Czech Technical University in Prague,
Faculty of Electrical Engineering

Bachelor's thesis

Connections on Differentiable Manifolds

Author
M. Šimák
Department of Radio Engineering

Supervisor
doc. RNDr. J. Velebil, Ph.D.
Department of Mathematics

May 2022

[This page intentionally left blank.]

BACHELOR‘S THESIS ASSIGNMENT

I. Personal and study details

Student's name: **Šimák Martin** Personal ID number: **483523**

Faculty / Institute: **Faculty of Electrical Engineering**

Department / Institute: **Department of Radioelectronics**

Study program: **Open Electronic Systems**

Branch of study: **Open Electronic Systems**

II. Bachelor’s thesis details

Bachelor’s thesis title in English:

Connections on Differentiable Manifolds

Bachelor’s thesis title in Czech:

Konexe na diferencovatelných varietách

Guidelines:

The shape of geometry of a differentiable manifold is given only after specifying an affine connection, or, equivalently, a covariant derivative. In the thesis, the basic ways of endowing a manifold with this additional structure will be investigated. The core of the thesis will focus on the relationship among various ways of specifying a connection.
The style and presentation of the thesis will be theoretical, maintaining mathematical standards.

Bibliography / sources:

[1] Theodore Frankel, The geometry of physics: An introduction, 3rd ed, Cambridge University Press, 2011.
[2] Jeffrey M. Lee, Manifolds and differential geometry, American mathematical Society, 2009.
[3] Jozsef Szilasi, David Cs Kertesz, Rezso L. Lovas, Connections, sprays and Finsler structures, World Scientific Publishing Company, 2013.

Name and workplace of bachelor’s thesis supervisor:

doc. RNDr. Jiří Velebil, Ph.D. Department of Mathematics FEE

Name and workplace of second bachelor’s thesis supervisor or consultant:

Date of bachelor’s thesis assignment: **19.01.2022** Deadline for bachelor thesis submission: **20.05.2022**

Assignment valid until: **30.09.2023**

doc. RNDr. Jiří Velebil, Ph.D.	doc. Ing. Stanislav Vítek, Ph.D.	prof. Mgr. Petr Páta, Ph.D.
Supervisor’s signature	Head of department’s signature	Dean’s signature

III. Assignment receipt

The student acknowledges that the bachelor’s thesis is an individual work. The student must produce his thesis without the assistance of others, with the exception of provided consultations. Within the bachelor’s thesis, the author must state the names of consultants and include a list of references.

.

Date of assignment receipt Student’s signature

Declaration

I declare that I completed the presented thesis independently and that all used sources are quoted in accordance with the Methodological instructions that cover the ethical principles for writing an academic thesis.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze, ...

....
Martin Šimák

Acknowledgements

First and foremost, I would like to thank my supervisor, Jiří Velebil, for his outstanding guidance and his zeal throughout the creation of this thesis. Further, I would like to thank my family and my colleagues, namely Erik Rapp and Karolína Veselá, for their invaluable support.

Poděkování

Především bych rád poděkoval svému vedoucímu, Jiřímu Velebilovi, za jeho vynikající vedení a za jeho zapálenost a vstřícnost během tvorby této práce. Dále bych rád poděkoval své rodině a svým kolegům, zejména Eriku Rappovi a Karolíně Veselé, za jejich neocenitelnou podporu.

Title:
Connections on Differentiable Manifolds

Author: Martin Šimák

Study programme: Open Electronic Systems

Supervisor: doc. RNDr. Jiří Velebil, Ph.D., Department of Mathematics FEE

Abstract:
One of the aims of this thesis is to fully introduce a comprehensive differentiable structure on a smooth manifold. To achieve that, we first inspect various aspects of its definition. Further, we introduce a tangent structure on said manifolds, allowing us to locally speak of derivatives and vector fields. This construction naturally results in definitions of a covariant derivative and a connection form, where the former serves as a generalisation of a derivative in a direction and the latter as a tool for "glueing" tangent spaces together. As a climax of the thesis, we show that the two previously mentioned additional structures on the ambient manifold are equivalent.

Keywords: smooth manifolds, tangent structure, covariant derivative, connection

Název práce:
Konexe na diferencovatelných varietách

Autor: Martin Šimák

Studijní program: Otevřené elektronické systémy

Vedoucí: doc. RNDr. Jiří Velebil, Ph.D., katedra matematiky FEL

Abstrakt:
Jedním z cílů této práce je v plném měřítku uvést rozsáhlou diferenciální strukturu na hladké varietě. Abychom toho dosáhli, prozkoumáme nejprve jednotlivé aspekty její definice. Dále na varietách uvedeme tečnou strukturu, jenž nám umožňuje lokálně hovořit o derivacích a vektorových polích. Tato konstrukce přirozeně vyúsťuje v definice kovariantní derivace a konexe, přičemž první z pojmů slouží jakožto zobecnění derivace ve směru a druhý jakožto nástroj pro „slepování“ tečných prostorů. Jako vyvrcholení práce ukážeme, že tyto dvě dodatečné struktury na varietě jsou ekvivalentní.

Klíčová slova: hladké variety, tečná struktura, kovariantní derivace, konexe

Contents

Introduction

In the year 1905, Albert Einstein published his renowned paper titled *"On the Electrodynamics of Moving Bodies"* in which he reconciled *Newton's laws of motion* with electrodynamics and thus laid the foundation for a whole new era in physics. In this paper—and further in the following development of the *general theory of relativity* in 1915—humanity had begun to understand that our universe was not flat.

One of the major discoveries Einstein's relativity brings is that spacetime has a non-trivial curvature that can be directly related to the energy and momentum of matter and radiation. This discovery naturally inspired a great number of mathematicians to find a new tool that would elegantly grasp the nature of the universe. These efforts led to a rapid development of *differential geometry*. Such an outburst of interest in a fairly new branch of mathematics—the notion of a *topological space* was distilled by Felix Hausdorff only in 1914—can be at least partially attributed to this phenomenal discovery.

However, it is worth mentioning that although Einstein's work was definitely a massive inspiration for academic contributions in the area of calculus on curved geometries, his observation was definitely not the first. In this regard, we can mention that in the 19th century, William Kingdon Clifford in his philosophical writings coined the expression of *mindstuff*, i.e., a geometry underlying the fabric of space and the curvature of which manifests itself as gravity. The rejection of a purely flat geometry reaches even further into the history as the renowned mathematician Pierre-Simon Laplace attempted to find the curvature of space already at the turn of the 19th century. Another example of earlier thoughts on the fundamental hypotheses of geometry is the classic habilitation dissertation *On the hypotheses which lie at the bases of geometry* by Bernhard Riemann from 1868, translated by William Kingdon Clifford in 1873.

Synopsis of the thesis. In **Chapter 1**, we aim to explore the possible differential structure of *smooth manifolds*. Although this part could be considered to be standard and left out as a pre-requisite knowledge, we prefer to inspect the introduction of the smooth manifold delicately to avoid any confusion.

Chapter 2 introduces a *tangent structure* on a manifold. This construction allows us to locally define notions such as *derivatives* and *vector fields*. While all of the aforementioned should evoke an elementary course of multivariable calculus, on smooth manifolds, we will need to be more careful when defining them.

Further on, in **Chapter 3**, we finally introduce the notions of a covariant derivative and a connection. In this chapter, we focus on the development of intuition on simple examples rather than a proper definition. The abstraction is instilled in **Chapter 4** where we take the conception introduced in the previous chapter and we put it into a more general setting. In this chapter, we also further explore the notions' properties and additional constructions.

Ultimately, in **Chapter 5**, we discuss the main result of this thesis: the equivalence of a covariant derivative and a connection form. Apart from that, we also present some directions of possible generalisation of the thesis.

Although the results of Chapters 4 and 5 are known, the exposition and most of the proof-work was done by me.

1. Introduction to smooth manifolds

In the first chapter of this thesis, we introduce the concept of a smooth manifold. To readers experienced in topology, definition of the aforementioned might come across as trivial, but we consider it very important to specify how exactly we approach manifolds in general, since the underlying definitions might differ from text to text.

Apart from the smooth manifold itself, we will also introduce other key concepts of differential topology, such as charts, atlases, paracompactness, etc.

Our presentation of these notions follows closely the exposition given in the book [7].

1.1 Charts and atlases

Before we begin, let us discuss the goal of this section. We would like to introduce a structure on a set that endows it with the property of "looking like" $\mathbb{R}^n$ locally. In general, this will not be necessarily achievable globally on the whole set but we will manage to arrive at a reasonable compromise using charts and atlases.

Definition 1.1.1. Let M be a set. A *chart* on M is a pair $(U, \mathbf{x})$ consisting of $U \subseteq M$ and an injective map $\mathbf{x} : U \to \mathbb{R}^n$ such that $\mathbf{x}[U] = \{\mathbf{x}(a) \mid a \in U\}$ is an open set in $\mathbb{R}^n$. The composite $x^i = \mathrm{pr}_i \circ \mathbf{x} : M \to \mathbb{R}$ is often called the *i-th coordinate function.*

Note 1.1.2. Since $\mathbf{x}$, by definition, has $\mathbb{R}^n$ as a codomain, we can always write its action upon a point $p \in M$ as $\mathbf{x} : p \mapsto \mathbf{x}(p) = \big(x^1(p), \ldots, x^n(p)\big)$. Whenever we wish to emphasise the individual coordinate functions, we write $\big(x^1, \ldots, x^n\big)$ or $\big(x^i\big)$ instead of $\mathbf{x}$.

Remark 1.1.3. Note that we have defined charts on a plain set M. Therefore, we do not require U in $(U, \mathbf{x})$ to be open since we do not have the topological structure on M just yet.

Definition 1.1.4. A collection $\mathcal{A} \coloneqq \{(U_\alpha, \mathbf{x}_\alpha) \mid \alpha \in A\}$ of $\mathbb{R}^n$-valued charts on a given set M is called a *smooth* ($\mathbb{R}^n$*-valued*) *atlas* on M, if the following three conditions hold:

(a) $\bigcup_{\alpha \in A} U_\alpha = M$, i.e., the collection $\{U_\alpha \mid \alpha \in A\}$ covers M.

(b) For every $\alpha, \beta \in A$, the sets $\mathbf{x}_\alpha[U_\alpha \cap U_\beta] = \{\mathbf{x}_\alpha(a) \mid a \in U_\alpha \cap U_\beta\}$ are open in $\mathbb{R}^n$.

(c) Whenever $U_\alpha \cap U_\beta \neq \emptyset$, the map $\mathbf{x}_\beta \circ \mathbf{x}_\alpha^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is smooth (i.e., it has all derivatives). We also say that $(U_\alpha, \mathbf{x}_\alpha), (U_\beta, \mathbf{x}_\beta)$ are *smooth-related.*

Remark 1.1.5. The above notation $\mathbf{x}_\beta \circ \mathbf{x}_\beta^{-1}$ is a shorthand for the precise, yet clumsy,

$$\mathbf{x}_\beta\restriction_{U_\alpha \cap U_\beta} \circ \mathbf{x}_\alpha^{-1}\restriction_{\mathbf{x}_\alpha[U_\alpha \cap U_\beta]},$$

where $f\restriction_A$ denotes the standard domain restriction of f to A.

Definition 1.1.6. Suppose $\mathcal{A}_1$ and $\mathcal{A}_2$ are smooth $\mathbb{R}^n$-valued atlases on a set M. Define

$$\mathcal{A}_1 \equiv \mathcal{A}_2 \quad \iff \quad \mathcal{A}_1 \cup \mathcal{A}_2 \text{ is a smooth } \mathbb{R}^n\text{-valued atlas on } M.$$

Claim 1.1.7. *Let $\equiv$ be the relation defined above. Then the following hold:*

(i) *$\equiv$ is an equivalence relation on the set of all smooth $\mathbb{R}^n$-valued atlases on a set M.*

(ii) *Given an atlas $\mathcal{A}$, the union $\bigcup\{\mathcal{B} \mid \mathcal{B} \equiv \mathcal{A}\}$ is an atlas again. Moreover, this atlas is equivalent to $\mathcal{A}$. Let us call it the* maximal atlas *generated by $\mathcal{A}$ and denote it $\mathcal{A}^{\max}$.*

(iii) *Let $[\mathcal{A}]_{\equiv} = \{\mathcal{B} \mid \mathcal{B} \equiv \mathcal{A}\}$ be the equivalence class of an atlas $\mathcal{A}$ under the relation $\equiv$. The correspondence $[\mathcal{A}]_{\equiv} \mapsto \mathcal{A}^{\max}$ is a bijection.*

Proof. Ad (i): For a relation to be an equivalence, it must hold that the relation is reflexive, symmetric and transitive.

(a) For reflexivity, we need to show that $\mathcal{A} \equiv \mathcal{A}$. This fact is trivial, since $\mathcal{A}$ is a smooth atlas and $\mathcal{A} \cup \mathcal{A} = \mathcal{A}$ holds, thus $\mathcal{A} \equiv \mathcal{A}$ is obvious.

(b) Symmetry means that $\mathcal{A}_1 \equiv \mathcal{A}_2$ if and only if $\mathcal{A}_2 \equiv \mathcal{A}_1$. This follows naturally from the symmetry of the relation $\cup$.

(c) Transitivity: if $\mathcal{A}_1 \equiv \mathcal{A}_2$ and $\mathcal{A}_2 \equiv \mathcal{A}_3$, then $\mathcal{A}_1 \equiv \mathcal{A}_3$. This fact is easily checked by going through the defining properties.

Ad (ii): First, we should note that $\mathcal{A}^{\max}$ surely is an atlas in the first place. This is trivially guaranteed by the fact that $\mathcal{A}^{\max}$ is a union of elements that are already atlases themselves. Next, let us verify that $\mathcal{A}^{\max}$ is equivalent to $\mathcal{A}$. Thanks to the reflexivity of $\equiv$, it holds that $\mathcal{A} \subseteq \mathcal{A}^{\max}$, i.e., $\mathcal{A} \cup \mathcal{A}^{\max} = \mathcal{A}^{\max}$, which is an atlas based on the preceding discussion.

Ad (iii): The bijective correspondence follows easily from the fact that both $[\mathcal{A}]_{\equiv}$ and $\mathcal{A}^{\max}$ are defined by collecting atlases equivalent to $\mathcal{A}$ and their unions, respectively. ■

Definition 1.1.8. Let $(U, \mathbf{x})$ be a chart on M and let $\mathcal{A}$ be a smooth atlas on M. We say that $(U, \mathbf{x})$ is *compatible* with $\mathcal{A}$, provided that $\mathcal{A} \cup \{(U, \mathbf{x})\}$ is an atlas on M.

Claim 1.1.9. *Let $\mathcal{A} = \{(U_\alpha, \mathbf{x}_\alpha) \mid \alpha \in A\}$ be a smooth $\mathbb{R}^n$-valued atlas on a set M. Then the following holds:*

(i) *Let $(U, \mathbf{x})$ and $(V, \mathbf{y})$ are charts compatible with $\mathcal{A}$, and such that $U \cap V \neq \emptyset$. Then the charts $(U \cap V, \mathbf{x}\restriction_{U\cap V})$ and $(U \cap V, \mathbf{y}\restriction_{U\cap V})$ are compatible with $\mathcal{A}$, hence they both belong to the maximal atlas generated by $\mathcal{A}$.*

(ii) *Let $(U, \mathbf{x})$ be a chart compatible with $\mathcal{A}$, let O be an open subset of $\mathbf{x}[U]$ and let V denote $\mathbf{x}^{-1}[O]$. Then $(V, \mathbf{x}\restriction_V)$ is a chart compatible with $\mathcal{A}$.*

Proof.

(i) For $(U \cap V, \mathbf{x}\restriction_{U\cap V})$ to be compatible with $\mathcal{A}$, it must hold that $\mathcal{A} \cup \{(U \cap V, \mathbf{x}\restriction_{U\cap V})\}$ is an atlas over M, i.e., that it satisfies the conditions (a)—(c) in Definition 1.1.4.

   (a) The collection $\{U_\alpha \mid \alpha \in A\} \cup \{U \cap V\}$ obviously covers M.

(b) For every $\alpha \in A$, it holds that the sets $\mathbf{x}_\alpha[U_\alpha \cap U]$ and $\mathbf{x}_\alpha[U_\alpha \cap V]$ are open in $\mathbb{R}^n$. Further, since

$$\mathbf{x}_\alpha[U_\alpha \cap U \cap V] = \mathbf{x}_\alpha[U_\alpha \cap U] \cap \mathbf{x}_\alpha[U_\alpha \cap V]$$

holds and since finite intersections of open sets are open, we immediately obtain that $\mathbf{x}_\alpha[U_\alpha \cap U \cap V]$ is open in $\mathbb{R}^n$, for every $\alpha \in A$. Hence, the chart $(U \cap V, \mathbf{x}\restriction_{U\cap V})$ is compatible with $\mathcal{A}$.

(c) Whenever $U_\alpha \cap U \cap V \neq \emptyset$, the maps $\mathbf{x}\restriction_{U\cap V} \circ \mathbf{x}_\alpha^{-1}$ and $\mathbf{x}_\alpha \circ \mathbf{x}\restriction_{U\cap V}^{-1}$ are smooth. This condition holds trivially, since both restrictions and compositions of smooth maps are also smooth.

Since for $(U \cap V, \mathbf{y}\restriction_{U\cap V})$, the reasoning is analogous, we consider this proved.

(ii) For the second part, we construct a very similar proof to the above. Let us show that $\mathcal{A} \cup \{(V, \mathbf{x}\restriction_V)\}$ is an atlas:

(a) The collection $\{U_\alpha \mid \alpha \in A\} \cup \{V\}$ covers M for obvious reasons.

(b) For each $\alpha \in A$, the sets $\mathbf{x}_\alpha[U_\alpha \cap V]$ and $\mathbf{x}\restriction_V [U_\alpha \cap V]$ are open in $\mathbb{R}^n$ due to the fact that $(U, \mathbf{x})$ is compatible with $\mathcal{A}$ and that V is the inverse image of an open set in $\mathbb{R}^n$ under $\mathbf{x}$.

(c) Whenever $U_\alpha \cap V \neq \emptyset$, the maps $\mathbf{x}\restriction_V \circ \mathbf{x}_\alpha^{-1}$ and $\mathbf{x}_\alpha \circ \mathbf{x}\restriction_V^{-1}$ are smooth because we are once again dealing with restrictions and compositions of smooth maps.

■

1.2 Topology

Now, let us turn our focus to topology. To introduce our desired construction—a smooth manifold—we will have to review some of the basic definitions such as the topology and open sets, basis, etc. Apart from these elementary notions, we will also demand that our smooth manifold is also Hausdorff and paracompact which are delicate properties we will talk about later. An excellent survey of general topology is the book [3].

Definition 1.2.1. Let X be a set and let τ be a family of subsets of X such that the following conditions hold:

(a) $\emptyset \in \tau$ and $X \in \tau$.

(b) If $U_1 \in \tau, \ldots, U_n \in \tau$, then $U_1 \cap \cdots \cap U_n \in \tau$.

(c) If $U_i \in \tau$, $i \in I$, then $\bigcup_{i\in I} U_i \in \tau$.

Such family τ is called *topology*, elements of τ are called *open sets* and we say that the pair (X, τ) forms a *topological space*.

Remark 1.2.2. The notion of a topological space is a vast generalisation of the familiar notion of a metric space.

More in detail: any metric space (X, d) becomes a topological space (X, τ_d) by declaring $U \subseteq X$ to be in τ_d (i.e., declaring U to be open in (X, τ_d)) whenever the following holds:

given $p \in U$, there exists $r > 0$ such that $B_r(p) = \{x \in X \mid d(p,x) < r\} \subseteq U$.

Then it is easy to prove that τ_d is indeed a topology on X.

Hence, the notion of a topological space axiomatises the well-known property of open sets in a metric space: an open set is a set, where one can "wiggle around" any of the points a bit without leaving the open set.

Of course, we cannot talk about any notion of distance in a general topological space. In fact, it is easy to come up with an example of a topological space (X, τ) such that $\tau \neq \tau_d$ for any metric d on X. Consider, for example, $X = \{a, b\}$ and $\tau = \{\emptyset, X\}$.

