Geometry in Physics*

Winter Semester 2018/19

Stefan Teufel
Mathematisches Institut
Uni Tübingen

February 4, 2019

* These lecture notes are not a substitute for attending the lectures, but should be used only in parallel! If you find typos or more serious errors, please let me know.

Contents

1 Manifolds

I assume that you know calculus for functions and vector fields on the "flat" euclidean space $\mathbb{R}^n$. In this course you learn how to do calculus on "curved" spaces, called manifolds, that still look locally like some possibly curved piece of $\mathbb{R}^n$. In order to translate concepts like differentiation and integration to manifolds, we need to understand which mathematical structures of $\mathbb{R}^n$ underlie the respective concepts and how we can lift them to these more general spaces. In the process of doing this, we will gain a more geometric view of calculus that is necessary to understand most of modern physics.

Recall that one way to define **continuity** of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is to require that for every open set $O \subset \mathbb{R}^m$ its pre-image $f^{-1}(O) := \{x \in \mathbb{R}^n \mid f(x) \in O\} \subset \mathbb{R}^m$ under f is open as well. More generally, a map $f : X \to Y$ between topological spaces is continuous, if pre-images of open sets under f are open, i.e. if for every open set $O \subset Y$ the set $f^{-1}(O) \subset X$ is open. In other words, to speak of continuity, we need the structure of a topological space.

1.1 Definition. **Topological space**

A **topological space** is a pair $(M, \mathcal{O})$, where M is a set and $\mathcal{O}$ is a set of subsets of M. (The elements $O \in \mathcal{O}$ of $\mathcal{O}$ are thus subsets of M, called the *open sets.*) One requires that

(i) $\emptyset$ and M are open, i.e. $\emptyset \in \mathcal{O}$ and $M \in \mathcal{O}$,

(ii) arbitrary unions of open sets are open,

(iii) the intersection of two open sets is open.

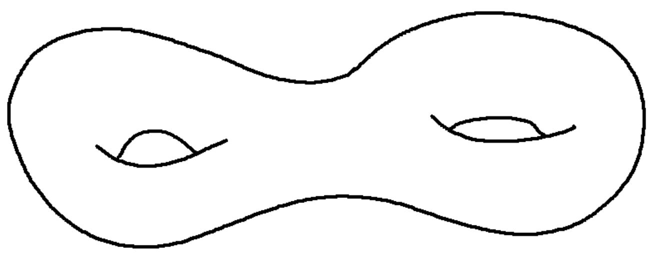
A topological space $(M, \mathcal{O})$ has the **Hausdorff property** if $\mathcal{O}$ separates points in M, i.e. if for any pair $x, y \in M$ of different points $x \neq y$ there exist open sets $V, U \in \mathcal{O}$ such that $x \in V$, $y \in U$, but $V \cap U = \emptyset$.

A topological space $(M, \mathcal{O})$ is called **second countable** if there exists a countable basis for the topology of M, that is a countable set $\mathcal{B} \subset \mathcal{O}$ such that any open set can be written as a union of sets in $\mathcal{B}$. $\diamond$

1.2 Example. **Euclidean space $\mathbb{R}^n$**

Consider euclidean space $\mathbb{R}^n$ with the euclidean metric $d : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$, $(x, y) \mapsto d(x, y) := \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$. Equipped with the topology induced by the metric[1], $\mathbb{R}^n$ is a second countable Hausdorff space: For $x \neq y \in \mathbb{R}^n$ and $\varepsilon := d(x, y)/2$ the two balls $B_\varepsilon(x)$ and $B_\varepsilon(y)$ are disjoint open sets containing x resp. y. Moreover, the open balls $B_\varepsilon(x)$ with rational radii $\varepsilon \in \mathbb{Q}$ and rational centers $x = (x_1, \ldots, x_n) \in \mathbb{Q}^n$, form a countable basis for the topology. $\diamond$

Heuristically speaking, a topological manifold is a topological space that "looks locally like" euclidean space $\mathbb{R}^n$, but might have a completely different shape globally. A curved surface as depicted here is an example of a topological manifold with a topology that looks locally like that of $\mathbb{R}^2$.



[1] In a metric space X a set $U \subset X$ is called open, if for any $x \in U$ there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subset U$, i.e. such that the ball $B_\varepsilon(x) := \{y \in X \mid d(y, x) < \varepsilon\}$ of radius ε around x lies still completely in U.

To make the notion of “topological spaces that look locally alike” precise, one uses the concept of homeomorphisms.

1.3 Definition. **Continuous maps and homeomorphisms**

A map $f : M \to N$ between topological spaces $(M, \mathcal{O})$ and $(N, \mathcal{P})$ is called **continuous** if preimages of open sets are open, i.e. if $U \in \mathcal{P}$ implies $f^{-1}(U) \in \mathcal{O}$. It is called a **homeomorphism**, if it is bijective (i.e. one-to-one) and bi-continuous (i.e. continuous with continuous inverse). $\diamond$

1.4 Definition. **Topological manifold**

A second countable topological Hausdorff space M is called a **topological manifold** of dimension n, if any point $x \in M$ has a neighbourhood that is homeomorphic to an open set in $\mathbb{R}^n$. The latter means that for any point $x \in M$ there exist an open set $V \subset M$ with $x \in V$, an open set $U \subset \mathbb{R}^n$, and a homeomorphism $f : V \to U$. $\diamond$

Remark. Don’t worry if you are not familiar with topological spaces, these concepts will not be central to the following. Just remember that in a topological space we have the notion of open sets and concepts derived from there, like continuous functions, compactness etc. The additional conditions on the topology in the above definition of a topological manifold ensure that there are not too few open sets (Hausdorff) and not too many (second countable). Note also that the properties of being Hausdorff and second countable do not follow from the property of being locally homeomorphic to the second countable Hausdorff space $\mathbb{R}^n$, see also example 1.8 (f). $\diamond$

Since differentiability is a local property, we can use the fact that a manifold looks locally like a piece of $\mathbb{R}^n$ to also lift the differentiable structure of $\mathbb{R}^n$ to manifolds. Recall for the following definition that a map $f : V \to U$ between open sets $V \subset \mathbb{R}^n$ and $U \subset \mathbb{R}^m$ is in $C^r(V, U)$ or short a C^r-map, if it is r-times continuously differentiable. It is a C^r**-diffeomorphism**, if it is bijective, a C^r-map, and the inverse $f^{-1} : U \to V$ is also a C^r-map.

1.5 Definition. **Charts**

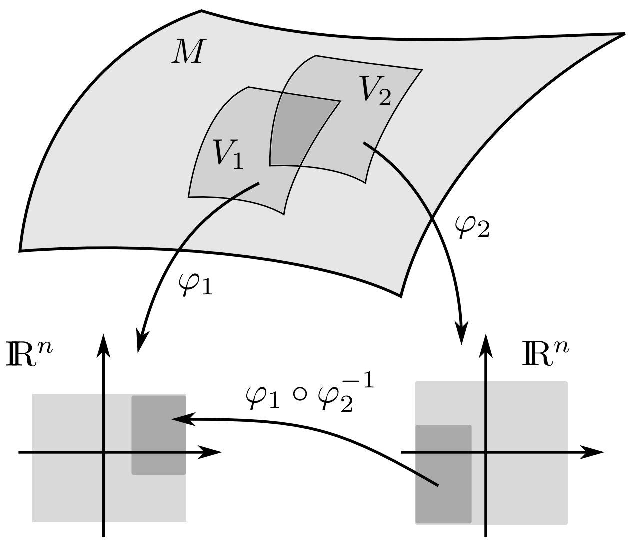
Let M be a topological manifold of dimension n. A **chart** on M is a pair (V, φ) of an open set $V \subset M$ and a homeomorphism $\varphi : V \to U \subset \mathbb{R}^n$ onto an open subset $U = \varphi(V)$ of $\mathbb{R}^n$. Charts are also called **coordinate charts**, because a chart (V, φ) allows one to label points $x \in V$ by “coordinate vectors” $\varphi(x) \in \mathbb{R}^n$.

Two charts (V_1, φ_1) and (V_2, φ_2) on M are called C^r**-compatible**, if either $V_1 \cap V_2 = \emptyset$, or if the composition

$$\varphi_1 \circ \varphi_2^{-1} : \varphi_2(V_1 \cap V_2) \to \varphi_1(V_1 \cap V_2)$$

is a C^r-diffeomorphism of open sets of $\mathbb{R}^n$. The maps $\varphi_1 \circ \varphi_2^{-1}$ and $\varphi_2 \circ \varphi_1^{-1}$ are called **transition maps**. $\diamond$

Remark. Note that for any two charts (V_1, φ_1) and (V_2, φ_2) the transition maps $\varphi_1 \circ \varphi_2^{-1}$ and $\varphi_2 \circ \varphi_1^{-1}$ are necessarily homeomorphisms because φ_1 and φ_2 are homeomorphisms. $\diamond$



1.6 Definition. C^r**-atlas**

A set of pairwise C^r-compatible charts $\mathcal{A} = \{(V_j, \varphi_j) \mid j \in J\}$ that cover M, i.e. $M = \bigcup_{j \in J} V_j$, is called a C^r**-atlas**.

Two C^r-atlases are **equivalent**, if any two charts in the atlases are C^r-compatible. $\diamond$

Remark. It is straightforward to check that equivalence of atlases is really an equivalence relation. The only non-obvious property to check is transitivity. Let $\mathcal{A}_j$, $j = 1, 2, 3$, be atlases such that $\mathcal{A}_1$ and $\mathcal{A}_2$ are compatible and $\mathcal{A}_2$ and $\mathcal{A}_3$ are compatible. To see that also $\mathcal{A}_1$ and $\mathcal{A}_3$ are compatible, let $(V_1, \varphi_1) \in \mathcal{A}_1$ and $(V_3, \varphi_3) \in \mathcal{A}_3$. Then

$$\varphi_1 \circ \varphi_3^{-1} : \varphi_3(V_1 \cap V_3) \to \varphi_1(V_1 \cap V_3)$$

is a homeomorphism and it suffices to check differentiability locally. But for any $(V_2, \varphi_2) \in \mathcal{A}_2$ with $V_1 \cap V_2 \cap V_3 \neq \emptyset$ we have that

$$\varphi_1 \circ \varphi_3^{-1} = (\varphi_1 \circ \varphi_2^{-1}) \circ (\varphi_2 \circ \varphi_3^{-1})$$

is the composition of two C^r-functions and thus itself a C^r-function on $\varphi_3(V_1 \cap V_2 \cap V_3)$.

Thus every atlas lies in a unique equivalence class of atlases. ◇

1.7 Definition. **Differentiable manifold**

A topological manifold M together with an equivalence class of C^r-atlases is called a **differentiable manifold** (or more precisely a C^r-manifold). This equivalence class of atlases is often called a **differentiable structure** for M. ◇

Remark. In view of the previous remark it suffices to provide one atlas in order to specify a differentiable manifold. More precisely, let M be a second countable Hausdorff space. Then the structure of a topological manifold and a unique differentiable structure on M are defined by providing one atlas of compatible charts, i.e. open sets $V_i \subset M$ with $\cup_i V_i = M$ and homeomorphisms $\varphi_i : V_i \to \varphi_i(V_i) \subset \mathbb{R}^n$ with $\varphi_i \circ \varphi_j^{-1} \in C^r$ for all i, j. Moreover, on a second countable differentiable manifold one can always find a countable atlas.[2] ◇

1.8 Examples. (a) Any open subset O of euclidean space $\mathbb{R}^n$ is a differentiable manifold in a natural sense: just pick the single chart (O, id) as an atlas. Indeed, any open subset O of a differentiable manifold M is again a differentiable manifold in a natural sense: just restrict the charts of M to $O \subset M$.

(b) The unit circle $S^1 := \{x \in \mathbb{R}^2 \mid \|x\| = 1\} \subset \mathbb{R}^2$ with the relative topology[3] is a 1-dimensional topological manifold. The four charts $(V_j, \varphi_j)_{j=1,\dots,4}$ with

$$\begin{aligned}
\varphi_1 : V_1 = \{x_1 > 0\} \to (-1, 1), \quad & x \mapsto \varphi_1(x) = x_2 \\
\varphi_2 : V_2 = \{x_1 < 0\} \to (-1, 1), \quad & x \mapsto \varphi_2(x) = x_2 \\
\varphi_3 : V_3 = \{x_2 > 0\} \to (-1, 1), \quad & x \mapsto \varphi_3(x) = x_1 \\
\varphi_4 : V_4 = \{x_2 < 0\} \to (-1, 1), \quad & x \mapsto \varphi_4(x) = x_1
\end{aligned}$$

provide the local homeomorphisms to $\mathbb{R}$ and define a differentiable structure on S^1.

(c) Note that definition 1.7 does not require M to be embedded into some ambient space, like S^1 into $\mathbb{R}^2$. We can, for example, define the "same" differentiable manifold S^1 by equipping

[2] To see this let $\mathcal{A} = \{(V_i, \varphi_i)\}_{i \in I}$ be any atlas for M and $\{U_k\}_{k \in \mathbb{N}}$ a countable base for the topology of M. Then for every $k \in \mathbb{N}$ there is an $i(k) \in I$ such that $U_k \subset V_{i(k)}$. Hence, $\{(U_k, \varphi_{i(k)}|_{U_k})\}_{k \in \mathbb{N}}$ is an equivalent countable atlas for M.

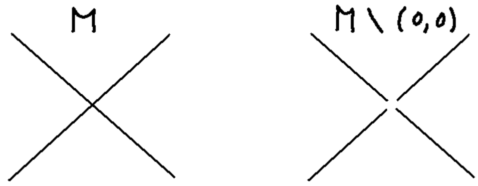
[3] Let $(M, \mathcal{O})$ be a topological space and $N \subset M$. Then the relative topology on N is $\mathcal{P} := \{U \subset N \mid \exists O \in \mathcal{O} : U = O \cap N\}$.

the topological space $\mathbb{R}/\mathbb{Z}$ [4]

$$\varphi_1 : \mathbb{R}/\mathbb{Z} \setminus \{[0]\} \to (0,1) \quad \text{and} \quad \varphi_2 : \mathbb{R}/\mathbb{Z} \setminus \{[\tfrac{1}{2}]\} \to (-\tfrac{1}{2}, \tfrac{1}{2})$$

that map $[x] \in \mathbb{R}/\mathbb{Z}$ to its representative in $[0,1)$ or $[-\frac{1}{2}, \frac{1}{2})$ respectively. More precisely, the manifold obtained in this way is diffeomorphic to S^1, cf. definition 1.13.

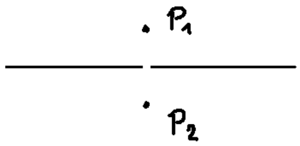
(d) The n^2 elements of an $n \times n$-matrix define a point in $\mathbb{R}^{n^2}$. We can thus identify the set $\mathrm{Mat}(n)$ of $n \times n$-matrices with $\mathbb{R}^{n^2}$ and thereby obtain for $\mathrm{Mat}(n)$ the structure of a differentiable manifold. The invertible matrices A, i.e. $\{A \in \mathrm{Mat}(n) \,|\, \det A \neq 0\}$, form an open subset of $\mathrm{Mat}(n)$ and thus, according to (a), again a differentiable manifold, the group $GL(n)$.



(e) $M = \{x \in \mathbb{R}^2 \,|\, |x_1| = |x_2|\}$ with the induced topology is not a topological manifold. This is because there is no connected open set containing $(0,0) \in M$ that can be mapped homeomorphically to an open connected set (i.e. an interval) in $\mathbb{R}$. To see this note that every connected open set containing $(0,0) \in M$ dissociates into four disconnected pieces when removing the point $(0,0)$, while an open interval in $\mathbb{R}$ dissociates into two pieces when removing a single point. But the number of connected components is invariant under homeomorphisms.

(f) The requirement that the topology of a manifold is Hausdorff is not redundant, i.e. it does not follow from being locally homeomorphic to $\mathbb{R}^n$ (whose topology is indeed Hausdorff). **Example:** Let $M = (\mathbb{R} \setminus \{0\}) \cup \{p_1\} \cup \{p_2\}$ be equipped with the two charts

$$\varphi_j : \mathbb{R}\setminus\{0\} \cup \{p_j\} \to \mathbb{R}\,, \quad \varphi_j(x) = \begin{cases} x & \text{if } x \neq p_j \\ 0 & \text{if } x = p_j\,. \end{cases}$$



Then the transition maps $\varphi_1 \circ \varphi_2^{-1} = \varphi_2 \circ \varphi_1^{-1} = \mathrm{id}|_{\mathbb{R}\setminus\{0\}}$ are diffeomorphisms, but the induced topology[5] on M turns φ_1 and φ_2 into homeomorphisms but it is not Hausdorff: $\{p_1\}$ and $\{p_2\}$ have no disjoint neighbourhoods. Note that in the induced topology a set $B \subset M$ is open, iff there exist open sets $O_1, O_2 \subset \mathbb{R}$ such that $B = \varphi_1^{-1}(O_1) \cup \varphi_2^{-1}(O_2)$.

(g) Given two manifolds M_1 and M_2 one can form the **product manifold** $M_1 \times M_2$. To this end one equips the cartesian product $M_1 \times M_2$ with the product topology and then covers this space with product charts of the form $(V_1, \varphi_1) \times (V_2, \varphi_2) = (V_1 \times V_2, (\varphi_1, \varphi_2))$, where (V_1, φ_1) and (V_2, φ_2) are charts from atlases of M_1 resp. M_2.

[4] Let M be a topological space and $f : M \to N$ surjective. Then the quotient topology on N is given by defining $U \subset N$ to be open iff its preimage $f^{-1}(U) \subset M$ is open. If $\sim$ is an equivalence relation on M, then the quotient $M/\sim$ is the set of equivalence classes $[x] := \{y \in M \,|\, y \sim x\}$ and the projection $p : M \to M/\sim$, $x \mapsto [x]$, is surjective. Hence, with respect to the quotient topology a set $U \in M/\sim$ is open if $\cup_{[x]\in U}[x] \subset M$ is open. The notation $\mathbb{R}/\mathbb{Z}$ stands for $\mathbb{R}/\sim$ where the equivalence relation is defined by $x \sim y \Leftrightarrow x - y \in \mathbb{Z}$. I.e. two points in $\mathbb{R}$ are equivalent if they lie on the same orbit of the canonical group action of $\mathbb{Z}$ acting on $\mathbb{R}$. The equivalence classes are all of the form $[x] = \{\ldots, x-2, x-1, x, x+1, x+2, \ldots\}$ and in any interval of the form $[a, a+1)$ their lies exactly one representative of each equivalence class.

[5] Let M be a set, $(N, \mathcal{P})$ a topological space, and $f : M \to N$ a map. Then f induces a topology on M be putting $\mathcal{O}_f := \{f^{-1}(U) \,|\, U \in \mathcal{P}\}$. It is the smallest topology on M such that f is a continuous function. In case of several maps $f_i : M \to N$, as in the example, the induced topology is the smallest topology on M containing all $\mathcal{O}_{f_i}$.

1.9 Remarks. (a) In the physics literature a chart is often called a local coordinate system. The inverse φ^{-1} of a chart φ is also called a local parametrisation.

(b) The fact that the differentiable structure of a manifold is defined through an equivalence class of atlases, and not through a single one, shows that there are no preferred coordinate charts on a manifold. All coordinate systems compatible with the differentiable structure are on an equal footing.

(c) Note that a differentiable manifold does not yet have a metric structure. Distances between points are not defined.

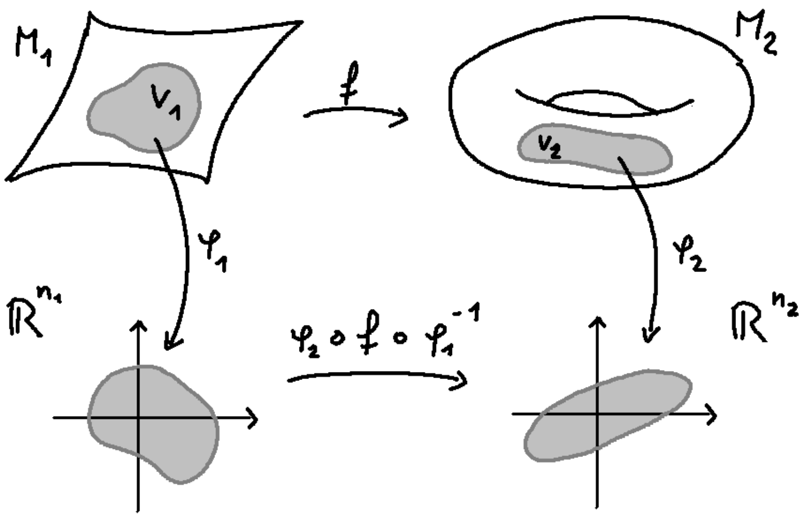
1.10 Convention. From now on, by a chart on a differentiable manifold we always mean a chart that is compatible with the differentiable structure, i.e. a chart from an atlas within the equivalence class of atlases defining the differentiable structure.

1.11 Definition. Differentiable maps

Let M_1 and M_2 be C^r-manifolds with $\dim M_1 = n_1$ and $\dim M_2 = n_2$. A continuous map $f : M_1 \to M_2$ is **p-times differentiable** (where $p \leq r$), if for any chart (φ_1, V_1) on M_1 and for any chart (φ_2, V_2) on M_2 the map

$$\varphi_2 \circ f \circ \varphi_1^{-1} : \mathbb{R}^{n_1} \supset \varphi_1\left(V_1 \cap f^{-1}(V_2)\right) \to \mathbb{R}^{n_2}$$

is p-times continuously differentiable. The set of all p-times differentiable functions $f : M_1 \to M_2$ is denoted by $C^p(M_1, M_2)$.



1.12 Convention. Whenever we speak about a differentiable manifold in the following, we always mean a C^r-manifold with $r \geq 1$ large enough for the statements in the corresponding context to make sense. The notion **smooth manifold** is synonymous for C^∞-manifold. ◇

1.13 Definition. C^r-diffeomorphisms and diffeomorphic manifolds

A **C^r-diffeomorphism** f of two differentiable manifolds M_1 and M_2 is a bijection $f : M_1 \to M_2$ such that $f \in C^r(M_1, M_2)$ and $f^{-1} \in C^r(M_2, M_1)$. Two differentiable manifolds M_1 and M_2 are called **diffeomorphic**, if there exists a diffeomorphism $f : M_1 \to M_2$. ◇

1.14 Example. Every chart (V, φ) of a manifold M is a diffeomorphism of the manifold $V \subset M$ to $\varphi(V) \subset \mathbb{R}^n$ (cf. example 1.8 (a)). ◇

1.15 Remark. In the exercises it is shown that to a given topological manifold one can find different differentiable structures that are nevertheless diffeomorphic. $\diamond$

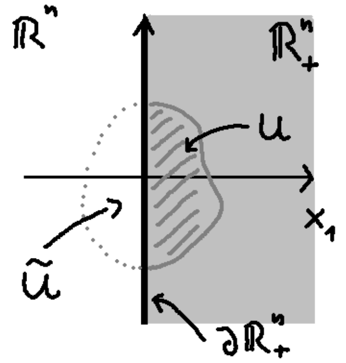
Heuristically it is clear that a sufficiently smooth boundary of an n-dimensional differentiable manifold is itself a $(n-1)$-dimensional differentiable manifold. As an example think of the circle S^1 (a one-dimensional manifold) being the boundary of the unit disk in $\mathbb{R}^2$ (a two-dimensional manifold). The precise definition of a manifold with boundary rests on a very simple idea: instead of using $\mathbb{R}^n$ as the "local model", we use the half-space

$$\mathbb{R}^n_+ := \{x \in \mathbb{R}^n \,|\, x_1 \geq 0\} \quad \text{with } (n-1)\text{-dimensional boundary} \quad \partial\,\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \,|\, x_1 = 0\}\,.$$

That is, charts now map into $\mathbb{R}^n_+$ instead of $\mathbb{R}^n$. Recall that by definition a set $U \subset \mathbb{R}^n_+$ is open with respect to the induced topology (also called relatively open), if there exists an open set $\tilde{U} \subset \mathbb{R}^n$ such that $U = \tilde{U} \cap \mathbb{R}^n_+$.

1.16 Definition. Topological manifold with boundary

A second-countable topological Hausdorff space M is called a **topological manifold with boundary** of dimension n, if any point $x \in M$ has a neighbourhood that is homeomorphic to an (relatively!) open set in $\mathbb{R}^n_+$.
A **chart** for a topological manifold with boundary M is a pair (V, φ), where $V \subset M$ is an open set and $\varphi : V \to \varphi(V) \subset \mathbb{R}^n_+$ a homeomorphism . $\diamond$



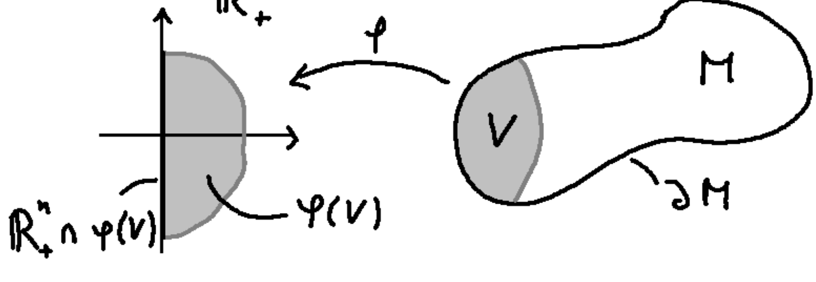
In order to define also differentiable structures on manifolds with boundary, we first have to clarify what it means for a function defined on $\mathbb{R}^n_+$ to be differentiable at a point x on the boundary $\partial\mathbb{R}^n_+$: we say that a map $f : U \to \mathbb{R}^m$ on a relatively open subset of $U \subset \mathbb{R}^n_+$ is r-times continuously differentiable, if there exist an open set $\tilde{U} \subset \mathbb{R}^n$ containing U and a r-times continuously differentiable map $\tilde{f} : \tilde{U} \to \mathbb{R}^m$, such that $\tilde{f}|_U = f$.
From here on C^r-compatibility and atlases can be defined as in definition 1.5 and 1.6 also for charts taking values in $\mathbb{R}^n_+$.

1.17 Definition. Differentiable manifold with boundary

A topological manifold with boundary M together with an equivalence class of C^r-atlases is called a **differentiable manifold with boundary** (or more precisely a C^r-manifold with boundary).
The **boundary** of M is

$$\partial M := \bigcup_i \varphi_i^{-1}\left(\varphi_i(V_i) \cap \partial\mathbb{R}^n_+\right)\,,$$

where $\{(V_i, \varphi_i)\}$ is an atlas. $\diamond$



1.18 Remark. Note that ∂M is well defined, since for any diffeomorphism $f : \mathbb{R}^n_+ \supset V \to U \subset \mathbb{R}^n_+$ of relatively open sets $V, U \subset \mathbb{R}^n_+$ (and thus, in particular, for the transition maps) it holds that $x \in V \cap \partial\mathbb{R}^n_+$ if and only if $f(x) \in \partial\mathbb{R}^n_+$[6]. To see this, let $x \in V \cap (\mathbb{R}^n_+ \setminus \partial\mathbb{R}^n_+)$ be a point in the interior. By Taylor's theorem $f(x+h) = f(x) + Df|_x h + o(\|h\|)$, and thus, since the differential $Df|_x$ at x is an isomorphism, there exists an open neighbourhood O of x such that $f(O)$ is open in $\mathbb{R}^n$. Hence, $f(x) \in U \cap (\mathbb{R}^n_+ \setminus \partial\mathbb{R}^n_+)$. $\diamond$

1.19 Example. The closed interval $M = [a, b]$ with the charts $(V_1 = [a, b)$, $\varphi_1 : x \mapsto x - a)$, and $(V_2 = (a, b]$, $\varphi_2 : x \mapsto b - x)$, is a differentiable manifold with boundary $\partial M = \{a\} \cup \{b\}$. $\diamond$

1.20 Remark. The boundary ∂M of a manifold with boundary as just defined can differ from its topological boundary as a subset of some other topological space. For example, as a manifold the circle S^1 has no boundary, but as a subset of the plane $\mathbb{R}^2$ the topological boundary of S^1 is S^1 itself. $\diamond$

1.21 Consequence. Let M be a differentiable manifold with boundary ∂M. Then $M \setminus \partial M$ and ∂M inherit the structure of manifolds without boundary of dimension $\dim(M \setminus \partial M) = n$ resp. $\dim(\partial M) = n - 1$.

Proof. Let (V_i, φ_i) with $\varphi_i : V_i \to \mathbb{R}^n_+$ be an atlas for M. Then

$$(V_i \cap (M \setminus \partial M),\ \varphi_i|_{V_i \cap (M \setminus \partial M)})$$

is an atlas for $M \setminus \partial M$ where none of the charts hits a boundary point in $\partial\mathbb{R}^n_+$. To obtain an atlas for ∂M, put $U_i = V_i \cap \partial M$ and $\tilde{\varphi}_i : U_i \to \partial\mathbb{R}^n_+ \cong \mathbb{R}^{n-1}$, $\tilde{\varphi}_i = \varphi_i|_{U_i}$. Then $(U_i, \tilde{\varphi}_i)$ is an atlas for ∂M. $\square$

1.22 Convention. Note that differentiable manifolds without boundary as in definition 1.7 can be seen as a special case of differentiable manifolds with boundary as in definition 1.17, where the boundary just happens to be empty. Therefore, with the exception of the beginning of chapter 2, we will not distinguish the two concepts: a "manifold" may have or may not have a boundary. $\diamond$

German nomenclature

atlas = Atlas
chart = Karte
continuous = stetig
diffeomorphism = Diffeomorphismus
equivalent = äquivalent
homeomorphism = Homöomorphismus
open set = offene Menge
topology = Topologie
boundary = Rand
compatible = verträglich
countable = abzählbar
differentiable = differenziebar
euclidian space = Euklidischer Raum
manifold = Mannigfaltigkeit
topological space = topologischer Raum
transition map = Kartenwechsel

[6] The same statement also applies to homoeomorphisms, but requires more advanced techniques to prove it.

2 The tangent bundle

For an n-dimensional manifold M embedded into an ambient $\mathbb{R}^m$, like the sphere S^2 embedded in $\mathbb{R}^3$, it is easy to picture the tangent space at any point $x \in M$ as an n-dimensional euclidean space sitting in $\mathbb{R}^m$ with the origin attached to the point x and being "tangent" to M. While this is a very useful geometric picture to keep in mind, it does not provide a mathematical definition of tangent spaces for general manifolds. There are indeed different but in the end equivalent approaches to such a general definition of tangent vectors and spaces, some more geometric, others more algebraic. We will start with a geometric approach based on the idea that a tangent vector to a manifold can be understood as a velocity vector associated with a trajectory on the manifold. To simplify the discussion, we will initially define tangent vectors and tangent spaces only for manifolds without boundary.

A C^1-curve on a differentiable manifold M is a C^1-map $c : I \to M$ from an open interval $I \subset \mathbb{R}$ into M, i.e. a map $c \in C^1(I, M)$. For any $x \in M$ let

$$C_x := \left\{c \in C^1(I, M) \,|\, I \subset \mathbb{R} \text{ open}, \, 0 \in I, \, c(0) = x\right\} , \tag{2.1}$$

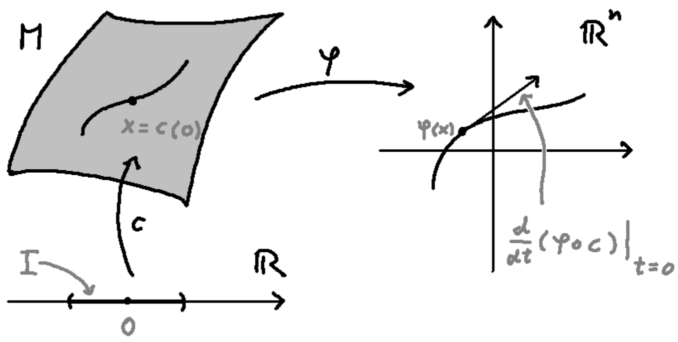
which will be called the set of curves through x for short. Two curves c_1 and c_2 in C_x are said to be equivalent, $c_1 \sim c_2$ for short, if in one (and thus in any) chart (V, φ) with $x \in V$ it holds that[1]

$$\tfrac{\mathrm{d}}{\mathrm{d}t}(\varphi \circ c_1)|_{t=0} \quad = \quad \tfrac{\mathrm{d}}{\mathrm{d}t}(\varphi \circ c_2)|_{t=0} \,. \tag{2.2}$$

Thus two curves are equivalent if they pass through x with the same "velocity". The fact that (2.2) holds either for all or for no chart follows from the chain rule: Let φ_1 and φ_2 be charts, then

$$\tfrac{\mathrm{d}}{\mathrm{d}t}(\varphi_1 \circ c)|_{t=0} = \tfrac{\mathrm{d}}{\mathrm{d}t}(\varphi_1 \circ \varphi_2^{-1} \circ \varphi_2 \circ c)|_{t=0} = D(\varphi_1 \circ \varphi_2^{-1})|_{\varphi_2(x)} \, \tfrac{\mathrm{d}}{\mathrm{d}t}(\varphi_2 \circ c)|_{t=0} \,. \tag{2.3}$$

Thus, when changing the coordinate system, the velocity vector $\frac{\mathrm{d}}{\mathrm{d}t}(\varphi_2 \circ c)|_{t=0}$ of a curve c at x with respect to a chart φ_2 is mapped to the velocity vector $\frac{\mathrm{d}}{\mathrm{d}t}(\varphi_1 \circ c)|_{t=0}$ with respect to the chart φ_1 by acting on it with the Jacobian $D(\varphi_1 \circ \varphi_2^{-1})$ of the transition function evaluated at the point $\varphi_2(x)$. Being the Jacobian of a diffeomorphism, $D(\varphi_1 \circ \varphi_2^{-1})$ is an isomorphism. Moreover, it is independent of the curve c and thus the equality (2.2) holds either in all charts or in none.



[1] It will be convenient in the following to write the evaluation of a derivative Df at a certain point p on a manifold M not as $Df(p)$, but as $Df|_p$. To be consistent we use this notation also for derivatives of functions on $\mathbb{R}$, e.g. $\frac{\mathrm{d}}{\mathrm{d}t}(\varphi \circ c_1)|_{t=0} = (\varphi \circ c_1)'(0)$ in (2.2).

2.1 Definition. **Tangent vectors and tangent spaces**

A **tangent vector** v to a differentiable manifold M at a point $x \in M$ is an equivalence class $[c]_x$ of curves $c \in C_x$ under the equivalence relation (2.2). The set of all tangent vectors to M at x is called the **tangent space** to M at x and denoted by T_xM. ◇

2.2 Proposition. **The tangent space is a vector space**

For any chart (V, φ) with $x \in V$ the map

$$T_x\varphi : T_xM \to \mathbb{R}^n \,, \qquad [c]_x \mapsto \tfrac{\mathrm{d}}{\mathrm{d}t}(\varphi \circ c)|_{t=0}$$

is well defined and a bijection. Moreover, the vector space structure induced by $(T_x\varphi)^{-1} : \mathbb{R}^n \to T_xM$ on T_xM is independent of φ and thus turns T_xM in a natural way into **a real vector space** of dimension $\dim T_xM = n = \dim M$.

Proof. The map $T_x\varphi$ is well defined (meaning that its value does not depend on the chosen representative $c \in [c]_x$) and injective by definition of T_xM. To see that it is also surjective, note that for any given vector $w \in \mathbb{R}^n$ the curve $c(t) := \varphi^{-1}(\varphi(x) + tw)$ (defined on a sufficiently small interval around zero) satisfies

$$\tfrac{\mathrm{d}}{\mathrm{d}t}(\varphi \circ c)|_{t=0} = w \,.$$

For the last statement let φ_1 and φ_2 be charts at x. Then with (2.3) we have

$$T_x\varphi_1([c]_x) = \tfrac{\mathrm{d}}{\mathrm{d}t}(\varphi_1 \circ c)|_{t=0} = D(\varphi_1 \circ \varphi_2^{-1})|_{\varphi_2(x)} \tfrac{\mathrm{d}}{\mathrm{d}t}(\varphi_2 \circ c)|_{t=0} = D(\varphi_1 \circ \varphi_2^{-1})|_{\varphi_2(x)}\, T_x\varphi_2([c]_x)$$

and thus also

$$T_x\varphi_1^{-1} = T_x\varphi_2^{-1}\,(D(\varphi_1 \circ \varphi_2^{-1})|_{\varphi_2(x)})^{-1} \,,$$

where $D(\varphi_1 \circ \varphi_2^{-1})|_{\varphi_2(x)}$ is a vector space isomorphism of $\mathbb{R}^n$. Hence, for $v, u \in T_xM$ we have

$$\begin{aligned}
v +_{\varphi_1} u &:= T_x\varphi_1^{-1}\,(T_x\varphi_1(v) + T_x\varphi_1(u)) \\
&= T_x\varphi_1^{-1}\left(D(\varphi_1 \circ \varphi_2^{-1})|_{\varphi_2(x)}\, T_x\varphi_2(v) + D(\varphi_1 \circ \varphi_2^{-1})|_{\varphi_2(x)}\, T_x\varphi_2(u)\right) \\
&= T_x\varphi_1^{-1} D(\varphi_1 \circ \varphi_2^{-1})|_{\varphi_2(x)}\,(T_x\varphi_2(v) + T_x\varphi_2(u)) \\
&= T_x\varphi_2^{-1}\,(T_x\varphi_2(v) + T_x\varphi_2(u)) \\
&= v +_{\varphi_2} u \,.
\end{aligned}$$

□

2.3 Remark. According to the previous proposition, every chart (V, φ) for M with $x \in V$ yields a vector space isomorphism $T_x\varphi : T_xM \to \mathbb{R}^n$. In the following we will treat $T_x\varphi$ also notationally as a linear map between vector spaces and write $T_x\varphi\, v$ instead of $T_x\varphi(v)$.

The components v^i of $v \in T_xM$ with respect to φ are by definition

$$v^i := \langle e_i, T_x\varphi\, v\rangle_{\mathbb{R}^n} \qquad \text{and thus} \qquad T_x\varphi\, v \;=\; \sum_{i=1}^n v^i e_i \,,$$

where $(e_1, \ldots, e_n)$ denotes the canonical basis of $\mathbb{R}^n$. Clearly the components v^i of a tangent vector v depend on the chosen chart φ. Since there are no preferred charts on a general manifold M (in physics one would often say that there is no preferred coordinate system), there is also no preferred basis of T_xM and also no natural scalar product. Moreover, there is no canonical way to identify tangent spaces T_xM and T_yM at different points $x \neq y$.

However, in the special case of an open set $M \subset \mathbb{R}^n$ we can identify T_xM for any $x \in M$ in a natural way with $\mathbb{R}^n$ using the canonical chart $T_x\mathrm{id}$. ◇

2.4 Remark. **Derivations at points**

An alternative, more algebraic way to define tangent vectors is via derivations. Given a tangent vector $v \in T_xM$ at $x \in M$, i.e. an equivalence class of curves through x, one can define the directional derivative of a function $f \in C^1(M, \mathbb{R}) =: C^1(M)$ at the point x in the direction v by

$$D_v(f) := \tfrac{\mathrm{d}}{\mathrm{d}t}(f \circ c_v)|_{t=0} ,$$

where $c_v \in v$ is any curve in the equivalence class v. This is well defined, since using a chart (V, φ) we find that (again by the chain rule for functions on euclidean spaces) that

$$\begin{aligned} D_v(f) &= \tfrac{\mathrm{d}}{\mathrm{d}t}(f \circ c_v)|_{t=0} = \tfrac{\mathrm{d}}{\mathrm{d}t}\big((f \circ \varphi^{-1}) \circ (\varphi \circ c_v)\big)|_{t=0} \\ &= D(f \circ \varphi^{-1})|_{\varphi(x)} \tfrac{\mathrm{d}}{\mathrm{d}t}(\varphi \circ c_v)|_{t=0} \\ &= \big\langle \nabla(f \circ \varphi^{-1})|_{\varphi(x)}, T_x\varphi\, v \big\rangle_{\mathbb{R}^n} \end{aligned} \tag{2.4}$$

is given by the directional derivative of the function $f \circ \varphi^{-1} : \mathbb{R}^n \supset \varphi(V) \to \mathbb{R}$ in the direction $T_x\varphi\, v \in \mathbb{R}^n$, where the latter depends only on v and not on the representative c_v.
Moreover, this computation also shows that the map $C^1(M) \to \mathbb{R}$, $f \mapsto D_v(f)$, is linear and satisfies the product rule, i.e. that for $f, g \in C^1(M)$ we have

$$D_v(fg) = D_v(f)\, g(x) + f(x)\, D_v(g) . \tag{2.5}$$

On a smooth manifold, a linear map $D : C^\infty(M) \to \mathbb{R}$ is called a **derivation at** $x \in M$, if it satisfies (2.5). The set of derivations at a point x is naturally a real vector space with $(D + \tilde{D})(f) := D(f) + \tilde{D}(f)$ and $(aD)(f) := aD(f)$ for $a \in \mathbb{R}$. We have just shown that every tangent vector $v \in T_xM$ defines a derivation D_v at x and that the map $v \mapsto D_v$ is linear and injective. One can show that this map is also surjective and thus defines a vector space isomorphism between T_xM and the space of derivations at x. Hence one can naturally identify tangent vectors at x with derivations at x. Actually, the latter concept is often taken as the definition of a tangent vector. ◇

With the previous two remarks and proposition 2.2 in mind, the following notation for coordinate bases of tangent spaces appears natural.

