Analysis on Manifolds

(lecture notes)

Spring Term 2024

Andreas Čap

Institut für Mathematik, Universität Wien, Oskar–Morgenstern–Platz 1, A–1090 Wien

Email address: Andreas.Cap@univie.ac.at

Contents

Preface

The first version of these lecture notes was complied when I was teaching the course "Analysis on Manifolds" in spring term 2020. After about 2 weeks of the semester, Austria went into the first lock-down caused by the COVID19-crisis and Universities switched to distance teaching without much preparation. In those circumstances, self learning from the lecture notes supported by additional texts on the material that were called "informal remarks" seemed like the best way to continue the course. Thus the written notes that were prepared during the semester became the main source of information for the students, and it seemed to me that this worked out very well. I want to thank the participants of the course for several very helpful suggestions and questions on the material which went into the notes right away.

Unexpectedly, I had to teach the same course in spring term 2021, still with quite a lot of online teaching involved. The current version is for the course in spring term 2024, where things are back to normal (?). The lecture notes are the essential material for the course.

The material covered in the notes constitutes the fundamentals of what is often called "differential geometry": The general theory of abstract smooth manifolds and smooth mappings, vector fields, tensor fields, differential forms, and integration theory. The point of view is mainly analytic, truly geometric aspects (i.e. situations in which there are local invariants like curvature) only occur in examples. I decided to start with a short discussion of submanifolds of $\mathbb{R}^n$, which, on the one hand, provides a large number of examples. On the other hand, I consider including submanifolds as important from the point of view of motivation. For submanifolds the notion of tangent spaces is much simpler and more intuitive, which is very helpful as a preparation for the definition of tangent spaces used for abstract manifolds. Submanifolds are also used to outline the fundamental principle I have follows throughout the text of using definitions that do not depend on choices (e.g. of local coordinates) and use such choices only to explicitly compute quantities that are known in advance to be well defined.

The material covered in the course is more or less standard. I have decided to include a relatively careful discussion of Lie derivatives which emphasized the relation to compatibility of tensor fields with the flow of a vector field. Moreover, there is a rather careful discussion of integration of densities (which avoids the need of orientability) with an emphasis on the volume density of a Riemann metric.

Apart from my own experience with the material and earlier lecture notes of mine I have mainly used two books for preparing the notes, namely Peter Michor's book [**Michor**] and Jack Lee's book [**Lee**]. I want to thank students form various semesters who provided me with feedback on the course and the notes and with small corrections. Particular thanks go to my colleague Michael Kunzinger who informed me that the orignal version of Section 2.11 was not correct. This error has been corrected for the 2024 version of the course.

CHAPTER 1

Manifolds

The aim of this first chapter is to develop the fundamentals of the theory of abstract smooth manifolds. The development will follow the books [**Michor**] (which does things very quickly and condensed) and [**Lee**] (which contains more detailed information and covers more material than the course). In contrast to both books, we will start with a discussion of submanifolds of $\mathbb{R}^n$. This provides lots of examples of smooth manifolds as well as motivation, which is important for understanding the abstract concepts. Many of the fundamental concepts, in particular the construction of tangent spaces and the tangent bundle, are much simpler in the setting of submanifolds. At the same time, this introduction will exhibit the shortcomings of the concept of submanifolds and the motivation for the abstract concepts to be introduced afterwards.

Motivation: Submanifolds of $\mathbb{R}^n$

In classical analysis, differentiation theory is developed on open subsets of $\mathbb{R}^n$. The fundamental purpose of any notion of a manifold is to extend differential calculus to a broader class of "spaces". In the theory of submanifolds, these "spaces" are taken to be more general subsets of $\mathbb{R}^n$. Throughout these lecture notes, "smooth" will mean C^∞, i.e. infinitely differentiable.

1.1. Submanifolds and smooth maps. The basic idea for the definition of a smooth submanifold is rather easy. First, one observes that $\mathbb{R}^k \subset \mathbb{R}^n$ is a (in general non-open) subset on which differential calculus can be introduced without problems. Second, one observes that differentiation is a local operation, so the definition should focus on local properties of a subset. Finally, there is the concept of a diffeomorphism between open subsets of $\mathbb{R}^n$ as a bijective smooth map, whose inverse is smooth, too, and the image of subset under a diffeomorphism should be as nice as the original subset. This readily leads to the definition of a smooth submanifold of $\mathbb{R}^n$. One can then introduce a concept of smooth maps between submanifolds via the classical notion for maps between open subsets.

Definition 1.1. (1) A subset $M \subset \mathbb{R}^n$ is called a *smooth submanifold of dimension* k if for any $x \in M$, there are open subsets $U, V \subset \mathbb{R}^n$ with $x \in U$ and there is a diffeomorphism $\Phi : U \to V$ such that $\Phi(U \cap M) = V \cap \mathbb{R}^k$. Here we view $\mathbb{R}^k$ as the subset of $\mathbb{R}^n$ consisting of all points whose last $n-k$ coordinates are zero. The diffeomorphism Φ is called a *local trivialization* for M.

(2) For a smooth submanifold $M \subset \mathbb{R}^n$ and $m \in \mathbb{N}$, a map $f : M \to \mathbb{R}^m$ is called *smooth* if for any $x \in M$, there is an open subset $U \subset \mathbb{R}^n$ with $x \in U$ and a smooth (in the usual sense of analysis) map $\tilde{f} : U \to \mathbb{R}^m$ such that $\tilde{f}|_{U\cap M} = f|_{M\cap U}$.

(3) Let $M \subset \mathbb{R}^n$ and $N \subset \mathbb{R}^m$ be submanifolds. Then a map $f : M \to N$ is called *smooth*, if f is smooth as a map $M \to \mathbb{R}^m$.

There are several immediate consequences of these definitions. Of course, $\mathbb{R}^k \subset \mathbb{R}^n$ is a k-dimensional smooth submanifold (take $U = V = \mathbb{R}^n$, $\Phi = \mathrm{id}$) and (2) and (3) lead to the usual smooth maps. On the other hand, open subsets of $\mathbb{R}^n$ obviously are smooth submanifolds of dimension n, and (2) and (3) recover the usual concept of smoothness in this case, too.

As a subset of $\mathbb{R}^n$, any submanifold M inherits a topology (open subsets are the intersections of open subsets of $\mathbb{R}^n$ with M). From the definitions it follows readily that any open subset of a k-dimensional submanifold is itself a k-dimensional submanifold. Finally, it is also a direct consequence of the definitions that for a smooth submanifold $M \subset \mathbb{R}^n$ of dimension k, an open subset $W \subset \mathbb{R}^n$ with $M \subset W$ and a diffeomorphism $F : W \to \tilde{W}$ onto another open subset $\tilde{W} \subset \mathbb{R}^n$, also $F(M)$ is a smooth submanifold of dimension k (see exercises).

Example 1.1. Verifying the defining properties of a submanifold usually is a rather annoying task, and we will soon meet much more efficient ways to verify that a subset $M \subset \mathbb{R}^n$ is a smooth submanifold. Thus we only present one basic example here, namely the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$. Let us first construct a local trivialization around $e_1 = (1, 0, \dots, 0) \in S^{n-1}$ and for this purpose write points in $\mathbb{R}^n$ as (t, y) with $t \in \mathbb{R}$ and $y \in \mathbb{R}^{n-1}$. Then $U := \{(t, y) : t > |y|\} \subset \mathbb{R}^n$ is open and contains e_1. Likewise, $V := \{(z, s) : z \in \mathbb{R}^{n-1}, |z| < 1, s \in \mathbb{R}, s > -1\}$ is an open subset of $\mathbb{R}^n$. Then we define $\Phi : U \to \mathbb{R}^n$ by $\Phi(t, y) := (\frac{1}{t}y, |(t, y)| - 1)$ and $\Psi : V \to \mathbb{R}^n$ by $\Psi(z, s) := (\lambda, \lambda z)$, where $\lambda = \lambda(z, s) := \frac{s+1}{\sqrt{1+|z|^2}}$. These are evidently smooth and a short computation shows that Φ has values in V, Ψ has values in U and the maps are inverse to each other. Since the second component of $\Phi(t, y)$ evidently vanishes if and only $(t, y) \in S^{n-1} \subset \mathbb{R}^n$, we have indeed constructed a local trivialization for S^{n-1} around e_1.

Now for a general point $x \in S^{n-1}$, there is an orthogonal linear map A on $\mathbb{R}^n$ such that $Ae_1 = x$. Then one defines $U_x := A(U)$ and clearly $\Phi_x := \Phi \circ A^{-1} : U_x \to V$ defines a local trivialization of S^{n-1} around x (with inverse $A \circ \Psi$).

Remark 1.1. For getting the right perspective it is important to realize that we do not intend to study the geometry of submanifolds in this course, but only do analysis on them. So in our example of the sphere $S^{n-1} \subset \mathbb{R}^n$ it is not important that we have used the round sphere. Any ellipsoid or (much more generally) the image of S^{n-1} under any diffeomorphism of $\mathbb{R}^n$ would be as good from our current point of view.

1.2. Tangent spaces and tangent maps. The basic idea of differentiation is to find a linear approximation of a map in the neighborhood of a point (and then see further how this depends on the point). For the case of maps between open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, these approximations are just linear maps between these vector spaces. (Although already in classical analysis it is better to consider these as copies of the ambient vector spaces attached to the points in the open subsets.) In the case of submanifolds, one first has to define appropriate vector spaces on which such linear approximations can be defined, and usually one will obtain different spaces for different points. Still the situation is relatively easy, since all these spaces can be realized as linear subspaces of the ambient $\mathbb{R}^n$.

Now it would be easy to obtain such a space using (the derivative of) a local trivialization. However, taking a trivialization represents a choice and using a definition of this type would require a verification that the result is independent of this choice. To avoid this, it is an important principle in the theory of manifolds to use objects that have an existence independent of choices as much as possible and use ingredients that involve choices only to establish properties and/or to explicitly compute things. We follow this principle here. Observe that Definition 1.1 in particular gives us a notion of smooth curves in a submanifold $M \subset \mathbb{R}^n$. One simply takes smooth maps from an open interval $I \subset \mathbb{R}$ to $\mathbb{R}^n$ that have values in $M \subset \mathbb{R}^n$.

Definition 1.2. Let $M \subset \mathbb{R}^n$ be a smooth submanifold and $x \in M$ a point. Then we define the *tangent space T_xM to M at x* to be the subset of $\mathbb{R}^n$ formed by the derivatives $c'(0)$ of all smooth curves $c : I \to M$, where $I \subset \mathbb{R}$ is an open interval with $0 \in I$ and c satisfies $c(0) = x$.

This definition of course makes sense for any subset of $\mathbb{R}^n$, but in general it does not lead to a linear subspace of $\mathbb{R}^n$. On the other hand, if we can prove that the definition leads to a linear subspace, then there is only one reasonable definition for the derivative of a smooth map $f : M \to N$ between submanifolds: If one wants the chain rule to hold and to recover the usual derivative for maps defined on open subsets, then the derivative of f in $x = c(0)$ has to send $c'(0)$ to $(f \circ c)'(0)$. Initially, it is unclear whether this is well defined and leads to a linear map, but we can easily prove that all that really works.

Theorem 1.2. *(1) For any k-dimensional submanifold $M \subset \mathbb{R}^n$ and any point $x \in M$, the tangent space T_xM is a linear subspace of $\mathbb{R}^n$ of dimension k.*

(2) Suppose that $f : M \to N$ is a smooth map between submanifolds. Then for each point $x \in M$, there is a unique linear map $T_xf : T_xM \to T_{f(x)}N$ such that for any smooth curve $c : I \to M$ as in Definition 1.2, we get $T_xf(c'(0)) = (f \circ c)'(0)$.

(3) If $f : M \to N$ and $g : N \to P$ are smooth maps between submanifolds, then $g \circ f : M \to P$ is smooth and for each $x \in M$, we have the chain rule $T_x(g \circ f) = T_{f(x)}g \circ T_xf : T_xM \to T_{(g\circ f)(x)}P$.

Proof. (1) We take a local trivialization $\Phi : U \to V$ for M with $x \in U$ and claim that the derivative $D\Phi(x) : \mathbb{R}^n \to \mathbb{R}^n$ restricts to a linear isomorphism $T_xM \to \mathbb{R}^k$. If $c : I \to M$ is a smooth curve as in Definition 1.2, we may assume $c(I) \subset U$ (shrink I if necessary). Then $\Phi \circ c : I \to \mathbb{R}^n$ has the property that its last $n - k$ coordinates are identically zero, so the same holds for $(\Phi \circ c)'(0) = D\Phi(c(0))(c'(0))$. This shows that $D\Phi(x)$ maps T_xM into $\mathbb{R}^k$. Conversely for $v \in \mathbb{R}^k$ we can choose an open interval $I \subset \mathbb{R}$ with $0 \in I$ such that $\Phi(x) + tv \in V$ (and thus in $V \cap \mathbb{R}^k$) for all $t \in I$. Then $c(t) := \Phi^{-1}(\Phi(x) + tv)$ is a smooth curve $c : I \to M$ as in Definition 1.2 and $(\Phi \circ c)'(0) = D\Phi(c(0))(c'(0)) = v$, which completes the proof.

(2) By definition, there is an open subset U of $\mathbb{R}^n$ with $x \in U$ and a smooth function $\tilde{f} : U \to \mathbb{R}^m$ such that $\tilde{f}|_{U\cap M} = f|_{M\cap U}$. The point about the proof is that T_xf can be obtained as a restriction of $D\tilde{f}(x)$. Taking a smooth curve $c : I \to M$ as in Definition 1.2, we may again assume that $c(I) \subset U \cap M$. Then $\tilde{f} \circ c$ is a smooth curve in $\mathbb{R}^m$ and since c has values in $U \cap M$, $\tilde{f} \circ c = f \circ c$. In particular, this has values in N, so $(\tilde{f}\circ c)'(0) = (f\circ c)'(0)$ is a well defined vector in $T_{f(x)}N$. But the usual chain rule shows that $(f\circ c)'(0) = (\tilde{f}\circ c)'(0) = D\tilde{f}(c(0))(c'(0))$. This shows that $(f\circ c)'(0)$ depends only on $c'(0)$, so $T_xf : T_xM \to T_{f(x)}N$ is well defined and it is linear since it coincides with the restriction of $D\tilde{f}(x)$.

(3) This is also proved via the smooth extensions. By definition, we have open subsets $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ with $x \in U$ and $f(x) \in V$ and smooth maps $\tilde{f} : U \to \mathbb{R}^m$ and $\tilde{g} : V \to \mathbb{R}^p$. Replacing U by $U \cap \tilde{f}^{-1}(V)$, we may assume that $\tilde{f}(U) \subset V$, and thus $\tilde{g}\circ\tilde{f} : U \to \mathbb{R}^p$ is a smooth function. For $y \in U\cap M$, we get $\tilde{g}(\tilde{f}(y)) = \tilde{g}(f(y)) = g(f(y))$, since $f(y) \in V \cap N$. Since this works in any point x, we conclude that $g \circ f$ is smooth. Moreover, locally around x, $\tilde{g} \circ \tilde{f}$ is a smooth extension, so from the proof of part (2) we know that $T_x(g \circ f)$ coincides with $D(\tilde{g} \circ \tilde{f})(x)|_{T_xM}$. By the classical chain rule, the

derivative equals $D\tilde{g}(\tilde{f}(x)) \circ D\tilde{f}(x)$. On T_xM, the second map restricts to T_xf and has values in $T_{f(x)}N$, on which the first map equals $T_{f(x)}g$, and the result follows. $\square$

In particular, in the case of smooth maps defined on open subsets, one recovers the usual derivative.

Now we can also extend the concept of diffeomorphisms to submanifolds. A *diffeomorphism* between M and N is a smooth bijective map $f : M \to N$ such that also the inverse $f^{-1} : N \to M$ is smooth. More generally, we say that $f : M \to N$ is a *diffeomorphism locally around* $x \in M$ if there is an open subset $U \subset M$ with $x \in U$ such that $f(U) \subset N$ is open and $f : U \to f(U)$ is a diffeomorphism (which makes sense since U and $f(U)$ are submanifolds). Finally, we call f a *local diffeomorphism* if it is a diffeomorphism locally around each point $x \in M$. From the chain rule it follows readily that if $f : M \to N$ is a diffeomorphism locally around $x \in M$ then $T_xf \circ T_{f(x)}f^{-1} = \mathrm{id}_{T_{f(x)}N}$ and $T_{f(x)}f^{-1} \circ T_xf = \mathrm{id}_{T_xM}$, so T_xf is a linear isomorphism. In particular, we see that this is only possible if M and N have the same dimension. Moreover, for a local diffeomorphism, all tangent maps are linear isomorphisms.

1.3. Simpler descriptions. As mentioned already, it is often rather tedious to verify the defining property of a submanifold directly. There actually are simpler characterizations that we discuss next. Motivating these simpler conditions is rather easy. In the definition of a local trivialization $\Phi : U \to V$ for a submanifold $M \subset \mathbb{R}^n$, it is natural to split the target space $\mathbb{R}^n$ (that contains V) as a product $\mathbb{R}^k \times \mathbb{R}^{n-k}$. Accordingly, we get $\Phi = (\Phi_1, \Phi_2)$, where $\Phi_1 : U \to \mathbb{R}^k$ and $\Phi_2 : U \to \mathbb{R}^{n-k}$ are smooth maps. The defining property of Φ just says that $M \cap U = (\Phi_2)^{-1}(\{0\})$, so we have a realization of $M \cap U$ as the zero locus of a smooth function with values in $\mathbb{R}^{n-k}$. In addition, we know that for each $y \in M \cap U$ the derivative $D\Phi(y)$ is a linear isomorphism, which of course implies that $D\Phi_2(y)$ is surjective. This is often phrased as the fact that $M \cap U$ is a *regular zero locus*.

Similarly, we can restrict the inverse Φ^{-1} to the open subset $W := V \cap \mathbb{R}^k$ of $\mathbb{R}^k$. This defines a smooth map $\psi := \Phi^{-1}|_W : W \to \mathbb{R}^n$ which is a bijection onto $M \cap U$. The fact that the derivative of Φ^{-1} in each point is a linear isomorphism of course implies that $D\psi(w) : \mathbb{R}^k \to \mathbb{R}^n$ is injective for each $w \in W$. Moreover, ψ is continuous as a map $W \to M \cap U$ and the map Φ_1 from above is a continuous inverse to ψ, so ψ is actually a homeomorphism $W \to M \cap U$. Such a map ψ is called a *regular local parametrization* for M.

Now we can prove that either of these two parts of a local trivialization is sufficient to make a subset $M \subset \mathbb{R}^n$ into a smooth submanifold. In addition, each of the two descriptions comes with a corresponding description of the tangent spaces of M.

THEOREM 1.3. *Let $M \subset \mathbb{R}^n$ be a subset endowed with the induced topology. Then the following conditions are equivalent:*

(1) M is a k-dimensional submanifold of $\mathbb{R}^n$.

(2) ("M admits local realizations as a regular zero locus") For each $x \in M$, there is an open subset $U \subset \mathbb{R}^n$ with $x \in U$ and a smooth function $F : U \to \mathbb{R}^{n-k}$ such that

- *$M \cap U = F^{-1}(\{0\})$*
- *For each $y \in M \cap U$, $DF(y) : \mathbb{R}^n \to \mathbb{R}^{n-k}$ is surjective.*

In this case, for each $y \in U \cap M$, we get $T_yM = \ker(DF(y))$.

(3) ("M admits local regular parametrizations") For each $x \in M$, there are open subsets $V \subset M$ with $x \in V$ and $W \subset \mathbb{R}^k$ and a smooth function $\psi : W \to \mathbb{R}^n$ such that

- *ψ defines a homeomorphism $W \to V$.*

- *For each $z \in W$, $D\psi(z) : \mathbb{R}^k \to \mathbb{R}^n$ is injective.*

In this case, for each $z \in W$, we get $T_{\psi(z)}M = \operatorname{im}(D\psi(z))$.

Proof. Our above considerations show that (1)⇒(2) and (1)⇒(3) hold.

(2)⇒(1): For a point $x \in M$ with U and F as in (2), we construct a local trivialization for M around x. As we have noted in 1.1, the image of a submanifold under a diffeomorphism is again a submanifold. Thus we may first apply a translation to assume without loss of generality that $x = 0 \in \mathbb{R}^n$. Second, $\ker(DF(0))$ is a k-dimensional subspace of $\mathbb{R}^n$ and by applying an appropriate orthogonal linear map, we may assume without loss of generality that $\ker(DF(0)) = \mathbb{R}^k \subset \mathbb{R}^n$. Let us denote by $\pi : \mathbb{R}^n \to \mathbb{R}^k$ the obvious linear projection $\pi(a_1, \dots, a_n) := (a_1, \dots, a_k)$. Now we consider the map $\Phi := (\pi|_U, F) : U \to \mathbb{R}^k \times \mathbb{R}^{n-k} \cong \mathbb{R}^n$. Its derivative in 0 is clearly given by $D\Phi(0)(v) = (\pi(v), DF(0)(v))$. But if $DF(0)(v) = 0$, then $v \in \ker(DF(0)) = \mathbb{R}^k$ and hence $\pi(v) = v$. This shows that $D\Phi(0)$ has trivial kernel and thus is a linear isomorphism.

By the inverse function theorem, there is an open neighborhood $\tilde{U}$ of 0 in U such that Φ restricts to a diffeomorphism from $\tilde{U}$ onto an open neighborhood V of $(0, 0)$. But for $y \in \tilde{U}$, we by construction have $y \in \tilde{U} \cap M$ if and only if $F(y) = 0$ which in turn is equivalent to $\Phi(y) \in V \cap \mathbb{R}^k$.

In addition, if $c : I \to \mathbb{R}^n$ is a smooth curve through y as in Definition 1.2, then we may assume $c(I) \subset U$, so $F \circ c : I \to \mathbb{R}^{n-k}$ is defined. But since c has values in M, this is identically zero and differentiating we conclude that $0 = DF(y)(c'(0))$. This shows that $T_yM \subset \ker(DF(y))$ and since these both are k-dimensional subspaces of $\mathbb{R}^n$, they have to agree.

(3)⇒(1): Again we take $x \in M$ and V, W and ψ as in (3), as well as the point $z_0 \in W$ such that $x = \psi(z_0)$. As in the above step we can translate M in such a way that $x = 0$ and then apply a rotation to assume without loss of generality that the k-dimensional subspace $\operatorname{im}(D\psi(z_0))$ coincides with $\mathbb{R}^k \subset \mathbb{R}^n$. Now we view $\mathbb{R}^{n-k}$ as the subspace of $\mathbb{R}^n$ for which the first k-coordinates are zero, and define $\Psi : W \times \mathbb{R}^{n-k} \to \mathbb{R}^n$ by $\Psi(z, y) := \psi(z) + y$. Differentiating at $(z_0, 0)$, we get $D\Psi(z_0, 0)(v_1, v_2) = D\psi(z_0)(v_1) + v_2$, so by construction $D\Psi(z_0, 0)$ is surjective and hence a linear isomorphism. Thus there is a neighborhood $\tilde{W} \subset W \times \mathbb{R}^{n-k}$ of $(z_0, 0)$ in $\mathbb{R}^n$ such that Ψ restricts to a diffeomorphism from $\tilde{W}$ onto an open subset $\tilde{U} \subset \mathbb{R}^n$.

This is not enough to ensure that the inverse of Ψ defines a local trivialization and we have to use the condition that ψ is a homeomorphism. Since $\tilde{W} \cap \mathbb{R}^k$ is open in W, $\psi(\tilde{W} \cap \mathbb{R}^k)$ is open in M. Hence there is an open subset $\tilde{V} \subset \mathbb{R}^n$ such that $\tilde{V} \cap M = \psi(\tilde{W} \cap \mathbb{R}^k)$. Then $U := \tilde{V} \cap \tilde{U}$ is an open subset of $\mathbb{R}^n$ that contains x, $\Psi^{-1}(U)$ is open in $\mathbb{R}^n$ and $\Psi^{-1}|_U : U \to \Psi^{-1}(U)$ is a diffeomorphism. If for $y \in U$, we have $\Psi^{-1}(y) = (z, 0)$, then $y = \Psi(z, 0) = \psi(z) \in M$. Conversely, for $y \in U \cap M \subset \tilde{V} \cap M$, there is a unique element $z \in \tilde{W} \cap \mathbb{R}^k$ such that $\psi(z) = \Psi(z, 0) = y$, so Ψ^{-1} indeed is a local trivialization around x.

Finally, for $z \in W$ and $v \in \mathbb{R}^k$ there is an open interval $I \subset \mathbb{R}$ containing 0 such that $z + tv \in W$ for all $t \in I$. But then $c : I \to M$, $c(t) = \psi(z + tv)$ is a curve as in Definition 1.2 with $c(0) = \psi(z)$. Since $c'(0) = D\psi(z)(v)$, we conclude that $\operatorname{im}(D\psi(z)) \subset T_{\psi(z)}M$, and since both are k-dimensional subspaces of $\mathbb{R}^n$, they have to agree. □

Example 1.3. (1) We can now get a much easier argument why $S^{n-1} \subset \mathbb{R}^n$ is a smooth submanifold. Define $F : \mathbb{R}^n \to \mathbb{R}$ as $F(x) := \langle x, x \rangle - 1$, so $S^{n-1} = F^{-1}(\{0\})$. To prove regularity, we use the chain rule to obtain $DF(x)(v) = \frac{d}{dt}|_{t=0} F(x + tv) = 2\langle x, v \rangle$.

In particular, for $x \neq 0$, we get $DF(x)(x) \neq 0$ so S^{n-1} can be globally realized as a regular zero locus. This works in the same way for ellipsoids and similar subsets.

(2) Take an open subset $U \subset \mathbb{R}^k$ and a smooth function $f : U \to \mathbb{R}^{n-k}$ and let $M \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$ be the graph of f. Thus $M = \{(x, f(x)) : x \in U\}$ and we consider $\psi : U \to \mathbb{R}^k \times \mathbb{R}^{n-k}$, $\psi(x) := (x, f(x))$. Of course, this is smooth and bijective onto M. Moreover, the first projection onto $\mathbb{R}^k$ restricts to a continuous map $M \to U$ which is evidently inverse to ψ. Thus we have found a global regular parametrization for M, which thus is a k-dimensional submanifold of $\mathbb{R}^n$. Since the first projection is smooth, we also see that $\psi : U \to M$ is a diffeomorphism, which illustrates the fact that we are on the level of analysis rather than geometry here.

(3) For $k \leq n$ consider the space $(\mathbb{R}^n)^k$ of k-tuples of vectors in $\mathbb{R}^n$. It is most convenient to view this as the space of $n \times k$ matrices by interpreting a matrix as a collection of column vectors. Identify this with $\mathbb{R}^{kn}$ and let $M \subset \mathbb{R}^{kn}$ be the subspace of k-tuples $(a_1, \dots, a_k)$ of vectors that are orthonormal, i.e. satisfy $\langle a_i, a_j \rangle = \delta_{ij}$. We claim that this is a submanifold of dimension $k(2n-k-1)/2$. To prove this, we denote by V the vector space of symmetric $k \times k$-matrices, which has dimension $k(k+1)/2$. Denoting by $\mathbb{I}$ the $k \times k$-unit matrix, we define $F : \mathbb{R}^{kn} \to V$ by $F(A) := A^t A - \mathbb{I}$. If $A = (a_1, \dots, a_k)$ then the matrix $A^t A$ has entries $\langle a_i, a_j \rangle$, so $M = F^{-1}(\{0\})$. To compute the derivative of F we again use the chain rule to write

$$DF(A)(B) = \tfrac{d}{ds}|_{s=0} F(A + sB) = \tfrac{d}{ds}|_{s=0} (A^t + sB^t)(A + sB) - \mathbb{I} = A^t B + B^t A.$$

Now for $C \in V$ put $B := \frac{1}{2} AC$. Since C is symmetric, we get $B^t = \frac{1}{2} C A^t$ and since $A^t A = \mathbb{I}$, we see that $DF(A)(B) = C$. Hence F is regular, which proves the claim.

Observe that the global realization of M as a zero locus implies that M is a closed subset of $\mathbb{R}^{nk}$. Moreover, for $A \in M$, any coefficient of A has norm ≤ 1. Thus M is bounded and hence compact by the Heine-Borel theorem. So we have found an example of a compact submanifold. Note that for $k = n$, we obtain the subspace $O(n)$ of all orthogonal $n \times n$-matrices in the space of all $n \times n$-matrices. This is a group under matrix multiplication, which obviously is a smooth map, so it is a fundamental example of a compact Lie group.

1.4. Tangent bundle and tangent map. As a first application of the simpler description of submanifolds, we show how to collect the derivatives of a smooth map in individual points together to define a smooth map. To do this, one first has to collect the tangent spaces of a submanifold in different points together in such a way that one again obtains a submanifold.

DEFINITION 1.4. (1) For a smooth submanifold $M \subset \mathbb{R}^n$ we define the *tangent bundle* $TM \subset \mathbb{R}^n \times \mathbb{R}^n$ of M as the subset $\{(x, v) : x \in M, v \in T_x M\}$.

(2) Let $M \subset \mathbb{R}^n$ and $N \subset \mathbb{R}^m$ be submanifolds and let $f : M \to N$ be a smooth map. Then we define the *tangent map* $Tf : TM \to TN$ of f by $Tf(x, v) := (f(x), T_x f(v))$.

PROPOSITION 1.4. *(1) For a smooth submanifold $M \subset \mathbb{R}^n$ of dimension k, the tangent bundle TM is a smooth submanifold of $\mathbb{R}^{2n}$ of dimension $2k$. The first projection $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ induces a smooth map $p = p_M : TM \to M$.*

(2) For a smooth map $f : M \to N$ between submanifolds, the tangent map $Tf : TM \to TN$ is smooth, too, and it satisfies $p_N \circ Tf = f \circ p_M$.

(3) For smooth maps $f : M \to N$ and $g : N \to P$ between submanifolds, we have the chain rule $T(g \circ f) = Tg \circ Tf$.

PROOF. (1) Take a point $(x, v) \in TM$. Then we know that there is an open subset $U \subset \mathbb{R}^n$ with $x \in U$ and a smooth function $F : U \to \mathbb{R}^{n-k}$ such that $M \cap U = F^{-1}(\{0\})$.

Now we define $\tilde{U} := U \times \mathbb{R}^n \subset \mathbb{R}^{2n}$ and consider the smooth map $\tilde{F} : \tilde{U} \to \mathbb{R}^{n-k} \times \mathbb{R}^{n-k}$ by $\tilde{F}(y, w) := (F(y), DF(y)(w))$. Now $\tilde{F}(y, w) = (0, 0)$ is equivalent to $y \in F^{-1}(\{0\}) = U \cap M$ and $w \in \ker(DF(y)) = T_yM$ and thus to $(y, w) \in \tilde{U} \cap TM$. Assuming this, we compute

$$D\tilde{F}(y, w)(v_1, v_2) = (DF(y)(v_1), D^2F(y)(w, v_1) + DF(y)(v_2)).$$

This readily implies that $D\tilde{F}(y, w)$ is surjective, which completes the proof of the first part. The second part is clear, since the first projection is a global smooth extension of p.

(2) Let us again take $(x, v) \in TM$. Then by assumption, there is an open subset $U \subset \mathbb{R}^n$ and a smooth map $\tilde{f} : U \to \mathbb{R}^m$ such that $\tilde{f}|_{U\cap M} = f|_{M\cap U}$. Similarly as above, we define a smooth map $\varphi : U \times \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^m$ by $\varphi(y, w) := (\tilde{f}(y), D\tilde{f}(y)(w))$. For $(y, w) \in (U \times \mathbb{R}^n) \cap TM$, we then get $\varphi(y, w) = (f(y), T_yf(w)) = Tf(y, w)$, compare with the proof of Theorem 1.2. Thus φ is a smooth extension of Tf on an open neighborhood of (x, v) and smoothness follows. The last claim is obvious from the definition of Tf.

(3) This is an obvious consequence of the chain rule from Theorem 1.2. □

1.5. Local parametrizations and smooth maps. At this stage, we have defined an analog of the derivative for smooth functions between submanifolds. Having this at hand, we move towards the passage to abstract manifolds. In the description of submanifolds via local parametrizations, the ambient space $\mathbb{R}^n$ already plays only a relatively small role. We shall see next that local parametrizations can also be used to characterize smoothness of maps between submanifolds in a way that eliminates the need to use smooth extensions to open subsets in the ambient space. To derive this, we first prove a lemma, which is of independent interest.

LEMMA 1.5. *Let $\psi : U \to V \subset M$ be a local parametrization for a smooth submanifold $M \subset \mathbb{R}^n$. Then ψ is a diffeomorphism from U onto V. Conversely, any diffeomorphism from an open subset of $\mathbb{R}^k$ onto an open subset of M is a local parametrization for M.*

PROOF. By definition, ψ is smooth as a map to $\mathbb{R}^n$ and hence also as a map to M and to V. So it remains to show that $\psi^{-1} : V \to U$ is smooth, too. For a point $x \in V$ take $z = \psi^{-1}(x)$. In the proof of Theorem 1.3, we have seen that there exists an open neighborhood $\tilde{W}$ of $(z, 0)$ in $U \times \mathbb{R}^{n-k}$ and a diffeomorphism Ψ from $\tilde{W}$ onto an open neighborhood of x in $\mathbb{R}^n$ that restricts to ψ on $\tilde{W} \cap (U \times \{0\})$. The inverse $\Phi := \Psi^{-1}$ can be decomposed as (Φ_1, Φ_2) and then $\Phi_1 : \Psi(\tilde{W}) \to \mathbb{R}^k$ is a smooth extension of ψ^{-1} to an open neighborhood of x in $\mathbb{R}^n$, thus showing that ψ^{-1} is smooth.

For the converse assume that $U \subset \mathbb{R}^k$ and $V \subset M$ are open and that $\psi : U \to V$ is a diffeomorphism. Then ψ is smooth as a map to V and hence as a map to $\mathbb{R}^n$. Moreover, the inverse of ψ is smooth and thus continuous, so ψ is a homeomorphism $U \to V$. Finally, we know that the tangent maps of a diffeomorphism are linear isomorphisms, which shows that for any $z \in U$, $D\psi(z)$ is injective as a map to $\mathbb{R}^n$. Thus ψ satisfies all properties of a local parametrization. □

Using this, the characterization of smooth maps follows rather easily.

PROPOSITION 1.5. *Let $M \subset \mathbb{R}^n$ and $N \subset \mathbb{R}^m$ be smooth submanifolds, which we consider as topological spaces with the induced topologies. Then for a continuous map $f : M \to N$, the following conditions are equivalent.*

(1) f is smooth

(2) For any $x \in M$, there are local parametrizations $\varphi : U \to M$ and $\psi : V \to N$ with $x \in \varphi(U)$ and $f(x) \in \psi(V)$ such that $\psi^{-1} \circ f \circ \varphi$ is smooth as a map from the open subset $\varphi^{-1}(f^{-1}(\psi(V))) \subset \mathbb{R}^k$ to $\mathbb{R}^\ell$.

(3) For any local parametrizations $\varphi : U \to M$ and $\psi : V \to N$ such that $f^{-1}(\psi(V)) \cap \varphi(U) \neq \emptyset$ the map $\psi^{-1} \circ f \circ \varphi$ is smooth as in (2).

Proof. If we assume that f is smooth then for any local parametrization $\varphi : U \to M$, $f \circ \varphi : U \to N$ is smooth as a composition of smooth functions. Thus also its restriction to any open subset of U is smooth. On the other hand, Lemma 1.5 shows that $\psi^{-1} : \psi(V) \to V$ is smooth, so also $\psi^{-1} \circ f \circ \varphi$ is smooth on its domain of definition, so (1) implies (3). Evidently, (3) implies (2), so it remains to show that (2) implies (1).

In the setting of (2), we know from Lemma 1.5 that $\varphi^{-1} : \varphi(U) \to U$ is smooth and thus the same holds for its restriction to any open subset of $\varphi(U)$. But now we can write the restriction of f to $f^{-1}(\psi(V))$ as $\psi \circ (\psi^{-1} \circ f \circ \varphi) \circ \varphi^{-1}|_{f^{-1}(\psi(V))}$, so we conclude that $f|_{f^{-1}(\psi(V))}$ is smooth. By definition, this means that there is a smooth extension of this restriction to an open neighborhood of x in $\mathbb{R}^n$. Since x is arbitrary, this implies that f is smooth. □

So for verifying smoothness of maps, the ambient space also is not really necessary. But then it becomes visible, that there are natural examples of spaces that admit nice local parametrizations, but for which it is unclear how to realize them as subsets of $\mathbb{R}^n$. As a simple example, let us consider the sphere $S^{n-1} \subset \mathbb{R}^n$, and define an equivalence relation on S^{n-1} by declaring each $x \in S^{n-1}$ to be equivalent to itself and to its antipodal point $-x$. Then let $\mathbb{R}P^{n-1}$ be the set of equivalence classes endowed with the quotient topology and let $\pi : S^{n-1} \to \mathbb{R}P^{n-1}$ be the obvious projection. The nice fact about this is that any 1-dimensional linear subspace of $\mathbb{R}^n$ intersects S^{n-1} in two antipodal points, so one can also view $\mathbb{R}P^{n-1}$ as the space of all these linear subspaces.

By definition of the quotient topology, a subset $U \subset \mathbb{R}P^{n-1}$ is open if and only if $\pi^{-1}(U) \subset S^{n-1}$ is open. In particular, if we take an open hemisphere in S^{n-1}, then its image under π will be an open subset of $\mathbb{R}P^{n-1}$. Hence we see that from appropriate local parametrizations of S^{n-1}, we can easily construct the analogs of local parametrizations of $\mathbb{R}P^{n-1}$. Indeed, it is clear that $\mathbb{R}P^{n-1}$ locally "looks like" S^{n-1} so it should be possible to extend differential calculus from S^{n-1} to $\mathbb{R}P^{n-1}$. Now it turns out that $\mathbb{R}P^{n-1}$ can be realized as a submanifold of $\mathbb{R}^N$ for large enough N, but all the relation to S^{n-1} and to linear subspaces of $\mathbb{R}^n$ is lost in such a picture. Consequently, there is also not too much intuitive meaning to the tangent spaces as subspaces in $\mathbb{R}^N$ and similar concepts. Finally, it is not obvious whether such a realization is unique and so on. Consequently, it is preferable to completely dispense with the concept of an ambient space, which is realized by the concept of an abstract manifold.

Abstract manifolds and smooth maps

1.6. Topological manifolds and smooth structures. The basic idea for the definition of an abstract manifold is now rather easy to guess. One takes a sufficiently nice topological space and looks for the analogs of local parametrizations. As we have seen in 1.5, such a parametrization is just a diffeomorphism from an open subset of a Euclidean space, so we may as well use the inverses of local parametrizations ("local charts") as basic ingredients. Since initially there is no notion of smoothness, these are just defined to be homeomorphisms. However, there is a simple compatibility condition, which makes sure that conditions analogous to Proposition 1.5 have the same meaning in different charts. For the definition, we only need enough charts to cover the space, but we don't want to give these charts a specific role, which leads to a more involved definition.

DEFINITION 1.6. (1) An n-dimensional *topological manifold* M is a second countable, Hausdorff topological space, which is locally Euclidean in the sense that each point $x \in M$ has an open neighborhood which is homeomorphic to an open subset of $\mathbb{R}^n$.

(2) A *chart* on a topological manifold M is a pair (U, u), where $U \subset M$ is an open subset and u is a homeomorphism from U onto an open subset $u(U) \subset \mathbb{R}^n$.

(3) For $k \in \mathbb{N} \cup \{\infty\}$, two charts (U_α, u_α) and (U_β, u_β) are called *C^k-compatible* if $U_{\alpha\beta} := U_\alpha \cap U_\beta$ is either empty or $u_{\alpha\beta} := u_\alpha \circ u_\beta{}^{-1}$ is a C^k-diffeomorphism between the open subsets $u_\beta(U_{\alpha\beta})$ and $u_\alpha(U_{\alpha\beta})$ of $\mathbb{R}^n$.

(4) A *C^k-atlas* $\mathcal{A}$ on a topological manifold M is a collection $\{(U_\alpha, u_\alpha) : \alpha \in I\}$ of mutually C^k-compatible charts on M such that $M = \cup_{\alpha \in I} U_\alpha$.

(5) Two C^k-atlases $\mathcal{A}$ and $\mathcal{B}$ on a topological manifold M are called *equivalent* if and only if each chart of $\mathcal{A}$ is C^k-compatible with each chart of $\mathcal{B}$.

(6) A *C^k-structure* on a topological manifold M is an equivalence class of C^k-atlases on M. A *C^k-manifold* is a topological manifold M together with a C^k-structure on M.

Observe that by definition, the union of any family of equivalent atlases is an atlas that is equivalent to any member of the family. In particular, a C^k-structure on M is equivalent to a maximal atlas defined by the union of all the atlases in the equivalence class. Here maximality of an atlas $\mathcal{A}$ means that any chart that is compatible with all the charts of $\mathcal{A}$ is already contained in $\mathcal{A}$. In what follows a *chart on a C^k-manifold M* will mean one of the charts of the maximal atlas corresponding to the chosen C^k-structure on M. Assume that (U, u) is such a chart and $V \subset U$ is open. Then $u(V) \subset \mathbb{R}^n$ is open and $u|_V : V \to u(V)$ is a homeomorphism, and from the definitions it follows readily, that $(V, u|_V)$ is a chart, too.

REMARK 1.6. (1) There are various ways to phrase the restrictions on the underlying topology of a topological manifold. Notice in particular (see exercises) that the Hausdorff property does not follow from the property of being locally Euclidean. Another usual formulation is to require the topology to be metrizable and separable (which implies that it is Hausdorff and second countable by standard results of topology). On the other hand, the conditions in part (1) of Definition 1.6 imply that the topology is metrizable and it is also well known that second countability implies separability. Observe that both separability and second countability imply that M has at most countably many connected components. Finally, since $\mathbb{R}^n$ is locally compact the same holds for any topological manifold.

(2) Initially, it is not clear that a topological manifold has a well defined dimension. This follows from algebraic topology, which implies that if $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ are non-empty open subsets which are homeomorphic, then $n = m$. This would still allow different connected components of M to have different dimensions, but one usually excludes this possibility by definition. Once one is in a differentiable setting, things become much easier, since for a diffeomorphism between open subsets the derivative in each point has to be a linear isomorphism.

(3) We have allowed differentiability of class C^k in the definition only for completeness. We will actually only work with class C^∞ and use "*smooth*" as an equivalent wording for C^∞. Thus we speak about smooth atlases and smooth structures, etc. This is no real restriction, since one can prove in general that any topological manifold that

admits a C^1-structure also admits a C^1-equivalent smooth structure (and even an analytic structure, see below). Moreover, if two smooth structures on M are C^1-equivalent, then they are C^∞-equivalent.

(4) Actually, the concepts in parts (3) to (6) of Definition 1.6 also make sense for the class C^ω of real analytic functions (i.e. those that can be locally written as convergent power series). Finally, it is also possible to replace $\mathbb{R}^n$ by $\mathbb{C}^n$ in the definition and then consider holomorphic analogs of the conditions in (3)–(6). This leads to the concept of *complex manifolds* on which there is a notion of holomorphic functions. We will not work with real analytic and complex manifolds in this course.

(5) It can be proved in general that a topological manifold of dimension $n \leq 3$ always does admit a smooth structure. This is not true in higher dimensions, there are (many) topological manifolds that do not admit any C^1-structure.

In Definition 1.6 we have chosen to start with a topology on M. This is because in many applications one knows the "right" topology on a space that one wants to identify as a smooth manifold in advance. However, it is also possible to start with just a set and an atlas and to also construct a topology on M from that atlas. Since we will need this for some constructions, we formulate it explicitly.

Lemma 1.6. *Let M be a set and suppose we have given a family $\{(U_\alpha, u_\alpha) : \alpha \in I\}$ of subsets $U_\alpha \subset M$ and bijections $u_\alpha : U_\alpha \to u_\alpha(U_\alpha)$ onto open subsets of $\mathbb{R}^n$ such that $M = \cup_{\alpha \in I} U_\alpha$ and*

- *The index set I is finite or countable.*
- *For any two points $x, y \in M$ with $x \neq y$, there either is an index α such that $x, y \in U_\alpha$ or there are indices α, β such that $U_\alpha \cap U_\beta = \emptyset$ and $x \in U_\alpha$ and $y \in U_\beta$.*
- *For any two indices $\alpha, \beta \in I$ with $U_{\alpha\beta} := U_\alpha \cap U_\beta \neq \emptyset$, the sets $u_\alpha(U_{\alpha\beta})$ and $u_\beta(U_{\alpha\beta})$ are open in $\mathbb{R}^n$ and the compositions $u_{\beta\alpha} := u_\beta \circ {u_\alpha}^{-1} : u_\alpha(U_{\alpha\beta}) \to \mathbb{R}^n$ and $u_{\alpha\beta} : u_\beta(U_{\alpha\beta}) \to \mathbb{R}^n$ are smooth.*

Then there is a unique topology on M making it into a topological manifold and such that $\{(U_\alpha, u_\alpha) : \alpha \in I\}$ is a smooth atlas for M.

Proof. This mainly is a sequence of elementary verifications, more details will be done in the exercises if needed. We first observe that $u_{\alpha\beta}$ and $u_{\beta\alpha}$ are inverse smooth bijections between the subsets $u_\alpha(U_{\alpha\beta})$ and $u_\beta(U_{\alpha\beta})$ so in particular, they are inverse homeomorphisms. Now we define $\mathcal{T}$ to be the set of those subsets $U \subset M$ such that for each $\alpha \in I$, $u_\alpha(U \cap U_\alpha)$ is open in $\mathbb{R}^n$. One directly verifies that this defines a topology on M for which each of the sets U_α is open in M. The second condition then easily implies that this topology is Hausdorff.

For $V \subset u_\alpha(U_\alpha)$, one obtains $u_\beta({u_\alpha}^{-1}(V) \cap U_\beta) = u_{\beta\alpha}(V \cap u_\alpha(U_{\alpha\beta}))$ and if V is open, this is open, too. Conversely, for an open subset $V \subset U_\alpha$ the image under u_α is open by definition, so $u_\alpha : U_\alpha \to u_\alpha(U_\alpha)$ is a homeomorphism. Finally, denoting open balls in $\mathbb{R}^n$ by $B_r(x)$, the set

$$\{(\alpha, y, k) : \alpha \in I, y \in \mathbb{Q}^n : B_{1/k}(y) \subset u_\alpha(U_\alpha)\}$$

is countable by the first condition. Defining $V_{(\alpha,y,k)} := {u_\alpha}^{-1}(B_{1/k}(y))$ we thus obtain a countable family of open subsets of M, which are easily seen to be a basis for the topology on M. Hence M is a topological manifold and then by construction $\{(U_\alpha, u_\alpha) : \alpha \in I\}$ is a smooth atlas on M. $\square$

1.7. Examples. (1) The single chart (U, id_U) defines a smooth atlas on any open subset $U \subset \mathbb{R}^n$, thus making it into a smooth manifold of dimension n.

(2) Let $M \subset \mathbb{R}^n$ be a k-dimensional submanifold as defined in 1.1. Then for a family of local parametrizations whose images cover M, the inverses define a smooth atlas for M. Moreover, any two atlases obtained in this way are visibly equivalent. Hence M canonically inherits the structure of a smooth manifold.

(3) Suppose that M is a smooth manifold and $U \subset M$ is an open subset. Then we claim that U is canonically a smooth manifold. To see this, take any smooth atlas $\{(U_\alpha, u_\alpha) : \alpha \in I\}$ for M. For any α, $U \cap U_\alpha$ is open in U_α, so $u_\alpha(U \cap U_\alpha)$ is open in $\mathbb{R}^n$ and u_α restricts to a homeomorphism on this subset. This shows that, with the induced subspace topology, U is a topological manifold and that the restrictions of the charts define a smooth atlas for U. From the definitions, it follows readily that starting from an equivalent atlas for M, one arrives at an equivalent atlas for U.