Definition 1.2.3. A collection $\mathcal{B}$ of subsets of a set X is called the *basis*, if the following hold:

(a) $\mathcal{B}$ covers X, i.e., $\bigcup_{B \in \mathcal{B}} B = X$.

(b) If $B_1 \in \mathcal{B}$, $B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, then there exists a basis element $B \in \mathcal{B}$ such that $x \in B \subseteq B_1 \cap B_2$.

Claim 1.2.4. *Let $\mathcal{B}$ be a basis of subsets of X. Then the collection τ of arbitrary unions of elements of $\mathcal{B}$ forms a topology on X. This topology is called the topology* generated *by $\mathcal{B}$.*

PROOF. Using the defining properties of a basis, we shall verify that the conditions (a)–(c) of Definition 1.2.1 indeed hold when taking unions of the elements of $\mathcal{B}$.

(a) Both $\emptyset$ and all of X can be formed as specific unions of elements of $\mathcal{B}$. The former can be constructed as a union of zero elements and the latter by taking the union of all elements of $\mathcal{B}$ which makes up the whole of X, thanks to (a) in Definition 1.2.3.

(b) Any arbitrary finite intersection can be constructed by repeated application of the second property of basis elements.

(c) This condition holds trivially as we take all possible unions into account.

■

Claim 1.2.5. *If $\mathcal{A} = \{(U_\alpha, \mathbf{x}_\alpha) \mid \alpha \in A\}$ is a maximal smooth atlas on M, then the collection $\{U_\alpha \mid \alpha \in A\}$ is a basis for a topology on M (called the* topology induced by *$\mathcal{A}$).*

PROOF. First, we will show that $\{U_\alpha \mid \alpha \in A\}$ forms a basis on M. The first condition of covering M holds immediately when reviewing the definition of an atlas. The second condition is assured by Claim 1.1.9. Now that have a basis on M, we can generate a topology on M by taking arbitrary unions of elements of the basis as we have seen in Claim 1.2.4. ■

Claim 1.2.6. *Let M be a set equipped with a maximal atlas $\mathcal{A} = \{(U_\alpha, \mathbf{x}_\alpha) \mid \alpha \in A\}$ and a topology τ induced by $\mathcal{A}$. This topology can be characterised as follows: $V \subseteq M$ is open if and only if $\mathbf{x}_\alpha[U_\alpha \cap V]$ is an open subset of $\mathbb{R}^n$ for all charts $(U_\alpha, \mathbf{x}_\alpha)$ of $\mathcal{A}$.*

PROOF. To prove the logical equivalence used in the body of the claim, we shall break it down into two implications:

- " $\Longrightarrow$ ": Let $V \subseteq M$ be an open set within the topology τ. Following Claim 1.2.5, it can be written as $V = \bigcup_{\alpha \in I} U_\alpha$ for some $(U_\alpha, \mathbf{x}_\alpha)$ in $\mathcal{A}$. Therefore $\mathbf{x}_\alpha[U_\alpha \cap V]$ is indeed an open subset of $\mathbb{R}^n$ for every $\alpha \in A$ from the definition of an atlas.

- " $\Longleftarrow$ ": Let $\mathbf{x}_\alpha[U_\alpha \cap V]$ be an open subset of $\mathbb{R}^n$ for all charts $(U_\alpha, \mathbf{x}_\alpha)$ of $\mathcal{A}$. This implication is a simple consequence of what we have seen in Claim 1.1.9. For every $O_\alpha = \mathbf{x}_\alpha[U_\alpha \cap V]$, we get an inverse image $V_\alpha = \mathbf{x}_\alpha^{-1}[O_\alpha]$, which, according to the referred claim, is a chart when accompanied by the restriction $\mathbf{x}_\alpha\restriction_{V_\alpha}$ and thus belongs to the maximal atlas. If we do this to all the intersections in $\mathbb{R}^n$, we get a collection of open sets in M under the topology τ, which was generated by the atlas $\mathcal{A}$. However, the topology τ was naturally created by taking arbitrary unions of elements of $\{U_\alpha \mid \alpha \in A\}$. This indeed means that V as a union of V_α belongs to τ as well and therefore is an open set.

$\blacksquare$

As we said at the beginning, a manifold should locally "look like" $\mathbb{R}^n$. Now we have the tools to say it precisely.

Definition 1.2.7. A topological space (X, τ) is called *locally Euclidean* if each point $x \in X$ has an open neighbourhood that is homeomorphic to an open subset of $\mathbb{R}^n$.

Note 1.2.8. The above definition of a locally Euclidean space assumes the definition of a homeomorphism as a mapping $f : X \to Y$ between two topological spaces (X, τ) and (Y, σ), which is bijective and continuous both ways, i.e., both f and f^{-1} are continuous. In topology, continuous maps "reflect" openness, i.e., if $V \subseteq Y$ is an open set under the topology σ on Y, then the inverse image $f^{-1}(V) \subseteq X$ is an open set under the topology τ on X.

Remark 1.2.9. The definition of a smooth manifold that we are heading towards should formalise the notion of a topological space which locally "looks like" $\mathbb{R}^n$. There are spaces, however, that locally "look like" $\mathbb{R}^n$, yet they do not have an important property of a Euclidean space.

Consider the space $X = (\mathbb{R} \times \{0\}) \cup (\mathbb{R} \times \{1\})$ with the topology τ of a subspace of $\mathbb{R}^2$ (the space (X, τ) consists of two parallel lines in the plane). Now define an equivalence relation $\sim$ on X by putting $(x, 0) \sim (x, 1)$ for all $x \in \mathbb{R} \setminus \{0\}$.

If we endow the quotient set $X/_\sim$ with the quotient topology[1] $\tau_\sim$ (i.e., declare $U \subseteq X/_\sim$ open, whenever $\pi^{-1}(U)$ is open in X, where $\pi : X \to X/_\sim$ is the quotient map), then the space $(X/_\sim, \tau_\sim)$ is not convenient for us since e.g., the points $[(0, 0)]_\sim$ and $[(0, 1)]_\sim$ cannot be distinguished in $(X/_\sim, \tau_\sim)$ by disjoint open sets, yet every point of $X/_\sim$ has an open neighbourhood that is homeomorphic to $\mathbb{R}$.

Thus, the space $(X/_\sim, \tau_\sim)$ from the above remark lacks the property that any pair of its distinct points can be "separated" by two disjoint open sets. In a slogan, such a property means that any two distinct points are "quite far away". We formulate the property precisely now.

Definition 1.2.10. A topological space (X, τ) is called *Hausdorff* (also that it *satisfies the T_2 axiom*), if for all $x \neq y$ in X there exist disjoint open sets U, V such that $x \in U$ and $y \in V$.

[1] Since the notion of a quotient topology serves as a mere example in this thesis, we do not elaborate on it any further. For more information about quotient topologies, see e.g., [3].

Thus, Remark 1.2.9 gives an example of a locally Euclidean space that is not Hausdorff. It is quite easy to see why such property is convenient for us. For a topological space to be Hausdorff it simply means that we can make a clear distinction (via separation by open neighbourhoods) between two different points of that space. This proves especially convenient when treating further notions such as unambiguity of a limit.

There are two other desirable properties a manifold should have: it should be second countable and/or paracompact. Let us mention that these additional properties are there for comfort only: non-paracompact manifolds are studied as well, see [5].

Definition 1.2.11. A topological space (X, τ) is called *second countable*, if τ has an at most countable basis.

Definition 1.2.12. Let (X, τ) be a topological space.

- A collection $\{U_i \mid i \in I\}$ is an *open cover* of X, if $\bigcup_{i\in I} U_i = X$ and every U_i is an open set.
- An open cover $\{V_j \mid j \in J\}$ is a *refinement* of an open cover $\{U_i \mid i \in I\}$, if for every $j \in J$ there exists an $i \in I$ such that $V_j \subseteq U_i$.
- A cover $\{U_i \mid i \in I\}$ is called *locally finite*, provided that for every $x \in X$, there exists an open set U such that $x \in U$ and the set $\{i \in I \mid U \cap U_i \neq \emptyset\}$ is finite.

Definition 1.2.13. A topological space (X, τ) is called *paracompact*, if the following holds: every open cover of X has a locally finite refinement to an open cover.

Roughly speaking: paracompactness of a manifold ensures that one can perform certain constructions locally and then "glue" them together in a non-conflicting, pleasing way. This is particularly useful when defining, e.g., *integration* on a manifold. We do not develop integral calculus in this text. For details, see, e.g., [7].

The following result shows why second countability and paracompactness are also topologically desirable. We state the theorem without proof. For details, we refer to [3].

Theorem 1.2.14. *Suppose (X, τ) is a Hausdorff, locally Euclidean space. Then the following conditions are equivalent:*

(i) *(X, τ) is paracompact.*

(ii) *(X, τ) is a metrisable space.*

(iii) *Each connected component of (X, τ) is second countable.*

(iv) *Each connected component of (X, τ) is σ-compact (i.e., it is a union of at most countably many compact spaces).*

(v) *Each connected component of (X, τ) is separable (i.e., it contains an at most countable dense subset).*

As the final step towards our definition of a smooth manifold, we formulate the following claim which shows how to attain the "nice" properties we have defined earlier (being Hausdorff, second countable and paracompact) via specific requirements on our smooth structure.

Claim 1.2.15. *Let $\mathcal{A}$ be a smooth atlas on a set M. Then the following holds:*

(i) *Suppose that, for any $p \neq q$ in M, we have either charts $(U_\alpha, \mathbf{x}_\alpha)$, $(U_\beta, \mathbf{x}_\beta)$ with $U_\alpha \cap U_\beta = \emptyset$ and $p \in U_\alpha$, $q \in U_\beta$, or there exists a chart $(U, \mathbf{x})$ such that $p \in U$ and $q \in U$. Then the topology on M given by $\mathcal{A}$ is Hausdorff.*

(ii) *If $\mathcal{A}$ is countable (or if it has a countable subatlas), then the topology τ on M induced by $\mathcal{A}$ is second countable.*

(iii) *If the collection $\{U_\alpha \mid \alpha \in A\}$ of chart domains of $\mathcal{A}$ is such that for every $\alpha_0 \in A$, the set $\{\alpha \in A \mid U_\alpha \cap U_{\alpha_0} \neq \emptyset\}$ is at most countable, then the topology on M induced by $\mathcal{A}$ is paracompact.*

PROOF.

(i) Since the first condition indirectly copies the definition of a Hausdorff space, we will only tend to the latter one. If there exists a chart $(U, \mathbf{x})$ such that both p and q belong to U, then due to the injectivity of $\mathbf{x}$, we obtain $\mathbf{x}(p) \neq \mathbf{x}(q)$. Since $\mathbb{R}^n$ is a Hausdorff space, we can separate by open sets: $\mathbf{x}(p) \in U_p$ and $\mathbf{x}(q) \in U_q$, where $U_p \cap U_q = \emptyset$ and both U_p and U_q are open. According to Claim 1.1.9, we can then follow the inverse images $\mathbf{x}^{-1}[U_p]$ and $\mathbf{x}^{-1}[U_q]$ and consider them open since they form compatible chart domains. Once again, thanks to the injectivity of $\mathbf{x}$, we have two disjoint open sets separating p and q.

(ii) This claim is trivial since the topology τ is generated by taking arbitrary unions of chart domains from $\mathcal{A}$. Therefore the set τ contains a countable amount of basis elements.

(iii) We refer to [7] for the full (rather technical) proof.

■

Now, we arrive at the point where we know what all of our desirable properties really mean and we are ready to define the final concept of this chapter—the definition of a smooth manifold.

Definition 1.2.16. An *n-dimensional smooth manifold* is a pair $(M, \mathcal{A})$, where M is a set and $\mathcal{A}$ is a smooth ($\mathbb{R}^n$-valued) atlas on M, such that the topology on M induced by $\mathcal{A}$ is Hausdorff and paracompact.

Note 1.2.17. An n-dimensional smooth manifold is often referred to as a smooth n-manifold. It is also common to call just the set M a manifold when specification of the atlas is unnecessary for the sake of brevity.

Definition 1.2.18. Let M and N be smooth manifolds, and let $F : M \to N$ be a map. We say that F is *smooth* if for every $p \in M$ there exist smooth charts $(U, \mathbf{x})$ containing p and $(V, \mathbf{y})$ containing $F(p)$ such that $F[U] \subseteq V$ and the composite $\mathbf{y} \circ F \circ \mathbf{x}^{-1} : \mathbf{x}[U] \to \mathbf{y}[V]$ is smooth.

Notation 1.2.19. In the following, we will often use the symbol $C^\infty(M)$ to denote the set of all smooth, real-valued maps on M.

1.3 Examples of smooth manifolds

Before we proceed to further explore the general theory of smooth structures on manifolds, let us show some examples of what we can classify as a smooth manifold.

Example 1.3.1 [Euclidean Spaces]. For every $n \in \mathbb{N}_0$, the Euclidean space $\mathbb{R}^n$ is a smooth manifold $(\mathbb{R}^n, \{(\mathbb{R}^n, \mathrm{id}_{\mathbb{R}^n})\})$ of dimension n. Notice that this smooth structure is not canonically given and is called the *standard smooth structure on* $\mathbb{R}^n$.

Example 1.3.2 [Spheres in Euclidean spaces]. Let $\|-\|$ be the Euclidean norm on $\mathbb{R}^{n+1}$, $n \geq 0$. The unit sphere

$$\mathbb{S}^n := \left\{x \in \mathbb{R}^{n+1} \mid \|x\| = 1\right\}$$

is indeed Hausdorff and second-countable as it is a topological subspace of $\mathbb{R}^{n+1}$. To inspect its structure, let us denote for each $i \in \{1, \dots, n+1\}$

$$\begin{aligned} U_i^+ &:= \left\{\left(x^1, \dots, x^{n+1}\right) \in \mathbb{R}^{n+1} \mid x^i > 0\right\}, \\ U_i^- &:= \left\{\left(x^1, \dots, x^{n+1}\right) \in \mathbb{R}^{n+1} \mid x^i < 0\right\}. \end{aligned}$$

Now consider the continuous function $f : \mathbb{B}^n \to \mathbb{R}$

$$f(u) = \sqrt{1 - \|u\|^2},$$

where $\mathbb{B}^n = \{x \in \mathbb{R}^n \mid \|x\| < 1\}$. Now it is easy to see that $U_i^+ \cap \mathbb{S}^n$ and $U_i^- \cap \mathbb{S}^n$ are the graphs of functions

$$\begin{aligned} x^i &= f\left(x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^{n+1}\right), \\ x^i &= -f\left(x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^{n+1}\right) \end{aligned}$$

respectively. Therefore, we can conclude that each subset $U_i^\pm \cap \mathbb{S}^n$ is locally Euclidean of dimension n and that the maps $\mathbf{x}_i^\pm : U_i^\pm \cap \mathbb{S}^n \to \mathbb{B}^n$ given by

$$\mathbf{x}_i^\pm\left(x^1, \dots, x^{n+1}\right) = \left(x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^{n+1}\right)$$

are graph coordinates for $\mathbb{S}^n$. Now let us take an arbitrary composition $\mathbf{x}_i^\pm \circ \left(\mathbf{x}_j^\pm\right)^{-1}$, where we can assume that $i < j$. Explicitly, we obtain

$$\mathbf{x}_i^\pm \circ \left(\mathbf{x}_j^\pm\right)^{-1}\left(x^1, \dots, x^n\right) = \left(x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^{j-1}, \pm\sqrt{1 - \|x\|^2}, x^{j+1}, \dots, x^n\right).$$

In the case of $i > j$, we obtain a similar relation. Also, consider that if $i = j$, we get

$$\mathbf{x}_i^+ \circ \left(\mathbf{x}_i^-\right)^{-1} = \mathbf{x}_i^- \circ \left(\mathbf{x}_i^+\right)^{-1} = \mathrm{id}_{\mathbb{B}^n}.$$

As we can see, the collection $\left\{\left(U_i^\pm \cap S^n, \mathbf{x}_i^\pm\right)\right\}$ forms a smooth atlas on $\mathbb{S}^n$, turning it into a smooth manifold of dimension n.

Remark 1.3.3. Example 1.3.2 shows that one can define manifolds in $\mathbb{R}^{n+1}$ by using certain functions. Namely, we have used the function

$$\|-\| : \mathbb{R}^{n+1} \to \mathbb{R}$$

to conclude that the inverse image

$$\|-\|^{-1}(1) = \{x \in \mathbb{R}^{n+1} \mid \|x\| = 1\}$$

is a manifold.

This is an instance of a general procedure that is supported by the so-called *Regular Value Theorem*. We state the theorem later, see 2.3.3 below.

2. Tangent spaces

In this chapter, we solidify the terminology and definitions of core concepts of the differential structure on manifolds such as the derivative at a point and the derivative of a smooth map between two manifolds.

Thus, in Section 2.1, we introduce the concept of a tangent vector at a point of a manifold. We also show that these vectors quite naturally form a vector space of the same dimension as the ambient manifold. Moreover, smooth maps are able to carry tangent vectors to tangent vectors by means of a derivative of a smooth map.

The point of view that derivatives of smooth maps are linear maps is central to modern differential geometry. In addition, it leads quite naturally to the concept of a tangent bundle of a manifold, i.e., of tangent spaces "glued together", see Section 2.4. The tangent bundle of a manifold is crucial in an easy treatment of vector fields, curves, and, ultimately, covariant derivatives.

While we postpone the notion of a covariant derivative to Chapter 3, we also introduce vector fields, curves and their velocities in Sections 2.5 and 2.6.

For further details on the material of this chapter, we refer to [7] again.

2.1 Tangent vectors at a point

We would like to build a notion similar to the definition of a directional derivative in $\mathbb{R}^n$. We will see that such a concept will lead quite naturally to the concept of a tangent vector. Let us see an example of what we mean by that.

Example 2.1.1. Let us fix a point p in the Euclidean plane $\mathbb{R}^2$. "Standing" at p, there are various directions, in which we can "look". Any such direction can be visualised as a non-zero vector $v_p = (v^1, v^2)^T$ having its foot at p.

If a smooth function, say $f : \mathbb{R}^2 \to \mathbb{R}$, is given, we can compute the derivative of f at p in direction v by computing the expression

$$\frac{\partial f}{\partial x}(p) \cdot v^1 + \frac{\partial f}{\partial y}(p) \cdot v^2.$$

If we denote the above expression by $v_p(f)$, then it is easy to see that the assignment $f \mapsto v_p(f)$ yields a function that

(a) is linear, i.e., the equality

$$v_p(af + bg) = av_p(f) + bv_p(g)$$

holds for all $a, b \in \mathbb{R}$ and all smooth functions $f : \mathbb{R}^2 \to \mathbb{R}$, $g : \mathbb{R}^2 \to \mathbb{R}$.

(b) satisfies the so-called *Leibniz rule*

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

for all smooth functions f, g.

Conversely, if an assignment $f \mapsto v_p(f)$ satisfies conditions (a) and (b) for all smooth maps f, it is easy to extract a vector $(v^1, v^2)^T$ by putting

$$\begin{aligned} v^1 &= v_p(f) \quad \text{for } f(x, y) = x, \\ v^2 &= v_p(g) \quad \text{for } g(x, y) = y. \end{aligned}$$

Hence, the vector v_p can be identified with a map $f \mapsto v_p(f)$ satisfying (a) and (b) above.

The above example serves as a motivation for the following definition: we define a tangent vector at a point to be a "derivative".

Definition 2.1.2. Let M be a smooth manifold of dimension n and let $p \in M$. A linear map $v_p : C^\infty(M) \to \mathbb{R}$ is called a *derivation at* p, if it satisfies the so-called *Leibniz rule* (or, *product rule*)

$$v_p(fg) = g(p)v_p(f) + f(p)v_p(g) \quad \text{for all } f, g \in C^\infty(M).$$

The set of all derivations at p is called a *tangent space at* p and is denoted by T_pM.

