^k_j + \delta^i_1 \Gamma^k_{ir}(\alpha(t))\delta^r_j = 0\,,$$

this entails $\Gamma^k_{1j}(\alpha(t)) = \Gamma^k_{j1}(\alpha(t)) = 0$ for $j = 2, \ldots, n$.
The proof of (c) is consequence of the following further result. Since α is a geodesic, in coordinates $y^1, \ldots, y^n$, the curve $y^j = 0$ if $j = 2, \ldots, n$ and $y^1 = \lambda$ is a geodesic. With the same procedure as above, one find that $\Gamma^k_{11}(\alpha(t)) = 0$. $\Box$

Definition 6.3. Let M be a differentiable manifold endowed with a smooth Riemannian or Lorentzian metric g, and $\alpha : (a, b) \to M$ a differentiable curve with $\dot\alpha(t) \neq 0$ for every $t \in (a, b)$ and of time-like type if g is Lorentzian. Take $t_0 \in (a, b)$. The coordinate system defined as in (6.7) is called **(Riemannian) normal coordinate system about**, or **adapted to**, γ. $\diamondsuit$

6.1.3 The equivalence principle in General Relativity.

As previously said (see the example 5.1), a *spacetime* is a (C^∞) four-dimensional manifold M, equipped with a smooth metric Φ with Lorentzian signature $(-,+,+,+)$. A nonsingular smooth curve $M \ni \gamma = \gamma(t)$ with $t \in (a,b)$ is called *spacelike*, *timelike* or *null* (equivalently *lightlike*) if the tangent vector $\dot\gamma$ satisfies, respectively, $(\dot\gamma|\dot\gamma) > 0$, $(\dot\gamma|\dot\gamma) < 0$ and $(\dot\gamma|\dot\gamma) = 0$. *Causal curves* are those smooth curves which are piece-wisely timelike or null indifferently, these curves describes the stories of physical objects (material points) evolving in the universe. Causal geodesics are the stories of free-falling objects, i.e. experiencing the background gravitational force only. The gravitational force is described by the Lorentzian metric Φ.
Consider a free-falling observer $\mathscr{I}$ represented by the (C^∞) timelike geodesic $\gamma : (a,b) \to M$. We assume, as in general relativity, that the affine parameter t coincides with the *proper time* measured by a clock at rest with the observer. This is true provided the tangent vector $\dot\gamma$ fulfills $(\dot\gamma|\dot\gamma) = 1$ (see the example 5.1, we are assuming that $c = 1$, c being the speed of the light). Fix a point $p = \gamma(t_0)$ and consider a normal coordinate system (t, x^1, x^2, x^3) about γ including p in its domain. The property $\Gamma^i_{jk}(\gamma(t)) = 0$ has an important physical consequence, in General Relativity, called *equivalence principle.* If the considered observer $\mathscr{I}$ studies the stories (timelike geodesics) $\alpha = \alpha(t)$ of free-falling bodies leaving $\mathscr{I}$ at some t_i using the coordinates $(y^1, y^2, y^3, y^4) = (t, x^1, x^2, x^3)$, he finds that they are represented as straight non-accelerated motions for sufficiently small times about t_1. In other words, *locally, the gravitational effect has been canceled out 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. The equations for α have the form

$$\frac{d^2y^i}{d\lambda^2} = -\Gamma^i_{jk}(\alpha)\frac{dy^j}{d\lambda}\frac{dy^k}{d\lambda}, \quad i,j,k = 1,2,3,4\,,$$

where λ is any affine parameter for the geodesic α. We arrange λ such that $\lambda = 0$ individuates the event $\gamma(0)$, where α departs from γ. Taylor expansion about $\lambda = 0$ yields:

$$y^i(\lambda) = \lambda\frac{dy^j}{d\lambda}|_{\lambda=0} + \frac{\lambda^2}{2}\frac{d^2y^j}{d\lambda^2}|_{\lambda=0} + \lambda^2 O^j(\lambda)\,,$$

where $O^j(\lambda) \to 0$ as $\lambda \to 0$. However, in the considered coordinates, in view of (c) in Theorem 6.2, we have $\Gamma^i_{jk}(\alpha(0)) = 0$, so that:

$$\frac{d^2y^j}{d\lambda^2}|_{\lambda=0} = -\Gamma^i_{jk}(\alpha(0))\frac{dy^j}{d\lambda}|_{\lambda=0}\frac{dy^k}{d\lambda}|_{\lambda=0} = 0\,,$$

and thus

$$y^i(\lambda) = \lambda\frac{dy^j}{d\lambda}|_{\lambda=0} + \lambda^2 O^j(\lambda)\,. \qquad (6.10)$$

In particular, for the first coordinate $y^1 \equiv t$, we find

$$t(\lambda) = \lambda\frac{dt}{d\lambda}|_{\lambda=0} + \lambda^2 O^1(\lambda)\,. \qquad (6.11)$$

Since the geodesics α and γ are both timelike, it must be

$$\frac{dt}{d\lambda}|_{\lambda=0} \neq 0$$

(we leave the proof of the reader as an exercise). This fact implies that, about $t = 0$, we can use t as a parameter for the smooth curve α. The spatial components of (6.10) read

$$x^\alpha(\lambda) = \lambda \frac{dx^\alpha}{d\lambda}|_{\lambda=0} + \lambda^2 O^\alpha(\lambda) \,, \quad \alpha = 1, 2, 3 \,. \tag{6.12}$$

The velocity of α respect to γ for $t = 0$ has components

$$v^\alpha = \frac{\frac{dx^\alpha}{d\lambda}|_{\lambda=0}}{\frac{dt}{d\lambda}|_{\lambda=0}} \,, \quad \alpha = 1, 2, 3 \,,$$

so that, inserting it in (6.12) and taking (6.11) into account, we find with trivial manipulations,

$$x^\alpha(t) = tv^\alpha + t^2 O_1^\alpha(t) \,, \quad \alpha = 1, 2, 3 \,. \tag{6.13}$$

where $O_1^\alpha(t) \to 0$ as $t \to 0$. We have fount that, up to terms of third order in t, the motion of α is given by a constant-speed motion, so that the acceleration due to the gravitational field, has been suppressed by an appropriate choice of the coordinate. Actually, the way followed by Einstein was just the opposite: He first noticed the possibility to cancel out the effect of the gravitational field locally by an appropriate choice of the reference frame as a consequence of the fact that, in an assigned gravitational field, the Newtonian equations for the motion of test bodies do not depend on the masses of the test bodies[2]. This fact, and relying on the structure of special theory of relativity, lead him to postulate that the spacetime is a Lorentzian manifold where the Lorentzian metric accounts for the gravitational field and the stories of free falling bodies are (causal) geodesics.

6.2 Fermi's transport in Lorentzian manifolds.

Consider a differentiable curve $\gamma : (a, b) \to M$, M being a spacetime as in the previous subsection. We further assume that the curve is timelike, i.e., $(\dot\gamma(t)|\dot\gamma(t)) < 0$ everywhere along the curve. We finally assume that t denotes the length parameter and thus $(\dot\gamma(t)|\dot\gamma(t)) = -1$. t is the proper time associated with the particle which admits γ as its worldline (see example 5.1). It is possible to define a smooth vector field along the curve itself, i.e., the restriction $(a, b) \ni t \mapsto V_{\gamma(t)} \in T_{\gamma(t)}M$ of a a differentiable vector field defined in a neighborhood of γ. For the moment we also suppose that $V_{\gamma(t)} \in \Sigma_{\gamma(t)}$, $\Sigma_{\gamma(t)}$ denoting the subspace of $T_{\gamma(t)}(M)$ made of the vectors u with $(u|\dot\gamma(t)) = 0$. From a physical point of view, in the Lorentzian case, $V_{\gamma(t)}$ is a vector in the rest space $\Sigma_{\gamma(t)}$ at time t (see example 5.1) of the observer associated with the world line γ. For instance V could be the *spin* of a particle whose world line is γ itself.

---

[2] This is equivalent to say that the gravitational mass and the inertial mass coincide.

We want to formalize the idea of vectors V which *do not rotate* in $\Sigma_{\gamma(t)}$ during their evolution along the worldline *preserving metrical properties.*
As $T_{\gamma(t)}M$ is orthogonally decomposed as $L(\dot\gamma(t)) \otimes \Sigma_{\gamma(t)}$, the only possible infinitesimal deformations of $V_{\gamma(t)}$ during an infinitesimal interval of time t must take place in the linear space spanned by $\dot\gamma$. If $V_{\gamma(t)}$ does not satisfy $V_{\gamma(t)} \in \Sigma_{\gamma(t)}$, a direct generalization of the said condition is that the orthogonal projection of $V_{\gamma(t)}$ onto $\Sigma_{\gamma(t)}$ does not rotate in the sense said above: its infinitesimal evolution involves deformations along $\dot\gamma$ only. The second condition about the preservation of metrical structures means that $(V_{\gamma(t)}|V_{\gamma(t)})$ is preserved in the evolution along γ. Notice that $\dot\gamma$ naturally satisfies both requirements.
The non-rotating and metric preserving conditions can be generalized to set of vectors $\{V_{(a)\gamma(t)}\}_{a\in A}$: the annotating condition is formulated exactly as above for each vector separately, while the metric preserving property means that the scalar products $(V_{(a)\gamma(t)}|(V_{(b)\gamma(t)})$, with $a, b \in A$, are preserved during evolution along the line for $t \in (a, b)$.
In formulae, interpreting $\nabla_{\dot\gamma(t)}$ as said in **3.4**, if V is any differentiable contravariant vector field defined in an open neighborhood of $\gamma((a, b))$ and $V(t) := V(\gamma(t))$, the *non rotation constraint* reads:

$$\nabla_{\dot\gamma(t)} \left[V(t) + (V(t)|\dot\gamma(t))\, \dot\gamma(t)\right] = \alpha(t)\dot\gamma(t)\,, \tag{6.14}$$

for some suitable function α.