(4) Consider smooth manifolds M and N and the product space $M \times N$. For charts (U, u) for M and (V, v) for N, the product $U \times V$ is open in $M \times N$ and $u \times v : U \times V \to u(U) \times v(V)$ is a homeomorphism. This shows that $M \times N$ is a topological manifold, whose dimension is the sum of the dimensions of the two factors. Moreover, starting with smooth atlases for the factors, one easily constructs a smooth atlas for the product, and equivalent atlases on the factors lead to equivalent atlases for the product.

(5) It is easy to make the space $\mathbb{R}P^{n-1}$ from Section 1.5 into a smooth manifold (see exercises). Alternatively, this is a special case of the following example.

(6) To discuss a substantial example, we consider the so-called *Grassmann manifold* $Gr(k, n)$, which is defined to be the space of all k-dimensional linear subspaces of $\mathbb{R}^n$. To endow this space with a topology, we start with the space $\mathcal{V}(k, n)$ of k-tuples of linearly independent vectors in $\mathbb{R}^n$. This can be viewed as a subset of the space of $(n \times k)$-matrices with real entries. For a matrix A with linearly independent columns, we can choose k rows in A such that the corresponding $(k \times k)$-submatrix of A has non-zero determinant. Then all matrices for which that submatrix has non-zero determinant form an open neighborhood of A contained in $\mathcal{V}(k, n)$, so $\mathcal{V}(k, n)$ is an open subset of $\mathbb{R}^{nk}$. Now we define an equivalence relation $\sim$ on $\mathcal{V}(k, n)$ by defining $A \sim B$ if and only if the columns of A and B span the same linear subspace of $\mathbb{R}^n$. Clearly the set of equivalence classes is $Gr(k, n)$, so there is a natural surjection $q : \mathcal{V}(k, n) \to Gr(k, n)$ which induces a topology on $Gr(k, n)$.

Now fix a linear subspace $E \in Gr(k, n)$ and a complementary subspace $F \subset \mathbb{R}^n$. Then $\mathbb{R}^n = E \oplus F$ and we denote by π_E and π_F the corresponding projections, so in particular $\ker(\pi_E) = F$. Then we consider the subset

$$U := \{Z \in Gr(k, n) : Z \cap F = \{0\}\}.$$

Now we can find a linear map $\mathbb{R}^n \to \mathbb{R}^k$ whose kernel is F. (Choose bases for E and F, which together form a basis of $\mathbb{R}^n$, send the basis vectors of E to the standard basis of $\mathbb{R}^k$ and those of F to 0.) This is represented by a $(k \times n)$-matrix C and we get $A \in q^{-1}(U)$ if and only if $\det(CA) \neq 0$. Thus $q^{-1}(U)$ is open, so U is open by definition of the quotient topology. For $Z \in U$, the restriction $\pi_E|_Z$ has trivial kernel by construction, so π_E restricts to a linear isomorphism $Z \to E$. Now we define $u(Z) \in L(E, F)$, the space of linear maps from E to F, by $u(Z) := \pi_F \circ (\pi_E|_Z)^{-1}$. Conversely, for a linear map $g : E \to F$, we define a linear subspace $Z \subset \mathbb{R}^n$ as $Z := \{v + g(v) : v \in E\}$. One immediately verifies that this has dimension k, lies in U and that the construction defines an inverse to u, so $u : U \to L(E, F)$ is bijective.

Now fix bases $\{v_i\}$ for E and $\{w_j\}$ for F as above and take $(z_1, \dots, z_k) \in q^{-1}(U)$. Then the $(n \times n)$-matrices A with columns v_i and w_j and B with columns z_i and w_j are both invertible. The product $A^{-1}B$ then has a block form $\begin{pmatrix} C & 0 \\ D & \mathbb{I} \end{pmatrix}$ with blocks of sizes k and $n-k$ and with $\mathbb{I}$ denoting a unit matrix. The first k columns contain the coefficients of the z_i in the expansion as linear combinations of the v_i and the w_j. But this says that C and D are the matrix representations of $\pi_E|_Z$ and $\pi_F|_Z$ in the given bases, respectively. So for $Z = q(z_1, \dots, z_k)$ the matrix expansion of $u(Z)$ is DC^{-1}. This shows that $u \circ q : q^{-1}(U) \to L(E, F)$ is continuous (and indeed smooth and even real analytic), so u is continuous. Conversely, for $g \in L(E, F)$, we can realize $u^{-1}(g)$ as $q(v_1 + g(v_1), \dots, v_k + g(v_k))$, so this is continuous, too, and thus u is a homeomorphism.

We have actually seen now that $u \circ q$ is smooth, while u^{-1} can be written as $q \circ \varphi$, where φ is evidently smooth. But this implies that for a second chart (V, v), we can write the chart-change $v \circ u^{-1}$ as the restriction of $v \circ q \circ \varphi$ to an open subset of $u(U)$, so this is smooth. So it only remains to verify that the topology on $Gr(k, n)$ is Hausdorff and second countable. But for two subspaces $E_1, E_2 \in Gr(k, n)$, we can clearly find a linear subspace $F \subset \mathbb{R}^n$ that is complementary to both E_1 and E_2. This implies that E_1 and E_2 both are contained in the domain U of the chart defined by (E_1, F). Since U is homeomorphic to $L(E_1, F)$, we find disjoint open neighborhoods of these two points. On the other hand, $Gr(k, n)$ can be covered by finitely many charts. For example for any k-element subset X of the standard basis of $\mathbb{R}^n$, we can take the subspaces E_X spanned by these vectors and F_X spanned by the remaining elements of the standard basis. Taking the preimages of balls with rational center and radius under the chart maps clearly gives rise to a countable basis for the topology on $Gr(k, n)$.

1.8. Smooth maps. The main reason for using charts rather than local parametrizations is that charts immediately give rise to *local coordinates* on a manifold. Indeed, if (U, u) is a chart for M, then we can write the map $u : U \to \mathbb{R}^n$ in components as $u = (u^1, \dots, u^n)$ and each $u^i : U \to \mathbb{R}$ is continuous. Given a function $f : M \to \mathbb{R}$, we can then consider $f \circ u^{-1} : u(U) \to \mathbb{R}$ and this is called the *coordinate representation of* f with respect to the chart U. This clearly generalizes without problems to functions with values in $\mathbb{R}^m$.

A similar idea applies to maps between smooth manifolds, but one has to be a bit careful with domains of definition in this case. So assume that $F : M \to N$ is a function between smooth manifolds. Then we can proceed as above, using charts (U, u) for M and (V, v) for N provided that $F(U) \subset V$. Under this assumption, we can simply consider $v \circ F \circ u^{-1} : u(U) \to \mathbb{R}^m$, where m is the dimension of N. As above, this is called the (local) coordinate representation of F with respect to the two charts. The condition that $F(U) \subset V$ is not a big deal if F is continuous. In this case, $F^{-1}(V)$ is open in M and hence for any open subset $U \subset M$, $U \cap F^{-1}(V)$ is open in U and as observed in 1.6, we can restrict charts to open subsets.

In analogy to Proposition 1.5 we can now define smoothness of maps between manifolds via smoothness of coordinate representations.

Definition 1.8. Let M and N be smooth manifolds and let $F : M \to N$ be a map.

(1) F is called *smooth* if and only if for any point $x \in M$, there are charts (U, u) for M and (V, v) for N such that $x \in U$, $F(U) \subset V$ and such that $v \circ F \circ u^{-1} : u(U) \to \mathbb{R}^m$ is smooth as a map on the open subset $u(U) \subset \mathbb{R}^n$.

(2) F is called a *diffeomorphism* if and only if F is smooth and bijective and the inverse map $F^{-1} : N \to M$ is smooth, too.

(3) One says that F is a *diffeomorphism locally around a point* $x \in M$ if F is smooth and there is an open subset $U \subset M$ with $x \in U$ such that $F(U)$ is open in N and $F|_U : U \to F(U)$ is a diffeomorphism. The map F is called a *local diffeomorphism* if it is a diffeomorphism locally around each point of M.

The following are simple consequences of the definitions:

- A composition of two smooth maps is smooth.
- A composition of two diffeomorphisms is a diffeomorphism.
- Restrictions of smooth maps to open subsets are smooth.
- If for a map $F : M \to N$, there is an open covering $\{U_\alpha : \alpha \in I\}$ of M such that $F|_{U_\alpha} : U_\alpha \to N$ is smooth for each α, then F is smooth.
- For any chart (U, u), the map $u : U \to u(U)$ is a diffeomorphism.

Next, we claim that a smooth map $F : M \to N$ is continuous. Indeed, for $x \in M$ we find charts (U, u) for M with $x \in U$ and (V, v) for N with $F(U) \subset V$ such that $v \circ F \circ u^{-1} : u(U) \to v(V)$ is smooth and thus continuous. Thus we can write $F|_U$ as $v^{-1} \circ (v \circ F \circ u^{-1}) \circ u$ and since u and v^{-1} are continuous, we conclude that $F|_U$ is continuous. Thus we can cover M by open sets U_i such that $F|_{U_i}$ is continuous for all i, so $F : M \to N$ is continuous.

Given a smooth map $F : M \to N$ and any chart (V, v) for N, we conclude that $F^{-1}(V)$ is open and the restriction of $v \circ F$ to any open subset of $F^{-1}(V)$ is smooth. This in turn implies that for any chart (U, u) for M, $v \circ F \circ u^{-1}$ is smooth on the open subset $u(U \cap F^{-1}(V)) \subset \mathbb{R}^n$.

Finally, suppose that $W \subset M$ is open and that w is a diffeomorphism onto an open subset $w(W) \subset \mathbb{R}^n$. Then w is a homeomorphism and our last observation says that for each chart (U, u) for M, $w \circ u^{-1} : u(U \cap W) \to w(W)$ is smooth, while smoothness of w^{-1} implies that $u \circ w^{-1} : w(U \cap W) \to u(U)$ is smooth. But this says that (W, w) is compatible to any chart on M and hence itself is a chart on M. Thus charts are exactly the diffeomorphisms from open subsets of M onto open subsets of $\mathbb{R}^n$.

Example 1.8. (1) From Proposition 1.5 we conclude that in the case of submanifolds (and hence in particular of open subsets in $\mathbb{R}^n$), we recover the concept of smoothness from Definition 1.2.

(2) From the charts in Example (4) of 1.7 it is evident that on a product $M \times N$ of smooth manifolds the projections $\pi_M : M \times N \to M$ and $\pi_N : M \times N \to N$ are smooth maps.

(3) The construction of charts in Example (6) of 1.7 shows that the map $q : \mathcal{V}(k, n) \to Gr(k, n)$ used there is smooth.

Remark 1.8. Having the notion of diffeomorphism at hand, we can now also discuss the question of uniqueness of smooth structures on topological manifolds. There are simple (and rather misleading) examples that show that there are smooth structures on simple topological manifolds like $\mathbb{R}$, which are different from the standard structure, say the one induced by the single chart $u : \mathbb{R} \to \mathbb{R}$, $u(x) = x^3$. However, from our above considerations we see that u actually is a diffeomorphism to the standard structure, so these are not really different.

Still it may happen that a topological manifold admits several non-diffeomorphic smooth structures ("*exotic smooth structures*"). For example, by results of J. Milnor, there are 28 different smooth structures on S^7, and from dimension 7 on, many spheres carry (finitely many) exotic smooth structures. More drastically, on the topological manifold $\mathbb{R}^4$, there are uncountably many different smooth structures, while in all other dimensions $\mathbb{R}^n$ has just one smooth structure up to diffeomorphism. Fascinating as they are, these results do not really imply that one will run into choices of smooth structures on well known spaces. The proofs of existence of exotic smooth structures usually consist of constructing a smooth manifold in some way and then it is hard work to show that this manifold is homeomorphic to some well known space and that the structure is not diffeomorphic to the standard one.

1.9. Smooth functions and partitions of unity. We next study real valued smooth functions on smooth manifolds. Our first main aim here is the proof of a result that is technically very important for many constructions with smooth functions. The space of smooth functions $M \to \mathbb{R}$ will be denoted by $C^\infty(M, \mathbb{R})$. From Definition 1.8, we see that $f : M \to \mathbb{R}$ is smooth if and only if for each $x \in M$ there is a chart (U, u) for M with $x \in U$ such that $f \circ u^{-1} : u(U) \to \mathbb{R}$ is smooth. This immediately implies that $C^\infty(M, \mathbb{R})$ is a vector space and an associative algebra under point-wise operations. Likewise, if $f : M \to \mathbb{R}$ is smooth such that $f(x) \neq 0$ for all $x \in M$, then $\frac{1}{f}$ is smooth, too. Recall that for a real valued function f defined on a topological space X, the *support* $\operatorname{supp}(f)$ of f is defined as the closure $\overline{\{x : f(x) \neq 0\}}$. Otherwise put, the complement of $\operatorname{supp}(f)$ is the maximal open subset of X on which f vanishes identically.

We also need a few notions from topology. On the one hand, there are several weakenings of the concept of compactness. In particular, a topological space X is called a *Lindelöff space* if any open covering $\{U_i : i \in I\}$ of X admits a countable subcovering, i.e. there is a sequence $(i_n)_{n \in N}$ in I such that $\cup_{n \in \mathbb{N}} U_{i_n} = X$. It is a simple result of general topology (see exercises) that any second countable space is a Lindelöff space.

On the other hand, for a topological space X, a family $\{A_i : i \in I\}$ of subsets $A_i \subset X$ is called *locally finite* if each point $x \in X$ has an open neighborhood U in X, which intersects only finitely many of the sets A_i.

Now suppose that M is a smooth manifold and we have given a family $\{\varphi_i : i \in I\}$ of smooth functions $\varphi_i : M \to \mathbb{R}$ such that the family $\{\operatorname{supp}(\varphi_i) : i \in I\}$ of supports is locally finite. Then for each $x \in M$, only finitely many of the numbers $\varphi_i(x)$ are non-zero, so $\sum_{i \in I} \varphi_i(x)$ is well defined. In this way, we get a function $M \to \mathbb{R}$, which we write as $\sum_{i \in I} \varphi_i$. Now given $x \in M$, there is an open neighborhood U of x in M on which only finitely many of the φ_i are not identically zero. This shows that $(\sum_{i \in I} \varphi_i)|_U$ equals a finite sum of smooth functions and thus is smooth, too. In this way, we obtain an open covering of M by sets on which $\sum_{i \in I} \varphi_i$ is smooth, so we see from Section 1.8 that it is a smooth function $M \to \mathbb{R}$.

Finally, we recall from analysis that given $r_1, r_2 \in \mathbb{R}$ with $0 < r_1 < r_2$, there is a smooth function $h : \mathbb{R} \to \mathbb{R}$ with values in $[0, 1]$ such that $h(t) = 1$ for $t \leq r_1$ and $h(t) = 0$ for $t \geq r_2$ ("cutoff function"). To construct such a function, one mainly needs a smooth function $f : \mathbb{R} \to \mathbb{R}$ such that $f(t) = 0$ for all $t \leq 0$ and $f(t) > 0$ for all $t > 0$, for example $f(t) = e^{-1/t}$ for $t > 0$. Then one just defines $h(t) := \frac{f(r_2 - t)}{f(r_2 - t) + f(t - r_1)}$ and checks that this has the required properties. Using these observations, we can now formulate the result:

THEOREM 1.9. *Let M be a smooth manifold and let $\{U_i : i \in I\}$ be an open covering of M. Then there is a family $\{\varphi_n : n \in \mathbb{N}\}$ of smooth functions on M such that*

(i) *For all $x \in M$ and $n \in \mathbb{N}$, $\varphi_n(x) \geq 0$.*
(ii) *For each $n \in \mathbb{N}$ there is $i_n \in I$ such that $\operatorname{supp}(\varphi_n) \subset U_{i_n}$.*
(iii) *The family $\{\operatorname{supp}(\varphi_n) : n \in \mathbb{N}\}$ of supports is locally finite.*
(iv) *The sum $\sum_{n \in \mathbb{N}} \varphi_n$ is the constant function 1.*

Proof. **Step 1**. For $i \in I$ and $x \in U_i$ there is a function $f_{x,i} : M \to \mathbb{R}$ with non-negative values such that $f_{x,i}(x) > 0$ and $\operatorname{supp}(f_{x,i}) \subset U_i$:

Choose a chart (U, u) for M with $x \in U$ and (without loss of generality) $u(x) = 0 \in \mathbb{R}^n$. Since $u(U \cap U_i)$ is open, there is $\epsilon > 0$ such that $B_\epsilon(0) \subset u(U \cap U_i)$. Using a cutoff function for $r_1 = \epsilon/2$ and $r_2 = 2\epsilon/3$ we get a smooth function $h : \mathbb{R}^n \to \mathbb{R}$ with values in $[0, 1]$ which is identically one on $\overline{B}_{\epsilon/2}(0)$ and has support contained in $B_\epsilon(0)$. Now we define $f_{x,i} : M \to \mathbb{R}$ as $h \circ u$ on U and as 0 on $M \setminus U$. Since $u^{-1}(\operatorname{supp}(h))$ is a closed set contained in U, its complement in M is open and, together with U, defines an open covering of M. But $f_{x,i}$ is clearly smooth on U and on $M \setminus u^{-1}(\operatorname{supp}(h))$ (since it is identically zero there), so $f_{x,i} : M \to \mathbb{R}$ is smooth.

Step 2. Pass to a countable and locally finite family:

By construction, $\{y : f_{x,i}(y) \neq 0\} \subset U_i$ and these sets form an open covering of M. Since M is a Lindelöff space, countably many of these open sets cover M, and we denote the corresponding functions by f_n for $n \in \mathbb{N}$ and the indices by i_n, so $\operatorname{supp}(f_n) \subset U_{i_n}$. For $n \in \mathbb{N}$, we next define $W_n := \{x \in M : f_n(x) > 0, f_k(x) < 1/n \quad \forall k < n\}$. Taking a smooth function $f : \mathbb{R} \to [0, \infty)$ such that $f(t) > 0$ iff $t > 0$ as above, we then define

$$g_n(x) := f_n(x)f(1/n - f_1(x)) \cdots f(1/n - f_{n-1}(x)).$$

As a product of finitely many smooth functions with non-negative values, this is smooth and has non-negative values. Moreover, $g_n(x) > 0$ if and only if $x \in W_n$, so $\operatorname{supp}(g_n) = \overline{W_n} \subset U_{i_n}$. For $x \in M$, there by construction is an index $n \in \mathbb{N}$ such that $f_n(x) > 0$. Denoting by n_0 the minimal index with this property, we get $x \in W_{n_0}$ and hence $M = \cup_{n\in\mathbb{N}} W_n$. But for $\alpha := f_{n_0}(x) > 0$, the set $U := \{y \in M : f_{n_0}(y) > \alpha/2\}$ is an open neighborhood of x in M. Taking $N \in \mathbb{N}$ with $1/N < \alpha/2$ we see that for $y \in U$ and $n \geq N$, we get $f_{n_0}(y) > 1/n$. If $n > n_0$, this implies $g_n(y) = 0$ and hence $U \cap \operatorname{supp}(g_n) = \emptyset$. Thus the family $\overline{W_n}$ is locally finite.

Step 3. Construct the functions φ_n.

Since the family $\operatorname{supp}(g_n)$ is locally finite, we know that $g(x) := \sum_{n\in\mathbb{N}} g_n(x)$ defines a smooth function. By constructions all summands are non-negative and at least one of them is positive, so $g(x) > 0$ for all x. Hence $1/g$ is a smooth function on M, and for $n \in \mathbb{N}$, we define $\varphi_n(x) := \frac{g_n(x)}{g(x)}$. Since $\operatorname{supp}(\varphi_n) = \operatorname{supp}(g_n) = \overline{W_n}$ by construction, we see that this family satisfies all claimed properties. □

A family of functions (without restrictions on the index set) which satisfies conditions (i), (iii), and (iv) is called a *partition of unity*. Condition (ii) is phrased as the fact that the partition of unity is *subordinate to the open covering* $\{U_i : i \in I\}$ of M. A typical application of partitions of unity is the following result.

Corollary 1.9. *Let M be a smooth manifold, $U \subset M$ open and $A \subset M$ closed, such that $A \subset U$. Then for any smooth function $f : U \to \mathbb{R}$, there is a smooth function $\tilde{f} : M \to \mathbb{R}$ such that $\tilde{f}|_A = f|_A$.*

Proof. Putting $U_1 := U$ and $U_2 := M \setminus A$, $\{U_1, U_2\}$ is an open covering of M, so Theorem 1.9 gives us a family $\{\varphi_n : n \in \mathbb{N}\}$ of smooth functions. Define $K \subset \mathbb{N}$ to be the set of those n for which $\operatorname{supp}(\varphi_n) \subset U_2$, and define $\varphi : M \to \mathbb{R}$ by $\varphi := \sum_{n\notin K} \varphi_n$. Then this is a smooth function with values in $[0, 1]$ and we claim that $\operatorname{supp}(\varphi) \subset \cup_{n\notin K} \operatorname{supp}(\varphi_n) \subset U_1$. The second inclusion is clear, since $n \notin K$ implies $\operatorname{supp}(\varphi_n) \subset U_1$. To see the first inclusion, it suffices to show that $\cup_{n\notin K} \operatorname{supp}(\varphi_n)$ is closed, since it evidently contains the set $\{y : \varphi(y) > 0\}$. But for $x \notin \operatorname{supp}(\varphi_n)$ for all $n \notin K$, we find an open neighborhood V of x that intersects only finitely many of the sets $\operatorname{supp}(\varphi_n)$. Since the latter are closed, the complement of their union in V is open and thus an open neighborhood of x contained in the complement of $\cup_{n\notin K}\operatorname{supp}(\varphi_n)$. Finally, φ can also be written as $1-\sum_{n\in K}\varphi_n$. But for $n\in K$, φ_n vanishes identically on A, so φ is identically one on A.

Now given f, we can define $\tilde{f}$ as $f\varphi$ on U and as zero on $M\setminus\operatorname{supp}(\varphi)$. Since these are two open subsets that together cover M and on which $\tilde{f}$ is evidently smooth, the claim follows. □

Tangent spaces and the tangent bundle

The definition of the tangent spaces for an abstract manifold is significantly more complicated than in the case of submanifolds. It would be rather easy to define tangent spaces via local charts, but this would lead to a notion in which one has to make a choice in order to express a tangent vector. Consequently, in the further development there is the constant need to check independence of that choice, which makes the approach unsatisfactory. So we will rather put a bit of effort into a definition of tangent vectors that does not need choices, and for which it is evident that the tangent space is a vector space. These tangent vectors can then be described in local charts.

1.10. Tangent spaces in $\mathbb{R}^n$. The basis for the general definition is an alternative description of the tangent spaces of $\mathbb{R}^n$ (or of submanifolds). The definition for submanifolds in 1.2 does not carry over to abstract manifolds, since there is no evident notion of the derivative of a smooth curve in a point. One could use a definition via an equivalence relation on curves (see Section 1.13 below) but this again comes with the need of choosing representatives. The idea that works for abstract manifolds is to view a tangent vector at a point as a "direction into which smooth real valued functions can be differentiated". It turns out the the operators of directional derivatives in a point can be easily characterized algebraically, and one can then *define* a tangent vector as an operator of that type.

For a point $a\in\mathbb{R}^n$, a vector $v\in\mathbb{R}^n$ and $f\in C^\infty(\mathbb{R}^n,\mathbb{R})$ put $v_a(f):=Df(a)(v)$. This clearly defines a linear map $v_a:C^\infty(\mathbb{R}^n,\mathbb{R})\to\mathbb{R}$ and the product rule shows that $v_a(fg)=v_a(f)g(a)+f(a)v_a(g)$. We call a linear map with this property a *a derivation at* a. From the definition it follows readily that the derivations at a form a linear subspace of the space $L(C^\infty(\mathbb{R}^n,\mathbb{R}),\mathbb{R})$ of linear maps. The following result provides crucial motivation.

Lemma 1.10. *For any point $a\in\mathbb{R}^n$ the map $v\mapsto v_a$ defines a linear isomorphism from $\mathbb{R}^n$ onto the space of all derivations at a.*

Proof. Since $Df(a):\mathbb{R}^n\to\mathbb{R}$ is linear, we see that $v\mapsto v_a$ is linear. For the ith coordinate function x^i, we obtain $v_a(x^i)=v^i$, the ith component of v. Thus $v_a=w_a$ implies $v=w$ and we can complete the proof by showing that any derivation at a is of the form v_a for some $v\in\mathbb{R}^n$.

So let $D:C^\infty(\mathbb{R}^n,\mathbb{R})\to\mathbb{R}$ be a derivation at a. For the constant function 1, we get $1=1\cdot 1$ and thus $D(1)=D(1\cdot 1)=2D(1)$, so D vanishes on 1 and hence on all constant functions. Now the result follows from a simple instance of Taylor's theorem. We can write $f(x)=f(a)+\int_0^1\frac{d}{dt}f(a+t(x-a))dt$. Computing the derivative as $\sum_i\frac{\partial f}{\partial x^i}(a+t(x-a))(x^i-a^i)$, we can take the sum and the factors x^i-a^i out of the integral. Defining $h_i(x):=\int_0^1\frac{\partial f}{\partial x^i}(a+t(x-a))dt$, we thus obtain $f(x)=f(a)+\sum_{i=1}^n h_i(x)(x^i-a^i)$. Inserting this, we get

$$D(f)=0+\textstyle\sum_i D(h_i\cdot(x^i-a^i))=\sum_i(D(h_i)(a^i-a^i)+h_i(a)D(x^i-a^i)).$$

In the last term, we can drop a_i and form the vector v with components $v^i = D(x^i)$. Observing that $h_i(a) = \frac{\partial f}{\partial x^i}(a)$, we conclude that $D(f) = \sum_i \frac{\partial f}{\partial x^i}(a)D(x^i) = v_a(f)$. □

Guided by this result, one could try to define tangent spaces on manifolds via derivations at points on $C^\infty(M,\mathbb{R})$. This works out, but it would make some of the subsequent proofs rather clumsy. Hence it is better to do yet another technical step to localize things.

1.11. Germs. This is a notion of "functions defined locally around a point" which is important in many areas of mathematics. The concept makes sense for many classes of functions, but we will only consider smooth ones. Given a smooth manifold M and a point $x \in M$, we consider the set of all pairs (U, f), where $U \subset M$ is an open subset with $x \in U$ and $f : U \to \mathbb{R}$ is a smooth function. So these are just smooth functions defined on some open neighborhood of x. On this set, we define a relation by $(U, f) \sim (V, g)$ if and only if there is an open subset $W \subset U \cap V$ with $x \in W$ such that $f|_W = g|_W$. One immediately verifies that this is an equivalence relation. The set of equivalence classes is denoted by $C^\infty_x(M,\mathbb{R})$ and called the space of all *germs at x of smooth functions.* Observe that all representatives of a germ have the same value at x. Thus any germ has a well defined value in x, but there are no well defined values in other points.

PROPOSITION 1.11. *(1) Point-wise addition and multiplication of functions induce well defined operations on $C^\infty_x(M,\mathbb{R})$ making it into a commutative associative algebra over $\mathbb{R}$.*

(2) Let $F : M \to N$ be a smooth map between manifolds. Then for each point $x \in M$, composition with F induces an algebra homomorphism $F^*_x : C^\infty_{F(x)}(N,\mathbb{R}) \to C^\infty_x(M,\mathbb{R})$. If F is a diffeomorphism or the inclusion of an open subset, then F^*_x is an isomorphism of algebras for all $x \in M$.

PROOF. (1) Given representatives (U, f) and (V, g) for two germs, one adds respectively multiplies the restrictions of the functions to $U \cap V$ and takes the class of the result to define the operations. A simple direct check shows that this is well defined. Having that, given finitely many germs, one can always choose representatives defined on the same open subset of M, and doing this all claimed properties of the operations are obviously satisfied.

(2) Given a germ at $F(x)$, choose a representative (U, f). Since F is continuous, $F^{-1}(U)$ is an open subset in M that contains x, so we can form the class of $(F^{-1}(U), f \circ F)$ in $C^\infty_x(M,\mathbb{R})$. Clearly, this is independent of the choice of representative, so we obtain a well defined map F^* as claimed. On the level of representatives, composition with F is obviously compatible with point-wise operations, so we conclude that F^* is an algebra homomorphism.

If F is a diffeomorphism, one puts $G := F^{-1}$. On the level of representatives, composition with G is inverse to composition with F and thus $G^*_{F(x)}$ is inverse to F^*_x. Finally, consider the inclusion $i : V \hookrightarrow M$ of an open subset. Then i^* is induced by sending a representative (U, f) for a germ at x to $(U \cap V, f|_{U\cap V})$, so this clearly induces an isomorphism. □

A smooth function $f \in C^\infty(M,\mathbb{R})$ determines a germ at any point $x \in M$, namely the class of (M, f). This leads to an algebra homomorphism $C^\infty(M,\mathbb{R}) \to C^\infty_x(M,\mathbb{R})$. For smooth functions, this homomorphism is always surjective (which is the reason why we could work with derivations on $C^\infty(M,\mathbb{R})$). Given a representative (U, f) for a germ at x, we can use balls in a chart around x to construct an open subset $V \subset M$ such that $x \in V$ and $\overline{V} \subset U$. By Corollary 1.9, there is a smooth function $\tilde{f} : M \to \mathbb{R}$, which agrees with f on $\overline{V}$ and thus on V and hence determines the same germ at x as f. This result becomes wrong already for real analytic functions.

1.12. Tangent spaces and tangent maps in a point for abstract manifolds. Motivated by Lemma 1.10, we now define tangent spaces on abstract manifolds. For simplicity, in what follows we ignore the domain of definition for representatives of germs and simply denote them in the same way as functions.

Definition 1.12. Let M be a smooth manifold, let $x \in M$ be a point and let $C^\infty_x(M,\mathbb{R})$ be the algebra of germs of smooth functions at x. Then we define the *tangent space* T_xM to M at x to be the space of linear maps $X_x : C^\infty_x(M,\mathbb{R}) \to \mathbb{R}$, which are a derivation at x in the sense that $X_x(fg) = X_x(f)g(x) + f(x)X_x(g)$.

We have defined T_xM as a subset of the space $L(C^\infty_x(M,\mathbb{R}),\mathbb{R})$ which is a vector space under point-wise operations. But it is obvious that a linear combination of derivations at x is again a derivation at x, so T_xM actually is a linear subspace and hence canonically a vector space. To see that this has the properties one would expect, we have to study tangent maps. Keeping in mind that derivations at x represent directional derivatives, the definition of tangent maps is again forced by the chain rule (read backward this time), which says that $D(f \circ F)(x)(v) = Df(F(x))(DF(x)(v))$.

Theorem 1.12. *Let M and N be smooth manifolds and let $F : M \to N$ be a smooth map. For a point $x \in M$ let $F^* = F^*_x : C^\infty_{F(x)}(N,\mathbb{R}) \to C^\infty_x(M,\mathbb{R})$ be the homomorphism from Proposition 1.11.*

(1) For $X_x \in T_xM$, $X_x \circ F^*_x : C^\infty_{F(x)}(N,\mathbb{R}) \to \mathbb{R}$ is a derivation at $F(x)$. Denoting this by $T_xF(X_x)$, we obtain a linear map $T_xF : T_xM \to T_{F(x)}N$.

(2) If F is a diffeomorphism or the embedding of an open subset, then T_xF is a linear isomorphism for each $x \in M$. In particular, if M has dimension n, then for each $x \in M$, the vector space T_xM has dimension n.

(3) If $G : N \to P$ is another smooth map, then we get the chain rule

$$T_x(G \circ F) = T_{F(x)}G \circ T_xF : T_xM \to T_{G(F(x))}P.$$

In particular, if F is a diffeomorphism with inverse G then $T_{F(x)}G$ is inverse to T_xF.

Proof. (1) From Proposition 1.11, we know that $F^*(fg) = F^*(f)F^*(g)$ and clearly $F^*(f)(x) = f(F(x))$, and thus $X_x \circ F^*$ is a derivation at $F(x)$. Linearity of T_xF then is obvious.

(2) Again from Proposition 1.11, we know that F^* is an isomorphism of algebras, which implies that also its inverse is an isomorphism of algebras. The proof of (1) shows that composition with this inverse maps derivations at $F(x)$ to derivations at x, and hence defines an inverse to T_xF. For the last statement, let (U,u) be a chart for M containing x. Since U is open in M, we see that $T_xM \cong T_xU$ and since $u : U \to U(U)$ is a diffeomorphism, this is in turn isomorphic to $T_{u(x)}u(U)$. But from Lemma 1.10 (whose proof applies to germs without problems), we know that this coincides with the usual tangent space which has dimension n.

(3) On representatives, the map $(G \circ F)^*$ is induced by $f \mapsto f \circ G \circ F$, which immediately implies that $(G \circ F)^* = F^* \circ G^*$. But then

$$T_x(G \circ F)(X_x) = (X_x \circ F^*) \circ G^* = T_xF(X_x) \circ G^* = T_{F(x)}G(T_xF(X_x)),$$

and the chain rule follows. For the second statement, one only has to observe that for the identity map, also id^* is the identity and so all tangent maps are the identity map. $\square$

As we have observed already, any $f \in C^\infty(M, \mathbb{R})$ determines a germ at x, so a tangent vector $X_x \in T_xM$ also defines a linear map $C^\infty(M, \mathbb{R}) \to \mathbb{R}$. By construction, this is again a derivation at x in the sense that $X_x(fg) = X_x(f)g(x) + f(x)X_x(g)$. As announced already, one can also view the tangent space as the space of all those derivations.

PROPOSITION 1.12. *Let M be a smooth manifold and $x \in M$ a point. Let $D : C^\infty(M, \mathbb{R}) \to \mathbb{R}$ be a linear map such that $D(fg) = D(f)g(x) + f(x)D(g)$. Then there is a unique $X_x \in T_xM$ such that $D(f) = X_x(f)$.*

PROOF. To see this, we have to show that $D(f) = D(g)$ if f and g have the same germ at x. By definition, this means that there is an open neighborhood U of $x \in M$ such that $f|_U = g|_U$, so $f - g$ vanishes identically on U. From 1.9, we know that there is a smooth function $\varphi : M \to \mathbb{R}$ with values in $[0, 1]$ such that $\varphi(x) = 1$ and $\operatorname{supp}(\varphi) \subset U$. But then $\varphi(f - g)$ vanishes identically, so linearity of D implies that

$$0 = D(\varphi(f - g)) = D(\varphi)(f - g)(x) + \varphi(x)D(f - g) = 0 + D(f) - D(g),$$

so $D(f) = D(g)$. $\square$

1.13. Description via curves. A description of tangent vectors via curves is also available for abstract manifolds. This also shows that in the case of submanifolds of $\mathbb{R}^n$ we recover the notions of tangent spaces and tangent maps from 1.2. Let $I \subset \mathbb{R}$ be an open interval with $0 \in I$ and $c : I \to M$ a smooth map to a smooth manifold M and put $c(0) = x \in M$. If $f : U \to \mathbb{R}$ represents a germ at x, $f \circ c$ is a smooth real valued function defined locally around 0 on $\mathbb{R}$, so we can form $(f \circ c)'(0) \in \mathbb{R}$. Evidently, this defines an element of T_xM, which we denote by $c'(0)$. Similarly, one gets $c'(t) \in T_{c(t)}M$ for all $t \in I$, but this will not be needed at the moment. Observe that by the chain rule and Lemma 1.10 this recovers the usual derivative of curves if M is an open subset of $\mathbb{R}^n$.

Next assume that $F : M \to N$ is a smooth map. Then for a representative f of a germ at $F(x)$ we by definition get

$$T_xF(c'(0))(f) = c'(0)(f \circ F) = (f \circ F \circ c)'(0) = (F \circ c)'(0)(f).$$

Thus $T_xF(c'(0)) = (F \circ c)'(0)$, which is the expected behavior of the tangent map. Now we can apply this to a local chart (U, u) for M with $x \in U$. We get $T_xu(c'(0)) = (u \circ c)'(0)$ and in the proof of Theorem 1.12 we have observed that T_xu is a linear isomorphism. This shows that any element of T_xM can be written as $c'(0)$ for an appropriate curve c. Indeed, for $v \in \mathbb{R}^n$ the curve $c_v(t) := u^{-1}(u(x) + tv)$ is defined on an open neighborhood of zero and satisfies $T_xu(c_v'(0)) = v$.

On the other hand, we see that for two curves c_1, c_2 with $c_1(0) = c_2(0) = x$ as above, we have $c_1'(0) = c_2'(0)$ if and only if for one or equivalently any chart (U, u) for M with $x \in U$, we obtain $(u \circ c_1)'(0) = (u \circ c_2)'(0)$. This clearly defines an equivalence relation on such curves and the set of equivalence classes is isomorphic to T_xM. Moreover, if M is a submanifold of $\mathbb{R}^m$, then of course $(u \circ c_1)'(0) = (u \circ c_2)'(0)$ holds if and only if, viewed as curves to $\mathbb{R}^m$, c_1 and c_2 have the same derivative at 0. Thus, for submanifolds of $\mathbb{R}^m$ we recover the tangent spaces and tangent maps from 1.2. In the case of abstract manifolds, dealing with equivalence classes cannot be avoided by forming a derivative as a curve in some ambient space, however.

1.14. Tangent vectors in local charts. It is now easy to interpret tangent vectors in local charts. For a local chart (U, u) for M and $x \in U$, $T_x u : T_x M \to T_{u(x)}\mathbb{R}^n \cong \mathbb{R}^n$ is a linear isomorphism. In the picture of derivations, the elements of the standard basis for $\mathbb{R}^n$ correspond to the partial derivatives of a function in a point. Mapping them to $T_x M$ with $T_{u(x)}u^{-1}$, the results act on a smooth function $f : M \to \mathbb{R}$ as the partial derivatives of $f \circ u^{-1}$ in the point $u(x)$. Now $f \circ u^{-1}$ is exactly the local coordinate representation of f with respect to the chart (U, u) from 1.8. Therefore, these elements are usually denoted by $\frac{\partial}{\partial u^i}|_x$, where, as before, we write u in components as $u = (u^1, \dots, u^n)$.

It is also easy to relate these pictures for different charts. Take two charts (U_α, u_α) and (U_β, u_β) with $x \in U_{\alpha\beta} := U_\alpha \cap U_\beta$. Recall that $u_\alpha(U_{\alpha\beta})$ and $u_\beta(U_{\alpha\beta})$ are open subsets in $\mathbb{R}^n$, and we have the chart change $u_{\alpha\beta} := u_\alpha \circ {u_\beta}^{-1} : u_\beta(U_{\alpha\beta}) \to u_\alpha(U_{\alpha\beta})$ which is a diffeomorphism. On $U_{\alpha\beta}$, we by construction have $u_\alpha = u_{\alpha\beta} \circ u_\beta$ and hence $T_x u_\alpha = T_{u_\beta(x)}u_{\alpha\beta} \circ T_x u_\beta$. Moreover, we know that $T_{u_\beta(x)}u_{\alpha\beta}$ is just the ordinary derivative $Du_{\alpha\beta}(u_\beta(x))$. Alternatively, we can write the relation as $(T_x u_\beta)^{-1} = (T_x u_\alpha)^{-1} \circ Du_{\alpha\beta}(u_\beta(x))$. Denoting the vector in $\mathbb{R}^n$ that $X_x \in T_x M$ corresponds to under u_α by $v_\alpha = (v_\alpha^1, \dots, v_\alpha^n)$ and similarly for u_β, we get

$$v_\alpha^i = \textstyle\sum_j \partial_j u_{\alpha\beta}^i(u_\beta(x))v_\beta^j, \tag{1.1}$$

where ∂_j denotes the jth partial derivative. Thus we see that we can interpret elements in the tangent space $T_x M$ as equivalence classes of pairs $((U, u), v)$, where (U, u) is a chart with $x \in U$ and v is vector in $\mathbb{R}^n$. Then formula (1.1) describes an equivalence relation on the set of these pairs and $T_x M$ is the set of equivalence classes.

Alternatively, we may compute

$$T_x u_\alpha \left(\tfrac{\partial}{\partial {u_\beta}^i}|_x\right) = Du_{\alpha\beta}(u_\beta(x))(e_i) = \partial_i u_{\alpha\beta}(u_\beta(x)),$$

which leads to

$$\tfrac{\partial}{\partial {u_\beta}^i}|_x = \textstyle\sum_j \partial_i u_{\alpha\beta}^j(u_\beta(x))\tfrac{\partial}{\partial {u_\alpha}^j}|_x. \tag{1.2}$$

The difference between these two rules just expresses the familiar fact from linear algebra that the transformation between two bases of a vector space and the transformation between coordinate vectors between these two bases are represented by inverse matrices.

1.15. Tangent bundle and tangent maps. We can also rephrase the computations above as saying that $T_x u_\alpha \circ (T_x u_\beta)^{-1} = Du_{\alpha\beta}(u_\beta(x))$ and this visibly depends smoothly on x. Thus we are led to a natural idea for obtaining a tangent bundle for abstract manifolds. Given a smooth manifold M, we define TM *as a set* to be the disjoint union of the tangent spaces $T_x M$ for all points in M. This comes with a natural map $p : TM \to M$ that sends $T_x M$ to x for each point $x \in M$. Using Lemma 1.6 and the above considerations, we can make this into a manifold.

THEOREM 1.15. *For any smooth manifold M, the space TM can be naturally made into a smooth manifold such that $p : TM \to M$ is a smooth map.*

PROOF. Start with a countable atlas $\{(U_\alpha, u_\alpha) : \alpha \in I\}$ for M. For each α, consider the set $p^{-1}(U_\alpha)$ of tangent spaces at points in U_α. Define $Tu_\alpha : p^{-1}(U_\alpha) \to u_\alpha(U_\alpha) \times \mathbb{R}^n$ by $Tu_\alpha(X_x) := (u_\alpha(x), T_x u_\alpha(X_x))$ for $x \in U_\alpha$ and $X_x \in T_x M$. We claim that the family $\{(p^{-1}(U_\alpha), Tu_\alpha) : \alpha \in I\}$ satisfies the conditions of Lemma 1.6. Clearly, $u_\alpha(U_\alpha) \times \mathbb{R}^n$ is open in $\mathbb{R}^{2n}$ and since $T_x u_\alpha$ is a bijection $T_x M \to \mathbb{R}^n$, we see that Tu_α is bijective. It is also clear that the sets $p^{-1}(U_\alpha)$ cover TM.

Next, $p^{-1}(U_\alpha) \cap p^{-1}(U_\beta) = p^{-1}(U_{\alpha\beta})$, which is mapped to the open set $u_\alpha(U_{\alpha\beta}) \times \mathbb{R}^n$ by Tu_α. Finally, $Tu_\alpha \circ (Tu_\beta)^{-1}$ by construction sends (z, v) to $(u_{\alpha\beta}(z), Du_{\alpha\beta}(z)(v))$

which is evidently smooth. The condition ensuring the Hausdorff property in Lemma 1.6 is not satisfied in general, but we can easily see that the topology on TM induced as in the proof of Lemma 1.6 is Hausdorff: If two points of TM do not lie in a common chart, then they are in different tangent spaces, and so their base points can be separated by open subsets in M, whose preimages under p are disjoint open subsets of TM. Hence Lemma 1.6 provides a smooth manifold structure on TM for which $\{(p^{-1}(U_\alpha), Tu_\alpha) : \alpha \in I\}$ is an atlas. In the charts $(p^{-1}(U_\alpha), Tu_\alpha)$ for TM and (U_α, u_α) for M, p is just given as the first projection, so $p : TM \to M$ is smooth.

Finally, one has to check that this is independent of the choice of the countable atlas we started with. But for any chart (U, u) on M, the above considerations show that $(p^{-1}(U), Tu)$ is compatible with each of the charts $(p^{-1}(U_\alpha), Tu_\alpha)$. This shows that starting from an equivalent countable atlas for M, one arrives at an equivalent atlas for TM, so the smooth structure on TM is canonical. □

Now we can go ahead and define tangent maps in general. Suppose that $F : M \to N$ is a smooth map between manifolds. Then by definition, TM and TN are just the union of the tangent spaces and we define $TF : TM \to TN$ by the fact that on T_xM, TF is given by T_xF. In particular, this means that $TF(T_xM) \subset T_{F(x)}N$ or, otherwise put, that $p \circ TF = F \circ p$, where we denote the projections on both tangent bundles by p. This also implies that $TF^{-1}(p^{-1}(V)) = p^{-1}(F^{-1}(V))$ for any open subset $V \subset N$.

PROPOSITION 1.15. *(1) For a smooth map $F : M \to N$, also the tangent map $TF : TM \to TN$ is smooth. More precisely, local coordinate representations of TF with respect to charts on TM and TN induced by charts of M and N are given by the local coordinate representation of F and its derivative.*

(2) For smooth maps $F : M \to N$ and $G : N \to P$, we have the chain rule $T(G \circ F) = TG \circ TF$.

PROOF. (1) For a point $x \in M$ we know that there are charts (U, u) for M with $x \in U$ and (V, v) for N such that $U \subset F^{-1}(V)$. Then $p^{-1}(U) \subset TF^{-1}(p^{-1}(V))$, and we have the smooth local coordinate representation $v \circ F \circ u^{-1} : u(U) \to v(V)$ of F. For $z \in u(U)$ and $x = u^{-1}(z)$, we already know that $T_z(v \circ F \circ u^{-1}) = T_{F(x)}v \circ T_xF \circ T_zu^{-1}$ and that $T_zu^{-1} = (T_xu)^{-1}$. Finally, since $v \circ F \circ u^{-1}$ maps between open subsets of $\mathbb{R}^n$'s, the tangent map coincides with the ordinary derivative. But this exactly says that for $(z, w) \in u(U) \times \mathbb{R}^n$ we get

$$Tv \circ TF \circ (Tu)^{-1}(z, w) = (v \circ F \circ u^{-1}(z), D(v \circ F \circ u^{-1})(z)(w)).$$

This proves both that TF is smooth and the claim about local coordinate representations. The statement in (2) is then a direct consequence of part (3) of Theorem 1.12. □

EXAMPLE 1.15. Consider the product $M \times N$ of two smooth manifolds M and N, which is a smooth manifold by part (4) of 1.7. Moreover, we know that the projections $\pi_M : M \times N \to M$ and $\pi_N : M \times N \to N$ are smooth, so we have their tangent maps $T\pi_M : T(M \times N) \to TM$ and $T\pi_N : T(M \times N) \to TN$. Using these as components, we obtain a map $(T\pi_M, T\pi_N) : T(M \times N) \to TM \times TN$. But now for charts (U, u) for M and (V, v) for N, we have obtained a chart $(U \times V, u \times v)$ for $M \times N$, and this in turn defines a chart $(p^{-1}(U \times V), T(u \times v))$ for $T(M \times N)$. On the other hand, the induced charts $(p^{-1}(U), Tu)$ for TM and $(p^{-1}(V), Tv)$ for TN together give rise to a chart $(p^{-1}(U) \times p^{-1}(V), Tu \times Tv)$ for $TM \times TN$. But from the definitions it is obvious that $(T\pi_M \times T\pi_N)(p^{-1}(U \times V)) \subset p^{-1}(U) \times p^{-1}(V)$ and that $(Tu \times Tv) \circ (T\pi_M \times T\pi_N) \circ$

$T(u \times v)^{-1}$ simply sends $((x, y), (w_1, w_2))$ to $((x, w_1), (y, w_2))$. This easily implies that $T\pi_M \times T\pi_N$ is bijective and admits local smooth inverses, so it is a diffeomorphism.

Special smooth maps

To conclude this chapter we study the general concept of submanifolds and a few special types of smooth maps.

