March 25, 2021

# Differential Geometry

lecture notes, summer term 2020, University of Hamburg

David Lindemann

Department of Mathematics and Center for Mathematical Physics
University of Hamburg, Bundesstraße 55, D-20146 Hamburg, Germany
david.lindemann@uni-hamburg.de

## Contents



# 1 Smooth manifolds and vector bundles

## 1.1 Basic definitions

In the field of differential geometry one is concerned with geometric objects that look locally like $\mathbb{R}^n$ for some $n \in \mathbb{N}$. In the following we will clarify exactly what this should mean and explain the reason for the term "differential" in differential geometry.

**Remark 1.1.** Recall the definition of a **topological space**. Let $M, N$ be topological spaces. A map $f : M \to N$ is called **continuous** if for all $U \subset N$ open, $f^{-1}(U) \subset M$ is open. A continuous map is called a **homeomorphism** if it has an inverse, that is if it is bijective as a map between sets, and the inverse is continuous. A **basis of the topology** of a topological space $M$ is a collection of open sets $\mathcal{B}$, so that for all $U \subset M$ open there exist an index set $I$ and corresponding open sets $B_i$ each contained in $\mathcal{B}$, such that $\cup_{i \in I} B_i = B$. Note that $I$ might be uncountable.

The study of topological spaces in full generality is not the topic of this course. We need to introduce two additional properties that topological spaces might fulfil in order to define the kind of objects we will study, that is smooth manifolds.

**Definition 1.2.** Let $M$ be a topological space. $M$ is called **Hausdorff**[1] if for any two distinct points $p, q \in M$, $p \neq q$, we can find $U, V \subset M$ open, such that $p \in U$, $q \in V$, and $U \cap V$. This means that we can separate any distinct points in $M$ with disjoint open sets. $M$ is said to fulfil the **second countability axiom** (or, simply, are second countable) if its topology has a countable basis.

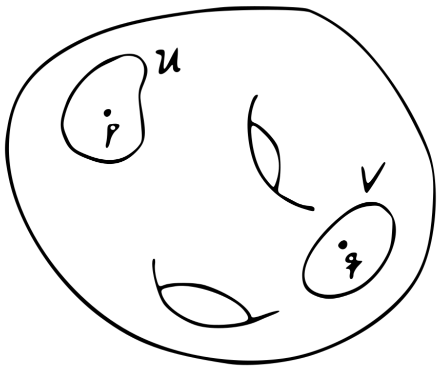


**Figure 1:** Open sets $U$ and $V$ in a Hausdorff space separating two points $p$ and $q$.

If the reader is new to general topology and the above definitions seem confusing, consider the following well known examples of Hausdorff topological spaces that are second countable. These are also basically the only examples the reader has to keep in mind for this course.

**Example 1.3.** For any $n \in \mathbb{N}_0$, $\mathbb{R}^n$ equipped with its standard topology induced by the Euclidean norm is Hausdorff and second countable. A choice for a countable basis of the topology is given by

$$\mathcal{B} := \{B_r(p) \mid r \in \mathbb{Q}_{>0},\ p \in \mathbb{Q}^n\}.$$

[Exercise: Prove that $\mathcal{B}$ is, in fact, a basis of the norm topology on $\mathbb{R}^n$.]

**Exercise 1.4.** Prove that $\mathcal{B}$ in Example 1.3 is, in fact, a basis of the norm topology on $\mathbb{R}^n$

Now that we have introduced all topological perquisites, we will give a precise meaning to the term "locally looks like" that we have used before.

**Definition 1.5.** Let $M$ be a Hausdorff topological space that is second countable. An **$n$-dimensional smooth atlas on $M$**,

$$\mathcal{A} = \{(\varphi_i, U_i) \mid i \in A\},$$

is a collection of tuples $(\varphi_i, U_i)$, each consisting of an open set $U_i \subset M$ and a homeomorphism

$$\varphi_i : U_i \to \varphi_i(U_i) \subset \mathbb{R}^n, \tag{1.1}$$

such that

[1]Felix Hausdorff (1868 – 1942)

(i) $\bigcup_{i \in A} U_i = M$, that is the $U_i$ form a covering of $M$,

(ii)
$$\varphi_i \circ \varphi_j^{-1} : \varphi_j(U_i \cap U_j) \to \varphi_i(U_i \cap U_j) \tag{1.2}$$
is smooth for all $i, j \in A$ with $U_i \cap U_j \neq \emptyset$.

Maps of the form (1.1) together with their domains are called **charts** on $M$ and the maps in (1.2) are corresponding **transition functions**. Any two charts $(\varphi_i, U_i)$, $(\varphi_j, U_j)$, not necessarily from the same atlas, are called **compatible** if the corresponding transition function $\varphi_i \circ \varphi_j^{-1}$ and its inverse are smooth.

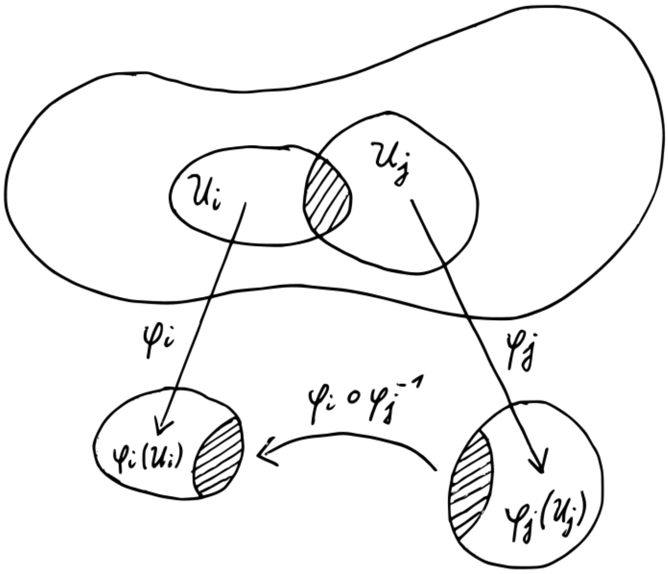


**Figure 2:** Two charts $(\varphi_i, U_i)$ and $(\varphi_j, U_j)$ with $U_i \cap U_j \neq \emptyset$.

In the following we will simply speak of atlases and drop the prefix "$n$-dimensional smooth", unless it is of specific value for a statement. Now consider the following questions. Firstly assume that you are given two different atlases $\mathcal{A}$ and $\mathcal{B}$ on $M$. What is a good notion for compatibility of these two atlases? A reasonable idea is to require that their charts are compatible in the sense of (1.2). Secondly there should always be the question whether or not there is a canonical choice for some sort of structure, in this setting that of an atlas. This leads us to the following definition:

**Definition 1.6.** Two atlases $\mathcal{A} = \{(\varphi_i, U_i) \mid i \in A\}$ and $\mathcal{B}$ $\{(\varphi_i, U_i) \mid i \in B\}$ on a second countable Hausdorff topological space $M$ are called **equivalent** if

$$\mathcal{A} \cup \mathcal{B} := \{(\varphi_i, U_i) \mid i \in A \cup B\}$$

is an atlas on $M$. This is equivalent to the requirement that the transition function $\varphi_i \circ \varphi_j^{-1}$ (1.2) for all $i, j \in A \cup B$ are smooth. For $\mathcal{A}$ and $\mathcal{B}$ equivalent we write $[\mathcal{A}] = [\mathcal{B}]$. An atlas $\mathcal{A}$ on $M$ is called **maximal** if for all atlases $\mathcal{A}'$ on $M$ equivalent to $\mathcal{A}$ it holds that $\mathcal{A}' \subset \mathcal{A}$.

Now we have all tools at hand to define the notion of a smooth manifold:

**Definition 1.7.** A second countable Hausdorff topological space $M$ together with a maximal $n$-dimensional smooth atlas $\mathcal{A}$ is called **smooth manifold** of dimension $n$.

In the following we will always assume that smooth manifolds are of dimension $n \geq 1$.

**Remark 1.8.** If one left out the requirement of second countability, the definition of a smooth manifold would still be usable for effectively every local statement about smooth manifolds. This approach is for example taken in [O]. However some global constructions might not work, in particular those involving a countable partition of unity (cf. Exercise 2.12) which might not exist. An example of an analogue of a smooth manifold that is not second countable is the so-called "long line" [SS].

We will call the process of defining a maximal atlas on $M$, defining the structure of a smooth manifold on $M$. A caveat of the above definition is that it is not in any way clear how to completely specify or write down a maximal atlas, at least not if $n > 0$. The following Lemmas 1.9 and 1.11 provide a solution for this problem.

**Lemma 1.9.** Let $\mathcal{A}$ be an atlas on a second countable Hausdorff topological space $M$. Then $\mathcal{A}$ is contained in a maximal atlas, i.e. there exists a maximal atlas $\overline{\mathcal{A}}$ on $M$, such that $\mathcal{A} \subset \overline{\mathcal{A}}$.

*Proof.* The set of atlases equivalent to $\mathcal{A}$, $\mathrm{Eq}(\mathcal{A})$, is a partially ordered set with respect to the inclusion. By Zorn's[2] lemma $\mathrm{Eq}(\mathcal{A})$ contains a maximal element $\overline{\mathcal{A}}$ which by construction is an atlas and fulfils all requirements of a maximal atlas. $\square$

**Remark 1.10.** The precise statement of Zorn's lemma is that every partially ordered set $(S, \leq)$ has a maximal element. This means that there exists $s_{\max} \in S$, such that either $s \leq s_{\max}$, or neither $s \leq s_{\max}$ nor $s \geq s_{\max}$. Note that $s_{\max}$ is in general **not unique**.

Remark 1.10 raises the question whether a maximal atlas containing any given atlas is uniquely determined. The answer is yes, and the proof feels a bit like we were cheating.

**Lemma 1.11.** Each atlas is contained in a unique maximal atlas.

*Proof.* Let $\mathcal{A} = \{(\varphi_i, U_i) \mid i \in A\}$ be an $n$-dimensional smooth atlas on a second countable Hausdorff topological space $M$. We define

$$\overline{\mathcal{A}} := \{(\varphi, U) \mid \varphi : U \to \varphi(U) \text{ is a chart on } M,\ \varphi \text{ and } \varphi_i \text{ are compatible } \forall i \in A\}.$$

We now write $\overline{\mathcal{A}} = \left\{(\varphi_i, U_i) \,\middle|\, i \in \overline{A}\right\}$ and claim that it is both a maximal atlas and unique in the stated sense. Firstly note that $\mathcal{A} \subset \overline{\mathcal{A}}$ and, hence, $\bigcup_{i \in \overline{A}} U_i = M$. Next we need to show that for any $i, j \in \overline{A}$, $\varphi_i \circ \varphi_j^{-1} : \varphi_j(U_i \cap U_j) \to \varphi_i(U_i \cap U_j)$ is smooth. Being smooth is a local property, so we fix any point $p \in \varphi_j(U_i \cap U_j)$ and choose a chart $(\varphi, U)$ in $\mathcal{A}$, such that $p \in \varphi(U)$. Then we choose $V \subset \varphi(U) \cap \varphi_j(U_i \cap U_j)$, $V \subset \mathbb{R}^n$ open, such that $p \in \varphi^{-1}(V)$, observe that

$$\varphi_i \circ \varphi_j^{-1} = (\varphi_i \circ \varphi^{-1}) \circ (\varphi \circ \varphi_j^{-1})$$

coincide on $V$. Since the right-hand-side of the above equation is a composition of by construction of $\overline{\mathcal{A}}$ smooth maps, it follows that $\varphi_i \circ \varphi_j^{-1}$ is smooth as well. This shows that $\overline{\mathcal{A}}$ is indeed an $n$-dimensional smooth atlas on $M$ and that $\mathcal{A} \subset \overline{\mathcal{A}}$. Lastly suppose that $\overline{\mathcal{A}}$ is not maximal. Then there exists an atlas $\mathcal{A}'$ on $M$ that is equivalent to $\overline{\mathcal{A}}$ and there exists a chart $(\varphi, U)$ in $\mathcal{A}'$ that is not contained in $\overline{\mathcal{A}}$. By $\mathcal{A} \subset \overline{\mathcal{A}}$ this means that even though $(\varphi, U)$ is compatible with every chart in $\mathcal{A}$ it is not contained in $\overline{\mathcal{A}}$. This is a contradiction to the construction of $\overline{\mathcal{A}}$. This shows that $\overline{\mathcal{A}}$ is maximal and finishes the proof. $\square$

[2] Max August Zorn (1906 – 1993)

**Remark 1.12.** We have seen in Lemma 1.9 that it is **sufficient** to specify an atlas $\mathcal{A}$ on a second countable Hausdorff topological space $M$ in order to define the structure of a smooth manifold on $M$ without the need of requiring maximality of $\mathcal{A}$. Furthermore, we have proven in Lemma 1.11 that there exists a **unique** maximal atlas on $M$ that is equivalent to $\mathcal{A}$, meaning that there is no possibility to choose an other structure of a smooth manifold on $M$ for which $\mathcal{A}$ is an atlas. This justifies calling a second countable Hausdorff topological space equipped with **any** atlas, be it maximal or not, a smooth manifold.

It is however not clear at this point whether for a given smooth manifold $M$ with maximal atlas $\overline{\mathcal{A}}$ there might exist some other maximal atlas $\overline{\mathcal{B}}$ on the underlying topological space $M$ that is not equivalent to $\overline{\mathcal{A}}$. This is in general a very difficult question. There are some examples where this question has been answered, see the so-called exotic spheres [M] and for 4-dimensional smooth manifolds cf. [Sc].

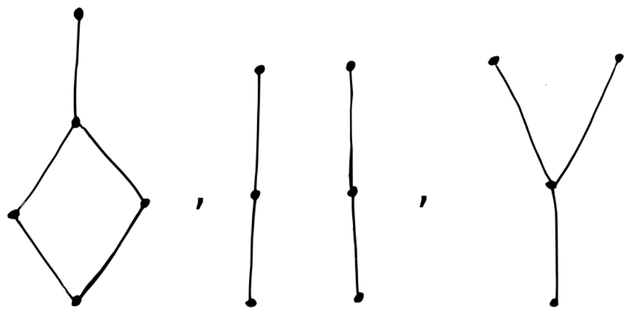

**Figure 3:** Which of the three partially ordered sets (higher means $\geq$) is a good representation for the equivalence classes of atlases?

Next we should ask ourselves how "good" we might expect a choice of an atlas for a given smooth manifold to look like, meaning an atlas contained in the by definition provided maximal atlas. Can we always choose a countable atlas, that is an atlas containing only a countable number of charts, that is equivalent to our given maximal atlas? Can we always choose a finite atlas if our manifold is connected? The answer is yes to both, but the latter is much more difficult to prove than the former, for the non-compact case see [So].

**Exercise 1.13.**

(i) Show that every smooth manifold $M$ with maximal atlas $\overline{\mathcal{A}}$ has a countable atlas that is equivalent to $\overline{\mathcal{A}}$. [Hint: Use that $M$ is second countable.]

(ii) Show that every connected compact smooth manifold with maximal atlas $\overline{\mathcal{A}}$ has a finite atlas that is equivalent to $\mathcal{A}$.

An important analytical tool that we will need in this course is shrinking the chart neighbourhoods of a given atlas.

**Definition 1.14.** Let $\mathcal{A} = \{(\varphi_i, U_i) \mid i \in A\}$ be an atlas on a smooth manifold $M$. Another atlas on $M$, $\widetilde{\mathcal{A}} = \left\{(\widetilde{\varphi}_i, \widetilde{U}_i) \;\middle|\; i \in \widetilde{A}\right\}$, is called a **refinement of** $\mathcal{A}$ if for all $i \in \widetilde{A}$ there exists $j \in A$, such that $\widetilde{U}_i \subset U_j$ and $\widetilde{\varphi}_i = \varphi_j|_{\widetilde{U}_i}$.

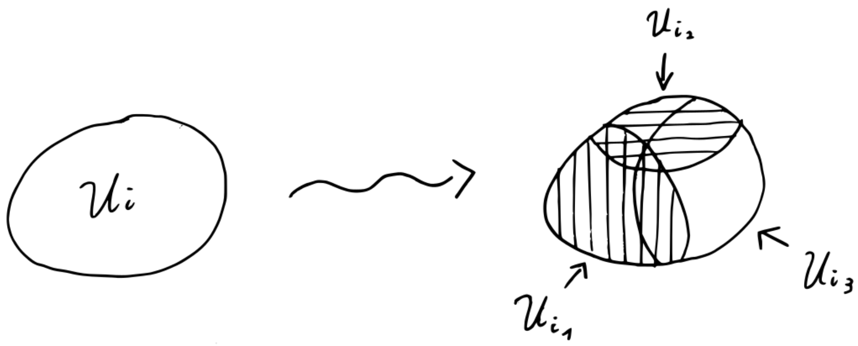


**Figure 4:** Refining a chart neighbourhood $U_i$ into three proper subsets $U_{i_1}$, $U_{i_2}$, and $U_{i_3}$.

**Exercise 1.15.** Check that any atlas is equivalent to any possible refinement of itself.

We have now finished setting up the theoretical framework for the basic definitions of smooth manifolds, so next we should study some examples.

**Example 1.16.** The probably easiest example of a smooth manifold is $\mathbb{R}^n$ equipped with the atlas containing the sole chart $(\mathrm{id}, \mathbb{R}^n)$, where id is the identity map

$$\mathrm{id} = (u^1, \dots, u^n), \quad u^i : (p_1, \dots, p_n) \mapsto p_i, \tag{1.3}$$

with domain the whole $\mathbb{R}^n$. It is also immediate that for any $U \subset \mathbb{R}^n$ open, $U$ equipped with $(\mathrm{id}, U)$ is a smooth manifold.

**Definition 1.17.** The maps $u^i$, $1 \leq i \leq n$, in (1.3) are called **canonical coordinates** on any open subset $U \subset \mathbb{R}^n$.

A similar notation is used for general smooth manifolds.

**Definition 1.18.** Let $M$ be an $n$-dimensional smooth manifold and $(\varphi, U)$ be a chart on $M$. With the notation

$$\varphi = (u^1 \circ \varphi, \dots, u^n \circ \varphi), \quad u^i \circ \varphi : U \to \mathbb{R},\ 1 \leq i \leq n,$$

the maps $x^i := u^i \circ \varphi$, $1 \leq i \leq n$, are called **local coordinates** on $M$, and $\varphi$ is called **local coordinate system**.

Now that we have setup our basic theoretical framework, it is time to look at some non-trivial examples of smooth manifolds to get a better feeling for what one needs to validate to confirm that a given space with an atlas is in fact a smooth manifold.

**Example 1.19.**

(i) Let $S^n = \{x = (x^1, \dots, x^{n+1}) \in \mathbb{R}^{n+1} \mid \|x\| = 1\}$ denote the unit $n$-sphere, equipped with the subspace topology. Let $p_\pm = (0, \dots, 0, \pm 1)$ denote the north $(+)$ and south $(-)$ pole. An atlas on $S^n$ is given by the two charts

$$\sigma_+ : S^n \setminus \{p_+\} \to \mathbb{R}^n, \quad x \mapsto \frac{x}{1 - x^{n+1}},$$
$$\sigma_- : S^n \setminus \{p_-\} \to \mathbb{R}^n, \quad x \mapsto \frac{x}{1 + x^{n+1}}.$$

The first thing to check is that the two chart cover $S^n$, which follows from $S^n \setminus p_+ \cup S^n \setminus p_- = S^n$. Next, we need to check that all transition functions are smooth maps. We find that

$$\sigma_{\pm}^{-1} = \left( \frac{2x}{1+\|x\|^2}, \pm \frac{\|x\|^2 - 1}{1+\|x\|^2} \right).$$

Further calculating yields

$$\sigma_+ \circ \sigma_-^{-1} = \sigma_- \circ \sigma_+^{-1} : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n \setminus \{0\}, \quad x \mapsto \frac{x}{\|x\|^2},$$

meaning that the transition functions coincide and are given by the **inversion on the unit** $(n-1)$**-sphere** which is self-inverse and smooth.

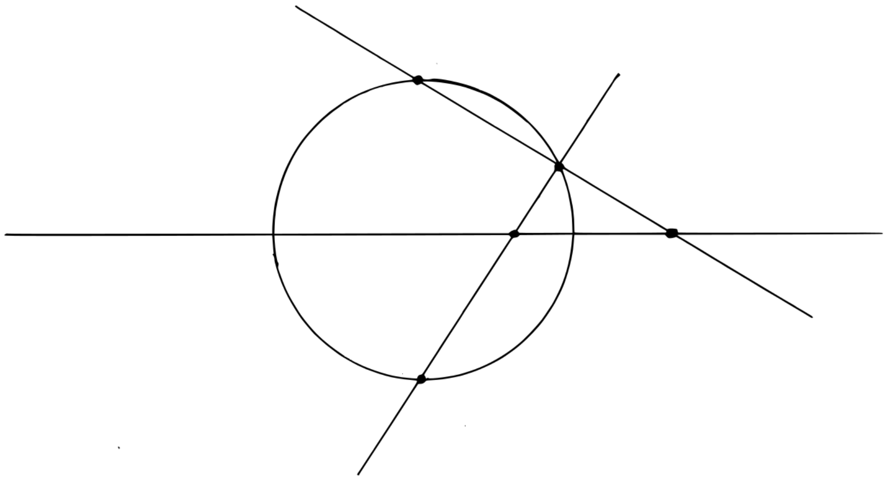

**Figure 5:** A sketch of the stereographic projection on $S^1$.

(ii) The $n$-dimensional **real projective space** $\mathbb{R}P^n$ is defined as the set of lines in $\mathbb{R}^{n+1}$. Formally, $\mathbb{R}P^n$ is the set of equivalence classes

$$\mathbb{R}P^n = \left\{ [x^1 : \ldots : x^{n+1}] \;\middle|\; x = (x^1, \ldots, x^{n+1}) \in \mathbb{R}^{n+1} \setminus \{0\} \right\},$$

where $[x] = [y]$ if $x = cy$ for some $c \in \mathbb{R} \setminus \{0\}$. This precisely means that the non-zero vectors $x$ and $y$ span the same line. One can check that $\mathbb{R}P^n$ equipped with the quotient topology induced by the canonical projection $\pi : \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{R}P^n$, $x \mapsto [x]$, [3] is Hausdorff (draw a sketch!) and second countable. An atlas on $\mathbb{R}P^n$ is given by $(\varphi_i, U_i)$, $1 \le i \le n$,

$$\begin{aligned} &\varphi_i : \pi\left(\mathbb{R}^{n+1} \setminus \{x^i = 0\}\right) \to \mathbb{R}^n, \\ &[x^1 : \ldots : x^{i-1} : x^i : x^{i+1} : \ldots : x^{n+1}] \mapsto \left( \frac{x^1}{x^i} : \ldots : \frac{x^{i-1}}{x^i} : \widehat{x^i} : \frac{x^{i+1}}{x^i} : \ldots : \frac{x^{n+1}}{x^i} \right), \end{aligned}$$

where " $\widehat{\ }$ " means that the element is supposed to be left out so that we end up with an $n$-vector. In order to check that the charts cover $\mathbb{R}P^n$ it suffices to check that

[3] The term "quotient topology" means that the open sets, in this case of $\mathbb{R}P^n$, are defined to be the images of open sets in the domain of the projection, in this case $\mathbb{R}^n \setminus \{0\}$. The notation of the elements in $\mathbb{R}P^n$ with the ":" is traditional.

$\bigcup_{1\leq i\leq n} \mathbb{R}^{n+1}\setminus\{x^i \neq 0\} = \mathbb{R}^{n+1}\setminus\{0\}$. Next we need to check that all transition functions are smooth. Observe that the range of each $\varphi_i$ is $\mathbb{R}^n$ for all $1 \leq i \leq n$, and that for all $i \neq j$

$$\varphi_j\left(\pi\left(\mathbb{R}^{n+1}\setminus\{x^i = 0\}\right) \cap \pi\left(\mathbb{R}^{n+1}\setminus\{x^j = 0\}\right)\right) = \begin{cases} \mathbb{R}^n\setminus\{x^i = 0\}, & i < j, \\ \mathbb{R}^n\setminus\{x^{i-1} = 0\}, & i > j. \end{cases}$$

Furthermore we have for all $1 \leq j \leq n$

$$\varphi_j^{-1}((x^1,\dots,x^n)) = [x^1 : \ldots : x^{j-1} : 1 : x^j : \ldots : x^n].$$

Hence we obtain for all $i < j$

$$\begin{aligned} &\varphi_i \circ \varphi_j^{-1} : \mathbb{R}^n\setminus\{x^i = 0\} \to \mathbb{R}^n\setminus\{x^{j-1} = 0\}, \\ &(x^1,\dots,x^{n+1}) \mapsto \left(\frac{x^1}{x^i},\dots,\widehat{\frac{x^i}{x^i}},\dots,\frac{x^{j-1}}{x^i},\frac{1}{x^i},\frac{x^j}{x^i},\dots,\frac{x^n}{x^i}\right), \end{aligned}$$

and for $i > j$ we find a similar formula. We see that all transition functions are indeed smooth and conclude that $\mathbb{R}P^n$ with the provided atlas is indeed a smooth manifold. The local coordinate systems $\varphi_i$ are called **inhomogeneous coordinates** on $\mathbb{R}P^n$.

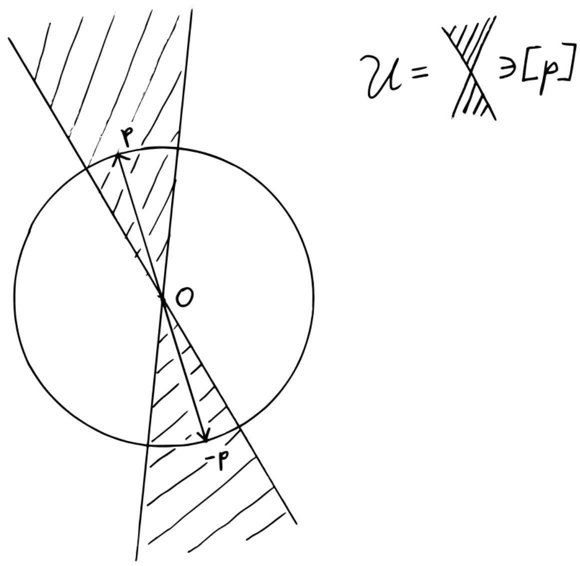


**Figure 6:** A subset $U$ of $\mathbb{R}P^1$ is a set of lines through the origin $0 \in \mathbb{R}^2$.

(iii) Let $U \subset \mathbb{R}^n$ be open and let $f : U \to \mathbb{R}$ be a smooth map. Then the **graph of** $f$, $\mathrm{graph}(f) := \{(x, f(x)) \mid x \in U\} \subset \mathbb{R}^{n+1}$ is an $n$-dimensional smooth manifold with an atlas consisting of a single chart $\varphi : \mathrm{graph}(f) \to U$, $(x, f(x)) \mapsto x$.

(iv) For a given smooth manifold $M$ with atlas $\mathcal{A} = \{(\varphi_i, U_i) \mid i \in A\}$, any open subset $U \subset M$ equipped with the restriction of the atlas $\mathcal{A}$ to $U$, $\mathcal{A}|_U := \{(\varphi_i, U_i \cap U) \mid i \in A\}$, is a smooth manifold.

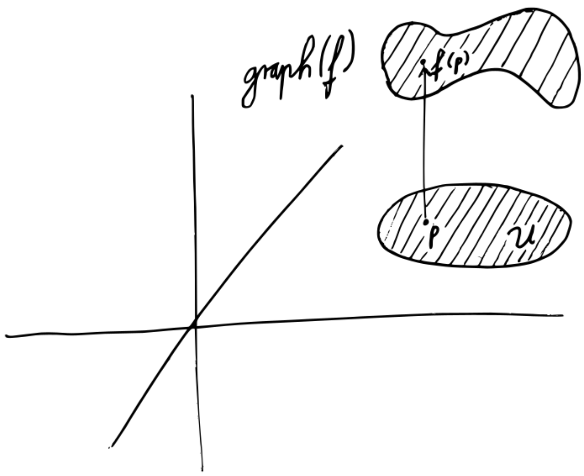


**Figure 7:** A sketch of the graph of a function $f : U \to \mathbb{R}$.

An other important class of smooth manifolds are so-called smooth submanifolds of a given smooth manifold. We will define that concept in full generality later, see Definition 1.58, but we already know from real analysis what a smooth submanifold of $\mathbb{R}^n$ is. Recall the following theorem from real analysis.

**Theorem 1.20** (Implicit Function Theorem (IFT))**.** Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, $(x, y) \mapsto f(x, y)$, be a smooth map and assume that $f(p) = 0$ for a point $p = (x_0, y_0) \in \mathbb{R}^n \times \mathbb{R}^m$ and furthermore that the Jacobi matrix of $f$ with respect to $y$ at $p$,

$$d_y f|_p = \begin{pmatrix} \frac{df_1}{dy^1}(p) & \dots & \frac{df_1}{dy^m}(p) \\ \vdots & \ddots & \vdots \\ \frac{df_m}{dy^1}(p) & \dots & \frac{df_m}{dy^m}(p) \end{pmatrix},$$

is invertible. Then there exists an open set $U \subset \mathbb{R}^n$ containing $x_0$ and an open set $V \subset \mathbb{R}^m$ containing $y_0$, such that there exists a unique smooth map $g : U \to V$ fulfilling

$$f(x, y) = 0,\ x \in U,\ y \in V \quad \Leftrightarrow \quad y = g(x).$$

In particular we have $g(x_0) = y_0$.

**Definition 1.21.** An $m < n$**-dimensional smooth submanifold of** $\mathbb{R}^n$ is a subset $M \subset \mathbb{R}^n$, such that for all $p \in M$ there exists an open set $U \subset \mathbb{R}^n$ containing $p$ and a smooth map $f : U \to \mathbb{R}^{n-m}$ with Jacobi matrix of maximal rank $n - m$ for all points in $U$ fulfilling

$$M \cap U = \{x \in U \mid f(x) = 0\}. \tag{1.4}$$

With the help of the implicit function theorem 1.20 it follows that locally up to re-ordering of coordinates on $\mathbb{R}^n$, any smooth $m < n$-dimensional submanifold $M$ of $\mathbb{R}^n$ can be written as a graph of a smooth map $g : V \to \mathbb{R}^{n-m}$, $V \subset \mathbb{R}^m$ open. This in particular implies that, after possibly reordering the coordinates on $\mathbb{R}^n$, there exists locally near every point $p$ in $M$ a smooth invertible map with smooth inverse

$$F : U \to \mathbb{R}^n, \tag{1.5}$$

$p \in U$ and $U \subset \mathbb{R}^n$ open, such that

$$F|_{U \cap M} : (x^1, \dots, x^n) \mapsto (x^1, \dots, x^m, 0, \dots, 0). \tag{1.6}$$

We call $F$ a locally defining function of $M$, which is motivated by $p \in U \cap M$ if and only if $u^{m+1}(F(p)) = \dots = u^n(F(p)) = 0$ after possibly shrinking $U$.

**Exercise 1.22.** Prove the above statements.

Now that we have recalled the definition of smooth submanifolds of $\mathbb{R}^n$, we need to ask ourselves if it is compatible with our general definition of smooth manifolds, Definition 1.7.

**Exercise 1.23.** Show that smooth submanifolds of $\mathbb{R}^n$ are smooth manifolds. [Hint: Use that the inclusion map is smooth and construct new local coordinates on the ambient space $\mathbb{R}^n$ with the help of (1.6).]

Note that there are subsets of $\mathbb{R}^n$ which are not smooth submanifolds but can still be equipped with an atlas and, hence, are smooth manifolds.

**Exercise 1.24.** Show that the boundary of the unit cube $[0,1]^n \subset \mathbb{R}^n$ is not a smooth submanifold of $\mathbb{R}^n$ but can be equipped with a smooth atlas. Find an explicit example of such an atlas.

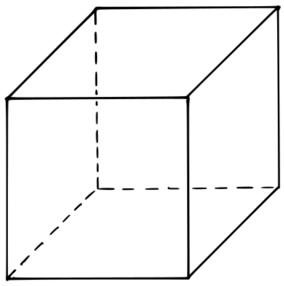

**Figure 8:** A cube.

Yet another way to produce examples of smooth manifolds are products of smooth manifolds.

**Lemma 1.25.** Let $M$ with atlas $\mathcal{A} = \{(\varphi_i, U_i) \mid i \in A\}$ be an $m$-dimensional smooth manifold and $N$ with atlas $\mathcal{B} = \{(\psi_i, V_i) \mid i \in B\}$ be an $n$-dimensional smooth manifold. Then the Cartesian product of $M$ and $N$, $M \times N$, equipped with the product topology and the product atlas $\mathcal{A} \times \mathcal{B} := \{(\varphi_i \times \psi_j, U_i \times V_j) \mid i \in A,\ j \in B\}$, is an $(m+n)$-dimensional smooth manifold.

*Proof.* This follows immediately from the definition of the product maps

$$\varphi_i \times \psi_j : U_i \times V_j \to \varphi_i(U_i) \times \psi_j(V_j) \subset \mathbb{R}^m \times \mathbb{R}^m \cong \mathbb{R}^{n+m}, \quad (p,q) \mapsto (\varphi_i(p), \psi_j(q)).$$

$\square$

**Exercise 1.26.** Show that under the additional assumption that $\mathcal{A}$ and $\mathcal{B}$ are maximal in Lemma 1.25, the product atlas $\mathcal{A} \times \mathcal{B}$ is not necessarily maximal.

We now know what a smooth manifold is and we have seen some examples and counterexamples. Next we will define smooth maps between manifolds. In the language of category theory, these are the the homomorphism in the category of smooth manifolds.

**Definition 1.27.** Let $M$ and $N$ be smooth manifolds of dimension $m = \dim(M)$ and $n = \dim(N)$. A continuous map $f : M \to N$ is called **smooth** if for all charts $(\varphi, U)$ of $M$, $(\psi, V)$ of $N$, the map

$$\psi \circ f \circ \varphi^{-1} : \varphi(U \cap f^{-1}(V)) \to \psi(V), \tag{1.7}$$

is a smooth map between open sets in $\mathbb{R}^n$ and $\mathbb{R}^m$. By the term "$f$ in local coordinates" we mean exactly the above formula (1.7) for a choice of charts $(\varphi, U)$ and $(\psi, V)$.

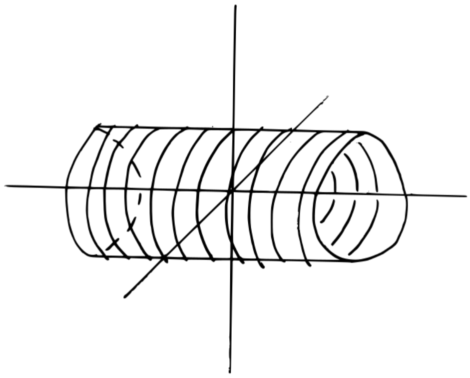

**Figure 9:** A rough sketch of the cylinder $S^1 \times (-1, 1)$.

**Definition 1.28.** By $C^\infty(M)$ we denote the **$\mathbb{R}$-vector space of smooth $\mathbb{R}$-valued functions** on a smooth manifold $M$, that is all smooth maps $f : M \to \mathbb{R}$ in the sense of Definition 1.27. If $U \subset M$ is open, $U$ is by restriction of the atlas of $M$ a smooth manifold itself. By $C^\infty(U)$ we denote all smooth functions $f : U \to \mathbb{R}$

**Exercise 1.29.** Show that for $U \subset M$ open with $U \neq M$, the restriction map $C^\infty(M) \to C^\infty(U)$, $f \mapsto f|_U$, is in general not surjective (find a counterexample). Ask yourself what kind of difficulties might arise if one wanted to prove that the restriction map for any such $U$, $M$, is never surjective.

**Example 1.30.** An example of smooth functions on open subsets of a smooth manifold $M$ are the local coordinates $x^i : U \to \mathbb{R}$ of a given chart $(\varphi, U)$, cf. Definition 1.18. This is the reason why the $x^i$ are also called **local coordinate functions**.

**Definition 1.31.**

(i) Let $M$, $N$ be smooth manifolds. A smooth map $f : M \to N$ is called a **diffeomorphism** if it is invertible and its inverse is smooth.

(ii) Two smooth manifolds $M$ and $N$ are called **diffeomorphic** if there exists a diffeomorphism $f : M \to N$.

**Remark 1.32.** There exist no two diffeomorphic smooth manifolds with different dimensions. This follows from the fact that every diffeomorphism is automatically a homeomorphism of the underlying topological spaces and, hence, locally a homeomorphism between open sets in $\mathbb{R}^m$ and $\mathbb{R}^n$. It follows from [Bro] that then $m = n$.

**Exercise 1.33.**

(i) Show that $S^1$ and $\mathbb{R}P^1$ are diffeomorphic.

(ii) Let $M$ be a second countable Hausdorff topological space and let $\overline{\mathcal{A}}$ and $\overline{\mathcal{B}}$ be inequivalent maximal atlases on $M$. Prove that $M$ equipped with $\overline{\mathcal{A}}$ is not diffeomorphic to $M$ equipped with $\overline{\mathcal{B}}$.

## 1.2 Tangent spaces and differentials

So far we have introduced the "geometric" part of differential geometry in the sense that we have learned what the objects are that we will be studying, namely smooth manifolds. We have however not made sense of the "differential" part yet, which is what we will do next.

**Remark 1.34.** Recall the definition of tangent vectors in $\mathbb{R}^n$ that you know from real analysis. A tangent vector at a point $p \in \mathbb{R}^n$ is defined to be an equivalence class of smooth curves through $p$, $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$, $\gamma(0) = p$, where

$$[\gamma] = [\widetilde{\gamma}] \quad :\Leftrightarrow \quad \gamma'(0) = \widetilde{\gamma}'(0)$$

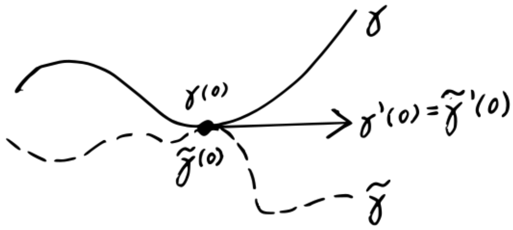


**Figure 10:** Two curves $\gamma, \widetilde{\gamma}$ with $\gamma(0) = \widetilde{\gamma}(0)$ that are in the same class.

So-defined tangent vectors act on locally near $p$ defined smooth functions $f \in C^\infty(U)$, $p \in U$, $U \subset \mathbb{R}^n$ open, via

$$[\gamma]f := \left.\frac{d(f \circ \gamma)}{dt}\right|_{t=0} \tag{1.8}$$

which is precisely the directional derivative of $f$ at $p$ in the direction $\gamma'(0)$. Note that the value of $[\gamma]f$ depends, aside from $\gamma'(0)$, only on the values of $f$ on an arbitrary small open neighbourhood of $p$ in $\mathbb{R}^n$.

Furthermore recall that a tangent vector $[\gamma]$ at $p$ is called tangential to a smooth $m < n$-dimensional submanifold $M$ of $\mathbb{R}^n$ if for any locally defining function $F : M \cap U \to \mathbb{R}^n$, cf. equations (1.5) and (1.6), we have

$$dF_p \cdot \gamma'(0) \in \mathbb{R}^m \times \{0\}.$$

In the above equation, $dF_p$ denotes the Jacobi matrix of $F$ at $p$ and the the statement of the equation just means that the last $n - m$ entries of $dF_p \cdot \gamma'(0)$ all vanish. Equivalently, $[\gamma]$ is tangential to $M$ if it fulfils

$$df_p \cdot \gamma'(0) = 0$$

for a smooth map $f : U \to \mathbb{R}^{n-m}$ with Jacobi matrix of maximal rank with $M \cap U = \{x \in U \mid f(x) = 0\}$ for some open neighbourhood $U \subset \mathbb{R}^n$ of $p$.

The tangent space of $\mathbb{R}^n$ at $p \in \mathbb{R}^n$ is the collection of all tangent vectors at $p$ and isomorphic to $\{p\} \times \mathbb{R}^n \cong \mathbb{R}^n$. The tangent space of $\mathbb{R}^n$ is the disjoint union of the tangent spaces at all points and is thereby given by $\mathbb{R}^{2n} \cong \mathbb{R}^n \times \mathbb{R}^n$. An element $(p, v)$ in the tangent space has a base point $p$ and a direction $v$ which is the tangent vector.

We want to define tangent vectors and the tangent space for general smooth manifolds. The constructions should coincide (i.e. be isomorphic in some sense to be explained later, cf. Lemma **??**) with the above definition when considered for $\mathbb{R}^n$ viewed as a smooth manifold.

make Lemma: $J(\mathbb{R}, M) \cong TM$

**Definition 1.35.** Let $M$ be a smooth manifold. A **tangent vector** $v$ **at** $p \in M$ is a linear map

$$v : C^\infty(M) \to \mathbb{R}$$

that fulfils the Leibniz rule[4]

$$v(fg) = g(p)v(f) + f(p)v(g)$$

for all $f, g \in C^\infty(M)$.

[4] Compare this to the Leibniz rule you know from real analysis!

The set of tangent vectors at any fixed point $p \in M$ forms a real vector space:

**Definition 1.36.** The **tangent space at** $p \in M$,

$$T_pM := \{v : C^\infty(M) \to \mathbb{R} \mid v \text{ tangent vector at p}\},$$

is the real vector space of all tangent vectors $v$ at $p \in M$. This means that $cv$ and $v + w$ are tangent vectors for all $c \in \mathbb{R}$, $v, w \in T_pM$, with

$$(cv)(f) = c \cdot v(f), \quad (v + w)(f) = v(f) + w(f),$$

for all $f \in C^\infty(M)$.

A central property of tangent vectors at $p \in M$ is that for any $f \in C^\infty$, $v(f)$ depends only on the values of $f$ on an arbitrary small open neighbourhood of $p$. In order to prove this statement in Proposition 1.42 we will need some technical tools.

**Definition 1.37.** A **smooth partition of unity** of a smooth manifold $M$ is a set of smooth functions on $M$

$$\{f_i : M \to [0, 1] \mid i \in I\},$$

where $I$ is an index set (e.g. $\mathbb{N}$ or $\mathbb{R}$), such that for all $x \in M$

$$\sum_{i \in I} f_i(x) = 1.$$

A smooth partition of unity is called **locally finite** if

$$\{i \in I \mid f_i(x) \neq 0\}$$

is finite for all $x \in M$. If $\{U_i \subset M \mid i \in I\}$ is an open cover of $M$ and $\mathrm{supp}(f_i) = \overline{\{x \in M \mid f_i(x) \neq 0\}}$ fulfils

$$\mathrm{supp}(f_i) \subset U_i$$

for all $i \in I$, then the smooth partition of unity is called **subordinate to the open cover** $\{U_i \subset M \mid i \in I\}$.

**Proposition 1.38.** Let $M$ be a smooth manifold and $\{U_i,\ i \in I\}$ an open cover of $M$. Then there exists a locally finite countable partition of unity on $M$ subordinate to the open cover $\{U_i,\ i \in I\}$.

*Proof.* Exercise. [Hint: You might use your knowledge from real analysis and assume that the statement of this proposition is true for $M = \mathbb{R}^n$, $n \geq 1$. Also recall the existence of a countable atlas on any given manifold that is equivalent to the defining maximal atlas, see Exercise 1.13.] □

**Definition 1.39.** Let $M$ be a smooth manifold and $U \subset M$ open. Let $V$ be a subset of $M$ with non-empty interior that is compactly embedded in $U$. Then a **bump function** with respect to the given data is a compactly supported smooth function $b \in C^\infty(M)$, such that

$$b|_{\overline{V}} \equiv 1, \quad \mathrm{supp}(b) \subset U. \tag{1.9}$$

**Proposition 1.40.** Let $M$, $U$, $V$ be as in Definition 1.39 arbitrary but fixed. Then there exists a bump function $b$ fulfilling (1.9).

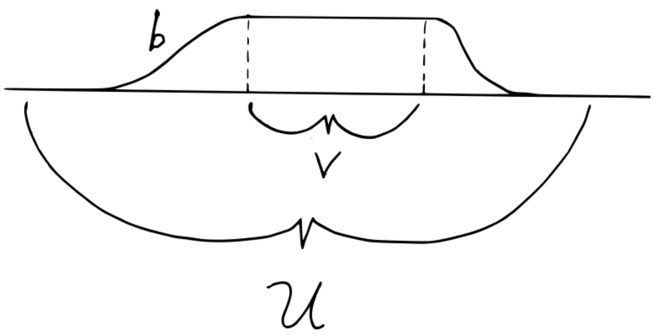


**Figure 11:** A bump function $b$ w.r.t. $V$ and $U$.

*Proof.* We know from real analysis that this statement is true for $M = \mathbb{R}^n$. By using Proposition 1.38 it follows for arbitrary smooth manifolds as well. $\square$

**Exercise 1.41.**

(i) Fill in the details of the proof of Proposition 1.40.

(ii) Let $U \subset M$ be any open subset of a smooth manifold $M$, $V \subset U$ a compactly embedded set with non-empty interior, and $b \in C^\infty(M)$ a bump function with respect to this data. Let $F : U \to \mathbb{R}^n$, $n \geq 1$, be a smooth map. Show that

$$bF : M \to \mathbb{R}^n, \quad (bF)(p) = b(p)F(p) \ \forall p \in U, \ (bF)(p) = 0 \ \forall p \in M \setminus U$$

is smooth (the above globally on $M$ defined map is called the **trivial extension** of $bF : U \to \mathbb{R}^n$ to $M$).

Now that we can use the existence of bump functions on smooth manifolds, we can continue our study of tangent vectors.

**Proposition 1.42.** Let $v \in T_pM$ be any tangent vector.

(i) Let $f, g \in C^\infty(M)$ and assume that for som open neighbourhood $U \subset M$ of $p \in M$, $f|_U = g|_U$. Then $v(f) = v(g)$.

(ii) Let $f \in C^\infty(M)$ be a smooth function that is locally constant near $p \in M$, meaning that there exists an open neighbourhood $U \subset M$ of $p$, such that $f|_U \equiv c$ for some $c \in \mathbb{R}$. Then $v(f) = 0$.

*Proof.* By the linearity of tangent vectors we have $v(f) = v(g)$ if and only if $v(f - g) = 0$. Thus, in order to prove (i) it suffices to show that if $v(f) = 0$ for some $f \in C^\infty(M)$ then $f$ must already vanish near $p$, meaning that $f|_U \equiv 0$ for some open neighbourhood $U \subset M$ of $p$. Let $V \subset U$ be open and compactly embedded in $U$ with $p \in V$ and fix a bump function $b \in C^\infty(M)$, such that

$$b|_V \equiv 1, \quad \operatorname{supp}(b) \subset U.$$

Then $bf \equiv 0$ on $M$. By using the Leibniz rule for tangent vectors and $v(0) = 0$ by the linearity of $v$ we obtain

$$0 = v(0) = v(bf) = f(p)v(b) + b(p)v(f) = 0 + v(f).$$

Hence, $v(f) = 0$ as claimed.

We can now use (i) and find that for any locally constant function $f|_U \equiv c$ for some open neighbourhood $U \subset M$ of $p$, the value of $v(f)$ is the same as $v(c)$, where we view $c \in \mathbb{R}$ as the constant function on $M$ with value $c$. We calculate

$$v(f) = v(c) = cv(1) = cv(1 \cdot 1) = c(1 \cdot v(1) + 1 \cdot v(1)) = 2cv(1) = 2v(f).$$

This shows that $v(f) = 0$. □

**Remark 1.43.** Proposition 1.42 shows that the action of tangent vectors at a point only depends on the local form of the functions near that point. This is sometimes phrased as "tangent vectors are local objects" and allows us to define the action of tangent vectors on functions that are only defined locally. Let $v \in T_pM$, $U \subset M$ an open neighbourhood of $p$, and let $f \in C^\infty(U)$. Since the the action of $v$ on globally defined functions only depends on their behaviour near $p$, it is reasonable to define

$$v : f \mapsto v(bf),$$

for any bump function $b$ with $p$ contained in the interior of its support, so that $\operatorname{supp}(b) \subset U$ and such that there exists $V \subset U$ compactly embedded with nonempty interior fulfilling $p \in V$, $V \subset \operatorname{supp}(b)$, and $b|_{\overline{V}} \equiv 1$. Note that $v(bf)$ does not depend on the choice of such a bump function $b$. On the other hand, any open subset $U \subset M$ is a smooth manifold itself by restricting any atlas on $M$. This means that for any $p \in U$, the tangent space $T_pU$ is well-defined. For any $\widetilde{v} \in T_pU$, we can define its action on $C^\infty(M)$ by

$$\widetilde{v}(f) := \widetilde{v}(f|_U).$$

By Proposition 1.42 we know that $\widetilde{v}(f)$ does in fact only depend on the behaviour of $f$ on *any* open neighbourhood of $p$ in $U$, which is then automatically an open neighbourhood of $p$ in $M$. This means that we can canonically identify $T_pM$ and $T_pU$ for all $U \subset M$ open and all $p \in U$. From now on we will simply write $v(f)$ for $v \in T_pM$ and $f \in C^\infty(U)$ any locally defined smooth function with $p \in U$.

The above motivates a slightly different definition of tangent vectors in $T_pM$ as linear maps on the **germ of smooth functions at** $p \in M$

$$\mathcal{F}_p := \{f \in C^\infty(U) \mid p \in U,\ U \subset M \text{ open}\}/_\sim$$

where $f \sim g$ if and only if there exists $U \in M$ open and contained in the domain of definition of both $f$ and $g$, such that $f|_U = g|_U$. The notion of a germ comes from sheaf theory. For an introduction see [Bre]. One can show that $\mathcal{F}_p$ is an $\mathbb{R}$-algebra and then define tangent vectors $v \in T_pM$ as linear maps

$$v : \mathcal{F}_p \to \mathbb{R}$$

fulfilling the Leibniz rule $v([f][g]) = g(p)v([f]) + f(p)v([g])$. This definition is equivalent to our definition with the same reasoning as for why $T_pM$ and $T_pU$ for $U \subset M$ open can be identified. This approach is used in [G].

While we have explained how a tangent vector should behave and have seen that it depends only on the local behaviour of functions, we do not yet have a convenient way to write down actual examples of tangent vectors. For that we introduce specific tangent vectors that generalize partial derivatives we know from real analysis to smooth manifolds in any given local coordinate system.

**Definition 1.44.** Let $\varphi = (x^1, \dots, x^n)$ be a local coordinate system on a smooth manifold $M$. The tangent vector $\frac{\partial}{\partial x^i}\Big|_p \in T_pM$ is defined as

$$\frac{\partial}{\partial x^i}\bigg|_p (f) := \frac{\partial f}{\partial x^i}(p) := \frac{\partial (f \circ \varphi^{-1})}{\partial u^i}(\varphi(p))$$

for all $f \in C^\infty(M)$.

**Lemma 1.45.** $\left.\frac{\partial}{\partial x^i}\right|_p$ is a well-defined tangent vector.

*Proof.* Linearity follows from the fact that partial derivatives in $\mathbb{R}^n$ are linear with respect to scalar multiplication. For the Leibniz rule we recall that partial differentiation in $\mathbb{R}^n$ fulfils the Leibniz rule and calculate for any two $f, g \in C^\infty(M)$

$$\begin{aligned}\frac{\partial(f\cdot g)}{\partial x^i}(p) &= \frac{\partial((f\cdot g)\circ\varphi^{-1})}{\partial u^i}(\varphi(p)) = \frac{\partial((f\circ\varphi^{-1})\cdot(g\circ\varphi^{-1}))}{\partial u^i}(\varphi(p)) \\ &= g(p)\frac{\partial(f\circ\varphi^{-1})}{\partial u^i}(\varphi(p)) + f(p)\frac{\partial(g\circ\varphi^{-1})}{\partial u^i}(\varphi(p)) = g(p)\frac{\partial f}{\partial x^i}(p) + f(p)\frac{\partial g}{\partial x^i}(p).\end{aligned}$$

□

Important examples of $\left.\frac{\partial}{\partial x^i}\right|_p$ acting on smooth function are derivatives of coordinate functions.

**Example 1.46.** Let $\varphi = (x^1, \ldots, x^n)$ be a local coordinate system on a smooth manifold $M$ covering $p \in M$. Then

$$\frac{\partial x^j}{\partial x^i}(p) = \delta_i^j$$

for all $1 \le i \le n$, $1 \le j \le n$. This follows from $(x^j \circ \varphi^{-1})(u^1, \ldots, u^n) = u^j$.

Using Definition 1.44 we can write down any tangent vector $v \in T_pM$ in a fixed local coordinate system $(\varphi, U)$ with $p \in U$ as a linear combinations of the $\left.\frac{\partial}{\partial x^i}\right|_p$'s.

**Proposition 1.47.** For all $p \in M$ and any local chart $(\varphi = (x^1, \ldots, x^n), U)$ with $p \in U$, the set of tangent vectors $\left\{\left.\frac{\partial}{\partial x^i}\right|_p,\ 1 \le i \le n\right\}$ is basis of $T_pM$.

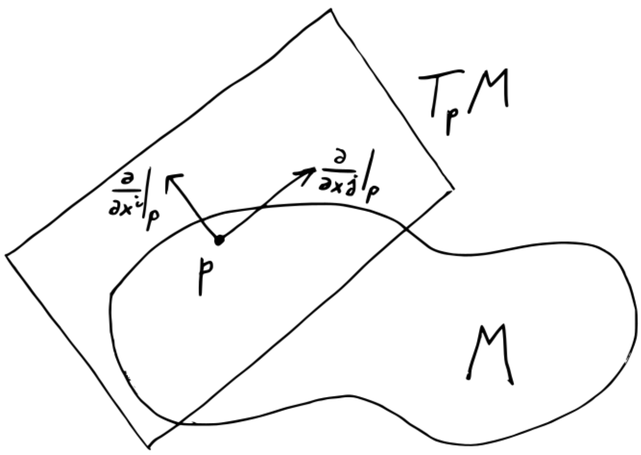


**Figure 12:** The tangent space $T_pM$ at $p \in M$ is the linear span of the $\left.\frac{\partial}{\partial x^i}\right|_p$'s.