Remark 6.2.
(1) $V(t) + (V(t)|\dot\gamma(t))\dot\gamma(t)$ is the orthogonal projection of V onto $\Sigma_{\gamma(t)}$. Indeed as $T_{\gamma(t)}M = L(\dot\gamma(t)) \otimes \Sigma_{\gamma(t)}$,

$$V(t) = c(t)\dot\gamma(t) + X(t)\ ,$$

where $X(t) \in \Sigma_{\gamma(t)}$ is the wanted projection. Since $\Sigma_{\gamma(t)} = L(\dot\gamma(t))^\perp$,

$$(V(t), \dot\gamma(t)) = c(t)(\dot\gamma(t)|\dot\gamma(t)) = -c(t)$$

and thus

$$X(t) = V(t) + (V(t)|\dot\gamma(t))\dot\gamma(t)\ .$$

(2) We have interpreted the infinitesimal deformations of a vector $U(t)$ during an infinitesimal interval of time $dt = h$ as $dU = \nabla_{\dot\gamma(t)}U dt$ making explicit use of the Levi-Civita connection. As explained in **3.5**, up to an infinitesimal function of order h^2, $h\nabla_{\dot\gamma(t)}U$ is the difference of vectors in $T_{\gamma(t)}M$,

$$\mathcal{P}_\alpha^{-1}(\gamma(t), \gamma(t+h))U(\gamma(t+h)) - U(\gamma(t))\ ,$$

where α is a geodesic from $\gamma(t)$ to $\gamma(t+h)$ (which in general is different from γ) and $\mathcal{P}_\alpha(\alpha(u), \alpha(v)) : T_{\alpha(u)} \to T_{\alpha(v)}$ is the isometric vector-space isomorphism induced by Levi-Civita's connection by means of parallel transport along α (see the remark (3) after proposition 5.3.) The existence of the geodesic α is assured if h is sufficiently small (see the remark (5) after proposition 5.3).

It is possible to get a mathematical formulation of the non-rotating condition more precise than (6.14). Expanding (6.14) we get

$$\nabla_{\dot\gamma(t)}V(t) + (\nabla_{\dot\gamma(t)}V(t)|\dot\gamma(t))\dot\gamma(t) + (V(t)|\nabla_{\dot\gamma(t)}\dot\gamma(t))\dot\gamma(t) + (V(t)|\dot\gamma(t))\nabla_{\dot\gamma(t)}\dot\gamma(t) = \alpha(t)\dot\gamma(t) \tag{6.15}$$

Taking the scalar product with $\dot\gamma(t)$ and using $(\dot\gamma(t)|\dot\gamma(t)) = -1$ we obtain

$$(\nabla_{\dot\gamma(t)}V(t)|\dot\gamma(t)) - (\nabla_{\dot\gamma(t)}V(t)|\dot\gamma(t)) - (V(t)|\nabla_{\dot\gamma(t)}\dot\gamma(t)) = -\alpha(t) \tag{6.16}$$

and thus

$$(V(t)|\nabla_{\dot\gamma(t)}\dot\gamma(t)) = \alpha(t)\,.$$

That identity used in the right-hand side of (6.14) produces the more precise equation

$$\nabla_{\dot\gamma(t)}\left[V(t) + (V(t)|\dot\gamma(t))\,\dot\gamma(t)\right] = (V(t)|\nabla_{\dot\gamma(t)}\dot\gamma(t))\dot\gamma(t)\,. \tag{6.17}$$

Equivalently:

$$\nabla_{\dot\gamma(t)}V(t) + \nabla_{\dot\gamma(t)}\left[(V(t)|\dot\gamma(t))\,\dot\gamma(t)\right] - (V(t)|\nabla_{\dot\gamma(t)}\dot\gamma(t))\dot\gamma(t) = 0\,,$$

or

$$\nabla_{\dot\gamma(t)}V(t) + (V(t)|\dot\gamma(t))\nabla_{\dot\gamma(t)}\dot\gamma(t) + (\nabla_{\dot\gamma(t)}V(t)|\dot\gamma(t))\dot\gamma(t) = 0\,. \tag{6.18}$$

This identity, which is the mathematical formulation of the non-rotating property, can be rewritten in a more suitable form which allows one to use the *metric preserving property*:

$$\nabla_{\dot\gamma(t)}V(t) + (V(t)|\dot\gamma(t))\nabla_{\dot\gamma(t)}\dot\gamma(t) - (V(t)|\nabla_{\dot\gamma(t)}\dot\gamma(t))\dot\gamma(t) + \left[\frac{d}{dt}(V(t)|\dot\gamma(t))\right]\dot\gamma(t) = 0\,. \tag{6.19}$$

Both $\dot\gamma$ and V satisfy the metric preserving property and thus it also holds

$$\frac{d}{dt}(V(t)|\dot\gamma(t)) = 0 \tag{6.20}$$

As a consequence (6.19) reduces to

$$\nabla_{\dot\gamma(t)}V(t) + (V(t)|\dot\gamma(t))\nabla_{\dot\gamma(t)}\dot\gamma(t) - (V(t)|\nabla_{\dot\gamma(t)}\dot\gamma(t))\dot\gamma(t) = 0\,. \tag{6.21}$$

We have found that if V satisfies both the non-rotating condition and the metric preserving condition, it satisfies (6.21). However if vectors satisfy (6.21) their scalar products along γ are preserved as shown below, moreover $\dot\gamma$ itself satisfies (6.21) and thus (6.20) holds true. We conclude that (6.21) implies both (6.19), which states the *non-rotating property*, and the *metric preserving property*. (6.21) is the wanted equation.

Definition 6.4. (**Fermi's Transport of a vector along a curve**.) Let M be a Lorentzian manifold and $\gamma : [a, b] \to M$ a timelike (i.e. $(\dot\gamma(t)|\dot\gamma(t) < 0$ for all $t \in [a, b]$) differentiable curve where t is the length parameter (i.e., the proper time). A differentiable vector field V defined in a neighborhood of $\gamma([a,b])$ is said to be **Fermi transported** along γ if

$$\nabla_{\dot\gamma(t)}V(\gamma(t))) + (V(\gamma(t))|\dot\gamma(t))\nabla_{\dot\gamma(t)}\dot\gamma(t) - (V(\gamma(t))|\nabla_{\dot\gamma(t)}\dot\gamma(t))\dot\gamma(t) = 0$$

for all $t \in [a,b]$. ♢

Proposition 6.1. *The notion of Fermi transport along a curve $\gamma : [a,b] \to M$ defined in definition 6.4 enjoys the following properties.*
(1) *It is* **metric preserving**, *i.e, if $t \mapsto V(\gamma(t)$ and $t \mapsto V'(\gamma(t)$ are Fermi transported along γ,*

$$t \mapsto (V(\gamma(t))|V'(\gamma(t)))$$

is constant in $[a,b]$.
(2) *$t \mapsto \dot\gamma(t)$ is Fermi transported along γ.*
(3) *If γ is a geodesic with respect to Levi-Civita's connection, the notions of parallel transport and Fermi transport along γ coincide.*♢

Proof. (1) Using the fact that the connection is metric one has:

$$\frac{d}{dt}(V(\gamma(t))|V'(\gamma(t))) = (\nabla_{\dot\gamma}V(\gamma(t)|V'(\gamma(t))) + (V(\gamma(t))|\nabla_{\dot\gamma}V'(\gamma(t)))\,. \tag{6.22}$$

Making use of the equation of Fermi's transport,

$$\nabla_{\dot\gamma(t)}U(\gamma(t)) = -(U(\gamma(t))|\dot\gamma(t))\nabla_{\dot\gamma(t)}\dot\gamma(t) + (U(\gamma(t))|\nabla_{\dot\gamma(t)}\dot\gamma(t))\dot\gamma(t)\,,$$

for both V and V' in place of U, the terms in the right-hand side of (6.22) cancel out each other. The proof of (2) is direct by noticing that

$$(\dot\gamma(t)|\dot\gamma(t)) = -1$$

and

$$(\dot\gamma(t)|\nabla_{\dot\gamma}\dot\gamma(t)) = \frac{1}{2}\frac{d}{dt}(\dot\gamma(t)|\dot\gamma(t)) = -\frac{1}{2}\frac{d}{dt}1 = 0\,.$$