1.16. Submanifolds. The following definition generalizes Definition 1.1 in an obvious way.

DEFINITION 1.16. Let M be a smooth manifold of dimension n. A subset $N \subset M$ is called a *k-dimensional submanifold* of M if for each $x \in N$, there is a chart (U, u) for M with $x \in U$ such that $u(U \cap N) = u(U) \cap \mathbb{R}^k$. As before, we view $\mathbb{R}^k \subset \mathbb{R}^n$ as the subspace of points for which the last $n - k$ coordinates are zero.

Charts as in the definition are referred to as *submanifold charts* for N. Loosely speaking, a submanifold looks like a linear subspace in appropriate charts. (One only requires existence of submanifold charts, nothing is required about what the intersection with N looks like in other charts. It is a nice exercise to observe that this image will always be a k-dimensional submanifold of $\mathbb{R}^n$.) As one might expect at this point, a submanifold $N \subset M$ is itself a manifold.

PROPOSITION 1.16. *Let M be a smooth manifold and $N \subset M$ a submanifold. Then N itself is a manifold.*

PROOF. We endow N with the subspace topology, which makes it into a second countable Hausdorff space. By definition, we find a family of submanifold charts $\{(U_\alpha, u_\alpha) : \alpha \in I\}$ such that $N \subset \cup_{\alpha \in I} U_\alpha$. For each $\alpha \in I$, $U_\alpha \cap N$ is open in N and the restriction $u_\alpha|_{U_\alpha \cap N}$ defines a homeomorphism onto to subset $u_\alpha(U_\alpha) \cap \mathbb{R}^k$, which by definition is an open subset of $\mathbb{R}^k$. To prove that $\{(U_\alpha \cap N, u_\alpha|_{U_\alpha \cap N}) : \alpha \in I\}$ is an atlas for N, it suffices to see that the chart changes are smooth. Observe first that $(U_\alpha \cap N) \cap (U_\beta \cap N) = U_{\alpha\beta} \cap N$. Now the smooth map $u_{\alpha\beta} : u_\beta(U_{\alpha\beta}) \to u_\alpha(U_{\alpha\beta})$ restricts to a smooth map from the open subset $u_\beta(U_{\alpha\beta}) \cap \mathbb{R}^k$ to $\mathbb{R}^n$. But by construction, this has values in the open subset $u_\alpha(U_{\alpha\beta}) \cap \mathbb{R}^k$ of $\mathbb{R}^k$, and hence is smooth as a map $u_\beta(U_{\alpha\beta}) \cap \mathbb{R}^k \to u_\beta(U_{\alpha\beta}) \cap \mathbb{R}^k$, and this is the chart change between the restricted charts.

The same argument shows that for any submanifold chart (U, u) for N, the chart $(U \cap N, u|_{U \cap N})$ is compatible with each chart of our atlas. Hence the smooth structure on N does not depend on the choice of the initial family of submanifold charts and thus is canonical. $\square$

If one considers a submanifold $N \subset M$ as an abstract manifold, then the inclusion $i : N \to M$ is smooth and a topological embedding (since $i : N \to i(N) \subset M$ is a homeomorphism by definition). Therefore, one often calls such objects *embedded submanifolds* in particular to distinguish from weaker notions that we will discuss below.

1.17. Inverse function theorem and constant rank. We have already noted in Theorem 1.12 that for a diffeomorphism $F : M \to N$, any tangent map T_xF is a linear isomorphism. Of course, this also holds if F is a diffeomorphism locally around x. Now we first observe that the inverse function theorem also holds for maps between manifolds. Indeed, let $F : M \to N$ be a smooth map and $x \in M$ a point such that $T_xF : T_xM \to T_{F(x)}N$ is a linear isomorphism. Then we can find charts (U, u) for M

and (V, v) for N such that $x \in U \subset F^{-1}(V)$ and such that $v \circ F \circ u^{-1} : u(U) \to v(V)$ is smooth. We know already that $D(v \circ F \circ u^{-1})(u(x)) = T_{F(x)}v \circ T_xF \circ (T_xu)^{-1}$, so this is a linear isomorphism. Hence we can apply the classical inverse function theorem to obtain an open subset $W \subset u(U)$ such that $v \circ F \circ u^{-1}$ restricts to a diffeomorphism from W onto an open subset of $v(V)$. Composing with v^{-1} and u on appropriate subsets, which are diffeomorphisms onto their images, we conclude that F restricts to a diffeomorphism from the open subset $u(W) \subset M$ onto its image. In particular, this shows that local diffeomorphisms from Definition 1.8 are exactly those smooth maps, for which each tangent map is a linear isomorphism.

For a smooth map $F : M \to N$ between general manifolds (of different dimensions) and a point $x \in M$, the only invariant of the linear map $T_xF : T_xM \to T_{F(x)}N$ is its rank, i.e. the dimension of its image. The unifying feature of the special types of smooth maps we will discuss here is that they have *constant rank*, i.e. the ranks of the tangent maps T_xF are the same for all points $x \in M$. Of course, this also applies to local diffeomorphisms, where the rank of T_xF equals $\dim(M) = \dim(N)$ in each point. Now it turns out that one can nicely adapt charts to maps of constant rank.

Theorem 1.17. *Let M and N be smooth manifolds of dimension n and m and let $F : M \to N$ be a smooth map such that for each $x \in M$, the tangent map T_xF has rank k (so $k \leq \min\{m, n\}$). Then for each $x \in M$ there are local charts (U, u) for M and (V, v) for N with $x \in U$ and $F(x) \in V$ for which the local coordinate representation of F has the form $(z^1, \dots, z^n) \mapsto (z^1, \dots, z^k, 0, \dots, 0)$.*

Proof. For $x \in M$ we start with a chart (U_1, u_1) for M with $x \in U_1$ and $u_1(x) = 0$. The kernel of $T_0(F \circ {u_1}^{-1})$ by definition is a subspace of $\mathbb{R}^n$ of dimension $n - k$, and composing u_1 with an appropriate linear map if necessary, we may assume that $\ker(T_0(F \circ {u_1}^{-1})) \cap \mathbb{R}^k = \{0\}$. Similarly, we choose a chart (V_1, v_1) for N with $F(x) \in V_1$, $v_1(F(x)) = 0$ and such that the image of $T_x(v_1 \circ F)$ is the k-dimensional subspace $\mathbb{R}^k \subset \mathbb{R}^m$. By construction, $D(v_1 \circ F \circ {u_1}^{-1})(0) : \mathbb{R}^n \to \mathbb{R}^m$ restricts to a linear isomorphism $\mathbb{R}^k \to \mathbb{R}^k$. Writing coordinates in $\mathbb{R}^n$ as z^i and writing $v_1 \circ F \circ {u_1}^{-1}$ in components as $(f^1, \dots, f^m)$, we conclude that the matrix $A := \left(\frac{\partial f^i}{\partial z^j}(0)\right)_{i,j=1,\dots,k}$ is invertible.

Now consider $\varphi : u_1(U_1) \to \mathbb{R}^n$ with components $\varphi^i = f^i$ for $i \leq k$ and $\varphi^j = z^j$ for $j > k$. Hence $\varphi(0) = 0$ and in blocks with sizes k and $n-k$, the derivative $D\varphi(0)$ has the form $\left(\begin{smallmatrix} A & * \\ 0 & \mathbb{I} \end{smallmatrix}\right)$. Since A is invertible, this is invertible, too, so φ restricts to a diffeomorphism between open neighborhoods of zero. Let $U_2 \subset M$ be the preimage of such an open neighborhood under u_1 and define $u_2 := \varphi \circ u_1|_{U_2}$. Then (U_2, u_2) is a chart for M and by construction, writing $v_1 \circ F \circ {u_2}^{-1}$ in components, we obtain $(z^1, \dots, z^k, g^{k+1}, \dots, g^m)$ for certain functions g^j. This implies that in *any* point $z \in u_2(U_2)$ the first k rows of the derivative $D(v_1 \circ F \circ {u_2}^{-1})(z)$ are the first k unit vectors. However, by assumption the tangent maps of F all have rank k, while the tangent maps of v_1 and ${u_2}^{-1}$ are all invertible, so we conclude that $D(v_1 \circ F \circ {u_2}^{-1})(z)$ has to have rank k for each z. Since the first k rows are linearly independent, the other $m - k$ rows must be linear combinations of these, so we conclude that $\frac{\partial g^i}{\partial z^j}(z) = 0$ for $i, j > k$. But this implies that the functions g^i do not depend on the last $n-k$ coordinates, so $g^j(z^1, \dots, z^n) = g^j(z^1, \dots, z^k, 0, \dots, 0)$ on $u_2(U_2)$. Now we consider

$$W := \{y = (y^1, \dots, y^m) \in v_1(V_1) : (y^1, \dots, y^k, 0, \dots, 0) \in u_2(U_2)\}$$

which evidently is an open subset of $\mathbb{R}^m$. Now define $\psi : W \to \mathbb{R}^m$ by putting $\psi^i := y^i$ for $i \leq k$ and $\psi^j(y) := y^j - g^j(y^1, \dots, y^k, 0, \dots, 0)$ for $j > k$. The derivative $D\psi(0)$

evidently is lower triangular with all entries on the main diagonal equal to 1, so this is invertible. Thus ψ restricts to a diffeomorphism between open neighborhoods of zero in $\mathbb{R}^m$. Now we define V the be the preimage of such a neighborhood under v_1 and define $v := \psi \circ v_1 : V \to \mathbb{R}^m$, so (V, v) is a chart for N with $F(x) \in V$. Finally, we define $U := U_2 \cap F^{-1}(V)$ and $u := u_2|_U$, so (U, u) is a chart for M with $x \in U$. But then

$$v \circ F \circ u^{-1}(z) = \psi(z^1, \dots, z^k, g^{k+1}(z), \dots, g^m(z)),$$

so the first k components are just $(z^1, \dots, z^k)$. For $j > k$, the jth component is $g^j(z) - g^j(z_1, \dots, z_k, 0, \dots, 0)$ and we have observed already that these vanish. □

In particular, we see that this gives a generalized version of the result on local regular zero sets in Theorem 1.3.

Corollary 1.17. *Let $F : M \to N$ be a smooth map between manifolds of dimension n and m such that T_xF has rank k for all $x \in M$. Then for each $y \in F(M) \subset N$, the preimage $F^{-1}(\{y\}) \subset M$ is a submanifold of dimension $n - k$.*

Proof. Suppose that $x \in M$ is such that $F(x) = y$, and take charts (U, u) for M and (V, v) for N as in Theorem 1.17 with $x \in U$, $u(x) = 0$ and $y \in V$. Then we see that for $x_1 \in U$, we have $F(x_1) = F(x)$ if and only if the first k coordinates of $u(x_1)$ are zero, so renumbering coordinates, this provides a submanifold chart around x. □

1.18. Submersions. A smooth map $F : M \to N$ between smooth manifolds is called a *submersion* if all its tangent maps are surjective. Of course, a submersion $F : M \to N$ can only exist if $\dim(M) \geq \dim(N)$ and in the case of equal dimension submersions are exactly local diffeomorphisms. Of course, if F is a submersion then T_xF has rank $\dim(N)$ for each x, so F has constant rank. There are some additional features compared to general constant rank maps. In particular, surjective submersions should be viewed as the right concept of "quotient maps" for smooth manifolds.

Proposition 1.18. *Let $F : M \to N$ be a smooth submersion between smooth manifolds of dimensions n and $m \leq n$. Then we have*

(1) For $x \in M$ and each chart (V, v) for N with $F(x) \in V$, there is a local chart (U, u) for M with $x \in U$ such that the local coordinate representation of F is given by $(z^1, \dots, z^n) \mapsto (z^1, \dots, z^m)$.

(2) For each $y \in F(M) \subset N$, $F^{-1}(\{y\}) \subset M$ is a submanifold of dimension $n - m$.

(3) For each $y \in F(M)$ there is an open neighborhood W of y in N and a smooth map $\sigma : W \to M$ such that $F \circ \sigma = \mathrm{id}_W$. ("$F$ admits local smooth sections.")

(4) If F is surjective, then for any manifold P a map $G : N \to P$ is smooth if and only if $G \circ F : M \to P$ is smooth. ("Universal property of surjective submersions")

Proof. (1) follows directly from the proof of Theorem 1.17, noting that there is no need to modify the chart on N in the case of a submersion. Likewise, (2) is just a specialization of Corollary 1.17.

To prove (3) choose a chart (V, v) for N with $y \in V$ and a point $x \in M$ such that $F(x) = y$. Then take a chart (U, u) as in (1) with $x \in U$ and $U \subset F^{-1}(V)$. Writing $u(x) = (z^1, \dots, z^n)$, there is an open neighborhood $\widetilde{W}$ of $(z^1, \dots, z^m) \in \mathbb{R}^m$ such that for $w \in \widetilde{W}$ we have $(w, z^{m+1}, \dots, z^n) \in u(U)$. Since the latter point is mapped to w by $v \circ F \circ u^{-1}$, we see that $v(y) \in \widetilde{W} \subset v(V)$, so $W := v^{-1}(\widetilde{W})$ is an open neighborhood of y in N. Now define $\sigma : W \to M$ by $\sigma(\tilde{y}) := u^{-1}(v(\tilde{y}), z^{m+1}, \dots, z^n)$. This is evidently smooth and has the required property.

(4): If G is smooth, then $G \circ F$ is smooth as a composition of smooth maps. Conversely, assume that $G \circ F$ is smooth and take a point $y \in N$. Then by (3), there is an open neighborhood W of y in N and a smooth map $\sigma : W \to M$ such that $F \circ \sigma = \mathrm{id}_W$. But then we can write $G|_W$ as $(G \circ F) \circ \sigma$, and this is again smooth as a composition of smooth maps. As observed in 1.8 this implies that G is smooth. □

1.19. Immersions. This is the concept dual to submersions. A smooth map $F : M \to N$ is an immersion if for each $x \in M$ the tangent map T_xF is injective. Of course, an immersion can only exist if $\dim(N) \geq \dim(M)$ and for equal dimensions, immersions are exactly local diffeomorphisms. It is also clear that for a submanifold $N \subset M$, the inclusion $i : N \to M$ is an immersion. Now we can prove that locally any immersion is of that form and characterize the immersions, for which this is true globally.

PROPOSITION 1.19. *Let $F : M \to N$ be an immersion between manifolds M and N of dimensions n and $m \geq n$. Then we have:*

(1) For each $x \in M$ and any chart (U, u) with $x \in U$, there is a chart (V, v) for N with $F(x) \in V$ such that the local coordinate representation $v \circ F \circ u^{-1}$ has the form $(z^1, \dots, z^n) \mapsto (z^1, \dots, z^n, 0, \dots, 0)$.

(2) For each $x \in M$, there is an open neighborhood U of x in M such that $F(U)$ is a submanifold in N and such that $F|_U : U \to F(U)$ is a diffeomorphism.

(3) If F defines a homeomorphism from M onto $F(M)$, then $F(M) \subset N$ is a submanifold and $F : M \to F(M)$ is a diffeomorphism.

PROOF. (1) This easily follows from the inverse function theorem (see exercises).

(2) Take charts (U, u) and $(\tilde{V}, \tilde{v})$ as in (1) such that $F(U) \subset \tilde{V}$. (Replace U by $U \cap F^{-1}(\tilde{V})$ if needed.) Then the image of $\tilde{v} \circ F \circ u^{-1}$ is just $u(U) \subset \mathbb{R}^n \subset \mathbb{R}^m$. Since $u(U)$ is open in $\mathbb{R}^n$, there is an open set $W \subset \mathbb{R}^m$ such that $W \cap \mathbb{R}^n = u(U)$. Define $V := \tilde{v}^{-1}(\tilde{v}(\tilde{V}) \cap W) \subset \tilde{V}$ and put $v = \tilde{v}|_V$. Then $v(V) \cap \mathbb{R}^n = u(U)$, so (V, v) is a global submanifold chart for $F(U)$. But with respect to the induced chart of $F(U)$ and the chart (U, u), the local coordinate representation of F is the identity map, so $F : U \to F(U)$ is a diffeomorphism.

(3) Given $x \in M$, take a neighborhood U as in (2) and a submanifold chart $(\tilde{V}, \tilde{v})$ for $F(U) \subset N$ with $F(x) \in \tilde{V}$. By assumption, $F(U)$ is open in $F(M)$, so we find an open subset $W \subset N$ such that $W \cap F(M) = F(U)$. Putting $V := \tilde{V} \cap W$ and $v := \tilde{v}|_V$, we obtain $F(M) \cap V = F(U) \cap V$, so we see that (V, v) is a submanifold chart for $F(M) \subset N$. By assumption $F : M \to F(M)$ is bijective and from (2) we know that it is a diffeomorphism locally around each point $x \in M$. This shows that the inverse is smooth, so $F : M \to F(M)$ is a diffeomorphism. □

In view of this result, the images of injective immersions are often called *immersed submanifolds*. General injective immersions show rather subtle behavior, which can already be seen from the case of immersions $I \to M$, where I is an open interval in $\mathbb{R}$. By definition such an immersion is just a smooth curve c, which is regular in the sense that $c'(t) \neq 0$ for all $t \in I$. Taking $I = (0, 1)$ is is easy to construct an example for which c is not a topological embedding by arranging c in such a way that $\lim_{t \to 1} c(t) = c(t_0)$ for some $t_0 \in (0, 1)$. Likewise, by parametrizing a figure eight by an open interval in two different ways, one obtains immersions $i_1, i_2 : I \to \mathbb{R}^2$ with $i_1(I) = i_2(I) =: J$ but such that for functions $f : J \to \mathbb{R}$ smoothness of $f \circ i_1$ is not equivalent to smoothness of $f \circ i_2$. This shows that the term "immersed submanifold" is rather problematic (since there are two ways of making J into an immersed submanifold, which lead to different "structures"), but it is quite common.

An interesting generalization of submanifolds is provded by the concept of *initial submanifolds*. A typical example of this situation is again provided by a regularly parametrized curve which at the first sight looks very badly behaved. We consider S^1 as the set $\{z \in \mathbb{C} : |z| = 1\}$ and for a positive real number α, we take the smooth curve $t \mapsto (e^{it}, e^{i\alpha t})$ in $S^1 \times S^1$, which is a torus. If α is rational, this defines a closed curve and one easily shows that this is an embedded submanifold in $S^1 \times S^1$. If α is irrational, however, one easily shows that this defines an injection $c : \mathbb{R} \to S^1 \times S^1$. With a bit of work, one shows that $c(\mathbb{R})$ is a dense subset of $S^1 \times S^1$. Indeed, whenever on takes a local submanifold chart (V, v) for $c(I)$ for a sufficiently small open interval $I \subset \mathbb{R}$, the curve $c(\mathbb{R})$ "returns" to V infinitely many times, coming arbitrarily close to $c(I)$. This in particular shows that the topology on $c(\mathbb{R})$ induced from $S^1 \times S^1$ is very different from the standard topology on $\mathbb{R}$.

Similar things happen for initial submanifolds: Take a subset $N \subset M$. For a point $x \in N$ and an open subset $U \subset M$, define $C_x(U \cap N)$ to be the set of those points $y \in U \cap N$ for which there is a smooth curve $c : I \to U$ with values in $U \cap N$ that connects y to x. One then calls N an *initial submanifold* of dimension k in M if for each $x \in N$ there is a local chart (U, u) for M with $x \in U$ such that $u(C_x(U \cap N)) = u(U) \cap \mathbb{R}^k$. One can then take the set N and endow it with a new topology, which has as a basis the subsets of $C_x(U \cap N)$ which are open under u for such charts. Calling the result $\tilde{N}$, the resulting topology is finer than the subspace topology on N induced from M, so in particular, it is Hausdorff. Further, one shows that each connected component of $\tilde{N}$ is second countable, so if $\tilde{N}$ has at most countably many connected components, then its topology satisfies the conditions required for a topological manifold. But then one can use the restrictions of charts as above to the subsets $C_x(U \cap N)$ as an atlas on $\tilde{N}$, making it into a smooth manifold of dimension k. Then the obvious map $i : \tilde{N} \to N$ which "puts the set back into M" is continuous and smooth. As in part (3) of Proposition 1.19, one shows that N is a true submanifold of M if $i : \tilde{N} \to N$ is a homeomorphism.

This "inclusion" then has a nice universal property: Let P be any manifold and let $F : P \to M$ be a map which has values in $N \subset M$. Then by construction, there is a map $\tilde{F} : P \to \tilde{N}$ such that $F = i \circ \tilde{F}$. Obviously, smoothness of $\tilde{F}$ implies that F is smooth as a map to N, and the universal property says that the converse holds, too, so if F is smooth as a map to M, then also $\tilde{F}$ is smooth. A proof of this fact and more information on initial submanifolds can be found in Sections 2.10–2.15 of [**Michor**].

Remark 1.19. There are general results on existence of immersions and embeddings into Euclidean spaces. The starting point for these are results on existence of finite atlases. It is clear that any compact manifold admits a finite atlas. More generally, one can observe that from two charts (U, u) and (V, v) with $U \cap V = \emptyset$, one can easily construct a chart defined on $U \cup V$. Then one can use topological dimension theory to prove that any manifold admits a finite atlas (consisting of disconnected charts in general). Appropriately extending the local coordinates to functions defined globally on M, and taking these as components, one obtains an embedding of M into $\mathbb{R}^{mn}$, where m is the number of charts in the atlas. It is then relatively easy to reduce the dimension by projecting to appropriate hyperplanes, as long as it is at least $2n + 1$. With a some more effort, one proves that any manifold of dimension n admits an embedding into $\mathbb{R}^{2n}$ ("Whitney embedding theorem"). Further, there is an immersion into $\mathbb{R}^{2n-1}$.

Having this in mind, one could actually only discuss submanifolds of $\mathbb{R}^N$, but in many cases there is no natural embedding into some $\mathbb{R}^N$ and then advantages like the intuitive understanding of tangent spaces are lost.

CHAPTER 2

Vector fields

We now turn to the study of the first type of geometric (or analytic) objects on smooth manifolds. These are rather easy to understand intuitively in different ways and give rise to several fundamental operations on manifolds. They also have disadvantages, however. In particular, a smooth map between two manifolds in general does not induce a mapping between vector fields on the two manifolds. This will work better for differential forms that will be discussed later.

2.1. Basic concepts. Given a manifold M, we have the tangent bundle TM, which is a smooth manifold, too. Thus we can study smooth maps $M \to TM$ which associate to each point of M a tangent vector at some point of M. Composing with the natural projection $p : TM \to M$ such a map gives rise to a smooth map on M. It is particularly natural to consider functions which associate to each $x \in M$ a tangent vector at the point x, i.e. for which the induced map $M \to M$ is the identity. This leads to the concept of vector fields. As we have noted above, a smooth map between manifolds does not act on vector fields. However, there is an obvious notion of vector fields being compatible with a smooth map.

Definition 2.1. Let M be a smooth manifold with tangent bundle $p : TM \to M$.

(1) A *vector field* on M is a smooth map $\xi : M \to TM$ such that $p \circ \xi = \mathrm{id}$, i.e. such that $\xi(x) \in T_xM$ for all $x \in M$. The set of all vector fields on M is denoted by $\mathfrak{X}(M)$. A *local vector field* defined on an open subset $U \subset M$ is a smooth map $\xi : U \to TM$ such that $p \circ \xi = \mathrm{id}_U$ (which implies that ξ has values in $p^{-1}(U) = TU$).

(2) Let $F : M \to N$ be a smooth map between smooth manifolds with tangent map TF. Then we call two vector fields $\xi \in \mathfrak{X}(M)$ and $\eta \in \mathfrak{X}(N)$ *F-related* and write $\xi \sim_F \eta$ iff $TF \circ \xi = \eta \circ F$. More explicitly, for each $x \in M$, we require that $\eta(F(x)) = T_xF(\xi(x))$.

If $\xi \in \mathfrak{X}(M)$ is a vector field, then $\xi(x)$ lies in the vector space T_xM for each $x \in M$. So we expect that one can define point-wise addition of vector fields. Similarly, it should be possible to multiply vector fields by real numbers. There is no reason why this shouldn't work for numbers depending smoothly on the point, so there should be point-wise multiplication of vector fields by smooth functions. To see that this really works out, we have to understand the smoothness condition on vector fields.

Recall that any chart (U, u) for M gives rise to an induced chart $(p^{-1}(U), Tu)$ for TM. Now for $\xi \in \mathfrak{X}(M)$, we of course have $\xi^{-1}(p^{-1}(U)) = U$ and so

$$Tu \circ \xi \circ u^{-1} : u(U) \to Tu(p^{-1}(U)) = u(U) \times \mathbb{R}^n$$

is a smooth map. But by definition, p corresponds to the first projection, so this maps $z \in u(U)$ to $(z, \varphi(z))$ for some smooth function $\varphi : u(U) \to \mathbb{R}^n$. Now take two vector fields $\xi_1, \xi_2 \in \mathfrak{X}(M)$ with corresponding functions φ_1 and φ_2. For $x \in M$, $Tu|_{T_xM} = T_xu$ and this is a linear map, so $\xi_1(x)+\xi_2(x)$ is mapped to $(u(x), \varphi_1(u(x))+\varphi_2(u(x)))$. Thus we see the point-wise sum $\xi_1 + \xi_2$ corresponds to $\varphi_1 + \varphi_2$ and hence is smooth too. Similarly, for $f \in C^\infty(M, \mathbb{R})$, the point-wise product $f\xi$ corresponds to the point-wise product $(f \circ u^{-1})\varphi$, so this is also smooth. Otherwise put, the point-wise operations make $\mathfrak{X}(M)$ into a vector space and into a module over the algebra $C^\infty(M,\mathbb{R})$ of smooth functions. Observe that it is also no problem to define the support of a vector field to be the closure $\overline{\{x : \xi(x) \neq 0\}}$.

These observations also imply directly how vector fields are represented in local charts:

Proposition 2.1. *Let M be a smooth manifold of dimension n with tangent bundle $p : TM \to M$.*

(1) For a chart (U,u) for M and each $i = 1,\dots,n$, associating to each $x \in U$ the tangent vector $\frac{\partial}{\partial u^i}|_x$ from 1.14 defines a local smooth vector field $\frac{\partial}{\partial u^i}$ on U.

(2) For a vector field $\xi \in \mathfrak{X}(M)$ and a chart (U,u) there are smooth functions $\xi^i : U \to \mathbb{R}$ for $i = 1,\dots,n$ such that $\xi|_U = \sum_i \xi^i \frac{\partial}{\partial u^i}$.

(3) Consider two charts (U_α,u_α) and (U_β,u_β) with $U_{\alpha\beta} \neq \emptyset$ with chart change $u_{\alpha\beta} : u_\beta(U_{\alpha\beta}) \to u_\alpha(U_{\alpha\beta})$. Then the functions ξ^i_α and ξ^j_β corresponding to the two charts are related by

$$\xi^i_\alpha(x) = \textstyle\sum_j \partial_j u^i_{\alpha\beta}(u_\beta(x))\xi^j_\beta(x). \tag{2.1}$$

(4) Let $U \subset M$ be open and $A \subset M$ be closed with $A \subset U$ in M. Then for any local vector field ξ defined on U, there is a vector field $\tilde{\xi} \in \mathfrak{X}(M)$ such that $\tilde{\xi}|_A = \xi|_A$.

Proof. By construction $\frac{\partial}{\partial u^i}(x) = Tu^{-1}(u(x),e_i)$, so this is evidently smooth and (1) follows. From above, we know that $(Tu \circ \xi \circ u^{-1})(z) = (z,\varphi(z))$ for a smooth function $\varphi : u(U) \to \mathbb{R}^n$. Denoting the components of φ by φ^i, we see that the functions $\xi^i := \varphi^i \circ u$ have the property required in (2). Part (3) then immediately follows from formula (1.1) in Section 1.14.

For part (4), we recall from the proof of Corollary 1.9 that there is a smooth function $\varphi : M \to \mathbb{R}$ with values in $[0,1]$ such that $\operatorname{supp}(\varphi) \subset U$, which is identically one on A. Now we define $\tilde{\xi}(x) = \varphi(x)\xi(x)$ for $x \in U$ and $\tilde{\xi}(x) = 0$ for $x \notin U$. But then $\tilde{\xi}$ is clearly smooth on U and on $M \setminus \operatorname{supp}(\varphi)$ (where it is identically zero). Since these two sets form an open covering of M, this implies that $\tilde{\xi}$ is smooth. □

The vector fields $\frac{\partial}{\partial u^i}$ are called the *coordinate vector fields* for the chart U. Similarly to part (3), their behavior under a change of charts can be immediately read off formula (1.2) from Section 1.14. Generalizing part (4) we see immediately that partitions of unity can be used to glue together locally defined vector fields to global vector fields.

The "transformation law" (2.1) for the components of a vector field is the basis for a chart dependent definition of vector fields, which is often used, in particular in texts oriented towards physics. Roughly, one says that in a chart a vector field is given by an n-tuple of smooth functions, and these tuples have to obey the right transformation law under a chart change. This is a legitimate approach, but it carries the danger of misunderstandings, in particular, once one starts to construct operations on vector fields in this way. It is no problem, to define a local vector field on the domain of a chart by choosing n smooth functions and then forming the corresponding linear combination of coordinate vector fields. But to talk about vector fields properly in such a language, one would actually have to assign n-tuples of functions to *any* chart (or at least to all charts in an atlas).

Example 2.1. Let us give an interesting example of a global vector field here. Consider an odd-dimensional sphere, say $S^{2n-1} \subset \mathbb{R}^{2n}$. Identifying $\mathbb{R}^{2n}$ with $\mathbb{C}^n$, the standard inner product on $\mathbb{R}^{2n}$ gets identified with the real part of the standard Hermitian inner product on $\mathbb{C}^n$. Thus for $x \in S^{2n-1}$ the tangent space T_xS^{2n-1} can be identified with the space of those $y \in \mathbb{C}^n$ for which the Hermitian inner product with x is purely imaginary. In particular $ix \in \mathbb{C}^n$ has this property, so the map $x \mapsto ix$ defines a vector field ξ on S^{2n-1}, which evidently satisfies $\xi(x) \neq 0$ for all $x \in S^{2n-1}$.

The interest in this examples comes from the so-called hairy ball theorem from algebraic topology. This says that, for even n, any smooth (or even continuous) vector field on S^n sends at least one point to zero.

2.2. Vector fields as derivations. By definition, the value of a vector field ξ in a point x is a tangent vector. By definition, this acts on smooth functions as a derivation at x, which is interpreted as a directional derivative in the point x in direction $\xi(x)$. Given $f \in C^\infty(M,\mathbb{R})$, we thus get $\xi(x)(f) \in \mathbb{R}$ and it it is a natural idea to look at this as a function of x. Otherwise put, for a vector field $\xi \in \mathfrak{X}(M)$ and a smooth function $f : M \to \mathbb{R}$, we define $\xi(f) : M \to \mathbb{R}$ by $\xi(f)(x) := \xi(x)(f)$ where, as in 1.12, we view $\xi(x) \in T_xM$ as a linear map $C^\infty(M,\mathbb{R}) \to \mathbb{R}$. This is called the *derivative of f in direction ξ* and it is an analog of a directional derivative with the direction depending on the point. (Without additional structures, there is no reasonable concept of directions which do not depend on a point.)

PROPOSITION 2.2. *Let M be a smooth manifold and $\xi \in \mathfrak{X}(M)$ a vector field on M. Then we have*

(1) Let (U,u) be a chart for M and expand $\xi|_U = \sum_i \xi^i \frac{\partial}{\partial u^i}$ as in Proposition 2.1. For $f \in C^\infty(M,\mathbb{R})$, the local coordinate representation $\xi(f) \circ u^{-1}$ with respect to the chart is given by

$$\xi(f) \circ u^{-1} = \textstyle\sum_i (\xi^i \circ u^{-1})\partial_i(f \circ u^{-1}). \tag{2.2}$$

(2) For any $f \in C^\infty(M,\mathbb{R})$, the function $\xi(f)$ is smooth, too. Moreover, ξ defines a linear map $C^\infty(M,\mathbb{R}) \to C^\infty(M,\mathbb{R})$, which is a derivation *in the sense that $\xi(fg) = \xi(f)g + f\xi(g)$ for all $f, g \in C^\infty(M,\mathbb{R})$.*

(3) Conversely, if $D : C^\infty(M,\mathbb{R}) \to C^\infty(M,\mathbb{R})$ is a derivation, then there is a unique vector field $\xi \in \mathfrak{X}(M)$ such that $D(f) = \xi(f)$ for all $f \in C^\infty(M,\mathbb{R})$.

(4) If $F : M \to N$ is a smooth map and $\eta \in \mathfrak{X}(N)$ is another vector field, then $\xi \sim_F \eta$ if and only if $\xi(f \circ F) = \eta(f) \circ F$ holds for all $f \in C^\infty(N,\mathbb{R})$.

PROOF. (1) For $x \in U$, we have $\xi(x) = \sum_i \xi^i(x) \frac{\partial}{\partial u^i}|_x$ and from 1.14 we know that acting on f, this tangent vector produces $\sum_i \xi^i(x)\partial_i(f \circ u^{-1})(u(x))$, which is exactly the claimed result.

(2) From part (1), we see that $\xi(f) \circ u^{-1}$ is smooth and since this holds for any chart, we conclude that $\xi(f) \in C^\infty(M,\mathbb{R})$. Since we know that $\xi(x) : C^\infty(M,\mathbb{R}) \to \mathbb{R}$ is linear and addition and scalar multiplication of smooth functions are point-wise, we readily see that ξ is linear as a map $C^\infty(M,\mathbb{R}) \to C^\infty(M,\mathbb{R})$. Finally, $\xi(x)(fg) = \xi(x)(f)g(x) + f(x)\xi(x)(g)$, which exactly gives the claimed derivation property.

(3) Let D be a derivation, take a point $x \in M$ and consider the linear map $C^\infty(M,\mathbb{R}) \to \mathbb{R}$ defined by $f \mapsto D(f)(x)$. The derivation property of D immediately implies that this is a derivation at x. By Proposition 1.12, there is a unique tangent vector $X_x \in T_xM$ such that $D(f)(x) = X_x(f)$ holds for all $f \in C^\infty(M,\mathbb{R})$. Hence we can define a map $\xi : M \to TM$ by sending each x to the tangent vector X_x obtained in this way. Since $p \circ \xi = \mathrm{id}_M$ is obvious, it remains to show that ξ is smooth.

But for a chart (U,u), it follows from 1.14 how to expand X_x in the basis $\frac{\partial}{\partial u^i}|_x$: Since the coordinate functions u^j satisfy $\partial_i(u^j \circ u^{-1}) = \delta_i^j$, we conclude that $X_x = \sum_i X_x(u^i)\frac{\partial}{\partial u^i}|_x$. Given x, take an open neighborhood V of x such that $\bar{V} \subset U$. Then by Corollary 1.9, we know that for each i, there is a smooth function $\tilde{u}^i : M \to \mathbb{R}$ which coincides with u^i on V. Moreover, the proof of Proposition 1.12 shows that for $y \in V$, we get $D(\tilde{u}^i)(y) = X_y(u^i)$. This shows that $\xi|_V = \sum_i D(\tilde{u}^i)|_V \frac{\partial}{\partial u^i}|_V$ and since $D(\tilde{u}^i)$ is smooth for each i, we see that $\xi|_V$ is smooth. Since this works locally around each point, ξ is smooth.

(4) This is a simple computation. By definition $\xi \sim_F \eta$ is equivalent to $T_xF \cdot \xi(x) = \eta(F(x))$ for each $x \in M$. Letting the right hand side act on $f \in C^\infty(N, \mathbb{R})$, we get $\eta(f)(F(x)) = (\eta(f) \circ F)(x)$. The action of the left hand side on f by definition is given by $\xi(x)(f \circ F) = \xi(f \circ F)(x)$, and using uniqueness from (3), the claim follows. □

At this stage it will be helpful to eliminate the large number of summation signs, whence we introduce the so-called *Einstein summation convention*. It says that if in a formula there is an upper index and a lower index of the same name, then one has to sum over all values $1, \dots, n$ of the indices. To make things more complicated, the upper index in the denominator of $\frac{\partial}{\partial u^i}$ counts as a lower index. Using the convention, the expansion of a vector field with respect to coordinate vector fields reads as $\xi|_U = \xi^i \frac{\partial}{\partial u^i}$. This is a well chosen convention, but at least in the beginning one should carefully make clear to oneself what is actually meant by formulae involving a sum convention.

It is useful to note at this point that the action of vector fields on real valued functions extends without problems to functions with values in $\mathbb{R}^m$ or any finite dimensional vector space V. Given a smooth function $f : M \to V$, we can choose a basis $\{v_i\}$ of V, write $f(x) = \sum_i f_i(x)v_i$ for functions $f_i \in C^\infty(M, \mathbb{R})$. For a vector field $\xi \in \mathfrak{X}(M)$ one immediately concludes that $\xi(f)(x) := \sum_i \xi(f_i)(x)v_i$ does not depend on the choice of the basis $\{v_i\}$, so we obtain a well defined smooth function $\xi(f) : M \to V$. In particular, for $V = \mathbb{R}^m$ the action of vector fields is just component-wise.

2.3. Mapping vector fields. As mentioned already, a general smooth map $F : M \to N$ between smooth manifolds cannot be used to move vector fields from M to N or from N to M. One thing that can always be done is to take a vector field $\xi \in \mathfrak{X}(M)$ and consider $TF \circ \xi : M \to TN$. This is a smooth map for which the composition with p reproduces F. Such maps are referred to as *vector fields along* F, but the tools for vector fields that we are going to develop do not apply to these objects.

Given $F : M \to N$, consider the natural condition defining $\xi \sim_F \eta$, namely $\eta(F(x)) = T_xF(\xi(x))$. It is obvious that given $\eta \in \mathfrak{X}(N)$ it is not possible in general to find $\xi \in \mathfrak{X}(M)$ such that $\xi \sim_F \eta$. It may simply happen that $\eta(y)$ does not lie in the image of T_xF for all points x with $F(x) = y$. Similarly, given ξ it may happen that there are points x_1, x_2 with $F(x_1) = F(x_2)$ but $T_{x_1}F(\xi(x_1)) \neq T_{x_2}F(\xi(x_2))$, which makes it impossible to find $\eta \in \mathfrak{X}(N)$ such that $\xi \sim_F \eta$. On the other hand, it is rather evident that diffeomorphisms can be used to map vector fields. Indeed, there is a more general concept, which automatically extends to local diffeomorphisms.

Proposition 2.3. *Let $F : M \to N$ be a local diffeomorphism. Then for any vector field $\eta \in \mathfrak{X}(N)$, there is a unique vector field $F^*\eta \in \mathfrak{X}(M)$ such that $F^*\eta \sim_F \eta$.*

Proof. By definition, T_xF is invertible for each $x \in M$ and we define $F^*\eta(x) := (T_xF)^{-1}(\eta(F(x)))$. Moreover, for $x \in M$, there are open neighborhoods U of x in M and V of $F(x)$ in N such that $F|_U : U \to V$ is a diffeomorphism. Denoting by $G : V \to U$ the inverse we have $T_{F(y)}G = (T_yF)^{-1}$ for all $y \in U$. This shows that $F^*\eta|_U = TG \circ \eta \circ F$ which is obviously smooth, so $F^*\eta \in \mathfrak{X}(M)$. Knowing this, $F^*\eta \sim_F \eta$ and uniqueness are obvious. □

The vector field $F^*\eta$ is called the *pullback* of η by F. We can also phrase this as the fact that a local diffeomorphism $F : M \to N$ induces an operator $F^* : \mathfrak{X}(N) \to \mathfrak{X}(M)$.

Note that this in particular applies to open embeddings, where we simply get an obvious "restriction map" for vector fields.

Consider smooth maps $F : M \to N$ and $G : N \to P$ and vector fields $\xi \in \mathfrak{X}(M)$, $\eta \in \mathfrak{X}(N)$ and $\zeta \in \mathfrak{X}(P)$. Then $\xi \sim_F \eta$ and $\eta \sim_G \zeta$ together immediately imply $\xi \sim_{G\circ F} \zeta$. In particular, if F and G are both local diffeomorphisms, this shows that $(G \circ F)^*\zeta = F^*(G^*\zeta)$ for all $\zeta \in \mathfrak{X}(P)$.

Generalizing the case of local diffeomorphisms, we briefly discuss the concept of F-relatedness for immersions and submersions, where it is particularly interesting. Let us first consider an immersion $i : M \to N$, see 1.19. Then for each $x \in M$, $T_xi : T_xM \to T_{i(x)}N$ is injective, so we can view T_xM as a linear subspace of $T_{i(x)}N$. The natural setup here is to start with $\eta \in \mathfrak{X}(N)$. To have a chance to find $\xi \in \mathfrak{X}(M)$ such that $\xi \sim_i \eta$, we of course have to require that $\eta(i(x))$ lies in that subspace $T_xi(T_xM)$ for each $x \in M$. In particular in the setting of immersed submanifolds, this is usually phrased as the fact that η is *tangent to M along M*. If this is satisfied, then there is a unique $\xi(x) \in T_xM$ such that $T_xi(\xi(x)) = \eta(i(x))$. Thus we get a function $\xi : M \to TM$ and we only have to show that this is smooth to see that it defines a vector field such that $\xi \sim_i \eta$. But this is a local question, so we can prove this in charts in which the local coordinate representation of i is given by $(z^1, \dots, z^k) \mapsto (z^1, \dots, z^k, 0, \dots, 0)$, see Proposition 1.19. But in such a chart, we simply get $\xi^i = \eta^i$ for $i = 1, \dots, k$, so smoothness follows. In particular, for an immersed submanifold $i : M \to N$ and a vector field $\eta \in \mathfrak{X}(M)$, there is $\xi \in \mathfrak{X}(M)$ such that $\xi \sim_i \eta$ if and only if η is tangent to M along M.

Dually, suppose that $F : M \to N$ is a surjective submersion, so F and all tangent maps $T_xF : T_xM \to T_{F(x)}N$ are surjective. This immediately implies that for given $\xi \in \mathfrak{X}(M)$ there is at most one $\eta \in \mathfrak{X}(N)$ such that $\xi \sim_F \eta$, so the natural question is to characterize those ξ for which such an η exists. From Proposition 1.18, we know that for each $y \in N$, the subset $F^{-1}(\{y\}) \subset M$ is a smooth submanifold. The condition we are looking for of course is that the vectors $T_xF(\xi(x)) \in T_yN$ agree for all $x \in F^{-1}(\{y\})$. Such vector fields are called *projectable* and given such a vector field $\xi \in \mathfrak{X}(M)$, there is a unique function $\eta : N \to TN$ such that $\eta \circ F = TF \circ \xi$. But by Proposition 1.18, smoothness of $TF \circ \xi = \eta \circ F$ implies smoothness of η, so this is a vector field such that $\xi \sim_F \eta$. This is called the *projection* of the projectable vector field ξ.

The Lie bracket

The interpretation of vector fields as derivations on the algebra of smooth functions now leads to one of the fundamental operations of analysis on manifolds, the so-called Lie bracket of vector fields. The actual way how this comes up as well as the basic properties of this operation may be rather unexpected, though.

2.4. Definition and basic properties. The operation $(\eta, f) \mapsto \eta(f)$ of directional derivatives can be iterated, so for another vector field ξ, we ca form $\xi(\eta(f))$. But clearly, this will involve second derivatives of f, which is confirmed by the local formula (2.2) from Proposition 2.2. But it will also involve first derivatives of η in some sense. The terms depending on second derivatives of f can be removed by forming a commutator, which then leads to a bilinear differential operator on vector fields.

Theorem 2.4. *Let M be a smooth manifold.*

(1) For vector fields $\xi, \eta \in \mathfrak{X}(M)$, there is a unique vector field $[\xi, \eta] \in \mathfrak{X}(M)$ such that for all $f \in C^\infty(M, \mathbb{R})$, we have $[\xi, \eta](f) = \xi(\eta(f)) - \eta(\xi(f))$.

(2) This defines a bilinear operator $[\ ,\ ] : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$, called the Lie bracket, which is skew symmetric in the sense that $[\eta, \xi] = -[\xi, \eta]$ and satisfies the

Jacobi identity

$$[\xi,[\eta,\zeta]]=[[\xi,\eta],\zeta]+[\eta,[\xi,\zeta]]$$

for all $\xi,\eta,\zeta\in\mathfrak{X}(M)$.

(3) For $\xi,\eta\in\mathfrak{X}(M)$ *and* $f\in C^\infty(M,\mathbb{R})$*, we obtain* $[\xi,f\eta]=\xi(f)\eta+f[\xi,\eta]$.

(4) The Lie bracket is a local operator: If $\xi,\eta,\tilde\eta\in\mathfrak{X}(M)$ *are vector fields and* $U\subset M$ *is an open subset such that* $\eta|_U=\tilde\eta|_U$*, then* $[\xi,\eta]|_U=[\xi,\tilde\eta]|_U$.

(5) For a chart (U,u) *for* M *and the associated coordinate vector fields* $\frac{\partial}{\partial u^i}$*, we get* $[\frac{\partial}{\partial u^i},\frac{\partial}{\partial u^j}]=0$ *for all* i,j*. Moreover, in terms of the corresponding expansions* $\xi|_U=\xi^i\frac{\partial}{\partial u^i}$ *and* $\eta|_U=\eta^j\frac{\partial}{\partial u^j}$ *the local expression for* $[\xi,\eta]$ *is given by*

$$[\xi,\eta]|_U=\left(\xi^i\tfrac{\partial}{\partial u^i}(\eta^j)-\eta^i\tfrac{\partial}{\partial u^i}(\xi^j)\right)\tfrac{\partial}{\partial u^j}. \tag{2.3}$$

Proof. (1) In view of Proposition 2.2, we only have to prove that $f\mapsto\xi(\eta(f))-\eta(\xi(f))$ is a derivation on the algebra $C^\infty(M,\mathbb{R})$, which is a simple direct computation: Using that ξ is a derivation, we compute

$$\xi(\eta(fg))=\xi(\eta(f)g+f\eta(g))=\xi(\eta(f))g+\eta(f)\xi(g)+\xi(f)\eta(g)+f\xi(\eta(g)).$$

But the two middle summands are symmetric in ξ and η, so they will cancel with the corresponding summands in $-\eta(\xi(fg))$.

(2) From the definition it is obvious that $[\eta,\xi]=-[\xi,\eta]$. Since the operation $(\xi,f)\mapsto\xi(f)$ is bilinear over $\mathbb{R}$, it follows readily that also $(\xi,\eta)\mapsto[\xi,\eta]$ is bilinear over $\mathbb{R}$. To prove the Jacobi identity, it suffices to show that both sides act in the same way on a smooth function f, which again is a simple direct computation: Applying the right hand side to f, we obtain

$$\begin{aligned}[\xi,\eta](\zeta(f))-\zeta([\xi,\eta](f))+\eta([\xi,\zeta](f))-[\xi,\zeta](\eta(f))&=\xi(\eta(\zeta(f)))-\eta(\xi(\zeta(f)))\\ -\zeta(\xi(\eta(f)))+\zeta(\eta(\xi(f)))+\eta(\xi(\zeta(f)))-\eta(\zeta(\xi(f)))&-\xi(\zeta(\eta(f)))+\zeta(\xi(\eta(f))).\end{aligned}$$

Now the second and fifth and the third and last terms in the right hand side cancel, and the remaining four terms add up to $[\xi,[\eta,\zeta]](f)$.

(3) Again, we just evaluate on $g\in C^\infty(M,\mathbb{R})$ as follows:

$$[\xi,f\eta](g)=\xi(f\eta(g))-f\eta(\xi(g))=\xi(f)\eta(g)+f\xi(\eta(g))-f\eta(\xi(g)).$$

But the first summand in the right hand side is just $(\xi(f)\eta)(g)$, while the other two summands clearly add up to $(f[\xi,\eta])(g)$.