2.5 Definition. **Coordinate bases**

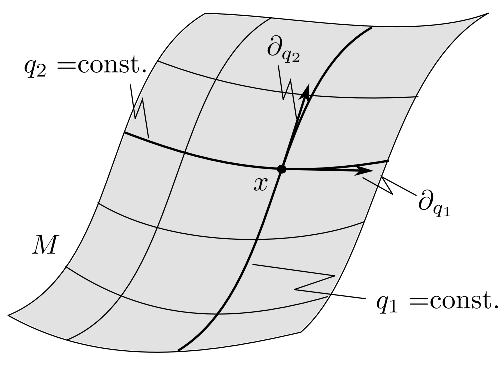
Let (V, φ) be a chart on M and let $x \in V$. Then $(\partial_{\varphi_1}, \ldots, \partial_{\varphi_n})$ with

$$\partial_{\varphi_j} := (T_x\varphi)^{-1} e_j \in T_xM \quad \text{for } j = 1 \ldots, n$$

is called the **coordinate basis** of T_xM with respect to the coordinate chart φ. Note that in the present context ∂_{φ_j} does not denote a differential operator, but a tangent vector. One can "act" by a tangent vector on a function by taking the directional derivative in the direction of the vector, however. ◇

2.6 Notation. When emphasising the "coordinate aspect" of a chart we will often use the letter q instead of φ for a coordinate chart $q : V \to \mathbb{R}^n$. Correspondingly the coordinate basis of T_xM is then $(\partial_{q_1}, \ldots, \partial_{q_n})$.
Note that the coordinate tangent vector $\partial_{q_j} \in T_xM$ is the tangent vector defined by the curve

$$c : (-\varepsilon, \varepsilon) \to M , \quad t \mapsto q^{-1}\big((q_1(x), \ldots, q_j(x) + t, \ldots, q_n(x))\big) .$$



2.7 Remark. **The tangent space at a boundary point**
Up to now we defined tangent spaces only at interior points of a manifold. Our definition of tangent vectors in terms of equivalence classes of curves through a point needs to be slightly modified in order to cope with points on the boundary. This could be done by considering also curves that start or end at the boundary point, i.e. by replacing C_x in (2.1) by

$$\tilde{C}_x := \left\{ c \in C^1(I, M) \,|\, I = [0, \varepsilon) \text{ or } I = (-\varepsilon, 0] \text{ for some } \varepsilon > 0,\, c(0) = x \right\} .$$

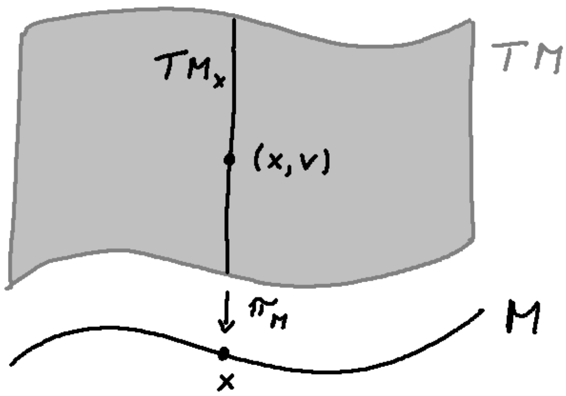
Alternatively we could take the algebraic viewpoint as a definition and identify the tangent space at a boundary point $x \in \partial M$ with the space of derivations at that point. We skip the details and just note that all the statements of this section naturally carry over to tangent spaces at boundary points. In particular, the tangent space at a boundary point of an n-dimensional manifold with boundary is also an n-dimensional real vector space that can be identified (non-uniquely) with $\mathbb{R}^n$ using a chart containing that point. ◇

2.8 Definition. **The tangent bundle**

The **tangent bundle** TM of M is the disjoint union of the tangent spaces

$$TM := \bigcup_{x \in M} (\{x\} \times T_x M) = \{(x, v) \,|\, x \in M\,,\ v \in T_x M\}$$

Points in TM are thus pairs (x, v) with $x \in M$ (the base point) and $v \in T_x M$ (the tangent vector). The map $\pi_M : TM \to M$ that maps a point (x, v) in the total space TM to the point x in the base space M is called the projection onto the base space. The second component of the pre-image $\pi_M^{-1}(\{x\}) = \{x\} \times T_x M$ is called the **fibre** over $x \in M$.



2.9 Example. For an open set $M \subset \mathbb{R}^n$ we can identify TM in a natural way with $M \times \mathbb{R}^n$. Since $M \times \mathbb{R}^n$ is an open subset of $\mathbb{R}^{2n}$ and thus a manifold, we can in this case equip the tangent bundle TM in a natural way with the structure of a manifold. We will soon see that also for general differentiable manifolds of dimension n the tangent bundle TM is again a manifold of dimension $2n$. ◇

We now define the derivative (called differential) of a smooth map between differentiable manifolds. Recall that for a smooth map $f : \mathbb{R}^n \to \mathbb{R}^m$ the differential of f at a point $x \in \mathbb{R}^n$ is the **linear approximation** $Df|_x : \mathbb{R}^n \to \mathbb{R}^m$ to f at x. Linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$ are given by matrices and the differential $(Df|_x)_{ij} = \frac{\partial f_i}{\partial x_j}(x)$ is the Jacobian matrix. For a map $f : M_1 \to M_2$ between manifolds the linear approximation to f at $x \in M_1$ is a map between the tangent spaces $T_x M_1$ and $T_{f(x)} M_2$. With our definition of tangent vectors as equivalence classes of curves the following definition is very natural.

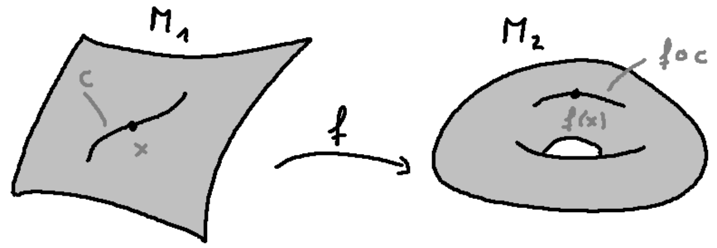
2.10 Definition. **The differential of a smooth map and the pushforward**
Let $f \in C^1(M_1, M_2)$. Its **differential**

$$Df : TM_1 \to TM_2$$

maps the point $(x, [c]_x) \in TM_1$ to the point $(f(x), [f \circ c]_{f(x)}) \in TM_2$, i.e.

$$Df\big((x, [c]_x)\big) \;:=\; \big(f(x), [f \circ c]_{f(x)}\big) .$$

The differential Df of a smooth map is often also called the **pushforward** of tangent vectors under f, as it “pushes” tangent vector on M_1 to tangent vectors on M_2. Other notations for Df are $\mathrm{d}f$, Tf or f_*.



2.11 Proposition. **Linearity of the differential**

The differential Df of a function $f \in C^1(M_1, M_2)$ is well defined and its restriction

$$Df|_x : T_xM_1 \to T_{f(x)}M_2\,, \quad v \mapsto Df|_x v := (Df(x,v))_2\,,$$

to a single tangent space is a linear map. Here $(Df(x,v))_2$ denotes the second component of $Df(x,v) \in \{f(x)\} \times T_{f(x)}M_2$.

Proof. It is clear that f maps C^1-curves c in M_1 to C^1-curves $f \circ c$ in M_2. By using charts it is easy to see that equivalent curves through $x \in M_1$ are mapped to equivalent curves through $f(x) \in M_2$ and that $Df|_x : T_xM_1 \to T_{f(x)}M_2$ is a linear map:

$$\begin{aligned} (Df(x,v))_2 &= T_{f(x)}\varphi_2^{-1} \tfrac{\mathrm{d}}{\mathrm{d}t}\left(\varphi_2 \circ f \circ c_v\right)|_{t=0} = T_{f(x)}\varphi_2^{-1} \tfrac{\mathrm{d}}{\mathrm{d}t}\left(\varphi_2 \circ f \circ \varphi_1^{-1} \circ \varphi_1 \circ c_v\right)|_{t=0} \\ &= T_{f(x)}\varphi_2^{-1} D(\varphi_2 \circ f \circ \varphi_1^{-1})|_{\varphi_1(x)} \tfrac{\mathrm{d}}{\mathrm{d}t}\left(\varphi_1 \circ c_v\right)|_{t=0} \qquad (2.6) \\ &= T_{f(x)}\varphi_2^{-1} D(\varphi_2 \circ f \circ \varphi_1^{-1})|_{\varphi_1(x)}\, T_x\varphi_1\, v\,. \end{aligned}$$ □

Note that for $f : \mathbb{R}^n \supset M \to \mathbb{R}^m$ using the canonical chart id it is common to identify the differential $Df|_x = (T_{f(x)}\mathrm{id}^{-1}) D(\mathrm{id} \circ f \circ \mathrm{id}^{-1})|_x\, (T_x\mathrm{id})$ with the Jacobian $D(\mathrm{id} \circ f \circ \mathrm{id}^{-1})|_x$.

2.12 Definition. **The tangent bundle as a manifold**

We can equip the tangent bundle TM with the structure of a differentiable manifold of dimension $2n$ by covering it with the **natural atlas**: Let $\mathcal{A} = \{(V_i, \varphi_i)\}$ be an atlas for M, then

$$T\mathcal{A} := \{(TV_i, D\varphi_i)\}$$

is an atlas of TM, where the charts

$$D\varphi_i : TM \supset TV_i \to D\varphi_i(TV_i) = \varphi_i(V_i) \times \mathbb{R}^n \subset T\mathbb{R}^n = \mathbb{R}^{2n} \qquad (2.7)$$

map into $\mathbb{R}^{2n}$. Note that in this case also the topology on TM is defined by the charts $D\varphi_i$.[2] ⋄

2.13 Consequence. **The differential of a smooth map is smooth**

Since fibre wise $D\varphi|_x = T_x\varphi$, (2.6) implies that the differential $Df : TM_1 \to TM_2$ of a smooth map is itself a smooth map between manifolds, cf. definition 1.11 and the following diagram

$$\begin{array}{ccc} TM_1 \supset TV_1 & \xrightarrow{Df} & TV_2 \subset TM_2 \\ D\varphi_1 = (\varphi_1, T\varphi_1) \downarrow & & \downarrow D\varphi_2 = (\varphi_2, T\varphi_2) \\ \varphi_1(V_1) \times \mathbb{R}^n & \xrightarrow{D(\varphi_2 \circ f \circ \varphi_1^{-1})} & \varphi_2(V_2) \times \mathbb{R}^m\,. \end{array}$$

[2] To see that the topology induced by the charts $D\varphi_i$ is Hausdorff and second countable, note that for two different points (x_1, v_1) and (x_2, v_2) in TM we have either $x_1 \neq x_2$, or $x_1 = x_2$ and $v_1 \neq v_2$. In the first case there are disjoint open sets $U_1, U_2 \subset V_1$ containing x_1 respectively x_2. Then $D\varphi_i^{-1}(U_1 \times \mathbb{R}^n)$ and $D\varphi_i^{-1}(U_2 \times \mathbb{R}^n)$ are disjoint open sets containing (x_1, v_1) respectively (x_2, v_2). In the second case with $x_1 = x_2 =: x$ there are disjoint open sets $U_1, U_2 \subset \mathbb{R}^n$ containing $D\varphi_i|_x v_1$ respectively $D\varphi_i|_x v_2$. Again the preimages $D\varphi_i^{-1}(V \times U_1)$ and $D\varphi_i^{-1}(V \times U_2)$ are disjoint open sets containing (x_1, v_1) respectively (x_2, v_2). To constrcut a countable base for the induced toplogy on TM let $\{U_j\}_j$ be a countable base for the toplogy of M (which is second countable by assumption) and $\{W_k\}_k$ a countable base for the topology of $\mathbb{R}^n$. Then $\{D\varphi_i^{-1}(U_j \times W_k)\}_{i,j,k}$ is a countable base for TM when we chose a countable atlas $\mathcal{A}$.

Remark. In classical mechanics the configuration space is usually a manifold M. Then the tangent bundle TM corresponds to the space of configurations and velocities, i.e. a point $y \in TM$, $y = (x, v)$ is a pair of a configuration $x = \pi_M y$ and a velocity $v \in T_x M$. We will see that the Lagrangian is a function on TM (but not the Hamiltonian!). $\diamond$

2.14 Remark. Locally TM is diffeomorphic to $M \times \mathbb{R}^n$, since every **bundle chart** yields such a diffeomorphism, cf. (2.7). However, globally TM and $M \times \mathbb{R}^n$ need not be diffeomorphic. $\diamond$

2.15 Definition. **Parallelisable manifolds**

A differentiable manifold M is called **parallelisable** (and TM **trivialisable**) if there exists a diffeomorphism $\phi : TM \to M \times \mathbb{R}^n$ such that $\phi|_{T_x M} : T_x M \to \{x\} \times \mathbb{R}^n$ is a vector space isomorphism for all $x \in M$.

The terminology is motivated by the fact that such a trivialisation ϕ allows for an identification of different tangent spaces,

$$T_x M \overset{\phi}{\leftrightarrow} \{x\} \times \mathbb{R}^n \overset{\cong}{\leftrightarrow} \{y\} \times \mathbb{R}^n \overset{\phi}{\leftrightarrow} T_y M\,.$$

Note that this identification is not canonical but depends on the choice of ϕ. $\diamond$

2.16 Proposition. **The chain rule**

For differentiable maps $f : M_1 \to M_2$ and $g : M_2 \to M_3$ between differentiable manifolds the chain rule

$$D(g \circ f) = Dg \circ Df$$

holds.

Proof. $D(g \circ f)([c]_x) = [g \circ f \circ c]_{g\circ f(x)} = Dg([f \circ c]_{f(x)}) = Dg(Df([c]_x))$. $\square$

2.17 Definition. **Immersions, submersions, embeddings**

Let M_1 and M_2 be differentiable manifolds.

(a) $f : M_1 \to M_2$ is called an **immersion** or **immersive**, if $f \in C^1$ and $Df|_x$ is injective for all $x \in M_1$.

(b) $f : M_1 \to M_2$ is called a **submersion** or **submersive**, if $f \in C^1$ and $Df|_x$ is surjective for all $x \in M_1$.

(c) $f : M_1 \to M_2$ is called an **embedding**, if f is an injective immersion that is also a homeomorphism onto its range.

2.18 Examples. Let $M_1 = \mathbb{R}^n$ and $M_2 = \mathbb{R}^m$.

(a) Let $n < m$ and $f : \mathbb{R}^n \to \mathbb{R}^m$, $x = (x_1, \dots, x_n) \mapsto (x_1, \dots, x_n, 0, \dots, 0)$. Then the $m \times n$-matrix

$$Df \equiv \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix}$$

has full rank equal to n and is therefore injective. Hence, f is an immersion. Moreover, the map f is injective and continuously invertible on its range and therefore an embedding.

(b) Let $n > m$ and $f : \mathbb{R}^n \to \mathbb{R}^m$, $x = (x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_m)$. Then the $m \times n$-matrix

$$Df \equiv \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \vdots & & \vdots \\ \vdots & & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 1 & 0 & \cdots & 0 \end{pmatrix}$$

has full rank equal to m and is therefore surjective. Hence, f is a submersion.

(c) Let $n = 1$, $m > 1$, and $f : \mathbb{R} \to \mathbb{R}^m$ a smooth curve. The map f is immersive iff its velocity vector satisfies $f'(t) \neq 0$ for all $t \in \mathbb{R}$. If the curve intersects itself, i.e. $f(t) = f(s)$ for $t \neq s$, then f is not an embedding.

2.19 Definition. <u>**Submanifolds**</u>

Let M be a n-dimensional differentiable manifold. Then $N \subset M$ is called a **k-dimensional submanifold** of M if for every $x \in N$ there exists a chart (V, φ) for M with

$$\varphi(y) = (q_1, \ldots, q_k, \underbrace{0, \ldots, 0}_{n-k}) \qquad \text{for all } y \in N \cap V.$$

2.20 Remark. Note that a k-dimensional submanifold is, in particular, a k-dimensional manifold: the charts in definition 2.21 become charts for N by dropping the last $n - k$ components.

2.21 Proposition. The following assertions are equivalent:

(a) $N \subset M$ is a k-dimensional submanifold.

(b) N is locally the image of an embedding of a piece of $\mathbb{R}^k$. More precisely, for every point $x \in N$ there exists an (relatively) open neighbourhood $V \subset N$ of x, an open set $U \subset \mathbb{R}^k$, and an embedding

$$f : U \to M \quad \text{with} \quad f(U) = V.$$

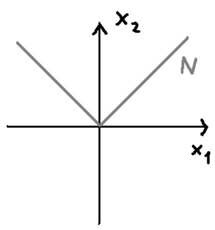
(c) N is locally a level set of a submersion into $\mathbb{R}^{n-k}$. More precisely, for every $x \in N$ there exist an open neighbourhood $V \subset M$ of x and a submersion $F : V \to \mathbb{R}^{n-k}$ such that

$$N \cap V = \{y \in V \mid F(y) = 0\}.$$

Proof. The proof is a somewhat tedious application of the inverse function theorem resp. of the implicit function theorem. □

2.22 Example. The sphere $S^2 = \{x \in \mathbb{R}^3 \mid \|x\| = 1\}$ is a two-dimensional submanifold of $\mathbb{R}^3$. This can be seen most easily by using condition (c) above. Let $F : \mathbb{R}^3 \to \mathbb{R}$, $x \mapsto \|x\|^2 - 1$. Then F is smooth, $\{F(x) = 0\} = S^2$, and $DF|_x = \langle 2x, \cdot \rangle_{\mathbb{R}^3} \neq 0$ for $x \in S^2$.

2.23 Example. Let $M = \mathbb{R}^2$, then $N = \{x \in M \mid x_2 = |x_1|\}$ is **not** a submanifold, but it can be equipped with the structure of a manifold. E.g. the atlas $(V = N, \varphi : (x_1, x_2) \mapsto x_1)$ turns the set N into a manifold that is diffeomorphic to $\mathbb{R}$.



2.24 Remark. <u>**Whitney embedding theorem**</u>

A famous theorem of Hassler Whitney states that any smooth n-dimensional manifold (Hausdorff and second-countable) can be smoothly embedded into $\mathbb{R}^{2n}$. Thus any abstract manifold is actually diffeomorphic to a of submanifold of some $\mathbb{R}^m$.

We now introduce a new type of functions on manifolds namely vector fields. A vector field is a map that selects at each point of a manifold a tangent vector at that point in a smooth way.

2.25 Definition. **Vector fields**

A C^p-map $X : M \to TM$ with $\pi_M \circ X = \mathrm{id}_M$ is called a **C^p-vector field**. We denote the set of C^∞-vector fields by $\mathcal{T}_0^1(M)$.

It will often be convenient to identify for a vector field $X \in \mathcal{T}_0^1(M)$ its value $X(x) \in \{x\} \times T_xM$ at $x \in M$ with its part in T_xM without making this explicit in the notation by projecting on the second factor. ◇

The idea behind the notation $\mathcal{T}_0^1(M)$ will become clear later on: tangent vector fields are just tensor fields of type $(1, 0)$.

2.26 Remark. In the exercises you will show that M is parallelisable if and only of there exist a **global frame** for the tangent bundle, that is vector fields $X_1, \ldots, X_n \in \mathcal{T}_0^1(M)$ such that $(X_1(x), \ldots, X_n(x))$ is a basis of T_xM for each $x \in M$. ◇

While the differential of any smooth map between manifolds defines a map between the tangent bundles, only diffeomorphisms allow for mapping also vector fields to vector fields.

2.27 Definition. **The pushforward of vector fields**

A diffeomorphism $\Phi : M_1 \to M_2$ allows to map ("push forward") vector fields on M_1 to vector fields on M_2. The map

$$\Phi_* : \mathcal{T}_0^1(M_1) \to \mathcal{T}_0^1(M_2)\,, \qquad X \mapsto \Phi_* X = D\Phi \circ X \circ \Phi^{-1}$$

is called the **pushforward** and can be most easily understood through the following commutative diagram:

$$\begin{array}{ccc} M_1 & \overset{\Phi^{-1}}{\leftarrow} & M_2 \\ X \downarrow & & \downarrow \Phi_* X \\ TM_1 & \overset{D\Phi}{\rightarrow} & TM_2\,. \end{array}$$

2.28 Remark. **Coordinate representation of a vector field**

Using a coordinate chart (V, φ) for M, the restriction of a vector field $X \in \mathcal{T}_0^1(M)$ to V can be mapped to a vector field on $\varphi(V) \subset \mathbb{R}^n$ using the pushforward φ_*,

$$\varphi_* X : \mathbb{R}^n \supset \varphi(V) \to T\varphi(V) = \varphi(V) \times \mathbb{R}^n\,, \quad q \mapsto (q, v(q)) \quad \text{with} \quad v(q) = \sum_{j=1}^n v^j(q)\, e_j\,,$$

where $v^j(q)$ are the components of the vector $X(x) \in T_xM$ at $x = \varphi^{-1}(q)$ with respect to the coordinate basis $(\partial_{q_1}, \ldots, \partial_{q_n})$, cf. definition 2.5. Note the twofold role of the chart φ here: it provides the coordinates $q = \varphi(x)$ on the patch V and also the coordinate bases of the tangent spaces.

When using charts one often writes only the vector part $v(q)$ of the vector field X. ◇

In remark 2.4 we saw that a tangent vector v at a point $x \in M$ defines a derivation (i.e. a linear map $D_v : C^\infty(M) \to \mathbb{R}$ satisfying the product rule) at that point by taking the directional derivative of a function at that point. A vector field X now provides a tangent vector, and hence a derivation, at every point of a manifold and thus a map $C^\infty(M) \to C^\infty(M)$.

2.29 Definition. <u>**The Lie derivative of a function**</u>

For a vector field $X \in \mathcal{T}^1_0(M)$ the map $L_X : C^\infty(M) \to C^\infty(M)$ with

$$f \mapsto L_X(f) \quad := \quad I \circ Df \circ X$$

is called the **Lie derivative** of f with respect to X. Here $I : T\mathbb{R} = \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ denotes the projection onto the second factor. One also writes just $X(f)$ instead of $L_X(f)$. $\diamond$

2.30 Remark. Within a chart (V, φ) we write $\varphi_* X(q) = (q, v(q))$ and according to (2.4) we have

$$L_X(f)(\varphi^{-1}(q)) = \sum_{i=1}^n v^i(q) \frac{\partial f}{\partial q_i}(q)\,,$$

where $\frac{\partial f}{\partial q_i}(q)$ is a short hand for $\partial_i (f \circ \varphi^{-1})|_q$, i.e. for the partial derivative of the pull-back of f to coordinate space.

2.31 Proposition. <u>**Properties of the Lie derivative**</u>

The Lie derivative has the following properties:

(a) $L_X(f + g) = L_X(f) + L_X(g)$

(b) $L_X(f \cdot g) = f L_X(g) + g L_X(f)$

(c) $L_{\alpha X + \beta Y}(f) = \alpha L_X(f) + \beta L_Y(f)$

for all $f, g, \alpha, \beta \in C^\infty(M)\,,\ X, Y \in \mathcal{T}^1_0(M)$.

Proof. Homework assignments. $\square$

2.32 Remark. <u>**Derivations**</u>

A map $L : C^\infty(M) \to C^\infty(M)$ satisfying

(i) $L(\alpha f + g) = \alpha L f + L g$

(ii) $L(f \cdot g) = f L g + g L f$

for all $f, g \in C^\infty(M)$ and $\alpha \in \mathbb{R}$ is called a **derivation**. Every derivation L is associated with a unique vector field $X \in \mathcal{T}^1_0(M)$ such that $L = L_X$, cf. remark 2.4.

2.33 Corollary. <u>**The commutator of vector fields**</u>

Let $X, Y \in \mathcal{T}^1_0(M)$. There exists a unique vector field $Z \in \mathcal{T}^1_0(M)$ such that for all $f \in C^\infty(M)$

$$[L_X, L_Y] f := L_X L_Y f - L_Y L_X f = L_Z f\,.$$

The vector field Z is also written as $Z = [X, Y]$ and called the **commutator of X and Y**.

Proof. Homework assignments. $\square$

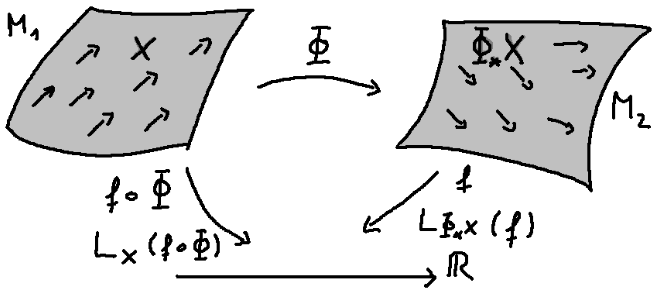
2.34 Remark. <u>**Naturality of the Lie derivative**</u>

Let $\Phi : M_1 \to M_2$ be a diffeomorphism, $f \in C^\infty(M_2)$, and $X \in \mathcal{T}^1_0(M_1)$. Since the diagram

$$\begin{array}{ccccc} M_1 & \xrightarrow{\Phi} & M_2 & & \\ X \downarrow & & \downarrow \Phi_* X & & \\ TM_1 & \xrightarrow{D\Phi} & TM_2 & \xrightarrow{Df} & T\mathbb{R} \end{array}$$

is commutative according to the definition of $\Phi_* X$, it follows that

$$\begin{aligned} L_X(f \circ \Phi) &= I \circ D(f \circ \Phi) \circ X = I \circ Df \circ D\Phi \circ X = I \circ Df \circ \Phi_* X \circ \Phi \\ &= L_{\Phi_* X}(f) \circ \Phi\,. \end{aligned}$$



The fact that an operation behaves "natural" under diffeomorphisms is often called naturality. Here the Lie derivative transforms in the natural way: one can either pull back the function f to M_1 or push forward the vector field X to M_2.

German nomenclature

bundle chart = Bündelkarte
commutator = Kommutator
derivation = Derivation
embedding = Einbettung
immersion = Immersion
parallelisable = parallelisierbar
representation = Darstellung
submanifold = Untermannigfaltigkeit
tangent bundle = Tangentialbündle
tangent vector = Tangentialvektor
vector field = Vektorfeld
chain rule = Kettenregel
coordinates = Koordinaten
differential = Differential, Ableitung
fibre = Faser
Lie derivative = Lieableitung
pushforward = Pushforward
smooth curve = glatte Kurve
submersion = Submersion
tangent space = Tangentialraum
trivialisable = trivialisierbar

3 The cotangent bundle

3.1 Reminder. The dual of a vector space

For a real vector space V of dimension $n \in \mathbb{N}$ its **dual space** $V^* := \mathcal{L}(V, \mathbb{R})$ is defined as the space of linear maps from V to $\mathbb{R}$. Hence, also V^* is an n-dimensional real vector space. The elements of V^* are often called linear functionals and for $v^* \in V^*$ and $u \in V$ one writes

$$v^*(u) =: (v^*, u) =: (v^* \,|\, u)\,,$$

even though the **dual pairing** $(v^*|\, u)$ is **not** a scalar product.

3.2 Definition. The cotangent space

Let M be a differentiable manifold and $x \in M$. The dual space $T_x^*M := (T_xM)^*$ of the tangent space T_xM is called the **cotangent space** of M at x. The elements of T_x^*M are called **cotangent vectors**, **covectors**, or **1-forms**.

3.3 Remark. A scalar product $\langle\cdot,\cdot\rangle$ on a vector space V provides a natural identification of V and V^*, namely $V \ni v \mapsto \langle v, \cdot\rangle \in V^*$. Without scalar product it is still true that $\dim V^* = \dim V$, but there is no canonical isomorphism.

3.4 Example. The differential of a function

For $f : \mathbb{R}^n \to \mathbb{R}$ one usually considers the gradient $\nabla f(x)$ at a point $x \in \mathbb{R}^n$ to be a vector. Without further structure, however, the differential of a function $f : M \to \mathbb{R}$ on a manifold is a covector:
For $f \in C^\infty(M, \mathbb{R})$ the differential (cf. definition 2.10)

$$Df|_x : T_xM \to T_{f(x)}\mathbb{R} = \mathbb{R}$$

is linear and thus $Df|_x \in T_x^*M$. The differential of a real-valued function is usually written as $\mathrm{d}f|_x$ and also called the exterior derivative of f.

As we saw in the previous chapter, the action of $\mathrm{d}f|_x$ on a tangent vector $v \in T_xM$ is just the directional derivative of f at the point $x \in M$ in the direction v. For $v = [c]$ we have

$$\mathrm{d}f|_x(v) = \tfrac{\mathrm{d}}{\mathrm{d}t} f(c(t))|_{t=0}\,.$$

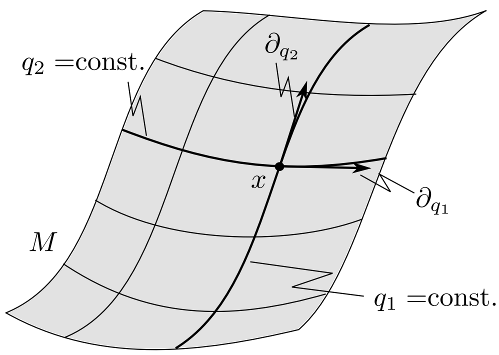
We can thus think of the dual pairing $(\mathrm{d}f|\, v)$ either as a linear action of the covector $\mathrm{d}f$ on the vector v or vice-versa as the linear action of the vector v as a derivation operating on the function f.

3.5 Remark. Coordinate 1-forms

In a chart (V, φ), $\varphi : M \supset V \to \mathbb{R}^n$, $x \mapsto \varphi(x)$ we can think of the components q_i for $i = 1, \ldots, n$, as functions $q_i : V \to \mathbb{R}$, $x \mapsto q_i(x) := \varphi(x)_i$. The corresponding coordinate 1-forms

$\mathrm{d}q_i|_x \in T_x^*M$, $i = 1, \ldots, n$, form a basis of the corresponding cotangent space T_x^*M, since with (2.4)

$$\big(\mathrm{d}q_i|\,\partial_{q_j}\big) = \langle\underbrace{\nabla(q_i \circ \varphi^{-1})}_{\equiv e_i}|_{\varphi(x)}, e_j\rangle = e_i \cdot e_j = \delta_{ij}\,.$$



Thus the coordinate basis $(\mathrm{d}q_i)_{i=1,\ldots,n}$ of T_x^*M is the dual basis to the coordinate basis $(\partial_{q_i})_{i=1,\ldots,n}$ of T_xM. Somewhat formally the above computation reads

$$\big(\mathrm{d}q_i|\,\partial_{q_j}\big) = \frac{\partial}{\partial q_j}\,q_i = \delta_{ij}\,.$$

Moreover, given a covector $\omega = \sum_{j=1}^n \omega_j \mathrm{d}q_j$ and a vector $v = \sum_{i=1}^n v^i \partial_{q_i}$ expressed with respect to the corresponding coordinate bases from a chart φ, then, by linearity in both arguments, the dual pairing takes the form

$$(\omega|\,v) = \big(\sum_{j=1}^n \omega_j \mathrm{d}q_j\big|\sum_{i=1}^n v^i\partial_{q_i}\big) = \sum_{j=1}^n\sum_{i=1}^n \omega_j v^i\big(\mathrm{d}q_j\big|\partial_{q_i}\big) = \sum_{j=1}^n \omega_j v^j\,. \tag{3.1}$$

3.6 Remark. <u>**The double dual**</u>

While there is no canonical identification of T_xM with its dual space T_x^*M, the double dual $T_x^{**}M := (T_x^*M)^*$ can be canonically identified with T_xM: Let $v \in T_xM$ then

$$\ell_v : T_x^*M \to \mathbb{R}\,, \quad u^* \mapsto \ell_v(u^*) := (u^*|\,v)$$

is a linear map and therefore $\ell_v \in T_x^{**}M$. The map $\iota : T_xM \to T_x^{**}M$, $v \mapsto \ell_v$, is a vector space isomorphism. To see this note that ι is clearly linear, that $\ker(\iota) = \{0\}$ implies that ι is injective, and thus because of finite dimension $\dim T_xM = \dim T_x^{**}M < \infty$ also surjective. $\diamond$

We thus keep in mind that a vector acts as a linear functional on covectors and a covector acts as a linear functional on vectors.

While we saw that a smooth map $f : M_1 \to M_2$ can be naturally used to push forward tangent vectors from T_xM_1 to $T_{f(x)}M_2$ via its differential $f_* = Df|_x$, the natural direction for covectors is the pullback from $T_{f(x)}^*M_2$ to $T_x^*M_1$.

3.7 Definition. <u>**The pullback of covectors**</u>

Let $f : M_1 \to M_2$ be a smooth map between manifolds and $x \in M_1$. The linear map

$$f_x^* : T_{f(x)}^*M_2 \to T_x^*M_1\,, \quad u^* \mapsto f_x^*u^* \quad \text{with} \quad (f_x^*u^*)(v) := u^*(f_*v) \text{ for all } v \in T_xM_1$$

is called the pullback of covectors under f.

3.8 Definition. <u>**The cotangent bundle**</u>

The **cotangent bundle** T^*M of M is the disjoint union of the cotangent spaces

$$T^*M := \bigcup_{x\in M}(\{x\} \times T_x^*M) = \{(x, v*)\,|\,x \in M\,,\ v^* \in T_x^*M\}\,.$$

As in the case of TM, also T^*M can be equipped with a canonical differentiable structure: Let $\mathcal{A} = (V_i, \varphi_i)$ be an atlas of M, then

$$T^*\mathcal{A} := (T^*V_i, (\varphi_i, \varphi_i^{-1*}))$$

is an atlas of T^*M. Here

$$\begin{aligned}(\varphi_i, \varphi_i^{-1*}) : T^*V_i &\to T^*\varphi_i(V_i) \subset T^*\mathbb{R}^n \\ (x, u^*) &\mapsto (\varphi_i(x), (\varphi_i^{-1})^* u^*)\,.\end{aligned}$$

3.9 Definition. **Covector fields or 1-forms**

A C^p-map $\omega : M \to T^*M$ with $\pi_M \circ \omega = \mathrm{id}_M$ is called a **C^p-covector field** or **1-form**. We denote the set of C^∞-covector fields by $\mathcal{T}_1^0(M)$.

As for vector fields we will often identify for a covector field $\omega \in \mathcal{T}_1^0(M)$ its value $\omega(x) \in \{x\}\times T_x^*M$ at $x \in M$ with its part in T_x^*M without making this explicit in the notation by projecting on the second factor. $\diamond$

3.10 Example. **The differential of a function as a 1-form**

Let $f \in C^\infty(M)$, then the map

$$\mathrm{d}f : M \to T^*M\,, \quad x \mapsto \mathrm{d}f|_x \in T_x^*M$$

defines a covector field $\mathrm{d}f \in \mathcal{T}_1^0(M)$.

3.11 Definition. **The pullback of 1-forms**

A smooth map $f : M_1 \to M_2$ allows to map ("pull back") covector fields on M_2 to covector fields on M_1. The map

$$f^* : \mathcal{T}_1^0(M_2) \to \mathcal{T}_1^0(M_1)\,, \qquad \omega \mapsto f^*\omega \quad \text{with} \quad (f^*\omega)(x) := f_x^*\,(\omega(f(x))$$

is called the **pullback**.

3.12 Remark. While one needs a diffeomorphism Φ in order to push-forward a vector field (cf. definition 2.27), one can pull back a covector field with any smooth map f.

3.13 Remark. **Coordinate representation of a 1-form resp. differential**

Using a coordinate chart (V, φ) for M, the restriction of a covector field $\omega \in \mathcal{T}_1^0(M)$ to V can be mapped to a covector field on $\varphi(V) \subset \mathbb{R}^n$ using the pullback $(\varphi^{-1})^*$,

$$(\varphi^{-1})^*\omega : \mathbb{R}^n \supset \varphi(V) \to T^*\varphi(V) = \varphi(V) \times \mathbb{R}^n\,, \quad q \mapsto (q, u^*(q)) \quad \text{with} \quad u^*(q) = \sum_{j=1}^n \omega_j(q)\, e_j\,,$$

where $\omega_j(q)$ are the components of the covector $\omega(x) \in T_x^*M$ with respect to the coordinate basis $(\mathrm{d}q_1, \dots, \mathrm{d}q_n)$, i.e. $\omega(x) = \sum_j \omega_j(\varphi^{-1}(q))\mathrm{d}q_j$.

Using once more (2.4) and recalling (3.1) (cf. homework assignment 15), we find that for $\omega = \mathrm{d}f$ the components with respect to the coordinate basis are $\omega_j(q) = \frac{\partial f}{\partial q_j}(q)$, i.e.

$$\mathrm{d}f = \sum_{j=1}^n \frac{\partial f}{\partial q_j}\,\mathrm{d}q_j\,. \tag{3.2}$$

Recall that $\frac{\partial f}{\partial q_j}(q)$ is a short hand for $\partial_j(f \circ \varphi^{-1})|_{\varphi(x)}$.

This implies immediately the **product rule**: Let $f, g \in C^\infty(M)$, then

$$\mathrm{d}(fg) = g\,\mathrm{d}f + f\,\mathrm{d}g\,. \qquad \diamond$$

3.14 Proposition. The pullback of a differential

Let $f \in C^\infty(M_1, M_2)$ and $g \in C^\infty(M_2)$. Then

$$f^*\mathrm{d}g = \mathrm{d}(g \circ f) =: \mathrm{d}(f^*g)\,,$$

where we also introduce the pullback of a function as $f^*g := g \circ f$.

$$\begin{array}{ccc} M_1 & \xrightarrow{f} & M_2 \\ f^*g \downarrow & & \downarrow g \\ \mathbb{R} & = & \mathbb{R} \end{array}$$

Proof. By the chain rule for differentials we have for all $x \in M_1$ and $v \in T_xM$ that

$$\mathrm{d}(g \circ f)|_x\, v = \mathrm{d}g|_{f(x)}\, Df|_x\, v = \mathrm{d}g|_{f(x)}(Df|_x\, v) = (f^*\mathrm{d}g)|_x\, v\,. \qquad \square$$

3.15 Example. Polar coordinates on $\mathbb{R}^2$

Polar coordinates on $\mathbb{R}^2$ are most easily defined by the map

$$\begin{aligned} \Phi : (0,\infty) \times (-\pi,\pi) &\;\to\; \mathbb{R}^2 \setminus \{x \in \mathbb{R}^2 \,|\, x_2 = 0 \text{ and } x_1 \le 0\} \\ (r,\theta) &\;\mapsto\; (r\cos\theta, r\sin\theta)\,. \end{aligned}$$

It is straight forward to check that Φ is indeed a diffeomorphism between open subsets of $\mathbb{R}^2$ and we might think of Φ^{-1} as a coordinate chart for (parts of) $\mathbb{R}^2$.

On the target $\mathbb{R}^2$ we have the basis 1-forms $\mathrm{d}x_1$ and $\mathrm{d}x_2$. In order to express them in terms of the coordinate 1-forms $\mathrm{d}r$ and $\mathrm{d}\theta$ we apply proposition 3.14, the product rule, and (3.2):

$$\begin{aligned} \Phi^*(\mathrm{d}x_1) &= \mathrm{d}(x_1 \circ \Phi) = \mathrm{d}(r\,\cos\theta) = \cos\theta\,\mathrm{d}r + r\,\mathrm{d}(\cos\theta) = \cos\theta\,\mathrm{d}r - r\,\sin\theta\,\mathrm{d}\theta \\ \Phi^*(\mathrm{d}x_2) &= \mathrm{d}(x_2 \circ \Phi) = \mathrm{d}(r\,\sin\theta) = \sin\theta\,\mathrm{d}r + r\,\mathrm{d}(\sin\theta) = \sin\theta\,\mathrm{d}r + r\,\cos\theta\,\mathrm{d}\theta\,. \end{aligned} \qquad \diamond$$

We will see later that 1-forms are the natural geometric objects that can be integrated along 1-dimensional (oriented) submanifolds, i.e. along curves. Without giving further details at this point, let us just give the definition of the integral of a 1-form along a curve as a first hint.

3.16 Definition. The line integral of a 1-form

Let M be a smooth manifold, $I = [a,b] \subset \mathbb{R}$ an interval, $\gamma \in C^\infty(I, M)$ a smooth curve, and $\omega \in \mathcal{T}^0_1(M)$ a 1-form. Then the **integral of ω along γ** is the number

$$\int_\gamma \omega := \int_I \gamma^*\omega := \int_a^b (\gamma^*\omega \,|\, e)(t)\,\mathrm{d}t\,,$$

where $\gamma^*\omega$ is the pullback of ω to I under γ and $e : I \to I \times \mathbb{R}$, $t \mapsto (t,1)$, is the unit vector field on I. The dual pairing $(\gamma^*\omega \,|\, e) \in C^\infty(I)$ between $\gamma^*\omega \in \mathcal{T}^0_1(I)$ and $e \in \mathcal{T}^1_0(I)$ is to be taken pointwise and defines a smooth function on I that is integrated in the standard Riemannian sense.

So the idea is to pull back the 1-form to parameter space and interpret the integral there as a usual Riemann integral. You will show in the homewrok assignments that the value of a line integral according to this definition is invariant under reparametrisations of the curve up to a sign coming from a possible change of orientation. Moreover, you will show the fundamental theorem of calculus, namely that for $\omega = \mathrm{d}f$

$$\int_\gamma \mathrm{d}f = f(\gamma(b)) - f(\gamma(a))\,.$$

German nomenclature

cotangent space = Kotangentialraum
double dual = Bidual
line integral = Wegintegral
product rule = Produktregel
1-form = 1-Form
differential = Differential
dual space = Dualraum
polar coordinates = Polarkoordinaten
pullback = Rückzug

4 Tensors

As explained in the previous chapter, covectors are by definition linear maps from the underlying vector space V into $\mathbb{R}$ and vectors can be understood in a canonical way as linear maps from the dual space V^* into $\mathbb{R}$. In a nutshell, **tensors** are multi-linear maps on cartesian products of the form $V^* \times \cdots \times V^* \times V \times \cdots \times V$. For example, a scalar product is a bilinear map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$. Also the (signed euclidean) area of a parallelogram spanned by two vectors $v, u \in \mathbb{R}^2$ is a bilinear function of the two vectors,

$$\text{area} : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}\,, \quad (u, v) \mapsto \text{area}(u, v) := u \wedge v := u_1 v_2 - u_2 v_1\,,$$

since

$$\begin{aligned}
\text{area}(\alpha u, v) = \text{area}(u, \alpha v) &= \alpha\, \text{area}(u, v) &&\text{for all } u, v \in \mathbb{R}^2 \text{ and } \alpha \in \mathbb{R}\\
\text{area}(u + w, v) &= \text{area}(u, v) + \text{area}(w, v) &&\text{for all } u, v, w \in \mathbb{R}^2\\
\text{area}(u, v + w) &= \text{area}(u, v) + \text{area}(u, w) &&\text{for all } u, v, w \in \mathbb{R}^2\,.
\end{aligned}$$

Such functions of several vectors or covectors that are linear in each argument are also called multi-linear forms or tensors. Multi-linear functions of tangent vectors and covectors to manifolds appear naturally in different geometrical and physical contexts. In the following we will discuss basic definitions and properties of tensors for arbitrary finite dimensional real vector spaces V, independently of the context of differentiable manifolds. But you can keep in mind that later on the tangent space T_xM at a point x in a differentiable manifold M will take the role of the vector space V.

4.1 Definition. **Tensors**

Let V be a n-dimensional vector space and V^* its dual. A multi-linear map

$$t : \underbrace{V^* \times \cdots \times V^*}_{r\text{-copies}} \times \underbrace{V \times \cdots \times V}_{s\text{-copies}} \to \mathbb{R}$$

is called a **tensor of type** (r, s) and we write, similar to the notation used for the dual pairing,

$$t(v_1^*, \ldots, v_r^*\,;\, v_1, \ldots, v_s) =: (t \,|\, v_1^*, \ldots, v_r^*\,;\, v_1, \ldots, v_s)\,.$$

For tensors t_1 and t_2 of the same type and α_1, $\alpha_2 \in \mathbb{R}$ one defines

$$(\alpha_1 t_1 + \alpha_2 t_2 \,|\, \ldots) := \alpha_1 (t_1 \,|\, \ldots) + \alpha_2 (t_2 \,|\, \ldots)\,,$$

and thereby equips the space of tensors of a given type with the structure of a vector space that will be denoted by V^r_s. In particular, $V^* = V^0_1$ and $V = V^1_0$. ◇

Given, for example, two covectors $v_1^*, v_2^* \in V^*$, one can define the bilinear map

$$v_1^* \otimes v_2^* : V \times V \to \mathbb{R}\,, \qquad (u_1, u_2) \mapsto v_1^* \otimes v_2^*(u_1, u_2) := (v_1^* | \, u_1)\, (v_2^* | \, u_2)\,,$$

called the tensor product of v_1^* and v_2^*. In the same way one can define tensor products of arbitrary tensors in order to define new tensors of higher order.

