Differential Geometry

Shubham Dwivedi*
Universität Hamburg

July 16, 2026

These are the lecture notes on differential geometry for a course taught during the summer semester 2026 at the University of Hamburg. I thank all my students for a very active participation in the course, for asking very good questions and overall discussions. These notes are based on topics whose discussions can be found in many excellent sources like (but not limited to) [Lee13], [Bär], [Wen], [Tu11], [Wal] and [Bau].

*Please email any typos or errors to shubham.dwivedi@uni-hamburg.de

1. Motivations for studying differential geometry

Before diving deep into the study of differential geometry, let us try to discuss informally what a study of differential geometry comprises of. Suppose $\Sigma^2 \subset \mathbb{R}^3$ is a "smooth surface" and let $x, y \in \Sigma^2$ and we would like to measure "distance" between x and y. Naturally, we would like to measure the distance between two point on Σ *along* Σ. For the time being, suppose we have a path $\gamma : [a, b] \to \mathbb{R}^3$ with $\gamma([a.b]) \subset \Sigma$, $\gamma(a) = x$ and $\gamma(b) = y$. We can then define the "length of γ" by

$$l(\gamma) = \int_a^b |\dot{\gamma}(t)| dt = \int_a^b \sqrt{\langle \dot{\gamma}(t), \dot{\gamma}(t) \rangle} dt \tag{1.1}$$

where $\dot{\gamma}(t) = \frac{d}{dt}\gamma(t)$, $\langle \cdot, \cdot \rangle$ is the Euclidean inner product and $|v| = \sqrt{\langle v, v \rangle}$. Having defined the length, we can define the distance between x and y as

$$d(x, y) := \inf_{P(x,y)} l(\gamma), \tag{1.2}$$

where $P(x, y) = \{\gamma \mid \gamma \text{ is a path between } x \text{ and } y\}$.

This gives rise to the following question:

Question 1.1. *Given a smooth surface Σ^2 and $x, y \in \Sigma$, does $\exists$ a smooth path from x to y that has the shortest possible length? If yes, then is such a path unique?*

We will see later in the course that the answer to both the questions is yes if x and y are close enough. Such a shortest path is called a **geodesic** (**Geodäte**) and is characterized by a second-order ordinary differential equation (gewöhnliche Differentialgelichung zweiter ordnung.) Such geodesics are the analogoues of "straight lines" on Σ even when no *straight path* might exist on Σ. For instance, you might already know that the shortest path between any two points on the surface of a Sphere (Sphäre), or in other words geodesics, are "great circles" (Großkreise).

One thing which you might have noticed is that the main concept which we needed to define the length of a path and in turn, the distance between two points is that of the Euclidean inner product $\langle , \rangle$. We needed to know the value of $\langle v, w \rangle$ for those $v, w \in \mathbb{R}^3$ which are tangent to Σ at any given point. Again, you might know from your previous courses that $\langle , \rangle$ allows us to compute **angles** (**Winkel**)θ between two vectors v, w by the formula

$$\langle v, w \rangle = |v||w| \cos \theta.$$

Thus, we can not only define length of any smooth path along Σ but also the angle between two smooth paths if they intersect. The procedure of being able to take inner product of two "tangent vectors" (Tangentialvektoren) makes Σ a 2**-dimensional Riemannian manifold** (**2-dimensionale Riemannsche Mannigfaltigkeit**) and the inner product will be called a **Riemannian metric** (**Riemannsche metrik**). As with any new mathematical object/structures, a natural question arises that suppose $\Sigma_1, \Sigma_2 \subset \mathbb{R}^3$ are two smooth surfaces and $f : \Sigma_1 \to \Sigma_2$ a smooth bijective map (bijektive Abbildung) such that $f^{-1} : \Sigma_2 \to \Sigma_2$ is also smooth. Such a function f is called a **diffeomorphism** (**Diffeomorphismus**) and Σ_1 and Σ_2 are called **diffeomorphic** (**diffeomorph**) to each other. In addition, if f preserves all distances and angles then it is called an **isometry** (**Isometrie**) and Σ_1 and Σ_2 are then **isometric** (**isometrisch**) to each other.

Question 1.2. *If Σ_1 and Σ_2 are two diffeomorphic surfaces, is there any way to know if they are also isometric?*

For example, consider the two different spheres in $\mathbb{R}^3$. In Figure 1, we have the standard embedding of sphere in $\mathbb{R}^3$ and in Figure 2 we have a slightly nonstandard embedding with a part of sphere elongated so as to look like a cylinder. Intutively, it look like the two spheres are diffeomorphic to each other. Again, intuitively, it feels that they should *not* be isometric as the length of paths on the parts of the spheres in the elongated portion might be different. But how do we make this idea rigorous?

One way to make this idea rigorous is to introduce the concept of parallel transport (Paralleltransport) of vectors along closed paths. In the figures, Figure 1 has a "rounder" sphere and the parallel transport of a vector v along the drawn closed path gives a different vector on return. On the contrary, the shaded part of the sphere in Figure 2 looks flat and the parallel transport of the vector v is the same vector in the end. This will be made rigorous in the coming lectures and we will show that this property is related to the **curvature** (**Krümmung**) of the sphere at a point. This will also allow us to prove that the spheres in both the figures are indeed not isometric!

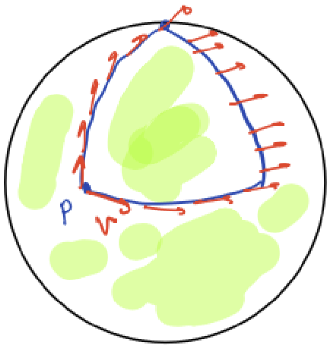


Figure 1: Standard embedding of $S^2 \subset \mathbb{R}^3$ with parallel transport of a vector along a closed path leading to a different vector on return.

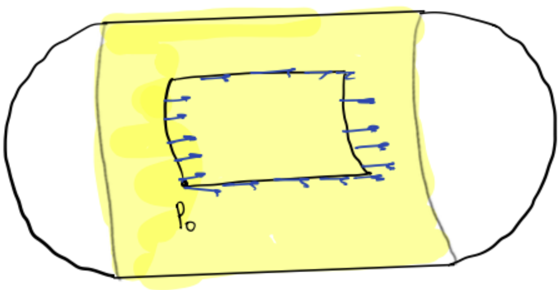


Figure 2: Another embedding of $S^2 \subset \mathbb{R}^3$ but now the yeellow shaded region is locally flat. Parallel transport of a vector returns to the same vector along a closed path.

We will spend a lot of time in defining and studying the properties of curvature where we will prove that it is a mathematical object called a **tensor field** (**Tensorfeld**). We will also see the notion of *local flatness*, i.e., parts of surfaces which locally look like parts of $\mathbb{R}^2$. For instance, the cylinder in Figure 3 shows a cylinder which is locally flat. In fact, this begs the question that can the *whole* sphere in Figure 2 be elongated so that the sphere is not curved at all?

This will be answered after we will prove the following amazing result.

Theorem 1.3 (Gauss–Bonnet theorem)**.** *If M is a compact, 2-dimensional Riemannian manifold with boundary ∂M and if K is the Gauss curvature of M and k is the geodesic curvature of ∂M, then*

$$\int_M K dA + \int_{\partial M} k ds = 2\pi\chi(M), \tag{1.3}$$

where $\chi(M)$ is the Euler characteristic of M.

We will use the Gauss–Bonnet theorem to prove that there cannot exist a Riemannian metric on S^2 which is everywhere locally flat.

We now start studying the topics formally.

2. Manifolds

2.1 Topological manifolds

We first recall some basic notions from topology.

Definition 2.1. Let M be a set. A subset $\mathcal{O} \subset \mathcal{P}(M)$ is called a **topology** on M if

1. $\emptyset, M \in \mathcal{O}$.
2. Arbitrary unions of elements of $\mathcal{O}$ is in $\mathcal{O}$, ie., if $\{U_i\}_{i\in I} \in \mathcal{O}$ then so is $\bigcup\limits_{i\in I} U_i$.
3. Finite intersections of elements in $\mathcal{O}$ are in $\mathcal{O}$, i.e., if $U_1, U_2 \in \mathcal{O}$ then $U_1 \cap U_2 \in \mathcal{O}$. We call the pair $(M, \mathcal{O})$ a **topological space. (topologischen Raum)**

Elements of $\mathcal{O}$ are called **open sets** (**offene Mengen**). A subset $\subset M$ is called **closed** (abgeschlossen) is $M \setminus C \in \mathcal{O}$. Let $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ be two topological spaces. A map $f : M \to N$ is called **continuous** (**Stetig**), if

$$f^{-1}(V) \in \mathcal{O}_M \;\; \forall \;\; V \in \mathcal{O}_N.$$

This means, that preimages of open sets in N are open in M. A bijective continuous map $f : M \to N$ with f^{-1} also being continuous is called a **homeomorphism** (**Homöomorphismus**), in which case M and N are said to be **homoemorphic** (**homömorph**).

We can now define what a topological manifold is.