The proof of (3) is trivial noticing that if γ is a geodesic $\nabla_{\dot\gamma(t)}\dot\gamma(t) = 0$ and (6.22) reduces to the equation of the parallel transport

$$\nabla_{\dot\gamma(t)}U(\gamma(t)) = 0\,.$$

□

Remark 6.3.
(1) If $\gamma : [a,b] \to M$ is fixed, the Fermi's transport condition

$$\nabla_{\dot\gamma(t)}V(\gamma(t)) = (V(\gamma(t))|\nabla_{\dot\gamma(t)}\dot\gamma(t))\dot\gamma(t) - (V(\gamma(t))|\dot\gamma(t))\nabla_{\dot\gamma(t)}\dot\gamma(t)$$

can be used as a differential equation. Expanding both sides in local coordinates $(x^1, \ldots, x^n)$ 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, the initial vector $V(\gamma(a))$ determines V uniquely along the whole curve (see the analog comment for the case of the parallel transport). In a certain sense, one may view the solution $t \mapsto V(t)$ as the "transport" and "evolution" of the initial condition $V(\gamma(a))$ along γ itself.
The global existence and uniqueness theorem has an important consequence. If $\gamma : [a, b] \to M$ is fixed and $u, v \in [a, b]$ with $u \neq v$, the notion of parallel transport along γ produces an vector space isomorphism $\mathcal{F}_\gamma[\gamma(u), \gamma(v)] : T_{\gamma(u)} \to T_{\gamma(v)}$ which associates $V \in T_{\gamma(u)}$ with that vector in $T_{\gamma(u)}$ which is obtained by Fermi's transporting V in $T_{\gamma(u)}$. Notice that $\mathcal{F}_\gamma[\gamma(u), \gamma(v)]$ also preserves the scalar product by property (1) of proposition 6.1, i.e., it is an isometric isomorphism.
(2) The equation of Fermi transport of a vector X in a n-dimensional Lorentz manifold M can be re-written

$$\nabla_{V(t)} X(\gamma(t)) = (X(\gamma(t))|A(t))V(t) - (X(\gamma(t))|V(t))A(t) ,$$

where we have introduced the n-**velocity** $V(t) := \dot{\gamma}(t)$ and the n-**acceleration** $A(t) := \nabla_{\dot{\gamma}(t)}\dot{\gamma}(t)$ of a worldline γ parametrized by the proper time t. These vectors have a deep physical meaning if $n = 4$ (i.e., M ia a *spacetime*). Notice that $(A(t)|V(t)) = 0$ for all t and thus if $A \neq 0$, it turns out to be *spacelike* because V is timelike by definition.
(3) The non-rotating property of Fermi transport can be viewed from another point of view. Consider the proper Lorentz group $SO(1, 3)$ represented by real 4×4 matrices $\Lambda : \mathbb{R}^4 \to \mathbb{R}^4$ $\Lambda = [\Lambda^i_j]$, $i, j = 0, 1, 2, 3$. Here the coordinate x^0 represents the *time coordinate* and the remaining three coordinates are the space coordinates. It is known that every $\Lambda \in SO(1, 3)$ can uniquely be decomposed as

$$\Lambda = \Omega P ,$$

where $\Omega, P \in SO(1, 3)$ 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 contains rotations and represents the coordinate transformation between a pair of pseudo-orthonormal reference frames (in Minkowski spacetime) 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 $O(h) \to 0$ as $h \to 0$. 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 reabsorbed in the coefficients A_i) are called *infinitesimal pure Lorentz transformations.*

Then consider a differentiable timelike curve $\gamma : [0, \epsilon) \to M$ starting from p in a four dimensional Lorentzian manifold M and fix a pseudo-orthonormal basis in T_pM, 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's transport along γ. We want to investigate the following issue.

What is the Lorentz transformation which relates the basis $\{e_i(t)\}_{i=0,\ldots,3}$ with the basis of Fermi transported elements $\{e_i(t+h)\}_{i=0,\ldots,3}$ in the limit $h \to 0$?

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 $\{e_i(t)\}_{i=0,\ldots,3}$ with the basis $\{e_i(t+h)\}_{i=0,\ldots,3}$ we have to transport, by means of parallel transport, the latter basis in $\gamma(t)$. In other words we want to find the Lorentz transformation between $\{e_i(t)\}_{i=0,\ldots,3}$ and $\{\mathcal{P}_\alpha^{-1}[\gamma(t), \gamma(t+h)]e_i(t+h)\}_{i=0,\ldots,3}$, α being the geodesic joining $\gamma(t)$ and $\gamma(t+h)$ for h small sufficiently. We define

$$e'_i(t+h) := \mathcal{P}_\alpha^{-1}[\gamma(t), \gamma(t+h)]e_i(t+h) .$$

By the discussion in **3.5** we have

$$e'_i(t+h) - e_i(t) = h\nabla_{\dot\gamma(t)} e_i(t) + hO(h) .$$

Using the equation of Fermi transport we get

$$e'_i(t+h) - e_i(t) = h(e_i(t)|A(t))e_0(t) - h(e_i(t)|e_0(t))A(t) + hO(h) , \tag{6.23}$$

where $A(t) = \nabla_{\dot\gamma(t)}\dot\gamma(t)$ is the 4-acceleration of the worldline γ itself and $O(h) \to 0$ as $h \to 0$. Notice that $(A(t)|e_0(t)) = 0$ by the remark (2) above and thus

$$A(t) = \sum_{i=1}^{3} A_i(t)e_i(t) , \tag{6.24}$$

for some triple of functions A_1, A_2, A_3. If $\eta_{ab} = \mathrm{diag}(-1, 1, 1, 1)$ and taking (6.24) and the pseudo orthonormality of the basis $\{e_i(t)\}_{i=0,\ldots,3}$ into account, (6.23) can be re-written

$$e'_i(t+h) = e_i(t) + h(A_i(t)e_0(t) - \eta_{i0}A(t)) + hO(h) . \tag{6.25}$$

If we expand $e'_i(t+h)$ in components refereed to the basis $\{e_i(t)\}_{i=0,\ldots,3}$, (6.25) becomes

$$(e'_i(t+h))^j = \delta_i^j + h(A_i(t)\delta_0^j(t) - \eta_{i0}A_j(t)) + hO^j(h)\,, \tag{6.26}$$

where one should remind that $A_0 = 0$. As $(e_i(t))^j = \delta_i^j$, (6.26) can be re-written

$$e'_i(t+h) = I + h\left(\sum_{j=1}^{3} A_j K_j\right) e_i(t) + hO(h)\,. \tag{6.27}$$

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 (except for terms $hO(h)$) if $A = 0$, i.e., if the curve is a timelike geodesic.

Chapter 7

Curvature.

We remind the reader that a (pseudo) Riemannian manifold is called **locally flat** (see definition 4.2), if it admits an atlas made of local **canonical charts** (U, ψ), that is the coeffients $g_{ij}(p)$, representing the metric, are constant when varying $p \in U$ and assume the canonical diagonal values $\pm\delta_{ij}$ with sign determined by the segnature of the metric. In general the mentioned atlas does not contain *global canonical chartes*, 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 h induced by the standared metric of $\mathbb{R}^2$: (C, h) is locally flat but not globally flat.
Let M be a Riemannian manifold which is locally flat. As the metric tensor is constant in canonical coordinates defined in a neighborhood U of any $x \in M$, the Levi-Civita connection is represented by trivial connection coefficients in those coordinates: $\Gamma^k_{ij} = 0$. As a consequence, in those coordinates it holds

$$\nabla_i \nabla_j Z^k = \frac{\partial^2 Z^k}{\partial x^i \partial x^j} = \frac{\partial^2 Z^k}{\partial x^j \partial x^i} = \nabla_j \nabla_i Z^k \ ,$$

for every differentiable vector field Z defined in U. In other words, the covariant derivatives commute on differentiable vector fields defined on U:

$$\nabla_i \nabla_j Z^k = \nabla_j \nabla_i Z^k$$

Notice that, by the intrinsic nature of covariant derivatives, that identity holds in any coordinate system in the neighborhood U of $p \in M$, not only in those coordinates where the connection coefficients vanish. Since p is arbitrary, we have proved that *the local flatness of* (M, Φ) *implies local commutativity of (Levi-Civita) covariant derivatives on vector fields on* M. This fact completely characterizes locally flat (pseudo) Riemannian manifolds because the converse proposition holds true too as we prove at the end of this chapter. Therefore a (pseudo) Riemannian manifold can be considered "curved" whenever local commutativity of (Levi-Civita) covariant derivatives fails to be satisfied. Departing from (pseudo) Riemannian manifolds, investigation about commutativity of covariant derivatives naturally leads to a very important tensor R, called the *curvature tensor* (field). Commutativity of the covariant derivatives in M turns out to be equivalent to

$R = 0$ in M. Actually, coming back to manifolds equipped with Levi-Civita's connection, it is possible to prove a stronger result, i.e., the condition $R = 0$ locally is equivalent to the local flatness of the manifold. The next sections are devoted to these topics and straightforward extensions to cases of non metric connections.

7.1 Curvature tensor and Riemann's curvature tensor.

To introduce (Riemann's) curvature tensor let us consider the commutativity property of covariant derivative once again.