*Proof.* First we show that the set of the $\left.\frac{\partial}{\partial x^i}\right|_p$'s is a linearly independent set of tangent vectors at $p$. Suppose that there exists $(c^1, \ldots, c^n) \in \mathbb{R}^n \neq 0$, such that

$$v_0 := \sum_{i=1}^n c^i \left.\frac{\partial}{\partial x^i}\right|_p$$

vanishes identically as a linear map $v_0 : C^\infty(M) \to \mathbb{R}$. By assumption there exists at least one $1 \le j \le n$, such that $c^j \neq 0$. But then

$$v_0(x^j) = c^j \neq 0$$

which is a contradiction to $v_0 = 0$.

Next we need to show that every tangent vector $T_pM$ can be written as a linear combination of the $\frac{\partial}{\partial x^i}\Big|_p$'s. Assume without loss of generality that $\varphi(U) = B_r(0)$ for some $r > 0$ with $\varphi(p) = 0$, that is the Euclidean unit ball at the origin with radius $r$. This can always be achieved by shrinking $U$ and translating $\varphi(U)$ if necessary. For any smooth function $g$ on $\varphi(U)$ it follows from the fundamental theorem of calculus[5] that with

$$g_i(q) := \int_0^1 \frac{\partial g}{\partial u^i}(tq)\, dt$$

for all $q \in \varphi(U)$ we have

$$g = g(0) + \sum_{i=1}^n g_i u^i$$

on $\varphi(U)$. In particular, we obtain for any $f \in C^\infty(U)$ with $g = f \circ \varphi^{-1}$

$$f = g \circ \varphi = f(p) + \sum_{i=1}^n f_i x^i$$

where $f_i = g_i \circ \varphi$. By acting with the tangent vector $\frac{\partial}{\partial x^i}\Big|_p$, $1 \le i \le n$, on both sides of the above equation we obtain[6]

$$f_i(p) = \frac{\partial f}{\partial x^i}(p).$$

Hence, we get using $x^i(p) = 0$ for $1 \le i \le n$ for $v \in T_pM$ fixed

$$v(f) = 0 + \sum_{i=1}^n \Big( v(f_i)x^i(p) + f_i(p)v(x^i) \Big) = \sum_{i=1}^n \frac{\partial f}{\partial x^i}(p) v(x^i).$$

Since $f$ was arbitrary this shows that the tangent vectors $v$ and $\sum_{i=1}^n v(x^i) \frac{\partial}{\partial x^i}\Big|_p$ coincide. Hence, $v$ can be written as a linear combination of the proposed basis vectors. This finishes the proof. □

**Corollary 1.48.** The dimensions of a smooth manifold $M$ and its tangent space $T_pM$ coincide for all $p \in M$.

**Example 1.49.** As an example how the coordinate tangent vectors $\frac{\partial}{\partial x^i}\Big|_p$ change for different coordinates consider the following example. Let $f \in C^\infty(M)$ be any smooth function and $(x^1, \dots, x^n)$ be local coordinates covering $p \in M$. Then $(y^1, \dots, y^n) := (2x^1, x^2, \dots, x^n)$ are also local coordinates covering $p$. The vectors $\frac{\partial}{\partial x^i}\Big|_p$ and $\frac{\partial}{\partial y^i}\Big|_p$ coincide for $2 \le i \le n$, but for $i = 1$ we have

$$\frac{\partial}{\partial x^1}\Big|_p = 2\, \frac{\partial}{\partial y^1}\Big|_p.$$

See Figure 13 for a sketch of this example.

[5] If you do not see this, apply the fundamental theorem of calculus to $t \mapsto g(tq)$ for $q \in \varphi(U)$ fixed.

[6] Verifying this is a good exercise.

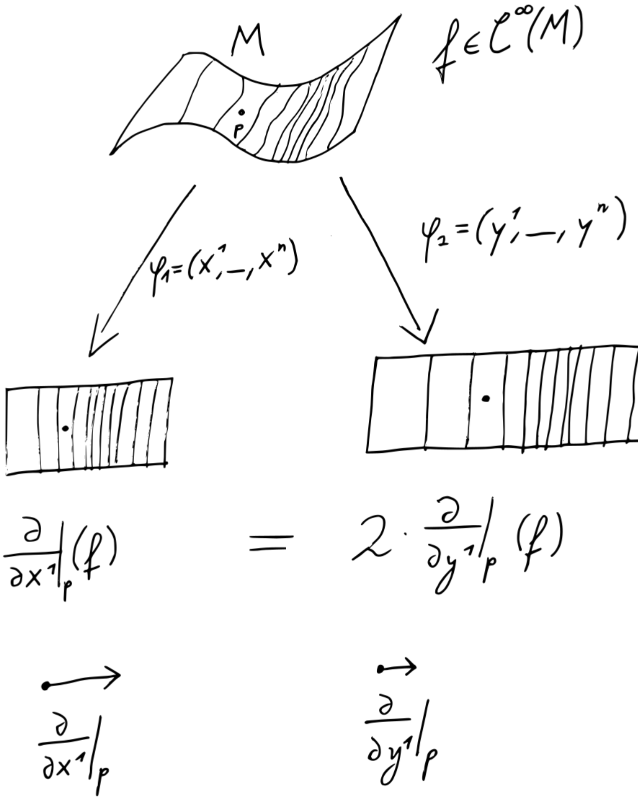


**Figure 13:** The lines in $M$, respectively the images of the charts, are supposed to be level sets of $f$.

We now know the properties of tangent vectors and how they can be written locally, meaning that we can now properly calculate with them in fixed local coordinates. This allows us to define an analogue of the Jacobi matrix for smooth manifolds.

**Definition 1.50.** Let $M$ be a smooth manifold of dimension $m$ and $N$ be a smooth manifold of dimension $n$.

(i) The **differential at a point** $p \in M$ **of a smooth function** $f \in C^\infty(M)$ is defined as the linear map

$$df_p : T_pM \to \mathbb{R}, \quad v \mapsto v(f).$$

In a given local coordinate system $\varphi = (x^1, \ldots, x^m)$ on $M$ that covers $p$, $df_p$ is of the form

$$df_p : \left.\frac{\partial}{\partial x^i}\right|_p \mapsto \frac{\partial f}{\partial x^i}(p).$$

(ii) The **differential at a point** $p \in M$ **of a smooth map** $F : M \to N$ in given local coordinate systems $\varphi = (x^1, \ldots, x^m)$ on $M$ and $\psi = (y^1, \ldots, y^n)$ on $N$ covering $p \in M$ and $F(p) \in N$, respectively, is defined as the linear map

$$dF_p : T_pM \to T_{F(p)}N, \quad \left.\frac{\partial}{\partial x^i}\right|_p \mapsto \sum_{j=1}^{n} \frac{\partial F^j}{\partial x^i}(p) \left.\frac{\partial}{\partial y^j}\right|_{F(p)},$$

where we have used the notation

$$F^j := y^j \circ F.$$

The **rank** of $F$ at $p$ is the rank of the linear map $dF_p : T_pM \to T_{F(p)}N$, which coincides with the rank of the **Jacobi matrix of** $F$ **at** $p$ in the local coordinate systems $\varphi$, $\psi$,

$$\left(\frac{\partial F^j}{\partial x^i}(p)\right)_{ji} \in \mathrm{Mat}(n \times m, \mathbb{R}).$$

In the above equation, $j$ is the row and $i$ is the column of the matrix.

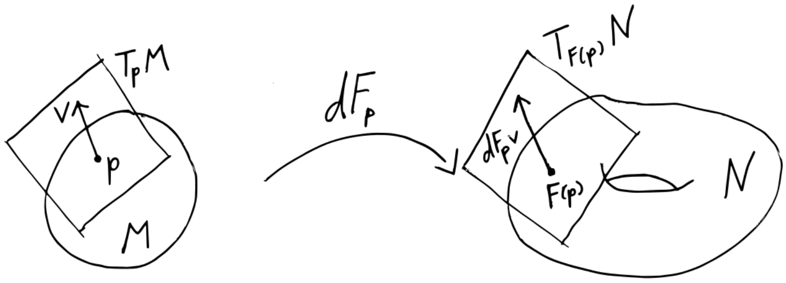


**Figure 14:** Sketch of the differential at $p \in M$ of a smooth map $F : M \to N$.

**Example 1.51.** Let $\varphi = (x^1, \ldots, x^n)$ be a local coordinate system on a smooth manifold $M$ covering $p \in M$. Then $d\varphi_p : T_pM \to T_{\varphi(p)}\mathbb{R}^n \cong \mathbb{R}^n$ is of the form

$$d\varphi_p = (dx^1, \ldots, dx^n)_p = (dx^1_p, \ldots, dx^n_p), \quad dx^j_p\left(\left.\frac{\partial}{\partial x^i}\right|_p\right) = \delta^j_i \quad \forall 1 \le i, j \le n.$$

We will usually omit the base point and simply write $dx^i := dx^i_p$ if it is clear from either the context or the tangent vector's base point that $dx^i$ acts on.

**Remark 1.52.** Note that Definition 1.50 (i) is a special case of (ii) (using the canonical coordinate $u^1$ on $\mathbb{R}$).

Similar to real analysis, the differential of smooth maps between smooth manifolds fulfils the following chain rule.

**Lemma 1.53.** Let $M, N, P$ be smooth manifolds and $F : M \to N$, $G : N \to P$, smooth maps. Then

$$d(G \circ F)_p = dG_{F(p)} \circ dF_p$$

for all $p \in M$.

*Proof.* For any $v \in T_pM$ and $f \in C^\infty(P)$ we have

$$d(G \circ F)_p(v)(f) = v(f \circ G \circ F) = dF_p(v)(f \circ G) = dG_{F(p)}(dF_p(v))(f).$$

□

**Definition 1.54.**

(i) A smooth map between smooth manifolds $F : M \to N$ is called an **immersion** if $dF_p : T_pM \to T_{F(p)}N$ is injective for all $p \in M$.

(ii) $F : M \to N$ is called a **submersion** if $dF_p : T_pM \to T_{F(p)}N$ is surjective for all $p \in M$.

(iii) An immersion $F : M \to N$ is called an **embedding** if $F$ is injective and an homeomorphism onto its image $F(M) \subset N$ equipped with the subspace topology.

(iv) A smooth map $F : M \to N$ between smooth manifolds of the same dimension is called a **local diffeomorphism** if for all $p \in M$ there exists an open neighbourhood of $p$, $U \subset M$, such that $F|_U : U \to N$ is a diffeomorphism onto its image.

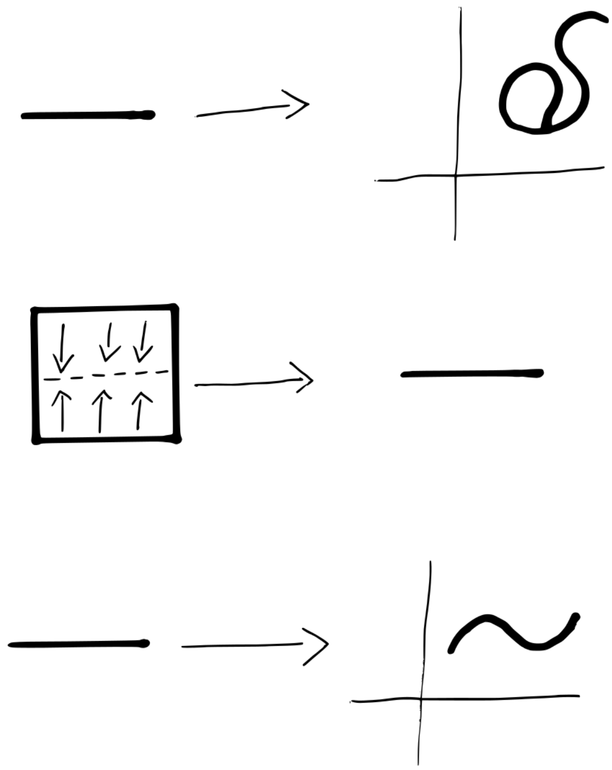

**Figure 15:** An immersion, an embedding, and a submersion. Which is which?

Suppose that we are given a smooth map $F : M \to N$, $\dim(M) = \dim(N)$, and want to check if it is a local diffeomorphism. At first this sounds fairly complicated, but luckily we can use the following result.

**Theorem 1.55.** Let $F : M \to N$ be a smooth map between two smooth manifolds of the same dimension $n$ and let $p \in M$ be arbitrary. Then $dF_p : T_pM \to T_{F(p)}N$ is a linear isomorphism if and only if there exists an open neighbourhood $U \subset M$ of $p$, such that $F|_U$ is a diffeomorphism onto its image.

*Proof.* Let $(\varphi, U)$ and $(\psi, V)$ be local charts covering $p \in M$ and $F(p) \in N$, respectively. Observe that, by definition, $dF_p$ is a linear isomorphism if any only if its Jacobi matrix in any given local coordinates is invertible. On the other hand, there exists an open neighbourhood $U \subset M$ of $p$, such that $F|_U$ is a diffeomorphism onto its image if and only if there exist open sets $U', V' \subset \mathbb{R}^n$ with $\varphi(p) \in U'$, $\psi(F(p)) \in V'$, such that

$$\psi \circ F \circ \varphi^{-1} : U' \to V'$$

is a diffeomorphism. We can without loss of generality assume that $\varphi(U) = U'$ and $\psi(V) = V'$. Hence, the "$\Rightarrow$"-direction of the statement of this theorem follows from the inverse function theorem[7]. The "$\Leftarrow$"-direction follows from the fact that invertible smooth maps with smooth inverse in the real analysis setting have pointwise invertible Jacobi matrix. $\square$

**Corollary 1.56.** $F : M \to N$ is a local diffeomorphism if and only if $dF_p$ is a linear isomorphism for all $p \in M$.

**Exercise 1.57.**

(i) Find explicit examples of an immersion that is not injective, and an injective immersion that is not an embedding.

(ii) Show that for all $n \in \mathbb{N}$, the map

$$\pi : S^n \to \mathbb{R}P^n, \quad (x^1, \dots, x^{n+1}) \mapsto [x^1 : \dots : x^{n+1}],$$

is a local diffeomorphism but not a diffeomorphism. In the above equation we view $S^n$ as a subset of $\mathbb{R}^{n+1}$.

### 1.3 Submanifolds

We already know what a smooth submanifold of $\mathbb{R}^n$ is. Using Definition 1.54 we can now define what a smooth submanifold in our more general setting should be.

**Definition 1.58.** Let $N$ be an $n$-dimensional and $M$ be an $m$-dimensional smooth manifold. Let further $F : M \to N$ be a smooth map.

(i) $F(M) \subset N$ is called an **embedded smooth submanifold** if $F$ is an embedding.

(ii) In the special case that $F$ is the inclusion map $\iota : M \hookrightarrow N$, we will say that $M \subset N$ is a **smooth submanifold** if the inclusion is an embedding.

(iii) If $M \subset N$ is a smooth submanifold, the number $\dim(N) - \dim(M)$ is called the **codimension** of $M$ in $N$. Smooth submanifolds of codimension 1 are called **hypersurfaces**.

We will mainly be concerned with smooth submanifolds that are given as subsets of the ambient manifold. The first thing one should ask is how to obtain a the structure of a smooth manifold on a submanifold and if it coincides with the initial manifold structure.

**Proposition 1.59.** Let $M \subset N$, $\dim(M) = m < n = \dim(N)$, be a smooth submanifold and let $p \in M$ be arbitrary. Then there exists a chart[8] $(\varphi = (x^1, \dots, x^n), U)$ on $N$, such that $U \cap M$ is an open neighbourhood of $p$ in $M$ and

$$x^{m+1}(q) = \dots = x^n(q) = 0$$

for all $q \in U \cap M$. The first $m$ entries of $\varphi$ are a local coordinate system on $M$ near $p$.

*Proof.* Fix $p \in M \subset N$ and choose local coordinates $(x^1, \dots, x^n)$ on $N$ and $(y^1, \dots, y^m)$ on $M$ covering $p$. Since $M$ is a submanifold of $N$, the differential of the inclusion map $\iota : M \to N$ at $p$, $d\iota_p$, is injective and its Jacobi matrix

$$\left(\frac{\partial x^i}{\partial y^j}(p)\right)_{ij} \in \mathrm{Mat}(n \times m, \mathbb{R})$$

[7] Which, in turn, follows from the implicit function theorem. Note, however, that one usually proves the implicit function theorem using the inverse function theorem, see e.g. [R]

[8] This means that this chart is compatible with the given maximal atlas on $N$.

has rank $m$. After possibly reordering the $x^i$-coordinate functions, we can assume without loss of generality that its first $m$ rows are linearly independent. By the implicit function theorem this means that the first $m$ coordinate functions on $N$ form, by restriction, a coordinate system on an open set $V \subset M$ containing $p$. Furthermore, after possibly shrinking $V$, we obtain again by the implicit function theorem that $(q^1, \dots, q^n) \in \iota(V) \subset N$ if and only if $x^k(q) = f^k(x^1(q), \dots, x^m(q))$ for uniquely defined smooth functions $f^k : (x^1, \dots, x^m)(V) \to \mathbb{R}$ for all $m+1 \leq k \leq n$. Choose an open subset $U \subset N$, so that the local coordinates $(x^1, \dots, x^n)$ are defined on $V$, $U \cap N = V$, and define on $U$ smooth functions

$$F^k := x^k - f^k(x^1, \dots, x^m), \quad m+1 \leq k \leq n.$$

In the last step we will define new coordinates on $N$ fulfilling the statement of this proposition as follows. Define

$$\varphi : U \to \mathbb{R}^n, \quad \varphi = (x^1, \dots, x^m, F^{m+1}, \dots, F^n).$$

The Jacobi matrix of $\varphi$ at $p$ with respect to the coordinates $(x^1, \dots, x^n)$ is of the form

$$\begin{pmatrix} \mathrm{id}_{\mathbb{R}^m} & 0 \\ A & \mathrm{id}_{\mathbb{R}^{n-m}} \end{pmatrix}$$

for some real-valued matrix $A \in \mathrm{Mat}((n-m) \times m, \mathbb{R})$. The above Jacobi matrix is in particular invertible, showing that $\varphi$ is a local diffeomorphism. Furthermore

$$\varphi|_{U \cap M} = \varphi|_V = (x^1|_V, \dots, x^m|_V, 0, \dots, 0).$$

Hence, the first $m$ entries of the restriction of $\varphi$ to $V$ form a local coordinate system on $M$ near $p$ and thereby fulfil the claims of this proposition. □

**Definition 1.60.** Local coordinates as in Proposition 1.59 for a submanifold $M \subset N$ near a given point $p \in M$ are called **adapted coordinates**.

An important consequence of Proposition 1.59 is that the smooth structure of a manifold that can be realized as a smooth submanifold coincides with the smooth structure obtained by adapted coordinates.

**Corollary 1.61.** Any smooth manifold $M$ that can be realized as a submanifold of some ambient manifold $N$ is diffeomorphic to $M$, viewed as a topological subspace of $N$, equipped with any atlas consisting only of adapted coordinates.

Note that adapted coordinates relate the definition of smooth submanifolds of $\mathbb{R}^n$ to the more general Definition 1.58, cf. equation (1.6). Furthermore observe that Corollary 1.61 also means that if we can cover a topological subspace[9] of $M$ by adapted coordinates it will automatically be a submanifold of $M$. For a more detailed explanation of the latter see [L2, Thm. 5.8]. One way to construct explicit examples of submanifolds is via pre-images of regular values of smooth maps between smooth manifolds.

**Definition 1.62.** Let $M$ and $N$ be smooth manifolds and let $F : M \to N$ be a smooth map. A point $p \in M$ is called **regular point** of $F$ if $dF_p : T_pM \to T_{F(p)}N$ is surjective. A point $q \in N$, such that $F^{-1}(q) \subset M$ consists only of regular points, is called **regular value** of $F$. Points in $M$ that are not regular points of $F$ are called **critical points** of $F$, and points in $N$ such that the pre-image under $F$ in $M$ contains at least one critical point of $F$ are called **critical values** of $F$.

[9] Careful, it needs to be second countable and Hausdorff.

Note that for a smooth map $F : M \to N$ to have regular values it is a necessary condition that $\dim(M) \geq \dim(N)$.

**Proposition 1.63.** Let $M$ and $N$ be smooth manifolds with $\dim(M) = m \geq n = \dim(N)$. Let $F : M \to N$ be smooth and let $q \in N$ be a regular value of $F$. Then the level set

$$F^{-1}(q) \subset M$$

is an $(m-n)$-dimensional smooth submanifold of $M$. The structure of a smooth manifold on $F^{-1}(q)$ is uniquely determined by requiring that the inclusion is smooth.

For the proof of the above proposition we need the following definition and two theorems, cf. [L2, Thm. 4.12, Thm. 5.12].

**Definition 1.64.** Let $F : M \to N$ be a smooth map between smooth manifolds and let $(\varphi, U)$ and $(\psi, V)$ be local charts of $M$ and $N$, respectively, such that $F(U) \subset V$. Let further $\dim(M) = m$ and $\dim(N) = n$. The **coordinate representation of** $F$ in the local coordinate systems $\varphi$ and $\psi$ is defined to be the smooth map

$$\widehat{F} : \varphi(U) \to \varphi(V), \quad \widehat{F}(u^1, \ldots, u^m) := (\psi \circ F \circ \varphi^{-1})(u^1, \ldots, u^m).$$

**Theorem 1.65.** Let $M$ be an $m$-dimensional and $N$ be an $n$-dimensional smooth manifold. Let $F : M \to N$ be a smooth map of constant rank $r$. Then for each $p \in M$ there exist local charts $(\varphi, U)$ of $M$ with $p \in U$ and $(\psi, V)$ of $N$ with $F(p) \in V$, such that $F(U) \subset V$ and that the coordinate representation of $F$ is of the form

$$\widehat{F}(u^1, \ldots, u^r, u^{r+1}, \ldots, u^m) = (u^1, \ldots, u^r, 0, \ldots, 0).$$

*Proof.* For a detailed proof see [L2, Thm. 4.12]. The case $r = 0$ is left as an exercise. Assume that $r \geq 1$. The proof works as follows. For $p \in M$ fixed we choose local coordinates $M$ covering $p$ and of $N$ covering $F(p)$. Since the statement of this theorem is local, by switching to local coordinates we find that in order to prove it it suffices to consider the special case $M \subset \mathbb{R}^m$ open and $N \subset \mathbb{R}^n$ open. This shows that this theorem is equivalent to the rank theorem known from real analysis, see (in a slightly different formulation) [R, Thm. 9.32]. $\square$

**Theorem 1.66.** Let $M$ and $N$ be smooth manifolds and let $F : M \to N$ be a smooth map of constant rank $r$. Each level set $F^{-1}(q) \subset M$, $q \in N$, is a smooth submanifold of codimension $r$ in $M$.

*Proof.* Let $q \in N$ and $p \in F^{-1}(q)$ be fixed. Using Theorem 1.65 we chose charts $(\varphi = (x^1, \ldots, x^m), U)$ of $M$ with $p \in U$ and $(\psi, V)$ of $N$ with $q \in V$ fulfilling $\varphi(p) = 0$ and $\psi(q) = 0$, such that the coordinate representation $\widehat{F}$ of $F$ is of the form

$$\widehat{F} : \varphi(U) \to \varphi(V), \quad \widehat{F}(u^1, \ldots, u^r, u^{r+1}, \ldots, u^m) = (u^1, \ldots, u^r, 0, \ldots, 0).$$

Then $(\psi \circ F)(x^1, \ldots, x^m) = (x^1, \ldots, x^r, 0, \ldots, 0)$ and, hence,

$$F^{-1}(q) \cap U = \{p \in U \mid x^1(p) = \ldots = x^r(p) = 0\}.$$

We see that such a coordinate choice, up to reordering of the coordinate functions, on $M$ yields adapted coordinates on $F^{-1}(q) \cap U$. Since the rank of $F$ is constant we can cover $F^{-1}(q)$ with so constructed adapted coordinates and obtain that it is, in fact, a smooth submanifold of $M$. $\square$

Lastly we will need the following fact.

**Proposition 1.67.** Let $M$ and $N$ be smooth manifolds and let $F : M \to N$ be a smooth map. Suppose that $p \in M$ is a regular point of $F$. Then there exists an open neighbourhood $U \subset M$ of $p$, such that all points in $U$ are regular points of $F$. In particular this means that the set of regular points of $F$ is open in $M$.

*Proof.* Exercise. [Hint: Use local coordinates to reduce the proof to the case $M \subset \mathbb{R}^m$ open and $N \subset \mathbb{R}^n$ open.] $\square$

*Proof of Proposition 1.63.* Let $F : M \to N$ be smooth and $q \in N$ a regular value of $F$. By Proposition 1.67 the set

$$\mathrm{reg}(F) := \{p \in M \mid p \text{ regular point of } F\}$$

is open in $M$ and thereby a smooth submanifold of $M$. We further have $F^{-1}(q) \subset \mathrm{reg}(F)$. The restriction of $F$ to $\mathrm{reg}(F)$,

$$F|_{\mathrm{reg}(F)} : \mathrm{reg}(F) \to N,$$

is by Definitions 1.54 and 1.62 a submersion and thereby of constant rank equal to $\dim(N)$. Using Theorem 1.66 it follows that $F^{-1}(q) \subset \mathrm{reg}(F)$ is a smooth submanifold. Since the composition of the inclusions $F^{-1}(q) \subset \mathrm{reg}(F)$ and $\mathrm{reg}(F) \subset M$ is still the inclusion and thereby in particular still a smooth embedding it follows that $F^{-1}(q) \subset M$ is a smooth submanifold. Since $\mathrm{reg}(F) \subset M$ is open it follows with Theorem 1.66 that $\dim(F^{-1}(q)) = m - n$. $\square$

## 1.4 Vector bundles and sections

At this point, we understand what tangent vectors at a specific given point are. The next step is to study vector fields, that is maps that assign to points in a smooth manifold tangent vectors in the respective tangent spaces. In the general setting of smooth manifold these objects are more involved than in the case we know from real analysis.

**Remark 1.68.** Recall that a smooth vector field on $\mathbb{R}^n$ is a smooth vector valued function

$$X : \mathbb{R}^n \to \mathbb{R}^n, \quad p \mapsto X_p.$$

We think of points $(p, X_p) \in \mathbb{R}^n \times \mathbb{R}^n$ as tangent vectors $X_p$ with basepoint $p$. An example of a smooth vector field on $\mathbb{R}^n$ is the position vector field $X : p \mapsto p$ for all $p \in \mathbb{R}^n$. Observe that vector fields on $\mathbb{R}^n$, similar to tangent vectors, act on functions via

$$X(f) : \mathbb{R}^n \to \mathbb{R}, \quad p \mapsto [\gamma]f,$$

where $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ is any smooth curve, such that

$$\gamma(0) = p, \quad \gamma'(0) = X(p),$$

cf. equation (1.8). Note that for any smooth vector field $X$ on $\mathbb{R}^n$ and any $f \in C^\infty(\mathbb{R}^n)$, $X(f) \in C^\infty(\mathbb{R}^n)$. One might also write $X(f) = df(X) : p \mapsto df_p(X_p)$.

**Definition 1.69.** A **vector bundle** $E \to M$ **of rank** $k \in \mathbb{N}$ over a smooth manifold $M$ is a smooth manifold $E$ together with a smooth **projection map** $\pi : E \to M$, such that

(i) the **fibre** $E_p := \pi^{-1}(p)$ is an $k$-dimensional real vector space for all $p \in M$,

(ii) for all $p \in M$ there exists an open neighbourhood $U \subset M$ of $p$ and a diffeomorphism $\psi : \pi^{-1}(U) \to U \times \mathbb{R}^k$, such that $\psi|_{E_q} : E_q \to q \times \mathbb{R}^k \cong \mathbb{R}^k$ is a linear isomorphism for all $q \in U$ and the diagram

$$\begin{tikzcd} \pi^{-1}(U) \arrow[rr, "\psi"] \arrow[dr, "\pi"'] & & U \times \mathbb{R}^k \arrow[dl, "\mathrm{pr}_U"] \\ & U & \end{tikzcd}$$

commutes. The map $\mathrm{pr}_U$ denotes the canonical projection onto the first factor.

$E$ is called the **total space**, $M$ is called the **basis**, and the map $\psi$ is called a **local trivialization** of the vector bundle $E \to M$.

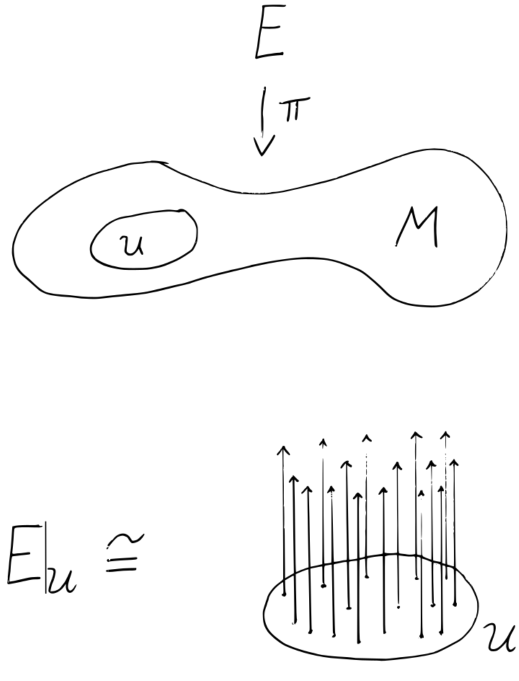


**Figure 16:** Locally, $E|_U \cong U \times \mathbb{R}^k$.

Vector bundles provide the setting for an analogue to vector valued functions on smooth manifolds.

**Definition 1.70.** Let $E \to M$ be a vector bundle. A **local section in** $E \to M$ is a smooth map

$$s : U \to E$$

with $U \subset M$ open, such that $\pi \circ s = \mathrm{id}_U$. This precisely means that $s(p) \in E_p$ for all $p \in U$. If $U = M$, $s$ is called a **(global) section**. The set of local sections in $E \to M$ on $U \subset M$ is denoted by $\Gamma(E|_U)$ and the set of global sections by $\Gamma(E)$, where $E|_U$ denotes the vector bundle $\pi^{-1}(U) \to U$. The **support** of a section (or, analogously, local section) in a vector bundle $s \in \Gamma(E)$ is defined to be the set

$$\mathrm{supp}(s) := \overline{\{p \in M \mid s(p) \neq 0\}}.$$

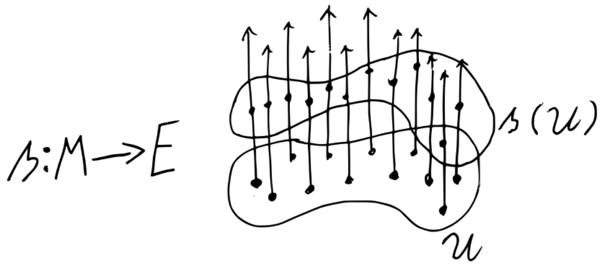


**Figure 17:** A sketch of a section.

**Exercise 1.71.**

(i) Show that $\Gamma(E)$ is a $C^\infty(M)$-module. Also show that $\Gamma(E|_U)$ is a $C^\infty(U)$-module for all $U \subset M$ open.

(ii) Show that for $k > 0$ the restriction map $\Gamma(E) \to \Gamma(E|_U)$ for $U \subset M$ open and precompact, such that the boundary of $U$, $\partial U$, is nonempty and a smooth hypersurface in $M$, is not surjective.

**Definition 1.72.** Let $\psi : \pi^{-1}(U) \to U \times \mathbb{R}^k$ and $\phi : \pi^{-1}(V) \to V \times \mathbb{R}^k$ be two local trivializations of a vector bundle $E \to M$. Assume that $U \cap V \neq \emptyset$. Then the smooth map

$$\psi \circ \phi^{-1} : (U \cap V) \times \mathbb{R}^k \to (U \cap V) \times \mathbb{R}^k$$

is called **transition function**[10]. For $p \in M$ fixed, $(\psi \circ \phi^{-1})(p, \cdot)$ is called **transition function at** $p$.

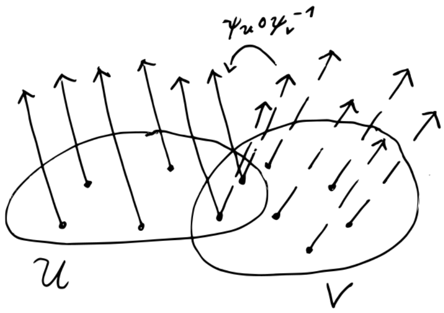


**Figure 18:** How a transition function w.r.t. $(\psi_U, U)$ and $(\psi_V, V)$ from the overlap of $V \times \mathbb{R}^k$ and $U \times \mathbb{R}^k$ to itself can be imagined.

**Lemma 1.73.** Transition functions $\psi \circ \phi^{-1}$ as in Definition 1.72 are of the form

$$\psi \circ \phi^{-1} : (p, v) \mapsto (p, A(p)v), \quad A(p) \in \mathrm{GL}(k),$$

for all $p \in U \cap V$, $v \in \mathbb{R}^n$. The map

$$A : U \cap V \to \mathrm{GL}(k), \quad p \mapsto A(p),$$

is smooth.

[10]Note that a change of coordinates in a smooth manifold is also called a transition function. Always make sure to clarify which kind of transition functions you are dealing with.

*Proof.* The diagram

$$\begin{tikzcd}
U \cap V \times \mathbb{R}^k \arrow[r, "\phi^{-1}"] \arrow[dr, "\mathrm{pr}_{U\cap V}"'] & \pi^{-1}(U \cap V) \arrow[r, "\psi"] \arrow[d, "\pi"] & U \cap V \times \mathbb{R}^k \arrow[dl, "\mathrm{pr}_{U\cap V}"] \\
& U \cap V &
\end{tikzcd}$$

commutes and, hence, it follows that $\psi \circ \phi^{-1}$ sends $(p, v)$ to $(p, A(p, v))$ for some smooth function $A : U \cap V \times \mathbb{R}^k \to \mathbb{R}^k$. Smoothness follows from the diffeomorphism property of $\phi$ and $\psi$. We need to show that for $p$ fixed, $A(p, \cdot) : \mathbb{R}^k \to \mathbb{R}^k$ is an invertible linear map. This follows from the fact that fibrewise $\phi$ and $\psi$ are linear isomorphisms. $\square$

An important tool to construct vector bundles is from, heuristically speaking, a given set transition functions.

**Proposition 1.74** ("Vector bundle chart lemma")**.** Let $M$ be a smooth manifold and assume that for every $p \in M$, $E_p$ is a real vector space of fixed dimension $k$. Define a set

$$E := \bigsqcup_{p \in M} E_p$$

together with a map $\pi : E \to M$, $\pi(v) = p$ for all $v \in E_p$ and all $p \in M$. Assume that $\{U_i,\ i \in I\}$ is an open cover of $M$ and for each $i \in I$,

$$\phi_i : \pi^{-1}(U_i) \to U_i \times \mathbb{R}^k$$

is a bijection with the property that the restriction $\phi_i : E_p \to \{p\} \times \mathbb{R}^k \cong \mathbb{R}^k$ is a linear isomorphism for all $p \in U_i$. Further assume that for all $i, j \in I$ with $U_i \cap U_j \neq \emptyset$ there exists a smooth map $\tau_{ij} : U_i \cap U_j \to \mathrm{GL}(k)$, such that $\phi_i \circ \phi_j^{-1} : (U_i \cap U_j) \times \mathbb{R}^k \to (U_i \cap U_j) \times \mathbb{R}^k$ is of the form

$$\phi_i \circ \phi_j^{-1}(p, v) = (p, \tau_{ij}(p)v).$$

Then there exists a unique topology and maximal atlas on $E$, such that $\pi : E \to M$ is a vector bundle of rank $k$ and the $\phi_i$, $i \in I$, are local trivializations.

*Proof.* The proof follows [L2, Lem. 10.6]. Without loss of generality assume that we can find an atlas $\{(\varphi_i, U_i) \mid i \in I\}$ on $M$. This can always be achieved by shrinking the $U_i$ if necessary and, on possible new overlaps $U_i \cap U_j$, set $\tau_{ij} = \mathrm{id}_{\mathbb{R}^k}$. Now we can explicitly construct an atlas on the total space $E$. Define for $i \in I$

$$\psi_i : \pi^{-1}(U_i) \to \varphi_i(U_i) \times \mathbb{R}^k, \quad v \mapsto (\varphi_i \times \mathrm{id}_{\mathbb{R}^k})(\phi_i(v)).$$

In order for $\{(\psi_i, \pi^{-1}(U_i)) \mid i \in I\}$ to be a smooth atlas on $E$, we need to show that the transition functions (as in transition functions of a smooth atlas, cf. Definition 1.5) are smooth. We check that

$$\psi_i(\pi^{-1}(U_i) \cap \pi^{-1}(U_j)) = \varphi_i(U_i \cap U_j) \times \mathbb{R}^k$$

for all $i, j \in I$, and we find

$$\psi_i \circ \psi_j^{-1} = (\varphi_i \times \mathrm{id}_{\mathbb{R}^k}) \circ (\phi_i \circ \phi_j^{-1}) \circ (\varphi_j^{-1} \times \mathrm{id}_{\mathbb{R}^k}) : \varphi_j(U_i \cap U_j) \times \mathbb{R}^k \to \varphi_i(U_i \cap U_j) \times \mathbb{R}^k.$$

Since, by assumption, $\tau_{ij}(p)$ is invertible and depends smoothly on $p \in U_i \cap U_j$, $\phi_i \circ \phi_j^{-1}$ is a diffeomorphism for all $i, j \in I$ such that $U_i \cap U_j \neq \emptyset$. Since the $\varphi_i$ form a smooth atlas on $M$, each $\varphi_i \times \mathrm{id}_{\mathbb{R}^k}$ is a diffeomorphism. Hence, $\psi_i \circ \psi_j^{-1} : \varphi_j(U_i \cap U_j) \times \mathbb{R}^k \to \varphi_i(U_i \cap U_j) \times \mathbb{R}^k$ is also a diffeomorphism for all $i, j \in I$ such that $U_i \cap U_j \neq \emptyset$. By defining the open sets on $E$ as the preimages of open sets under $\psi_i$, $i \in I$, we obtain that the so-defined topology is second countable and Hausdorff by the assumption that $M$ is a smooth manifold (and $\mathbb{R}^k$ is, of course, second countable and Hausdorff as well). It follows that $\mathcal{B} := \{(\psi_i, \pi^{-1}(U_i)) \mid i \in I\}$ is a smooth atlas on the total space $E$. Then all the maps $\phi_i$, $i \in I$, are automatically smooth and, since $\phi_i : E_p \to \{p\} \times \mathbb{R}^k$ is a linear isomorphism by assumption, form a covering of local trivializations which turns $E \to M$ into a vector bundle of rank $k$. The uniqueness of the smooth manifold structure on $E$ now follows from the assumption that all $\phi_i$ are diffeomorphisms onto their image and, thus, every smooth atlas on $E$ with that property must, by construction, be a refinement of $\mathcal{B}$ and thus be contained in the same maximal smooth atlas as $\mathcal{B}$. □

Now we have all the tools at hand that we need to define the tangent bundle of a smooth manifold.

**Definition 1.75.** Let $M$ be an $n$-dimensional smooth manifold. The **tangent bundle**[11] $TM \to M$ of $M$ is a vector bundle of rank $n$ with total space $TM := \bigsqcup_{p \in M} T_pM$ and projection $\pi(v) = p$ for all $v \in T_pM$.

At this point, however, we still need to explain the structure of a smooth manifold on the total space of the tangent bundle $TM$ and we need to show that it actually is a vector bundle.

**Proposition 1.76.** The tangent bundle $TM$ of any given manifold is, in fact, a vector bundle of rank $n$.

*Proof.* We need to explain the topology on $TM$, find an atlas, and show that we can locally trivialize it as a vector bundle. Fix a countable atlas (cf. Exercise 1.13)

$$\mathcal{A} = \{(\varphi_i = (x_i^1, \ldots, x_i^n), U_i) \mid i \in A\}$$

on $M$. Since $\pi$ is assumed to be smooth and hence continuous, the pre-images $\{\pi^{-1}(U_i) \mid i \in A\}$ form an open cover of $TM$. Taking pre-images under $\pi$ of a basis of the topology on $M$ is not sufficient to explain the topology on $TM$. For $i \in A$ consider the maps

$$\begin{aligned} \psi_i &: \pi^{-1}(U_i) \to \varphi_i(U_i) \times \mathbb{R}^n, \\ \psi_i &: v \mapsto (\varphi_i(\pi(v)), v(x_i^1), \ldots, v(x_i^n)) = (\varphi_i(\pi(v)), d\varphi_i(v)), \end{aligned} \tag{1.10}$$

and observe that each $\psi_i$ is a bijection. We can think of the above maps as candidates for a local trivialization that, via a chart on the manifold $M$ itself, has its target space changed as in in the following diagram

$$\begin{tikzcd} \pi^{-1}(U_i) \arrow[dd, bend left=60, "\psi_i"] \\ U_i \times \mathbb{R}^n \arrow[d, "\varphi_i \times \mathrm{id}_{\mathbb{R}^n}"] \\ \varphi_i(U_i) \times \mathbb{R}^n \end{tikzcd} \tag{1.11}$$

We define a basis of the topology on $TM$ as

$$\{\psi_i^{-1}(V) \mid i \in A,\ V \subset \varphi_i(U_i) \times \mathbb{R}^n \text{ open}\}$$

[11] Also simply called **tangent space**, without the “at $p$” part.

which is precisely the coarsest topology on $TM$, such that all maps $\psi_i$, $i \in A$, are homeomorphisms. Since $A$ is a countable set and the topology in each $U_i$ has a countable basis, it follows that the so-defined topology on $TM$ is countable. To see that it is also Hausdorff, consider for $p \neq q \in TM$ the points $\psi_i(p)$ and $\psi_j(q)$ for fitting $i, j \in A$. If $U_i \cap U_j = \emptyset$, we can separate $p$ and $q$ by $\pi^{-1}(U_i)$ and $\pi^{-1}(U_j)$. For $U_i \cap U_j \neq \emptyset$, observe that each space $\varphi_i(U_i) \times \mathbb{R}^n$ is Hausdorff and we can thus find open neighbourhoods of $\psi_i(p)$ and $\psi_j(q)$ in $\varphi_i(U_i \cap U_j) \times \mathbb{R}^n$ that separate these points. The pre-images under $\psi_i$ of these sets will then separate $p$ and $q$. Next consider the transition functions (thought of as change of coordinates) of the $\psi_i$'s. For $U_i \cap U_j \neq \emptyset$ we have (recall Example 1.51)

$$\begin{aligned} \psi_i \circ \psi_j^{-1} : \varphi_j(U_i \cap U_j) \times \mathbb{R}^n &\to \varphi_i(U_i \cap U_j) \times \mathbb{R}^n, \\ (u, w) &\mapsto ((\varphi_i \circ \varphi_j^{-1})(u), d(\varphi_i \circ \varphi_j^{-1})_u(w)). \end{aligned} \tag{1.12}$$

Since the transition functions $\varphi_i \circ \varphi_j^{-1}$ are smooth it follows that the countable set

$$\{(\psi_i, \pi^{-1}(U_i)) \mid i \in A\}$$

$\{(\psi_i, \pi^{-1}(U_i)) \mid i \in A\}$ defines a countable atlas on $TM$. The vector bundle structure on $TM$ is explained by the local trivializations $(\varphi_i^{-1} \times \mathrm{id}_{\mathbb{R}^n}) \circ \psi_i$, $i \in A$, cf. (1.11). □

**Remark 1.77.** Compare the proofs of Propositions 1.74 and 1.76 for similarities and differences. Note that by Proposition 1.74 the structure of a vector bundle in $TM \to M$ is uniquely determined by requiring that the transition functions are given by (1.12). Also note that the transition functions of the local trivializations $(\varphi_i^{-1} \times \mathrm{id}_{\mathbb{R}^n}) \circ \psi_i$ are given by

$$(\varphi_i^{-1} \times \mathrm{id}_{\mathbb{R}^n}) \circ \psi_i \circ \Big((\varphi_j^{-1} \times \mathrm{id}_{\mathbb{R}^n}) \circ \psi_j\Big)^{-1} = (\mathrm{id}_M, d(\varphi_i \circ \varphi_j^{-1})),$$

that is the matrix parts are differentials of the transition functions of the charts on $M$.

**Remark 1.78.** Let $M$ be a smooth manifold and $\varphi = (x^1, \ldots, x^n)$ a local coordinate system covering $p \in M$. Let $v \in T_pM$,

$$v = \sum_{i=1}^n v^i \left.\frac{\partial}{\partial x^i}\right|_p .$$

Observe that with $\psi$ as in (1.10)

$$\psi(v) = (x^1(p), \ldots, x^n(p), v^1, \ldots, v^n),$$

meaning that the vector part of $\psi(v)$ consists of the prefactors of $v$ in the basis $\left\{ \left.\frac{\partial}{\partial x^i}\right|_p,\ 1 \leq i \leq n \right\}$.

**Exercise 1.79.** Consider $S^n$ with atlas containing the stereographic projections as in Example 1.19 (i). Explicitly calculate the corresponding transition functions (1.12) in the tangent bundle $TS^n$.

Given a smooth manifold, one might ask how "bad" the tangent bundle might look like. For this question we first need to clarify when two vector bundles are considered isomorphic.

**Definition 1.80.** Let $\pi_E : E \to M$ and $\pi_F : F \to M$ be vector bundles over a smooth manifold $M$. Then a **smooth vector bundle homomorphism**[12] is a smooth map between the total spaces

$$f : E \to F,$$

[12] a.k.a. "smooth vector bundle map"

such that the diagram

$$\begin{tikzcd} E \arrow[rr, "f"] \arrow[dr, "\pi_E"'] & & F \arrow[dl, "\pi_F"] \\ & M & \end{tikzcd}$$

commutes and $f$ is fibrewise linear. The last condition means that for each $p \in M$,

$$f|_{E_p} : E_p \to F_p$$

is a linear map.

**Definition 1.81.** Two vector bundles $\pi_1 : E_1 \to M$ and $\pi_2 : E_2 \to M$ are **isomorphic** if there exists a diffeomorphism $F : E_1 \to E_2$ that is a smooth vector bundle map, so that

$$F|_{E_{1p}} : E_{1p} \to E_{2p}$$

is a linear isomorphism for all $p \in M$.

**Definition 1.82.** A vector bundle of rank $k$, $E \to M$, is called **trivializable** if it is isomorphic to $M \times \mathbb{R}^k \to M$ equipped with the canonical projection onto $M$.

**Lemma 1.83.** Assume that $E \to M$ is trivializable. Then there exists a nowhere vanishing section $s \in \Gamma(E)$.

*Proof.* Exercise. □

The best case scenario we can expect for the tangent bundle of a smooth manifold is that it is trivializable, which is in general not true. An example of a smooth manifold with non-trivializable tangent bundle is $S^2$. This follows from the "hairy ball theorem"[13] [M]. There are however non-obvious examples of manifolds with trivializable tangent bundle.

**Exercise 1.84.** Show that $TS^1$ is trivializable and, hence, as a smooth manifold isomorphic to the cylinder $S^1 \times \mathbb{R}$. Draw a sketch of the isomorphism.

We now have all tools at hand to define vector fields on smooth manifolds.

**Definition 1.85.** Sections in the tangent bundle of a smooth manifold, $\Gamma(TM)$, are called **vector fields**. For $X \in \Gamma(TM)$ we will denote the value of $X$ at $p \in M$ by $X_p$. For $U \subset M$ open, we will call elements of $\Gamma(TM|_U)$ **local vector fields**, or simply vector fields if the setting does not explicitly use the locality property. We will use the notations

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

and

$$\mathfrak{X}(U) := \Gamma(TM|_U)$$

for $U \subset M$ open.

**Remark 1.86.** For a smooth manifold $M$ and $U \subset M$ open, the two vector spaces $T_pU$ and $T_pM$ are canonically isomorphic via restriction of charts for all $p \in U$. In the following we will omit using $T_pU$ and instead write $T_pM$, e.g. if we want to denote the action of a tangent vector on a function $f \in C^\infty(U)$, $v(f)$, we will write $v \in T_pM$ and not $v \in T_pU$.

[13]German: "Satz vom Igel"

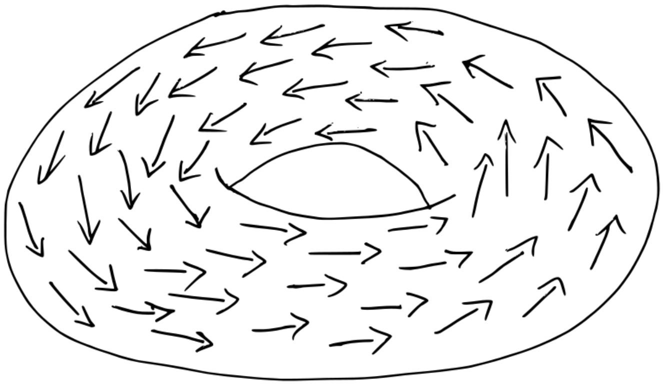

**Figure 19:** A vector field on the 2-torus.

**Remark 1.87.** Vector fields, similar to tangent vectors, act on $C^\infty(M)$ by

$$X(f)(p) := X_p(f) = df(X_p).$$

Thus we may write $X(f) = df(X) \in C^\infty(M)$. On the other hand, a map of the form $X : M \to TM$, $p \mapsto X_p \in T_pM$, is a vector field if and only if $X(f) : p \mapsto df_p(X_p)$ is smooth for all $f \in C^\infty(M)$.

Recall that in Proposition 1.47 we have shown that in local coordinates $(x^1, \dots, x^n)$ the tangent vectors $\frac{\partial}{\partial x^i}\Big|_p$, $1 \leq i \leq n$, form a basis of $T_pM$. We want to have a similar result for the local form of vector fields in $\Gamma(TM|_U)$ for $U$ the chart neighbourhood of the local coordinates $x^i$.

**Definition 1.88.** Let $(\varphi = (x^1, \dots, x^n), U)$ be a chart on a smooth manifold $M$. The corresponding **coordinate vector fields** are defined as

$$\frac{\partial}{\partial x^i} \in \mathfrak{X}(U), \quad \frac{\partial}{\partial x^i} : p \mapsto \frac{\partial}{\partial x^i}\Big|_p.$$

**Proposition 1.89.** Let $(\varphi = (x^1, \dots, x^n), U)$ be a chart on a smooth manifold $M$ and $X \in \mathfrak{X}(U)$. With $X^i := X(x_i) \in C^\infty(U)$ we have

$$X = \sum_{i=1}^n X^i \frac{\partial}{\partial x^i}.$$

On the other hand for any choice of smooth functions $f^i \in C^\infty(U)$, $1 \leq i \leq n$,

$$\sum_{i=1}^n f^i \frac{\partial}{\partial x^i} \in \mathfrak{X}(U).$$

*Proof.* The first claim follows from the fact that for any $p \in U$ fixed, $X_p = \sum\limits_{i=1}^n X_p(x^i) \frac{\partial}{\partial x^i}\Big|_p$, which follows from Proposition 1.47. The second claim follows from the fact that each $\frac{\partial}{\partial x^i}$ is a vector field on $U$ and Exercise 1.71 (i). □

**Exercise 1.90.**

(i) Prove the statements in Remark 1.87.

(ii) Construct a vector field $X \in \mathfrak{X}(S^2)$ with precisely one bald spot, meaning that there exists precisely one $p \in S^2$, such that $X_p = 0$.

There is an alternative, but equivalent, way of introducing vector fields on smooth manifolds, see [O]. Recall that in Definition 1.35 we have initially defined tangent vectors to be linear maps from $C^\infty(M)$ to $\mathbb{R}$ satisfying a Leibniz rule. Vector fields can be introduced similarly using the concept of derivations from differential algebra.

**Definition 1.91.** Let $A$ be an algebra over a field $K$. A **derivation** of a $A$ is a $K$-linear map $D : A \to A$ that fulfils the Leibniz rule

$$D(ab) = D(a)b + aD(b)$$

for all $a, b \in A$. The set of all derivations of $A$ is denoted by $\mathrm{Der}(A)$. If $A$ is commutative, $\mathrm{Der}(A)$ is an $A$ module.

Recall that the smooth functions on a manifold, $C^\infty(M)$, form an $\mathbb{R}$-algebra.

**Proposition 1.92.** Let $M$ be a smooth manifold. Then vector fields on $M$ are precisely the derivations of $C^\infty(M)$, meaning that $\mathfrak{X}(M)$ and $\mathrm{Der}(C^\infty(M))$ are isomorphic as $C^\infty(M)$ modules.

*Proof.* The map

$$\iota : \mathfrak{X}(M) \to \mathrm{Der}(C^\infty(M)), \quad X \mapsto (f \mapsto X(f)),$$

is a $C^\infty(M)$ module map. Injectivity of $\iota$ follows from $X = 0$ if and only if $X(f) = 0$ for all $f \in C^\infty(M)$ (cf. proof of Proposition 1.47 if you have problems seeing that fact). For surjectivity we define for a given derivation $D$ a vector field $X^D$ via

$$D \mapsto X^D, \quad X^D_p(f) = D(f)(p),$$

for all $p \in M$ and all $f \in C^\infty(M)$. By Remark 1.87 we know that $X^D$ is in fact a smooth vector field. The map $D \mapsto X^D$ is precisely the inverse of $\iota$. □

We now know the algebraic properties of vector fields as derivations and we know how to write down and calculate with vector fields locally. The following lemma describes explicitly how vector fields behave under a change of coordinates.