(4) Since the Lie bracket is bilinear, the claimed statement is evidently equivalent to the fact that if $\eta|_U$ is identically zero, then also $[\xi,\eta]|_U$ is identically zero. Assuming that $\eta|_U=0$, then for a point $x\in U$, the proof of Corollary 1.9 shows that there is a smooth function $\varphi:M\to\mathbb{R}$ such that $\varphi(x)=1$ and $\operatorname{supp}(\varphi)\subset U$. But then $\varphi\eta$ is identically zero, so bilinearity implies that $0=[\xi,\varphi\eta]$. Expanding this according to (3) and evaluating in x, we conclude that $0=\xi(\varphi)(x)\eta(x)+\varphi(x)[\xi,\eta](x)=0+1\cdot[\xi,\eta](x)$.

(5) By definition, for a smooth function $f:M\to\mathbb{R}$, $\frac{\partial}{\partial u^i}(f)\circ u^{-1}$ is the ith partial derivative of the local coordinate representation $f\circ u^{-1}$. Expressing $[\frac{\partial}{\partial u^i},\frac{\partial}{\partial u^j}](f)$ accordingly, symmetry of the second partial derivatives implies $[\frac{\partial}{\partial u^i},\frac{\partial}{\partial u^j}]=0$.

Taking the local coordinate representation $\eta|_U=\eta^j\frac{\partial}{\partial u^j}$, we can extend the functions η^j and the coordinate vector fields to globally defined objects that we denote by the same symbols. Given a point $x\in U$ we can do this without changing them in a neighborhood of x. Thus part (4) implies that locally around x, $[\xi,\eta]$ equals $[\xi,\eta^j\frac{\partial}{\partial u^j}]$. Since x is arbitrary, we get $[\xi,\eta]|_U=[\xi,\eta^j\frac{\partial}{\partial u^j}]|_U$ (again using part (4)). In the same way, we can replace ξ by $\xi^i\frac{\partial}{\partial u^i}$. But then parts (2) and (3) imply

$$[\xi^i\tfrac{\partial}{\partial u^i},\eta^j\tfrac{\partial}{\partial u^j}]=\big(\xi^i\tfrac{\partial}{\partial u^i}(\eta^j)\big)\tfrac{\partial}{\partial u^j}+\eta^j[\xi^i\tfrac{\partial}{\partial u^i},\tfrac{\partial}{\partial u^j}].$$

In the last summand, we can apply skew symmetry and (3) to rewrite it as

$$-\big(\eta^j\tfrac{\partial}{\partial u^j}(\xi^i)\big)\tfrac{\partial}{\partial u^i}+\eta^j\xi^i[\tfrac{\partial}{\partial u^i},\tfrac{\partial}{\partial u^j}],$$

and we have just observed that the last term vanishes. □

On submanifolds of $\mathbb{R}^n$, there is an alternative description of the Lie bracket. Observe that for a smooth, k-dimensional submanifold $M\subset\mathbb{R}^n$, one may view a vector field $\eta\in\mathfrak{X}(M)$ as a smooth map $M\to\mathbb{R}^n$ (which has the property that $\eta(x)\in T_xM$ for each $x\in M$). Hence given another vector field $\xi\in\mathfrak{X}(M)$ we can differentiate $\eta:M\to\mathbb{R}^n$ in direction ξ, compare with Section 2.2, to obtain $\xi(\eta):M\to\mathbb{R}^n$. The result is not a vector field, however, there is no reason why $\xi(\eta)(x)$ should lie in T_xM. But it turns out that the $\xi(\eta)(x)-\eta(\xi)(x)$ always lies in T_xM and coincides with $[\xi,\eta](x)$ (see exercises).

Example 2.4. (1) Consider the open subset $\{(x^1,x^2,x^3):x^3>0\}$ on $\mathbb{R}^3$ and the coordinate vector fields $\frac{\partial}{\partial x^i}$ for the standard coordinates. Put $\xi:=x^3\frac{\partial}{\partial x^1}-x^1\frac{\partial}{\partial x^3}$ and $\eta:=x^3\frac{\partial}{\partial x^2}-x^2\frac{\partial}{\partial x^3}$. Then the coordinate formula in part (5) of Theorem 2.4 readily shows that

$$[\xi,\eta]=-x^1\tfrac{\partial}{\partial x^2}+x^2\tfrac{\partial}{\partial x^1}=\tfrac{x^2}{x^3}\xi-\tfrac{x^1}{x^3}\eta.$$

Hence in this case, $[\xi,\eta](x)$ lies in the span of $\xi(x)$ and $\eta(x)$ for each point x.

(2) On $\mathbb{R}^3$, consider the vector fields $\xi=\frac{\partial}{\partial x^1}-x^2\frac{\partial}{\partial x^3}$ and $\eta:=\frac{\partial}{\partial x^2}+x^1\frac{\partial}{\partial x^3}$. As above, one computes that $[\xi,\eta]=2\frac{\partial}{\partial x^3}$, so in this case, the values of ξ, η and $[\xi,\eta]$ are linearly independent in each point. We will interpret the difference to example (1) more geometrically in Example 2.5 below.

2.5. Naturality of the Lie bracket. In an approach which defines vector fields via local charts, one would use the formula in part (5) of Theorem 2.4 to define the Lie bracket of vector fields. Of course, this requires a verification that working in different charts leads to the same result on the intersection of the domains of the charts. This can be done via a direct computation using the transformation law (2.1) from Proposition 2.1. While in our approach, such a verification is not needed, it is a nice exercise to carry out this computation.

The fact that the local formula leads to the same result in all charts is closely related to naturality of the Lie bracket under local diffeomorphisms. While this may not look as exciting at the first glance it is an extremely strong property. In fact we will only meet very few operations with similar properties, which makes each of them a cornerstone for analysis on manifolds. In fact, the Lie bracket has an even stronger naturality property, since it is compatible with the notion of F-relatedness of vector fields:

Theorem 2.5. *Let $F:M\to N$ be a smooth map between manifolds and let $\xi_1,\xi_2\in\mathfrak{X}(M)$ and $\eta_1,\eta_2\in\mathfrak{X}(N)$ be vector fields. If $\xi_1\sim_F\eta_1$ and $\xi_2\sim_F\eta_2$, then $[\xi_1,\xi_2]\sim_F[\eta_1,\eta_2]$. In particular, if F is a local diffeomorphism, then $[F^*\eta_1,F^*\eta_2]=F^*([\eta_1,\eta_2])$.*

Proof. Take a smooth function $f\in C^\infty(N,\mathbb{R})$. Then by part (4) of Proposition 2.2, $\xi_2\sim_F\eta_2$ implies $\xi_2(f\circ F)=\eta_2(f)\circ F$. Applying ξ_1 and using F-relatedness to η_1, we get

$$\xi_1(\xi_2(f\circ F))=\xi_1(\eta_2(f)\circ F)=\eta_1(\eta_2(f))\circ F.$$

In the same way $\xi_2(\xi_1(f \circ F)) = \eta_2(\eta_1(f)) \circ F$ and then part (4) of Proposition 2.2 implies the first claim. The second claim then immediately follows in view of Proposition 2.3. □

EXAMPLE 2.5. We can now give a more geometric explanation for the different behaviors of the Lie bracket in the examples from 2.4. In the setting of Example (1) from 2.4 we can take a point x with $x^3 > 0$, and look at the sphere of radius $|x|$ around zero. So this is an embedded submanifold M and of course T_xM is the orthocomplement of x. Now we had $\xi(x) = x^3\frac{\partial}{\partial x^1} - x^1\frac{\partial}{\partial x^3} \in T_xM$ and similarly $\eta(x) \in T_xM$. In the terminology from 2.3, wee see that both ξ and η are tangent to M along M. Denoting by i the inclusion of M, we have seen in 2.3 that there are vector fields $\tilde{\xi}, \tilde{\eta} \in \mathfrak{X}(M)$ such that $\tilde{\xi} \sim_i \xi$ and $\tilde{\eta} \sim_i \eta$. So by Theorem 2.5, we get $[\tilde{\xi}, \tilde{\eta}] \sim_i [\xi, \eta]$, which in particular implies $[\xi, \eta](x) \in T_xM$ for each $x \in M$. Since ξ and η are linearly independent in each point, their values span the tangent spaces to M, so we see without any computation that $[\xi, \eta](x)$ must be a linear combination of $\xi(x)$ and $\eta(x)$.

In contrast to this, we have defined in Example (2) of 2.4 two vector fields $\xi, \eta \in \mathfrak{X}(\mathbb{R}^3)$ such that $\xi(x)$, $\eta(x)$ and $[\xi, \eta](x)$ are linearly independent in each point $x \in \mathbb{R}^3$. Hence our considerations show that there does not exists any (local) two-dimensional submanifold $M \subset \mathbb{R}^3$ such that for each $x \in M$, the tangent space T_xM is spanned by $\xi(x)$ and $\eta(x)$.

Integral curves and flows

We now move to a second natural interpretation of vector fields, namely as describing first order ODEs on manifolds.

2.6. Integral curves. For a smooth curve $c : I \to M$ defined on an open interval $I \subset \mathbb{R}$ and each $t \in I$, we have $c'(t) \in T_{c(t)}M$. Given a vector field $\xi \in \mathfrak{X}(M)$, it thus makes sense to look for *integral curves* for ξ, i.e. smooth curves $c : I \to M$ such that $c'(t) = \xi(c(t))$ for each $t \in I$.

LEMMA 2.6. *Let $\xi \in \mathfrak{X}(M)$ be a vector field on a smooth manifold M.*

(1) For $x \in M$ there is an unique maximal open interval $J_x \subset \mathbb{R}$ containing 0 and a unique integral curve $c_x : J_x \to M$ for ξ with $c_x(0) = x$.

(2) For $s \in J_x$ we obtain $J_{c_x(s)} = \{t - s : t \in J_x\}$ and $c_{c_x(s)}(t) = c_x(t + s)$ for all t.

PROOF. (1) Let (U, u) be a local chart for M and let $\xi|_U = \xi^i\frac{\partial}{\partial u^i}$ be the corresponding local coordinate representation of ξ. By definition, this means that $T_yu(\xi(y)) = (\xi^1(y), \dots, \xi^n(y))$. For a smooth curve $\tilde{c} : I \to u(U)$ with components $\tilde{c}^i$, the curve $u^{-1} \circ \tilde{c}$ thus is an integral curve of ξ if and only if $(\tilde{c}^i)'(t) = (\xi^i \circ u^{-1})(\tilde{c}(t))$. This is a system of first order ODEs with smooth coefficients, so it has unique local solutions for any choice of initial condition in $u(U)$ by the Picard-Lindelöff theorem.

Thus we conclude that integral curves exist through each point. Moreover, suppose that we have two integral curves $c_1 : I_1 \to M$ and $c_2 : I_2 \to M$ defined on open intervals I_1, I_2 such that $I_1 \cap I_2 \neq \emptyset$. Then if $c_1(t_0) = c_2(t_0)$ for some $t_0 \in I_1 \cap I_2$, then $c_1|_{I_1 \cap I_2} = c_2|_{I_1 \cap I_2}$ and hence they can be pieced together to an integral curve defined on $I_1 \cup I_2$. This easily implies that local integral curves can be pieced together to maximal integral curves and hence part (1).

(2) Of course, for $s \in J_x$, $t \mapsto c_x(t + s)$ is an integral curve for ξ which maps 0 to $c_x(s)$ and is defined on $\{t \in \mathbb{R} : t + s \in J_x\}$. Thus, the latter interval has to be contained in $J_{c_x(s)}$ and on this sub-interval $c_{c_x(s)}(t) = c_x(t + s)$. But this in particular

shows that $-s \in J_{c_x(s)}$ and that $c_{c_x(s)}(-s) = x$. So we conclude in the same way that $\{t \in \mathbb{R} : t - s \in J_{c_x(s)}\} \subset J_x$ which implies the claim. $\square$

EXAMPLE 2.6. (1) For $v \in \mathbb{R}^n$ the constant map v defines a vector field on $\mathbb{R}^n$. Of course, the affine line $t \mapsto x + tv$ is an integral curve through x. Thus in the case $J_x = \mathbb{R}$ for each $x \in \mathbb{R}^n$.

But of course, we can also view the constant map v as a vector field on any open subset $U \subset \mathbb{R}^n$. Then for each $x \in U$, $\{t \in \mathbb{R} : x + tv \in U\}$ is open in $\mathbb{R}$ and J_x in this case is the connected component of this subset that contains zero. So it may easily happen that maximal integral curves are defined on finite intervals only.

(2) For any linear map $A : \mathbb{R}^n \to \mathbb{R}^n$, one may view $x \mapsto Ax$ as a vector field on $\mathbb{R}^n$. It is a standard result from the theory of ordinary differential equations that that an integral curve through x is given via the matrix exponential as $c(t) = e^{tA}x$. Here $e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$ and this sum is absolutely convergent for all $t \in \mathbb{R}$ and defines a linear map $\mathbb{R}^n \to \mathbb{R}^n$. So again integral curves are defined on all of $\mathbb{R}$.

One can analyze the behavior of the integral curves by using that the matrix exponential is compatible with conjugation, i.e. that $e^{tBAB^{-1}} = Be^{tA}B^{-1}$. This means that after an appropriate change of basis, one may assume that the matrix A is in (the real version of) Jordan normal form. In low dimensions, this reduces things to a small number of basic cases that can be studied explicitly.

(3) Up to now, the examples in which integral curves are defined for finite time only were rather artificial, but this occurs naturally in many cases. Consider for example the map $x \mapsto x^2 \frac{\partial}{\partial x}$ as a vector field on $\mathbb{R}$. Then $c : (-\infty, 1) \to \mathbb{R}$, $c(t) := \frac{1}{1-t}$ visibly satisfies $c'(t) = c(t)^2$ and thus is an integral curve. But for $t \to 1$ this curve evidently tends to $+\infty$, so there is no way to extend it to a bigger interval. Hence in this case $J_1 = (-\infty, 1)$ even though we are working on the full space $\mathbb{R}$.

2.7. Flows. Fixing a vector field $\xi \in \mathfrak{X}(M)$, we next define $\mathcal{D}(\xi) \subset M \times \mathbb{R}$ as the set of all (x, t) such that $t \in J_x$, where J_x is the maximal interval obtained in part (1) of Lemma 2.6. Moreover, we define the *flow of* ξ as a map $\mathrm{Fl}^\xi : \mathcal{D}(\xi) \to M$ by $\mathrm{Fl}^\xi(x, t) := \mathrm{Fl}^\xi_t(x) := c_x(t)$, where c_x is the maximal integral curve for ξ obtained in part (1) of Lemma 2.6.

THEOREM 2.7. *For a vector field $\xi \in \mathfrak{X}(M)$ on a smooth manifold M we have*

(1) The subset $\mathcal{D}(\xi) \subset M \times \mathbb{R}$ defined above is an open neighborhood of $M \times \{0\}$ and thus a manifold.

(2) The flow map $\mathrm{Fl}^\xi : \mathcal{D}(\xi) \to M$ is smooth and it satisfies

$$\mathrm{Fl}^\xi_{t+s}(x) = \mathrm{Fl}^\xi_t(\mathrm{Fl}^\xi_s(x)), \tag{2.4}$$

whenever both sides are defined. Moreover, if the right hand side is defined then the left hand side is defined and the converse holds if t and s have the same sign.

(3) For $(x, t) \in \mathcal{D}(\xi)$, there is an open neighborhood U of x in M such that Fl^ξ_t restricts to a diffeomorphism from U onto an open subset of M.

PROOF. Equation (2.4) in part (2) just reads as $c_x(t + s) = c_{c_x(s)}(t)$, so this has been proved in part (2) of Lemma 2.6. The latter also easily implies the last part of (2), since this just relates the conditions that $s \in J_x$, $t \in J_{c_x(s)}$ and $t + s \in J_x$.

(1) and smoothness in (2): The Picard-Lindelöff theorem (including smooth dependence on the initial conditions) applied in charts implies that for any $x \in M$ there is an open neighborhood U of x in M and $\epsilon > 0$ such that $U \times (-\epsilon, \epsilon) \subset \mathcal{D}(\xi)$ and Fl^ξ

restricts to a smooth map $U \times (-\epsilon, \epsilon) \to M$. Thus it remains to show that $\mathcal{D}(\xi)$ is open and that Fl^ξ is smooth on all of $D(\xi)$. Consider the subset $W \subset \mathcal{D}(\xi)$ consisting of all (x,t) such that there is an open neighborhood U of x in M and an open interval $J \subset \mathbb{R}$ with $0, t \in J$ such that $U \times J \subset \mathcal{D}(\xi)$ and Fl^ξ is smooth on $U \times J$. Then from above we know that W is an open neighborhood of $M \times \{0\}$ in $M \times \mathbb{R}$ and we can complete this part of the proof by showing that $W = \mathcal{D}(\xi)$.

With a view towards contradiction, suppose that $(x_0, \tau) \in \mathcal{D}(\xi) \setminus W$. We do the case $\tau > 0$, the case $\tau < 0$ is closely similar. Putting $t_0 := \inf\{t \in [0,\infty) : (x_0, t) \notin W\}$, we see from above that $0 < t_0 \leq \tau$ and since $\tau \in J_{x_0}$, we conclude that $t_0 \in J_{x_0}$. Putting $y_0 := \mathrm{Fl}^\xi_{t_0}(x_0)$, we know from above that there is an open neighborhood V of y_0 in M and $\epsilon > 0$ such that $V \times (-\epsilon, \epsilon) \subset \mathcal{D}(\xi)$ and Fl^ξ is smooth on $V \times (-\epsilon, \epsilon)$.

By construction, we can find elements $t_1 < t_0$ arbitrarily close to t_0 such that $(x_0, t_1) \in W$, so in particular we may require $t_1 + \epsilon > t_0$. By continuity of c_x we can also require that $c_{x_0}(t_1) = \mathrm{Fl}^\xi_{t_1}(x_0) \in V$. Since $(x_0, t_1) \in W$, we find an open neighborhood U of x_0 in M and a an open interval $J \subset \mathbb{R}$ with $[0, t_1] \subset J$ such that Fl^ξ is defined and smooth on $U \times J$. Replacing U by $U \cap (\mathrm{Fl}^\xi_{t_1})^{-1}(V)$, we may assume that $\mathrm{Fl}^\xi_{t_1}(U) \subset V$. Now define $\tilde{J} := J \cup [t_1, t_1 + \epsilon)$, which by construction is an open interval containing 0 and t_0. Define a map $F : U \times \tilde{J} \to M$ by

$$F(x,t) := \begin{cases} \mathrm{Fl}^\xi(x,t) & t \in J \\ \mathrm{Fl}^\xi(\mathrm{Fl}^\xi(x,t_1), t - t_1) & t \in (t_1 - \epsilon, t_1 + \epsilon) \cap \tilde{J} \end{cases}$$

Note that the first line defines a smooth map on $U \times J$ while in the second line we have $\mathrm{Fl}^\xi(x, t_1) \in V$ and $t - t_1 \in (-\epsilon, \epsilon)$, so this is smooth as a composition of two smooth maps. Moreover, the equation in (2) shows that the two definitions agree on their overlap, so we obtain a smooth map $U \times \tilde{J} \to M$. But the equation in (2) also shows that F agrees with Fl^ξ on this open set, which implies $(x_0, t_0) \in W$, thus contradicting the construction of t_0.

(3) Having proved (1) and (2), we readily see that for fixed $t \in \mathbb{R}$, the set $M_t := \{x \in M : (x,t) \in \mathcal{D}(\xi)\}$ is open in M and $\mathrm{Fl}^\xi_t : M_t \to M$ is smooth. From part (2) of Lemma 2.6, we see that $-t \in J_{c_x(t)}$, so $\mathrm{Fl}^\xi_t(M_t) \subset M_{-t}$. Similarly, we conclude that Fl^ξ_{-t} is a smooth map on M_{-t} and has values in M_t. By the equation in part (2) and the obvious observation that $\mathrm{Fl}^\xi_0 = \mathrm{id}_M$, we conclude that these maps are inverse to each other, which implies the claim. □

It is rather easy to understand the implications of relatedness for flows:

COROLLARY 2.7. *Let $F : M \to N$ be a smooth map between smooth manifolds and suppose that $\xi \in \mathfrak{X}(M)$ and $\eta \in \mathfrak{X}(N)$ satisfy $\xi \sim_F \eta$. Then $(F \times \mathrm{id}_{\mathbb{R}})(\mathcal{D}(\xi)) \subset \mathcal{D}(\eta)$ and for each $(x,t) \in \mathcal{D}(\xi)$ we get $\mathrm{Fl}^\eta_t(F(x)) = F(\mathrm{Fl}^\xi_t(x))$.*

In particular, if F is a local diffeomorphism, then for each vector field $\eta \in \mathfrak{X}(N)$, we get $F \circ \mathrm{Fl}^{F^*\eta}_t = \mathrm{Fl}^\eta_t \circ F$ whenever both sides are defined, and if the left hand side is defined, then the right hand side is defined.

PROOF. For a smooth curve $c : I \to M$, we get $(F \circ c)'(t) = T_{c(t)}F(c'(t))$. If c is an is an integral curve for ξ, then this equals $T_{c(t)}F(\xi(c(t))) = \eta(F(c(t)))$. Thus $F \circ c$ is an integral curve for η. Applying this to $c_x : J_x \to M$ for $x \in M$, we conclude that $J_x \subset J_{F(x)}$ (where we use the same notation for the maximal intervals for both vector fields). This implies both the relation between the domains of definition and the relation between the flows claimed in the first part. The second part then follows immediately from Proposition 2.3. □

2.8. Completeness. A vector field ξ on a smooth manifold M is called *complete* iff all its integral curves are defined for all times, i.e. if $\mathcal{D}(\xi) = M \times \mathbb{R}$. Then by Theorem 2.7, $\mathrm{Fl}^\xi_t : M \to M$ is smooth for all t and Fl^ξ_{-t} is a smooth inverse, so each $\mathrm{Fl}^\xi_t : M \to M$ is a diffeomorphism. Moreover, the equation in part (2) of Theorem 2.7 just says that $\mathrm{Fl}^\xi_{t+s} = \mathrm{Fl}^\xi_t \circ \mathrm{Fl}^\xi_s$, so $t \mapsto \mathrm{Fl}^\xi_t$ is a homomorphism from the abelian group $(\mathbb{R}, +)$ to the group of all diffeomorphisms of M (which evidently is a group under composition). Otherwise put $\mathrm{Fl}^\xi : \mathbb{R} \times M \to M$ defines a group action of $(\mathbb{R}, +)$ on M by diffeomorphisms. The flows of general vector fields should be thought of a "locally defined analog" of this situation. There are nice sufficient conditions for completeness.

Proposition 2.8. *Let ξ be a vector field on a smooth manifold M.*
(1) If there is $\epsilon > 0$ such that $M \times [-\epsilon, \epsilon] \subset \mathcal{D}(\xi)$, then ξ is complete.
(2) If ξ has compact support, then ξ is complete. In particular, any vector field on a compact manifold is complete.

Proof. (1) By assumption, $[-\epsilon, \epsilon] \subset J_x$ for all $x \in M$. Applying part (2) of Lemma 2.6 to $c_x(\epsilon)$ and $c_x(-\epsilon)$ we conclude that $[-2\epsilon, 2\epsilon] \subset J_x$ for all x. Iterating this argument, we see that $[-2^n\epsilon, 2^n\epsilon] \subset J_x$ for all x and $n \in \mathbb{N}$, which implies $J_x = \mathbb{R}$ for all x.

(2) If $\xi(x) = 0$ then the constant curve $c_x(t) = x$ is an integral curve through x, so $J_x = \mathbb{R}$ in this case. Putting $K := \mathrm{supp}(\xi) \subset M$, we conclude that $(M \setminus K) \times \mathbb{R} \subset \mathcal{D}(\xi)$. On the other hand, since $\mathcal{D}(\xi)$ is an open neighborhood of $K \times \{0\}$ in $M \times \mathbb{R}$, it is well known from general topology that $K \times (-\epsilon, \epsilon) \subset \mathcal{D}(\xi)$ for some $\epsilon > 0$. But this implies that $M \times [-\epsilon/2, \epsilon/2] \subset \mathcal{D}(\xi)$ and the result follows from (1). □

Example 2.8. Without compactness assumptions, completeness is a rather subtle property. In particular, the set of complete vector fields is neither a linear subspace of $\mathfrak{X}(M)$ nor closed under the Lie bracket. This can be shown by simple examples on $M = \mathbb{R}^2$. Define $\xi := x^2 \frac{\partial}{\partial x^1}$ and $\eta = (x^1)^2 \frac{\partial}{\partial x^2}$. The flows for these two vector fields are easily found: $\mathrm{Fl}^\xi_t(x^1, x^2) = (x^1 + tx^2, x^2)$ and $\mathrm{Fl}^\eta_t(x^1, x^2) = (x^1, x^2 + t(x^1)^2)$, so essentially they behave as the constant vector field in Example (1) of 2.6.

For the sum $\xi + \eta$, the system of differential equations describing the flow becomes $(x^1)'(t) = x^2(t)$ and $(x^2)'(t) = x^1(t)^2$, so we obtain $(x^1)''(t) = x^1(t)^2$. One immediately verifies that $t \mapsto (1 - t/\sqrt{6})^{-2}$ solves the second equation on the interval $(-\infty, \sqrt{6})$ and there is no extension to a larger interval, so $\xi + \eta$ is not complete.

Likewise, for the Lie bracket we get $[\xi, \eta] = 2x^1x^2 \frac{\partial}{\partial x^2} - (x^1)^2 \frac{\partial}{\partial x^1}$. Thus the equation for the first component of an integral curve is $(x^1)'(t) = -x^1(t)^2$ and as in Example (3) of 2.6 one concludes that the solution of this equation is not defined on all of $\mathbb{R}$.

2.9. The flow box theorem. As a simple application of the theory of flows, we can show that locally around points in which they are non-zero, all vector fields "look the same". More precisely, locally around a point in which it is non-zero, any vector field can be realized as a coordinate vector field. The idea for the proof is geometrically very transparent: One locally chooses an appropriate submanifold which is transversal to the vector field and then "flows out" from this.

Proposition 2.9. *Let ξ be a vector field on a smooth manifold M and let $x \in M$ be a point such that $\xi(x) \neq 0$. Then there is a local chart (U, u) for M with $x \in U$ such that $\xi|_U = \frac{\partial}{\partial u^1}$.*

In particular, for $y \in U$ and an open interval I around 0 such that $u(y)+te_1 \in u(U)$ for all $t \in I$, the flow of ξ is given by $\mathrm{Fl}^\xi_t(y) = u^{-1}(u(y)+te_1)$ for $t \in I$.

Proof. Choose some chart (V, v) for M with $x \in V$ such that $v(x) = 0 \in \mathbb{R}^n$ and $T_xv(\xi(x)) = e_1$, the first unit vector. (The first condition can be achieved by composing with a translation and since $\xi(x) \neq 0$ the second condition can be achieved by further composing with a linear map.) Now view $\mathbb{R}^{n-1}$ as the linear subspace of $\mathbb{R}^n$ of all vectors with first component 0 and consider the intersection $v(V) \cap \mathbb{R}^{n-1} \subset \mathbb{R}^{n-1}$. Now $(t, z) \mapsto (v^{-1}(0, z), t)$ is a smooth map $\mathbb{R} \times (v(V) \cap \mathbb{R}^{n-1}) \to M \times \mathbb{R}$, and of course $\mathbb{R} \times (v(V) \cap \mathbb{R}^{n-1}) \subset \mathbb{R}^n$ is an open subset. By Theorem 2.8, $\mathcal{D}(\xi) \subset M \times \mathbb{R}$ is open and $\mathrm{Fl}^\xi : \mathcal{D}(\xi) \to M$ is smooth, so there is an open neighborhood W of 0 in $\mathbb{R}^n$ on which $w(t, z) := \mathrm{Fl}^\xi_t(v^{-1}(0, z))$ defines a smooth map $W \to M$.

Now for each $(t_0, z_0) \in W$, there is $\epsilon > 0$ such that $(t, z_0) \in W$ for $t \in (t_0 - \epsilon, t_0 + \epsilon)$. By construction $t \mapsto w(t, z_0)$ is an integral curve for ξ, which implies that

$$T_{(t_0,z_0)}w(e_1) = \xi(w(t_0, z_0)). \tag{2.5}$$

On the other hand, $(0, z) \in W$ for z in some open neighborhood of zero in $\mathbb{R}^{n-1}$ and $w(0, z) = v^{-1}(0, z)$. This in particular shows that $T_{(0,0)}w$ coincides with $T_{(0,0)}v^{-1}$ on the subspace spanned by the last $n-1$ basis vectors. But from (2.5), we also see that these two maps coincide on multiples of the first basis vector. Thus $T_{(0,0)}w = T_{(0,0)}v^{-1}$ and hence is a linear isomorphism. By the inverse function theorem, there is an open neighborhood of zero on which w restricts to a diffeomorphism onto an open neighborhood U of x in M, and we let u be the inverse of this restriction. Then $u : U \to u(U)$ is a diffeomorphism, so we know from 1.8 that (U, u) is a chart on M. Finally, equation (2.5) shows that for $y \in U$ we get $T_yu(\xi(y)) = e_1$, which exactly says that $\xi|_U = \frac{\partial}{\partial u^1}$. The last statement then immediately follows from Example (1) of 2.6 by applying Corollary 2.7 to the diffeomorphism $w : u(U) \to U$. □

Excursion: Flows and Lie brackets

To conclude the chapter, we discuss the relations between the two points of view on vector fields that we have developed so far. This will also provide the first instance of a general way to construct an action of vector fields on geometric objects via the so-called Lie derivative.

2.10. The Lie derivative of vector fields. Via the flow map, a vector field ξ on a manifold M gives rise to local diffeomorphisms around each point. Along such a local diffeomorphism, we can pull back another vector field η. So locally around $x \in M$ and for $t \in \mathbb{R}$ sufficiently close to 0, we can consider $(\mathrm{Fl}^\xi_t)^*\eta$, and this defines a local vector field. In particular, for small enough t, we have $((\mathrm{Fl}^\xi_t)^*\eta)(x) \in T_xM$. Now one should expect that this depends smoothly on t and thus one obtains a smooth curve in the finite dimensional vector space T_xM. The derivative of this curve at $t = 0$ can be naturally interpreted as a an element of T_xM. Since this can be done in each point $x \in M$, we obtain a map $\zeta : M \to TM$ such that $p \circ \zeta = \mathrm{id}_M$. Finally, by the smooth dependence of all ingredients on all variables, one should expect that ζ is smooth and thus defines a vector field on M.

Now we can verify that this strategy actually works out and recovers the Lie bracket of vector fields.

Theorem 2.10. *Let $\xi, \eta \in \mathfrak{X}(M)$ be vector fields on a smooth manifold M.*

(1) For each $x \in M$, $t \mapsto ((\mathrm{Fl}^\xi_t)^*\eta)(x)$ is a smooth curve in T_xM which is defined on an open interval around 0. *Hence we can define*

$$\mathcal{L}_\xi\eta(x) := \tfrac{d}{dt}|_{t=0}((\mathrm{Fl}^\xi_t)^*\eta)(x) \in T_xM.$$

(2) The construction in (1) defines a smooth vector field $\mathcal{L}_\xi\eta \in \mathfrak{X}(M)$, and indeed $\mathcal{L}_\xi\eta = [\xi,\eta]$.

Proof. We actually have to prove first that we obtain a well defined vector field and then identify it with the Lie bracket. For a point $x \in M$, we know from Theorem 2.7 that there is an open neighborhood U of x in M and $\epsilon > 0$ such that $U \times (-\epsilon,\epsilon) \subset \mathcal{D}(\xi)$. Hence $(y,t) \mapsto (\mathrm{Fl}^\xi(y,t), -t)$ defines a smooth map $U \times (-\epsilon,\epsilon) \to M \times \mathbb{R}$ and we have observed in 2.7 that this has values in $\mathcal{D}(\xi)$. Next, for $(z,s) \in \mathcal{D}(\xi)$, we can consider $(\eta(z), 0) \in T_zM \times T_s\mathbb{R} \cong T_{(z,s)}(M \times \mathbb{R})$. This actually lies in $p^{-1}(\mathcal{D}(\xi)) = T\mathcal{D}(\xi)$ and we obtain a smooth map $\mathcal{D}(\xi) \to T\mathcal{D}(\xi)$. Finally, $T\,\mathrm{Fl}^\xi : T\mathcal{D}(\xi) \to TM$ is smooth as the tangent map of a smooth map. Composing these three smooth maps, we obtain a smooth map $\Phi : U \times (-\epsilon,\epsilon) \to TM$, which by construction satisfies

$$\Phi(y,t) = T_{(\mathrm{Fl}^\xi_t(y),-t)}\,\mathrm{Fl}^\xi(\eta(\mathrm{Fl}^\xi_t(y)),0) = T_{\mathrm{Fl}^\xi_t(y)}\,\mathrm{Fl}^\xi_{-t}(\eta(\mathrm{Fl}^\xi_t(y))). \tag{2.6}$$

(The last equality holds because a tangent vector of the form $(X,0)$ in a point $(z,-t)$ can be realized as the derivative at $s = 0$ of a curve of the form $c(s) = (c_1(s), -t)$. Then $\mathrm{Fl}^\xi(c(s)) = \mathrm{Fl}^\xi_{-t}(c_1(s))$ and differentiating at $s = 0$, the claim follows.) Applying (2.6) for the fixed point $y = x$, we observe that the right hand side has values in T_xM for all $t \in (-\epsilon,\epsilon)$ and this completes the proof of (1).

Possibly shrinking U, we may assume that it is the domain of a chart (U,u) for M. Then the last n components of $Tu \circ \Phi : U \times (-\epsilon,\epsilon) \to u(U) \times \mathbb{R}^n$ are smooth functions $\varphi^i : U \times (-\epsilon,\epsilon) \to \mathbb{R}$, $i = 1,\dots,n$ and by definition we obtain $\Phi(y,t) = \sum_i \varphi^i(y,t)\frac{\partial}{\partial u^i}(y)$. Now it follows readily from the definitions that for each i the last partial derivative $\frac{\partial\varphi^i}{\partial t}$ defines a smooth function $U \times (-\epsilon,\epsilon) \to \mathbb{R}$, too. By construction, we obtain

$$\mathcal{L}_\xi\eta(y) = \textstyle\sum_i \frac{\partial\varphi^i}{\partial t}(y,0)\frac{\partial}{\partial u^i}(y), \tag{2.7}$$

which shows that $\mathcal{L}_\xi\eta$ is a smooth vector field on U and hence on M.

To complete the proof, we have to show that $\mathcal{L}_\xi\eta(x) = [\xi,\eta](x)$ for all $x \in M$. Let us first consider the case that $\xi(x) \neq 0$. Then Theorem 2.9 shows that we can find a local chart (U,u) around x such that $\xi|_U = \frac{\partial}{\partial u^1}$. For $y \in U$ and t sufficiently close to 0, we then get $u(\mathrm{Fl}^\xi_t y) = u(y) + te_1$. Using Proposition 1.15, this shows that the local coordinate representation of $T\,\mathrm{Fl}^\xi_{-t}$ with respect to (U,u) is given by $(z,v) \mapsto (z - te_1, v)$. The expansion $\eta^j\frac{\partial}{\partial u^j}$ of η by definition means that the local coordinate representation of $\eta : M \to TM$ with respect to the charts (U,u) and $(p^{-1}(U), Tu)$ is given by $z \mapsto (z, \eta^1(u^{-1}(z)),\dots,\eta^n(u^{-1}(z)))$. Together with the above, this exactly says that, in a neighborhood of x and for t sufficiently close to 0, we obtain $\varphi^i(y,t) = \eta^i(u^{-1}(u(y)+te_1))$ for $i = 1,\dots n$ in the notation from above.

Hence the component functions in (2.7) are just the first partial derivatives of the functions $\eta^i \circ u^{-1}$. But these are just the functions $\frac{\partial}{\partial u^1}(\eta^i)$, which are the components of $[\frac{\partial}{\partial u^1}, \eta]$ by part (5) of Theorem 2.4. This completes the argument in the case $\xi(x) \neq 0$. By continuity of $\mathcal{L}_\xi\eta$ and $[\xi,\eta]$, the two fields then have to coincide on $\mathrm{supp}(\xi)$. But $V := M \setminus \mathrm{supp}(\xi)$ is an open subset of M on which ξ vanishes identically. Hence on V, Fl^ξ_t exists and is the identity for all t, which readily implies that $\mathcal{L}_\xi\eta|_V = 0$, while $[\xi,\eta]|_V = 0$ by part (4) of Theorem 2.4. □

For later developments it will be useful to see that also the action of vector fields on smooth functions fits into the the picture of a Lie derivative: There is an obvious way how to pull back a smooth function $f : N \to \mathbb{R}$ along a local diffeomorphism $F : M \to N$, by simply putting $F^*f := f \circ F$. (Indeed, this extends to arbitrary smooth maps, but this is not important at this stage.) In particular, given $f \in C^\infty(M, \mathbb{R})$ and $\xi \in \mathfrak{X}(M)$, then locally around each $x \in M$ and for t sufficiently close to 0, we can consider $(\mathrm{Fl}^\xi_t)^*f : M \to \mathbb{R}$. By definition, in terms of the integral curve c_x through x, we get $(\mathrm{Fl}^\xi_t)^*f(x) = f(c_x(t))$. Differentiating this at $t = 0$, we get $T_xf(\xi(x)) = \xi(f)(x)$. Hence we can define the Lie derivative of f by $\mathcal{L}_\xi f(x) := \frac{d}{dt}|_{t=0}((\mathrm{Fl}^\xi_t)^*f(x))$ and observe that it defines a smooth function $\mathcal{L}_\xi f : M \to \mathbb{R}$, which coincides with $\xi(f)$.

2.11. Commuting vector fields. We next want to characterize the condition that for two vector fields $\xi, \eta \in \mathfrak{X}(M)$, the Lie bracket $[\xi, \eta]$ vanishes identically. This first needs a small extension of Theorem 2.10:

LEMMA 2.11. *Let $\xi, \eta \in \mathfrak{X}(M)$ be vector fields on a smooth manifold M and let $x \in M$ and $t_0 \in \mathbb{R}$ be such that $(x, t_0) \in \mathcal{D}(\xi)$. Then for all y in some neighborhood of x and t close enough to t_0, $t \mapsto (\mathrm{Fl}^\xi_t)^*\eta(y)$ defines a smooth curve in T_yM and*

$$\tfrac{d}{dt}|_{t=t_0}((\mathrm{Fl}^\xi_t)^*\eta)(y) = \mathcal{L}_\xi((\mathrm{Fl}^\xi_{t_0})^*\eta)(y) = ((\mathrm{Fl}^\xi_{t_0})^*(\mathcal{L}_\xi\eta))(y).$$

PROOF. From Theorem 2.7, we know there is a neighborhood U of x in M and $\epsilon > 0$ such that $U \times (t_0 - \epsilon, t_0 + \epsilon) \subset \mathcal{D}(\xi)$ and then smoothness follows as in the proof of Theorem 2.10. From Theorem 2.7, we also know that $\mathrm{Fl}^\xi_{t_0+s} = \mathrm{Fl}^\xi_{t_0} \circ \mathrm{Fl}^\xi_s$ and as observed in Section 2.3 this implies that $(\mathrm{Fl}^\xi_{t_0+s})^*\eta = (\mathrm{Fl}^\xi_s)^*((\mathrm{Fl}^\xi_{t_0})^*\eta)$. Differentiating this with respect to s at $s = 0$ in a point and using Theorem 2.10, the first claimed equality follows.

We can also write $\mathrm{Fl}^\xi_{t_0+s} = \mathrm{Fl}^\xi_s \circ \mathrm{Fl}^\xi_{t_0}$ and inserting this, we obtain $(\mathrm{Fl}^\xi_{t_0+s})^*\eta = (\mathrm{Fl}^\xi_{t_0})^*((\mathrm{Fl}^\xi_s)^*\eta)$. Evaluating the right hand side in y, we get $T_{\tilde{y}}\,\mathrm{Fl}^\xi_{-t_0}((\mathrm{Fl}^\xi_s)^*\eta(\tilde{y}))$, where $\tilde{y} = \mathrm{Fl}^\xi_{t_0}(y)$. Differentiating with respect to s at $s = 0$, we can differentiate through the constant linear map $T_{\tilde{y}}\,\mathrm{Fl}^\xi_{-t_0}$, thus obtaining $T_{\tilde{y}}\,\mathrm{Fl}^\xi_{-t_0}(\mathcal{L}_\xi\eta(\tilde{y}))$. This gives the second claimed equality. □

To formulate the desired characterization, we say that the flows of two vector fields $\xi, \eta \in \mathfrak{X}(M)$ commute if and only if for each $x \in M$ and any open intervals $I, J \subset \mathbb{R}$ containing zero the following is satisfied. If for one of the expressions $\mathrm{Fl}^\xi_t(\mathrm{Fl}^\eta_s(x))$ and $\mathrm{Fl}^\eta_s(\mathrm{Fl}^\xi_t(x))$ exists for all $(t, s) \in I \times J$, then the other expression exists for all t and s and they are equal. Using this terminology, we state

THEOREM 2.11. *For two vector fields ξ and η on a smooth manifold M, the following conditions are equivalent:*

(1) $[\xi, \eta] = 0$.

(2) For any $(x, t_0) \in \mathcal{D}(\xi)$ we get $(\mathrm{Fl}^\xi_t)^*\eta = \eta$ locally around x and for t in an open interval containing 0 and t_0.

(3) For any $(x, t_0) \in \mathcal{D}(\eta)$ we get $(\mathrm{Fl}^\eta_t)^*\xi = \xi$ locally around x and for t in an open interval containing 0 and t_0.

(4) The flows of ξ and η commute.

PROOF. (1) $\Rightarrow$ (2): Take $(x, t_0) \in \mathcal{D}(\xi)$. Since $\mathcal{D}(\xi)$ is open in $M \times \mathbb{R}$ and $\{x\} \times [0, t_0]$ is compact, there is an open neighborhood U of $x \in M$ and an open interval $I \subset \mathbb{R}$ with $0, t_0 \in I$ such that $U \times I \subset \mathcal{D}(\xi)$. Hence $(\mathrm{Fl}^\xi_t)^*\eta(y)$ is defined for all $t \in I$ and $y \in U$.

But by Lemma 2.11, the curve $t \mapsto ((\mathrm{Fl}^\xi_t)^*\eta)(y)$ in T_yM has vanishing derivative and thus is constant. Since it clearly gives $\eta(y)$ for $t = 0$, the result follows.

(2) $\Rightarrow$ (1): Since $(x, 0) \in \mathcal{D}(\xi)$ for all $x \in M$, we see that $(\mathrm{Fl}^\xi_t)^*\eta = \eta$ on some neighborhood of x and for t sufficiently close to 0. Thus we can differentiate at $t = 0$ and (1) follows from Theorem 2.10.

Since condition (1) is symmetric in ξ and η, the equivalence of (1) and (3) follows in the same way.

To show that these equivalent conditions imply (4), let us assume that we have x, I and J such that $\mathrm{Fl}^\eta_s(\mathrm{Fl}^\xi_t(x))$ exists for all $(t, s) \in I \times J$. Fix s and put $c(t) := \mathrm{Fl}^\eta_s(\mathrm{Fl}^\xi_t(x))$, so $c(0) = \mathrm{Fl}^\eta_s(x)$ and $c'(t) = T_{\mathrm{Fl}^\xi_t(x)} \mathrm{Fl}^\eta_s(\xi(\mathrm{Fl}^\xi_t(x)))$. Of course $(c(t), -s) \in \mathcal{D}(\eta)$ for any t and hence by (3), $(\mathrm{Fl}^\eta_{-s})^*\xi = \xi$ on some neighborhood of any of the points $c(t)$. This implies that $c'(t) = \xi(c(t))$ and hence $c(t) = \mathrm{Fl}^\xi_t(\mathrm{Fl}^\eta_s(x))$, so (4) is satisfied. If we assume the other combination of flows exists, we use condition (2) instead of (3).

So let us finally assume that (4) is satisfied. Taking $\epsilon > 0$ small enough, we can assume that $\mathrm{Fl}^\xi_t(\mathrm{Fl}^\eta_s(x))$ is defined for $|s|, |t| < \epsilon$ and by (4) this equals $\mathrm{Fl}^\eta_s(\mathrm{Fl}^\xi_t(x))$. Differentiating with respect to s at $s = 0$, we get $T_x \mathrm{Fl}^\xi_t(\eta(x)) = \eta(\mathrm{Fl}^\xi_t(x))$ for all $|t| < \epsilon$. Applying $T_{\mathrm{Fl}^\xi_t(x)} \mathrm{Fl}^\xi_{-t}$ we see that $\eta(x) = (\mathrm{Fl}^\xi_{-t})^*\eta(x)$ for $|t| < \epsilon$ and differentiating with respect to t at $t = 0$, $[\xi, \eta](x) = 0$ follows. □

Phrased in our current language, we have shown in Theorem 2.4 that the coordinate vector fields with respect to any local chart for a manifold M always commute with each other. Now we can show that apart from their point-wise linear independence, this is the only characterizing property of coordinate vector fields.

Corollary 2.11. *Let M be a smooth manifold of dimension n and let $\xi_1, \dots \xi_n$ be local vector fields defined on some open subset $V \subset M$ such that*

- *For each $x \in V$, $\{\xi_1(x), \dots, \xi_n(x)\}$ is a basis for T_xM.*
- *For each i, j, we get $[\xi_i, \xi_j] = 0$.*

Then for each $x \in V$, there is a chart (U, u) for M with $x \in U \subset V$ such that $\xi_i|_U = \frac{\partial}{\partial u^i}$ for each $i = 1, \dots, n$.

Proof. Consider the map $\varphi(t^1, \dots, t^n) := (\mathrm{Fl}^{\xi_1}_{t^1} \circ \dots \circ \mathrm{Fl}^{\xi_n}_{t^n})(x)$. From Theorem 2.7 it easily follows that this is well defined and smooth from some open neighborhood of 0 in $\mathbb{R}^n$ to V. By Theorem 2.11 the succession in which we compose the flow maps does not make any difference. Now we compute the ith partial derivative of φ as $\frac{d}{ds}|_{s=0}\varphi(t^1, \dots, t^i + s, \dots t^n)$. To do this, we write $\mathrm{Fl}^{\xi_i}_{t^i+s} = \mathrm{Fl}^{\xi_i}_s \circ \mathrm{Fl}^{\xi_i}_{t^i}$ and then commute the term $\mathrm{Fl}^{\xi_i}_s$ to the very left of the expression. This then shows that this ith partial derivative is given by $\xi_i(\varphi(t^1, \dots, t^n))$. By linear independence of the ξ_i, $T\varphi$ is a linear isomorphism in each point. In particular, there is an open neighborhood W of 0 such that φ restricts to a diffeomorphism from W onto an open neighborhood U of x in M. Putting $u := (\varphi|_W)^{-1}$, we obtain a chart (U, u) with $x \in U \subset V$, and the computation of partial derivatives above shows that $\xi_i|_U = \frac{\partial}{\partial u^i}$ for all i. □

2.12. Remarks on further developments. (1) Generalizing Theorem 2.11 there are improved results relating the non-commutativity of the flows of vector fields to their Lie brackets. Given two vector fields $\xi, \eta \in \mathfrak{X}(M)$ and a point $x \in M$, one considers the map

$$\alpha(t, s) := (\mathrm{Fl}^\xi_{-s} \circ \mathrm{Fl}^\eta_{-t} \circ \mathrm{Fl}^\xi_s \circ \mathrm{Fl}^\eta_t)(x).$$

This is defined and smooth on some open neighborhood of zero in $\mathbb{R}^2$, so $c(t) := \alpha(t,t)$ is a smooth curve through x defined on an interval containing zero. If ξ and η commute, then this clearly is the constant curve x. Now the first obvious idea is to consider $c'(0)$ but unfortunately, it turns out that $c'(0) = 0$. Indeed, $\alpha(t,0) = \alpha(0,s) = x$ for all t and s, and hence both partial derivatives and thus all directional derivatives of α in $(0,0)$ vanish. But one readily verifies that $c'(0) = 0$ implies that $f \mapsto (f \circ c)''(0)$ is a derivation on smooth functions and hence there is a well defined element $c''(0) \in T_xM$. Now it turns out that $c''(0) = 2[\xi,\eta](x)$ which gives a more quantitative version of the statement that the Lie bracket measures non-commutativity of the flows.