Remark 2.1.3. At any point $p \in M$, the tangent space T_pM is a vector space over $\mathbb{R}$ with the operations $+$ and $\cdot$ defined pointwise, i.e., for $v_p, w_p \in T_pM$ and $a \in \mathbb{R}$:

$$(v_p + w_p)(f) = v_p(f) + w_p(f), \qquad\qquad (av_p)(f) = av_p(f).$$

Example 2.1.4. An important example of a derivation at a point is the "partial" derivative with respect to a chart. More in detail, given a chart $(U, \mathbf{x})$ of an n-dimensional manifold M and a point $p \in U$, we define

$$\left.\frac{\partial}{\partial x^i}\right|_p : C^\infty(M) \to \mathbb{R}$$

as follows:

$$\left.\frac{\partial}{\partial x^i}\right|_p (f) = \partial_i \left(f \circ \mathbf{x}^{-1}\right)(\mathbf{x}(p)).$$

Above, on the right-hand side, we have used ∂_i to denote the usual partial derivative with respect to the i-th coordinate of the function $f \circ \mathbf{x}^{-1} : \mathbf{x}[U] \to \mathbb{R}$. It is then easy to show that $\partial/\partial x^i|_p$ is indeed a derivation at p. Notice that the above is chart-dependent and we should have written the more precise

$$\left.\frac{\partial}{\partial x^i}\right|_{(U,\mathbf{x}),p} .$$

Lemma 2.1.5. *Let M be a smooth manifold, let $p \in M$, $v_p \in T_pM$, and f, $g \in C^\infty$. Then the following hold:*

(i) *If f is a constant function, then $v_p(f) = 0$.*

(ii) *If $f(p) = g(p) = 0$, then $v_p(fg) = 0$.*

Proof.

(i) Assume, without loss of generality, that $f = 1$. This assumption is harmless since if the proposition holds for $f_1 = 1$, it immediately holds for $f = c \in \mathbb{R}$ as well because linearity of v_p yields $v_p(f) = v_p(cf_1) = cv_p(f_1) = 0$. For $f = 1$, the product rule gives

$$v_p(f) = v_p(ff) = f(p)v_p(f) + f(p)v_p(f) = 2v_p(f),$$

from which $v_p(f) = 0$ follows.

(ii) This is once again a simple consequence of the product rule:

$$v_p(fg) = \underbrace{f(p)}_{=0} v_p(g) + \underbrace{g(p)}_{=0} v_p(f) = 0.$$

■

The construction of a tangent space should give us the notion of something like a "linear approximation" much as it is in multivariable calculus in $\mathbb{R}^n$. Further, we will attempt to explore the way smooth maps affect tangent vectors.

Definition 2.1.6. Let M and N be smooth manifolds, and let $F : M \to N$ be a smooth map. For every $p \in M$, define a map

$$\begin{aligned} DF|_p : T_pM &\to T_{F(p)}N, \\ v_p &\mapsto (f \mapsto v_p(f \circ F)), \end{aligned} \tag{2.1.1}$$

where $f \in C^\infty(N)$. We call this map the *derivative* (or, *total derivative*) of F at p.

Note 2.1.7. Note that Formula 2.1.1 is well defined since when $f \in C^\infty(N)$ and $F : M \to N$, then the composite map $f \circ F \in C^\infty(M)$ and thus $v_p(f \circ F)$ makes perfect sense.

Every $DF|_p(v_p)$ is a derivation on N at the point $F(p)$. Linearity immediately follows from the fact that v_p is linear. To show that the product rule holds, consider that for any $f, g \in C^\infty(N)$, we have

$$\begin{aligned} DF|_p(v_p)(fg) &= v_p\left((fg) \circ F\right) = v_p\left((f \circ F)(g \circ F)\right) \\ &= (f \circ F)(p)v_p(g \circ F) + (g \circ F)(p)v_p(f \circ F) \\ &= f\left(F(p)\right) DF|_p(v_p)(g) + g\left(F(p)\right) DF|_p(v_p)(f). \end{aligned}$$

Claim 2.1.8 [Properties of Derivatives]. *Let M, N and P be smooth manifolds, let the maps $F : M \to N$ and $G : N \to P$ be smooth, and let $p \in M$. Then the following holds:*

(i) *$DF|_p : T_pM \to T_{F(p)}N$ is linear.*

(ii) *$D(G \circ F)|_p = DG|_{F(p)} \circ DF|_p : T_pM \to T_{(G\circ F)(p)}P$.*

(iii) *$D(\mathrm{id}_M)|_p = \mathrm{id}_{T_pM} : T_pM \to T_pM$.*

(iv) *If F is a diffeomorphism,[1] then $DF|_p : T_pM \to T_{F(p)}N$ is an isomorphism and it holds that $(DF|_p)^{-1} = D\big(F^{-1}\big)\big|_{F(p)}$.*

[1]Recall that the *diffeomorphism* from M to N is a smooth bijective map $F : M \to N$, which has a smooth inverse.

Proof. Let $v_p, w_p \in T_pM$ be derivations at p, let $a, b \in \mathbb{R}$, and let $f \in C^\infty(M)$, $h \in C^\infty(P)$. Then the following hold:

(i) $DF|_p(av_p + bw_p)(f) = (av_p + bw_p)(f \circ F) = aDF|_p(v_p)(f) + bDF|_p(w_p)(f)$.

(ii) $D(G \circ F)|_p(v_p)(h) = v_p(h \circ G \circ F) = DF|_p(v_p)(h \circ G) = (DG|_{F(p)} \circ DF|_p)(v_p)(h)$.

(iii) $D(\mathrm{id}_M)|_p(v_p)(f) = v_p(f \circ \mathrm{id}_M) = v_p(f) = \mathrm{id}_{T_pM}(v_p)(f)$.

(iv) Since according to (i), $DF|_p$ is linear, is suffices to prove that it always has an inverse. Applying (ii) and (iii), we obtain

$$\mathrm{id}_{T_pM} \overset{\text{(iii)}}{=} D(\mathrm{id}_M)|_p = D\left(F^{-1} \circ F\right)\big|_p \overset{\text{(ii)}}{=} D\left(F^{-1}\right)\big|_{F(p)} \circ DF|_p.$$

This directly proves that $(DF|_p)^{-1} = D\left(F^{-1}\right)\big|_{F(p)}$ and that $DF|_p$ is an isomorphism.

■

Now that we have properly defined the derivative and explored its basic properties, our first application of it will be to use the smooth structure we have introduced to relate the tangent space at any point of the n-dimensional manifold with $\mathbb{R}^n$. In fact, the derivations $\partial/\partial x^i|_p$, introduced in Example 2.1.4 form a basis of the space T_pM. This is a non-trivial result, hence we state it without proof (see, e.g., [7] for the full proof).

Claim 2.1.9. *Let M be a smooth manifold of dimension n. Then T_pM is a vector space of dimension n with*

$$\left.\frac{\partial}{\partial x^1}\right|_p, \dots, \left.\frac{\partial}{\partial x^n}\right|_p$$

being its basis.

The above claim (and the notation $\partial/\partial x^i|_p$) is yet one more trick that will allow us to work locally as if we were in a Euclidean space. We will see it in the next section.

We finish this section with the claim (again without a proof) that allows us to identify, for any $p \in U$ with U open in M, the spaces T_pM and T_pU. The full proof can be found in [7] again.

Claim 2.1.10. *Let M be a smooth manifold, let $U \subseteq M$ be open, and let $\iota : U \hookrightarrow M$ be the inclusion of U in M. Then for every $p \in M$, the derivative $D\iota|_p : T_pU \to T_pM$ is an isomorphism.*

2.2 The derivative in coordinates

So far, our research of tangent spaces has been purely theoretical and abstract. In this section, we will attempt to show how to perform a certain degree of computations with tangent vectors and derivatives. Eventually, we will arrive at an intersection of the differential calculus in Euclidean spaces and the notions we have defined on the tangent structure of a smooth manifold.

Recall from Example 2.1.4 that the operator $\partial/\partial x^i|_p$ acts upon $f \in C^\infty(M)$ as follows:

$$\left.\frac{\partial}{\partial x^i}\right|_p f = \frac{\partial \hat{f}}{\partial x^i}(\hat{p}),$$

where $\hat{f} = f \circ \mathbf{x}^{-1}$ and $\hat{p} = \mathbf{x}(p)$ denote the *coordinate representation* of f and p respectively.

Remark 2.2.1. Following Claim 2.1.9, a tangent vector $v_p \in T_pM$ can be written as a linear combination

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

where in the latter case, we have adapted the widely recognised *Einstein's summation convention*. The ordered basis $(\partial/\partial x^1|_p, \ldots, \partial/\partial x^n|_p)$, where x^i is the i-th coordinate function, is referred to as a *coordinate basis* for T_pM and the ordered n-tuple (v^i) is called the *components* of v with respect to the coordinate basis.

To compute the components of $v_p \in T_pM$ easily, we can utilise its action on the coordinate basis $\left(x^i\right)$. Specifically, if we want to compute the j-th component of v, we can simply put

$$v\left(x^j\right) = \left(v^i \left.\frac{\partial}{\partial x^i}\right|_p\right)\left(x^j\right) = v^i \frac{\partial x^j}{\partial x^i}(p) = v^j.$$

Example 2.2.2. Further, we would like to explore how derivatives of smooth functions look like in coordinates. Let us begin by introducing an example of a smooth map $F : U \to V$, where $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ are open. We will attempt to express $DF|_p : T_p\mathbb{R}^n \to T_{F(p)}\mathbb{R}^m$ as a matrix with respect to the standard bases in Euclidean spaces. Before we begin, let us mention the used notation, where $\left(x^i\right)$ denotes the coordinates in the domain, whereas $\left(y^i\right)$ denotes those in the codomain.

$$DF|_p\left(\left.\frac{\partial}{\partial x^i}\right|_p\right) f = \left.\frac{\partial}{\partial x^i}\right|_p (f \circ F) = \frac{\partial f}{\partial y^j}(F(p))\frac{\partial F^j}{\partial x^i}(p) = \left(\frac{\partial F^j}{\partial x^i}(p) \left.\frac{\partial}{\partial y^j}\right|_{F(p)}\right) f$$

and thus

$$DF|_p\left(\left.\frac{\partial}{\partial x^i}\right|_p\right) = \frac{\partial F^j}{\partial x^i}(p) \left.\frac{\partial}{\partial y^j}\right|_{F(p)}. \tag{2.2.1}$$

In terms of a matrix representation, we can express $DF|_p$ as

$$F'(p) = \begin{pmatrix} \dfrac{\partial F^1}{\partial x^1}(p) & \cdots & \dfrac{\partial F^1}{\partial x^n}(p) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F^m}{\partial x^i}(p) & \cdots & \dfrac{\partial F^m}{\partial x^n}(p) \end{pmatrix}.$$

This is of course the Jacobi matrix of F at p which is the matrix representation of the total derivative $DF|_p : \mathbb{R}^n \to \mathbb{R}^m$ familiar from multivariable calculus in Euclidean spaces. What we have attained is a direct correspondence of $DF|_p : \mathbb{R}^n \to \mathbb{R}^m$ to the total derivative $DF|_p$.

Now, let us reach out from the overly specific example back to the greater generality of smooth manifolds. As the following claim shows, the definition of a derivative has been created precisely to get a familiar result.

Claim 2.2.3. *Let M and N be smooth manifolds, and let $F : M \to N$ be smooth. Then the derivative $DF|_p : TM \to TN$ is represented in the coordinate bases by the Jacobi matrix of $\hat{F}$, the coordinate representation of F.*

Proof. Let p be a point of M contained in the smooth chart $(U, \mathbf{x})$ on M, and let $F(p)$ be contained in the smooth chart $(V, \mathbf{y})$ on N. Now we can write the coordinate representations $\hat{F} = \mathbf{y} \circ F \circ \mathbf{x}^{-1}$ and $\hat{p} = \mathbf{x}(p)$. Following our previous example, the derivative $D\hat{F}|_{\hat{p}}$ is represented by the Jacobi matrix of $\hat{F}$ at $\hat{p}$ with respect to the standard coordinate bases. Therefore, we can put

$$\begin{aligned}
DF|_p \left(\left. \frac{\partial}{\partial x^i} \right|_p \right) &= DF|_p \left(D\left(\mathbf{x}^{-1}\right)|_{\hat{p}} \left(\left. \frac{\partial}{\partial x^i} \right|_{\hat{p}} \right) \right) \\
&= D\left(\mathbf{y}^{-1}\right)|_{\hat{F}(\hat{p})} \left(D\hat{F}|_{\hat{p}} \left(\left. \frac{\partial}{\partial x^i} \right|_{\hat{p}} \right) \right) \\
&= D\left(\mathbf{y}^{-1}\right)|_{\hat{F}(\hat{p})} \left(\frac{\partial \hat{F}^j}{\partial x^i}(\hat{p}) \left. \frac{\partial}{\partial y^j} \right|_{\hat{F}(\hat{p})} \right) \\
&= \frac{\partial \hat{F}^j}{\partial x^i}(\hat{p})\, D\left(\mathbf{y}^{-1}\right)|_{\hat{F}(\hat{p})} \left(\left. \frac{\partial}{\partial y^j} \right|_{F(p)} \right) \\
&= \frac{\partial \hat{F}^j}{\partial x^i}(\hat{p}) \left. \frac{\partial}{\partial y^j} \right|_{F(p)},
\end{aligned}$$

where in the second equality, we have utilised the fact that $F \circ \mathbf{x}^{-1} = \mathbf{y}^{-1} \circ \hat{F}$, and in the last equality, we have emphasised the notation for coordinate vectors. ■

Remark 2.2.4. Observe that we have created a useful computation shorthand. In linear algebra, every linear map

$$g : L_1 \to L_2$$

between finitely-dimensional vector spaces can be "replaced" with a *matrix* of g with respect to chosen bases. Above, we have applied this principle to the linear map $T_pM \xrightarrow{DF|_p} T_{F(p)}N$ where we have chosen charts $(U, \mathbf{x})$, $(V, \mathbf{y})$ such that $F[U] \subseteq V$ and $p \in U$. Then the matrix $F'(p)$ of $\partial \hat{F}^j / \partial x^i(\hat{p})$ is the matrix of $DF|_p$ with respect to the bases $(\partial/\partial x^i|_p)$ and $(\partial/\partial y^j|_{F(p)})$, respectively. We will often loosely write

$$DF|_p(v_p) = F'(p) \cdot v$$

whenever

$$v_p = v^i \left. \frac{\partial}{\partial x^i} \right|_p \qquad \text{and} \qquad v = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}.$$

Remark 2.2.5. Let M be a smooth manifold, and let $\left(U, \left(x^i\right)\right)$ and $\left(V, \left(y^i\right)\right)$ be two overlapping smooth charts on M, i.e., $U \cap V \neq \emptyset$. As an interesting application of Claim 2.2.3, one can ask the following question: how are two different representations of a point $p \in U \cap V$ related? Naturally, any tangent vector at p can be represented with respect to either of the bases $\left(\partial/\partial x^i|_p\right)$ and $\left(\partial/\partial y^i|_p\right)$. Let us begin with the transition map $\mathbf{y} \circ \mathbf{x}^{-1}$. By (2.2.1), we can write its derivative as

$$D\left(\mathbf{y} \circ \mathbf{x}^{-1}\right)\big|_{\mathbf{x}(p)}\left(\left.\frac{\partial}{\partial x^i}\right|_{\mathbf{x}(p)}\right) = \frac{\partial y^j}{\partial x^i}(\mathbf{x}(p))\left.\frac{\partial}{\partial y^j}\right|_{\mathbf{y}(p)}.$$

With that in mind, we proceed by expressing the coordinate vector

$$\begin{aligned}
\left.\frac{\partial}{\partial x^i}\right|_p &= D\left(\mathbf{x}^{-1}\right)\big|_{\mathbf{x}(p)}\left(\left.\frac{\partial}{\partial x^i}\right|_{\mathbf{x}(p)}\right) \\
&= D\left(\mathbf{y}^{-1} \circ \mathbf{y} \circ \mathbf{x}^{-1}\right)\big|_{\mathbf{x}(p)}\left(\left.\frac{\partial}{\partial x^i}\right|_{\mathbf{x}(p)}\right) \\
&= D\left(\mathbf{y}^{-1}\right)\big|_{\mathbf{y}(p)} \circ D\left(\mathbf{y} \circ \mathbf{x}^{-1}\right)\big|_{\mathbf{x}(p)}\left(\left.\frac{\partial}{\partial x^i}\right|_{\mathbf{x}(p)}\right) \\
&= D\left(\mathbf{y}^{-1}\right)\big|_{\mathbf{y}(p)}\left(\frac{\partial y^j}{\partial x^i}(\mathbf{x}(p))\left.\frac{\partial}{\partial y^j}\right|_{\mathbf{y}(p)}\right) \\
&= \frac{\partial y^j}{\partial x^i}(\mathbf{x}(p)) D\left(\mathbf{y}^{-1}\right)\big|_{\mathbf{y}(p)}\left(\left.\frac{\partial}{\partial y^j}\right|_{\mathbf{y}(p)}\right) \\
&= \frac{\partial y^j}{\partial x^i}(\hat{p})\left.\frac{\partial}{\partial y^j}\right|_p,
\end{aligned} \tag{2.2.2}$$

where $\hat{p} = \mathbf{x}(p)$ as usual. When applied to the components of a tangent vector

$$v = v^i \left.\frac{\partial}{\partial x^i}\right|_p = \widetilde{v}^i \left.\frac{\partial}{\partial y^i}\right|_p,$$

we can conclude that the components of v transform as follows:

$$\widetilde{v}^j = \frac{\partial y^j}{\partial x^i}(\hat{p}) v^i. \tag{2.2.3}$$

2.3 Extrinsic manifolds

Although a smooth manifold M can be defined purely *intrinsically* (as a certain topological space), sometimes it is useful to treat a manifold *extrinsically*, i.e., as a "nice" subspace of an ambient space $\mathbb{R}^N$ (equipped with the usual Euclidean topology).

Hence, an *n-dimensional smooth manifold* in $\mathbb{R}^N$ is a subset M of $\mathbb{R}^N$ with the following property: for every point p in M, there exist

(i) an open subset V of $\mathbb{R}^N$, such that $p \in V$,

(ii) an open subset U of $\mathbb{R}^N$,

(iii) a homeomorphism $\varphi : U \to \varphi[U]$, such that $\varphi[U] = M \cap V$.

The pair $(p, \varphi^{-1} : \varphi[U] \to U)$ is then a *local chart* at p and the collection of local charts fulfils the usual compatibility requirements that make M a smooth manifold.

Example 2.3.1. For an open set $U \subseteq \mathbb{R}^n$ and a smooth map $f : U \to \mathbb{R}^m$, the *graph*

$$\mathrm{graph}(f) = \{(t, f(t)) \mid t \in U\} \subseteq \mathbb{R}^{n+m}$$

is a smooth n-dimensional manifold in $\mathbb{R}^N$, for $N = n + m$.

It is not always possible to exhibit a manifold in the form of a graph of a smooth function. Moreover, the definition of an extrinsic manifold may be quite difficult to work with.

Luckily, there is a way of producing manifolds via systems of *implicit* equations. This result, often called *The Regular Value Theorem*, is based on the well-known *Implicit Function Theorem*. We recall these theorems in this section.

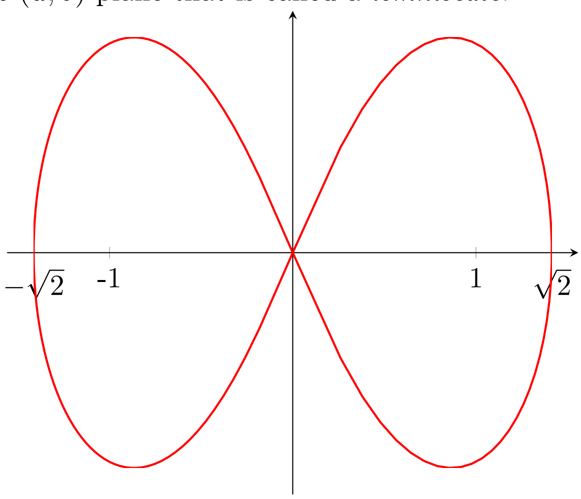
The Implicit Function Theorem. Let us start with the following example. Set

$$\varphi(u, v) = (u^2 + v^2)^2 - 2(u^2 - v^2).$$

Then φ is a smooth function of two variables and the equation

$$\varphi(u, v) = 0$$

defines a curve in the (u, v)-plane that is called a *lemniscate*.