**Lemma 1.93.** Let $M$ be a smooth manifold and let $(\varphi = (x^1, \ldots, x^n), U)$, $(\psi = (y^1, \ldots, y^n), V)$ be charts on $M$ such that $U \cap V \neq \emptyset$. For $X \in \mathfrak{X}(M)$ fixed, we have on $U \cap V$ the following forms of $X$ in local coordinates

$$X = \sum_{i=1}^n X(x^i) \frac{\partial}{\partial x^i}$$

and

$$X = \sum_{i=1}^n X(y^i) \frac{\partial}{\partial y^i}.$$

If we understand $d(\psi \circ \varphi^{-1}) : \varphi(U \cap V) \to \mathrm{GL}(n)$ as a matrix-valued function which associates each point $u \in \varphi(U \cap V)$ the Jacobi matrix of $\psi \circ \varphi^{-1}$ at $u$ we obtain

$$d(\psi \circ \varphi^{-1})_u \cdot \begin{pmatrix} X(x^1) \\ \vdots \\ X(x^n) \end{pmatrix}\Bigg|_{\varphi^{-1}(u)} = \begin{pmatrix} X(y^1) \\ \vdots \\ X(y^n) \end{pmatrix}\Bigg|_{\psi^{-1}(u)}$$

for all $u \in U \cap V$.

*Proof.* Follows from the definition of the Jacobi matrix and the coordinate vector fields. □

If the above formula looks difficult to you, calculate some examples for $M = \mathbb{R}^n$, $\varphi = \mathrm{id}_{\mathbb{R}^n}$, and $\psi : \mathbb{R}^n \to \mathbb{R}^n$ any diffeomorphism. Having defined vector fields and coordinate vector fields, we can now properly define differentials of smooth maps, compared to our pointwise Definition 1.50.

**Definition 1.94.** Let $M, N$ be smooth manifolds and $F : M \to N$ be a smooth map. The **differential of** $F$ is defined as the smooth map

$$dF : TM \to TN, \quad dF|_{\pi^{-1}(p)} = dF_p \quad \forall p \in M.$$

The above equation means that pointwise, $dF$ is given by its differential as in Definition 1.50. Thus, in local coordinates $(x^1, \ldots, x^m)$ of $M$ and $(y^1, \ldots, y^n)$ of $N$ with appropriate domain we have

$$dF\left(\frac{\partial}{\partial x^i}\right) = \sum_{j=1}^{n} \frac{\partial F^j}{\partial x^i} \frac{\partial}{\partial y^j}, \quad F^j = y^j \circ F, \quad \forall 1 \leq i \leq m.$$

The (non-pointwise) Jacobi matrix in given local coordinates is defined similarly by allowing the basepoint to vary. As a map from chart neighbourhoods in $M$ to $\mathrm{GL}(n)$, the Jacobi matrix in particular is a smooth map.

**Exercise 1.95.** Check using local coordinates on $TM$ and $TN$ that $dF$ and the Jacobi matrix as in Definition 1.94 are actually smooth as claimed.

On the vector fields on a smooth manifold we have the structure of a **Lie**[14] **algebra**. Before describing this concept in detail, consider for two derivations $X, Y \in \mathrm{Der}(A)$ of an algebra $A$ the commutator of $X$ and $Y$

$$[X, Y] := XY - YX.$$

**Exercise 1.96.** Show that $[X, Y] \in \mathrm{Der}(A)$.

By Proposition 1.92 there must be an analogue construction on the set of vector fields on a smooth manifold.

**Definition 1.97.** Let $V$ be a real vector space. A **Lie bracket** on $V$ is a skew-symmetric bilinear map

$$[\cdot, \cdot] : V \times V \to V, \quad (X, Y) \mapsto [X, Y]$$

that fulfils the[15] **Jacobi identity**

$$[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]]$$

for all $X, Y, Z \in V$. A vector space $V$ together with a Lie bracket is called **Lie algebra**.

**Exercise 1.98.**

(i) Prove that $[\cdot, \cdot]$ on $\mathrm{Der}(A)$ is a Lie bracket.

(ii) Show that the Jacobi identity in Definition 1.97 is equivalent to

$$\sum_{\text{cyclic}} [X, [Y, Z]] = 0,$$

for all $X, Y, Z \in V$, where $\sum\limits_{\text{cyclic}}$ stands for the cyclic sum.

[14] Sophus Lie (1842 – 1899)

[15] Careful: There is more than one Jacobi identity, e.g. graded Jacobi identities.

**Proposition 1.99.** The bilinear map on vector fields on a smooth manifold $M$

$$[\cdot,\cdot] : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M), \quad (X,Y) \mapsto [X,Y],$$
$$[X,Y](f) := X(Y(f)) - Y(X(f)) \quad \forall X,Y \in \mathfrak{X}(M)\ \forall f \in C^\infty(M),$$

is a Lie bracket on the real vector space $\mathfrak{X}(M)$.

*Proof.* Follows from Exercise 1.98 (i). □

Note that $[X,Y]_p(f) = X_p(Y(f)) - Y_p(X(f))$ for all $p \in M$, $f \in C^\infty(M)$, $X,Y \in \mathfrak{X}(M)$. From real analysis we know that partial derivatives commute. We can formulate a similar result for smooth manifolds with the help of the Lie algebra structure on $\mathfrak{X}(M)$.

**Lemma 1.100.** Let $M$ be a smooth manifold and let $(x^1,\dots,x^n)$ be local coordinates on $M$. Then

$$\left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right] = 0$$

for all $1 \le i \le n$, $1 \le j \le n$.

*Proof.* Exercise. □

**Exercise 1.101.** Show that $[X, fY] = df(X)Y + f[X,Y]$ for all $X,Y \in \mathfrak{X}(M)$, $f \in C^\infty(M)$.

**Definition 1.102.** Let $\phi : M \to N$ be a smooth map and let $X \in \mathfrak{X}(M)$. The smooth map

$$M \ni p \mapsto (d\phi(X))_p = d\phi_p(X_p) \in T_{\phi(p)}N$$

is called a **vector field along** $\phi$.

Note that $d\phi(X)$ sends smooth functions on $N$ to smooth functions on $M$ via

$$(d\phi(X))(f) = X(f \circ \phi) \in C^\infty(M)$$

for all $f \in C^\infty(N)$.

**Definition 1.103.** Let $I \subset \mathbb{R}$ be an interval (equipped with canonical coordinate $t$), $M$ a smooth manifold, and $\gamma : I \to M$ a smooth curve. The **velocity vector field** (or simply **velocity**) of $\gamma$ is the vector field along $\gamma$

$$\gamma' := d\gamma\left(\frac{\partial}{\partial t}\right), \quad t \mapsto \gamma'(t).$$

Note that the explicit form of $\gamma'(t)$ depends on the local coordinates $\varphi = (x^1,\dots,x^n)$ on $M$,

$$\gamma'(t) = \sum_{i=1}^n \frac{\partial \gamma^i}{\partial t}(t)\, \frac{\partial}{\partial x^i}\bigg|_{\gamma(t)} \in T_{\gamma(t)}M$$

for all $t \in I$, where $\gamma^i = x^i(\gamma)$ for all $1 \le i \le n$.

Now that we have defined the velocity of smooth curves in smooth manifolds, we can relate our initial definition of tangent vectors in $\mathbb{R}^n$ in Remark 1.34 to tangent vectors for general smooth manifolds as follows.

**Lemma 1.104.** Let $M$ be a smooth manifold, $v \in T_pM$, and $f \in C^\infty(M)$. Then

$$v(f) = \frac{\partial(f \circ \gamma)}{\partial t}(0)$$

for every smooth curve $\gamma : I \to M$, $0 \in I$, with $\gamma(0) = p$ and $\gamma'(0) = v$.

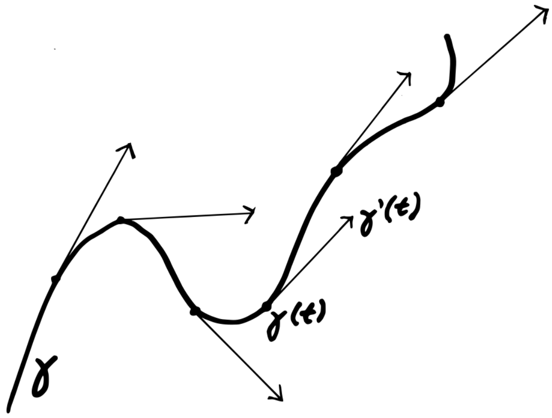


**Figure 20:** The velocity vector field of a curve $\gamma$.

*Proof.* Follows from the chain rule for differentials of smooth maps. $\square$

Recall the definition of an integral curve of a vector field on an open set of $\mathbb{R}^n$. There is, of course, a similar concept for general smooth manifolds.

**Definition 1.105.** Let $X \in \mathfrak{X}(M)$ be a smooth vector field on a smooth manifold $M$. An **integral curve** of $X$ at $p \in M$ is a smooth curve $\gamma : I \to M$, where $I \subset \mathbb{R}$ is an interval, $0 \in I$, such that $\gamma(0) = p$ and

$$\gamma'(t) = X_{\gamma(t)}$$

for all $t \in I$. An integral curve $\gamma : I \to M$ of $X$ is called **maximal** if there is no interval $\widetilde{I} \supset I$, such that $\widetilde{I} \setminus I \neq \emptyset$ and there exists an integral curve $\widetilde{\gamma} : \widetilde{I} \to M$ of $X$ with $\widetilde{\gamma}|_I = \gamma$. A vector field $X$ is called **complete** if every maximal integral curve $\gamma : I \to M$ is defined on $I = \mathbb{R}$.

If we omit the term "at $p$" for integral curves, we also drop the requirement $0 \in I$.

**Example 1.106.** Consider $X \in \mathfrak{X}(\mathbb{R}^2)$ given in canonical coordinates $(u^1, u^2) = (x, y)$ by

$$X = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}.$$

Its integral curves at any point $(x_0, y_0) \in \mathbb{R}^2$ are of the form

$$\gamma : t \mapsto \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

**Exercise 1.107.**

(i) Write down the vector field $X$ and its integral curves in Example 1.106 in polar coordinates. Is $X$ complete?

(ii) Construct a vector field on $\mathbb{R}^2 \setminus \{0\}$ with precisely one periodic maximal integral curve that is not a constant curve and no other periodic maximal integral curves.

**Remark 1.108.** While we have defined integral curves of vector fields, at this point we do not know if they always exists and whether they are unique or not. With the help of the theory of ordinary differential equations we obtain such results. Firstly note that locally, i.e. in any given local coordinates, the equation

$$\gamma' = X_\gamma$$

is an ordinary, in general non-linear, differential equation. Thus, for any given vector field $X \in \mathfrak{X}(M)$ and any $p \in M$ there exists an integral curve $\gamma : I \to M$ of $X$ at $p$. If $\gamma : I \to M$ and $\widetilde{\gamma} : \widetilde{I} \to M$ are two integral curves of $X$ at $p$, they coincide on $I \cap \widetilde{I}$ which, by definition, is never empty. For each $p \in M$, there exists a unique maximal integral curve of $X$ at $p$. Furthermore, the integral curves of $X$ at $p$ depend locally smoothly on $p \in M$. The proofs of these results need some care in case that an integral curve leaves a given coordinate neighbourhood, but they are otherwise identical to the case $M \subset \mathbb{R}^n$ open. For literature on the subject of ordinary differential equations and dynamical systems see e.g. [A1, A2]

In general it is a very difficult question whether a given vector field in $\mathfrak{X}(M)$ is complete, at least if $M$ is not compact. For compact smooth manifolds $M$ we have the following result.

**Proposition 1.109.** Vector fields on compact smooth manifolds are complete.

*Proof.* [A1] Chapter 2.6 together with the fact that since $M$ is compact, one can for any given atlas $\mathcal{A}$ on $M$ assume without loss of generality that $\mathcal{A}$ is finite, and we have that the closure of the chart neighbourhoods of $\mathcal{A}$ are compact in $M$. □

For $M$ not compact we still have the following result on vector fields with compact support.

**Proposition 1.110.** Let $X \in \mathfrak{X}(M)$ be a vector field with compact support, meaning that

$$\operatorname{supp}(X) = \overline{\{p \in M \mid X_p \neq 0\}} \subset M$$

is compact. Then $X$ is complete.

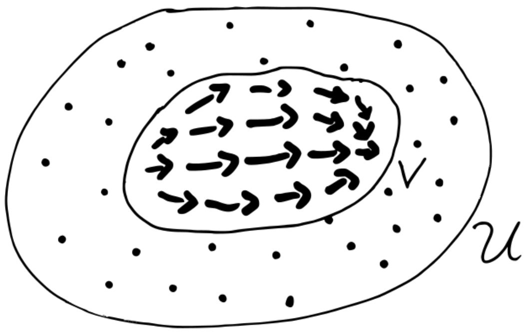


**Figure 21:** A sketch of a vector field with compact support $V \subset U$.

*Proof.* Exercise. [Hint: Try proving this for $M = \mathbb{R}^n$ first.] □

**Definition 1.111.** A **local one parameter group of diffeomorphisms** on a smooth manifold $M$ is a smooth map

$$\varphi : I \times U \to M, \quad (t,p) \mapsto \varphi_t(p),$$

such that $I \subset \mathbb{R}$ is an interval containing $0 \in \mathbb{R}$, $U \subset M$ is open, $\varphi_0 = \mathrm{id}_U$, $\varphi_t : M \to M$ is a diffeomorphism for all $t \in I$, and

$$\varphi_{s+t}(p) = \varphi_s(\varphi_t(p))$$

for all $p \in U$ and all $s, t \in I$ with $(s+t) \in I$ and $\varphi_t(p) \in U$. A **one parameter group of diffeomorphisms** is a local one parameter group of diffeomorphisms with $I = \mathbb{R}$ and $U = M$.

Local one parameter groups of diffeomorphisms on smooth manifolds are closely related to vector fields and their integral curves. For any given vector field on a smooth manifold we can attempt to consider all integral curves of $X$ "at once". This leads to the following concept.

**Definition 1.112.** A **local flow** of a vector field $X \in \mathfrak{X}(M)$ is a smooth map

$$\varphi : I \times U \to M, \quad (t, p) \mapsto \varphi_t(p),$$

for some interval $I \subset \mathbb{R}$ containing $0 \in \mathbb{R}$ and an open set $U \subset M$, such that $\varphi_0 = \mathrm{id}_U$ and for every $p \in U$ fixed, the smooth curve

$$t \mapsto \varphi_t(p)$$

is an integral curve of $X$. This just means that

$$\frac{\partial}{\partial t}(\varphi_t(p)) = X_{\varphi_t(p)}.$$

We say that a local flow $\varphi : I \times U \to M$ of $X$ is near a point $p \in M$ if $p \in U$. A local flow of $X$ is called **(global) flow** of $X$ if $I = \mathbb{R}$ and $U = M$.

**Lemma 1.113.** Every vector field on $M$ admits a local flow near any given point $p \in M$.

*Proof.* Let $X \in \mathfrak{X}(M)$ and $p \in M$ arbitrary but fixed. Choose a bump function $b : M \to \mathbb{R}$ such that on some open neighbourhood $U \subset M$ of $p$, $b|_U \equiv 1$. The maximal integral curves at $p$ of $bX$ are defined on $\mathbb{R}$ by Proposition 1.110 and depend smoothly on $p \in M$ by Remark 1.108. This already shows that vector fields with compact support admit a global flow $\varphi$. Fix $\varepsilon > 0$ and choose an open subset $V \subset U$, such that $V$ is an open neighbourhood of $p$ and for all $q \in V$ and all $t \in (-\varepsilon, \varepsilon)$ $\varphi_t(q) \in U$. Geometrically this mean that the set $V$ is not moved out of $U$ by the flow of $bX$ for $|t| < \varepsilon$. Since $X$ and $bX$ coincide on $U$, their integral curves at all $q \in V$ for $I = (-\varepsilon, \varepsilon)$ also coincide in $V$. Hence, the flow $\varphi$ of $bX$ restricted to $(-\varepsilon, \varepsilon) \times V$ is a local flow of $X$. $\square$

Observe that Definitions 1.111 and 1.112 have a certain similarity. They are connected as follows.

**Proposition 1.114.** Local flows of vector fields are local one parameter groups of diffeomorphisms.

*Proof.* It suffices to show that for a given vector field $X$ with two integral curves $\gamma : (a, b) \to M$ at $p = \gamma(0)$ and $\widetilde{\gamma} : (\widetilde{a}, \widetilde{b}) \to M$ with $\gamma(s) = \widetilde{\gamma}(0)$ for some $s \in (a, b)$ we have

$$\gamma(s + t) = \widetilde{\gamma}(t)$$

for all $t$, such that $(s + t) \in (a, b)$ and $t \in (\widetilde{a}, \widetilde{b})$. This means that $\widetilde{\gamma}$ extends $\gamma$ and follows from the fact that $t \mapsto \gamma(s + t)$ is an integral curve of $X$ (for $s$ small enough) and uniqueness of local solutions,

$$(\gamma(s + \cdot))'(t) \overset{\text{chain rule}}{=} \gamma'(s + t) = X_{\gamma(s+t)} = X_{(\gamma(s+\cdot))(t)}.$$

Hence, for a local flow $\varphi$ of $X$ near $p$ we obtain

$$\varphi_t(\phi_s(p)) = \varphi_t(\gamma(s)) = \widetilde{\gamma}(t) = \gamma(s + t) = \varphi_{s+t}(p).$$

$\square$

The following is an immediate consequence of Proposition 1.114.

**Corollary 1.115.** Assume that $X \in \mathfrak{X}(M)$ is complete. Then its flow is a one parameter group of diffeomorphisms.

In fact, one can prove that for any vector field $X \in \mathfrak{X}(M)$ the set

$$\bigcup_{p \in M} (I_p \times \{p\}) \subset (\mathbb{R} \times M)$$

is open, where $I_p$ is the uniquely determined interval for the maximal integral curve $\gamma : I_p \to M$ of $X$ starting at $\gamma(0) = p \in M$. For $X$ complete, $I_p = \mathbb{R}$ for all $p \in M$ and, hence, the maximal domain of definition of any local flow of $X$ is $\mathbb{R} \times M$, meaning $X$ has a global flow. For a proof see [G, S. 1.10.9].

**Example 1.116.** Translations in $\mathbb{R}^n$ are of the form $A_v : (p, v) \mapsto p + v$ where $v = (v^1, \dots, v^n)$ is the translation vector. Consider the constant vector field

$$\sum_{i=1}^{n} v^i \frac{\partial}{\partial u^i}$$

with global flow

$$\varphi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n, \quad (t, p) \mapsto p + tv.$$

We see that $\varphi_1(p) = A_v(p)$ for all $p \in \mathbb{R}^n$.

**Exercise 1.117.** Let $A \in \mathrm{SO}(2)$ be fixed. Find $X \in \mathfrak{X}(\mathbb{R}^2)$, such that its global flow $\varphi$ fulfils $\varphi_1(p) = Ap$ for all $p \in \mathbb{R}^2$. Can this always be achieved for any $A \in \mathrm{O}(2)$?

We have seen that local flows of vector fields are local one parameter groups of diffeomorphisms. The converse statement is also true.

**Definition 1.118.** Let $\varphi : I \times U \to M$ be a local one parameter group of diffeomorphisms. The **infinitesimal generator of** $\varphi$ is defined to be the map

$$U \ni p \mapsto \left.\frac{\partial}{\partial t}\right|_{t=0} (\varphi_t(p)) \in T_pM.$$

[Note: We have secretly used the upcoming Exercise 1.143, make sure you understand how and why.]

**Lemma 1.119.** Infinitesimal generators of local one parameter group of diffeomorphisms $\varphi : I \times U \to M$ are local vector fields in $\mathfrak{X}(U)$. Infinitesimal generators of one parameter groups of diffeomorphisms $\varphi : \mathbb{R} \times M \to M$ are complete.

*Proof.* Since any local one parameter group of diffeomorphisms is smooth, the map $X : p \mapsto X_p := \left.\frac{\partial}{\partial t}\right|_{t=0} (\varphi_t(p))$ is smooth, i.e. $X \in \mathfrak{X}(U)$. Hence, for any one parameter group of diffeomorphisms $\varphi : \mathbb{R} \times M \to M$, $X$ is a vector field in $\mathfrak{X}(M)$. Its integral curves at $p \in M$ are given by

$$t \mapsto \varphi_t(p)$$

and are defined for all $t \in \mathbb{R}$. This means that $X$ is complete. □

Recall the definition of the Lie bracket on $\mathfrak{X}(M)$, cf. Proposition 1.99. We know the algebraic motivation for it by considering vector fields as derivations of $C^\infty(M)$. But what does $[X, Y]$ for $X, Y \in \mathfrak{X}(M)$ stand for geometrically? To answer this question we must define the pushforward and pullback of vector fields under diffeomorphisms.

**Definition 1.120.** Let $F : M \to N$ be a diffeomorphism and let $X \in \mathfrak{X}(M)$, $Y \in \mathfrak{X}(N)$. The **pushforward** of $X$ under $F$ is the vector field $F_*X \in \mathfrak{X}(N)$ given by

$$(F_*X)_q := dF_{F^{-1}(q)}\left(X_{F^{-1}(q)}\right) \quad \forall q \in N.$$

The **pullback** of $Y$ under $F$ is the vector field $F^*Y \in \mathfrak{X}(M)$ given by

$$(F^*Y)_p := d(F^{-1})_{F(p)}\left(Y_{F(p)}\right) \quad \forall p \in M.$$

Note that $d(F^{-1})_{F(p)} = (dF_p)^{-1}$ for all $p \in M$.

**Exercise 1.121.** Verify that if $F : M \to N$ is a diffeomorphism and $\gamma$ is an integral curve of $X \in \mathfrak{X}(M)$, then $F \circ \gamma$ is an integral curve of $F_*X$. Formulate a version of this statement for local diffeomorphisms.

In order to explain the Lie bracket of vector fields geometrically, we need one more result about the local form of vector fields. Assume that $X \in \mathfrak{X}(M)$ does not vanish everywhere. Then near any point where $X$ does not vanish we can find local coordinates on $M$ in which $X$ has a particularly simple form.

**Proposition 1.122.** Let $X \in \mathfrak{X}(M)$ and $p \in M$, such that $X_p \neq 0$. Then there exist local coordinates on an open neighbourhood $U \subset M$ of $p$, such that $X$ is of the form

$$X_q = \left.\frac{\partial}{\partial x^1}\right|_q$$

for all $q \in U$.

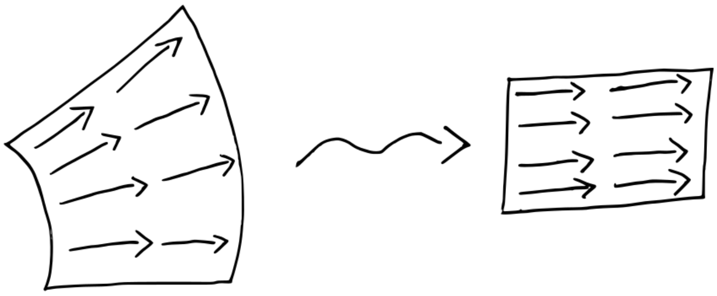

**Figure 22:** Locally rectifying a vector field.

*Proof.* Since $X$ is as a section in $TM$ it is in particular a continuous map and, hence, we can find an open neighbourhood $U$ of $p \in M$, such that $X_q \neq 0$ for all $q \in U$. Assume without loss of generality that $U$ is contained in a chart neighbourhood. Choose any local coordinate system $\phi = (y^1, \ldots, y^n)$ on $U$ and let $(u^1, \ldots, u^n)$ denote the canonical coordinates on $\mathbb{R}^n$. We can assume without loss of generality, after possibly shrinking $U$ and re-ordering the $y^i$'s, that $\phi_*(X) \in \mathfrak{X}(\phi(U))$ is transversal along the inclusion map $\{u^1 = 0\} \cap \phi(U) \hookrightarrow \mathbb{R}^n$ to $N := \{u^1 = 0\} \cap \phi(U)$, meaning that

$$(\phi_*X)_q \notin T_qN \cong T_q\{u^1 = 0\} \subset T_q\mathbb{R}^n$$

for all $q \in N$. Let, after again possibly shrinking $U$, $\Phi : I \times \phi(U) \to \mathbb{R}^n$ denote a local flow of $\phi_*X$. Since

$$(\phi_*X)_q = \left.\frac{\partial}{\partial t}\right|_{t=0} \Phi_t(q) \neq 0$$

by the transversality condition, we obtain with Theorem 1.55 after possibly shrinking $I$ that

$$F := \Phi|_{I\times N} : I \times N \to \Phi(I \times N)$$

is a diffeomorphism, where we understand $I$ as the "time" part so that $\Phi(t, q) := \Phi_t(q)$, and $\Phi(I \times N) \subset \mathbb{R}^n$ is open. Denoting the canonical coordinates in $I$ by $u^1$ and in $N$ by $(u^2, \ldots, u^n)$ (which is compatible with the canonical inclusion $I \times N \subset \mathbb{R}^n$), we in particular have

$$dF_{(u^1,u^2,\ldots u^n)}\left(\left.\frac{\partial}{\partial u^1}\right|_{(u^1,u^2,\ldots,u^n)}\right) = (\phi_* X)_{\Phi(u^1,u^2,\ldots,u^n)}$$

for all $(u^1, \ldots, u^n) \in I \times N$. Now we can define coordinates on $\phi^{-1}(\phi(U) \cap \Phi(I \times N)) \subset M$ by setting

$$\psi = (x^1, \ldots, x^n) := F^{-1} \circ \phi : \phi^{-1}(\phi(U) \cap \Phi(I \times N)) \to F^{-1}(\phi(U) \cap (I \times N)) \subset \mathbb{R}^n$$

and obtain for the local formula of $X$ in the local coordinate system $\psi$ and all $q \in \phi^{-1}(\phi(U) \cap \Phi(I \times N))$

$$X_q = \left.\frac{\partial}{\partial x^1}\right|_q .$$

□

In local coordinates as the ones constructed in Proposition 1.122, local flows look particularly simple.

**Corollary 1.123.** Any local flow of $X$ near $p$ as in Proposition 1.122 is, if $X_p \neq 0$, in the local coordinate system $\psi = (x^1, \ldots, x^n)$ of the form

$$\psi(\varphi_t(q)) = \psi(q) + te_1,$$

for all $q \in U$, where $e_1$ denotes the first unit vector in $\mathbb{R}^n$ in canonical coordinates, for $|t|$ small enough. Furthermore

$$d\varphi_t\left(\left.\frac{\partial}{\partial x^i}\right|_q\right) = \left.\frac{\partial}{\partial x^i}\right|_{\psi^{-1}(\psi(q)+te_1)}$$

for all $q \in U$ and $t$ small enough, where we understand the differential of $\varphi_t$ for $t$ fixed.

Next we will describe how the Lie algebra structure on vector fields is connected to their local flows. To do so we will need to introduce the following concept.

**Definition 1.124.** Let $M$ and $N$ be smooth manifolds $\phi : M \to N$ be a smooth map. Two vector fields $X \in \mathfrak{X}(M)$ and $\overline{X} \in \mathfrak{X}(N)$ are called **$\phi$-related** if $d\phi(X) = \overline{X}_\phi$. One then writes $X \sim_\phi \overline{X}$. Equivalently, $X \sim_\phi \overline{X}$ if $X(f \circ \phi) = \overline{X}(f) \circ \phi$ for all $f \in C^\infty(N)$.

We see that for $\phi : M \to N$ an embedding and any $X \in \mathfrak{X}(M)$, $d\phi(X)$, viewed as vector field along $\phi$, can be locally extended to a smooth vector field $\overline{X} \in \mathfrak{X}(N)$, such that $d\phi(X) = \overline{X}_\phi$. This means that, locally, we can find a $\phi$-related vector field to $X$. For the next lemma, the motivation is the case where $\phi$ is a transition function describing a change of coordinates on a smooth manifold.

**Lemma 1.125.** Let $\phi : M \to N$ be a smooth map, $X, Y \in \mathfrak{X}(M)$, and $\overline{X}, \overline{Y} \in \mathfrak{X}(N)$, such that $X \sim_\phi \overline{X}$ and $Y \sim_\phi \overline{Y}$. Then $[X, Y] \sim_\phi [\overline{X}, \overline{Y}]$.

*Proof.* Let $f \in C^\infty(N)$ be arbitrary. Then

$$\begin{aligned}
[X,Y](f \circ \phi) &= X(Y(f \circ \phi)) - Y(X(f \circ \phi)) \\
&= X(\overline{Y}(f) \circ \phi) - Y(\overline{X}(f) \circ \phi) \\
&= (\overline{X}(\overline{Y}(f)) - \overline{Y}(\overline{X}(f))) \circ \phi.
\end{aligned}$$

□

Lemma 1.125 means for $\phi$ a change of coordinates that the Lie algebra structure on vector fields is compatible with changing coordinates in the sense that their Lie brackets are also related by the same change of coordinates. Globally we obtain the following result.

**Corollary 1.126.** For any given diffeomorphism $F : M \to N$,

$$F_*[X,Y] = [F_*X, F_*Y]$$

for all $X, Y \in \mathfrak{X}(M)$ and

$$F^*[\overline{X}, \overline{Y}] = [F^*\overline{X}, F^*\overline{Y}]$$

for all $\overline{X}, \overline{Y} \in \mathfrak{X}(N)$.

*Proof.* Follows from Definition 1.120 and Lemma 1.125. □

**Remark 1.127.** In the case that $\phi$ is an embedding and $\dim(M) < \dim(N)$, Lemma 1.125 also implies that (locally and globally) $[\overline{X}, \overline{Y}] \circ \phi$ does **not** depend on the (local) extensions of $\overline{X} \circ \phi$ and $\overline{Y} \circ \phi$ to vector fields on in $N$ open neighbourhoods of points in $\phi(M)$.

Means: Vector fields tangent to a submanifold have Lie bracket tangent to submanifold.

Now we have cleared up all technical difficulties we can proof the following statement.

**Proposition 1.128.** Let $X, Y \in \mathfrak{X}(M)$ and for $p \in M$ arbitrary but fixed let $\varphi : I \times U \to M$ be a local flow of $X$ near $p$. Then

$$[X,Y]_p = \frac{\partial}{\partial t}\bigg|_{t=0} (\varphi_t^* Y)_p.$$

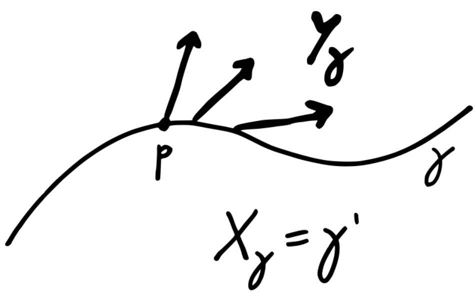


**Figure 23:** A sketch of $Y$ along an integral curve $\gamma$ through $p$ of $X$.

*Proof.* First observe that the right hand side of the above formula is actually a well-defined expression. This follows from $(\varphi_t^* Y)_p \in T_pM$ for all $t \in I$ and the fact that $T_pM$ is a real vector space. In the following, $d\varphi_t$ is to be understood as the differential of $\varphi_t$ for $t \in I$ fixed. First assume that $X_p \neq 0$. Without loss of generality we can, with the help of Proposition 1.122, assume that we have chosen local coordinates $(x^1, \ldots, x^n)$ on $U \subset M$ with $p \in M$, such that

$X_q = \frac{\partial}{\partial x^1}\Big|_q$ for all $q \in U$. Recall that, since $\varphi$ is a local one parameter group of diffeomorphisms, $\varphi_{-t} = \varphi_t^{-1}$ whenever defined. Hence, for $|t|$ small enough we have

$$(\varphi_t^* Y)_p = d(\varphi_t^{-1})_{\varphi_t(p)} \left(Y_{\varphi_t(p)}\right) = (d\varphi_{-t})_{\varphi_t(p)} \left(Y_{\varphi_t(p)}\right).$$

Observe that

$$(d\varphi_{-t})_{\varphi_t(p)} : \frac{\partial}{\partial x^i}\Big|_{\varphi_t(p)} \mapsto \frac{\partial}{\partial x^i}\Big|_{\psi^{-1}(\psi(\varphi_t(p))-te_1)} = \frac{\partial}{\partial x^i}\Big|_{\psi^{-1}(\psi(p)+te_1-te_1)} = \frac{\partial}{\partial x^i}\Big|_p.$$

In the local coordinates $(x^1, \dots, x^n)$, $Y$ is of the form

$$Y_q = \sum_{i=1}^n Y^i(q) \, \frac{\partial}{\partial x^i}\Big|_q$$

for all $q \in U$. Thus

$$(\varphi_t^* Y)_p = \sum_{i=1}^n Y^i(\varphi_t(p)) \, \frac{\partial}{\partial x^i}\Big|_p$$

and, hence,

$$\frac{\partial}{\partial t}\Big|_{t=0} (\varphi_t^* Y)_p = \sum_{i=1}^n dY^i \left(\frac{\partial}{\partial x^1}\Big|_p\right) \frac{\partial}{\partial x^i}\Big|_p$$

which coincides with $[X,Y]_p = \left[\frac{\partial}{\partial x^1}, Y\right]_p$ by Lemma 1.100 and Exercise 1.101.

Next assume that $X_p = 0$. If $X_q = 0$ for all $q$ in an open neighbourhood $U$ of $p$, the local flow of $X$ restricted to $U$ will be the identity for all $t \in I$. Hence,

$$\frac{\partial}{\partial t}\Big|_{t=0} (\varphi_t^* Y)_p = 0.$$

For any $f \in C^\infty(M)$ observe that $X(f)$ vanishes on $U$ and thus

$$[X,Y]_p(f) = X_p(Y(f)) - Y_p(X(f)) = 0.$$

Lastly assume that $X_p = 0$ and $X$ does not vanish identically on some open neighbourhood of $p$. Let $U \subset M$ be a compactly embedded open neighbourhood of $p$ and choose a sequence $\{p_n\}_{n\in\mathbb{N}}$, $\lim_{n\to\infty} p_n = p$, such that $X_{p_n} \neq 0$ and $p_n \neq p$ for all $n \in \mathbb{N}$. Then

$$[X,Y]_{p_n} = \frac{\partial}{\partial t}\Big|_{t=0} (\varphi_t^* Y)_{p_n}$$

for all $n \in \mathbb{N}$. Using the continuity in the base point of both sides of the above expression we take their respective limit as $n \to \infty$ and obtain that $[X,Y]_p = \frac{\partial}{\partial t}\Big|_{t=0} (\varphi_t^* Y)_p$ as claimed $\square$

Proposition 1.128 gives an answer to the question what the lie bracket of two vector fields should mean geometrically: $[X,Y]$ measures the infinitesimal change of $Y$ along integral curves of $X$ or, by skew-symmetry, the negative infinitesimal change of $X$ along integral curves of $Y$, both via the pullback. This motivates the following definition.

**Definition 1.129.** The **Lie derivative** of a vector field $Y \in \mathfrak{X}(M)$ with respect to[16] $X \in \mathfrak{X}(M)$ is defined as

$$\mathcal{L}_X(Y) := [X,Y] \in \mathfrak{X}(M).$$

[16] Or: "in direction of".

We will see in Section 2.2 that there is a different and very important alternative concept how to measure infinitesimal changes of vector fields or, more general, sections of vector bundles. Next we will study how to obtain, in a natural way, new vector bundles from given bundles. This will allow us to define what tensor fields should be, which are central objects in every flavour of differential geometry and in applications in physics.

**Definition 1.130.** Let $\pi_E : E \to M$ be a vector bundle of rank $k$. The **dual vector bundle** $\pi_{E^*} : E^* \to M$ is pointwise given by

$$\pi_{E^*}^{-1}(p) = E_p^* := \mathrm{Hom}_{\mathbb{R}}(E_p, \mathbb{R})$$

for all $p \in M$. The topology, smooth manifold structure, and bundle structure on $E^*$ are obtained as follows. Let $\{(\psi_i, V_i) \mid i \in A\}$ be a collection of local trivializations of the vector bundle $E$ of rank $k$, such that there exists an atlas $\mathcal{A} = \{(\varphi_i, \pi_E(V_i)) \mid i \in A\}$ of $M$.[17] Then $\mathcal{B} := \{((\varphi_i \times \mathrm{id}_{\mathbb{R}^k}) \circ \psi_i, V_i) \mid i \in A\}$ is an atlas on $E$. Recall that for any finite dimensional real vector space $W$, $(W^*)^*$ and $W$ are isomorphic via

$$W \ni v \mapsto (\omega \mapsto \omega(v)), \ \omega \in W^*.$$

The topology on $E^*$ is given by pre-images of open images of the **dual local trivializations** which are defined by

$$\widetilde{\psi_i} : \pi_{E^*}^{-1}(\pi_E(V_i)) \to \pi_E(V_i) \times \mathbb{R}^k, \quad \omega_p \mapsto (p, w),$$

where $w \in \mathbb{R}^k$ is the unique vector, such that $\omega_p(v_p) = \langle w, \mathrm{pr}_{\mathbb{R}^k}(\psi_i(v_p))\rangle$ for all $v_p \in \pi_E^{-1}(p)$ and $\langle \cdot, \cdot \rangle$ denotes the Euclidean scalar product on $\mathbb{R}^k$ induced by its canonical coordinates. The **dual atlas** $\mathcal{B}^*$ on $E^*$ is then defined by

$$\mathcal{B}^* := \{((\varphi_i \times \mathrm{id}_{\mathbb{R}^k}) \circ \widetilde{\psi_i}, V_i) \mid i \in A\}.$$

It follows that $E^* \to M$ is a vector bundle of rank $k$.

**Exercise 1.131.** Show that the transition functions of $E^* \to M$ fulfil

$$\widetilde{\psi_i} \circ \widetilde{\psi_j}^{-1} : (p, w) \mapsto \left(p, (A_p^{-1})^T w\right)$$

for all $p \in \pi_E(V_i)$, where $A : \pi_E(V_i) \to \mathrm{GL}(n)$ is given by the transition functions of $E \to M$,

$$\psi_i \circ \psi_j^{-1} : (p, v) \mapsto (p, A_p v).$$

**Exercise 1.132.** Show that $(E^*)^* \to M$ is isomorphic to $E \to M$ as a vector bundle for any vector bundle $E \to M$.

The most important example of a dual bundle is the dual to the tangent bundle of a smooth manifold (at least in this course).

**Definition 1.133.** The vector bundle $T^*M := (TM)^* \to M$ is called the **cotangent bundle** of $M$. Pointwise we denote $T_p^*M := (TM)_p^*$ for all $p \in M$. As for the tangent bundle we identify for any $U \subset M$ open and $p \in U$ the vector spaces $T_p^*U \cong T_p^*M$ via the inclusion map.

Similar to the tangent bundle, cf. Proposition 1.76, an atlas on $M$ induces an atlas on the total space $T^*M$ that is compatible with the bundle structure of $T^*M$ as the dual bundle of $TM$. We will specify how a given local coordinate system $\varphi = (x^1, \ldots, x^n)$ on an open set $U \subset M$ induces a local coordinate system on the total space $T^*M$. Let $\pi_{T^*M} : T^*M \to M$ denote the

[17] Exercise: Show that such a choice is always possible.

projection. In the tangent bundle case we used equation (1.10) to define the induced local charts. This definition uses the action of tangent vectors $v \in T_pM$, $p \in U$, on the coordinate functions $(x^1, \ldots, x^n)$. In our present case of the cotangent bundle, elements in $T_p^*M$ are linear maps

$$\omega \in T_p^*M, \quad \omega : T_pM \to \mathbb{R}, \quad \omega : v \mapsto \omega(v) \in \mathbb{R} \quad \forall v \in T_pM.$$

Thus we cannot let them act in any sensible way on coordinate functions. However, recall that the coordinate functions induce a basis of $T_pM$ for all $p \in U$, cf. Proposition 1.47. This motivates defining a local coordinate system on $T^*M$ by

$$\widetilde{\psi} : \pi_{T^*M}^{-1}(U) \to \varphi(U) \times \mathbb{R}^n, \quad \widetilde{\psi} : \omega \mapsto \left( \varphi(\pi_{T^*M}(\omega)), \omega\left( \left.\frac{\partial}{\partial x^1}\right|_p \right), \ldots, \omega\left( \left.\frac{\partial}{\partial x^n}\right|_p \right) \right). \tag{1.13}$$

This local coordinate system is dual to the local coordinate system $\psi$ on $\pi_{TM}^{-1}$ as in equation (1.10) in the sense that for all $\omega_p \in T_p^*M$ and all $v_p \in T_pM$

$$\omega_p(v_p) = \langle \mathrm{pr}_{\mathbb{R}^n}(\widetilde{\psi}(\omega_p)), \mathrm{pr}_{\mathbb{R}^n}(\psi(v_p)) \rangle, \tag{1.14}$$

where $\mathrm{pr}_{\mathbb{R}^n}$ denotes the canonical projection to the vector part and $\langle \cdot, \cdot \rangle$ denotes the Euclidean scalar product induced by the canonical coordinates on $\mathbb{R}^n$. The independence of the chosen local coordinate system of the right hand side of (1.14) follows from Exercise 1.131.

We have defined vector fields as sections in the tangent bundle of a smooth manifold. Sections in the cotangent bundle are of the same importance as vector fields when studying smooth manifolds.

**Definition 1.134.** Sections in $T^*M \to M$ are called **1-forms** and are denoted by $\Omega^1(M) := \Gamma(T^*M)$. For $U \subset M$ open, sections in $\Gamma(T^*M|_U)$ are denoted by $\Omega^1(U)$ and are called **local 1-forms**.

A straightforward way of obtaining explicit examples of 1-forms works as follows.

**Example 1.135.** Let $f \in C^\infty(M)$. Then the **differential**[18] **of** $f$, $df \in \Omega^1(M)$, is given by

$$df : p \mapsto df_p.$$

In local coordinates $(x^1, \ldots, x^n)$ we have $df\left(\frac{\partial}{\partial x^i}\right) = \frac{\partial f}{\partial x^i}$ for all $1 \leq i \leq n$. This implies that $df$ can locally be written as

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x^i}\, dx^i.$$

In particular it follows for $f = x^j$, $1 \leq j \leq n$, that the **coordinate 1-forms** $dx^j$ fulfil $dx^j\left(\frac{\partial}{\partial x^i}\right) \equiv \delta_i^j$ on the domain of definition of the local coordinates. This in in accordance with the pointwise version of this statement in Example 1.51.

**Lemma 1.136.** Let $M$ be a smooth manifold and let $\varphi = (x^1, \ldots, x^n)$ be a local coordinate system defined on an open set $U \subset M$. Then

$$\{dx^i_p \mid 1 \leq i \leq n\}$$

is a basis of $T_p^*M$ for all $p \in M$. It is precisely the dual basis to the basis $\left\{ \left.\frac{\partial}{\partial x^i}\right|_p \;\middle|\; 1 \leq i \leq n \right\}$ of $T_pM$. Any local 1-form $\omega \in \Omega^1(U)$ can be written as

$$\omega = \sum_{i=1}^n f_i\, dx^i \tag{1.15}$$

with uniquely determined smooth functions $f_i \in C^\infty(U)$ for $1 \leq i \leq n$.

[18] Cf. Definition 1.94.

*Proof.* The first two claims follow from Proposition 1.47 and Exercise 1.51. Next we observe that for all $f_i \in C^\infty(U)$, $1 \le i \le n$, the right hand side of equation (1.15) is a local section in $T^*M \to M$[19] by the construction of the smooth manifold structure on the total space $T^*M$ via charts of the form (1.13), which in particular implies that each $dx^i$, $1 \le i \le n$, is a local 1-form. On the other hand for a given local 1-form $\omega \in \Omega^1(U)$ define

$$\omega_i := \omega\left(\frac{\partial}{\partial x^i}\right)$$

for all $1 \le i \le n$. It now suffices to show that $\omega_i \in C^\infty(U)$ and, after that, to define $f_i := \omega_i$. $\omega_i$ being a local smooth function follows from observing that by equation (1.14), $\omega_i \circ \varphi^{-1}$ is precisely the $i$-th entry in the vector part of $\widetilde{\psi} \circ \omega \circ \varphi^{-1}$ and thereby by definition a smooth map. Uniqueness of the $f_i$ can be shown as follows. Suppose that locally

$$\omega = \sum_{i=1}^n f_i dx^i = \sum_{i=1}^n \widetilde{f}_i dx^i \tag{1.16}$$

such that for at least one $1 \le j \le n$, $f_j \neq \widetilde{f}_j$. Choose $p \in U$, such that $f_j(p) \neq \widetilde{f}_j(p)$. Then

$$\left(\sum_{i=1}^n f_i dx^i\right)\left(\left.\frac{\partial}{\partial x^j}\right|_p\right) = f_j(p) \neq \widetilde{f}_j(p) = \left(\sum_{i=1}^n \widetilde{f}_i dx^i\right)\left(\left.\frac{\partial}{\partial x^j}\right|_p\right)$$

which is a contradiction. □

In order to check whether a fibrewise map $\omega : M \to T^*M$ is a 1-form it suffices to check how it behaves when applied to vector fields. The converse statement also holds true.

**Lemma 1.137.** Let $\omega : M \to T^*M$, $\omega : p \mapsto T_p^*M$, be a fibrewise map. Then $\omega \in \Omega^1(M)$ if and only if for all $X \in \mathfrak{X}(M)$ the function $\omega(X) : M \to \mathbb{R}$, $p \mapsto \omega(X)(p)$ for all $p \in M$, is smooth.

*Proof.* Exercise. [Hint: Use Lemma 1.136 and bump functions.] □

Example 1.135 and Lemma 1.136 motivate viewing the coordinate 1-forms as dual objects to coordinate vector fields. Indeed we obtain the following more abstract statement reinforcing this point of view.

**Proposition 1.138.** $\Omega^1(M)$ is isomorphic as a $C^\infty(M)$-module to the $C^\infty(M)$-module dual to $\mathfrak{X}(M)$, i.e.

$$\Omega^1(M) \cong \mathrm{Hom}_{C^\infty(M)}(\mathfrak{X}(M), C^\infty(M)).$$

*Proof.* Let $\alpha \in \Omega^1(M)$. we have seen in the proof of Lemma 1.136 that for any choice of local coordinates $(x^1, \dots, x^n)$ on $M$, $\alpha\left(\frac{\partial}{\partial x^i}\right)$ is a local smooth function. Recall Proposition 1.89 and choose a locally finite countable partition of unity $\{b_i : U_i \to [0,1] \mid i \in I\}$ subordinate to a countable atlas $\{(\varphi_i = (x_i^1, \dots, x_i^n), U_i) \mid i \in I\}$ of $M$. Write $X(p) = \sum_{i \in I} b_i(p)X(p)$ and observe that this sum is finite for all fixed $p \in M$ and that $b_i X \in \mathfrak{X}(U_i)$ for all $i \in I$. We can write $b_i X$ in local coordinates as

$$b_i X = \sum_{j=1}^n b_i X(x_i^j)\frac{\partial}{\partial x_i^j} =: \sum_{j=1}^n b_i X_i^j \frac{\partial}{\partial x_i^j}.$$

[19] Recall that our definition of sections required them to be smooths maps.

Since all $b_i$, $i \in I$, are in particular bump functions, $b_i X_i^j \in C^\infty(U_i)$ can be trivially extended to a smooth function on $M$ for all $i \in I$ and all $1 \leq j \leq n$. Hence, $b_i X \in \mathfrak{X}(U_i)$ can be trivially extended to be a vector field on the whole manifold $M$ for all $i \in I$. We define

$$A_\alpha(X) := \alpha(X) = \sum_{i \in I} \alpha(b_i X) = \sum_{i \in I} \sum_{j=1}^{n} b_i \alpha_j X_i^j. \tag{1.17}$$

The sum on the right hand side of (1.17) is a locally finite sum of bump functions (defined on the respective $U_i$ which we trivially extent to $M$). This means that for all $p \in M$ fixed there exists an open neighbourhood $U \subset M$ of $p$, such that the set

$$\{(i,j) \in I \times \{1, \ldots, n\} \mid b_i \alpha_j X_i^j(q) \neq 0 \text{ for at least one } q \in U\}$$

is finite. Hence, the right hand side of (1.17) is indeed a smooth function defined on $M$ since it is, locally, the sum of finitely many smooth functions. This shows $\alpha$ defines a well-defined $C^\infty(M)$-linear map

$$A_\alpha : \mathfrak{X}(M) \to C^\infty(M), \quad X \mapsto \alpha(X), \; A(X)(p) := \alpha_p(X_p) \; \forall p \in M.$$

On the other hand let $A : \mathfrak{X}(M) \to C^\infty(M)$ be a $C^\infty(M)$-linear map. For a given $A \in \mathrm{Hom}_{C^\infty(M)}(\mathfrak{X}(M), C^\infty(M))$ define a fibre-preserving map

$$\alpha_A : M \to T^*M, \quad \alpha_A\big|_p(v) := A(X)(p)$$

for $X \in \mathfrak{X}(M)$ with $X_p = v$. We need to show that this definition does not depend on the choice of $X$ and that $\alpha_A$ is in fact a smooth map. Since $A(X+Y)(p) = A(X)(p) + A(Y)(p)$ for all $X, Y \in \mathfrak{X}(M)$ and all fixed $p \in M$ it suffices to show that $A(X)(p) = 0$ for all $X \in \mathfrak{X}(M)$ with $X_p = 0$. Let $(x^1, \ldots, x^n)$ be local coordinates on $U \subset M$ with $p \in U$ so that $X$ with $X_p = 0$ is locally of the form

$$X = \sum_{i=1}^{n} X^i \frac{\partial}{\partial x^i},$$

$X^i \in C^\infty(U)$ for all $1 \leq i \leq n$. Then $X_p = 0$ precisely means that $X^i(p) = 0$ for all $1 \leq i \leq n$. Since $A(fX)(p) = f(p)A(X)(p)$ for all $f \in C^\infty(M)$, we can choose a bump function $b \in C^\infty(M)$, so that $\mathrm{supp}(b) \subset U$ is compactly embedded and such that there exists a compactly embedded set $V \subset U$ with non-empty interior containing $p$ and $b|_V \equiv 1$. Then for all $1 \leq i \leq n$, $bX^i$ is also a bump function on $U \subset M$ and can thus be smoothly extended to $M$ (note that the $X^i$ might have "bad" behaviour when approaching $\partial U$, e.g. do not converge). Furthermore, $b\frac{\partial}{\partial x^i}$ can be extended to a globally defined vector field on $M$ for all $1 \leq i \leq n$ by setting

$$b \frac{\partial}{\partial x^i}\bigg|_q = 0$$

for all $q \in M \setminus U$. We will for simplicity write $bX^i \in C^\infty(M)$ and $b\frac{\partial}{\partial x^i} \in \mathfrak{X}(M)$ for all $1 \leq i \leq n$. We now calculate

$$\begin{aligned} A(X)(p) &= b^2(p)A(X)(p) = A(b^2X)(p) = A\left(\sum_{i=1}^{n}(bX^i)\left(b\frac{\partial}{\partial x^i}\right)\right)(p) \\ &= \left(\sum_{i=1}^{n} bX^i A\left(b\frac{\partial}{\partial x^i}\right)\right)(p) = \sum_{i=1}^{n} X^i(p) A\left(b\frac{\partial}{\partial x^i}\right)(p) = 0. \end{aligned}$$

It remains to show that $\alpha_A$ is smooth. This follows with the help of Lemma 1.137, the above result, and a similar construction using a locally finite partition of unity as for the other direction of the proof. One can further check that $A_{\alpha_A} = A$ and $\alpha_{A_\alpha} = \alpha$, that is that the two constructions are inverse to each other. □

**Exercise 1.139.** Work out the missing details of the "$\supset$" direction in the proof of Proposition 1.138.

As for vector fields, cf. Definition 1.120, we can define the pullback and pushforward of 1-forms under diffeomorphisms. Be wary of the differences!

**Definition 1.140.** Let $F : M \to N$ be a diffeomorphism and let $\alpha \in \Omega^1(M)$, $\beta \in \Omega^1(N)$. The **pushforward** of $\alpha$ under $F$ is the 1-form $F_*\alpha \in \Omega^1(N)$ given by

$$(F_*\alpha)_q := \alpha_{F^{-1}(q)} \circ d(F^{-1})_q \quad \forall q \in N.$$

The **pullback** of $\beta$ under $F$ is the 1-form $F^*\beta \in \Omega^1(M)$ given by

$$(F^*\beta)_p := \beta_{F(p)} \circ dF_p \quad \forall p \in M.$$

The above compositions denote compositions of linear maps.

**Remark 1.141.** The pullback of a 1-form $\beta \in \Omega^1(N)$ is well-defined for any smooth map $F : M \to N$.

Next we will study some constructions on how to obtain new vector bundles from given vector bundles.