Lemma 7.1. *Let M be a differentiable manifold equipped with a torsion-free affine connection ∇ (e.g. Levi-Civita's connection with respect to some metric on M).*
Covariant derivatives of contravariant vector fields commute in M, i.e.,

$$\nabla_i \nabla_j Z^k = \nabla_j \nabla_i Z^k \,. \tag{7.1}$$

in every local coordinate system, for all differentiable contravariant vector fields Z and all coordinate indices i, j, k, if and only if

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z = 0 \,, \tag{7.2}$$

for all differentiable vector fields X, Y, Z in M.$\diamondsuit$

Proof. If X, Y are differentiable vector fields (7.1) 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

$$X^i \nabla_i Y^j \nabla_j Z^k - Y^j \nabla_j X^i \nabla_i Z^k - (X^i (\nabla_i Y^j) \nabla_j Z^k - Y^i (\nabla_i X^j) \nabla_j Z^k) = 0 \,,$$

and finally

$$X^i \nabla_i Y^j \nabla_j Z^k - Y^j \nabla_j X^i \nabla_i Z^k - (X^i (\nabla_i Y^j) - Y^i (\nabla_i X^j)) \nabla_j Z^k = 0 \,.$$

Using proposition 5.1, the above identity can be re-written in the implicit form

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z = 0 \,.$$

(7.2) is equivalent to (7.1) because the latter implies the former as shown and the former implies the latter under the specialization $X = \frac{\partial}{\partial x^i}$ and $Y = \frac{\partial}{\partial x^j}$. Notice that $[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] = 0$.

Proposition 7.1. *Let M be a differentiable manifold equipped with an affine connection ∇ (not necessarily tosion free).*
(a) *There is a (unique) differentiable tensor field R such that, for every $p \in M$ the tensor R_p belongs to $T_p^*M \otimes T_p^*M \otimes T_p^*M \otimes T_pM$ and*

$$R_p(X_p, Y_p, Z_p) = \left(\nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X,Y]} Z\right)_p . \tag{7.3}$$

(b) *In local coordinates,*

$$R_{ijk}{}^l = \frac{\partial \Gamma^l_{ik}}{\partial x^j} - \frac{\partial \Gamma^l_{jk}}{\partial x^i} + \Gamma^r_{ik}\Gamma^l_{jr} - \Gamma^r_{jk}\Gamma^l_{ir} , \tag{7.4}$$

where

$$(R_p)_{ijk}{}^l := \left\langle R_p\left(\frac{\partial}{\partial x^i}|_p, \frac{\partial}{\partial x^j}|_p, \frac{\partial}{\partial x^k}|_p\right), dx^l_p \right\rangle .$$

♢

Proof. By direct inspection one finds that, if $R_{ijk}{}^l$ is defined as in the right-hand side of (7.4), and $(R(X,Y,Z))_p$ is defined as in the right-hand side of (7.3), then:

$$(R(X,Y,Z))^l_p = R_{ijk}{}^l X^i_p Y^j_p Z^k_p ,$$

because, in particular, all derivatives of the fields X, Y, Z cancels each other. Such an identity proves, in particular, that the mapping which associates triples of differentiable contravariant vector fields on M, X, Y, Z, to the contravariant vector

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

depends only on the values attained at p by X, Y, Z. Since the map is multi $\mathbb{R}$-linear, it defines a tensor in $R_p \in T_p^*M \otimes T_p^*M \otimes T_p^*M \otimes T_pM$. By construction the components of R_p are given in (7.4) and (7.3) holds true. Finally, by construction, R_p is differentiable when varying p. □

Remark 7.1. Notice that, in the hypotheses, we have not assumed that the connection is Levi-Civita's one nor that it is torsion free.

Definition 7.1. (**Curvature tensor and Riemann's curvature tensor**.) *The differentiable tensor field R associated to the affine connection ∇ on a differentiable manifold M as indicated in proposition 7.1 is called* **curvature tensor** *(field)* **associated with** ∇. *If ∇ is Levi-Civita's connection obtained by a metric Φ, R is called* **Riemann's curvature tensor** *(field)* **associated with** Φ.

From now on we adopt the following usual notations: $R(X,Y,Z)$ indicates the vector field which coincides with $R_p(X_p, Y_p, Z_p)$ at every point $p \in M$. Moreover $R(X,Y)Z := R(X,Y,Z)$, in other words $R(X,Y)$ denotes the differential operator acting on differentiable contravariant vector fields

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

7.1.1 Flatness and curvature tensor.

To conclude we state a general proposition concerning, in particular, the interplay between flatness and curvature tensor. This statement will be completed later into a more general proposition.

Proposition 7.2. *Let M be a differentiable manifold equipped with a torsion-free affine connection ∇. The following facts are equivalent.*
(a) *Covariant derivatives of differentiable tensor fields Ξ commute i.e.,*

$$\nabla_i \nabla_j \Xi^A = \nabla_j \nabla_i \Xi^A ,$$

in every local coordinate frame;
(b) *covariant derivatives of differentiable contravariant vector fields X commute;*
(c) *covariant derivatives of differentiable covariant vector fields ω commute;*
(d) *the curvature tensor associated with ∇ vanishes everywhere in M, i.e., $R = 0$ in M.*

Moreover, if for every point $p \in M$ there is a local chart (U, ψ) with $p \in U$ and such that the connection coefficients of ∇ associated with (U, ψ) vanish in U the following pair of facts hold;
(e) *the curvature tensor vanishes everywhere in M;*
(f) *the covariant derivatives of differentiable tensor fields commute.*
In particular, if (M, Φ) is (pseudo) Riemannian and ∇ is Levi-Civita's connection, whenerver (M, Φ) is locally flat the following pair of facts hold;
(e)' *Riemann's curvature tensor vanishes everywhere in M;*
(f)' *Levi-Civita's covariant derivatives of differentiable tensor fields commute.*$\diamondsuit$

Proof. It is clear that (a) implies (b) and (c) and, together (b) and (c) imply (a) by linearity and property **(6)** of covariant derivatives (see below proposition 5.2). Finally (b) can be shown to be equivalent to (c) by direct use of properties **(5)** and **(7)** .
Let us prove the equivalence of (b) and (d). lemma 5.1 proves that $\nabla_i \nabla_j Z^k = \nabla_j \nabla_i Z^k$ for all Z is equivalent to $\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z = 0$ for all X, Y, Z. In other words $\nabla_i \nabla_j Z^k = \nabla_j \nabla_i Z^k$ for all Z is equivalent to the fact that the multi-linear mapping associated to R at each point of M vanishes (notice that lemma 2.5 must be used to achieve such a conclusion). This is equivalent to $R = 0$ in M.
The last statement is a straightforward consequence of (7.4): in the coordinates of the chart (U, ψ) all the coefficients R^i_{jkl} must vanish, but since they define a tensor, they vanish in every coordinate system, i.e., $R = 0$ in M. As a consequence covariant derivatives of differentiable tensor fields X commute because of the equivalence of (d) and (b).
In the case of the Levi-Civita connection notice that local flatness implies that for each $p \in M$ there is a coordinate patch defined about p where the coefficients of the metric are constant and thus Levi-Civita connection coefficients vanish and one reduces to the previously considered case. $\square$

Exercises 7.1.   .

1. Prove that
$$\nabla_i\nabla_j\omega_k - \nabla_j\nabla_i\omega_k = R_{ijk}{}^l\omega_l\ .$$

2. Prove that, in the general case, **Ricci's identity** holds:
$$\nabla_i\nabla_j\Xi^{i_1\cdots i_p}{}_{j_1\cdots j_q} - \nabla_j\nabla_i\Xi^{i_1\cdots i_p}{}_{j_1\cdots j_q} = -\sum_{u=1}^{p} R_{ijs}{}^{i_u}\Xi^{i_1\cdots s\cdots i_p}{}_{j_1\cdots j_q} + \sum_{u=1}^{p} R_{ijj_u}{}^{s}\Xi^{i_1\cdots i_p}{}_{j_1\cdots s\cdots j_q}\ .$$

7.2 Properties of curvature tensor. Bianchi's identity.

The curvature tensor enjoys a set of useful properties which we go to summarize in the proposition below. In the (pseudo) Riemannian case, these properties are very crucial in physics because they play a central rôle in relativistic theories as we specify below.

Proposition 7.3. *The curvature tensor associated with an affine connection ∇ on a differentiable manifold M enjoys the following properties where X, Y, Z, W are arbitrary differentiable contravariant vector fields on M.*
(1)
$$R(X,Y)Z = -R(Y,X)Z \quad \textit{or equivalently} \quad R_{ijk}{}^l = -R_{jik}{}^l\ ;$$
(2) *If ∇ is torsion free,*
$$R(X,Y,Z) + R(Y,Z,X) + R(Z,X,Y) = 0 \quad \textit{or equivalently} \quad R_{ijk}{}^l + R_{jki}{}^l + R_{kij}{}^l = 0\ ;$$
(3) *if ∇ is metric [i.e.$\nabla\Phi = 0$ where locally $\Phi = g_{ij}dx^i\otimes dx^j$ is a (pseudo)metric on M],*
$$(R(X,Y)Z|W) = -\left(Z|R(X,Y)W\right) \quad \textit{or equivalently} \quad R_{ijkl} = -R_{ijlk}$$
where $R_{ijkl} := R_{ijk}{}^r g_{rl}$;
(4) *if ∇ is Levi-Civita's connection,* **Bianchi's identity** *holds*
$$\nabla_Y R(Z,W) + \nabla_Z R(W,Y) + \nabla_W R(Y,Z) = 0 \textit{ or equivalently } \nabla_k R_{ijp}{}^a + \nabla_i R_{jkp}{}^a + \nabla_j R_{kip}{}^a = 0.$$
(5) *if ∇ is Levi-Civita's connection,*
$$R_{ijkl} = R_{klij}\ .$$
$\diamondsuit$