4.2 Definition. <u>**The tensor product of tensors**</u>

Let V be a n-dimensional vector space and $t_1 \in V^r_s$ and $t_2 \in V^{r'}_{s'}$. The **tensor product** $t_1 \otimes t_2$ is the tensor of type $(r + r', s + s')$ defined by

$$(t_1 \otimes t_2 \,|\, u^*_1, \ldots, u^*_{r+r'}\,,\, u_1, \ldots, u_{s+s'}) = (t_1|\, u^*_1, \ldots, u^*_r\,;\, u_1, \ldots, u_s)\,(t_2|\, u^*_{r+1}, \ldots, u^*_{r+r'}\,;\, u_{s+1}, \ldots, u_{s+s'})\,.$$

Note that multi-linearity is obvious from this definition.

4.3 Remark. It follows directly from the definition that the map

$$\otimes : V^r_s \times V^{r'}_{s'} \to V^{r+r'}_{s+s'}$$

is associative and distributive, but not commutative. $\diamond$

While not every tensor can be written as a tensor product of vectors and covectors, linear combinations of such tensor products span the whole spaces V^r_s.

4.4 Proposition. Let $\{e_j\}$ and $\{e^*_i\}$ be bases of $V = V^1_0$ and $V^* = V^0_1$ respectively. Then every $t \in V^r_s$ can be uniquely written in the form

$$t = \sum_{(i),(j)} t^{j_1 \cdots j_r}_{i_1 \cdots i_s}\; e_{j_1} \otimes \cdots \otimes e_{j_r} \otimes e^*_{i_1} \otimes \cdots \otimes e^*_{i_s} \qquad (*)$$

with coefficients $t^{j_1 \cdots j_r}_{i_1 \cdots i_s} \in \mathbb{R}$. In the sum all indices $j_1, \ldots, j_r$ and $i_1, \ldots, i_s$ run from 1 to n. Thus, the n^{r+s} tensor products

$$e_{j_1} \otimes \cdots \otimes e_{j_r} \otimes e^*_{i_1} \otimes \cdots \otimes e^*_{i_s}\,, \qquad j_1, \ldots, j_r, i_1, \ldots, i_s = 1, \ldots, n$$

form a basis of V^r_s and the space V^r_s has dimension n^{r+s}.

Proof. Let $\{b^*_j\}$ and $\{b_i\}$ be the bases of V^* and V that are dual to $\{e_j\}$ and $\{e^*_i\}$, i.e.

$$(b^*_j|\, e_i) = \delta_{ij} \qquad \text{and} \qquad (e^*_i|\, b_j) = \delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}\,.$$

Note that the dual bases are unique because a linear map is uniquely specified when given its action on a basis. If we now define

$$t^{j_1 \cdots j_r}_{i_1 \cdots i_s} = t(b^*_{j_1}, \ldots, b^*_{j_r}, b_{i_1}, \ldots, b_{i_s})\,, \tag{4.1}$$

then $(*)$ holds on elements of the form $(b^*_{j_1}, \ldots, b^*_{j_r}, b_{i_1}, \ldots, b_{i_s})$ and by multi-linearity of both sides of $(*)$ on any tupel $(u^*_1, \ldots, u^*_r, u_1, \ldots, u_s)$. Uniqueness follows from linear independence of the different tensor products $e_{j_1} \otimes \cdots \otimes e_{j_r} \otimes e^*_{i_1} \otimes \cdots \otimes e^*_{i_s}$. $\square$

4.5 Remark. Note that we use the standard convention to write the coefficients of a tensor with vector indices (also called contravariant indices) as upper indices and covector indices (also called covariant indices) as lower indices. It is not uncommon in physics to call "tensor" the components $t^{j_1 \ldots j_r}_{i_1 \ldots i_s}$ of a tensor t and to say that the array of numbers $t^{j_1 \cdots j_r}_{i_1 \ldots i_s}$ is a tensor of type (r, s) because it transforms under basis changes in the appropriate way. $\diamond$

4.6 Definition. <u>**Metric tensor**</u>

(a) A tensor $g \in V^0_2$ that is symmetric, i.e. that satisfies

$$g(u, v) = g(v, u) \qquad \text{for all } u, v \in V^1_0\,,$$

and positive definite, i.e.

$$g(v, v) > 0 \qquad \text{for all } v \neq 0\,,$$

is called a **metric tensor** or **scalar product**.

(b) A tensor $g \in V_2^0$ with the property that

$$g(v,u) = 0 \quad \text{for all } u \in V \qquad \Rightarrow \qquad v = 0 \tag{4.2}$$

is called **non-degenerate**.

Any metric tensor g is non-degenerate, since $g(v,u) = 0$ for all u implies that, in particular, $g(v,v) = 0$ and thus by definiteness $v = 0$. Examples for (b) are **pseudo-metrics** (g symmetric and non-degenerate) or **symplectic forms** (g skew-symmetric and non-degenerate). ◇

4.7 Remark. Recall that a metric tensor allows us to define the "length" or norm of a vector $v \in V$ as $\|v\| := \sqrt{g(v,v)}$ and the angle between vectors as $g(v,u) =: \cos(\measuredangle(v,u))\, \|v\|\, \|u\|$. In particular, two vectors are orthogonal with respect to the metric g, if $g(u,v) = 0$.

4.8 Remark. With the help of a non-degenerate $g \in V_2^0$ one can identify a vector space V and its dual space V^*:

For $v \in V$ the map

$$g(v,\cdot) : V \to \mathbb{R}\,, \quad u \mapsto g(v,u)$$

is linear and hence defines an element of V^*. The map

$$\iota_g : V \to V^*\,, \quad v \mapsto \iota_g(v) := g(v,\cdot)$$

is again linear. Because of (4.2), its kernel contains only the zero vector and thus ι_g is an isomorphism. ◇

4.9 Notation. <u>**The index calculus**</u>

Let $(e_j)_{j=1,\dots,n}$ be a basis of V and let (from now on always) be $(e^i)_{i=1,\dots,n}$ (upper index!) the dual basis of V^* defined by

$$e^i(e_j) = \delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

To avoid confusion between indices of components of vectors and covectors, components of vectors have upper indices, while components of covectros have lower indices. For example, the basis representation of a vector $v \in V$ with respect to (e_i) is

$$v = \sum_{i=1}^n v^i e_i =: v^i e_i$$

and the basis representation of a covector $u \in V^*$ with respect to (e^i) is

$$u = \sum_{i=1}^n u_i e^i =: u_i e^i\,.$$

Note that in the expressions above we applied the **Einstein summation convention** where summation over indices appearing twice is understood implicitly.

Similarly, the basis representation of a metric tensor $g \in V_2^0$ is written as

$$g = \sum_{i,j=1}^n g_{ij}\, e^i \otimes e^j =: g_{ij}\, e^i \otimes e^j\,.$$

These conventions provide a very elegant calculus that simpifies computations in components considerably. Here are some examples: Let $u, v \in V$ with $u = u^n e_n$ and $v = v^n e_n$ and $w \in V^*$ with $w = w_n e^n$ then

(a) $w(v) = w_n e^n(v^m e_m) = w_n v^m e^n(e_m) = w_n v^n$

(b) $\langle u \,|\, v\rangle_g = g(u,v) = g_{ij}\,(e^i \otimes e^j)(u^m e_m, v^n e_n) = g_{ij}\,u^m\,v^n\,e^i(e_m)\,e^j(e_n) = g_{ij}\,u^i\,v^j = u^m\,g_{mn}\,v^n$

(c) From (a) and (b) it follows that $\iota_g(u) = g(u,\cdot)$ has the basis representation $\iota_g(u) = u_n e^n$ with components $u_n = u^m\,g_{mn}$.

(d) We denote by g^{ij} the entries of the inverse matrix of (g_{ij}), i.e.

$$g_{ij}\,g^{jn} = g^{ij}\,g_{jn} = \delta^i_n := \delta_{in}\,.$$

Thus in (c) we also have $u^n = u_m\,g^{mn}$ and $\iota_g(e_i) = g(e_i,\cdot) = g_{ij}e^j$. Hence, in general, $\iota_g(e_i) \neq e^i$.

(e) The space V^1_1 is canonically isomorphic to the space of endomorphism $\mathcal{L}(V)$ of V. Let $a \in V^1_1$, then

$$A : V \to V\,, \quad u \mapsto Au := \iota^{-1}(a(\cdot,u))$$

defines an endomorphism. Here we used the identification of V and V^{**} introduced in remark 3.6. Vice versa, an endomorphism $A : V \to V$ defines a tensor $a \in V^1_1$ through

$$a(v^*,u) := (v^* \,|\, Au)\,.$$

The maps $a \mapsto A$ and $A \mapsto a$ are clearly linear and inverse to each other.

Using the basis representation we find that the components a^j_i of $a \in V^1_1$ are just the matrix entries of the endomorphism A,

$$(e^i \,|\, Ae_j) = (e^i \,|\, \iota^{-1}(a(\cdot,e_j))) = a(e^i,e_j) = a^i_j \quad \Rightarrow \quad (Ae_j)^i = a^i_j \quad \Rightarrow \quad Ae_j = a^i_j e_i\,,$$

where we used the formula (4.1) to compute the components of a.

(f) In the same way the space V^1_1 is also canonically isomorphic to the space of endomorphism $\mathcal{L}(V^*)$ of V^*. Let $a \in V^1_1$, then

$$A^* : V^* \to V^*\,, \quad u^* \mapsto A^*u^* := a(u^*,\cdot)$$

defines an endomorphism. Given an element $a \in V^1_1$, the corresponding A^* is really the adjoint map of the corresponding A,

$$(v^* \,|\, Au) = a(v^*,u) = (A^*v^* \,|\, u)\,.$$

We now discuss how a change of basis affects the components of tensors. Let $A : V \to V$ be an isomorphism and define a new basis $(\hat{e}_i)$ of V by putting $\hat{e}_i := Ae_i$, with dual basis $(\hat{e}^i)$. The endomorphism $B : V^* \to V^*$ that connects the dual bases as $\hat{e}^i := Be^i$ is determined by

$$\delta_{ji} \overset{!}{=} (\hat{e}^i \,|\, \hat{e}_j) = (Be^i \,|\, Ae_j) =: (A^*Be^i \,|\, e_j) \quad \Rightarrow \quad B = (A^*)^{-1}\,. \tag{4.3}$$

If a^j_i is the matrix of A, i.e. $Ae_i = a^j_i e_j$, then $A^*e^j = a^j_i e^i$ and thus the matrix of b^j_k of B, i.e. $Be^j = b^j_k e^k$, must satisfy $b^j_k a^k_i = \delta_{ji}$. Hence, as a matrix, b^j_k is the inverse of a^j_k. However, b^j_k is the matrix of an endomorphism $B : V^* \to V^*$, while a^j_k is the matrix of an endomorphism $A : V \to V$.

The components of an arbitrary tensor $t \in V^r_s$ transform because of

$$t = t^{j_1\cdots j_r}_{i_1\cdots i_s}\; e_{j_1} \otimes \cdots \otimes e_{j_r} \otimes e^{i_1} \otimes \cdots \otimes e^{i_s} = \hat{t}^{l_1\cdots l_r}_{k_1\cdots k_s}\; \hat{e}_{l_1} \otimes \cdots \otimes \hat{e}_{l_r} \otimes \hat{e}^{k_1} \otimes \cdots \otimes \hat{e}^{k_s}$$

according to

$$\hat{t}^{l_1\cdots l_r}_{k_1\cdots k_s}\; a^{j_1}_{l_1} \cdots a^{j_r}_{l_r} b^{k_1}_{i_1} \cdots b^{k_s}_{i_s} = t^{j_1\cdots j_r}_{i_1\cdots i_s} \quad \text{or} \quad \hat{t}^{l_1\cdots l_r}_{k_1\cdots k_s} = t^{j_1\cdots j_r}_{i_1\cdots i_s}\; b^{l_1}_{j_1} \cdots b^{l_r}_{j_r} a^{i_1}_{k_1} \cdots a^{i_s}_{k_s}\,.$$

◇

4.10 Remark. Given a non-degenerate $g \in V_2^0$, one often identifies a vector u and the covector $\iota_g(u) = g(u, \cdot)$. Then, mostly in the physics literature, u^j are called the **contravariant** and u_j the **covariant** components of u, although u_j are actually the components of $\iota_g(u)$. ⋄

4.11 Definition. A non-degenerate $g \in V_2^0$ allows also for a canonical identification of any tensor space V_s^r with V_r^s and with V_{s+r}^0 and V_0^{s+r} by concatenating ι_g or ι_g^{-1} in the respective arguments. In particular,

$$I_g : V_s^r \to V_r^s\,, t \mapsto t \circ (\underbrace{\iota_g, \cdots, \iota_g}_{s \text{ copies}}, \underbrace{\iota_g^{-1}, \cdots, \iota_g^{-1}}_{r \text{ copies}})\,.$$

Within the index calculus this identification takes the following form. Let $t \in V_s^r$ with $t = t^{i_1 \ldots i_r}_{j_1 \ldots j_s} e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}$, then e.g.

$$\begin{aligned}
\tilde{t} &= t^{i_1 \ldots i_r}_{j_1 \ldots j_s} \iota_g(e_{i_1}) \otimes \cdots \otimes \iota_g(e_{i_r}) \otimes e^{j_1} \otimes \cdots \otimes e^{j_s} \\
&= t^{i_1 \ldots i_r}_{j_1 \ldots j_s}\; g_{i_1 n_1} \cdots g_{i_r n_r}\; e^{n_1} \otimes \cdots \otimes e^{n_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s} \\
&=: t_{j_1 \ldots j_s\, n_1 \ldots n_r}\; e^{n_1} \otimes \cdots \otimes e^{n_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s} \;\in V_{r+s}^0\,.
\end{aligned}$$

In general one can thus use g^{ij} and g_{ij} to raise resp. lower tensor indices and thereby change the tensor type from (r, s) to $(r+1, s-1)$ resp. $(r-1, s+1)$. ⋄

4.12 Definition. A non-degenerate bilinear map $g \in V_2^0$ can be lifted to a non-degenerate bilinear map on arbitrary tensors,

$$G : V_s^r \times V_s^r \to \mathbb{R}\,, \quad (t, \tilde{t}) \mapsto G(t, \tilde{t}) = (I_g(t) \,|\, \tilde{t})\,.$$

In index notation we have

$$G(t, \tilde{t}) = t^{i_1 \ldots i_r}_{j_1 \ldots j_s}\; \tilde{t}^{m_1 \cdots m_r}_{n_1 \cdots n_s}\; g_{i_1 m_1} \cdots g_{i_r m_r}\; g^{j_1 n_1} \cdots g^{j_s n_s}\,.$$

If g is a metric on V, then G is a metric on V_s^r. ⋄

4.13 Definition. <u>Contraction of tensors</u>

By **contracting** two indices, more precisely one upper, say the ℓth, and one lower, say the kth, of a tensor $t \in V_s^r$ with

$$t = t^{i_1 \cdots i_r}_{j_1 \cdots j_s} e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}$$

one obtains a tensor $\tilde{t} \in V_{s-1}^{r-1}$ by putting

$$\tilde{t} = t^{i_1 \cdots m \cdots i_r}_{j_1 \cdots m \cdots j_s}\, e_{i_1} \otimes \cdots \otimes \hat{e}_{i_\ell} \otimes \cdots \otimes e^{j_1} \otimes \cdots \otimes \hat{e}^{j_k} \otimes \cdots = \sum_{m=1}^{n} t(\cdots, e^m, \cdots, e_m, \cdots)\,.$$

Here the hat over a factor means that it is left out and in the sum on the right hand side the vector e^m is in the ℓth covector-slot and e_m is in the kth vector-slot. It is easy to check that the resulting tensor $\tilde{t}$ is independent of the chosen basis for V.
Unless $t^{i_1 \cdots i_r}_{j_1 \cdots j_s}$ is symmetric in all upper and lower indices, the resulting tensor $\tilde{t}$ depends on which indices are contracted.
For $a \in V_1^1$ the contraction $\operatorname{tr} a := a^i_i$ is called the trace of a and it is just the usual trace of the corresponding endomorphism $A : V \to V$. For a metric $g \in V_2^0$ and $t \in V_0^2$ resp. $\tilde{t} \in V_2^0$ the contractions $\operatorname{tr}_g t = t^{ij} g_{ij}$ resp. $\operatorname{tr}_g \tilde{t} = t_{ij}\, g^{ij}$ are called the metric traces. ⋄

We now define the tensor bundles over a manifold M by attaching locally the tensor spaces $T_x{}^r_s M$ associated with the tangent space $V = T_x M$.

4.14 Definition. **Tensor bundles**

We define $T^r_s M = \bigcup_{x\in M}(\{x\} \times T_x{}^r_s M) = \{(x,t)\,|\, x \in M\,,\; t \in T_x{}^r_s M\}$ as the bundle of tensors of type (r,s). Any atlas $\mathcal{A} = (V_i, \varphi_i)$ of M yields a natural atlas on $T^r_s M$ through

$$T^r_s \mathcal{A} = (T^r_s V_i, \tilde{\varphi}_i)\,.$$

Here

$$\tilde{\varphi}_i : T^r_s V_i \to T^r_s \varphi_i(V_i)$$

is defined by

$$\tilde{\varphi}_i(x, e_{k_1} \otimes \cdots \otimes e_{k_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}) = (\varphi_i(x), \varphi_{i*}e_{k_1} \otimes \cdots \otimes \varphi_{i*}e_{k_r} \otimes (\varphi_i^{-1})^* e^{j_1} \otimes \cdots \otimes (\varphi_i^{-1})^* e^{j_s})$$

and linearity in each fibre.

4.15 Definition. **Tensor fields**

A C^∞-map $t : M \to T^r_s M$ with $\pi_M \circ t = \mathrm{id}_M$ is called a **tensor field** and we denote the space of tensor fields by $\mathcal{T}^r_s(M)$. We set $\mathcal{T}^0_0(M) := C^\infty(M)$.

4.16 Remarks. (a) Locally any tensor field can be expressed in terms of coordinate basis fields:

$$t(x) = t^{i_1\cdots i_r}_{j_1\cdots j_s}(x)\; \partial_{q_{i_1}} \otimes \cdots \otimes \partial_{q_{i_r}} \otimes \mathrm{d}q^{j_1} \otimes \cdots \otimes \mathrm{d}q^{j_s}$$

with $t^{(i)}_{(j)} \in C^\infty(M)$. In the physics literature sometimes the component functions $t^{(i)}_{(j)}(x)$ are called tensor fields.

(b) If there exist n vector fields $e_i \in \mathcal{T}^1_0(M)$ that are pointwise linear independent, i.e. if M is parallelisable, then the covector fields $e^j \in \mathcal{T}^0_1(M)$ that are pointwise dual are pointwise linear independent in $\mathcal{T}^0_1(M)$ as well. By taking tensor products, one obtains basis sections of all tensor bundles. We thus have: if M is parallelisable, then all tensor bundles over M are trivialisable.

4.17 Definition. **Riemannian and pseudo-Riemannian metrics**

A non-degenerate and symmetric bilinear form $g \in \mathcal{T}^0_2(M)$ is called a **pseudo-Riemannian metric** and the pair (M, g) a **pseudo-Riemannian manifold**. If g is, in addition, fibre-wise positive definite, then it is called a **Riemannian metric** and (M, g) a **Riemannian manifold**.

4.18 Example. **Minkowski space**

On $M = \mathbb{R}^4$ a pseudo-Riemannian metric is given by

$$\eta = \eta_{ij}\, \mathrm{d}q^i \otimes \mathrm{d}q^j$$

with

$$\eta_{ij} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

It is called the Minkowski metric and the space M equipped with the Minkowski metric η is called Minkowski space. It is the space-time manifold of special relativity. ◇

4.19 Remark. **Abstract index notation**

Instead of using indices to label components of tensors in specific basis representations with respect to local charts, one can also use the indices to merely have a concise notation for indicating the type of a tensor and to have a shorthand for operations like the contraction that are difficult to express in the coordinate free notation.

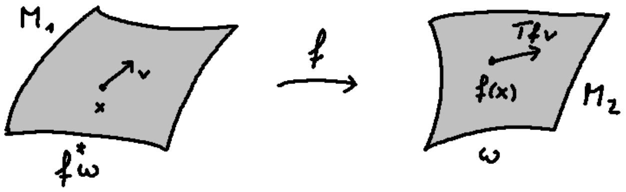
One would write for example for the Riemann tensor $R^a_{bcd} \in V^1_3$ and obtain the Ricci tensor by contracting to $R_{bd} := R^a_{bad} \in V^0_2$. Here the indices are merely a notational tool, not related to any basis and, in particular, have no numerical values. The notation R^a_{bad} then indicates that one contracts the first and the third argument of the tensor to obtain a new one. $\diamond$

We end this section with a few straight forward observations. First we can extend the pull back map to arbitary multilinear forms in an obvious way.

4.20 Definition. **The pull-back of multilinear forms**

Let $f : M_1 \to M_2$ be smooth and $\omega \in \mathcal{T}^0_p(M_2)$. Then $f^*\omega \in \mathcal{T}^0_p(M_1)$ is defined as

$$f^*\omega|_x(v_1, \dots, v_p) = \omega|_{f(x)}(f_*v_1, \dots, f_*v_p) \qquad \text{for all } v_1, \dots, v_p \in T_xM_1 \,. \qquad \diamond$$



Next we clarify how a diffeomorphism $\Phi : M_1 \to M_2$ between manifolds induces a diffeomorphism of corresponding tensor bundles of equal type. First observe that the fibre wise pull-back defines a diffeomorphism of the cotangent bundles,

$$T^*M_2 \to T^*M_1, \quad (y, \omega) \mapsto \left(\Phi^{-1}(y), \Phi^*\omega|_{\Phi^{-1}(y)}\right) \,.$$

For the inverse of this map we write

$$D^*\Phi : T^*M_1 \to T^*M_2 \,, \quad (x, \omega) \mapsto (\Phi(x), \Phi^{-1*}\omega|_{\Phi(x)}) \,.$$

It holds that

$$(D^*\Phi\omega \,|\, D\Phi v)_{\Phi(x)} = (\omega \,|\, v)_x \quad \text{for all } \omega \in T^*_xM_1,\ v \in T_xM_1 \,,$$

since $(D^*\Phi\omega)(D\Phi v) = \omega(D\Phi^{-1} \circ D\Phi\, v) = \omega(v)$.

We obtain diffeomorphisms of arbitrary tensor bundles through

$$D\Phi \otimes \cdots \otimes D\Phi \otimes D^*\Phi \otimes \cdots \otimes D^*\Phi : T^r_sM_1 \to T^r_sM_2 \,,$$

where one defines on products

$$\begin{aligned} D\Phi \otimes \cdots \otimes D\Phi \otimes D^*\Phi \otimes \cdots \otimes D^*\Phi(x,\, u_1 \otimes \cdots \otimes u_r \otimes \omega^1 \otimes \cdots \otimes \omega^s) \\ := \quad (\Phi(x), D\Phi u_1 \otimes \cdots \otimes D\Phi u_r \otimes D^*\Phi\omega^1 \otimes \cdots \otimes D^*\Phi\omega^s) \end{aligned}$$

and then extends this definition by linearity in every fibre $T_x{}^r_sM_1$.

4.21 Definition. **The push-forward of tensor fields**

A diffeomorphism $\Phi : M_1 \to M_2$ induces a map $\Phi_* : \mathcal{T}^r_s(M_1) \to \mathcal{T}^r_s(M_2)$, called the **push-forward** of tensor fields, that is defined by commutativity of the diagram

$$\begin{array}{ccc} M_1 & \xrightarrow{\Phi} & M_2 \\ t \downarrow & & \downarrow \Phi_*t \\ T^r_sM_1 & \xrightarrow{D\Phi\otimes\cdots\otimes D\Phi\otimes D^*\Phi\otimes\cdots\otimes D^*\Phi} & T^r_sM_2 \end{array}$$

i.e. for $t \in \mathcal{T}^r_s(M_1)$ we set

$$\Phi_*t = \underbrace{D\Phi \otimes \cdots \otimes D\Phi}_{r\text{-copies}} \otimes \underbrace{D^*\Phi \otimes \cdots \otimes D^*\Phi}_{s\text{-copies}} \circ t \circ \Phi^{-1} \,.$$

Again the chain rule $(\Phi \circ \Psi)_* = \Phi_*\Psi_*$ holds and the pull-back Φ^* can be extended to arbitrary tensor fields by defining $\Phi^* := (\Phi^{-1})_*$.

4.22 Examples. (a) For a function $f \in \mathcal{T}^0_0(M_1)$ we have $\Phi_* f = f \circ \Phi^{-1}$.

(b) For a vector field $X \in \mathcal{T}^1_0(M_1)$ the push-forward $\Phi_* X = D\Phi \circ X \circ \Phi^{-1}$ was already defined definition 2.27.

(c) The push-forward under a diffeomorphism commutes with the differential: for $f \in \mathcal{T}^0_0(M_1)$ we have $\Phi_* \mathrm{d}f = \mathrm{d}(\Phi_* f)$.

(d) Let (V, φ) be a chart on M, let e_i be the ith canonical unit vector field on $\mathbb{R}^n$, and let $e^i : \mathbb{R}^n \to \mathbb{R}$, $q = (q_1, \ldots, q_n) \mapsto q_i$ denote the ith coordinate function on $\mathbb{R}^n$. Then the coordinate 1-forms and the coordinate vector fields on $V \subset M$ are given by $\mathrm{d}q^i = \varphi^* \mathrm{d}e^i$ and $\partial_{q_j} = (\varphi^{-1})_* e_j$. From this the tranformation laws for ∂_{q_j} and $\mathrm{d}q^i$ under coordinate changes follow immediately: Let $\tilde\varphi$ be another chart on V, $\partial_{\tilde q_j} = (\tilde\varphi^{-1})_* e_j$ and $\mathrm{d}\tilde q^i = \tilde\varphi^* \mathrm{d}e^i$ the corresponding coordinate forms and vector fields, and $\Phi = \tilde\varphi \circ \varphi^{-1}$ the transition map in $\mathbb{R}^n$.

$$
\begin{tikzcd}
\partial_{q_j} & \in & \mathcal{T}^1_0(V) \arrow[dl, "\varphi_*"'] \arrow[dr, "\tilde\varphi_*"] & \ni & (D\Phi)^i_j\,\partial_{\tilde q_i} \\
& \mathcal{T}^1_0(U) \arrow[rr, "\Phi_*"] & & \mathcal{T}^1_0(\tilde U) & \\
e_j \arrow[uu, "\varphi_*^{-1}"] \arrow[rrrr, "\Phi_*"] \arrow[ur, phantom, "\in"] & & & & \Phi_* e_j = (D\Phi)^i_j\, e_i \arrow[uu, "\tilde\varphi_*^{-1}"'] \arrow[ul, phantom, "\ni"]
\end{tikzcd}
.
$$

Since the diagram is commutative,

$$\partial_{q_j} = (\varphi^{-1})_* e_j = (\varphi^{-1})_* (\Phi^{-1})_* \Phi_* e_j = (\tilde\varphi^{-1})_* (D\Phi)^i_j\, e_i = (D\Phi)^i_j \partial_{\tilde q_i}$$

and thus $\partial_{\tilde q_j} = (D\Phi^{-1})^i_j\, \partial_{q_i}$. It then follows from (4.3) that $\mathrm{d}\tilde q^j = (D\Phi)^j_i\, \mathrm{d}q^i$ and $\mathrm{d}q^j = (D\Phi^{-1})^j_i\, \mathrm{d}\tilde q^i$. ◇

While the dual pairing $(\,|\,)$ is invariant under diffeomorphisms,

$$(\Phi_* \omega \,|\, \Phi_* v) = (\omega \,|\, v)\,,$$

this does not hold for scalar products in general, but only for diffeomorphisms Φ that leave invariant the metric g.

4.23 Definition. <u>**Isometries and canonical transformations**</u>

Let M_1 and M_2 be smooth manifolds and let $g_1 \in \mathcal{T}^0_2(M_1)$ and $g_2 \in \mathcal{T}^0_2(M_2)$ be non-degenerate. A diffeomorphism $\Phi : M_1 \to M_2$ with $g_2 = \Phi_* g_1$, i.e.

$$g_2(\Phi_* v, \Phi_* u) \circ \Phi = g_1(v, u) \qquad \text{for all } v, u \in \mathcal{T}^1_0(M_1)\ ,$$

is called an **isometry**, if g_1 and g_2 are (pseudo-)Riemannian metrics, resp. **canonical transformation** or **symplectomorphism**, if g_1 and g_2 are symplectic forms.

German nomenclature

contravariant = kontravariant	covariant = kovariant
index = Index	isometry = Isometrie
metric tensor = metrischer Tensor, Metrik	non-degenerate = nicht entartet
pull-back = Rückzug	riemannian metric = Riemannsche Metrik
scalar product = Skalarprodukt	symplectic form = symplektische Form
tensor = Tensor	tensor product = Tensorprodukt

5 Differential forms and the exterior derivative

In this section we discuss special types of tensors, so called differential forms or k-forms. A k-form is a tensor of type $(0, k)$ that is alternating, i.e. skew-symmetric in all arguments. For $k = 1$ these are just the 1-forms or covectors. A k-form takes k vectors as arguments and defines a k-dimensional volume spanned by these k-vectors. Hence, as we will see in the next section, one can integrate k-forms over k-dimensional manifolds.

5.1 Definition. <u>**Alternating k-forms and the exterior product of covectors**</u>

Let V be a real n-dimensional vector space.

(a) A tensor $\omega \in V_k^0$, $1 \le k \le n$, is called **alternating**, if it is skew-symmetric in all arguments, i.e. if for all vectors $u_1, \ldots, u_k \in V$ and permutations $\pi \in S_k$ it holds that

$$\omega(u_{\pi(1)}, \ldots, u_{\pi(k)}) = \operatorname{sgn}(\pi)\, \omega(u_1, \ldots, , u_k)\,.$$

In particular, the exchange of two arguments changes the sign of ω.

The subspace of alternating tensors in V_k^0 is denoted by Λ_k and its elements are called **exterior forms**, **alternating k-forms**, or just **k-forms**. One defines $\Lambda_0 := V_0^0 := \mathbb{R}$.

(b) The **exterior product** (or **wedge product**) $v_1^* \wedge \cdots \wedge v_k^*$ of k covectors is defined as

$$(v_1^* \wedge \cdots \wedge v_k^* \,|\, u_1, \ldots, u_k) := \det\left((v_i^* \,|\, u_j)\right).$$

Skew-symmetry of the determinant under permutations of the columns implies that $v_1^* \wedge \cdots \wedge v_k^*$ is alternating and thus $v_1^* \wedge \cdots \wedge v_k^* \in \Lambda_k$. Analogously, skew-symmetry of the determinant under permutations of the rows implies that

$$v_{\pi(1)}^* \wedge \cdots \wedge v_{\pi(k)}^* = \operatorname{sgn}(\pi)\, v_1^* \wedge \cdots \wedge v_k^*\,. \tag{5.1}$$

Finally, using the Leibniz formula for the determinant, we note that

$$v_1^* \wedge \cdots \wedge v_k^* = \sum_{\pi \in S_k} \operatorname{sgn}(\pi)\, v_{\pi(1)}^* \otimes \cdots \otimes v_{\pi(k)}^* \tag{5.2}$$

5.2 Reminder. An alternating k-form ω defines the signed “volume” that is spanned by k vectors $v_1, \ldots, v_k$ in T_xM via

$$\omega(v_1, \ldots, v_k) =: \operatorname{vol}_k(v_1, \ldots, v_k)\,.$$

Here “volume” is to be understood in a dimension-dependent sense. For $k = 1$ it means length, for $k = 2$ area, and for $k = 3$ volume in the usual sense. In case you are not familiar with the connection between alternating forms and volume it is recommended that you read e.g. the section “Heuristics of volume measurement” in the book *Introduction to smooth manifolds* by John M. Lee. ◇

5.3 Proposition. <u>**A basis for Λ_k**</u>

Let $(e^j)_{j=1,\ldots,n}$ be a basis of V^* and $1 \le k \le n$, then the k-forms

$$e^{j_1} \wedge \cdots \wedge e^{j_k} \quad \text{with} \quad 1 \le j_1 < \cdots < j_k \le n$$

form a basis of the space of alternating k-forms $\Lambda_k \subset V_k^0$. Thus, $\dim(\Lambda_k) = \binom{n}{k}$, and, in particular, $\Lambda_n = \Lambda_0 = \mathbb{R}$ and $\Lambda_k = \{0\}$ for $k > n$.

Proof. First note that because of (5.1) all other wedge products $e^{j_1} \wedge \cdots \wedge e^{j_k}$ of k basis vectors than those with $1 \leq j_1 < \cdots < j_k \leq n$ are either zero (if the same basis vector appears twice) or are a linear multiple of the corresponding product with indices ordered increasingly.

Let $\omega \in \Lambda_k$ and let $(e_i)_{i=1,\dots,n}$ be the dual basis to $(e^j)_{j=1,\dots,n}$. Then for $u_1, \dots, u_k \in V$ we have $u_j = \sum_{i=1}^n e^i(u_j)\, e_i$ for $j = 1, \dots, k$ and

$$\begin{aligned}
\omega(u_1, \dots, u_k) &= \omega\left(\sum_{i_1=1}^n e^{i_1}(u_1)\, e_{i_1}, \dots, \sum_{i_k=1}^n e^{i_k}(u_k)\, e_{i_k}\right) \\
&= \sum_{i_1=1}^n \cdots \sum_{i_k=1}^n e^{i_1}(u_1) \cdots e^{i_k}(u_k)\, \omega\left(e_{i_1}, \dots, e_{i_k}\right) \\
&= \sum_{i_1=1}^n \cdots \sum_{i_k=1}^n (e^{i_1} \otimes \cdots \otimes e^{i_k} \,|\, u_1, \dots, u_k)\, \tfrac{1}{k!} \sum_{\pi \in S_k} \operatorname{sgn}(\pi)\, \omega\left(e_{i_{\pi(1)}}, \dots, e_{i_{\pi(k)}}\right) \\
&\overset{j_\ell := i_{\pi(\ell)}}{=} \tfrac{1}{k!} \sum_{j_1=1}^n \cdots \sum_{j_k=1}^n \omega\left(e_{j_1}, \dots, e_{j_k}\right) \sum_{\pi \in S_k} \operatorname{sgn}(\pi) (e^{j_{\pi^{-1}(1)}} \otimes \cdots \otimes e^{j_{\pi^{-1}(k)}} \,|\, u_1, \dots, u_k) \\
&= \tfrac{1}{k!} \sum_{j_1=1}^n \cdots \sum_{j_k=1}^n \omega\left(e_{j_1}, \dots, e_{j_k}\right) (e^{j_1} \wedge \cdots \wedge e^{j_k} \,|\, u_1, \dots, u_k) \\
&= \sum_{j_1=1}^{n-k+1} \sum_{j_2=j_1+1}^{n-k+2} \cdots \sum_{j_k=j_{k-1}+1}^{n} \omega\left(e_{j_1}, \dots, e_{j_k}\right) (e^{j_1} \wedge \cdots \wedge e^{j_k} \,|\, u_1, \dots, u_k)\,.
\end{aligned}$$

Thus

$$\omega = \sum_{i_1=1}^{n-k+1} \sum_{i_2=i_1+1}^{n-k+2} \cdots \sum_{i_k=i_{k-1}+1}^{n} \omega_{i_1 \cdots i_k}\, e^{i_1} \wedge \cdots \wedge e^{i_k}$$

with the skew-symmetric coefficients $\omega_{i_1 \cdots i_k} = \omega(e_{i_1}, \dots, e_{i_k})$. □

5.4 Remark. According to (4.1) we have the basis representation $\omega = \omega_{i_1 \cdots i_k} e^{i_1} \otimes \cdots \otimes e^{i_k}$ in V_k^0. When using the basis representation in Λ_k, we will often abbreviate $J = (j_1, \dots, j_k)$ with $1 \leq j_1 < \cdots < j_k \leq n$ for an ordered multi-index and $e^J := e^{j_1} \wedge \cdots \wedge e^{j_k}$. Then

$$\omega = \sum_J \omega_J\, e^J\,.$$

Sometimes one wants to avoid restricting the sum to ordered multi-indices, in particular when using the index calculus and the Einstein summation convention. Then one has

$$\omega = \frac{1}{k!} \sum_{(i)} \omega_{i_1 \cdots i_k}\; e^{i_1} \wedge \cdots \wedge e^{i_k} = \frac{1}{k!}\, \omega_{i_1 \cdots i_k}\; e^{i_1} \wedge \cdots \wedge e^{i_k}$$

where

$$\sum_{(i)} := \sum_{i_1=1}^n \cdots \sum_{i_k=1}^n$$

was abbreviated in the first expression and the Einstein summation convention was used in the second term. Be warned that some authors absorb the normalisation $\frac{1}{k!}$ into the definition of the coefficients $\omega_{i_1 \cdots i_k}$. ◇

5.5 Proposition. **The projection onto Λ_k**
Let $P_k : V_k^0 \to \Lambda_k$ be defined by

$$(P_k\, t)\,(u_1, \ldots, u_k) := \frac{1}{k!} \sum_{\pi \in S_k} \mathrm{sgn}(\pi)\ t(u_{\pi(1)}, \ldots, u_{\pi(k)}) \qquad \text{for all} \quad u_1, \ldots, u_k \in V\,.$$

Then P_k is a linear projection and it holds that

$$v_1^* \wedge \cdots \wedge v_k^* = k!\, P_k(v_1^* \otimes \cdots \otimes v_k^*)\,.$$

Proof. To see that P_k is a projection, just compute

$$\begin{aligned}
(P_k\, P_k t)(u_1, \ldots, u_k) &= \tfrac{1}{(k!)^2} \sum_{\pi, \pi' \in S_k} \mathrm{sgn}(\pi)\mathrm{sgn}(\pi')\ t(u_{\pi'(\pi(1))}, \ldots, u_{\pi'(\pi(k))}) \\
&\overset{\sigma := \pi' \circ \pi}{=} \tfrac{1}{(k!)^2} \sum_{\pi \in S_k} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma)\ t(u_{\sigma(1)}, \ldots, u_{\sigma(k)}) \\
&= \tfrac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma)\ t(u_{\sigma(1)}, \ldots, u_{\sigma(k)}) \ = \ (P_k t)(u_1, \ldots, u_k)\,.
\end{aligned}$$

The second claim follows from (5.2). $\square$

Since the tensor product $\omega \otimes \nu \in V_{k+p}^0$ of two alternating forms $\omega \in \Lambda_k$, $\nu \in \Lambda_p$, is in general not alternating, one introduces the so called exterior product of alternating forms. It is basically defined as the tensor product $\omega \otimes \nu$ followed by skew-symmetrisation, i.e. by projection onto the skew-symmetric subspace Λ_{k+p} of V_{k+p}^0.

5.6 Definition. **The exterior product of exterior forms**
We extend the exterior product (wedge product) to alternating forms by putting

$$\begin{aligned}
\wedge : \Lambda_k \times \Lambda_p &\to \Lambda_{k+p} \\
(\omega_1, \omega_2) &\mapsto \omega_1 \wedge \omega_2 := \frac{(k+p)!}{k!\, p!} P_{k+p}\,(\omega_1 \otimes \omega_2)\,.
\end{aligned}$$

5.7 Example. The wedge product of two 1-forms ω and ν is

$$\omega \wedge \nu = 2P_2(\omega \otimes \nu) = 2\,\tfrac{1}{2}(\omega \otimes \nu - \nu \otimes \omega) = \omega \otimes \nu - \nu \otimes \omega\,.$$

5.8 Proposition. The exterior product of exterior forms has the following properties:

$$\begin{aligned}
(\omega_1 \wedge w_2) \wedge \omega_3 &= \omega_1 \wedge (w_2 \wedge w_3) && \text{for} \quad \omega_1 \in \Lambda_k\,,\ \omega_2 \in \Lambda_p\,,\ \omega_3 \in \Lambda_q \quad \text{(associativity)} \\
(\omega_1 + \omega_2) \wedge \omega_3 &= \omega_1 \wedge \omega_3 + \omega_2 \wedge \omega_3 && \text{for} \quad \omega_1, \omega_2 \in \Lambda_k\,,\ \omega_3 \in \Lambda_q \quad \text{(distributivity)} \\
\omega_1 \wedge (\omega_2 + \omega_3) &= \omega_1 \wedge \omega_2 + \omega_1 \wedge \omega_3 && \text{for} \quad \omega_1 \in \Lambda_k\,,\ \omega_2, \omega_3 \in \Lambda_q \quad \text{(distributivity)} \\
\omega_1 \wedge \omega_2 &= (-1)^{kp}\, \omega_2 \wedge \omega_1 && \text{for} \quad \omega_1 \in \Lambda_k\,,\ \omega_2 \in \Lambda_p\,.
\end{aligned}$$

Proof. The first claim follows from

$$\begin{aligned}
(\omega_1 \wedge w_2) \wedge \omega_3 &= \frac{(k+p+q)!}{(k+p)!q!} \frac{(k+p)!}{k!\, p!} P_{k+p+q}(P_{k+p}(\omega_1 \otimes \omega_2) \otimes \omega_3) \\
&= \frac{(k+p+q)!}{k!\, p!\, q!} P_{k+p+q}(\omega_1 \otimes \omega_2 \otimes \omega_3)
\end{aligned}$$

and a similar computation for $\omega_1 \wedge (w_2 \wedge w_3)$. Note that the last step in the previous computation follows from a straightforward but tedious calculation as in the proof of proposition 5.5. Distributivity of the wedge product follows directly from distributivity of the tensor product. The final claim is a homework assignment. $\square$

5.9 Definition. **The inner product of exterior forms**

The **inner product** of exterior forms with respect to a non-degenerate $g \in V_2^0$ is the bilinear map

$$\begin{aligned} i : \Lambda_\ell \times \Lambda_k &\to \Lambda_{k-\ell}\,, \quad \ell \leq k \\ (\nu, \omega) &\mapsto i_\nu \omega \end{aligned}$$

that is inductively defined by the following rules

(i) $i_\nu \omega := G(\nu, \omega)$ for $\omega, \nu \in \Lambda_1$ (cf. defintion 4.12)

(ii) $i_\nu(\omega_1 \wedge \omega_2) := (i_\nu \omega_1) \wedge \omega_2 + (-1)^{k_1} \omega_1 \wedge i_\nu \omega_2$ for $\nu \in \Lambda_1$ and $\omega_i \in \Lambda_{k_i}$

(iii) $i_{\nu_1 \wedge \nu_2} = i_{\nu_2} \circ i_{\nu_1}$

together with bilinearity.