**Definition 1.142.** Let $\pi_E : E \to M$ be a vector bundle of rank $k$ and $\pi_F : F \to M$ a vector bundle of rank $\ell$ over an $n$-dimensional smooth manifold $M$. The **Whitney**[20] **sum of $E$ and $F$** is the the direct sum of the two vector bundles $\pi_{E\oplus F} : E \oplus F \to M$ with fibres

$$(E \oplus F)_p = \pi_{E\oplus F}^{-1}(p) := E_p \oplus F_p$$

for all $p \in M$. The structure of a vector bundle on $E \oplus F = \bigsqcup_{p\in M} (E_p \oplus F_p)$ is then explained by Proposition 1.74 and the requirement that the following maps are local trivializations of $E \oplus F$. Let $\{(\psi_i^E, V_i^E) \mid i \in I\}$ and $\{(\psi_i^F, V_i^F) \mid i \in I\}$ be coverings of local trivializations of $E$ and $F$, respectively, such that $U_i := \pi_E(V_i^E) = \pi_F(V_i^F)$ for all $i \in I$ and such that there exists an atlas $\mathcal{A} = \{(\varphi_i, U_i) \mid i \in I\}$ of $M$. We require now require that with

$$\begin{aligned} \phi_i^{-1} &:= (\psi_i^E \oplus \psi_i^F)^{-1} \circ (\Delta_M \times \mathrm{id}_{\mathbb{R}^{k+\ell}}) : U_i \times \mathbb{R}^{k+\ell} \cong U_i \times (\mathbb{R}^k \times \mathbb{R}^\ell) \to \bigsqcup_{p\in U_i} (E_p \oplus F_p), \\ (p, v, w) &\mapsto (\psi_i^E)^{-1}(p, v) \oplus (\psi_i^F)^{-1}(p, w) \quad \forall p \in U_i,\ v \in \mathbb{R}^k,\ w \in \mathbb{R}^\ell, \end{aligned} \tag{1.18}$$

where $\Delta_M : p \mapsto (p, p) \in M \times M$ denotes the diagonal embedding and

$$\mathbb{R}^k \times \mathbb{R}^\ell \ni (v, w) \mapsto \begin{pmatrix} v \\ w \end{pmatrix} \in \mathbb{R}^{k+\ell}$$

the linear isomorphism, all $\phi_i$, $i \in I$, are inverses of local trivializations covering $E \oplus F$. In order to use Proposition 1.74 we need to check that the transition functions have the required form. We obtain that for all $i, j \in I$, such that $U_i \cap U_j \neq \emptyset$,

$$\phi_i \circ \phi_j^{-1}(p, v, w) = (p, \tau_{ij}^E(p)v, \tau_{ij}^F(p)w),$$

where $\tau_{ij}^E$ and $\tau_{ij}^F$ are the transition functions of the local trivializations of $E$ and $F$, respectively. Lastly, we simply need to define

$$\tau_{ij}^{E\oplus F}(p) := \left(\begin{array}{c|c} \tau_{ij}^E(p) & 0 \\ \hline 0 & \tau_{ij}^F(p) \end{array}\right) \in \mathrm{GL}(k + \ell)$$

so that we can write $\phi_i \circ \phi_j^{-1}\left(p, \left(\begin{smallmatrix} v \\ w \end{smallmatrix}\right)\right) = \left(p, \tau_{ij}^{E\oplus F}(p) \left(\begin{smallmatrix} v \\ w \end{smallmatrix}\right)\right)$. Now all requirements in Proposition 1.74 are fulfilled and we conclude that $E \oplus F \to M$ is, indeed, a vector bundle of rank $k + \ell$.

[20] Hassler Whitney (1907 – 1989)

**Exercise 1.143.** Show that the vector bundles $T(M \times N) \to M \times N$ and

$$TM \oplus TN \to M \times N, \quad (v_p, w_q) \mapsto (p, q) \quad \forall v_p \in T_pM \ \forall w_q \in T_qN,$$

are isomorphic as vector bundles for any two smooth manifolds $M$ and $N$.

A construction similar to the Whitney sum is the tensor product of vector bundles. Recall that the tensor product of two real vector spaces $V_1$ of dimension $n$ and $V_2$ of dimension $m$ is a real vector space $V_1 \otimes V_2$ together with a bilinear map $\otimes : V_1 \times V_2 \to V_1 \otimes V_2$, such that for every real vector space $W$ and every bilinear map $F : V_1 \times V_2 \to W$, there exist a unique linear map $\widetilde{F} : V_1 \otimes V_2 \to W$ making the diagram

$$\begin{tikzcd} V_1 \times V_2 \arrow[d, "\otimes"] \arrow[dr, "F"] & \\ V_1 \otimes V_2 \arrow[r, dashed, "\widetilde{F}"] & W \end{tikzcd}$$

commute. The dimension of $V_1 \otimes V_2$ is $n \cdot m$. If $\{v_1^1, \dots, v_1^n\}$ and $\{v_2^1, \dots, v_2^m\}$ are bases of $V_1$ and $V_2$, respectively, we can construct a choice of basis for $V_1 \otimes V_2$ explicitly. A basis of $V_1 \otimes V_2$ is given by $\{v_1^i \otimes v_2^j,\ 1 \le i \le n,\ 1 \le j \le m\}$, and $\widetilde{F}$ for a bilinear map $F$ as above is given by

$$\widetilde{F} : v_1^i \otimes v_2^j \mapsto F(v_1^i, v_2^j)$$

on the basis vectors. By considering "$\otimes$" itself as a bilinear map from $V_1 \times V_2$ to $W = V_1 \otimes V_2$, we define[21] $v \otimes w$ for $v = \sum_{i=1}^{n} v^i v_1^i$, $w = \sum_{j=1}^{m} w^j v_2^i$, as

$$v \otimes w := \sum_{i=1}^{n} \sum_{j=1}^{m} v^i w^j \cdot v_1^i \otimes v_2^j.$$

An element $v \in V_1 \otimes V_2$ is called a **pure tensor** if it can be written as $v = v_1 \otimes v_2$ for some $v_1 \in V_1$ and $v_2 \in V_2$. In order to describe any linear map $L : V_1 \otimes V_2 \to W$ is suffices to know how it acts on pure tensors [Exercise: Prove the last statement.]

A very important example that you should keep in mind is the tensor product of a real vector space $V$ with its dual, that is $V \otimes V^*$.

**Exercise 1.144.**

(i) Show that the real vector space of endomorphisms $\mathrm{End}(V)$ and $V \otimes V^*$ are isomorphic as real vector spaces via

$$V \otimes V^* \ni v \otimes \omega \mapsto (u \mapsto \omega(u)v) \in \mathrm{End}(V).$$

(ii) Show that $V \otimes \mathbb{R} \cong V$ and $V \otimes W \cong W \otimes V$ for all real vector spaces $V$ and $W$.

For the **evaluation map**

$$\mathrm{ev} : V \times V^* \to \mathbb{R}, \quad (v, \omega) \mapsto \omega(v) \quad \forall v \in V,\ \omega \in V^*,$$

the induced map $\widetilde{\mathrm{ev}} : V \otimes V^* \to \mathbb{R}$ is called **contraction**. By saying that we contract $v \otimes \omega$ we simply mean sending it to $\omega(v)$. Further recall that $V_1 \otimes (V_2 \otimes V_3)$ and $(V_1 \otimes V_2) \otimes V_3$ are

[21] Make sure to understand why this is consistent with the definition of the tensor product.

isomorphic. In the following we will deal with objects that, pointwise, are elements of vector spaces of the form

$$\underbrace{V \otimes \ldots \otimes V}_{r \text{ times}} \otimes \underbrace{V^* \otimes \ldots \otimes V^*}_{s \text{ times}}.$$

A contraction of an element $v_1 \otimes \ldots \otimes v_r \otimes \omega_1 \otimes \ldots \otimes \omega_s \in V \otimes \ldots \otimes V \otimes V^* \otimes \ldots \otimes V^*$ will stand for a map of the form

$$v_1 \otimes \ldots \otimes v_r \otimes \omega_1 \otimes \ldots \otimes \omega_s \mapsto \omega_\beta(v_\alpha) \cdot v_1 \otimes \ldots \widehat{\otimes\, v_\alpha} \otimes \ldots v_r \otimes \omega_1 \otimes \ldots \widehat{\otimes\, \omega_\beta} \otimes \ldots \omega_s \quad (1.19)$$

for $1 \le \alpha \le r$ and $1 \le \beta \le s$ fixed, where "$\frown$" means that the element is supposed to be left out. This is precisely the induced map for the evaluation map in the $(\alpha, \beta)$-th entry. If $\alpha$ and $\beta$ are not further specified, any statement that contains such a contraction is supposed to hold for all possibilities of $\alpha$ and $\beta$.

We will now generalize the definition of a tensor product of vector spaces to vector bundles. Fibrewise, the two definitions coincide.

**Definition 1.145.** Let $\pi_E : E \to M$ be a vector bundle of rank $k$ and $\pi_F : F \to M$ be a vector bundle of rank $\ell$ and, as in Definition 1.142, let $\psi_i^E$ and $\psi_i^F$, $i \in I$, be local trivializations of $E$ and $F$, respectively, and $\mathcal{A}$ a fitting atlas of $M$ with charts $(\varphi_i, U_i)$, $i \in I$. The **tensor product of vector bundles** of $E$ and $F$, $\pi_{E\otimes F} : E \otimes F \to M$, is the vector bundle given fibrewise by

$$(E \otimes F)_p = \pi_{E\otimes F}^{-1}(p) := E_p \otimes F_p,$$

so that $E \otimes F := \bigsqcup_{p\in M} E_p \otimes F_p$. As in the construction of the Whitney sum of vector bundles, it suffices by Proposition 1.74 to construct local trivializations $\phi_i : \pi^{-1}(U_i) \to U_i \times \mathbb{R}^k \otimes \mathbb{R}^\ell \cong U_i \times \mathbb{R}^{k\ell}$ covering $E \otimes F$ with smooth vector parts of their transition functions in order to show that $E \otimes F$ is in fact a vector bundle. Analogous to equation (1.18) we set

$$\begin{aligned} \phi_i^{-1} &:= (\psi_i^E \otimes \psi_i^F)^{-1} \circ (\Delta_M \times \mathrm{id}_{\mathbb{R}^{k\ell}}) : U_i \times \mathbb{R}^{k\ell} \cong U_i \times (\mathbb{R}^k \otimes \mathbb{R}^\ell) \to \bigsqcup_{p\in U_i} (E_p \otimes F_p), \\ (p, v \otimes w) &\mapsto (\psi_i^E)^{-1}(p, v) \otimes (\psi_i^F)^{-1}(p, w) \quad \forall p \in U_i,\ v \in \mathbb{R}^k,\ w \in \mathbb{R}^\ell, \end{aligned} \quad (1.20)$$

where $\Delta_M : p \mapsto (p, p) \in M \times M$ again denotes the diagonal embedding and $\phi_i^{-1}$ on non-pure tensors is defined by linear extension for any $p \in U_i$ fixed. For the transition functions of the vector part in the change of local trivializations of $E \otimes F \to M$ we obtain for all $i, j \in I$, such that $U_i \cap U_j \neq \emptyset$,

$$\phi_i \circ \phi_j^{-1}(p, v \otimes w) = (p, \tau_{ij}^E(p)v \otimes \tau_{ij}^F(p)w),$$

where $\tau_{ij}^E$ and $\tau_{ij}^F$ are the transition functions of the local trivializations of $E$ and $F$, respectively. We check [Exercise!] that the linear extension of

$$\mathbb{R}^k \otimes \mathbb{R}^\ell \ni v \otimes w \mapsto \tau_{ij}^E(p)v \otimes \tau_{ij}^F(p)w \in \mathbb{R}^k \otimes \mathbb{R}^\ell$$

actually is an invertible linear map and conclude with Proposition 1.74 that $E \otimes F \to M$ is indeed a vector bundle of rank $k \cdot \ell$.

**Exercise 1.146.** The endomorphism bundle of a vector bundle $E \to M$ is defined as

$$\mathrm{End}(E) := E \otimes E^* \to M.$$

Describe the transition functions of $\mathrm{End}(E) \to M$ induced by given transition functions on $E \to M$.

**Definition 1.147.** Let $M$ be a smooth manifold and let $(r,s) \in \mathbb{N}_0 \times \mathbb{N}_0$ so that $r+s>0$. The vector bundle

$$T^{r,s}M := \underbrace{TM \otimes \ldots \otimes TM}_{r \text{ times}} \otimes \underbrace{T^*M \otimes \ldots \otimes T^*M}_{s \text{ times}} \to M$$

is called the **bundle of** $(r,s)$**-tensors** of $M$. In this notation, $T^{1,0}M = TM$ and $T^{0,1}M = T^*M$. The (local) sections in the bundle of $(r,s)$-tensors are called **(local)** $(r,s)$**-tensor fields**, or simply **tensor fields** if $(r,s)$ is clear from the context, and are denoted by

$$\mathcal{T}^{r,s}(M) := \Gamma(T^{r,s}M).$$

In local coordinates $(x^1,\ldots,x^n)$ on $U \subset M$, tensor fields $A \in \mathcal{T}^{r,s}(M)$ are of the form

$$\begin{aligned} A &= \sum_{\substack{1\le i_1,\ldots,i_r\le n\\ 1\le j_1,\ldots,j_r\le n}} A^{i_1\ldots i_r}{}_{j_1\ldots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_s}, \qquad (1.21)\\ A^{i_1\ldots i_r}{}_{j_1\ldots j_s} &\in C^\infty(U) \quad \forall 1 \le i_1,\ldots,i_r,j_1,\ldots,j_s \le n. \end{aligned}$$

The above local form of tensor fields is commonly called **index notation of tensor fields**. This is justified by the fact that locally $A$ is uniquely determined by the local smooth functions $A^{i_1\ldots i_r}{}_{j_1\ldots j_s}$ on chart neighbourhoods of an atlas of $M$. Note that the summation in (1.21) "pairs up" coinciding upper and lower indices. In physics literature, the summation signs are usually omitted. This type of notation is called the **Einstein summation convention**. We will not be using that convention a.k.a. notation but instead leave out the ranges of the summations from here on whenever they are clear from the context. For example, a vector field $X \in \mathfrak{X}(M)$ on an $n$-dimensional smooth manifold $M$ will then locally be written as

$$X = \sum X^i \frac{\partial}{\partial x^i}.$$

If $A \in \mathcal{T}^{r,s}(M)$ with $r>0$ and $s>0$ we can **contract** $A$ in the $i,j$-th index, $1 \le i \le r$, $1 \le j \le s$, which is pointwise in local coordinates defined as in (1.19), and obtain a tensor field in $\mathcal{T}^{r-1,s-1}(M)$.[22]

**Remark 1.148.** Recall Proposition 1.138 and the construction of the tensor product via its universal property. One can show that $\mathcal{T}^{r,s}(M)$ is as $C^\infty(M)$ module isomorphic to the $C^\infty(M)$-multilinear maps $\mathrm{Hom}_{C^\infty(M)}(\Omega^1(M)^{\times r} \times \mathfrak{X}(M)^{\times s}, C^\infty(M))$. The proof needs some knowledge about tensor products of modules, but essentially works as the proof of Proposition 1.138.

**Exercise 1.149.**

(i) Work out the transformation laws for $(r,s)$-tensor fields when changing coordinates. As an example consider linear (global) change of coordinates in $\mathbb{R}^n$, i.e.

$$\begin{pmatrix} u^1 \\ \vdots \\ u^n \end{pmatrix} = B \begin{pmatrix} w^1 \\ \vdots \\ w^n \end{pmatrix}$$

for $(u^1,\ldots,u^n)$ the canonical coordinates and $B \in \mathrm{GL}(n)$. [Even though it is a bit tedious, do **not** skip this exercise!]

(ii) Show that contraction of tensor fields is well-defined, i.e. show that for a tensor field in local coordinates first contracting and then changing coordinates yields the same expression as first changing coordinates and then contracting. This in particular means that the contraction of an **endomorphism field** in $\mathcal{T}^{1,1}(M)$ is a well-defined smooth function on $M$.

[22] $\mathcal{T}^{0,0}(M) := C^\infty(M)$, see Remark 1.150.

(iii) Check that the contraction of an endomorphism field $A$ is pointwise in local coordinates precisely the trace of the endomorphism $A_p : T_pM \to T_pM$.

We know what it means to transport vector fields and 1-forms via diffeomorphisms from one smooth manifold to another. For 1-forms we have seen that we can pull them back with respect to any smooth map, not just diffeomorphisms. These constructions work for general tensor fields as well by applying them entry-wise. Observe the following:

**Remark 1.150.** For any $a \in \mathcal{T}^{r,s}(M)$, $b \in \mathcal{T}^{R,0}(M)$, $c \in \mathcal{T}^{0,S}(M)$,

$$b \otimes a \in \mathcal{T}^{r+R,s}(M), \quad a \otimes c \in \mathcal{T}^{r,s+S}(M),$$

where the tensor product is understood over $C^\infty(M)$.[23] For the above reason we **identify $C^\infty(M)$ with $\mathcal{T}^{0,0}(M)$** so that a $(0,0)$-tensor field is simply a smooth function. For $f \in C^\infty(M)$, we set $f \otimes a := fa$ for all $a \in \mathcal{T}^{r,s}(M)$. reference!

**Definition 1.151.** Let $M$, $N$ be smooth manifolds and let $F : M \to N$ be a diffeomorphism. The **pushforward and pullback of tensor fields** under $F$ are the unique $\mathbb{R}$-linear maps

$$\begin{aligned} F_* &: \mathcal{T}^{r,s}(M) \to \mathcal{T}^{r,s}(N), \\ F^* &: \mathcal{T}^{r,s}(N) \to \mathcal{T}^{r,s}(M), \end{aligned}$$

such that

(i) $F_* : \mathcal{T}^{1,0}(M) \to \mathcal{T}^{1,0}(N)$ is the pushforward of vector fields, $F^* : \mathcal{T}^{1,0}(N) \to \mathcal{T}^{1,0}(M)$ is the pullback of vector fields,

(ii) $F_* : \mathcal{T}^{0,1}(M) \to \mathcal{T}^{0,1}(N)$ is the pushforward of 1-forms, $F^* : \mathcal{T}^{0,1}(N) \to \mathcal{T}^{0,1}(M)$ is the pullback of 1-forms,

(iii) $F_*(b \otimes a) = (F_*b) \otimes (F_*a)$ and $F^*(b \otimes a) = (F^*b) \otimes (F^*a)$ for all $a \in \mathcal{T}^{r,s}(M)$, $b \in \mathcal{T}^{R,0}(M)$,

(iv) $F_*(a \otimes c) = (F_*a) \otimes (F_*c)$ and $F^*(a \otimes c) = (F^*a) \otimes (F^*c)$ for all $a \in \mathcal{T}^{r,s}(M)$, $c \in \mathcal{T}^{0,S}(M)$.

For $f \in C^\infty(M)$, $g \in C^\infty(N)$, we set

$$F_*(f) := f \circ F^{-1}, \quad F^*g := g \circ F$$

so that $F_*(fa) = F_*(f)F_*(a)$ and $F^*(gb) = F^*(g)F^*(b)$ for all $f \in C^\infty(M)$, $g \in C^\infty(N)$, $a \in \mathcal{T}^{r,s}(M)$, $b \in \mathcal{T}^{r,s}(N)$.

The above definition might look worse than it actually is. If we are given some specific tensor field, say, an endomorphism field $A \in \mathcal{T}^{1,1}(M)$ which is in local coordinates on $U \subset M$ of the form

$$A = \sum A^i{}_j \frac{\partial}{\partial x^i} \otimes dx^j$$

and a diffeomorphism $F : M \to N$, all we need to do to calculate the local form of $F_*A$ is to choose fitting local coordinates on (or on a subset of) $F(U) \subset N$, and after possibly shrinking $U$ calculate $F_*\left(\frac{\partial}{\partial x^i}\right)$, $F_*(dx^j)$, for all $1 \leq i, j \leq n$. Then we can use the $\mathbb{R}$-linearity of the tensor product to get a local form of $F_*(A)$.

**Remark 1.152.** The pullback of $(0,s)$-tensors on $N$ is well-defined even if $F : M \to N$ is not a diffeomorphism.

[23] This means that $(fb) \otimes a = b \otimes (fa)$ for all $f \in C^\infty(M)$, the construction of the tensor product of modules is analogous to the construction of tensor products of vector spaces. For a reference see e.g. [].

**Lemma 1.153.** Contraction of tensor fields commute with the pushforward and with the pullback defined above.

*Proof.* By the $\mathbb{R}$-linearity and Definition 1.151, (i) and (ii), it suffices to show that

$$F_*(\alpha(X)) = F_*(\alpha)(F_*(X))$$

and

$$F^*(\beta(Y)) = F^*(\beta)(F^*(Y))$$

for all diffeomorphisms $F : M \to N$ and all $\alpha \in \Omega^1(M)$, $\beta \in \Omega^1(N)$, $X \in \mathfrak{X}(M)$, $Y \in \mathfrak{X}(N)$. This follows directly from Definitions 1.120 and 1.140. $\square$

**Exercise 1.154.**

(i) Show that the pushforward w.r.t. a diffeomorphism $F : M \to N$ is inverse to the pullback w.r.t. the inverse of the diffeomorphism $F^{-1} : N \to M$, independently of the type of tensor fields.

(ii) Determine all vector fields on $S^1$ that are invariant under the pushforward of all rotations in the ambient space $\mathbb{R}^2$ restricted to $S^1$. $X \in \mathfrak{X}(S^1)$ being invariant under $F : S^1 \to S^1$ means that $X_p = (F_*X)_p$ for all $p \in S^1$.

Recall the definition of the Lie derivative of vector fields. Geometrically, the Lie derivative is one way of measuring the infinitesimal change of a vector field along the local flow of another vector field. We can, analogously, define the Lie derivative of general tensor fields with respect to a given vector field.

**Definition 1.155.** Let $M$ be a smooth manifold, $X \in \mathfrak{X}(M)$ a vector field, and $A \in \mathcal{T}^{r,s}(M)$ a tensor field. Then the **Lie derivative of $A$ in direction of $X$**, $\mathcal{L}_X A \in \mathcal{T}^{r,s}(M)$, is defined as

$$(\mathcal{L}_X A)_p := \left.\frac{\partial}{\partial t}\right|_{t=0} (\varphi_t^* A)_p \quad \forall p \in M,$$

where $\varphi : I \times U \to M$ is any local flow of $X$ near $p \in M$.

Note that $(\varphi_t^* A)_p$ is, for $p$ fixed, for all $t \in I$ contained in the same vector space $T_pM \otimes \ldots \otimes T_pM \otimes T_p^*M \otimes \ldots \otimes T_p^*M$, thus $\mathcal{L}_X A$ is well-defined.

**Remark 1.156.** The above definition is consistent with the identification $\mathcal{T}^{0,0}(M) = C^\infty(M)$.

**Proposition 1.157.** The Lie derivative of tensor fields is a **tensor derivation**, i.e. it is compatible with all possible contractions and fulfils the Leibniz rule

$$\mathcal{L}_X(A \otimes B) = \mathcal{L}_X A \otimes B + A \otimes \mathcal{L}_X B$$

for all vector fields $X$ and all tensor fields $A$, $B$, such that $A \otimes B$ is defined.

*Proof.* To show compatibility with arbitrary contractions it suffices to show that this property holds true for endomorphism fields $A \in \mathcal{T}^{1,1}(M)$ for which there is precisely one possible contraction. All other possible cases will follow by induction and the Leibniz rule. We will prove first that the Leibniz rule is fulfilled. Let $p \in M$ be fixed and $A$, $B$, two tensor fields, such that $A \otimes B$ is defined. First assume that $(A \otimes B)_p = A_p \otimes B_p \neq 0$. Let $X \in \mathfrak{X}(M)$ be arbitrary but fixed and denote by $\varphi : I \times U \to M$ its local flow near $p$ with $U \subset M$ contained in a chart neighbourhood for some local coordinates. We can find an interval $(-\varepsilon, \varepsilon) \subset I$ for $\varepsilon > 0$

small enough, such that in the local coordinates on $U$ and the induced coordinates on the fitting $(r,s)$-tensor bundles $\psi$ and $\phi$, the pullbacks of $A$ and $B$ w.r.t. the local flow of $X$ are of the form

$$\psi((\varphi_t^* A)_p) = (p, a(t)v), \quad \phi((\varphi_t^* B)_p) = (p, b(t)w) \quad \forall t \in (-\varepsilon, \varepsilon).$$

In the above equation, $0 \neq v \in \mathbb{R}^{N_1}$ and $0 \neq w \in \mathbb{R}^{N_2}$ are fixed nonzero vectors and $N_1$, $N_2$, depend on the type of tensor field that $A$ and $B$ are. The expressions $a(t)$ and $b(t)$ stand for smooth and uniquely defined maps

$$a : (-\varepsilon, \varepsilon) \to \mathrm{GL}(N_1), \quad b : (-\varepsilon, \varepsilon) \to \mathrm{GL}(N_2),$$

with $a(0) = \mathrm{id}_{\mathbb{R}^{N_1}}$ and $b(0) = \mathrm{id}_{\mathbb{R}^{N_2}}$. Thus, in order to prove the Leibniz property, it suffices to show that for any finite dimensional real vector spaces $V$, $\dim(V) = N_1$, and $W$, $\dim(W) = N_2$, and any smooth maps $a$ and $b$ as above,

$$\left.\frac{\partial}{\partial t}\right|_{t=0} ((a(t)v) \otimes (b(t)w)) = (a'(0)v) \otimes w + v \otimes (b'(0)) \tag{1.22}$$

for all $v \in V$, $w \in W$. This follows from the defining universal property of the tensor product of vector spaces as follows. Let $L : V \times W \to \mathbb{R}$ be any bilinear map and $\widetilde{L} : V \otimes W \to \mathbb{R}$ the corresponding linear map, so that $L(a(t)v, b(t)w) = \widetilde{L}((a(t)v) \otimes (b(t)w))$ for all $v \in V$, $w \in W$, $t \in (-\varepsilon, \epsilon)$. By taking the $t$-derivative at $t = 0$ on both sides we obtain

$$\left.\frac{\partial}{\partial t}\right|_{t=0} \widetilde{L}((a(t)v) \otimes (b(t)w)) = \widetilde{L}((a'(0)v) \otimes w + v \otimes (b'(0)w)).$$

Since $L$ and thus $\widetilde{L}$ were arbitrary, the above statement hold in particular for all component functions. This shows (1.22) and, hence, proves the Leibniz property. To obtain the compatibility with contractions it is enough to consider $W = V^*$ and $L = \mathrm{ev}$ the evaluation map. Then $\widetilde{L}$ is precisely the contraction.

Next assume that $(A \otimes B)_p = 0$ and that there exists a convergent sequence $\{p_n\}_{n \in \mathbb{N}}$ with $p_n \to p$ as $n \to \infty$, such that $(A \otimes B)_{p_n} \neq 0$ for all $n \in N$. Then the statement of this proposition follows with a continuity argument similar to the one used in Proposition 1.128.

Lastly assume that $(A \otimes B)_p = 0$ and $A \otimes B$ vanishes identically on an open neighbourhood $U \subset M$ of $p$. Then $A$ or $B$ must already vanish identically on $U$. Without loss of generality we can assume that $U$ is a chart neighbourhood and choose a fitting bump function $b$ with $\mathrm{supp}(b) \subset U$ compactly embedded, so that the locally defined prefactors in the local forms of $A$ and $B$, multiplied with said bump function, are globally defined smooth functions. Now we use that $bA$ and $bB$ vanish identically and in some smaller open neighbourhood $V \subset U$ coincide with $A$ and $B$, respectively. Thus on $V$ if $bA \equiv 0$ we obtain $\mathcal{L}_X(A) = \mathcal{L}_X(bA) = \mathcal{L}_X(0) = 0$ and a similar identity for $B$ and $A \otimes B$. This finishes the proof. □

**Corollary 1.158.** $(\mathcal{L}_X \alpha)(Y) = X(\alpha(Y)) - \alpha([X,Y])$ for all $X, Y \in \mathfrak{X}(M)$ and all $\alpha \in \Omega^1(M)$.

**Exercise 1.159.** Show that $\mathcal{L}_X(df) = d(\mathcal{L}_X f)$ for all $f \in C^\infty(M)$, $X \in \mathfrak{X}(M)$.

# 2 Pseudo-Riemannian metrics, connections, and geodesics

## 2.1 Pseudo-Riemannian metrics and isometries

We start this section with quickly recalling some facts from linear algebra on finite-dimensional vector spaces equipped with a scalar product, that is a symmetric bilinear map with values in $\mathbb{R}$.

**Remark 2.1.** Let $V$ be a finite-dimensional real vector space. A **pseudo-Euclidean scalar product on** $V$ is a nondegenerate symmetric bilinear map

$$\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}.$$

Nondegenerate means that there exists no proper linear subspace $W \subset V$, such that $\langle\cdot,\cdot\rangle|_{W\times V} \equiv 0$. The **index of** $\langle\cdot,\cdot\rangle$ is defined as the number of its negative eigenvalues when written as a symmetric $\dim(V)\times\dim(V)$-matrix. The index does not depend on the choice of basis of $V$, this is **Sylvester's law of inertia**. If the index of the pseudo-Euclidean scalar product is zero, it is simply called **Euclidean scalar product**. A vector space equipped with a (pseudo)-Euclidean scalar product is called **(pseudo)-Euclidean vector space**. Prominent examples are $\mathbb{R}^n$ together with the Euclidean scalar product that is given by the dot-product, i.e.

$$\langle v,w\rangle = \sum_{i=1}^{n} v^i w^i,$$

and $\mathbb{R}^{n+1}$ together with the **Minkowski scalar product**

$$\langle v,w\rangle = -v^{n+1}w^{n+1} + \sum_{i=1}^{n} v^i w^i.$$

Note that in certain fields of theoretical physics one uses an overall sign in front of the Minkowski scalar product. The **length** of a vector $v \in V$ with respect to a pseudo-Euclidean scalar product $\langle\cdot,\cdot\rangle$ is defined as

$$\|v\| := \sqrt{|\langle v,v\rangle|}.$$

If $\langle\cdot,\cdot\rangle$ is a pseudo-Euclidean scalar product with positive index smaller than the dimension of the vector space, one says that a vector $v$ is **spacelike** if $\langle v,v\rangle > 0$, **timelike** if $\langle v,v\rangle < 0$, and **null** (or **lightlike**) if $\langle v,v\rangle = 0$. If $\langle\cdot,\cdot\rangle$ is Euclidean each nonzero vector has positive length. Let $A \in \mathrm{GL}(V)$ describe a change of basis in a pseudo-Euclidean vector space $(V,\langle\cdot,\cdot\rangle)$ in the sense that $A$ maps the new basis to the given one and assume that the representation matrix of $\langle\cdot,\cdot\rangle$ is given by the symmetric matrix $B$. Then in the new basis, the representation matrix of $\langle\cdot,\cdot\rangle$ is given by $A^T BA$.[24] Two pseudo-Euclidean vector spaces $(V,\langle\cdot,\cdot\rangle_V)$ and $(W,\langle\cdot,\cdot\rangle_W)$ are called **isometric** if there exist a linear isomorphism $A : V \to W$, such that $\langle\cdot,\cdot\rangle_V = \langle A\cdot, A\cdot\rangle_W$. $A$ is then called **(linear) isometry**. Two finite-dimensional pseudo-Euclidean vector spaces are isometric if and only if their dimension and index of the scalar product coincide. Note that any pseudo-Euclidean scalar product might be interpreted as an element in $\mathrm{Sym}^2(V^*)$ which denotes the set of symmetric two-tensors in $V^* \otimes V^*$.

**Exercise 2.2.** Show that for any pseudo-Euclidean vector space $(V,\langle\cdot,\cdot\rangle)$, $\langle\cdot,\cdot\rangle$ is completely determined by its value on the diagonal in $V \times V$, that is on vectors of the form $(v,v) \in V \times V$.

One can use Sylvester's law of inertia to prove the following fact from linear algebra.

**Proposition 2.3.** Let $(V,\langle\cdot,\cdot\rangle)$ be a pseudo-Euclidean vector space of dimension $n$ and let $\nu$ denote the index of $\langle\cdot,\cdot\rangle$. Then $(V,\langle\cdot,\cdot\rangle)$ is isometric to $(\mathbb{R}^n,\langle\cdot,\cdot\rangle_\nu)$, where

$$\langle v,v\rangle_\nu := \sum_{i=1}^{n-\nu} \left(v^i\right)^2 - \sum_{i=n-\nu+1}^{n} \left(v^i\right)^2.$$

Recall the definition of orthogonality from linear algebra:

[24] Compare this formula to the pullback of 1-forms.

**Definition 2.4.** Let $(V, \langle\cdot,\cdot\rangle)$ be a pseudo-Euclidean vector space and $W \subset V$ a pseudo-Euclidean linear subspace, meaning that $\langle\cdot,\cdot\rangle|_{W\times W}$ is a pseudo-Euclidean scalar product on $W$. Then the **orthogonal complement** $W^\perp \subset V$ of $W$ in $V$ with respect to $\langle\cdot,\cdot\rangle$ is given by

$$W^\perp := \{v \in V \mid \langle v, w\rangle = 0 \quad \forall w \in W\}.$$

$W^\perp$ is a linear subspace of $V$ of dimension $\dim(W^\perp) = \dim(V) - \dim(W)$ and

$$W \oplus W^\perp = V.$$

If $W \subset V$ is any linear subspace of $V$, we will also use the notation $W^\perp$ for its orthogonal complement. Two arbitrary vectors $v, w \in V$ are called **orthogonal** if $\langle v, w\rangle = 0$, and two linear subspaces $V_1, V_2$ of $V$ are called **orthogonal** to each other if $\langle v_1, v_2\rangle = 0$ for all $v_1 \in V_1$, $v_2 \in V_2$. A basis $\{v_1, \dots, v_n\}$ of $V$ is called **orthogonal basis** with respect to $\langle\cdot,\cdot\rangle$ if $\langle v_i, v_j\rangle = 0$ for all $1 \le i, j \le n$, $i \ne j$. An orthogonal basis is called **orthonormal basis** if additionally $\|v_i\| = 1$ for all $1 \le i \le n$.

The following exercise recaptures some additional facts from linear algebra.

**Exercise 2.5.**

(i) Prove that every pseudo-Euclidean vector space admits an orthonormal basis.

(ii) Show that the index $\nu$ of a pseudo-Euclidean scalar product coincides with the number of elements in $\{i \mid \langle v_i, v_i\rangle = -1\}$ for any given orthonormal basis $\{v_1, \dots, v_n\}$ of $(V, \langle\cdot,\cdot\rangle)$.

(iii) For any given pseudo-Euclidean vector space $(V, \langle\cdot,\cdot\rangle)$ and $W \subset V$ any linear subspace, prove:

  (a) $(W^\perp)^\perp = W$,
  (b) $W$ is a pseudo-Euclidean linear subspace $\Leftrightarrow W \cap W^\perp = \{0\} \Leftrightarrow V = W \oplus W^\perp$.

(iv) Linear isometries map orthonormal (orthogonal) bases to orthonormal (orthogonal) bases.

We want to translate the concept of pseudo-Euclidean vector spaces to smooth manifolds. More precisely we want to understand what it means to specify for each point $p$ in a given manifold $M$ a pseudo-Euclidean scalar product on $T_pM$, such that this assignment varies smoothly on $M$.

**Definition 2.6.** Let $M$ be a smooth manifold. A **pseudo-Riemannian metric with index** $0 \le \nu \le \dim(M)$ on $M$ is a symmetric $(0,2)$-tensor field $g \in \mathcal{T}^{0,2}(M)$, $g : p \mapsto g_p \in \mathrm{Sym}^2(T_p^*M)$, such that for all $p \in M$ $g_p$ is a pseudo-Euclidean scalar product of index $\nu$ on $T_pM$. This in particular means that

$$g(X, Y) = g(Y, X) \in C^\infty(M)$$

for all vector fields $X, Y \in \mathfrak{X}(M)$. If $\nu = 0$, $g$ is called **Riemannian metric**. In local coordinates $(x^1, \dots, x^n)$ on $U \subset M$, $g$ is of the form

$$g = \sum_{i,j=1}^n g_{ij}\, dx^i \otimes dx^j, \tag{2.1}$$

where

$$g_{ij} := g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) \in C^\infty(U) \tag{2.2}$$

for all $1 \le i, j \le n$. The symmetry condition for $g$ is equivalent to requiring that in all local coordinates $g_{ij} = g_{ji}$. This means that $(g_{ij})$, viewed as a $n \times n$-matrix valued smooth map on the coordinate domain, is at each point a symmetric matrix. If we write in local coordinates $X = \sum_{i=1}^{n} X^i \frac{\partial}{\partial x^i}$, $Y = \sum_{i=1}^{n} Y^i \frac{\partial}{\partial x^i}$, we obtain the local formula for $g(X, Y)$

$$g(X, Y) = \sum_{i,j=1}^{n} g_{ij} X^i Y^j,$$

which heuristically corresponds to plugging in $X$ in the left half and $Y$ in the right half of the tensor terms in $g$.

**Definition 2.7.** A smooth manifold $M$ equipped with a (pseudo)-Riemannian metric $g$ is called **(pseudo)-Riemannian manifold**.

**Remark 2.8.** Of particular importance in mathematics and physics are pseudo-Riemannian manifolds of index 0 and 1, that is **Riemannian manifolds** and **Lorentz manifolds**, respectively. The latter are the manifolds that are studied in general relativity, for an introduction see [O, Ch. 12]. Why would one want to study Riemannian manifolds in their full generality, aside from an explanation how the standard Riemannian metric on $\mathbb{R}^n$ induced by the Euclidean scalar product at each point transforms? The answer is manifold (this time, the latter is an adjective). First and foremost because it allows our studied geometrical objects to have **curvature**. We will study this topic extensively in Sections 3 and 4. Furthermore, Riemannian metrics give us a way to study volumes of submanifolds. This is not completely trivial, as it involves the construction of a so-called **volume form** from a given Riemannian metric, respectively its restriction to submanifolds, cf. Section **??**. For starters, it allows us to define the arc-length of a curve. fix this

**Definition 2.9.** Let $(M, g)$ be a Riemannian manifold and $\gamma : I \to M$ a smooth curve. Then the **arc-length**, or simply **length**, of $\gamma$ is defined as

$$L(\gamma) = \int_I \sqrt{g(\gamma', \gamma')} dt.$$

Note that $L(\gamma) = \infty$ is allowed.

Next we will study some explicit examples of pseudo-Riemannian manifolds.

**Example 2.10.**

(i) Any pseudo-Riemannian vector space $(V, \langle \cdot, \cdot \rangle)$ is, viewed as a smooth manifold with $g_p := \langle \cdot, \cdot \rangle$ for all $p \in V$[25], a pseudo-Riemannian manifold. If $V = \mathbb{R}^n$ equipped with its canonical coordinates and Euclidean scalar product at each tangent space, the induced Riemannian metric in canonical coordinates $(u^1, \dots, u^n)$ is given by

$$g = \sum_{i=1}^{n} du^i \otimes du^i.$$

(ii) Any smooth submanifold $M \subset \mathbb{R}^n$ equipped with

$$g \in \mathcal{T}^{0,2}(M), \quad g_p = \langle \cdot, \cdot \rangle|_{T_pM \times T_pM}$$

for all $p \in M$, that is the restriction of the Euclidean scalar product on $T_p\mathbb{R}^n \cong \mathbb{R}^n$ to the tangent space $T_pM$ of $M$ at $p$, is a Riemannian manifold.

(iii) More generally, any smooth submanifold of a Riemannian manifold is by restriction of the metric to the tangent bundle of the smooth submanifold a Riemannian manifold.

[25] Recall that $T_pV \cong V$ for all $p \in V$.

(iv) If $(M, g_M)$ and $(N, g_N)$ are pseudo-Riemannian manifolds with index $\nu_M$, $\nu_N$, respectively, the product $M \times N$ is a pseudo-Riemannian manifold of index $\nu_M + \nu_N$. The corresponding **product metric** on $M \times N$ is given by

$$g_{M\times N} := g_M + g_N, \quad g_{M\times N}((v,w),(v,w)) = g_M(v,v) + g_N(w,w),$$

for all $(v,w) \in TM \oplus TN \cong T(M \times N)$.

Example 2.10 (iii) motivates the following definition.

**Definition 2.11.** Let $(N, \overline{g})$ be a pseudo-Riemannian manifold and $M \subset N$ a smooth submanifold. $M$ is called **pseudo-Riemannian submanifold** of $N$ if

$$g := \overline{g}|_{TM\times TM}$$

is a pseudo-Riemannian metric on $M$. In the above equation, the restriction to $TM \times TM$ means that we restrict the basepoint of $\overline{g}$ to $M \subset N$ and the vectors we are allowed to plug in to vectors in $TM \subset TN$.

**Exercise 2.12.**

(i) Show that any smooth manifold can be equipped with a Riemannian metric. [Hint: Use a countable smooth partition of unity subordinate to a countable atlas on $M$.]

(ii) Show that not every manifold can be equipped with a pseudo-Riemannian, not Riemannian, metric. This is to be understood to also exclude the index $\nu = \dim(M)$. ["Hint": This exercise is very difficult.]

(iii) Show that every $n \geq 2$-dimensional pseudo-Riemannian manifold $N$ with metric $\overline{g}$ of index $1 \leq \nu \leq n-1$ has smooth submanifolds that are not pseudo-Riemannian submanifolds.

The pseudo-Riemannian manifold-analogue to isometries of pseudo-Euclidean vector spaces is as follows.

**Definition 2.13.** Let $(M, g)$ and $(N, h)$ be pseudo-Riemannian manifolds and $F : M \to N$ a diffeomorphism. Then $F$ is called an **isometry** if $F^*h = g$ or, equivalently, $F_*g = h$. One checks that the first condition is equivalent to

$$g_p(X_p, Y_p) = h_{F(p)}(dF_p(X_p), dF_p(Y_p))$$

for all $X, Y \in \mathfrak{X}(M)$ and all $p \in M$, meaning that fibrewise $dF_p$ is a linear isometry. The two pseudo-Riemannian manifolds $(M, g)$ and $(N, h)$ are then called **isometric**. The isometries $F : M \to M$ with respect to $g$ form a group, **the isometry group of** $(M, g)$, which is denoted by $\mathrm{Isom}(M, g)$.

**Example 2.14.**

(i) Every orthogonal transformation $A \in \mathrm{O}(n+1)$ is, by definition, an isometry of $\mathbb{R}^{n+1}$ equipped with the standard Riemannian metric given pointwise by the Euclidean scalar product $\langle\cdot,\cdot\rangle$.

(ii) Since each $A \in \mathrm{O}(n+1)$ restricts to a diffeomorphism of the unit sphere $S^n \subset \mathbb{R}^{n+1}$, it is an isometry of $(S^n, \langle\cdot,\cdot\rangle|_{TS^n\times TS^n})$. The Riemannian metric $\langle\cdot,\cdot\rangle|_{TS^n\times TS^n}$ is sometimes called the **round metric**.

(iii) Consider the upper half plane $H := \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$ equipped with the Riemannian **Poincaré metric**

$$g = \frac{1}{y^2}(dx^2 + dy^2).$$

$(H, g)$ is called the **Poincaré half-plane model**. When viewed as a subset of $\mathbb{C}$ via $H \ni (x, y) \mapsto x + iy \in \mathbb{C}$, one obtains an **isometric action**[26] of

$$\mathrm{PSL}(2, \mathbb{R}) = \mathrm{SL}(2, \mathbb{R})/_\sim, \quad A \sim B :\Leftrightarrow A = \pm B,$$

on $H \subset \mathbb{C}$ defined by

$$\mu : \mathrm{PSL}(2, \mathbb{R}) \times H \to H, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z := \frac{az + b}{cz + d}.$$

**Exercise 2.15.** Prove the statement in Example 2.14 (iii) and show that the group action $\mu : \mathrm{PSL}(2, \mathbb{R}) \times H \to H$ is transitive.

A change of coordinates on $M$ induces a fibrewise change of bases in $T^{r,s}M$ for all $r + s > 0$. We obtain the following result for local forms of pseudo-Riemannian metrics under a change of coordinates.

**Lemma 2.16.** Let $(M, g)$ be a pseudo-Riemannian manifold and let $\varphi = (x^1, \ldots, x^n)$ and $\psi = (y^1, \ldots, y^n)$ be local coordinate systems on $U \subset M$ and $V \subset M$, respectively, such that $U \cap V \neq \emptyset$. Denote on $U \cap V$

$$g = \sum_{i,j} g_{ij}\, dx^i \otimes dx^j = \sum_{i,j} \widetilde{g}_{ij}\, dy^i \otimes dy^j. \tag{2.3}$$

The local coordinate systems $\varphi$ and $\psi$ are related by $(x^1, \ldots, x^n) = F(y^1, \ldots, y^n)$ on $U \cap V$, where $F : \psi(U \cap V) \to \varphi(U \cap V)$. Then the matrix valued maps $(g_{ij})$ and $(\widetilde{g}_{ij})$ in (2.3) are related by

$$(\widetilde{g}_{ij})|_p = dF^T_{\psi(p)} \cdot (g_{ij})|_{\varphi^{-1}(F(\psi(p)))} \cdot dF_{\psi(p)}.$$

*Proof.* Follows by considering coordinate representations of $(g_{ij})$ and $(\widetilde{g}_{ij})$, writing down the pullback of $(g_{ij})$ with respect to $F$, and comparing the prefactors. □

At first glance the above lemma might look more complicated than it actually is. Pointwise, the statement is precisely the transformation law for pseudo-Euclidean scalar products under a change of basis.

Recall the following construction from linear algebra.

**Definition 2.17.** Let $V$ be a real finite-dimensional vector space and $A \in \mathrm{End}(V) \cong V \otimes V^*$, so that for a basis $\{v_1, \ldots, v_n\}$ of $V$

$$A = \sum_{i,j=1}^{n} a^i{}_j\, v_i \otimes v^*_j.$$

The **trace** of $A$ is defined as

$$\mathrm{tr}(A) := \sum_{i=1}^{n} a^i{}_i.$$

[26]This means: A group action $\mu : \mathrm{PSL}(2, \mathbb{R}) \times H \to H$, where for every group element $A \in \mathrm{PSL}(2, \mathbb{R})$ fixed, the induced map $\mu(A, \cdot) : H \to H$ is an isometry.

**Exercise 2.18.** Check that the definition of the trace in Definition 2.17 is well-defined, meaning that it gives the same value for all choices of a basis of $V$.

**Example 2.19.**

(i) $\mathrm{tr}(\mathrm{id}_V) = \dim(V)$,

(ii) $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$ for all $A, B \in \mathrm{End}(V)$,

(iii) $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ for all $A, B \in \mathrm{End}(V)$,

(iv) $\mathrm{tr}(v \otimes \omega) = \omega(v)$ for all $v \in V$, $\omega \in V^*$.

One can furthermore show the following:

**Lemma 2.20.** Let $(V, \langle\cdot,\cdot\rangle)$ be a finite-dimensional pseudo-Euclidean vector space and $A \in \mathrm{End}(V)$. Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis of $V$ with respect to $\langle\cdot,\cdot\rangle$. Then

$$\mathrm{tr}(A) = \sum_{i=1}^{n} \varepsilon_i \langle e_i, Ae_i\rangle,$$

where $\varepsilon_i := \langle e_i, e_i\rangle \in \{-1, 1\}$ for all $1 \leq i \leq n$.

*Proof.* Exercise. □

One can for pseudo-Euclidean vector spaces further define natural (possibly indefinite) scalar products on $V^{\otimes r} \otimes (V^*)^{\otimes s}$ for all $r + s > 0$.

**Definition 2.21.** Let $(V, \langle\cdot,\cdot\rangle)$ be a pseudo-Euclidean vector space and $\{e_1, \ldots, e_n\}$ a basis of $V$. Let further $A \in V^{\otimes r} \otimes (V^*)^{\otimes s}$ and write

$$\begin{aligned}
\langle\cdot,\cdot\rangle &= \sum_{i,j=1}^{n} g_{ij}\, e_i^* \otimes e_j^*, \\
A &= \sum_{\substack{1\leq i_1,\ldots,i_r\leq n \\ 1\leq j_1,\ldots,j_r\leq n}} A^{i_1\ldots i_r}{}_{j_1\ldots j_s}\, e_{i_1} \otimes \ldots \otimes e_{i_r} \otimes e_{j_1}^* \otimes \ldots \otimes e_{j_s}^*.
\end{aligned}$$

Then

$$\langle A, A\rangle := \sum_{\substack{1\leq i_1,\ldots,i_r\leq n \\ 1\leq j_1,\ldots,j_r\leq n \\ 1\leq I_1,\ldots,I_r\leq n \\ 1\leq J_1,\ldots,J_r\leq n}} A^{i_1\ldots i_r}{}_{j_1\ldots j_s} \cdot A^{I_1\ldots I_r}{}_{J_1\ldots J_s} \cdot g_{i_1 I_1} \cdot \ldots \cdot g_{i_r I_r} \cdot g^{j_1 J_1} \cdot \ldots \cdot g^{j_s J_s} \tag{2.4}$$

defines a, possibly indefinite, symmetric bilinear form on $V^{\otimes r} \otimes (V^*)^{\otimes s}$. In the above formula the $g$-terms fulfil, when viewed as symmetric matrix-valued smooth maps,

$$(g^{ij}) := (g_{ij})^{-1}.$$

**Remark 2.22.** Formula (2.4) should make you ask one thing and realize another. Firstly you should ask why one would write down something like that. The formula (2.4) is, when generalized to smooth manifolds and tensor powers of the tangent bundle, used to define certain geometric invariants, e.g. the so-called **Kretschmann scalar**, see Remark 3.28. Secondly, the summation ranges in (2.4) should convince you that sometimes it might be a good idea to be a little bit imprecise to increase readability when the ranges are clear from the setting. In the following we will do just that.

As hinted in the above remark, Definitions 2.17 and 2.21 readily generalize to smooth manifolds and tensor bundles.

**Definition 2.23.** Let $(M, g)$ be a pseudo-Riemannian manifold, $A \in \mathcal{T}^{1,1}(M)$ an endomorphism field, $h \in \mathcal{T}^{0,2}(M)$ a symmetric $(0,2)$-tensor field, and $B \in \mathcal{T}^{r,s}(M)$ for $r + s > 0$ an arbitrary tensor field. Then the **trace of** $A$ is in local coordinates $(x^1, \ldots, x^n)$, so that $A = \sum A^i{}_j \frac{\partial}{\partial x^i} \otimes dx^j$, given by

$$\mathrm{tr}(A) := \sum_i A^i{}_i.$$

The above term is **invariant under coordinate change**, which follows from fibrewise invariance of the choice of basis in $T_pM$ and the fact that the coordinate cotangent vector at each point are precisely the dual to the coordinate tangent vectors at that point. This means that $\mathrm{tr}(A) \in C^\infty(M)$. The trace of $h$, given locally by $h = \sum_{i,j} h_{ij} dx^i \otimes dx^j$, with respect to $g$ is defined in local coordinates as

$$\mathrm{tr}_g(h) := \sum_{i,j} h_{ij}\, g^{ij}.$$

As for the endomorphism field, $\mathrm{tr}_g(h)$ is **invariant under coordinate change** which implies $\mathrm{tr}_g(h) \in C^\infty(M)$, but **not** invariant of the pseudo-Riemannian metric $g$. Furthermore, we define the induced pairing of $B$ with itself with respect to $g$ in the given local coordinates as

$$g(B, B) := \sum B^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} \cdot B^{I_1 \ldots I_r}{}_{J_1 \ldots J_s} \cdot g_{i_1 I_1} \cdot \ldots \cdot g_{i_r I_r} \cdot g^{j_1 J_1} \cdot \ldots \cdot g^{j_s J_s},$$

where

$$B = \sum B^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_s}$$

and $(g^{ij}) = (g_{ij})^{-1}$ at each point when viewed as symmetric matrix valued maps. As for the trace, the value of $g(B, B)$ does not depend on the choice of local coordinates which implies $g(B, B) \in C^\infty(M)$. Note that one can similarly define a symmetric pairing $g$ in the bundle $T^{r,s}M \to M$, which is an example of a possibly indefinite **bundle metric**.

**Example 2.24.** For any pseudo-Riemannian manifold $(M, g)$ of dimension $n$ we have

$$\mathrm{tr}_g(g) = g(g, g) \equiv n.$$

Pseudo-Riemannian metrics allow us to take any $(r, s)$-tensor field and change it into an $(r', s')$-tensor field if $r + s = r' + s'$. This process is reversible, and on the level of bundles known as **musical isomorphisms**.

**Proposition 2.25.** Let $(M, g)$ be a pseudo-Riemannian manifold. Then $T^{r,s}M \to M$ and $T^{r',s'}M \to M$ are isomorphic as vector bundles if $r + s = r' + s'$.

*Proof.* We first proof that $T^*M \to M$ and $TM \to M$ are isomorphic. Let

$$F : TM \to T^*M, \quad v \mapsto g(v, \cdot).$$

It is clear that $g(v, \cdot) \in T_p^*M$ for all $v \in T_pM$. Furthermore, the map $F$ is smooth, fibre-preserving, and at each point a linear isomorphism. Its inverse is given by

$$F^{-1} : T^*M \to TM, \quad \omega \mapsto g^{-1}(\omega, \cdot),$$

where we use the pointwise identification $(T_p^*M)^* = T_pM$ and $g^{-1}$ is given in local coordinates by

$$g^{-1} = \sum g^{ij} \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j}.$$

In order to show that $T^{r,s}M \to M$ and $T^{r',s'}M \to M$ are isomorphic for arbitrary $r, s, r', s'$ with $r + s = r' + s'$ one inductively uses entrywise isomorphisms. Note that there are usually choices involved which vector or covector parts to change into covector and vector parts, respectively. These choices correspond to which index is lowered or raised. Exceptions are e.g. going from $T^{1,1}M$ to $T^{0,2}M$ where there is only one possible choice of lowering an index. Care is also required when composing such isomorphisms as it might lead to "swapping" in the tensor powers of the vectors and covectors, where we recall that e.g. in $TM \otimes TM \to M$ swapping the fibres is an isomorphism of vector bundles. $\square$

The above Proposition 2.25 describes what is commonly, in particular in physics, known as lowering/raising indices. This is due to when one composes these isomorphisms with tensor fields, locally the prefactors' index locations change from up to down or the other way round. Check for example what happens to the used pseudo-Riemannian metric if one raises an index!