Proof. (1) is an immediate consequence of the definition of the curvature tensor given in proposition 7.1.
To prove (2) we start from the identity,
$$\nabla_{[i}\nabla_j\omega_{k]} := \frac{1}{3!}\left(\nabla_i\nabla_j\omega_k + \nabla_j\nabla_k\omega_i + \nabla_k\nabla_i\omega_j - \nabla_j\nabla_i\omega_k - \nabla_i\nabla_k\omega_j - \nabla_k\nabla_j\omega_i\right) = 0$$

which can be checked by direct inspection and using $\Gamma^r_{pq} = \Gamma^r_{qp}$. Then one directly finds by (7.4), $\nabla_i\nabla_j\omega_k - \nabla_j\nabla_k\omega_k = R_{ijk}{}^l\omega_l$ (see Exercise 7.1.1) and thus $\nabla_{[i}\nabla_j\omega_{k]} - \nabla_{[j}\nabla_i\omega_{k]} = R_{[ijk]}{}^l\omega_l$. And thus $R_{[ijk]}{}^l\omega_l = 0$. Since ω 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 7.1.2)

$$\nabla_i\nabla_j\Xi^{i_1\cdots i_p}{}_{j_1\cdots j_q} - \nabla_j\nabla_i\Xi^{i_1\cdots i_p}{}_{j_1\cdots j_q} = -\sum_{u=1}^{p} R_{ijs}{}^{i_u}\,\Xi^{i_1\cdots s\cdots i_p}{}_{j_1\cdots j_q} + \sum_{u=1}^{p} R_{ijj_u}{}^{s}\,\Xi^{i_1\cdots i_p}{}_{j_1\cdots s\cdots j_q}$$

to the case $\Xi = \Phi$ and using $\nabla_i g_{j_1j_2} = 0$.
(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 indices ijk and summing the results one gets

$$\begin{aligned}
&\left(X^a{}_{,ijk} - X^a{}_{,jik} - R_{ijp}{}^a X^p_{,k}\right) + \left(X^a{}_{,jki} - X^a{}_{,ikj} - R_{jkp}{}^a X^p_{,i}\right)\\
&+ \left(X^a{}_{,kij} - X^a{}_{,kji} - R_{kip}{}^a X^p_{,j}\right)\\
&= R_{ijp}{}^a{}_{,k}\, X^p + R_{jkp}{}^a{}_{,i}\, X^p + R_{kip}{}^a{}_{,j}\, X^p\,.
\end{aligned}$$

Using Ricci's identity (Exercise 7.1.2) and property (2) in the component form, one gets

$$X^a_{,p}\left(R_{ijk}{}^p + R_{jki}{}^p + R_{kij}{}^p\right) = 0$$

for every vector field X. Since that field is arbitrary one has

$$X^r_{,p}\left(R_{ijk}{}^p + R_{jki}{}^p + R_{kij}{}^p\right) = 0\,.$$

As a consequence it also holds

$$R_{ijp}{}^a{}_{,k}\, X^p + R_{jkp}{}^a{}_{,i}\, X^p + R_{kip}{}^a{}_{,j}\, X^p = 0\,.$$

Since X is arbitrary, we get Bianchi's identity (4). Contracting the inidice with thoe os 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).□

Exercises 7.2.

1. Prove that, at every point $p \in M$, R_{ijkl} has $n^2(n^2-1)/12$ independent components, R_{ijkl} being Riemann's tensor of a (pseudo) Riemannian manifold with dimension n. (*Hint. Use properties (1) and (2) and (3) above.*)

7.3 Ricci's tensor and Einstein's 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 to be useful in physics. 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) indices gives rise to a nontrivial tensor called **Ricci's tensor**:

$$Ric_{ij} := R_{ij} := R_{ikj}{}^{k} = R_{ki}{}^{k}{}_{j}.$$

By property (5) above one has the symmetry of Ric:

$$Ric_{ij} = Ric_{ji}\,.$$

The contraction of Ric produces the so-called **curvature scalar**

$$S := R := R^k_k\,.$$

Another relevant tensor is the so-called **Einstein's tensor** which plays a crucial role in General Relativity,

$$G_{ij} := Ric_{ij} - \frac{1}{2} g_{ij} S\,.$$

Einstein's tensor satisfies the equations

$$G_{ij},{}^{i} = 0$$

Let us prove those identities. 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 over 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:

$$\frac{1}{2}\nabla_i Ric^i_j + \frac{1}{2}\nabla_i Ric^i_j + \frac{1}{2}\nabla^i g_{ij} S = 0\,.$$

Those are the equations written above, since they can be rearranged into:

$$\nabla^i \left(Ric_{ij} + \frac{1}{2} g_{ij} S \right) = 0 \,.$$

Remark 7.2. The very celebrated **Einstein's equations** read

$$G_{ij} = kT_{ij} \,.$$

Above $k > 0$ is a constant and T is the so-called **stress-energy tensor** (field). That symmetric tensor field represents, in General Relativity, the mass-energy-momentum content of the material objects responsible for the gravity. Notice that the equations above hold at each point of the spacetime (a Lorentzian manifold). T satisfies another equations of the form

$$T_{ij},{}^{j} = 0 \,.$$

From a pure mathematical point of view, that identity must hold as a consequence of Einstein's equations and Ricci's identity. In the next section we prove that the local flatness of a (pseudo)Riemannian manifold, M, is equivalent to the fact that Riemann's tensor field vanishes everywhere in M. In General Relativity, the presence of gravity is mathematically defined as the non-flatness of the manifold (the spacetime). Equations of Einstein locally relate the tensor field G, instead of Riemann's one, with the content of matter in the spacetime. As a consequence the absence of matter does not imply that the Riemann tensor vanishes and the manifold is flat, i.e., there is no gravity. This fact is obvious from a physical point of view: gravity is present away from physical bodies because gravity propagates. However a flat spacetime must not have matter content because $R_{ijk}{}^{l} = 0$ implies $G_{ij} = 0$.

7.3.1 Weyl's tensor.

As we said above, in a (pseudo)Riemannian manifold M, Ricci's tensor and the curvature scalar are the only nonvanishing tensors which can be obtained from Riemann tensor using contractions. If $dimM =: n \geq 3$, 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 7.3 and produces the same tensors as* R_{ijkl} *under contractions.* That tensor is

$$D_{ijkl} := \frac{2}{n-2} g_{i[k} Ric_{l]j} - g_{j[k} Ric_{l]i} - \frac{2}{(n-1)(n-2)} S g_{i[k} g_{l]j} \,.$$

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 *con formal transformations.*

7.4 Flatness and Riemann's curvature tensor: the whole story.

In this section we establish a fundamental theorem concerning the whole interplay between Riemann curvature tensor and local flatness of a (pseudo)Riemannian manifold. By proposition 7.2, we know that the Riemann tensor must vanish whenever the manifold is (locally) flat. We aim to show that also the converse proposition holds true. In fact, Riemann's curvature tensor vanishes everywhere in a (pseudo)Riemannian manifold M *if and only* if M is locally flat. Actually we shall prove a more general proposition concerning torsion-free connection also different from Levi-Civita ones: In the general case of a torsion-free affine connection, the absence of curvature is equivalent to the existence, in a neighborhood of every point of the manifold, of a coordinate patch where the coefficient of the connection vanish.

Remark 7.3.
In the Levi-Civita case, this result has a remarkable consequence in physics since $R = 0$ if and only if there is no "geodesic deviation", i.e., there is no gravity in a spacetime. By this way one is allowed to physically identify *gravity* with *Riemannian curvature.*

7.4.1 Frobenius' theorem.

A lemma is necessary. That lemma is nothing but an elementary form of well-known Frobenius' theorem. Its proof can be found in any textbook of first order partial differential equations.

Lemma 7.2. *Let $U \subset \mathbb{R}^n$ an open set and let $F_{ij} : U \times \mathbb{R}^m \to \mathbb{R}$ be a set of C^∞ mapping, $i = 1, \ldots, n$, $j = 1, \ldots, m$. Consider the following system of differential equations*

$$\frac{\partial X_j}{\partial x^i} = F_{ij}(x^1, \ldots x^n, X_1, \ldots X_m)\,. \tag{7.5}$$

where $X_j = X_j(x^1, \ldots x^n)$ are real-valued C^∞ functions. For every point $p \in U$ and every set of initial conditions $X_j(p) = X_{j(0)}$, $j = 1, \ldots, m$, a C^∞ solution $\{X_j\}_{j=1,\ldots,m}$ exists in a neighborhood of p and it is unique therein if, for all $j = 1, \ldots, m$ the following **Frobenius'conditions** *hold.*

$$\begin{aligned}
&\frac{\partial F_{ij}(x^1, \ldots x^n, Y_1, \ldots Y_m)}{\partial x^k} + \sum_{r=1}^{j} \frac{\partial F_{ij}(x^1, \ldots x^n, Y_1, \ldots Y_m)}{\partial Y^r} F_{kr}(x^1, \ldots x^n, Y_1, \ldots Y_m) \\
= &\frac{\partial F_{kj}(x^1, \ldots x^n, Y_1, \ldots Y_m)}{\partial x^i} + \sum_{r=1}^{j} \frac{\partial F_{kj}(x^1, \ldots x^n, Y_1, \ldots Y_m)}{\partial Y^r} F_{ir}(x^1, \ldots x^n, Y_1, \ldots Y_m)
\end{aligned}$$

on $U \times \mathbb{R}^m$. ♢

Remark 7.4.
(1) Frobenius' conditions are nothing but the statement of Schwarz' theorem referred to the

solution $\{X_j\}_{j=1,\ldots,m}$,

$$\frac{\partial^2 X_j}{\partial x^i \partial x^k} = \frac{\partial^2 X_j}{\partial x^k \partial x^i} ,$$

written in terms of the functions F_{ij}, making use of the differential equation (7.5) itself.
(2) Actually the theorem could be proved with a weaker requirement about the smoothness of the involved functions (if each F_{ij} is C^2 the thesis holds true anyway and the fields X_j are C^3).