Observe that, save three points,[2] it seems to be the case that, *locally*, v is a function of u, i.e., $v = f(u)$. It may be hopeless to find f explicitly. But there exists a technique of finding enough derivatives of f.

Thus, suppose we have a point $(u_0, v_0)^T = p$ on the curve. Suppose that we can find f such that

$$\varphi(u, f(u)) = 0$$

in the neighbourhood of u_0. Thus, for an open set $U \subseteq \mathbb{R}$ containing u_0, we have smooth functions

$$\begin{array}{ccccc} U & \xrightarrow{\mathrm{graph}(f)} & \mathbb{R}^2 & \xrightarrow{\varphi} & \mathbb{R}, \\ u & \longmapsto & \begin{pmatrix} u \\ f(u) \end{pmatrix} & \longmapsto & \varphi(u, f(u)), \end{array}$$

[2] In what follows, let us suppose that $u_0 \notin \{-\sqrt{2}, 0, \sqrt{2}\}$ to avoid the points, where v cannot be easily expressed as a function of u.

such that $\varphi \circ \mathrm{graph}(f) = 0$. Hence, by the chain rule, we have

$$
\begin{aligned}
(\varphi \circ \mathrm{graph}(f))'(u) &= \varphi' \begin{pmatrix} u \\ f(u) \end{pmatrix} \cdot (\mathrm{graph}(f))'(u) \\
&= \left(\left.\frac{\partial \varphi}{\partial u}\right|_{(u,f(u))^T}, \left.\frac{\partial \varphi}{\partial v}\right|_{(u,f(u))^T} \right) \cdot \begin{pmatrix} 1 \\ f'(u) \end{pmatrix} = 0.
\end{aligned}
$$

Hence,

$$
f'(u) = -\frac{\left.\frac{\partial \varphi}{\partial u}\right|_{(u,f(u))^T}}{\left.\frac{\partial \varphi}{\partial v}\right|_{(u,f(u))^T}}.
$$

The Implicit Function Theorem is an obvious generalisation of the above. The full proof can be found e.g., in [1].

Theorem 2.3.2 [Implicit Function Theorem]. *Suppose $W \subseteq \mathbb{R}^n \times \mathbb{R}^m$ is open and suppose that*

$$
\varphi : W \to \mathbb{R}^m
$$

is a smooth function. Suppose, moreover, that there exists $(u_0, v_0) \in W$ such that $\varphi(u_0, v_0) = 0$ and such that the matrix

$$
D_v \varphi = \begin{pmatrix} \frac{\partial \varphi_1}{\partial v_1} & \cdots & \frac{\partial \varphi_1}{\partial v_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi_m}{\partial v_1} & \cdots & \frac{\partial \varphi_m}{\partial v_m} \end{pmatrix}
$$

is invertible at v_0. Then there exist open sets $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ such that

$$
(u_0, v_0) \in U \times V \subseteq W
$$

and such that for each $u \in U$, there exists a unique $f(u) \in V$ satisfying

$$
\varphi(u, f(u)) = 0.
$$

Moreover, the function $u \mapsto f(u)$ is smooth and

$$
Df(u) = -\left(D_v \varphi|_{(u,f(u))^T} \right)^{-1} \cdot D_u \varphi|_{(u,f(u))^T}
$$

holds, where

$$
D_u \varphi = \begin{pmatrix} \frac{\partial \varphi_1}{\partial u_1} & \cdots & \frac{\partial \varphi_1}{\partial u_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi_m}{\partial u_1} & \cdots & \frac{\partial \varphi_m}{\partial u_n} \end{pmatrix}.
$$

Theorem 2.3.3 [The Regular Value Theorem]. *Let $W \subseteq \mathbb{R}^N$ be an open set and let $\varphi : W \to \mathbb{R}^m$ be a smooth map. For any $c \in \mathbb{R}^m$, denote by*

$$
\varphi^{-1}(c) = \{x \in W \mid \varphi(x) = c\}
$$

and assume that c is a regular value *of φ, i.e., assume that the linear map*

$$\mathbb{R}^N \xrightarrow{D\varphi|_x} \mathbb{R}^m$$

has rank m for all $x \in \varphi^{-1}(c)$. If, moreover, $\varphi^{-1}(c) \neq \emptyset$, then

$$M = \varphi^{-1}(c)$$

is a smooth manifold in $\mathbb{R}^N$ of dimension $n = N - m$ (we also say that M has codimension *m in $\mathbb{R}^N$ in this case). Furthermore, the equality*

$$T_x M = \ker(D\varphi|_x)$$

holds for all $x \in M$.

Proof. Take $x \in M$. Then, after suitable renumbering of the axes of $\mathbb{R}^N$, we can assume that the last m columns of the matrix representation of $\mathbb{R}^N \xrightarrow{D\varphi|_x} \mathbb{R}^m$ are linearly independent.

Put $n = N - m$ and write $\mathbb{R}^N = \mathbb{R}^n \times \mathbb{R}^m$, $x = (u_0, v_0)$. By the Implicit Function Theorem, there are open sets $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ such that $(u_0, v_0) \in U \times V$ and such that for each $u \in U$, there is a unique $f(u) \in V$ such that

$$(u, f(u)) \in M,$$

i.e., such that $\varphi(u, f(u)) = c$. Thus,

$$M \cap (U \times V) = \mathrm{graph}(f)$$

is a smooth manifold of dimension n, since the function f is smooth.

The above argument is valid for all $x \in M$, hence M is a smooth manifold of dimension n.

To prove that $T_x M = \ker(D\varphi|_x)$, observe that

$$\begin{array}{ccccc} U & \xrightarrow{\mathrm{graph}(f)} & \mathbb{R}^N & \xrightarrow{\quad\varphi\quad} & \mathbb{R}^m, \\ u & \longmapsto & (u, f(u)) & \longmapsto & \underbrace{\varphi(u, f(u))}_{=c}. \end{array}$$

Therefore, the chain rule yields

$$D\varphi(\mathrm{graph}(f)(u)) \cdot D\,\mathrm{graph}(f)(u) = 0.$$

Setting $u = u_0$ hence $\mathrm{graph}(f)(u_0) = (u_0, f(u_0)) = x$, we obtain

$$D\varphi(x) \cdot D\,\mathrm{graph}(f)(u_0) = 0.$$

Since $T_x M = \mathrm{im}(D\,\mathrm{graph}(f)(u_0))$, we immediately obtain that

$$T_x M \subseteq \ker(D\varphi|_x).$$

However, both $T_x M$ and $\ker(D\varphi|_x)$ have dimension n:

(i) $T_x M$ does since M has dimension n,

(ii) $\ker(D\varphi|_x)$ does since $\mathrm{im}(D\varphi|_x)$ has dimension m and the equality

$$\dim(\ker(D\varphi|_x)) + \dim(\mathrm{im}(D\varphi|_x)) = N = n + m$$

holds.

Thus, the equality $T_x M = \ker(D\varphi|_x)$ holds for all $x \in M$. ■

2.4 Tangent bundles

Now, we are ready to state another very important definition: the tangent bundle. This construction allows us to refer to an object containing the information not only about the smooth manifold but also about its tangent spaces at all points. Recall that the vector spaces T_pM and T_qM are disjoint, whenever $p \neq q$, see Remark 2.1.3. The purpose of forming the tangent bundle is to "glue" these vector spaces together. Moreover, the bundle will become a manifold in a natural way.

Definition 2.4.1. Let M be a smooth manifold. The (disjoint) union

$$TM = \bigcup_{p \in M} T_pM$$

is called the *tangent bundle of M*.

Remark 2.4.2. Since the union $\bigcup_{p\in M} T_pM$ is disjoint, an element of TM is an element of T_pM, for a unique $p \in M$. Hence, there is a well-defined map

$$\begin{aligned} \pi_M : TM &\to M, \\ v_p &\mapsto p. \end{aligned}$$

We now define charts on TM as follows: for every chart $(U, \mathbf{x})$ on an n-dimensional smooth manifold M, we define a chart $(\widetilde{U}, \widetilde{\mathbf{x}})$ by putting

$$\widetilde{U} = \pi_M^{-1}[U] = \{v_p \in T_pM \mid p \in U\}.$$

Further, since every $v_p \in T_pM$ can be written as $v^i \partial/\partial x^i|_p$ (see Claim 2.1.9), we put

$$\widetilde{\mathbf{x}}(v_p) = \left(x^1(p), \dots, x^n(p), v^1, \dots, v^n\right).$$

Observe that, after we prove that the collection of all $(\widetilde{U}, \widetilde{\mathbf{x}})$ forms a smooth atlas, the set TM will become a smooth manifold of dimension $2n$.

Theorem 2.4.3. *Let $\mathcal{A} = \{(U_\alpha, \mathbf{x}) \mid \alpha \in A\}$ be a smooth atlas, turning M into a smooth n-dimensional manifold. Then the collection $\{(\widetilde{U}_\alpha, \widetilde{\mathbf{x}}_\alpha) \mid \alpha \in A\}$ turns TM into a smooth $2n$-dimensional manifold. Moreover, the map*

$$\pi_M : TM \to M$$

is smooth.

Proof. Note first that the image of $\widetilde{\mathbf{x}}$ is $\mathbf{x}[U] \times \mathbb{R}^n$ which is an open subset of $\mathbb{R}^{2n}$. Also note that the map $\widetilde{\mathbf{x}}$ is a bijection onto its image, because its inverse can be written explicitly as

$$(\widetilde{\mathbf{x}})^{-1}\left(x^1, \dots, x^n, v^1, \dots, v^n\right) = v^i \frac{\partial}{\partial x^i}\bigg|_{\mathbf{x}^{-1}(x)}.$$

Now that we have established charts on TM, let us check that they form a smooth atlas. Suppose that $(U, \mathbf{x})$ and $(V, \mathbf{y})$ are two charts on M, and let $(\widetilde{U}, \widetilde{\mathbf{x}})$ and $(\widetilde{V}, \widetilde{\mathbf{y}})$ be the corresponding charts on TM using the construction above. The sets

$$\widetilde{\mathbf{x}}[\widetilde{U} \cap \widetilde{V}] = \mathbf{x}[U \cap V] \times \mathbb{R}^n \qquad \text{and} \qquad \widetilde{\mathbf{y}}[\widetilde{U} \cap \widetilde{V}] = \mathbf{y}[U \cap V] \times \mathbb{R}^n$$

are clearly open in $\mathbb{R}^{2n}$. According to (2.2.3), the transition map $\widetilde{\mathbf{y}} \circ \widetilde{\mathbf{x}}^{-1}$ can be expressed explicitly as

$$\left(\widetilde{\mathbf{y}} \circ \widetilde{\mathbf{x}}^{-1}\right)\left(x^1, \ldots, x^n, v^1, \ldots, v^n\right) = \left(y^1(x), \ldots, y^n(x), \frac{\partial y^1}{\partial x^i}(x) v^i, \ldots, \frac{\partial y^n}{\partial x^i}(x) v^i\right)$$

which is clearly smooth. Thus, we obtain a smooth atlas on TM from an atlas on M which generates a topology on TM. It follows from Claim 1.2.15 that TM is Hausdorff and paracompact. This concludes the fact that TM is a smooth $2n$-dimensional manifold.

To see that π_M is smooth, note that with respect to charts $(U, \mathbf{x})$ on M and $(\widetilde{U}, \widetilde{\mathbf{x}})$ on TM, its coordinate representation is

$$\begin{aligned} \left(\mathbf{x} \circ \pi_M \circ \widetilde{\mathbf{x}}^{-1}\right)\left(x^1, \ldots, x^n, v^1, \ldots, v^n\right) &= \left(\mathbf{x} \circ \pi_M\right)\left(v^i \left.\frac{\partial}{\partial x^i}\right|_{\mathbf{x}^{-1}(x)}\right) \\ &= \left(\mathbf{x} \circ \mathbf{x}^{-1}\right)(x) \\ &= x. \end{aligned}$$

Hence, the coordinate representation is a mere projection $(x, v) \mapsto x$, which is smooth. ■

Remark 2.4.4 [The local trivialisation of TM]**.** Roughly, one might say that a smooth n-dimensional manifold M is "glued" together from "patches" of the form $\mathbf{x}[U] \subseteq \mathbb{R}^n$, where $(U, \mathbf{x})$ ranges through the charts $(U, \mathbf{x})$ of the smooth maximal atlas of M.

We show now that a similar thing can be said about the tangent bundle: for every chart $(U, \mathbf{x})$, there is a *trivialisation map*

$$\tau_{(U,\mathbf{x})} : \pi_M^{-1}[U] \to U \times \mathbb{R}^n.$$

More precisely, for every $v_p = v^i \partial/\partial x^i|_p$ in $\pi_M^{-1}[U]$, we define

$$\tau_{(U,\mathbf{x})}(v_p) = \left(p, \left(v^1, \ldots, v^n\right)\right).$$

It is easy to see that the triangle

$$\begin{tikzcd} \pi_M^{-1}[U] \arrow[rr, "\tau_{(U,\mathbf{x})}"] \arrow[dr, "\pi_M\restriction_{\pi_M^{-1}[U]}"'] & & U \times \mathbb{R}^n \arrow[dl, "\mathrm{proj}"] \\ & U & \end{tikzcd}$$

commutes, where proj is the projection onto the first component.

Moreover, the proof of Theorem 2.4.3 shows that $\tau_{(U,\mathbf{x})}$ is an isomorphism such that

$$\begin{aligned} \tau_{(U,\mathbf{x})}\restriction_{\pi_M^{-1}(p)} \colon T_pM &\to \{p\} \times \mathbb{R}^n, \\ v_p &\mapsto \left(p, \left(v^1, \ldots, v^n\right)\right) \end{aligned}$$

is an isomorphism of vector spaces. The above maps allow us to treat TM locally as $U \times \mathbb{R}^n$. We will use this trivialisation often in what follows.

Example 2.4.5 [The tangent map]**.** As a first example of trivialisation, we show how to treat the map

$$TM \xrightarrow{TF} TN$$

which is just consisting of the maps

$$T_pM \xrightarrow{DF|_p} T_{F(p)}N$$

for a smooth map $M \xrightarrow{F} N$. The map TF is called the *tangent map* of F.

Locally, we can describe TF as follows: identify, for every chart $(U, \mathbf{x})$, elements of $\pi_M^{-1}[U]$ with pairs (p, v), where $p \in U$ and $v \in \mathbb{R}^n$. Then the pair

$$\left(F(p), D\hat{F}|_{\mathbf{x}(p)} \cdot v\right)$$

is a pair in the trivialisation given by a chart $(V, \mathbf{y})$ of N such that $F(p) \in V$. To relax the notation further, we will write

$$(p, v) \mapsto (F(p), F'(p) \cdot v)$$

for the map TF in trivialisation given by $(U, \mathbf{x})$.

The above is quite useful. One can, for example, show that the diagram

$$\begin{tikzcd} TM \arrow[r, "TF"] \arrow[d, "\pi_M"'] & TN \arrow[d, "\pi_N"] \\ M \arrow[r, "F"'] & N \end{tikzcd}$$

commutes, just by "chasing the elements around in a trivialisation"

$$\begin{tikzcd} (p, v) \arrow[r, mapsto, "TF"] \arrow[d, mapsto, "\pi_M"'] & (F(p), F'(p) \cdot v) \arrow[d, mapsto, "\pi_N"] \\ p \arrow[r, mapsto, "F"'] & F(p) \end{tikzcd}$$

We will use such arguments a lot later, see Chapter 4.

Claim 2.4.6. *Let M and N be smooth manifolds, and let $TM \xrightarrow{\pi_M} M$ and $TN \xrightarrow{\pi_N} N$ be their respective tangent bundles. Then, for a smooth map $M \xrightarrow{F} N$, the map*

$$TF : TM \to TN$$

is itself a smooth map.

Proof. Let $(U, \mathbf{x})$ and $(V, \mathbf{y})$ be charts on the m-dimensional manifold M and the n-dimensional manifold N, respectively, such that $p \in U$, $F(p) \in V$. Further, let $(\widetilde{U}, \widetilde{\mathbf{x}})$ and $(\widetilde{V}, \widetilde{\mathbf{y}})$ be charts on TM and TN, respectively, constructed as in the proof of Theorem 2.4.3. Then, thanks to the trivialisation of TM and TN, we can write

$$\widetilde{\mathbf{x}}\big[\widetilde{U}\big] = \mathbf{x}[U] \times \mathbb{R}^m \qquad \text{and} \qquad \widetilde{\mathbf{y}}\big[\widetilde{V}\big] = \mathbf{y}[V] \times \mathbb{R}^n$$

which allows us to locally express $(\mathbf{x}(p), v) \in \widetilde{\mathbf{x}}[\widetilde{U}]$ as

$$\left(x^1(p), \ldots, x^n(p), v^1, \ldots, v^n\right).$$

Therefore, for action of the composite $\widetilde{\mathbf{y}} \circ TF \circ \widetilde{\mathbf{x}}^{-1}$, we can write

$$\left(\widetilde{\mathbf{y}} \circ TF \circ \widetilde{\mathbf{x}}^{-1}\right)(\mathbf{x}(p), v) = \left(\widetilde{\mathbf{y}} \circ DF|_p \circ \widetilde{\mathbf{x}}^{-1}\right)(\mathbf{x}(p), v) = \left(\mathbf{y}(F(p)), F'(p) \cdot v\right),$$

where $F'(p)$ is the Jacobi matrix of F at p. Since F is a smooth map, then the composite $\mathbf{y} \circ F$ is smooth as well. For the second component, we see that it is given by the Jacobi matrix, i.e., it is linear and therefore smooth. ■

We conclude this section with a list of basic properties of tangent maps. The proof is easy and we omit it.[3]

Claim 2.4.7. *Let M, N and P be smooth manifolds, and let $F : M \to N$ and $G : N \to P$ be smooth maps. Then the following hold:*

(a) $T(G \circ F) = TG \circ TF$.

(b) $T(\mathrm{id}_M) = \mathrm{id}_{TM}$.

(c) *If F is a diffeomorphism, then $TF : TM \to TN$ is also a diffeomorphism and it holds that $(TF)^{-1} = T\big(F^{-1}\big)$.*

2.5 Curves and velocity vectors

When we have introduced our definition of a tangent space, we had not discussed possible interpretations or different definitions. In fact, the definition of a tangent space can be approached from various viewpoints and one of them is considering a tangent vector to be an equivalence class of velocities of curves passing through the given point. We will not indulge in this definition much further but it is convenient to understand the notion of curves and their velocity vectors on smooth manifolds as it will prove rather useful further in the thesis.

Definition 2.5.1. Let M be a smooth manifold. A *curve* in M is a continuous map $\gamma : I \to M$, where $I \subseteq \mathbb{R}$ is an interval.[4]

Note 2.5.2. It is very important to note that the notion of a *curve* does not convey just a set of points in M. Instead, it is the entire map from an interval to M.

Definition 2.5.3. Let M be a smooth manifold, let $\gamma : I \to M$ be a curve on M, and let $t_0 \in I$. The *velocity vector of γ at t_0*, denoted by $\gamma'(t_0)$, is defined as follows:

$$\gamma'(t_0) = D\gamma|_{t_0} \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t_0} \right) \in T_{\gamma(t_0)}M,$$

where $\mathrm{d}/\mathrm{d}t|_{t_0}$ is the standard coordinate basis vector in $T_{t_0}\mathbb{R}$.

Remark 2.5.4. Based on the knowledge how derivatives act upon their targets, the velocity vector acts upon a function $f \in C^\infty(M)$ as

$$\gamma'(t_0)(f) = D\gamma|_{t_0} \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t_0} \right) f = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t_0} (f \circ \gamma) = (f \circ \gamma)'(t_0).$$

As we can see, the velocity vector $\gamma'(t_0)$ is the derivation at $\gamma(t_0)$ of a function taken along γ.[5]

[3] The proof is very similar to the proof of Claim 2.1.8.

[4] In most cases, we will require I to be an open interval of $\mathbb{R}$, but since it can be useful to consider it having one or two endpoints and since our definition works either way (and definitions relying on it can undergo only a slight modification in order to work), there is no harm in not specifying its openness.