**Remark 2.26.** The isomorphisms of vector bundles $T^{r,s}M \to T^{r+1,s-1}M$ are denoted by $\sharp$ (read: "sharp"), and the isomorphisms $T^{r,s}M \to T^{r-1,s+1}M$ are denoted by $\flat$ (read: "flat"). Hence the name "musical isomorphisms". One needs to make sure to be aware of **which index** is moved up or down if there is a choice! Note that, using the musical isomorphisms, we could have defined the trace of endomorphism fields $A \in \mathcal{T}^{1,1}(M)$ on a pseudo-Riemannian manifold $(M, g)$ as

$$\operatorname{tr}_g(A) = \sum_{ij} (\sharp A)^{ij} g_{ij}.$$

It is crucial to observe that, as for our definition of $\operatorname{tr}(A)$ in Definition 2.23, the above term $\operatorname{tr}_g(A)$ is invariant[27] of the pseudo-Riemannian metric $g$. Also note that usually one suppresses writing down $\sharp$ and $\flat$ and simply writes e.g. $A^{ij}$ instead of $(\sharp A)^{ij}$ since the location of the indices (i.e. "up" or "down") determine which one of them has been raised or lowered. It is however of **prime importance** to always be aware of **which metric** has been used to lower or raise indices!

Recall the gradient of smooth functions on $\mathbb{R}^n$. The gradient of a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ is a vector field given by

$$\operatorname{grad}(f) := \sum_{i=1}^{n} \frac{\partial f}{\partial u^i} \frac{\partial}{\partial u^i} \in \mathfrak{X}(\mathbb{R}^n).$$

There is an invariant generalization for that concept to pseudo-Riemannian manifolds using our above defined musical isomorphisms for which the above formula is precisely the to-be-defined gradient of $f$ on $\mathbb{R}^n$ with respect to the Riemannian metric given by the standard Euclidean scalar product (in **each** tangent space $T_p\mathbb{R}^n$).

**Definition 2.27.** Let $(M, g)$ be a pseudo-Riemannian manifold and $f \in C^\infty(M)$ a smooth function. The **gradient vector field of $f$ with respect to** $g$, $\operatorname{grad}_g(f) \in \mathfrak{X}(M)$, is defined as

$$\operatorname{grad}_g(f) := g^{-1}(df) \in \mathfrak{X}(M).$$

In local coordinates $(x^1, \ldots, x^n)$, $\operatorname{grad}_g(f)$ is of the form

$$\sum_{i,j=1}^{n} \frac{\partial f}{\partial x^i} g^{ij} \frac{\partial}{\partial x^j}.$$

Gradient vector fields are of critical importance in the study of pseudo-Riemannian submanifolds as we find the following description of tangent bundle of pseudo-Riemannian submanifolds.

[27] Ask yourself why this is true!

**Lemma 2.28.** Let $(\overline{M}, g)$ be a pseudo-Riemannian manifold, $M \subset \overline{M}$ a pseudo-Riemannian submanifold of codimension $k$, and identify $T_qM = \iota_*(T_qM) \subset T_q\overline{M}$ for all $q \in M$, where $\iota$ is the inclusion. For $p \in M$ fixed let[28] $f = (f^1, \ldots, f^k) : U \to \mathbb{R}^k$, $U \subset \overline{M}$ open, $p \in U$, be any smooth map of maximal rank such that

$$M \cap U = \{f = 0\} \subset \overline{M}.$$

Then

$$T_qM = \ker(df_q^1) \cap \ldots \cap \ker(df_q^k) \subset T_q\overline{M} \tag{2.5}$$

and

$$(T_qM)^\perp = \operatorname{span}_{\mathbb{R}}\{\operatorname{grad}_g(f^1)_q, \ldots, \operatorname{grad}_g(f^k)_q\} \subset T_q\overline{M} \tag{2.6}$$

for all $q \in M \cap U$. In particular, $T_qM \oplus (T_qM)^\perp = T_q\overline{M}$ for all $q \in M \cap U$.

*Proof.* Fix $q \in M \cap U$ and $v \in T_qM$. For any smooth curve $\gamma : I \to M \subset \overline{M}$, $\gamma'(t)$ is tangential to $M$ for all $t \in I$, which follows by using adapted coordinates. Choose a smooth curve $\gamma : (-\varepsilon, \varepsilon) \to M \subset \overline{M}$ fulfilling $\gamma'(0) = v$. Then for all $1 \leq i \leq k$,

$$df^i(v) = \left.\frac{\partial}{\partial t}\right|_{t=0} (f \circ \gamma) = \left.\frac{\partial}{\partial t}\right|_{t=0} (0) = 0.$$

This shows $T_qM \subset \ker(df_q^1) \cap \ldots \cap \ker(df_q^k)$. On the other hand, $f$ being of maximal rank implies that $df_q^1, \ldots, df_q^k$ are linearly independent. Hence, the intersection of their kernels fulfils

$$\dim(\ker(df_q^1) \cap \ldots \cap \ker(df_q^k)) = \dim(T_q\overline{M}) - k = \dim(T_qM).$$

Hence, (2.5) holds as claimed. For (2.6) one uses that $g$ is pointwise nondegenerate, hence each nonzero vector in $\operatorname{span}_{\mathbb{R}}\{\operatorname{grad}_g(f^1)_q, \ldots, \operatorname{grad}_g(f^k)_q\}$ is not contained in $T_qM = \ker(df_q^1) \cap \ldots \cap \ker(df_q^k)$. By $T_qM \oplus (T_qM)^\perp = T_q\overline{M}$ and by comparing dimensions, (2.6) follows. □

The above lemma tells us how to pointwise understand the tangent space of an ambient manifold of a submanifold as a combination of **tangent and normal parts**. How can we formulate this in a coordinate free, global statement? To do so we need to define bundles along submanifolds.

**Lemma 2.29.** Let $\pi_E : E \to \overline{M}$ be a vector bundle of rank $k$ and let $M$ be a submanifold of $\overline{M}$. Then

$$\pi_{E|_M} : E|_M \to M, \quad (E|_M)_p := \pi_{E|_M}^{-1}(p) := \pi_E^{-1}(p) \quad \forall p \in M, \quad E|_M := \bigsqcup_{p \in M} (E|_M)_p,$$

is a vector bundle of rank $k$ over $M$. It is called **vector bundle along** $M$.

*Proof.* In order to proof this statement it suffices to work in local coordinates. Without loss of generality assume that locally, $M$ is given by an open set in $\mathbb{R}^\ell$, $\ell \leq \dim(\overline{M})$, and that the inclusion $M \subset \overline{M}$ is of the form

$$\iota : (x^1, \ldots, x^\ell) \mapsto (x^1, \ldots, x^\ell, 0, \ldots, 0) \in \mathbb{R}^{\dim(\overline{M})}.$$

The rest of the proof consists of applying the vector bundle chart lemma to the restriction of, after possibly shrinking $U$, the transition functions of $E \to \overline{M}$ in local coordinates to $U \subset \mathbb{R}^{\dim(\overline{M})}$ and observing that the vector parts are, still, smooth. □

[28] If you are not convinced of the existence of such a function $f$ near any point: Prove its existence!

The above lemma might seem more complicated than it actually is. It means that locally, we can make the base space smaller in dimension but keep all possible vectors attached to that smaller set. The most important example for us is restricting the tangent and cotangent bundle of an ambient manifold to a submanifold. In this setting, we see that vector fields along the inclusion map are just sections of the tangent bundle of the ambient manifold along the submanifold. Lemmas 2.28 and 2.29 motivate the following definition.

**Definition 2.30.** Let $(\overline{M}, g)$ be a pseudo-Riemannian manifold and $M \subset \overline{M}$ a pseudo-Riemannian submanifold of codimension $k$. Then the **normal bundle of** $M$, $TM^\perp \to M$, is defined as

$$TM^\perp := \bigsqcup_{p \in M} (T_pM)^\perp,$$

with projection induced by the tangent bundle of $\overline{M}$ along $M$, $T\overline{M}|_M \to M$. In particular we have

$$T\overline{M}|_M = TM \oplus TM^\perp,$$

and the above direct sum is orthogonal with respect to $g$.

In Definition 2.30 above we have split up $T\overline{M}|_M$ into $TM \oplus TM^\perp$, so in particular we have at each point $p \in M$

$$T_p\overline{M}|_M = T_p\overline{M} = T_pM \oplus T_pM^\perp.$$

Fibrewise, $T_pM$ and $T_pM^\perp$ are subvector spaces of $T_p\overline{M}$. Is this a special case of a more general concept for bundles? As you might have guessed already, the answer is yes.

**Definition 2.31.** Let $\pi_E : E \to M$ be a vector bundle. Another vector bundle $\pi_F : F \to M$ is called **subbundle of** $E \to M$ if for all $p \in M$, $F_p$ is a linear subspace of $E_p$, the canonical injection

$$F \hookrightarrow E,$$

given fibrewise by the inclusion $F_p \subset E_p$, is an embedding, $\pi_F = \pi_E|_F$, and for all local trivializations $\phi$ of $E$ the restrictions $\phi|_F$ are local trivializations of $F$. This means that the bundle structure of $F \to M$ and the smooth manifold structure of the total space $F$ are induced by the bundle structure of $E \to M$ and the smooth manifold structure of the total space $E$, respectively.

In the sense of Definition 2.31, the vector bundles $TM \to M$ and $TM^\perp \to M$ of a pseudo-Riemannian submanifold $M \subset \overline{M}$ are both subbundles of $T\overline{M}|_M$.

**Exercise 2.32.** Give a rigorous proof of the above statement. [Hint: You will probably learn exactly the same and at the same time gain more geometrical insight if you prove this statement for surfaces in $\mathbb{R}^3$, while at the same time not having to fight with too many indices.]

The most prominent examples of gradient vector fields and their relation to the normal bundle that are usually used for introductory purposes are level sets of quadratic polynomials which fulfil a certain nondegeneracy condition.

**Example 2.33.**

(i) Let $f : \mathbb{R}^{n+1} \to \mathbb{R}$, $f(u^1, \ldots, u^n) = \sum\limits_i (u^i)^2$ and consider the ambient space $\mathbb{R}^{n+1}$ equipped with its standard Riemannian metric, denoted simply by $\langle \cdot, \cdot \rangle$. Then

$$S^n = \{f = 1\} \subset \mathbb{R}^{n+1}$$

is a Riemannian submanifold of $(\mathbb{R}^{n+1}, \langle\cdot,\cdot\rangle)$ with induced Riemannian metric

$$g := \langle\cdot,\cdot\rangle|_{TS^n \times TS^n}.$$

The normal bundle of $S^n$ realized as a submanifold of $(\mathbb{R}^{n+1}, \langle\cdot,\cdot\rangle)$, $TS^{n\perp}$, is spanned by the **position vector field** $\xi \in \mathfrak{X}(\mathbb{R}^{n+1})$ along $S^n$,

$$\xi : p \mapsto p \quad \forall p \in \mathbb{R}^{n+1},$$

where we have as usual identified $T_p\mathbb{R}^{n+1}$ with $\mathbb{R}^{n+1}$ for all $p \in \mathbb{R}^{n+1}$. The tangent bundle $TS^n$, viewed as a subbundle of $T\mathbb{R}^{n+1}|_{S^n}$, is thus fibrewise given by

$$T_pS^n = \ker(\langle\xi_p, \cdot\rangle) \subset T_p\mathbb{R}^{n+1}.$$

This means that a vector field $X$ along $S^n$ is tangential to $S^n$ if and only if $\langle\xi, X\rangle \equiv 0$. Note that the function $f$ used to define $S^n$ fulfils $f = \langle\xi, \xi\rangle$.

(ii) Next consider $\mathbb{R}^{n+1}$ but now equipped with a pseudo-Riemannian metric given in canonical coordinates by

$$\langle\cdot,\cdot\rangle_\nu := \sum_{i=1}^{n-\nu} du^i \otimes du^i - \sum_{i=n-\nu+1}^{n} du^i \otimes du^i$$

for $1 \le \nu \le n$ fixed. Let $\xi \in \mathfrak{X}(\mathbb{R}^{n+1})$ denote the position vector field and define $f : \mathbb{R}^{n+1} \to \mathbb{R}$, $f := \langle\xi, \xi\rangle$. Then the level sets $\{f = -1\}$[29] are called **hyperboloids**,

$$H^n_\nu := \left\{ \langle\xi,\xi\rangle = \sum_{i=1}^{n-\nu+1} \left(u^i\right)^2 - \sum_{i=n-\nu+2}^{n+1} \left(u^i\right)^2 = -1 \right\} \subset \mathbb{R}^{n+1}.$$

Hyperboloids in $(\mathbb{R}^{n+1}, \langle\cdot,\cdot\rangle_\nu)$ are $n$-dimensional pseudo-Riemannian manifolds with induced pseudo-Riemannian metric of index $\nu - 1$. As for $S^n$,

$$T_pH^n_\nu = \ker(\langle\xi_p, \cdot\rangle_\nu) \subset T_p\mathbb{R}^{n+1}$$

and

$$T_pH^{n\perp}_\nu = \mathbb{R}\xi_p,$$

where $\mathbb{R}\xi_p$ is another commonly used notation for the linear span of **one** vector, that is $\mathrm{span}_\mathbb{R}\{\xi_p\}$. In the case $n = 3$, $\nu = 1$, $H^3_1$ is known as **two-sheeted hyperboloid**, and for $n = 3$, $\nu = 2$, $H^3_2$ is the **one-sheeted hyperboloid**.

**Exercise 2.34.** Prove the claims in Example 2.33.

In Example 2.33 we used the term that a vector field spans a vector bundle. Conceptually, this belongs in the setting of frames of vector bundles, which generalize the concept of a basis of a vector space.

**Definition 2.35.** Let $E \to M$ be a vector bundle of rank $k$. A **(local) frame of** $E$ over $U \subset M$, $U$ open, is a set of $k$ (local) sections

$$\{s_i \in \Gamma(E|_U),\ 1 \le i \le k\},$$

such that for all $p \in U$ fixed, the vectors $s_i(p) \in E_p$, $1 \le i \le k$, are linearly independent. Equivalently,

$$\mathrm{span}_\mathbb{R}\{s_i(p) \in E_p \mid 1 \le i \le k\} = E_p$$

for all $p \in U$.

[29] Be aware of the sign!

**Exercise 2.36.** Show that every local section $s \in \Gamma(E|_U)$ in a vector bundle $E$ can be written as a $C^\infty(U)$-linear combination of the elements of a local frame of $E \to M$ over $U$. Check that these prefactors in $C^\infty(U)$ are uniquely determined for any given local section.

Local frames are very useful in order to check if subsets of a certain form of given vector bundles are subbundles.

**Lemma 2.37.** Let $E \to M$ be a vector bundle of rank $k$ and suppose that for some $\ell \in \mathbb{N}$ with $1 \leq \ell \leq k$ we are given a linear subspace $F_p \subset E_p$ of constant dimension $\ell$ for all $p \in M$. Then $\bigsqcup_{p\in M} F_p \to M$ is, with all data necessary induced by $E \to M$, a subbundle of $E \to M$ if and only if for every $p \in M$ we can find a local frame $\{s_1, \ldots, s_k\}$ of $E|_U \to U$, $U \subset M$ an open neighbourhood of $p$, such that for all $q \in U$, $\{s_1(q), \ldots, s_\ell(q)\}$ is a basis of $F_q$.

*Proof.* [L1, Lem. 10.32] □

Subbundles might look very complicated at first glance, but at least locally we can use the above lemma to always describe them as follows.

**Lemma 2.38.** Let $F \to M$ be a subbundle of rank $\ell$ of a vector bundle $E \to M$ of rank $k > \ell$. For any $p \in M$ we can find an open neighbourhood $U \subset M$ of $p$ and a local trivialization of $E \to M$ over $U$, $\phi : E|_U \to U \times \mathbb{R}^k$, such that

$$\phi(\iota(F|_U)) = U \times \{(v^1, \ldots, v^\ell, 0, \ldots, 0) \mid (v^1, \ldots, v^\ell) \in \mathbb{R}^\ell\} \subset U \times \mathbb{R}^k.$$

In the above equation, $\iota : F \hookrightarrow E$ denotes the inclusion map.

*Proof.* We use Lemma 2.37. Choose a local frame $\{s_1, \ldots, s_k\}$ of $E \to M$ over $U \subset M$, such that $\{s_1, \ldots, s_\ell\}$ is a local frame of $F \to M$ over $U$. The inverse of the smooth map

$$\eta : U \times \mathbb{R}^k \to E|_U, \quad (p, v^1, \ldots, v^k) \mapsto \sum_{i=1}^k v^i s_i(p)$$

is smooth and a local trivialization of $E \to M$ over $U$. This follows from the implicit function theorem. We obtain

$$\eta^{-1}(\iota(F|_U)) = U \times \{(v^1, \ldots, v^\ell, 0, \ldots, 0) \mid (v^1, \ldots, v^\ell) \in \mathbb{R}^\ell\},$$

so setting $\phi = \eta^{-1}$ finishes the proof. □

Lemma 2.38 means that locally up to vector bundle isomorphisms, subbundles of vector bundles look like the inclusion in the first $\ell$ factors of the vector parts in $U \times \mathbb{R}^\ell \to U$ into $U \times \mathbb{R}^k \to U$.

In the special case of the tangent bundle of an $n$-dimensional smooth manifold, a local frame of $TM \to M$ over $U \subset M$ open is a set of $n$ vector fields

$$\{X_i,\ 1 \leq i \leq n\}, \quad X_i \in \mathfrak{X}(U) \quad \forall 1 \leq i \leq n,$$

such that for all $p \in U$, $\{(X_i)_p,\ 1 \leq i \leq n\}$ is a set of linearly independent vectors. This in particular means that

$$\mathrm{span}_{\mathbb{R}}\{(X_i)_p,\ 1 \leq i \leq n\} = T_pU \quad \forall p \in U,$$

and by Exercise 2.36 we can for each local vector field $X \in \mathfrak{X}(U)$ find a uniquely determined set of local functions $f_i \in C^\infty(U)$, $1 \leq i \leq n$, such that $X = \sum_{i=1}^n f_i X_i$.

Next we will use the language of local frames and subbundles to split up the $(0,2)$-tensor bundle $T^{0,2}M \to M$ over a smooth manifold into symmetric and antisymmetric parts. This construction and similar constructions are important to properly understand our upcoming study of curvature and the exterior differential on $k$-form analogues for smooth manifolds. Recall the following fact from linear algebra.

**Lemma 2.39.** Let $V$ be a finite-dimensional real vector space with basis $\{v_1, \dots, v_n\}$. Then

$$V \otimes V \cong \mathrm{Sym}^2(V) \oplus \Lambda^2 V,$$

where $\mathrm{Sym}^2(V) := \mathrm{span}_{\mathbb{R}}\{v_i \otimes v_j + v_j \otimes v_i,\ 1 \le i,j \le n\}$ and $\Lambda^2 V := \mathrm{span}_{\mathbb{R}}\{v_i \otimes v_j - v_j \otimes v_i,\ 1 \le i,j \le n\}$.

When viewed as matrices, the direct sum in Lemma 2.39 corresponds to writing a square matrix as its symmetric and antisymmetric parts, which are uniquely determined. One writes

$$v_i v_j := \frac{1}{2}(v_i \otimes v_j + v_j \otimes v_i), \quad v_i \wedge v_j := v_i \otimes v_j - v_j \otimes v_i,$$

and has $v_i \otimes v_j = v_i v_j + \frac{1}{2} v_i \wedge v_j$ for all $1 \le i,j \le n$. In particular $v_i v_i = v_i \otimes v_i$. This concept translates to local frames of vector bundles. We obtain the following.

**Definition 2.40.** Let $M$ be a smooth manifold and let $(x^1, \dots, x^n)$ be local coordinates on $U \subset M$. The **bundle of symmetric $(0,2)$-tensors** on $M$ is the subbundle

$$\mathrm{Sym}^2(T^*M) \subset T^{0,2}M$$

with local frame over $U$ given by $\left\{dx^i dx^j = \frac{1}{2}(dx^i \otimes dx^j + dx^j \otimes dx^i),\ 1 \le i,j \le n\right\}$. Sections in $\mathrm{Sym}^2(T^*M) \to M$ are precisely symmetric $(0,2)$-tensor fields, which in particular include all possible pseudo-Riemannian metrics on $M$. On the other hand we have the **anti-symmetric $(0,2)$-tensors** on $M$,

$$\Lambda^2 T^*M \subset T^{0,2}M,$$

with local frame over $U$ given by $\{dx^i \wedge dx^j = dx^i \otimes dx^j - dx^j \otimes dx^i,\ 1 \le i,j \le n\}$. Sections in $\Lambda^2 T^*M \to M$ are called **2-forms** and are denoted by $\Omega^2(M)$. Local sections in $\Lambda^2 T^*M \to M$ over $U \subset M$, $U$ open, are denoted by $\Omega^2(U)$.

With Definition 2.40, a pseudo-Riemannian metric $g$ on $M$ can be written locally as

$$g = \sum_{i,j} g_{ij}\, dx^i \otimes dx^j = \sum_{i,j} g_{ij}\, dx^i dx^j.$$

Make sure you understand why the second equivalence in the above equation holds true! Also be aware that the rightmost sum in the above equation holds a certain error potential when going from tensor to matrix notation. For example, the pseudo-Riemannian metric $dxdy$ on $\mathbb{R}^2$ with canonical coordinates $(x,y)$ is in matrix notation given by

$$dx\, dy \text{ “}=\text{” } \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}.$$

Make absolutely sure to understand this.

Now suppose that we are given a smooth manifold $M$ and a symmetric $(0,2)$-tensor field $g \in \mathcal{T}^{0,2}(M)$. At each point $p \in M$, we can associate to $g_p$ a natural number called its index as follows.

**Definition 2.41.** The **index** of a symmetric $(0,2)$-tensor field $g \in \mathcal{T}^{0,2}(M)$ at $p \in M$ is defined as

$$\nu(p) := \text{ number of negative eigenvalues of } g_p,$$

where $g_p$ is viewed as symmetric matrix in local coordinates, i.e.

$$g_p = \sum_{ij} g_{ij}(p) dx^i \otimes dx^j.$$

**Exercise 2.42.** Verify that the index in Definition 2.41 is well defined, meaning that it does not depend on the choice of local coordinates.

How can we use the definition of the pointwise index of a symmetric $(0,2)$-tensor field $g$ to test if $g$ is a pseudo-Riemannian metric or not? The answer is as follows.

**Proposition 2.43.** Let $M$ be a connected smooth manifold and $g \in \mathcal{T}^{0,2}(M)$ a symmetric $(0,2)$-tensor field that is nondegenerate[30] in all fibres $T_pM$, $p \in M$. Then $g$ is a pseudo-Riemannian metric.

*Proof.* It suffices to show that the index of $g$, $\nu : M \to \mathbb{N}_0$, is continuous, where $\mathbb{N}_0$ is equipped with the discrete topology. In order to do so it suffices by using local charts and smooth curves to prove that the number of negative eigenvalues of any smooth function with values in the symmetric $n \times n$-matrices,

$$A : I \to \mathrm{Sym}^2((\mathbb{R}^*)^n), \quad t \mapsto A(t) \in \mathrm{Sym}^2((\mathbb{R}^*)^n),$$

such that $A(t)$ is nondegenerate for all $t \in I$, is locally constant. This follows from the continuity of the eigenvalues of $A(t)$ viewed each as functions[31] of $t$. To see this, consider the characteristic polynomial of $A(t)$ in dependence of $t \in I$,

$$P_t(\lambda) := \det(A(t) - \lambda \mathbb{1}).$$

$P_t(\lambda)$ can be written as

$$P_t(\lambda) = \sum_{i=0}^{n} a_i(t) \lambda^i,$$

where $a_i : I \to \mathbb{R}$ is smooth for all $0 \leq i \leq n$ and $a_n(t) \equiv (-1)^n$. Thus, the proof reduces to the continuous dependence of roots of a polynomial of fixed degree with smoothly varying prefactors and fixed highest order monomial. Since we already know that the eigenvalues must be real by the symmetry condition of $A(t)$, the result follows from [Z]. □

Next suppose that we are given just a smooth manifold and want to construct a pseudo-Riemannian metric. While this problem is usually difficult if the index of our metric is supposed to be positive (and not equal to the dimension of our manifold), for the Riemannian case, that is for vanishing index, we have the following nice result.

**Proposition 2.44.** Let $M$ be a smooth manifold. Then there exists a Riemannian metric $g$ on $M$.

*Proof.* This is Exercise 2.12 (i). We remark here that if $g$ is a Riemannian metric and $h$ is a symmetric $(0,2)$-tensor with compact support, then for $\varepsilon > 0$ small enough $g + h$ will still be a Riemannian metric. This means that our constructed metric is far from unique. □

[30] "nondegenerate" = at given point nondegenerate symmetric bilinear form

[31] More precisely: There exists a choice of $n$ nowhere vanishing continuous functions $\lambda_n : I \to \mathbb{R} \setminus \{0\}$, such that for all $t \in I$, the set $\{(1, \lambda_1(t)), \ldots, (n, \lambda_n(t))\}$ is precisely the set of (indexed) eigenvalues of $A(t)$.

Let us return to isometries of pseudo-Riemannian manifolds, cf. Definition 2.13. We already noted that they form a group, so it is reasonable to ask the following. It is clear that the identity is always an isometry, so how can we perturb it infinitesimally while preserving the isometry property? To answer that question we will use our knowledge of the Lie derivative and local flows of vector fields.

**Proposition 2.45.** Let $(M, g)$ be a pseudo-Riemannian manifold and let $X \in \mathfrak{X}(M)$. Suppose that for every local flow $\varphi : I \times U \to M$ of $X$, $\varphi_t : U \to M$ is an isometry for all $t \in I$. Then $\mathcal{L}_X g = 0$. The converse statement also holds true.

*Proof.* A local flow $\varphi : I \times U \to M$ of $X$ is an isometry of $(M, g)$ for all $t \in I$ if and only if

$$g_p(v, w) = g_{\varphi_t(p)}(d\varphi_t(v), d\varphi_t(w))$$

for all $t \in I$, $p \in M$, $v, w \in T_pM$. Hence,

$$\begin{aligned}
(\mathcal{L}_X g)(v, w) &= \left( \frac{\partial}{\partial t}\bigg|_{t=0} (\varphi_t^* g)_p \right)(v, w) \\
&= \frac{\partial}{\partial t}\bigg|_{t=0} \Big( g_{\varphi_t(p)}(d\varphi_t(v), d\varphi_t(w)) \Big) \\
&= \frac{\partial}{\partial t}\bigg|_{t=0} g_p(v, w) \\
&= 0.
\end{aligned}$$

Since $p \in M$, $v, w \in T_pM$ were arbitrary, this shows that $\mathcal{L}_X g = 0$.

For the other direction note that $d\varphi_{t_0} : T_pM \to T_{\varphi_{t_0}(p)}M$ is a linear isomorphism for all $t_0 \in I$ and, by the group property of local flows, we obtain $d\varphi_{t+t_0} = d\varphi_t d\varphi_{t_0}$ for $t$ small enough. We calculate for any $t_0 \in I$, $v, w \in T_pM$,

$$\begin{aligned}
0 &= (\mathcal{L}_X g)(d\varphi_{t_0}(v), d\varphi_{t_0}(w)) \\
&= \frac{\partial}{\partial t}\bigg|_{t=0} \Big( g_{\varphi_t(\varphi_{t_0}(p))}(d\varphi_t d\varphi_{t_0}(v), d\varphi_t d\varphi_{t_0}(w)) \Big) \\
&= \frac{\partial}{\partial t}\bigg|_{t=0} \Big( g_{\varphi_{t+t_0}(p)}(d\varphi_{t+t_0}(v), d\varphi_{t+t_0}(w)) \Big) \\
&= \frac{\partial}{\partial s}\bigg|_{s=t_0} \Big( g_{\varphi_s(p)}(d\varphi_s(v), d\varphi_s(w)) \Big).
\end{aligned}$$

This shows that the smooth function $I \ni s \mapsto g_{\varphi_s(p)}(d\varphi_s(v), d\varphi_s(w)) \in \mathbb{R}$ is constant for all $v, w \in T_pM$ and, hence, that the local flow of $X$ consists of isometries for any fixed time parameter. □

**Definition 2.46.** Vector fields as in Proposition 2.45, that is $\mathcal{L}_X g = 0$ for $X \in \mathfrak{X}(M)$, $(M, g)$ a pseudo-Riemannian manifold, are called **Killing**[32] **vector fields** of $(M, g)$.

Proposition 2.45 in particular means that Killing vector fields generate **local one parameter groups of isometries**. Killing vector fields, as a linear subspace of all vector fields, have the following algebraic structure.

**Lemma 2.47.** Let $(M, g)$ be a pseudo-Riemannian manifold. Killing vector fields form a Lie subalgebra of $(\mathfrak{X}(M), [\cdot, \cdot])$, meaning that for any Killing vector fields $X, Y \in \mathfrak{X}(M)$, $[X, Y]$ is also a Killing vector field.

[32] Wilhelm Karl Joseph Killing (1847 – 1923)

*Proof.* Exercise. [Hint: Use the Jacobi identity $\mathcal{L}_{[X,Y]}Z = \mathcal{L}_X(\mathcal{L}_Y Z) - \mathcal{L}_Y(\mathcal{L}_X Z)$ for all $X, Y, Z \in \mathfrak{X}(M)$.] □

In fact, one can show more if $M$ is Riemannian, but the proof of the following is beyond the scope of this course.

**Theorem 2.48.** Let $(M, g)$ be a connected Riemannian manifold of dimension $n$. The Lie algebra of Killing vector fields is finite-dimensional of dimension at most $\frac{1}{2}n(n+1)$.

*Proof.* [KN, Thm 3.3], in the corresponding chapter the structure of the isometry group as a Lie group acting on $M$ is also treated. □

Let us look at some explicit examples of Killing vector fields.

**Example 2.49.**

(i) Let $A$ be an $(n+1) \times (n+1)$ skew real matrix, that is $A^T = -A$. Then $e^{At} \in \mathrm{O}(n+1)$ for all $t \in \mathbb{R}$. Then the vector field $X \in \mathfrak{X}(S^n)$ given by

$$X_p = \frac{\partial}{\partial t}\bigg|_{t=0} \left(e^{At}p\right) \in T_p S^n$$

is a Killing vector field of the standard round metric on $S^n$, that is the restriction of the pointwise Euclidean scalar product in the ambient manifold $\mathbb{R}^{n+1}$. Note that $e^{A\cdot} : \mathbb{R} \times S^n \to \mathbb{R}^{n+1}$, $(t, v) \mapsto e^{At}v$ is the global flow of $X$.

(ii) Consider $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_\nu)$ for any $0 \leq \nu \leq n$ as in Example 2.33 and fix $(c^1, \ldots, c^n) \in \mathbb{R}^n$. Then $X \in \mathfrak{X}(\mathbb{R}^n)$, $X = \sum_i c^i \frac{\partial}{\partial u^i}$, is a Killing vector field.

(iii) Let $(M, g)$ and $(N, h)$ be pseudo-Riemannian manifolds, $X$ a Killing vector field on $(M, g)$, and $Y$ a Killing vector field on $(N, h)$. Then $X + Y$ is a Killing vector field on $(M \times N, g \oplus h)$.

Now suppose that we are given a pseudo-Riemannian manifold and do not know which vector fields are Killing vector fields. How do we approach this problem, at least locally?

**Lemma 2.50.** Let $(M, g)$ be a pseudo-Riemannian manifold. Then $X \in \mathfrak{X}(M)$ is a Killing vector field if and and only if it fulfils

$$\sum_{k=1}^{n} \left( X^k \frac{\partial g_{ij}}{\partial x^k} + \frac{\partial X^k}{\partial x^i} g_{jk} + \frac{\partial X^k}{\partial x^j} g_{ik} \right) = 0 \quad \forall 1 \leq i, j \leq n$$

for all local coordinates $(x^1, \ldots, x^n)$ on $M$.

*Proof.* Exercise. □

**Remark 2.51.** $k$-forms and corresponding bundle, exterior algebra structure, $d$-complex, cartans magic formula as exercise, induced volume form Note: Not yet decided how to implement, in full range far too much for this course, same for next remark.

**Remark 2.52.** Killing vector fields on induced volume form

**Remark 2.53.** We have seen the formal definition of isometries between pseudo-Riemannian manifolds and how to describe infinitesimal isometries of a given pseudo-Riemannian manifold. It is, however, in general a very difficult task to verify or disprove that two given pseudo-Riemannian manifolds are isometric. For a reasonable approach to this kind of problem we will need the definitions of the different curvatures of a pseudo-Riemannian manifold, but in order to introduce these we will need to study so-called connections in vector bundles, which is what we will do next.

## 2.2 Connections in vector bundles

The subject of this section, *connections in vector bundles*, is motivated as follows. Suppose that we are given a connected smooth manifold $M$. We know how to "connect" two points, namely by specifying a smooth curve starting at one and ending at the other. Next, suppose that we are given two tangent vectors $v, w \in TM$ that are contained in different fibres. How do we connect $v$ and $w$ or, more generally, their fibres? Since the total space $TM$ of the tangent bundle is a smooth manifold as well we can of course connect $v$ and $w$, viewed as points in $TM$, via a smooth curve in $TM$. But this does in a sense allow for too much freedom of choice, as we want in a some sense canonical way *connect* $T_{\pi(v)}M$ with $T_{\pi(w)}M$ via a linear isomorphism. The solution to this problem is to construct a so-called connection in $TM \to M$ with respect to a given pseudo-Riemannian metric, such that all of the latter identifications of fibres are not only linear isomorphisms, but also linear isometries. Furthermore we require that if we go around an infinitesimal parallelogram in $M$ and consider the identification of tangent spaces along the corresponding piecewise smooth curve, we should end up with the identity. In the following we will in detail describe these concepts and how to actually perform calculations with them. Before actually defining what a connection in a vector bundle is, we will study a motivating example.

**Remark 2.54.** INSERT PIC! Consider the smooth curve $\gamma : \mathbb{R} \to \mathbb{R}^2$,

$$\gamma(t) = \left(\begin{smallmatrix} t \\ 1 \end{smallmatrix}\right) \quad \forall t \in \mathbb{R}.$$

At $\gamma(0) = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$, let $v \in T_{\gamma(0)}\mathbb{R}^2$ be given by

$$v = \left(\gamma(0), \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)\right).$$

How would a reasonable approach to transport $v$ along $\gamma$ look like? Intuitively, we define a vector field along $\gamma$

$$X_\gamma : \mathbb{R} \to T\mathbb{R}^2, \quad X_{\gamma(t)} = (\gamma(t), v) \quad \forall t \in \mathbb{R},$$

so that $X_{\gamma(0)} = v$ and $X_{\gamma(t)}$ represents $v$ "transported" to $\gamma(t)$. Is this choice canonical in any meaningful sense? The answer is yes, but we will need to develop a lot of tools to see that. Observe that $X_\gamma$ is the restriction of $X := \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \in \mathfrak{X}(\mathbb{R}^2)$ to the image of $\gamma$, where $(x, y)$ are the canonical coordinates on $\mathbb{R}^2$. How does our above construction look like in different coordinates? In polar coordinates $(r, \varphi)$ on $\mathbb{R}^2 \setminus \{(x, 0),\ x \leq 0\}$ we have

$$X = (\cos(\varphi) + \sin(\varphi))\frac{\partial}{\partial r} + \frac{1}{r}(\cos(\varphi) - \sin(\varphi))\frac{\partial}{\partial \varphi}.$$

The curve $\gamma$ in polar coordinates is given by

$$\gamma(t) = \left(\begin{smallmatrix} r(\gamma(t)) \\ \varphi(\gamma(t)) \end{smallmatrix}\right) = \left(\begin{matrix} \sqrt{1+t^2} \\ \arctan\left(\sqrt{1+t^2}^{\,-1}\right) \end{matrix}\right).$$

Thus, when sketching $X$ along $\gamma$ in polar coordinates, our initial choice of transporting $v$ along $\gamma$ does not look, in any way, "parallel" any more. So how do we solve this problem? How can we define what transporting vectors along curves in a parallel way should be, and all that in a coordinate independent way? To do that we will need not only connections in vector bundles, but also need to single out a certain choice for a connection in the tangent bundle with the help of a pseudo-Riemannian metric, the so called Levi-Civita[33] connection.

First, we will introduce the most general concept of a connection in a vector bundle and then focus on the tangent bundle and its various tensor bundles.

[33] Tullio Levi-Civita (1873 – 1941)

**Definition 2.55.** Let $E \to M$ be a vector bundle. A **connection in** $E \to M$ is a bilinear map

$$\nabla : \mathfrak{X}(M) \times \Gamma(E) \to \Gamma(E), \quad (X, s) \mapsto \nabla_X s,$$

that is $C^\infty(M)$-linear[34] in the first entry, i.e.

$$\nabla_{fX} s = f \nabla_X s \quad \forall f \in C^\infty(M), X \in \mathfrak{X}(M), s \in \Gamma(E),$$

and fulfils the Leibniz rule

$$\nabla_X(fs) = X(f)s + f\nabla_X s \quad \forall f \in C^\infty(M), X \in \mathfrak{X}(M), s \in \Gamma(E).$$

The last condition can be written as $\nabla(fs) = s \otimes df + f\nabla s$. Note that a connection in $E \to M$ can be canonically extended to local sections.

The defining conditions of a connection hint at their interpretation as certain types of derivatives. In comparison with the Lie derivative (for $E = TM$), we see that it differs from a connection by the tensoriality property in the first argument. Recall that in general $\mathcal{L}_{fX} Y \neq f\mathcal{L}_X Y$ for vector fields $X, Y \in \mathfrak{X}(M)$. This fact points to preferring a connection over the Lie derivative for the concept of a derivative of vector fields since a derivative should ideally only depend on the direction in which we are differentiating and the local behaviour of the section we are taking the derivative of, and **not** on the local behaviour of our direction as part of a vector field. Next, we need to ask ourselves how to actually calculate with a connection. The answer lies in the use of local frames.

**Definition 2.56.** Let $\nabla$ be a connection in a vector bundle $E \to M$ of rank $\ell$ and let $\{s_1, \dots, s_\ell\}$ be a local frame of $E \to M$ over $U \subset M$ open, such that there exist local coordinates $(x^1, \dots, x^n)$ on $U \subset M$. This can always be achieved after possibly shrinking $U$. Let further $\dim(M) = n$. Define

$$\nabla s_i := \omega_i, \quad \omega_i(X) = \nabla_X s_i \quad \forall X \in \mathfrak{X}(M),$$

for all $1 \leq i \leq \ell$. Then each $\omega_i$ is an **$E$-valued 1-form**[35], that is $\omega_i \in \Gamma(E|_U \otimes T^*M|_U)$ for all $1 \leq i \leq \ell$. Thus we have

$$\omega_i = \sum_{j=1}^{n} \omega_{ij} \otimes dx^j$$

for all $1 \leq i \leq \ell$, where $\omega_{ij} \in \Gamma(E|_U)$ for all $1 \leq i \leq \ell$, $1 \leq j \leq n$. We can further write

$$\omega_{ij} = \sum_{k=1}^{\ell} \omega_{ij}^k s_k,$$

with $\omega_{ij}^k \in C^\infty(U)$ for all $1 \leq i \leq \ell$, $1 \leq j \leq n$, $1 \leq k \leq \ell$. Recall that for any local section $s \in \Gamma(E|_U)$ we can write $s = \sum_{i=1}^{k} f^i s_i$ with $f^i$, $1 \leq i \leq k$, uniquely determined for $s$. We obtain the general formula

$$\nabla s = \sum_{i=1}^{\ell} s_i \otimes df^i + \sum_{j=1}^{n} \sum_{i,k=1}^{\ell} f^i \omega_{ij}^k s_k \otimes dx^j. \tag{2.7}$$

On the other hand we might write

$$\nabla s_i = \omega_i = \sum_{k=1}^{\ell} s_k \otimes \omega_i^k$$

[34] a.k.a. "tensorial"

[35] That means: Plug in a (local) vector field, get a (local) section in $E$.

for all $1 \leq i \leq k$, where $\omega_i^k \in \Omega^1(U)$ for all $1 \leq i, k \leq \ell$. The $\omega_i^k$ are called **connection 1-forms** and determine the connection $\nabla$ in $E|_U \to U$ completely. We can view $(\omega_i^k)$ as an $(\ell \times \ell)$-matrix valued map where each entry is a local 1-form on $M$.

**Remark 2.57. Warning:** Connections, and with them the corresponding connection 1-forms, do **not** transform like tensors if $E$ is some tensor power of $TM$. The reason for that is that a connection itself is not tensorial in the second argument, so changing frames will lead to the new connection 1-forms to depend on the partial derivatives of the corresponding transformation. The explicit transformation behaviour will be studied in detail for connections in $TM \to M$.

The benefit of writing down a connection locally using its connection 1-form is the easy-to-formulate transformation behaviour when changing the local frame of $E$ (**not** the local coordinates on $M$). Changing the frame without changing the coordinates on the base space is not too important for our purposes, but nevertheless a nice exercise. Observe in particular that the transformation is not tensorial, that is not simply the pullback in the frame part.

**Exercise 2.58.** Let $\nabla$ be a connection in a vector bundle $E \to M$ of rank $\ell$. Let $\{s_1, \ldots, s_\ell\}$ and $\{\widetilde{s}_1, \ldots, \widetilde{s}_\ell\}$ be local frames of $E$ over a chart neighbourhood $U \subset M$, equipped with local coordinates $(x^1, \ldots, x^n)$, that are related by the $(\ell \times \ell)$-matrix valued smooth map

$$A : U \to \mathrm{GL}(\ell), \quad (s_1, \ldots, s_\ell) \cdot A = (\widetilde{s}_1, \ldots, \widetilde{s}_\ell).$$

Let $(\omega_i^k)$ denote the matrix of connection 1-forms with respect to the local frame $\{s_1, \ldots, s_\ell\}$ and $(\widetilde{\omega}_i^k)$ the matrix of connection 1-forms with respect to the local frame $\{\widetilde{s}_1, \ldots, \widetilde{s}_\ell\}$. Show that the two matrices of connection 1-forms are related by

$$(\widetilde{\omega}_i^k) = A^{-1} dA + A^{-1}(\omega_i^k) A.$$

In the above equation, $dA$ denotes the differential of the map $A : U \to \mathrm{GL}(\ell)$, where we identify $T\mathrm{GL}(\ell) \cong \mathrm{GL}(\ell) \times \mathrm{End}(\mathbb{R}^\ell)$[36].

Connections, just like tangent vectors, are local objects in the following sense.

**Lemma 2.59.** Let $\nabla$ be a connection in a vector bundle $E \to M$ of rank $\ell$. Let $U \subset M$ be open and suppose that for two vector fields $X, Y \in \mathfrak{X}(M)$ and two sections in $E \to M$, $s, \widetilde{s}$, we have

$$X|_U = Y|_U, \quad s|_U = \widetilde{s}|_U.$$

Then $\nabla_X s$ and $\nabla_Y \widetilde{s}$ coincide on $U$.

*Proof.* Note that $\nabla_X s|_U = \nabla_Y s|_U$, which follows from the tensoriality property in the first argument of any connection. It thus suffices to show that $\nabla_X s|_U = \nabla_X \widetilde{s}|_U$. Using Definition 2.56 we write, after possibly shrinking $U$, $s$ and $\widetilde{s}$ in a local frame $\{s_1, \ldots, s_\ell\}$ of $E|_U$,

$$s|_U = \sum_{i=1}^{\ell} f^i s_i, \quad \widetilde{s}|_U = \sum_{i=1}^{\ell} \widetilde{f}^i s_i,$$

with $f_i, \widetilde{f}_i \in C^\infty(U)$ for all $1 \leq i \leq \ell$. Equation (2.7) and $f^i = \widetilde{f}^i$ for all $1 \leq i \leq n$ by assumption now imply that $s$ and $\widetilde{s}$ coincide on $U$ imply that $\nabla_X s|_U = \nabla_X \widetilde{s}|_U$ holds true.

If one prefers to work without coordinates or frames, one can proceed as follows. By the linearity in the second argument, $\nabla_X s$ and $\nabla_X \widetilde{s}$ coincide in $U$ if and only if $\nabla_X (s - \widetilde{s})|_U \equiv 0$. Hence, it suffices to prove $\nabla_X s|_U = 0$ if $s|_U = 0$. Fix $p \in U$ and choose a bump function

[36] Recall that $\mathrm{GL}(\ell)$ is open in the real $(\ell \times \ell)$-matrices.

$b \in C^\infty(M)$ and an open neighbourhood of $p$, $V \subset U$, that is precompact in $U$, such that $b|_V \equiv 1$ and $\operatorname{supp}(b) \subset U$. Then by the Leibniz rule

$$0 = \nabla_X 0|_p = \nabla_X (bs)|_p = X(b)s|_p + b(p)\nabla_X s|_p = \nabla_X s|_p.$$

Since $p \in U$ was arbitrary, this finishes the proof. □

Lemma 2.59 means that $(\nabla_X s)(p)$ for any $p \in M$ depends only on $X_p \in T_pM$ and the restriction of $s$ to an arbitrary small open neighbourhood of $p$ in $M$.

Before continuing, we remark that there is a connection we are probably already aware of, although not under that name.

**Example 2.60.** Consider $\mathbb{R}^n$ with canonical coordinates $(u^1, \ldots, u^n)$ and induced global frame $\left\{\frac{\partial}{\partial u^1}, \ldots, \frac{\partial}{\partial u^n}\right\}$ of $T\mathbb{R}^n \to \mathbb{R}^n$. Vector fields on $\mathbb{R}^n$ can be, as we described before introducing vector fields on smooth manifolds, viewed as smooth vector valued functions. So a reasonable approach for a connection, defined in our choice of coordinates, is

$$\nabla_X Y := \sum_i X(Y^i)\frac{\partial}{\partial u^i} \in \mathfrak{X}(\mathbb{R}^n)$$

for all vector fields $X = \sum_i X^i \frac{\partial}{\partial u^i}$ and $Y = \sum_i Y^i \frac{\partial}{\partial u^i}$. This means that, in canonical coordinates, we differentiate $Y$ entrywise in $X$-direction. One verifies that the so-defined operator $\nabla$ in fact is a connection in $T\mathbb{R}^n \to \mathbb{R}^n$. This construction is, however, not coordinate-independent, meaning that in different coordinates, $\nabla_X Y$ will not be the entrywise differentiation of $Y$ in $X$-direction. Note that all connection 1-forms of the above connection identically vanish.

As described in the above example, we need to investigate how a connection in $TM \to M$, written in a choice of local coordinates and induced local frame, behaves under a change of coordinates. This problem is equivalent to understanding how the connection 1-forms transform under a change of coordinates and induced change of local frame of $TM \to M$. To do so we will introduce the so-called Christoffel symbols, which are commonly used to describe connections and, consequently, connection 1-forms in the tangent bundle of a smooth manifold. The difference to a connection in a general bundle is that a choice of coordinates on $M$ automatically gives us a local frame in $TM$.

**Definition 2.61.** Let $\nabla$ be a connection in $TM \to M$ and let $(x^1, \ldots, x^n)$ be local coordinates on $U \subset M$. Then in the induced local frame of $TM$,

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \sum_{k=1}^n \Gamma_{ij}^k \frac{\partial}{\partial x^k},$$

where $\Gamma_{ij}^k \in C^\infty(M)$, $1 \leq i, j, k \leq n$. The terms $\Gamma_{ij}^k$ are called **Christoffel[37] symbols** of the connection $\nabla$ with respect to the chosen local coordinates $(x^1, \ldots, x^n)$. The Christoffel symbols specify the connection $\nabla$ in $TM|_U \to U$ completely, meaning in particular that two connections in $TM \to M$ coincide if they have the same Christoffel symbols for all local coordinates on $M$. In comparison with the most general case, the Christoffel symbols are for the special case of the tangent bundle with induced local frame precisely the terms $\omega_{ij}^k$ in equation (2.7).

Note that, if one wants to be very precise, it is at this point not clear if every manifold admits a connection in its tangent bundle. This is either a not so easy exercise or a good excuse to consult [L1, Prop. 4.5]. The answer is yes, every manifold admits a connection in its tangent bundle, and the space of connections is, in a sense, very big, see also Definition 2.81.

Similar to Exercise 2.58, but with the difference that we now also change the local coordinates on the base manifold, we obtain the following transformation rule for Christoffel symbols.

[37] Elwin Bruno Christoffel (1829 – 1900)

**Lemma 2.62.** Let $M$ be an $n$-dimensional smooth manifold, $\nabla$ a connection in $TM \to M$. Let further $\varphi = (x^1, \dots, x^n)$ and $\psi = (y^1, \dots, y^n)$ local coordinate systems on an open set $U \subset M$. Let $\Gamma_{ij}^k$ denote the Christoffel symbols of $\nabla$ with respect to $\varphi$ and let $\widetilde{\Gamma}_{ij}^k$ denote the Christoffel symbols of $\nabla$ with respect to $\psi$. Then the following identity holds:

$$\Gamma_{ij}^k = \sum_\rho \frac{\partial^2 y^\rho}{\partial x^i \partial x^j} \frac{\partial x^k}{\partial y^\rho} + \sum_{\mu,\nu,\rho} \frac{\partial y^\mu}{\partial x^i} \frac{\partial y^\nu}{\partial x^j} \frac{\partial x^k}{\partial y^\rho} \widetilde{\Gamma}_{\mu\nu}^\rho.$$

*Proof.* Direct calculation using $\frac{\partial}{\partial x^i} = \sum_j \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j}$ and the corresponding inverse formula. □

Suppose that we are given a connection $\nabla$ in $TM \to M$. Then $\nabla$ induces a connection in all tensor powers $T^{r,s}M \to M$ of the tangent bundle by requiring compatibility with contractions.

**Lemma 2.63.** Let $\nabla$ be a connection in $TM \to M$. Then $\nabla$ induces a connection in each tensor bundle[38] $T^{r,s}M \to M$, $r \geq 0$, $s \geq 0$, such that

(i) the induced connection in $T^{1,0}M \cong TM \to M$ coincides with $\nabla$,

(ii) $\nabla f = df$ for all $f \in \mathcal{T}^{0,0}(M) = C^\infty(M)$,

(iii) the induced connection is a tensor derivation in the second argument, meaning that

$$\nabla(A \otimes B) = (\nabla A) \otimes B + A \otimes (\nabla B)$$

whenever the tensor field $A \otimes B$ is defined,

(iv) the induced connections commute with all possible contractions, meaning that for any contraction $C : \mathcal{T}^{r,s}(M) \to \mathcal{T}^{r-1,s-1}(M)$ we have

$$\nabla(C(A)) = C(\nabla(A))$$

for all tensor fields $A \in \mathcal{T}^{r,s}(M)$.

The so-defined connections in each tensor bundle $T^{r,s}M \to M$ are uniquely determined by the above properties.

*Proof.* We proceed as follows. First we define a candidate for a connection in each tensor bundle, then we show that it fulfils all of the above properties, and finally prove uniqueness. In order to define a connection in $T^{r,s}M \to M$ it suffices to specify what it does on sections that can be, locally, written as pure tensor products of $r$ local vector fields and $s$ local 1-forms. For $T^{1,0}M \to M$, we simply take $\nabla$ to be our initial connection, which thereby automatically fulfils (i) and for $f \in \mathcal{T}^{0,0}(M) = C^\infty(M)$ we set $\nabla f = df$, thereby fulfilling (ii). Now we define $\nabla$ in $T^{0,1}M \to M$ in such a way, that (iii) and (iv) will be satisfied. Set for all 1-forms $\omega \in \Omega^1(M)$

$$(\nabla_X \omega)(Y) := X(\omega(Y)) - \omega(\nabla_X Y)$$

for all vector fields $X, Y \in \mathfrak{X}(M)$. After checking [Exercise!] that this defines a connection in $T^{0,1}M \to M$, we proceed as initially mentioned and obtain a connection in $T^{r,s}M \to M$ for all $r \geq 0$, $s \geq 0$ by requiring (iii) to hold on pure and, hence by linear extension, on all tensor fields. After checking that this really does define a connection [Again, exercise!] it remains to check that (iv) holds. This can be done inductively using (iii) after checking that it holds for the only possible contraction in $T^{1,1}M \to M$, which on pure tensor fields is of the form

$$C(X \otimes \omega) = \omega(X)$$

[38] Conventionally denoted by the same symbol $\nabla$.

for all $X \in \mathfrak{X}(M)$, $\omega \in \Omega^1(M)$, and analogously for local sections. We find for all $X, Y \in \mathfrak{X}(M)$ and all $\omega \in \Omega^1(M)$

$$\nabla_Y(C(X \otimes \omega)) = \nabla_Y(\omega(X)) = Y(\omega(X))$$

which by definition of $\nabla$ in $T^{0,1}M \to M$ and our imposed condition (iii) coincides with

$$Y(\omega(X)) = (\nabla_Y \omega)(X) + \omega(\nabla_Y X) = C(X \otimes (\nabla_Y \omega) + (\nabla_Y X) \otimes \omega) = C(\nabla_Y(X \otimes \omega)).$$

It remains to show uniqueness. Suppose there is an other connection $\tilde{\nabla}$ fulfilling all requirements of this lemma. By linearity in the second argument it suffices to show that $\nabla$ and $\tilde{\nabla}$ coincide on local pure tensor fields. By (i) and (iii) it further suffices to show that $\nabla$ and $\tilde{\nabla}$ coincide in $T^{0,1}M = T^*M \to M$. This follows from (i), (ii), and (iv) by direct calculation of the left- and right-hand of $\tilde{\nabla}(C(A)) = C(\tilde{\nabla}(A))$ for $A = X \otimes \omega$ where $X$ is any local vector field and $\omega$ is any local 1-form. □

Observe that Lemma 2.63 is, formally, very similar to Proposition 1.157 about the Lie derivative of tensor fields.