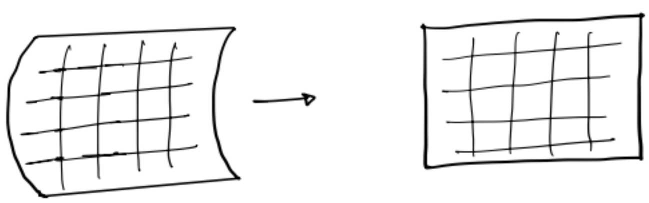


Figure 3:

Definition 2.2. Let $(M, \mathcal{O})$ be a topological space. Then M is called an **n-dimensional topological manifold** (**n-dimensionale topologische Mannigfaltigkeit**) if

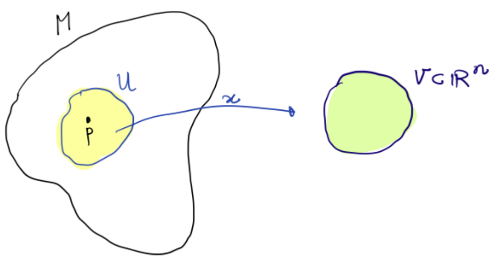
1. M is **Hausdorff** (**hausdorffsch**), ie., every pair of distinct points can be separated by open sets: given $p, q \in M$, there exists open sets $U \ni p$ and $V \ni q$ such that $U \cap V = \emptyset$.
2. The topology of M has a **countable basis** (**abzählbare Basis**), which means that there exists a countable subset $\mathcal{B} \subset \mathcal{O}$ such that for every $U \in \mathcal{O}$, there exists $B_i \in \mathcal{B}$, $i \in I$ with

$$U = \bigcup_{i \in I} B_i.$$

3. (most important) M is **locally homeomorphic** (**lokal homöonorph**) to $\mathbb{R}^n$ (same n as in the definition), that is, for all $p \in M$, there exists an open set $U \ni p$ and an open set $V \subset \mathbb{R}^n$ and a homeomorphsim between $x : U \to V$ between them.

   This basically means that any manifold locally looks an Euclidean space. A pictorial representation is shown in the figure below.

Definition 2.3. The homeomorphisms $x : U \to V$ in Definition 2.2 are called **charts** (**Karten**) or **local coordinate systems** of M (**lokale Koordinatensysteme**).



Example 2.4. Any Euclidean space $\mathbb{R}^n$ is a n-dimensional manifold with $U = M$, $V = \mathbb{R}^n$ and $x = id$, the identity map.

Example 2.5. Any open set $O \subset \mathbb{R}^n$ is an n-dimensional manifold with $U = O$, $V = \mathbb{R}^n$ and $x = id$.

Example 2.6. The n-dimensional sphere

$$S^n = \{y \in \mathbb{R}^n \mid ||y|| = 1\}$$

is a n-dimensional manifold. The existence of charts can be proved using the stereographic projection. If $s = \{-1, 0, \dots, 0\} \in \mathbb{R}^{n+1}$ denotes the south pole of the sphere then we define $U_1 = S^n \setminus \{s\}$ and let $V_1 = \mathbb{R}^n$. Define the map x by

$$x(y^0, \underbrace{y^1, \dots y^n}_{\hat{y}}) = \frac{2}{1 + y^0} \cdot \hat{y}.$$

Clearly the map x is continuous and bijective. The inverse map x^{-1} is given by

$$x^{-1}(z) = \frac{1}{4 + ||z||^2}(4 - ||z||^2, 4z)$$

and hence is also continuous. Thus, x is a chart of M.

Similarly, when we remove the north pole $n = \{1, 0, \dots 0\} \in bR^{n+1}$ and let $U_2 = S^n \setminus \{n\}$ and again choose $V_2 = \mathbb{R}^{n+1}$ then $x' : U_2 \to V_2$ given by

$$x'(y^0, \underbrace{y^1, \dots, y^n}_{\hat{y}}) = \frac{2}{1 - y^0} \cdot \hat{y},$$

is another chart, thus proving that S^n is an n-dimensional manifold.

Exercise 2.7. Can you cover S^n by a single chart? If yes, then what would that chart be? If no, then why not?

Exercise 2.8. Consider the double cone (Doppel-kegel) $M = \{(y^1, y^2, y^3) \in \mathbb{R}^3 \mid (y^1)^2 = (y^2)^2 + (y^3)^2\}$. Prove that M is *not* a 2-dimensional manifold by showing the failure of existence of charts.

Example 2.9. Let $U \subset \mathbb{R}^n$ be an open set and let $f : U \to \mathbb{R}^k$ be a continuous function. The **graph of** f (**Grafik von** f) is the set

$$\Gamma(f) = \{(y, f(y)) \mid y \in U\}.$$

Let $\pi_1 : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n$ be the projection onto the first factor and let $x : \Gamma(f) \to U$ be $x = \pi_{1|_{\Gamma(f)}}$ which is a continuous bijection with a continuous inverse. Thus, $\Gamma(f)$ is an n-dimensional topological manifold and is in fact, homeomorphic to U itself. The chart $(\Gamma(f), x)$ are called graph coordinates.

If you assume the fact that $\mathbb{R}^n$ and $\mathbb{R}^m$ can be homeomorphic if and only if $n = m$ then we have the following

Lemma 2.10. *(Topological Invariance of Dimension). The dimension of a topological manifold is a topological invariant of the manifold.*

We will see more examples of manifolds. Suppose we are given a chart or local coordinates system $x : U \to V$. Then at every point $p \in U$, we can look at $(x^1(p), \dots, x^n(p)) \in V \subset \mathbb{R}^n$, which we will called **coordinates of** p (**Koordinaten von** p).

Before discussing more examples and properties of manifolds, let's see if we can gather more information about them in lower dimensions.

0-dimension. If a manifold M is 0-dimensional then every point $p \in M$ has an open neighbourhood $U \ni p$ which is homemorphic to $\mathbb{R}^0 = \{0\}$. Thus, $\{p\}$ is an open subset of M for any p and hence M has the discrete topology. Moreover, the second requirement in Definition 2.2 forces M to be a countable space. Thus, we have proved the following

Proposition 2.11. *The only topological manifold of dimension 0 are countable spaces with discrete topology.*

We will prove the following classification about 1-dimensional manifolds in future lectures.

Proposition 2.12. *Any connected 1-dimensional manifold is homeomorphic to either S^1 or $\mathbb{R}$.*

Below giving the next example, we would like to understand better the realtion between the topology on M and the set of charts on M. Intuitively, it looks like we should be able to tell everything about the topology of M if we are given a collection of charts of M which cover it. This is the content of the next lemma.

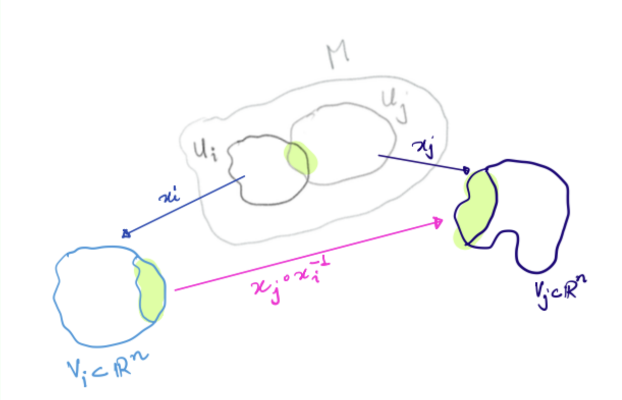
Lemma 2.13. *Every manifold has a unique topology with open sets being the domains of charts of the manifold.*

We prove this lemma in a series of propositions.

Proposition 2.14. *Let M be a set and I be an indexing set. Suppose for all $i \in I$, we have $U_i \subset M$, $V_i \subset \mathbb{R}^n$ and $x_i : U_i \to V_i$ are bijections. Suppose the following holds:*

1. $\bigcup_{i \in I} U_i = M$,
2. *for all $i \in I$, $x_i(U_i \cap U_j) \subset \mathbb{R}^n$ is open, and*
3. $x_j \circ x_i^{-1} : x_i(U_i \cap U_j) \to x_j(U_i \cap U_j)$ *is a continuous map on $\mathbb{R}^n$, for all $i, j \in I$ (see Figure 2.14).*

Then M carries a unique topology where the sets U_i are open sets and all the maps x_i are homeomorphisms.



If in addition, I is countable then the topology on M has a countable basis and if for any $p, q \in M$ $\exists\, i \in I$ with $p, q \in U_i$ then M is also Hausdorff.

Proof of Proposition 2.14. We first prove that if such a topology exists on M then it must be unique. We prove this by proving that if $\mathcal{O}$ is such a topology then a set $X \in \mathcal{O} \iff x_i(X \cap U_i)$ is open in $\mathbb{R}^n$ for all $i \in I$. Suppose $\mathcal{O}$ is a topology on M such that U_i are all open sets and x_i are all homeomorphisms. Let $X \in \mathcal{O}$. Then $X \cap U_i \in \mathcal{O}$ and hence $x_i(X \cap U_i)$ is open in $\mathbb{R}^n$ for all i. Conversely, suppose $X \subset M$ such that $x_i(U_i \cap X) \subset \mathbb{R}^n$ is open for all i. This means that $x_i \circ x_i^{-1}(U_i \cap X)$ is open in $\mathbb{R}^n$ and hence X is open in M. Thus, if a toplogy $\mathcal{O}$ given by the conditions of the proposition exists then it must be unique.

To prove the existence, we show that the set

$$\mathcal{O} = \{X \subset M \mid x_i(X \cap U_i) \subset \mathbb{R}^n \text{ open } \forall\, i \in I\}$$

describes a topology on M. Clearly, $\emptyset, M \in \mathcal{O}$ as $x_i(M \cap U_i) = x_i(U_i)$ is open in $\mathbb{R}^n$ by condition 2. If $X_j \in \mathcal{O},\ j \in J$ then, for all $i \in I$, we have

$$x_i\left(\left(\bigcup_{j \in J} X_j\right) \cap U_i\right) = x_i\left(\bigcup_{j \in J}(X_j \cap U_i)\right) = \bigcup_{j \in J} \underbrace{x_i(X_j \cap U_i)}_{\text{open in } \mathbb{R}^n}$$

and hence arbitrary unions of sets in $\mathcal{O}$ are in $\mathcal{O}$. You should prove the finite intersection part yourself. Thus, $\mathcal{O}$ described as above is indeed a topology on M.