The proof for this result is a rather tedious computation, see Theorem 3.16 in [**Michor**], which actually proves a much more general result.

(2) An important result related to our current discussions is the Frobenius theorem. This starts from the concept of a *distribution* (not related to generalized functions, just a different object with the same name) on a smooth manifold. By definition, a distribution of rank k is given by specifying in each point $x \in M$ a k-dimensional linear subspace $E_x \subset T_xM$. A (local) smooth *section* of E then is a (local) vector field ξ on M such that $\xi(x) \in E_x$ for all x in the domain of definition of ξ. A distribution $E \subset TM$ is called *smooth* if any point $x \in M$ has an open neighborhood U in M such that there are local smooth sections $\xi_1, \dots, \xi_k$ of E defined on U, such that $\xi_1(y), \dots, \xi_k(y)$ is a basis for E_y for all $y \in U$. Such a family is then called a (smooth) local *frame* for E.

The basic question that arises in this context is whether the subspaces forming the distribution can be realized as tangent spaces of submanifolds. More formally, the distribution E is called *integrable* if for each $x \in M$, there exists a k-dimensional submanifold $N \subset M$ with $x \in N$ such that $T_yN = E_y$ for all $y \in N$. From what we have done so far, we see that there is a simple necessary condition for integrability. If $N \subset M$ is an integral submanifold, then as discussed in Example 2.5, for vector fields $\xi, \eta \in \mathfrak{X}(M)$ that are tangent to N along N, also the Lie bracket $[\xi,\eta]$ is tangent to N along N. This implies that for an integrable distribution E and any two local sections ξ, η of E also the Lie bracket $[\xi,\eta]$ must be a (local) section of E. Smooth distributions with this property are called *involutive*.

At the first sight it looks like involutivity would be hard to check, since one has to check a condition for all local sections. Locally, however, this boils down to a finite problem: Suppose that we have a local frame $\{\xi_1, \dots, \xi_k\}$ for E defined on U. Then of course involutivity of E implies that $[\xi_i, \xi_j]$ is a (local) section of E for all i, j, but the key fact is that also the converse is true: One easily shows that any local section of E defined on U can be written as $\sum_i f_i\xi_i$ for smooth functions $f_i : U \to \mathbb{R}$. But then using the properties of the Lie bracket proved in Theorem 2.4, we see that we get

$$[\textstyle\sum_i f_i\xi_i, \sum_j g_j\xi_j] = \sum_{i,j} \left(f_i\xi_i(g_j)\xi_j - g_j\xi_j(f_i)\xi_i - f_ig_j[\xi_i,\xi_j]\right),$$

and the first two summands clearly are sections for each i and j.

Now the Frobenius theorem states that for smooth distributions involutivity implies integrability. In fact, one can construct local charts (U,u) for M with $u(U) = V \times W \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$ for open subsets V and W such that for each fixed $y \in W$, $u^{-1}(V \times \{y\})$ is an integral manifold for E. The proof of this result does not really go beyond what we have done so far. The basic idea is that one first constructs local frames for E consisting of commuting vector fields (which mainly needs linear algebra and observing that things depend smoothly on a point). Having found such frames, the charts are then constructed via a mix of the proofs of Proposition 2.9 and of Corollary 2.11. A complete proof can be found in Chapter 2 of my lecture notes [**Cap:Lie**] on Lie groups, which also contains several applications. The Frobenius theorem actually is one of the cornerstones of analysis on manifolds with a broad range of applications.

It should be mentioned that there are also extensions of the theory of distributions to the case where the dimension of the subspaces E_x is allowed to depend on x ("distributions of non-constant rank"), for which things become more subtle. This is discussed in the last part of Chapter 3 of [**Michor**].

CHAPTER 3

Tensor fields and differential forms

We now turn to a more general class of geometric objects. Initially, the definition of these objects gets its main input from multilinear algebra rather then from analysis or geometry. Therefore, the "best" way of formulating things depends quite a lot on the background of a reader on topics from multilinear algebra like tensor products. It is possible to basically avoid the use of tensor products by "converting" all objects to multilinear maps. This obscures many aspects, however, since it singles out one of several possible interpretations of a tensor product. Being able to fluently use tensor products and shift between different interpretations simplifies many arguments considerably. The notes try to follow an intermediate approach, tensor products will be used, but some details are added to arguments involving tensor products. It is important to notice that we only need tensor products of finite dimensional vector spaces, so a construction via free vector spaces is not really needed here.

It should also be mentioned here that at this point comparison to analysis on open subsets of $\mathbb{R}^n$ can become a bit difficult. On open subsets, all the objects we consider can be interpreted as smooth functions with values in appropriate vector spaces and people often are not very careful in distinguishing between different types of objects, say between the derivative of a real valued smooth function (which, as we shall see readily, is a one-form) and its gradient (which is a vector field in our interpretation).

3.1. One-forms. This is the simplest instance of the type of construction we have in mind. For a smooth manifold M, we want to look at a map ω, that associates to each $x \in M$ a linear map $\omega(x) : T_xM \to \mathbb{R}$. To talk about smoothness, there are at least two natural possibilities. On the one hand, given a vector field $\xi \in \mathfrak{X}(M)$, we can define $\omega(\xi) : M \to \mathbb{R}$ by $\omega(\xi)(x) := \omega(x)(\xi(x))$. Similarly, for an open subset $U \subset M$ and a local vector field ξ defined on U, we can form $\omega(\xi) : U \to \mathbb{R}$. On the other hand, ω defines a function $\hat{\omega} : TM \to \mathbb{R}$, defined by $\hat{\omega}(X) := \omega(p(X))(X)$, where $p : TM \to M$ is the projection. All these lead to the same concept of smoothness:

Lemma 3.1. *For a function ω as above, the following conditions are equivalent:*

(1) For any vector field $\xi \in \mathfrak{X}(M)$, the function $\omega(\xi) : M \to \mathbb{R}$ is smooth.

(2) For any open subset $U \subset M$ and any local vector field η defined on U, the function $\omega(\eta) : U \to \mathbb{R}$ is smooth.

(3) The function $\hat{\omega} : TM \to \mathbb{R}$ is smooth.

Proof. Since $\omega(\eta) = \hat{\omega} \circ \eta : U \to \mathbb{R}$, (3) implies (2). Conversely, assume that (2) is satisfied, take a chart (U, u) for M and the induced chart $(p^{-1}(U), Tu)$ for TM. Then for $(z, v) \in u(U) \times \mathbb{R}^n$ with $v = (v_1, \dots, v_n)$, we get $(Tu)^{-1}(z, v) = \sum_i v_i \frac{\partial}{\partial u^i}(u^{-1}(z))$. Hence

$$(\hat{\omega} \circ Tu^{-1})(z, v) = \textstyle\sum_i v_i \omega(\frac{\partial}{\partial u^i})(u^{-1}(z)),$$

so this is evidently smooth. Taking the charts from an atlas for M, we see that $\hat{\omega}$ has smooth local coordinate representations with respect to the induced atlas of TM and hence is smooth. Thus (2) implies (3).

Since (2) obviously implies (1), it remains to show that (1) implies (2). Given U and η, take a point $x \in U$. Then there is an open neighborhood V of x in M such that $\bar{V} \subset U$. By Proposition 2.1, there is a vector field $\tilde{\eta} \in \mathfrak{X}(M)$ such that $\tilde{\eta}|_V = \eta_V$, and then by construction $\omega(\tilde{\eta})|_V = \omega(\eta)|_V$. Thus $\omega(\eta)$ is smooth on an open neighborhood of x and since x was arbitrary, this completes the proof. $\square$

Definition 3.1. A *one-form* ω on a smooth manifold M is a map, which associates to each point $x \in M$ a linear map $\omega(x) : T_xM \to \mathbb{R}$ and satisfies the equivalent conditions of Lemma 3.1. The space of all one-forms on M is denoted by $\Omega^1(M)$.

From the definitions, it follows immediately that pointwise addition and multiplication by smooth functions makes $\Omega^1(M)$ into a vector space and a module over $C^\infty(M, \mathbb{R})$. It is also no problem to talk about local one-forms on a smooth manifold M, these are just one-forms on an open subset $U \subset M$. More surprisingly, for a smooth function $f : M \to \mathbb{R}$, $x \in M$ and $X \in T_xM$ we can define $df(x)(X_x) := T_xf(X_x) \in T_{f(x)}\mathbb{R} = \mathbb{R}$. Observing that $df(\xi) = \xi(f)$ for $\xi \in \mathfrak{X}(M)$, we see that we obtain $df \in \Omega^1(M)$. Thus we can naturally associate a one-form to a function, which is not possible with vector fields without choosing additional structures. Another indication for the usefulness of one-forms is that, in contrast to vector fields, they can be "moved" between manifolds by arbitrary smooth maps:

Proposition 3.1. *Let $F : M \to N$ be a smooth map between smooth manifolds. Then for any one-form $\omega \in \Omega^1(N)$, $(F^*\omega)(x)(X) := \omega(F(x))(T_xF(X))$ defines a one-form $F^*\omega \in \Omega^1(M)$. In particular, for $f \in C^\infty(N, \mathbb{R})$, we get $F^*df = d(f \circ F)$.*

Proof. Since $\omega(F(x)) : T_{F(x)}N \to \mathbb{R}$ and $T_xF : T_xM \to T_{F(x)}N$ are linear, $F^*\omega$ associates to each $x \in M$ a linear map $T_xM \to \mathbb{R}$. Concerning smoothness, we observe that in the notation of Lemma 3.1 we get $\widehat{F^*\omega} = \hat{\omega} \circ TF$, so smoothness follows. The last part readily follows from the chain rule $T_x(f \circ F) = T_{F(x)}f \circ T_xF$. $\square$

This immediately leads to the description of one-forms in local coordinates. Given a chart (U, u) for M, we have the local coordinates $u^i : U \to \mathbb{R}$ which define one-forms $du^i \in \Omega^1(U)$ for $i = 1, \dots, n$. In particular, for the coordinate vector fields, we get $du^i(\frac{\partial}{\partial u^j}) = \delta^i_j$. Given $\omega \in \Omega^1(M)$, we obtain smooth functions $\omega_i := \omega(\frac{\partial}{\partial u^i}) : U \to \mathbb{R}$, and we readily conclude that $\omega|_U = \omega_i du^i$. (Recall that we always use the summation convention.) In particular, for a vector field ξ with $\xi|_U = \xi^j \frac{\partial}{\partial u^j}$, we obtain $\omega(\xi)|_U = \omega_i \xi^i$. The behaviour under a change of charts can be deduced directly from the change of coordinate vector fields. For two charts (U_α, u_α) and (U_β, u_β) with $U_{\alpha\beta} \neq \emptyset$ let us denote the component functions of ω by ω^α_i and ω^β_j, respectively, and the chart changes by $u_{\alpha\beta} : u_\beta(U_{\alpha\beta}) \to u_\alpha(U_{\alpha\beta})$. Using formula (1.2) from Section 1.14, we conclude that

$$\omega^\beta_i = \omega(\tfrac{\partial}{\partial u^i_\beta}) = \omega\left((\partial_i u^j_{\alpha\beta})\tfrac{\partial}{\partial u^j_\alpha}\right) = (\partial_i u^j_{\alpha\beta})\omega^\alpha_j. \tag{3.1}$$

In the last equality, we have used that $\omega(f\xi) = f\omega(\xi)$ for $\xi \in \mathfrak{X}(M)$ and $f \in C^\infty(M, \mathbb{R})$, which is obvious from the point-wise definition of the action of one-forms on vector fields. Also, compare this carefully to formula (2.1) for vector fields from Proposition 2.1: Here the Jacobi-matrix of $u_{\alpha\beta}$ gives the coefficients in the expansion of the functions ω^β_i in terms of the ω^α_j, while there these were the coefficients of the functions ξ^i_α in terms of the functions ξ^j_β. Thus these coefficients are inverse matrices, which expresses the duality between vector fields and one-forms.

To get a more systematic description using linear algebra, take a point x in a smooth manifold M and define the *cotangent space T^*_xM of M at x* as the dual of the tangent space, i.e. $T_x^*M := (T_xM)^* = L(T_xM,\mathbb{R})$. For $\omega \in \Omega^1(M)$, we thus have $\omega(x) \in T_x^*M$. Our construction of the tangent bundle was based on the fact that any chart (U,u) for M with $x \in U$ gives rise to an identification of T_xM with $\mathbb{R}^n$ via T_xu. But since the construction of the dual space is natural (i.e. functorial), such an identification also induces an identification T_x^*M with $\mathbb{R}^{n*}$. Since the dual of a linear map goes in the opposite direction, we have to dualize $(T_xu)^{-1} : \mathbb{R}^n \cong T_{u(x)}\mathbb{R}^n \to T_xM$ to obtain $T_x^*u := ((T_xu)^{-1})^* : T_x^*M \to \mathbb{R}^{n*}$. This is easily made explicit: for $\lambda \in T_x^*M$, $T_x^*u(\lambda)$ by definition acts on $v \in \mathbb{R}^n$ as $\lambda((T_xu)^{-1}(v))$, so for $v = e_i$, one obtains $\lambda(\frac{\partial}{\partial u^i}(x))$.

For two compatible charts (U_α, u_α) and (U_β, u_β) with $x \in U_{\alpha\beta}$, one easily concludes that

$$T_x^*u_\alpha \circ (T_x^*u_\beta)^{-1} = (Du_{\beta\alpha}(u_\alpha(x)))^* : \mathbb{R}^{n*} \to \mathbb{R}^{n*}.$$

This of course depends smoothly on x, so we can imitate the construction of the tangent bundle from Theorem 1.15. We define a set T^*M to be the disjoint union of the cotangent spaces T_x^*M, which leads to a map $p : T^*M \to M$. For a chart (U,u) we piece the maps T_x^*u together to define $T^*u : p^{-1}(U) \to u(U) \times \mathbb{R}^{n*}$. Starting from a countable atlas (U_α, u_α), we obtain charts $(p^{-1}(U_\alpha), T^*u_\alpha)$ for T^*M,, which are smoothly compatible. Via Lemma 1.6 (and a separate argument for the Hausdorff property as in the case of TM), we conclude that T^*M is a smooth manifold and $p : T^*M \to M$ is a smooth map.

From the explicit description above, we conclude that for a map $\omega : M \to T^*M$ with $p \circ \omega = \mathrm{id}_M$, the local coordinate representation with respect to the charts (U,u) and $(p^{-1}(U), T^*u)$ is given by $z \mapsto (z, (\omega_1(u^{-1}(z)), \dots, \omega_n(u^{-1}(z))))$, where $\omega_i = \omega(\frac{\partial}{\partial u^i})$. This then directly implies that one-forms on M are exactly smooth maps $\omega : M \to T^*M$ such that $p \circ \omega = \mathrm{id}$. Thus we have arrived at a description closely similar to vector fields. There is much more symmetry in the situation: Recall that $X \in T_xM$ defines a linear map $T_x^*M \to \mathbb{R}$ by evaluating linear maps on X. In particular, given a map $\xi : M \to TM$, we can define $\xi(\omega)$ for each $\omega \in \Omega^1(M)$ (which equals $\omega(\xi)$) as well as $\hat{\xi} : T^*M \to \mathbb{R}$ by $\hat{\xi}(\lambda) := \lambda(\xi(p(\lambda)))$. One immediately verifies that smoothness of ξ is equivalent to smoothness of $\hat{\xi}$ and to smoothness of $\xi(\omega)$ for each $\omega \in \Omega^1(M)$.

REMARK 3.1. Observe that the dualization that takes place when going from TM to T^*M has opposite effects for the functorial properties of the construction and for the possibility to "transport" sections (i.e. vector fields respectively one-forms). Any smooth map $F : M \to N$ induces a smooth map $TF : TM \to TN$, but to define a pullback of vector fields, we have to require that each tangent map is invertible, i.e. that F is a local diffeomorphism. Now the dual to a tangent map $T_xF : T_xM \to T_{F(x)}N$ is $(T_xF)^* : T^*_{F(x)}N \to T_x^*M$. This allows us to define a pullback of one-forms for any F, since by definition $(F^*\omega)(x) = (T_xF)^*(\omega(F(x)))$. But these dual maps cannot be collected to define a map T^*F. To make the construction of the cotangent bundle functorial, we have to assume that $F : M \to N$ is a local diffeomorphism. As above, this allows us to define $T_x^*F := ((T_xF)^{-1})^* : T_x^*M \to T^*_{F(x)}N$ and these fit together to define a smooth map $T^*F : T^*M \to T^*N$ such that $p \circ T^*F = F \circ p$.

3.2. Tensor fields. Having prepared the case of one-forms carefully, we can now introduce the general concept of tensor fields. Given an vector space E and $k \in \mathbb{N}$, we write $\otimes^k E$ for the tensor product of k copies of E, with $\otimes^1 E = E$ and $\otimes^0 E = \mathbb{R}$. For a linear map $\varphi : E \to F$ between two vector spaces, we write $\otimes^k\varphi : \otimes^k E \to \otimes^k F$ for the induced map on tensor powers. On a tensor product of vectors, this just acts by applying φ to each of the vectors and forming the tensor product of resulting elements of F. Since all possible tensor products of k elements from a basis of E form a basis of $\otimes^k E$, this shows that the matrix entries of $\otimes^k\varphi$ with respect to this induced basis are obtained by polynomial expressions from the entries of the matrix of φ with respect to the original basis. So if φ depends smoothly on some parameters, then also $\otimes^k\varphi$ depends smoothly on these parameters. For tensor products of two different vector spaces, we use analogous notation and similar arguments apply.

Given a smooth manifold M, a point $x \in M$ and $(k,\ell) \in \mathbb{N}\times\mathbb{N}$, we can form the vector space $\otimes^k T_x^*M \otimes \otimes^\ell T_xM$. One possible interpretation of this space is as $k+\ell$-linear maps $(T_xM)^k \times (T_x^*M)^\ell \to \mathbb{R}$. Keep in mind that two such multilinear maps coincide if there values on all possible combinations of elements from some fixed bases of T_xM and T_x^*M agree. Given a map t that associates to each $x \in M$ an element $t(x) =: t_x \in \otimes^k T_x^*M \otimes \otimes^\ell T_xM$, k vector fields $\xi_1,\dots,\xi_k \in \mathfrak{X}(M)$ and ℓ one-forms $\omega^1,\dots,\omega^\ell \in \Omega^1(M)$, we get a function $t(\xi_1,\dots,\xi_k,\omega^1,\dots,\omega^\ell) : M \to \mathbb{R}$ whose value in x is $t_x(\xi_1(x),\dots,\xi_k(x),\omega^1(x),\dots,\omega^\ell(x))$.

Definition 3.2. A $\binom{\ell}{k}$-*tensor field* on a smooth manifold M is a map t that associates to each $x \in M$ an element $t_x \in \otimes^k T_x^*M \otimes \otimes^\ell T_xM$ such that for all $\xi_1,\dots,\xi_k \in \mathfrak{X}(M)$ and $\omega^1,\dots,\omega^\ell \in \Omega^1(M)$, the function $t(\xi_1,\dots,\xi_k,\omega^1,\dots\omega^\ell) : M \to \mathbb{R}$ is smooth. The space of all such tensor fields will be denoted by $\mathcal{T}_k^\ell(M)$.

As before, pointwise addition and multiplication by smooth functions makes $\mathcal{T}_k^\ell(M)$ into a vector space and a module over $C^\infty(M,\mathbb{R})$. There are obvious concepts of locally defined tensor fields and restriction of tensor fields to open subsets. Moreover, we can also insert vector fields and one-forms defined on an open subset U into (the restriction to U of) a tensor field to obtain a function defined on U. Similarly to the proof of Lemma 3.1, one shows that the resulting function is always smooth on U. The only fact needed here is that for $\omega \in \Omega^1(U)$ and $x \in U$, there exists $\tilde\omega \in \Omega^1(M)$ and an open neighborhood V of x such that $\tilde\omega|_V = \omega|_V$. This is proved using a bump function in exactly the same way as part (4) of Proposition 2.1.

We can also easily derive the description of tensor fields in local coordinates. Suppose that (U,u) is a chart for M and consider the associated coordinate vector fields $\frac{\partial}{\partial u^i}$ and one-forms du^j. Given $t \in \mathcal{T}_k^\ell(M)$ and any $k+\ell$-tuple $(i_1,\dots,i_k,j_1,\dots,j_\ell)$ of integers in $\{1,\dots,n\}$ we get a smooth function

$$t_{i_1\dots i_k}^{j_1\dots j_\ell} := t\left(\tfrac{\partial}{\partial u^{i_1}},\dots,\tfrac{\partial}{\partial u^{i_k}},du^{j_1},\dots,du^{j_\ell}\right) \in C^\infty(U,\mathbb{R}). \tag{3.2}$$

On the other hand, for each choice of indices as above we obtain a locally defined tensor field, which in hindsight we denote by

$$du^{i_1}\otimes\dots\otimes du^{i_k}\otimes\tfrac{\partial}{\partial u^{j_1}}\otimes\dots\otimes\tfrac{\partial}{\partial u^{j_\ell}} \in \mathcal{T}_k^\ell(U). \tag{3.3}$$

(We shall see later on that this is indeed a tensor product of coordinate vector fields and coordinate one-forms, but here we use this as a symbol.) This is defined by the fact that its value in x sends $X_1,\dots,X_k \in T_xM$ and $\lambda^1,\dots,\lambda^\ell \in T_x^*M$ to

$$(\textstyle\prod_r du^{i_r}(x)(X_r))(\prod_s \lambda^s(\tfrac{\partial}{\partial u^{j_s}}(x))) \in \mathbb{R}.$$

This definition readily implies that inserting arbitrary vector fields $\xi_1,\dots,\xi_r$ and one-forms $\omega^1,\dots,\omega^s$ defined on U into the tensor field (3.3), one obtains a product of component functions $\xi_1^{i_1}\dots\xi_r^{i_r}\omega_{j_1}^1\dots\omega_{j_s}^s$. Hence we have indeed defined a smooth tensor field on U. Moreover, this also easily implies that

$$t|_U = t_{i_1\dots i_k}^{j_1\dots j_\ell}du^{i_1}\otimes\dots\otimes du^{i_k}\otimes\tfrac{\partial}{\partial u^{j_1}}\otimes\dots\otimes\tfrac{\partial}{\partial u^{j_\ell}}, \tag{3.4}$$

since both sides coincide on all combinations of coordinate vector fields and coordinate one-forms. Observe that in (3.4) the summations convention applies to all $k+\ell$ indices, so this actually is a sum with $n^{k+\ell}$ terms.

The behavior of these component functions under a change of charts is also not difficult to describe, but the formulae get a bit long. For two charts (U_α, u_α) and (U_β, u_β) with $U_{\alpha\beta} \neq \emptyset$ let us define functions $A^i_j, B^r_s : U_{\alpha\beta} \to \mathbb{R}$ by $A^i_j(x) := \partial_j u^i_{\alpha\beta}(u_\beta(x))$ and $B^r_s(x) := \partial_s u^r_{\beta\alpha}(u_\alpha(x))$. By construction, the matrices $(A^i_j(x))$ and $(B^r_s(x))$ are inverse to each other, which reads as $A^i_r B^r_j = \delta^i_j$ in index notation. In our current notation, formula (1.2) reads as $\frac{\partial}{\partial u^i_\beta} = A^j_i \frac{\partial}{\partial u^j_\alpha}$. Applying du^r_α to this equation, we get $du^r_\alpha(\frac{\partial}{\partial u^i_\beta}) = A^r_i$, which implies $du^r_\alpha = A^r_i du^i_\beta$ and hence $du^i_\beta = B^i_j du^j_\alpha$. Now let us denote by t's the functions corresponding to (U_α, u_α) and by $\tilde{t}$'s the ones corresponding to (U_β, u_β). Using formula (3.2) to determine the functions $\tilde{t}^{j_1\dots j_\ell}_{i_1\dots i_k}$ and expanding using multilinearity, we get

$$\tilde{t}^{j_1\dots j_\ell}_{i_1\dots i_k} = A^{r_1}_{i_1} \dots A^{r_k}_{i_k} B^{j_1}_{s_1} \dots B^{j_\ell}_{s_\ell} t^{s_1\dots s_\ell}_{r_1\dots r_k}. \tag{3.5}$$

(Again, this involves a summation of $k+\ell$ indices.)

Even if this formula is much more complicated than the ones for vector fields and one-forms, it allows for an alternative interpretation which is completely parallel to what we did in these cases: Any chart (U, u) for M with $x \in U$ gives rise to the identification

$$\otimes^k T^*_x u \otimes \otimes^\ell T_x u : \otimes^k T^*_x M \otimes \otimes^\ell T_x M \to \otimes^k \mathbb{R}^{n*} \otimes \otimes^\ell \mathbb{R}^n.$$

For two compatible charts (U_α, u_α) and (U_β, u_β), the change of identifications is given by $\otimes^k (T^*_x u_\alpha \circ (T^*_x u_\beta)^{-1}) \otimes \otimes^\ell (T_x u_\alpha \circ (T_x u_\beta)^{-1})$, so this again depends smoothly on x. Hence we can follow the same strategy as in 1.15 and 3.1 to collect together all the spaces $\otimes^k T^*_x M \otimes \otimes^\ell T_x M$ into a manifold $\otimes^k T^* M \otimes \otimes^\ell TM$ that is endowed with a canonical smooth map $p : \otimes^k T^* M \otimes \otimes^\ell TM \to M$. For a map $t : M \to \otimes^k T^* M \otimes \otimes^\ell TM$ such that $p \circ t = \mathrm{id}$, the local coordinate expressions have as components the functions $t^{j_1\dots j_\ell}_{i_1\dots i_k}$, which shows that $\binom{\ell}{k}$-tensor fields on M are exactly the smooth functions with that property.

Example 3.2. (1) By definition, a $\binom{0}{1}$-tensor field on M is a one-form on M and from the end of Section 3.1, we conclude that a $\binom{1}{0}$-tensor field is a vector field on M.

(2) A $\binom{0}{2}$-tensor field g by definition associates to each point $x \in M$ a bilinear form $g_x : T_x M \times T_x M \to \mathbb{R}$ which is smooth in the sense that $x \mapsto g_x(\xi(x), \eta(x))$ is smooth for any $\xi, \eta \in \mathfrak{X}(M)$. If one in addition requires that each g_x is symmetric and positive definite and thus defines an inner product on $T_x M$, g is called a *Riemannian metric* on M. If one only requires g_x to be symmetric and non-degenerate (i.e. if $g_x(X, Y) = 0$ for some $X \in T_x M$ and all $Y \in T_x M$, then $X = 0$), then g is called a *pseudo-Riemannian metric* on M.

An element of $T^*_x M \otimes T^*_x M$ can also be viewed as defining a linear map $T_x M \to T^*_x M$ via $X \mapsto g_x(X, \_) : T_x M \to \mathbb{R}$. Thus given $g \in \mathcal{T}^0_2(M)$ and $\xi \in \mathfrak{X}(M)$, we can consider $x \mapsto g_x(\xi(x), \_) \in T^*_x M$, and this defines a one-form $g(\xi, \_) \in \Omega^1(M)$. If we deal with a pseudo-Riemannian metric, then non-degeneracy says that the induced map $T_x M \to T^*_x M$ has trivial kernel, so it must be a linear isomorphism in each point. Hence we can consider the point-wise inverses in $L(T^*_x M, T_x M) \cong T_x M \otimes T_x M$. Since inversion is a smooth map on invertible matrices, we conclude that these fit together to define a $\binom{2}{0}$-tensor field g^{-1} on M. In particular, to $\omega \in \Omega^1(M)$, we can associate a vector field $g^{-1}(\omega) \in \mathfrak{X}(M)$, and this defines an inverse to the map $\xi \mapsto g(\xi, \_)$. On the other hand, linear algebra shows that, viewed as a bilinear form on T_x^*M, g_x^{-1} is symmetric for each $x \in M$, and if g_x is positive definite, so is g_x^{-1}. Hence a Riemannian metric on M also defines inner products on the cotangent spaces of M.

(3) Similarly as in (2), there are different interpretations of $\binom{1}{1}$-tensor fields, which also follow just from linear algebra and observing the things depend smoothly on a point. Apart from the interpretation as bilinear maps $T_xM \times T_x^*M \to \mathbb{R}$, we can also identify $T_x^*M \otimes T_xM$ with $L(T_xM, T_xM)$, or with $L(T_x^*M, T_x^*M)$. (The resulting identification between $L(T_xM, T_xM)$ and $L(T_x^*M, T_x^*M)$ simply sends each map to the dual map.) Thus given a $\binom{1}{1}$-tensor field Φ, $\xi \in \mathfrak{X}(M)$ and $\omega \in \Omega^1(M)$, we can form $\Phi(\xi, \omega) \in C^\infty(M, \mathbb{R})$, $\Phi(\xi) \in \mathfrak{X}(M)$ and $\Phi(\omega) \in \Omega^1(M)$. This of course is a certain abuse of notation, that is justified by the fact that $\Phi(\xi, \omega) = \omega(\Phi(\xi)) = \Phi(\omega)(\xi)$. In local coordinates, the operations are given by

$$\Phi(\xi, \omega)|_U = \Phi_i^j \xi^i \omega_j \qquad (\Phi(\xi))^i = \Phi_j^i \xi^j \qquad (\Phi(\omega))_j = \Phi_j^i \omega_i.$$

3.3. Tensor products and contractions. These are two basic operations on tensor fields, which are defined point-wise, so understanding them again is mainly a matter of multilinear algebra. To formulate things, it will be helpful to first derive a description of tensor fields that will also be useful later on. Given $t \in \mathcal{T}_k^\ell(M)$, we can insert k vector fields and ℓ one-forms into t to obtain a smooth function. Thus t defines an operator $\mathfrak{X}(M)^k \times \Omega^1(M)^\ell \to C^\infty(M, \mathbb{R})$ and by construction, this is linear (over $\mathbb{R}$) in each entry. Now we can characterize among all these operators the ones induced by tensor fields. This will allows us to construct tensor fields in a coordinate-free way, by prescribing their actions on vector fields and one-forms. To formulate the characterization, we observe that vector fields, one-forms and smooth functions can all be multiplied (point-wise) by smooth functions, and thus for an operator Φ as above the concept of linearity over smooth functions in any variable makes sense.

LEMMA 3.3. *Let $\Phi : \mathfrak{X}(M)^k \times \Omega^1(M)^\ell \to C^\infty(M, \mathbb{R})$ be $k + \ell$-linear operator. Then Φ is induced by a tensor field $t \in \mathcal{T}_k^\ell(M)$ if and only if it is linear over $C^\infty(M, \mathbb{R})$ in each variable.*

PROOF. For $t \in \mathcal{T}_k^\ell(M)$, $\xi_1, \dots, \xi_k \in \mathfrak{X}(M)$ and $\omega^1, \dots, \omega^\ell \in \Omega^1(M)$, we by definition have

$$t(\xi_1, \dots, \xi_k, \omega^1, \dots, \omega^\ell)(x) = t_x(\xi_1(x), \dots, \xi_k(x), \omega^1(x), \dots, \omega^\ell(x))$$

for each $x \in M$. Replacing ξ_i by $f\xi_i$ for $f \in C^\infty(M, \mathbb{R})$, we get $(f\xi_i)(x) = f(x)\xi_i(x)$ and inserting this into t_x, we can take $f(x)$ out by multilinearity of t_x. So we see that $t(\dots, f\xi_i, \dots) = ft(\dots, \xi_i, \dots)$, and for the one-forms things work in the same way.

To prove the converse, we assume that Φ is linear over $C^\infty(M, \mathbb{R})$ in each variable, and claim that $\Phi(\xi_1, \dots, \omega^\ell)(x)$ depends only on $\xi_1(x), \dots, \xi_k(x) \in T_xM$ and on $\omega^1(x), \dots, \omega^\ell(x) \in T_x^*M$. Having proved this, we can use Φ to define a tensor field t as follows. Given $x \in M$, $X_1, \dots, X_k \in T_xM$ and $\lambda^1, \dots, \lambda^\ell \in T_x^*M$ we choose vector fields $\xi_i \in \mathfrak{X}(M)$ and one-forms $\omega^j \in \Omega^1(M)$ such that $\xi_i(x) = X_i$ and $\omega^j(x) = \lambda^j$ for all i and j. (Choosing a chart around x this is no problem locally, and then we get global extensions using a bump function.) Then we put

$$t_x(X_1, \dots, X_k, \lambda^1, \dots, \lambda^\ell) := \Phi(\xi_1, \dots, \xi_k, \omega^1, \dots, \omega^\ell)(x)$$

and observe that this is independent of all choices. Inserting vector fields and one-forms into t, we then recover Φ, so we see that t is smooth and that it induces Φ.

So it remains to prove the claim, and the argument is the same for all variables, so we do it for ξ_i. We want to show that $\Phi(\dots, \xi_i, \dots)(x) = \Phi(\dots, \tilde{\xi}_i, \dots)(x)$ provided that $\xi_i(x) = \tilde{\xi}_i(x)$. By inserting $\xi_i - \tilde{\xi}_i$, it suffices to prove that $\xi_i(x) = 0$ implies that $\Phi(\dots,\xi_i,\dots)(x) = 0$. As a first step, we prove that $\Phi(\dots,\xi_i,\dots)(x) = 0$ if there is an open subset $U \subset M$ with $x \in U$ such that ξ_i vanishes identically on U. To do this, we take a bump function $f \in C^\infty(M,\mathbb{R})$ with $\operatorname{supp}(f) \subset U$ and $f(x) = 1$. Then $f\xi_i$ is identically 0, so by multilinearity, we get $0 = \Phi(\dots, f\xi_i,\dots) = f\Phi(\dots,\xi_i,\dots)$. Evaluating the right hand side in x, we get $1 \cdot \Phi(\dots,\xi_i,\dots)(x)$, which completes this step. Assuming now that $\xi_i, \tilde{\xi}_i \in \mathfrak{X}(M)$ are such that $\xi_i|_U = \tilde{\xi}_i|_U$ then their difference vanishes on U and inserting into Φ the result vanishes in x, so $\Phi(\dots,\xi_i,\dots)(x) = \Phi(\dots,\tilde{\xi}_i,\dots)(x)$.

For the last step, we now assume that $\xi_i(x) = 0$, choose a chart (U,u) around x and take the expansion $\xi_i|_U = \xi_i^j \frac{\partial}{\partial u^j}$. Now we can extend the functions ξ_i^j to functions $f^j \in C^\infty(M,\mathbb{R})$ and the coordinate vector fields to $\eta_j \in \mathfrak{X}(M)$ without changing them on a smaller open neighborhood V of x. Thus $\xi_i|_V = (\sum_j f^j\eta_j)|_V$, and from the last step, we conclude

$$\Phi(\dots,\xi_i,\dots)(x) = \Phi(\dots,\textstyle\sum_j f^j\eta_j,\dots)(x) = \textstyle\sum_j f^j(x)\Phi(\dots,\eta_j,\dots)(x).$$

But $f^j(x) = \xi_i^j(x) = 0$ for all j, so the claim follows. □

Using this result, we can now express the operation on tensor fields obtained from the point-wise tensor product directly in terms of operators on vector fields and one-forms. Given $t \in \mathcal{T}^\ell_k(M)$ and $t' \in \mathcal{T}^{\ell'}_{k'}(M)$, we define an operator on $\mathfrak{X}(M)^{k+k'} \times \Omega^1(M)^{\ell+\ell'}$ by sending $\xi_1,\dots\xi_{k+k'}$ and $\omega^1,\dots,\omega^{\ell+\ell'}$ to

$$t(\xi_1,\dots,\xi_k,\omega^1,\dots,\omega^\ell) \cdot t'(\xi_{k+1},\dots,\xi_{k+k'},\omega^{\ell+1},\dots,\omega^{\ell+\ell'}).$$

This is smooth as a product of smooth functions, and obviously the expression is linear over $C^\infty(M,\mathbb{R})$ in each variable, so by Lemma 3.3 it defines $t \otimes t' \in \mathcal{T}^{\ell+\ell'}_{k+k'}(M)$. It is no problem to iterate tensor products and this is associative. Thus there is no need to put brackets in an iterated tensor product. Notice however, that the tensor product is not commutative in general. Obeserve also that, as claimed in Section 3.2, the "coordinate tensor fields" defined in formula (3.3) indeed are iterated tensor products of coordinate vector fields and coordinate one-forms.

The basis for the second operation is that for a vector space E with dual E^*, there is a natural bilinear map $E^* \times E \to \mathbb{R}$ which sends (λ, v) to $\lambda(v)$. This induces a linear map $E^* \otimes E \to \mathbb{R}$ and the construction easily extends to larger tensor products. In the setting of a point $x \in M$, choosing integers $k,\ell > 0$ and $1 \le r \le k$ and $1 \le s \le \ell$, one obtains a *contraction*

$$C^s_r : \otimes^k T^*_xM \otimes \otimes^\ell T_xM \to \otimes^{k-1}T^*_xM \otimes \otimes^{\ell-1}T_xM.$$

On a tensor product of elements, this takes the sth factor in the T_xM part, inserts it into the rth factor of the T^*_xM part and multiplies the resulting number with the tensor product of the remaining elements (in the original order). This is a bit clumsy to write out explicitly. A reasonable way is to use the convention that putting a hat over a symbol means omission. For example, taking $\lambda^1,\dots,\lambda^k \in T^*_xM$, we write $\lambda^1 \otimes \cdots \widehat{\lambda^r} \cdots \otimes \lambda^k$ for $\lambda^1 \otimes \cdots \otimes \lambda^{r-1} \otimes \lambda^{r+1} \otimes \cdots \otimes \lambda^k$. Then for these λ^j and $X_1,\dots,X_\ell \in T_xM$, we get

$$C^s_r(\lambda^1 \otimes \cdots \otimes \lambda^k \otimes X_1 \otimes \cdots \otimes X_\ell) = \lambda^r(X_s)\lambda^1 \otimes \cdots \widehat{\lambda^r} \cdots \otimes \lambda^k \otimes X_1 \otimes \cdots \widehat{X_s} \cdots \otimes X_\ell.$$

Using this description, one easily verifies that applying this in each point defines a smooth map $\otimes^k T^*M \otimes \otimes^\ell TM \to \otimes^{k-1}T^*M \otimes \otimes^{\ell-1}TM$. Composing this with a $\binom{\ell}{k}$-tensor field t on M, we obtain a $\binom{\ell-1}{k-1}$-tensor field, that we denote by $C^s_r(t)$.

To compute this in coordinates, recall from linear algebra the description of the basic contraction $E^* \otimes E \to \mathbb{R}$ in the picture of bilinear maps $E \times E^* \to \mathbb{R}$. Taking a basis $\{e_i\}$ for E with dual basis $\{e^j\}$ for E^* (i.e. $e^j(e_i) = \delta_i^j$), a bilinear map $\varphi : E \times E^* \to \mathbb{R}$ can be written as $\sum_{i,j} \varphi(e_i, e^j) e^i \otimes e_j$. Thus its contraction is given by $\sum_i \varphi(e_i, e^i)$ and one easily verifies that this is independent of the choice of basis. In the setting of tensor fields and a chart (U, u) such dual bases are formed by the elements $\frac{\partial}{\partial u^i}(x)$ and $du^i(x)$ for each $x \in U$. This shows for the local representations as in (3.4) and using "a" as a summation index, we get

$$(C_r^s(t))^{j_1 \dots j_{\ell-1}}_{i_1 \dots i_{k-1}} = t^{j_1 \dots j_{s-1} a j_s \dots j_{\ell-1}}_{i_1 \dots i_{r-1} a i_r \dots i_{k-1}}$$

(which finally explains the notation). Observe that viewing $E^* \otimes E$ as either $L(E, E)$ or $L(E^*, E^*)$ the contraction just corresponds to taking the trace of an endomorphism.

Example 3.3. Several things we have done so far admit an interpretation in terms of tensor products and contractions, which often is technically very helpful. For example given $\xi \in \mathfrak{X}(M) = \mathcal{T}_0^1(M)$ and $\omega \in \Omega^1(M) = \mathcal{T}_1^0(M)$, we can form $\omega \otimes \xi =\in \mathcal{T}_1^1(M)$. For the (unique possible) contraction we then get $C_1^1(\omega \otimes \xi) = \omega(\xi)$. Note that in this case, there is no difference between $\omega \otimes \xi$ and $\xi \otimes \omega$.

Similarly, for pseudo-Riemannian metric $g \in \mathcal{T}_2^0(M)$ as discussed in Example (2) of 3.2 and $\xi \in \mathfrak{X}(M)$, we get $g \otimes \xi \in \mathcal{T}_2^1(M)$. Now there are two possible contractions C_1^1 and C_2^1 of this, but the fact that each g_x is symmetric implies that they coincide in this case and equal $g(\xi, \_)$. Also the insertion $g^{-1}(\omega)$ of a one-form ω into the inverse metric can be written as $C_1^1(g^{-1} \otimes \omega) = C_1^2(g^{-1} \otimes \omega)$ with $g^{-1} \otimes \omega \in \mathcal{T}_1^2(M)$. In these examples, there is also no difference between $g \otimes \xi$ and $\xi \otimes g$ or between $g^{-1} \otimes \omega$ and $\omega \otimes g^{-1}$.

In the setting of Example (3) of 3.2, we have $\Phi \in \mathcal{T}_1^1(M)$, $\xi \in \mathfrak{X}(M)$ and $\omega \in \Omega^1(M)$. In this case, there already is a contraction of Φ, namely $C_1^1(\Phi) \in C^\infty(M, \mathbb{R})$. Further, we can form $\Phi \otimes \xi \in \mathcal{T}_1^2(M)$, $\Phi \otimes \omega \in \mathcal{T}_2^1(M)$ and $\Phi \otimes \omega \otimes \xi \in \mathcal{T}_2^2(M)$. Now $C_1^2(\Phi \otimes \xi) = \Phi(\xi)$ and $C_2^1(\Phi \otimes \omega) = \Phi(\omega)$ in the notation from Example (3) of 3.3, while $C_1^1(\Phi \otimes \xi)$ is the product of the smooth function $C_1^1(\Phi)$ with ξ and similarly for $C_1^1(\Phi \otimes \omega)$. Finally, we can form iterated contractions of $\Phi \otimes \omega \otimes \xi$ to obtain $\Phi(\xi, \omega)$ as either $C_1^1(C_1^2(\Phi \otimes \omega \otimes \xi))$ or $C_1^1(C_2^1(\Phi \otimes \omega \otimes \xi))$ (but here the notation we introduced is not optimal any more). Here $\Phi \otimes \omega \otimes \xi = \Phi \otimes \xi \otimes \omega$, but the the position of Φ with respect to either of the other two factors does matter (as seen from the different contractions).

As a generalization of the last observation, observe that for $t \in \mathcal{T}_k^\ell(M)$, $\xi_1, \dots, \xi_k \in \mathfrak{X}(M)$ and $\omega^1, \dots, \omega^\ell \in \Omega^1(M)$, we can always obtain $t(\xi_1, \dots, \xi_k, \omega^1, \dots, \omega^\ell)$ via a sequence of contractions from $t \otimes \omega^1 \otimes \dots \otimes \omega^\ell \otimes \xi_1 \otimes \dots \otimes \xi_k$.

3.4. Excursion: Functoriality and Lie derivatives. We discuss these topics for general tensor fields only briefly, the case of differential forms will be discussed in more detail below. Let $F : M \to N$ be a local diffeomorphism between smooth manifolds. Then for each $x \in M$, the tangent map $T_x F : T_x M \to T_{F(x)} N$ is a linear isomorphism and we can also form its dual map $(T_x F)^* : T^*_{F(x)} N \to T^*_x M$. Combining this with $(T_x F)^{-1}$ we get, for any $k, \ell \in \mathbb{N}$, linear isomorphisms

$$\otimes^k (T_x F)^* \otimes \otimes^\ell (T_x F)^{-1} : \otimes^k T^*_{F(x)} N \otimes \otimes^\ell T_{F(x)} N \to \otimes^k T^*_x M \otimes \otimes^\ell T_x M. \tag{3.6}$$

Given a tensor field $t \in \mathcal{T}_k^\ell(N)$, we can apply this map to $t(F(x))$ to obtain an element $(F^* t)(x) \in \otimes^k T^*_x M \otimes \otimes^\ell T_x M$. Now for each $x \in M$, there is an open neighborhood U of x in M such that $F(U) =: V \subset N$ is open and $F|_U : U \to V$ is a diffeomorphism. It then follows easily that the maps from (3.6) fit together over V to define a smooth map

$p^{-1}(V) \to \otimes^k T^*M \otimes \otimes^\ell TM$, which in turn implies that F^*t is smooth on a neighborhood of x, and hence $F^*t \in \mathcal{T}^\ell_k(M)$. This is called the *pullback of t along F*. Alternatively, we can use the map $T^*_xF = ((T_xF)^{-1})^* : T^*_xM \to T^*_{F(x)}N$ from Section 3.1 to express the definition as

(3.7)
$$(F^*t)_x(X_1,\dots,X_k,\lambda^1,\dots,\lambda^\ell) = t_{F(x)}(T_xF(X_1),\dots,T_xF(X_k),T^*_xF(\lambda^1),\dots,T^*_xF(\lambda^\ell))$$

for $X_1,\dots,X_k \in T_xM$ and $\lambda^1,\dots,\lambda^\ell \in T^*_xM$. In particular, this shows that for vector fields and one-forms, we recover the operations introduced in Sections 2.3 and 3.1, respectively. This point-wise formula also show that the pullback is linear, and compatible with multiplication by smooth functions, i.e. $F^*(t_1+t_2) = F^*t_1 + F^*t_2$ and $F^*(ft) = (f\circ F)(F^*t)$ for $f \in C^\infty(N,\mathbb{R})$.

Using this interpretation and the point-wise definition of the operations from Section 3.3, we immediately conclude that pullbacks are compatible with these operations: For tensor fields $t \in \mathcal{T}^\ell_k(N)$ and $t' \in \mathcal{T}^{\ell'}_{k'}(N)$, we obtain $F^*(t\otimes t') = (F^*t)\otimes(F^*t')$. Similarly, in case that $k,\ell > 0$, for chosen integers $1 \le r \le k$ and $1 \le s \le \ell$, we obtain $F^*(C^s_r(t)) = C^s_r(F^*t)$. For example, for $\xi \in \mathfrak{X}(N)$ and $\omega \in \Omega^1(N)$, we know that $\omega(\xi) = C^1_1(\omega\otimes\xi)$. Now of course $F^*(\omega(\xi)) = \omega(\xi)\circ F$ and we have just seen that this coincides with $C^1_1(F^*(\omega\otimes\xi)) = C^1_1((F^*\omega)\otimes(F^*\xi)) = (F^*\omega)(F^*\xi)$.

Having this at hand, we can now follow the ideas from Section 2.10 to define a Lie derivative for general tensor fields. This will be easier than the developments there since we can build on the case of vector fields and work with explicit formulae. Let us start with a fixed vector field $\eta \in \mathfrak{X}(M)$ and consider its flow Fl^η_s which, locally around each $x \in M$, is defined for $s \in \mathbb{R}$ sufficiently close to 0. Given a tensor field $t \in \mathcal{T}^\ell_k(M)$ we have $c(s) := ((\mathrm{Fl}^\eta_s)^*t)(x) \in \otimes^k T^*_xM \otimes \otimes^\ell T_xM$ which is defined for $s \in \mathbb{R}$ sufficiently close to 0. To verify that this is a smooth curve, we have to treat the case of one-forms separately: For $\omega \in \Omega^1(M)$ and $X \in T_xM$, we by definition have $((\mathrm{Fl}^\eta_s)^*\omega)(x)(X) = \omega(\mathrm{Fl}^\eta_s(x))(T_x\mathrm{Fl}^\eta_s(X))$, and this visibly depends smoothly on s. We also know from 2.10 that for each vector field $\xi \in \mathfrak{X}(M)$, $s \mapsto ((\mathrm{Fl}^\eta_s)^*\xi)(x)$ is a smooth curve in T_xM.