**Exercise 2.64.** Let $\nabla$ be a connection in $TM \to M$ and $\Gamma^k_{ij}$ its Christoffel symbols in local coordinates $(x^1, \ldots, x^n)$. Find a formula for the analogue of the Christoffel symbols of the induced connection $\nabla$ in $T^*M \to M$ with respect to the local frame given by the local coordinate 1-forms.

**Remark 2.65.** Differentiation of tensor fields with respect to a connection induced by a connection in the tangent bundle is sometimes called **covariant differentiation**. $\nabla_X A$ for $X \in \mathfrak{X}(M)$, $A \in \mathfrak{T}^{r,s}(M)$, is then called **covariant derivative of $A$ in direction $X$**. When talking about covariant derivatives make sure to always specify the corresponding connection.

ADD REMARK OF , ; IN INDEX NOTATION!

A central usage of connections is a preferred way to "transport", that is smoothly change, vectors along a curve in the base manifold. In order to properly introduce this concept, we need to study how to in a covariant manner differentiate vector fields, or more generally tensor fields, along curves. This is to be read in the way that we want to give meaning to expressions of the form

$$\nabla_{\gamma'} A$$

where $\gamma : I \to M$ is a smooth curve in a smooth manifold $M$ and $A$ is a tensor field that is **only parametrised along** $\gamma(I) \subset M$. Recall the definition of vector fields along curves, Definition 1.102, and keep it in mind for the differences when compared with what is to come next.

**Definition 2.66.** An $(r, s)$**-tensor field** $A = A_\gamma$, $r, s \geq 0$, on a smooth manifold $M$ **along a curve** $\gamma : I \to M$ is a smooth map

$$A_\gamma : I \to T^{r,s}M, \quad t \mapsto A_{\gamma(t)} \in T^{r,s}_{\gamma(t)}M.$$

If $\gamma$ is an embedding and thus $\gamma(I)$ is a submanifold of $M$, an $(r, s)$-tensor field $A_\gamma$ along $\gamma$ is simply a parametrisation[39] of a smooth section in $T^{r,s}M|_{\gamma(I)} \to \gamma(I)$.

Note that Definition 2.66 relaxes Definition 1.102 for vector fields, that is $(1, 0)$-tensor fields, as it allows the vector field to not be just the velocity vector field of the curve up to multiplication with a smooth map. In the following, "vector field along a curve" is to be understood in the sense of Definition 2.66.

A tensor field along a curve need not be the restriction of a tensor field of the ambient manifold to the image of the curve. This is the general case, as the curve is allowed to have self-intersections. We have however the following local result.

[39] i.e. $t \mapsto A_{\gamma(t)}$

**Lemma 2.67.** Let $\gamma : I \to M$ be a smooth curve and suppose that $\gamma'(t_0) \neq 0$. Let further $A_\gamma$ be an $(r,s)$-tensor field along $\gamma$. Then there exists an open interval $I' \subset I$ with $t_0 \in I'$, such that $A_\gamma|_{I'}$ is the restriction of an $(r,s)$-tensor field $\overline{A} \in \mathcal{T}^{r,s}(M)$ to $\gamma(I')$.

*Proof.* Following the proof of Proposition 1.122, we can assume without loss of generality that, after restriction to $I'$ and a suitable choice of local coordinates $\varphi = (x^1, \dots, x^n)$ on $U \subset M$ with $\gamma(I') \subset U$, $\gamma$ is of the form $t \mapsto \varphi^{-1}(t, 0, \dots, 0)$, so that $x^1(\gamma(t)) = t$ and $x^i(\gamma(t)) = 0$ for $2 \leq i \leq n$. Hence, if

$$A_{\gamma(t)} = \sum f^{i_1 \dots i_r}{}_{j_1 \dots j_s}(t) \frac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_s},$$

with all $f^{i_1 \dots i_r}{}_{j_1 \dots j_s} : I' \to \mathbb{R}$ smooth, the tensor field

$$\overline{A} = \sum \left( f^{i_1 \dots i_r}{}_{j_1 \dots j_s} \circ x^1 \right) \frac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_s}$$

fulfils the requirements of this lemma. □

A tensor field $\overline{A}$ as in Lemma 2.67 is called local extension of a given tensor field along a curve. This in particular means that we can, at least locally, extend vector fields along curves to vector fields of the ambient manifold.

Suppose that we are given a vector field $X$ along a curve $\gamma$ and a connection in the tangent bundle $TM \to M$ of the ambient manifold $M$. How can we, covariantly w.r.t. $\nabla$, differentiate $X$ in direction of $\gamma'$? Be aware that $X$ need not be the restriction of a vector field on $M$ to the image of $\gamma$, and in general in our sense the restriction of $TM \to M$ to the image of $\gamma$ need not be a vector bundle at all. The latter is in particular the case if the image of $\gamma$ is not a smooth submanifold of $M$. Thus what we are looking for is not a connection in the sense of Definition 2.55.

**Proposition 2.68.** Let $M$ be a smooth manifold and $\nabla$ a connection in $TM \to M$. Let $\gamma : I \to M$ be a smooth curve and denote the set of vector fields along $\gamma$ by $\Gamma_\gamma(TM)$. Then there exists a unique $\mathbb{R}$-linear map

$$\tfrac{\nabla}{dt} : \Gamma_\gamma(TM) \to \Gamma_\gamma(TM),$$

such that

$$\tfrac{\nabla}{dt}(fX) = \frac{\partial f}{\partial t} X + f \tfrac{\nabla}{dt} X$$

for all $f \in C^\infty(I)$ and all $X \in \Gamma_\gamma(TM)$ and, if $X = \overline{X}|_{\gamma(I)}$,

$$\tfrac{\nabla}{dt} X = \nabla_{\gamma'} \overline{X}$$

for all $t \in I$.

*Proof.* First suppose that a map $\frac{\nabla}{dt}$ fulfilling the requirements exists. We show that it is then unique. If $\gamma'(t_0) = 0$, we set $\frac{\nabla}{dt} X|_{t=t_0} = 0$ for all $X \in \Gamma_\gamma(TM)$. This is compatible with the tensoriality in the first argument of any connection, so that $\left( \frac{\nabla}{dt} X \right)\Big|_{t=t_0} = 0$ for all vector fields $X$ along $\gamma$ that are restrictions $X = \overline{X}_\gamma$ of vector fields $\overline{X} \in \mathfrak{X}(M)$. If $\gamma'(t_0) \neq 0$, we use the locality property of connections, cf. Lemma 2.59, and Lemma 2.67 with analogous coordinate choice $(x^1, \dots, x^n)$[40] to obtain the local formula

$$\left( \tfrac{\nabla}{dt} X \right)\Big|_{t=t_0} = \left( \nabla_{\gamma'} \overline{X} \right)\Big|_{t=t_0} = \sum_{k=1}^n \left( \frac{\partial X^k}{\partial t}(t_0) + \sum_{i,j=1}^n \frac{\partial \gamma^i}{\partial t}(t_0) X^j(t_0) \Gamma^k_{ij}(\gamma(t_0)) \right) \frac{\partial}{\partial x^k}\Big|_{t=t_0} \tag{2.8}$$

[40] And restricting the domain of definition of $\gamma$ if necessary so that its image is contained in the coordinate domain.

for all $X = \sum\limits_{k=0}^{n} X^k(t)\frac{\partial}{\partial x^k}$ and corresponding local extension $\overline{X} = \sum\limits_{k=1}^{n}\left(X^k \circ x^1\right)\frac{\partial}{\partial x^k}$. This shows that if the operator $\frac{\nabla}{dt}$ exists, it is uniquely determined by the connection $\nabla$. On the other hand observe that formula (2.8) defines by the locality property of connections an operator $\frac{\nabla}{dt}$ fulfilling the requirements of this proposition. To see this one must check the right hand side of (2.8) actually transforms as a connection and is thus independent of the chosen local extension $\overline{X}$ of $X$, which is a lengthy but not difficult calculation. □

**Definition 2.69.** The linear differential operator $\nabla_{\gamma'}$ is called **covariant derivative along** $\gamma$.

**Remark 2.70.** In local coordinates $(x^1, \ldots, x^n)$ on $U \subset M$, the covariant derivative along $\gamma : I \to M$[41] of $X \in \Gamma_\gamma(TM)$ locally of the form $X = \sum\limits_{k=1}^{n} X^k \frac{\partial}{\partial x^k}$, $X^k = X^k(t) \in C^\infty(I)$ for all $1 \leq k \leq n$, is given by the local formula

$$\nabla_{\gamma'} X = \sum_{k=1}^{n}\left(\frac{\partial X^k}{\partial t} + \sum_{i,j=1}^{n}\frac{\partial \gamma^i}{\partial t} X^j \Gamma_{ij}^k\right)\frac{\partial}{\partial x^k}. \tag{2.9}$$

Note that for the sake of readability we simply wrote $\Gamma_{ij}^k$ instead of $\Gamma_{ij}^k \circ \gamma$ in the above equation.

**Exercise 2.71.** Formulate and prove an analogous statement to Proposition 2.68 for any vector bundle $E \to M$, including in particular the tensor bundles $T^{r,s}M \to M$, where the sections in $E \to M$ along a curve $\gamma : I \to M$ are defined analogously to Definition 2.66.

Proposition 2.68 and Exercise 2.71 in particular imply compatibility of $\nabla_{\gamma'}$ with contractions.

**Corollary 2.72.** Let $A$ be an $(r,s)$-tensor field on a smooth manifold $M$ along $\gamma : I \to M$ with $r, s \geq 1$. Let $C : \mathcal{T}^{r,s}(M) \to \mathcal{T}^{r-1,s-1}(M)$ be any contraction and note that $C$ canonically extends to $(r,s)$-tensor fields along curves. Then $C(\nabla_{\gamma'} A) = \nabla_{\gamma'}(C(A))$.

**Remark 2.73.** We will not use the notation $\frac{\nabla}{dt}$ and simply write $\nabla_{\gamma'}$ instead, even if we plug in a vector field along $\gamma$ that is, globally, not the restriction of a vector field in the ambient manifold. The reason is firstly that the notation $\frac{\nabla}{dt}$ obfuscates the curve that we are working with, and secondly that locally if $\gamma'(t) \neq 0$, we can assume that any vector field along $\gamma$ is the restriction of a vector field in the ambient manifold, together with a parametrisation, see Lemma 2.67.

Proposition 2.68 allows us to answer the question posed in Remark 2.54 from the beginning of this section.

**Definition 2.74.** Let $X \in \Gamma_\gamma(TM)$ be a vector field along a smooth curve $\gamma : I \to M$ and let $\nabla$ be a connection in $TM \to M$. $X$ is called **parallel along** $\gamma$, or simply **parallel**, if $\nabla_{\gamma'} X = 0$.

Using Exercise 2.71 and Corollary 2.72, one can similarly define parallel sections of an arbitrary vector bundle $E \to M$ along curves. For $E = T^*M$ we obtain that a 1-form along a smooth curve $\gamma : I \to M$, $\omega \in \Gamma_\gamma(T^*M)$, is parallel along $\gamma$ with respect to a connection $\nabla$ in $TM \to M$ if and only if

$$\frac{\partial(\omega(X))}{\partial t} - \omega(\nabla_{\gamma'} X) = 0$$

for all $X \in \Gamma_\gamma(TM)$.

Next, suppose that we are given a connection in $TM \to M$, a smooth curve $\gamma : I \to M$, and a vector $v \in T_{\gamma(t_0)}M$, $t_0 \in I$. Does this data specify a preferred way to define a vector field $X$ along $\gamma$ with **initial value** $X_{\gamma(0)} = v$? This is, formulated using our newly gained knowledge, the question we asked at the beginning of this section. The answer is as follows.

[41] w.l.o.g. $\gamma(I) \subset U$

**Theorem 2.75.** Let $\nabla$ be a connection in $TM \to M$, $\gamma : I \to M$, $t_0 \in I$, a smooth curve with non-vanishing velocity, and $v \in T_{\gamma(t_0)}M$. Then there exists a unique vector field along $\gamma$, $X \in \Gamma_\gamma(TM)$, such that $X$ is parallel along $\gamma$ and $X_{\gamma(t_0)} = v$. This means that $X$ is the unique solution to the initial value problem

$$\nabla_{\gamma'} X = 0, \quad X_{\gamma(t_0)} = v. \tag{2.10}$$

*Proof.* It follows from Remark 2.70 that locally, (2.10) is an ordinary differential equation which thus has, locally, a unique solution [A1]. Here locally means restricted to a coordinate domain. So we need to deal with cases where $\gamma(I)$ is not covered by a single chart. This detail of the proof is left as an exercise to the reader. Alternatively see the proof of [L1, Thm. 4.11]. $\square$

**Exercise 2.76.** Formulate and prove Theorem 2.75 for any vector bundle, not just the tangent bundle.

**Example 2.77.** Let $\gamma : \mathbb{R} \to \mathbb{R}^2$ and $X = X_\gamma \in \Gamma_\gamma(T\mathbb{R}^2)$ as in Remark 2.54. Let $\nabla$ be the connection in $T\mathbb{R}^2 \to \mathbb{R}^2$ defined by setting its Christoffel symbols in canonical coordinates all equal to 0. Then $X$ is parallel along $\gamma$, i.e. $\nabla_{\gamma'} X = 0$, meaning that $X$ solves the initial value problem of parallelly transporting $v = (\gamma(0), \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right))$ along $\gamma$.

Suppose that we are given a connection in $TM \to M$ and somehow know how to, at least locally, solve every possible parallel transport equation with respect to any smooth curve and all initial values. Can we use this data to recover our connection? What does this tell us about the geometric interpretation of connections in general? To answer these questions, first observe the following property of parallel translations.

**Lemma 2.78.** Let $\nabla$ be a connection in $TM \to M$ and let $\gamma : I \to M$ be a smooth curve. Consider parallel translations along $\gamma$ as maps

$$P_{t_0}^t(\gamma) : T_{\gamma(t_0)}M \to T_{\gamma(t)}M,$$

mapping initial values $v \in T_{\gamma(t_0)}M$, $t_0 \in I$, of the differential equation $\nabla_{\gamma'} X = 0$, to the value of its unique solution $X$ at $t \in I$, namely $X_{\gamma(t)} \in T_{\gamma(t)}M$. Then $P_{t_0}^t(\gamma)$ is a linear isomorphism for all $t_0, t \in I$.

*Proof.* Linearity of $P_{t_0}^t(\gamma)$ follows by observing that whenever $X$ solves $\nabla_{\gamma'} X = 0$ for initial value $v \in T_{\gamma t_0}M$, $cX$ is also parallel along $\gamma$ for all $c \in \mathbb{R}$ and is the unique solution of the parallel transport equation for initial value $cv \in T_{\gamma(t_0)}M$. To see that $P_{t_0}^t(\gamma)$ is invertible, fix $t \in I$ and let $\widetilde{\gamma}(s) := \gamma(t - s)$. Then the parallel transport with respect to $\widetilde{\gamma}$ from $s = 0$ to $s = t - t_0$, $P_0^{t-t_0}(\widetilde{\gamma}) : T_{\gamma(t)}M \to T_{\gamma(t_0)}M$ is precisely the inverse of $P_{t_0}^t(\gamma) : T_{\gamma(t_0)}M \to T_{\gamma(t)}M$, which follows from $\nabla_{\widetilde{\gamma}'}(X \circ (t - s)) = 0$ for $X$ being the unique solution of $\nabla_{\gamma'} X = 0$ with fixed initial value in $T_{\gamma(t_0)}M$. $\square$

Using Lemma 2.78, we can describe a connection in $TM \to M$ completely using its parallel transport solutions.

**Proposition 2.79.** Let $\nabla$ be a connection in $TM \to M$ and $X, Y \in \mathfrak{X}(M)$. For $p \in M$ arbitrary let $\gamma : (-\varepsilon, \varepsilon) \to M$, $\varepsilon > 0$, be an integral curve of $X$ with $\gamma(0) = p$, and let $P_{t_0}^t$ denote the corresponding parallel transport maps. Then

$$(\nabla_X Y)_p = \left.\frac{\partial}{\partial t}\right|_{t=0} P_t^0(\gamma) Y_{\gamma(t)}.$$

*Proof.* First note that $t \mapsto P_t^0(\gamma)Y_{\gamma(t)}$ is smooth, which follows from the smoothness of the local prefactors of the defining differential equation in local coordinates by setting (2.9) to zero and, of course, replacing $\gamma$ and $X$ appropriately. Here the smooth manifold structure in $T_pM$ is given by its linear isomorphy to $\mathbb{R}^n$. Furthermore, $P_t^0(\gamma)Y_{\gamma(t)} \in T_pM$ for all $t \in (-\varepsilon, \varepsilon)$, so it makes sense to take its time derivative. Choose a basis $\{v_1, \ldots, v_n\}$ of $T_pM$, e.g. via local charts. For each $1 \leq i \leq n$, $V_i = V_i|_{\gamma(t)} := P_0^t(\gamma)v_i$ defines a parallel vector field along $\gamma$, i.e. $\nabla_{\gamma'}V_i = 0$. Hence, $\{V_1, \ldots, V_n\}$ is a parallel frame of $TM \to M$ along $\gamma|_{(-\varepsilon,\varepsilon)}$, meaning that each vector field along $\gamma$ that is the restriction of a vector field on the ambient manifold can be written as a $C^\infty((-\varepsilon, \varepsilon))$-linear combination of the $V_i$'s. Thus we can in particular write

$$Y_\gamma = \sum_{i=1}^n f^i V_i,$$

$f^i \in C^\infty((-\varepsilon, \varepsilon))$ for all $1 \leq i \leq n$. After recalling that $(\nabla_X Y)_p = \nabla_{\gamma'} Y_\gamma|_{t=0}$ (cf. (2.8)), we calculate

$$\nabla_{\gamma'} Y_\gamma|_{t=0} = \sum_{i=1}^n \left( \frac{\partial f^i}{\partial t} V_i + f^i \nabla_{\gamma'} V_i \right)\bigg|_{t=0} = \sum_{i=1}^n \frac{\partial f^i}{\partial t}(0) v_i.$$

On the other hand, we have for all $t \in (-\varepsilon, \varepsilon)$

$$P_t^0(\gamma)Y_{\gamma(t)} = P_t^0(\gamma) \left( \sum_{i=1}^n f^i(t) V_i|_{\gamma(t)} \right) = f^i(t) v_i, \tag{2.11}$$

where we used that $P_t^0(\gamma) = \left(P_0^t(\gamma)\right)^{-1}$ and that, by construction, $V_i$ is precisely the parallel extension of $v_i$ along $\gamma$ for all $1 \leq i \leq n$. Taking the $t$-derivative at $t = 0$ of the right hand side of (2.11) finishes the proof. $\square$

Together with the tensoriality in the first argument of $\nabla_X Y$, Proposition 2.79 implies the following, maybe at first sight surprising, fact.

**Corollary 2.80.** Let $\nabla$ be a connection in $TM \to M$ and $X, Y \in \mathfrak{X}(M)$. Then $(\nabla_X Y)_p$ depends only on $X_p$, any choice of smooth curve $\gamma : (-\varepsilon, \varepsilon) \to M$, $\varepsilon > 0$, with $\gamma'(0) = X_p$, and $Y_\gamma$, that is $Y$ along $\gamma$.

We will later use Corollary 2.80 in order to define induced connections on submanifolds. Note at this point that Corollary 2.80 implies that, after having chosen $\gamma$, we might change $Y$ outside of the image of $\gamma$ as we like and still will not change the value of $(\nabla_X Y)_p$.

Lastly before turning our focus on connections on pseudo-Riemannian manifolds we should ask ourselves how two, possibly different, connections in the tangent bundle of a smooth manifold are related. This will allow us to make (at least a limited) sense of the term "affine space of connections".

**Definition 2.81.** Let $M$ be a smooth manifold and let $\nabla^1, \nabla^2$ be connections in $TM \to M$. Then the **difference tensor** $A \in \mathcal{T}^{1,2}(M)$ of $\nabla^1$ and $\nabla^2$ is defined as

$$A(X, Y) := \nabla^1_X Y - \nabla^2_X Y$$

for all $X, Y \in \mathfrak{X}(M)$.

**Exercise 2.82.** Show that the difference tensor as in Definition 2.81 is, in fact, a tensor field.

The above definition and exercise can be interpreted as follows. In order to describe the space of all connections in the tangent bundle of a given smooth manifold $M$, one first chooses a reference connection $\nabla$ in $TM \to M$. Then every other connection $\widetilde{\nabla}$ in $TM \to M$ can be written as

$$\widetilde{\nabla}_X Y = \nabla_X Y + A(X,Y),$$

where $A \in \mathcal{T}^{1,2}(M)$ is a $(1,2)$-tensor field, namely the difference tensor of $\widetilde{\nabla}$ and $\nabla$. This means that we can view the space of all connections in $TM \to M$ as an **affine space**, given by $\nabla + \mathcal{T}^{1,2}(M)$, where $\nabla$ is the chosen reference point and $\mathcal{T}^{1,2}(M)$ the (infinite dimensional if $\dim(M) > 0$) real vector space.

**Remark 2.83.** We understand what a connection, in particular in $TM \to M$ and its tensor bundles, is. But we have not yet singled out a preferred one. The first step for doing so does not depend on any extra data on a given manifold $M$, but still has a nice geometric interpretation.

**Definition 2.84.** The **torsion tensor** $T \in \mathcal{T}^{1,2}(M)$ of a connection $\nabla$ in $TM \to M$ is given by

$$T(X,Y) := \nabla_X Y - \nabla_Y X - [X,Y]$$

for all $X, Y \in \mathfrak{X}(M)$. The connection $\nabla$ is called **torsion-free** if $T \equiv 0$.

**Exercise 2.85.**

(i) Show that the torsion tensor of a connection in $TM \to M$ is, in fact, a tensor field.

(ii) Show that a connection in $TM \to M$ is torsion-free if and only if all its Christoffel symbols in all local coordinates fulfil $\Gamma^k_{ij} = \Gamma^k_{ji}$.

(iii) Show that two connections in $TM \to M$ have the same torsion tensor if and only if their difference tensor is symmetric in its covector part.

Note that one consequence of Exercise 2.85 (iii) is that requiring torsion-freeness does not determine a connection in $TM \to M$ uniquely.

Next we should ask ourselves what torsion-freeness means geometrically. The following explanation is taken from [G, Bem. 2.6.2].

**Remark 2.86.** Consider for $n \geq 2$ the connection in $T\mathbb{R}^n \to \mathbb{R}^n$ given in Example 2.60, fix $p \in \mathbb{R}^n$, and choose two linearly independent vectors $v, w \in T_p\mathbb{R}^n$. Let further $\varepsilon > 0$ and

$$\gamma_v := t \mapsto p + tv, \quad \gamma_w := t \mapsto p + tw.$$

For any $t > 0$, the four vectors

$$v, w, P^1_0(\gamma_v)w, P^1_0(\gamma_w)v$$

can be interpreted as the **edges of a parallelogram**. What is the proper analogue for this picture for general smooth manifolds $M$ and connections in $TM \to M$? The answer lies in making $t > 0$ infinitesimally small and using Proposition 2.79. We fix $p \in M$ and local coordinates $\varphi = (x^1, \dots, x^n)$ on $U \subset M$, $p \in U$. For $1 \leq k \leq n$ and $\varepsilon > 0$ small enough, consider the smooth curves

$$\gamma_k : (-\varepsilon, \varepsilon) \to M, \quad x^\ell(\gamma_k(t)) = \delta^\ell_k t \quad \forall 1 \leq k, \ell \leq n,$$

so that $\gamma_k' = \frac{\partial}{\partial x^k}$. For any $i \neq j$, we obtain using Proposition 2.79

$$\begin{aligned} T\left(\tfrac{\partial}{\partial x^i}, \tfrac{\partial}{\partial x^j}\right)\Big|_p &= \left(\nabla_{\frac{\partial}{\partial x^i}} \tfrac{\partial}{\partial x^j} - \nabla_{\frac{\partial}{\partial x^j}} \tfrac{\partial}{\partial x^i}\right)\Big|_p \\ &= \frac{\partial}{\partial t}\Big|_{t=0} \left(P^0_t(\gamma_i)\tfrac{\partial}{\partial x^j} - P^0_t(\gamma_j)\tfrac{\partial}{\partial x^i}\right) \end{aligned}$$

$$= \lim_{\substack{t\to 0\\ t>0}} \frac{\frac{\partial}{\partial x^i}\Big|_p + P_t^0(\gamma_i)\frac{\partial}{\partial x^j} - \frac{\partial}{\partial x^j}\Big|_p - P_t^0(\gamma_j)\frac{\partial}{\partial x^i}}{t}.$$

Hence, the "infinitesimal" parallelograms spanned by any two different coordinate vectors and their parallel translations close, meaning that there is no "gap" when gluing the "infinitesimal" edges together. Note that since our chosen coordinates were arbitrary, it is easy to generalise this statement for any two linear independent vectors in $T_pM$, and not just for the coordinate vectors.

Recall that the infinitesimal symmetries of a pseudo-Riemannian manifold are given by Killing vector fields, cf. Definition 2.46. How can we connect this concept with connections in the tangent bundle of a pseudo-Riemannian manifold? The answer lies in parallel translations and requiring that for any fixed start- and end-points they must be linear isometries. It is, however, not immediately obvious how to implement this.

**Definition 2.87.** Let $(M, g)$ be a pseudo-Riemannian manifold. A connection $\nabla$ in $TM \to M$ is called **metric** if $\nabla g = 0$, that is

$$X(g(Y,Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \tag{2.12}$$

for all $X, Y, Z \in \mathfrak{X}(M)$.

What does it mean in a geometric sense that a connection is metric? The answer is, in fact, our initial ansatz for a sensible restriction of which connections to consider in the tangent bundle of a given pseudo-Riemannian manifold.

**Proposition 2.88.** A connection in $TM \to M$ on a pseudo-Riemannian manifold $(M, g)$ is metric if and only if its parallel transport maps $P_{t_0}^t(\gamma) : T_{\gamma(t_0)}M \to T_{\gamma(t)}M$ are linear isometries for all smooth curves $\gamma : I \to M$.

*Proof.* All possible[42] parallel transport maps $P_{t_0}^t(\gamma)$ are linear isometries if and only if for all such $P_{t_0}^t(\gamma)$ and all $v, w \in T_pM$, $p = \gamma(t_0)$, the map

$$t \mapsto g_{\gamma(t)}(P_{t_0}^t(\gamma)v, P_{t_0}^t(\gamma)w)$$

is constant. By considering affine reparametrisations of curves by $t \to t + c$ for constant $c \in \mathbb{R}$, it is easy to see that this holds if and only if

$$\frac{\partial}{\partial t}\Big|_{t=t_0} g_{\gamma(t)}(P_{t_0}^t(\gamma)v, P_{t_0}^t(\gamma)w) = 0 \tag{2.13}$$

for all parallel translations $P_{t_0}^t(\gamma)$. By considering $P_{t_0}^t(\gamma)v$ and $P_{t_0}^t(\gamma)w$ as vector fields along $\gamma$, it now follows from the tensor derivation property of any connection $\nabla$ that if $\nabla$ is metric, the left hand side of (2.13) must always vanish. If one has problems seeing that, one might want to formally replace $\frac{\partial}{\partial t}\big|_{t=0}$ by $\nabla_{\gamma'}|_{t=0}$.

For the other direction suppose that (2.13) holds for all parallel translations. Let $X, Y, Z \in \mathfrak{X}(M)$. One then, similarly to the proof of Proposition 2.79, fixes $p \in M$ and constructs a local parallel frame of $TM$ along a curve $\gamma$ fulfilling $\gamma'(0) = X_p$. We now write $Y_\gamma$ and $Z_\gamma$ in that parallel frame. The last step is, using these local forms, writing out

$$X_p(g(Y,Z)) = \nabla_{\gamma'}(g_\gamma(Y_\gamma, Z_\gamma))|_{t=0}$$

using the tensor derivation property of $\nabla$ and Proposition 2.79. By the imposed equality (2.13) it then follows that (2.12) must hold. Since $X, Y, Z$ and $p$ were arbitrary, it follows that $\nabla$ is indeed a metric connection. □

[42] As in corresponding to a smooth curve $\gamma : I \to M$.

**Remark 2.89.** How does a metric connection $\nabla$ on a pseudo-Riemannian manifold $(M, g)$ relate to the notion of Killing vector fields? Observe that for a given Killing vector field $X \in \mathfrak{X}(M)$ and any integral curve $\gamma : (-\varepsilon, \varepsilon) \to M$ of $X$, the parallel translation $P_0^t(\gamma)\gamma'(0)$ of $\gamma'(0)$ along $\gamma$ with respect to $\nabla$ is a linear isomorphism for all $t \in (-\varepsilon, \varepsilon)$. However, after restricting the local flow $\varphi : (-\varepsilon, \varepsilon) \times U \to M$ that contains $\gamma$ of $X$ if necessary, it does in general not hold that $P_0^t(\gamma)\gamma'(0) = d\varphi_t(\gamma'(0))$, since $P_0^t(\gamma)\gamma'(0)$ need not be tangent to $\gamma$ for $t \neq 0$. If however $P_0^t(\gamma)\gamma'(0)$ is tangent to $\gamma$ for all $t \in (-\varepsilon, \varepsilon)$, the equality must hold simply by the fact that both sides of the equation are linear isometries for all $t \in (-\varepsilon, \varepsilon)$. This train of thought leads to the definition of **geodesics**, cf. Definition 2.98, and in extension to the **geodesic flow**, cf. Definition 2.102.

**Definition 2.90.** Let $(M, g)$ be a pseudo-Riemannian manifold. A connection $\nabla$ in $TM \to M$ is called **Levi-Civita connection** if it is metric and torsion-free.

As always when defining some abstract object, we need to ask ourselves if it exists, and furthermore "how" unique it is. In the case of a Levi-Civita connection we obtain the following characterisation.

**Proposition 2.91.** Let $(M, g)$ be a pseudo-Riemannian manifold. Then there exists a unique Levi-Civita connection in $TM \to M$.

*Proof.* Follows from the following Proposition 2.92. □

In order to prove Proposition 2.91 we will use the following convenient result which can furthermore be used to find an explicit local formula for the Levi-Civita connections.

**Proposition 2.92.** Let $(M, g)$ be a pseudo-Riemannian manifold and $\nabla$ a connection in $TM \to M$. Then $\nabla$ is the Levi-Civita connection of $(M, g)$ if and only if it satisfies the **Koszul formula**

$$\begin{aligned} 2g(\nabla_X Y, Z) &= X(g(Y, Z)) + Y(g(X, Z)) - Z(g(X, Y)) \\ &\quad - g(X, [Y, Z]) - g(Y, [X, Z]) + g(Z, [X, Y]) \end{aligned} \tag{2.14}$$

for all $X, Y, Z \in \mathfrak{X}(M)$. Furthermore, the Koszul formula determines the connection uniquely.

*Proof.* First we will show that $\nabla$, defined by the Koszul formula, is in fact a connection. If we have shown this, it follows from the fibrewise nondegeneracy of $g$ that it is uniquely determined as $2g(\nabla_X Y, Z) - 2g(\widetilde{\nabla}_X Y, Z) = 0$ for all $X, Y, Z \in \mathfrak{X}(M)$ if $\widetilde{\nabla}$ is any other connection satisfying the Koszul formula. Instead of writing out the details, we will specify what to do and leave the actual calculations to the reader as this is a tedious, but rewarding exercise. Bilinearity in both arguments of $\nabla$ is easily checked. Next, we need to show that $\nabla_{fX} Y = f\nabla_X Y$ for all $f \in C^\infty(M)$. Since $g$ is fibrewise nondegenerate, this is equivalent to showing that $g(\nabla_{fX} Y, Z) = fg(\nabla_X Y, Z)$ for all $X, Y, Z \in \mathfrak{X}(M)$. Now we write out the right hand side of the Koszul formula and use the identity for the Lie derivative of vector fields[43] $\mathcal{L}_{fX} = f\mathcal{L}_X + X \otimes df$. Next we want to prove $\nabla_X(fY) = f\nabla_X Y + X(f)Y$ for all $X, Y \in \mathfrak{X}(M)$. We proceed as in the previous step and consider $g(\nabla_X(fY), Z)$ instead so we need to check that $g(\nabla_X(fY), Z) = fg(\nabla_X Y, Z) + X(f)g(Y, Z)$. Again, this amounts to using the afore mentioned identity of the Lie derivative of vector fields and the right hand side of the Koszul formula. At this point we have shown that $\nabla$ is, in fact, a connection in $TM \to M$. In order to show that $\nabla$ is torsion free we can use local coordinates $(x^1, \ldots, x^n)$ on $U \subset M$. Since the torsion tensor $T$ of $\nabla$ is a tensor field and $g$ is fibrewise nondegenerate, it suffices to show that

$$g\left(T\left(\tfrac{\partial}{\partial x^i}, \tfrac{\partial}{\partial x^j}\right), \tfrac{\partial}{\partial x^k}\right) = 0$$

[43] Careful when considering e.g. the Lie derivative of 1-forms!

for all $1 \leq i, j, k \leq n$. Since $T\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} - \nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^i}$, we can write

$$g\left(T\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right), \frac{\partial}{\partial x^k}\right) = g\left(\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right) - g\left(\nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^k}\right),$$

and replace each term using the right hand side of the Koszul formula. It turns out that it does, in fact, vanish for all $1 \leq i, j, k \leq n$ and, hence, that $\nabla$ is torsion free. Alternatively, one simply observes (and uses) that the terms 1+2, 3, and 4+5 in the right hand side of the Koszul formula are symmetric in $Y$ and $Z$. In order to show that $\nabla$ is metric, that is $\nabla g = 0$, we use that this property is equivalent to $Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$ for all $X, Y, Z \in \mathfrak{X}(M)$. Writing out the right hand side using the Koszul formula yields the result. Again, one might alternatively use that the terms 2+3, 4, and 5+6 in the right hand side of the Koszul formula are skew in $Y$ and $Z$. □

**Lemma 2.93.** The Christoffel symbols of the Levi-Civita connection on a pseudo-Riemannian manifold $(M, g)$ with respect to local coordinates $(x^1, \ldots, x^n)$ are given by

$$\Gamma_{ij}^k = \frac{1}{2} \sum_{\ell=1}^{n} \left(\frac{\partial g_{j\ell}}{\partial x^i} + \frac{\partial g_{i\ell}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^\ell}\right) g^{\ell k}$$

for all $1 \leq i, j, k \leq n$.

*Proof.* Follows by inserting the local coordinate vector fields into the Koszul formula. □

**Example 2.94.** The Christoffel symbols of the Levi-Civita connection of $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$, cf. Example 2.60, vanish identically in canonical coordinates. In polar coordinates $(r, \varphi)$ on $\mathbb{R}^2 \setminus \{(x, 0) \in \mathbb{R}^2 \mid x \leq 0\}$, the Christoffel symbols with respect to the standard Riemannian metric, given by $dx^2 + dy^2$ in canonical coordinates $(x, y)$, are of the form

$$\Gamma_{\varphi\varphi}^r = -r, \quad \Gamma_{r\varphi}^{\varphi} = \Gamma_{\varphi r}^{\varphi} = \frac{1}{r}, \quad 0 \text{ else.}$$

Recall the definition of the Hessian matrix of a smooth function on $\mathbb{R}^n$. Using the language of connections and traces with respect to pseudo-Riemannian metrics, we can now describe a coordinate-free way for these constructions.

**Definition 2.95.** Let $\nabla$ be a connection in $TM \to M$. The **covariant Hessian** of a smooth function $f \in C^\infty(M)$ is defined as the $(0, 2)$-tensor field

$$\nabla^2 f := \nabla(\nabla f) = \nabla df \in \mathcal{T}^{0,2}(M).$$

If $(M, g)$ is a pseudo-Riemannian manifold and $\nabla$ is the Levi-Civita connection, we can take the trace of the covariant Hessian with respect to $g$ and obtain the **Laplace-Beltrami operator** on smooth functions $f \in C^\infty(M)$ given by

$$\Delta f := \mathrm{tr}_g(\nabla^2 f).$$

Note that there are different definitions of the Laplace-Beltrami operator, e.g. one involving the divergence of vector fields, cf. for example [J].

**Exercise 2.96.** Show that the covariant Hessian of a connection $\nabla$ in $TM \to M$ is symmetric in the sense that $\nabla^2 f \in \Gamma(\mathrm{Sym}^2(T^*M))$ for all $f \in C^\infty(M)$ if and only if $\nabla$ is torsion-free.

## 2.3 Geodesics

The **acceleration** of a smooth curve $\gamma : I \to \mathbb{R}^n$ in canonical coordinates is defined as

$$\gamma'' : I \to T\mathbb{R}^n, \quad t \mapsto \left(\gamma(t), \tfrac{\partial^2 \gamma^1}{\partial t^2}, \dots, \tfrac{\partial^2 \gamma^n}{\partial t^2}\right),$$

which is a vector field along $\gamma$. How can we define $\gamma''$ in a coordinate free way, such that it gives the above formula on $\mathbb{R}^n$ for the Levi-Civita connection of the standard Riemannian metric $\sum\limits_{i=1}^{n} (du^i)^2$?

**Definition 2.97.** Let $M$ be a smooth manifold, $\nabla$ a connection in $TM \to M$, and $\gamma : I \to M$ a smooth curve. Then $\nabla_{\gamma'}\gamma' \in \Gamma_\gamma(TM)$ is called the **acceleration** of $\gamma$ (with respect to $\nabla$).

**Definition 2.98.** A smooth curve $\gamma : I \to M$ is called **geodesic** with respect to a given connection $\nabla$ in $TM \to M$ if its acceleration vanishes, that is $\nabla_{\gamma'}\gamma' = 0$.

The above definition of a geodesic might be formulated as follows: A curve is a geodesic if and only if its velocity vector field is parallel. It follows from (2.9) that in local coordinates $(x^1, \dots, x^n)$ of $M$, $\nabla_{\gamma'}\gamma'$ is of the form

$$\nabla_{\gamma'}\gamma' = \sum_{k=1}^{n} \left( \frac{\partial^2 \gamma^k}{\partial t^2} + \sum_{i,j=1}^{n} \frac{\partial \gamma^i}{\partial t} \frac{\partial \gamma^j}{\partial t} \Gamma^k_{ij} \right) \frac{\partial}{\partial x^k}, \tag{2.15}$$

where the terms $\Gamma^k_{ij} = \Gamma^k_{ij} \circ \gamma$ are the Christoffel symbols of the given connection $\nabla$ evaluated along $\gamma$, and $\gamma^k = x^k \circ \gamma$ as usual. Hence, $\gamma$ is a geodesic if and only if in all local coordinates covering a nonempty subset of the image of $\gamma$ it holds that

$$\frac{\partial^2 \gamma^k}{\partial t^2} + \sum_{i,j=1}^{n} \frac{\partial \gamma^i}{\partial t} \frac{\partial \gamma^j}{\partial t} \Gamma^k_{ij} = 0 \quad \forall 1 \leq k \leq n. \tag{2.16}$$

An alternative common notation for the above equation which uses the Einstein summation convention is

$$\ddot{x}^k + \dot{x}^i \dot{x}^j \Gamma^k_{ij} = 0 \tag{2.17}$$

for all $1 \leq k \leq n$. The above equation comes with a certain error potential as the terms $x^i$ in (2.17) denote the components of the curve and not the coordinate functions. Hence, caution is advised when using (2.17) instead of (2.16).

**Exercise 2.99.** Verify that for the Levi-Civita connection $\nabla$ of $\left(\mathbb{R}^n, \sum\limits_{i=1}^{n} (du^i)^2\right)$, the formula $\nabla_{\gamma'}\gamma' = \gamma''$, as explained in the motivation in the beginning of this section, holds true.

A nice property of geodesics with respect to metric connections is the following. It allows to quickly check if a suspected geodesic can be excluded without having to calculate all Christoffel symbols and its second derivatives.

**Lemma 2.100.** Let $\gamma : I \to M$ be a geodesic on a pseudo-Riemannian manifold $(M, g)$ with respect to a metric connection $\nabla$. Then $g(\gamma', \gamma') : I \to \mathbb{R}$ is constant.

*Proof.* We calculate

$$\frac{\partial(g(\gamma', \gamma'))}{\partial t} = \nabla_{\gamma'}(g(\gamma', \gamma')) = (\nabla_{\gamma'} g)(\gamma', \gamma') + 2g(\nabla_{\gamma'}\gamma', \gamma') = 0.$$

□

Note that Lemma 2.100 in particular holds for the Levi-Civita connection. In the Riemannian case, the **speed** of a curve $\gamma$ can be defined as $\sqrt{g(\gamma',\gamma')}$. The above lemma implies that geodesics in Riemannian manifolds have constant speed.

At this point we have neither proved existence of geodesics nor examined their uniqueness. Locally, we obtain the following result.

**Proposition 2.101.** Let $M$ be a smooth manifold and $\nabla$ a connection in $TM \to M$. Let further $p \in M$ and $v \in T_pM$. Then there exists $\varepsilon > 0$ and a smooth curve $\gamma : (-\varepsilon, \varepsilon) \to M$ with $\gamma(0) = p$ and $\gamma'(0) = v$, such that $\gamma$ is a geodesic. If $\gamma_1 : I_1 \to M$ and $\gamma_2 : I_2 \to M$ are geodesics on $M$ such that $I_1 \cap I_2 \neq \emptyset$ and for some point $t_0 \in I_1 \cap I_2$, $\gamma_1(t_0) = \gamma_2(t_0)$ and $\gamma_1'(t_0) = \gamma_2'(t_0)$, then $\gamma_1|_{I_1\cap I_2} = \gamma_2|_{I_1\cap I_2}$.

*Proof.* It suffices to prove this proposition in local coordinates. The differential equation for a geodesic in local coordinates (2.16) is a nonlinear system of second order ordinary differential equations. In order to turn this system of $n$ second order ODEs into a first order system of ODEs, we increase the number of time-dependent variables to $2n$. Locally this means instead of trying to solve (2.16) we choose local coordinates $\varphi = (x^1, \dots, x^n)$ on $U \subset M$, $p \in U$, use, for the sake of readability, the alternative notation (2.17), and consider the system of equations

$$\begin{aligned} \dot{x}^k &= v^k, \\ \dot{v}^i &= -\dot{x}^i\dot{x}^j\Gamma^k_{ij}, \end{aligned} \tag{2.18}$$

for $1 \leq k \leq n$ with fitting initial values. The first thing we need to clarify are the symbols $v^k$. These are precisely the induced coordinates on $TU \subset TM$, so that $v^k(V) = V(x^k)$ for all $V \in T_qM$ with $q \in U$. In the above equation (2.18), the $x^k$ and $v^k$ are, however, to be read as components of a curve $(x = x(t), v = v(t)) : I \to d\varphi(TU)$, which is the price we have to pay for improved readability. In these local coordinates, (2.18) can be viewed as the local version of an integral curve equation of a vector field on $TU$, $G \in \mathfrak{X}(TU)$, that is a smooth section in $TTU \to TU$. To see this first observe that since the $x^k$ and $v^k$ are coordinate functions on $TU$, they induce coordinates on $TTU$. The corresponding local frame in $TTU \to TU$ is given by

$$\left\{\tfrac{\partial}{\partial x^1}, \dots, \tfrac{\partial}{\partial x^n}, \tfrac{\partial}{\partial v^1}, \dots, \tfrac{\partial}{\partial v^n}\right\}.$$

One can imagine each $\frac{\partial}{\partial x^k}$ as being "horizontal" and each $\frac{\partial}{\partial v^k}$ as being "vertical".[44] Using the Einstein summation convention, $G$ is given by insert drawing

$$G = v^k\tfrac{\partial}{\partial x^k} - v^iv^j\Gamma^k_{ij}\tfrac{\partial}{\partial v^k}. \tag{2.19}$$

Note that by covering $M$ with charts, $G$ uniquely extends to a vector field on $TM$, that is $G \in \mathfrak{X}(TM)$, via the induced charts on $TM$. Since any integral curve of $G$, locally given by $(x,v) : I \to d\varphi(TU)$, $t \mapsto (x(t), v(t))$, in particular fulfils $\dot{x} = v$, it is precisely the velocity vector field of the curve $x : I \to \varphi(U)$, $t \mapsto x(t)$. This means that the projection of any integral curve of $G$ to $M$ via the bundle projection $\pi : TM \to M$ is a geodesic. By Remark 1.108 or alternatively e.g. [A1] the statement of this proposition follows. $\square$

In fact, since we were going to cite [A1] anyway, we could have simply "proved" the above proposition without constructing the vector field $G$ on the tangent bundle. It is however an important ingredient in the construction of normal coordinates which we will introduce later, cf. Definition 2.126. After getting over the initial shock of double tangent bundles, it also gives a nice geometric way to in a sense obtain an infinitesimal generator of all geodesics with respect to a given connection, at once.

[44] Careful: This is only to gain a visual intuition, there are actual related definitions of "horizontal" and "vertical" in this setting.

**Definition 2.102.** The (local) flow of the vector field $G \in \mathfrak{X}(TM)$ which is in local coordinates given by (2.19) is called **geodesic flow** with respect to $\nabla$.

Local uniqueness of geodesics allows us to define a maximality property for geodesics.

**Definition 2.103.** A geodesic $\gamma : I \to M$ is called **maximal** if there exists no strictly larger interval $\widetilde{I} \supset I$ and a geodesic $\widetilde{\gamma} : \widetilde{I} \to M$, such that $\widetilde{\gamma}|_I = \gamma$. This means that $\gamma$ cannot be extended to a larger domain while still keeping its geodesic property. A smooth manifold with connection $\nabla$ in $TM \to M$ is called **geodesically complete** if every maximal geodesic is defined on $I = \mathbb{R}$. A pseudo-Riemannian manifold $(M, g)$, respectively the metric $g$, is called **geodesically complete** if its Levi-Civita connection is complete.

Assume that we are given a geodesic $\gamma : I \to M$ on $M$ with respect to a connection $\nabla$. How can we reparametrise $\gamma$ while at the same time preserving its geodesic property?

**Lemma 2.104.** Let $\gamma : I \to M$ be a geodesic with non-vanishing speed with respect to $\nabla$ and $f : I \to I'$ a diffeomorphism. Then $\gamma \circ f$ is a geodesic with non-vanishing speed if and only if $f$ is affine-linear, that is of the form $f(t) = at + b$ for $a \in \mathbb{R} \setminus \{0\}$, $b \in \mathbb{R}$.

*Proof.* We find using the local formula (2.15) and the chain rule

$$\nabla_{(\gamma \circ f)'}(\gamma \circ f)' = f'' \cdot \gamma' \circ f + (f')^2 \cdot (\nabla_{\gamma'}\gamma') \circ f = f'' \cdot \gamma' \circ f,$$

where the last equality comes from the assumption that $\gamma$ is a geodesic. Hence, $\gamma \circ f$ is a geodesic if and only if $f'' = 0$, that is if $f = at + b$ with $a \in \mathbb{R} \setminus \{0\}$, $b \in \mathbb{R}$. $\square$

**Corollary 2.105.** Maximal geodesics are unique up to affine reparametrisation.

**Exercise 2.106.** Let $\gamma : I \to M$ be a geodesic with initial value $\gamma(0) = p$, $\gamma'(0) = v \in T_pM$. Show that for all $a \in \mathbb{R}$, $t \mapsto \gamma_a(t) := \gamma(at)$ is a geodesic with initial value $\gamma_a(0) = p$, $\gamma_a'(0) = av$, defined on a fitting interval $I'$.

**Corollary 2.107.** A geodesic in a pseudo-Riemannian manifold with respect to the Levi-Civita connection with nonvanishing velocity can always be parametrised to be of unit speed, that is either $g(\gamma', \gamma') \equiv 1$ or $g(\gamma', \gamma') \equiv -1$.

Before coming to a geometrically more intuitive way why geodesics should be studied, we will take a look at some examples of geodesics.

**Example 2.108.**

(i) Each maximal geodesics of $\mathbb{R}^n$ equipped with the canonical connection in Example 2.60 with initial condition $\gamma(0) = p \in \mathbb{R}^n$, $\gamma'(0) = v \in T_p\mathbb{R}^n \cong \mathbb{R}^n$, is of the form

$$\gamma : \mathbb{R} \to \mathbb{R}^n, \quad t \mapsto p + tv.$$

This in particular means that the canonical connection on $\mathbb{R}^n$ is geodesically complete.

(ii) Consider $S^n \subset \mathbb{R}^{n+1}$ with induced metric $g = \langle \cdot, \cdot \rangle|_{TS^n \times TS^n}$, where $\langle \cdot, \cdot \rangle$ denotes the standard Riemannian metric on $\mathbb{R}^n$. The maximal geodesics of $(S^n, g)$ with respect to the Levi-Civita connection are great circles, that is

$$\gamma : \mathbb{R} \to S^n, \quad t \mapsto e^{At}p$$

for $\gamma(0) = p \in S^n$, $\gamma'(0) = Ap$, $A \in \mathrm{Mat}(n \times n)$ skew. This, again, means that the Levi-Civita connection of $(S^n, g)$ is geodesically complete.

**Exercise 2.109.** Prove the claims in Example 2.108.

The geodesic flow of a given connection allows us to locally identify small enough open neighbourhoods of points $p \in M$ with open neighbourhoods in the corresponding tangent space $T_pM$.

**Definition 2.110.** Let $M$ be a smooth manifold with connection $\nabla$ in $TM \to M$. An **open neighbourhood of the zero section** in $TM \to M$ is an open set $V \subset TM$ such that for all $p \in M$, $V_p := T_pM \cap V$ is an open neighbourhood of the origin $0 \in T_pM$. Note that the smooth manifold structure and topology on $T_pM$ are induced by the local trivialisations of $TM \to M$ and the corresponding fibrewise isomorphisms $T_pM \cong \mathbb{R}^n$.

We want to use open neighbourhoods of the zero section for our upcoming construction of the exponential map in Definition 2.112. To do so, we need the following additional result.

**Lemma 2.111.** Let $\nabla$ be a connection in $TM \to M$. Then there exists an open neighbourhood of the zero section $V \subset TM$, such that for all $v \in V_p \subset V$, the maximal geodesic $\gamma_v$ with initial condition $\gamma_v(0) = p$, $\gamma_v'(0) = v$, has domain containing the compact interval $[0, 1]$.

*Proof.* Exercise 2.106 implies that if $\gamma_v$ is defined on at least $[0, 1]$, then $\gamma_{rv}$ for $r \in [0, 1]$ is also defined on at least $[0, 1]$. We have seen in the proof of Proposition 2.101 that geodesics can be viewed as projections of integral curves of a vector field on $TM$. Thus, by identifying[45] $M$ with the image of the zero section in $TM$ and using Exercise 2.106, in order to prove this proposition it in fact suffices to show that for all $p \in M \subset TM$ we can find $\varepsilon_p > 0$ and an open neighbourhood[46] $W_p$ of $p$ in $TM$, such that all integral curves of $G$ (2.19) starting in $W_p$ are defined on at least $[0, \varepsilon_p]$. This follows from the fact that $G$ is a smooth vector field. If $\varepsilon_p < 1$, we rescale $W_p$ fibrewise with scaling factor $\varepsilon_p$, so that we can assume without loss of generality that all integral curves of $G$ starting in $W_p$ are defined on at least $[0, 1]$. Doing this procedure for all $p \in M \subset TM$, we obtain our desired open neighbourhood $V \subset TM$ of the zero section in $TM \to M$ by setting

$$V := \bigcup_{p \in M} W_p.$$

$\square$

**Definition 2.112.** Let $V \subset TM$ be an open neighbourhood of the zero section in $TM \to M$ such that for all $v \in V$, the unique maximal geodesic $\gamma_v$ with respect to $\nabla$ with initial condition $\gamma_v(0) = p$, $\gamma_v'(0) = v$, is defined on $[0, 1]$. The **exponential map** with respect to $\nabla$ is defined as

$$\exp : V \to M, \quad v \mapsto \gamma_v(1).$$

The **exponential map at** $p \in M$ $\exp_p : V_p \to M$ is the restriction of exp to $V_p = V \cap T_pM$.

The exponential map of a given connection can be used to construct useful local coordinates, in particular when the connection is the Levi-Civita connection with respect to a pseudo-Riemannian metric. This will be a later step, see Definition 2.126.

**Proposition 2.113.** Let $M$ be a smooth manifold and $\nabla$ a connection in $TM \to M$. For all $p \in M$, the exponential map at $p$ is a local diffeomorphism near $0 \in T_pM$.

[45] If you have problems with this step try drawing a sketch first.

[46] Note: $W_p$ is **not** a subset of the fibre $T_pM$!

*Proof.* We will show $d\exp_p = \mathrm{id}_{T_pM}$, which together with Theorem 1.55 will complete the proof. As in Definition 2.112 let $\gamma_v$ denote the maximal geodesic with chosen initial value $\gamma_v(0) = p$, $\gamma_v'(0) = v$ for $v \in T_pM$. We use Exercise 2.106 and obtain

$$d\exp_p(v) = \frac{\partial}{\partial t}\Big|_{t=0} \exp_p(tv) = \frac{\partial}{\partial t}\Big|_{t=0} \gamma_{tv}(1) = \frac{\partial}{\partial t}\Big|_{t=0} \gamma_v(t) = \gamma_v'(0) = v.$$

Since $v \in T_pM$ was arbitrary the claim follows. Note that, strictly speaking, we identified $T_0T_pM$ with $T_pM$ for the domain of $d\exp_p$ via the canonical isomorphism $(0, v) = v$. $\square$

Note that exp is defined on $TM$ if $\nabla$ is geodesically complete. This however does not mean that in this case there exists $p \in M$, such that $\exp_p$ is a diffeomorphism.