7.4.2 The crucial theorem.

We can state and prove the crucial theorem.

Theorem 7.1. *Let M be a manifold equipped with a torsion free smooth affine connection ∇. The following facts are equivalent.*
(a) *Every point $p \in M$ there is a local chart (U, ψ) with $p \in U$ such that the connection coefficients of ∇ referred to (U, ψ) vanish everywhere in U;*
(b) *Curvature tensor vanishes everywhere in M;*
(c) *Covariant derivatives of contravariant vector fields in M commute;*
(d) *Covariant derivatives of covariant vector fields in M commute;*
(e) *Covariant derivatives of tensor fields in M commute.*
If (M, Φ) is (pseudo) Riemannian, and ∇ is the Levi-Civita connection associated with Φ, each condition above is equivalent to the local flatness condition for (M, Φ).$\diamondsuit$

Proof. By proposition 7.2 we know that (a) implies (b) and that (b), (c), (d) and (e) are equivalent. We only have to show that (b) implies (a). In other words we have to show that if the curvature tensor vanishes everywhere, there is an open neighborhood of each $p \in M$ where the connection coefficients vanish or, in the (pseudo) Riemannian case, canonical coordinates can be defined therein where the metric takes its constant diagonal form. To this end fix any $p \in M$ and take vector basis in T_pM, $e_1, \cdots, e_n$, this basis has to be taken (pseudo) orthonormal in the case of a Levi-Civita connection. The proof consists of two steps.
(A) First of all, we prove that there are n differentiable (C^∞) contravariant vector fields $X_{(1)}, \ldots, X_{(n)}$ defined in a sufficiently small neighborhood U of p such that $(X_{(a)})_p = e_a$ and $\nabla X_{(a)} = 0$ for $a = 1, \ldots, n$. In the (pseudo) Riemannian case each scalar product $(X_{(a)}|X_{(b)})$ turns out to be constant in U because

$$\frac{\partial}{\partial x^r}(X_{(a)}|X_{(b)}) = (\nabla_{\frac{\partial}{\partial x^r}} X_{(a)}|X_{(b)}) + (X_{(a)}|\nabla_{\frac{\partial}{\partial x^r}} X_{(b)}) = 0 ,$$

where $x^1, \ldots, x^n$ are arbitrary coordinates defined on U. Hence, in the (pseudo) Riemannian case, the vector fields $X_{(1)}, \ldots, X_{(n)}$ give rise to a orthonormal basis at each point of U.
(B) As a second step, we finally prove that there is a coordinate system $y^1, \ldots, y^n$ defined in U, such that

$$(X_{(a)})_q = \frac{\partial}{\partial y^a}|_q ,$$

for every $q \in U$ and $i = 1, \ldots, n$. As a consequence, everywhere in U:

$$\Gamma^i_{jk} = \left\langle \nabla_{\frac{\partial}{\partial y^j}} \frac{\partial}{\partial y^k}, dy^i \right\rangle = \left\langle \nabla_{\frac{\partial}{\partial y^j}} X_{(k)}, dy^i \right\rangle = \left\langle 0, dy^i \right\rangle = 0\,.$$

In the (pseudo) Riemannian case these coordinates are canonical by construction: the metric takes everywhere constantly its diagonal canonical form because everywhere

$$\left(\frac{\partial}{\partial y^i} \middle| \frac{\partial}{\partial y^j} \right) = \left(X_{(i)} \middle| X_{(j)} \right)\,.$$

This will prove the thesis.

Proof of (A). The condition $\nabla X = 0$ (we omit the index $_{(a)}$ for the sake of simplicity), using a local coordinate system about p reads

$$\frac{\partial X^i}{\partial x^r} = -\Gamma^i_{rj} X^j\,.$$

Lemma 7.2 assures that a solution locally exist (with fixed initial condition) if, in a neighborhood of p,

$$-\frac{\partial \Gamma^i_{rj}}{\partial x^s} X^j + \Gamma^i_{rj} \Gamma^j_{sq} X^q$$

equals

$$-\frac{\partial \Gamma^i_{sj}}{\partial x^r} X^j + \Gamma^i_{sj} \Gamma^j_{rq} X^q$$

for all the values of i, r, s. Using the absence of torsion ($\Gamma^i_{kl} = \Gamma^i_{lk}$) and (7.4), the given condition can be rearranged into

$$R_{srj}{}^i X^j = 0\,,$$

which holds because $R = 0$ in M. To conclude, using the found result, in a sufficiently small neighborhood U of p we can define the vector fields $X_{(1)}, \ldots X_{(n)}$ as said above. Notice that thise fields are linearly independent: If there were constants $c^1, \ldots, c^n$ with

$$\left(\sum_{a=1}^n c^a X_{(a)} \right)_q = 0$$

for some $q \in U$, since $\nabla \left(\sum_{a=1}^n c^a X_{(a)} \right) = 0$ everywhere on U, we would have for $q = p$

$$\sum_{a=1}^n c^a e_a = 0$$

(indeed the equation $\nabla \left(\sum_{a=1}^n c^a X_{(a)} \right) = 0$ with initial condition $(\sum_{a=1}^n c^a X_{(a)})(q) = 0$, just in view of the validity of Frobenius conditions would admit the *unique* solution $\sum_{a=1}^n c^a X_{(a)} = 0$ everywhere un U) which is impossible by hypotheses.

Proof of (B). Fix coordinates $x^1, \ldots, x^n$ on U and, for every $q \in U$, consider the dual basis $\omega^{(1)}|_q, \ldots, \omega^{(n)}|_q \in T_q^*M$ of the basis $X_{(1)}|_q, \ldots, X_{(n)}|_q \in T_qM$. The fields $\omega^{(a)}$ satisfy $\nabla\omega^{(a)} = 0$. Indeed, for every vector field Z one has:

$$0 = Z(\delta_a^b) = \nabla_Z \langle X_{(a)}, \omega^{(b)} \rangle = \langle \nabla_Z X_{(a)}, \omega^{(b)} \rangle + \langle X_{(a)}, \nabla_Z \omega^{(b)} \rangle = 0 + \langle X_{(a)}, \nabla_Z \omega^{(b)} \rangle \,.$$

We have found that $\langle X_{(a)}, \nabla_Z \omega^{(b)} \rangle = 0$ and thus $\nabla \omega^{(b)} = 0$ because, for every vector field Z:

$$\nabla_Z \omega^{(b)} = \sum_a \langle X_{(a)}, \nabla_Z \omega^{(b)} \rangle \omega^{(a)} = 0 \,,$$

since the vectors $X_{(a)}$ form a basis with dual basis given by the co-vectors $\omega^{(a)}$. We seek for n differentiable functions $y^a = y^a(x^1, \ldots, x^n)$, $a = 1, \ldots, n$ defined on U (or in a smaller open neighborhood of p contained in U) such that

$$\frac{\partial y^a}{\partial x^i} = \omega_i^{(a)} \,, \tag{7.6}$$

for $i = 1, \ldots, n$. Once again lemma 7.2 assures that these functions exists provided

$$\frac{\partial \omega_i^{(a)}}{\partial x^r} = \frac{\partial \omega_r^{(a)}}{\partial x^i}$$

for $a, i, r = 1, \ldots, n$ in a neighborhood of p. Using 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 differentiable functions $y^a = y^a(x^1, \ldots, x^n)$, $a = 1, \ldots, n$ satisfy

$$det \left[\frac{\partial y^a}{\partial x^i} \right] \neq 0.$$

This is because, from (7.6), $det \left[\frac{\partial y^a}{\partial x^i} \right] = 0$ 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 $y^a = y^a(x^1, \ldots, x^n)$, $a = 1, \ldots, n$ define a local coordinate system about p. To conclude, we notice that

$$\langle X_{(a)}, \omega^{(b)} \rangle = \delta_a^b$$

can be re-written in view of (7.6)

$$X_{(a)}^i \frac{\partial y^b}{\partial x^j} \left\langle \frac{\partial}{\partial x^i}, dx^j \right\rangle = \delta_a^b \,,$$

that is

$$X_{(a)}^i \frac{\partial y^b}{\partial x^i} = \delta_a^b \,.$$

Therefore:

$$X^i_{(a)} = \frac{\partial x^i}{\partial y^a} \,.$$

is valid in a neighborhood of p. In other words, for each point q in a neighborhood of p,

$$(X_{(a)})_q = \frac{\partial}{\partial y^a}|_q \,.$$

This concludes the proof of (B). $\square$

7.5 Geodesic deviation and local flatness of a spacetime

As previously defined (see the example 5.1), a *spacetime* is a (C^∞) four-dimensional manifold M, equipped with a smooth metric g with Lorentzian signature $(-,+,+,+)$. A nonsingular smooth curve $M \ni \gamma = \gamma(t)$ with $t \in (a,b)$ is called *spacelike*, *timelike* or *null* (equivalently *lightlike*) if the tangent vector $\dot\gamma$ satisfies, respectively, $(\dot\gamma|\dot\gamma) > 0$, $(\dot\gamma|\dot\gamma) < 0$ and $(\dot\gamma|\dot\gamma) = 0$. *Causal curves* are those smooth curves which are piece-wisely timelike or null indifferently, these curves describes the stories of physical objects (material points) evolving in the universe. Causal geodesics are the stories of free-falling objects, i.e. experiencing the background gravitational force only. The gravitational force is described by the metric. A 3-dimensional embedded submanifold is also called **hupersurface** and it is said to be *spacelike* provided that its normal vector is timelike.
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 starting from the remark that, form 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 that a generic spacetime is (locally) flat if and only if there is no geodesic deviation for *timelike geodesics* or, equivalently, there is no geodesic deviation for *null geodesics.*