In components one finds for $\omega = \frac{1}{k!}\, \omega_{j_1 \cdots j_k} e^{j_1 \cdots j_k}$ and $\nu = \frac{1}{\ell!} \nu_{i_1 \cdots i_\ell} e^{i_1 \cdots i_\ell}$ that

$$i_\nu \omega = \frac{1}{\ell!(k-\ell)!} \nu^{j_1 \cdots j_\ell} \omega_{j_1 \cdots j_\ell j_{\ell+1} \cdots j_k} e^{j_{\ell+1}} \wedge \cdots \wedge e^{j_k}\,.$$

In particular, for $k = \ell$ we have $i_\nu \omega = \frac{1}{k!} G(\nu, \omega)$. ◇

5.10 Definition. **The canonical volume form**

Let V be a n-dimensional real vector space, $(e_1, \ldots, e_n)$ a basis of V, $g \in V_2^0$ non-degenerate, and $g = g_{ij}\, e^i \otimes e^j$. The **canonical volume form** associated with $g \in V_2^0$ is defined as

$$\varepsilon := \sqrt{|\det(g_{ij})|}\, e^1 \wedge e^2 \wedge \cdots \wedge e^n\,.$$

You will show as a homework assignment that ε depends only on the orientation of the chosen basis. By definition, two bases $(e_1, \ldots, e_n)$ and $(\tilde{e}_1, \ldots, \tilde{e}_n)$ have the same orientation, if the matrix associated with the basis change has positive determinant, i.e. if $\det(e^i(\tilde{e}_j)) > 0$. If $(e_1, \ldots, e_n)$ and $(\tilde{e}_1, \ldots, \tilde{e}_n)$ have the same orientation, then $\varepsilon = \tilde{\varepsilon}$, otherwise $\varepsilon = -\tilde{\varepsilon}$. ◇

Since Λ_k and Λ_{n-k} both have dimension $\binom{n}{k}$, we can identify the two spaces using a non-degenerate $g \in V_2^0$ and an orientation of V, i.e. a choice of the sign of its associated volume form ε.

5.11 Definition. **The Hodge isomorphism**

Let V be a n-dimensional oriented real vector space, $g \in V_2^0$ non-degenerate, and ε the corresponding volume form. For $0 \leq k \leq n$ the linear bijection

$$\begin{aligned} * : \Lambda_k &\to \Lambda_{n-k}\,, \\ \omega &\mapsto *\omega := i_\omega \varepsilon \end{aligned}$$

is called **Hodge isomorphism**, **Hodge duality**, or **Hodge star operator**.

5.12 Remark. **Properties of $*$ and ε**

Let $g \in V_2^0$ be non-degenerate and $g = g_{ij} e^i \otimes e^j$ its representation with respect to a basis (e^j) of V^*.

(i) The coefficients of ε are

$$\varepsilon_{j_1 \cdots j_n} = \begin{cases} 0 & \text{if} \quad j_l = j_k \quad \text{for} \quad l \neq k \\ \operatorname{sgn}(\pi) \cdot \sqrt{|\det g|} & \text{if} \quad (j_1, \ldots, j_n) = \pi(1, \ldots, n) \quad \text{for some } \pi \in S_n \end{cases}$$

where $\det g := \det(g_{ij})$. (Note that $\det g$ depends on the chosen basis, since g_{ij} does not transform like the matrix of a linear map!)

(ii) For $1 \in \Lambda_0$ it holds that

$$*1 = i_1\varepsilon := \varepsilon\,,$$

and vice-versa that

$$\begin{aligned} *\varepsilon = i_\varepsilon\varepsilon &= \tfrac{1}{n!}\,\varepsilon_{j_1\cdots j_n}\,\varepsilon_{i_1\cdots i_n}\,g^{j_1 i_1}\cdots g^{j_n i_n} \\ &= |\det g|\,\tfrac{1}{n!}\sum_{\pi,\pi'\in S_n}\operatorname{sgn}(\pi)\operatorname{sgn}(\pi')g^{\pi(1)\pi'(1)}\cdots g^{\pi(n)\pi'(n)} \\ &= |\det g|\det g^{-1} =: (-1)^s\,, \end{aligned}$$

where $\det g^{-1} := \det(g^{ij})$ and thus $s = 0$ if g is a metric.

(iii) If g is symmetric, then $* \circ *|_{\Lambda_k} = (-1)^{k(n-k)+s}\,\mathrm{id}_{\Lambda_k}$. (Homework assignment)

(iv) $i_\nu * \omega = i_\nu i_\omega\varepsilon = i_{\omega\wedge\nu}\varepsilon = *(\omega\wedge\nu)$

5.13 Definition. **Differential forms on manifolds**

On a smooth manifold M we denote the space of tensor fields $\omega \in \mathcal{T}^0_k(M)$ that take values in the alternating k-forms on T_xM by $\Lambda_k(M)$ and call $\omega \in \Lambda_k(M)$ a **differential k-form** or just **k-form**. In particular, $\Lambda_0(M) = C^\infty(M)$ and $\Lambda_1(M) = \mathcal{T}^0_1(M)$.

Note also that the fibre-wise wedge product defines a bilinear map

$$\wedge : \Lambda_k(M)\times\Lambda_p(M)\to\Lambda_{k+p}(M)\,,\quad (\omega,\nu)\mapsto\omega\wedge\nu \quad \text{with} \quad (\omega\wedge\nu)(x) = \omega(x)\wedge\nu(x)\,.$$

5.14 Proposition. **Pull-backs of exterior products**

Let $\omega\in\Lambda_k(M)$, $\nu\in\Lambda_p(M)$, and $f : N\to M$ smooth. Then

$$f^*(\omega\wedge\nu) = f^*\omega\wedge f^*\nu\,,$$

and

$$f^*\left(\sum_J \omega_J\,\mathrm{d}q^{j_1}\wedge\cdots\wedge\mathrm{d}q^{j_k}\right) = \sum_J(\omega_J\circ f)\,\mathrm{d}(q^{j_1}\circ f)\wedge\cdots\wedge\mathrm{d}(q^{j_k}\circ f)\,. \tag{5.3}$$

Proof. Both claims follow from the definition 4.20 and proposition 3.14 about the pull-back of a differential. □

We saw that the exterior derivative of a function $f\in\Lambda_0$ is a 1-form $\mathrm{d}f\in\Lambda_1$. We will now generalise the exterior derivative to a map $\mathrm{d} : \Lambda_k(M)\to\Lambda_{k+1}(M)$.

5.15 Definition. **The exterior derivative in a chart**

Let $\omega\in\Lambda_p(U)$ for an open subset $U\subset\mathbb{R}^n$ and

$$\omega = \sum_I \omega_I\,\mathrm{d}e^{i_1}\wedge\cdots\wedge\mathrm{d}e^{i_p} \quad \text{with} \quad \omega_I\in C^\infty(U)\,.$$

Then its **exterior derivative** $\mathrm{d}\omega\in\Lambda_{p+1}(U)$ is defined by

$$\mathrm{d}\omega := \sum_{(i)}\mathrm{d}\omega_I\wedge\mathrm{d}e^{i_1}\wedge\cdots\wedge\mathrm{d}e^{i_p}\,.$$

For a smooth manifold M, $\omega\in\Lambda_p(M)$, and (V,φ) a chart, the **exterior derivative** $\mathrm{d}\omega\in\Lambda_{p+1}(M)$ is defined locally by $\mathrm{d}\omega|_V := \varphi^*\mathrm{d}(\varphi_*\omega)$, i.e. for

$$\omega|_V = \sum_I\omega_I\,\mathrm{d}q^{i_1}\wedge\cdots\wedge\mathrm{d}q^{i_p} \quad \text{with} \quad \omega_I\in C^\infty(M)$$

we define

$$\mathrm{d}\omega|_V := \sum_I\mathrm{d}\omega_I\wedge\mathrm{d}q^{i_1}\wedge\cdots\wedge\mathrm{d}q^{i_p}\,. \qquad \diamond$$

5.16 Proposition. The exterior derivative: global definition
For two charts (V_1, φ_1) and (V_2, φ_2) on M and $\omega \in \Lambda_p(M)$ it holds that

$$\varphi_1^*(\mathrm{d}(\varphi_{1*}\omega)|_{\varphi_1(V_1\cap V_2)}) = \varphi_2^*(\mathrm{d}(\varphi_{2*}\omega)|_{\varphi_2(V_1\cap V_2)})$$

and thus the exterior derivative $\mathrm{d}\omega \in \Lambda_{p+1}(M)$ is uniquely defined by the local expressions. More generally, it holds for any diffeomorphism $\Phi : M \to N$, $\omega \in \Lambda_p(N)$, and chart (V, φ) on M that

$$\Phi^*(\mathrm{d}\omega|_{\Phi(V)}) = \mathrm{d}\Phi^*\omega|_V$$

and thus $\Phi^*\mathrm{d}\omega = \mathrm{d}\Phi^*\omega$.

Proof. Let (V, φ) be a chart on M and $(\tilde V, \tilde\varphi) = (\Phi(V), \varphi \circ \Phi^{-1})$ its push-forward to $\Phi(V) \subset N$. For $\omega = \sum_I \omega_I \, \mathrm{d}\tilde q^I \in \Lambda_p(N)$ we compute that

$$\begin{aligned}
\mathrm{d}\Phi^*\omega|_V &\overset{(5.3)}{=} \mathrm{d}\left(\sum_I (\omega_I \circ \Phi)\, \Phi^*(\mathrm{d}\tilde q^{i_1}) \wedge \cdots \wedge \Phi^*(\mathrm{d}\tilde q^{i_p})\right)\\
&\overset{3.14}{=} \mathrm{d}\left(\sum_I (\omega_I \circ \Phi)\, \mathrm{d}q^{i_1} \wedge \cdots \wedge \mathrm{d}q^{i_p}\right)\\
&\overset{\text{Def.}}{=} \sum_I \mathrm{d}(\omega_I \circ \Phi) \wedge \mathrm{d}q^{i_1} \wedge \cdots \wedge \mathrm{d}q^{i_p}\\
&\overset{3.14}{=} \sum_I \Phi^*(\mathrm{d}\omega_I) \wedge \Phi^*(\mathrm{d}\tilde q^{i_1}) \wedge \cdots \wedge \Phi^*(\mathrm{d}\tilde q^{i_p}) \overset{(5.3)}{=} \Phi^*(\mathrm{d}\omega|_{\Phi(V)})\,.
\end{aligned}$$

Applying this general result to the diffeomorphism $\Phi : \varphi_2(U) \to \varphi_1(U)$, $\Phi = \varphi_1 \circ \varphi_2^{-1}$ on $U := V_1 \cap V_2$ yields

$$\begin{aligned}
\varphi_1^*(\mathrm{d}(\varphi_{1*}\omega)|_{\varphi_1(U)}) &= \varphi_2^*(\varphi_2^{-1})^*\varphi_1^*(\mathrm{d}(\varphi_{1*}\omega)|_{\varphi_1(U)}) = \varphi_2^*\Phi^*(\mathrm{d}(\varphi_{1*}\omega)|_{\varphi_1(U)})\\
&= \varphi_2^*(\mathrm{d}(\Phi^*\varphi_{1*}\omega)|_{\varphi_2(U)}) = \varphi_2^*(\mathrm{d}(\varphi_{2*}\omega)|_{\varphi_2(U)})\,,
\end{aligned}$$

where we used that because of the chain rule we have $\Phi^* = (\varphi_2^{-1})^*\varphi_1^* = \varphi_{2*}(\varphi_1^{-1})_*$. □

5.17 Proposition. Properties of the exterior derivative
From the definition we can conclude the following properties of the exterior derivative:

(a) $\mathrm{d}(\omega_1 + \omega_2) = \mathrm{d}\omega_1 + \mathrm{d}\omega_2$ for all $\omega_i \in \Lambda_p(M)$
(b) $\mathrm{d}(f\omega) = \mathrm{d}f \wedge \omega + f\mathrm{d}\omega$ for all $f \in C^\infty$ and $\omega \in \Lambda_p$
(c) $\mathrm{d}(\omega_1 \wedge \omega_2) = \mathrm{d}\omega_1 \wedge \omega_2 + (-1)^p\, \omega_1 \wedge \mathrm{d}\omega_2$ for all $\omega_1 \in \Lambda_p$, $\omega_2 \in \Lambda_k$
(d) $\mathrm{d}(\mathrm{d}\omega) = 0$ for all $\omega \in \Lambda_p$

Proof. (a) is obvious and (b) follows directly from the definition, and (c) follows from the observation that one has to commute the exterior derivatives of the coefficients of ω_2 through the p-form ω_1 when comparing the right side with the left side of the claimed equality. Finally, (d) results from the commutativity of partial derivatives,

$$\mathrm{d}(\mathrm{d}\omega)|_V = \sum_I \sum_{j,k} \frac{\partial^2\omega_I}{\partial q_k \partial q_j}\, \mathrm{d}q^k \wedge \mathrm{d}q^j \wedge \mathrm{d}q^I = \sum_I \sum_{j<k} \left(\frac{\partial^2\omega_I}{\partial q_k \partial q_j} - \frac{\partial^2\omega_I}{\partial q_j \partial q_k}\right) \mathrm{d}q^k \wedge \mathrm{d}q^j \wedge \mathrm{d}q^I = 0\,.$$

□

5.18 Proposition. Naturality of the exterior derivative

Let $f : M \to N$ be smooth and $\omega \in \Lambda_p(N)$, then

$$f^*\mathrm{d}\omega = \mathrm{d}(f^*\omega)\,.$$

Proof. Homework assignment (Look at the computation in the proof of proposition 5.16 and apply what we learned in proposition 5.17). □

5.19 Remark. Restriction of differential forms

Let $N \subset M$ be a submanifold and $\psi : N \to M$ the natural injection. For $\omega \in \Lambda_p(M)$

$$\psi^*\omega \in \Lambda_p(N)$$

is called its restriction to N. According to proposition 5.18, restriction and exterior derivative commute,

$$\psi^*\mathrm{d}\omega = \mathrm{d}(\psi^*\omega)\,.$$

5.20 Example. Vector differential operators in $\mathbb{R}^3$ and the exterior derivative

Let $M = \mathbb{R}^3$ with the Euclidean metric $g_{ij} = \delta_{ij}$. We can identify vector fields and 1-forms using the canonical isomporphism ι_g (cf. remark 4.8) and for the components with respect to cartesian coordinates we have $v_i = v^i$. Let $f \in C^\infty(\mathbb{R}^3)$ and $v \in \mathcal{T}^1_0(\mathbb{R}^3)$. Then we have the following relations between vector differential operators in $\mathbb{R}^3$ and the exterior derivative:

$$\begin{aligned}
\mathrm{d}f &= \frac{\partial f}{\partial q_i}\mathrm{d}q^i = (\mathrm{grad} f)_i\,\mathrm{d}q^i = \iota_g(\mathrm{grad}\, f)\\
*\mathrm{d}\,\iota_g(v) &= *(\mathrm{d}(v_i \mathrm{d}q^i)) \;=\; *(\mathrm{d}v^i \wedge \mathrm{d}q^i) \;=\; *(\frac{\partial v_i}{\partial q_k}\,\mathrm{d}q^k \wedge \mathrm{d}q^i) \;=\; \varepsilon_{jki}\frac{\partial v_i}{\partial q_k}\mathrm{d}q^j \;=\; (\mathrm{curl}\, v)_j\,\mathrm{d}q^j\\
&= \iota_g(\mathrm{curl}\, v)\\
\mathrm{d}\,\iota_g(v) &= *\mathrm{d}(v^i \tfrac{1}{2}\varepsilon_{ijk}\,\mathrm{d}q^j \wedge \mathrm{d}q^k) \;=\; *(\tfrac{1}{2}\varepsilon_{ijk}\,\mathrm{d}v^i \wedge \mathrm{d}q^j \wedge \mathrm{d}q^k) \;=\; *(\tfrac{1}{2}\varepsilon_{ijk}\frac{\partial v^i}{\partial q_l}\mathrm{d}q^l \wedge \mathrm{d}q^j \wedge \mathrm{d}q^k)\\
&= \tfrac{1}{2}\varepsilon_{ijk}\,\frac{\partial v^i}{\partial q_l}\,\varepsilon_{ljk} \;=\; \sum_i \frac{\partial v_i}{\partial q_i} \;=\; \mathrm{div}\, v\,.
\end{aligned}$$

The properties of the exterior derivative from proposition 5.17 can now be translated to the differential operators. From (b) we conclude:

- For $p = k = 0$ it follows from $\mathrm{d}(f \cdot g) = f\mathrm{d}g + g\mathrm{d}f$ that $\mathrm{grad}\,(f \cdot g) = f\,\mathrm{grad}\, g + g\,\mathrm{grad}\, f$
- For $p = 0, k = 1$ it follows from $\mathrm{d}(f\omega) = \mathrm{d}f \wedge \omega + f\mathrm{d}\omega$ that $\mathrm{curl}\,(fv) = \mathrm{grad}\, f \times v + f\,\mathrm{curl}\, v$

From (c) we conclude:

- For $p = 0$ it follows from $\mathrm{d}(\mathrm{d}f) = 0$ that $\mathrm{curl}\,\mathrm{grad}\, f = 0$
- For $p = 1$ it follows from $0 = *\mathrm{d}\mathrm{d}\omega = *\mathrm{d}**\mathrm{d}\omega$ that $\mathrm{div}\,\mathrm{curl}\, v = 0$.

5.21 Definition. Closed and exact forms

A p-form ω is called **closed** if $\mathrm{d}\omega = 0$. It is called **exact** if $\omega = \mathrm{d}\nu$ for some $\nu \in \Lambda_{p-1}(M)$, i.e. if it has a primitive.

Because of proposition 5.17 (c) we have

$$\omega \text{ is exact} \quad\Rightarrow\quad \omega \text{ is closed.}$$

The converse holds, in general, only locally. ◇

We state two versions of the so called Poincaré lemma, the classical one and a slightly more general statement.

5.22 Theorem. Poincaré lemma (version 1)
Let $\omega \in \Lambda_p(M)$ be closed, i.e. $\mathrm{d}\omega = 0$. Let $V \subset M$ be open and diffeomorphic to a star-shaped domain of $\mathbb{R}^n$. Then there exists $\nu \in \Lambda_{p-1}(V)$ such that $\omega|_V = \mathrm{d}\nu$.

Proof. The following proof reduces the problem to the Poincaré lemma in $\mathbb{R}^n$. Let $\varphi : V \to U \subset \mathbb{R}^n$ be a diffeomorphism to the star-shaped domain $U \subset \mathbb{R}^n$. Then $\tilde{\omega} := \varphi_*\omega$ is a closed p-form on U and according to the Poincaré lemma on $\mathbb{R}^n$ (which we will show below) there exists a primitive $\tilde{\nu} \in \Lambda_{p-1}(U)$ with $\tilde{\omega} = \mathrm{d}\tilde{\nu}$. But then $\nu := \varphi^*\tilde{\nu}$ is a primitive for ω since $\mathrm{d}\varphi^*\tilde{\nu} = \varphi^*\mathrm{d}\tilde{\nu} = \varphi^*\tilde{\omega} = \omega$.
We now show the Poincaré lemma on $\mathbb{R}^n$. Assume without loss of generality that $U \subset \mathbb{R}^n$ is star-shaped with respect to the origin. We define the map $P^p : \Lambda_p(U) \to \Lambda_{p-1}(U)$ by setting for arbitrary $\omega \in \Lambda_p(U)$ with $\omega = \sum_I \omega_I \mathrm{d}q^I$

$$P^p\omega := \sum_I \sum_{\alpha=1}^{p} (-1)^{\alpha-1} \left(\int_0^1 t^{p-1} \omega_I(tq)\, q_{i_\alpha}\, \mathrm{d}t \right) \mathrm{d}q^{I^\alpha}\,.$$

Recall the notation $I = (i_1, \ldots, i_p)$ for an ordered p-Tupel $1 \le i_1 < \cdots < i_p \le n$ and $\mathrm{d}q^I = \mathrm{d}q^{i_1} \wedge \cdots \wedge \mathrm{d}q^{i_p} \in \Lambda_p(U)$. Moreover, we also abbreviated $\mathrm{d}q^{I^\alpha} = \mathrm{d}q^{i_1} \wedge \cdots \wedge \widehat{\mathrm{d}q^{i_\alpha}} \wedge \cdots \wedge \mathrm{d}q^{i_p} \in \Lambda_{p-1}(U)$, where the hat means that the corresponding factor is omitted.
We will show that $\omega = \mathrm{d}P^p\omega + P^{p+1}\mathrm{d}\omega$. For closed ω this implies that $\omega = \mathrm{d}(P^p\omega)$, i.e. the existence of a primitive. We first compute

$$\begin{aligned}
\mathrm{d}P^p\omega &= \sum_{\ell=1}^{n} \sum_I \sum_{\alpha=1}^{p} (-1)^{\alpha-1} \left(\int_0^1 t^{p-1} \frac{\partial(\omega_I(tq)\, q_{i_\alpha})}{\partial q_\ell} \mathrm{d}t \right) \mathrm{d}q^\ell \wedge \mathrm{d}q^{I^\alpha} \\
&= \sum_I \sum_{\ell \notin I} \sum_{\alpha=1}^{p} (-1)^{\alpha-1} \left(\int_0^1 t^p \frac{\partial \omega_I}{\partial q_\ell}(tq)\, q_{i_\alpha}\, \mathrm{d}t \right) \mathrm{d}q^\ell \wedge \mathrm{d}q^{I^\alpha} \\
&\quad + \sum_I \sum_{\alpha=1}^{p} \left(\int_0^1 t^{p-1} \frac{\partial(\omega_I(tq)\, q_{i_\alpha})}{\partial q_{i_\alpha}} \mathrm{d}t \right) \mathrm{d}q^I\,.
\end{aligned}$$

On the other hand

$$\mathrm{d}\omega = \sum_I \sum_{\ell \notin I} \frac{\partial \omega_I}{\partial q_\ell} \mathrm{d}q^\ell \wedge \mathrm{d}q^I$$

and thus

$$P^{p+1}\mathrm{d}\omega = \sum_I \sum_{\ell \notin I} \left\{ \left(\int_0^1 t^p \frac{\partial \omega_I}{\partial q_\ell}(tq) q_\ell \mathrm{d}t \right) \mathrm{d}q^I + \sum_{\alpha=1}^{p} (-1)^\alpha \left(\int_0^1 t^p \frac{\partial \omega_I}{\partial q_\ell}(tq) q_{i_\alpha} \mathrm{d}t \right) \mathrm{d}q^\ell \wedge \mathrm{d}q^{I^\alpha} \right\}.$$

For the sum of the two terms we finally get

$$\begin{aligned}
\mathrm{d}P^p\omega + P^{p+1}\mathrm{d}\omega &= \sum_I \int_0^1 \left(\sum_{\ell=1}^{n} t^p \frac{\partial \omega_I}{\partial q_\ell}(tq) q_\ell + p t^{p-1} \omega_I(tq) \right) \mathrm{d}t\ \mathrm{d}q^I \\
&= \sum_I \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} \Big(t^p \omega_I(tq) \Big) \mathrm{d}t\ \mathrm{d}q^I = \sum_I \omega_I \mathrm{d}q^I = \omega\,.
\end{aligned}$$

□

5.23 Theorem. Poincaré lemma (version 2)
Let $\omega \in \Lambda_p(M)$ be closed and let $V \subseteq M$ be a submanifold that is contractible in M, i.e. there exists a smooth function $F : [0,1] \times V \to M$ such that $F(0,\cdot) = \mathrm{id}_V$ and $F(1,\cdot) \equiv x_0$ for some $x_0 \in M$. Then there exists $\nu \in \Lambda_{p-1}(V)$ such that $\omega|_V = \mathrm{d}\nu$.

Proof. The proof uses the so called homotopy operator (cf. P^p in the previous proof) and will be developed in the homework assignments. $\square$

German nomenclature

alternating form = alternierende Form	canonical volume form = kanonische Volumenform
closed form = abgeschlossene Differentialform	determinant = Determinante
differential form = Differentialform	exact form = exakte Differentialform
exterior derivative = äußere Ableitung	exterior form = äußere Form
exterior product = äußeres Produkt	skew-symmetric = schiefsymmetrisch
wedge product = Dachprodukt, Hutprodukt	

6 Integration on manifolds

In this chapter we discuss integration of n-forms over n-dimensional manifolds. Using pull-backs, this will, in particular, also result in a notion of integration of p-forms over p-dimensional submanifolds.

The sign of the integral of a differential n-form is only fixed after choosing an orientation of the manifold.

6.1 Definition. <u>**Orientation**</u>

In an **oriented atlas** $(V_i, \varphi_i)_{i\in I}$ all charts have the same orientation, i.e. for all transition functions $\Phi_{ij} := \varphi_i \circ \varphi_j^{-1}$ it holds that $\det D\Phi_{ij} > 0$.

An **oriented manifold** is a manifold with an oriented atlas. A manifold that allows for an oriented atlas is called **orientable**.

6.2 Remarks. (a) Not every manifold is orientable. For example the Möbius strip is not orientable.

(b) If a manifold is orientable, there are exactly two different orientations, i.e. equivalence classes of atlases with the same orientation. Given an orientation on a manifold, that is an oriented atlas, any chart from the same equivalence class of atlases is called positively oriented, all others are called negatively oriented.

(c) A nowhere vanishing volume form ω defines an orientation: If in the coordinate representation $\omega = \omega(q)\,\mathrm{d}q^1 \wedge \cdots \wedge \mathrm{d}q^n$ with respect to a chart $\varphi(x) = q$ it holds that $\omega(q) > 0$, then we say that the chart φ is positively oriented with respect to ω, otherwise it is negatively oriented.

To see that this really defines an orientation, let φ and $\tilde\varphi$ be two charts and let $\Phi = \tilde\varphi \circ \varphi^{-1}$. Because of $\mathrm{d}\tilde q^j = (D\Phi)^j_i \mathrm{d}q^i$ we have with $\tilde q = \Phi(q)$ that

$$\begin{aligned}\omega = \tilde\omega(\tilde q)\,\mathrm{d}\tilde q^1 \wedge \cdots \wedge \mathrm{d}\tilde q^n &= (\tilde\omega \circ \Phi)(q)\, D\Phi^1_{j_1}|_q \cdots D\Phi^n_{j_n}|_q\, \mathrm{d}q^{j_1} \wedge \cdots \wedge \mathrm{d}q^{j_n} \\ &= (\tilde\omega \circ \Phi)(q) \det(D\Phi|_q)\,\mathrm{d}q^1 \wedge \cdots \wedge \mathrm{d}q^n \\ &= \omega(q)\,\mathrm{d}q^1 \wedge \cdots \wedge \mathrm{d}q^n\end{aligned}$$

and thus

$$\omega(q) = (\tilde\omega \circ \Phi)(q) \det(D\Phi|_q)\,. \tag{6.1}$$

Hence $\omega(q)$ and $\tilde\omega(\tilde q)$ have the same sign if and only if $\det(D\Phi|_q) > 0$. ◇

6.3 Definition. <u>**Compact support**</u>

The **support** of a tensor field $t \in \mathcal{T}^r_s(M)$ is the set

$$\operatorname{supp} t := \overline{\{x \in M \mid t(x) \neq 0\}} \subset M\,.$$

We say that $t \in \mathcal{T}^r_s(M)$ is **compactly supported** if $\operatorname{supp} t$ is a compact set. ◇

6.4 Definition. <u>**The integral on the domain of a single chart**</u>

Let (V, φ) be a chart from an oriented atlas of M and let $\omega \in \Lambda_n(M)$ have compact support in V. Then one defines

$$\int_M \omega \;=\; \int_V \omega \;:=\; \int_{\varphi(V)} \varphi_*\omega \;:=\; \int_{\mathbb{R}^n_+} \omega(q)\,\mathrm{d}^n q\,,$$

where

$$\varphi_*\omega \;=:\; \omega(q)\, e^1 \wedge \cdots \wedge e^n \in \Lambda_n(\mathbb{R}^n_+)$$

and $\mathrm{d}^n q$ denotes the n-dimensional Lebesgue measure on $\mathbb{R}^n$. ◇

6.5 Proposition. Up to orientation, $\int_M \omega$ is independent of the chosen chart.

Proof. Let φ and $\tilde\varphi$ be two charts on V with the same orientation and let $\Phi = \tilde\varphi \circ \varphi^{-1}$. The change of variables formula for integrals on $\mathbb{R}^n$ implies

$$\int_{\varphi(V)} \varphi_*\omega \;:=\; \int \omega(q)\,\mathrm{d}^n q \;\overset{(6.1)}{=}\; \int (\tilde\omega \circ \Phi)(q)\det(D\Phi|_q)\,\mathrm{d}^n q \;\overset{\text{c.o.v.}}{=}\; \int \tilde\omega(\tilde q)\,\mathrm{d}^n \tilde q \;=\; \int_{\tilde\varphi(V)} \tilde\varphi_*\omega\,. \;\square$$

In order to integrate also forms that are not supported in the domain of a single chart, one decomposes the integration with the help of a partition of unity.

6.6 Definition. **Partition of unity**

Let M be a smooth manifold and $\mathcal{A} = (V_i, \varphi_i)_{i\in I}$ an atlas. A family of smooth functions $(\chi_i)_{i\in I}$, $\chi_i : M \to [0,1]$, is called an **partition of unity adapted to** $\mathcal{A}$ if the following properties hold.

(a) Each χ_i is supported in the single chart V_i, $\mathrm{supp}\chi_i \subset V_i$.

(b) The partition is locally finite, i.e. each $x \in M$ has a neighbourhood U such that $U \cap \mathrm{supp}\chi_i \neq \emptyset$ only for finitely many $i \in I$.

(c) The functions χ_i sum up to one everywhere: $\sum_{i\in I} \chi_i(x) = 1$ for all $x \in M$.

Note that because of (b) the sum in (c) contains for any $x \in M$ only finitely many terms. ◇

6.7 Remark. **Existence of a partition of unity**

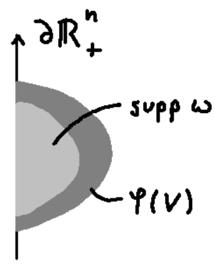
If one assumes that a manifold is second countable, as we did in the definition of a manifold, then one can show that for any atlas $(V_i, \varphi_i)_{i\in I}$ an adapted partition of unity exists.

Moreover, if $K \subset M$ is compact, there exists a finite subcover $(V_{i_j})_{j=1,\ldots,m}$ and an adapted partition of unity $(\chi_j)_{j=1,\ldots,m}$ such that $\sum_j \chi_j(x) = 1$ for all $x \in K$. ◇

In order to avoid additional technicalities, we will only discuss integration over compact regions. More precisely, we will integrate volume forms $\omega \in \Lambda_n(M)$ with compact support.

6.8 Remark. **Behaviour at the boundary**

The fact that the closed set $\mathrm{supp}\,\omega$ is contained in the open set V does not imply that ω vanishes on the boundary ∂M of M. For example, $V = [0, \frac{1}{2})$ is a relatively open subset of $M = [0, \infty)$ and $K = [0, \frac{1}{4}]$ a compact subset of V.



6.9 Definition. **The integral of volume forms**

Let M be a smooth oriented manifold and (V_i, φ_i) a positively oriented atlas. Let $\omega \in \Lambda_n(M)$ have compact support. Then

$$\int_M \omega \;:=\; \sum_{j=1}^m \int_{V_j} \chi_j \omega\,,$$

where (χ_j) is a partition of unity adapted to the finite cover of $\mathrm{supp}\,\omega$ by charts (V_j) such that $\sum_j \chi_j(x) = 1$ for $x \in \mathrm{supp}\,\omega$. The summands on the right hand side are integrals as defined in definition 6.4. ◇

6.10 Proposition. The value of $\int_M \omega$ does neither depend on the choice of the atlas nor on the choice of the partition of unity.

Proof. The independence from the choice of the charts has already been demonstrated in proposition 6.5. Let $(\tilde{\chi}_j)$ be another partition of unity adapted to (V_j) with $\sum_j \tilde{\chi}_j(x) = 1$ for $x \in \operatorname{supp}\omega$. Then

$$\begin{aligned}\sum_{j=1}^{m} \int_{\varphi_j(V_j)} \varphi_{j*}(\chi_j \omega) &= \sum_{j=1}^{m} \int_{\varphi_j(V_j)} \varphi_{j*}(\chi_j \sum_{i=1}^{m} \tilde{\chi}_i \omega) \\ &= \sum_{j,i=1}^{m} \int_{\varphi_j(V_j \cap V_i)} \varphi_{j*}(\chi_j \tilde{\chi}_i \omega) = \sum_{j,i=1}^{m} \int_{\varphi_i(V_j \cap V_i)} \varphi_{i*}(\chi_j \tilde{\chi}_i \omega) \\ &= \sum_{i=1}^{m} \int_{\varphi_i(V_i)} \varphi_{i*}(\sum_{j=1}^{m} \chi_j \tilde{\chi}_i \omega) = \sum_{i=1}^{m} \int_{\varphi_i(V_i)} \varphi_{i*}(\tilde{\chi}_i \omega)\,,\end{aligned}$$

where we used proposition 6.5 in the third equality. □

6.11 Example. Let $M = [a, b] \subset \mathbb{R}$ and $f \in C_0^\infty(M)$ (this does **not** imply that $f(a) = f(b) = 0$ since M is compact). Then $\mathrm{d}f \in \Lambda_1(M)$ and $\operatorname{supp} \mathrm{d}f \subset \operatorname{supp} f$ is compact. In the canonical chart $\varphi = \mathrm{id}_{\mathbb{R}}$ one has

$$\int_M \mathrm{d}f = \int_a^b \frac{\partial f}{\partial x}\,\mathrm{d}x = f(b) - f(a) =: \int_{\partial M} f\,.$$

6.12 Definition. <u>**The integral on submanifolds**</u>

Let M be an n-dimensional smooth manifold, N a p-dimensional smooth manifold, and $\psi : N \to M$ smooth. (If $N \subset M$ is a submanifold, then $\psi : N \to M$ is just the inclusion map.) Let $\omega \in \Lambda_p(M)$ have compact support. Then one defines

$$\int_N \omega := \int_N \psi^* \omega\,.$$

6.13 Definition. <u>**Orientation of the boundary**</u>

If M is an oriented manifold, then also its boundary ∂M is orientable and an orientation on M induces an orientation on ∂M: Let (V_i, φ_i) be a positively oriented atlas on M, then we define $(V_i|_{\partial M}, \varphi_i|_{\partial M})$ to be negatively oriented. Put differently: if $\mathrm{d}q^1 \wedge \cdots \wedge \mathrm{d}q^n$ is a positive volume form on M, then $-\mathrm{d}q^2 \wedge \cdots \wedge \mathrm{d}q^n$ is a positive volume form on ∂M. ⋄

We now formulate and prove the generalisation of the fundamental theorem of calculus to manifolds with boundary.

6.14 Theorem. <u>**Stokes theorem**</u>

Let M be an oriented manifold with boundary ∂M and let $\omega \in \Lambda_{n-1}(M)$ be compactly supported. Then

$$\int_M \mathrm{d}\omega = \int_{\partial M} \omega\,,$$

where ∂M inherits the orientation of M as in definition 6.13. ⋄

For integration on submanifolds this implies the following result.

6.15 Corollary. **Stokes theorem for submanifolds**
Let M be a smooth manifold and $N \subset M$ an oriented submanifold of dimension p. Let $\omega \in \Lambda_{p-1}(M)$ be compactly supported. Then

$$\int_N \mathrm{d}\omega = \int_{\partial N} \omega\,,$$

where again ∂N inherits the orientation of N.

The result holds as well if one replaces the submanifold $N \subset M$ by the image of p-dimensional manifold N under a smooth map $\psi : N \to M$. ◇

6.16 Remarks. (a) An important special case of Stokes theorem is the observation that the integral of a compactly supported exact form over a manifold without boundary vanishes.

(b) On the on hand, the requirement in Stokes theorem that the integrand is compactly supported serves to avoid technical discussions of convergence of the integral. However, it is also important at the boundary: Consider, for example, $M = (a, b)$ with $\partial M = \emptyset$ and the function $f(x) = x$. Then

$$\int_a^b \mathrm{d}f = b - a \neq \int_{\partial M} f = 0\,,$$

which does not contradict Stokes theorem since f is not compactly supported. On $M = [a, b]$ we have $\partial M = \{a, b\}$ and $f(x) = x$ is compactly supported.

(c) The fact that $\mathrm{d}\mathrm{d}\omega = 0$ for every $\omega \in \Lambda_p(M)$ corresponds to the fact that a boundary has no boundary, i.e. that $\partial\partial N = \emptyset$ for any manifold N:

$$0 = \int_N \mathrm{d}\mathrm{d}\omega = \int_{\partial N} \mathrm{d}\omega = \int_{\partial\partial N} \omega = 0\,.$$

Proof. **of Stokes theorem.**
Let $(V_i, \varphi_i)_{i=1,\dots,m}$ be a finite cover of $\operatorname{supp}\omega$ by positively oriented charts and (χ_i) an adapted partition of unity with $\sum_i \chi_i(x) = 1$ for all $x \in \operatorname{supp}\omega$. Abbreviating $\omega = \sum_i \chi_i\omega =: \sum_i \omega_i$ we have

$$\int_M \mathrm{d}\omega = \sum_i \int_{V_i} \mathrm{d}\omega_i$$

and it suffices to show that $\int_{V_i} \mathrm{d}\omega_i = \int_{\partial V_i} \omega_i$ for $i = 1, \dots, m$. In a boundary chart ω_i has the form

$$\omega_i = \sum_{j=1}^n a_j(q)\,\mathrm{d}q^1 \wedge \cdots \wedge \widehat{\mathrm{d}q^j} \wedge \cdots \wedge \mathrm{d}q^n\,,$$

and with $\Psi^*\mathrm{d}q^1 = 0$ it holds that

$$\omega_i\Big|_{\partial M} = \Psi^*\omega_i = a_1(0, q_2, \dots, q_n)\,\mathrm{d}q^2 \wedge \cdots \wedge \mathrm{d}q^n\,.$$

Computing

$$\mathrm{d}\omega_i = \sum_{j=1}^n \frac{\partial a_j(q)}{\partial q_j}(-1)^{j-1}\mathrm{d}q^1 \wedge \cdots \wedge \mathrm{d}q^n\,,$$

we find that

$$
\begin{aligned}
\int_M \mathrm{d}\omega_i &= \sum_{j=1}^n \int_M \frac{\partial a_j}{\partial q_j} (-1)^{j-1} \, \mathrm{d}q^1 \wedge \cdots \wedge \mathrm{d}q^n \\
&= \sum_{j=1}^n \int_0^\infty \mathrm{d}q_1 \int_{-\infty}^\infty \mathrm{d}q_2 \cdots \int_{-\infty}^\infty \mathrm{d}q_n \frac{\partial a_j(q)}{\partial q_j} (-1)^{j-1} \\
&= -\int_{-\infty}^\infty \mathrm{d}q_2 \cdots \int_{-\infty}^\infty \mathrm{d}q_n \, a_1(0, q_2, \cdots, q_n) \\
&= -\int_{\mathbb{R}^{n-1}} a_1(0, q_2, \cdots, q_n) \, \mathrm{d}q^2 \wedge \cdots \wedge \mathrm{d}q^n \;=\; \int_{\partial M} \omega_i \, . \qquad \square
\end{aligned}
$$

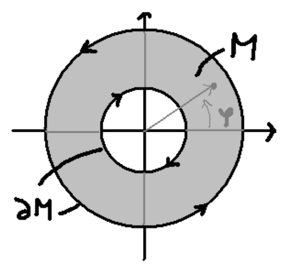
6.17 Examples. (a) Consider the annulus $M = \{(x,y) \in \mathbb{R}^2 \mid \frac{1}{2} \leq x^2 + y^2 \leq 1\}$ and the 1-form $\omega = \frac{-y\mathrm{d}x + x\mathrm{d}y}{x^2+y^2}$ ("$= \mathrm{d}\theta$"). Then $\mathrm{d}\omega = 0$, hence $\int_M \mathrm{d}\omega = 0$, and also

$$
\int_{\partial M} \omega = \int_{x^2+y^2=1} \omega + \int_{x^2+y^2=\frac{1}{2}} \omega = \int_0^{2\pi} \mathrm{d}\theta - \int_0^{2\pi} \mathrm{d}\theta = 2\pi - 2\pi = 0 \, .
$$

Note that while locally ω is the differential of the angle function θ, it can't be exact on all of M: $\omega = \mathrm{d}\nu$ would imply that

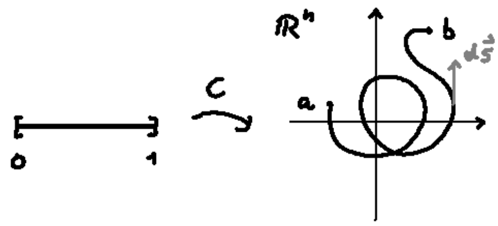
$$
2\pi = \int_{S^1} \omega = \int_{S^1} \mathrm{d}\nu = \int_{\partial S^1} \nu = 0
$$

since $\partial S^1 = \emptyset$.



(b) The integral theorems from calculus are special cases of Stokes theorem (cf. example 5.20).

(i) Let $c : N = [0,1] \to M = \mathbb{R}^n$ be a curve in $\mathbb{R}^n$ and let $f \in C_0^\infty(\mathbb{R}^n)$.