Now we want to show that in this topology, U_i are open sets. But this is clear by condition 2. We now prove that x_i are homeomorphisms for all $i \in I$. Let $x_i : U_i \to V_i \subset \mathbb{R}^n$ be a map for some $i \in I$. We only know part 3. Let $Y \subset V_i$ be open. Then, for any $j \in I$, we have

$$\begin{aligned} x_j(x_i^{-1}(Y) \cap U_j) &= x_j(x_i^{-1}(Y \cap x_i(U_i \cap U_j))) \\ &= \underbrace{(x_j \circ x_i^{-1})}_{\text{continuous}} \underbrace{Y \cap \underbrace{x_i(U_i \cap U_j)}_{\text{open by part 2.}}}_{\text{open}}, \end{aligned}$$

is open in $\mathbb{R}^n$ and hence by the definition of $\mathcal{O}$, $x_i^{-1}(Y) \in \mathcal{O}$. Similarly, you can prove that x_i^{-1} is also continuous, which proves the proposition.

Now suppose the indexing set $I' \subset I$ is countable which covers M. Since the topology is unique, I and I' give the same topology on M. Now the topology of each $V_i \subset \mathbb{R}^n$ has a countable basis B_i. Then $x_i^{-1}(B_i)$is a countable basis of the topology of U_i and thus, $\bigcup\limits_{i \in I'} x_i^{-1} B_i$ is a countable basis of M.

Finally, Let $p, q \in M$ with $p \neq q$ and let $i \in I$ such that $p, q \in U_i$. Since $\mathbb{R}^n$ is Hausdorff, we can choose $V_1, V_2 \subset V_i$ open with $x_i(p) \in V_1$, $x_i(q) \in V_2$ and $V_1 \cap V_2 = \emptyset$. Then $x_i^{-1}(V_1)$ and $x_i^{-1}(V_2)$ separate p and q. □

Example 2.15 (Real Projective Space). We define the real projective space (reeller projektiver Raum) by

$$\mathbb{RP}^n = \{L \subset \mathbb{R}^{n+1} \mid L \text{ is one-dimensional subspace}\}, \tag{2.1}$$

that is, $\mathbb{RP}^n$ is the set of lines in $\mathbb{R}^{n+1}$. We use the quotient topology (Quotiententopologie) on $\mathbb{RP}^n$ by using the quotient map $\pi : \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{RP}^n$ which sends each point p in $\mathbb{R}^{n+1} \setminus 0$ to the subspace spanned by p, in other words, we use the equivalence relation $p \sim \lambda p, \lambda \neq 0 \in \mathbb{R}$.

We exhibit charts for $\mathbb{RP}^n$. For each $i = 1, \ldots, n+1$, let

$$\overline{U}_i \subset \mathbb{R}^{n+1} \setminus \{0\} \;\text{ where }\; p^i \neq 0,$$

and let

$$U_i = \pi(\overline{U}_i).$$

By the properties of quotient maps, $U_i \subset \mathbb{RP}^n$ is open and we define $\phi_i : U_i \to \mathbb{R}^n$ by

$$\phi_i\left([p^1, \ldots, p^{n+1}]\right) = \left(\frac{p^1}{p^i}, \ldots, \frac{p^{i-1}}{p^i}, \frac{p^{i+1}}{p^i}, \ldots, \frac{p^{n+1}}{p^i}\right).$$

This map is well-defined. Since $\phi_i \circ \pi$ is continuous, so is ϕ_i and in fact, it is a homeomorphism with the the continuous inverse given by

$$\phi_i^{-1}(u^1, \ldots, u^n) = [u^1, \ldots, u^{i-1}, 1, u^i, \ldots u^n].$$

Geometrically, $\phi([p]) = u$ means $(u, 1)$ is the point in $\mathbb{R}^{n+1}$ where the line $[p]$ intersects the hyperplane where $p^i = 1$. Now the sets $U_1, ..., U_{n+1}$ cover $\mathbb{RP}^n$, we see that (U_i, ϕ_i) form a chart of $\mathbb{RP}^n$, making it into a n-dimensional topological manifold.

Example 2.16 (Product manifolds). If M_1 and M_2 are topological manifolds of dimensions n_1 and n_2 respectively, then $M_1 \times M_2$ is again a topological manifold of dimension $n_1 + n_2$.

A simple example of the above phenomenon is the n-**torus** for $n \in \mathbb{N}$. It is denoted by $T^n = \underbrace{S^1 \times S^1 \times \cdots \times S^1}_{n-\text{times}}$ and is a topological n-manifold.

3. Smooth structures

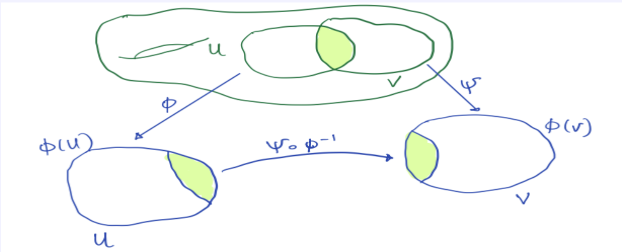
We recall the definition of smooth maps between Euclidean spaces. Recall that if $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ are open subsets then a map $f : U \to V$ is called **smooth** or $\mathbf{C}^\infty$ if each of its component functions has continuous partial derivatives of all orders. If, moreover, f is also a bijection and f^{-1} is also smooth then f is said to be a **smooth diffeomorphism**. Since any topological manifold locally looks like an open subset of an Euclidean space, it makes sense to try to define smooth maps on a topological manifold.

We would like to give a definition of when a map $f : M \to \mathbb{R}$ is smooth where M is a topological manifold. For doing this, we would need to put more structure on the manifold. We start with the following

Definition 3.1. Let M be a topological manifold and let (U, ϕ) and (V, ψ) be two charts such that $U \cap V \neq \emptyset$. The map

$$\psi \circ \phi^{-1} : \phi(U \cap V) \to \psi(U \cap V) \tag{3.1}$$

is called the **transition map from** ϕ **to** ψ.



We make the following

Definition 3.2. Let $\phi : U \to \tilde{U}$ and $\psi : V \to \tilde{V}$ be two charts on an n-dimensional manifold M. They are called **smoothly compatible** or C^∞**-compatible** if

$$\psi \circ \phi^{-1} : \phi(U \cap V) \to \psi(U \cap V) \tag{3.2}$$

is a C^∞diffeomorphism.

Having understood the definition of a smoothly compatibe charts , we now define what a atlas and a smooth atlas is.

Definition 3.3. A set of charts $\phi_i : U_i \to \tilde{U}_i$, $i \in I$, of M is called an **atlas** of M, if the domains of the charts cover M, i.e.,

$$\bigcup_{i \in I} U_i = M.$$

An atlas $\mathcal{A}$ is called a **smooth atlas** or a C^∞**-atlas** if any two charts in $\mathcal{A}$ are C^∞-compatible.

For proving that an atlas is a smooth atlas, we would only need to verify that each transition map $\psi \circ \phi^{-1}$ is smooth whenever (U, ϕ) and (V, ψ) are charts in $\mathcal{A}$ as $\psi \circ \phi^{-1}$ is obviously a diffeomorphism because $(\psi \circ \phi^{-1})^{-1} = \phi \circ \psi^{-1}$ and the latter is one of the transition maps which is smooth.

3.1 Examples of smooth atlas

Example 3.4. Let $M = U \subset \mathbb{R}^n$ be open. Then $\mathcal{A} = \{id : U \to U\}$ is a C^∞-atlas.

Example 3.5. For $M = S^n$, we recall the charts from Example 2.6. Consider $\mathcal{A} = \{(x_1 : U_1 \to V_1),\ (x_2 : U_2 \to V_2)\}$ where $U_1 = S^n \setminus \{s\}$ and $U_2 = S^n \setminus \{n\}$. We also recall that

$$\begin{aligned} x_1(y) &= \frac{2}{1+y^0}\hat{y}, \quad \text{where } y = (y^0, \hat{y}) \in \mathbb{R}^{n+1}, \\ x_1^{-1}(z) &= \frac{1}{4+||z||^2}(4-||z||^2, 4z),\ z \in S^n, \\ x_2(y) &= \frac{2}{1-y^0}\hat{y}. \end{aligned}$$

Thus, for any $v \in x_1(U_1 \cap U_2)$, we have

$$\begin{aligned} x_2 \circ x_1^{-1}(v) &= x_2\left(\frac{4-||v||^2}{4+||v||^2}, \frac{4v}{4+||v||^2}\right) \\ &= \frac{2}{1-\frac{4-||v||^2}{4+||v||^2}} \frac{4v}{4+||v||^2} \\ &= \frac{4v}{||v||^2}. \end{aligned}$$

Note that $v \in \mathbb{R}^n \setminus \{0\}$. Thus, $x_2 \circ x_1^{-1}$ is C^∞. Similarly, $x_1 \circ x_2^{-1}$ is also smooth, making x_1 and x_2 C^∞-compatible and $\mathcal{A}$ a C^∞-atlas for S^n.