Now taking a chart (U,u) for M with $x \in U$, we know from Section 3.2 that we can write
$$t|_U = t^{j_1\dots j_\ell}_{i_1\dots i_k} du^{i_1}\otimes\dots\otimes du^{i_k}\otimes\tfrac{\partial}{\partial u^{j_1}}\otimes\dots\otimes\tfrac{\partial}{\partial u^{j_1}}.$$
Pulling this back along Fl^η_s and evaluating in x, we get a sum of terms which are products of the smooth functions $t^{j_1\dots j_\ell}_{i_1\dots i_k}(\mathrm{Fl}^\eta_s(x))$ with the tensor product of the pullbacks of the individual one-forms du^{i_r} and pullbacks of the individual vector fields $\frac{\partial}{\partial u^{j_s}}$ both evaluated at x. The tensor products of the resulting smooth curves of course define a smooth curve in $\otimes^k T^*_xM\otimes\otimes^\ell T_xM$ and a linear combination of smooth curves is smooth. Hence we see that the curve $c(s) := ((\mathrm{Fl}^\eta_s)^*t)(x)$ is smooth and we can define

$$\mathcal{L}_\eta t(x) := \tfrac{d}{ds}|_{s=0}((\mathrm{Fl}^\eta_s)^*t)(x) \in \otimes^k T^*_xM\otimes\otimes^\ell T_xM. \tag{3.8}$$

To see that $\mathcal{L}_\eta t$ is again a smooth tensor field, we describe its action on vector fields and one-forms explicitly, and again we have to treat the case of one-forms separately. For $\omega \in \Omega^1(M)$, $\xi \in \mathfrak{X}(M)$, and $x \in M$, we get from above that
$$\omega(\xi)(\mathrm{Fl}^\eta_s(x)) = C(((\mathrm{Fl}^\eta_s)^*\omega)(x)\otimes((\mathrm{Fl}^\eta_s)^*\xi)(x)),$$
where $C : T^*_xM\otimes T_xM \to \mathbb{R}$ is the basic contraction. Differentiating the left hand side at $s=0$, we of course get the directional derivative $\eta(x)(\omega(\xi))$. For the right hand side,

we can differentiate through the linear map C and use that bilinearity of the tensor product implies

$$\tfrac{d}{ds}|_{s=0}(c_1(s)\otimes c_2(s)) = c_1'(0)\otimes c_2(0) + c_1(0)\otimes c_2'(0).$$

Using that $\frac{d}{ds}|_{s=0}((\mathrm{Fl}^\eta_s)^*\xi)(x) = [\eta,\xi](x)$, we conclude that the right hand side equals $\mathcal{L}_\eta\omega(x)(\xi(x))+\omega(x)([\eta,\xi](x))$. Using that this holds in each point x, we can now simply rearrange terms to obtain

$$(\mathcal{L}_\eta\omega)(\xi) = \eta(\omega(\xi)) - \omega([\eta,\xi]). \tag{3.9}$$

Since the right hand side is evidently smooth, this shows that $\mathcal{L}_\eta\omega$ indeed defines a smooth one-form on M.

If you feel uneasy with this computation, there also is an alternative approach, namely using equation (3.9) to *define* a one-form $\mathcal{L}_\eta\omega$. In view of Proposition 3.3 we only have to verify that the right hand side of (3.9) is linear over smooth function in the variable ξ. But this immediately follows from the derivation property of η and part (3) of Theorem 2.4. However, for many applications, the relation of this operation to the flow of η is crucial. Finally, let us observe that equation(3.9) easily implies that $\mathcal{L}_\eta(f\omega) = \eta(f)\omega + f\mathcal{L}_\eta\omega$.

This line of argument extends to the case of $t\in\mathcal{T}^\ell_k(M)$ using the last observation in Section 3.3, namely that $t(\xi_1,\dots,\xi_k,\omega^1,\dots,\omega^\ell)$ can be obtained via a sequence of contractions from $t\otimes\xi_1\otimes\dots\otimes\omega^\ell$. From Section 3.3 we know that both the tensor product and the sequence of contractions commute with the pullback along local flows of η. Hence we can write $t(\xi_1,\dots,\xi_k,\omega^1,\dots,\omega^\ell)(\mathrm{Fl}^\eta_s(x))$ as a linear map (the sequence of contractions) applied to the tensor product of $((\mathrm{Fl}^\eta_s)^*t)(x)$, and factors of the form $((\mathrm{Fl}^\eta_s)^*\xi_i)(x)$ and $((\mathrm{Fl}^\eta_s)^*\omega^j)(x)$. Differentiating with respect to s at $s=0$, we again get $\eta(x)(t(\xi_1,\dots,\xi_k,\omega^1,\dots,\omega^\ell))$ in the left hand side. In the right hand side, we can differentiate through the linear map and get a sum of terms in which exactly one of the curves in the tensor product is differentiated. Now in the first summand, $\mathcal{L}_\eta t(x)$ occurs together with the vector fields and one-forms. For the vector fields, differentiation just leads to $[\eta,\xi_i](x)$ while for the one forms we get $\mathcal{L}_\eta\omega^j(x)$ as computed above. Rearranging terms, we obtain a formula at the point x and the we can leave out x to conclude that $(\mathcal{L}_\eta t)(\xi_1,\dots,\xi_k,\omega^1,\dots,\omega^\ell)$ can be written as

$$\begin{aligned}\eta(t(\xi_1,\dots,\omega^\ell)) &- \textstyle\sum_{i=1}^k t(\xi_1,\dots,[\eta,\xi_i],\dots,\xi_k,\omega^1,\dots,\omega^\ell)\\ &- \textstyle\sum_{j=1}^\ell t(\xi_1,\dots,\xi_k,\omega^1,\dots,\mathcal{L}_\eta\omega^j,\dots,\omega^\ell).\end{aligned} \tag{3.10}$$

Since the right hand side is smooth, this shows that $\mathcal{L}_\eta t\in\mathcal{T}^\ell_k(M)$. The alternative approach described in the case of one-forms can also be applied here, namely using formula (3.10) to define $\mathcal{L}_\eta t$. By Lemma 3.3, we only have to shows that formula (3.10) is linear over smooth functions in any of the ξ_i and any of the ω^j. But this immediately follows from the fact that η acts as a derivation, the formula for $[\eta,f\xi_i]$ from Theorem 2.4 and the formula for $\mathcal{L}_\eta(f\omega^j)$ that we derived above. However, for many applications the relation to the pullback along flows of η is crucial.

Differential forms

Although they are probably more difficult to understand intuitively than vector fields, differential forms are the most versatile geometric objects available on manifolds. They can be pulled back along arbitrary smooth maps and carry a rich algebraic structure. Moreover, there are several natural operations available in this setting, which combine to an efficient calculus. As we shall see later on, they are also related to integration on manifolds and to algebraic topology.

3.5. Definition, pull back and wedge product. A *differential form of degree k* or a k-form on a manifold M is a $\binom{0}{k}$-tensor field $\omega \in \mathcal{T}^0_k(M)$ such the for each $x \in M$ the k-linear map $\omega_x : (T_xM)^k \to \mathbb{R}$ is alternating. Recall the latter condition can either be phrased as the fact that $\omega_x(X_1, \dots, X_k) = 0$ whenever $X_i = X_j$ for some $i \neq j$. Equivalently, $\omega_x(X_{\sigma_1}, \dots, X_{\sigma_k}) = \operatorname{sgn}(\sigma)\omega_x(X_1, \dots, X_k)$ for any permutation $\sigma \in \mathfrak{S}_k$ of k letters and all X_i. The space of all k-forms on M will be denoted by $\Omega^k(M)$. Observe that for $k = 1$ the condition of being alternating is vacuous, so we recover notation from Section 3.1. We will also extend the definition to $k = 0$ by putting $\Omega^0(M) := C^\infty(M, \mathbb{R})$. By construction, $\Omega^k(M)$ is a linear subspace of $\mathcal{T}^0_k(M)$ and closed under multiplication by smooth functions.

From linear algebra, we know that any k-linear map is uniquely determined by its values on (k-tuples of) elements of a basis, and if the map is alternating, all these basis elements have to be different in order to lead to a non-zero result. Applying this to ω_x, we conclude that $\Omega^k(M) = 0$ for $k > n = \dim(M)$. We then define $\Omega^*(M) := \oplus_{k=0}^n \Omega^k(M)$. Finally, we observe that Lemma 3.3 shows that we can identify $\Omega^k(M)$ with the space of those k-linear maps $\mathfrak{X}(M)^k \to C^\infty(M, \mathbb{R})$ which are alternating and linear over $C^\infty(M, \mathbb{R})$ in each variable. (Of course, for an alternating map, it suffices to verify linearity over $C^\infty(M, \mathbb{R})$ in one variable.)

In Section 3.1, we have seen that one-forms can be pulled back along arbitrary smooth maps, and we want to generalize this to $\binom{0}{k}$-tensor fields and hence to differential forms. To prepare this, take a manifold M and associate to each $x \in M$ a k-linear map $t_x : (T_xM)^k \to \mathbb{R}$. Then we can combine these to a map $\hat{t} : \otimes^k TM \to \mathbb{R}$ and similar to the proof of Lemma 3.1, one immediately verifies that $\hat{t}$ is smooth if and only if t defines a $\binom{0}{k}$-tensor field on M. Now suppose we have given a smooth map $F : M \to N$ and $t \in \mathcal{T}^0_k(N)$. Then one simply puts

$$(F^*t)(x)(X_1, \dots, X_k) := t(F(x))(T_xF(X_1), \dots, T_xF(X_k)) \tag{3.11}$$

for $X_1, \dots, X_k \in T_xM$. Now of course the map $TF : TM \to TN$ gives rise to a smooth map $\otimes^k TF : \otimes^k TM \to \otimes^k TN$ and by construction $\widehat{F^*t} = \hat{t} \circ \otimes^k TF$. Thus we conclude that $F^*t \in \mathcal{T}^0_k(M)$. Of course, if $t(y)$ is alternating for each $y \in N$, then $(F^*t)(x)$ is alternating for each $x \in M$, so for $\omega \in \Omega^k(N)$, we get $F^*\omega \in \Omega^k(M)$.

The basic algebraic structure on differential forms is a point-wise issue of multilinear algebra, but we discuss it in the picture relevant for us right away. Recall that a k-linear map $t_x : (T_xM)^k \to \mathbb{R}$ can be alternated by defining

$$\operatorname{Alt}(t_x)(X_1, \dots, X_k) := \textstyle\sum_{\sigma \in \mathfrak{S}_k} \operatorname{sgn}(\sigma) t_x(X_{\sigma_1}, \dots, X_{\sigma_k}). \tag{3.12}$$

If t_x itself is alternating, then $\operatorname{Alt}(t_x) = k!t_x$. Evidently, we can apply the alternation point–wise to obtain $\operatorname{Alt}(t) \in \Omega^k(M)$ from $t \in \mathcal{T}^0_k(M)$. In particular, given one-forms $\varphi^1, \dots, \varphi^k \in \Omega^1(M)$ we define

$$\varphi^1 \wedge \dots \wedge \varphi^k := \operatorname{Alt}(\varphi^1 \otimes \dots \otimes \varphi^k) \in \Omega^k(M). \tag{3.13}$$

By construction the action of this on tangent vectors $X_1, \dots, X_k \in T_xM$ can be written as the determinant of the matrix $(\varphi^i(x)(X_j))_{i,j=1,\dots,k}$. We shall see later on that $\varphi^1 \wedge \dots \wedge \varphi^k$ actually is the iterated wedge product (as defined below) of the one-forms φ^j. However, at this point, the expression should just be viewed as a notation for a specific k-form.

This immediately allows us to deduce the description of k-forms in local coordinates. For a chart (U,u) for M with $x\in U$ and $\omega\in\Omega^k(M)\subset T^0_k(M)$, we know that

$$\omega|_U=\omega_{i_1\dots i_k}du^{i_1}\otimes\dots\otimes du^{i_k},$$

with a sum over all k-tuples of indices in the right hand side and where the function $\omega_{i_1\dots i_k}$ is obtained by inserting appropriate coordinate vector fields into ω. In particular, this implies that $\omega_{i_1\dots i_k}=0$ if two of the indices agree. Alternating both sides of the equation, the left hand side gives $k!\omega|_U$, while in the right hand side the smooth function $\omega_{i_1\dots i_k}$ can be taken out of the alternation. Hence we conclude that

$$\omega|_U=\tfrac{1}{k!}\omega_{i_1\dots i_k}du^{i_1}\wedge\dots\wedge du^{i_k}.$$

But now permuting the indices either in the functions $\omega_{i_1\dots i_k}$ or in the coordinate one-forms, one picks up the sign of the permutation. Hence the sum over all possible permutations of these k values of the indices just gives $k!$ times the single expression for the case in which the indices occur in (strictly) increasing order. This leads to an alternative presentation of a k-form in local coordinates as

$$\omega|_U=\sum_{1\le i_1<\dots<i_k\le n}\omega_{i_1\dots i_k}du^{i_1}\wedge\dots\wedge du^{i_k} \tag{3.14}$$

for the same functions $\omega_{i_1\dots i_k}:U\to\mathbb{R}$ as above.

Similarly to the case of one-forms in (3.13), we can try to define a product of two differential forms by alternating their tensor product. However, it is necessary to carefully introduce numerical factors in order to ensure associativity. It turns out that the right choice for $\omega\in\Omega^k(M)$ and $\tau\in\Omega^\ell(M)$ is to define the *wedge-product* as

$$\omega\wedge\tau:=\tfrac{1}{k!\ell!}\operatorname{Alt}(\omega\otimes\tau)\in\Omega^{k+\ell}(M). \tag{3.15}$$

As a first indication that this is a good choice observe that for $k=0$, we get $\omega=f\in C^\infty(M,\mathbb{R})$ and $\omega\otimes\tau=f\tau$. Now of course $\operatorname{Alt}(f\tau)=f\operatorname{Alt}(\tau)=\ell!f\tau$, and thus $f\wedge\tau=f\tau$ as one would hope. Now we can prove that the wedge product has good properties in general.

THEOREM 3.5. *Let M be a smooth manifold. Then the wedge-product on $\Omega^*(M)$ is bilinear over $\mathbb{R}$ and over $C^\infty(M,\mathbb{R})$. It is associative and graded commutative in the sense that for $\omega\in\Omega^k(M)$ and $\tau\in\Omega^\ell(M)$ one gets $\tau\wedge\omega=(-1)^{k\ell}\omega\wedge\tau$. Moreover, the wedge product is compatible with pull backs, i.e. $F^*(\omega\wedge\tau)=(F^*\omega)\wedge(F^*\tau)$ for any smooth map F.*

PROOF. Since the tensor product is bilinear and the alternation is linear, the wedge product is bilinear over $\mathbb{R}$. The point-wise definition of the wedge product then readily implies bilinearity over $C^\infty(M,\mathbb{R})$ as well as the compatibility with pullbacks. Making the definition explicit, we see that

$$(\omega\wedge\tau)(x)(X_1,\dots,X_{k+\ell})=\tfrac{1}{k!\ell!}\textstyle\sum_{\sigma\in\mathfrak{S}_{k+\ell}}\operatorname{sgn}(\sigma)\omega_x(X_{\sigma_1},\dots,X_{\sigma_k})\tau_x(X_{\sigma_{k+1}},\dots,X_{\sigma_{k+\ell}})$$

for $X_1,\dots,X_{k+\ell}\in T_xM$. Looking at the right hand side, we conclude that if we change the order of ω and τ, this only changes the "right" sign that should be associated to a permutation, and it always changes by $(-1)^{k\ell}$, since we have to exchange each of the first k entries with each of the last ℓ entries. This proves the claim on graded commutativity.

Thus it remains to verify associativity, which is slightly subtle. The crucial step is to prove that for $\varphi^1,\dots,\varphi^k,\psi^1,\dots,\psi^\ell$ we get

$$(\varphi^1\wedge\dots\wedge\varphi^k)\wedge(\psi^1\wedge\dots\wedge\psi^\ell)=\varphi^1\wedge\dots\wedge\psi^\ell. \tag{3.16}$$

(The right hand side and the terms in brackets in the left hand side are defined in (3.13), while between the two bracket terms, we have the product defined in (3.15).) Indeed, having shown this, we readily conclude that $(\omega \wedge \tau) \wedge \rho = \omega \wedge (\tau \wedge \rho)$ in case that each of the forms is an iterated wedge product of one-forms as in (3.13). By bilinearity, this extends to forms which can be written as linear combinations of such iterated wedge products. Since the operation is point-wise, it suffices to be of that form locally, and we have noted above that this is always the case.

Thus we are left with proving (3.16), which is a point-wise statement. Hence let us take tangent vectors $X_1, \dots, X_{k+\ell} \in T_xM$. Inserting these into the alternation of $(\varphi^1_x \wedge \dots \wedge \varphi^k_x) \otimes (\psi^1_x \wedge \dots \wedge \psi^\ell_x)$, we obtain

$$\textstyle\sum_\sigma \operatorname{sgn}(\sigma)(\varphi^1_x \wedge \dots \wedge \varphi^k_x)(X_{\sigma_1}, \dots, X_{\sigma_k})(\psi^1_x \wedge \dots \wedge \psi^\ell_x)(X_{\sigma_{k+1}}, \dots, X_{\sigma_{k+\ell}}).$$

But to evaluate each of the factors in the product, we again have to alternate the entries and then insert into one of the one-forms. The result corresponds to one of the summands in the expansion of $(\varphi^1_x \wedge \dots \wedge \psi^\ell_x)(X_1, \dots, X_{k+\ell})$ for some permutation (built up from σ and the additional permutations of k respectively ℓ letters). But this also shows that each of the summands in this expansion occurs exactly $k!\ell!$ times in this process, which is exactly the factor divided by in (3.15). □

3.6. The exterior derivative. This is probably the most important operation in the whole field of analysis on manifolds. Following the spirit of this course, we will define the operator by a global formula, i.e. by describing the action on vector fields. While this formula looks rather involved, it directly shows that the operation is independent of the choice of local coordinates. One can also derive the properties of the operation from this global formula, but it is easier to first find an expression in local coordinates and study the operation via this. The expression in local coordinates is simpler than the global formula and better suited for explicit computations.

There is a relatively simple motivation for the global formula. We just outline this here, the details will be done in the exercises: Let us consider an open subset $U \subset \mathbb{R}^n$. For each $x \in U$, we can naturally identify T_xM with $\mathbb{R}^n$, so a k-form $\omega \in \Omega^k(U)$ can be simply viewed as a smooth map ω from U to the vector space of all k-linear, alternating maps $(\mathbb{R}^n)^k \to \mathbb{R}$. For the moment, we write L^k_a for that space. Now we can differentiate $\omega : U \to L^k_a$, so for $x \in U$, $D\omega(x)$ is a linear map $\mathbb{R}^n \to L^k_a$. But then for each $x \in U$, $(X_0, \dots, X_k) \mapsto (D\omega(x)(X_0))(X_1, \dots, X_k)$ is a $(k+1)$-linear map $(\mathbb{R}^n)^{k+1} \to \mathbb{R}$. Now we can make this alternating and then the result defines a smooth function $d\omega : U \to L^{k+1}_a$. Since the expression is already alternating in $X_1, \dots, X_k$, we don't have to use the full alternation as described in Section 3.5, it suffices to form $\sum_i (-1)^i (D\omega(x)(X_i))(X_0, \dots, \widehat{X_i}, \dots, X_k)$ with the hat indicating omission.

Having constructed this, one can compute how the map $d\omega$ acts on $k+1$ vector fields. Initially, this involves directional derivatives of vector fields viewed as functions with values in $\mathbb{R}^n$ in the direction of other vector fields. However with a bit of manipulation one shows (see exercises) that these terms can be combined into an expression which only uses the action of vector fields on smooth real valued functions and Lie brackets of vector fields. The resulting formula then makes sense on any manifold, and this is what we use to define the exterior derivative.

Let M be a smooth manifold, $\omega \in \Omega^k(M)$ a k-form on M and take $k+1$ vector fields $\xi_0, \dots, \xi_k \in \mathfrak{X}(M)$. Then we consider the following expression in which a hat over an argument indicates that the argument has to be omitted:

$$\begin{aligned}\sum_{i=0}^{k}(-1)^i\xi_i(\omega(\xi_0,\dots,\widehat{\xi_i},\dots,\xi_k))\\ +\sum_{i<j}(-1)^{i+j}\omega([\xi_i,\xi_j],\xi_0,\dots,\widehat{\xi_i},\dots,\widehat{\xi_j},\dots,\xi_k)\end{aligned}\tag{3.17}$$

In the first terms, we plug all but one of the vector fields into ω and then use the remaining vector field to differentiate the resulting smooth function. In the second terms we plug in the Lie bracket of two of the vector fields together with the remaining $k-1$ vector fields into ω. The signs showing up in the formula come from the alternation in the construction. Anyway, (3.17) evidently is a smooth function $M\to\mathbb{R}$.

Observe that for $\omega=f\in\Omega^0(M)=C^\infty(M,\mathbb{R})$ this specializes to $\xi_0\mapsto\xi_0(f)=df(\xi_0)$, so this recovers the one-form df from Section 3.1. For $\omega\in\Omega^1(M)$, we get

$$\xi_0(\omega(\xi_1))-\xi_1(\omega(\xi_0))-\omega([\xi_0,\xi_1]),$$

which still is rather simple. For higher degrees the formula gets more complicated, but our main use for it will be conceptual anyway.

Lemma 3.6. *Let M be a smooth manifold and let $\omega\in\Omega^k(M)$ be a k-form on M. Then there is a unique $k+1$-form $d\omega\in\Omega^{k+1}(M)$ such that for any choice of $k+1$ vector fields $\xi_0,\dots,\xi_k\in\mathfrak{X}(M)$, the smooth function $d\omega(\xi_0,\dots,\xi_k)$ is given by* (3.17).

Proof. We use Lemma 3.3 in the version for differential forms. So we have to show that the expression in (3.17) is alternating and linear over smooth functions in one variable. To prove that the expression is alternating it suffices to show that it vanishes if two neighboring entries are equal. (This implies that it changes sign if we exchange two neighboring entries and such changes generate all permutations.) So let us assume that for some fixed $i_0\in\{0,\dots,k-1\}$, we have $\xi_{i_0}=\xi_{i_0+1}=\xi$. Then in all terms of the first sum in (3.17) with $i\neq i_0,i_0+1$, both ξ_{i_0} and ξ_{i_0+1} are inserted into ω, so the corresponding summand vanishes identically. On the other hand the summands for $i=i_0$ and $i=i_0+1$ agree, apart form the fact that they get opposite signs. So we conclude that the first sum in (3.17) vanishes identically.

For the second sum in (3.17) we have to distinguish more cases. Let us first fix some $i<i_0$. Then the terms for $j=i_0$ and $j=i_0+1$ are identical apart from an opposite sign, while all other terms vanish because both ξ_{i_0} and ξ_{i_0+1} are inserted into ω. Hence the terms with $i<i_0$ do not contribute and for $i>i_0+1$, there is no contribution since again both ξ_{i_0} and ξ_{i_0+1} are inserted into ω. So it remains to discuss the summands with $i=i_0$ and $i=i_0+1$. The summand with $i=i_0$ and $j=i_0+1$ vanishes by skew symmetry of the Lie bracket. But for $j>i_0+1$ the summand for $i=i_0$ and j agrees with the one from $i=i_0+1$ and j apart from an opposite sign.

Knowing that the expression in (3.17) is alternating, it suffices to show that replacing ξ_0 by $f\xi_0$ in (3.17) for a smooth function $f\in C^\infty(M,\mathbb{R})$, the result simply gets multiplied by f. In the first sum, f simply comes out of the summand with $i=0$. For $i>0$, we can take f out of ω and then in addition to the terms, in which f simply comes out, we obtain

$$\sum_{i>0}(-1)^i\xi_i(f)\omega(\xi_0,\dots,\widehat{\xi_i},\dots,\xi_k).$$

For the second sum, f simply comes out of all summands with $i>0$. For $i=0$, we just use $[f\xi_0,\xi_j]=-[\xi_j,f\xi_0]=-\xi_j(f)\xi_0+f[\xi_0,\xi_j]$. For the second summand in the right hand side, the f then comes out of ω. For the first summand in the right hand side, we can take the smooth function $-\xi_j(f)$ out of ω and thus these terms produce

$$\sum_{j>0}-(-1)^{0+j}\xi_j(f)\omega(\xi_0,\dots,\widehat{\xi_j},\dots,\xi_k),$$

which exactly cancels with the contribution from above. □

Observe that this lemma in particular says that the value of $d\omega(\xi_0,\dots,\xi_k)$ in a point $x\in M$ depends only on the values $\xi_0(x),\dots,\xi_k(x)\in T_xM$. In particular, it is possible to apply $d\omega$ to tangent vectors, although so far we have to extend them to vector fields in order to compute the result.

Definition 3.6. For a k-form ω on a smooth manifold M, the $k+1$-form $d\omega$ obtained in Lemma 3.6 is called the *exterior derivative of* ω.

At this point, we can view the map $\omega\mapsto d\omega$ as defining an operator $d:\Omega^k(M)\to\Omega^{k+1}(M)$, the *exterior derivative* for any smooth manifold M and any integer k. From the definition equation (3.17) it is obvious that d is linear over $\mathbb{R}$ (but not over $C^\infty(M,\mathbb{R})$ as we shall see). We can now prove the important properties of this operation as well as its description in local coordinates.

Theorem 3.6. *The exterior derivative d on a smooth manifold M has the following properties for $\omega\in\Omega^k(M)$, $\tau\in\Omega^\ell(M)$ and $f\in\Omega^0(M)=C^\infty(M,\mathbb{R})$:*

(1) d is a local operator, so for an open subset $U\subset M$, $d\omega|_U$ depends only on $\omega|_U$.

(2) d is a graded derivation, i.e. we get $d(f\omega)=df\wedge\omega+fd\omega$ and, more generally, $d(\omega\wedge\tau)=(d\omega)\wedge\tau+(-1)^k\omega\wedge d\tau$.

(3) d is a differential, i.e. $d(d\omega)=0$.

(4) For a local chart (U,u) for M with $\omega|_U=\omega_{i_1\dots i_k}du^{i_1}\wedge\dots\wedge du^{i_k}$ we get

$$d\omega|_U=d(\omega_{i_1\dots i_k})\wedge du^{i_1}\wedge\dots\wedge du^{i_k}=\tfrac{\partial}{\partial u^{i_0}}(\omega_{i_1\dots i_k})du^{i_0}\wedge\dots\wedge du^{i_k}.$$

(5) d is compatible with pullbacks along arbitrary smooth maps, so for any smooth map $F:N\to M$, we get $d(F^*\omega)=F^*(d\omega)$.

Observe that in part (4) we are using standard Einstein sum convention, so the indices are not necessarily ordered. An explicit expression with ordered indices is not as easy to write out, since in the right hand side one would have to move du^{i_0} to the right position (that depends on all the i_j) at the expense of a sign.

Proof. (1) Let us first assume that ω vanishes identically on U. Inserting arbitrary vector fields in ω, the resulting function vanishes identically on U and hence the same is true for any derivative in the direction of a further vector field. Hence all summands in the defining equation (3.17) for $d\omega$ vanish identically on U, so $d\omega|_U=0$. Applying this to the difference of two forms that agree on U, linearity of d implies the claim.

We next prove the first claim in (2) by inserting $f\omega$ into the defining equation (3.17). For the second sum, f simply comes out, while for the terms in the first sum, we get $\xi_i(f\omega(\cdots))=\xi_i(f)\omega(\cdots)+f\xi_i(\omega(\cdots))$. Using $\xi_i(f)=df(\xi_i)$, this shows that

$$d(f\omega)(\xi_0,\dots,\xi_k)=fd\omega(\xi_0,\dots,\xi_k)+\textstyle\sum_i(-1)^idf(\xi_i)\omega(\xi_0,\dots,\widehat{\xi_i},\dots,\xi_k).$$

Skew symmetry of ω immediately implies that the sum coincides with $(df\wedge\omega)(\xi_0,\dots,\xi_k)$ as defined in (3.15).

Next, we claim that for a chart (U,u) for M and any choice of indices $i_1,\dots,i_k\in\{1,\dots,n\}$ the form $du^I:=du^{i_1}\wedge\dots\wedge du^{i_k}\in\Omega^k(U)$ hast the property that $d(du^I)=0$. To prove this, it suffices to show that $d(du^I)$ vanishes upon insertion of $k+1$ coordinate vector fields, since the coordinate vector fields form a basis for the tangent space in each point of U. But the Lie bracket of two coordinate vector fields always vanishes, while inserting k coordinate vector fields in du^I, we always get a constant function. This now implies that all the summands in the defining equation (3.17) vanish identically, which proves the claim.

Having these facts at hand, (4) becomes easy: By (1), we know that we can compute $d\omega|_U$ from the local coordinate representation, which is a sum of terms of the form $\omega_I du^I$. But then the first expression in (4) follows from the first property in (2) and $d(du^I) = 0$. The second expression then readily follows from the local coordinate expression $d(\omega_{i_1\dots i_k}) = d(\omega_{i_1\dots i_k})(\frac{\partial}{\partial u^i})du^i$.

The second property in (2) can be proved locally, and by bilinearity of the wedge product, it suffices to consider the case that $\omega = f du^I$ and $\tau = g du^J$ for some $f, g \in C^\infty(M, \mathbb{R})$ and sets I and J of indices. Using the properties of the wedge product, we get $\omega \wedge \tau = fg(du^I \wedge du^J)$. If nonzero, $du^I \wedge du^J$ is a wedge product of coordinate one forms, so $d(\omega \wedge \tau) = d(fg) \wedge du^I \wedge du^J$. Now $d(fg) = (df)g + fdg$, and inserting this and using the properties of the wedge product, we get

$$df \wedge du^I \wedge (g du^J) + (-1)^k f du^I \wedge dg \wedge du^J = (d\omega) \wedge \tau + (-1)^k \omega \wedge d\tau.$$

(3) Again we can verify this locally and restrict to $\omega = f du^I$ for which $d\omega = df \wedge du^I$ and then by (2) $d^2\omega = d^2 f \wedge du^I$ since $d(du^I) = 0$. But computing $d^2 f(\xi_0, \xi_1)$ directly from (3.17), we get

$$\xi_0(df(\xi_1)) - \xi_1(df(\xi_0)) - df([\xi_0, \xi_1]) = \xi_0(\xi_1(f)) - \xi_1(\xi_0(f)) - [\xi_0, \xi_1](f) = 0.$$

(5) For $f \in C^\infty(M, \mathbb{R}) = \Omega^0(M)$, we by definition have $F^*f = f \circ F$, so the last statement in Proposition 3.1 reads as $F^*df = d(F^*f)$. The proof of the general result can be done locally around a point $y \in N$, so we can take a chart (U, u) for M with $F(y) \in U$ and work on $F^{-1}(U) \subset N$. By linearity, we may again restrict to the case that $\omega = f du^I$ and then $d\omega = df \wedge du^I$ by part (4). Now we already know that $F^*du^i = d(u^i \circ F)$, so part (3) shows that $d(F^*du^i) = 0$. Now by Theorem 3.5, F^*du^I is the wedge product of the forms $F^*du^{i_j}$ so part (2) implies that $d(F^*du^I) = 0$. Again by Theorem 3.5, $F^*\omega = (F^*f)(F^*du^I)$ so applying (2), we get

$$d(F^*\omega) = d(F^*f) \wedge F^*(du^I) + 0 = F^*df \wedge F^*du^I = F^*d\omega.$$

□

Remark 3.6. From our proof we can simply deduce that, as a linear operator $\Omega^k(M) \to \Omega^{k+1}(M)$, the exterior derivative is characterized by a few simple properties. Namely, assume that we require the graded derivation property from (2) and $d^2 = 0$ from (3), as well as the fact that $df(\xi) = \xi(f)$. Then via (2), we can adapt the usual argument via a bump function to show that d is a local operator in the sense of part (1). But in our proof we have seen that (1), (2) and (3) readily lead to the coordinate formula for d from (4).

Example 3.6. We discuss here examples of one-forms and their exterior derivative on (open subsets of) $\mathbb{R}^3$, which nicely connects to topics we have studied already. Via local coordinates, a similar discussion applies on general smooth manifolds of dimension three, and parts extend to general dimensions. Suppose that M is a smooth manifold of dimension n and consider $\omega \in \Omega^1(M)$. For each point $x \in M$, $\omega(x) : T_xM \to \mathbb{R}$ is a linear map, so if $\omega(x) \neq 0$, then its kernel is a hyperplane in T_xM. Specializing to an open subset $U \subset \mathbb{R}^3$, we can write $\omega = \omega_i dx^i$ (using summation convention) for smooth functions $\omega_i : U \to \mathbb{R}$. Here the kernel in points, in which ω is non-zero, is a plane in $\mathbb{R}^3$, so we can compare to the planes spanned by two vector fields $\xi, \eta \in \mathfrak{X}(U)$ as in Examples 2.4 and 2.5 and realize them in this way.

Let us first take U to be the set of all points with positive last coordinate and $\xi = x^3\frac{\partial}{\partial x^1} - x^1\frac{\partial}{\partial x^3}$ and $\eta = x^3\frac{\partial}{\partial x^2} - x^2\frac{\partial}{\partial x^3}$ as in Example (1) of 2.4. Then $\omega(\xi) = x^3\omega_1 - x^1\omega_3$, so to have this vanishing we need $x^3\omega_1 = x^1\omega_3$ and similarly $\omega(\eta) = 0$

leads to $x^3\omega_2 = x^2\omega_3$. Since x^3 is nowhere vanishing on U, we can take any nowhere vanishing smooth function $f : U \to \mathbb{R}$, put $\omega_3 = f$, $\omega_1 = \frac{x^1}{x^3}f$ and $\omega_2 = \frac{x^2}{x^3}f$ in order to get $\omega \in \Omega^1(U)$ such that $\ker(\omega(x))$ is the span of $\xi(x)$ and $\eta(x)$ for all $x \in U$.

Taking f to be the constant function 1, we get $\omega = \frac{x^1}{x^3}dx^1 + \frac{x^2}{x^3}dx^2 + dx^3$. Now $d\frac{x^1}{x^3} = \frac{1}{x^3}dx^1 - \frac{x^1}{(x^3)^2}dx^3$, and similarly for $\frac{x^2}{x^3}$. Skew symmetry of the wedge product then leads to $d\omega = \frac{x^1}{(x^3)^2}dx^1 \wedge dx^3 + \frac{x^2}{(x^3)^2}dx^2 \wedge dx^3$. At this point, one can verify explicitly that $d\omega(\xi, \eta)$ vanishes identically, which follows from abstract reasoning, since by (3.17), $\omega(\xi) = 0$ and $\omega(\eta) = 0$ implies $d\omega(\xi, \eta) = -\omega([\xi, \eta])$ and Example 2.4 (1). Alternatively, we can write the expression for $d\omega$ as $\omega \wedge \tau$ where $\tau = \frac{1}{x^3}dx^3$, which also clearly implies $d\omega(\xi, \eta) = 0$ as well as $\omega \wedge d\omega = 0$ (by skew symmetry of the wedge product).

These observations also tell us how we can simplify matters. Since x^3 is nowhere vanishing, $x^3\omega$ has the same kernel in each point as ω. But we know that $d(x^3\omega) = dx^3 \wedge \omega + x^3 d\omega = 0$, which is also immediately seen directly from $x^3\omega = \sum_i x^i dx^i$. Indeed, we can write $x^3\omega = dg$, where $g(x) = \frac{1}{2}\sum_i (x^i)^2$. This provides another argument for $d(x^3\omega) = 0$ and, as we have observed in Example 2.5, the planes spanned by $\xi(x)$ and $\eta(x)$ are exactly the tangent spaces to the level sets of g.

In Example (2) of 2.4, we considered $\xi = \frac{\partial}{\partial x^1} - x^2\frac{\partial}{\partial x^3}$ and $\eta = \frac{\partial}{\partial x^2} + x^1\frac{\partial}{\partial x^3}$ on $\mathbb{R}^3$. In this case, a simple choice for $\omega \in \Omega^1(M)$ such that $\ker(\omega(x))$ is the span of $\xi(x)$ and $\eta(x)$ for each $x \in \mathbb{R}^3$ is $\omega = x^2dx^1 - x^1dx^2 + dx^3$. This readily leads to $d\omega = -2dx^1 \wedge dx^2$ and hence $d\omega(\xi, \eta) = -2$, which is consistent with the computation of $[\xi, \eta]$ in 2.4. This in turn shows that $\omega \wedge d\omega = -2dx^1 \wedge dx^2 \wedge dx^3$, so this is nowhere vanishing.

Indeed, the last property depends only on the kernel of ω and not on the form itself: Any other one-form with the same kernel can be written as $f\omega$ for a nowhere vanishing function $f \in C^\infty(\mathbb{R}^3, \mathbb{R})$ and this leads to $d(f\omega) = df \wedge \omega + fd\omega$ and thus to $(f\omega) \wedge d(f\omega) = f^2(\omega \wedge d\omega)$, so again this is nowhere vanishing.

The relation between the exterior derivative and submanifolds that becomes visible in these examples also admits a direct explanation in terms of the calculus of differential forms. This applies to forms of any degree and is very useful in many situations. Suppose that M is a smooth manifold and $N \subset M$ a submanifold. Denoting by $i : N \to M$ the inclusion, we get the map $i^* : \Omega^k(M) \to \Omega^k(N)$ for any k. By definition, for $\omega \in \Omega^k(M)$, the form $i^*\omega$ is just the restriction of ω to N in an obvious sense, i.e. one evaluates ω in points of N and restricts the corresponding multilinear map to vectors tangent to N. (Therefore, the operator i^* is often suppressed from the notation, and one simply views ω also as a form on N.) But now compatibility of the exterior derivative with pullbacks shows that $i^*d\omega = di^*\omega$. Hence if we know that $i^*\omega = 0$ (i.e. that ω vanishes along N if all its entries are tangent to N), then $d\omega$ must have the same property. Hence this provides a systematic way to obtain obstructions against the existence of certain submanifolds.

3.7. Excursion: Insertion operators and Lie derivatives. Here we briefly discuss how to extend the exterior derivative and the Lie derivative that we already know from 3.4 to a calculus on differential forms. In the case of differential forms, formula (3.10) from Section 3.4 simplifies. For $\omega \in \Omega^k(M)$, we get

$$(\mathcal{L}_\eta\omega)(\xi_1, \dots, \xi_k) = \eta(\omega(\xi_1, \dots, \xi_k)) - \textstyle\sum_{i=1}^k \omega(\xi_1, \dots, [\eta, \xi_i], \dots, \xi_k). \tag{3.18}$$

This is evidently skew symmetric and a short computation shows that it is linear over smooth functions in all entries, so $\mathcal{L}_\eta\omega \in \Omega^k(M)$ by the version of Lemma 3.3 for differential forms. This can be taken as a definition of the Lie derivative, but for many applications it is important to know the relation to pullbacks along flows as discussed in Section 3.4. The general results from that section (or alternatively direct computations based on (3.18)) show that $\mathcal{L}_\eta(\omega \wedge \tau) = (\mathcal{L}_\eta\omega) \wedge \tau + \omega \wedge (\mathcal{L}_\eta\tau)$.

The second type of operation is even easier. Given $\eta \in \mathfrak{X}(M)$ and $\omega \in \Omega^k(M)$, we define a map $i_\eta\omega$ on $(\mathfrak{X}(M))^{k-1}$ by

$$(i_\eta\omega)(\xi_1, \dots, \xi_{k-1}) := \omega(\eta, \xi_1, \dots, \xi_{k-1}) \tag{3.19}$$

Obviously, the right hand side is in $C^\infty(M, \mathbb{R})$ and $i_\eta\omega$ is alternating an linear over smooth functions in each variable, so $i_\eta\omega \in \Omega^{k-1}(M)$. Thus we have a linear operator $i_\eta : \Omega^k(M) \to \Omega^{k-1}(M)$ (where $\Omega^{-1}(M) = \{0\}$), called the *insertion operator* associated to $\eta \in \mathfrak{X}(M)$. A direct computation via the definition of the wedge product shows that for $\omega \in \Omega^k(M)$ and $\tau \in \Omega^\ell(M)$, one gets $i_\eta(\omega \wedge \tau) = (i_\eta\omega) \wedge \tau + (-1)^k\omega \wedge (i_\eta\tau)$. In particular, $i_\eta(f\omega) = f i_\eta\omega$, which is actually obvious from (3.19).

Now there is a number of useful identities relating these operations, which are collected in the following result:

PROPOSITION 3.7. *Let M be a smooth manifold. Then for vector fields $\xi, \eta \in \mathfrak{X}(M)$ and $\omega \in \Omega^k(M)$, we have the following identities:*

(1) ("Cartan's magic formula") $\mathcal{L}_\xi\omega = i_\xi(d\omega) + d(i_\xi\omega)$
(2) $d(\mathcal{L}_\xi\omega) = \mathcal{L}_\xi(d\omega)$
(3) $\mathcal{L}_{[\xi,\eta]}\omega = \mathcal{L}_\xi(\mathcal{L}_\eta\omega) - \mathcal{L}_\eta(\mathcal{L}_\xi\omega)$
(4) $i_{[\xi,\eta]}\omega = \mathcal{L}_\xi(i_\eta\omega) - i_\eta(\mathcal{L}_\xi\omega)$.
(5) $i_\eta(i_\xi\omega) = -i_\xi(i_\eta\omega)$.

All of these identities can be proved by direct computations from the definitions of the operations. In some cases, say for (4) and (5), this is rather easy and a good exercise, in other cases it is more tedious. Understanding the relation between Lie derivatives and pullbacks along flows, (2) is a simple consequence of naturality of d, and so on.

There is a neat algebraic way to interpret these (and further identities), which also leads to a simple uniform way of proving them. We can view $\Omega^*(M) = \oplus_{k=0}^n \Omega^k(M)$ as a graded algebra and write $|\omega| = k$ for $\omega \in \Omega^k(M)$. The graded commutativity property of the wedge product then reads as $\tau \wedge \omega = (-1)^{|\omega|\cdot|\tau|}\omega \wedge \tau$. Now one defines a *graded derivation of degree* r on $\Omega^*(M)$ as a linear map $D : \Omega^*(M) \to \Omega^*(M)$ such that $D(\Omega^k(M)) \subset \Omega^{k+r}(M)$ and such that $D(\omega \wedge \tau) = D(\omega) \wedge \tau + (-1)^{r|\omega|}\omega \wedge D(\tau)$. Here we agree that $\Omega^s(M) = \{0\}$ if $s \notin \{0, \dots, \dim(M)\}$. In this language, d, $\mathcal{L}_\xi$ and i_ξ are graded derivations of degree 1, 0 and -1, respectively.

Similarly to the discussion in Section 2.2, the composition of two graded derivations is not a graded derivation, but if D_1 has degree r and D_2 has degree s, then the *graded commutator* $[D_1, D_2] := D_1 \circ D_2 - (-1)^{rs} D_2 \circ D_1$ is a graded derivation of degree $r + s$. In this language, the statements in Proposition 3.7 read as $\mathcal{L}_\xi = [d, i_\xi]$, $[d, \mathcal{L}_\xi] = 0$, $\mathcal{L}_{[\xi,\eta]} = [\mathcal{L}_\xi, \mathcal{L}_\eta]$, $i_{[\xi,\eta]} = [\mathcal{L}_\xi, i_\eta]$, and $[i_\xi, i_\eta] = 0$, respectively. Thus, all these statements boil down to the equality of two graded derivations of the same degree. But then it is easy to show that any graded derivation on $\Omega^*(M)$ is a local operator, so studying these one can work with local coordinate expressions. Using this, it is easy to show that two graded derivations on $\Omega^*(M)$ agree provided that they agree on $\Omega^0(M)$ and on all one-forms of the form df with $f \in \Omega^0(M)$. But in these cases, all the above identities are easy to verify directly.

Excursion: Symplectic manifolds and classical mechanics

To conclude the chapter, we briefly discuss a "geometric structure" that is closely related to differential forms and is of interest both in mathematics and in physics.

3.8. Symplectic manifolds. Recall from linear algebra that skew symmetric bilinear forms $b : V \times V \to \mathbb{R}$ on a real vector space V are classified by their rank, which automatically is even. Most conceptually, this can be viewed in terms of the null-space $N = \{v \in V : \forall w \in V : b(v, w) = 0\}$. Clearly $N \subset V$ is a linear subspace and the rank $\mathrm{rk}(b)$ of b is $n - \dim(N)$, where $n := \dim(V)$. In particular, b has the maximal possible rank n if and only if it is non-degenerate as a bilinear form. Alternatively, $rk(b)$ can be viewed as the rank of the $n \times n$-matrix $(b(v_i, v_j))$ for any basis $\{v_1, \dots, v_n\}$ of V. Skew symmetry of this matrix then implies that $\mathrm{rk}(b)$ is even. In particular, non-degenerate skew symmetric bilinear forms only exist in even dimensions and in each even dimension, there is only one such form up to isomorphism. One possible realization of this is the imaginary part of the standard Hermitian form on $\mathbb{C}^n \cong \mathbb{R}^{2n}$.

Definition 3.8. Let M be a smooth manifold of even dimension $2n$. A *symplectic form* or a *symplectic structure* on M is a two-form $\omega \in \Omega^2(M)$ such that

- For each $x \in M$, $\omega(x) : T_xM \times T_xM \to \mathbb{R}$ is non-degenerate.
- $d\omega = 0$

The pair (M, ω) is referred to as a symplectic manifold.

If there exists a one-form $\varphi \in \Omega^1(M)$ such that $\omega = d\varphi$, then the symplectic structure ω is called *exact* and (M, ω) is called an exact symplectic manifold.

Let us postpone the question of why one should look at such objects for the moment, but use the calculus of differential forms to derive some consequences:

Theorem 3.8. *Let (M, ω) be a symplectic manifold dimension $2n$.*

(1) For each $x \in M$, $\omega(x)$ induces a linear isomorphism $T_xM \to T_x^*M$. The inverses of these isomorphisms fit together to define a $\binom{2}{0}$-tensor field $\Pi \in \mathcal{T}_0^2(M)$.

(2) Mapping ξ to $i_\xi\omega$ induces an isomorphism $\mathfrak{X}(M) \to \Omega^1(M)$, whose inverse is induced by Π in a similar way.

(3) For any smooth function $f \in C^\infty(M, \mathbb{R})$, there is a unique vector field $H_f \in \mathfrak{X}(M)$ such that $i_{H_f}\omega = df \in \Omega^1(M)$. In addition, we get $\mathcal{L}_{H_f}\omega = 0$.

(4) Putting $\{f, g\} := -\omega(H_f, H_g) = dg(H_f) = -df(H_g)$ defines a skew symmetric bilinear operation on $C^\infty(M, \mathbb{R})$, which satisfies $\{f, gh\} = \{f, g\}h + g\{f, h\}$ for all $f, g, h \in C^\infty(M, \mathbb{R})$.

Proof. (1) Non-degeneracy of $\omega(x)$ exactly says that for each $0 \neq X \in T_xM$, the linear functional $\omega(x)(X, \_)$ is non-zero. Viewed as a linear map $T_xM \to T_x^*M$, $\omega(x)$ thus is injective and hence bijective. The inverse of this linear isomorphism can be viewed as $\Pi_x \in \otimes^2 T_xM$. Writing ω as a tensor field in local coordinates around x as $\omega_{ij} du^i \otimes du^j$, we get $\Pi_x = B^{ij} \frac{\partial}{\partial u^i}|_x \otimes \frac{\partial}{\partial u^j}|_x$, where (B^{ij}) is the inverse matrix to $(\omega_{ij}(x))$. Since matrix inversion is a smooth map, we see that $x \mapsto \Pi_x$ is smooth in local coordinates.