Then

$$
\begin{aligned}
\int_N \mathrm{d}f &= \int_N \frac{\partial f}{\partial q_j} \mathrm{d}q^j = \int_C \mathrm{grad} f \cdot \mathrm{d}\vec{s} \\
&= \int_{\partial N} f = f(c(1)) - f(c(0)) \, .
\end{aligned}
$$

(ii) Let N be 2-dimensional submanifold of $M = \mathbb{R}^3$ with boundary and $\omega = \omega_j \mathrm{d}q^j$ a compactly supported 1-form. Then, on the one hand,

$$
\int_{\partial N} \omega = \int_{\partial N} \omega_j \, \mathrm{d}q^j = \int_{\partial N} \vec{\omega} \cdot \mathrm{d}\vec{s},
$$

and on the other hand

$$
\begin{aligned}
\int_N \mathrm{d}\omega &= \int_N *(*\mathrm{d}\omega) = \int_N (\mathrm{curl}\, \vec{\omega})_j *\mathrm{d}q^j \\
&= \int_N \mathrm{curl}\, \vec{\omega} \cdot \begin{pmatrix} \mathrm{d}q^2 \wedge \mathrm{d}q^3 \\ \mathrm{d}q^3 \wedge \mathrm{d}q^1 \\ \mathrm{d}q^1 \wedge \mathrm{d}q^2 \end{pmatrix} =: \int_N \mathrm{curl}\, \vec{\omega} \cdot \mathrm{d}\vec{F} \, .
\end{aligned}
$$

We thus obtain the classical Stokes theorem in $\mathbb{R}^3$:

$$
\int_{\partial N} \vec{\omega} \cdot \mathrm{d}\vec{s} = \int_N \mathrm{curl}\, \vec{\omega} \cdot \mathrm{d}\vec{F} \, .
$$

(iii) Let N be a n-dimensional submanifold of $M = \mathbb{R}^n$ with boundary and $\omega = *(\omega_j \mathrm{d}q^j)$ a compactly supported $(n-1)$-form. Then

$$\int_{\partial N} \omega \;=\; \int_{\partial N} \omega_j *\mathrm{d}q^j \;=\; \int_{\partial N} \vec{\omega} \cdot \mathrm{d}\vec{F}$$

and

$$\int_N \mathrm{d}\omega \;=\; \int_N * * \mathrm{d} * \omega_j \,\mathrm{d}q^j \;=\; \int_N * \operatorname{div} \vec{\omega} \;=\; \int_N \operatorname{div} \vec{\omega}\ \varepsilon \;=\; \int_N \operatorname{div} \vec{\omega}\ \mathrm{d}V\,.$$

We thus obtain the Gauß theorem in $\mathbb{R}^n$:

$$\int_{\partial N} \vec{\omega} \cdot \mathrm{d}\vec{F} \;=\; \int_N \operatorname{div} \vec{\omega}\ \mathrm{d}V\,.$$

◇

We now introduce the notion of smoothly deforming submanifolds (or smooth images of manifolds) into each other.

6.18 Definition. Diffeotopy (=smooth homotopy)

Let $N_0 = \psi_0(N)$ and $N_1 = \psi_1(N)$ be smooth images of a p-dimensional manifold N in the n-dimensional manifold M, i.e. $\psi_0 : N \to M$ and $\psi_1 : N \to M$ are smooth. The maps ψ_0 and ψ_1 are called **diffeotopic** (or **smoothly homotopic**), if there exists a smooth function $F : [0,1] \times N \to M$ such that

$$\psi_0 = F \circ \iota_0 : N \to N_0 \qquad \text{and} \qquad \psi_1 = F \circ \iota_1 : N \to N_1\,.$$

Here ι_0 and ι_1 denotes the inclusion of N in $\{0\} \times N$ respectively in $\{1\} \times N$, i.e. $\iota_\ell : N \to [0,1] \times N$, $x \mapsto \iota_\ell(x) = (\ell, x)$. The map F is then called a **diffeotopy**. If N has a boundary, one requires in addition that $\psi_0|_{\partial N} = \psi_1|_{\partial N} = F(t,\cdot)|_{\partial N}$ for all $t \in (0,1)$, and refers to a **diffeotopy with fixed boundary**.

An further corollary of Stokes theorem is now the following statement: the integral of a closed differential form over the smooth image of a p-dimensional manifold doesn't change, if one smoothly deforms the latter while keeping the image of its boundary fixed.

6.19 Theorem. Invariance of the integral of closed forms under diffeotopies

Let M be an n-dimensional smooth manifold, N a p-dimensional smooth orientable manifold, and $\psi_0 : N \to M$ and $\psi_1 : N \to M$ smooth and diffeotopic. For every closed $\omega \in \Lambda_p(M)$ with compact support it holds that

$$\int_{N_0} \omega = \int_{N_1} \omega\,.$$

Proof. Homework assignment. □

We end this section with a few simple observations that will be useful in applications.

6.20 Proposition. Invariance of the integral under diffeomorphisms

Let $\Phi : M_1 \to M_2$ be a diffeomorphism and let $\omega \in \Lambda_n(M_2)$ be compactly supported. Then

$$\int_{M_1} \Phi^* \omega = \int_{M_2} \omega\,.$$

Proof. See the proof of proposition 6.5. □

6.21 Corollary. Let $\Phi : M \to M$ be a diffeomorphism and $\Omega \in \Lambda_n(M)$ an invariant volume form, i.e. $\Phi^*\Omega = \Omega$. Then it holds for all functions $f \in C_0^\infty(M)$ that

$$\int_M f\Omega = \int_M (f \circ \Phi)\Omega$$

Proof. This follows from proposition 6.20 and

$$\Phi^*(f\Omega) = (f \circ \Phi)\Phi^*\Omega = (f \circ \Phi)\Omega\,. \qquad \square$$

6.22 Remark. **The integral of measurable resp. integrable functions**

The integral defined in definition 6.9 can be extended to measurable functions in a straightforward way: Let $\Omega \in \Lambda_n(M)$ be a positive volume form and let $f : M \to [0, \infty)$ be measurable. Then one defines

$$\begin{aligned}\int_M f\Omega &= \sum_i \int_{\varphi_i(V_i)} \varphi_{i*}(\chi_i f\Omega) = \sum_i \int_{\varphi_i(V_i)} (\chi_i f \circ \varphi_i^{-1})\varphi_{i*}\Omega \\ &= \sum_i \int_{\varphi_i(V_i)} (\chi_i f \circ \varphi_i^{-1})(q)\,\Omega(q)\,\mathrm{d}^n q\,,\end{aligned}$$

where the last integral is again a Lebesgue integral on $\mathbb{R}^n$. One calls $f : M \to \mathbb{R}$ integrable if $\int_M |f|\Omega < \infty$ and defines for integrable $f \in L^1(M, \Omega)$ its integral as $\int_M f\Omega := \int_M f^+\Omega - \int_M f^-\Omega$.

German nomenclature

compact support = kompakter Träger
orientation = Orientierung
volume form = Volumenform
integral = Integral
partition of unity = Zerlegung der Eins

7 Integral curves and flows

Similar to the situation on $\mathbb{R}^n$ also vector fields on manifolds define ordinary differential equations (ODEs): Let $I \subset \mathbb{R}$ be an open interval and $u : I \to M$ a smooth curve in a manifold M. At each point $u(t) \in M$ on the curve the tangent vector defined by the curve is

$$\dot{u}(t) = [u(\cdot + t)]_{u(t)} = (Du \circ e)(t) \in T_{u(t)}M \,,$$

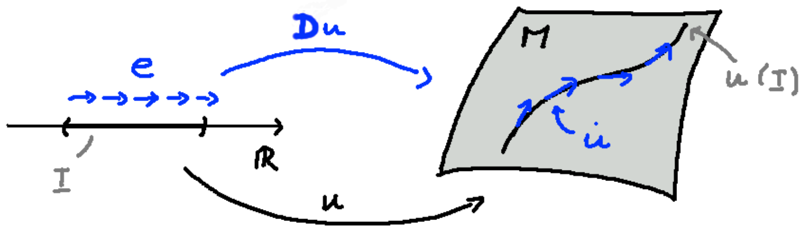
where $e : I \to TI$, $t \mapsto (t, 1)$ denotes the canonical unit vector field on I.

7.1 Definition. Integral curve
A smooth curve $u \in C^\infty(I, M)$ with $I \subset \mathbb{R}$ an open interval is called **integral curve** for the vector field $X \in \mathcal{T}_0^1(M)$ if

$$\dot{u} \;\; := \;\; Du \circ e \;\; = \;\; X \circ u \,, \tag{7.1}$$

i.e. if the following diagram is commutative:

$$\begin{array}{ccc} I & \xrightarrow{u} & M \\ e \downarrow & & \downarrow X \\ TI & \xrightarrow{Du} & TM \end{array}$$



7.2 Remark. Using a coordinate chart φ with

$$\varphi \circ u : t \mapsto \begin{pmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{pmatrix} \qquad \text{and} \qquad \varphi_* X : \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix} \mapsto \left(\begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix}, \begin{pmatrix} v_1(q) \\ \vdots \\ v_n(q) \end{pmatrix} \right)$$

we can apply $D\varphi$ on both sides of (7.1) to obtain the familiar form

$$\dot{u}_j(t) = v_j\,(u_1(t), \dots, u_n(t)) \quad , \, j = 1\,, \dots, n\,,$$

of a first order ODE on $\mathbb{R}^n$. Hence we can at least locally transfer the results for ODEs in $\mathbb{R}^n$ also to ODEs on manifolds.

7.3 Reminder. Let

$$X : \mathbb{R}^n \supset U \to \mathbb{R}^n$$

be a vector field on an open subset U of $\mathbb{R}^n$ and

$$\dot{u}(t) = X(u(t)) \tag{7.2}$$

the corresponding first order ODE. Then we have the following implications:

(a) If X is continuous, then to each $x_0 \in U$ there exists an $\varepsilon > 0$ and a differentiable curve $u_{x_0} : (-\varepsilon, \varepsilon) \to U$ with $u_{x_0}(0) = x_0$ that solves (7.2) (Peano existence theorem).

(b) If X is locally Lipschitz continuous, then the solution u_{x_0} from (a) is unique (Picard-Lindelöf theorem).

(c) If $X \in C^p(U, \mathbb{R}^n)$, then the solution map $\Phi : (t, x_0) \mapsto \Phi(t, x_0) := u_{x_0}(t)$ is p-times continuously differentiable as a function of the initial data, i.e. $\Phi(t, \cdot) \in C^p(U)$ for all t in the existence interval.

7.4 Theorem. Existence, uniqueness and differentiability of local solutions

Let $X \in \mathcal{T}^1_0(M)$ be a smooth vector field on M. For every $x \in \mathring{M} := M \setminus \partial M$ there exist $\varepsilon > 0$, an open neighbourhood U of x, and a unique map

$$\begin{array}{rcl} \Phi : (-\varepsilon\,, \varepsilon) \times U & \to & M \\ (t, x_0) & \mapsto & \Phi(t\,, x_0)\,, \end{array}$$

such that

(a) for every $x_0 \in U$ the curve

$$\Phi_{x_0} : (-\varepsilon, \varepsilon) \to M\,, \quad t \mapsto \Phi_{x_0}(t) := \Phi(t, x_0)$$

is an integral curve of X through x_0, i.e. $\dot{\Phi}_{x_0} = X \circ \Phi_{x_0}$ and $\Phi_{x_0}(0) = x_0$,

and

(b) for every $t \in (-\varepsilon, \varepsilon)$ the map

$$\Phi_t : U \to M\,, \quad x_0 \mapsto \Phi_t(x_0) := \Phi(t, x_0)$$

is a diffeomorphism from U onto an open subset of M .

Proof. The statements follow immediately from the corresponding results on $\mathbb{R}^n$ when we restrict to a chart. □

Also the existence of unique maximal solutions that necessarily leave every compact set can be shown exactly as in the case of $\mathbb{R}^n$.

7.5 Theorem. Existence and behaviour of maximal solutions

Let $X \in \mathcal{T}^1_0(M)$.

(a) For every initial point $x_0 \in \mathring{M}$ there exists a unique maximal solution of $\dot{u}_{x_0} = X \circ u_{x_0}$. More precisely, there exist an open interval $I_{x_0} \subset \mathbb{R}$ and a unique $u_{x_0} : I_{x_0} \to M$ with $\dot{u}_{x_0} = X \circ u_{x_0}$ and $u_{x_0}(0) = x_0$. The solution is maximal and unique in the following sense: for every other solution $\tilde{u} : \tilde{I} \to M$ with $\dot{\tilde{u}} = X \circ \tilde{u}$ and $\tilde{u}(0) = x_0$ it holds that $\tilde{I} \subset I_{x_0}$ and $u_{x_0}|_{\tilde{I}} = \tilde{u}$.

Moreover, the set $D := \{(t, x_0) \in \mathbb{R} \times \mathring{M} \,|\, t \in I_{x_0}\}$ is open and

$$\Phi^X : D \to M\,, \quad (t, x_0) \mapsto \Phi^X(t, x_0) := u_{x_0}(t)$$

is called the **maximal flow**. For any $t \in \mathbb{R}$ the map

$$\Phi^X_t : \{x_0 \in \mathring{M} \,|\, (t, x_0) \in D\} =: D_t \to M\,, \quad x_0 \mapsto \Phi^X(t, x_0)\,,$$

is a diffeomorphism onto its range.

(b) Let $x_0 \in \mathring{M}$ and $I_{x_0} =: (t^-(x_0), t^+(x_0))$ be the existence interval of the maximal solution $u_{x_0} : I_{x_0} \to M$ of $\dot{u} = X \circ u$ starting at x_0. Let $K \subset \mathring{M}$ be compact. If $t^+(x_0) < \infty$, then there exists $0 < \tau < t^+(x_0)$ such that

$$u_{x_0}(t) \notin K \qquad \text{for all} \quad t \in (\tau, t^+(x_0))\,.$$

Put differently, a maximal integral curve can not "end" inside a compact set that doesn't contain boundary points. Thus, if a solution does not exist for all times, then it must either run to infinity in finite time or hit the boundary of M.

7.6 Definition. Global flow

Let $\Phi : \mathbb{R} \times M \to M$ be a smooth map such that for every $t \in \mathbb{R}$ the map

$$\Phi_t : M \to M\,, \quad x \mapsto \Phi(t, x)\,,$$

is a diffeomorphism with $\Phi_0 = \mathrm{id}_M$ and such that

$$\Phi_{t_1} \circ \Phi_{t_2} = \Phi_{t_1+t_2} \qquad \text{for all} \quad t_1, t_2 \in \mathbb{R}\,. \tag{7.3}$$

Then Φ is called a **global flow**. ◇

7.7 Definition. Complete vector fields and global flows

Let $X \in \mathcal{T}^1_0(M)$ and $\partial M = \emptyset$. If $I_{x_0} = \mathbb{R}$ holds for all $x_0 \in M$ then $D = \mathbb{R} \times M$ and, according to theorem 7.5 (a), $\Phi^X_t : M \to M$ is a diffeomorphism for all $t \in \mathbb{R}$, and, by the uniqueness of integral curves,

$$\Phi^X_{t_1} \circ \Phi^X_{t_2} = \Phi^X_{t_1+t_2} \qquad \text{for all} \quad t_1, t_2 \in \mathbb{R}\,. \tag{7.4}$$

Hence Φ^X is a global flow and X is called **complete**. ◇

For simplicity we discuss only global flows in the following. But all results hold also for local flows where one has to restrict the domains appropriately, as indicated in theorem 7.5.

7.8 Definition. Closed manifold

A compact manifold without boundary is called a **closed manifold**. ◇

7.9 Corollary. Vector fields on closed manifolds are complete

Let M be closed. Then every vector field $X \in \mathcal{T}^1_0(M)$ on M is complete and the corresponding flow Φ^X is a global flow.

Proof. This is a direct consequence of theorem 7.5 (b). □

7.10 Proposition. A global flow is always the flow of a complete vector field X

Let Φ be a global flow on M. There exists a unique vector field $X \in \mathcal{T}^1_0(M)$ such that $\Phi = \Phi^X$.

Proof. For $x \in M$ define the curve $\Phi_x : \mathbb{R} \to M$, $t \mapsto \Phi(t, x)$, and the vector field

$$X : M \to TM\,, \quad x \mapsto X(x) := (D\Phi_x \circ e)(0)\,.$$

Since $\Phi_x(0) = x$, we have $(D\Phi_x \circ e)(0) = X(x) = (X \circ \Phi_x)(0)$. To conclude that $\Phi = \Phi^X$ it remains to check that $(D\Phi_x \circ e)(t) = (X \circ \Phi_x)(t)$ for all $t \in \mathbb{R}$. However, by the flow property (7.3) we have for $T_t : \mathbb{R} \to \mathbb{R}$, $s \mapsto s + t$ and $y := \Phi_t(x)$ that

$$(\Phi_x \circ T_t)(s) = \Phi(s + t, x) = \Phi_{s+t}(x) \overset{(7.3)}{=} (\Phi_s \circ \Phi_t)(x) = \Phi_s(\Phi_t(x)) = \Phi_s(y) = \Phi_y(s)\,,$$

i.e. $\Phi_x \circ T_t = \Phi_y$. Using that T_t is the flow of e, we find

$$\begin{aligned}
(X \circ \Phi_x)(t) &= (X \circ \Phi_x \circ T_t)(0) = (X \circ \Phi_y)(0) = (D\Phi_y \circ e)(0) \\
&= (D(\Phi_x \circ T_t) \circ e)(0) = (D\Phi_x \circ DT_t \circ e)(0) \\
&= (D\Phi_x \circ e \circ T_t)(0) = (D\Phi_x \circ e)(t)\,. \qquad \square
\end{aligned}$$

7.11 Definition. **Killing vector fields and Hamiltonian vector fields**

Let $g \in \mathcal{T}_2^0(M)$ be either a (pseudo-)metric or a symplectic form and $X \in \mathcal{T}_0^1$ a complete vector field. If the flow maps Φ_t^X are isometries resp. canonical transformations, then X is called a **Killing vector field** (named after Wilhelm Killing, a German mathematician) resp. a **Hamiltonian vector field**. $\diamond$

7.12 Example. **Linear drift on $\mathbb{R}^n$**

Let $M = \mathbb{R}^n$, $L : (q_1, \ldots, q_n) \mapsto ((q_1, \ldots, q_n), (1, 0, 0, \ldots, 0))$. Then Φ^L is a global flow, called the **linear drift**, and is explicitly given by

$$\Phi^L(t, q) = (q_1 + t, q_2, \ldots, q_n)\,.$$

The linear drift is obviously a Killing field for the euclidean metric on $\mathbb{R}^n$.

7.13 Proposition. **Integral curves and diffeomorphisms**

Let $\Psi : M_1 \to M_2$ be a diffeomorphism, $X \in \mathcal{T}_0^1(M_1)$ a vector field and $u : I \to M_1$ an integral curve of X. Then $\Psi \circ u : I \to M_2$ is an integral curve of $\Psi_* X$.

If X is complete, we thus have

$$\Psi \circ \Phi_t^X = \Phi_t^{\Psi_* X} \circ \Psi\,.$$

Proof.

$$D(\Psi \circ u) \circ e = D\Psi \circ Du \circ e = D\Psi \circ X \circ u = D\Psi \circ X \circ \Psi^{-1} \circ \Psi \circ u = \Psi_* X \circ (\Psi \circ u)\,. \qquad \square$$

The following theorem states that in appropriate coordinates every flow Φ^X has locally the form of a linear flow with the exception of the fixed points where $X = 0$. Note that for a point x with $X(x) = 0$ the unique integral curve through x is $u(t) \equiv x$ for all $t \in \mathbb{R}$.

7.14 Theorem. **Normal form of a flow away from the fixed points**

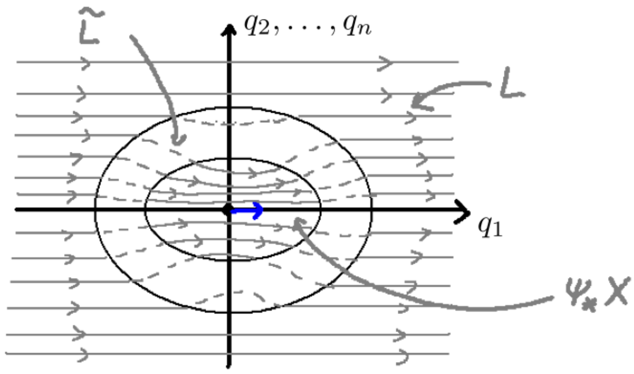
Let $X \in \mathcal{T}_0^1(M)$ and $x \in M$ with $X(x) \neq (x, 0)$. Then there exists a chart (V, φ) with $x \in V$ such that

$$\varphi_* X = L$$

and thus locally

$$\Phi_t^X = \varphi^{-1} \circ \Phi_t^L \circ \varphi\,.$$

Proof. Since $X(x) \neq (x, 0)$ there exists a chart (V_1, ψ) with $\psi(x) = 0 \in \mathbb{R}^n$ and $\psi_* X(0) = (1, 0, \ldots, 0)$. Since $\psi_* X \in \mathcal{T}_0^1(\psi(V_1))$ is continuous, there exists an open relatively compact neighbourhood $U_2 \subset \psi(V_1)$ of 0 on which the first component of $\psi_* X$ remains larger than $\frac{1}{2}$, i.e. $(\psi_* X)_1(q) > \frac{1}{2}$ for all $q \in U_2$.



Strategy: We first interpolate the vector fields $\psi_* X$ on $U \subset U_2$ and L on U_2^c in order to obtain a smooth vector field $\tilde{L}$ on all of $\mathbb{R}^n$. Next we show that L and $\tilde{L}$ are diffeomorphic, i.e. that there exists a diffeomorphism $\Omega : \mathbb{R}^n \to \mathbb{R}^n$ such that $\Omega_* \tilde{L} = L$. Then $\varphi = \Omega \circ \psi$ is the chart we are looking for since $\varphi_* X = \Omega_* \psi_* X = \Omega_* \tilde{L} = L$ on $V = \psi^{-1}(U)$.

So let $U \subset U_2$ be open with $0 \in U$ and pick a function $f \in C^\infty(\mathbb{R}^n)$ with

$$f(q) = \begin{cases} 1 & \text{if} \quad q \in U \\ 0 & \text{if} \quad q \in U_2^c \end{cases}$$

and $0 \leq f(q) \leq 1$. Then the interpolating vector field is defined as

$$\tilde{L} = f\, \psi_* X + (1 - f)\, L \in \mathcal{T}_0^1(\mathbb{R}^n)\,.$$

Clearly the flow $\Phi_t^{\tilde{L}}$ of $\tilde{L}$ is global as $\mathbb{R}^n$ has no boundary and no integral curve can escape to infinity in finite time. We now show that

$$\Omega = \lim_{t\to\infty} \Phi_{-t}^L \circ \Phi_t^{\tilde{L}}$$

exists and defines a diffeomorphism such that

$$\Omega_* \tilde{L} = L\,.$$

Since $(\Phi_t^{\tilde{L}}(q))_1 \geq q_1 + \frac{1}{2}\,t$, every integral curve leaves U_2 after a finite time. Hence on compact sets $K \subset \mathbb{R}^n$ the limit

$$\lim_{t\to\infty} \Phi_{-t}^L \circ \Phi_t^{\tilde{L}}\Big|_K = \Phi_{-\tau}^L \circ \Phi_\tau^{\tilde{L}}$$

is reached already after a finite time $\tau_0(K)$. Thus, Ω is well defined and a diffeomorphism. Moreover,

$$\Omega \circ \Phi_t^{\tilde{L}} \quad = \quad \lim_{s\to\infty} \Phi_{-s}^L \circ \Phi_s^{\tilde{L}} \circ \Phi_t^{\tilde{L}} \quad = \quad \lim_{s\to\infty} \Phi_t^L \circ \Phi_{-s-t}^L \circ \Phi_{s+t}^{\tilde{L}} \quad = \quad \Phi_t^L \circ \Omega$$

and the flows $\Phi^{\tilde{L}}$ and Φ^L are diffeomorphic. In particular, we also have $\Phi_x^{\tilde{L}} = \Omega^{-1} \circ \Phi_{\Omega(x)}^L$. Taking derivatives we find that for all $x \in \mathbb{R}^n$

$$\tilde{L} \circ \Phi_x^{\tilde{L}} = D\Phi_x^{\tilde{L}} \circ e = D\Omega^{-1} \circ D\Phi_{\Omega(x)}^L \circ e = D\Omega^{-1} \circ L \circ \Phi_{\Omega(x)}^L\,.$$

At $t = 0$ this gives $\tilde{L} = D\Omega^{-1} \circ L \circ \Omega = \Omega^* L$. □

7.15 Remark. **Linearisation of a vector field at a fixed point**

Away from the fixed points the flow of a vector field is diffeomorphic to the linear drift. The local behaviour near a fixed point can be analysed by looking at the **linearisation** of the vector field at such a fixed point.

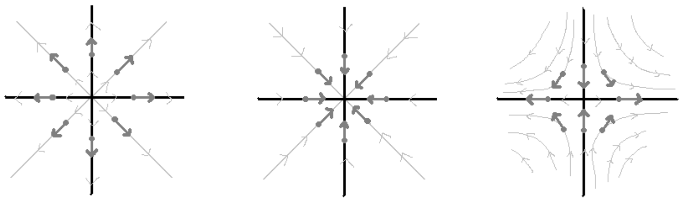
Let $X \in \mathcal{T}_0^1(M)$ and $x_0 \in M$ with $X(x_0) = (x_0, 0)$. In a chart φ with $\varphi(x_0) = 0$ it then holds that

$$X_\varphi(q) \quad = \quad \underbrace{X_\varphi(0)}_{=0} + DX_\varphi|_0\, q + \mathcal{O}(\|q\|^2) \quad = \quad DX_\varphi|_0\, q + \mathcal{O}(\|q\|^2)\,,$$

where we abbreviated $X_\varphi(q) := (I \circ \varphi_* X)(q)$. Close to $q = 0$ we can thus approximate $X_\varphi(q)$ by its linearisation $DX_\varphi|_0\, q$. Qualitatively, the behaviour close to the fixed point is determined by the eigenvalues of $DX_\varphi|_0$ and their (geometric) multiplicities. These are independent of the chosen chart since for different charts φ and ψ the differentials $DX_\varphi|_0$ and $DX_\psi|_0$ are similar matrices.

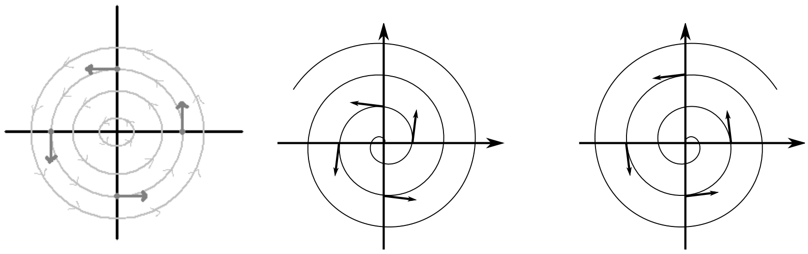
We illustrate the different possible types of fixed points only for $n = 2$:

(a) $DX|_0$ diagonalisable with real eigenvalues and
(i) $\lambda_1, \lambda_2 > 0:$ (ii) $\lambda_1, \lambda_2 < 0:$ (iii) $\lambda_1 > 0, \lambda_2 < 0:$

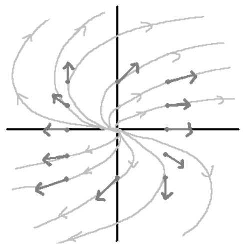


(b) $DX|_0$ has two complex conjugate eigenvalues $\lambda_{1/2} = a \pm \mathrm{i}b$, e.g. $DX|_0 = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$:

(i) $b > 0, a = 0:$ (ii) $b > 0, a > 0:$ (iii) $b > 0, a < 0:$



(c) $DX|_0$ has one real eigenvalue of geometric multiplicity 1, e.g. $DX|_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$:



German nomenclature

complete vector field = vollständiges Vektorfeld
fixed point = Fixpunkt
solution = Lösung
integral curve = Integralkurve
flow = Fluss

8 The Lie derivative

The exterior derivative of differential forms is defined on any smooth manifold without involving further structure. If one tries to differentiate other tensor fields, e.g. vector fields, in the conventional way by taking the limit of a difference quotient, the following question comes up: how can one compare the values $\tau(x)$ and $\tau(y)$ of a tensor field $\tau \in \mathcal{T}^r_s(M)$ at nearby points? In contrast to the familiar situation on $\mathbb{R}^n$, $\tau(x)$ and $\tau(y)$ are elements of different vector spaces, namely of $T_x{}^r_s M$ and $T_y{}^r_s M$, and thus an expression of the form

$$D_v\tau|_x = \frac{\mathrm{d}}{\mathrm{d}t}(\tau\circ c_v)(t)|_{t=0} = \lim_{h\to 0}\frac{\tau(c_v(h)) \overset{?}{-} \tau(x)}{h}\,,\quad v = [c_v]_x\,,$$

makes no sense. One could, of course, use charts and define, e.g., for a vector field $X \in \mathcal{T}^1_0(M)$

$$DX = D(X^i\partial_{q_i}) := \frac{\partial X^i}{\partial q_j}\,\mathrm{d}q^j\otimes\partial_{q_i}\,.$$

However, this definition is **not** independent of the chosen chart and does **not** define a global tensor field in $\mathcal{T}^1_1(M)$.

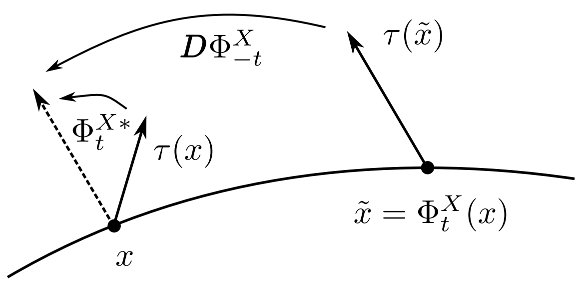
In order to differentiate tensor fields in the usual sense we need further structure that allows for a coordinate independent identification of neighbouring tangent spaces.

- A **connection** ∇ (on the tangent bundle) allows one to define the derivative $\nabla_v X|_x$ of a vector field X in the direction $v \in T_xM$ at the point $x \in M$. Every metric on M induces a connection, the so called Levi-Civita connection. Hence, on Riemannian manifolds we have a canonical way of differentiating vector- and other tensor fields. We will study connections in section 10.
- Alternatively, a vector field X yields itself an identification of neighbouring tangent spaces by using the corresponding flow Φ^X_t. The pull-back Φ^{X*}_t under the flow map is an isomorphism between $T_y{}^r_s M$ at $y = \Phi^X_t(x)$ and $T_x{}^r_s M$ and the limit of the corresponding difference quotient defines the Lie derivative. Note, however, that we now take the derivative in the direction of a vector field, not in the direction of a single tangent vector.

8.1 Definition. <u>**The Lie derivative**</u>

For $X \in \mathcal{T}^1_0(M)$ the **Lie derivative** of a tensor field $\tau \in \mathcal{T}^r_s(M)$ is defined by

$$L_X\tau := \lim_{t\to 0}\frac{\Phi^{X*}_t\tau - \tau}{t} = \frac{\mathrm{d}}{\mathrm{d}t}\Phi^{X*}_t\tau\,|_{t=0}\,.$$



Note that the pull back Φ^{X*}_t acts fibre wise as a linear map $\Phi^{X*}_t|_x$ on the vector space $T_x{}^r_s$ and that thus $\frac{\mathrm{d}}{\mathrm{d}t}(\Phi^{X*}_t\tau)(x) = (\frac{\mathrm{d}}{\mathrm{d}t}\Phi^{X*}_t|_x)\,\tau(x)$.

8.2 Remark. For $r = s = 0$ we recover

$$L_X f = \frac{\mathrm{d}}{\mathrm{d}t} f \circ \Phi_t^X \,|_{t=0} = \mathrm{d}f(X) = (\mathrm{d}f \,|X)\,,$$

i.e. the Lie derivative of functions from definition 2.29. In this specific case $(L_X f)(y)$ depends only on the value $X(y)$ of X at the point y. For $r + s > 0$, in contrast, the value $(L_X\tau)(y)$ of $L_X\tau$ at y depends on the behaviour of X on a whole neighbourhood of $y \in M$. ◇

8.3 Lemma. **L_X commutes with the flow maps Φ_t^X**

Let $X \in \mathcal{T}_0^1(M)$ and $\tau \in \mathcal{T}_s^r(M)$. Then

$$L_X \, \Phi_t^{X*} \, \tau(x) = \Phi_t^{X*} \, L_X \, \tau(x)$$

for all $t \in I_x$. If X is complete, it thus holds for all $t \in \mathbb{R}$ that

$$L_X \, \Phi_t^{X*} \, \tau = \Phi_t^{X*} \, L_X \, \tau\,.$$

Proof. By the flow property $\Phi_s^X \circ \Phi_t^X = \Phi_{s+t}^X = \Phi_t^X \circ \Phi_s^X$ we have that $\Phi_s^{X*}\,\Phi_t^{X*} = \Phi_{t+s}^{X*} = \Phi_t^{X*}\,\Phi_s^{X*}$, which restricted to the fibre $T_{x\,s}^{\,r}$ are just compositions of linear maps $\Phi_s^{X*}|_x \Phi_t^{X*}|_x = \Phi_{t+s}^{X*}|_x = \Phi_t^{X*}|_x \Phi_s^{X*}|_x$. Hence,

$$\begin{aligned} L_X \, \Phi_t^{X*}|_x \, \tau(x) &= \left(\frac{\mathrm{d}}{\mathrm{d}s}\Phi_s^{X*}|_x\right)|_{s=0}\Phi_t^{X*}|_x \, \tau(x) = \left(\frac{\mathrm{d}}{\mathrm{d}s}\,\Phi_{s+t}^{X*}\right)|_{s=0}\tau(x) \\ &= \Phi_t^{X*}|_x \left(\frac{\mathrm{d}}{\mathrm{d}s}\Phi_s^{X*}|_x\right)|_{s=0}\tau(x) = \Phi_t^{X*}|_x L_X \, \tau(x)\,. \end{aligned}$$

□

8.4 Proposition. **Properties of the Lie derivative**

Let $t_1, t_2 \in \mathcal{T}_s^r(M)$, $t_3 \in \mathcal{T}_{s'}^{r'}(M)$, $t_4 \in \mathcal{T}_r^s(M)$, and $X \in \mathcal{T}_0^1(M)$. Then

(i) $L_X(t_1 + t_2) = L_X t_1 + L_X t_2$

(ii) $L_X(t_1 \otimes t_3) = L_X t_1 \otimes t_3 + t_1 \otimes L_X t_3$

(iii) $L_X(t_1 \,|\, t_4) = (L_X t_1 \,|\, t_4) + (t_1 \,|\, L_X t_4)$

Proof. For Φ_t^{X*} it holds that

(i) $\Phi_t^{X*}(t_1 + t_2) = \Phi_t^{X*} t_1 + \Phi_t^{X*} t_2$

(ii) $\Phi_t^{X*}(t_1 \otimes t_3) = \Phi_t^{X*} t_1 \otimes \Phi_t^{X*} t_3$

(iii) $\Phi_t^{X*}(t_1 \,|\, t_4) = (\Phi_t^{X*} t_1 \,|\, \Phi_t^{X*} t_4)\,.$

Evaluating in each case the derivative with respect to t at $t = 0$ yields the corresponding claim. □

8.5 Remark. Let $g \in \mathcal{T}_2^0(M)$ be a metric or symplectic form and denote by $\langle\cdot|\cdot\rangle_g := G(\cdot,\cdot)$ the bilinear form of definition 4.12. For isometries resp. canonical transformations Φ it holds by definition that

$$\Phi^* g = g \quad \text{and thus} \quad \Phi^*\langle t_1 \,|\, t_2\rangle_g = \langle \Phi^* t_1 \,|\, \Phi^* t_2\rangle_g \quad \text{for tensor fields } t_1 \text{ and } t_2 \text{ of the same type}\,.$$

Hence, for Killing resp. Hamiltonian vector fields X is holds that

$$L_X\langle t_1 \,|\, t_2\rangle_g = \langle L_X t_1 \,|\, t_2\rangle_g + \langle t_1 \,|\, L_X t_2\rangle_g\,.$$

Warning: for general vector fields this is not true, since also g itself must be "differentiated"! ◇

8.6 Proposition. **Naturality of the Lie derivative**
Let $\Psi : M_1 \to M_2$ be a diffeomorphism, $X \in \mathcal{T}_0^1(M_1)$, $\omega \in \Lambda_k(M_1)$, and $\tau \in \mathcal{T}_s^r(M_1)$. Then

$$\Psi_* \, L_X \tau = L_{\Psi_* X} \, \Psi_* \, \tau$$

and

$$\mathrm{d} L_X \omega = L_X \mathrm{d}\omega \,.$$

Proof. According to proposition 7.13 we have

$$\Psi \circ \Phi_t^X = \Phi_t^{\Psi_* X} \circ \Psi \,.$$

Hence

$$\Psi_* (\Phi_t^X)_* \tau = (\Phi_t^{\Psi_* X})_* \, \Psi_* \, \tau \,,$$

and thus, after taking a derivative,

$$\Psi_* \, L_X \tau = L_{\Psi_* X} \, \Psi_* \, \tau \,.$$

Since the exterior derivative commutes with the pull-back $(\Phi_t^X)^*$, $\mathrm{d}(\Phi_t^X)^*\omega = (\Phi_t^X)^* \mathrm{d}\omega$, it also commutes with L_X. □

8.7 Definition. **The Lie bracket of vector fields**
The Lie bracket of two vector fields $X, Y \in \mathcal{T}_0^1(M)$ is the vector field

$$[X, Y] := L_X Y \,. \qquad \diamond$$

8.8 Proposition. **Properties of the Lie bracket**
For $t \in \mathcal{T}_s^r(M)$ and $X, Y, Z \in \mathcal{T}_0^1(M)$ we have

$$L_{[X,Y]} t = (L_X L_Y - L_Y L_X) t \,. \tag{8.1}$$

As a consequence, $[X, Y] = -[Y, X]$ and the Jacobi identity holds:

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \,.$$

Proof. On $\mathcal{T}_0^0(M)$ we have with proposition 8.4 (iii), remark 8.2, and proposition 8.6 that

$$L_{[X,Y]} f \overset{8.2}{=} (\mathrm{d}f | [X, Y]) = (\mathrm{d}f | L_X Y) \overset{8.4,8.6}{=} L_X (\mathrm{d}f | Y) - (\mathrm{d} L_X f | Y) = L_X L_Y f - L_Y L_X f \,.$$

It follows that $[X, Y] = -[Y, X]$ since $L_{-Z} = -L_Z$ and since a vector field is uniquely specified by its action on $\mathcal{T}_0^0$. By the same reasoning the Jacobi identity follows from the following computation:

$$\begin{aligned}
&([X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]])(f) = \\
&= \; L_X L_{[Y,Z]} f - L_{[Y,Z]} L_X f + L_Y L_{[Z,X]} f - L_{[Z,X]} L_Y f + L_Z L_{[X,Y]} f - L_{[X,Y]} L_Z f \\
&= \; L_X L_Y L_Z f - L_X L_Z L_Y f - L_Y L_Z L_X f + L_Z L_Y L_X f \\
&\quad + L_Y L_Z L_X f - L_Y L_X L_Z f - L_Z L_X L_Y f + L_X L_Z L_Y f \\
&\quad + L_Z L_X L_Y f - L_Z L_Y L_X f - L_X L_Y L_Z f + L_Y L_X L_Z f \; = \; 0 \,.
\end{aligned}$$

With proposition 8.6 we get

$$L_{[X,Y]} \mathrm{d}f = L_X L_Y \mathrm{d}f - L_Y L_X \mathrm{d}f$$

and thus, with proposition 8.4 (i), the equality (8.1) also on $\mathcal{T}_1^0(M)$. With proposition 8.4 (ii) and (iii) (8.1) holds on all $\mathcal{T}_s^r(M)$. □

8.9 Examples. Let X be given within a coordinate chart as $X = X^i \partial_{q_i}$.

(a) For $f \in \mathcal{T}_0^0$ we have $L_X f = (\mathrm{d}f \mid X) = \frac{\partial f}{\partial q_i} X^i =: f_{,i} X^i$.

(b) For $\omega = \omega_i \mathrm{d}q^i \in \mathcal{T}_1^0$ we have

$$\begin{aligned} L_X \omega &= (L_X \omega_i)\mathrm{d}q^i + \omega_i \mathrm{d}(L_X q^i) \;=\; \omega_{i,k}\, X^k \mathrm{d}q^i + \omega_i \mathrm{d}(X^i) \\ &= \omega_{i,k}\, X^k \mathrm{d}q^i + \omega_i X^i_{,k} \mathrm{d}q^k \;=\; (\omega_{i,k}\, X^k + \omega_k\, X^k_{,i})\mathrm{d}q^i \,. \end{aligned}$$

(c) Since $L_{[X,Y]} f = X^j \partial_j Y^i \partial_i f - Y^j \partial_j X^i \partial_i f = (X^j Y^i_{,j} - Y^j X^i_{,j}) f_{,i}$, the commutator has the coordinate expression

$$[X,Y] = (X^j Y^i_{,j} - Y^j X^i_{,j})\partial_{q_i} = ((\mathrm{d}Y^i | X) - (\mathrm{d}X^i | Y))\partial_{q_i} \,. \qquad \diamond$$

8.10 Lemma. Let $X \in \mathcal{T}_0^1(M)$. Then

$$\Phi_t^{X*}|_x X(x) = X(x) \tag{8.2}$$

for all $t \in I_x$. Thus, for complete vector fields it holds that $\Phi_t^{X*} X = X$ for all $t \in \mathbb{R}$.

Proof. For $t = 0$ the equality (8.2) holds, since $\Phi_0^{X*}|_x = \mathrm{id}$. Moreover,

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi_t^{X*}|_x X(x) = \left(\frac{\mathrm{d}}{\mathrm{d}s}\Phi_{t+s}^{X*}|_x\right)|_{s=0} X(x) = \Phi_t^{X*}|_x \left(\frac{\mathrm{d}}{\mathrm{d}s}\Phi_s^{X*}|_x\right)|_{s=0} X(x) = \Phi_t^{X*}|_x L_X X = 0 \,,$$

as $L_X X = [X,X] = -[X,X] = -L_X X$ implies that $L_X X = 0$. Hence the left hand side of (8.2) is independent of $t \in I_x$ and thus the equality holds for all times. $\square$

The next statement establishes a close connection between the Lie bracket of two vector fields X and Y and the corresponding flows: the flows commute (statement (iii)) if and only if each vector field is invariant under the flow of the other (statement (ii)) if and only if the Lie bracket $[X,Y]$ vanishes.

8.11 Proposition. **Commuting flows and vector fields**

Let $X, Y \in \mathcal{T}_0^1(M)$ be complete vector fields and Φ_t^X and Φ_s^Y the corresponding flows. Then the following assertions are equivalent:

(i) $[X,Y] = 0$

(ii) $\Phi_t^{X*} Y = Y$ for all $t \in \mathbb{R}$ and $\Phi_s^{Y*} X = X$ for all $s \in \mathbb{R}$

(iii) $\Phi_s^X \circ \Phi_t^Y = \Phi_t^Y \circ \Phi_s^X$ for all $s, t \in \mathbb{R}$.

Proof. Homework assignment. $\square$

Finally, we discuss the relationship between the action of the Lie derivative on differential forms and the exterior derivative. To this end, we need the following definition.