7.5.1 Collecting some useful notions and results.

Since they will be useful shortly, we recall here some properties of the curvature tensor discussed in sections 7.1, 7.2, 7.4.2. It is point-wisely defined as the unique three-linear funtional $R_p : T_pM \otimes T_pM \otimes T_pM \to T_pM$ with

$$R_p(X_p, Y_p)Z_p = (\nabla_X \nabla_Y Z)_p - (\nabla_Y \nabla_X Z)_p - (\nabla_{[X,Y]}Z)_p \,, \tag{7.7}$$

for every $p \in M$ and for every triple of C^2 vector fields X, Y, Z, where ∇ is a (C^1) affine connection on a (C^2) manifold M. The *curvature tensor* is called *Riemann tensor* when it is the curvature tensor of the Levi-Civita connection associated with a metric. The curvature tensor R fulfills the following algebraic properties, valid for every $p \in M$ and all $X_p, Y_p, Z_p \in T_pM$ which will play a role in the rest of the paper.

$$R_p(X_p, Y_p)Z_p + R_p(Y_p, X_p)Z_p = 0 \,, \tag{7.8}$$

$$R_p(X_p, Y_p)Z_p + R_p(Y_p, Z_p)X_p + R_p(Z_p, X_p)Y_p = 0 \,, \quad \text{provided } \nabla \text{ is torsionfree} \,. \tag{7.9}$$

In view of definition 4.5 we say that:

Definition 7.2. A spacetime (M, Φ) is **locally flat** if it admits a covering $\{U_i\}_{i\in\mathscr{I}}$ made of open subsets, such that every spacetime $(U_i, \Phi\restriction_{U_i})$ is isometric to a spacetime $(V_i, \eta^4\restriction_{V_i})$, V_i being an open set in Minkowski spacetime $\mathbb{M}^4$ and η^4 being the standard Minkowski metric of $\mathbb{M}^4$ (see **3** in Examples 4.1).

In view of Theorem 7.1, we have the following lemma.

Lemma 7.3. *A spacetime (M, Φ) is locally flat if and only if the Riemann tensor R of its Levi-Civita connection vanishes everywhere.*

7.5.2 Geodesic deviation and the notion of gravitation in General Relativity.

Let us finally introduce the notion of *geodesic deviation* starting with a definition.

Definition 7.3. Consider, in the spacetime (M, Φ), a pair $(\{\gamma_s\}_{s\in I}, J)$, such that $I, J \subset \mathbb{R}$ are open nonempty intervals, for every fixed $s \in I$, $J : t \mapsto \gamma_s(t)$ is a geodesic, t being a (common) affine parameter, and the map $I \times J \ni (s,t) \mapsto \gamma_s(t)$ is smooth. Defining $T := \frac{\partial}{\partial t}$ and $S := \frac{\partial}{\partial s}$, assume that

$$T_{\gamma_s(t)}\,, S_{\gamma_s(t)} \quad \text{are linearly independent and} \quad [T,S]_{\gamma_s(t)} = 0\,, \quad \text{for every } (s,t) \in I \times J\,. \tag{7.10}$$

Such a pair $(\{\gamma_s\}_{s\in I}, J)$ will be called a **smooth congruence of geodesics**.

The constraint (7.10) assures that, as is physically expected, one can adapt a coordinate system to the smooth class of geodesics, at least locally, such that two coordinates just coincide with t and s. For nonnull geodesics, t can be chosen as the proper length parameter for spacelike geodesics, or the proper time for timelike geodesics. At least when S is spacelike, $\nabla_T S$ defines the relative speed, referred to the parameter t, between infinitesimally close geodesics (say γ_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. Starting form $\nabla_T(\nabla_T S)$, employing the definition (7.7), applying (7.10), and taking the geodesic equation $\nabla_T T = 0$ into account, one finds the **geodesic deviation equation**:

$$\nabla_T(\nabla_T S) = R(S,T)T\,. \tag{7.11}$$

Let us restrict, from now on, to smooth congruences of *causal geodesics* with *spacelike* vectors S, since they have a dynamical interpretation, describing the stories of of free falling bodies, $\nabla_T(\nabla_T S)$ being the relative acceleration. The presence of tidal forces on free falling bodies, represented by the left-hand side of (7.11), cannot be canceled by means of a suitable choice of the reference frame, but it is a property of the geometry of the spacetime. Thus, the presence of *geodesic deviation for a smooth congruences of causal geodesics with spacelike S* can be used to give a sensible, relativistic, definition of gravitation, which is not affected by the *equivalence principle* 6.1.3. In locally flat spacetimes, where $R = 0$, there is no geodesic deviation – so that gravitation disappears. It is interesting to study if the absence of geodesic deviation *for causal geodesics* – i.e. the absence of gravitation – implies the local flatness of the spacetime. It is important to remark that the full information about the curvature may be obtained from the equation of geodesic deviation (7.11), if considering also smooth congruences of spacelike geodesics. However, from the experimentalist's viewpoint, $\nabla_T(\nabla_T S)$ can hardly be measured along spacelike geodesics, excluding particular cases of spacetimes as static ones, and referring to a very special choice of the field S. For this reason we stick to smooth congruences of causal geodesics with spacelike S only. The following theorem shows that, actually, geodesic deviation of timelike geodesics, or equivalently, geodesic deviation of null geodesics, encodes all information on the curvature.

Theorem 7.2. *Consider a spacetime* (M, Φ)*. The following facts are equivalent.*

(a) (M, Φ) *is locally flat;*

(b) *for every smooth congruence of geodesics* $(\{\gamma_s\}_{s\in I}, J)$ *such that,* γ_s *is a timelike geodesic and* $S_{\gamma_s(t)}$ *is spacelike* $\forall (s,t) \in I \times J$*, there is no geodesic deviation, i.e.* $(\nabla_T(\nabla_T S))_{\gamma_s(t)} = 0$*, for all* $(s,t) \in I \times J$*;*

(c) *for every smooth congruence of geodesics* $(\{\gamma_s\}_{s\in I}, J)$ *such that,* γ_s *is a null geodesic and* $S_{\gamma_s(t)}$ *is spacelike* $\forall (s,t) \in I \times J$*, there is no geodesic deviation, i.e.* $(\nabla_T(\nabla_T S))_{\gamma_s(t)} = 0$*, for all* $(s,t) \in I \times J$*.*

Proof. In view of Lemma 7.3 and of Eq.(7.11), (a) implies both (b) and (c). Let us demonstrate that (b) implies (a). The idea is to prove, making use of (b), (7.8) and (7.9), that for each point $p \in M$, it holds $R_p(X_p, Y_p)Z_p = 0$ for every choice of vectors $X_p, Y_p, Z_p \in T_pM$. This is equivalent to say that the Riemann tensor vanishes everywhere on M. At this point, Lemma 7.3 implies (a). Let us proceed step-by-step along this way. Fix $p \in M$ and assume that (b) is valid. The following lemma holds, whose proof stays at the end of this proof.

Lemma 7.4. *If* (M, Φ) *is a spacetime,* $p \in M$*, let* $T_p \in T_pM \setminus \{0\}$ *and* $S_p \in T_pM \setminus \{0\}$ *be, respectively timelike and spacelike, vectors with* $(T_p|S_p) = 0$*. There is a smooth congruence of geodesics* $(\{\gamma_s\}_{s\in I}, J)$ *as in (b) of Theorem 7.2, fulfilling* $T_{\gamma_{s_0}(t_0)} = T_p$*, and* $S_{\gamma_{s_0}(t_0)} = S_p$*, for some* $(s_0, t_0) \in I \times J$*.*