[5] This is one of the points where we can employ the property of I having an endpoint. If t_0 happens to be an endpoint of I, this still holds, provided that we interpret $\mathrm{d}/\mathrm{d}t|_{t_0}$ as a one-sided derivative or the derivative of any smooth extension of $f \circ \gamma$ to an open subset of $\mathbb{R}$.

Remark 2.5.5. Now, we can ask the question of how does $\gamma'(t_0)$ look like when expanded into the coordinate basis on $T_{\gamma(t_0)}M$. Let $(U, \mathbf{x})$ be a smooth chart on M containing $\gamma(t_0)$. For t sufficiently close to t_0, the coordinate representation of γ is $\gamma(t) = (\gamma^1(t), \ldots, \gamma^n(t))$. Further, the derivative yields

$$\gamma'(t_0) = \frac{\mathrm{d}\gamma^i}{\mathrm{d}t}(t_0) \left.\frac{\partial}{\partial x^i}\right|_{\gamma(t_0)} .$$

This formula is similar to the one in Euclidean spaces.

Now—because we have promised not to indulge in the definition of tangent spaces via curves on manifolds—we state the next proposition without proof. It should serve predominantly as an interesting viewpoint giving us further geometric intuition behind tangent vectors.

Claim 2.5.6. *Let M be a smooth manifold, and let $p \in M$. Every $v_p \in T_pM$ is the velocity vector of some smooth curve in M passing through p.*

Claim 2.5.7. *Let M and N be smooth manifolds, let $F : M \to N$ be a smooth map, and let $\gamma : I \to M$ be a smooth curve. For any $t_0 \in I$, the velocity at t_0 of the composite curve $F \circ \gamma : I \to N$ is given by*

$$(F \circ \gamma)'(t_0) = DF|_{\gamma(t_0)} \left(\gamma'(t_0)\right) .$$

Proof. From the definition of the velocity of a curve:

$$(F \circ \gamma)'(t_0) = D(F \circ \gamma)|_{t_0} \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t_0} \right) = DF|_{\gamma(t_0)} \circ D\gamma|_{t_0} \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t_0} \right) = DF|_{\gamma(t_0)} \left(\gamma'(t_0)\right) .$$

■

Clearly, Claim 2.5.7 tells us how to compute the velocity of a composite curve using the derivative. However, it is often much more convenient to use this proposition the other way round and use it to compute the derivative $DF|_p$ at some point $p \in M$. For any $v_p \in T_pM$, we can compute $DF|_p(v_p)$ using a smooth curve γ whose initial tangent vector is v. Such a curve always exists, thanks to Claim 2.5.6. Then we apply Claim 2.5.7 to $F \circ \gamma$. The next corollary summarizes our discussion.

Corollary 2.5.8. *Let M and N be smooth manifolds, let $p \in M$ and $v_p \in T_pM$, and let $F : M \to N$ be smooth. Then for every smooth curve $\gamma : I \to M$ such that $0 \in I$, $\gamma(0) = p$, and $\gamma'(0) = v$, the following holds:*

$$DF|_p(v) = (F \circ \gamma)'(0).$$

2.6 Vector fields and frames

An insightful reader may remark that so far, we have been generalising various notions we know from calculus in Euclidean spaces onto smooth manifolds. Another notion we extend in a similar manner are vector fields. In the setting of multivariable calculus, a vector field on an open subset $U \subseteq \mathbb{R}^n$ is simply a continuous map from U to $\mathbb{R}^n$, often visualised as attaching an "arrow" to each point of U. In the context of smooth manifolds, we think of a vector field as of a special kind of a continuous map X from M to its tangent bundle TM.

Definition 2.6.1. Let M be a smooth manifold. A *rough vector field* on M is a section of the map $\pi_M : TM \to M$, i.e., a map $X : M \to TM$ with the property that

$$\pi_M \circ X = \mathrm{id}_M, \tag{2.6.1}$$

or equivalently, $X_p \in T_pM$ for each $p \in M$.

Note 2.6.2. As we can see already in the definition, we usually write the action of X upon p as $p \mapsto X_p$ instead of $X(p)$.

To further expand the nomenclature regarding vector fields, let us introduce the following notions:

- A *vector field* is a continuous rough vector field.
- A *smooth vector field* is a vector field, which is smooth as a map from M to TM with respect to the smooth structure on TM we have solidified in Theorem 2.4.3.
- The *support* $\mathrm{supp}(X)$ of X is the closure of the set $\{p \in M \mid X_p \neq 0_p\}$.
- A vector field is said to be *compactly supported* if its support is a compact set, i.e., if every open cover of $\mathrm{supp}(X)$ has a finite subcover.
- If M is a smooth n-manifold, $X : M \to TM$ is a rough vector field, and $(U, \mathbf{x})$ is a smooth chart on M, we can expand the action of X upon a point $p \in M$ using the coordinate basis vectors of T_pM as follows:

  $$X_p = X^i(p) \left.\frac{\partial}{\partial x^i}\right|_p .$$

  This defines n *coordinate functions* $X^i : U \to \mathbb{R}$ of X in $(U, \mathbf{x})$.

Claim 2.6.3. *Let M be a smooth n-manifold, let $(U, \mathbf{x})$ be an arbitrary smooth chart on M, and let $X : M \to TM$ be a rough vector field. Then the restriction $X\restriction_U$ is smooth if and only if its coordinate functions with respect to $(U, \mathbf{x})$ are smooth.*

Proof. Given the smooth structure on TM, let $(x^1, \dots, x^n, v^1, \dots, v^n)$ be the coordinates on $\pi^{-1}[U] \subseteq TM$. Then the coordinate representation of X on U is

$$\hat{X}(x) = \left(x^1, \dots, x^n, X^1(x), \dots, X^n(x)\right),$$

where X^i is the i-th component function of X in x^i-coordinates. Immediately, we can see that smoothness of X on U is equivalent to smoothness of its component functions. ∎

Example 2.6.4 [Coordinate Vector Fields]**.** Let M be a smooth n-manifold, and let $(U, \mathbf{x})$ be an arbitrary smooth chart. Then the assignment

$$p \mapsto \left.\frac{\partial}{\partial x^i}\right|_p$$

determines a vector field on U. We call this vector field the *i-th coordinate vector field* and denote it $\partial/\partial x^i$. Naturally, it is smooth since its component functions are constant.

Remark 2.6.5. The set $\mathfrak{X}(M)$ of all smooth vector fields on a smooth manifold M is a vector space under the pointwise addition and scalar multiplication

$$(X + Y)_p = X_p + Y_p \qquad \text{and} \qquad (aX)_p = aX_p.$$

The zero element is the zero vector field, which for every $p \in M$ yields $0_p \in T_pM$. Furthermore, for $X \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$, we define $fX : M \to TM$ by

$$(fX)_p = f(p)X_p,$$

turning $\mathfrak{X}(M)$ into a module over the ring $C^\infty(M)$.[6]

We will now proceed by taking a slight detour which culminates in an interesting characterisation of smooth manifolds. Before that, we will show that the representation of vector fields using coordinate vector fields is convenient since their values form a basis for the tangent space at each point but that it is not the only option. Let us start with a bulk definition of several notions.

Definition 2.6.6. Let M be a smooth n-manifold.

- An ordered k-tuple $(X_1, \ldots, X_k)$ of rough vector fields defined on a subset $A \subseteq M$ is said to be *linearly independent* if $((X_1)_p, \ldots, (X_k)_p)$ is a linearly independent k-tuple in T_pM for every $p \in A$. Furthermore, it is said to *span the tangent bundle* on A, if the k-tuple $((X_1)_p, \ldots, (X_k)_p)$ spans T_pM at every $p \in A$.
- A *local frame for M* is defined as an ordered n-tuple of rough vector fields $(E_1, \ldots, E_n)$ defined on an open subset $U \subseteq M$, which is linearly independent and which spans the tangent bundle on U. Therefore, the vectors $((E_1)_p, \ldots, (E_n)_p)$ form a basis for T_pM at every $p \in U$. Furthermore, if $U = M$, it is called a *global frame* and if each of the vector fields E_i is smooth, we call it a *smooth frame*.

Example 2.6.7. Let $(U, \mathbf{x})$ be a smooth chart for a smooth manifold M. Then the coordinate vector fields form a smooth local frame $(\partial/\partial x^i)$ on U. We call this frame the *coordinate frame* on U.

Remark 2.6.8. When working in $\mathbb{R}^n$, a very special type of frame exists which can prove rather useful when discussing geometric problems. A k-tuple of vector fields $(E_1, \ldots, E_k)$ defined on a subset $A \subseteq \mathbb{R}^n$ is said to be *orthonormal* if for every $p \in A$, the vectors $((E_1)_p, \ldots, (E_k)_p)$ are orthonormal with the respect to the Euclidean dot product under the usual identification of $T_p\mathbb{R}^n$ with $\mathbb{R}^n$. A frame—whether local or global—consisting of orthonormal vector fields is called an *orthonormal frame*.

Remark 2.6.9. Regarding local frames, several tools for constructing them have been developed, e.g., the Gram-Schmidt Algorithm for Frames, see [8]. Therefore, local frames are rather common. However, global frames are not that easy to come by. A smooth manifold is said to be *parallelisable* if it admits a smooth global frame. It can be shown that the elementary examples of smooth manifolds such as $\mathbb{R}^n$ or $\mathbb{S}^1$ are indeed parallelisable. Despite this naive initial evidence, most smooth manifolds are not parallelisable. Not even the (still fairly elementary) example of the 2-sphere $\mathbb{S}^2$ is parallelisable. This specific matter is subject of the renowned *Hairy ball theorem*, see e.g., [7]. Generally, it can be shown that parallelisability of a smooth manifold M is closely related to the question of whether its tangent bundle is diffeomorphic to $M \times \mathbb{R}^n$ or not.

[6] A module is a "vector space over the structure that is not a field". See [11] or [7].

3. Covariant derivatives and connections

One of the important questions in the study of manifolds is the following: what does it mean to "transport" a tangent vector "parallelly" along a curve on a manifold? While the previous chapters introduced manifolds, tangent vectors and curves, we have no way how to answer the above question yet. For that, we need an *additional* structure on our smooth manifold. This extra structure can be essentially given in two conceptually different ways that can be proved to be equivalent. It is the aim of this thesis to prove this equivalence. The two ways, informally, are as follows:

(a) A covariant derivative on a manifold.

Roughly speaking, a covariant derivative is a generalisation of a directional derivative. By giving a covariant derivative, we can thus define "to be parallelly transported" as having a zero covariant derivative in its own direction.

In fact, the above leads to another important concept of differential geometry—that of a *geodesic curve*. Geodesic curves (also known simply as *geodesics*) are those curves that are "as straight as possible".

(b) A connection form on a manifold.

Although a proper definition of a connection form is a bit technical, its underlying idea is again very simple. One looks at the tangent bundle and tries to "connect" individual tangent spaces at points that are "close to each other". In this manner, one can speak about the "same vector in different tangent spaces". Having established such a connection, the definition of a "parallel transport" of a vector is simple: the vector is required to "stay the same".

As we have already mentioned, both approaches above are equivalent to each other. We prove this equivalence in Chapter 5 below.

In the current chapter, we give examples of covariant derivatives on a plane and on a two-dimensional sphere. We also indicate how the connection form works on a simple example of an open region of a Euclidean plane.

3.1 Covariant derivative in $\mathbb{E}^2_{x,y}$

Introducing and explaining the covariant derivative is very easy on a flat space $\mathbb{E}^2_{x,y}$, i.e., a Euclidean plane with coordinates x, y. At each point (x_0, y_0), we have a tangent space (which is a plane once again) spanned by $e_x = \partial R/\partial x|_{(x_0,y_0)}$ and $e_y = \partial R/\partial y|_{(x_0,y_0)}$, where $R(x,y) = (x,y)$ is the *position* (or *radius*) *vector.*

The notion of a covariant derivative $\nabla_U V$ is supposed to compute the rate of change of a *tangent vector field* V in the direction of a vector field U. Consider

$$V : (x,y) \mapsto 3e_x - 2e_y$$

and let us compute its rate of change in the direction of the x and y axes:

$$\begin{aligned}\frac{\partial}{\partial x} V\bigg|_{(x_0,y_0)} &= \left(\frac{\partial}{\partial x}(3e_x) - \frac{\partial}{\partial x}(2e_y)\right)\bigg|_{(x_0,y_0)} \\ &= 3\underbrace{\frac{\partial}{\partial x}(e_x)\bigg|_{(x_0,y_0)}}_{=0} - 2\underbrace{\frac{\partial}{\partial x}(e_y)\bigg|_{(x_0,y_0)}}_{=0} \\ &= 0.\end{aligned}$$

For y, we get an analogous result. This is pretty much what we have expected. For a non-constant vector field

$$V(x,y) = v^1(x,y)e_x + v^2(x,y)e_y,$$

we need to be slightly more careful:

$$\begin{aligned}\frac{\partial}{\partial x} V\bigg|_{(x_0,y_0)} &= \frac{\partial}{\partial x}\left(v^1 e_x + v^2 e_y\right)\bigg|_{(x_0,y_0)} \\ &= \frac{\partial}{\partial x}\left(v^1 e_x\right)\bigg|_{(x_0,y_0)} + \frac{\partial}{\partial x}\left(v^2 e_y\right)\bigg|_{(x_0,y_0)} \\ &= \frac{\partial v^1}{\partial x}\bigg|_{(x_0,y_0)} e_x + v^1 \underbrace{\frac{\partial}{\partial x}(e_x)\bigg|_{(x_0,y_0)}}_{=0} + \frac{\partial v^2}{\partial x}\bigg|_{(x_0,y_0)} e_y + v^2 \underbrace{\frac{\partial}{\partial x}(e_y)\bigg|_{(x_0,y_0)}}_{=0} \\ &= \frac{\partial v^1}{\partial x}\bigg|_{(x_0,y_0)} e_x + \frac{\partial v^2}{\partial x}\bigg|_{(x_0,y_0)} e_y.\end{aligned}$$

Analogously, we obtain

$$\frac{\partial}{\partial y} V\bigg|_{(x_0,y_0)} = \frac{\partial v^1}{\partial y}\bigg|_{(x_0,y_0)} e_x + \frac{\partial v^2}{\partial y}\bigg|_{(x_0,y_0)} e_y.$$

Hence—in coordinates x, y—derivatives of vectors fields in directions of the axes are precisely the partial derivatives.

Apparently, *in flat space*, the covariant derivative is just the ordinary partial derivative, where we differentiate both the vector field component functions and the basis vectors in tangent spaces. Our result can be written in a more elegant way, denoting $c^1 = x$, $c^2 = y$, $e_1 = e_x$, and $e_2 = e_y$ and employing Einstein's summation convention:

$$\frac{\partial}{\partial c^i} V\bigg|_{(x_0,y_0)} = \frac{\partial}{\partial c^i}\left(v^j e_j\right)\bigg|_{(x_0,y_0)} = \frac{\partial v^k}{\partial c^i}\bigg|_{(x_0,y_0)} e_k, \tag{3.1.1}$$

where we have harmlessly renamed the indices $j \leftrightarrow k$.

3.2 Covariant derivative on a sphere

In the previous section, we have learned that the covariant derivative can be given the following interpretation:

$$\nabla_u v = \text{"the rate of change of a vector field } v \text{ in direction } u \\ \text{with the normal component subtracted"}$$

since the tangent vector to a point of a plane "remains" in the plane and thus its normal component is zero.

We will now pass onto a "more exotic" manifold, while maintaining the above idea for a covariant derivative. Namely, let

$$\mathbb{S}^2 = \{(x, y, z) \in \mathbb{E}^3_{x,y,z} \mid x^2 + y^2 + z^2 = 1\}$$

be the unit sphere, considered a smooth regular submanifold od $\mathbb{E}^3_{x,y,z}$. We can give it the traditional parametrisation:

$$\begin{aligned}\Phi : (0, \pi) \times (0, 2\pi) &\to \mathbb{E}^3_{x,y,z},\\ (u, v) &\mapsto (X(u, v), Y(u, v), Z(u, v)),\end{aligned}$$

where

$$X(u, v) = \cos(v)\sin(u), \quad Y(u, v) = \sin(v)\sin(u), \quad Z(u, v) = \cos(u).$$

Note that, strictly speaking, the image of $(0, \pi) \times (0, 2\pi)$ under Φ is not the whole of $\mathbb{S}^2$. However, we will ignore this nuisance for the time being. At each point $\Phi(u_0, v_0)$ of $\mathbb{S}^2$, there are two tangent vectors: $e_u = \partial\Phi/\partial u|_{(u_0,v_0)}$ and $e_v = \partial\Phi/\partial v|_{(u_0,v_0)}$. Their exact formulas can be computed using the chain rule:

$$\begin{aligned}e_u = \left.\frac{\partial\Phi}{\partial u}\right|_{(u_0,v_0)} &= \left.\left(\frac{\partial X}{\partial u}\frac{\partial\Phi}{\partial X} + \frac{\partial Y}{\partial u}\frac{\partial\Phi}{\partial Y} + \frac{\partial Z}{\partial u}\frac{\partial\Phi}{\partial Z}\right)\right|_{(u_0,v_0)}\\ &= \cos(v_0)\cos(u_0)\left.\frac{\partial\Phi}{\partial X}\right|_{(u_0,v_0)} + \sin(v_0)\cos(u_0)\left.\frac{\partial\Phi}{\partial Y}\right|_{(u_0,v_0)} - \sin(u_0)\left.\frac{\partial\Phi}{\partial Z}\right|_{(u_0,v_0)}\\ &= \cos(v_0)\cos(u_0)e_x + \sin(v_0)\cos(u_0)e_y - \sin(u_0)e_z,\end{aligned}$$

$$\begin{aligned}e_v = \left.\frac{\partial\Phi}{\partial v}\right|_{(u_0,v_0)} &= \left.\left(\frac{\partial X}{\partial v}\frac{\partial\Phi}{\partial X} + \frac{\partial Y}{\partial v}\frac{\partial\Phi}{\partial Y} + \frac{\partial Z}{\partial v}\frac{\partial\Phi}{\partial Z}\right)\right|_{(u_0,v_0)}\\ &= -\sin(v_0)\sin(u_0)\left.\frac{\partial\Phi}{\partial X}\right|_{(u_0,v_0)} + \cos(v_0)\sin(u_0)\left.\frac{\partial\Phi}{\partial Y}\right|_{(u_0,v_0)} + 0\left.\frac{\partial\Phi}{\partial Z}\right|_{(u_0,v_0)}\\ &= -\sin(v_0)\sin(u_0)e_x + \cos(v_0)\sin(u_0)e_y + 0e_z,\end{aligned}$$

where we have used that $\partial\Phi/\partial x|_{(u_0,v_0)} = e_x$, $\partial\Phi/\partial y|_{(u_0,v_0)} = e_y$ and $\partial\Phi/\partial z|_{(u_0,v_0)} = e_z$ for the orthonormal basis e_x, e_y, e_z with respect to the identification of $T_{\Phi(u_0,v_0)}\mathbb{E}^3_{x,y,z}$ with $\mathbb{E}^3_{x,y,z}$.

Since $T_{\Phi(u_0,v_0)}\mathbb{S}^2$ is spanned by the two vectors e_u, e_v above, once can consider the vector $n = e_u \times e_v$, to form the basis (e_u, e_v, n) of $\mathbb{E}_{x,y,z}$. Observe that $\langle e_u | e_v\rangle = 0$, where $\langle - | - \rangle$ denotes the standard scalar product in $\mathbb{R}^3$. Hence (e_u, e_v, n) is an orthogonal basis of $\mathbb{R}^3$.

Let us use the above computation to establish what it means that a curve is "as straight as possible". As a criterion for straightness, we will demand the velocity of such a curve to be constant, i.e., taking the second derivative alongside the curve should be zero. Note that since we are working extrinsically, taking derivatives of curves within the manifold can yield objects that do not belong in the tangent space. Thus, we will decompose them into a *tangential* and a *normal component*.