Example 3.6. Let us recall the real projective space $\mathbb{RP}^n$ and its charts from Example 2.15. One can compute that for say, $i > j$ and $(u^1, ..., u^n) \in (U_i \cap U_j)$, we have

$$\phi_j \circ \phi_i^{-1}(u^1, ..., u^n) = \left(\frac{u^1}{u^j}, \cdots, \frac{u^{j-1}}{u^j}, \frac{u^{j+1}}{u^j}, \cdots, \frac{u^{i-1}}{u^j}, \frac{1}{u^j}, \frac{u^{i+1}}{u^j}, \cdots, \frac{u^n}{u^j}\right),$$

which are just rational functions and hence C^∞. Thus, we get a smooth atlas for $\mathbb{RP}^n$.

We will see more examples later. But for now, recall that we introduced this notion of smooth atlas because we wanted to define when a function $f : M \to \mathbb{R}$ is differentiable. We would like to say that $f : M \to \mathbb{R}$ is smooth if $f \circ \phi^{-1}$ is smooth in the sense of ordinary calculus for each coordinate chart (U, ϕ) in the atlas. There is a small problem with this definition. It can happen that may different atlases give rise to the "same" smooth strcuture on M. To avoid any redundancy, we make the following definition.

Definition 3.7. A C^∞-atlas $\mathcal{A}_{max}$ is called a **maximal atlas** or a **differentiable structure on** M, if every chart that is smoothly compatiblle with all charts in $\mathcal{A}_{max}$ is already contained in $\mathcal{A}_{max}$. In other words,

$$\mathcal{A}_{max} := \{\text{charts } (U, \phi) \text{ in atlas } \mathcal{A} \text{ of } M \mid \phi \text{ is } C^\infty - \text{compatible with all charts in } \mathcal{A}\}. \tag{3.3}$$

Clearly, $\mathcal{A}_{max}$ is also a C^∞-atlas. If ϕ, ψ are two charts of M which are C^∞-compatible with all charts in $\mathcal{A}$ then ϕ and ψ are C^∞-compatible with each other because if $x \in \phi(U \cap V)$ then there exists a chart $\tau : \tilde{U} \to \tilde{\tilde{U}}$ in $\mathcal{A}$ such that $\phi^{-1}(x) \in \tilde{\tilde{U}}$. So, near x, we have

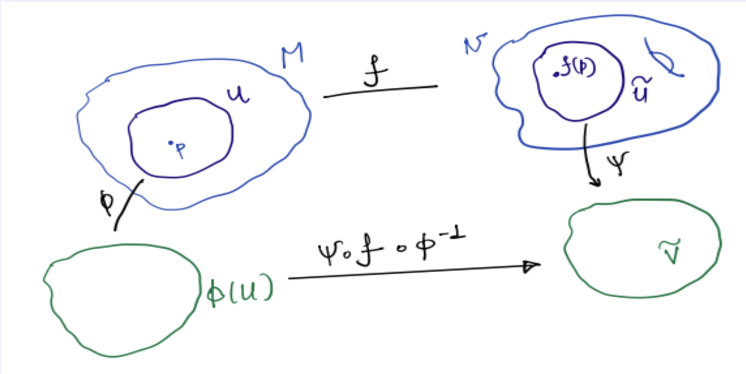
$$\psi \circ \phi^{-1} = \underbrace{(\psi \circ \tau^{-1})}_{C^\infty} \circ \underbrace{(\tau \circ \phi^{-1})}_{C^\infty}.$$

Definition 3.8. Let M^n be a topological manifold. A pair $(M, \mathcal{A}_{max})$, where $\mathcal{A}_{max}$ is a maximal atlas on M, is called an n**-dimensional differentiable manifold.**

With this definition in hand we can finally define the differentiability of functions.

Definition 3.9. Let M and N be differentiable manifolds, $p \in M$ and let $k \in \mathbb{N} \cup \{\infty\}$. A continuous map $f : M \to N$ is called **k-times continuously differentiable at p** or **C^k at p** if $\exists$ a chart $\phi : U \to V \in \mathcal{A}_{max}(M)$ with $p \in U$ and another chart $\psi : \tilde{U} \to \tilde{V} \in \mathcal{A}_{max}(N)$ with $f(p) \in \tilde{U}$ and $f(U) \subset \tilde{U}$ such

$$\psi \circ f \circ \phi^{-1} : \phi(U) \to \tilde{V} \text{ is } C^k. \tag{3.4}$$



Definition 3.10. Let M, N be differentiable manifolds. A homeomorphism $f : M \to N$ is called a **C^k-diffeomorphism** if both f and f^{-1} are C^k. In such a case, M and N are called **C^k-diffeomorphic**.

Before seeing examples of smooth maps, let's see more examples of smooth manifolds.

Remark 3.11. We emphasize that a smooth structure is an additional data which we associate to a topological manifold M. So when we say that M is a "smooth manifold" then it is already understood that there is a smooth maximal atlas prescribed. It can happen that a given topological manifold has many different smooth structures. For example, consider the homeomorphism

$$\psi : \mathbb{R} \to \mathbb{R}, \quad \psi(x) = x^3.$$

The atlas $(\mathbb{R}, \psi)$ is a smooth structure on $\mathbb{R}$ but it is **not C^∞-compatible** with the standard smooth structure $(\mathbb{R}, id)$ on $\mathbb{R}$ because

$$id \circ \psi^{-1}(y) = y^{\frac{1}{3}}$$

is not smooth at the point $\{0\}$. Thus, the two smooth structures are different.

Exercise 3.12. Let M be a nonempty topological manifold of dimension $n \geq 1$. Suppose M has a smooth structure. Show that M has uncountably many distinct smooth structures.

Remark 3.13. In fact, it can happen that a topological manifold has no smooth structures at all. The first example was given by Kervaire [Ker60] on a compact 10-dimensional manifold.

Although, it seems tempting to define a smooth structure by explicitly describing a maximal smooth atlas, it is not always convenient to do so because that can contain many charts and so things might be harder to check explicitly. This can be remedied by the following result.

Proposition 3.14. *Let M be a topological manifold.*

1. *Every smooth atlas $\mathcal{A}$ of M is contained in a unique maximal smooth atlas $\mathcal{A}_{max}$. Such an $\mathcal{A}_{max}$ is also called the* **smooth structure determined by** *$\mathcal{A}$.*

2. *Two smooth atlases of M determine the same smooth structure if and only if their union is a smooth atlas.*

Proof. 1. We have already proved the existence of such $\mathcal{A}_{max}$. To prove uniqueness, note that if $\mathcal{B}$ is any other maximal smooth atlas containing $\mathcal{A}$ then each of its charts is C^∞-compatible with each chart in $\mathcal{A}$ and hence by definition $\mathcal{B} \subset \mathcal{A}_{max}$. Similarly, $\mathcal{A}_{max} \subset \mathcal{B}$ and hence $\mathcal{A}_{max} = \mathcal{B}$.

2. Exercise. □

3.2 Some more examples

Example 3.15. Let M^0 be a 0-dimensional manifold. We know that it a countable space with discrete topology. The charts are $\{(p, \phi) \mid p \in M, \phi : \{p\} \to \mathbb{R}^0\}$ and they trivially satisfy the compatibility conditions and thus M^0 has a unique smooth structure.

Example 3.16. The smooth structure on $\mathbb{R}^n$ determined by the atlas $(\mathbb{R}^n, id)$ is called the **standard smooth structure on** $\mathbb{R}^n$**.**

Example 3.17. Let V^n be a finite dimensional real vector space. Since any norm on V determines a topology and all the norms are equivalent, we use this topology to view V as a topological manifold. We show below that V is a smooth manifold. Suppose $\{e_1, \ldots, e_n\}$ is an ordered basis of V and define the isomorphism $\phi : \mathbb{R}^n \to V$ by

$$\phi(x) = \sum_{i=1}^{n} x^i e_i.$$

This map is a homeomorphism and hence (V, ϕ^{-1}) is a chart.

Suppose $\{e'_1, \ldots, e'_n\}$ is any other basis and $\phi'(x) = \sum_{i=1}^n x^i e'_i$ is the corresponding isomorphism. We know from linear algebra that there exists an invertible matrix $[A^i_j]$ such that $e_i = \sum_{j=1}^n A^j_i e'_j$ for each i. The transition map then is given by

$$\phi'^{-1} \circ \phi(x) = x', \quad \text{where} \quad x' \in \mathbb{R}^n \ \text{ and } \ \sum_{j=1}^{n} x^{j'} e'_j = \sum_{i=1}^{n} x^i e_i = \sum_{i,j=1}^{n} x^i A^j_i e'_j.$$

In other words, any map sending $x \mapsto x'$ is an invertible linear map and hence is a diffeomorphism, thus proving that any two charts are C^∞–compatible. The collection of all such charts define a smooth structure on V and makes it into a smooth manifold.

Example 3.18. Let $M(n \times m, \mathbb{R})$ denote the space of real $n \times m$ matrices. Since this is a real vector space of dimension nm, from the previous example, it is a smooth nm-dimensional manifold.

Example 3.19 ($\mathrm{GL}(n, \mathbb{R})$)**.** The space of $n \times n$ real invertible matrices is called the **general linear group** $\mathrm{GL}(n, \mathbb{R})$. Since this is an open subset of the set of all matrices, it is a n^2-dimensional smooth manifold.

Example 3.20. If $M_1, \ldots, M_k$ are smooth manifolds of dimensions $n_1, \ldots, n_k$, respectively then we already know that $M_1 \times \cdots \times M_k$ is a topological manifold of dimension $n_1 + n_2 + \cdots + n_k$. The charts are of the form $(U_1 \times \cdots \times U_k, \phi_1 \times \cdots \times \phi_k)$. We notice that if there two such charts then

$$(\psi_1 \times \cdots \times \psi_k) \circ (\phi_1 \times \cdots \times \phi_k)^{-1} = (\psi_i \circ \phi_1^{-1}) \times \cdots \times (\psi_k \circ \phi_k^{-1}),$$

making $M_1 \times \cdots \times M_k$ into a smooth manifold. For example, the n-torus T^n is a smooth manifold.