(2) For $\xi \in \mathfrak{X}(M)$, we get $i_\xi\omega \in \Omega^1(M)$. Conversely, for $\Pi \in \mathcal{T}_0^2(M)$ and $\varphi \in \Omega^1(M)$, we can form the contraction $C_1^1(\Pi \otimes \varphi) \in \mathcal{T}_0^1(M) = \mathfrak{X}(M)$. The considerations in (1) easily imply that these two constructions are inverse to each other.

(3) Uniqueness of H_f is clear by injectivity of the map $\xi \mapsto i_\xi\omega$ from (2). Moreover, for $f \in C^\infty(M, \mathbb{R})$ we can form $df \in \Omega^1(M)$ and then $C_1^1(\Pi \otimes df) \in \mathfrak{X}(M)$ has the desired

property by (2). To compute $\mathcal{L}_{H_f}\omega$, we use Cartan's magic formula from Proposition 3.7 and $d\omega = 0$. This gives $\mathcal{L}_{H_f}\omega = d(i_{H_f}\omega) = d(df) = 0$.

(4) By definition of the vector fields and skew symmetry of ω, we get $-\omega(H_f, H_g) = dg(H_f) = -df(H_g)$. Linearity of d immediately implies that $f \mapsto H_f$ is a linear map, which shows that $\{\ ,\ \}$ is bilinear, while skew symmetry is obvious. By definition, we get $i_{f\xi}\omega = fi_\xi\omega$ and $i_{\xi+\eta}\omega = i_\xi\omega + i_\eta\omega$. This shows that inserting $gH_h + hH_g \in \mathfrak{X}(M)$ into ω, we get $hdg + gdh = d(gh)$. Hence $H_{gh} = gH_h + hH_g$ and from this, the last claimed property follows immediately. □

The tensor field Π is called the *Poisson tensor* associated to the symplectic form ω, while the vector field H_f is called the *Hamiltonian vector field* associated to $f \in C^\infty(M, \mathbb{R})$. The operation $\{\ ,\ \}$ on $C^\infty(M, \mathbb{R})$ is called the *Poisson bracket*. It is not too difficult to show that the Poisson bracket defines a Lie algebra structure on $C^\infty(M, \mathbb{R})$ and that $H_{\{f,g\}} = [H_f, H_g]$, so one obtains a homomorphism to the Lie algebra of vector fields. Let us next discuss an example, which is also crucial for the applications of symplectic structures to physics. These applications also motivate the terminology we use.

EXAMPLE 3.8. The simplest version of this is to take $M := U \times \mathbb{R}^n$ for an open subset $U \subset \mathbb{R}^n$ and use coordinates q^i ("positions") and p_i ("momenta") for $i = 1, \dots, n$ on the two factors. The put $\varphi = \sum_i p_i dq^i \in \Omega^1(M)$ and $\omega = -d\varphi = \sum_i dq^i \wedge dp_i \in \Omega^2(M)$. This is obviously non-degenerate and thus defines an exact symplectic structure on M.

This generalizes vastly in a surprising way: Let N be any smooth manifold ("configuration space") of dimension n and put $M := T^*N$, the cotangent bundle ("phase space"). Then there is a canonical one-form $\varphi \in \Omega^1(M)$ constructed as follows. There is a canonical projection from $M = T^*N$ to N which we denote by π here. Now a point in M by definition is an element of some cotangent space of N, i.e. a linear map $\lambda : T_{\pi(\lambda)}N \to \mathbb{R}$. Now given $X \in T_\lambda M$, we can form $T_\lambda\pi(X) \in T_{\pi(\lambda)}N$, so it makes sense to form $\varphi(\lambda)(X) := \lambda(T_\lambda\pi(X))$. To see that this is smooth, we use appropriate charts following the above notation (which is standard in physics).

Take a chart for N defined on $U \subset N$ and denote the corresponding local coordinates by q^i. Then any point $\lambda \in \pi^{-1}(U) \subset M$ can be uniquely expanded in the basis $dq^i(\pi(\lambda))$ and we denote by $p_i = p_i(\lambda)$ the coefficients in this expansion. Then we can use the functions q^i (actually these are $q^i \circ \pi$ but it is standard to drop this from the notation) and p_i for $i = 1, \dots, n$ as the local coordinates on the manifold M. From this description, it follows readily that (with our slight abuse of notation) we can write $\varphi(\lambda) = \sum_i p_i(\lambda)dq^i(\lambda)$. Thus we get $\varphi|_{\pi^{-1}(U)} = \sum_i p_i dq^i$ and, as above, for $\omega = -d\varphi$, we get $\omega|_{\pi^{-1}(U)} = \sum_i dq^i \wedge dp_i$, so (M, ω) is an exact symplectic manifold. (I am not aware of an explanation from physics why momenta should be linear functionals rather than vectors but this is how a generalization works.)

The natural construction of the symplectic structure on $M = T^*N$ has a very remarkable consequence: If $F : N \to N$ is any diffeomorphism, we get an induced diffeomorphism $\tilde{F} := T^*F : T^*N \to T^*N$ such that $\pi \circ \tilde{F} = F \circ \pi$. The latter equation implies that $T_{\tilde{F}(\lambda)}\pi \circ T_\lambda\tilde{F} = T_{\pi(\lambda)}F \circ T_\lambda\pi$. Using this, a short computation shows that $\tilde{F}^*\varphi = \varphi$, which in turn gives $\tilde{F}^*\omega = -F^*(d\varphi) = -d(F^*\varphi) = \omega$. Thus any diffeomorphism of N induces a diffeomorphism of T^*N compatible with the symplectic structure, and there is an infinite dimensional family of these.

REMARK 3.8. There are further indications that there are many diffeomorphisms preserving a symplectic structure. As we have seen in part (3) of Theorem 2.8 any

Hamiltonian vector field H_f has the property that $\mathcal{L}_{H_f}\omega = 0$. With the interpretation via flows discussed in Section 3.4 (see also Theorem 2.11) one easily concludes that $(\mathrm{Fl}^{H_f}_t)^*\omega = \omega$ whenever the flow is defined. In fact there are further strong results showing that symplectic structures are very flexible (at least locally). In particular, the so-called *Darboux theorem* (see Theorem 22.13 in [**Lee**] or Theorem 31.15 in [**Michor**]) states that for a symplectic manifold (M, ω) and a point $x \in M$, one can always find local coordinates q^i and p_i around x in which ω has the form $\sum_i(dq^i \wedge dp_i)$. In particular, all symplectic manifolds are locally isomorphic. However, global properties of symplectic structures are a highly interesting topic with lots of current research activity.

3.9. Hamiltonian mechanics. Now we can discuss the motivation for looking at symplectic manifolds and indicate how the tools from Proposition 3.8 can be applied. We go to the physics setting that $N \subset \mathbb{R}^n$ is an open subset and we consider $M = N \times \mathbb{R}^n$ with coordinates q^i and p_i and the symplectic form $\sum_i dq^i \wedge dp_i$. One interprets N as the possible positions of some mechanical system and the second components as recording momentum (mass times velocity). Introducing the momentum as an independent variable implements the common trick to reduce the order of differential equations (in this case from second order to first order) by introducing derivatives as new variables. The time evolution of the system will be described by a curve in M, whose components we simply denote by $q^i(t)$ and $p_i(t)$ for $i = 1, \dots, n$. The interpretation of the p_i as momenta then says that $p_i(t) = m \cdot (q^i)'(t)$ for some constant m.

Now we assume that we have some force-field on N (say an electrical field acting on a charged particle) that can be described by a potential. This means that there is a smooth function $V : N \to \mathbb{R}$ such that the force experienced at the point q is $-grad(V)(q)$, the negative of the gradient of V. Now by Newton's second law (which is the only real physics ingredient in the whole discussion) this force causes an acceleration given by the quotient of the force by mass. So this leads to the equation $(q^i)''(t) = -\frac{1}{m}\frac{\partial V}{\partial q^i}(q(t))$, which can be equivalently written as $(p_i)'(t) = -\frac{\partial V}{\partial q^i}(q(t))$.

Combining this, we get a system of first order ODEs in the variables q^i and p_i that govern the system. We can now show that these are the equations for the integral curves of the Hamiltonian vector field H_E associated to a smooth function $E : M \to \mathbb{R}$. As an additional bonus, this function has a beautiful interpretation in physics terms, namely as the total energy (i.e. kinectic energy plus potential energy): Putting $E = \sum_i \frac{(p_i)^2}{2m} + V(q)$, we get $dE = \sum_i \frac{p_i}{m} dp_i + \sum_i \frac{\partial V}{\partial q^i} dq^i$. Making the ansatz $H_E = \sum_i (a^i \frac{\partial}{\partial q^i} + b^i \frac{\partial}{\partial p_i})$ and computing $i_{H_E}\omega$, we readily conclude that $a_i = \frac{p_i}{m}$ and $b_i = -\frac{\partial V}{\partial q^i}$, which proves the claim.

This has a number of important consequences. First $dE(H_E) = -\omega(H_E, H_E) = 0$, so E is constant along flow lines of H_E (conservation of energy). More generally, for any smooth function $f : M \to \mathbb{R}$ such that $\{f, E\} = 0$, we have $H_E(f) = 0$, so again f is constant along flow lines of H_E. Thus one obtains a systematic approach to identify constants of motion.

Guided by this example, symplectic manifolds are considered as the natural arena for studying classical mechanics. The symplectic nature of classical mechanics is reflected even in the setting of open subsets of $\mathbb{R}^n$ by the concept of what physicists call a *canonical transformation*. This is a diffeomorphism of $M = T^*N$ which pulls back the canonical symplectic form to itself. As we have seen in 3.8 above, any diffeomorphism of N (a "point transformation") induces a canonical transformation, but in fact, canonical transformations form a much larger class. In physics terms, a canonical transformation may mix positions and momenta, which is an unexpected idea from the point of view of physics. Such transformations may be used to establish isomorphisms between different systems that are highly non-obvious from a naive point of view, and these can be used to study their physical properties.

CHAPTER 4

Integration and de-Rham cohomology

In the last part of the course, we discuss two topics that are related to differential forms. The first topic is integration on manifolds. The emphasis here is not on integrating a very general class of functions, but to develop a theory of integration for smooth objects with compact support that is independent of coordinates and natural under diffeomorphisms. An extension to non-compact support and to non-smooth objects is then possible in a second step similar to what is classically done on $\mathbb{R}^n$ (approximation by compactly supported objects, completion of spaces of smooth objects with respect to integral norms, and/or measure theory). The requirement of naturality for the integral brings in a completely new aspect, which is already fully visible in the smooth case. In particular, functions are not the appropriate object to be integrated, and there are two possible replacements. One may either work with so-called densities or restrict to oriented manifolds, on which differential forms of maximal degree can be integrated.

Integration of differential forms is also the setting in which the most fundamental theorem about integration is formulated, namely Stokes's theorem. This needs an extension of the concept of a manifold, namely manifolds with boundary. Stokes's theorem connects integrals of exterior derivatives of forms over a manifold with boundary to integrals over the boundary (which is again a manifold). The classical theorems from analysis concerning boundary integrals can all be deduced rather easily from the general version.

Finally, we discuss de-Rham cohomology, which provides a connection to algebraic topology. This mainly builds on differential forms and the exterior derivative, but in some points also integration becomes an important ingredient, which is why we haven't discussed it earlier.

4.1. Densities. The problem of finding a notion of integral that is independent of coordinates is already there for objects having support contained in the intersection of two charts. Equivalently on can assume that the support is contained in an open subset, which is the domain of two charts. Indeed, as we shall see that understanding this case, a general integral can be built up via some technicalities involving partitions of unity.

To understand what has to happen in this special case, recall how multiple integrals behave under a change of variables. Consider an open subset $U \subset \mathbb{R}^n$ and a smooth function f with compact support that is contained in U. To avoid problems with notation, we simply denote the integral of f over U by $\int_U f$. For another open subset $V \subset \mathbb{R}^n$ and a diffeomorphism $\Phi : V \to U$, $f \circ \Phi$ is a smooth function with compact support that is contained in V. On the other hand, for each $x \in V$, $D\Phi(x) : \mathbb{R}^n \to \mathbb{R}^n$ is a linear isomorphism, which has a well defined determinant. Hence we can consider $x \mapsto f(\Phi(x))|\det(D\Phi(x))|$, which is a smooth function with compact support contained in $\mathrm{supp}(f \circ \Phi) \subset V$. The transformation law for multiple integrals then states that

$$\int_U f = \int_{\Phi(V)} f = \int_V (f \circ \Phi)|\det(D\Phi)|. \tag{4.1}$$

As we shall see, the absolute value that occurs in this formula is responsible for all the trouble in integration theory on manifolds. However, this clearly has to be there, e.g. since integrals of non-negative functions should always be non-negative.

This also suggests a brute-force solution to the problem of integration. On introduces a class of geometric objects, called *densities*, which have the right transformation law in order to be integrated. There are conceptual ways to do this (in the language of natural vector bundles). Since these are beyond the scope of this course and essentially we are only interested on one class of examples, we use a coordinate based "definition" here. If you feel uneasy with this kind of definition, just view it as a wording used to describe the examples occurring below.

Definition 4.1. A *density* ν on a smooth manifold M is described in a chart (U_α, u_α) by a smooth function $\nu_\alpha : U_\alpha \to \mathbb{R}$ in such a way that for two charts (U_α, u_α) and (U_β, u_β) with $U_{\alpha\beta} \neq \emptyset$ and chart change $u_{\alpha\beta} : u_\beta(U_{\alpha\beta}) \to u_\alpha(U_{\alpha\beta})$, the corresponding functions are related by

$$\nu_\beta(x) = \nu_\alpha(x)|\det(Du_{\alpha\beta}(u_\beta(x)))|. \tag{4.2}$$

Observe first that it is no problem to restrict densities to open subsets and to talk about locally defined densities. Next, $|\det(Du_{\alpha\beta}(u_\beta(x)))| > 0$, so if $\nu_\alpha(x) \neq 0$ for one chart, then it is non-zero for any chart. Hence it makes sense to say that ν is non-zero in a point, and hence a density has a well defined support $\operatorname{supp}(\nu)$. But clearly, ν does not have a well defined value at x if it is non-zero at x. Moreover, from the definition it is clear that densities can be added and multiplied by smooth functions point-wise and we use the usual notation $\nu + \mu$ and $f\nu$ for these operations. Hence densities form a vector space and a module over $C^\infty(M, \mathbb{R})$. However, in this case this module behaves like a one-dimensional vector space: Suppose that ν is density on M and x is a point in which ν is non-zero. Then ν is non-zero on an open neighborhood U of x in M. If μ is a density defined on U, then for $y \in U$ we can use a chart (U_α, u_α) with $y \in U_\alpha$ and consider $f(y) := \mu_\alpha(y)/\nu_\alpha(y) \in \mathbb{R}$. From (4.2) we see that this is independent of the choice of charts, so we get a smooth function $f : U \to \mathbb{R}$ such that $\mu = f\nu$. So if there are global, nowhere vanishing densities on M and we fix one density ν with this property, then $f \mapsto f\nu$ identifies $C^\infty(M, \mathbb{R})$ with the space of densities. This is exactly what happens for the central example of densities that we discuss next.

Example 4.1. As in Example (2) of 3.2, we consider a $\binom{0}{2}$-tensor field $g \in \mathcal{T}^0_2(M)$ on on a smooth manifold M, such that $g_x : T_xM \times T_xM \to \mathbb{R}$ is symmetric for each $x \in M$. For a chart (U_α, u_α), this is described by the functions $g^\alpha_{ij} = g(\frac{\partial}{\partial u^i_\alpha}, \frac{\partial}{\partial u^j_\alpha}) : U_\alpha \to \mathbb{R}$. The values $g^\alpha_{ij}(x)$ are exactly the components of the (symmetric) matrix associated to the symmetric bilinear form g_x with respect to the basis formed by the coordinate vector fields. In particular, we can form the determinant $\det(g^\alpha_{ij}(x))$ which gives rise to a smooth function $U_\alpha \to \mathbb{R}$. If we assume that g is a Riemannian metric i.e. that g_x is positive definite for each x, then $\det(g^\alpha_{ij}(x)) > 0$ for each x. Hence $\nu_\alpha(x) := \sqrt{\det(g^\alpha_{ij}(x))}$ defines a (nowhere vanishing) smooth function $\nu_\alpha : U_\alpha \to \mathbb{R}$.

If (U_β, u_β) is another chart such that $U_{\alpha\beta} \neq \emptyset$, then formula (3.5) from Section 3.2 shows that $g^\beta_{ij} = A^k_i A^\ell_j g^\alpha_{k\ell}$, where $A^r_s(x) = \partial_s u^r_{\alpha\beta}(u_\beta(x))$ for $x \in U_{\alpha\beta}$. In matrix language, these relations read as $(g^\beta_{ij}) = A(g^\alpha_{ij})A^t$ and as $A(x) = Du_{\alpha\beta}(u_\beta(x))$, respectively. (Observe that by symmetry, it doesn't matter which index of g corresponds to rows and which to columns.) But now of course $\det(A^t) = \det(A)$ and $\sqrt{\det(A)^2} = |\det(A)|$, and this readily shows that ν_β and ν_α are related by (4.2).

But this exactly shows that our construction leads to a density ν that is defined on all of M and nowhere vanishing by construction. This is called the *volume density* of the Riemannian metric g. (In fact, there is a very similar construction in the case that g is a pseudo-Riemannian metric, since $\det(g_{ij})$ always has constant sign.) As discussed above, a Riemannian metric on a manifold M thus gives rise to an identification of $C^\infty(M,\mathbb{R})$ with the space of densities. As we shall see below, this leads to a well defined integral for (compactly supported) smooth functions on Riemannian manifolds.

4.2. Integration of densities. Getting to a well defined integral for densities (with compact support) now mainly is a bit of technical trickery based on partitions of unity. We first derive a simple consequence of the existence of partitions of unity. (At this point, it would not be necessary to start from an atlas, but this will be useful later on.)

LEMMA 4.2. *Let M be a smooth manifold, $K \subset M$ a compact subset and let $\mathcal{A}$ be an atlas for M. Then we can find finitely many charts (U_α, u_α), $\alpha \in \{1,\dots,N\}$ from $\mathcal{A}$ and smooth functions $\varphi_\alpha \in C^\infty(M,\mathbb{R})$ with values in $[0,1]$ such that* $\operatorname{supp}(\varphi_\alpha) \subset U_\alpha$ *and $\sum_{\alpha=1}^N \varphi_\alpha$ is identically one on K.*

PROOF. Since K is compact, we can find finitely many charts in $\mathcal{A}$ such that $K \subset \cup_{\alpha=1}^N U_\alpha$. Adding $U_0 := M \setminus K$, we obtain an open covering of M, so by Theorem 1.9, there is a partition of unity $\{\tilde\varphi_i : i \in \mathbb{N}\}$ subordinate to this covering. Put $A_1 := \{i : \operatorname{supp}(\tilde\varphi_i) \subset U_1\}$ and $\varphi_1 := \sum_{i\in A_1} \tilde\varphi_i$. Since the family of supports is locally finite, this is smooth. Moreover, one easily verifies (see exercises) that the union of a locally finite family of closed sets is closed. This shows that $\cup_{i\in A_1} \operatorname{supp}(\tilde\varphi_i)$ is a closed set, and this certainly contains all points in which φ_1 is non-zero. Hence $\operatorname{supp}(\varphi_1) \subset \cup_{i\in A_1} \operatorname{supp}(\varphi_i) \subset U_1$.

Similarly, putting $A_2 := \{i \in \mathbb{N} \setminus A_1 : \operatorname{supp}(\tilde\varphi_i) \subset U_2\}$ and $\varphi_2 := \sum_{i\in A_2} \tilde\varphi_i$, we obtain a smooth function with $\operatorname{supp}(\varphi_2) \subset U_2$. Continuing inductively we obtain smooth functions φ_α for $\alpha \in \{1,\dots,N\}$, and it remains to verify that $\sum \varphi_\alpha$ is identically one on K. But if $j \notin A_1 \cup \dots \cup A_N$, we must by construction have $\operatorname{supp}(\tilde\varphi_j) \subset U_0$, so this is identically zero on K. Hence on K, $\sum \varphi_\alpha$ coincides with $\sum_{i\in\mathbb{N}} \tilde\varphi_i = 1$. □

Using this we can now define the integral of a density ν with compact support as follows. By Lemma 4.2, we find finitely many charts (U_α, u_α) for M and functions φ_α such that $\operatorname{supp}(\nu) \subset \cup_\alpha U_\alpha$ and $\sum_\alpha \varphi_\alpha$ is identically one on $\operatorname{supp}(\nu)$. With respect to the chart (U_α, u_α), ν is represented by a smooth function $\nu_\alpha : U_\alpha \to \mathbb{R}$. Now $(\varphi_\alpha \nu_\alpha) \circ u_\alpha^{-1} : u_\alpha(U_\alpha) \to \mathbb{R}$ is a smooth function with compact support, which can be integrated over $\mathbb{R}^n$ without problems (and has finite integral). Thus we may define

$$\int_M \nu := \sum_{\alpha=1}^N \int_{\mathbb{R}^n} (\varphi_\alpha \nu_\alpha) \circ u_\alpha^{-1}. \tag{4.3}$$

It is not clear that this is well defined, but it is not so difficult to verify this:

THEOREM 4.2. *The expression in* (4.3) *is independent of the choice of the charts (U_α, u_α) and of the functions φ_α. Thus we obtain a well defined integral, which defines a surjective linear map from the space of densities with compact support to $\mathbb{R}$.*

PROOF. Suppose that (V_i, v_i) and ψ_i with $i = 1,\dots,L$ is another choice of charts and functions. Putting $K := \operatorname{supp}(\nu)$, we see that for each fixed α, we get $\varphi_\alpha|_K = \sum_i \psi_i \varphi_\alpha|_K$ and $\psi_i \varphi_\alpha$ has support in $U_\alpha \cap V_i$. By linearity of the integral of functions on $\mathbb{R}^n$, we conclude that

$$\sum_\alpha \int_{\mathbb{R}^n} (\varphi_\alpha \nu_\alpha) \circ u_\alpha^{-1} = \sum_{\alpha,i} \int_{\mathbb{R}^n} (\psi_i \varphi_\alpha \nu_\alpha) \circ u_\alpha^{-1}. \tag{4.4}$$

In the same way, we can start from (V_i, v_i), and the functions $\psi_i\nu$, which leads to

$$\sum_i \int_{\mathbb{R}^n} (\psi_i\nu_i) \circ v_i^{-1} = \sum_{i,\alpha} \int_{\mathbb{R}^n} (\varphi_\alpha\psi_i\nu_i) \circ v_i^{-1}. \tag{4.5}$$

Now both in (4.4) and in (4.5) it suffices to sum over those pairs (α, i) for which $U_\alpha \cap V_i \neq \emptyset$ since $\mathrm{supp}(\varphi_\alpha\psi_i)$ is contained in this subset. Fixing one such pair (α, i), this also shows that we can change the domain of integration to $u_\alpha(U_\alpha \cap V_i)$ in (4.4) and to $v_i(U_\alpha \cap V_i)$ in (4.5).

Now consider the chart change $\Phi := u_\alpha \circ v_i^{-1} : v_i(U_\alpha \cap V_i) \to u_\alpha(U_\alpha \cap V_i)$. Now $u_\alpha(U_\alpha \cap V_i) = \Phi(v_i(U_\alpha \cap V_i))$, so by (4.1), we get

$$\int_{u_\alpha(U_\alpha\cap V_i)} (\psi_i\varphi_\alpha\nu_\alpha) \circ u_\alpha^{-1} = \int_{v_i(U_\alpha\cap V_i)} ((\psi_i\varphi_\alpha\nu_\alpha) \circ u_\alpha^{-1} \circ \Phi)|\det(D\Phi)|.$$

But on $v_i(U_\alpha \cap V_i)$, we get $u_\alpha^{-1} \circ \Phi = v_i^{-1}$ and for all $x \in U_\alpha \cap V_i$ the transformation law (4.2) in Definition 4.1 shows that $\nu_i(x) = \nu_\alpha(x)|\det(D\Phi(v_i(x)))|$. Thus we conclude that the summands in (4.4) and (4.5) all coincide, which shows that the integral is well defined.

Linearity of the integral is obvious from the definition and to see that it is surjective, it suffices to construct one density with non-zero integral. This can easily be done via Example 4.1. Given a manifold M, we fix some chart (U, u) for M and a bump function φ with support contained in U. Then $\varphi^2 \sum_i (du^i \otimes du^i)$ is a smooth $\binom{0}{2}$-tensor field on U, which can be smoothly extended to all of M by zero. Applying the construction from example 4.1 we get a density ν on M with support contained in U. The coordinate representation of our tensor field with respect to U is $\varphi\delta_{ij}$ by construction. Since φ is non-negative, the function representing of ν in the chart (U, u) thus is $\sqrt{\varphi^{2n}} = \varphi^n$. But we already know that we can compute $\int_M \nu$ as $\int_{\mathbb{R}^n} \varphi^n \circ u^{-1}$, and this is evidently positive. □

Orientations and integration of differential forms

Integration of densities works very well, but apart from the case of (pseudo-)Riemannian geometry discussed in Example 4.1, there is not much connection to the concepts discussed so far. Such a connection is provided by integration theory for differential forms, in particular via Stokes's theorem. To integrate differential forms, one has to introduce an additional structure to deal with the problem of the absolute value that occurs in (4.1).

4.3. Orientations. The main observation needed to get integration theory for differential forms started is that on a manifold M of dimension n, the behavior of n-forms under a chart-change is very close to the one of densities. Indeed given a chart (U, u) for M and $\omega \in \Omega^n(M)$, the local coordinate representation from formula (3.14) simplifies to $\omega|_U = \omega_{1\dots n} du^1 \wedge \dots \wedge du^n$ so this is described by the single smooth function $\omega_{1\dots n}$. To obtain the behavior of this function under a change of charts, it is easier to argue directly than to specialize (3.5). By construction, $\omega_{1\dots n} = \omega(\frac{\partial}{\partial u^1}, \dots, \frac{\partial}{\partial u^n})$ so we can obtain the transformation law from formula (1.2) from Section 1.14 that applies to the coordinate vector fields. For two charts (U_α, u_α) and (U_β, u_β) with $U_{\alpha\beta} \neq \emptyset$, this shows that to pass from $\omega^\alpha_{1\dots n}$ to $\omega^\beta_{1\dots n}$ we have to hit each of the entries of ω with the linear isomorphism $Du_{\alpha\beta}(u_\beta(x))$. But the from linear algebra, it is known that for any n-linear, alternating map on an n-dimensional vector space, this causes a multiplication by the determinant. Thus we conclude that

$$\omega^\beta_{1\dots n}(x) = \omega^\alpha_{1\dots n}(x) \det(Du_{\alpha\beta}(u_\beta(x))), \tag{4.6}$$

so apart from the missing absolute value, this is like formula (4.2) for densities.

Hence we conclude that for an n-form with support contained in the domain of a chart, we have an integral, which is independent of the choice of chart up to its sign. The only question is now to deal with this sign issue, having done that, we can define an integral for compactly supported n-forms exactly as for densities in Section 4.2. In particular, there would be no problem, if we can work with an atlas such that the determinants of the derivatives of all chart changes are positive. To move towards this, recall from linear algebra that on a vector space V of dimension $n \geq 1$, one can introduce an equivalence relation on the set of ordered bases of V. One declares two ordered bases as equivalent if the matrix describing the change between the two bases has positive determinant. There are exactly two equivalence classes for this relation, and an *orientation* on V is given by choosing one of these two classes. Having made this choice, the bases in that class are called *positively oriented*, while the others are called *negatively oriented*. Given a smooth manifold M of dimension $n \geq 1$, we can now try to choose an orientation on each of the tangent spaces T_xM of M. There is a natural compatibility condition available in this setting as follows.

Definition 4.3. Let M be a smooth manifold of dimension $n \geq 1$.

(1) An *orientation* on M is given by a choice of orientation for each of the tangent spaces T_xM of M which is consistent in the following sense: For any connected open subset $U \subset M$ and local vector fields $\xi_1, \dots, \xi_n$ defined on U such that $\{\xi_1(x), \dots, \xi_n(x)\}$ is a basis of T_xM for each $x \in U$, all these bases have the same orientation (positive or negative).

(2) The manifold M is called *orientable* if it admits an orientation, a manifold endowed with an orientation is called an *oriented manifold*.

(3) An *oriented atlas* for M is an atlas $\mathcal{A}$ such that for any two charts (U_α, u_α) and (U_β, u_β) from $\mathcal{A}$ and each $x \in U_{\alpha\beta}$, we have $\det(Du_{\alpha\beta}(u_\beta(x))) > 0$.

(4) Two oriented atlases $\mathcal{A}$ and $\mathcal{B}$ of M are said to be *oriented equivalent* if and only if their union is an oriented atlas.

The consistency condition in (1) can be interpreted as saying that the orientations of the tangent spaces depend continuously on the base point. Indeed, it does not really depend on the specific family of vector fields. Given two families, say $\xi_1, \dots, \xi_n$ and $\eta_1, \dots, \eta_n$ as in the definition, there are smooth functions $a_{ij} : U \to \mathbb{R}$ such that $\eta_i = \sum_j a_{ij}\xi_j$. By construction the matrix $(a_{ij}(x))$ is invertible for each $x \in U$, and since U is connected, this implies that $\det(a_{ij}(x))$ is either positive for all x or negative for all x. So if the consistency condition is satisfied for $\xi_1, \dots, \xi_n$, then it is also satisfied for $\eta_1, \dots, \eta_n$. Having this observation at hand, we can nicely characterize orientability and describe the possible orientations.

Proposition 4.3. *Let M be a smooth manifold of dimension $n \geq 1$.*
(1) The following conditions are equivalent

- *(i) M is orientable.*
- *(ii) There exists a form $\omega \in \Omega^n(M)$ such that $\omega(x) \neq 0$ for all $x \in M$.*
- *(iii) M admits an oriented atlas.*

(2) If the conditions from (1) are satisfied, then the possible orientations of M are in bijective correspondence with the set of oriented equivalence classes of oriented atlases. If M is connected, then there are exactly two possible orientations, in general the number of orientations is 2^k, where k is the number of connected components of M.

PROOF. (1): We first show that (i) implies (iii). If M is orientable, we choose an orientation. Then by definition for a connected chart (U, u) for M, the ordered bases $\{\frac{\partial}{\partial u^1}(x), \dots, \frac{\partial}{\partial u^n}(x)\}$ for $x \in U$ are either all positively oriented or all negatively oriented. Depending on this, we call the chart (U, u) positively oriented or negatively oriented. Moreover, if (U, u) is negatively oriented, then we can just flip the sign of the first local coordinate u^1 to obtain a positively oriented chart. Starting from an atlas for M that consists of connected charts, we can therefore construct an atlas consisting of positively oriented charts. Observe that the notion of a positively oriented chart extends to disconnected charts without problems.

Now, if (U_α, u_α) and (U_β, u_β) are positively oriented charts, then by formula (1.2) from Section 1.4 the derivative $Du_{\alpha\beta}(u_\beta(x))$ just represents the base-change between the bases of T_xM coming from the coordinate vector fields. By assumption, these both are positively oriented, so $\det(Du_{\alpha\beta}(u_\beta(x))) > 0$ for each $x \in U_{\alpha\beta}$. Hence any atlas consisting entirely of positively oriented charts is an oriented atlas.

$(iii) \Rightarrow (ii)$: Let $\{(U_\alpha, u_\alpha) : \alpha \in I\}$ be an oriented atlas for M. Then by Theorem 1.9, there is a subordinate partition $\{\varphi_i : i \in \mathbb{N}\}$ of unity and for each i we choose α_i such that $\operatorname{supp}(\varphi_i) \subset U_{\alpha_i}$. Putting $\omega_i := \varphi_i du^1_{\alpha_i} \wedge \dots \wedge du^n_{\alpha_i} \in \Omega^n(U_{\alpha_i})$, we see that for $x \in U_{\alpha_i}$ we get $\omega_i(x) \neq 0$ if and only if $\varphi_i(x) \neq 0$. Hence $\operatorname{supp}(\omega_i) \subset U_{\alpha_i}$, so we can extend it by zero to a form defined on all of M. Moreover, the family $\operatorname{supp}(\omega_i)$ is locally finite, so $\omega := \sum_{i \in \mathbb{N}} \omega_i \in \Omega^n(M)$ is a well defined n-form on M.

Since we started from an oriented atlas, ω_i has the property that inserting the coordinate vector fields of any of the charts (U_α, u_α) into ω_i (in increasing order), the result will be non-negative. Moreover, for each $x \in M$, there exists an index i such that $\varphi_i(x) > 0$ and then $\omega_i(x)$ takes a strictly positive value on the coordinate vector fields for $(U_{\alpha_i}, u_{\alpha_i})$. But this shows that $\omega(x) \neq 0$.

$(ii) \Rightarrow (i)$: Suppose that $\omega \in \Omega^n(M)$ is nowhere vanishing. Then for each $x \in M$, inserting the elements of an ordered basis of T_xM into $\omega(x)$, one obtains a non-zero number, and we say that the basis is positively oriented if that number is positive. Since inserting local vector fields into ω one obtains smooth functions, the resulting orientations are clearly consistent, so we get an orientation on M is this way.

(2): As observed in (1), there is the notion of positively oriented charts on an oriented manifold M, by requiring that the coordinate vector fields form a positively oriented basis for the tangent space in each point. Moreover, from (1) we know that there is an oriented atlas $\mathcal{A}$ for M consisting of charts that all are positively oriented. Now it follows from the definitions that an oriented atlas $\mathcal{B}$ is oriented equivalent to $\mathcal{A}$ if and only if all its charts are positively oriented. This establishes the first claim in (2).

For the second claim, assume first that M is connected and fix an orientation. Then as in (1), we find $\omega \in \Omega^n(M)$ such that $\omega(x)$ is positive on positively oriented ordered bases for T_xM. Now let us keep the form ω but switch to a different orientation on M. Then there is at least one point $x \in M$ such that $\omega(x)$ takes a negative value on all positively oriented ordered bases for T_xM and we look at the set $A \subset M$ of all such points. For $y \in A$, we can take a connected chart (U, u) with $y \in U$ and its coordinate vector fields to see that $U \subset A$, so A is open in M. But $M \setminus A$ admits a similar description with positive values on positively oriented ordered bases, so this is open, too. Hence by connectedness $M = A$, so there is just one other possible orientation.

In the general case, it is clear that the connected components of M are connected manifolds, and from the definitions it follows easily that the choice of an orientation on M is equivalent to the choice of an orientation on each component (and there is no relation between the components). □

In some cases, orientations are derived from other structures. A nice example is provided by symplectic manifolds as discussed in Section 3.8. The basis for this is a result from multilinear algebra, namely that for a non-degenerate alternating bilinear map b on a real vector space V of dimension $2n$ the wedge-product of n copies of b is nonzero. (Indeed, this is an equivalent characterization of non-degeneracy.) Applying this point-wise, we readily see that for a symplectic manifold (M, ω) of dimension $2n$, the form $\omega^n := \omega \wedge \dots \wedge \omega \in \Omega^{2n}(M)$ is nowhere vanishing. This form determines a canonical orientation on M. Indeed, the condition that $d\omega = 0$ is not relevant for these considerations.

Remark 4.3. (1) Orientability is a global issue. On the domain of a chart, one always finds an orientation. Given orientations on open subsets U and V such that $U \cap V$ is connected, one may always swap the orientation on one of the two factors to obtain an orientation on $U \cup V$. In this way, one may extend orientations step by step. But examples like the Möbius band show that one may run into the following situation: $M = U \cup V$ for connected open subsets U and V such that $U \cap V$ has two connected components and for any choice of orientations on U and V, the orientations agree on one component of the intersection and disagree on the other. So there is no hope to find a global orientation in such cases.

(2) Orientability is not a very strong restriction in general. Given a non-orientable manifold M, there is an orientable manifold $\widehat{M}$ and a local diffeomorphism $q : \widehat{M} \to M$ such that $q^{-1}(\{x\})$ has exactly two elements for each $x \in M$. This falls into the theory of covering maps from algebraic topology, whence $\widehat{M}$ is called the *orientable covering* of M. This often allows to reduce problems to the orientable case.

The construction of $\widehat{M}$ roughly is as follows. For each $x \in M$, the n-linear alternating maps $(T_xM)^n \to \mathbb{R}$ form a one-dimensional vector space, that is usually denoted by $\Lambda^n T_x^*M$. Now we can remove zero and then identify maps which are positive multiples of each other to get "two points sitting over x". Now as in 3.1 and 3.3, we define $\widehat{M}$ as the disjoint union of these sets for all $x \in M$, which also defines $q : \widehat{M} \to M$. Then a chart (U, u) identifies $\Lambda^n T_x^*M$ with $\mathbb{R}$ for all points $x \in U$ and this leads to an identification of $q^{-1}(U)$ with $U \times \{1, -1\}$, and we can use u as a chart map on either of the two components. Still as in 3.1 and 3.3 one shows that things work out well with chart changes and one can even obtain an oriented atlas in this way, see the last part of chapter 15 of [**Lee**] for details.

(3) A nowhere-vanishing n-form on an n-dimensional manifold M is called a *volume-form* on M. In particular, the considerations from Example 4.1 show that on an oriented Riemannian n-manifold (M, g) there is a canonical volume form $\mathrm{vol}(g) \in \Omega^n(M)$. Explicitly, for an oriented chart (U, u) for M such that $g|_U = g_{ij} du^i \otimes du^j$, we get $\mathrm{vol}(g)|_U = \sqrt{\det(g_{ij})} du^1 \wedge \dots \wedge du^n$.

4.4. Integration of differential forms. Having the background on orientations at hand, we can now construct an integral for (compactly supported) n-forms completely parallel to the case of densities in Section 4.2. Consider an oriented manifold M and a form $\omega \in \Omega^n(M)$ with compact support. Applying Lemma 4.2 to some oriented atlas compatible with the orientation of M, we find finitely many positively oriented charts (U_α, u_α), $\alpha = 1, \dots, N$ such that $\mathrm{supp}(\omega) \subset \cup U_\alpha$ and functions φ_α with values in $[0, 1]$ such that $\mathrm{supp}(\varphi_\alpha) \subset U_\alpha$ and $\sum_\alpha \varphi_\alpha$ is identically one on $\mathrm{supp}(\omega)$. For each α, we have the function $\omega^\alpha_{1\dots n}:U_\alpha\to\mathbb{R}$ from the local coordinate representation of ω, and we define

$$\int_M\omega:=\sum_{\alpha=1}^N\int_{\mathbb{R}^n}(\varphi_\alpha\omega^\alpha_{1\dots n})\circ u_\alpha^{-1}. \tag{4.7}$$

(The summation convention does not apply to α in the right hand side.)

In contrast to the case of densities, where we haven't discussed the actions of diffeomorphisms at all, we want to prove compatibility of the integral with diffeomorphisms here. This needs another definition.

Definition 4.4. Let M and N be oriented smooth manifolds and let $F:M\to N$ be a diffeomorphism. Then we call F *orientation preserving* if for each $x\in M$ the tangent map $T_xF:T_xM\to T_{F(x)}N$ is orientation preserving. We call F *orientation reversing* if T_xF is orientation reversing for each $x\in M$.

If M and N are connected, then any diffeomorphism $F:M\to N$ is either orientation preserving or orientation reversing, in general there may be different behavior on different connected components. Now we can formulate the main result on integration of differential forms:

Theorem 4.4. *(1) On an oriented manifold M, the expression in* (4.7) *is independent of the choice of the (positively oriented) charts (U_α,u_α) and of the functions φ_α. Thus we obtain a well defined integral, which defines a surjective linear map from the space of compactly supported n-forms to $\mathbb{R}$.*

(2) Let $F:M\to N$ be a diffeomorphism between oriented smooth manifolds of dimension n, which is either orientation preserving or orientation reversing. Then for any compactly supported form $\omega\in\Omega^n(N)$ we get $\int_M F^*\omega=\int_N\omega$ if F is orientation preserving and $\int_M F^*\omega=-\int_N\omega$ if F is orientation reversing.

Proof. (1) Since the derivative of the chart change between two positively oriented charts always has positive determinant, we can use exactly the arguments as in the proof of Theorem 4.2 to prove that the integral is well defined. Linearity then is obvious, and for surjectivity, we can use $\varphi du^1\wedge\dots\wedge du^n$ for a positively oriented chart (U,u) and a bump function φ with compact support contained in U.

(2) Note that $\operatorname{supp}(F^*\omega)=F^{-1}(\operatorname{supp}(\omega))$ which is compact since F is a homeomorphism, so it is no problem to form $\int_M F^*\omega$. Given ω, we choose positively oriented charts (U_α,u_α) for N and functions φ_α to compute $\int_N\omega$ as in (4.7). If F is orientation preserving, then for each α, $(F^{-1}(U_\alpha),u_\alpha\circ F)$ is a positively oriented chart for M and $F^*du^i_\alpha=d(u^i_\alpha\circ F)$. This readily shows that $\omega|_{U_\alpha}=\omega^\alpha_{1\dots n}du^1_\alpha\wedge\dots\wedge du^n_\alpha$ implies that $F^*\omega|_{F^{-1}(U_\alpha)}=(\omega^\alpha_{1\dots n}\circ F)d(u^1_\alpha\circ F)\wedge\dots\wedge d(u^n_\alpha\circ F)$. Since $(u_\alpha\circ F)^{-1}=F^{-1}\circ u_\alpha^{-1}$ we readily conclude from (4.7) that $\int_M F^*\omega=\int_N\omega$.

If F is orientation reversing, then we can, for each α, obtain a positively oriented chart from $(F^{-1}(U_\alpha),u_\alpha\circ F)$ by swapping the sign of one coordinate. This readily shows that the coordinate representation of $F^*\omega$ with respect to the resulting chart is given by the function $-\omega^\alpha_{1\dots n}\circ F$ and using this, the claim follows as in the orientation-preserving case. □

Manifolds with boundary and Stokes's theorem

A key feature of integration theory for differential forms is that there is an interplay between integration and the exterior derivative. For usual manifolds, this boils down to the fact that for an oriented manifold M of dimension n and a form $\omega\in\Omega^{n-1}(M)$ with compact support, one has $\int_M d\omega=0$. The simplest version of this is $M=\mathbb{R}$, where $\int_{\mathbb{R}} df = 0$ for a compactly supported smooth function f follows easily from the fundamental theorem of calculus. But in this case, there is a better version, in which one integrates over $[a, b]$ rather then $\mathbb{R}$ and then $\int_a^b df = f(b) - f(a)$, so the integral is determined by a "boundary term".

There is a general version of this, which however requires an extension of the notion of a smooth manifold, which allows for a boundary. The necessary generalization are not too difficult, so we will discuss this rather briefly.

4.5. Manifolds with boundary. For $n \geq 1$, define the n-dimensional *half-space* $\mathcal{H}^n$ as $\{x \in \mathbb{R}^n : x^1 \leq 0\}$. (The choice of the first coordinate and of non-positive values is just a convention, this choice avoids the occurrence of signs in Stokes's theorem.) We'll refer to points with $x^1 < 0$ as *interior points* of $\mathcal{H}^n$ and to points with $x^1 = 0$ as *boundary points*. This terminology obviously extends to open (for the induced topology) subsets $U \subset \mathcal{H}^n$, but here it may happen that U entirely consists of interior points.

It is no problem to defines smoothness for a map $F : U \to V$ between open subsets of half spaces. One requires that for each $x \in U$, there is a an open subset $\tilde{U} \subset \mathbb{R}^n$ and a smooth function $\tilde{F} : \tilde{U} \to \mathbb{R}^n$, which coincides with F on $U \cap \tilde{U}$. Hence it is no problem to consider diffeomorphisms between such open subsets, and the inverse function theorem immediately implies that such a diffeomorphism maps interior points to interior points and boundary points to boundary points. (Results from algebraic topology imply that indeed the same is true for homeomorphisms between open subsets of half spaces.)

Now one defines a *topological manifold with boundary* of dimension n as a second countable Hausdorff space M such that each point $x \in M$ has an open neighborhood in M that is homeomorphic to an open subset of $\mathcal{H}^n$. Then there are obvious analogs of all the further notions from Definition 1.6 based on charts (U_α, u_α) for which $u_\alpha(U_\alpha)$ is open in $\mathcal{H}^n$. This leads to the notion of a *smooth manifold with boundary.* From above, we conclude that if $u_\alpha(x)$ is a boundary point for one chart (U_α, u_α) with $x \in U_\alpha$, then the same holds for any compatible chart. Hence on a smooth manifold with boundary, there is a well defined notion of *interior points* and of *boundary points*. The set of all boundary points of M is denoted by $\partial M \subset M$ and called the *boundary* of M.

Observe that the set of interior points of $\mathcal{H}^n$ clearly is diffeomorphic to $\mathbb{R}^n$, so any open subset of $\mathbb{R}^n$ is diffeomorphic to an open subset of $\mathcal{H}^n$ that consists of interior points only. Thus the notions from Section 1.6 are the special cases of the ones developed here in which $\partial M = \emptyset$, so any manifold is a manifold with boundary. Of course, an open subset $U \subset \mathcal{H}^n$ is a manifold with boundary, for which ∂U is the set of boundary points. Another obvious example of a manifold with boundary is $M := \{x \in \mathbb{R}^n : |x| \leq 1\}$ for which ∂M is the sphere S^{n-1}.

Next, the notions of smoothness of functions on and maps between manifolds with boundary can be defined via local coordinate representations as for manifolds in Section 1.8. Observe that by our definition of smoothness on open subsets of $\mathcal{H}^n$ this means that one always requires existence of smooth extensions to open neighborhoods of boundary points in $\mathbb{R}^n$. Theorem 1.9 on partitions of unity extends to manifolds with boundary without problems.

For an open subset $U \subset \mathcal{H}^n$, the set of boundary points of U is an open subset of $\mathbb{R}^{n-1}$ (embedded into $\mathbb{R}^n$ as points with zero first coordinate). Hence for a chart (U_α, u_α) for M, we can consider the open subset $U_\alpha \cap \partial M$ of ∂M and $u_\alpha|_{U_\alpha \cap \partial M}$ is a homeomorphism onto the set of boundary points of $u_\alpha(U_\alpha)$. This easily implies that ∂M is a topological manifold in the sense of Definition 1.6. Moreover, any diffeomorphism of open subsets of $\mathcal{H}^n$ restricts to a diffeomorphism between the sets of boundary points. This easily implies that starting from an atlas for M, one obtains compatible charts and hence an atlas for ∂M. It is easy to see that the resulting smooth structure on ∂M is independent of the choice of atlas, so ∂M is canonically a smooth manifold (without boundary). By construction, the inclusion $i : \partial M \to M$ is a smooth map, so for smooth functions or maps defined on M, the restrictions to ∂M are smooth, too.

The concept of germs of smooth functions extends to manifolds with boundary without problems. Parallel to Lemma 1.10, one easily shows that for any $a \in \mathcal{H}^n$ one can canonically identify the space of derivations at a of germs of smooth functions with $\mathbb{R}^n$. (For interior points, this is clear, but by definition of smoothness, it extends to boundary points without problems.) Using this, one simply extends the definition of tangent spaces, all results from Section 1.12, and the coordinate description from Section 1.14 to manifolds with boundary.

Here a new feature arises in a boundary point $x \in \partial M$. Suppressing the inclusion from the notation, we can view $T_x\partial M$ as a linear subspace of T_xM. In view of our conventions for half spaces, for any chart (U, u) around x, this linear subspace is spanned by the tangent vectors $\frac{\partial}{\partial u^i}|_x$ for $i = 2, \dots, n$. Moreover, vectors in $T_xM \setminus T_x\partial M$ (which are automatically non-zero) can be either *inward pointing* or *outward pointing*. Writing such a vector as a linear combination of the basis elements $\frac{\partial}{\partial u^i}|_x$ determined by a chart (U, u), then according to our convention, the inward pointing vectors are those, for which the coefficient of $\frac{\partial}{\partial u^1}|_x$ is negative.