Let us start with the computation of the velocity of a curve $\gamma : I \to \mathbb{S}^2$. Such a curve can be interpreted as $I \ni \lambda \mapsto (u(\lambda), v(\lambda)) \mapsto \Phi(u(\lambda), v(\lambda)) \in \mathbb{S}^2$. The velocity of γ at a point $\gamma(\lambda_0)$ is

$$\left.\frac{\mathrm{d}\gamma}{\mathrm{d}\lambda}\right|_{\lambda_0} = \frac{\partial u}{\partial\lambda}\left.\frac{\partial\Phi}{\partial u}\right|_{\lambda_0} + \frac{\partial v}{\partial\lambda}\left.\frac{\partial\Phi}{\partial v}\right|_{\lambda_0} = \frac{\partial u}{\partial\lambda}(\lambda_0)e_u + \frac{\partial v}{\partial\lambda}(\lambda_0)e_v.$$

Note that the velocity vector is always in $T_{\gamma(\lambda_0)}S^2$. Hence, we have

$$\nabla_{\frac{\mathrm{d}}{\mathrm{d}\lambda}}\gamma = \frac{\mathrm{d}\gamma}{\mathrm{d}\lambda}$$

since the normal component is zero. Now, using the product rule, the acceleration, i.e., the second derivative, of γ at point λ_0 is given by

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\mathrm{d}\gamma}{\mathrm{d}\lambda}\right)\bigg|_{\lambda_0} &= \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\partial u}{\partial\lambda}\frac{\partial\Phi}{\partial u} + \frac{\partial v}{\partial\lambda}\frac{\partial\Phi}{\partial v}\right)\bigg|_{\lambda_0} \\ &= \frac{\partial^2 u}{\partial\lambda^2}\frac{\partial\Phi}{\partial u}\bigg|_{\lambda_0} + \frac{\partial u}{\partial\lambda}\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\partial\Phi}{\partial u}\right)\bigg|_{\lambda_0} + \frac{\partial^2 v}{\partial\lambda^2}\frac{\partial\Phi}{\partial v}\bigg|_{\lambda_0} + \frac{\partial v}{\partial\lambda}\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\partial\Phi}{\partial v}\right)\bigg|_{\lambda_0}.\end{aligned}$$

Further, observe that it is not yet entirely clear what the quantities

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\partial\Phi}{\partial u}\right)\bigg|_{\lambda_0} \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\partial\Phi}{\partial v}\right)\bigg|_{\lambda_0}$$

represent and where they "live". Proceeding with the computation of the above quantities using the chain rule, we obtain the following formula:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\mathrm{d}\gamma}{\mathrm{d}\lambda}\right)\bigg|_{\lambda_0} &= \frac{\partial^2 u}{\partial\lambda^2}\frac{\partial\Phi}{\partial u}\bigg|_{\lambda_0} + \frac{\partial u}{\partial\lambda}\left(\frac{\partial u}{\partial\lambda}\frac{\partial^2\Phi}{\partial u^2} + \frac{\partial v}{\partial\lambda}\frac{\partial^2\Phi}{\partial u\partial v}\right)\bigg|_{\lambda_0} + \\ &+ \frac{\partial^2 v}{\partial\lambda^2}\frac{\partial\Phi}{\partial v}\bigg|_{\lambda_0} + \frac{\partial v}{\partial\lambda}\left(\frac{\partial u}{\partial\lambda}\frac{\partial^2\Phi}{\partial v\partial u} + \frac{\partial v}{\partial\lambda}\frac{\partial^2\Phi}{\partial v^2}\right)\bigg|_{\lambda_0}.\end{aligned}$$

Putting $u^1 = u$ and $u^2 = v$, we can rewrite the above into a more concise

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\mathrm{d}\gamma}{\mathrm{d}\lambda}\right)\bigg|_{\lambda_0} = \frac{\partial^2 u^i}{\partial\lambda^2}\frac{\partial\Phi}{\partial u^i}\bigg|_{\lambda_0} + \frac{\partial u^i}{\partial\lambda}\frac{\partial u^j}{\partial\lambda}\frac{\partial^2\Phi}{\partial u^i\partial u^j}\bigg|_{\lambda_0} \tag{3.2.1}$$

in compliance with Einstein's summation convention. A keen reader will surely notice that we have utilized the fact that $\partial^2\Phi/\partial u\partial v = \partial^2\Phi/\partial v\partial u$ holds.

We will now expand the 3-dimensional vectors $\partial^2\Phi/\partial u^i\partial u^j|_{\lambda_0}$ in the basis (e_{u^1}, e_{u^2}, n), where n is the unit normal vector to the tangent space given by $n = e_{u^1}\times e_{u^2}$. Therefore, we obtain

$$\frac{\partial^2\Phi}{\partial u\partial v}\bigg|_{\lambda_0} = \Gamma^k_{ij}e_{u^k} + L_{ij}n, \tag{3.2.2}$$

where the coefficients Γ^k_{ij} are called the *Christoffel symbols* and L_{ij} the *second fundamental form.*

Merging (3.2.1) with (3.2.2) and using the fact that $\partial\Phi/\partial u^k|_{\lambda_0} = e_{u^k}$, we are now able to split the acceleration vector $\mathrm{d}^2\gamma/\mathrm{d}\lambda^2|_{\lambda_0}$ into the tangential and the normal components:

$$\begin{aligned}\frac{\mathrm{d}^2\gamma}{\mathrm{d}\lambda^2}\bigg|_{\lambda_0} &= \frac{\partial^2 u^k}{\partial\lambda^2}e_{u^k} + \frac{\partial u^i}{\partial\lambda}\frac{\partial u^j}{\partial\lambda}\left(\Gamma^k_{ij}e_{u^k} + L_{ij}n\right) \\ &= \left(\frac{\partial^2 u^k}{\partial\lambda^2} + \Gamma^k_{ij}\frac{\partial u^i}{\partial\lambda}\frac{\partial u^j}{\partial\lambda}\right)e_{u^k} + L_{ij}\frac{\partial u^i}{\partial\lambda}\frac{\partial u^j}{\partial\lambda}n.\end{aligned}$$

Since the acceleration vector of a geodesic should change only in its normal part, all of the coefficients of the tangential part should be zero:

$$\frac{\partial^2 u^k}{\partial\lambda^2} + \Gamma^k_{ij}\frac{\partial u^i}{\partial\lambda}\frac{\partial u^j}{\partial\lambda} = 0, \quad \text{for all } k. \tag{3.2.3}$$

The above are the *geodesic equations* and curves $\gamma : I \to \mathbb{S}^2$ that fulfil these equations are called *geodesics.*

Example 3.2.1. Now, let us figure out the geodesic equations on the sphere $\mathbb{S}^2$, parametrised by Φ. We will first compute the Christoffel symbols on the sphere. We denote by $\langle - | - \rangle$ the standard scalar product in $\mathbb{R}^3$ and we will use the orthogonal basis

$$\begin{aligned} e_u &= \cos(v_0)\cos(u_0)e_x + \sin(v_0)\cos(u_0)e_y - \sin(u_0)e_z, \\ e_v &= -\sin(v_0)\sin(u_0)e_x + \cos(v_0)\sin(u_0)e_y \end{aligned}$$

of $T_{\Phi(u_0,v_0)}\mathbb{S}^2$ and the vector $n = e_u \times e_v$. By computing the matrix

$$\begin{pmatrix} \langle e_u|e_u\rangle & \langle e_u|e_v\rangle \\ \langle e_v|e_u\rangle & \langle e_v|e_v\rangle \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2(u_0) \end{pmatrix}$$

we observe that the basis (e_u, e_v, n) is orthogonal but not orthonormal. Since the Christoffel symbols Γ_{ij}^k, $k = 1, 2$, are mere coordinates of the orthogonal projection of $\partial^2\Phi/(\partial u^i \partial u^j)$ to $T_{\Phi(u_0,v_0)}\mathbb{S}^2 = \operatorname{span}(e_{u^1}, e_{u^2})$, we have

$$\Gamma_{ij}^k = \frac{\left\langle \dfrac{\partial^2\Phi}{\partial u^i \partial u^j} \middle| \, e_{u^k} \right\rangle}{\langle e_{u^k} | \, e_{u^k} \rangle}, \quad k = 1, 2.$$

For example, the Christoffel's symbol Γ_{11}^1 is calculated as follows:

$$\begin{aligned} \Gamma_{11}^1 &= \left\langle \frac{\partial^2\Phi}{\partial u^2} \middle| \frac{\partial\Phi}{\partial u} \right\rangle = \left\langle \begin{pmatrix} -\cos(v_0)\sin(u_0) \\ -\sin(v_0)\sin(u_0) \\ -\cos(u_0) \end{pmatrix} \middle| \begin{pmatrix} \cos(v_0)\cos(u_0) \\ \sin(v_0)\cos(u_0) \\ -\sin(u_0) \end{pmatrix} \right\rangle \\ &= -\cos^2(v_0)\sin(u_0)\cos(u_0) - \sin^2(v_0)\sin(u_0)\cos(u_0) + \sin(u_0)\cos(u_0) \\ &= -\sin(u_0)\cos(u_0)\underbrace{\left[\sin^2(v_0) + \cos^2(v_0)\right]}_{=1} + \sin(u_0)\cos(u_0) = 0. \end{aligned}$$

Via such routine computations, we can see that

$$\begin{aligned} \Gamma_{11}^1 = \Gamma_{12}^1 = \Gamma_{21}^1 = 0 \quad &\text{and} \quad \Gamma_{22}^1 = -\frac{1}{2}\sin(2u_0), \\ \Gamma_{11}^2 = \Gamma_{22}^2 = 0 \quad &\text{and} \quad \Gamma_{12}^2 = \Gamma_{21}^2 = \cot(u_0). \end{aligned}$$

Hence, the geodesic equations on the sphere are given by

$$\begin{aligned} \frac{\partial^2 u}{\partial\lambda^2} - \frac{1}{2}\sin(2u)\frac{\partial v}{\partial\lambda}\frac{\partial v}{\partial\lambda} &= 0, \\ \frac{\partial^2 v}{\partial\lambda^2} + 2\cot(u)\frac{\partial u}{\partial\lambda}\frac{\partial v}{\partial\lambda} &= 0. \end{aligned}$$

For example, a part of the "equator" of $\mathbb{S}^2$ is a geodesic on $\mathbb{S}^2$, i.e., a curve, where $u = \pi/2$. Let us verify this claim: For $u = \pi/2$, the geodesic equation take the shape of

$$\begin{aligned} 0 &= 0, \\ \frac{\partial^2 v}{\partial\lambda^2} &= 0. \end{aligned}$$

The solution for this single equation is $v(\lambda) = k_v\lambda + v_0$ for some real numbers k_v, v_0. When we take the usual parameterisation of a sphere into account, we get

$$\begin{aligned} \Phi(u(\lambda), v(\lambda)) &= (\cos(v(\lambda))\sin(u(\lambda)), \sin(v(\lambda))\sin(u(\lambda)), \cos(u(\lambda))) \\ &= (\cos(k_v\lambda + v_0), \sin(k_v\lambda + v_0), 0). \end{aligned}$$

As we can see, a part of the equator spanned by the variable parameter $k_v\lambda + v_0$, where $\lambda \in I \subseteq \mathbb{R}$, is a geodesic on $\mathbb{S}^2$ for $u = \pi/2$.

Example 3.2.2. Consider the curve

$$\lambda \mapsto (u(\lambda), v(\lambda)) \mapsto \Phi(u(\lambda), v(\lambda)) \in \mathbb{S}^2,$$

where the parameter $\lambda \in (0, \pi/2)$ and $(u(\lambda), v(\lambda)) = (\pi/2, \lambda)$. Mapping Φ is the usual (smooth) parameterisation of $\mathbb{S}^2$, i.e., altogether, the image of this curve forms a quarter of the sphere's equator. Now, take the vector field

$$v(\lambda) = \cos\left(u^2(\lambda)\right) e_{u^1} + \sin\left(u^2(\lambda)\right) e_{u^2}$$

and compute the covariant derivative of v in the direction of the curve. That is, we compute the partial derivative with the normal component subtracted. This gives

$$\nabla_{\frac{\mathrm{d}}{\mathrm{d}\lambda}} v = \nabla_{\frac{\partial}{\partial u^2}} v = -\sin\left(u^2(\lambda)\right) e_{u^1} + \cos\left(u^2(\lambda)\right) e_{u^2},$$

where the first relation holds since we are working with a curve, where $u^2(\lambda) = \lambda$. It is worth noting that $v(0) = e_{u^1}$ and $v(\pi/2) = e_{u^2}$, i.e., the vector field v "rotates" in the tangent planes.

3.3 Covariant derivative in the plane, once more

The study of a covariant derivative on $\mathbb{S}^2$ from the previous section raises the question whether the "flat" manifold $\mathbb{E}^2_{x,y}$ fits into the picture as well. Indeed, Equation 3.1.1 can be rewritten as

$$\frac{\partial}{\partial c^i}(v) = \frac{\partial v^k}{\partial c^i} e_k + v^j \Gamma^k_{ij} e_k$$

by simply introducing $\Gamma^k_{ij} = 0$ for all k, i, j.

However, the situation changes rather dramatically when we move to *different coordinates* in our plane, for example to *polar coordinates*, i.e.,

$$\begin{aligned} \Phi : (0, \infty) \times (0, 2\pi) &\to \mathbb{E}^2_{x,y}, \\ (r, \varphi) &\mapsto (X(r, \varphi), Y(r, \varphi)), \end{aligned}$$

where

$$X(r, \varphi) = r\cos(\varphi) \qquad \text{and} \qquad Y(r, \varphi) = r\sin(\varphi).$$

Again, it is worth noting that the image of $(0, \infty) \times (0, 2\pi)$ under Φ is not the whole of $\mathbb{E}^2_{x,y}$ but let us ignore that for the time being. At each point $\Phi(r_0, \varphi_0)$, we have the tangent vectors b_{r_0} and b_{φ_0} forming the basis of the tangent space at $\Phi(r_0, \varphi_0)$. More specifically, the vectors take the form of

$$\begin{aligned} b_{r_0} = \left.\frac{\partial \Phi}{\partial r}\right|_{(r_0,\varphi_0)} &= \left.\left(\frac{\partial X}{\partial r}\frac{\partial \Phi}{\partial X} + \frac{\partial Y}{\partial r}\frac{\partial \Phi}{\partial Y}\right)\right|_{(r_0,\varphi_0)} \\ &= \frac{\partial X}{\partial r}(r_0, \varphi_0) e_x + \frac{\partial Y}{\partial r}(r_0, \varphi_0) e_y = \begin{pmatrix} \cos(\varphi_0) \\ \sin(\varphi_0) \end{pmatrix}, \end{aligned}$$

$$\begin{aligned} b_{\varphi_0} = \left.\frac{\partial \Phi}{\partial \varphi}\right|_{(r_0,\varphi_0)} &= \left.\left(\frac{\partial X}{\partial \varphi}\frac{\partial \Phi}{\partial X} + \frac{\partial Y}{\partial \varphi}\frac{\partial \Phi}{\partial Y}\right)\right|_{(r_0,\varphi_0)} \\ &= \frac{\partial X}{\partial \varphi}(r_0, \varphi_0) e_x + \frac{\partial Y}{\partial \varphi}(r_0, \varphi_0) e_y = \begin{pmatrix} -r_0\sin(\varphi_0) \\ r_0\cos(\varphi_0) \end{pmatrix}. \end{aligned}$$

It is worth noting that b_{φ_0} has its length depending on φ_0. Now, if we consider a vector field

$$v(r,\varphi) = v^1(r,\varphi)b_r + v^2(r,\varphi)b_\varphi,$$

then

$$\begin{aligned}
\frac{\partial}{\partial r}(v) &= \frac{\partial}{\partial r}(v^1 b_r) + \frac{\partial}{\partial r}(v^2 b_\varphi) = \frac{\partial v^1}{\partial r} b_r + v^1 \frac{\partial}{\partial r}(b_r) + \frac{\partial v^2}{\partial r} b_\varphi + v^2 \frac{\partial}{\partial r}(b_\varphi),\\
\frac{\partial}{\partial \varphi}(v) &= \frac{\partial}{\partial \varphi}(v^1 b_r) + \frac{\partial}{\partial \varphi}(v^2 b_\varphi) = \frac{\partial v^1}{\partial \varphi} b_r + v^1 \frac{\partial}{\partial \varphi}(b_r) + \frac{\partial v^2}{\partial \varphi} b_\varphi + v^2 \frac{\partial}{\partial \varphi}(b_\varphi).
\end{aligned}$$

Observe that the expression contains a change of basis vectors in tangent spaces. Specifically

$$\begin{aligned}
\frac{\partial}{\partial r}(b_r) &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \frac{\partial}{\partial \varphi}(b_r) = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{pmatrix},\\
\frac{\partial}{\partial r}(b_\varphi) &= \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{pmatrix}, \quad \frac{\partial}{\partial \varphi}(b_\varphi) = \begin{pmatrix} -r\cos(\varphi) \\ -r\sin(\varphi) \end{pmatrix}
\end{aligned}$$

which we need to express in the basis b_r, b_φ again. For this, we need the transformation matrix

$$\begin{aligned}
T_{(e_x,e_y)\mapsto(b_r,b_\varphi)} &= \left(T_{(b_r,b_\varphi)\mapsto(e_x,e_y)}\right)^{-1} = \begin{pmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi) \end{pmatrix}^{-1}\\
&= \begin{pmatrix} \cos(\varphi)/r & -\sin(\varphi)/r \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}.
\end{aligned}$$

Hence

$$\frac{\partial}{\partial r}(b_\varphi) = \frac{\partial}{\partial \varphi}(b_r) = 0b_r + \frac{1}{r}b_\varphi = \frac{1}{r}b_\varphi \qquad \text{and} \qquad \frac{\partial}{\partial \varphi}(b_\varphi) = -rb_r + 0b_\varphi = -rb_r$$

because

$$\begin{aligned}
\begin{pmatrix} \cos(\varphi)/r & -\sin(\varphi)/r \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{pmatrix} &= \begin{pmatrix} 0 \\ 1/r \end{pmatrix},\\
\begin{pmatrix} \cos(\varphi)/r & -\sin(\varphi)/r \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \begin{pmatrix} -r\sin(\varphi) \\ -r\cos(\varphi) \end{pmatrix} &= \begin{pmatrix} -r \\ 0 \end{pmatrix}.
\end{aligned}$$

Thus, altogether, we obtain

$$\begin{aligned}
\frac{\partial}{\partial r}(v) &= \frac{\partial v^1}{\partial r} b_r + \left(\frac{\partial v^2}{\partial r} + \frac{1}{r}v^2\right) b_\varphi,\\
\frac{\partial}{\partial \varphi}(v) &= \left(\frac{\partial v^1}{\partial \varphi} - rv^2\right) b_r + \left(\frac{\partial v^2}{\partial \varphi} + \frac{1}{r}v^1\right) b_\varphi.
\end{aligned}$$

If we put $p^1 = r$, $p^2 = \varphi$, $b_1 = b_r$, and $b_2 = b_\varphi$, we have

$$\frac{\partial}{\partial p^i}(v) = \left(\frac{\partial v^k}{\partial p^i} + v^j \Gamma^k_{ij}\right) b_k,$$

where the Christoffel's symbols *in polar coordinates* are

$$\begin{aligned}
\Gamma^1_{11} &= 0, \quad \Gamma^1_{22} = -r, \quad \Gamma^1_{12} = \Gamma^1_{21} = 0,\\
\Gamma^2_{11} &= 0, \quad \Gamma^2_{22} = 0, \quad \Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r}.
\end{aligned}$$

3.4 The idea behind connections

Although the construction of the tangent bundle TM of a manifold M glues together tangent spaces T_pM for $p \in M$, there is *no notion* of how to pass from T_pM to T_qM, whenever $p \neq q$.

More precisely, there is no linear isomorphism between spaces T_pM and T_qM that would be given by M itself. The collection of such isomorphisms would then allow us to speak about the "same vector" in different tangent spaces.

The notion of a *connection* on M is a new additional structure that allows us to speak about the "same vector but in different tangent spaces". The definition of a connection can be given in the spirit of differential geometry: we seek an isomorphism between T_pM and T_qM, whenever p and q are "infinitesimally close" to each other.

The exposition in this section was much inspired by Chapter XVII.16 of the book [2].