We can write a version of Proposition 2.14 for smooth manifolds as follows.

Lemma 3.21. *Let M be a set and suppose we have a collection $\{U_i\}_{i\in I}$, I is countable, of subsets of M with maps $\phi_i : U_i \to \mathbb{R}^n$ such that the following holds.*

(a) $\underset{i\in I}{U_i} = M$.

(b) For all i, ϕ_i is a bijection between U_i and an open subset $V_i \subset \mathbb{R}^n$.

(c) For all i, j, the sets $\phi_i(U_i \cap U_j)$ and $\phi_j(U_i \cap U_j)$ are open in $\mathbb{R}^n$.

(d) If $U_i \cap U_j \neq \emptyset$ for some i, j, then the map $\phi_j \circ \phi_i^{-1} : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$ is smooth.

(e) For points $p, q \in M$ either there exists some $U_i \ni p, q$ or there exists distinct U_i, U_j with $p \in U_1$ and $q \in U_j$.

Then M has a unique smooth structure such that each (U_i, ϕ_i) is a smooth chart.

Proof. The proof is similar to the proof of Proposition 2.14 with the smoothness condition guaranteed by point (d) above. □

We can use the previous lemma to give an example of smooth manifolds which generalize the real projective spaces.

Example 3.22 (Grassmann manifolds). Let V^n be a real vector space and let

$$G_k(V) = \{W \subset V \mid W \text{ is a } k - \text{dimensional subspace}\}. \tag{3.5}$$

The set $G_k(V)$ is a smooth manifold of dimension $k(n-k)$ and is called the Grassmannian. Clearly, $G_1(\mathbb{R}^{n+1}) = \mathbb{RP}^n$.

We show that $G_k(V)$ is a smooth manifold. Let P and Q be complementary subspaces of V of dimensions k and $n-k$, respectively, so that $V = P \oplus Q$. If $f : P \to Q$ is a linear map then $\Gamma(f) \subset V$ is a k-dimensional subspace with

$$\Gamma(f) = \{v + fV \mid v \in P\}.$$

One can check that $\Gamma(f) \cap Q = \{0\}$. Conversely, if $S \subset V$ is any subspace with trivial intersection with Q then $S = \Gamma(f)$ where $f = (\pi_{Q_{|S}}) \circ (\pi_{P_{|S}})^{-1}$ where $\pi_T : V \to T$ is the projection.

Having established this, we let

$$L(P, Q) = \{L : P \to Q \mid L \text{ linear}\},$$

which is a vector space and let $U_Q \subset G_k(V)$ with U_Q being those k-dimensional subspaces of V which have trivial intersection with Q. Since $P \in U_Q$ so it is non-empty. The assignment $f \mapsto \Gamma(f)$ gives a map

$$\Gamma : L(P, Q) \to U_Q,$$

and is in fact, a bijection. Let us denote $\Gamma^{-1} = \phi$. Since $L(P, Q) \cong M((n-k) \times k, \mathbb{R}) \cong \mathbb{R}^{k(n-k)}$, (U_Q, ϕ) is a chart for $G_k(V)$ and part (b) of Lemma 3.21 is satisfied.

If (P', Q') is another pair of subspaces as before and $\phi' : U_{Q'} \to L(P', Q')$ is the corresponding chart, then we first of all notice that $\phi(U_Q \cap U_{Q'}) \subset L(P, Q)$ is the set of all linear maps $f : P \to Q$ such that $\Gamma(f) \cap Q' = \{0\}$. This set $\phi(U_Q \cap U_{Q'})$ is open in $L(P, Q)$. This can be seen as follows. Let $f \in \phi(U_Q \cap U_{Q'})$ and this means that $\Gamma(f) \cap Q' = \{0\}$. If we denote by $I_f : P \to V$, $I_f(p) = p + fv$ then it is abijection from P to $\Gamma(f)$. Moreover, $\ker \pi_{P'} = Q'$ so using the linear algebraic fact that

$$\text{rank}(A \circ B) \le \text{rank}(B) \text{ with equality } \iff \text{Im}(B) \cap \ker A = \{0\},$$

we see that $\pi_{P'} \circ I_f$ has full rank and since the corresponding matrix depend continuously on f, we see that the set of all such f with $\Gamma(f) \cap Q' = \{0\}$ is open. Thus, part (c) of Lemma 3.21 is also true.

We now show that the transition maps are smooth and leave the rest of the details as an exercise. We want to show that $\phi' \circ \phi^{-1}$ is a smooth map on $\phi(U_Q \cap U_{Q'})$. Let $f \in \phi(U_Q \cap U_{Q'}) \subset L(P, Q)$ be arbitrary. Let S denote the subspace $\Gamma(f) \subset V$. If $f' = \phi' \circ \phi^{-1}(f)$, then, $f' = (\pi_{Q'_{|S}}) \circ (\pi_{P'_{|S}})^{-1}$. Somehow, we would like to relate this map to f. This can be done as follows: note that $I_f : P \to S = \Gamma(f)$ is an isomorphism so

$$f' = (\pi_{Q'_{|S}}) \circ I_f \circ (I_f)^{-1} \circ (\pi_{P'_{|S}})^{-1} = (\pi_{Q'} \circ I_f) \circ (\pi_{P'} \circ I_f)^{-1}.$$

Now we can choose any basis for V and write a basis for spaces $P, P', Q.Q'$ and then let $A : P \to P'$, $B : P \to Q'$, $C : Q \to P'$ and $D : Q \to Q'$ be the linear maps with

$$A = \pi_{P'_{|P}},\ B = \pi_{Q'_{|P}},\ C = \pi_{P'_{|Q}},\ \text{and } D = \pi_{Q'_{|Q}}.$$

So, for instance, for $p \in P$, we get

$$(\pi_{P'} \circ I_f)p = (A + Cf)p,\ (\pi_{Q'} \circ I_f)p = (B + Df)p,$$

and hence $f' = (B + Df)(A + Cf)^{-1}$. But then we are done because all these are matrices whose entries are just rational functions depending on f and as a result, the matrix entries of f' depend smoothly on those of f. Since f was smooth then so is the transition map. Thus, part (d) of Lemma 3.21 is also satisfied.

Part (a) also follows because if say, $\{e_1, \dots, e_n\}$ is any basis of V and $F \subset \{1, \dots, n\}$ with $|F| = k$. Then if $V_F =$ linear span$\{e_1, i \in F\}$ and we let $U_F = U_{V_F}$ with corresponding $\phi_{V_F} = \phi_F$ then the finite collection $\{(U_F, \phi_F) \mid F \subset \{1, \dots, n\},\ |F| = k\}$ is a smooth atlas of $G_k(V)$. Similarly, one can check that $G_k(V)$ is indeed Hausdorff and hence the Grassmannian is a smooth manifold.

XXXXXX

4. Properties of smooth maps

After having understood what smooth maps, let us try to state and prove some basic properties. We will see more properties of smooth maps as the course progresses.

Proposition 4.1 (Smoothness is a local property). *Let M, N be smooth manifolds and let $f : M \to N$ be a map.*

(a) If every point $p \in M$ has a neighbourhood $U \ni p$ such that f_U is smooth then f is smooth on M.

(b) Conversely, if f is smooth then its restriction to every open set $U \subset M$ is smooth.

Proof. It is clear from the definition. □

Proposition 4.2. *Let M, N, P be smooth manifolds. Then the following holds.*

1. *Every constant function is a smooth map.*
2. *The identity map $id_M : M \to M$ is smooth.*
3. *If $U \subset M$ is open then the inclusion map $i : U \hookrightarrow M$ is smooth.*
4. *If $f : M \to N$ is smooth and $g : N \to P$ is smooth then $g \circ f : M \to P$ is smooth.*

Proof. We just prove (1) and (4) and leave the rest as exercises. Let $p \in M$ and suppose (U, ϕ) is a chart containing p. If $c : M \to N$ is a constant function with $c(p) = c \in N$, $\forall\, p \in M$ and (V, ψ) is a chart containing c, then

$$\psi \circ c \circ \phi^{-1}(x) = \psi(c) \in \psi(V),$$

and hence is a constant function between subsets of Euclidean spaces and so is smooth.

For proving (4), suppose (V, ψ) is a chart containing $f(p)$ and (W, τ) is a chart containing $g(f(p))$ with $g(V) \subset W$ and $\tau \circ g \circ \psi^{-1} : \psi(V) \to \tau(W)$ is smooth. Note that $f^{-1}(V) \ni p$ is an open set in M and hence by the smoothness of f, we know that that $\exists$ a smooth chart (U, ϕ) of M such that $p \in U \subset f^{-1}(V)$ and $\psi \circ f \circ \phi^{-1}$ is smooth. Moreover, $g(f(U)) \subset g(V) \subset W$ and

$$\tau \circ (g \circ f) \circ \phi^{-1} = \tau \circ (g \circ \psi^{-1} \circ \psi \circ f) \circ \phi^{-1} = (\tau \circ g \circ \psi^{-1}) \circ (\psi \circ f \circ \phi^{-1}),$$

is smooth and hence the composite function $g \circ f$ is smooth. □

We also have the following

Proposition 4.3. *Let $M_1, \dots, M_k$ and N be smooth manifolds. If $\pi_i : M_1 \times \cdots \times M_k \to M_i$ is the projection map on the M_i factor then a map $f : N \to M_1 \times \cdots \times M_k$ is smooth if and only if $f_i = \pi_i \circ f : N \to M_i$ is smooth.*

Proof. Exercise. □

Let's see an example. Let S^1 denote the unit circle in $\mathbb{R}^2$. Define the map $f : \mathbb{R} \to S^1$ by $f(t) = \exp 2\pi i t$. If we use the angle coordinate θ for S^1 then with respect to any chart in this coordinate, we have

$$\psi \circ f \circ \phi^{-1}(t) = 2\pi t + c, \ \ c \text{ constant}.$$

Thus, f is a smooth map. Similarly, $f^n : \mathbb{R}^n \to T^n$ given by $f^n(x^1, \dots, x^n) = (e^{2\pi i x^1}, \dots, e^{2\pi i x^n})$ is a smooth map.