8.12 Definition. **Inner product of a vector field and a differential form**

Let $X \in \mathcal{T}_0^1(M)$ and $\omega \in \Lambda_p(M)$, then the inner product $i_X\omega \in \Lambda_{p-1}(M)$ is defined by

$$i_X\omega(v_1, \ldots, v_{p-1})\,|_y = \omega(X(y), v_1, \ldots, v_{p-1}) \,.$$

Hence $(i_X\omega)_{i_1 \cdots i_{p-1}} = X^j \omega_{j i_1 \cdots i_{p-1}}$. $\diamond$

8.13 Theorem. **The Lie derivative of differential forms: Cartan's formula**
For $X \in \mathcal{T}^1_0(M)$ and $\omega \in \Lambda_p(M)$, $0 \le p \le n$, it holds that

$$L_X\omega = i_X \mathrm{d}\omega + \mathrm{d}(i_X\omega)\,.$$

On differential forms we thus have

$$L_X = i_X \circ \mathrm{d} + \mathrm{d} \circ i_X\,.$$

Proof. For $f \in \mathcal{T}^0_0(M)$ we have according to remark 8.2 that $L_X f = \mathrm{d}f(X) = i_X \mathrm{d}f$. And $i_X f = 0$ holds by definition. For $\omega = \omega_j \mathrm{d}q^j \in \Lambda_1(M)$ proposition 8.4 (ii) yields that

$$L_X\omega_j \mathrm{d}q^j = (L_X\omega_j)\mathrm{d}q^j + \omega_j L_X \mathrm{d}q^j = \mathrm{d}\omega_j(X)\mathrm{d}q^j + \omega_j \mathrm{d}L_X q^j = \mathrm{d}\omega_j(X)\mathrm{d}q^j + \omega_j \mathrm{d}(\mathrm{d}q^j(X))\,.$$

On the other hand

$$\begin{aligned}(i_X\mathrm{d} + \mathrm{d}i_X)\omega_j\mathrm{d}q^j &= \mathrm{d}\omega_j(X)\mathrm{d}q^j - \mathrm{d}q^j(X)\mathrm{d}\omega_j + \mathrm{d}q^j(X)\mathrm{d}\omega_j + \omega_j\mathrm{d}(\mathrm{d}q^j(X))\\ &= \mathrm{d}\omega_j(X)\mathrm{d}q^j + \omega_j\mathrm{d}(\mathrm{d}q^j(X))\,.\end{aligned}$$

Thus, $L_X = i_X \circ \mathrm{d} + \mathrm{d} \circ i_X$ on Λ_0 and on Λ_1. Again with proposition 8.4 (ii) it holds for $\omega \in \Lambda_1$ and $\nu \in \Lambda_k$ that

$$L_X(\omega \wedge \nu) = L_X\omega \wedge \nu + \omega \wedge L_X\nu$$

and

$$\begin{aligned}(i_X\mathrm{d} + \mathrm{d}i_X)(\omega \wedge \nu) &= i_X(\mathrm{d}\omega \wedge \nu - \omega \wedge \mathrm{d}\nu) + \mathrm{d}(\omega(X)\nu - \omega \wedge i_X\nu)\\ &= i_X\mathrm{d}\omega \wedge \nu + \mathrm{d}\omega \wedge i_X\nu - \omega(X)\mathrm{d}\nu + \omega \wedge i_X\mathrm{d}\nu\\ &\quad + \mathrm{d}\omega(X) \wedge \nu + \omega(X)\mathrm{d}\nu - \mathrm{d}\omega \wedge i_X\nu + \omega \wedge \mathrm{d}i_X\nu\\ &= (i_X\mathrm{d} + \mathrm{d}i_X)\omega \wedge \nu + \omega \wedge (i_X\mathrm{d} + \mathrm{d}i_X)\nu\,.\end{aligned}$$

By induction it follows that L_X and $\mathrm{d}i_X + i_X\mathrm{d}$ agree also on $\omega = \frac{1}{p!}\sum_{(i)}\omega_{(i)}\mathrm{d}q^{i_1} \wedge \cdots \wedge \mathrm{d}q^{i_p}$. □

The following proposition computes the rate of change of the volume of a set $V(t) := \Phi^X_t(V)$ that is transported along the flow of a vector field X either in terms of the Lie derivative or in terms of an integral over the boundary of V. It can be seen as a variant of the divergence theorem of Gauß.

8.14 Proposition. **A variant of Gauß' theorem**
Let $\omega \in \Lambda_n(M)$, $V \subset M$ an n-dimensional submanifold, and $X \in \mathcal{T}^1_0(V)$ with compact support. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Vol}^\omega(\Phi^X_t(V))\Big|_{t=0} := \frac{\mathrm{d}}{\mathrm{d}t}\int_{\Phi^X_t(V)}\omega\Big|_{t=0} = \int_V L_X\omega = \int_{\partial V} i_X\omega$$

and, as a consequence, if $\partial M = \emptyset$ then

$$\int_M L_X\omega = 0\,.$$

Proof. The first equality is the infinitesimal version of proposition 6.20. The second equality follows from Cartan's formula and Stokes theorem. □

German nomenclature

to commute = kommutieren, vertauschen	Lie bracket = Lie-Klammer
Lie derivative = Lie-Ableitung	

9 Vector bundles

We have seen already plenty examples of vector bundles, namely all the tensor bundles over a smooth manifold M. A general vector bundle is a base space (a smooth manifold) M where at each point $x \in M$ a finite dimensional real or complex vector space is attached and all these fibres are glued together in a “smooth way”.

9.1 Definition. **Real vector bundles**

A **smooth vector bundle of rank** k over an n-dimensional smooth manifold M is a smooth $(n+k)$-dimensional manifold E together with a smooth surjective map $\pi : E \to M$ with the following properties:

(a) For each $x \in M$ the set $E_x := \pi^{-1}(\{x\}) \subset E$ has the structure of a real k-dimensional vector space, called the **fibre** over x.

(b) For each $x \in M$ there exists a neighbourhood $U \subset M$ of x and a diffeomorphism $\Phi : \pi^{-1}(U) \to U \times \mathbb{R}^k$ such that $\pi_1 \circ \Phi = \pi$ where $\pi_1 : U \times \mathbb{R}^k \to U$ is the projection on the first factor. Moreover, $\Phi|_x : E_x \to \{x\} \times \mathbb{R}^k \cong \mathbb{R}^k$ is a vector space isomorphism.

The manifold E is called the **total space**, M is called the **base space**, π its **projection**, and each map Φ as above is called a **local trivialisation**. If there exists a local trivialisation Φ defined on all of M, i.e. $\Phi : E \to M \times \mathbb{R}^k$, such a map is called a **global trivialisation** and the bundle E is called **trivialisable**. ◇

9.2 Examples. (a) For any manifold M and k-dimensional real vector space V one can define the **trivial bundle** $E = M \times V$ with $\pi(x, v) = x$ and $\Phi = (\mathrm{id}_M, I)$ for some isomorphism $I : V \to \mathbb{R}^k$.

(b) The tangent bundle TM, the cotangent bundle T^*M, and actually all the tensor bundles $T^r_s M$ are vector bundles. For TM the fibres are the tangent spaces T_xM and local trivialisations are given by the bundle charts of definition 2.12. ◇

We now discuss several basic schemes that allow us to construct new vector bundles.

9.3 Lemma. (a) Given a vector bundle $\pi : E \to M$ and a submanifold $N \subset M$, then the restriction $E|_N := \pi_M^{-1}(N)$ with $\pi_N : E|_N \to N$, $\pi_N = \pi|_N$, is a vector bundle.

(b) Given two vector bundles $\pi_1 : E_1 \to M_1$ and $\pi_2 : E_2 \to M_2$ also the product

$$\pi_1 \times \pi_2 : E_1 \times E_2 \to M_1 \times M_2\,, \quad (p_1, p_2) \mapsto (\pi_1(p_1), \pi_2(p_2))$$

is a vector bundle with fibres $\pi_1^{-1}(\{x\}) \times \pi_2^{-1}(\{y\})$.

Proof. (a) is obvious and for (b) note that from local trivialisations $\Phi_1 : \pi_1^{-1}(U_1) \to U_1 \times \mathbb{R}^k$ for E_1 and $\Phi_2 : \pi_2^{-1}(U_2) \to U_2 \times \mathbb{R}^l$ for E_2 one obtains a local trivialisation $\Phi_1 \times \Phi_2 : \pi_1^{-1}(U_1) \times \pi_2^{-1}(U_2) \to (U_1 \times U_2) \times \mathbb{R}^{k+l}$ for $E_1 \times E_2$. □

9.4 Definition. **The direct sum of vector bundles**

Let E_1 and E_2 be vector bundles over M. Then the **direct sum** $\pi : E_1 \oplus E_2 \to M$ is the bundle with total space

$$E_1 \oplus E_2 := \{(p_1, p_2) \in E_1 \times E_2 \,|\, \pi_1(p_1) = \pi_2(p_2)\}$$

and the projection $\pi : E_1 \oplus E_2 \to M$, $\pi((p_1, p_2)) := \pi_1(p_1) = \pi_2(p_2)$. Hence the fibre $\pi^{-1}(\{x\})$ over $x \in M$ is the direct sum $E_{1,x} \oplus E_{2,x}$. To check that this really defines a vector bundle, note that $\pi : E_1 \oplus E_2 \to M$ is just the restriction of the product $E_1 \times E_2$ to the submanifold $M = \{(x, x) \in M \times M\} \subset M \times M$ and apply lemma 9.3. $\diamond$

9.5 Definition. **Inner product on a vector bundle**

An **inner product** on a vector bundle $\pi : E \to M$ is a smooth map

$$\langle \cdot, \cdot \rangle : E \oplus E \to \mathbb{R}$$

such that its restriction to each fibre $E_x \oplus E_x$ is an inner product, i.e. a positive definite symmetric bilinear form. $\diamond$

9.6 Example. A metric tensor $g \in \mathcal{T}_2^0(M)$ defines an inner product on the tangent bundle TM. $\diamond$

9.7 Definition. **Vector subbundle**

Let $\pi : E \to M$ be a vector bundle and $F \subset E$ a submanifold such that for each $x \in M$ the intersection $F_x := F \cap E_x$ is a p-dimensional linear subspace of the vector space E_x and such that $\pi|_F : F \to M$ defines a rank-p vector bundle. Then $\pi|_F : F \to M$ is called a **subbundle** of E. $\diamond$

9.8 Example. **The tangent bundle of a submanifold**

Let $N \subset M$ be a p-dimensional submanifold of the n-dimensional manifold M. Then the restriction $TM|_N$ is a rank-n vector bundle over N and the tangent bundle TN of N can be naturally seen as a rank-p subbundle of $TM|_N$ as follows. Let $\psi : N \to M$ denote the natural injection, then $D\psi(TN) \subset TM|_N$ is a subbundle. $\diamond$

Often one needs to construct new vector bundles where the total space initially comes without the structure of a differentiable manifold. Then the following construction can be helpful.

9.9 Proposition. **Construction of vector bundles**

Let M be a smooth manifold, E a set, and $\pi : E \to M$ a surjective map. Let $(V_i, \varphi_i)_{i \in I}$ be an atlas for M and suppose we are given bijective maps $\Phi_i : \pi^{-1}(V_i) \to V_i \times \mathbb{R}^k$ satisfying $\pi_1 \circ \Phi_i = \pi$ such that on $V := V_i \cap V_j$ it holds that $\Phi_i \circ \Phi_j^{-1} : V \times \mathbb{R}^k \to V \times \mathbb{R}^k$ is of the form

$$\Phi_i \circ \Phi_j^{-1}((x, y)) = (x, A_{ij}(x)y) \tag{9.1}$$

for some for some smooth map $A_{ij} : V \to GL(k, \mathbb{R})$.

Then E has a unique structure of a rank-k vector bundle over M for which the maps Φ_i are local trivialisations.

Proof. For each $x \in M$ a vector space structure on $E_x := \pi^{-1}(\{x\})$ is defined by choosing $i \in I$ with $x \in V_i$ and declaring the bijective map $\Phi_i|_x : E_x \to \{x\} \times \mathbb{R}^k$ to be a linear isomorphism. This vector space structure is independent of the choice of i since by (9.1) $\Phi_i|_x \circ \Phi_j|_x^{-1} = A(x)$ is an isomorphism of $\mathbb{R}^k$.

Declaring the compositions $(\varphi_i, \mathrm{id}) \circ \Phi_i : \pi^{-1}(V_i) \to \varphi_i(V_i) \times \mathbb{R}^k \subset \mathbb{R}^n \times \mathbb{R}^k$ to be homeomorphism resp. charts, we obtain a topology resp. a differentiable structure on E. Now, by definition, the maps Φ_i are diffeomorphisms and define local trivialisations of the bundle E. $\square$

9.10 Remark. Note that for a given vector bundle E the local trivialisations always satisfy (9.1).

9.11 Definition. The dual bundle

Given a vector bundle E over M one can define the **dual bundle** E^* with total space

$$E^* = \{(x, v) \mid x \in M\,,\ v \in E_x^*\}$$

and projection $\tilde{\pi}((x, v)) = x$, where E_x^* denotes the dual space of the fibre E_x of E over x. Given a local trivialisation $\Phi : \pi^{-1}(U) \to U \times \mathbb{R}^k$ of E, let

$$\Psi : \tilde{\pi}^{-1}(U) \to \mathbb{R}^k\,, \quad \Psi((x, v)) := (x, w) \quad \text{with} \quad w = ((\Phi|_x)^\dagger)^{-1} v\,,$$

where $(\Phi|_x)^\dagger : (\mathbb{R}^k)^* \cong \mathbb{R}^k \to E_x^*$ denotes the adjoint map defined by

$$((\Phi|_x)^\dagger(w) \mid v) := (w \mid \Phi|_x(v))\,.$$

For two such maps Ψ_1 and Ψ_2 we have that

$$(\Psi_1 \circ \Psi_2^{-1})(x, w) = (\Phi_1^\dagger)^{-1}|_x\, \Phi_2^\dagger|_x w = (\Phi_2|_x\, \Phi_1^{-1}|_x)^\dagger w = A^T(x) w$$

where $\Phi_2|_x \Phi_1^{-1}|_x y =: A(x) y$. Hence, according to porposition 9.9, the Ψ_i define local trivialisations for a unique vector bundle E^*. ◇

9.12 Example. The cotangent bundle T^*M is the dual bundle of the tangent bundle TM. ◇

9.13 Definition. The tensor product of vector bundles

Let E and F be vector bundles over M. Then the **tensor product bundle** $\pi : E \otimes F \to M$ is the bundle with fibre $E_x \otimes F_x$ over $x \in M$, i.e. with total space

$$E \otimes F = \{(x, v) \mid x \in M\,,\ v \in E_x \otimes F_x\}$$

and projection $\pi((x, v)) = x$. We can again apply proposition 9.9. Choose local trivialisations $\Phi_{E,i} : \pi_E^{-1}(V_i) \to V_i \times \mathbb{R}^n$ and $\Phi_{F,i} : \pi_F^{-1}(V_i) \to V_i \times \mathbb{R}^m$ for each domain $V_i \subset M$ of an appropriate atlas. Then the fibre-wise tensor product maps

$$\Psi_i := \Phi_{E,i} \otimes \Phi_{F,i} : \pi_1^{-1}(V_i) \otimes \pi_2^{-1}(V_i) \to V_i \times (\mathbb{R}^n \otimes \mathbb{R}^m)$$

are again fibre-wise and $\Psi_i \circ \Psi_j^{-1}$ are fibre-wise isomorphisms and thus satisfy (9.1). ◇

9.14 Example. The tensor bundles $T_s^r M$ arise as the tensor products of the tangent and the cotangent bundles TM and T^*M,

$$T_s^r M = \underbrace{TM \otimes \cdots \otimes TM}_{r \text{ copies}} \otimes \underbrace{T^*M \otimes \cdots \otimes T^*M}_{s \text{ copies}}\,.$$

9.15 Definition. Sections and frames

Let $\pi : E \to M$ be a rank-k vector bundle.

(a) A smooth function $S : M \to E$ with $\pi \circ S = \mathrm{id}_M$ is called a **section** (or **global section**) of E. A smooth function $S : M \supset U \to E$ defined on an open set $U \subset M$ with $\pi \circ S = \mathrm{id}_U$ is called a **local section** of E. We write $\Gamma(E)$ for the space of global sections of E and $\Gamma(E|_U)$ for the space of local sections on $U \subset M$.

(b) A family of k local sections $(S_1, \ldots, S_k)$ such that $(S_1(x), \ldots, S_k(x))$ is a basis of E_x for each $x \in U$ is called a **local frame**. If $U = M$, then $(S_1, \ldots, S_k)$ is called a **global frame**. The sections S_j in a frame are sometimes called **basis sections**.

9.16 Examples. (a) Vector fields on a manifold M are global sections of the tangent bundle, i.e. $\Gamma(TM) = \mathcal{T}^1_0(M)$.

(b) More generally, tensor fields are global sections of the tensor bundles, i.e. $\Gamma(T^r_s M) = \mathcal{T}^r_s(M)$.

(c) Functions $f \in C^\infty(M)$ are sections of the trivial bundle $M \times \mathbb{R}$.

(d) For any vector bundle $\pi : E \to M$ the zero-section $S(x) = 0$ for all $x \in M$ is an embedding of M into E.

(e) Any chart on a manifold M yields local frames $(\partial_{q_1}, \ldots, \partial_{q_n})$ and $(\mathrm{d}q^1, \ldots, \mathrm{d}q^n)$ of the tangent bundle TM resp. the cotangent bundle T^*M, so called **coordinate frames**.

9.17 Remark. Given a local frame $(S_1, \ldots, S_k)$ of a vector bundle E, any section $Y \in \Gamma(E)$ can be written as a linear combination of the elements of the frame since the latter form a basis of E_x at each point $x \in U \subset M$. We employ again the Einstein summation convention and write

$$Y = Y^j S_j \qquad \text{on } U\,,$$

where $Y^j \in C^\infty(U)$ for $j = 1, \ldots, k$. ◇

9.18 Proposition. A vector bundle $\pi : E \to M$ is trivialisable if and only if it admits a global frame.

Proof. Let $\Phi : E \to M \times \mathbb{R}^k$ be a global trivialisation and $(e_1, \ldots, e_k)$ the canonical basis of $\mathbb{R}^k$. Then $(S_1(x), \ldots, S_k(x)) := (\Phi|_x^{-1} e_1, \ldots, \Phi|_x^{-1} e_k)$ is a global frame.
Let conversely $(S_1, \cdots, S_k)$ be a global frame for E. Then

$$\Phi : E \to M \times \mathbb{R}^k\,, \quad (x, v) =: (x, v^i S_i(x)) \mapsto (x, (v^1, \ldots, v^k))$$

is a global trivialisation. □

9.19 Example. The tangent bundle TS^2 of the two-sphere S^2 is not trivialisable since, as was shown as a homework assignment, every smooth vector field on S^2 has at least one zero. Thus, no global frame for TS^2 can exist.

9.20 Definition. **The endomorphism bundle**
Let $\pi : E \to M$ be a vector bundle. Then the **endomorphism bundle** $\mathrm{End}(E)$ is the vector bundle over M whose fibre over $x \in M$ is $\mathrm{End}(E)_x = \mathcal{L}(E_x)$, i.e. the space of linear endomorphisms of the vector space E_x. The local trivializations $\Phi : \pi^{-1}(U) \to U \times \mathbb{R}^{k^2}$ are defined using the matrix representation with respect to local frames $(S_1, \ldots, S_k)$.
Since for any vector space V the space of endomorphisms $\mathcal{L}(V)$ is isomorphic to $V \otimes V^*$, we have that $\mathrm{End}(E) \cong E \otimes E^*$. ◇

German nomenclature

fibre = Faser	frame = Rahmen
section = Schnitt	subbundle = Unterbündel
total space = Totalraum	trivial = trivial
vector bundle = Vektorbündel	

10 Connections on vector bundles

Loosely speaking, a connection provides a way of taking directional derivatives of a section of a vector bundle. In particular, it allows to define the concept of "constant" or "parallel" sections and, along curves, to define "parallel transport".

10.1 Definition. **<u>Connections on vector bundles</u>**
Let $\pi : E \to M$ be a vector bundle and $\Gamma(E)$ the space of its smooth sections. A **connection** ∇ in E is a map

$$\nabla : \mathcal{T}_0^1(M) \times \Gamma(E) \to \Gamma(E)\,, \qquad (X, S) \mapsto \nabla_X S\,,$$

with the following properties:

(a) $X \mapsto \nabla_X S$ is linear over $C^\infty(M)$, i.e.

$$\nabla_{fX+gY} S = f\,\nabla_X S \;+\; g\,\nabla_Y S$$

for all $f, g \in C^\infty(M)$, $X, Y \in \mathcal{T}_0^1(M)$, and $S \in \Gamma(E)$.

(b) $S \mapsto \nabla_X S$ is $\mathbb{R}$-linear, i.e.

$$\nabla_X(\alpha S + \beta \tilde{S}) = \alpha\,\nabla_X S \;+\; \beta\,\nabla_X \tilde{S}$$

for all $\alpha, \beta \in \mathbb{R}$, $X \in \mathcal{T}_0^1(M)$, and $S, \tilde{S} \in \Gamma(E)$.

(c) ∇ satisfies the product rule

$$\nabla_X(fS) = (\mathrm{d}f|X)\,S + f\,\nabla_X S$$

for all $f \in C^\infty(M)$, $X \in \mathcal{T}_0^1(M)$, and $S \in \Gamma(E)$.

The section $\nabla_X S \in \Gamma(E)$ is called the **covariant derivative** of S in the direction of X. If $\nabla_X S = 0$ for all $X \in \mathcal{T}_0^1(M)$ then S is called **parallel** or **constant** with respect to ∇. $\diamond$

10.2 Example. **<u>The trivial connection on a trivial bundle</u>**
Let $E = M \times V$ be a trivial vector bundle. Then sections $S : M \to M \times V$, $x \mapsto S(x) =: (x, s(x))$ are in one-to-one correspondence with smooth functions $s : M \to V$, $x \mapsto s(x)$. But for such functions we can define the derivative in the direction of a tangent vector $v \in T_xM$ as in the case of real-valued functions (cf. remark 2.4) pointwise as

$$D_v(s) := \frac{\mathrm{d}}{\mathrm{d}t}(s \circ c_v)(t)|_{t=0} \qquad \text{for } v \in T_xM \text{ with } v = [c_v]_x.$$

Defining $(\nabla_X S)(x) := D_{X(x)}(s)$ we obtain the so called **trivial connection** on the trivial bundle $E = M \times V$. For $V = \mathbb{R}$ this is just the action of vector fields on functions, i.e. $\nabla_X S = (\mathrm{d}s|X)$, or equivalently but with different notation, $\nabla_X S = L_X s$. The properties of the Lie derivative acting on functions were collected in proposition 2.31 and show that ∇_X defines indeed a connection for this case. But the same argument leading to proposition 2.31 shows that the trivial connection is a connection also for vector valued functions. $\diamond$

10.3 Example. As a special case of example 10.2 we consider the exterior derivative on functions as a connection on the trivial bundle $M \times \mathbb{R}$, i.e. for $f \in C^\infty(M) = \Gamma(M \times \mathbb{R})$ set

$$\mathrm{d}_X f := (\mathrm{d}f | X) .$$

Given any 1-form $\omega \in \mathcal{T}_1^0(M)$, also the map

$$\nabla_X f := \mathrm{d}_X f + \omega(X) f$$

defines a connection on $M \times \mathbb{R}$, denoted by $\mathrm{d} + \omega$. To see this, note that properties (a) and (b) of definition 10.1 are clearly satisfied. The product rule is an easy computation:

$$\nabla_X(fg) = (\mathrm{d}(fg)|X) + \omega(X)fg = (\mathrm{d}f|X)g + f(\mathrm{d}g|X) + f\omega(X)g = (\mathrm{d}f|X)g + f\nabla_X g .$$

Actually, as we will show later on, the set of all connections on $C^\infty(M)$ is given by

$$\{\mathrm{d} + \omega \,|\, \omega \in \Lambda_1(M)\} . \qquad \diamond$$

Although this is not completely obvious from the definition, the value $\nabla_X S(x)$ of $\nabla_X S$ at a point x depends only on the value $X(x)$ of X at x. So $\nabla_X S(x)$ is really a directional derivative and, in contrast to the Lie derivative, not a derivative along a vector field.

10.4 Proposition. **Locality of the covariant derivative**

Let ∇ be a connection on a vector bundle $\pi : E \to M$. Let $X, \tilde{X} \in \mathcal{T}_0^1(M)$ and $S, \tilde{S} \in \Gamma(E)$ be such that for some $x \in M$ there exists an open neighbourhood $U \subset M$ of x with

$$S|_U = \tilde{S}|_U \quad \text{and} \quad X(x) = \tilde{X}(x) .$$

Then $\nabla_X S(x) = \nabla_{\tilde{X}} \tilde{S}(x)$.

Proof. By linearity it suffices to show that $\nabla_X S(x) = 0$ whenever $X(x) = 0$ or $S|_U = 0$.
First assume that $S|_U = 0$ and $x \in U$. Now choose a bump function $f \in C_0^\infty(M)$ with $\mathrm{supp} f \subset U$ and $f(x) = 1$. Then $fS = 0$ on all of M and therefore

$$\nabla_X(fS) = \nabla_X(0 \cdot fS) = 0 \cdot \nabla_X(fS) = 0 .$$

On the other hand, by the product rule,

$$\nabla_X(fS) = (\mathrm{d}f \,|\, X)S + f\nabla_X S = f\nabla_X S ,$$

since in the first factor either $\mathrm{d}f$ or S vanishes. Hence, $\nabla_X S(x) = f(x)\nabla_X S(x) = 0$.
Now let $X|_U = 0$. Then $fX = 0$ on all of M and analogously

$$f\nabla_X S = \nabla_{fX} S = \nabla_{0 \cdot fX} S = 0 \cdot \nabla_{fX} S = 0$$

implies $\nabla_X S(x) = f(x)\nabla_X S(x) = 0$.
Up to now we showed that $\nabla_X S(x)$ depends only on S and X in a neighbourhood of U of x. Hence, we can compute $\nabla_X S(x)$ using a local coordinate frame $(\partial_1, \ldots, \partial_n)$ on U. Let $X = X^i \partial_i$, then

$$\nabla_X S(x) = \nabla_{X^i \partial_i} S(x) = X^i(x) \nabla_{\partial_i} S(x) .$$

Hence $\nabla_X S(x) = 0$ whenever $X(x) = 0$. $\square$

We emphasise once more that the preceding proposition tells us that $\nabla_X S(x)$ depends only on S in a neighbourhood of x and on $X(x)$. Hence, we can compute $\nabla_X S(x)$ using a local frame.

10.5 Definition. Christoffel symbols
Let $(Z_1, \ldots, Z_n)$ be a local frame for TM, $(S_1, \ldots, S_k)$ a local frame (on the same set $U \subset M$) for a vector bundle $\pi : E \to M$, and ∇ a connection on E. Then $\nabla_{Z_i} S_\beta \in \Gamma(E|_U)$ can again be represented in terms of the basis sections S_α and the coefficients $\Gamma^\alpha_{i\beta} \in C^\infty(U)$ in

$$\nabla_{Z_i} S_\beta =: \Gamma^\alpha_{i\beta} S_\alpha$$

are called **Christoffel symbols** of ∇ with respect to these frames. $\diamond$

10.6 Lemma. A connection is determined by its Christoffel symbols
Let $(Z_1, \ldots, Z_n)$ be a local frame for the tangent bundle $TU \subset TM$ of a manifold M, $(S_1, \ldots, S_k)$ a local frame (on the same set $U \subset M$) for a vector bundle $\pi : E \to M$, ∇ a connection on E, and $\Gamma^\alpha_{i\beta}$ its Christoffel symbols with respect to these frames. Let $X \in \mathcal{T}^1_0(M)$ and $Y \in \Gamma(E)$ be locally represented as $X = X^j Z_j$ and $Y = Y^\alpha S_\alpha$. Then, on U,

$$\nabla_X Y = \Big((\mathrm{d}Y^\alpha | X) + X^i Y^\beta \Gamma^\alpha_{i\beta} \Big) S_\alpha \,. \tag{10.1}$$

Proof. The proof is a simple computation using the defining properties of a connection:

$$\begin{aligned}
\nabla_X Y &= \nabla_X \Big(\sum_\alpha Y^\alpha S_\alpha \Big) \overset{10.1(b)}{=} \sum_\alpha \nabla_X (Y^\alpha S_\alpha) \\
&\overset{10.1(c)}{=} \sum_\alpha \Big((\mathrm{d}Y^\alpha | X) S_\alpha + Y^\alpha \nabla_{\sum_i X^i Z_i} S_\alpha \Big) \\
&\overset{10.1(a)}{=} \sum_\alpha \Big((\mathrm{d}Y^\alpha | X) S_\alpha + Y^\alpha \sum_i X^i \nabla_{Z_i} S_\alpha \Big) = \sum_\alpha \Big((\mathrm{d}Y^\alpha | X) S_\alpha + Y^\alpha \sum_{i,\beta} X^i \Gamma^\beta_{i\alpha} S_\beta \Big) .
\end{aligned}$$

We did not use the Einstein convention here to highlight that interchanging the summation with the action of ∇ is indeed justified. From now on we will take this fact for granted again and use the Einstein summation convention again. $\square$

Conversely, given local frames $(Z_1, \ldots, Z_n)$ for TM and $(S_1, \ldots, S_k)$ for E on $U \subset M$, any choice of $n \cdot k^2$ smooth functions $\Gamma^\alpha_{i\beta} \in C^\infty(U)$, $i = 1, \ldots, n$, $\alpha, \beta = 1, \ldots, k$ determines a connection by the formula (10.1).

10.7 Lemma. Let $(Z_1, \ldots, Z_n)$ be a local frame for the tangent bundle $TU \subset TM$ of a manifold M and $(S_1, \ldots, S_k)$ a local frame (on the same set $U \subset M$) for a vector bundle $\pi : E \to M$. Then for any choice of $\Gamma^\alpha_{i\beta} \in C^\infty(U)$, $i = 1, \ldots, n$, $\alpha, \beta = 1, \ldots, k$, the expression

$$\nabla_X Y := \Big((\mathrm{d}Y^\alpha | X) + X^i Y^\beta \Gamma^\alpha_{i\beta} \Big) S_\alpha$$

defines a connection on $E|_U$, where $X \in \mathcal{T}^1_0(M)$ and $Y \in \Gamma(E)$ with $X = X^j Z_j$ and $Y = Y^\alpha S_\alpha$.

Proof. The properties (a) and (b) of definition 10.1 are clearly satisfied. And the product rule follows from the product rule for the differential: for $f \in C^\infty(U)$ we have

$$\begin{aligned}
\nabla_X (fY) &= \Big((\mathrm{d}(fY^\alpha) | X) + f X^i Y^\beta \Gamma^k_{i\beta} \Big) S_\alpha \\
&= \Big((\mathrm{d}f | X) Y^\alpha + f (\mathrm{d}Y^\alpha | X) + f X^i Y^\beta \Gamma^k_{i\beta} \Big) S_\alpha \\
&= (\mathrm{d}f | X) Y + f \, \nabla_X Y \,.
\end{aligned}$$

$\square$

10.8 Lemma. **Tensor characterisation lemma**

Let $\pi_j : E_j \to M$, $j = 1, \dots, m$, be vector bundles over M. A map

$$\tau : \Gamma(E_1) \times \cdots \times \Gamma(E_m) \;\to\; C^\infty(M)$$

is induced by a section $F \in \Gamma(E_1^* \otimes \cdots \otimes E_m^*)$ in the sense that

$$\tau(Y_1, \dots, Y_m)(x) = (F(x) \,|\, Y_1(x), \dots, Y_m(x)) \tag{10.2}$$

if and only if it is $C^\infty(M)$-linear in all arguments, i.e.

$$\tau(Y_1, \dots, Y_j + f\tilde{Y}, \dots, Y_m) = \tau(Y_1, \dots, Y_m) + f\,\tau(Y_1, \dots, \tilde{Y}, \dots, Y_m)$$

for all $f \in C^\infty(M)$, $Y_j \in \Gamma(E_j)$, and $j = 1, \dots, m$.

Proof. Let $F \in \Gamma(E_1^* \otimes \cdots \otimes E_m^*)$ be given. Then $F(x) \in E_{1,x}^* \otimes \cdots \otimes E_{m,x}^*$ and thus the map τ defined by (10.2) is $C^\infty(M)$-linear in all arguments. On the other hand, for a given $C^\infty(M)$-multi-linear map τ, the map

$$F(x) : E_{1,x} \times \cdots \times E_{m,x} \to \mathbb{R}\,, \quad (v_1, \dots, v_m) \mapsto \tau(Y_1, \dots, Y_m)(x)$$

for any choice of $Y_j \in \Gamma(E_j)$ with $Y_j(x) = v_j$ is well defined and multi-linear: By a bump function argument completely analogous to the one in the proof of proposition 10.4 one can show that $\tau(Y_1, \dots, Y_m)(x)$ depends only on the Y_j in a neighbourhood of x. Then, using local sections we see that for $Y_j = Y_j^i S_{j,i}$

$$\tau(Y_1, \dots, Y_m)(x) = \tau(Y_1^{i_1} S_{1,i_1}, \dots, Y_m^{i_m} S_{m,i_m})(x) = Y_1^{i_1}(x) \cdots Y_m^{i_m}(x) \tau(S_{1,i_1}, \dots, S_{m,i_m})(x)$$

depends only on the values of Y_i at x. □

10.9 Definition. **The total covariant derivative**

Let ∇ be a connection on E and $Y \in \Gamma(E)$. Then $\nabla Y \in \Gamma(T^*M \otimes E)$ defined by the $C^\infty(M)$-bilinear map

$$\nabla Y : \Gamma(E^*) \times \mathcal{T}_0^1(M) \;\to\; C^\infty(M)\,, \qquad (S, X) \mapsto \nabla Y(S, X) := (\nabla_X Y | S)$$

is called the **total covariant derivative** of Y. ◇

10.10 Notation. When one writes the components of a total covariant derivative ∇Y of a section $Y \in \Gamma(E)$, one uses a semicolon to separate the index that results from differentiation: for $Y = Y^i\, S_i$ one writes

$$\nabla Y = Y^i_{;k}\;\, S_i \otimes \mathrm{d}q^k \in \Gamma(E \otimes T^*M)\,.$$

10.11 Proposition. **The difference of two connections**

Let ∇ and $\tilde{\nabla}$ be connections on a vector bundle $\pi : E \to M$. Then there exists a unique endomorphism-valued 1-form ω, i.e. a section of the vector bundle $T^*M \otimes \mathrm{End}(E)$, such that

$$\nabla_X Y - \tilde{\nabla}_X Y = \omega(X) Y\,,$$

where $(\omega(X)Y)(x) := \omega(X)(x)Y(x)$ denotes the pointwise action of the linear map $\omega(X)(x) \in \mathcal{L}(E_x)$ on $Y(x) \in E_x$.

Conversely, given a connection ∇ and an $\omega \in \Gamma(T^*M \otimes \mathrm{End}(E))$, then the map

$$\nabla + \omega : \mathcal{T}_0^1(M) \times \Gamma(E) \to \Gamma(E)\,, \quad (X, Y) \mapsto \nabla_X Y + \omega(X) Y$$

defines a connection on E.

Proof. By the tensor characterisation lemma we need to show that the map

$$\nabla - \tilde{\nabla} : \mathcal{T}_0^1(M) \times \Gamma(E) \times \Gamma(E^*) \to C^\infty(M)\,, \quad (X, Y, Z) \mapsto \Big(\nabla_X Y - \tilde{\nabla}_X Y \,|\, Z\Big)$$

is $C^\infty(M)$-linear in all arguments. This holds by definition for X and Z and for Y it follows from the product rule,

$$\Big(\nabla_X(fY) - \tilde{\nabla}_X(fY) \,|\, Z\Big) = \Big((\mathrm{d}f|X)Y + f\nabla_X Y - (\mathrm{d}f|X)Y - f\tilde{\nabla}_X Y \,|\, Z\Big) = f\,\Big(\nabla_X Y - \tilde{\nabla}_X Y \,|\, Z\Big)\,.$$

To see that $\nabla + \omega$ is a connection for every $\omega \in \Gamma(T^*M \otimes \mathrm{End}(E))$, note that (a) and (b) of definition 10.1 are obvious, and the product rule (c) follows easily,

$$\nabla_X(fY) + \omega(X)fY = (\mathrm{d}f|X)Y + f\nabla_X Y + f\omega(X)Y = (\mathrm{d}f|X)Y + f(\nabla_X Y + \omega(X)Y)\,. \quad \square$$

10.12 Remark. Proposition 10.11 shows that the set of connections on a vector bundle E is not a linear set. For example, if ∇ and $\tilde{\nabla}$ are connections, then neither $\nabla + \tilde{\nabla}$ nor $\frac{1}{2}\nabla$ is a connection, since neither $\nabla + \tilde{\nabla} - \nabla = \tilde{\nabla}$ nor $\frac{1}{2}\nabla - \nabla = -\frac{1}{2}\nabla$ are endomorphism-valued 1-forms. Alternatively one can check directly that neither of them satisfies the product rule.

Instead, given a connection ∇ on E, the set of connections is

$$\{\nabla + \omega \,|\, \omega \in \Gamma(T^*M \otimes \mathrm{End}(E))\}\,,$$

i.e. it is an affine space over the vector space $\Gamma(T^*M \otimes \mathrm{End}(E))$. ◇

In the following chapters we will be mainly concerned with connections on the tangent bundle.

10.13 Definition. Affine connections

A connection

$$\nabla : \mathcal{T}_0^1(M) \times \Gamma(TM) \to \Gamma(TM)$$

on the tangent bundle TM of a smooth manifold M is called an **affine connection** or **linear connection** or just a **connection on** M. (Note that the terminology in the literature is not at all uniform in this respect.) We used the different notations for the same object $\Gamma(TM) = \mathcal{T}_0^1(M)$ in order to emphasise the different roles played by the two factors. ◇

10.14 Remark. Note once more that an affine connection ∇ is **not** a tensor field of type $(1, 2)$ since it is not linear over $C^\infty(M)$ in the second argument, but instead satisfies the product rule. However, given a connection ∇ and a tensor field $A \in \mathcal{T}_2^1(M) = \Gamma(T^*M \otimes \mathrm{End}(TM))$, also

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

is a connection on M. Actually, by proposition 10.11, the set of all affine connections is precisely the affine space $\{\nabla + A \,|\, A \in \mathcal{T}_2^1(M)\}$ over the vector space $\mathcal{T}_2^1(M)$. ◇

10.15 Definition. Christoffel symbols for an affine connection

For an affine connection ∇ it suffices to choose a local frame $(Z_1, \dots, Z_n)$ of the tangent bundle $TU \subset TM$ in order to define the Christoffel symbols

$$\nabla_{Z_i} Z_j =: \Gamma_{ij}^k Z_k\,.$$

Note that a chart on $U \subset M$ provides a local coordinate frame $(\partial_1, \dots, \partial_n)$ and thus corresponding Cristoffel symbols $\nabla_{\partial_i}\partial_j = \Gamma_{ij}^k \partial_k$. ◇

10.16 Proposition. Every manifold admits an affine connection.

Proof. Let $\mathcal{A} = (V_\alpha, \varphi_\alpha)$ be an atlas for M and (χ_α) an adapted partition of unity. Then on each coordinate patch V_α we can choose an arbitrary family of smooth functions $\Gamma^k_{\alpha,ij} \in C^\infty(V_\alpha)$ in order to define a connection ∇^α with the help of the frame $(\partial_{q_1^\alpha}, \ldots, \partial_{q_n^\alpha})$ of TV_α as in lemma 10.7. Now we define

$$\nabla_X Y := \sum_\alpha \chi_\alpha \nabla^\alpha_X Y \,.$$

Again it is obvious that properties (a) and (b) of definition 10.1 are inherited from the corresponding properties of ∇^α. For the product rule we find that

$$\nabla_X(fY) = \sum_\alpha \chi_\alpha \nabla^\alpha_X(fY) = \sum_\alpha \chi_\alpha((\mathrm{d}f|X)Y + f\,\nabla^\alpha_X Y) = (\mathrm{d}f|X)Y + f\,\nabla_X Y \,. \qquad \square$$

A connection on the tangent bundle TM of a manifold M can be canonically lifted to a connection on all tensor bundles $T^r_s M$.

10.17 Proposition. **The lift of an affine connection to tensor bundles**
Given be an affine connection $\widetilde{\nabla}$ on M, there exists a unique family of connections $\nabla = (\nabla^r_s)$ on the tensor bundles $T^r_s M$ with the following properties:

(a) On $T^0_0 M$, $\nabla_X f = \mathrm{d}_X f := (\mathrm{d}f|X)$.

(b) On TM, $\nabla = \widetilde{\nabla}$.

(c) ∇ obeys the product rule with respect to tensor products,

$$\nabla_X(F \otimes G) = (\nabla_X F) \otimes G + F \otimes (\nabla_X G) \,,$$

where F and G are tensor fields of arbitrary type.

(d) ∇ commutes with all contractions: if "tr" denotes the contraction of any pair of indices (one upper and one lower), then

$$\nabla_X(\mathrm{tr}\, F) = \mathrm{tr}(\nabla_X F) \,.$$

This connection satisfies the following additional properties:

(i) Let $\omega \in \Lambda_1(M)$ and $Y \in \mathcal{T}^1_0(M)$. Then

$$(\nabla_X \omega)(Y) = \mathrm{d}_X\, \omega(Y) - \omega(\nabla_X Y) \,.$$

(ii) For any tensor $F \in \mathcal{T}^r_s(M)$, vector fields $Y_j \in \mathcal{T}^1_0(M)$, and 1-forms $\omega^i \in \Lambda_1(M)$, it holds that

$$\begin{aligned}(\nabla_X F)(\omega^1, \ldots, \omega^r, Y_1, \ldots, Y_s) \;&=\; \mathrm{d}_X F(\omega^1, \ldots, \omega^r, Y_1, \ldots, Y_s) \\ &\quad - \sum_{i=1}^r F(\omega^1, \ldots, \nabla_X \omega^i, \ldots, \omega^r, Y_1, \ldots, Y_s) \\ &\quad - \sum_{j=1}^s F(\omega^1, \ldots, \omega^r, Y_1, \ldots, \nabla_X Y_j, \ldots, Y_s) \,.\end{aligned}$$

Proof. We only sketch the strategy of the proof. By (a) and (b) the action of ∇ on functions and vector fields is fixed. Now we use (i) to define the action of ∇ on 1-forms, $(\nabla_X \omega \,|\, Y) := \mathrm{d}_X(\omega \,|\, Y) - (\omega \,|\, \nabla_X Y)$, and then (ii) to define the action of ∇ on arbitrary tensor fields. Now one needs to check that ∇ really defines a connection on each tensor bundle and that the properties (c) and (d) are also satisfied. Uniqueness follows, finally, by showing that the properties (a)–(d) imply (i) and (ii). $\square$

10.18 Remark. Let ∇ be an affine connection. Then in a local coordinate chart its action on 1-forms is given by

$$\nabla_X \omega = \left(X^i \partial_i \omega_j - X^i \omega_k \Gamma^k_{ij} \right) \mathrm{d}q^j \,.$$

Proof. Homework assignment. □

10.19 Lemma. Let ∇ be an affine connection and $F \in \mathcal{T}^r_s(M)$. Then the components of its total covariant derivative ∇F with respect to a local coordinate basis are

$$F^{j_1 \dots j_r}_{i_1 \dots i_s ; k} = \partial_k F^{j_1 \dots j_r}_{i_1 \dots i_s} + \sum_{m=1}^{r} F^{j_1 \dots p \dots j_r}_{i_1 \dots i_s} \Gamma^{j_m}_{kp} - \sum_{m=1}^{s} F^{j_1 \dots j_r}_{i_1 \dots p \dots i_s} \Gamma^p_{k i_m} \,.$$

Proof. This follows from lemma 10.6, proposition 10.17 (ii), and remark 10.18. □

While in general there need not exist parallel sections for a given connection (except from the zero section), one can always construct parallel sections along curves by so called parallel transport. In order to define parallel transport we need to introduce some more terminology.