An open subset of a plane. Let M be an open subset of the plane, and let p and $q = p + h$ be two points of M. Assume that, for all h in a sufficiently small open neighbourhood V of the point $(0,0)^T$ in the plane, we have a linear isomorphism

$$F(h) : \mathbb{R}^2 \to \mathbb{R}^2,$$

such that the mapping

$$(p, u) \mapsto (p + h, F(h)^{-1} \cdot u),$$

provides us with an isomorphism between T_pM and $T_{p+h}M$. Moreover, we want the dependence $h \mapsto F(h)$ to be smooth and we want that $F(0) = \mathrm{id} : \mathbb{R}^2 \to \mathbb{R}^2$.

More in detail, we use the linear isomorphism $F(h)$ to provide an identification of vectors in $T_{p+h}M$ with vectors in T_pM for "small enough" h.

In order to conform with the traditional notation, we are going to use $F(h)^{-1}$ to construct a smooth mapping

$$\begin{aligned} \Phi_p : V \times \mathbb{R}^2 &\to U \times \mathbb{R}^2, \\ (h, u) &\mapsto (p + h, F(h)^{-1} \cdot u). \end{aligned}$$

Hence if h is "infinitesimally close" to $(0,0)^T$, then the value $\Phi_p(h, u)$ will return the vector $F(h)^{-1} \cdot u$ as the "copy" of vector u in $T_{p+h}M$.

To express the above slogan precisely, we need to consider the derivative of Φ_p at $((0,0)^T, u)$, which is a linear map

$$D\Phi_p|_{((0,0)^T,u)} : T_{((0,0)^T,u)}(V \times \mathbb{R}^2) \to T_{\Phi_p((0,0)^T,u)}(U \times \mathbb{R}^2)$$

that is "infinitesimally close" to the identity mapping.

The value of the above derivative at (k, v) is the pair[1]

$$(k, v - DF|_0(k, u)).$$

The pair (k, v) should be interpreted as follows: k is the vector in V that represents the "move" from p to $p+h$, whereas v is the "change" of vector u. The second component of

[1]We omit the lengthy calculations that lead to this result. We only mention that they follow the standard approach: split $\Phi_p : V \times \mathbb{R}^2 \to U \times \mathbb{R}^2$ into two functions $\Phi_{p,1} : V \times \mathbb{R}^2 \to U$, $\Phi_{p,2} : V \times \mathbb{R}^2 \to \mathbb{R}^2$ and find $D\Phi_p$ in the form of a "Jacobi matrix".

the pair $(k, v - DF|_0(k,u))$ then represents the change of v that is "caused" by "passing" u to a "nearby" tangent space.

The crucial part of the result is the appearance of the derivative $DF|_0$ at point (k,u) of the smooth mapping

$$\begin{aligned} F : V &\to \mathrm{Iso}(\mathbb{R}^2, \mathbb{R}^2), \\ h &\mapsto F(h), \end{aligned}$$

where we denoted by $\mathrm{Iso}(\mathbb{R}^2, \mathbb{R}^2)$ the manifold[2] of linear isomorphisms of $\mathbb{R}^2$ with itself.

The above derivative is therefore a linear map

$$DF|_0 : T_0 V \to T_{F(0)} \mathrm{Iso}(\mathbb{R}^2, \mathbb{R}^2)$$

which, due to the fact that $T_0 V \cong \mathbb{R}^2$ and $T_{F(0)} \mathrm{Iso}(\mathbb{R}^2, \mathbb{R}^2) \cong \mathrm{Lin}(\mathbb{R}^2, \mathbb{R}^2)$, is a bilinear map

$$\begin{aligned} DF|_0 : \mathbb{R}^2 \times \mathbb{R}^2 &\to \mathbb{R}^2, \\ (k,v) &\mapsto DF|_0(k,u). \end{aligned}$$

Conversely, if we give a bilinear map

$$\Gamma_p : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2,$$

we can define $F(h) : \mathbb{R}^2 \to \mathbb{R}^2$ to be the linear mapping

$$\mathrm{id}_{\mathbb{R}^2} + \Gamma_p(h, -).$$

Then, for h "small enough", the map $F(h)$ will be a linear isomorphism, since then $F(h)$ is "not far away" from the identity mapping.[3]

Thus, we can define a mapping

$$C_p((p,k),(p,u)) = ((p,u),(k, -\Gamma_p(k,u)))$$

that we will, for now, call the *connection form*.

We will show in the next chapter that connection forms and covariant derivatives are two equivalent descriptions of the same phenomenon, namely, of "transportation of a vector through different tangent spaces".

A final remark: our notation Γ_p resembles the notation for Christoffel symbols of Section 4.1. This is no coincidence; we prove later that they carry essentially the same information.

[2]It can be proved, using determinants, that $\mathrm{Iso}(\mathbb{R}^2, \mathbb{R}^2)$ is an open set in the complete normed space $\mathrm{Lin}(\mathbb{R}^2, \mathbb{R}^2)$ of all linear maps from $\mathbb{R}^2$ to $\mathbb{R}^2$. Hence, $\mathrm{Iso}(\mathbb{R}^2, \mathbb{R}^2)$ is a smooth manifold of dimension 4.

[3]We allude here for the generalisation of the well-known fact that $(1+x)^{-1} = \sum_{k=0}^{\infty} (-1)^k x^k$ for $|x| < 1$. See e.g., [1].

4. Abstract covariant derivatives and abstract connections

So far, we have worked with covariant derivatives in the concrete setting of examples. We have also indicated how the connection on a tangent bundle could provide us with essentially the same information as a covariant derivative.

To prove that covariant derivatives and connections are indeed the same thing, we need to start investigating both concepts abstractly. Thus, in this chapter, we define both covariant derivatives and connections as mappings on a tangent bundle, which satisfy additional properties.

4.1 The abstract covariant derivative

Recall that we have defined vector fields on a smooth manifold M to be (smooth) maps of the form $M \xrightarrow{X} TM$ such that

$$\begin{tikzcd} & TM \arrow[d, "\pi_M"] \\ M \arrow[ur, "X"] \arrow[r, "\mathrm{id}_M"'] & M \end{tikzcd}$$

commutes, see Definition 2.6.1.

Clearly, such a notion can be considered only locally which leads us to considering a bit more general notion of a vector field. Namely, for a chart $(U, \mathbf{x})$, we can consider smooth maps $U \xrightarrow{X} TM$ such that the triangle

$$\begin{tikzcd} & TM \arrow[d, "\pi_M"] \\ U \arrow[ur, "X"] \arrow[r, "\text{inclusion}"'] & M \end{tikzcd}$$

commutes.

Due to local trivialisation of TM (see Remark 2.4.4) we can say that X has the form $p \mapsto (p, x(p))$, where $p \mapsto x(p)$ is a smooth map from U to $\mathbb{R}^n$. In fact, one can go further and use the even more relaxed notation of Example 2.4.5.

Remark 4.1.1. Analogously to Remark 2.6.5, the set of all vector fields forms a *module* over $C^\infty(U)$ and we denote it by $\mathfrak{X}(U)$.

Definition 4.1.2. Let M be a smooth manifold of dimension n and let $(U, \mathbf{x})$ be a chart on M. A *covariant derivative* on U is a map

$$\begin{aligned} \nabla : \mathfrak{X}(U) \times \mathfrak{X}(U) &\to \mathfrak{X}(U), \\ (X, Y) &\mapsto \nabla_X Y \end{aligned}$$

such that the following conditions hold for every $f \in C^\infty(U)$:

(D1) $\nabla_{X_1 + X_2} Y = \nabla_{X_1} Y + \nabla_{X_2} Y$.

(D2) $\nabla_{fX} Y = f \nabla_X Y$.

(D3) $\nabla_X (Y_1 + Y_2) = \nabla_X Y_1 + \nabla_X Y_2$.

(D4) $\nabla_X (fY) = Xf \cdot Y + f \cdot \nabla_X Y$.

Above, Xf is the vector field $p \mapsto (p, f(p) \cdot X_p)$.

Remark 4.1.3. In the chart $(U, \mathbf{x})$, we have vector fields $U \to TM$, given by

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

which we can denote by $\partial / \partial x^i$, $i = 1, \ldots, n$. Further, in chart $(U, \mathbf{x})$, we can put

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma^k_{ij} \frac{\partial}{\partial x^k}.$$

Moreover, it then follows from (D1)—(D4) above that, for the vector fields

$$X = X^i \frac{\partial}{\partial x^i} \qquad \text{and} \qquad Y = Y^j \frac{\partial}{\partial x^j},$$

we can write

$$\begin{aligned}
\nabla_X Y &= \nabla_{X^i \frac{\partial}{\partial x^i}} \left(Y^j \frac{\partial}{\partial x^j} \right) = X^i \left(\nabla_{\frac{\partial}{\partial x^i}} \left(Y^j \frac{\partial}{\partial x^j} \right) \right) \\
&= X^i \left(\frac{\partial Y^j}{\partial X^i} \frac{\partial}{\partial x^j} + Y^j \nabla_{\frac{\partial}{\partial x^i}} \left(\frac{\partial}{\partial x^j} \right) \right) = X^i \left(\frac{\partial Y^j}{\partial X^i} \frac{\partial}{\partial x^j} + Y^j \Gamma^k_{ij} \frac{\partial}{\partial x^k} \right) \\
&= X^i \left(\frac{\partial Y^k}{\partial X^i} + Y^j \Gamma^k_{ij} \right) \frac{\partial}{\partial x^k},
\end{aligned}$$

where in the last adjustment, we have harmlessly renamed index j to k.

Remark 4.1.4. The above remark allows us to define *geodesics* on M in a manner similar to Section 3.2. More in detail: if we put $\gamma : t \mapsto \gamma(t)$ together with the requirement of self-parallelism

$$\nabla_{\frac{\mathrm{d}}{\mathrm{d}t}} \dot{\gamma} = 0,$$

we obtain the relations

$$\frac{\mathrm{d}^2 x^k}{\mathrm{d}t^2} + \Gamma^k_{ij} \frac{\mathrm{d}x^i}{\mathrm{d}t} \frac{\mathrm{d}x^j}{\mathrm{d}t} = 0.$$

In the rest of this section, we prove some basic properties of a covariant derivative. We keep the chart $(U, \mathbf{x})$ fixed in all what follows.

Definition 4.1.5. Let M be a smooth manifold, and let ∇ and $\widetilde{\nabla}$ be two covariant derivatives on U. Define the *difference* of the covariant derivatives ∇ and $\widetilde{\nabla}$ as $D(X, Y) := \nabla_X Y - \widetilde{\nabla}_X Y$.

Claim 4.1.6. *We can easily see that the difference $D(X, Y)$ is bilinear:*

(a) *For any $f, g \in C^\infty(U)$ it holds that*

$$\begin{aligned}
D(fX_1 + gX_2, Y) &= \nabla_{fX_1 + gX_2} Y - \widetilde{\nabla}_{fX_1 + gX_2} Y \\
&\overset{\text{(D1)}}{=} \nabla_{fX_1} v + \nabla_{gX_2} Y - \widetilde{\nabla}_{fX_1} Y - \widetilde{\nabla}_{gX_2} Y \\
&\overset{\text{(D2)}}{=} f \nabla_{X_1} Y + g \nabla_{X_2} Y - f \widetilde{\nabla}_{X_1} Y - g \widetilde{\nabla}_{X_2} Y \\
&= fD(X_1, Y) + gD(X_2, Y).
\end{aligned}$$

(b) *For any $f, g \in C^\infty(U)$ it holds that*

$$\begin{aligned}
D(X, fY_1 + gY_2) &= \nabla_X(fY_1 + gY_2) - \widetilde{\nabla}_X(fY_1 + gY_2) \\
&\overset{\text{(D3)}}{=} \nabla_X(fY_1) + \nabla_X(gY_2) - \widetilde{\nabla}_X(fY_1) - \widetilde{\nabla}_X(gY_2) \\
&\overset{\text{(D4)}}{=} X(f)Y_1 + f\nabla_X(Y_1) + X(g)Y_2 + g\nabla_X(Y_2) \\
&\quad - X(f)Y_1 - f\widetilde{\nabla}_X(Y_1) - X(g)Y_2 - g\widetilde{\nabla}_X(Y_2) \\
&= fD(X, Y_1) + gD(X, Y_2).
\end{aligned}$$

Remark 4.1.7. Conversely to Claim 4.1.6, if B is any bilinear map on smooth functions, and if ∇ is any covariant derivative, it is easy to see that the map $\widetilde{\nabla} := \nabla - B$ is a covariant derivative again: For any $f \in C^\infty(U)$ it holds that

(D1)

$$\begin{aligned}
\widetilde{\nabla}_{X_1+X_2}Y &= \nabla_{X_1+X_2}Y - B(X_1 + X_2) \\
&= \nabla_{X_1}Y + \nabla_{X_2}Y - B(X_1, Y) - B(X_2, Y) \\
&= \widetilde{\nabla}_{X_1}Y - \widetilde{\nabla}_{X_2}Y.
\end{aligned}$$

(D2)

$$\begin{aligned}
\widetilde{\nabla}_{fX}Y &= \nabla_{fX}Y - B(fX, Y) \\
&= f\nabla_X Y - fB(X, Y) \\
&= f\widetilde{\nabla}_X Y.
\end{aligned}$$

(D3)

$$\begin{aligned}
\widetilde{\nabla}_X(Y_1 + Y_2) &= \nabla_X(Y_1 + Y_2) - B(X, Y_1 + Y_2) \\
&= \nabla_X Y_1 + \nabla_X Y_2 - B(X, Y_1) - B(X, Y_2) \\
&= \widetilde{\nabla}_X Y_1 + \widetilde{\nabla}_X Y_2.
\end{aligned}$$

(D4)

$$\begin{aligned}
\widetilde{\nabla}_X(fY) &= \nabla_X(fY) - B(X, fY) \\
&= X(f)Y + f\nabla_X Y - fB(X, Y) \\
&= X(f)Y + f\widetilde{\nabla}_X Y.
\end{aligned}$$

Remark 4.1.8. Recall that every bilinear[1] map can be decomposed into its symmetric and alternating parts, i.e., put $D(X, Y) = S(X, Y) + A(X, Y)$, where

$$\begin{aligned}
S(X, Y) &= \frac{1}{2}\left(D(X, Y) + D(Y, X)\right), \\
A(X, Y) &= \frac{1}{2}\left(D(X, Y) - D(Y, X)\right).
\end{aligned}$$

Claim 4.1.9. *Let M be a smooth manifold of dimension n, and let ∇ and $\widetilde{\nabla}$ be two covariant derivatives on a chart $(U, \mathbf{x})$. Then the following are equivalent:*

(i) *∇ and $\widetilde{\nabla}$ have the same geodesics.*

[1]This is a trivial claim and in fact holds for all *multilinear* maps, see [2]. Bilinear map is then just a special case.

(ii) $\nabla_X X = \widetilde{\nabla}_X X$ *holds for all* X.

(iii) S *(i.e., the symmetric part of the difference* $D = \nabla - \widetilde{\nabla}$*) vanishes.*

PROOF. Let $x : \mathbb{R} \to U$, $t \mapsto x(t)$, be a curve. Now, let us break the equivalence into separate implications:

- (i) $\implies$ (ii): Recall that γ is a geodesic for ∇, if

$$\frac{\mathrm{d}^2\gamma^k}{\mathrm{d}t^2} + \Gamma^k_{ij}\frac{\mathrm{d}\gamma^i}{\mathrm{d}t}\frac{\mathrm{d}\gamma^j}{\mathrm{d}t} = 0$$

  holds. Analogously, γ is a geodesic fo $\widetilde{\nabla}$, if

$$\frac{\mathrm{d}^2\gamma^k}{\mathrm{d}t^2} + \widetilde{\Gamma}^k_{ij}\frac{\mathrm{d}\gamma^i}{\mathrm{d}t}\frac{\mathrm{d}\gamma^j}{\mathrm{d}t} = 0$$

  holds. Above, Γ^k_{ij} and $\widetilde{\Gamma}^k_{ij}$ are the Christoffel symbols belonging to ∇ and $\widetilde{\nabla}$, respectively.

  By the existence and uniqueness theorem for ordinary differential equations, see e.g., [6], for each $p \in U$ and each $v_p \in T_pM$, there is a geodesic such that $\gamma(p) = 0$ and $\mathrm{d}\gamma/\mathrm{d}t|_0 = v_p$. Thus, $\Gamma^k_{ij} = \widetilde{\Gamma}^k_{ij}$.

- (ii) $\implies$ (i): As we have established in Claim 2.5.6, at every t_0, any tangent vector X at the point $x(t_0)$ can be interpreted as a velocity vector $x'(t_0)$ of some smooth curve in U, passing through $x(t_0)$. Given the fact that $\nabla_X X = \widetilde{\nabla}_X X$ for any X, the geodesic equations yielded by putting[2] $\nabla_X X = 0$ must be the same for both covariant derivatives ∇ and $\widetilde{\nabla}$.

- (ii) $\implies$ (iii): Assumption $\nabla_X X = \widetilde{\nabla}_X X$ immediately also yields that $D(X, X) = 0$. Then

$$\begin{aligned} D(X+Y, X+Y) &= 0, \\ \underbrace{D(X,X)}_{=0} + D(X,Y) + D(Y,X) + \underbrace{D(Y,Y)}_{=0} &= 0. \end{aligned}$$

  Thus

$$D(Y, X) = -D(X, Y),$$

  i.e., D is an alternating bilinear form and hence its symmetrical part is equal to zero.

- (iii) $\implies$ (ii): The assumption $S = 0$ can be interpreted as $\nabla - \widetilde{\nabla} = D = A$. This means that $\nabla_X X - \widetilde{\nabla}_X X = D(X, X) = A(X, X) = 0$. Therefore $\nabla_X X = \widetilde{\nabla}_X X$.

■

Definition 4.1.10. Let M be a smooth manifold, and let ∇ be a covariant derivative on U. The map

$$T_\nabla(X, Y) := \nabla_X Y - \nabla_Y X - [X, Y],$$

where $[X, Y]f = X(f)Y - Y(f)X$ is the *Lie bracket* of vector fields, is called the *torsion* of the covariant derivative ∇.

[2]Thanks to the fact that $X = x'(t) = Dx|_{t_0}(\mathrm{d}/\mathrm{d}t|_{t_0})$, this is the exact same case of self-parallelism as we have seen in Section 3.2.

Remark 4.1.11. Lie bracket is an antisymmetric map:

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

Using this, we can see that also torsion is an antisymmetric map:

$$T_\nabla(Y,X) = \nabla_Y X - \nabla_X Y - [Y,X] = -(\nabla_X Y - \nabla_Y X - [X,Y]) = -T_\nabla(X,Y).$$

Remark 4.1.12. For torsion, it generally holds that

$$\begin{aligned} T_\nabla - T_{\widetilde{\nabla}} &= \nabla_X Y - \nabla_Y X - [X,Y] - \widetilde{\nabla}_X Y + \widetilde{\nabla}_Y X + [X,Y] \\ &= D(X,Y) - D(Y,X) \\ &= 2A, \end{aligned}$$

where A is the alternating part of the difference $D = \nabla - \widetilde{\nabla}$.

Theorem 4.1.13. *Let M be a smooth manifold. For any covariant derivative ∇ on U, there exists a unique torsion-free covariant derivative $\widetilde{\nabla}$ on U, such that ∇ and $\widetilde{\nabla}$ have the same geodesics.*

Proof. Set $\widetilde{\nabla} = \nabla - \frac{1}{2}T_\nabla$. Then

$$\begin{aligned} T_{\widetilde{\nabla}}(X,Y) &= \widetilde{\nabla}_X Y - \widetilde{\nabla}_Y X - [X,Y] \\ &= \nabla_X Y - \frac{1}{2}T_\nabla(X,Y) - \nabla_Y X + \frac{1}{2}T_\nabla(Y,X) - [X,Y] \\ &= T_\nabla(X,Y) - \frac{1}{2}T_\nabla(X,Y) + \frac{1}{2}T_\nabla(Y,X) \\ &= T_\nabla(X,Y) - \frac{1}{2}T_\nabla(X,Y) - \frac{1}{2}T_\nabla(X,Y) \\ &= 0, \end{aligned}$$

where in the fourth row we have utilised the result of Remark 4.1.11. ■

Remark 4.1.14. Let us make a comment on why torsion-free covariant derivatives are desirable. A non-vanishing torsion of ∇ means—very vaguely—that certain "loops" do not close. A geometry with non-zero torsion can therefore behave rather unexpectedly. Models of physics usually demand torsion-free covariant derivatives. See, e.g., [4] for the discussion of torsions. Our Theorem 4.1.13 states that one can *always* modify a local covariant derivative to a torsion-free one *without* changing the geodesics.