Exercise 4.4. Prove that the natural quotient map $\pi : \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{RP}^n$ is a smooth map.

5. Tangent vectors and Tangent spaces

Recall that the definition of a smooth manifold as well as a smooth map was motivated from the point of view of defining what a function f being *differentiable* could mean. We again go back to our study of differential calculus and recall that the derivative of a map was the *best linear approximation* of the map. In Euclidean spaces, a linear approximation makes sense both theoretically and pictorially. What could be the appropriate notion of a linear approximation on a manifold? Let's start with the simplest possible case.

5.1 Tangent vector and tangent spaces

We first recall that if $p \in M$ then by a curve on M passing through p, we mean a map $\gamma : (-\varepsilon, \varepsilon) \to M$ with $\gamma(0) = p$. If the map is differentiable/smooth then we say that the curve is differentiable/smooth.

Definition 5.1. A **tangent vector** on M at the point $p \in M$ is an **equivalence class** of differentiable curves $[\gamma]$ passing through p where $\gamma_1 \sim \gamma_2$ if for a chart $\phi : U \to V$ with $p \in U$, we have

$$\frac{d}{dt}(\phi \circ \gamma_1)_{|t=0} = \frac{d}{dt}(\phi \circ \gamma_2)_{|t=0}. \tag{5.1}$$

We denote $[\gamma] = \gamma'(0)$.

It's obvious that $\sim$ is indeed an equivalence relation. We also note that this definition is independent of the choice of chart. If (U', ψ) is another chart containing p then we use the chain rule to compute

$$\frac{d}{dt}(\psi \circ \gamma)_{|t=0} = \frac{d}{dt}\left((\psi \circ \phi^{-1}) \circ (\phi \circ \gamma)\right)_{|t=0} = D(\psi \circ \phi^{-1}) \circ \frac{d}{dt}(\phi \circ \gamma)_{|t=0},$$

where $D(\psi \circ \phi^{-1})(\phi(p)) : \mathbb{R}^n \to \mathbb{R}^n$ is the derivative of the transition map $\psi \circ \phi^{-1}$ at the point $\phi(p)$ and is an invertible linear map since $\psi \circ \phi^{-1}$ is a smooth map with a smooth inverse. Thus, the condition

$$\frac{d}{dt}(\phi \circ \gamma_1)_{|t=0} = \frac{d}{dt}(\phi \circ \gamma_2)_{|t=0}$$

is equivalent to the condition

$$\frac{d}{dt}(\psi \circ \gamma_1)_{|t=0} = \frac{d}{dt}(\psi \circ \gamma_2)_{|t=0}.$$

Definition 5.2. Let M be a smooth manifold and $p \in M$. The set

$$T_pM = \{\gamma'(0) \mid \gamma \text{ is a curve passing through } p\}$$

is called the **tangent space** of M at p.

Even though the definition of a tangent space is very geometric, it doesn't give much idea about the type of space T_pM is. The next proposition shows that it is a n-dimensional vector space.

Proposition 5.3. *Let M^n be a smooth manifold and $p \in M$ with chart (U, ϕ) containing p. The tangent space T_pM has a unique vector space structure such that the map*

$$d\phi_p : T_pM \to \mathbb{R}^n, \ [\gamma] \mapsto (\phi \circ \gamma)'(0) \tag{5.2}$$

is an isomorphism of vector spaces. Thus, for every point $p \in M$, the tangent space T_pM at that point is a n-dimensional real vector space.

Proof. The map in (5.2) is clearly well-defined and injective. For surjectivity, we'd need to show that if $v \in \mathbb{R}^n$ then there exists a curve whose tangent vector is v. Define $\gamma(t) = \phi^{-1}(\phi(p) + tv)$ and choose $\varepsilon > 0$ small enough so that $\phi(p) + tv \in \phi(U)$. Then

$$d\phi_p(\gamma'(0)) = \frac{d}{dt}\left(\phi \circ \phi^{-1}(\phi(p) + tv)\right)_{|t=0} = \frac{d}{dt}(\phi(p) + tv)_{|t=0} = v.$$

The map is also a linear isomorphism and in fact, the vector space structure is explicitly described as follows: let $a, b \in \mathbb{R}$ and $\gamma_1'(0), \gamma_2'(0) \in T_pM$ then

$$a\gamma_1'(0) + b\gamma_2'(0) = (d\phi_p)^{-1}\left(ad\phi_p(\gamma_1'(0)) + bd\phi_p(\gamma_2'(0))\right).$$

So the only remaining thing to prove is that the vector space structure on T_pM is independent of the chart. To see this, let $(\psi : \tilde{U} \to \tilde{V})$ be another chart with $p \in \tilde{U}$. We want to show that the map $d\psi_p : T_pM \to \mathbb{R}^n$ is linear with respect to the vector space structure induced by ϕ. But we have already seen that

$$d\psi_p = \underbrace{D(\psi \circ \phi^{-1})_{\phi(p)}}_{\text{invertible linear}} \circ \underbrace{d\phi_p}_{\text{invertible linear}},$$

and hence is a vector space isomorphism. □

We can think of T_pM as the linear approximation to M at the point p.

We can now define the differential of a differentiable map between smooth manifolds.

Lemma 5.4. *Let M and N be smooth manifolds and let $p \in M$. Let $f : M \to N$ be a smooth map near p. Then the map*

$$df_p : T_pM \to T_{f(p)}N, \ \gamma'(0) \mapsto (f \circ \gamma)'(0), \tag{5.3}$$

is a well-defined linear map between vector spaces T_pM and $T_{f(p)}N$.
The map df_p is called the **differential of f at the point** *p.*

Proof. Let (U, ϕ) be a chart of M containing p and let $(\tilde{U}, \psi)$ be a chart of N containing $f(p)$. We compute, using the chain rule,

$$\begin{aligned} d\psi_{f(p)}((f \circ \gamma)'(0)) &= (\psi \circ f \circ \gamma)'(0) \\ &= \left((\psi \circ f \circ \phi^{-1}) \circ (\phi \circ \gamma)\right)'(0) \\ &= D(\psi \circ f \circ \phi^{-1})_{\phi(p)} \cdot \left((\phi \circ \gamma)'(0)\right) \\ &= D(\psi \circ f \circ \phi^{-1})_{\phi(p)} \cdot d\phi_p(\gamma'(0)). \end{aligned}$$

The previous equation implies that

$$df_p = \left(d\psi_{f(p)}\right)^{-1} \circ D(\psi \circ f \circ \phi^{-1})_{\phi(p)} \cdot d\phi_p,$$

and hence is well-defined, that is, independent of the choice of curve γ and is a linear map. □

We notice that if $U \subset M$ is open then the differential of the inclusion map $i : U \hookrightarrow M$ is the natural isomorphism $di : T_pU \to T_pM$ given by

$$\gamma'(0) \mapsto (i \circ \gamma)'(0) = \gamma'(0),$$

and hence $T_pU = T_pM$.

Exercise 5.5. Prove that if M, N are two smooth manifolds and $p \in M, q \in N$, then

$$T_{(p,q)}(M \times N) \cong T_pM \times T_qN$$

in a canonical way.

Recall that V is a vector space then its dual space is denoted by V^* and is the space of linear functionals on V, i.e., scalar-valued linear maps on V. In other words,

$$V^* = Hom(V, \mathbb{R}), \quad \text{where} \quad Hom(V, W) = \{f : V \to W \mid f \text{ linear}\}.$$

Thus, we define the **cotangent space** at $p \in M$ as

$$T_p^*(M) = Hom(T_pM, \mathbb{R}) = (T_pM)^*. \tag{5.4}$$

The previously discussed spaces T_pM and T_p^*M, when taken together for all points of M, themselves have a manifold structure and are useful examples of mathematical objects called **vector bundles**. We will see more about them much later but for now, let's discuss tangent and cotangent bundles.

5.2 Tangent and cotangent bundle

Definition 5.6. The **tangent bundle** TM of a smooth manifold M is the (disjoint) union of all of its tangent spaces:

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

• The map $\pi : TM \to M$ which maps a tangent space $T - pM$ to the point p is called the **projection map.** Clearly, π is surjective.
• The subset of TM which consists of the zero vectors $0 \in T_pM, \ p \in M$ is called the **zero-section** of TM.
• The tangent space $T_pM \subset TM$ is called the **fiber** over $p \in M$.

Thus, every fiber of the tangent bundle is an n-dimensional real vector space. Even though, TM looks like an arbitrary set right now, the result below proves that not only it has a topology inherited from M, but it is a smooth manifold in its own right.

Remark 5.7. An arbitray point of TM looks like (p, v) with $p \in M$ and $v \in T_pM$ where the latter is the tangent space at the point p.