**Exercise 2.114.**

(i) Show that for any $p \in \mathbb{R}^n$, $\exp_p$ defined on $T_p\mathbb{R}^n$ with respect to the canonical connection in Example 2.60 is a diffeomorphism.

(ii) Show that if $M$ is compact and $\nabla$ is any connection in $TM \to M$, $\exp_p$ is never a diffeomorphism for all $p \in M$, independent of its domain $V_p \subset T_pM$.

Instead of trying to solve the geodesic equations, we can also study the weaker requirement that the velocity vector field of a curve and its acceleration are linearly dependent. This leads to the following definition.

**Definition 2.115.** A smooth curve $\gamma : I \to M$ is called **pregeodesic** with respect to a connection in $TM \to M$ if it has a reparametrisation as a geodesic, that is if there exists a diffeomorphism $f : I' \to I$, such that $\gamma \circ f$ is a geodesic.

Pregeodesics fulfil an equation similar to the geodesic equation.

**Lemma 2.116.** Any given pregeodesic $\gamma : I \to M$ with respect to a connection $\nabla$ in $TM \to M$ fulfils $\nabla_{\gamma'}\gamma' = F\gamma'$ for some smooth function $F : I \to \mathbb{R}$.

*Proof.* Let $f : I' \to I$ be a diffeomorphism such that $\gamma \circ f$ is a geodesic. Then, by definition, $\nabla_{(\gamma\circ f)'}(\gamma \circ f)'$=0. Writing out the left hand side of the equation with the help of the chain rule and (2.15) yields the desired result. More specifically one obtains $F = -\frac{f''}{(f')^2} \circ f^{-1}$. $\square$

While one might guess that setting the (covariant) acceleration of a curve to zero yields something interesting, we do not yet have explained a **geometric reason** why one would study geodesics. In the following we will see that geodesics are, in fact, critical points of the energy functional. In the Riemannian case, geodesics are furthermore (local!) minimizers of the length, viewed as a functional.

**Definition 2.117.** Let $(M, g)$ be a pseudo-Riemannian manifold and let $\gamma : [a, b] \to M$ be a smooth curve. The **energy functional** evaluated at $\gamma$, or simply **energy of** $\gamma$, is given by

$$E(\gamma) := \frac{1}{2}\int_a^b g(\gamma', \gamma')dt.$$

We see that in the Riemannian case the definition of the energy $E(\gamma)$ of a curve is similar to the definition of its length $L(\gamma)$, cf. Definition 2.9.

**Remark 2.118.** Note that the energy and length functionals can also be defined for **piecewise smooth curves**, that is continuous curves $\gamma : I \to M$ defined on a compact interval $I = [a, b]$, such that there exists a finite subdivision $I_i = [a_i, a_{i+1}]$, $1 \le i \le m$, $a_1 = a$, $a_{m+1} = b$, of $I$ and smooth curves $\gamma_i : I_i \to M$ for each $1 \le i \le m$, such that $\gamma|_{I_i} = \gamma_i$ for all $1 \le i \le m$. This approach is taken in [L1] (for Riemannian manifolds) and [O]. It involves some technical subtleties. We will only work with smooth curves. Admittedly, this is obfuscating parts of the larger picture, however it is from a geometric intuition point of view almost equivalent. If one understands our slightly restricted approach well, generalizing it will be a rather easy task. The interested reader is highly encouraged to take a look at the topic as presented in [O, Ch. 10]. We will proceed similarly to a mix of [Bae, Ch. 2.6] and [L1]

While we are talking about (nonlinear) functionals, we did not specify any further structure on their domains, e.g. the structure of a Banach manifold. In fact, we will not be concerned with questions like an optimal domain for $E$ or $L$ as for our purposes we only need to be concerned with what type of perturbations for fixed $\gamma$ we allow.

**Definition 2.119.** Let $\gamma : [a, b] \to M$ be a smooth curve and $\varepsilon > 0$. A smooth family of curves[47] $\eta : (-\varepsilon, \varepsilon) \times [a, b] \to M$ is called **variation of** $\gamma$ if $\eta(0, t) = \gamma(t)$ for all $t \in [a, b]$. $\eta$ is called **variation with fixed endpoints of** $\gamma$ if $\eta(s, a) = \gamma(a)$ and $\eta(s, b) = \gamma(b)$ for all $s \in (-\varepsilon, \varepsilon)$. The vector field $V$ along $\gamma$, $V_{\gamma(t)} = \frac{\partial \eta}{\partial s}(0, t) \in T_{\gamma(t)}M$, is called **variational vector field** of $\eta$.

Note that for any variation with fixed endpoints we have $V_{\gamma(a)} = V_{\gamma(b)} = 0$. Can we obtain every vector field along a given smooth curve $\gamma : [a, b] \to M$ as a variational vector field? The answer is positive.

**Lemma 2.120.** Let $\gamma : [a, b] \to M$ be a smooth curve. Let further $V \in \Gamma_\gamma(TM)$. Then there exists a variation $\eta$ of $\gamma$, such that $V$ is the variational vector field of $\eta$. If $V_{\gamma(a)} = V_{\gamma(b)} = 0$, $\eta$ can be chosen to be a variation with fixed endpoints.

*Proof.* Fix a Riemannian metric $g$ on $M$ with Levi-Civita connection $\nabla$. Let $\exp : V \to M$ denote the corresponding exponential map. We now define a variation of $\gamma$ via

$$\eta : (-\varepsilon, \varepsilon) \times [a, b] \to M, \quad \eta(s, t) := \exp(sV_{\gamma(t)}),$$

for $\varepsilon > 0$ small enough. Note that we can always find such an $\varepsilon$ by the compactness of $[a, b]$ and the smoothness of $V$. If $V$ vanishes at $\gamma(a)$ and $\gamma(b)$, $\eta$ has the property $\eta(s, a) = \gamma(a)$ and $\eta(s, b) = \gamma(b)$ for all $s \in (-\varepsilon, \varepsilon)$. Furthermore we check with a calculation as the one in the proof of Proposition 2.113

$$\frac{\partial \eta}{\partial s}(0, t) = V_{\gamma(t)}$$

for all $t \in [a, b]$. Hence, $\eta$ fulfils the required properties of this lemma. □

Next we will determine the so-called first variation of the energy. The obtained formula will show that geodesics with respect to the Levi-Civita connection of a pseudo-Riemannian manifold are indeed critical points of the energy functional.

**Lemma 2.121.** Let $(M, g)$ be a pseudo-Riemannian manifold with Levi-Civita connection $\nabla$. Then the first **variation of the energy** at a smooth curve $\gamma : [a, b] \to M$ with respect to a given variational vector field $V \in \Gamma_\gamma(TM)$ with a choice of corresponding variation of $\gamma$, $\eta : (-\varepsilon, \varepsilon) \times [a, b] \to M$, $\eta : (s, t) \mapsto \eta(s, t)$, is given by $\frac{\partial}{\partial s}\Big|_{s=0} E(\eta(s, \cdot))$ and fulfils

$$\frac{\partial}{\partial s}\Big|_{s=0} E(\eta(s, \cdot)) = -\int_a^b g(V, \nabla_{\gamma'}\gamma')dt + g(V_{\gamma(b)}, \gamma'(b)) - g(V_{\gamma(a)}, \gamma'(a)).$$

[47] Read: $\eta$ is in particular a smooth map, "family of curves" is the geometric picture one should have in mind.

In the special case that $V$ vanishes at the start- and end-point of $\gamma$, we have have $\frac{\partial}{\partial s}\Big|_{s=0} E(\eta(s,\cdot)) = -\int\limits_a^b g(V, \nabla_{\gamma'}\gamma')dt$.

*Proof.* In this proof we will use the Einstein summation convention. Let $\eta' = \eta'(s,t) = \frac{\partial \eta}{\partial t}$ denote the velocity vector field of the family of smooth curves $\eta$ for $s$ fixed. Also recall Exercise 2.71. We calculate

$$\frac{\partial}{\partial s}\Big|_{s=0} E(\eta(s,\cdot)) = \frac{1}{2}\int\limits_a^b \frac{\partial}{\partial s}\Big|_{s=0} g(\eta', \eta')dt = \frac{1}{2}\int\limits_a^b \nabla_V(g(\eta',\eta'))dt = \int\limits_a^b g(\gamma', \nabla_V \eta')dt.$$

For the last equality we have used that $\nabla$ is metric. For the next step we need to prove that $\nabla_V \eta' = \nabla_{\gamma'} V$. To do so we will use local coordinates. Fix $p \in \gamma([a,b])$ and choose local coordinates $(x^1, \ldots, x^n)$ on an open neighbourhood of $p \in M$. Denote $V^s = V^s_{\eta(s,t)} := \frac{\partial \eta^k}{\partial s}\frac{\partial}{\partial x^k}$, so that $V^0 = V$. It now suffices to show that $(\nabla_{V^s}\eta')|_{s=0} = \nabla_{\gamma'} V$. Using formula (2.9) (while not confusing $s$ and $t$), we find

$$\nabla_{V^s}\eta' = \left(\frac{\partial^2 \eta^k}{\partial s \partial t} + \frac{\partial \eta^i}{\partial s}\frac{\partial \eta^j}{\partial t}\Gamma^k_{ij}\right)\frac{\partial}{\partial x^k}$$

and

$$\nabla_{\gamma'} V = \left(\frac{\partial^2 \eta^k}{\partial t \partial s} + \frac{\partial \eta^i}{\partial t}\frac{\partial \eta^j}{\partial s}\Gamma^k_{ij}\right)\Bigg|_{s=0}\frac{\partial}{\partial x^k}.$$

The connection is torsion-free, which as we have seen in Exercise 2.85 is equivalent to the Christoffel symbols being symmetric in the lower indices. Hence, the above local formulas for $(\nabla_{V^s}\eta')|_{s=0}$ and $\nabla_{\gamma'}V$ indeed coincide. Since $p \in \gamma([a,b])$ was arbitrary we deduce that the equality holds for all $t \in [a,b]$. Thus we obtain using partial integration

$$\begin{aligned}
\int\limits_a^b g(\gamma', \nabla_V \eta')dt &= \int\limits_a^b g(\gamma', \nabla_{\gamma'} V)dt \\
&= \int\limits_a^b \left(\frac{\partial}{\partial t} g(\gamma', V) - g(\nabla_{\gamma'}\gamma', V)\right)dt \\
&= g(V_{\gamma(b)}, \gamma'(b)) - g(V_{\gamma(a)}, \gamma'(a)) - \int\limits_a^b g(V, \nabla_{\gamma'}\gamma')dt.
\end{aligned}$$

Reordering the above equation finishes the proof. □

**Corollary 2.122.** Geodesics defined on a compact interval with respect to the Levi-Civita connection of a pseudo-Riemannian are **critical points** of the energy functional in the sense that the first variation of the energy with respect to all variations with fixed end points vanishes.

One can also prove the converse statement.

**Lemma 2.123.** A curve in a pseudo-Riemannian manifold defined on a compact interval is a geodesic with respect to the Levi-Civita connection if it is a critical point of the energy functional in the sense of Corollary 2.122.

*Proof.* Exercise. [Hint: Use bump functions.] □

**Exercise 2.124.** Find a formula for the first variation of the length of a curve in a Riemannian manifold. Are geodesics also critical points of the length functional in our sense?

**Remark 2.125.** In Riemannian geometry, one can show that geodesics with respect to the Levi-Civita connection and with compact domain are not just critical points of the energy and length functional, but also local minimisers. This means that for every variation with fixed endpoints $\eta : (-\varepsilon, \varepsilon) \times [a, b] \to M$ of a geodesic $\gamma : [a, b] \to M$ in $(M, g)$, $E(\eta(s, \cdot)) \geq E(\gamma)$ for $\varepsilon$ small enough. We will not prove this here as it would conflict with the time constrains of this course, for references see [L1, Ch. 6] and [Bae], or [G] (in German).

We return to the exponential map. We have seen in Proposition 2.113 that the exponential map is a local diffeomorphism near every point of a given manifold with connection in its tangent bundle. Hence, we can use the exponential map to define local coordinates near every given point. In the case of pseudo-Riemannian manifolds equipped with their respective Levi-Civita connection, these kind of coordinates are of particular interest.

**Definition 2.126.** Let $M$ be a smooth manifold with connection $\nabla$ in its tangent bundle. Suppose that $V \subset T_pM$ is a star-shaped open neighbourhood of the origin, such that $\exp_p : V \to \exp_p(V)$ is a diffeomorphism. Then $U = \exp_p(V)$ is an open neighbourhood of $p \in M$ and is called **normal neighbourhood** of $p \in M$. Let $U \subset M$ be such a normal neighbourhood of a point $p \in M$. Then the exponential map at $p$ can be used to define local coordinates $(x^1, \ldots, x^n)$ near $p$ as follows. Choose a basis $\{v_1, \ldots, v_n\}$ of $T_pM$ and define coordinates implicitly via

$$\exp_p\left(\sum_{i=1}^{n} x^i(q) v_i\right) = q$$

for all $q \in U$. This just means that the $x^i$ are the prefactor functions of $\exp_p^{-1}$ written in the basis $\{v_1, \ldots, v_n\}$. Smoothness of the $x^i$ follows from the implicit function theorem. If $(M, g)$ is a pseudo-Riemannian manifold with Levi-Civita connection $\nabla$, normal coordinates at $p \in M$ with respect to an orthonormal basis $\{v_1, \ldots, v_n\}$ of $T_pM$ are called **Riemannian normal coordinates**[48] at $p \in M$. If $(M, g)$ is Riemannian and $V = B_r(0) = \{v \in T_pM \mid g_p(v, v) < r\}$ for some $r > 0$, the corresponding domain of the Riemannian normal coordinates $B^g_r(p) := \exp_p(B_\varepsilon(0))$ is called **geodesic ball of radius** $r$ **centred at** $p$ in $M$. The upper index $g$ indicates the corresponding Riemannian metric.

**Exercise 2.127.** Let $(M, g)$ be a connected pseudo-Riemannian manifold. Then any two points of $M$ can be connected by a piecewise smooth curve (cf. Remark 2.118), such that every smooth segment of that curve is a geodesic.

The converse of the statement in the above exercise holds of course too, see (after trying to solve the exercise yourself) [O, Ch. 3, Lem. 32].

Riemannian normal coordinates have the property that near their reference point, several geometric data has a particularly simple form.

**Proposition 2.128.** Let $(M, g)$ be a pseudo-Riemannian manifold with Levi-Civita connection $\nabla$ and let $\varphi = (x^1, \ldots, x^n)$ be Riemannian normal coordinates near $p \in M$ corresponding to a choice of orthonormal basis $\{v_1, \ldots, v_n\}$ of $T_pM$. Then $g$, written in its local form (2.1) with local smooth functions $g_{ij}$ as in (2.2), fulfils

$$g_{ij}(p) = \varepsilon_{ij}$$

[48] Yes, also if $M$ is **not** Riemannian but pseudo-Riemannian.

for all $1 \leq i, j \leq n$, where $\varepsilon_{ij} = g(v_i, v_j)$. The Christoffel symbols of $\nabla$ and all partial derivatives of the local smooth functions $g_{ij}$ vanish at $p$, that is

$$\Gamma^k_{ij}(p) = 0, \quad \frac{\partial g_{ij}}{\partial x^k}(p) = 0,$$

for all $1 \leq i, j, k \leq n$. If $\gamma : (-\varepsilon, \varepsilon) \to M$, $\gamma(0) = p$, $\gamma'(0) = w \in T_pM$, is a geodesic starting at $p \in M$ such that its image is contained in the domain of $\varphi$, $\varphi \circ \gamma$ is of the form

$$\varphi(\gamma(t)) = tw$$

for all $t \in (-\varepsilon, \varepsilon)$.

*Proof.* For $g_{ij}(p) = \varepsilon_{ij}$ we show that $v_k = \frac{\partial}{\partial x^k}\Big|_p$ for all $1 \leq k \leq n$. Since $x^k(p) = 0$ for all $1 \leq k \leq n$ by construction, we obtain (after, as before, identifying $T_0T_pM \cong T_pM$)

$$\sum_{k=1}^{n} d\exp_p\Big|_0 (v_k) \otimes dx^k|_p = \mathrm{id}_{T_pM}. \tag{2.20}$$

On the other hand we know by Proposition 2.113 that $d\exp_p|_0(v_k) = v_k$ for all $1 \leq k \leq n$. Applying both sides of (2.20) to $\frac{\partial}{\partial x^k}\Big|_p$ proves our claim and we deduce that $g_{ij}(p) = \varepsilon_{ij}$.

Next, note that $\frac{\partial g_{ij}}{\partial x^k}(p) = 0$ implies with the help of Lemma 2.93 that all Christoffel symbols at $p$ must also vanish. We will however first show the latter and use it to prove the former. We first show that the local form of geodesics $\varphi \circ \gamma$ is of the claimed form. By construction of the exponential map, $\gamma(t) = \exp_p(tw)$ for all $t \in (-\varepsilon, \varepsilon)$. Writing $w = \sum_{k=1}^{n} w^k v_k$, we have by definition of Riemannian normal coordinates

$$\gamma(t) = \exp_p(tw) = \exp_p\left(\sum_{k=1}^{n} tw^k v_k\right) = \exp_p\left(\sum_{k=1}^{n} x^k(\gamma(t))v_k\right),$$

showing that $\varphi(\gamma(t)) = tw$ for all $t \in (-\varepsilon, \varepsilon)$ as claimed. Writing down the geodesic equation for $\gamma$ in our local coordinates as in (2.17) at $p$ with $\ddot{x}^k(0) = \frac{\partial^2(x^k(\gamma))}{\partial t^2}(0) = 0$ and $\dot{x}^k(0) = w^k$ for all $1 \leq k \leq n$ shows that

$$\sum_{i,j=1}^{n} \Gamma^k_{ij}(p) w^i w^j = 0$$

for all $1 \leq k \leq n$. Since this holds for arbitrary initial condition for the geodesic $\gamma'(0) = w \in T_pM$, this proves that for each fixed $1 \leq k \leq n$, $(\Gamma^k_{ij}(p))_{ij}$ viewed as symmetric bilinear form on $T_pM \times T_pM$ must vanish identically. Hence, $\Gamma^k_{ij}(p) = 0$ for all $1 \leq i, j, k \leq n$.

For the last claim of this proposition, that is the vanishing partial derivatives of each $g_{ij}$ at $p$, observe that $\nabla$ being metric implies

$$\frac{\partial g_{ij}}{\partial x^k} = \frac{\partial}{\partial x^k}\left(g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)\right) = g\left(\sum_{\ell=1}^{n} \Gamma^\ell_{ki} \frac{\partial}{\partial x^\ell}, \frac{\partial}{\partial x^j}\right) + g\left(\sum_{\ell=1}^{n} \Gamma^\ell_{kj} \frac{\partial}{\partial x^\ell}, \frac{\partial}{\partial x^i}\right)$$

for all $1 \leq k \leq n$. Evaluating the above equation at $p$ and using that all Christoffel symbols vanish at $p$ yields the desired result. $\square$

**Remark 2.129. Warning:** In Proposition 2.128 we have seen that with the right choice of coordinates, any pseudo-Riemannian metric and Levi-Civita connection can be brought to a very simple form **at a chosen point**. While this works of course for every point in the manifold, this does **not** mean that every pseudo-Riemannian metric is **locally** of the form $g_{ij} = \varepsilon_{ij}$ on some open neighbourhood of our chosen reference point, this can in general only be achieved at said point! Otherwise, every manifold would be flat, and comparing with Section 3 shows that this is clearly not the case.

For the next corollary recall the definition of the Landau symbols and Taylor expansions.

**Corollary 2.130.** The Taylor expansion of $g_{ij}$ in Riemannian normal coordinates $(x^1, \dots, x^n)$ at their reference point $p \in M$ is of the form

$$g_{ij} = \varepsilon_{ij} + O\left(\sum_{k=1}^{n} (x^k)^2\right).$$

Before proving two, in practice very helpful, facts about geodesic completeness, we will review other results in the setting of geodesics, variations, and the exponential map of Riemannian manifolds that we cannot prove in detail due to time constrains. An excellent reference for the following remark is [L1, Ch. 6] and [L1, Ch. 10].

**Remark 2.131.**

(i) We have seen that geodesics with compact domain in Riemannian manifolds are critical points of the energy functional, and solving Exercise 2.124 shows that they are in fact also critical points of the length functional. One can, however, show more and prove that they are not just any type of critical point but **local minimisers**, meaning that for any variation $\eta$ of $\gamma$ with fixed endpoints, $E(\eta(s,\cdot)) \geq E(\gamma)$ and $L(\eta(s,\cdot)) \geq L(\gamma)$ for $s$ small enough. For a reference see [L1, Ch. 6].

(ii) An other way to study Riemannian manifolds is in the context of **metric geometry**. In fact, every Riemannian metric $g$ on a smooth manifold $M$ induces the structure of a **metric space** on $M$, which in turn induces a topology on $M$. It turns out that the induced topology on $M$ coincides, independently of the Riemannian metric $g$, with the initial topology on $M$.

(iii) We have interpreted variations of curves as a family of curves depending on one parameter. In the case that $\gamma : I \to M$ is a geodesic and $\eta : (-\varepsilon, \varepsilon) \times I \to M$ is a variation of $\gamma$, are there choices for $\eta$, such that every $\eta(s,\cdot) : I \to M$ is a geodesic, not just $\eta(0,\cdot) = \gamma$? The answer is yes, and the corresponding variational vector fields are called **Jacobi fields**.

In the setting of metric geometry a natural question that arises is whether $M$ is complete as a metric space. It turns out that this is closely related to geodesic completeness:

**Theorem 2.132** (Hopf-Rinow)**.** Let $(M, g)$ be a Riemannian manifold. Then the following are equivalent:

(i) $(M, g)$ is geodesically complete.

(ii) $M$ with the induced metric[49] from the Riemannian metric $g$ is complete as a metric space.

(iii) Every closed and bounded[50] subset of $M$ is compact.

*Proof.* For this version of the theorem see [O, Ch. 5, Thm. 21] or, alternatively, [G, Thm. 2.10.2] (in German), [L1, Thm. 6.13], [Bae, Thm. 5.2.2]. □

**Example 2.133.**

(i) Every compact Riemannian manifold is geodesically complete. This is an immediate consequence of Theorem 2.132. In particular, compact submanifolds of $\mathbb{R}^n$ are geodesically complete.

[49] As in metric space.

[50] W.r.t. the induced metric.

(ii) $\mathbb{R}^n$ equipped with the pseudo-Riemannian metric induced by $\langle\cdot,\cdot\rangle_\nu$, cf. Proposition 2.3, is geodesically complete. Any bounded open submanifold of $\mathbb{R}^n$ with the restriction of these metrics is incomplete.

(iii) The hyperbolic upper half plane, $H = \{(x,y) \in \mathbb{R}^2 \mid y > 0\}$ equipped with the Riemannian metric $g = \frac{dx^2+dy^2}{y^2}$, is geodesically complete.

Next we will discuss two methods for (realistically) checking whether a given Riemannian manifold is complete. In particular they do not necessarily involve solving any geodesic equation explicitly.

**Lemma 2.134.** A Riemannian manifold $(M,g)$ is geodesically complete if and only if every curve with image not contained in any compact set has infinite length.

*Proof.* If $(M,g)$ is geodesically complete, a curve $\gamma$ that is not contained in any compact set is by Theorem 2.132 in particular not contained in the closure of the geodesic ball $B^g_r(\gamma(t_0))$ for any $t_0$ in the domain of $\gamma$ and any $r > 0$. Hence, $\gamma$ has infinite length.

If $(M,g)$ is geodesically incomplete, we can find an inextensible geodesic $\gamma : [0,a) \to M$, $a > 0$, of unit speed. Suppose that $\gamma([0,a))$ is contained in a compactum $K \subset M$. Then $\gamma$ converges in $K$ and can thus be extended as a geodesic, which is a contradiction. □

**Lemma 2.135.**

(i) Let $M$ be a smooth manifold and $g, h$ Riemannian metrics on $M$. Assume that for all $p \in M$ and all $v \in T_pM$, $h_p(v,v) \geq g_p(v,v)$, or $h \geq g$ for short. If $(M,g)$ is geodesically complete, $(M,h)$ is also geodesically complete.

(ii) Let $(M,g)$ be a Riemannian manifold. If there exists $R > 0$, such that $\overline{B^g_R(p)}$ is compactly embedded in $M$ for all $p \in M$, then $(M,g)$ is geodesically complete.

*Proof.* Exercise. □

**Remark 2.136.** While geodesic completeness is defined for pseudo-Riemannian manifolds with arbitrary index, its main importance lies in Riemannian geometry. A concept of similar importance in Lorentz geometry is **global hyperbolicity**, cf. [BGP].

# 3 Curvature

In this section we will study different notions of curvature of pseudo-Riemannian manifolds. For surfaces in $\mathbb{R}^3$, one has a heuristic idea what curvature should mean, i.e. looking locally like a part of an affine plane should mean not curved, looking like a parabola, a hyperboloid, or a monkey saddle should mean curved in some sense. How do we formalise this, even in the aforementioned case, in a coordinate-free way? How does this relate to the notion of **Gauß-curvature** of surfaces in $\mathbb{R}^3$? First, we will introduce the pseudo-Riemannian curvature tensor by considering parallel translation around infinitesimal parallelograms. The reason why we want to study curvature is as follows. As of now, we do not have effective tools to check if two given pseudo-Riemannian manifolds might or might not be isometric. For example, while we have shown that near every point we can find Riemannian normal coordinates turning the metric **at that point** into a very simple form, it is not clear yet why this might not be possible **locally**. This means that we do not know yet if we can always find local coordinates $\varphi$ on a subset $U \subset M$ a pseudo-Riemannian manifold near every fixed point so that the pseudo-Riemannian metric has coordinate representation $(\varphi(U), \langle\cdot,\cdot\rangle_\nu)$. We will see that this is in general not true.

**Definition 3.1.** Let $(M,g)$ be a pseudo-Riemannian manifold with Levi-Civita connection $\nabla$. The **Riemann**[51] **curvature tensor** of $(M,g)$ is defined as

$$R : \mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M), \quad R(X,Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z, \tag{3.1}$$

for all $X, Y, Z \in \mathfrak{X}(M)$. In the above formula, we understand $\nabla_X \nabla_Y Z$ as $\nabla_X(\nabla_Y Z)$, analogously for $X$ and $Y$ interchanged.

**Exercise 3.2.**

(i) Check that the Riemann curvature tensor is, in fact, a tensor field, i.e. $R \in \mathcal{T}^{1,3}(M)$.

(ii) Formally replace "$\nabla$" in the right hand side of (3.1) with "$\mathcal{L}$", that is the Lie derivative. Verify that the so-defined expression vanishes identically.

(iii) Show that the Riemann curvature tensor vanishes identically if $\dim(M) = 1$.

Before studying the Riemann curvature tensor any further, we must ask ourselves which geometric picture motivates its definition in the first place. Compare the following construction with Remark 2.86 and make sure to understand the difference.

**Lemma 3.3.** Let $(M,g)$ be a pseudo-Riemannian manifold with Levi-Civita connection $\nabla$. For any $X \in \mathfrak{X}(M)$, denote for every $p \in M$ by $P_0^t(X) : T_pM \to T_{\gamma(t)}M$ the parallel transport map with respect to $\nabla$ along the integral curve $\gamma : (-\varepsilon, \varepsilon) \to M$ of $X$ with $\gamma(0) = p$ for $\varepsilon > 0$ small enough. For $p \in M$ fixed we have $P_0^t(X) = P_0^t(\gamma)$. Let $(x^1, \dots, x^n)$ be local coordinates on $U \subset M$. Then

$$R\left(\left.\frac{\partial}{\partial x^i}\right|_p, \left.\frac{\partial}{\partial x^j}\right|_p\right) \left.\frac{\partial}{\partial x^k}\right|_p = \left.\frac{\partial}{\partial s}\right|_{s=0} \left.\frac{\partial}{\partial t}\right|_{t=0} P_0^s\left(\frac{\partial}{\partial x^i}\right)^{-1} P_0^t\left(\frac{\partial}{\partial x^j}\right)^{-1} P_0^s\left(\frac{\partial}{\partial x^i}\right) P_0^t\left(\frac{\partial}{\partial x^j}\right) \left.\frac{\partial}{\partial x^k}\right|_p \tag{3.2}$$

for all $1 \leq i, j, k \leq n$ and all $p \in U$. The Riemann curvature tensor is the unique $(1,3)$-tensor field fulfilling (3.2) in all local coordinates.

*Proof.* In their coordinate representations, $P_0^t\left(\frac{\partial}{\partial x^j}\right) : T_pM \to T_{\gamma(t)}M$ for $t \in (-\varepsilon, \varepsilon)$, $\varepsilon$ small enough, and the other parallel translations are smooth maps of the form

$$\widehat{P_0^t\left(\frac{\partial}{\partial x^j}\right)} : (-\varepsilon, \varepsilon) \to \mathrm{GL}(n),$$

where $\mathrm{GL}(n)$ being the codomain follows from Lemma 2.78. The above map should be understood as mapping prefactors of vectors in $T_pM$ written in the coordinate basis to prefactors of vectors in $T_{\gamma(t)}M$, again written in the coordinate basis. This means for the right hand side of (3.2) that the partial derivatives behave according to the product rule of matrix valued curves, i.e. for all $A, B : (-\varepsilon, \varepsilon) \to \mathrm{GL}(n)$ smooth with $A(0) = B(0) = \mathbb{1}$ and all $v \in \mathbb{R}^n$ we have

$$\begin{aligned}
&\left.\frac{\partial}{\partial s}\right|_{s=0} \left.\frac{\partial}{\partial t}\right|_{t=0} A(s)^{-1} B(t)^{-1} A(s) B(t) v \\
&= \left(\left.\frac{\partial}{\partial s}\right|_{s=0} A(s)^{-1}\right) \left(\left.\frac{\partial}{\partial t}\right|_{t=0} B(t)^{-1}\right) v + \left(\left.\frac{\partial}{\partial t}\right|_{t=0} B(t)^{-1}\right) \left(\left.\frac{\partial}{\partial s}\right|_{s=0} A(s)\right) v \\
&= \left(\left.\frac{\partial}{\partial s}\right|_{s=0} A(s)\right) \left(\left.\frac{\partial}{\partial t}\right|_{t=0} B(t)\right) v - \left(\left.\frac{\partial}{\partial t}\right|_{t=0} B(t)\right) \left(\left.\frac{\partial}{\partial s}\right|_{s=0} A(s)\right) v.
\end{aligned}$$

Note that $\left.\frac{\partial}{\partial s}\right|_{s=0} A(s) \in \mathrm{End}(\mathbb{R}^n)$, meaning that the derivative is in general not invertible. Using fix first line + explanation!

[51] As for Riemannian normal coordinates, $R$ is always called the Riemann curvature tensor, even if $(M,g)$ is not Riemannian.

Proposition 2.79 we thus obtain

$$\begin{aligned}&\left.\frac{\partial}{\partial s}\right|_{s=0}\left.\frac{\partial}{\partial t}\right|_{t=0} P_0^s\left(\tfrac{\partial}{\partial x^i}\right)^{-1} P_0^t\left(\tfrac{\partial}{\partial x^j}\right)^{-1} P_0^s\left(\tfrac{\partial}{\partial x^i}\right) P_0^t\left(\tfrac{\partial}{\partial x^j}\right) \left.\tfrac{\partial}{\partial x^k}\right|_p \\ &= \left.\left(\nabla_{\frac{\partial}{\partial x^i}}\nabla_{\frac{\partial}{\partial x^j}}\tfrac{\partial}{\partial x^k} - \nabla_{\frac{\partial}{\partial x^j}}\nabla_{\frac{\partial}{\partial x^i}}\tfrac{\partial}{\partial x^k}\right)\right|_p\end{aligned}$$

which proves equation (3.2). In order to show that the Riemann curvature tensor is indeed the unique tensor field fulfilling the above, we only need to check that for any local functions $X^1, Y^1, Z^1, \dots, X^n, Y^n, Z^n \in C^\infty(U)$ and all $p \in U$, $\sum_{i,j,k} X^i(p)Y^j(p)Z^k(p)R\left(\left.\frac{\partial}{\partial x^i}\right|_p, \left.\frac{\partial}{\partial x^j}\right|_p\right) \left.\frac{\partial}{\partial x^k}\right|_p$ and $\left.(R(X,Y)Z)\right|_p$ as in (3.1) coincide. This follows immediately from the tensoriality of $R$ proven in Exercise 3.2 (i). ☐

Lemma 3.3 means that the Riemann curvature tensor measures the infinitesimal change of vectors parallelly transported around infinitesimal parallelograms. It is important to note that firstly, we allow pairs of opposite edges of said parallelograms to become small independently, meaning that $R$ involves second order differentiation. Secondly note that this already points to the importance of planes in tangent spaces, that is two-dimensional linear subspaces, for the interpretation of the curvature, see Lemma 3.18.

**Lemma 3.4.** In local coordinates $(x^1, \dots, x^n)$ the Riemann curvature tensor of a pseudo-Riemannian manifold $(M,g)$ has components

$$R^\ell{}_{ijk} := dx^\ell\left(R\left(\tfrac{\partial}{\partial x^i}, \tfrac{\partial}{\partial x^j}\right)\tfrac{\partial}{\partial x^k}\right),$$

so that locally[52] $R = \sum_{i,j,k,\ell} R^\ell{}_{ijk} \frac{\partial}{\partial x^\ell} \otimes dx^i \otimes dx^j \otimes dx^k$. The local functions $R^\ell{}_{ijk}$ are given by

$$R^\ell{}_{ijk} = \frac{\partial \Gamma^\ell_{jk}}{\partial x^i} - \frac{\partial \Gamma^\ell_{ik}}{\partial x^j} + \sum_{m=1}^n \left(\Gamma^\ell_{im}\Gamma^m_{jk} - \Gamma^\ell_{jm}\Gamma^m_{ik}\right)$$

for all $1 \le i,j,k,\ell \le n$.

*Proof.* Direct calculation. ☐

The Riemann curvature tensor fulfils various symmetry and anti-symmetry identities.

**Lemma 3.5.** Let $(M,g)$ be a pseudo-Riemannian manifold with Riemann curvature tensor $R$. Then

(i) $R(X,Y) = -R(Y,X)$,

(ii) $g(R(X,Y)Z,W) = -g(Z,R(X,Y)W)$,

(iii) $R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0$ (**first** or **algebraic Bianchi identity**),

(iv) $g(R(X,Y)Z,W) = g(R(Z,W)X,Y)$,

(v) $(\nabla_X R)(Y,Z) + (\nabla_Y R)(Z,X) + (\nabla_Z R)(X,Y) = 0$ (**second** or **differential Bianchi identity**)

for all $X,Y,Z,W \in \mathfrak{X}(M)$.

[52] **Warning:** The order of the indices $i,j,k,\ell$ is not standardised. I have probably seen every possible combination, so be careful to understand which index refers to which input in other lecture notes, papers, and books.

*Proof.* (i) follows from (3.1). Note that this means that $R$ can be viewed as an $\mathrm{End}(TM)$-valued 2-form, cf. Remark 2.51.

In order to show (ii) it suffices to show that $g(R(X,Y)Z,Z) = 0$ for all $X,Y,Z \in \mathfrak{X}(M)$, that is precisely (ii) for $Z = W$, which then implies

$$\begin{aligned} 0 &= \frac{1}{2}\left(g(R(X,Y)(Z+W),Z+W) - g(R(X,Y)(Z-W),Z-W)\right) \\ &= g(R(X,Y)Z,W) - g(R(X,Y)W,Z). \end{aligned}$$

We might further assume that $[X,Y] = 0$, e.g. by setting $X$ and $Y$ locally to coordinate vector fields, since $R$ is a tensor field. We obtain using that $\nabla$ is a metric connection

$$\begin{aligned} g(R(X,Y)Z,Z) &= g(\nabla_X\nabla_Y Z,Z) - g(\nabla_Y\nabla_X Z,Z) \\ &= X(g(\nabla_Y Z,Z)) - g(\nabla_Y Z,\nabla_X Z) - Y(g(\nabla_X Z,Z)) + g(\nabla_X Z,\nabla_Y Z) \\ &= \tfrac{1}{2}(X(Y(g(Z,Z))) - Y(X(g(Z,Z)))) = \tfrac{1}{2}[X,Y](g(Z,Z)) = 0 \end{aligned}$$

for all $X,Y,Z$ with $[X,Y] = 0$ which, as explained above, proves (ii).

For the third identity (iii) we might, as before, assume using the tensoriality of $R$ that $[X,Y] = [Y,Z] = [Z,X] = 0$. Also observe that we might write (iii) as $\sum_{\text{cycl.}} R(X,Y)Z = 0$. We find for all $X,Y,Z \in \mathfrak{X}(M)$ with pairwise vanishing Lie bracket

$$\begin{aligned} \sum_{\text{cycl.}} R(X,Y)Z &= \sum_{\text{cycl.}} \nabla_X\nabla_Y Z - \sum_{\text{cycl.}} \nabla_Y\nabla_X Z \\ &= \sum_{\text{cycl.}} \nabla_Y\nabla_Z X - \sum_{\text{cycl.}} \nabla_Y\nabla_X Z \\ &= \sum_{\text{cycl.}} \nabla_Y[Z,X] = 0, \end{aligned}$$

where we have used that $\nabla$ is torsion free.

(iv) is a combinatorial exercise[53] and follows using the identities (i), (ii), and (iii).

For last identity (v) we will use Riemannian normal coordinates. Fix $p \in M$ and let $(x^1,\dots,x^n)$ be Riemannian normal coordinates at $p$ on an open neighbourhood $U \subset M$ of $p$. We want to show that for all

$$X_p = \sum_{i=1}^n X^i \left.\frac{\partial}{\partial x^i}\right|_p,\quad Y_p = \sum_{i=1}^n Y^i \left.\frac{\partial}{\partial x^i}\right|_p,\quad Z_p = \sum_{i=1}^n Z^i \left.\frac{\partial}{\partial x^i}\right|_p,\qquad X^1,Y^1,Z^1,\dots,X^n,Y^n,Z^n \in \mathbb{R},$$

it holds that

$$(\nabla_{X_p}R)(Y_p,Z_p) + (\nabla_{Y_p}R)(Z_p,X_p) + (\nabla_{Z_p}R)(X_p,Y_p) = 0. \tag{3.3}$$

Let $X,Y,Z \in \mathfrak{X}(U)$ be the constant extensions of $X_p, Y_p, Z_p$ in the sense that

$$X = \sum_{i=1}^n X^i \frac{\partial}{\partial x^i},\quad Y_p = \sum_{i=1}^n Y^i \frac{\partial}{\partial x^i},\quad Z_p = \sum_{i=1}^n Z^i \frac{\partial}{\partial x^i}.$$

Note that $[X,Y] = [Y,Z] = [Z,X] = 0$. Equation (3.3) follows if we can show that

$$\left((\nabla_X R)(Y,Z)W + (\nabla_Y R)(Z,X)W + (\nabla_Z R)(X,Y)W\right)|_p = 0$$

for all local vector fields $W \in \mathfrak{X}(U)$. Recall that all Christoffel symbols in our chosen local coordinates vanish at $p$, cf. Proposition 2.128. Hence, $\nabla_A B|_p = 0$ for all possible combinations of $A,B \in \{X,Y,Z\}$. We thus obtain for all cyclic permutations of $X,Y,Z$ and all $W \in \mathfrak{X}(U)$

$$(\nabla_Z R)(X,Y)W|_p = (\nabla_Z(R(X,Y)W) - R(\nabla_Z X,Y)W - R(X,\nabla_Z Y)W - R(X,Y)\nabla_Z W)|_p$$

[53] "Exercise" as in you should try proving this yourself.

$$= \left(\nabla_Z(R(X,Y)W) - R(X,Y)\nabla_Z W\right)\big|_p .$$

Hence by using that $p \in M$ was arbitrary, the tensoriality of $R$, and that $W \in \mathfrak{X}(U)$ was arbitrary we obtain

$$(\nabla_Z R)(X,Y) = [\nabla_Z, R(X,Y)] = [\nabla_Z, [\nabla_X, \nabla_Y]]$$

for all $X, Y, Z \in \mathfrak{X}(M)$ with pairwise vanishing Lie bracket. In the above equation, $[\cdot,\cdot]$ means the formal commutator of differential operators. By writing out

$$\sum_{\text{cycl.}} [\nabla_Z, [\nabla_X, \nabla_Y]] = \sum_{\text{cycl.}} (\nabla_X \nabla_Y \nabla_Z - \nabla_X \nabla_Z \nabla_Y - \nabla_Y \nabla_Z \nabla_X + \nabla_Z \nabla_Y \nabla_X),$$

we find that $\sum\limits_{\text{cycl.}} [\nabla_Z, [\nabla_X, \nabla_Y]] = 0$ and, hence, we conclude that (v) holds true. $\square$

The Riemann curvature tensor behaves well under isometries.

**Lemma 3.6.** Let $F : (M,g) \to (N,h)$ be an isometry and let $R^M$ and $R^N$ denote the Riemann curvature tensors of $(M,g)$ and $(N,h)$, respectively. Then

$$F_* \left( R^M(X,Y)Z \right) = R^N(F_*X, F_*Y)F_*Z$$

for all $X, Y, Z \in \mathfrak{X}(M)$.

*Proof.* It suffices to show that $F_* \nabla^M_X Y = \nabla^N_{F_*X}(F_*Y)$ for all $X, Y \in \mathfrak{X}(M)$, where $\nabla^M$ and $\nabla^N$ denote the Levi-Civita connections of $(M,g)$ and $(N,h)$, respectively. This is a lengthy, but not difficult calculation using the Koszul formula (2.14) for $\nabla^M$ and $\nabla^N$, the fact that we are dealing with an isometry, which in particular means that $F_* : \mathfrak{X}(M) \to \mathfrak{X}(N)$ is bijective, and Corollary 1.126. $\square$

**Definition 3.7.** A pseudo-Riemannian manifold with vanishing Riemann curvature tensor is called **flat**.

The easiest examples of flat pseudo-Riemannian manifolds are $(\mathbb{R}^n, \langle\cdot,\cdot\rangle_\nu)$, $0 \leq \nu \leq n$. There exist other not so obvious examples, e.g. the Riemannian cylinder $\mathbb{R} \times S^1$ and the torus $T^2 = S^1 \times S^1$, both equipped with the respective product metric.

Suppose that we are given a flat Riemannian manifold $(M,g)$. In our definition this means that its Riemann curvature tensor vanishes identically. This is, in fact, a sufficient condition to show that locally, $(M,g)$ is isometric to $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$, where $\langle\cdot,\cdot\rangle$ stands for the standard Riemannian metric given in canonical coordinates by $\sum\limits_{i=1}^n (du^i)^2$.

**Theorem 3.8.** An $n$-dimensional Riemannian manifold $(M,g)$ is flat if and only if it is locally isometric to $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$, meaning that for all $p \in M$ there exists an open neighbourhood $U \subset M$ of $p$ and an isometry $F : (U,g) \to (F(U), \langle\cdot,\cdot\rangle)$, $F(U) \subset \mathbb{R}^n$ open.

*Proof.* Lemma 3.6 implies that local isometry to $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$ implies flatness, i.e. $R \equiv 0$. The other direction of this proof requires more work, for details see [L1, Thm. 7.3] (with slightly different conventions). The idea is to construct a commuting[54] orthonormal local frame of $TM \to M$ near every given point. The key ingredient is that parallel transport of vectors at, say, $p \in M$, to a close enough point $q \in M$ does not depend on the chosen curve starting at $p$ and ending at $q$ if it is required to be contained in a small enough open neighbourhood of both $p$ and $q$. This follows from a similar argument as in Lemma 3.3. $\square$

[54] w.r.t. the Lie bracket

In Lemma 3.3 we have seen how to interpret the Riemann curvature tensor geometrically as infinitesimal change of parallel transport of tangent vectors around infinitesimal parallelograms. There is an alternative, more algebraically flavoured, interpretation of the Riemann curvature tensor involving covariant derivatives of second order.

**Definition 3.9.** Let $(M, g)$ be a pseudo-Riemannian manifold with Levi-Civita connection $\nabla$. Then for all $X, Y, Z \in \mathfrak{X}(M)$,

$$\nabla^2_{X,Y} Z := (\nabla_X(\nabla Z))(Y) = \nabla_X \nabla_Y Z - \nabla_{\nabla_X Y} Z \tag{3.4}$$

is called the **second covariant derivative** of $Z$ in direction $X, Y$.

Note that the second covariant derivative $\nabla^2_{X,Y}$ of vector fields is in general not symmetric in $X$ and $Y$. This holds true even if $X = \frac{\partial}{\partial x^i}$ and $Y = \frac{\partial}{\partial x^j}$ are both coordinate vector fields with respect to local coordinates $(x^1, \ldots, x^n)$, contrary to partial derivatives with respect to local coordinates, cf. Lemma 1.100. Also compare the second covariant derivative of vector fields with the definition of the covariant Hessian in Definition 2.95.

**Exercise 3.10.** Check that $(\nabla_X(\nabla Z))(Y) = \nabla_X \nabla_Y Z - \nabla_{\nabla_X Y} Z$ in (3.4) actually holds true for all $X, Y, Z \in \mathfrak{X}(M)$. Note that $\nabla_X(\nabla Z)$ is to be understood as the covariant derivative of the endomorphism field $\nabla Z \in \Gamma(\mathrm{End}(TM))$ in direction $X$.

By using second covariant derivatives, we can rewrite the Riemann curvature tensor as follows.

**Lemma 3.11.** The Riemann curvature tensor of a pseudo-Riemannian manifold $(M, g)$ with Levi-Civita connection $\nabla$ fulfils

$$R(X, Y)Z = \nabla^2_{X,Y} Z - \nabla^2_{Y,X} Z \tag{3.5}$$

for all $X, Y, Z \in \mathfrak{X}(M)$.

*Proof.* Using torsion-freeness of $\nabla$ we obtain

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

for all $X, Y, Z \in \mathfrak{X}(M)$. Equation (3.5) now follows by writing it out using (3.4). □

Lemma 3.11 means that the Riemann curvature tensor measure "how much" second order covariant derivatives of vector fields do not commute. In the case that the given pseudo-Riemannian manifold is flat we see that second order covariant derivatives in directions $X, Y$, respectively $Y, X$, do in fact commute. We can interpret this as a coordinate free Schwartz's Theorem for flat pseudo-Riemannian manifolds.

The Riemann curvature tensor of a pseudo-Riemannian manifold $(M, g)$ is a $(1, 3)$-tensor field. Another common definition is to instead define the Riemann curvature tensor as a $(0, 4)$-tensor field $\widetilde{R} \in \mathcal{T}^{0,4}(M)$ given by

$$\widetilde{R}(X, Y, Z, W) := g(R(X, Y)Z, W)$$

for all $X, Y, Z, W \in \mathfrak{X}(M)$. It is clear that $R$ can be recovered from $\widetilde{R}$ by raising the fitting index. In local coordinates $(x^1, \ldots, x^n)$, $\widetilde{R}$ is of the form

$$\widetilde{R} = \sum_{i,j,k,\ell} R_{ijk\ell}\, dx^i \otimes dx^j \otimes dx^k \otimes dx^\ell, \tag{3.6}$$

where $R_{ijk\ell} = \sum_m g_{\ell m} R^m{}_{ijk}$. Recall that we have seen in Proposition 2.128 that the first partial derivatives of the prefactors $g_{ij}$ of the metric $g$ in Riemannian normal coordinates $(x^1, \ldots, x^n)$ vanish at the reference point. Using the above definition of the curvature as a $(0, 4)$-tensor field, we obtain the following result for the second partial derivatives of the $g_{ij}$ at the reference point.

**Lemma 3.12.** Let $(M, g)$ be a pseudo-Riemannian manifold with Levi-Civita connection $\nabla$ and let $\varphi = (x^1, \ldots, x^n)$ be Riemannian normal coordinates at $p \in M$ corresponding to a choice of orthonormal basis $\{v_1, \ldots, v_n\}$ of $T_pM$. Let further $R_{ijk\ell}$ as in (3.6). Then the local prefactors $g_{ij}$ as in (2.2) of $g$ fulfil

$$\frac{\partial^2 g_{ij}}{\partial x^k \partial x^\ell}(p) = \frac{2}{3} R_{ijk\ell}(p).$$

for all $1 \le i, j, k, \ell \le n$.

*Proof.* [Bae, Prop. 3.1.12], try to solve it yourself before looking it up! □

The next curvature type we will study is the so-called sectional curvature which assigns to each nondegenerate plane in a given tangent space a real number. For a pseudo-Riemannian manifold $(M, g)$, a plane $\Pi \subset T_pM$ being nondegenerate means that $g_p|_{\Pi \times \Pi}$ is a pseudo-Euclidean scalar product.

**Definition 3.13.** Let $(M, g)$ be a pseudo-Riemannian manifold with Riemann curvature tensor $R$. Let $\Pi \subset T_pM$ be a nondegenerate plane spanned by linearly independent vectors $v, w \in T_pM$. The **sectional curvature** of $\Pi$ is defined by

$$K(\Pi) := K(v, w) := \frac{g(R(v, w)w, v)}{g(v, v)g(w, w) - g(v, w)^2}.$$

The first thing we need to investigate if the sectional curvature is well-defined, that is we need to show that $K(\Pi)$ is independent of the basis vectors $v, w$ of $\Pi$.

**Lemma 3.14.** *$K$ only depends on the plane $\Pi$, not on the choice of basis vectors $v, w$ of $\Pi$.*

*Proof.* Let $\{V, W\}$ be another basis of $\Pi$. Then we can write $v = aV + bW$, $w = cV + dW$ for $a, b, c, d \in \mathbb{R}$ and since both $\{v, w\}$ and $\{V, W\}$ are a basis of $\Pi$ we have that

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \neq 0.$$

A short calculation shows $g(R(v, w)w, v) = (ad - bc)^2 g(R(V, W)W, V)$ and $g(v, v)g(w, w) - g(v, w)^2 = (ad - bc)^2 (g(V, V)g(W, W) - g(V, W)^2)$ which proves our claim. Note that the latter also proves that $\Pi$ is nondegenerate if and only if $g(v, v)g(w, w) - g(v, w)^2 \neq 0$. □

Sectional curvature can be interpreted as a coordinate free generalisation of the Gauß-curvature of surfaces in $\mathbb{R}^3$. In order to properly understand this we need more knowledge about induced structure on submanifolds which is the topic of the next section, cf. Remark **??**.

**Definition 3.15.** A pseudo-Riemannian manifold $(M, g)$ is of **constant curvature** if its sectional curvatures coincide at every point for every nondegenerate plane in the corresponding tangent space.

**Exercise 3.16.**

(i) Find an explicit local formula for $K\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)$ in terms of the Christoffel symbols and their derivatives. How does this formula look like in Riemannian normal coordinates?

(ii) Show that $(S^n, \langle \cdot, \cdot \rangle|_{TS^n \times TS^n})$ has positive constant curvature.

(iii) Show that the hyperboloids $H^n_\nu$ as in Example 2.33 (ii) have constant curvature and determine the sign.

It turns out that knowing all sectional curvatures at all points is equivalent to knowing the Riemann curvature tensor. In order to prove this, we will define so-called abstract curvature tensors.

**Definition 3.17.** A $(1,3)$-tensor

$$F \in T_p^{1,3}M, \quad F : (u,v,w) \mapsto F(u,v)w \in T_pM \quad \forall u,v,w \in T_pM,$$

on a pseudo-Riemannian manifold $(M,g)$ is called **abstract curvature tensor** if it fulfils the identities[55]

(i) $F(v,w) = -F(w,v)$,

(ii) $g(F(v,w)V,W) = -g(V,F(v,w)W)$,

(iii) $\sum_{\text{cycl.}} F(u,v)w = 0$

for all $u,v,w,V,W \in T_pM$.

Note that similarly to the identity Lem. 3.5 (iv) for the Riemann curvature tensor one can show for abstract curvature tensors that $g_p(F(v,w)V,W) = g_p(F(V,W)v,w)$ for all $v,w,V,W$.

**Lemma 3.18.** Let $(M,g)$ be a pseudo-Riemannian manifold with Riemann curvature tensor $R$ and assume that for $p \in M$ fixed and an abstract curvature tensor $F \in T_p^{1,3}M$

$$K(v,w) = \frac{g(F(v,w)w,v)}{g(v,v)g(w,w) - g(v,w)^2} \tag{3.7}$$

for all linearly independent $v,w \in T_pM$ spanning a nondegenerate plane in $T_pM$. Then $F = R_p$.