10.20 Definition. <u>**Sections along curves**</u>

Let $\pi : E \to M$ be a vector bundle, $I \subset \mathbb{R}$ an interval, and $u : I \to M$ a smooth curve. Then a smooth map $Y : I \to E$ such that $\pi \circ Y = u$ is called a **section along** u. If there exists a section $\tilde{Y} \in \Gamma(E)$ such that $Y = \tilde{Y} \circ u$, then Y is called **extendible**.

The space of sections along u is denoted by $\Gamma(u)$. For the space of (r, s)-tensor fields along a curve u we write $\mathcal{T}^r_s(u)$. ◇

10.21 Example. Any section $\tilde{Y} \in \Gamma(E)$ defines a section Y along u through $Y := \tilde{Y} \circ u$. By definition, such a Y is extendible. ◇

10.22 Example. <u>**The velocity vector field along a curve**</u>

Let $u : I \to M$ be a smooth curve. Then the velocity vector field $\dot{u} := Du \circ e$ is a vector field along u, cf. definition 7.1. If u has self-intersections then, typically, $\dot{u}$ is not extendible, as u can pass through a self-intersection with different velocities at different times. ◇

Let ∇ be a connection on the vector bundle $\pi : E \to M$ and $u : I \to M$ a smooth curve. For any extendible section Y along u we can define its covariant derivative along u by choosing an extension $\tilde{Y}$ with $Y = \tilde{Y} \circ u$ and setting

$$D_t Y : I \to E \,, \quad t \mapsto (\nabla_{\dot{u}(t)} \tilde{Y})(u(t)) \,,$$

i.e. by taking at each point $u(t)$ on the curve the covariant derivative of $\tilde{Y}$ in the direction $\dot{u}(t)$. Intuitively we expect that this derivative does not depend on the choice of the extension $\tilde{Y}$ and that a similar map should also exist for non-extendible sections along u.

10.23 Proposition. <u>**Covariant derivative along a curve**</u>

Let ∇ be a connection on the vector bundle $\pi : E \to M$ and $u : I \to M$ a smooth curve. Then there exists a unique operator

$$D_t : \Gamma(u) \to \Gamma(u)$$

satisfying the following properties:

(a) D_t is $\mathbb{R}$-linear, i.e. $D_t(aX + bY) = aD_t X + bD_t Y$ for all $a, b \in \mathbb{R}$ and $X, Y \in \Gamma(u)$.

(b) D_t satisfies the product rule: $D_t(fY) = \dot{f}Y + fD_t Y$ for all $f \in C^\infty(I)$ and $Y \in \Gamma(u)$.

(c) For extendible $Y \in \Gamma(u)$ and any extension $\tilde{Y}$ of Y it holds that $(D_t Y)(t) = (\nabla_{\dot{u}(t)} \tilde{Y})(u(t))$.
$D_t Y$ is called the **covariant derivative of Y along u**.

Proof. We only sketch the proof. First assume that a map D_t with the claimed properties exists. Using the product rule (b) one concludes in the usual way that $D_t Y(t_0)$ depends only on the values of Y in a neighbourhood of t_0. The properties (a), (b), and (c) imply that with respect to a local frame $(\tilde{S}_1, \ldots, \tilde{S}_k)$ defined on a neighbourhood U of $u(t)$ we have with $Y = Y^\alpha S_\alpha$ and $S_\alpha := \tilde{S}_\alpha \circ u$ that

$$\begin{aligned} D_t Y(t) &\overset{(a),(b)}{=} \dot{Y}^\alpha(t) S_\alpha(t) + Y^\alpha(t) D_t S_\alpha(t) \\ &\overset{(c)}{=} \dot{Y}^\alpha(t) \tilde{S}_\alpha(u(t)) + Y^\alpha(t) (\nabla_{\dot{u}(t)} \tilde{S}_\alpha)(u(t)) \\ &= \Big(\dot{Y}^\alpha(t) + Y^\beta(t) \dot{u}^i(t) \Gamma^\alpha_{i\beta}(u(t)) \Big) \tilde{S}_\alpha(u(t)) \,. \end{aligned} \tag{10.3}$$

This proves, in particular, uniqueness. To prove existence, we use the coordinate expression (10.3) as a local definition of D_t and need to check that it really satisfies (a), (b), and (c). For (a) and (b) this is obvious. For (c) we just revert the above computation and find with (10.1) that for any extension $\tilde{Y}$ of Y we have

$$\begin{aligned} (\nabla_{\dot{u}(t)} \tilde{Y})(u(t)) &\overset{(10.1)}{=} \Big((\mathrm{d}\tilde{Y}^\alpha(u(t)) \,|\, \dot{u}(t)) + \tilde{Y}^\beta(u(t)) \dot{u}^i(t) \Gamma^\alpha_{i\beta}(u(t)) \Big) \tilde{S}_\alpha(u(t)) \\ &= \Big(\tfrac{\mathrm{d}}{\mathrm{d}s} \tilde{Y}^\alpha(u(s))|_{s=t} + Y^\beta(t) \dot{u}^i(t) \Gamma^\alpha_{i\beta}(u(t)) \Big) \tilde{S}_\alpha(u(t)) \\ &= \Big(\tfrac{\mathrm{d}}{\mathrm{d}s} Y^\alpha(s)|_{s=t} + Y^\beta(t) \dot{u}^i(t) \Gamma^\alpha_{i\beta}(u(t)) \Big) \tilde{S}_\alpha(u(t)) = D_t Y(t) \,. \end{aligned}$$

Hence we have local existence of D_t on sets where a frame exists. By uniqueness the local expressions agree on intersections of local domains and thus we have also global existence. □

10.24 Lemma. D_t and reparametrisations of curves

Let ∇ be a connection on the vector bundle $\pi : E \to M$, $u : I \to M$ a smooth curve, and $Y \in \Gamma(u)$. For a diffeomorphism $\Phi : \tilde{I} \to I$ of open intervals let $\tilde{u} := u \circ \Phi$ be the reparametrised curve $\tilde{u} : \tilde{I} \to M$ and $\tilde{Y} := Y \circ \Phi \in \Gamma(\tilde{u})$. Then

$$D_t \tilde{Y} = \Phi' \cdot D_t Y \circ \Phi \,.$$

Proof. This follows from the observation that in (10.3) we have $\frac{\mathrm{d}}{\mathrm{d}s} \tilde{Y}^\alpha(s) = \dot{Y}^\alpha(\Phi(s))\, \Phi'(s)$ and also $\frac{\mathrm{d}}{\mathrm{d}s} \tilde{u}^i(s) = \dot{u}^i(\Phi(s))\, \Phi'(s)$. □

10.25 Definition. Parallel sections

A section Y along a curve $u : I \to M$ of a vector bundle $\pi : E \to M$ is called **parallel** with respect to a connection ∇ on E, if

$$D_t Y \equiv 0 \,.$$

10.26 Proposition. Parallel transport

Let ∇ be a connection on a vector bundle $\pi : E \to M$ and $u : I \to M$ a smooth curve.

Let $t_0 \in I$ and $Y_0 \in E_{u(t_0)}$. Then there exists a unique parallel section Y along u with $Y(t_0) = Y_0$. The map $T_{t,t_0} : E_{u(t_0)} \to E_{u(t)}$, $Y_0 \mapsto T_{t,t_0} Y_0 := Y(t)$, is a linear isomorphism and called the **parallel transport** map.

The parallel transport does not depend on the parametrisation of the curve: Let $\Phi : \tilde{I} \to I$ be a diffeomorphism of open intervals, $\tilde{u} := u \circ \Phi$ the reparametrised curve, and $\tilde{T}_{s,s_0}$ its parallel transport map. Then

$$\tilde{T}_{s,s_0} = T_{\Phi(s),\Phi(s_0)} \,.$$

Proof. According to (10.3), within a coordinate patch the condition $D_t Y = 0$ is a homogeneous first order linear ODE,

$$\dot{Y}^\alpha(t) = -\dot{u}^i(t)\Gamma^\alpha_{i\beta}(u(t))\, Y^\beta(t) =: A^\alpha_\beta(t)\, Y^\beta(t)\,,$$

where $A^\alpha_\beta(t)$ is a smooth matrix-valued function of t. Hence, given an initial value $Y^\alpha(t_0)$ it has a unique smooth solution and the solution map T_{t,t_0} (propagator) is for each $t \in I$ an isomorphism. Uniqueness allows to patch the local solutions within coordinate charts together.

The last claim is a consequence of uniqueness and lemma 10.24. □

10.27 Definition. **Geodesics**

Let ∇ be an affine connection on M. Then a curve $u : I \to M$ is called a **geodesic** for the connection ∇ if its velocity vector field is parallel, i.e. if

$$D_t \dot{u} \equiv 0\,.$$

10.28 Remark. One possible rephrasing of the definition of a geodesic is a "curve of constant speed": the velocity vector field $\dot{u}$ is constant with repsect to the given connection. One could also say that the "acceleration" $D_t\dot{u}$ vanishes. ◇

In order to prove existence and uniqueness of geodesics, given a starting point $x_0 \in M$ and an initial velocity $v_0 \in T_{x_0}M$, we first observe that, using again (10.3), the geodesic equation $D_t\dot{u} = 0$ takes the following form within a coordinate chart (V, φ):

$$\ddot{u}^k(t) + \dot{u}^i(t)\,\dot{u}^j(t)\,\Gamma^k_{ij}(u(t)) = 0\,. \tag{10.4}$$

Note that by definition $u^k(t) := (\varphi(u(t))^k$ and $\dot{u}(t) =: (\dot{u}^k(t)\partial_{q_k})(u(t))$, and thus indeed $\dot{u}^k(t) = \frac{\mathrm{d}}{\mathrm{d}t}u^k(t)$. (10.4) is a system of second order ODEs for the components $u^k(t)$ of $(\varphi \circ u)(t)$ on $\varphi(V)$. Introducing $v^k(t) := \dot{u}^k(t)$, (10.4) is equivalent to the system of first order ODEs

$$\dot{u}^k(t) = v^k(t) \quad \text{and} \quad \dot{v}^k(t) = -v^i(t)\,v^j(t)\,\Gamma^k_{ij}(u(t)) \quad \text{on } \varphi(V)\times\mathbb{R}^n. \tag{10.5}$$

Now we could either use local existence and uniqueness of solutions to this ODE to prove local unique existence of geodesics and patch the local solutions together in order to obtain maximal solutions. Or we observe that (10.5) defines actually a vector field on the tangent bundle TM. To see this first recall that any chart (V, φ) on M defines a bundle chart

$$D\varphi : TV \to \varphi(V)\times\mathbb{R}^n\,, \quad (x, v) \mapsto D\varphi(x, v) =: (q(x), \dot{q}(x, v))$$

on the tangent bundle TV. Note that the dot in $\dot{q}$ is not a time derivative but $\dot{q}$ is just the name for a generic point in the second factor of $\varphi(V)\times\mathbb{R}^n$. Using these local coordinates and the corresponding coordinate vector fields, we can define a vector field $\mathbf{X}$ on TV by

$$\mathbf{X}(x, v) := v^k\partial_{q_k} \,-\, v^j\, v^i\,(\Gamma^k_{ij}\circ\pi_M)(x)\,\partial_{\dot{q}_k}\,.$$

By construction, $u : I \to V$ is a geodesic if and only if the curve $\dot{u} : I \to TV$, $\dot{u} = Du \circ e$ is an integral curve to $\mathbf{X}$. To see this, note that for a curve

$$w : I \to TV\,, \quad t \mapsto w(t) = (u(t), v^k(t)\partial_{q_k})$$

we have that

$$\dot{w} : I \to T(TV)\,, \quad t \mapsto \Big((u(t), v^k(t)\partial_{q_k}), \dot{u}^k(t)\partial_{q_k} + \dot{v}^k(t)\partial_{\dot{q}_k}\Big)$$

and

$$\mathbf{X}(w(t)) = \mathbf{X}(u(t), v^k(t)\partial_{q_k}) = \Big((u(t), v^k(t)\partial_{q_k}), v^k(t)\partial_{q_k} \,-\, v^j(t)\, v^i(t)\, \Gamma^k_{ij}(u(t))\, \partial_{\dot{q}_k}\Big) \,.$$

Hence, comparing the coefficients, we see that $\dot{w} = X(w(t))$ is equivalent to (10.5).

Since the geodesic equations are independet of the choice of coordinates, $\mathbf{X}$ is defined actually on all of TM and $u : I \to M$ is a geodesic if and only if $\dot{u}$ is an integral curve to $\mathbf{X}$.

10.29 Proposition. **Existence of unique maximal geodesics and the geodesic flow**

Let ∇ be an affine connection on M. For each $x_0 \in M$ and $v_0 \in T_{x_0}M$ there exists a unique maximal geodesic $\gamma_{x_0,v_0} : I_{x_0,v_0} \to M$ with $\dot{\gamma}_{x_0,v_0}(0) = (x_0, v_0)$. The corresponding flow $\Phi^{\mathbf{X}}$ on TM is called the **geodesic flow**.

Moreover, for any $\alpha \in \mathbb{R} \setminus \{0\}$ it holds that

$$I_{x_0,\alpha v_0} = \tfrac{1}{\alpha}\, I_{x_0,v_0} \quad \text{and} \quad \gamma_{x_0,\alpha v_0}(t) = \gamma_{x_0,v_0}\,(\alpha t) \,.$$

Hence, if one starts at the same point in the same direction with a different velocity, one runs through the same curve just with a different speed.

Proof. For existence and uniqueness just apply theorem 7.5 to the vector field $\mathbf{X}$ on TM. To prove the second claim, either check directly that also $t \mapsto \gamma_{x_0,v_0}\,(\alpha t)$ is a geodesic by inserting it into the differential equation. Or note that on a geodesic the velocity field is obtained by parallel transport of the initial velocity vector along the curve and that parallel transport is linear in the fibres. $\square$

10.30 Remark. Note that TM is never compact. Thus we don't get easily a global existence result for geodesics of general connections. Indeed, one can easily construct examples where the integral curves to $\mathbf{X}$ run to "infinity in the fibres" of TM in finite time.

In the next chapter we will discuss the affine connection induced by a Riemannian metric. For this so called Levi-Civita connection, however, we will obtain global existence of geodesics on compact manifolds M since the "length" of the velocity vector remains constant along a Riemannian geodesic. $\diamond$

By following the geodesic starting in a point x with a velocity v for a fixed time, say $t = 1$, one obtains a map from (a subset of) TM to M, the so called exponential map.

10.31 Definition. **The exponential map**

Let ∇ be an affine connection on M and $\gamma_{x,v} : I_{x,v} \to M$ the maximal geodesic of ∇ starting at x with velocity v. The **exponential map**

$$\exp_x : T_xM \supset \mathcal{E}_x \to M \,, \quad v \mapsto \exp_x(v) := \gamma_{x,v}(1) \,,$$

is defined on $\mathcal{E}_x := \{v \in T_xM \,|\, 1 \in I_{x,v}\}$, and

$$\exp : TM \supset \mathcal{E} \to M \,, \quad (x, v) \mapsto \exp_x(v) \,,$$

is defined on $\mathcal{E} := \{(x, v) \in TM \,|\, v \in \mathcal{E}_x\}$. $\diamond$

10.32 Proposition. **Properties of the exponential map**

(a) $\mathcal{E}$ is an open subset of TM containing the zero-section and $\exp : \mathcal{E} \to M$ is smooth.

(b) Each set $\mathcal{E}_x$ is star shaped with respect to 0, i.e. $v \in \mathcal{E}_x$ implies that also $\alpha v \in \mathcal{E}_x$ for all $\alpha \in [0, 1]$.

(c) For $t \in I_{x,v}$ it holds that

$$\gamma_{x,v}(t) = \exp_x(tv)\,.$$

Proof. According to theorem 7.5 the maximal domain $D_t := \{(x,v) \in TM \,|\, t \in I_{x,v}\}$ of the geodesic flow

$$\Phi_t^{\mathbf{X}} : D_t \to TM$$

is open for any $t \in \mathbb{R}$ and $\Phi_t^{\mathbf{X}}$ is a diffeomorphism onto its range. Now (a) follows from noting that $\mathcal{E} = D_1$ and $\exp = \pi_M \circ \Phi_1^{\mathbf{X}}$. Claims (b) and (c) follow immediately from the rescaling property of geodesics shown in proposition 10.29. $\square$

For the next lemma we require the manifolds version of the inverse function theorem.

10.33 Theorem. **Inverse function theorem**

Let $f : M_1 \to M_2$ be a smooth map between manifolds M_1 and M_2. If $Df|_{T_xM_1}$ is invertible at a point $x \in M_1$, then there exist open neighbourhoods U of x and V of $f(x)$ such that $f : U \to V$ is a diffeomorphism.

Proof. Just apply the inverse function theorem for maps on $\mathbb{R}^n$ within local charts. $\square$

10.34 Lemma. **Normal neighbourhood lemma**

Let M be a manifold with affine connection ∇. For each $x \in M$ there exists a star shaped open neighbourhood $\mathcal{V}_x \subset \mathcal{E}_x$ of $0 \in T_xM$ such that

$$\exp_x : \mathcal{V}_x \to \exp_x(\mathcal{V}_x)$$

is a diffeomorphism. Such a neighbourhood $\exp_x(\mathcal{V}_x)$ is called a **normal neighbourhood** of x.

Proof. It suffices to show that the differential $D\exp_x$ is invertable at $0 \in T_xM$, i.e. that the map

$$D\exp_x : T_0(T_xM) \cong T_xM \to T_{\exp_x(0)}M = T_xM$$

is invertible. However, $\exp_x$ maps the curve $c(t) = tv$ in T_xM to the curve $\exp_x(tv) = \gamma_{x,v}(t)$ in M. By definition of the geodesic $\gamma_{x,v}$ we have that $\frac{\mathrm{d}}{\mathrm{d}t}\gamma_{x,v}(t)|_{t=0} = v$, and thus $D\exp_x$ restricted to $T_0(T_xM)$ is actually the identity and, in particular, invertible. $\square$

German nomenclature

connection = Zusammenhang
exponential map = Exponentialabbildung
inverse function = Umkehrfunktion
section along a curve = Schnitt entlang einer Kurve
covariant derivative = Kovariante Ableitung
geodesic = Geodäte
parallel transport = Paralleltransport

11 The Riemannian connection

For a (pseudo-)Riemannian manifold (M, g), i.e. a manifold M with a (pseudo-)metric $g \in \mathcal{T}_2^0(M)$, there exists a natural connection on its tangent bundle TM, the so called Levi-Civita connection. It is natural in the sense that it is the unique connection that is compatible with the metric structure and symmetric.

11.1 Definition. **Compatability**

An affine connection ∇ on a (pseudo-)Riemannian manifold (M, g) is called **compatible** with a (pseudo-)metric $g \in \mathcal{T}_2^0(M)$, if

$$\nabla_X g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$$

for all $X, Y, Z \in \mathcal{T}_0^1(M)$. Recall that for an affine connection we put $\nabla_X f := L_X f = (\mathrm{d}f|X)$. ◇

Note that in general a connection ∇ need not be compatible with a given metric g, since when taking the derivative of the function $g(Y, Z)$ also g itself is differentiated.

11.2 Lemma. Let ∇ be an affine conncetion on a (pseudo-)Riemannian manifold (M, g). Then ∇ is compatible with g if and only if $\nabla g \equiv 0$.

Proof. According to proposition 10.17 (ii) we have that

$$\nabla_X g(Y, Z) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z) = (\nabla_X g)(Y, Z)\,.$$ □

11.3 Remark. If ∇ is compatible with g, then it also follows that

(a) For any curve u and any vector fields X, Y along u

$$\frac{\mathrm{d}}{\mathrm{d}t}\, g(X, Y) = g(D_t X, Y) + g(X, D_t Y)\,.$$

(b) If X, Y are parellel vector fields along a curve u, then $g(X, Y)$ is constant along u.

(c) Along any curve u, the parallel transport map $T_{t,t_0} : T_{u(t_0)}M \to T_{u(t)}M$ is an isometry.

Actually it is not difficult to see that the validity of either (a), or (b), or (c) in turn implies that ∇ is compatible with g. ◇

11.4 Remark. As a consequence of remark 11.3 (b), geodesics with respect to a affine connection ∇ that is compatible with a metric g are constant speed curves. That means that $\|\dot{u}(t)\|^2 := g(\dot{u}(t), \dot{u}(t))(u(t))$ is independent of t. Thus, on a compact manifold without boundary the corresponding geodesic flow exists globally.

11.5 Definition. **Symmetry and torsion**

Let ∇ be an affine conncetion on a manifold M. The **torsion map**

$$\tau : \mathcal{T}_0^1(M) \times \mathcal{T}_0^1(M) \to \mathcal{T}_0^1(M)\,, \qquad \tau(X, Y) := \nabla_X Y - \nabla_Y X - [X, Y]\,,$$

defines (by lemma 10.8 and a simple computation) a $(1, 2)$-tensor field on M, the **torsion tensor** of ∇. The connection ∇ is called **symmetric** or **torsion free** if $\tau \equiv 0$, i.e. if

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

for all vector fields X, Y. ◇

11.6 Remark. In local coordinates the symmetry condition reads

$$\begin{aligned}(\nabla_X Y - \nabla_Y X - [X,Y])^i &= (\mathrm{d}Y^i|X) + X^l Y^k \Gamma^i_{lk} - (\mathrm{d}X^i|Y) - Y^l X^k \Gamma^i_{lk} - ((\mathrm{d}Y^i|X) - (\mathrm{d}X^i|Y)) \\ &= X^l Y^k (\Gamma^i_{lk} - \Gamma^i_{kl}) \overset{!}{=} 0\,.\end{aligned}$$

Thus, an affine connection is symmetric if and only if its Christoffel symbols with respect to one (and thus to any) coordinate basis are symmetric in the lower indices. ◇

11.7 Theorem. **Fundamental lemma of Riemannian geometry**
Let (M, g) be a (pseudo-)Riemannian manifold. There exists a unique affine connection ∇ on M that is compatible with g and symmetric.
This connection is called the **Riemannian connection** or the **Levi-Civita connection** of g. It satisfies the **Koszul formula**

$$\begin{aligned}g(\nabla_X Y, Z) &= \tfrac{1}{2}\Big((\mathrm{d}g(Y,Z)|X) + (\mathrm{d}g(Z,X)|Y) - (\mathrm{d}g(X,Y)|Z) \\ &\qquad - g(Y,[X,Z]) - g(Z,[Y,X]) + g(X,[Z,Y])\Big)\end{aligned} \tag{11.1}$$

and its Christoffel symbols with respect to a local coordinate frame are

$$\Gamma^l_{ij} = \tfrac{1}{2}\, g^{lk}\Big(g_{jk,i} + g_{ki,j} - g_{ij,k}\Big)\,.$$

Proof. The strategy of the proof is as follows. One first assumes that an affine connection with the desired properties exists and derives the Koszul formula. From this we can conclude that if such a connection exists, then it is unique. Finally we show that (11.1) actually defines a connection with the desired properties, which proves existence.
Assume that ∇ is an affine connection that is compatible with g and symmetric. Then for $X, Y, Z \in \mathcal{T}^1_0(M)$ three versions of the compatibility condition are

$$\begin{aligned}(\mathrm{d}g(Y,Z)|X) &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \\ (\mathrm{d}g(Z,X)|Y) &= g(\nabla_Y Z, X) + g(Z, \nabla_Y X) \\ (\mathrm{d}g(X,Y)|Z) &= g(\nabla_Z X, Y) + g(X, \nabla_Z Y)\,,\end{aligned}$$

where we permuted X, Y, Z cyclically. Applying the symmetry condition on the last term in each line this becomes

$$\begin{aligned}(\mathrm{d}g(Y,Z)|X) &= g(\nabla_X Y, Z) + g(Y, \nabla_Z X) + g(Y,[X,Z]) \\ (\mathrm{d}g(Z,X)|Y) &= g(\nabla_Y Z, X) + g(Z, \nabla_X Y) + g(Z,[Y,X]) \\ (\mathrm{d}g(X,Y)|Z) &= g(\nabla_Z X, Y) + g(X, \nabla_Y Z) + g(X,[Z,Y])\,.\end{aligned}$$

Adding the first two lines and subtracting the last one yields

$$\begin{aligned}&(\mathrm{d}g(Y,Z)|X) + (\mathrm{d}g(Z,X)|Y) - (\mathrm{d}g(X,Y)|Z) \\ &\qquad = 2\,g(\nabla_X Y, Z) + g(Y,[X,Z]) + g(Z,[Y,X]) - g(X,[Z,Y])\,,\end{aligned}$$

where we used also symmtry of g. Solving for $g(\nabla_X Y, Z)$, this yields the Koszul formula (11.1). Since the right hand side of (11.1) does not depend on ∇, this proves that for any two symmetric connections ∇ and $\tilde{\nabla}$ that are compatible with g we have

$$g(\nabla_X Y, Z) = g(\tilde{\nabla}_X Y, Z) \qquad \text{for all } X, Y, Z \in \mathcal{T}^1_0(M).$$

As a (pseudo-)metric, g is non-degenerate. Thus, $\nabla_X Y = \tilde{\nabla}_X Y$ for all $X, Y \in \mathcal{T}_0^1(M)$, which proves uniqueness of ∇.

Existence is now proved by using the Koszul formula as a definition of ∇. Since we already have uniqueness, it suffices to construct a connection with the desired properties within local coordinate charts. Replacing X, Y, Z in (11.1) by coordinate vector fields we find that

$$g(\nabla_{\partial_i}\partial_j, \partial_k) = \tfrac{1}{2}\Big((\mathrm{d}g(\partial_j, \partial_k)|\partial_i) + (\mathrm{d}g(\partial_k, \partial_i)|\partial_j) - (\mathrm{d}g(\partial_i, \partial_j)|\partial_k)\Big)\,.$$

Recalling the definitions

$$g_{ij} := g(\partial_i, \partial_j) \qquad \text{and} \qquad \nabla_{\partial_i}\partial_j =: \Gamma^l_{ij}\partial_l\,,$$

we obtain

$$\Gamma^l_{ij} g_{lk} = \tfrac{1}{2}\Big(g_{jk,i} + g_{ki,j} - g_{ij,k}\Big) \quad \text{and hence} \quad \Gamma^l_{ij} = \tfrac{1}{2}\, g^{lk}\Big(g_{jk,i} + g_{ki,j} - g_{ij,k}\Big)\,. \tag{11.2}$$

It remains to show that these Christoffel symbols really define a connection with the desired properties. Symmetry in the lower indices, and thus according to remark 11.6 also symmetry of the connection, is evident by the symmetry of g, i.e. $g_{ij} = g_{ji}$. To show also compatibility with g, i.e. $\nabla g = 0$, we use again lemma 10.19:

$$g_{ij;k} = g_{ij,k} - \Gamma^l_{ki} g_{lj} - \Gamma^l_{kj} g_{il}\,.$$

On the other hand, using (11.2) twice, we obtain

$$\Gamma^l_{ki} g_{lj} + \Gamma^l_{kj} g_{il} = \tfrac{1}{2}\Big(g_{ji,k} + g_{jk,i} - g_{ik,j} + g_{ji,k} + g_{ki,j} - g_{kj,i}\Big) = g_{ij,k}$$

and thus $g_{ij;k} = 0$. □

11.8 Proposition. Naturality of the Riemannian connection

Let $\Phi : M \to \tilde{M}$ be an isometry of (pseudo-)Riemannian manifolds (M, g) and $(\tilde{M}, \tilde{g})$.

(a) For all $X, Y \in \mathcal{T}_0^1(M)$ it holds that

$$\Phi_*(\nabla_X Y) = \tilde{\nabla}_{\Phi_* X}(\Phi_* Y)\,.$$

(b) For any smooth curve $u : I \to M$ and vector field $Y \in \mathcal{T}_0^1(u)$ along u

$$\Phi_*(D_t Y) = \tilde{D}_t(\Phi_* Y)\,,$$

where $\Phi_* Y : I \to T\tilde{M}$ is defined as $\Phi_* Y := D\Phi \circ Y$.

(c) Let $u : I \to M$ be a geodesic on M, then $\tilde{u} := \Phi \circ u$ is a geodesic on $\tilde{M}$.

Proof. Define the pull-back connection

$$\overline{\nabla} : \mathcal{T}_0^1(M) \times \Gamma(TM) \to \Gamma(TM)\,, \quad \overline{\nabla}_X Y := \Phi^*(\tilde{\nabla}_{\Phi_* X}(\Phi_* Y))\,.$$

One now checks easily that $\overline{\nabla}$ defines a connection on TM. It is compatible with g,

$$\begin{aligned} \overline{\nabla}_X g(Y, Z) = \mathrm{d}_X g(Y, Z) &= \mathrm{d}_X(\Phi^*\tilde{g})(Y, Z) = \Phi^*\mathrm{d}_{\Phi_* X}\tilde{g}(\Phi_* Y, \Phi_* Z) \\ &= \Phi^*\tilde{g}(\tilde{\nabla}_{\Phi_* X}\Phi_* Y, \Phi_* Z) + \Phi^*\tilde{g}(\Phi_* Y, \tilde{\nabla}_{\Phi_* X}\Phi_* Z) \\ &= g(\Phi^*\tilde{\nabla}_{\Phi_* X}\Phi_* Y, Z) + g(Y, \Phi^*\tilde{\nabla}_{\Phi_* X}\Phi_* Z) \\ &= g(\overline{\nabla}_X Y, Z) + g(Y, \overline{\nabla}_X Z)\,, \end{aligned}$$

and, by a similar computation, also symmetric. Hence, $\overline{\nabla} = \nabla$. For (b) proceed analogously using the characterization of D_t from proposition 10.23 and (a). Finally, (c) follows from

$$\tilde{D}_t\dot{\tilde{u}} = \tilde{D}_t(D(\Phi \circ u) \circ e) = \tilde{D}_t(D\Phi \circ Du \circ e) = \tilde{D}_t(\Phi_*\dot{u}) = \Phi_*(D_t\dot{u}) = 0\,.$$ □

11.9 Proposition. **Naturality of the exponential map**

Let $\Phi : M \to \tilde{M}$ be an isometry of (pseudo-)Riemannian manifolds (M, g) and $(\tilde{M}, \tilde{g})$ and denote by exp and $\widetilde{\exp}$ the exponential maps with respect to the corresponding Riemannian connections. Then

$$\Phi \circ \exp_x = \widetilde{\exp}_{\Phi(x)} \circ \Phi_*|_{\mathcal{E}_x}$$

Proof. Homework assignment. □

11.10 Definition. **Geodesic balls and spheres**

Let (M, g) be a Riemannian manifold. For $\varepsilon > 0$ and $x \in M$ let $B_\varepsilon(0) := \{v \in T_xM \,|\, g_x(v, v) < \varepsilon^2\}$ be the ε-ball around zero in T_xM. For ε small enough $B_\varepsilon(0) \subset \mathcal{V}_x$ and then $\exp_x(B_\varepsilon(0)) \subset M$ is called a **geodesic ball** around x with radius ε. If $\partial B_\varepsilon(0) := \{v \in T_xM \,|\, g_x(v, v) = \varepsilon^2\}$ is contained in $\mathcal{V}_x$, then $\exp_x(\partial B_\varepsilon(0)) \subset M$ is called a **geodesic sphere** around x with radius ε.

11.11 Definition. **Riemannian normal coordinates**

Let $(w_1, \ldots, w_n)$ be an orthonormal basis of T_xM with respect to a (pseudo-)Riemannian metric $g \in \mathcal{T}^0_2(M)$. For the induced isomorphism with $\mathbb{R}^n$ we write

$$W : \mathbb{R}^n \to T_xM\,, \quad (q^1, \ldots, q^n) \mapsto q^j w_j\,.$$

Then on any normal neighbourhood U of x one can define **(Riemannian) normal coordinates** centred at x by

$$\varphi : U \to \mathbb{R}^n\,, \quad \varphi := W^{-1} \circ \exp_x^{-1}\,.$$

11.12 Proposition. **Properties of normal coordinates**

Let (M, g) be a (pseudo-)Riemannian manifold and let (U, φ) be normal coordinates centred at $x \in M$. Then

(a) $\varphi(x) = (0, \ldots, 0)$.

(b) $g_{ij}(x) = \delta_{ij}$.

(c) $\Gamma^k_{ij}(x) = 0$ for all $i, j, k = 1, \ldots, n$.

(d) For $v = v^j \partial_{q_j} \in T_xM$ the coordinate representation of the geodesic $\gamma_{x,v}$ has the simple form

$$(\varphi \circ \gamma_{x,v})(t) = (tv^1, \ldots, tv^n)\,,$$

as long as $\gamma_{x,v}$ stays in U.

Proof. (a) is obvious, since $\exp_x^{-1}(x) = 0$. For (b) note that $\partial_{q_i}(0) = w_i$ and thus $g_{ij}(0) := g(\partial_{q_i}, \partial_{q_j})(0) = g|_{T_xM}(w_i, w_j) = \delta_{ij}$, since $(w_1, \ldots, w_n)$ is, by definition of normal coordinates, an orthonormal basis of T_xM. For (d) note that $\gamma_{x,v}(t) = \exp_x(tv)$ and thus $u(t) := (\varphi \circ \gamma_{x,v})(t) = W^{-1}(tv) = (tv^1, \ldots, tv^n)$. Since $\ddot{u}^k(0) = 0$, the geodesic equation (10.5) implies that $v^i v^j \Gamma^k_{ij}(x) = 0$ for all $i, j, k = 1, \ldots, n$ and $v \in T_xM$. Thus also $\Gamma^k_{ij}(x) = 0$ for all $i, j, k = 1, \ldots, n$. □

Be warned that (b) and (c) only hold at the point x and that there is, in general, no coordinate system such that $g_{ij} = \delta_{ij}$ or $\Gamma^k_{ij} = 0$ holds in an open neighbourhood of x. Also (d) holds in general only for geodesics passing through x, so called radial goedesics. The coordinate representation of other geodesics passing through U need not be a straight line.

German nomenclature

normal coordinates = Normalkoordinaten   geodesic ball = geodätischer Ball
torsion = Torsion

12 Curvature

Let $\pi : E \to M$ be a vector bundle with connection ∇. Given a vector $s_0 \in E_x$ one can ask whether there exists a section $S \in \Gamma(E|_U)$ defined in a neighbourhood $U \subset M$ of the point x such that $S(x) = s_0$ and such that S is parallel, i.e. $\nabla S = 0$. We already know that for one-dimensional manifolds (i.e. curves) the answer is yes and that S is indeed unique and given by the parallel transport. On vector bundles over higher dimensional manifolds this is only true if the curvature of the connection vanishes. Hence, curvature can be seen as an obstruction to the existence of parallel sections.

Let us briefly motivate the definition of curvature given below. Assume the above setting where M is a two dimensional manifold. Choose a chart φ centred at x, i.e. such that $\varphi(x) = 0$. To construct a parallel section S with $S(x) = s_0$ we proceed in the following way: We first parallel transport s_0 along the q_1-coordinate line going through x and then, starting on each point on this coordinate line, parallel translate the resulting vector along the corresponding q_2-coordinate line. By smooth dependence on initial data this indeed yields a local section $S \in \Gamma(E|_U)$ with $S(x) = s_0$. By construction, it holds that $\nabla_{\partial_{q_1}} S = 0$ on the q_1-coordinate line through x and $\nabla_{\partial_{q_2}} S = 0$ on all of U. While this section S is the only candidate for a parallel section on U with $S(x) = s_0$, in general it is not true that $\nabla_{\partial_{q_1}} S = 0$ on all of U. If we could show that $\nabla_{\partial_{q_1}} S$ is constant along the q_2-coordinate lines, i.e. that

$$\nabla_{\partial_{q_2}} \nabla_{\partial_{q_1}} S = 0\,,$$

then $\nabla_{\partial_{q_1}} S = 0$ on the q_1-coordinate line through x would imply that $\nabla_{\partial_{q_1}} S = 0$ on all of U. Since $\nabla_{\partial_{q_1}} \nabla_{\partial_{q_2}} S = 0$, the condition for the local existence of parallel sections thus reads

$$\nabla_{\partial_{q_2}} \nabla_{\partial_{q_1}} S = \nabla_{\partial_{q_1}} \nabla_{\partial_{q_2}} S\,.$$

12.1 Definition. **The curvature map**

Let $\pi : E \to M$ be a vector bundle with connection ∇. The map

$$\mathcal{R} : \mathcal{T}_0^1(M) \times \mathcal{T}_0^1(M) \times \Gamma(E) \to \Gamma(E)\,, \quad (X, Y, S) \mapsto \mathcal{R}(X, Y, S) := \nabla_X \nabla_Y S - \nabla_Y \nabla_X S - \nabla_{[X,Y]} S$$

is called the **curvature map**. It is obviously skew-symmetric in the first two arguments, $\mathcal{R}(Y, X, S) = -\mathcal{R}(X, Y, S)$. $\diamond$

The next proposition shows, that the map R is indeed tensorial and thus defines, given two tangent vectors $X(x)$ and $Y(x)$ at $x \in M$, an endomorphism of the fibre E_x.

12.2 Remark. This endomorphism can be thought of being the parallel transport map along an infinitesimal curve that is obtained by first going in direction Y, then in direction X, then back in direction $-Y$, then $-X$, and finally, to close the loop, in direction $[X, Y]$. Let us briefly discuss this view on curvature on a formal level:

First denote the parallel transport map along the integral curves of a vector field X for time t by $\mathrm{e}^{t\nabla_X}$, i.e.

$$\mathrm{e}^{t\nabla_X} : \Gamma(E) \to \Gamma(E)\,, \quad S \mapsto (\mathrm{e}^{-t\nabla_X} S)(x) := T_{t,0} S(\Phi^X_{-t}(x))\,,$$

where $T_{t,0} : E|_{\Phi^X_{-t}(x)} \to E|_x$ is the parallel transport map along the integral curve of X that passes through x at time t and through $\Phi^X_{-t}(x)$ at time zero. The notation as the exponential of a differential operator is motivated by the fact that the "time-dependent" section $S(t) := \mathrm{e}^{-t\nabla_X} S$ satisfies the differential equation $\frac{\mathrm{d}}{\mathrm{d}t} S(t) = -\nabla_X S(t)$. With this notation the parallel transport map around the (almost) closed loop described above is given by the map

$$H(t) := \mathrm{e}^{-t^2\nabla_{[X,Y]}}\mathrm{e}^{t\nabla_X}\mathrm{e}^{t\nabla_Y}\mathrm{e}^{-t\nabla_X}\mathrm{e}^{-t\nabla_Y}\,.$$

Clearly, $\lim_{t\to 0} H(t) = \mathrm{Id}$. However, if we expand $H(t)$ in powers of t by (formally) using the Baker-Campbell-Hausdorff formula repeatedly, e.g.

$$\mathrm{e}^{t\nabla_X}\mathrm{e}^{t\nabla_Y} = \mathrm{e}^{t(\nabla_X+\nabla_Y)+\frac{t^2}{2}[\nabla_X,\nabla_Y]+\mathcal{O}(t^3)} \quad \text{and} \quad \mathrm{e}^{-t\nabla_X}\mathrm{e}^{-t\nabla_Y} = \mathrm{e}^{-t(\nabla_X+\nabla_Y)+\frac{t^2}{2}[\nabla_X,\nabla_Y]+\mathcal{O}(t^3)}\,,$$

we find that

$$H(t) = \mathrm{e}^{t^2([\nabla_X,\nabla_Y]-\nabla_{[X,Y]})+\mathcal{O}(t^3)} = \mathrm{Id} + t^2([\nabla_X,\nabla_Y] - \nabla_{[X,Y]}) + \mathcal{O}(t^3)$$

and thus

$$\lim_{t\to 0}\frac{H(t)S - S}{t^2} = \mathcal{R}(X,Y,S)\,.$$

12.3 Proposition. **The curvature map is tensorial**

The curvature map $\mathcal{R}$ is induced by a tensor field $R \in \Gamma(T^*M \otimes T^*M \otimes E^* \otimes E)$, i.e. for all $X, Y \in \mathcal{T}^1_0(M)$, $S \in \Gamma(E)$, and $T \in \Gamma(E^*)$

$$(\mathcal{R}(X,Y,S)\,|\,T)(x) = (R(x)\,|\,X(x), Y(x), S(x), T(x))\,.$$

R is called the curvature tensor.