In view of Lemma 7.4 and Eq.(7.11), one has that, for every $p \in M$, $R_p(S_p, T_p)T_p = 0$ for all $T_p, S_p \in T_pM$, respectively timelike and spacelike, with $(T_p|S_p) = 0$. To extend this result to all possible arguments of R, we start noticing that $R_p(S_p, T_p)T_p = 0$ is still valid if dropping the requirements S_p spacelike and $(T_p|S_p) = 0$. Indeed, if $S_p \in T_pM$ is generic, we can decompose it as $S_p = S'_p + cT_p$, where $c \in \mathbb{R}$ and S'_p is spacelike with $(T_p|S'_p) = 0$, 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 Eq.(7.9). Summarizing, (b) implies that $R_p(S_p, T_p)T_p = 0$ for all $T_p, S_p \in T_pM$ with T_p timelike. Let us show that this last constraint can be dropped, too. To this goal, fix $S_p \in T_pM$ arbitrarily and consider the bi-linear map $T_pM \ni T_p \mapsto F_{S_p}(T_p) := R_p(S_p, T_p)T_p$. 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 $F_{S_p} \restriction_{V_p^{(+)}} = 0$ in view of the discussion above. Since F_{S_p} is analytic (it being a polynomial) and defined on the connected open domain T_pM, it must vanish everywhere on T_pM. Summarizing, we have obtained that $R_p(S_p, T_p)T_p = 0$ for every vectors $T_p, S_p \in T_pM$. To conclude, let us prove that the identity above holds true if replacing the latter T_p with a generic vector Z_p. Starting from $R_p(S_p, T_p)T_p = 0$, assuming $T_p = U_p + V_p$ and $T_p = U_p - V_p$, subtracting side-by-side the obtained results, taking bi-linearity into account, one finds:

$$R_p(S_p, U_p)V_p + R_p(S_p, V_p)U_p = 0, \tag{7.12}$$

which is valid for every $S_p, U_p, V_p \in T_pM$. Identity (7.9) can be specialized here as:

$$R_p(S_p, U_p)V_p + R_p(U_p, V_p)S_p + R_p(V_p, S_p)U_p = 0\,. \tag{7.13}$$

Summing side-by-side (7.12) and (7.13), taking Eq.(7.8) 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 as $2R_p(S_p, U_p)V_p - R_p(U_p, S_p)V_p = 0$, where we employed Eq.(7.12) (with different names of the vectors). Using Eq.(7.8) 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 $S_p, U_p, V_p \in T_pM$, so that $R_p = 0$ as wanted. This concludes the proof that (b) implies (a), in view of Lemma 7.3.

Let us finally demonstrate that (c) implies (a) by reducing to the proof of the implication (b) $\Rightarrow$ (a). Fix $p \in M$ and assume that (c) is valid. The following lemma holds, whose proof stays at the end of this proof.

Lemma 7.5. *If (M, Φ) is a spacetime, $p \in M$, let $T_p \in T_pM \setminus \{0\}$ and $S_p \in T_pM \setminus \{0\}$ be, respectively timelike and spacelike and with $(T_p|T_p) = -(S_p|S_p)$. Defining the null vectors $N^\pm := T_p \pm S_p$, there are smooth congruences of geodesics $(\{\gamma_s^\pm\}_{s\in I^\pm}, J^\pm)$ as in (c) of Theorem 7.2, fulfilling $T_{\gamma^\pm_{s_0}(t_0)} = N_p^\pm$, and $S_{\gamma^\pm_{s_0}(t_0)} = S_p$, for some $(s_0^\pm, t_0^\pm) \in I^\pm \times J^\pm$.*

In view of the lemma and of Eq.(7.11), one has that, for every $p \in M$, $R_p(S_p, N_p^\pm)N_p^\pm = 0$ for all $N_p^\pm, S_p \in T_pM$ as in the hypotheses of the lemma. Consequently, if T_p and S_p, respectively timelike and spacelike, satisfies $(T_p|T_p) = -(S_p|S_p)$, it holds:

$$R(S,T)T = R(S_p, N_p^+ - S_p)(N_p^+ - S_p) = -R_p(S_p, S_p)N_p^+ - R_p(S_p, N_p^+)S_p = -R_p(S_p, N_p^+)S_p\,,$$

where we have used Eq. (7.8), and also

$$R(S,T)T = R(S_p, N_p^- + S_p)(N_p^- + S_p) = R_p(S_p, S_p)N_p^- + R_p(S_p, N_p^-)S_p = R_p(S_p, N_p^-)S_p\,.$$

Summing the two expressions found for $R(S,T)T$ we have that, using Eq. (7.8) again:

$$2R(S,T)T = R_p(S_p, N_p^-)S_p - R_p(S_p, N_p^+)S_p = -2R_p(S_p, S_p)S_p = 0\,.$$

We have found that $R_p(S_p, T_p)T_p = 0$ for all $T_p, S_p \in T_pM$, respectively timelike and spacelike, with $(T_p|S_p) = 0$ (the requirement $(T_p|T_p) = -(S_p|S_p)$ may be dropped in view of multi-linearity of R_p). Henceforth the proof goes on as in the proof of (b) $\Rightarrow$ (a) given above.

Proof of Lemma 7.4. Take two spacelike orthogonal vectors $U_p, V_p \in T_pM$ such that they are also orthogonal with T_p and S_p and consider a normal coordinate system $D \ni (x^0, x^1, x^2, x^3) \mapsto \exp\left(x^0 T_p + x^1 S_p + x^2 U_p + x^3 V_p\right)$. $D \subset \mathbb{R}^4$ is a sufficiently small neighborhood of the origin of $\mathbb{R}^n$. Since $\frac{\partial}{\partial x^0}|_p = T_p$ is timelike, the vectors $\frac{\partial}{\partial x^0}$ have to be timelike by continuity, restricting the domain D sufficiently about the origin. With this restriction, the embedded submanifold Σ through p, defined by $(x^0, x^1, x^2, x^3) \in D$ and $x^0 = 0$, turns out to be timelike, x^1, x^2, x^3 are coordinates on Σ and $\frac{\partial}{\partial x^0}|_{(0,x^1,x^2,x^3)}$ is the (timelike) normal vector at each point of Σ. Finally consider the system of normal coordinates t, s, s^2, s^3 about Σ, defined by $G \ni (t, s, s^2, s^3) \mapsto \exp_{q(s,s^2,s^3)}\left(t\frac{\partial}{\partial x^0}|_{q(s,s^2,s^3)}\right)$, where $q(s, s^2, s^3)$ in the right-hand side indicates the point on Σ with coordinates $x^1 = s, x^2 = s^2, x^3 = s^3$ and the exponential map is that in M. The open set G is a sufficiently small neighborhood of the origin of $\mathbb{R}^4$ which, obviously, can always be taken of the form $J \times I \times I \times I$, where $I, J \subset \mathbb{R}$ are open intervals containing the origin of $\mathbb{R}$. The wanted smooth congruence of geodesics is $\gamma_s(t) := \exp_{q(s,0,0)}\left(t\frac{\partial}{\partial x^0}|_{q(s,0,0)}\right)$. The vectors $S_{\gamma_s(t)} = \frac{\partial}{\partial s}$ and $T_{\gamma_s(t)} = \frac{\partial}{\partial t}$ are linearly independent and their commutator vanishes because they are tangent to a coordinate system. $T_{\gamma_s(t)}$ is timelike, since it is the tangent vector to geodesics with timelike initial tangent vector $\frac{\partial}{\partial x^0}|_{(0,s,0,0)}$. As requested, it also trivially arises that $T_{\gamma_0(0)} = T_p$ and $S_{\gamma_0(0)} = S_p$. As S_p is spacelike, $S_{\gamma_s(t)} = \frac{\partial}{\partial s}$ has to be spacelike everywhere by continuity (shrinking I and J if necessary). $\Box$

Proof of Lemma 7.5. Starting from T_p and S_p, construct the coordinates x^0, x^1, x^2, x^3 about the spacelike hypersufrace Σ, defined by $x^0 = 0$, exactly as in the proof of Lemma 7.4. Next, at each point $q \in \Sigma$, define the null vectors $N^\pm_{q(x^1,x^2,x^3)} := \sqrt{\frac{(\partial_{x^1}|\partial_{x^1})}{(\partial_{x^0}|\partial_{x^0})}}\frac{\partial}{\partial x^0}|_{q(x^1,x^2,x^3)} \pm \frac{\partial}{\partial x^1}$. Notice that those fields are well-defined and coincides to $N^\pm_p$ if $q = p$. From now on we focus on the N^+ case only, the other case being closely similar. Since N^+ is nowhere tangent to Σ by construction, one may define a system of null-Riemannian coordinates t, s, s^2, s^3 about Σ, defined by the expression $G \ni (t, s, s^2, s^3) \mapsto \exp_{q(s,s^2,s^3)}\left(tN^+_{q(s,s^2,s^3)}\right)$, where $q(s, s^2, s^3)$ in the right-hand side indicates the point on Σ with coordinates $x^1 = s, x^2 = s^2, x^3 = s^3$ and the exponential map is that in M. The open set G is a sufficiently small neighborhood of the origin of $\mathbb{R}^4$ which, obviously, can always be taken of the form $J^+ \times I^+ \times I^+ \times I^+$, where $I^+, J^+ \subset \mathbb{R}$ are open intervals containing the origin of $\mathbb{R}$. The wanted smooth congruence of geodesics is $\gamma^+_s(t) := \exp_{q(s,0,0)}\left(tN^+_{q(s,0,0)}\right)$. The vectors $S_{\gamma^+_s(t)} = \frac{\partial}{\partial s}$ and $T_{\gamma^+_s(t)} = \frac{\partial}{\partial t}$ are linearly independent and their commutator vanishes because they are tangent to a coordinate system. $T_{\gamma^+_s(t)}$ is null, since it is the tangent vector to geodesics with null initial tangent vector $N^+_{(0,s,0,0)}$. As requested, it also trivially arises that $T_{\gamma^+_0(0)} = N^+_p$ and $S_{\gamma^+_0(0)} = S_p$. As S_p is spacelike, by continuity (shrinking I^+ and J^+ if necessary), $S_{\gamma_s(t)} = \frac{\partial}{\partial s}$ has to be spacelike everywhere it is defined. $\Box$

$\Box$