*Proof.* One checks that $F - R_p$ is an abstract curvature tensor. By linearity it thus suffices to show that $K(v,w) = 0$ for all linearly independent $v,w \in T_pM$ implies $F = 0$ in (3.7). As discussed in the proof of Lemma 3.14, the denominator $g(v,v)g(w,w) - g(v,w)^2$ does not vanish by the requirement that $\Pi = \operatorname{span}_{\mathbb{R}}\{v,w\}$ is nondegenerate. Hence, it suffices to show that $g(F(v,w)w,v) = 0$ for all $v,w$ spanning a nondegenerate plane implies $F = 0$. Note that this is in fact equivalent to showing $g(F(v,w)w,v) = 0$ for all $v,w \in T_pM$. In the Riemannian and index equal to the dimension of $M$ cases, every plane is nondegenerate, so this claim is trivial. So supposed that the index of $g$ is between $1$ and $n-1$. To see that the claim also holds in these cases, first note that

$$T_pM \times T_pM \ni (v,w) \mapsto g(F(v,w)w,v) \in \mathbb{R} \tag{3.8}$$

is continuous. Suppose that $g(F(v,w)w,v) = 0$ for all $v,w \in T_pM$ spanning a nondegenerate plane in $T_pM$ holds true but there are $V,W \in T_pM$ spanning a degenerate plane in $T_pM$, such that $g(F(V,W)W,V) \neq 0$

$$g(F(V,W)W,V) \neq 0. \tag{3.9}$$

To show that this is a contradiction to the continuity of (3.8) we show that there exist vectors $V_\varepsilon$ and $W_\varepsilon$ in arbitrary small open neighbourhoods of $V,W$ in $T_pM$, respectively, such that $\Pi_\varepsilon := \operatorname{span}_{\mathbb{R}}\{V_\varepsilon, W_\varepsilon\}$ is nondegenerate. A plane $\Pi_\varepsilon$ spanned by linearly independent $V_\varepsilon, W_\varepsilon$ is nondegenerate if and only if $g(V_\varepsilon,V_\varepsilon)g(W_\varepsilon,W_\varepsilon) - g(V_\varepsilon,W_\varepsilon)^2 \neq 0$. This follows from the fact that in the basis $V_\varepsilon, W_\varepsilon$ of $\Pi_\varepsilon \subset T_pM$, the representing symmetric $2\times 2$-matrix of $g|_{\Pi_\varepsilon \times \Pi_\varepsilon}$ is given by

$$\begin{pmatrix} g(V_\varepsilon,V_\varepsilon) & g(V_\varepsilon,W_\varepsilon) \\ g(V_\varepsilon,W_\varepsilon) & g(W_\varepsilon,W_\varepsilon) \end{pmatrix}.$$

[55] Compare these to Lemma 3.5 (i)–(v).

If $g(V,V)=0$, that is $V$ is a timelike vector, $\mathrm{span}_{\mathbb{R}}\{V,W\}$ being degenerate is equivalent to $g(V,W)=0$. In this case choose a vector $\xi \in T_pM$, such that $g(V,\xi)\neq 0$. If $g(V,V)<0$, choose $\xi \in T_pM$ with $g(\xi,\xi)>0$, if $g(V,V)>0$, choose $\xi\in T_pM$ with $g(\xi,\xi)<0$. In any of the three described cases set $V_\varepsilon = V$ and $W_\varepsilon = W+\varepsilon\xi$. Then

$$g(V_\varepsilon,V_\varepsilon)g(W_\varepsilon,W_\varepsilon)-g(V_\varepsilon,W_\varepsilon)^2 = c\varepsilon + \left(g(V,V)g(\xi,\xi)-g(V,\xi)^2\right)\varepsilon^2$$

for some $c\in\mathbb{R}$. If $c\neq 0$, we see that there exists $\delta>0$ such that for all $\varepsilon$ with $0<|\varepsilon|<\delta$, $g(V_\varepsilon,V_\varepsilon)g(W_\varepsilon,W_\varepsilon)-g(V_\varepsilon,W_\varepsilon)^2$ is either positive (if $c>0$) or negative (if $c<0$). If $c=0$, we have by construction $g(V,V)g(\xi,\xi)-g(V,\xi)^2<0$ in any case of $g(V,V)=0$, $g(V,V)<0$, or $g(V,V)>0$. Hence, we again obtain that there exists $\delta>0$ such that for all $\varepsilon$ with $0<|\varepsilon|<\delta$, $g(V_\varepsilon,V_\varepsilon)g(W_\varepsilon,W_\varepsilon)-g(V_\varepsilon,W_\varepsilon)^2$ is negative. We conclude that in any case the corresponding plane $\Pi_\varepsilon$ is nondegenerate. Since $\varepsilon$ is allowed to be arbitrary small and (3.9) is supposed to hold, this is a contradiction to the assumption that $g(F(v,w)w,v)=0$ for all $v,w$ spanning a nondegenerate plane and the continuity of (3.8).

The next step is to show that $g(F(v,w)w,v)=0$ for all $v,w\in T_pM$ implies $F=0$. First note that $g(F(v,w)u,v)$ is symmetric in $u,w$ for all $u,v,w\in T_pM$. This follows from

$$g(F(v,w)u,v)=g(F(u,v)v,w)=g(F(v,u)w,v)$$

for all $u,v,w\in T_pM$. We further obtain

$$0=g(F(v,w+u)(w+u),v)=2g(F(v,w)u,v)=-2g(F(v,w)v,u)$$

for all $u,v,w\in T_pM$, which shows $F(v,w)v=0$ for all $v,w\in T_pM$. Hence we have for all $u,v,w\in T_pM$

$$0=F(v+u,w)(v+u)=F(v,w)u+F(u,w)v,$$

which is equivalent to $F(v,w)u=F(w,u)v$. Using that we obtain

$$0=\sum_{\text{cycl.}}F(v,w)u=3F(v,w)u$$

for all $u,v,w\in T_pM$, proving that indeed $F=0$. $\square$

As a consequence of Lemma 3.18 we obtain the following formula for the Riemann curvature tensor of pseudo-Riemannian manifolds with constant curvature.

**Corollary 3.19.** Let $(M,g)$ be a pseudo-Riemannian manifold with constant sectional curvature $K=c\in\mathbb{R}$. Then the Riemann curvature tensor of $(M,g)$ fulfils

$$R(X,Y)Z=c(g(Y,Z)X-g(X,Z)Y) \tag{3.10}$$

for all $X,Y,Z\in\mathfrak{X}(M)$.

*Proof.* We check that for every point $p\in M$, the right hand side of (3.10) restricted to $T_pM\times T_pM\times T_pM$ defines an abstract curvature tensor fulfilling

$$K(v,w)=c$$

for all $v,w$ spanning a nondegenerate plane in $T_pM$. Lemma 3.18 now implies that the equality (3.10) holds at $p\in M$. Since $p\in M$ was arbitrary this finishes the proof. $\square$

Now we will introduce the Ricci[56] curvature which is obtained by contracting the Riemann curvature tensor.

[56] Gregorio Ricci-Curbastro (1853 – 1925)

**Definition 3.20.** Let $(M, g)$ be a pseudo-Riemannian manifold with Riemann curvature tensor $R$. The **Ricci curvature** $\mathrm{Ric} \in \mathcal{T}^{0,2}(M)$ is defined as

$$\mathrm{Ric}(X, Y) := \mathrm{tr}(R(\cdot, X)Y)$$

for all $X, Y \in \mathfrak{X}(M)$ where

$$R(\cdot, X)Y \in \mathcal{T}^{1,1}(M), \quad R(\cdot, X)Y : Z \mapsto R(Z, X)Y.$$

In local coordinates $(x^1, \ldots, x^n)$, the Ricci curvature is of the form

$$\mathrm{Ric} = \sum_{i,j=1}^{n} \mathrm{Ric}_{ij}\, dx^i \otimes dx^j = \sum_{i,j=1}^{n} \left( \sum_{k=1}^{n} R^k{}_{kij} \right) dx^i \otimes dx^j.$$

**Exercise 3.21.**

(i) Show that Ric is symmetric, that is $\mathrm{Ric}(X, Y) = \mathrm{Ric}(Y, X)$ for all $X, Y \in \mathfrak{X}(M)$.

(ii) Determine a local formula for each $\mathrm{Ric}_{ij}$ in terms of the Christoffel symbols.

(iii) Find a formula for Ric for pseudo-Riemannian manifolds of constant curvature.

The Ricci curvature plays a prominent role in general relativity and, as indicated by the name, the study of the Ricci flow. On a pseudo-Riemannian manifold $(M, g_0)$ The Ricci flow is (up to a multiplication of the right hand side with a constant) formally defined as a system of second order ODEs

$$\frac{\partial g_t}{\partial t} = -2\mathrm{Ric}^{g_t},$$

with fitting initial condition. Here $g_t$ is a smooth family of pseudo-Riemannian metrics and $\mathrm{Ric}^{g_t}$ denotes the Ricci curvature of $(M, g_t)$. This is a very actively studied field of modern mathematics, for an introduction see [CK, T].

In case that $\mathrm{Ric} = \lambda g$ for a pseudo-Riemannian manifold $(M, g)$ and some real number $\lambda \in \mathbb{R}$, $(M, g)$ is called **Einstein manifold**. Einstein manifolds are subject of active research in both physics, where they originate as critical points of the total scalar curvature functional, and mathematics. For a (mathematical) introduction see [Bes].

**Exercise 3.22.** Find a solution of the Ricci flow equation for $(M, g)$ Einstein. What can you say about the maximal existence interval for the time variable $t$?

The Ricci curvature can be used to define a scalar curvature invariant as follows.

**Definition 3.23.** The **scalar curvature** of a pseudo-Riemannian manifold $(M, g)$ is defined as

$$S := \mathrm{tr}_g(\mathrm{Ric}) \in C^\infty(M).$$

Note that Exercise 3.21 ensures that $S$ is in fact well defined.

**Exercise 3.24.** Find a local formula of the scalar curvature in terms of the Christoffel symbols.

Why would one want to study scalar curvature invariants like the scalar curvature in the first place? The answer is that they are useful when trying to prove that two given pseudo-Riemannian manifolds are not isometric.

**Lemma 3.25.** The number of isolated local minima and maxima of the scalar curvature of a pseudo-Riemannian manifold is invariant under isometries.

*Proof.* Let $(M, g)$ and $(N, h)$ be two isometric pseudo-Riemannian manifolds with scalar curvature $S_M, S_N$, respectively, and let $F : M \to N$ be an isometry. It follows from Lemma 3.6 that $S_N = S_M \circ F$. Since $F$ is in particular a diffeomorphism the claim of this lemma follows. □

Hence, if we have two pseudo-Riemannian manifold $(M, g)$ and $(N, h)$ (with the same index, otherwise this would be a trivial statement) but do not know any of the topological invariants of $M$ and $N$ like their respective fundamental or homology groups, but still want to show that they are not isometric, we might as a reasonable ansatz try to study the extrema of their respective scalar curvature functions. It is not enough to check that their scalar curvatures are different at some points as we can see in the following exercise.

**Exercise 3.26.**

(i) Let $(M, g)$ be a pseudo-Riemannian manifold and $r > 0$. Let $S^g$ denote the scalar curvature of $(M, g)$ and $S^{rg}$ denote the scalar curvature of $(M, rg)$. Show that $S^{rg} = r^{-1} S^g$.

(ii) Show that the 2-torus as the product $T^2 = S^1 \times S^1$ equipped with the product metric[57] and the 2-torus embedded in $\mathbb{R}^3$ via $F : T^2 \to \mathbb{R}^3$ equipped with the induced metric $\langle \cdot, \cdot \rangle|_{TT^2 \times TT^2}$ are never isometric, independent of the chosen embedding $F$.

The scalar curvature can also be calculated using the sectional curvatures for a good choice of basis of the tangent spaces.

**Lemma 3.27.** Let $(M, g)$ be an $n \geq 2$-dimensional pseudo-Riemannian manifold. For $p \in M$ fixed let $\{v_1, \ldots, v_n\}$ be an orthonormal basis of $T_pM$. Then

$$S(p) = \sum_{i \neq j} K(v_i, v_j).$$

*Proof.* Exercise. [Hint: Simply write out the right hand side of the equation.] □

**Remark 3.28.** Another commonly studied scalar curvature invariant of pseudo-Riemannian manifolds is the so-called **Kretschmann**[58] **scalar** which is for a pseudo-Riemannian manifold $(M, g)$ given by $g(R, R) \in C^\infty(M)$, cf. Definition 2.23.

# 4 Pseudo-Riemannian submanifolds

## 4.1 Induced structures

In this section we will deepen our studies of pseudo-Riemannian submanifolds. Throughout this sections, $\overline{M}$ will denote a smooth manifold which is supposed to be the ambient manifold of a smooth submanifold $M \subset \overline{M}$. Recall that this means that $M$ is an embedded submanifold of $\overline{M}$ via the inclusion map $\iota : M \to \overline{M}$. Also note that in this notation $\overline{M}$ is **not** the topological closure of $M$. Similarly, we will denote the geometrical structures on the ambient manifold $\overline{M}$ with bars, e.g. the Levi-Civita connection will be denoted by $\overline{\nabla}$.

**Definition 4.1.** The induced metric of a pseudo-Riemannian submanifold $M$ in $(\overline{M}, \overline{g})$ given by the restriction $g := \overline{g}|_{TM \times TM}$ is called **first fundamental form**.

Recall Definition 2.30.

[57] $S^1 \subset \mathbb{R}^2$ with induced metric $\langle \cdot, \cdot \rangle|_{TS^1 \times TS^1}$.
[58] Erich Justus Kretschmann (1887 – 1973)

**Definition 4.2.** Let $(M, g)$ be a pseudo-Riemannian submanifold of $(\overline{M}, \overline{g})$. We identify $\mathfrak{X}(M)$ with tangential sections of $T\overline{M}|_M \cong TM \oplus TM^\perp \to M$, where tangential means that the normal part of these sections vanishes identically. Sections in the subbundle $TM^\perp \to M$ are called **normal sections** and are denoted by $\mathfrak{X}(M)^\perp$. We further define **tangential and normal projections** as bundle maps

$$\tan : TM \oplus TM^\perp \to TM, \quad \mathrm{nor} : TM \oplus TM^\perp \to TM^\perp,$$

given fibrewise by $\tan(v + \xi) = v$ and $\mathrm{nor}(v + \xi) = \xi$ for all $v \in T_pM$ and all $\xi \in T_pM^\perp$.

The following holds for any smooth submanifolds $M \subset \overline{M}$ without any assumption on additional geometric data like a pseudo-Riemannian metric.

**Lemma 4.3.** Let $M \subset \overline{M}$ be a smooth submanifold. Let further $X, Y \in \mathfrak{X}(M)$ be arbitrary and $\overline{X}, \overline{Y} \in \mathfrak{X}(\overline{M})$ be arbitrary extensions of $X, Y$ to $\overline{M}$, i.e. $\overline{X}_p = X_p$ and $\overline{Y}_p = Y_p$ for all $p \in M$. Then $[\overline{X}, \overline{Y}]_p \in T_pM$ for all $p \in M$.

*Proof.* Follows from Lemma 1.125 with $\phi = \iota$, see also Remark 1.127. □

Whenever $M \subset \overline{M}$ is a pseudo-Riemannian submanifold, the above Lemma 4.3 justifies the notation $[X, Y] \in \mathfrak{X}(M)$ for all $X, Y \in \mathfrak{X}(M)$, where $\mathfrak{X}(M)$ is viewed as the set of sections of the tangent part of the bundle $T\overline{M}|_M \cong TM \oplus TM^\perp \to M$ along $M \subset \overline{M}$.

Next we need to study the Levi-Civita connection of the ambient manifold and show that it, in a certain sense, defines a connection in both $TM \to M$ and $TM^\perp \to M$.

**Lemma 4.4.** Let $M$ be a pseudo-Riemannian submanifold of $(\overline{M}, \overline{g})$ and let $\overline{\nabla}$ denote the Levi-Civita connection of $(\overline{M}, \overline{g})$. Let $X \in \mathfrak{X}(M)$ and $Y \in \Gamma(T\overline{M}|_M)$ with arbitrary extensions $\overline{X}, \overline{Y} \in \mathfrak{X}(\overline{M})$, respectively. Then

$$\overline{\nabla}_{\overline{X}} \overline{Y}\Big|_M \in \Gamma(T\overline{M}|_M)$$

is independent of the chosen extensions $\overline{X}$ and $\overline{Y}$.

*Proof.* Any integral curve of $\overline{X}$ starting in $M$ will have image contained in $M$ and can thus be viewed as an integral curve of $X$. Using Corollary 2.80 finishes the proof. □

Lemma 4.4 in particular justifies the notation $\overline{\nabla}_X Y$ for all $X, Y \in \mathfrak{X}(M)$. Furthermore, it implies the following.

**Corollary 4.5.** The Levi-Civita connection of a pseudo-Riemannian manifold $(\overline{M}, \overline{g})$ with pseudo-Riemannian submanifold $(M, g)$ induces a connection in $TM \oplus TM^\perp \to M$.

Note that it is at this point not clear if $\overline{\nabla}$ also induces a connection in the tangent bundle $TM \to M$ and the normal bundle $TM^\perp \to M$. The answer is yes to both as we will see next.

**Proposition 4.6.** Let $(M, g)$ be a pseudo-Riemannian submanifold of $(\overline{M}, \overline{g})$. Then the Levi-Civita connection $\nabla$ of $(M, g)$ is precisely the tangent part of the Levi-Civita connection $\overline{\nabla}$ of $(\overline{M}, \overline{g})$ restricted to $\mathfrak{X}(M) \times \mathfrak{X}(M)$, i.e.

$$\nabla_X Y = \tan \overline{\nabla}_X Y$$

for all $X, Y \in \mathfrak{X}(M)$.

*Proof.* First we need to show that $\tan\overline{\nabla}$ restricted to $\mathfrak{X}(M) \times \mathfrak{X}(M)$ is, in fact, a connection. This follows from the fibrewise linearity of $\tan : T_p\overline{M} \to T_pM$ and Lemma 4.4. Note that this easily implies that $\tan\overline{\nabla}A$ is well-defined for all tensor fields $A \in \mathfrak{T}^{r,s}(M)$. Next we need to check that $\tan\overline{\nabla}$ is metric and torsion-free. Let $X, Y, Z \in \mathfrak{X}(M)$ be arbitrary and let $\overline{X}, \overline{Y}, \overline{Z} \in \mathfrak{X}(\overline{M})$ be respective extensions. Since $g = \overline{g}|_{TM\times TM}$ we obtain for all $p \in M$ by the locality property of tangent vectors and Lemma 4.4

$$\begin{aligned}
(\tan\overline{\nabla}_X g)(Y,Z)|_p &= \Big(X(g(Y,Z)) - g(\tan\overline{\nabla}_X Y, Z) - g(Y, \tan\overline{\nabla}_X Z)\Big)\Big|_p \\
&= \Big(\overline{X}(\overline{g}(\overline{Y},\overline{Z})) - \overline{g}(\tan\overline{\nabla}_{\overline{X}}\overline{Y},\overline{Z}) - \overline{g}(\overline{Y},\tan\overline{\nabla}_{\overline{X}}\overline{Z})\Big)\Big|_p \\
&= \Big(\overline{X}(\overline{g}(\overline{Y},\overline{Z})) - \overline{g}(\overline{\nabla}_{\overline{X}}\overline{Y},\overline{Z}) - \overline{g}(\overline{Y},\overline{\nabla}_{\overline{X}}\overline{Z})\Big)\Big|_p .
\end{aligned}$$

In the last equivalence we have used that $\overline{X}_p, \overline{Y}_p, \overline{Z}_p$ are tangent to $M$, hence the normal part of e.g. $\overline{\nabla}_{\overline{X}}\overline{Y}$ at $p$ does not chance the value of $\overline{g}(\overline{\nabla}_{\overline{X}}\overline{Y},\overline{Z})$. Since $\overline{\nabla}$ is metric and $X, Y, Z \in \mathfrak{X}(M)$ and $p \in M$ were arbitrary, it follows that $\tan\overline{\nabla}g = 0$ as claimed. For the torsion-freeness, let $X, Y \in \mathfrak{X}(M)$ be arbitrary with arbitrary extensions $\overline{X}, \overline{Y} \in \mathfrak{X}(\overline{M})$, respectively. Using Lemmas 4.4, 4.3, and the fibrewise linearity of tan we obtain for all $p \in M$

$$\begin{aligned}
\Big(\tan\overline{\nabla}_X Y - \tan\overline{\nabla}_Y X - [X,Y]\Big)\Big|_p &= \Big(\tan\overline{\nabla}_{\overline{X}}\overline{Y} - \tan\overline{\nabla}_{\overline{Y}}\overline{X} - [\overline{X},\overline{Y}]\Big)\Big|_p \\
&= \tan\Big(\overline{\nabla}_{\overline{X}}\overline{Y} - \overline{\nabla}_{\overline{Y}}\overline{X} - [\overline{X},\overline{Y}]\Big)\Big|_p \\
&= 0,
\end{aligned}$$

where the last equality follows from the torsion-freeness of $\overline{\nabla}$. Summarizing, we have shown that $\tan\overline{\nabla}$ is a metric and torsion free connection in $TM \to M$ and, hence, coincides with the Levi-Civita connection $\nabla$ of $(M,g)$. □

Proposition 4.6 might be a surprising result at first glance, as it allows us to calculate covariant derivatives in pseudo-Riemannian submanifolds with respect to the Levi-Civita connection using only tangential projections and the Levi-Civita connection of the ambient manifold. The latter is usually easier to handle, in particular if $(\overline{M},\overline{g}) = (\mathbb{R}^n, \langle\cdot,\cdot\rangle_\nu)$. By a result from Nash[59] [N], this is actually the most general case if the ambient manifold is $\mathbb{R}^n$ equipped with the standard Riemannian metric $\langle\cdot,\cdot\rangle$. The reader is encouraged to take a look at the related publication since even though it is a very hard result, it is very neatly structured and well readable. The following is a special case of [N, Thm. 3].

**Theorem 4.7.** Let $(M,g)$ be a Riemannian manifold of dimension $n$. Then there exists an isometric embedding of $(M,g)$ into any open subset $U \subset \mathbb{R}^m$ for $m = \frac{3}{2}n^3 + 7n^2 + \frac{11}{2}n$ equipped with the standard Riemannian metric $\langle\cdot,\cdot\rangle$.

The above theorem means that studying Riemannian manifolds is equivalent to studying (automatically Riemannian) submanifolds of $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$. However, in reality, constructing an explicit embedding from, say, $\mathbb{R}P^n$ equipped with some Riemannian metric into any given open subset $U \subset \mathbb{R}^m$ for fitting $m$ is far from trivial.

**Exercise 4.8.** Find a formula for the Christoffel symbols of the Levi-Civita connection of a pseudo-Riemannian submanifold in adapted coordinates.

We have seen how to obtain a connection in the tangent bundle of a pseudo-Riemannian submanifold $(M,g)$ which turned out to be the best possible case, that is the Levi-Civita connection. The Levi-Civita connection of the ambient manifold $(\overline{M},\overline{g})$ also induces a connection in the normal bundle $TM^\top \to M$.

[59] John Nash (1928 – 2015)

**Proposition 4.9.** Let $(M, g)$ be a pseudo-Riemannian submanifold of $(\overline{M}, \overline{g})$ and let $\overline{\nabla}$ denote the Levi-Civita connection of $(\overline{M}, \overline{g})$. Then

$$\overline{\nabla}^{\mathrm{nor}} := \mathrm{nor}\overline{\nabla} : \mathfrak{X}(M) \times \mathfrak{X}(M)^{\perp} \to \mathfrak{X}(M)^{\perp}, \quad (X, \xi) \mapsto \mathrm{nor}\overline{\nabla}_X \xi,$$

for all $X \in \mathfrak{X}(M)$ and all $\xi \in \mathfrak{X}(M)^{\perp}$ is a connection in $TM^{\perp} \to M$, called the **normal connection**.

*Proof.* Follows from the fibrewise linearity of nor and Lemma 4.4. □

Recall that we have called the metric $g = \overline{g}|_{TM \times TM}$ of a pseudo-Riemannian submanifold $(M, g)$ of $(\overline{M}, \overline{g})$ the first fundamental form.

**Definition 4.10.** Let $(M, g)$ be a pseudo-Riemannian submanifold of $(\overline{M}, \overline{g})$ with Levi-Civita connection $\overline{\nabla}$ in $T\overline{M} \to \overline{M}$. The **second fundamental form** of $M$ is defined as

$$\mathrm{II} : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)^{\perp}, \quad \mathrm{II}(X, Y) := \mathrm{nor}\overline{\nabla}_X Y$$

for all $X, Y \in \mathfrak{X}(M)$.

**Lemma 4.11.** The second fundamental form is a symmetric $TM^{\perp}$-valued $(0, 2)$-tensor field, that is a section in $TM^{\perp} \otimes \mathrm{Sym}^2(T^*M) \to M$.

*Proof.* It is clear that $II$ is $C^{\infty}(M)$-linear in its first argument. Hence it suffices to show that $II$ is symmetric in order to also obtain the $C^{\infty}(M)$-linearity in the second argument. For the symmetry we check that for all $X, Y \in \mathfrak{X}(M)$ using the fibrewise linearity of nor, the torsion-freeness of $\overline{\nabla}$, and Lemma 4.3

$$II(X, Y) - II(Y, X) = \mathrm{nor}\left(\overline{\nabla}_X Y - \overline{\nabla}_Y X\right) = \mathrm{nor}[X, Y] = 0.$$

□

**Corollary 4.12.** The Levi-Civita connection $\nabla$ and second fundamental form $II$ of a pseudo-Riemannian submanifold $(M, g)$ of $(\overline{M}, \overline{g})$ fulfil the **Gauß equation**[60]

$$\overline{\nabla}_X Y = \nabla_X Y + II(X, Y) \tag{4.1}$$

for all $X, Y \in \mathfrak{X}(M)$.

**Exercise 4.13.**

(i) Find a formula for the second fundamental form in coordinates in terms of the Christoffel symbols of $(M, g)$ and $(\overline{M}, \overline{g})$.

(ii) For $U \subset \mathbb{R}^2$ open consider a smooth function $f : U \to \mathbb{R}$. The graph of $f$ is a Riemannian submanifold of $(\mathbb{R}^3, \langle \cdot, \cdot \rangle)$. Find a formula for the second fundamental form of $\left(\mathrm{graph}(f), g = \langle \cdot, \cdot \rangle|_{T\mathrm{graph}(f) \times T\mathrm{graph}(f)}\right)$ in terms of the first and second partial differentials of $f$.

The covariant derivatives of normal fields along pseudo-Riemannian manifolds also split in a certain manner. For this we need to define the following.

[60] There are **many** similar equations with that name, e.g. in affine and centro-affine differential geometry, so one has to make sure to clarify the context.

**Definition 4.14.** Let $(M, g)$ be a pseudo-Riemannian submanifold of $(\overline{M}, \overline{g})$. For all $\xi \in \mathfrak{X}(M)^{\perp}$, the $g$-symmetric endomorphism field $S^{\xi} \in \mathcal{T}^{1,1}(M)$ defined by the **Weingarten equation**

$$\overline{g}(II(X,Y), \xi) = g(S^{\xi}X, Y) \tag{4.2}$$

for all $X, Y \in \mathfrak{X}(M)$ is called **Weingarten**[61] **map** (alternatively **shape operator**).

Note that $S^{\xi}$ is well-defined for each $\xi \in \mathfrak{X}(M)^{\perp}$ by the fibrewise nondegeneracy of $g$.

**Exercise 4.15.** Determine a coordinate description of the Weingarten map in adapted coordinates.

**Proposition 4.16.** Let $(M, g)$ be a pseudo-Riemannian submanifold of $(\overline{M}, \overline{g})$. The Weingarten map fulfils the **Weingarten equation**

$$\overline{\nabla}_X \xi = -S^{\xi}X + \overline{\nabla}^{\mathrm{nor}}_X \xi \tag{4.3}$$

for all $X \in \mathfrak{X}(M)$ and all $\xi \in \mathfrak{X}(M)^{\perp}$.

*Proof.* Follows by writing out $\overline{g}(\overline{\nabla}_X \xi, \overline{Y})$ for $\overline{Y} \in \mathfrak{X}(\overline{M})$ arbitrary and using Definition 4.14, Proposition 4.9, and the metric property of $\overline{\nabla}$. □

## 4.2 Curvature of pseudo-Riemannian submanifolds

We have seen how to relate obtain the Levi-Civita connection and several related geometric data of a pseudo-Riemannian submanifold from the geometric information of its ambient pseudo-Riemannian manifold. We will now develop formulas for the various curvature tensors of pseudo-Riemannian submanifolds. We start with relating the respective Riemann curvature tensors.

**Proposition 4.17.** Let $(M, g)$ be a pseudo-Riemannian submanifold of $(\overline{M}, \overline{g})$. The Riemann curvature tensors $R$ of $(M, g)$ and $\overline{R}$ of $(\overline{M}, \overline{g})$ are related by the **Gauß equation for Riemann curvature tensors**

$$\overline{g}(\overline{R}(X,Y)Z, W) = g(R(X,Y)Z, W) + \overline{g}(\mathrm{II}(X,Z), \mathrm{II}(Y,W)) - \overline{g}(\mathrm{II}(Y,Z), \mathrm{II}(X,W)) \tag{4.4}$$

for all $X, Y, Z, W \in \mathfrak{X}(M)$.

*Proof.* Since (4.4) is a tensor equation, we might without loss of generality assume $[X, Y] = 0$. Observe that for all $X, Y, Z, W \in \mathfrak{X}(M)$, equation (4.3) implies

$$\begin{aligned}
\overline{g}(\overline{\nabla}_X \overline{\nabla}_Y Z, W) &= \overline{g}(\tan \overline{\nabla}_X \overline{\nabla}_Y Z, W) \\
&= \overline{g}(\nabla_X(\tan \overline{\nabla}_Y Z), W) + \overline{g}(\tan \overline{\nabla}_X(\mathrm{nor} \overline{\nabla}_Y Z), W) \\
&= \overline{g}(\nabla_X \nabla_Y Z, W) + \overline{g}(\tan \overline{\nabla}_X(\mathrm{II}(Y,Z)), W) \\
&= \overline{g}(\nabla_X \nabla_Y Z, W) - \overline{g}(S^{\mathrm{II}(Y,Z)}X, W) \\
&= \overline{g}(\nabla_X \nabla_Y Z, W) - \overline{g}(\mathrm{II}(X,W), \mathrm{II}(Y,Z)).
\end{aligned}$$

By doing the same calculation as above with $X$ and $Y$ interchanged and using $\overline{R}(X,Y) = \overline{\nabla}_X \overline{\nabla}_Y - \overline{\nabla}_Y \overline{\nabla}_X$ for $[X,Y] = 0$ we obtain the claimed formula (4.4). □

Since the first fundamental form of a pseudo-Riemannian submanifold is simply the restriction of the pseudo-Riemannian metric of the ambient manifold, Proposition 4.17 and Definition 3.13 imply the following for the sectional curvature of pseudo-Riemannian submanifolds.

[61] Julius Weingarten (1836 – 1910)

**Corollary 4.18.** Let $(M, g)$ be a pseudo-Riemannian submanifold of $(\overline{M}, \overline{g})$. For every nondegenerate plane spanned by $v, w \in T_pM$ in $T_pM \subset T_p\overline{M}$, the sectional curvatures $K$ of $(M, g)$ and $\overline{K}$ of $(\overline{M}, \overline{g})$ are related by

$$\overline{K}(v, w) = K(v, w) - \frac{\overline{g}(\mathrm{II}(v, v), \mathrm{II}(w, w)) - \overline{g}(\mathrm{II}(v, w), \mathrm{II}(v, w))}{g(v, v)g(w, w) - g(v, w)^2}. \tag{4.5}$$

Equation (4.5) is particularly useful if the ambient manifold $(\overline{M}, \overline{g})$ has constant curvature as it allows to calculate the sectional curvatures of any pseudo-Riemannian submanifold without having to calculate its Riemann curvature tensor.

**Example 4.19.** Consider the unit $n$-sphere $S^n \subset \mathbb{R}^{n+1}$ equipped with the restriction of the standard Riemannian metric. The sectional curvatures of the ambient Riemannian manifold $(\mathbb{R}^{n+1}, \langle\cdot,\cdot\rangle)$ all vanish and since we are in the Riemannian setting, every plane in $T_pS^n$ is nondegenerate for all $p \in S^n$. Without loss of generality we can assume for a given plane in $T_pS^n \subset T_p\mathbb{R}^n$ that it is spanned by two orthonormal vectors $v, w \in T_pM$. Hence, the sectional curvature of $S^n$ of that plane is given by

$$K(v, w) = \langle \mathrm{II}(v, v), \mathrm{II}(w, w)\rangle - \langle \mathrm{II}(v, w), \mathrm{II}(v, w)\rangle.$$

In the next step we need to determine II for which we will use (4.2). First note that $TS^{n\perp} \to S^n$ is a rank 1 vector bundle. A global frame is given by the position vector field $\xi \in \mathfrak{X}(S^n)^\perp$, $\xi_p = p$ for all $p \in S^n$. In order to determine the Weingarten map, we use (4.3), calculate

$$\overline{\nabla}_X \xi = X \tag{4.6}$$

for all $X \in \mathfrak{X}(S^n)$ and, hence, obtain $S^\xi = -\mathrm{id}_{TS^n}$. Since II is a $TS^{n\perp}$-valued symmetric $(0, 2)$-tensor, (4.6) implies with (4.2)

$$\langle II(X, Y), \xi\rangle = -\langle X, Y\rangle$$

for all $X, Y \in \mathfrak{X}(S^n)$, showing that $II(X, Y) = -\langle X, Y\rangle\xi$. Hence, $K(v, w) = 1$, and since the initial plane spanned by $v, w$ was arbitrary this shows that the unit $n$-sphere $S^n$ has constant curvature with value 1.

**Exercise 4.20.**

(i) Show that $S^n$ embedded as a sphere of radius $r > 0$ in $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$ has positive constant curvature $\frac{1}{r^2}$.

(ii) Show that $H_1^n$ as a Riemannian submanifold of the $n+1$-dimensional Minkowski space $(\mathbb{R}^{n+1}, \langle\cdot,\cdot\rangle_1)$ as in Example 2.33 (ii) has negative constant curvature.

**Theorem 4.21.** Gauss theorem without proof maybe?

**Example 4.22.** counterexample to existence of Lorentz metric needs euler char, check bookmark maybe not?

## 4.3 Geodesics of pseudo-Riemannian submanifolds

Next we will use the results about curvature of pseudo-Riemannian submanifolds to find tools for determining geodesics of pseudo-Riemannian submanifolds.

**Proposition 4.23.** Let $(M, g)$ be a pseudo-Riemannian submanifold of $(\overline{M}, \overline{g})$ with respective Levi-Civita connections $\nabla$ and $\overline{\nabla}$. A smooth curve $\gamma : I \to M$ is a geodesic with respect to $\nabla$ if and only if $\overline{\nabla}_{\gamma'}\gamma'$ is normal at every point, i.e. $\overline{\nabla}_{\gamma'}\gamma' \in \Gamma_\gamma(TM^\perp)$.

*Proof.* By the Gauß equation (4.1) we have

$$\overline{\nabla}_{\gamma'}\gamma' = \nabla_{\gamma'}\gamma' + \mathrm{II}(\gamma',\gamma').$$

Since $\mathrm{II}(\gamma',\gamma')$ is precisely the normal part of $\overline{\nabla}_{\gamma'}\gamma'$, the claim follows. $\square$

Proposition 4.23 usually makes life easier in case the ambient manifold is the flat Euclidean space. Try the following exercise first without our new results, that is by choosing coordinates, and then with the help of Proposition 4.23.

**Exercise 4.24.** Let $S^n \subset \mathbb{R}^{n+1}$ be the unit $n$-sphere with induced Riemannian metric. The geodesics in $S^n$ are curves of constant speed contained in the great circles.

Another interesting thing to consider for pseudo-Riemannian submanifolds $(M,g)$ of $(\overline{M},\overline{g})$ is geodesics of its ambient manifold with initial condition in $TM \subset T\overline{M}$, meaning $\gamma(0) = p \in M$ and $\gamma'(0) = v \in T_pM \subset T_p\overline{M}$. It is not automatically true that geodesics with such initial conditions stay in the submanifold, see Exercise 4.24. In the following, geodesic refers to geodesic with respect to the respective Levi-Civita connection.

**Definition 4.25.** A pseudo-Riemannian submanifold $(M,g)$ of a pseudo-Riemannian manifold $(\overline{M},\overline{g})$ is called **totally geodesic** if all geodesics $\gamma : I \to \overline{M}$ of $(\overline{M},\overline{g})$ starting in $M$ with initial velocity tangent to $M$ stay in $M$ for all time, i.e. $\gamma(I) \subset M$.

Using the Gauß equation (4.1) we obtain the following description of totally geodesic submanifolds.

**Lemma 4.26.** A pseudo-Riemannian submanifold $(M,g)$ of a pseudo-Riemannian manifold $(\overline{M},\overline{g})$ is totally geodesic if and only if its second fundamental form vanishes identically.

*Proof.* If $\mathrm{II} \equiv 0$, (4.1) implies that $\overline{\nabla}_{\gamma'}\gamma' = \nabla_{\gamma'}\gamma'$ for all smooth curves $\gamma : I \to M$. Hence, $\gamma$ is a geodesic in $(M,g)$ if and only if it is a geodesic in $(\overline{M},\overline{g})$.

Now suppose that $\mathrm{II} \not\equiv 0$. Fix $p \in M$ and $v \in T_pM \subset T_p\overline{M}$, such that $\mathrm{II}(v,v) \neq 0$. For $\varepsilon > 0$ small enough let $\gamma : (-\varepsilon,\varepsilon) \to M \subset \overline{M}$ be a geodesic in $(M,g)$ with $\gamma'(0) = v$. By (4.1) it follows that $\gamma$ is not a geodesic in $(\overline{M},\overline{g})$. Hence, $(M,g)$ is not totally geodesic. $\square$

In order to check if a pseudo-Riemannian submanifold is totally geodesic we can use the following equivalent conditions.

**Proposition 4.27.** Let $(M,g)$ be a pseudo-Riemannian submanifold of $(\overline{M},\overline{g})$. $(M,g)$ is totally geodesic if and only if one of the following equivalent statements hold:

(i) Every geodesic in $(M,g)$ is a geodesic in $(\overline{M},\overline{g})$.

(ii) For every geodesic $\gamma : I \to \overline{M}$ with $0 \in I$, $I$ open, and initial conditions $\gamma(0) = p$, $\gamma'(0) = v \in T_pM \subset T_p\overline{M}$ there exists $\varepsilon > 0$, such that $\gamma((-\varepsilon,\varepsilon)) \subset M$.

(iii) For every smooth curve $\gamma : [a,b] \to M \subset \overline{M}$ the parallel transport in $(M,g)$,

$$P_a^b(\gamma) : T_{\gamma(a)}M \to T_{\gamma(b)}M$$

coincides with the parallel transport in $(\overline{M},\overline{g})$,

$$\overline{P}_a^b(\gamma) : T_{\gamma(a)}\overline{M} \to T_{\gamma(b)}\overline{M}$$

restricted to $T_{\gamma(a)}M \subset T_{\gamma(a)}\overline{M}$.

*Proof.* (i) is by definition of totally geodesic pseudo-Riemannian submanifolds, cf. Definition 4.25, equivalent to $(M, g)$ being totally geodesic. (iii) is equivalent to $(M, g)$ being totally geodesic by Lemma 4.26.

To see that (i) and (ii) are equivalent, suppose first that (i) holds and let $\gamma : I \to \overline{M}$ be a geodesic in $(\overline{M}, \overline{g})$ with $0 \in I$, $I$ open, and $\gamma'(0) = v \in T_pM$. On the other hand, let $\widetilde{\gamma} : \widetilde{I} \to M$ be a geodesic in $(M, g)$ with $0 \in \widetilde{I}$, $\widetilde{I}$ open, and $\widetilde{\gamma}'(0) = v$. By assumption, $\widetilde{\gamma}$ is also a geodesic in $(\overline{M}, \overline{g})$ and by the uniqueness property of maximal geodesics coincides with $\gamma$ on $I \cap \widetilde{I}$. Choosing $\varepsilon > 0$ small enough so that $(-\varepsilon, \varepsilon) \subset I \cap \widetilde{I}$ proves the claim.

Next suppose that (ii) holds. We have already seen that (i) is equivalent to the vanishing of the second fundamental form II of $M \subset \overline{M}$. Let $p \in M$ and $v \in T_pM$ be arbitrary. Let further $\gamma : (-\varepsilon, \varepsilon) \to \overline{M}$ for any $\varepsilon > 0$ small enough be a geodesic in $(\overline{M}, \overline{g})$ with $\gamma(0) = p$, $\gamma'(0) = v$, such that $\gamma((-\varepsilon, \varepsilon)) \subset M$. By the Gauß equation (4.1) we have

$$0 = \overline{\nabla}_{\gamma'}\gamma'\Big|_{t=0} = \nabla_{\gamma'}\gamma'\big|_{t=0} + \mathrm{II}_p(v, v)$$

and by the fibrewise direct sum of the splitting $T\overline{M}|_M = TM \oplus TM^{\perp}$ and $p$ and $v$ being arbitrary, we obtain $\mathrm{II} \equiv 0$ as required. □

A nice tool to construct totally geodesic pseudo-Riemannian submanifolds is by studying fixpoints of isometries of their ambient manifold.

**Proposition 4.28.** Let $(\overline{M}, \overline{g})$ be a pseudo-Riemannian manifold and let $F \in \mathrm{Isom}(\overline{M}, \overline{g})$ be an isometry of $(\overline{M}, \overline{g})$. Suppose that a connected component $M$ of $\mathrm{Fix}(F) := \{p \in \overline{M} \mid F(p) = p\}$ is a pseudo-Riemannian submanifold of $(\overline{M}, \overline{g})$. Then $M$ is totally geodesic.

*Proof.* $F$, restricted to $M$, is the identity, i.e.

$$F|_M = \mathrm{id}_M.$$

Hence, $dF$ restricted to the subbundle $TM \subset T\overline{M}|_M$ is also the identity, meaning that $dF(v) = v$ for all $v \in T_pM \subset T_p\overline{M}$ and all $p \in M$. In order to show that $M$ is totally geodesic, it suffices to show by Proposition 4.27 (ii) that isometries map geodesics of $(\overline{M}, \overline{g})$ to geodesics of $(\overline{M}, \overline{g})$, meaning that for $\varepsilon > 0$ small enough any geodesic of $(\overline{M}, \overline{g})$, $\gamma : (-\varepsilon, \varepsilon) \to \overline{M}$ with $\gamma(0) = p \in M$, $\gamma'(0) = v \in T_pM$, will be contained in $M$ by construction. This follows from $F_*\overline{\nabla}_{\overline{X}}\overline{Y} = \overline{\nabla}_{F_*\overline{X}}F_*\overline{Y}$ for all $\overline{X}, \overline{Y} \in \mathfrak{X}(\overline{M})$, see the proof of Lemma 3.6. □

While it is easy to produce "more" examples of totally geodesic pseudo-Riemannian submanifold if we are given one by restricting to open subsets, in the case that a totally geodesic pseudo-Riemannian submanifold is geodesically complete we have the following uniqueness statement.

**Proposition 4.29.** Let $M$ and $N$ be connected totally geodesic geodesically complete pseudo-Riemannian submanifolds of $(\overline{M}, \overline{g})$. If there exists $p \in M \cap N$, such that $T_pM = T_pN$ as linear subspaces of $T_p\overline{M}$, we already have $M = N$.

*Proof.* We will show $M \subset N$, the other direction follows by symmetry of the arguments. Let $\gamma : [a, b] \to M$ be a geodesic in $M$ from $p = \gamma(a)$ to $q := \gamma(b)$. Note that $M$ being totally geodesic implies $\gamma$ is also a geodesic in $\overline{M}$. By the assumption of geodesic completeness of $N$, there exists a unique geodesic

$$\widetilde{\gamma} : \mathbb{R} \to N$$

with $\widetilde{\gamma}'(a) = \gamma'(a)$. Since $N \subset \overline{M}$ is totally geodesic, $\widetilde{\gamma}$ is also a geodesic in $\overline{M}$. Hence, by the uniqueness of maximal geodesics $\gamma$ and $\widetilde{\gamma}|_{[a,b]}$ coincide, showing in particular $q \in N$. By

Proposition 4.27 (iii) and the linear isometry property of $P_a^b(\gamma) : T_pM \to T_qM$ it follows that $T_qM = T_qN$. By the connectedness of $M$ and $N$ and Exercise 2.127 we conclude that this argument holds for all $q \in M$, showing that $q \in N$. □

Propositions 4.28 and 4.29 allow us to obtain the following.

**Corollary 4.30.** The connected totally geodesic geodesically complete Riemannian submanifolds of $\mathbb{R}^n$ with standard Riemannian metric are the affine $m \leq n$-spaces, that is smooth submanifolds of the form

$$M = p + V, \quad V \subset \mathbb{R}^n \text{ } m\text{-dimensional linear subspace.}$$

**Exercise 4.31.** Show that the connected totally geodesic geodesically complete $m \geq 1$-dimensional Riemannian submanifolds of $S^n \subset \mathbb{R}^{n+1}$ equipped with the round metric[62] are precisely the **great spheres**

$$M = S^n \cap V, \quad V \subset \mathbb{R}^{n+1} \text{ } (m+1)\text{-dimensional linear subspace.}$$

Next we specialise our studies to hypersurfaces, that is submanifolds of codimension one.

**Remark 4.32.** Let $(M, g)$ be a pseudo-Riemannian hypersurface in $(\overline{M}, \overline{g})$, that is a smooth hypersurface such that the restriction $g = \overline{g}|_{TM \times TM}$ is a pseudo-Riemannian metric on $M$. The normal bundle $TM^\perp \to M$ is of rank 1, meaning that a local frame consists of a single local vector field $\xi \in \mathfrak{X}(U)^\perp$, $U \subset M$, that is at each point orthogonal to $TM$. If $TM^\perp \to M$ is trivial, that is if there exists a global frame or equivalently a vector field $\xi \in \mathfrak{X}(M)^\perp$ spanning $TM^\perp$, $M$ is called **orientable**. This in particular means that $\xi$ is nowhere vanishing. There are in fact other equivalent definitions of orientability, one being the existence of a global volume form on $M$, cf. Remark 2.51.

**Definition 4.33.** Let $(\overline{M}, \overline{g})$ be a pseudo-Riemannian manifold with orientable pseudo-Riemannian hypersurface $(M, g)$. An orthogonal vector field $\xi \in \mathfrak{X}(M)^\perp$ is called **unit normal** if $\overline{g}(\xi, \xi) \equiv 1$ or $\overline{g}(\xi, \xi) \equiv -1$.

**Exercise 4.34.** In the setting of Definition 4.33, show that a connected hypersurface admitting a unit normal admits precisely two unit normals related by a sign flip.

In the case of $M \subset \overline{M}$ compact, one differentiates between the **outward pointing** and **inward pointing** unit normal. For hypersurfaces the Gauß equations are of a rather simple form.

**Proposition 4.35.** Let $(M, g)$ be an oriented pseudo-Riemannian hypersurface of $(\overline{M}, \overline{g})$ and let $\xi \in \mathfrak{X}(M)^\perp$ be a unit normal with $\overline{g}(\xi, \xi) \equiv \varepsilon \in \{-1, 1\}$. Then the second fundamental form of $M$ is of the form

$$\mathrm{II} = \xi \otimes \widetilde{g},$$

where $\widetilde{g} \in \Gamma(\mathrm{Sym}^2(T^*M))$ is given by

$$\widetilde{g}(X, Y) = \varepsilon g(S^\xi X, Y)$$

for all $X, Y \in \mathfrak{X}(M)$ with $S^\xi$ the Weingarten map. The Gauß equations for the curvature (4.4) and the sectional curvature equation (4.5) are given by

$$\overline{g}(\overline{R}(X, Y)Z, W) = g(R(X, Y)Z, W) + \varepsilon(g(S^\xi X, Z)g(S^\xi Y, W) - g(S^\xi Y, Z)g(S^\xi X, W)),$$
$$\overline{K}(v, w) = K(v, w) - \varepsilon \frac{g(S^\xi v, v)g(S^\xi w, w) - g(S^\xi v, w)^2}{g(v, v)g(w, w) - g(v, w)^2},$$

for all $X, Y, Z, W \in \mathfrak{X}(M)$ and all $v, w \in T_pM$ spanning a nondegenerate plane.

[62] a.k.a. the restriction of the standard Riemannian metric

*Proof.* The fact that $\xi$ is by assumption nowhere vanishing implies that we can write $\mathrm{II} = \xi \otimes \widetilde{g}$. Furthermore, $\overline{g}(\mathrm{II}(X,Y),\xi) = \varepsilon\widetilde{g}(X,Y)$ for all $X, Y \in \mathfrak{X}(M)$ means that $\widetilde{g}$ is uniquely determined. By the Weingarten equation (4.2) we obtain

$$\overline{g}(\mathrm{II}(X,Y),\xi) = g(S^\xi X, Y)$$

for all $X, Y \in \mathfrak{X}(M)$ and, hence, $\widetilde{g}(X,Y) = \varepsilon g(S^\xi X, Y)$ as claimed. For the other two equations in this proposition observe that

$$\overline{g}(\mathrm{II}(X,Z),\mathrm{II}(Y,W)) = \varepsilon\widetilde{g}(X,Z)\widetilde{g}(Y,W) = \varepsilon g(S^\xi X, Z)g(S^\xi Y, W),$$

for all $X, Y, Z, W \in \mathfrak{X}(M)$, which follows from $\varepsilon = \varepsilon^{-1}$ and our previous results. The rest of this proof is just writing out the formulas (4.4) and (4.5) and is left as an exercise for the reader. $\square$

**Example 4.36.** Let $(M, g)$ be an orientable Riemannian hypersurface in $(\mathbb{R}^{n+1}, \langle\cdot,\cdot\rangle)$. For $p \in M$ and $U \subset \mathbb{R}^{n+1}$ a small enough open neighbourhood of $p$, choose $f \in C^\infty(U)$ of maximal rank, such that

$$M \cap U = \{f = 0\}.$$

After a possible overall sign flip of $f$, we can assume without loss of generality that the unit normal of $M$ is given locally on $M \cap U$ by

$$\xi = \frac{\mathrm{grad}_{\langle\cdot,\cdot\rangle}(f)}{\sqrt{\langle \mathrm{grad}_{\langle\cdot,\cdot\rangle}(f), \mathrm{grad}_{\langle\cdot,\cdot\rangle}(f)\rangle}} = \frac{\mathrm{grad}_{\langle\cdot,\cdot\rangle}(f)}{\|\mathrm{grad}_{\langle\cdot,\cdot\rangle}(f)\|}.$$

The second fundamental form of $M \cap U$ is then given by

$$\mathrm{II}(X,Y) = -\frac{1}{\|\mathrm{grad}_{\langle\cdot,\cdot\rangle}(f)\|}\overline{\nabla}^2 f(X,Y)$$

for all $X, Y \in \mathfrak{X}(M \cap U)$, where $\overline{\nabla}^2 f$ denotes the Hessian of $f$, cf. Definition 2.95.

## References


[A1] V.I. Arnold, *Ordinary Differential Equations*, Springer Universitext (1992).

[A2] V.I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer GTM **60** (1978).

[Bae] C. Bär, *Differential Geometry*, lecture notes (2013).

[Bes] A. L. Besse, *Einstein Manifolds* (1987), Ergebnisse der Mathematik und ihrer Grenzgebiete Vol. **10**.

[BGP] C. Bär, N. Ginoux, and F. Pfäffle, *Wave Equations on Lorentzian Manifolds and Quantization*, ESI Lectures in Mathematics and Physics.

[Bre] G.E. Bredon, *Sheaf Theory*, second edition, Springer GTM **170** (1997).

[Bro] L.E.J. Brouwer, *Beweis der Invarianz des n-dimensionalen Gebiets*, Math. Ann. **71**, 305–313 (1911).

[CK] B. Chow and D. Knopf, *The Ricci Flow: An Introduction* (2004), AMS Mathematical Surveys and Monographs Vol. **110**.

[G] O. Goertsches, *Differentialgeometrie*, lecture notes (2014) (in German).

[J] J. Jost, *Riemannian Geometry and Geometric Analysis*, 7th edition (2017), Springer Universitext.

[KN] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* **Vol. I**, Wiley Classics Library (1996).

[L1] J.M. Lee, *Riemannian Manifolds – An Introduction to Curvature*, Springer GTM 176 (1997).

[L2] J.M. Lee, *Introduction to Smooth Manifolds*, Springer GTM 218 (2003).

[M] J. Milnor, *Analytic proofs of the "hairy ball theorem" and the Brouwer fixed-point theorem*, The American Mathematical Monthly, Vol. **85** (1978).

[N] J. Nash, *The Imbedding Problem for Riemannian Manifolds*, Annals of Mathematics, Second Series, Vol. **63**, No. **1**.

[O] B. O'Neill, *Semi-Riemannian Geometry With Applications to Relativity* (1983), Pure and Applied Mathematics, Vol. **103**, Academic Press, NY.

[R] W. Rudin, *Principles of Mathematical Analysis* (1953), McGraw-Hill Education, 3rd edition.

[Sc] A. Scorpan, *The Wild World of 4-Manifolds* (2005), AMS.

[Sh] R.W. Sharpe, *Differential Geometry: Cartan's Generalization of Klein's Erlangen Program*, Springer GTM **166** (1997).

[So] A. Solecki, *Finite atlases on manifolds* (1974), Commentat. Math. Vol. **17**, No **2**.

[SS] L.A. Steen, J.A. Seebach Jr., *Counterexamples in Topology* (1995), second edition, Dover Books on Mathematics.

[T] P. M. Topping, *Lectures on the Ricci flow* (2006), L.M.S. Lecture note series **325** C.U.P.

[Z] M. Zedek, *Continuity and Location of Zeroes of Linear Combinations of Polynomials*, Proc. Amer. Math. Soc. **16** (1965).

Put exercises at the end of each chapter, not necessarily from exercise sheets.

## Todo list