Proof. According to lemma 10.8 we need to show $C^\infty(M)$-linearity of

$$(X,Y,S,T) \mapsto (\mathcal{R}(X,Y,S)\,|\,T) = (\nabla_X\nabla_Y S - \nabla_Y\nabla_X S - \nabla_{[X,Y]}S\,|\,T)$$

in all arguments. For T this is obvious. For X (and completely analogous for Y) we find

$$\begin{aligned}
\mathcal{R}(fX,Y,S) &= \nabla_{fX}\nabla_Y S - \nabla_Y\nabla_{fX}S - \nabla_{[fX,Y]}S\\
&= f\nabla_X\nabla_Y S - \nabla_Y f\nabla_X S - \nabla_{f[X,Y]-(\mathrm{d}f|Y)X}S\\
&= f\nabla_X\nabla_Y S - f\nabla_Y\nabla_X S - (\mathrm{d}f|Y)\nabla_X S - f\nabla_{[X,Y]}S + (\mathrm{d}f|Y)\nabla_X S\\
&= f(\nabla_X\nabla_Y S - \nabla_Y\nabla_X S - \nabla_{[X,Y]}S) = f\mathcal{R}(X,Y,S)\,,
\end{aligned}$$

and for S

$$\begin{aligned}
\mathcal{R}(X,Y,fS) &= \nabla_X\nabla_Y fS - \nabla_Y\nabla_X fS - \nabla_{[X,Y]}fS\\
&= \nabla_X f\nabla_Y S + \nabla_X(\mathrm{d}f|Y)S - \nabla_Y f\nabla_X S - \nabla_Y(\mathrm{d}f|X)S - f\nabla_{[X,Y]}S - (\mathrm{d}f|[X,Y])S\\
&= f\mathcal{R}(X,Y,S) + (\mathrm{d}f|X)\nabla_Y S + (\mathrm{d}f|Y)\nabla_X S + (\mathrm{d}(\mathrm{d}f|Y)|X)S\\
&\quad -(\mathrm{d}f|Y)\nabla_X S - (\mathrm{d}f|X)\nabla_Y S - (\mathrm{d}(\mathrm{d}f|X)|Y)S - (\mathrm{d}f|[X,Y])S\\
&= f\mathcal{R}(X,Y,S) + (L_X L_Y f)S - (L_Y L_X f)S - ((L_X L_Y - L_Y L_x)f)S\\
&= f\mathcal{R}(X,Y,S)\,.
\end{aligned}$$

□

12.4 Lemma. **Coordinate representation of the curvature tensor**
Let $(S_1, \ldots, S_k)$ be a local frame of E and $(\partial_{q_1}, \ldots, \partial_{q_n})$ a coordinate frame on M. Then

$$R = R_{ij\alpha}^{\beta}\, \mathrm{d}q^i \otimes \mathrm{d}q^j \otimes S^{\alpha} \otimes S_{\beta}$$

with

$$R_{ij\alpha}^{\beta} = \Gamma_{j\alpha,i}^{\beta} - \Gamma_{i\alpha,j}^{\beta} + \Gamma_{j\alpha}^{\gamma}\Gamma_{i\gamma}^{\beta} - \Gamma_{i\alpha}^{\gamma}\Gamma_{j\gamma}^{\beta}\,.$$

Proof. Since $[\partial_i, \partial_j] = 0$, we find that

$$\begin{aligned}
\mathcal{R}(\partial_i, \partial_j, S_\alpha) &= \nabla_{\partial_i}\nabla_{\partial_j} S_\alpha - \nabla_{\partial_j}\nabla_{\partial_i} S_\alpha = \nabla_{\partial_i}\Gamma_{j\alpha}^{\beta} S_\beta - \nabla_{\partial_j}\Gamma_{i\alpha}^{\beta} S_\beta \\
&= \Gamma_{j\alpha,i}^{\beta} S_\beta + \Gamma_{j\alpha}^{\beta}\Gamma_{i\beta}^{\gamma} S_\gamma - \Gamma_{i\alpha,j}^{\beta} S_\beta - \Gamma_{i\alpha}^{\beta}\Gamma_{j\beta}^{\gamma} S_\gamma \\
&= (\Gamma_{j\alpha,i}^{\beta} - \Gamma_{i\alpha,j}^{\beta} + \Gamma_{j\alpha}^{\gamma}\Gamma_{i\gamma}^{\beta} - \Gamma_{i\alpha}^{\gamma}\Gamma_{j\gamma}^{\beta}) S_\beta\,. \qquad \square
\end{aligned}$$

We now specialise to affine connections again.

12.5 Proposition. **The algebraic Bianchi identity**
Let ∇ be a symmetric connection on TM. Then

$$\mathcal{R}(X, Y, Z) + \mathcal{R}(Y, Z, X) + \mathcal{R}(Z, X, Y) = 0\,.$$

Proof.

$$\begin{aligned}
&\nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z + \nabla_Y\nabla_Z X - \nabla_Z\nabla_Y X - \nabla_{[Y,Z]}X + \nabla_Z\nabla_X Y - \nabla_X\nabla_Z Y - \nabla_{[Z,X]}Y \\
&\quad = \nabla_X(\nabla_Y Z - \nabla_Z Y) + \nabla_Y(\nabla_Z X - \nabla_X Z) + \nabla_Z(\nabla_X Y - \nabla_Y X) \\
&\qquad - \nabla_{[Y,Z]}X - \nabla_{[Z,X]}Y - \nabla_{[X,Y]}Z \\
&\quad = \nabla_X[Y, Z] - \nabla_{[Y,Z]}X + \nabla_Y[Z, X] - \nabla_{[Z,X]}Y + \nabla_Z[X, Y] - \nabla_{[X,Y]}Z \\
&\quad = [X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0\,,
\end{aligned}$$

where we used symmetry of ∇ twice and the Jacobi identity. $\square$

12.6 Proposition. **Symmetries of the Riemannian curvature tensor**
Let (M, g) be a (pseudo-)Riemannian manifold. Then the curvature tensor R of the Riemannian connection is a $(1, 3)$-tensor field and we define the $(0, 4)$-tensor field

$$\mathrm{Rm}(X, Y, Z, W) := g(\mathcal{R}(X, Y, Z), W)$$

and call it the **Riemannian curvature tensor**. It has the following symmetries:

(a) $\mathrm{Rm}(X, Y, Z, W) = -\mathrm{Rm}(Y, X, Z, W)$
(b) $\mathrm{Rm}(X, Y, Z, W) = -\mathrm{Rm}(X, Y, W, Z)$
(c) $\mathrm{Rm}(X, Y, Z, W) = \mathrm{Rm}(Z, W, X, Y)$

Proof. (a) is just the skew-symmetry of $\mathcal{R}$ in its first arguments. For (b) we will show that $\mathrm{Rm}(X, Y, V, V) = 0$ for all vector fields, which then implies that

$$0 = \mathrm{Rm}(X, Y, Z + W, Z + W) = \mathrm{Rm}(X, Y, Z, W) + \mathrm{Rm}(X, Y, W, Z)\,.$$

Using compatibility with the metric and symmetry of the latter we have that

$$\begin{aligned}
L_X L_Y g(V, V) &= L_X(g(\nabla_Y V, V) + g(V, \nabla_Y V)) = 2g(\nabla_X\nabla_Y V, V) + 2g(\nabla_Y V, \nabla_X V) \\
L_Y L_X g(V, V) &= L_Y(g(\nabla_X V, V) + g(V, \nabla_X V)) = 2g(\nabla_Y\nabla_X V, V) + 2g(\nabla_X V, \nabla_Y V) \\
L_{[X,Y]} g(V, V) &= 2g(\nabla_{[X,Y]} V, V)\,.
\end{aligned}$$

Subtracting the second and third equation from the first and dividing by 2 yields

$$0 = g(\nabla_X \nabla_Y V, V) - g(\nabla_Y \nabla_X V, V) - g(\nabla_{[X,Y]} V, V) = \mathrm{Rm}(X, Y, V, V) .$$

To prove (c), we write out four times the algebraic Bianchi identity,

$$\begin{aligned}
\mathrm{Rm}(Y,Z,X,W) + \mathrm{Rm}(Z,X,Y,W) + \mathrm{Rm}(X,Y,Z,W) &= 0 \\
-\mathrm{Rm}(W,X,Y,Z) - \mathrm{Rm}(Y,W,X,Z) - \mathrm{Rm}(X,Y,W,Z) &= 0 \\
-\mathrm{Rm}(X,Z,W,Y) - \mathrm{Rm}(W,X,Z,Y) - \mathrm{Rm}(Z,W,X,Y) &= 0 \\
\mathrm{Rm}(Y,Z,W,X) + \mathrm{Rm}(W,Y,Z,X) + \mathrm{Rm}(Z,W,Y,X) &= 0 .
\end{aligned}$$

Now add up the four lines. Applying (b) four times makes all the terms in the first two columns cancel. Then applying (a) and (b) in the last column yields $2\mathrm{Rm}(X,Y,Z,W) - 2\mathrm{Rm}(Z,W,X,Y) = 0$, which is equivalent to (c). □

12.7 Remark. In components with respect to a basis frame Rm is obtained from R by

$$\mathrm{Rm}_{ijkl} = R_{ijk}{}^{m} g_{ml} .$$

Note that it is sometimes useful to keep track of the order of all indices, not just for lower and upper indices independently. Then it is clear from the notation which index has been raised or lowered using the metric. The symmetries of Rm now read

$$\begin{aligned}
\mathrm{Rm}_{ijkl} &= -\mathrm{Rm}_{jikl} \\
\mathrm{Rm}_{ijkl} &= -\mathrm{Rm}_{ijlk} \\
\mathrm{Rm}_{ijkl} &= \mathrm{Rm}_{klij} \\
0 &= \mathrm{Rm}_{ijkl} + \mathrm{Rm}_{kijl} + \mathrm{Rm}_{jkil} ,
\end{aligned}$$

where the last line is the algebraic Bianchi identity. ⋄

12.8 Proposition. <u>**Invariance of the curvature tensor under isometries**</u>
The Riemannian curvature tensor is invariant under (local) isometries. More precisely, if $\Phi : (M, g) \to (\tilde{M}, \tilde{g})$ is a (local) isometry, then

$$\Phi^* \widetilde{\mathrm{Rm}} = \mathrm{Rm}$$

and

$$\tilde{\mathcal{R}}(\Phi_* X, \Phi_* Y, \Phi_* Z) = \Phi_* \mathcal{R}(X, Y, Z) .$$

Proof. With proposition 11.8 we have

$$\begin{aligned}
\tilde{\mathcal{R}}(\Phi_* X, \Phi_* Y, \Phi_* Z) &= (\tilde{\nabla}_{\Phi_* X} \tilde{\nabla}_{\Phi_* Y} - \tilde{\nabla}_{\Phi_* Y} \tilde{\nabla}_{\Phi_* X} - \tilde{\nabla}_{\Phi_* [X,Y]}) \Phi_* Z \\
&= \Phi_* (\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}) Z \\
&= \Phi_* \mathcal{R}(X, Y, Z) ,
\end{aligned}$$

and thus also

$$\begin{aligned}
\widetilde{\mathrm{Rm}}(\Phi_* X, \Phi_* Y, \Phi_* Z, \Phi_* W) &= \tilde{g}(\tilde{\mathcal{R}}(\Phi_* X, \Phi_* Y, \Phi_* Z), \Phi_* W) = \tilde{g}(\Phi_* \mathcal{R}(X, Y, Z), \Phi_* W) \\
&= \Phi_* g(\mathcal{R}(X, Y, Z), W) = \Phi_* \mathrm{Rm}(X, Y, Z, W) .
\end{aligned}$$

□

12.9 Definition. Flat manifolds

A Riemannian manifold is called **flat**, if it is locally isometric to euclidean space. ◇

12.10 Theorem. A Riemannian manifold is flat if and only if its curvature tensor vanishes.

Proof. Proposition 12.8 implies that the curvature tensor of a flat manifold vanishes. For the converse implication one can show that with respect to normal coordinates the metric g is euclidean if the curvature tensor vanishes (see e.g. chapter 7 in *John Lee, Riemannian manifolds: An introduction*). □

12.11 Proposition. The differential Bianchi identity

The total covariant derivative of the Riemannian curvature tensor satisfies

$$\nabla\mathrm{Rm}(X,Y,Z,V,W) + \nabla\mathrm{Rm}(X,Y,V,W,Z) + \nabla\mathrm{Rm}(X,Y,W,Z,V) = 0\,,$$

and in components

$$\mathrm{Rm}_{ijkl;m} + \mathrm{Rm}_{ijmk;l} + \mathrm{Rm}_{ijlm;k} = 0\,.$$

Proof. By proposition 12.6 the claim is equivalent to

$$\nabla\mathrm{Rm}(Z,V,X,Y,W) + \nabla\mathrm{Rm}(V,W,X,Y,Z) + \nabla\mathrm{Rm}(W,Z,X,Y,V) = 0\,.$$

By multi-linearity it suffices to prove the identity when X,Y,Z,V,W are elements of a coordinate frame. Let $(\partial_1,\ldots,\partial_n)$ be the coordinate frame of normal coordinates centered at $x\in M$. Then $[\partial_i,\partial_j]=0$, and, by proposition 11.12 (c), $(\nabla_{\partial_i}\partial_j)(x)=0$. Hence, for X,Y,Z,V,W taken from this frame,

$$\nabla\mathrm{Rm}(Z,V,X,Y,W)(x) = \nabla_W g(\mathcal{R}(Z,V,X),Y)(x) = g(\nabla_W\nabla_Z\nabla_V X,Y)(x) - g(\nabla_W\nabla_V\nabla_Z X,Y)(x)\,.$$

Thus, when evaluated at x,

$$\begin{aligned}
&\nabla\mathrm{Rm}(Z,V,X,Y,W) + \nabla\mathrm{Rm}(V,W,X,Y,Z) + \nabla\mathrm{Rm}(W,Z,X,Y,V) = \\
&\quad= g(\nabla_W\nabla_Z\nabla_V X,Y) - g(\nabla_W\nabla_V\nabla_Z X,Y) + g(\nabla_Z\nabla_V\nabla_W X,Y) - g(\nabla_Z\nabla_W\nabla_V X,Y) \\
&\qquad + g(\nabla_V\nabla_W\nabla_Z X,Y) - g(\nabla_V\nabla_Z\nabla_W X,Y) \\
&\quad= g(\mathcal{R}(W,Z,\nabla_V X),Y) + g(\mathcal{R}(Z,V,\nabla_W X),Y) + g(\mathcal{R}(V,W,\nabla_Z X),Y) \\
&\quad= 0\,,
\end{aligned}$$

since $\nabla_V X = \nabla_W X = \nabla_Z X = 0$. □

By taking traces of the Riemannian curvature tensor one obtains somewhat simpler objects that still encode parts of the geometric information contained in Rm. Recall that Rm was obtained from the curvature tensor R of the Riemannian connection by lowering the last index using the metric.

12.12 Definition. The Ricci and the scalar curvature

Let R be the curvature tensor of a (pseudo-)Riemannian manifold (M,g). Then the **Ricci curvature tensor** is the $(0,2)$-tensor obtained by contracting the first and the last argument of R,

$$\mathrm{Ric}(X,Y) := \mathrm{tr}\,R(\cdot,X,Y,\cdot)\,.$$

In local coordinates the definition reads

$$\mathrm{Ric}_{ij} = R_{kij}{}^{k} = g^{kl}\mathrm{Rm}_{kijl}\,.$$

The **scalar curvature** is the metric trace of the Ricci tensor,

$$S := \mathrm{tr}_g\mathrm{Ric}\,,\quad \text{i.e. locally}\quad S = g^{ij}\mathrm{Ric}_{ij}\,.$$

12.13 Lemma. The Ricci tensor is symmetric, i.e. $\mathrm{Ric}(X,Y)=\mathrm{Ric}(Y,X)$, and can be expressed in the following ways:

$$\mathrm{Ric}_{ij}=R_{kij}{}^{k}=R_{ik}{}^{k}{}_{j}=-R_{ki}{}^{k}{}_{j}=-R_{ikj}{}^{k}\,.$$

Proof. Homework assignment. □

12.14 Definition. **The divergence operator**

Let F be a tensor field on a (pseudo-)Riemannian manifold. The **divergence** $\mathrm{div}F$ of F with respect to the kth argument is the (metric)trace of ∇F with respect to the kth and the last argument. In particular, for $X\in\mathcal{T}^1_0(M)$ and $\omega\in\mathcal{T}^0_1(M)$

$$\mathrm{div}\,X=\mathrm{tr}\nabla X=X^j_{;j}\qquad\text{and}\qquad\mathrm{div}\,\omega=\mathrm{tr}_g\nabla\omega=g^{ij}\omega_{i;j}\,.$$

12.15 Proposition. **Contracted Bianchi identity**

It holds that

$$\mathrm{div}\,\mathrm{Ric}=\tfrac{1}{2}\nabla S\,,\quad\text{i.e. locally}\quad g^{jk}\mathrm{Ric}_{ij;k}=\tfrac{1}{2}S_{;i}\,.$$

Proof. Recall the differential Bianchi identity

$$\mathrm{Rm}_{ijkl;m}+\mathrm{Rm}_{ijmk;l}+\mathrm{Rm}_{ijlm;k}=0\,.$$

Metric contraction of the index pairs i,l and j,k yields the claim. □

We end this section with a short introduction to Riemannian geometry of submanifolds.

Let $N\subset M$ be a submanifold of a Riemannian manifold (M,g) and denote by $\psi:N\hookrightarrow M$ the inclusion map. Then $\tilde g:=\psi^*g$ is a metric on N and turns N into a Riemannian manifold. Sometimes $\tilde g$ is called the **first fundamental form**. In problem 52 of the homework assignments it is shown that the restriction

$$\tilde\nabla:\mathcal{T}^1_0(N)\times\Gamma(TN)\to\Gamma(TN)\,,\quad(\tilde X,\tilde Y)\mapsto\tilde\nabla_{\tilde X}\tilde Y:=P_N\nabla_{\tilde X}\tilde Y$$

of the Riemannian connection ∇ on TM to TN is indeed the Riemannian connection of $(N,\tilde g)$. Here $\tilde X,\tilde Y\in\Gamma(TN)\subset\Gamma(TM|_N)$ and $\nabla_{\tilde X}\tilde Y$ is well defined by problem 48. Recall that $P_N\in\mathrm{End}(TM|_N)$ where, for $x\in N$, $P_N(x)$ is the orthogonal (w.r.t. g) projection within T_xM onto the tangent space $T_xN\subset T_xM$. We denote by

$$T^\perp N:=\{(x,v)\in TM|_N\,|\,x\in N\text{ and }v\in T_xN^\perp\}$$

the normal bundle to N. Note that $TM|_N=TN\oplus T^\perp N$. One thus has the decomposition

$$\nabla_{\tilde X}\tilde Y=P_N\nabla_{\tilde X}\tilde Y+P_N^\perp\nabla_{\tilde X}\tilde Y=\tilde\nabla_{\tilde X}\tilde Y+I\!I(\tilde X,\tilde Y)\,.$$

Here

$$I\!I:\Gamma(TN)\times\Gamma(TN)\to\Gamma(T^\perp N)\,,\quad I\!I(\tilde X,\tilde Y):=P_N^\perp\nabla_{\tilde X}\tilde Y$$

is called the **second fundamental form**.

12.16 Proposition. **The second fundamental form**

The second fundamental form $I\!I$ is a symmetric vector-valued 2-form, taking values in the normal bundle.

Proof. Symmetry of $I\!I$ follows form the symmetry of the Riemannian connection and the fact that for two vector fields $\tilde{X}$ and $\tilde{Y}$ that are tangent to N also $[\tilde{X},\tilde{Y}]$ is tangent to N,

$$I\!I(\tilde{X},\tilde{Y}) = P_N^\perp \nabla_{\tilde{X}}\tilde{Y} = P_N^\perp \nabla_{\tilde{Y}}\tilde{X} - P_N^\perp[\tilde{X},\tilde{Y}] = P_N^\perp \nabla_{\tilde{Y}}\tilde{X} = I\!I(\tilde{Y},\tilde{X})\,.$$

Thus $I\!I$ is also $C^\infty(N)$-linear in both arguments and hence a vector-valued 2-form. □

12.17 Proposition. The Weingarten equation

Let $\tilde{X},\tilde{Y} \in \Gamma(TN)$ and $V \in \Gamma(T^\perp N)$. Then on N it holds that

$$g(\nabla_{\tilde{X}}V,\tilde{Y}) = -g(V, I\!I(\tilde{X},\tilde{Y}))\,.$$

Proof. Since $g(V,\tilde{Y}) = 0$ on N, we find

$$\begin{aligned} 0 &= L_{\tilde{X}}\, g(V,\tilde{Y}) = g(\nabla_{\tilde{X}}V,\tilde{Y}) + g(V,\nabla_{\tilde{X}}\tilde{Y}) \\ &= g(\nabla_{\tilde{X}}V,\tilde{Y}) + g(V,\tilde{\nabla}_{\tilde{X}}\tilde{Y}) + g(V, I\!I(\tilde{X},\tilde{Y})) \\ &= g(\nabla_{\tilde{X}}V,\tilde{Y}) + g(V, I\!I(\tilde{X},\tilde{Y}))\,. \end{aligned}$$ □

12.18 Proposition. The Gauß equation

For $\tilde{X},\tilde{Y},\tilde{Z},\tilde{W} \in \mathcal{T}_0^1(N)$ it holds that

$$\mathrm{Rm}(\tilde{X},\tilde{Y},\tilde{Z},\tilde{W}) = \widetilde{\mathrm{Rm}}(\tilde{X},\tilde{Y},\tilde{Z},\tilde{W}) + g(I\!I(\tilde{X},\tilde{Z}), I\!I(\tilde{Y},\tilde{W})) - g(I\!I(\tilde{Y},\tilde{Z}), I\!I(\tilde{X},\tilde{W}))$$

and in coordinates adapted to N

$$\mathrm{Rm}_{ijkl} = \widetilde{\mathrm{Rm}}_{ijkl} + g_{mp} I\!I^m_{ik} I\!I^p_{jl} - g_{mp} I\!I^m_{jk} I\!I^p_{il}$$

Proof. Using the definition of $I\!I$ and the Weingarten equation, we find that

$$\begin{aligned} g(\mathcal{R}(\tilde{X},\tilde{Y},\tilde{Z}),\tilde{W}) &= g(\nabla_{\tilde{X}}\nabla_{\tilde{Y}}\tilde{Z},\tilde{W}) - g(\nabla_{\tilde{Y}}\nabla_{\tilde{X}}\tilde{Z},\tilde{W}) - g(\nabla_{[\tilde{X},\tilde{Y}]}\tilde{Z},\tilde{W}) \\ &= g(\nabla_{\tilde{X}}(\tilde{\nabla}_{\tilde{Y}}\tilde{Z} + I\!I(\tilde{Y},\tilde{Z})),\tilde{W}) - g(\nabla_{\tilde{Y}}(\tilde{\nabla}_{\tilde{X}}\tilde{Z} + I\!I(\tilde{X},\tilde{Y})),\tilde{W}) \\ &\quad - g(\tilde{\nabla}_{[\tilde{X},\tilde{Y}]}\tilde{Z},\tilde{W}) + \underbrace{g(I\!I([\tilde{X},\tilde{Y}],\tilde{Z}),\tilde{W})}_{=0} \\ &= g(\tilde{\nabla}_{\tilde{X}}\tilde{\nabla}_{\tilde{Y}}\tilde{Z},\tilde{W}) + g(\nabla_{\tilde{X}} I\!I(\tilde{Y},\tilde{Z}),\tilde{W}) - g(\tilde{\nabla}_{\tilde{Y}}\tilde{\nabla}_{\tilde{X}}\tilde{Z},\tilde{W}) - g(\nabla_{\tilde{Y}} I\!I(\tilde{X},\tilde{Y}),\tilde{W}) \\ &\quad - g(\tilde{\nabla}_{[\tilde{X},\tilde{Y}]}\tilde{Z},\tilde{W}) \\ &= \tilde{g}(\tilde{\mathcal{R}}(\tilde{X},\tilde{Y},\tilde{Z}),\tilde{W}) - g(I\!I(\tilde{Y},\tilde{Z}), I\!I(\tilde{X},\tilde{W})) + g(I\!I(\tilde{X},\tilde{Z}), I\!I(\tilde{Y},\tilde{W}))\,. \end{aligned}$$ □

German nomenclature

curvature = Krümmung   flat manifold = flache Mannigfaltigkeit

13 Symplectic forms and Hamiltonian flows

In this section we collect basic definitions and results about Hamiltonian vector fields and flows on symplectic manifolds. These flows play a central role in classical Hamiltonian mechanics, a very elegant formulation of classical mechanics.

13.1 Definition. Symplectic manifolds

A **symplectic form** on a manifold M is a closed non-degenerate 2-form ω on M, i.e. $\omega \in \Lambda_2(M)$ with $\mathrm{d}\omega = 0$. The pair (M, ω) is called a **symplectic manifold**. $\diamond$

13.2 Remarks. (a) Recall the following result from linear algebra: Let V be an n-dimensional real vector space and ω a skew-symmetric bilinear form with $\operatorname{rank}\omega = r$, where the rank of ω is the basis-independent rank of the representing matrix $J_{ij} := \omega(e_i, e_j)$. Then $r = 2m$ for some $m \in \mathbb{N}_0$ and there exists a basis in which J has the form

$$J = \begin{pmatrix} 0 & \mathrm{id}_{m\times m} & 0 \\ -\mathrm{id}_{m\times m} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathcal{M}(n \times n, \mathbb{R}).$$

Thus, in particular, non-degenerate skew-symmetric bilinear forms exist only on even-dimensional vector spaces. And symplectic manifolds must also be even-dimensional.

(b) If ω is exact, then (M, ω) is called **exact symplectic**.

(c) On $M = T^*\mathbb{R}^n$ the canonical symplectic form is $\omega_0 := \sum_{j=1}^n \mathrm{d}q^j \wedge \mathrm{d}p^j$, where $(q,p) \in T^*\mathbb{R}^n \cong \mathbb{R}^n \times \mathbb{R}^n$ are cartesian coordinates.

Recall from remark 4.8 that any non-degenerate 2-form ω induces an isomorphism between $\mathcal{T}_0^1(M)$ and $\mathcal{T}_1^0(M)$:

$$X \in \mathcal{T}_0^1(M) \mapsto X^* = \omega(X, \cdot) = \mathrm{i}_X\omega \in \mathcal{T}_1^0(M).$$

13.3 Definition. Hamiltonian vector fields

A vector field $X \in \mathcal{T}_0^1(M)$ on a symplectic manifold (M, ω) is called a **Hamiltonian vector field**, if $\omega(X, \cdot)$ is an exact 1-form, resp. **locally Hamiltonian**, if $\omega(X, \cdot)$ is closed.

For $H \in C^\infty(M)$ the vector field X_H associated through ω with $\mathrm{d}H$, i.e.

$$\omega(X_H, \cdot) = \mathrm{d}H\,,$$

is called the **Hamiltonian vector field generated by** H and H is called the **Hamiltonian function**.

Note that the map $C^\infty(M) \to \mathcal{T}_0^1(M)$, $H \mapsto X_H$ is $\mathbb{R}$-linear. $\diamond$

13.4 Example. For $T^*\mathbb{R}^n \cong \mathbb{R}^n_q \times \mathbb{R}^n_p$ and a symplectic form $\omega \in \Lambda_2(T^*\mathbb{R}^n)$ let $J_{ij} = \omega(e_i, e_j)$. For $H \in C^\infty(T^*\mathbb{R}^n)$ it follows that

$$\begin{aligned} \omega(X_H, Y) = \mathrm{d}H(Y) \qquad \forall Y \in \mathcal{T}_0^1(T^*\mathbb{R}^n) \quad &\Leftrightarrow \quad X_H^i J_{ij} Y^j = Y^j \partial_j H \qquad \forall Y \in \mathcal{T}_0^1(T^*\mathbb{R}^n) \\ &\Leftrightarrow \quad J_{ij} X_H^i = \partial_j H \qquad \forall\, j = 1, \ldots, 2n \\ &\Leftrightarrow \quad J^T X_H = \nabla H \\ &\Leftrightarrow \quad X_H = (J^T)^{-1} \nabla H\,. \end{aligned}$$

If J has the canonical form $J_0 = \begin{pmatrix} 0 & \mathrm{id} \\ -\mathrm{id} & 0 \end{pmatrix}$, then $(J_0^T)^{-1} = J_0$ implies for X_H the usual Hamiltonian equations of motion,

$$X_H^j = \frac{\partial H}{\partial p_j} \quad \text{and} \quad X_H^{j+n} = -\frac{\partial H}{\partial q_j}\,, \qquad j = 1, \dots, n\,.$$

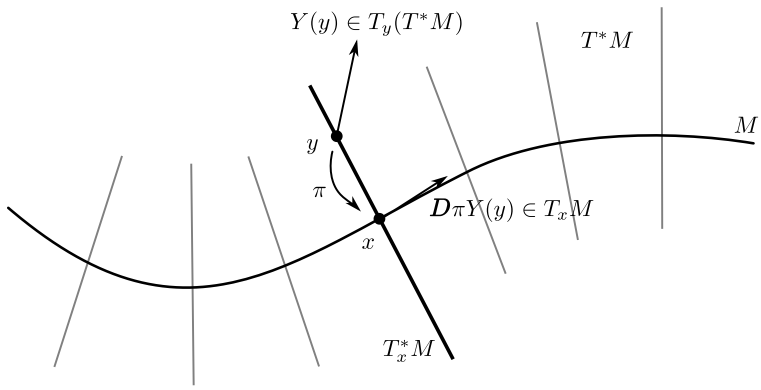
We will next understand in which sense ω_0 is canonical on $T^*\mathbb{R}^n$. To this end we show that on any cotangent bundle T^*M one can define a canonical symplectic form ω_0 without involving additional structure. We first define a canonical 1-form Θ_0 on T^*M and then ω_0 as its exterior derivative.
Let $y = (x, v^*) \in T^*M$ (i.e. $x \in M$ and $v^* \in T_x^*M$) and $Y \in \mathcal{T}_0^1(T^*M)$. Then put

$$\Big(\Theta_0(y)\,|\,Y(y)\Big)_{T_y^*(T^*M),\,T_y(T_xM)} := \Big(v^*|D\pi_M(Y(y)\Big)_{T_x^*M,\,T_xM}$$

where $\pi_M : T^*M \to M$ is the projection $(x, v^*) \mapsto x$ on the base point and hence its differential is a map

$$D\pi_M : T(T^*M) \to TM\,.$$

So $\Theta_0(y)$ acts on $Y(y)$ as v^* acts on $T\pi(Y(y))$.



13.5 Definition. **The canonical symplectic form**
The 1-form $\Theta_0 \in \mathcal{T}_1^0(T^*M)$ is called the **canonical 1-form** on the cotangent bundle. The form $\omega_0 = -\mathrm{d}\Theta_0 \in \mathcal{T}_2^0(T^*M)$ is called the **canonical symplectic form** on the **phase space** T^*M. M is called the **configuration space**. ◇

13.6 Remark. To see that ω_0 is indeed non-degenerate, we compute its local coordinate expression. In a bundle chart $D^*\varphi : T^*V \to T^*\varphi(V) \cong \varphi(V) \times \mathbb{R}^n \subset \mathbb{R}_q^n \times \mathbb{R}_p^n$ we denote the coordinates in the first factor by q and in the second factor by p. A point $(q,p) \in T^*M$ thus denotes the covector $p_i \mathrm{d}q^i \in T_q^*M$ at the point $q \in M$.
A vector field $Y \in \mathcal{T}_0^1(T^*M)$ can thus be locally written as

$$Y(q,p) = v^i(q,p)\,\partial_{q_i} + w^i(q,p)\,\partial_{p_i}\,,$$

where the index i runs from 1 to n. We have that

$$D\pi\, Y(q,p) = v^i(q,p)\,\partial_{q_i}$$

and

$$\Big(\Theta_0(q,p)\,|\,Y(q,p)\Big) \;=\; \big(p_j \mathrm{d}q^j\,|\,v^i(q,p)\,\partial q_i\big) = p_i\, v^i(q,p) = \big(p_j \mathrm{d}q^j\,|\,Y(q,p)\big)\,.$$

Hence, $\Theta_0 = p_j \mathrm{d}q^j$, where it is important to keep in mind that $\Theta_0 \in \mathcal{T}^0_1(T^*M)$ is a 1-form on the manifold T^*M and not on M! However, Θ_0 vanishes on all vectors tangent to the fibres of T^*M.
For the canonical symplectic form we thus find

$$\omega_0 = -\mathrm{d}\Theta_0 = -\mathrm{d}p^i \wedge \mathrm{d}q^i = \mathrm{d}q^i \wedge \mathrm{d}p^i\,,$$

and ω_0 is indeed non-degenerate and defines a symplectic form. Its coefficient matrix with respect to the coordinate vector fields $(\partial_{q_1}, \ldots, \partial_{q_n}, \partial_{p_1}, \ldots, \partial_{p_n})$ has the form

$$\begin{pmatrix} 0 & \mathrm{id}_{n\times n} \\ -\mathrm{id}_{n\times n} & 0 \end{pmatrix}$$

on the whole coordinate patch. $\diamond$

13.7 Theorem. Darboux' theorem
Let (M, ω) be a symplectic manifold of dimension $2n$ and $x \in M$. Then there exists a chart (V, φ) with $x \in V$ such that with $\varphi : V \to \mathbb{R}^{2n}$, $\varphi(x) =: (q(x), p(x))$, it holds on all of V that

$$\omega = \sum_{i=1}^{n} \mathrm{d}q^i \wedge \mathrm{d}p^i\,.$$

A chart in which ω has this form is called a **canonical chart** oder **Darboux chart**.

Proof. We will prove a slightly more general statement:

13.8 Lemma. Moser's trick
Let $U_0, U_1 \subset \mathbb{R}^{2n}$ be open neighbourhoods of $0 \in \mathbb{R}^{2n}$ and ω_0, ω_1 symplectic forms on U_0 resp. U_1. There exists an open neighbourhood $U \subset U_0 \cap U_1$ of zero and a diffeomorphism $F : U \to U$ such that $\omega_0|_U = F^*(\omega_1|_U)$ and $F(0) = 0$.

The statement of Darboux' theorem now follows by choosing any local chart $(\tilde{V}, \tilde{\varphi})$ and making the following identifications: $U_0 = \tilde{\varphi}(\tilde{V})$, $U_1 = \mathbb{R}^{2n}$, $\omega_0 = \tilde{\varphi}_*\omega$ and $\omega_1 = \sum_{i=1}^n \mathrm{d}q^i \wedge \mathrm{d}p^i$. Then the desired chart (V, φ) is given by $V = \tilde{\varphi}^{-1}(U)$ and $\varphi = F \circ \tilde{\varphi}|_V$. $\square$

Proof. of Moser's Trick. According to remark 13.2, there is a linear transformation A of $\mathbb{R}^{2n}$ such that $A_*\omega_0(0) = \omega_1(0)$. To make the notation in the proof less heavy, we can thus assume without loss of generality that $\omega_0(0) = \omega_1(0)$.
The diffeomorphism $F = F_1$ is now constructed as the solution of the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} F_t = X_t \circ F_t$$

with initial datum $F_0 = \mathrm{id}$ and a time dependent vector field X_t that remains to be constructed in such a way, that

$$L_{X_t}\omega_t = \omega_0 - \omega_1 \quad \text{with} \quad \omega_t := (1-t)\omega_0 + t\omega_1 \tag{13.1}$$

holds. Then it follows that $F_t^*\omega_t = \omega_0$ for all $t \in [0, 1]$, since $F_0^*\omega_0 = \omega_0$ and

$$\frac{\mathrm{d}}{\mathrm{d}t} F_t^*\omega_t = F_t^*(L_{X_t}\omega_t + \frac{\mathrm{d}}{\mathrm{d}t}\omega_t) = 0\,.$$

In order to find a vector field X_t with the property (13.1), we first note that $\omega_t|_U$ is a symplectic form for all $t \in [0, 1]$ if we chose the neighbourhood U of zero sufficiently small. This is because $\omega_t(0) = \omega_0(0) = \omega_1(0)$ is constant and thus non-degenerate for all t.

If we choose a contractable U, then on U the closed form $\omega_0 - \omega_1$ is exact. Hence, $\omega_0 - \omega_1 = \mathrm{d}\theta$ for some 1-form θ. On the other hand, also ω_t is closed and according to Cartan's formula (theorem 8.13) we have $L_{X_t}\omega_t = \mathrm{d}i_{X_t}\omega_t$. We thus need to achieve

$$i_{X_t}\omega_t = \theta + \mathrm{d}f$$

for some $f \in \Lambda_0(U)$ and all $t \in [0,1]$, in order to obtain (13.1). However, since ω_t is non-degenerate on U there exists such an X_t for every smooth function f. By choosing f appropriately, we obtain $\theta(0) + \mathrm{d}f(0) = 0$ and hence $X_t(0) = 0$ and $F_t(0) = 0$ for all $t \in [0,1]$. □

13.9 Remark. Thus, on **every** symplectic manifold there are local coordinates such that ω has the normal form $\omega = \sum_{i=1}^n \mathrm{d}q^i \wedge \mathrm{d}p^i$ not only point wise, but in an open neighbourhood.

In the previous section we saw that the analogous statement does not hold for Riemannian manifolds (M,g). While for each $x \in M$ there exists a chart such that $g(x) = \sum_{i=1}^n \mathrm{d}q^i \otimes \mathrm{d}q^i$, a chart such that

$$g|_V = \mathrm{d}q^i \otimes \mathrm{d}q^i$$

for some open neighbourhood V of x exists if and only if the Riemannian curvature Rm associated with g vanishes on some neighbourhood of x, i.e. if M is flat around x. ◇

We now come to a central result of Hamiltonian mechanics, Liouville's theorem. It states that Hamiltonian flows, i.e. flows of Hamiltonian vector fields, are symplectomorphisms, i.e. leave invariant the symplectic form.

13.10 Theorem. Liouville

Let (M,ω) be a symplectic manifold of dimension $\dim M = 2n$, and let X_H be the Hamiltonian vector field generated by $H \in C^\infty(M)$. Then the flow of X_H satisfies

$$\Phi_t^{X_H*}\omega = \omega$$

and hence also

$$\Phi_t^{X_H*}(\underbrace{\omega \wedge \cdots \wedge \omega}_{k\,\text{copies}}) = \underbrace{\omega \wedge \cdots \wedge \omega}_{k\,\text{copies}} \quad \text{for } 1 \le k \le n\,.$$

Proof. It holds that

$$L_{X_H}\omega = (i_{X_H}\mathrm{d} + \mathrm{d}i_{X_H})\omega = i_{X_H}\underbrace{\mathrm{d}\omega}_{=0} + \underbrace{\mathrm{dd}H}_{=0} = 0\,.$$

Because of $\Phi_0^{X_H} = \mathrm{Id}$ we have $\Phi_0^{X_H*}\omega = \omega$ and with

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi_t^{X_H*}\omega = \Phi_t^{X_H*}L_{X_H}\omega = 0$$

the statement follows. □

13.11 Definition. The Liouville measure

The volume form

$$\Omega := \frac{(-1)^{\frac{(n-1)n}{2}}}{n!}\underbrace{\omega \wedge \omega \wedge \cdots \wedge \omega}_{n\,\text{copies}} \in \Lambda_{2n}(M)$$

is called the **Liouville measure**. In every canonical chart it has the form

$$\Omega = \mathrm{d}q^1 \wedge \cdots \wedge \mathrm{d}q^n \wedge \mathrm{d}p^1 \wedge \cdots \wedge \mathrm{d}p^n\,.$$

13.12 Corollary. **Invariance of the Liouville measure**
A Hamiltonian flow $\Phi_t^{X_H}$ leaves invariant Ω, i.e.

$$\Phi_t^{X_H*}\Omega = \Omega\,.$$

One says that Hamiltonian flows are volume preserving. ◇

13.13 Definition. **The Poisson bracket**
Let (M,ω) be a symplectic manifold and $f,g \in C^\infty(M)$. The **Poisson bracket** of f and g is the function

$$\{f,g\} \;:=\; \omega(X_f,X_g) \in C^\infty(M).$$

It holds that

$$\{f,g\} \;=\; \omega(X_f,X_g) \;=\; i_{X_g}\omega(X_f,\cdot) \;=\; i_{X_g}\mathrm{d}f \;=\; L_{X_g}f \;=\; -L_{X_f}g\,.$$

13.14 Remark. In a canonical chart we have

$$\{f,g\} = \sum_{i=1}^n \Big(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}\Big).$$

For the coordinate function the canonical relations hold,

$$\{q_i,q_j\} = \{p_i,p_j\} = 0 \quad \text{and} \quad \{q_i,p_j\} = \delta_{ij}\,.$$

13.15 Proposition. **The Liouville equation**
Let $g,H \in C^\infty(M)$ and let Φ_t^H be the Hamiltonian flow of X_H. Then $g(t) := g\circ\Phi_t^H$ solves the Liouville equation

$$\frac{\partial}{\partial t}\,g(t) = \{g(t),H\}\,.$$

Whenever $\{g,H\} = 0$, then $g(t) = g(0)$ for all $t \in \mathbb{R}$. In particular, $H(t) = H\circ\Phi_t^H = H$.

Proof. Homework assignment. □

13.16 Definition. **Symplectic maps and canonical transformations**
Let (M,ω) and (N,σ) be symplectic manifolds. A smooth map $\Psi : M \to N$ is called **symplectic**, if

$$\Psi^*\sigma = \omega\,.$$

A symplectic diffeomorphism is called a **symplectomorphism** or **canonical transformation**.

13.17 Proposition. **Canonical transformations of Hamiltonian vector fields**
Let $\Psi : M \to N$ be a canonical transformation and $f \in C^\infty(N)$. Then

$$\Psi^* X_f = X_{\Psi^* f}\,.$$

Proof. Since ω is non-degenerate, this follows from

$$\begin{aligned}\omega(X_{\Psi^*f},Y) \;&=\; \mathrm{d}(\Psi^*f)(Y) \;=\; (\Psi^*\mathrm{d}f)(Y) \;=\; \mathrm{d}f(\Psi_*Y) \;=\; \sigma(X_f,\Psi_*Y)\\ &=\; (\Psi_*\omega)(X_f,\Psi_*Y) \;=\; \omega(\Psi^*X_f,Y)\,.\end{aligned}$$ □

13.18 Corollary. The Lie bracket of Hamiltonian vector fields

The commutator of Hamiltonian vector fields is again a Hamiltonian vector fields, more precisely

$$[X_g, X_f] = X_{\{f,g\}}\,.$$

Proof. Let Φ_t be the (local) flow of X_g. Then proposition 13.17 implies

$$\Phi_t^* X_f = X_{\Phi_t^* f}$$

evaluating the derivative at $t = 0$ we find

$$L_{X_g} X_f = X_{L_{X_g} f}\,.$$

Here we also used that the map $f \mapsto X_f$ is linear. By definition 13.13 we have $L_{X_g} f = \{f, g\}$ and by definition 8.7 $L_{X_g} X_f = [X_g, X_f]$. □

13.19 Corollary. The form of the Hamiltonian equations of motion is invariant under canonical transformations: Let $H \in C^\infty(M)$ and let (q, p) be a canonical chart on M. According to example 13.4 in such a chart it holds that

$$X_H = \begin{pmatrix} \frac{\partial H}{\partial p} \\ -\frac{\partial H}{\partial q} \end{pmatrix} := \frac{\partial H}{\partial p_i}\,\partial_{q_i} - \frac{\partial H}{\partial q_i}\,\partial_{p_i}\,.$$

Let now $\Psi : M \to N$ be a canonical transformation and $K := \Psi_* H = H \circ \Psi^{-1}$. The $(Q, P) = (q, p) \circ \Psi^{-1}$ is a canonical chart on N and in this chart

$$\Psi_* X_H = X_K = \begin{pmatrix} \frac{\partial K}{\partial P} \\ -\frac{\partial K}{\partial Q} \end{pmatrix}.$$

Proof. The statement that $\Psi_* X_H = X_K$ was shown in proposition 13.17. It remains to check that (Q, P) is a canonical chart. This follows from

$$\sigma(\partial_{Q_i}, \partial_{P_j}) = \sigma(\Psi_*\partial_{q_i}, \Psi_*\partial_{p_j}) = \Psi^*\sigma(\partial_{q_i}, \partial_{p_j}) = \omega(\partial_{q_i}, \partial_{p_j})\,.$$ □

German nomenclature

Hamiltonian vector field = Hamiltonsches Vektorfeld	Liouville measure = Liouvillemaß
Poisson bracker = Poissonklammer	symplectic form = symplektische Form