For an open subset $U \subset \mathcal{H}^n$, we can still identify the union of all tangent spaces with $U \times \mathbb{R}^n$, which is an open subset in $\mathcal{H}^{2n}$. Using this, one can collect the tangent spaces of a manifold M with boundary into a set TM endowed with an obvious projection $p : TM \to M$ and make this into a manifold with boundary $p^{-1}(\partial M)$. Using this, one can define vector fields, tensor fields and differential forms on manifolds with boundary parallel to what we have done in Chapters 2 and 3 with some minor changes. What one certainly has to be careful about is domains of definitions of flows. These are influenced by conditions like the well defined concepts that a vector field $\xi \in \mathfrak{X}(M)$ is tangent to ∂M or inward pointing along ∂M, and so on. What we mainly will need here is the exterior derivative, which extends without problem. Similarly as discussed in Example 3.6, we observe that for $\omega \in \Omega^k(M)$, one has $i^*\omega \in \Omega^k(\partial M)$, and one usually suppresses i^* from the notation and simply views ω as a form on M and on ∂M at the same time. Finally, the concepts of orientations and of oriented atlases makes sense for manifolds with boundary, and the following observation is crucial for integration theory:

Lemma 4.5. *Suppose that M is an orientable manifold with boundary of dimension $n \geq 2$. Then any choice of orientation on M gives rise to an orientation of the manifold ∂M, so ∂M is orientable, too. This correspondence has the property that for a positively oriented chart (U, u) for M, also the induced chart $(U \cap \partial M, u|_{U\cap\partial M})$ is positively oriented.*

Proof. Let us consider two charts (U_α, u_α) and (U_β, u_β) for M and a point $x \in U_{\alpha\beta} \cap \partial M$ (so we assume that this intersection is non-empty). Then the chart change $u_{\alpha\beta} : u_\beta(U_{\alpha\beta}) \to \mathcal{H}^n$ maps boundary points to boundary points. In particular, the derivative $Du_{\alpha\beta}(u_\beta(x))$ has to be of the block form $\begin{pmatrix} \lambda & 0 \\ v & A \end{pmatrix}$ with block sizes 1 and $n-1$, so $\lambda \in \mathbb{R}$ and $v \in \mathbb{R}^{n-1}$ and A is the derivative of the chart change between the two induced charts at x. But by assumption $u_{\alpha\beta}$ also has to send points with negative first coordinate to points with negative first coordinate, so $\lambda > 0$. Thus $\det(Du_{\alpha\beta}(x))$

and $\det(A)$ have the same sign. Hence for an oriented atlas for M the induced atlas for ∂M is oriented and for two such atlases which are oriented equivalent, the induced atlases are oriented equivalent, too. By Proposition 4.3, this implies that an orientation on M induces an orientation on ∂M, an then the last claimed property is satisfied by construction. □

It is not difficult to describe the induced orientation on ∂M explicitly: We have the linear subspace $T_x\partial M \subset T_xM$ and of course adding any vector in $T_xM \setminus T_x\partial M$ to a basis of $T_x\partial M$, one obtains a basis of T_xM. Then from the construction in the proof we conclude that an ordered basis of $T_x\partial M$ is positively oriented if and only if adding an outward pointing vector in $T_xM \setminus T_x\partial M$ as the first element, one obtains a positively oriented basis of T_xM.

4.6. Stokes's theorem. Given a manifold M with boundary, let us write $\Omega^k_c(M)$ for the space of k-forms on M with compact support. By part (1) of Theorem 3.6, the exterior derivative is a local operator, so if $\omega \in \Omega^k(M)$ vanishes identically on some open subset $U \subset M$, then also $d\omega|_U$ vanishes identically. This implies that $\operatorname{supp}(d\omega) \subset \operatorname{supp}(\omega)$, so d maps $\Omega^k_c(M)$ to $\Omega^{k+1}_c(M)$. In particular, for $\omega \in \Omega^{n-1}_c(M)$, we have $d\omega \in \Omega^n_c(M)$, so if M is oriented, we can form $\int_M d\omega \in \mathbb{R}$. But on the other hand, we can restrict ω to ∂M, which is also oriented, and then form $\int_{\partial M} \omega$. Hence we can formulate Stokes's theorem:

Theorem 4.6 (Stokes). *For an oriented smooth manifold M with boundary ∂M and any form $\omega \in \Omega^{n-1}_c(M)$, we have $\int_M d\omega = \int_{\partial M} \omega$. In case of a manifold without boundary, we get $\int_M d\omega = 0$ for any $\omega \in \Omega^{n-1}_c(M)$.*

Proof. As observed in Section 4.4, there are finitely many positively oriented charts (U_α, u_α), $\alpha = 1, \dots, N$ such that $\operatorname{supp}(\omega) \subset \cup_{\alpha=1}^N U_\alpha$ and smooth functions φ_α such that $\operatorname{supp}(\varphi_\alpha) \subset U_\alpha$ and $\sum_\alpha \varphi_\alpha$ is identically one on $\operatorname{supp}(\omega)$. As we have noted above $\operatorname{supp}(d(\varphi_\alpha\omega)) \subset \operatorname{supp}(\varphi_\alpha\omega) \subset U_\alpha$ and by construction $d\omega = \sum_{\alpha=1}^N d(\varphi_\alpha\omega)$. Linearity of the integral then shows that $\int_M d\omega = \sum_{\alpha=1}^N \int_M d(\varphi_\alpha\omega)$. On the other hand denoting by $i : \partial M \to M$ the inclusion we see that for $i^*\omega \in \Omega^{n-1}(\partial M)$ we get $\operatorname{supp}(i^*\omega) \subset \operatorname{supp}(\omega) \cap \partial M$. Hence this is contained in $\cup_\alpha \partial U_\alpha$, where $\partial U_\alpha := U_\alpha \cap \partial M$ and $\sum_{\alpha=1}^N (\varphi_\alpha|_{\partial M})$ is identically one on $\operatorname{supp}(i^*\omega)$. By Lemma 4.5, the charts $(\partial U_\alpha, u_\alpha|_{\partial U_\alpha})$ are positively oriented for the induced orientation on ∂M, so $\int_{\partial M} \omega = \sum_{\alpha=1}^N \int_{\partial M} (\varphi_\alpha|_{\partial M}) i^*\omega$.

This shows that it suffices to prove that $\int_M d\omega = \int_{\partial M} i^*\omega$ in the case that the support of ω is contained in one positively oriented chart (U, u) for M, and we restrict to this case from now on. Slightly simplifying the notation for the local coordinate expression for the rest of this proof, we get

$$\omega|_U = \textstyle\sum_{k=1}^n \omega_k du^1 \wedge \dots \wedge \widehat{du^k} \wedge \dots \wedge du^n.$$

Now we know that for $x \in \partial U$, a tangent vector X_x lies in the subset $T_x\partial M \subset T_xM$ if and only if expanding it in terms of the $\frac{\partial}{\partial u^i}|_x$, the coefficient of $\frac{\partial}{\partial u^1}|_x$ is zero. This shows that $i^*du^1 = 0$ and hence $i^*\omega = \omega_1 du^2 \wedge \dots \wedge du^n$. So by definition, we get $\int_{\partial M} i^*\omega = \int_{u(U)\cap\mathbb{R}^{n-1}} \omega_1 \circ u^{-1}$. Observe that the integral remains unchanged if we replace the domain of integration by all of $\mathbb{R}^{n-1}$.

Applying part (4) of Theorem 3.6, we also see that

$$d\omega|_U = \textstyle\sum_{k=1}^n (-1)^{k-1} \frac{\partial}{\partial u^k}(\omega_k) du^1 \wedge \dots \wedge du^n$$

(with the sign coming from the fact that we have moved du^k to the kth position in the wedge product). Hence by definition we get

$$\int_M d\omega = \sum_{k=1}^n (-1)^{k-1} \int_{u(U)} \partial_k(\omega_k \circ u^{-1}).$$

Now by construction, the integral remains unchanged if we replace the domain of integration by $\mathcal{H}^n = (-\infty, 0] \times \mathbb{R}^{n-1}$. Then by Fubini's theorem, the integral over $\mathcal{H}^n$ can be computed as an iterated integral over the individual variables and the result is independent of the order in which the integrals are performed. Now for the summand involving $\partial_k(\omega_k \circ u^{-1})$, we first integrate over the kth variable. For $k > 1$, this gives us an integral of the form

$$\int_{-\infty}^{\infty} \tfrac{\partial}{\partial t}(\omega_k \circ u^{-1})(y_1, \dots, y_{k-1}, t, y_{k+1}, \dots, y_n) dt.$$

Since the function in the interior has compact support, this may be computed as an integral over some large closed interval and then the integral vanishes by the fundamental theorem of calculus. Hence all these summands to not contribute to the integral. For $k = 1$, we get the integral

$$\int_{-\infty}^{0} \tfrac{\partial}{\partial t}(\omega_1 \circ u^{-1})(t, y_2, \dots, y_n) dt = (\omega_1 \circ u^{-1})(0, y_2, \dots, y_n).$$

Thus we conclude that $\int_M d\omega = \int_{\{0\}\times\mathbb{R}^{n-1}} \omega_1 \circ u^{-1}$, which we know from above coincides with $\int_{\partial M} i^*\omega$. □

Even the version of Stokes's theorem for manifolds without boundary has very interesting applications. Suppose that (M, ω) is a symplectic manifold of dimension $2n$ as in Section 3.8. As we have seen in Section 4.3, the form $\omega^n = \omega \wedge \dots \wedge \omega \in \Omega^{2n}(M)$ is nowhere vanishing and hence determines an orientation on M. If M is compact, then we can form $\int_M \omega^n$ and since the coordinate representation of ω^n with respect to any oriented chart is a positive function, $\int_M \omega^n > 0$. By Stokes's theorem and compactness of M, we conclude that there cannot be any form $\alpha \in \Omega^{2n-1}(M)$ such that $\omega^n = d\alpha$. But now suppose that the symplectic structure ω were exact, i.e. $\omega = d\beta$ for some $\beta \in \Omega^1(M)$. Then we can define $\alpha := \beta \wedge \omega^{n-1} = \beta \wedge \omega \wedge \dots \wedge \omega$ and the compatibility of d with the wedge product from Theorem 3.6 shows that $d\alpha = d\beta \wedge \omega^{n-1} = \omega^n$, which is a contradiction. Hence in contrast to the example of cotangent bundles discussed in Section 3.8, a symplectic structure on a compact manifold can never be exact.

Remark 4.6. Initially, integration and Stokes's Theorem applies only to top degree forms on a manifold, but it can be easily used to study lower degree forms. For example let S be a compact, oriented smooth k-dimensional manifold without boundary and let $i : S \to M$ be an embedding into some n-dimensional manifold. Then for $\omega \in \Omega^k(M)$, we can form $i^*\omega \in \Omega^k(S)$ and then $\int_S i^*\omega \in \mathbb{R}$. Now if there is some $\tau \in \Omega^{k-1}(M)$ such that $\omega = d\tau$ then $i^*\omega = i^*d\tau = di^*\tau$ by Theorem 3.6 and hence $\int_S i^*\omega$ vanishes by Stokes's theorem. Conversely, if for given ω we can find some S such that $\int_S i^*\omega \neq 0$, we can conclude that ω cannot be of the form $d\tau$ for some $\tau \in \Omega^{k-1}(M)$.

4.7. Excursion: vector analysis and classical integral theorems. The classical operations of gradient, divergence and rotation for vector fields on open subsets of $\mathbb{R}^3$ actually are all instances of the exterior derivative in disguise. Realizing this, the general version of Stokes's theorem in Theorem 4.6 easily leads to the classical integral theorems from analysis. Some of these ideas admit generalizations, often related to Riemannian geometry and we'll discuss things in this general setting.

As we have observed in Example 3.3, a Riemannian metric g on a smooth manifold M gives rise to an isomorphism between vector fields and one-forms on M, which sends

$\xi \in \mathfrak{X}(M)$ to $\eta \mapsto g(\xi, \eta)$. In particular, given a smooth function $f \in C^\infty(M, \mathbb{R})$, there is a unique vector field $\operatorname{grad}(f) \in \mathfrak{X}(M)$, called the *gradient of* f, such that $g(\operatorname{grad}(f), \eta) = df(\eta)$. In particular, this can be done on an open subset $U \subset \mathbb{R}^n$ with respect to metric given by the standard inner product on $\mathbb{R}^n \cong T_xU$ for each $x \in U$. Then $\operatorname{grad}(f)$ has components $\partial_i f$, so this is just the classical gradient, which thereby is derived from the exterior derivative.

On the other hand, suppose that M is an n-dimensional manifold, and we fix a top-degree form $\omega \in \Omega^n(M)$. This obviously gives rise to a map from $C^\infty(M, \mathbb{R})$ to $\Omega^n(M)$ by sending f to $f\omega$. Moreover, for $\xi \in \mathfrak{X}(M)$, we can form $i_\xi\omega \in \Omega^{n-1}(M)$. Both this operations are point-wise and it follows from multilinear algebra that they define isomorphisms at a point x provided that $\omega(x) \neq 0$. Hence if ω is a volume form on M (i.e. $\omega(x) \neq 0$ for all $x \in M$), it leads to an identification of $C^\infty(M, \mathbb{R})$ with $\Omega^n(M)$ and of $\mathfrak{X}(M)$ with $\Omega^{n-1}(M)$. In particular, for $\xi \in \mathfrak{X}(M)$, there is a unique function $\operatorname{div}(\xi) \in C^\infty(M, \mathbb{R})$, called the *divergence of* ξ such that $di_\xi\omega = \operatorname{div}(\xi)\omega$.

Recall from Section 4.5 that on an oriented Riemannian manifold, there is a canonical volume form $\operatorname{vol}(g) \in \Omega^n(M)$ associated to g. Hence any vector field on an oriented Riemannian manifold has a divergence, which can be interpreted as a smooth function. In the case of an open subset $U \subset \mathbb{R}^n$ and the metric coming from the inner product, this becomes easy: The volume form in this case is just $dx^1 \wedge \dots \wedge dx^n$ (which gives the determinant on each tangent space). Hence for a vector field $\sum \xi^i \frac{\partial}{\partial x^i}$, $i_\xi \operatorname{vol}(g)$ is given by

$$\textstyle\sum_k (-1)^{k-1}\xi^k dx^1 \wedge \dots \wedge \widehat{dx^k} \wedge \dots \wedge dx^n.$$

Thus $\operatorname{div}(\xi) = \sum_k \frac{\partial \xi^k}{\partial x^k}$, which is the classical divergence. In this setting, Stokes's theorem leads to the *divergence theorem*, which says that the integral of a divergence over a compact oriented Riemannian manifold without boundary always vanishes.

In dimensions $n = 2$ and $n = 3$, this already allows us to identify all the spaces $\Omega^k(M)$ on an oriented Riemannian manifold M of dimension n with either $C^\infty(M, \mathbb{R})$ or with $\mathfrak{X}(M)$. For $n = 2$, we have encoded the available exterior derivatives into $f \mapsto \operatorname{grad}(f)$ and $\xi \mapsto \operatorname{div}(\xi)$. Now this may suggest that $d \circ d = 0$ implies $\operatorname{div}(\operatorname{grad}(f)) = 0$ for any $f \in C^\infty(M, \mathbb{R})$, which of course isn't true (since this gives the Laplacian of f). The reason for this is that we have used two different identifications of $\mathfrak{X}(M)$ with $\Omega^1(M)$ which send ξ to $g(\xi, \_)$ and $i_\xi \operatorname{vol}(g)$, respectively. In particular, for an open subset $U \subset \mathbb{R}^2$, $\xi^1 \frac{\partial}{\partial x^1} + \xi^2 \frac{\partial}{\partial x^2}$ corresponds to $\xi^1 dx^1 + \xi^2 dx^2$ in the first identification and to $-\xi^2 dx^1 + \xi^1 dx^2$ in the second identification.

In dimension 3, this does not arise, but we have the additional exterior derivative $d : \Omega^1(M) \to \Omega^2(M)$ to interpret. Given $\xi \in \mathfrak{X}(M)$, there exists a unique vector field $\operatorname{rot}(\xi) \in \mathfrak{X}(M)$ such that $d(g(\xi, \_)) = i_{\operatorname{rot}(\xi)}(\operatorname{vol}(g))$. In the case of an open subset $U \subset \mathbb{R}^3$, one immediately derives the usual formula related to the cross-product. Here the fact that $d \circ d = 0$ immediately leads to $\operatorname{rot}(\operatorname{grad}(f)) = 0$ and $\operatorname{div}(\operatorname{rot}(\xi)) = 0$. As in dimension 2, $\operatorname{div}(\operatorname{grad}(f))$ gives the Laplacian of f.

To discuss an example of a classical integral theorem, consider a compact manifold M of dimension 3 with boundary ∂M, which is contained in some open subset $U \subset \mathbb{R}^3$. Then $T_xM = \mathbb{R}^3$ for each $x \in M$ and hence we get an orientation on M and the induced orientation on ∂M. Restricting the volume form $\operatorname{vol}(g)$ to M causes no problem, and the inner product on $\mathbb{R}^3$ also gives rise to a volume form (or area form) on ∂M, which is classically denoted by dA. What one requires is that this form sends a positively oriented orthonormal basis to $1 \in \mathbb{R}$, which is well defined and pins down the form uniquely. The area form can most easily be described in terms of an outward pointing unit normal $\mathfrak{n}$ for ∂M, i.e. the function $\mathfrak{n} : \partial M \to \mathbb{R}^3$, which assigns to each $x \in \partial M$ the outward pointing vector of length one that is orthogonal to $T_x\partial M \subset \mathbb{R}^3$. This turns out to be smooth and for $x \in \partial M$ one obtains $dA(x)$ by inserting $\mathfrak{n}(x)$ into $\mathrm{vol}(g)(x)$.

Now take a vector field $\xi \in \mathfrak{X}(U)$. As discussed above, this gives rise to $\omega = i_\xi \mathrm{vol}(g) \in \Omega^2(U)$, which can be restricted to both M and ∂M, and by Theorem 4.6, $\int_M d\omega = \int_{\partial M} \omega$. Interpreting the integral as being defined on functions, the left hand side by definition gives $\int_M \mathrm{div}(\xi)$. For the right hand side, consider a point $x \in \partial M$, and split $\xi(x) = \langle \xi(x), \mathfrak{n}(x)\rangle \mathfrak{n}(x) + Y$ with $Y \in T_x\partial M$. Inserting Y into $\mathrm{vol}(g)(x)$, one obtains a form that evidently vanishes on $T_x\partial M \times T_x\partial M$. But this shows that the restriction of $\omega(x)$ to $T_x\partial M \times T_x\partial M$ coincides with $\langle \xi(x), \mathfrak{n}(x)\rangle dA$. Hence we obtain $\int_{\partial M} \langle \xi, \mathfrak{n}\rangle dA = \int_M \mathrm{div}(\xi)$, which is called *Gauß' theorem*.

In physics terms, ξ is interpreted as describing a field, and the theorem says that the "normal flow of ξ through ∂M" can be computed via the integral of $\mathrm{div}(\xi)$ over M. Vanishing of the boundary integral interpreted as the fact that there are no "sources" of this field inside M. From above, we conclude that this always is the case if $\xi = \mathrm{rot}(\eta)$ for some $\eta \in \mathfrak{X}(U)$. On the other hand, if $\xi = \mathrm{grad}(f)$ (which means that the field can be described by a potential, compare to Section 3.9), this is the case if $\Delta(f) = 0$. There is a (simpler) analog for compact 2-dimensional manifolds with boundary contained in an open subset $U \subset \mathbb{R}^2$, which is called *Green's theorem.*

There is also a result in vector analysis called Stokes's theorem. This deals with a compact oriented manifold M of dimension 2 with boundary ∂M contained in an open subset $U \subset \mathbb{R}^3$. For the applications of this result in physics it is important the the vector fields we are talking about are defined on all of U, while M should rather be viewed as a "test surface". Thus we also use the identifications of differential forms with vector fields or functions coming from the (three-dimensional) open subset U and not from the (two-dimensional) manifold M. Observe that by Theorem 3.6, restricting differential forms to M is compatible with the exterior derivative.

Given a vector field $\xi \in \mathfrak{X}(U)$, we get the one-form $\omega = g(\xi, \ ) \in \Omega^1(U)$ and by Theorem 4.6 we get $\int_M d\omega = \int_{\partial M} \omega$. In the simplest case, $\partial M \cong S^1$, so this can be interpreted as a closed, regular curve in $\mathbb{R}^3$, and the right hand side leads to the classical version of a loop integral. For a parametrization $c : [a, b] \to \partial M$ (for which $c(a)$ is the only double point), this can be written as $\int_a^b \langle \xi(c(t)), c'(t)\rangle dt$. In physics terms, this is interpreted as the energy needed to "move a test particle through the field" along the closed curve. From above, we know that $d\omega = i_{\mathrm{rot}(\xi)} \mathrm{vol}(g)$ and, as above, we conclude that the loop integral can be computed as $\int_M \langle \mathrm{rot}(\xi), \mathfrak{n}\rangle dA$. In particular, this always vanishes for fields that can be described by a potential, i.e. if $\xi = \mathrm{grad}(f)$. Indeed the relation to closed curves in this theorem is why the operator rot got its name. In this setting, one obtains simple consequences of the result, which are very interesting. For example, the integrals of $\langle \mathrm{rot}(\xi), \mathfrak{n}\rangle dA$ over two surfaces with the same boundary have to agree.

Excursion: De-Rham cohomology

The last topic we discuss in the course provides a connection between differential forms and the topology of manifolds. Here one should think about "topology" as properties of manifolds that are unchanged by "deformations", the difference between smoothness and continuity is not as important initially. While from the discussion here it should be visible that one obtains relatively robust notions, much more background is needed to see that there is an actual connection to algebraic topology, which in addition concerns fairly advanced parts of that field.

4.8. Basic notions. The basis for de-Rham cohomology is the fact that $d^2 = d \circ d = 0$ as observed in Theorem 3.6. The standard terminology here is to call a form $\omega \in \Omega^k(M)$ *closed* if $d\omega = 0$ and *exact* if it is of the form $d\tau$ for some $\tau \in \Omega^{k-1}(M)$. In this language, $d^2 = 0$ says that any exact form is closed. By linearity of d, both the sets $Z^k(M)$ of closed k-forms and the set $B^k(M)$ of exact k-forms are linear subspaces of $\Omega^k(M)$, and $B^k(M) \subset Z^k(M)$. Thus we can form the quotient vector space $H^k(M) := Z^k(M)/B^k(M)$, which is called the kth *de-Rham cohomology* of M. One then puts $H^*(M) := \oplus_{k=0}^n H^k(M)$ and calls this the *total de-Rham cohomology of* M.

The best way to view $H^k(M)$ is as a space of equivalence classes $[\omega]$ for $\omega \in \Omega^k(M)$ with $d\omega = 0$, where $[\tilde{\omega}] = [\omega]$ if and only if there is a form $\alpha \in \Omega^{k-1}(M)$ such that $\tilde{\omega} = \omega + d\alpha$. Addition and scalar multiplication on $H^k(M)$ is simply defined via representatives. Using this language, we can easily sort out some basic properties:

Proposition 4.8. *Let M be a smooth manifold and let $\omega \in Z^k(M)$ and $\tau \in Z^\ell(M)$ be closed forms.*

(1) Putting $[\omega] \wedge [\tau] := [\omega \wedge \tau]$ induces a well defined map $H^k(M) \times H^\ell(M) \to H^{k+\ell}(M)$. This defines a multiplication on $H^*(M)$, making it into an associative algebra which is graded commutative in the sense that $[\tau] \wedge [\omega] = (-1)^{k\ell}[\omega] \wedge [\tau]$.

(2) For another smooth manifold N, and a smooth map $F : N \to M$, putting $F^\#([\omega]) := [F^*\omega]$ defines an algebra homomorphism $F^\# = H^*(F) : H^*(M) \to H^*(N)$. If F is a diffeomorphism, then this is an isomorphism of algebras.

Proof. (1) Since d is a graded derivation (part (2) of Theorem 3.6), we see that $\omega \wedge \tau \in Z^{k+\ell}(M)$, so it makes sense to form $[\omega \wedge \tau]$. This also shows that for $\alpha \in \Omega^{k-1}(M)$ we get $d\alpha \wedge \tau = d(\alpha \wedge \tau)$, which together with bilinearity of the wedge product shows that $[\omega \wedge \tau]$ depends only on $[\omega]$. Similarly it depends only on $[\tau]$, so we get a well defined map as claimed. By Theorem 3.5 is is bilinear and hence makes $H^*(M)$ into an algebra, whose properties follow directly from Theorem 3.5.

(2) By part (5) of Theorem 3.6, pullbacks of closed forms are closed and pullbacks of exact forms are exact, so there is a well defined linear map $F^\#$ between the quotient spaces. Compatibility of the wedge product with pullbacks shows that this defines a homomorphism of algebras. If F is a diffeomorphism, then F^{-1} induces an inverse to this homomorphism, which completes the argument. □

The first indication that cohomology is robust and may be related to topology is what happens in degree zero. By definition, $H^0(M) = \{f \in C^\infty(M, \mathbb{R}) : df = 0\}$, so in local coordinates, these are the functions for which all partial derivatives vanish identically. Of course, this means that they have to be constant in a neighborhood of each point of M and hence on any connected component of M. Conversely, if f is constant on each connected component of M, then $df = 0$, so $H^0(M) = \mathbb{R}^{b_0(M)}$, where $b_0(M)$ is the number of connected components of M. If M has several connected components, each of them is a smooth manifold M_i and M is the disjoint union $\sqcup_{i \in I} M_i$ of the M_i. It is then easy to see that $H^k(M)$ is the product $\prod_{i \in I} H^k(M_i)$ for each k (so if I is finite, this coincides with the direct sum).

Generalizing the number $b_0(M)$, one defines $b_k(M) \in \mathbb{N} \cup \{\infty\}$ to be the dimension of $H^k(M)$ and calls it the kth *Betti number* of M. It turns out these Betti numbers are often finite, in particular, it turns out that they are always finite if M is compact, see Section 4.11.

As an example for an application of de-Rham cohomology, let us consider a compact manifold M of even dimension $2n$. If there exists a symplectic form $\omega \in \Omega^2(M)$ on M, then $d\omega = 0$, so we can consider $[\omega] \in H^2(M)$. From Section 4.6 we know that ω cannot be exact, so $[\omega]$ has to be non-zero. This immediately shows that a compact manifold M such that $H^2(M) = \{0\}$ cannot admit any symplectic structure. In particular, it turns out that for the sphere S^k, the only nontrivial cohomology groups are H^0 and H^k, so for $n \geq 2$, the sphere S^{2n} does not admit a symplectic structure. This can be refined, since from Section 4.6 we known that $\omega^n = \omega \wedge \cdots \wedge \omega$ cannot be exact either. Thus if M admits a symplectic structure, then there even must be a class in $H^2(M)$, whose nth power is non-zero in $H^{2n}(M)$. This easily implies that all the lower powers must be non-zero, too. Hence for M to admit a symplectic structure, $H^{2i}(M)$ must be non-trivial for all $i = 1, \dots, n$.

4.9. Homotopy invariance and the Poincaré lemma. The crucial step towards robustness and topological significance of the de-Rham cohomology is related to the concept of *homotopy*, which is central for algebraic topology. In our setting we need a smooth version of this concept.

Definition 4.9. Let $F, G : M \to N$ be two smooth maps between smooth manifolds. A *homotopy* from F to G is a smooth map $H : M \times I \to N$, where $I \subset \mathbb{R}$ is an open interval containing $[0, 1]$, such that $H(x, 0) = F(x)$ and $H(x, 1) = G(x)$ for each $x \in M$. If such a homotopy exists, we call F and G *homotopic*.

One shows that being homotopic is an equivalence relation on the set $C^\infty(M, N)$ of smooth maps from M to N. Reflexivity and symmetry of the relation are fairly obvious, for transitivity, one has to work with cutoff functions. Using those, one shows that one may always work with homotopies defined on $M \times \mathbb{R}$, which even satisfy $H(x, t) = F(x)$ for all x and $t < 1/4$ and $H(x, t) = G(x)$ for all x and $t > 3/4$. Such homotopies can then be pieced together smoothly without big problems.

The main result on homotopy invariance of de-Rham cohomology then reads as follows.

Theorem 4.9. *Let $F, G : M \to N$ be homotopic smooth maps between smooth manifolds. Then F and G induce the same homomorphisms in cohomology, i.e. $H^*(F) = H^*(G) : H^*(N) \to H^*(M)$.*

Sketch of proof. Let $I \subset \mathbb{R}$ be an open interval containing $[0, 1]$ and for $t \in I$ consider the smooth map $j_t : M \to M \times I$ defined by $j_t(x) := (x, t)$. Then it suffices to construct, for each k, a map $h = h_k : \Omega^k(M \times I) \to \Omega^{k-1}(M)$ such that for each $\omega \in \Omega^k(M \times I)$, we get $(j_1)^*\omega - (j_0)^*\omega = h_{k+1}(d\omega) + dh_k(\omega)$. Indeed, for a homotopy H from F to G, we then get $F = H \circ j_0$ and $G = H \circ j_1$ and for $\tau \in Z^k(N)$ we compute as follows, using $dH^*\tau = 0$.

$$G^*(\tau) = (j_1)^*(H^*\tau) = (j_0)^*(H^*\tau) + dh_k(H^*\tau) = F^*\tau + dh_k(H^*\tau),$$

and hence $G^\#[\tau] = F^\#[\tau]$.

Denote by s the obvious coordinate on I, which also defines a coordinate on $M \times I$, and by $\partial_s \in \mathfrak{X}(M \times I)$ the corresponding coordinate vector field. Then for $\omega \in \Omega^k(M \times I)$, $x \in M$ and $t \in I$, we can form $(j_t)^*(i_{\partial_s}\omega)(x)$. This lies in the finite dimensional vector space $\Lambda^{k-1}T^*_xM$ of $(k-1)$-linear alternating maps $(T_xM)^{k-1} \to \mathbb{R}$. One easily verifies that this depends smoothly on t and then defines $(h\omega)(x) := \int_0^1 (j_t)^*(i_{\partial_s}\omega)(x)dt$. Next, one shows that this depends smoothly on x and thus $h\omega \in \Omega^{k-1}(M)$ and verifies directly that $dh\omega(x) = \int_0^1 (j_t)^*(di_{\partial_s}\omega)(x)dt$, which mainly needs commuting d with the integral. Using this and Cartan's magic formula from Proposition 3.7, we obtain

$$hd\omega(x) + dh\omega(x) = \int_0^1 (j_t)^*(\mathcal{L}_{\partial_s}\omega)(x)dt \tag{4.8}$$

Since the flow of ∂_s is just translation in the I-coordinate, one finally concludes that $(j_t)^*(\mathcal{L}_{\partial_s}\omega)(x) = \frac{d}{dt}j_t^*\omega(x)$, and then the fundamental theorem of calculus implies that the right hand side of (4.8) gives $(j_1)^*\omega(x) - (j_0)^*\omega(x)$ which completes the argument. □

As a corollary, we can compute the cohomology of $\mathbb{R}^n$ and of subsets $U \subset \mathbb{R}^n$ which are *star-shaped* in the sense that they contain a point $x_0 \in U$ such that for each $x \in U$, the line-segment $\{tx_0 + (1-t)x : t \in [0,1]\}$ joining x_0 to x is contained in U. This in turn implies that any closed form on a smooth manifold M is locally exact.

Corollary 4.9 (Lemma of Poincaré). *(1) For a star shaped open subset $U \subset \mathbb{R}^n$, we have $H^0(U) = \mathbb{R}$ and $H^k(U) = \{0\}$ for $k > 0$.*

(2) For a smooth manifold M, a closed form $\omega \in Z^k(M)$ with $k > 0$ and any point $x \in M$, there are an open neighborhood V of $x \in M$ and a form $\alpha \in \Omega^{k-1}(V)$ such that $\omega|_V = d\alpha$.

Proof. (1) By assumption U is path connected and thus connected so $H^0(U) \cong \mathbb{R}$. Let $\psi : \mathbb{R} \to \mathbb{R}$ be a smooth function with values in $[0,1]$ such that $\psi(t) = 0$ for all $t \leq 1/4$ and $\psi(t) = 1$ for all $t \geq 3/4$, compare to Section 1.9. Then the smooth map $H : U \times \mathbb{R} \to \mathbb{R}^n$ defined by $H(y,t) := \psi(t)x_0 + (1-\psi(t))y$ has values in U and defines a homotopy from id_U to the constant map $q(y) := x_0$. For $k > 0$ and $\omega \in \Omega^k(U)$ we of course have $q^*\omega = 0$. But if ω is closed, then by Theorem 4.9, this represents the same class in $H^k(U)$ as $(\mathrm{id}_U)^*\omega = \omega$, which proves the first claim.

(2) Take a chart (V,v) for M with $x \in V$ such $v(V) \subset \mathbb{R}^n$ is an open ball. By (1), $(v^{-1})^*\omega \in Z^k(v(V))$ can be written as $d\beta$ for some $\beta \in \Omega^{k-1}(v(V))$ and then $\alpha = v^*\beta \in \Omega^{k-1}(M)$ does the job. □

4.10. De-Rham cohomology and integration. Suppose that M is a compact, oriented manifold of dimension n. Then from Theorem 4.4 we know that the integral defines a surjective linear map $\int_M : \Omega^n(M) \to \mathbb{R}$. Moreover, by Stokes's theorem (Theorem 4.6), the integral of an exact form is always zero. Since on a manifold of dimension n, any n-form is automatically closed, we see that the integral actually induces a surjection $H^n(M) \to \mathbb{R}$. (This is sometimes called the *cohomological integral*.) It needs some effort to show that on a connected, compact oriented manifold M, the cohomological integral actually is a linear isomorphism. (In fact the proofs in [**Michor**] and [**Lee**] both use some facts about compactly supported cohomology that we will discuss in the next section.)

Theorem 4.10. *Let M be a connected manifold of dimension n. Then $H^n(M) \cong \mathbb{R}$ if M is compact and oriented (with the isomorphism induced by the integral) and $H^n(M) = \{0\}$ if M is either non-compact or non-orientable.*

This allows for immediate, rather strong applications: Consider a smooth map $F : M \to N$ between connected compact oriented manifolds of dimension n. Then we obtain a linear map $H^n(F) : H^n(N) \to H^n(M)$ and both spaces can be canonically identified with $\mathbb{R}$ by Theorem 4.10. Thus $H^n(F)$ is given by multiplication by a number, which is called the *mapping degree* of F and denoted by $\deg(F) \in \mathbb{R}$. Explicitly, this can be described as follows: Choose $\omega \in \Omega^n(N)$ such that $\int_N \omega = 1$ and then $\deg(F) = \int_M F^*\omega$.

It is a consequence of Theorem 4.10 that this is well defined: Indeed, for another choice $\tilde{\omega} \in \Omega^n(M)$ with $\int_N \tilde{\omega} = 1$, we obtain $\int_N(\tilde{\omega} - \omega) = 0$, which by Theorem 4.10 implies that $\tilde{\omega} - \omega = d\alpha$ for some $\alpha \in \Omega^{n-1}(N)$. But then $F^*\tilde{\omega} = F^*(\omega + d\alpha) = F^*\omega + dF^*\alpha$ by Theorem 3.6 and hence $\int_M F^*\tilde{\omega} = \int_M F^*\omega + \int_M d\alpha$, and $\int_M d\alpha = 0$ by Stokes's theorem.

Recall from Section 4.4 that a diffeomorphism $F : M \to N$ between connected oriented compact manifolds is either orientation preserving or orientation reversing. Theorem 4.4 then shows that $\deg(F) = 1$ if F is orientation preserving and $\deg(F) = -1$ if F is orientation reversing.

The mapping degree is a rather robust invariant: If $F, G : M \to N$ are homotopic, then $H^n(F) = H^n(G)$ by Theorem 4.9 and hence $\deg(F) = \deg(G)$, so this is stable under smooth deformations. On the other hand, suppose that $F : M \to N$ is not surjective. Then $F(M) \subset N$ is closed by compactness and hence $V := N \setminus F(M)$ is a non-empty open subset. Taking an oriented chart (U, u) for N with $U \subset V$, and $\omega = \varphi du^1 \wedge \cdots \wedge du^n$ for a bump function φ with support contained in U, we obtain $\int_N \omega > 0$, so multiplying by an appropriate constant we may assume $\int_N \omega = 1$. But then by construction $F^*\omega = 0$ and hence $\deg(F) = 0$. Hence we conclude that a diffeomorphism from a connected compact oriented manifold M (for example on S^n) to itself can never be homotopic to a map that is not surjective, so in particular it is not homotopic to a constant map. In particular, this applies to the identity map id_M.

In fact, it turns out that $\deg(F)$ always is an integer, which is one of the reasons for its robustness, but difficult to see from the definition. We can see this nicely assuming that there is a *regular value* $y \in N$ for F, i.e. a point such that for each $x \in M$ with $F(x) = y$ the tangent map $T_xF : T_xM \to T_{F(x)}N$ is a linear isomorphism. (Using differential topology, one can prove that the set of such regular values is always dense in N, but this is beyond the scope of this course.) If we assume that y is a regular value than each $x_i \in F^{-1}(\{y\})$ has an open neighborhood on which F restricts to a diffeomorphism. Hence $F^{-1}(\{y\}) \subset M$ is discrete and hence finite by compactness. Now for each x_i, one puts $\epsilon_i = 1$ if $T_{x_i}F$ is orientation preserving and $\epsilon_i = -1$ if $T_{x_i}F$ is orientation reversing. Then one easily shows that there are neighborhoods V of y in N and U_i of x_i in M for each i such that $U_i \cap U_j = \emptyset$ for $i \neq j$ and such that $F|_{U_i} : U_i \to V$ is a diffeomorphism for each i. As above, we can find $\omega \in \Omega^n(N)$ such that $\int_N \omega = 1$ with $\mathrm{supp}(\omega) \subset V$. Diffeomorphism invariance of the integral then easily implies that $\deg(F) = \int_M F^*\omega = \sum_i \int_{U_i} F^*\omega = \sum_i \epsilon_i$. Hence one can obtain the mapping degree of F by counting preimages of a regular value taking into account orientations.

The mapping degree also leads to a nice proof of the hairy-ball theorem that we have mentioned in Example 2.1. This says that on an even dimensional sphere S^n any vector field $\xi \in \mathfrak{X}(S^n)$ has at least one zero. The proof is based on the antipodal map $A : S^n \to S^n$, $A(x) = -x$. One can construct a volume form ω on S^n from the volume form on $\mathbb{R}^{n+1}$ as discussed in Section 4.7 and for this it is easy to see that $A^*\omega = (-1)^{n+1}\omega$. If n is even, then this implies that $\deg(A) = -1$.

Now if $\xi : S^n \to \mathbb{R}^{n+1}$ is any vector field on S^n, then $x + \xi(x) \neq 0$ for any x, so we can define $F : S^n \to S^n$ by $F(x) := \frac{x+\xi(x)}{|x+\xi(x)|}$. Of course, this is homotopic to the identity map via $H(x,t) := \frac{x+\psi(t)\xi(x)}{|x+\psi(t)\xi(x)|}$ for a cutoff function ψ and hence $\deg(F) = 1$. But if we in additon assume that $\xi(x) \neq 0$ for all $x \in S^n$, then $F(x) \neq x$ for all x and then F is homotopic to the antipodal map A via $H(x,t) = \frac{t(-x)+(1-t)F(x)}{|t(-x)+(1-t)F(x)|}$ (which makes sense only if $F(x) \neq x$). This implies that $1 = \deg(F) = \deg(A)$ which is a contradiction for even n.

4.11. Cohomology with compact supports and Poincaré duality. We conclude our discussion of cohomology by indicating how some of the ideas discussed so far can be extended beyond the setting of compact manifolds. At the same time, these ideas lead to Poincaré duality, which is the most important fact about the cohomology of manifolds.

The starting point for this is the observation that there is an analog of cohomology based on forms with compact support. As in Section 4.6, we write $\Omega^k_c(M)$ for the space of k-forms on M with compact support. There we have also noted that $\operatorname{supp}(d\omega) \subset \operatorname{supp}(\omega)$, so $d(\Omega^{k-1}_c(M)) \subset \Omega^k_c(M)$. Hence we can define the kth *de-Rham cohomology with compact support* $H^k_c(M)$ as the quotient of $\ker(d) \subset \Omega^k_c(M)$ by the subspace $d(\Omega^{k-1}_c(M))$ and then put $H^*_c(M) := \oplus_{k=0}^n H^k_c(M)$. In the picture of equivalence classes, we consider $[\omega]$ for $\omega \in \Omega^k_c(M)$ with $d\omega = 0$ and $[\omega] = [\tilde{\omega}]$ if and only if $\tilde{\omega} = \omega + d\alpha$ for some $\alpha \in \Omega^{k-1}_c(M)$. So this is more restrictive than $\tilde{\omega}$ and ω differing by an exact form.

From the definition of the wedge product, it follows that $\omega \wedge \tau$ vanishes in a point x if one of the two forms vanishes in x. This easily implies that $\operatorname{supp}(\omega \wedge \tau) \subset \operatorname{supp}(\omega) \cap \operatorname{supp}(\tau)$, so a wedge product of two forms has compact support if at least one of the factors has compact support. In particular, the wedge product again makes $H^*_c(M)$ into an an associative, graded commutative algebra.

The situation with pullbacks is less simple and needs a restriction. We call a smooth map $F : M \to N$ *proper* if for any compact subset $K \subset N$, the preimage $F^{-1}(K) \subset M$ is compact. This makes sure that for $\omega \in \Omega^k_c(N)$ one gets $F^*\omega \in \Omega^k_c(M)$, and then a proper smooth map F induces an algebra homomorphism $H^*_c(F) : H^*_c(N) \to H^*_c(M)$. This issue also affects the question of homotopy invariance, which we do not discuss in detail here. There is a notion of proper homotopy of proper smooth maps, see Section 12.5 of [**Michor**]. By essentially the same proof as for Theorem 4.9, proper homotopic proper maps induce the same homomorphism in compactly supported cohomology.

Also, $H^0_c(M)$ is slightly more subtle that $H^0(M)$. Indeed, a locally constant function $f : M \to \mathbb{R}$ has compact support if and only if it vanishes on any non-compact connected component of M. In particular, if M is connected then $H^0_c(M) = \mathbb{R}$ if M is compact and $H^0_c(M) = \{0\}$ if M is non-compact. For a disjoint union $M = \sqcup_{i\in I} M_i$ a form has compact support if and only if its restriction to each M_i has compact support and only finitely many of these restrictions are non-zero. This easily implies that $H^k_c(\sqcup_{i\in I} M_i) = \oplus_{i\in I} H^k_c(M_i)$ for each k.

The most important features of compactly supported cohomology come from its interplay with integration. On an oriented manifold M of dimension n, we can integrate compactly supported n-forms and form Theorem 4.4 we know that $\int_M : \Omega^n_c(M) \to \mathbb{R}$ is surjective. By Stokes's theorem (Theorem 4.6), we get $\int_M d\alpha = 0$ for any $\alpha \in \Omega^{n-1}_c(M)$, so the integral again factors to a surjective linear map $H^n_c(M) \to \mathbb{R}$, the *cohomological integral.* (Observe that Stokes's theorem does not say that the integral of $\omega \in \Omega^n_c(M)$ vanishes if ω is exact. Indeed, by Theorem 4.10 any n-form on a non-compact orientable n-manifold is exact.) It then turns out that for any connected oriented manifold M of dimension n, the cohomological integral defines a linear isomorphism $H^n_c(M) \to \mathbb{R}$ (which is usually obtained as a very special instance of Poincaré duality as discussed below). This then shows that the notion of mapping degree extends to proper smooth maps between connected oriented manifolds of the same dimension.

The ideas about integration apply in a much more general setting, however. Let M be an oriented manifold of dimension n, and for some $k = 0, \dots, n$ consider forms

$\omega \in \Omega^k(M)$ and $\tau \in \Omega_c^{n-k}(M)$. Then we can form $\omega \wedge \tau$ and as we have observed above $\operatorname{supp}(\omega \wedge \tau) \subset \operatorname{supp}(\tau)$ so this lies in $\Omega_c^n(M)$. Thus in this situation we can form $\int_M \omega \wedge \tau \in \mathbb{R}$. If we assume that τ is closed, then for $\alpha \in \Omega^{k-1}(M)$ we get $d\alpha \wedge \tau = d(\alpha \wedge \tau)$ and $\alpha \wedge \tau \in \Omega_c^{n-1}(M)$. Hence if also ω is closed, then $\int_M \omega \wedge \tau$ depends only on $[\omega] \in H^k(M)$. Similarly, if ω is closed then for $\beta \in \Omega_c^{n-k-1}(M)$ we get $\omega \wedge d\beta = (-1)^k d(\omega \wedge \beta)$ and $\omega \wedge \beta \in \Omega_c^{n-1}(M)$. Thus $\int_M \omega \wedge \tau$ also depends only on $[\tau] \in H_c^{n-k}(M)$, so we obtain a well defined bilinear map, the *Poincaré pairing*

$$\mathcal{P} : H^k(M) \times H_c^{n-k}(M) \to \mathbb{R} \qquad \mathcal{P}([\omega], [\tau]) := \int_M \omega \wedge \tau. \tag{4.9}$$

This allows us to formulate the Poincaré duality theorem, which is the most fundamental result on the topology of manifolds:

THEOREM 4.11. *On any oriented smooth manifold M of dimension n, the Poincaré pairing induces linear isomorphisms $H^k(M) \cong (H_c^{n-k}(M))^*$ for each k, so in particular, it is non-degenerate. If M is compact, then its Betti numbers satisfy $b_k = b_{n-k}$ for each k.*

As mentioned above, this in particular shows that $\dim(H_c^n(M)) = \dim(H^0(M))$, which is one in the case of connected M. Since we know that the cohomological integral defines a surjection $H_c^n(M) \to \mathbb{R}$ this implies that it is a linear isomorphisms and hence for $\omega \in \Omega_c^n(M)$, we get $\int_M \omega = 0$ if and only if there is $\alpha \in \Omega_c^{n-1}(M)$ such that $\omega = d\alpha$.

Poincaré duality has a number of interesting consequences. First, if M is compact, then the pairing acts on the ordinary cohomology $H^*(M)$. Applying it twice, we see that $H^k(M) \cong (H^{n-k}(M))^* \cong (H^k(M)^*)^*$. Since we are talking about algebraic duals here, being isomorphic to its bidual is only possible for finite-dimensional spaces, so we see that all cohomologies of a compact oriented manifold are finite dimensional. Next, for a compact manifold M of even dimension $n = 2m$, we can consider cohomology in the middle dimension m. Since $m = n - m$, the Poincaré pairing defines a non-degenerate bilinear form on the finite dimensional vector space $H^m(M)$. If m is even (and hence the dimension of M is divisible by 4), then this form is symmetric and thus has a well defined signature, which is called the *signature of M*. This is a fundamental topological invariant of such manifolds.

There are several other techniques of algebraic topology that can be formulated in the context of de-Rham cohomology. For example, for a covering of a smooth manifold M by two open subsets $U, V \subset M$ the *Mayer–Vietoris sequence* relates $H^*(M)$ to $H^*(U)$, $H^*(V)$ and $H^*(U \cap V)$. This for example allows to compute the cohomologies of spheres. Formulating the result precisely needs the concept of exact sequences, which is beyond the scope of this course, see Chapter 17 of [**Lee**] or Sections 11.6 – 11.10 of [**Michor**].

Bibliography

[Cap:Lie] A. Čap, *Lie Groups*, lecture notes available online at https://www.mat.univie.ac.at/∼cap/lectnotes.html .

[Lee] J.M. Lee, *Introduction to smooth manifolds* (second edition), Graduate Texts in Mathematics 218, Springer 2013.

[Michor] P.W. Michor, *Topics in Differential Geometry*, Graduate Studies in Mathematics 93, Amer. Math. Soc. 2008.