Remark 4.1.15. In Definition 4.1.2, we have defined covariant derivative in a specific chart $(U, \mathbf{x})$. One can obviously define a covariant derivative on the *entire manifold M* to be the map

$$\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$$

that satisfies (D1)—(D4) of Definition 4.1.2 with $U = M$.

Given such a derivative ∇ on M, one can define a covariant derivative $\widetilde{\nabla}$ on any chart $(U, \mathbf{x})$ such that the equality

$$(\nabla_X Y)\restriction_U = \widetilde{\nabla}_{(X\restriction_U)}(Y\restriction_U)$$

holds.

Conversely, having covariant derivatives $\nabla_{(U,\mathbf{x})}$ on all charts $(U, \mathbf{x})$, such that $\nabla_{(U,\mathbf{x})}$ and $\nabla_{(V,\mathbf{y})}$ are compatible whenever $U \cap V \neq \emptyset$, then one can define a covariant derivative on the whole manifold.

4.2 The Finsler bundle

Before we will be able to define connection forms abstractly, we have to introduce the so-called *Finsler bundle* T_2M of a smooth manifold M. Formally, the Finsler bundle is a smooth manifold T_2M equipped with two smooth projections

$$\begin{tikzcd} T_2M \arrow[r, "p_2"] \arrow[d, "p_1"'] & TM \\ TM & \end{tikzcd}$$

that make the square

$$\begin{tikzcd} T_2M \arrow[r, "p_2"] \arrow[d, "p_1"'] & TM \arrow[d, "\pi_M"] \\ TM \arrow[r, "\pi_M"'] & M \end{tikzcd}$$

commutative and that has a certain *universal property* making the above square a *pullback*. Before we state the property, let us show how a pullback of two maps between sets is formed.

Example 4.2.1. Let

$$\begin{tikzcd} & Y \arrow[d, "g"] \\ X \arrow[r, "f"'] & Z \end{tikzcd}$$

be a diagram of sets and mappings and let us consider the set $P = \{(x, y) \mid f(x) = g(y)\}$, equipped with the projections

$$\begin{aligned} p_1 : P &\to X, \\ (x, y) &\mapsto x, \end{aligned} \qquad \text{and} \qquad \begin{aligned} p_2 : P &\to Y, \\ (x, y) &\mapsto y. \end{aligned}$$

Then it is clear that the square

$$\begin{tikzcd} P \arrow[r, "p_2"] \arrow[d, "p_1"'] & Y \arrow[d, "g"] \\ X \arrow[r, "f"'] & Z \end{tikzcd}$$

commutes. Moreover: whenever *any* square

$$\begin{tikzcd} W \arrow[r, "p_Y"] \arrow[d, "p_X"'] & Y \arrow[d, "g"] \\ X \arrow[r, "f"'] & Z \end{tikzcd}$$

commutes, then there clearly is a *unique* map

$$\begin{aligned} h : W &\to P, \\ w &\mapsto (p_X(w), p_Y(w)), \end{aligned}$$

making both triangles

$$\begin{tikzcd} W \arrow[dr, "h"] \arrow[ddr, bend right, "p_X"'] & \\ & P \arrow[d, "p_1"] \\ & X \end{tikzcd} \qquad \begin{tikzcd} W \arrow[dr, "h"'] \arrow[drr, bend left, "p_Y"] & & \\ & P \arrow[r, "p_2"'] & Y \end{tikzcd}$$

commutative.

In fact, it is easy to see that the above *universal property* (i.e., the existence of a unique map h) determines the set P in the square

$$\begin{tikzcd} P \arrow[r, "p_2"] \arrow[d, "p_1"'] & Y \arrow[d, "g"] \\ X \arrow[r, "f"'] & Z \end{tikzcd}$$

up to a unique isomorphism (i.e., up to a bijection).

Definition 4.2.2. Define the *Finsler bundle* of a manifold M to be the pullback

$$\begin{tikzcd} T_2M \arrow[r, "p_2"] \arrow[d, "p_1"'] & TM \arrow[d, "\pi_M"] \\ TM \arrow[r, "\pi_M"'] & M \end{tikzcd}$$

in the realm of smooth manifolds and smooth maps.

Remark 4.2.3. Definition 4.2.2 is in the spirit of Category Theory. It can be unravelled into more elementary terms as follows:

- the points of T_2M are triples (p, v_p, w_p) with $p \in M$, $v_p \in T_pM$ and $w_p \in T_pM$.
- the topology of T_2M is that of a subspace of $TM \times TM$.
- the maps p_1 and p_2 are the obvious (smooth) projections

$$(p, v_p, w_p) \mapsto (p, v_p) \qquad \text{and} \qquad (p, v_p, w_p) \mapsto (p, w_p),$$

  respectively.

We use the universal property of the Finsler bundle to define a certain smooth map $TTM \xrightarrow{\tau_M} T_2M$ that we will use later.

Proposition 4.2.4. *For every manifold M, the square*

$$\begin{tikzcd} TTM \arrow[r, "T\pi_M"] \arrow[d, "\pi_{TM}"'] & TM \arrow[d, "\pi_M"] \\ TM \arrow[r, "\pi_M"'] & M \end{tikzcd}$$

commutes. Thus, there is a unique smooth map

$$TTM \xrightarrow{\tau_M} T_2M$$

making the triangles

$$
\begin{tikzcd}
TTM \arrow[dr, "\tau_M"] \arrow[ddr, bend right, "\pi_{TM}"'] & \\
& T_2M \arrow[d, "p_1"] \\
& TM
\end{tikzcd}
\qquad
\begin{tikzcd}
TTM \arrow[dr, "\tau_M"'] \arrow[drr, bend left, "T\pi_M"] & & \\
& T_2M \arrow[r, "p_2"'] & TM
\end{tikzcd}
$$

commutative.

Proof. The proposition will follow from the universal property of a pullback once we show that the square

$$
\begin{tikzcd}
TTM \arrow[r, "T\pi_M"] \arrow[d, "\pi_{TM}"'] & TM \arrow[d, "\pi_M"] \\
TM \arrow[r, "\pi_M"'] & M
\end{tikzcd}
$$

commutes. This is a special case of Example 2.4.5. ■

We are now ready to state the abstract definition of the connection form on a manifold.

Definition 4.2.5. A smooth map $C : T_2M \to TTM$ is called the *connection form*, if the following three properties hold:

(a) The diagram

$$
\begin{tikzcd}
T_2M \arrow[r, "C"] \arrow[dr, "\mathrm{id}"'] & TTM \arrow[d, "\tau_M"] \\
& T_2M
\end{tikzcd}
$$

commutes, where τ_M is the mapping of Proposition 4.2.4.

(b) The diagram

$$
\begin{tikzcd}
T_2M \arrow[rr, "C"] \arrow[dr, "p_1"'] & & TTM \arrow[dl, "\pi_{TM}"] \\
& TM &
\end{tikzcd}
$$

commutes and C is linear in every fibre.[3]

(c) The diagram

$$
\begin{tikzcd}
T_2M \arrow[rr, "C"] \arrow[dr, "p_2"'] & & TTM \arrow[dl, "T\pi_M"] \\
& TM &
\end{tikzcd}
$$

commutes and C is linear in every fibre.

[3]Recall that, for a map $f : X \to Y$, the *fibre* of an element $y \in Y$ under f is the inverse image of the singleton set $\{y\}$. Also, we omit the set notation and simply write $f^{-1}(y) = \{x \in X \mid f(x) = y\}$.

Remark 4.2.6. Definition 4.2.5 is best understood via the trivialisation technique. In $(U, \mathbf{x})$, we have

$$C : (p, v, w) \mapsto (p, v, w, -\Gamma_p(v, w))$$

due to condition (a). Conditions (b) and (c) then state that $\Gamma_p(v, w)$ is linear in v and w, respectively. After renaming, we have obtained the connection form

$$(p, k, u) \mapsto (p, k, u, -\Gamma_p(k, u))$$

familiar from Section 3.4.

Remark 4.2.7. There exists an equivalent description of the connection form that is more pleasant to work with. Namely, define a smooth map

$$TTM \xrightarrow{K} TM$$

in a local trivialisation by putting

$$((p, u), (k, v)) \mapsto (p, v + \Gamma_p(k, u))$$

and observe that it has the following two properties:

(a) The square

$$\begin{tikzcd} TTM \arrow[r, "K"] \arrow[d, "T\pi_M"'] & TM \arrow[d, "\pi_M"] \\ TM \arrow[r, "\pi_M"'] & M \end{tikzcd}$$

commutes and it is linear in every fibre. Indeed

$$\begin{tikzcd} ((p, u), (k, v)) \arrow[r, mapsto, "K"] \arrow[d, mapsto, "T\pi_M"'] & (p, v + \Gamma_p(k, u)) \arrow[d, mapsto, "\pi_M"] \\ (p, \underbrace{\pi'_M|_{(p,u)}(k, v)}_{=k}) \arrow[r, "\pi_M"'] & p \end{tikzcd}$$

(b) The square

$$\begin{tikzcd} TTM \arrow[r, "K"] \arrow[d, "\pi_{TM}"'] & TM \arrow[d, "\pi_M"] \\ TM \arrow[r, "\pi_M"'] & M \end{tikzcd}$$

commutes and it is linear in every fibre. Indeed

$$\begin{tikzcd} ((p, u), (k, v)) \arrow[r, mapsto, "K"] \arrow[d, mapsto, "\pi_{TM}"'] & (p, v + \Gamma_p(k, u)) \arrow[d, mapsto, "\pi_M"] \\ (p, u) \arrow[r, "\pi_M"'] & p \end{tikzcd}$$

5. Equivalence of covariant derivatives and connection forms

Recall from Chapter 4 that we have introduced two concepts that deal with "infinitesimally small changes" of vectors in the tangent bundle of a smooth manifold M:

(1) The covariant derivative $\nabla_U V$ describes changes of a vector field V "in the direction" of a vector field U.

(2) The connection form $K : TTM \to TM$ that arises from the fact that we can "connect" vectors in "nearby" fibres of the tangent bundle TM.

The goal of this chapter is to prove that the two concepts above are equivalent to each other. In the final section of the current chapter, we indicate the techniques that lead to an abstract treatment of "infinitesimally small changes", i.e., we show the style of reasoning that leads to the introduction of the so-called *tangent categories*.

5.1 Covariant derivatives yield connection forms

Recall from Remark 4.1.15 that a *covariant derivative* on a manifold M is a mapping

$$\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$$

that, for $f \in C^\infty(M)$, satisfies

(D1) $\nabla_{X_1+X_2} Y = \nabla_{X_1} Y + \nabla_{X_2} Y$.

(D2) $\nabla_{fX} Y = f \nabla_X Y$.

(D3) $\nabla_X (Y_1 + Y_2) = \nabla_X Y_1 + \nabla_X Y_2$.

(D4) $\nabla_X (fY) = Xf \cdot Y + f \cdot \nabla_X Y$.

We will now show that, given a covariant derivative on a manifold M, one can define a connection form K on M. The trick will be, of course, to define K locally, using the trivialisation technique.

The covariant derivative in a trivialisation gives a connection form. Let $(U, \mathbf{x})$ be a chart on M such that $\pi^{-1}[U] \cong U \times \mathbb{R}^n$, i.e., such that $\pi^{-1}[U]$ is trivial. Given smooth vector fields $M \xrightarrow{X} TM$, $M \xrightarrow{Y} TM$, then for any $p \in U$, we can write

$$X(p) = (p, F(p)), \qquad\qquad Y(p) = (p, G(p)),$$

where $U \xrightarrow{F} \mathbb{R}^n$ and $U \xrightarrow{G} \mathbb{R}^n$ are smooth. Then $(\nabla_X Y)(p) = (p, \delta(p))$, where $U \xrightarrow{\delta} \mathbb{R}^n$ is smooth and

$$\delta(p) = DG|_p(F(p)) + \Gamma_p(F(p), G(p)),$$

where $\Gamma_p : \mathbb{R}^n \to \mathbb{R}^n$ is a bilinear map. To see that this is indeed what we saw in Chapter 4, let us write

$$F(p) = f^i(p)e_i, \qquad\qquad G(p) = g^j(p)e_j,$$

where $(e_1, \dots, e_n)$ is the canonical basis of $\mathbb{R}^n$. Then

$$DG|_p(F(p)) = f^i(p)DG|_p(e_i) = f^i(p)\frac{\partial g^j}{\partial x^i}e_j$$

by linearity of $DG|_p$ and by the description of $DG|_p$ as a Jacobi matrix. Furthermore, bilinearity of Γ_p states that $\Gamma_p(F(p), G(p))$ has the form

$$\Gamma_p\left(f^i(p)e_i, g^j(p)e_j\right) = f^i(p)g^j(p)\Gamma_p(e_i, e_j).$$

Since $\Gamma_p(e_i, e_j)$ lies in $\mathbb{R}^n$, it has necessarily the form

$$\Gamma_p(e_i, e_j) = \Gamma^k_{ij}\Big|_p e_k \qquad \text{for some } \Gamma^k_{ij}\Big|_p . \tag{5.1.1}$$

Hence

$$\delta(p) = f^i(p)\left.\frac{\partial g^j}{\partial x^i}\right|_p \mathrm{e}_j + f^i(p)g^j(p)\,\Gamma^i_{ij}\big|_p e_k = \left(f^i(p)\left.\frac{\partial g^k}{\partial x^i}\right|_p + f^i(p)g^j(p)\,\Gamma^k_{ij}\Big|_p\right)e_k.$$

This is indeed the description of the covariant derivative from Chapter 4.

Now define

$$K(p, u, k, v) = (p, v + \Gamma_p(k, u))\,.$$

It is then trivial to see that K (locally) satisfies all of the required properties of a connection form.

As it can be easily seen from (5.1.1), conversely, any choice of coefficients $\Gamma^k_{ij}|_p$ yields a bilinear map Γ_p.

5.2 Connection forms yield covariant derivatives

Working in a local trivialisation $(U, \mathbf{x})$ again, any connection form $TTM \xrightarrow{K} TM$ has the form

$$K : (p, u, k, v) \mapsto (p, v + \Gamma_p(k, u))$$

for a bilinear map $\Gamma_p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, depending smoothly on p.

We will now consider two vector fields $M \xrightarrow{X} TM$, $M \xrightarrow{Y} TM$ and form the composite

$$M \xrightarrow{X} TM \xrightarrow{TY} TTM \xrightarrow{K} TM.$$

We claim that the above composite gives, in our trivialisation $(U, \mathbf{x})$, a covariant derivative.

By setting $X(p) = (p, F(p))$ and $Y(p) = (p, G(p))$ for all $p \in U$, we obtain smooth maps $U \xrightarrow{F} \mathbb{R}^n$, $U \xrightarrow{G} \mathbb{R}^n$ as in the previous section. Then the action of the above composite upon p can be written as

$$p \mapsto (p, F(p)) \mapsto ((p, G(p))), (F(p), DG|_p(F(p))) \mapsto (p, DG|_p(F(p)) + \Gamma_p(F(p), G(p)).$$

Thus, we have (locally) defined a covariant derivative.

That the above two processes are inverse to each other follows easily. Thus we can formulate the main result of this chapter.

Theorem 5.2.1. *Suppose M is a smooth manifold. Then to give a covariant derivative on M is to give a connection form on M.*

5.3 More general setting

There are at least two directions in which the results of this text can be generalised. The first one is entirely in the spirit of differential geometry, whereas the second uses quite abstract algebraic methods.

Connections and the derivatives in vector bundles. The tangent bundle TM of a smooth manifold M is a smooth manifold again and it comes equipped with a smooth map $\pi_M : TM \to M$ such that each fibre $\pi^{-1}(p) = T_pM$ is a vector space. A generalisation of this phenomenon leads to the notion of a *vector bundle* as a smooth map $\pi_{E,M} : E \to M$ such that every fibre $\pi^{-1}_{E,M}(p)$ is a vector space. There are, of course, certain additional *coherence conditions* that have to hold in order for $\pi_{E,M}$ to be a vector bundle. These extra conditions allow one to define and study covariant derivatives and connection forms on vector bundles. See, e.g., [7], [11].

Tangent categories. One of the results presented in Chapter 2 of this thesis, namely Claim 2.4.6, states that $TM \xrightarrow{TF} TN$ is a smooth map, whenever $M \xrightarrow{F} N$ is smooth. Moreover, Claim 2.4.7 says that $T(\mathrm{id}_M) = \mathrm{id}_{TM}$ and $T(G \circ F) = TG \circ TF$, whenever M is a smooth manifold and F, G are composable smooth maps. In the language of Category Theory, this means that T is an *endofunctor* of the *category* of all smooth manifolds and all smooth maps. Several other important concepts of this thesis have categorical nature, for example, the commutativity of the square in Example 2.4.5 states that π_M is a component of a natural transformation from the tangent functor T to the identity functor, the construction of the Finsler bundle T_2M from Definition 4.2.2 is given as a pullback, etc.

The first to introduce these concepts abstractly was Jiří Rosický [10]. Since his originating work on *tangent categories*, a lot of papers that develop differential geometry at this level of abstraction have appeared. Of particular relevance to this text is the recent paper [9] where covariant derivatives and connection forms are presented in tangent categories.

Conclusion

This thesis aimed to acquaint the reader with the subject matter of defining a differential structure on a smooth manifold. The text was constructed on a robust framework of multivariable calculus in Euclidean spaces where the issue is far simpler. However, we have successfully introduced a tangent structure which allowed us to define appropriate notions locally. As we have shown, this "piecewise construction" can be "glued together" using a connection. As a final result of this thesis, we have shown that a connection is equivalently given by a covariant derivative.

This theory serves a significant role in both theoretical and applied fields of physics. As we have already mentioned in the introduction of the thesis, there is a vast area of physics concerning the general theory of relativity which pronouncedly requires the notion of parallel transport. Apart from this obvious and already established example, we could name many more, e.g., in electrical engineering, transition of perpendicular frames along the direction of wave propagation in a waveguide. In contemporary science, it is clear that the assumption of a flat spacetime is very restrictive which immediately grants this branch of mathematics a worthy place among other brilliant discoveries.

Future work. In this thesis, we have established the equivalence of two additional structures on a smooth manifold: a *covariant derivative* and a *connection form.*

We have used rather traditional techniques of differential geometry to achieve this goal. We have a feeling, though, that the techniques of *tangent categories* are extremely interesting and that they could provide us with unexpected interrelations of various concepts of "classical" differential geometry. This must, however, be left for future work.

Bibliography

[1] Henri Cartan, *Differential calculus*, Hermann, 1983. 19, 36

[2] Jean Dieudonné, *Treatise on analysis*, volume 3, Academic Press, 1972. 35, 39

[3] Ryszard Engelking, *General topology*, Heldermann, Berlin, 1989. 5, 7, 8

[4] Theodore Frankel, *The geometry of physics: an introduction*, 3rd ed, Cambridge University Press, 2011. 41

[5] David Gauld, *Non-metrisable manifolds*, Springer, Singapore, 2014. 8

[6] Philip Hartman, *Ordinary differential equations*, John Wiley & Sons, New York, 1964. 40

[7] Jeffrey M. Lee, *Manifolds and differential geometry*, volume 107, American Mathematical Society, Rhode Island, 2009. 3, 8, 9, 11, 14, 27, 48

[8] John M. Lee, *Introduction to smooth manifolds*, Springer, New York, 2012. 27

[9] Rory B. B. Lucyshyn-Wright, On the geometric notion of connection and its expression in tangent categories, *Theory Appl. Categ.* 33 (2018), 832-866. `http://www.tac.mta.ca/tac/volumes/33/28/33-28.pdf`, Accessed: May 2022. 48

[10] Jiří Rosický, Abstract tangent functors, *Diagrammes* 12 (1984): JR1-JR11. `http://eudml.org/doc/91746`, Accessed: May 2022. 48

[11] József Szilasi, Dávid Cs Kertész, Rezső L. Lovas, *Connections, sprays and Finsler structures*, World Scientific Publishing Company, 2013. 27, 48

Index