Theorem 5.8. *Let M^n be a smooth manifold. The tangent bundle TM can be endowed naturally with a smooth structure from M making it into a $2n$-dimensional smooth manifold. With respect to the smooth structure, the projection map $\pi : TM \to M$, the inclusion map $i : M \hookrightarrow TM$ as the zero-section and the natural inclusions $i : T_pM \hookrightarrow TM, p \in M$ are all smooth maps.*

Proof. The first step in describing a smooth structure on TM is to provide an atlas.

Claim 5.9. *Let (U, ϕ) be a chart on M. Then this gives a chart $(TU, T\phi)$ on TM where $TU = \underset{p \in U}{T_pM}$ and $T\phi : TU \to \mathbb{R}^{2m}$ is defined using the linear isomorphisms $d\phi_p : T_pM \to \mathbb{R}^n$ by*

$$TU \supset T_pM \ni X \mapsto (\phi(p), d\phi_p(X)) \in \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}.$$

If (V, ψ) is another chart on M, then the transition map between the charts $(TU, T\phi)$ and $(TV, T\psi)$ is given by

$$T\psi \circ T\phi^{-1}(p, v) = (\psi \circ \phi^{-1}(p), D(\psi \circ \phi^{-1})(p)v).$$

Proof of Claim 5.9. Exercise.

We continue with the proof of the theorem. We endow TM with the unique maximal smooth atlas containing all the charts of the form $(TU, T\phi)$ where (U, ϕ) are charts on M, from the Claim 5.9. The second countability and Hausdorff property for TM follows from the corresponding properties of M and the fact that countable union of countable sets is again countable. Thus, TM is a $2n$-dimensional smooth manifold.

Consider the map $\pi : TM \to M$, $(p, T_pM) \mapsto p$. Consider a chart $(TU, T\phi) \ni (p, T_pM)$ with chart (U, ϕ) containing the point $\pi(p)$. Then

$$\phi \circ \pi \circ T\phi^{-1}(p, v) = \phi(p)$$

which is clearly a smooth map. Similarly, both the inclusions are smooth maps. □

Definition 5.10. The **cotangent bundle** of a manifold M is again a $2n$-dimensional smooth manifold as is defined as

$$T^*M = \bigcup_{p \in M} T_p^*M.$$

The cotangent bundle T^*M is also smooth manifold of dimension $2n$ and this can be proved in a similar manner as the case for the tangent bundle. We will learn more about the tangent and cotangent bundles later but we will keep using them.

We can now define the differential of a smooth map.

Definition 5.11. Let M and N be two smooth manifolds and let $f : M \to N$ be a smooth map. The **differential of** f, df is the map

$$df : TM \to TN : [\gamma] \mapsto [f \circ \gamma].$$

The restriction of df at every point is precisely the map df_p described in Lemma 5.4.

The proof of Lemma 5.4 also shows that if M, N and f are smooth then so is df.

Using the formulation of df described in the definition above as a map btween tangent bundles, we can now give a simole proof the very important chain rule.

Proposition 5.12 (Chain Rule). *For smooth maps $f : M \to N$, $g : N \to P$ between smooth manifolds, we have*

$$d(g \circ f) = dg \circ df : TM \to TP.$$

At every point $p \in M$, $f(p) \in N$, this reads as

$$d(g \circ f)_p = dg_{f(p)} \circ df_p. \tag{5.5}$$

Proof. We prove (5.5). Let $\gamma : (-\varepsilon, \varepsilon) \to M$ with $\gamma(0) = p$, we have

$$\begin{aligned}
d(g \circ f)_p(\gamma'(0)) &= \frac{d}{dt}((g \circ f) \circ \gamma)_{|t=0} \\
&= \frac{d}{dt}(g \circ (f \circ \gamma))_{|t=0} \\
&= dg_{f(p)}((f \circ \gamma)'(0)) \quad \text{(chain rule in } \mathbb{R}^n\text{)} \\
&= dg_{f(p)}(df_p(\gamma'(0))).
\end{aligned}$$

□

A simple corollary of the previous proposition is the following.

Corollary 5.13. *If $f : M \to N$ is a diffeomorphism then so is $df : TM \to TN$ and moreover,*

$$(df)^{-1} = d(f^{-1}) : TN \to TM.$$

Proof. Notice that $d(id_M) : TM \to TM$ is just the map id_{TM} and hence the result follows from using the chain rule to

$$id_{TM} = d(f \circ f^{-1}) = df \circ d(f^{-1}).$$

□

We also make the following definition.

Definition 5.14. Let $k \in \mathbb{N} \cup \{\infty\}$. A surjective map $f : M \to N$ between smooth manifolds is called a **local C^k-diffeomorphism** if for all $p \in M$ there exists a neighbourhood $U \ni p$ and a neighbourhood $V \ni f(p)$ such that $f_U : U \to V$ is a C^k-diffeomorphism.

Clearly, every C^k-diffeomorphism is a local C^k-diffeomorphism but the opposite is not necessarily true.

Example 5.15. Let $f : \mathbb{R} \to S^1$ be the map $f(t) = e^{2\pi it}$. Then f is a surjective map but is not injective, so it can never be a diffeomorphism. But it a local C^∞-diffeomorphism. Let $t_0 \in \mathbb{R}$ and consider $U = (t_0 - \pi, t_0 + \pi)$ and $V = S^1 \setminus \{-f(t_0)\}$.

Also notice that if f is a local diffeomorphism then $df_p : T_pM \to T_{f(p)}N$ is a linear isomorphism and as a result dim(T_pM) =dim$(T_{f(p)}M)$ and hence dim$M$ =dim(N).

6. Derivations and tangent vectors

In this lecture, we see another way of defining tangent vectors and also talk about directional derivatives which is the analogue of the notion of partial derivatives from multi-variable calculus. This will be useful in doing actual computations.

Notation. We will denote by $C^k(M) = \{f : M \to \mathbb{R} \mid f \text{ is } C^k\}$. We also denote by $C_p^\infty = \bigcup_{U \text{ open},\ p\in U} C^\infty(U)$.

Definition 6.1 (Directional derivatives)**.** Let M be a smooth manifold and let $p \in M$. Let $\gamma'(0) \in T_pM$ and $f \in C^\infty(M)$. We define

$$\partial_{\gamma'(0)} f = df_p(\gamma'(0)) = \frac{d}{dt}(f \circ \gamma)_{|t=0} \in \mathbb{R},$$

and call it the **directional derivative of** f in the direction of the tangent vector $\gamma'(0)$.

Definition 6.2 (Derivations)**.** A map $\partial : C_p^\infty \to \mathbb{R}$ is called a derivation at p if the following holds.

1. Locality: If $\tilde{U} \subset U$ is open, $p \in \tilde{U}$ and $f \in C^\infty(U)$ then
$$\partial f = \partial(f_{\tilde{U}}).$$
2. $\mathbb{R}$-linearity: if $a, b \in \mathbb{R}$ and $f, g \in C_p^\infty$ then
$$\partial(af + bg) = a\partial f + b\partial g.$$
3. Leibniz rule: If $f, g \in C_p^\infty$ then
$$\partial(fg) = g\partial f + f\partial g.$$

Clearly, the usual partial derivatives $\frac{\partial}{\partial x^i}$, $i = 1, ..., n$ in $\mathbb{R}^n$ are derivations. In fact, the definition was wade to prove the following.

Proposition 6.3. *The directional derivatives $\partial_{\gamma'(0)}$ on a smooth manifold are derivations.*

Proof. We just check the Leibniz rule which essentially follows from the Leibniz rule in $\mathbb{R}^n$. We have

$$\begin{aligned}
\partial_{\gamma'(0)}(fg) &= \frac{d}{dt}((fg) \circ \gamma)_{|t=0} \\
&= \frac{d}{dt}((f \circ \gamma)(g \circ \gamma))_{|t=0} \\
&= (g \circ \gamma)(0)\frac{d}{dt}(f \circ c)_{|t=0} + (f \circ \gamma)(0)\frac{d}{dt}(g \circ \gamma)_{|t=0} \qquad \text{(Leibiniz rule in Euclidean spaces)} \\
&= g(p)\partial_{\gamma'(0)} f + f(p)\partial_{\gamma'(0)} g.
\end{aligned}$$

□

We denote the set of **all derivations at** p **by** $\mathrm{Der}(C_p^\infty)$. This space is a real vector space via

$$(\alpha\partial_1 + \beta\partial_2)(f) = \alpha\partial_1 f + \beta\partial_2 f.$$

We want another way to understand tangent vectors and the next lemma exactly starts with that fact.

Lemma 6.4. *Let M be a smooth manifold and let $p \in M$. The map*

$$\partial : T_pM \to \mathrm{Der}(C_p^\infty), \qquad \gamma'(0) \mapsto \partial_{\gamma'(0)},$$

is linear.

Proof. Let $\phi : U \to V$ be a chart of M with $p \in U$. We will have to use the vector structure on T_pM to check the linearity of the map, which is clearly a well-defined map. That is, we have to show that

$$\partial \circ (d\phi|_p)^{-1}$$

is linear. Assume $v \in \mathbb{R}^n$ and put

$$\gamma(t) := \phi^{-1}(\phi(p) + tv).$$

For $f \in C_p^\infty$, we compute,

$$\begin{aligned}
\big(\partial \circ (d\phi|_p)^{-1}(v)\big)(f) &= df|_p\big((d\phi|_p)^{-1}(v)\big) \\
&= df|_p(\gamma'(0)) \\
&= \frac{d}{dt}(f \circ \gamma(t))\bigg|_{t=0} \\
&= \frac{d}{dt}\big(f \circ \phi^{-1}(\phi(p) + tv)\big)\bigg|_{t=0} \\
&= \big\langle D(f \circ \phi^{-1})|_{\phi(p)}, v\big\rangle .
\end{aligned}$$

which is an expression that is linear in v. $\square$

Thus, we can view tangent vectors are derivations and the way to think about it is that you take a tangent vector and take the directional derivative in its direction.

Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{R}^n$. Then $(d\phi|_p)^{-1}(e_1), \ldots, (d\phi|_p)^{-1}(e_n)$ form a basis of T_pM. What does this basis looks like? We compute,

$$\partial_{(d\phi|_p)^{-1}(e_j)}(f) = \big\langle D(f \circ \phi^{-1})|_{\phi(p)}, e_j\big\rangle = \frac{\partial(f \circ \phi^{-1})}{\partial x^j}\bigg|_{\phi(p)} =: \frac{\partial f}{\partial x^j}\bigg|_p$$

where we are abusing notation and still writing $f \circ \phi^{-1} = f$ which is the coordinate representation of f in the chart (U, ϕ).

This tells us that if (U, ϕ) is any chart on M and we write $(x^1, \ldots, x^n)$ as the coordinates on U then

For every chart (U, ϕ), we have the derivations

$$\frac{\partial}{\partial x^1}\bigg|_p, \ldots, \frac{\partial}{\partial x^n}\bigg|_p .$$

In fact, this also allows us to prove the following result which tells that the tangent vectors as equivalence classes is the same as derivations.

Lemma 6.5. *Let M be a differentiable manifold and let $p \in M$. Then the map*

$$\partial : T_pM \to \mathrm{Der}(C_p^\infty), \qquad \gamma'(0) \mapsto \partial_{\gamma'(0)},$$

is an **isomorphism.** *In particular, every derivation is a directional derivative and for every chart $\phi : U \to V$ with $p \in U$*

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

is a basis of $\mathrm{Der}(C_p^\infty)$.

Proof. We have already proved that ∂ is indeed a linear map between vector spaces and that given a chart, we get the derivations $\left.\frac{\partial}{\partial x^1}\right|_p, \ldots, \left.\frac{\partial}{\partial x^n}\right|_p$. We now show that

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

form a basis of $\mathrm{Der}(C_p^\infty)$. Namely, then we know that the linear map ∂ maps the basis

$$(d\phi|_p)^{-1}(e_1), \ldots, (d\phi|_p)^{-1}(e_n)$$

of T_pM onto the basis

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

of $\mathrm{Der}(C_p^\infty)$ and is hence an isomorphism.

a) **Linear Independence:** Let

$$\sum_{i=1}^n \alpha^i \left.\frac{\partial}{\partial x^i}\right|_p = 0.$$

We have to show:

$$\alpha^1 = \cdots = \alpha^n = 0.$$

Choose

$$f = x^j.$$

Then

$$0 = \sum_{i=1}^n \alpha^i \left.\frac{\partial x^j}{\partial x^i}\right|_p = \alpha^j \qquad \text{for } j = 1, \ldots, n.$$

b) **Spanning set:** Let $\delta \in \mathrm{Der}(C_p^\infty)$. Set

$$\alpha^j := \delta(x^j) = \delta(\phi^j), \qquad \text{for } j = 1, \ldots, n.$$

We will show that

$$\delta = \sum_{j=1}^n \alpha^j \cdot \left.\frac{\partial}{\partial x^j}\right|_p.$$

Exercise: Prove that any derivation vanish on constant functions.

Let $f \in C_p^\infty$, and assume, without loss of generality that

$$\phi(\tilde{\tilde{U}}) = B(\phi(p), r).$$

More precisely, since $f \in C^\infty(\tilde{U})$ with $p \in \tilde{U}$ open, so we choose a neighborhood $\tilde{\tilde{U}}$ of p with

$$\tilde{\tilde{U}} \subset \tilde{U} \cap U$$

and assume that its image is an open ball. If M were $\mathbb{R}^n$ then to produce a function we would have used Taylor's theorem. In fact,

Claim 6.6. *For any function $h \in C^\infty(B(q,r))$, there exist $g_1, \dots, g_n \in C^\infty(B(q,r))$ with*

(i)

$$h(x) = h(q) + \sum_{i=1}^{n} (x^i - q^i) g_i(x)$$

and

(ii)

$$\frac{\partial h}{\partial x^i}(q) = g_i(q).$$

Proof of the Claim: For $x \in B(q,r)$ set

$$w_x : [0,1] \to \mathbb{R}, \qquad w_x(t) := h(tx + (1-t)q).$$

Then,

$$\begin{aligned} h(x) - h(q) &= w_x(1) - w_x(0) \\ &= \int_0^1 \dot{w}_x(t)\, dt \qquad \text{(Fundamental theorem of Calculus)} \\ &= \int_0^1 \sum_{i=1}^{n} \frac{\partial h}{\partial x^i}\bigg|_{tx+(1-t)q} \cdot (x^i - q^i)\, dt \\ &= \sum_{i=1}^{n} (x^i - q^i) \underbrace{\int_0^1 \frac{\partial h}{\partial x^i}\bigg|_{tx+(1-t)q} dt}_{=: g_i(x)} . \end{aligned}$$

We now work backwards and define g_i as the function in the last line of the equation, thus proving (i) . Once we have (i), (ii) follows via differentiation. $\square$

We use the claim proved above with $h = f \circ \phi^{-1}$. Thus, there exist $g_1, \dots, g_n \in C^\infty(B(x(p), r))$ such that

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

and

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

It follows that

$$\begin{aligned}
\delta(f) &\overset{\text{locality}}{=} \delta(f|_{\tilde{U}}) \\
&= \delta\left(f(p) + \sum_{i=1}^{n} (\phi^i - \phi^i(p))(g_i \circ \phi)\right) \\
&= \sum_{i=1}^{n} \delta\big((\phi^i - \phi^i(p))(g_i \circ \phi)\big) \\
&\overset{\text{Leibniz rule}}{=} \sum_{i=1}^{n} \left(\delta(\phi^i - \phi^i(p)) \cdot g_i(\phi(p)) + \underbrace{(\phi^i - \phi^i(p))|_p}_{=0} \, \delta(g_i \circ \phi)\right) \\
&\overset{\text{linearity}}{=} \sum_{i=1}^{n} \delta(x^i) \, g_i(x(p)) \\
&= \sum_{i=1}^{n} \alpha^i \cdot \frac{\partial f}{\partial x^i}\bigg|_p .
\end{aligned}$$

□

From now on we identify T_pM with $\mathrm{Der}(C_p^\infty)$ via the isomorphism ∂. For example, we write for $\xi \in T_pM$

$$\xi = \sum_{i=1}^{n} \xi^i \, \frac{\partial}{\partial x^i}\bigg|_p$$

instead of

$$\partial_\xi = \sum_{i=1}^{n} \xi^i \, \frac{\partial}{\partial x^i}\bigg|_p \qquad \text{and} \qquad \xi = \sum_{i=1}^{n} \xi^i (d\phi|_p)^{-1}(e_i)$$

where

$$(\xi^1, \ldots, \xi^n)^\top = d\phi|_p(\xi).$$

Change of coordinates:

How do the coefficients $\xi^1, \ldots, \xi^n$ of a tangent vector change, if we replace the chart (U, ϕ) by another chart (V, ψ)?

Let $\xi \in T_pM$, let (U, ϕ) and (V, ψ) be charts, both containing p. We express ξ with respect to both charts

$$\xi = \sum_{i=1}^{n} \xi^i \, \frac{\partial}{\partial x^i}\bigg|_p = \sum_{j=1}^{n} \eta^j \, \frac{\partial}{\partial y^j}\bigg|_p .$$

Now we want to compute the coefficients ξ^i in terms of the η^j and vice versa. Using the Chain Rule, Prop 5.12, we compute

$$\begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix} = d\phi|_p(\xi) = (d\phi|_p)\left((d\psi|_p)^{-1}\begin{pmatrix} \eta^1 \\ \vdots \\ \eta^n \end{pmatrix}\right) = D(\phi \circ \psi^{-1})|_{\psi(p)} \begin{pmatrix} \eta^1 \\ \vdots \\ \eta^n \end{pmatrix} .$$

Similarly,

$$\begin{pmatrix} \eta^1 \\ \vdots \\ \eta^n \end{pmatrix} = D(\psi \circ \phi^{-1})|_{\phi(p)} \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix} .$$

Thus

$$\eta^j = \sum_{i=1}^{n} \frac{\partial(\psi^j \circ \phi^{-1})}{\partial x^i}\bigg|_{phi(p)} \cdot \xi^i \tag{1.6}$$

Let us look at the special case

$$\xi = \frac{\partial}{\partial x^i}\bigg|_p ,$$

that is,

$$(\xi^1, \dots, \xi^n)^\top = e_i.$$

We get

$$\begin{aligned}
\frac{\partial}{\partial x^i}\bigg|_p &= \sum_{j=1}^{n} \eta^j \, \frac{\partial}{\partial y^j}\bigg|_p \\
&= \sum_{j=1}^{n}\sum_{k=1}^{n} \xi^k \, \frac{\partial(y^j \circ x^{-1})}{\partial x^k}\bigg|_{x(p)} \cdot \frac{\partial}{\partial y^j}\bigg|_p \\
&= \sum_{j=1}^{n} \frac{\partial(y^j \circ x^{-1})}{\partial x^i}\bigg|_{x(p)} \cdot \frac{\partial}{\partial y^j}\bigg|_p ,
\end{aligned}$$

hence

$$\frac{\partial}{\partial x^i}\bigg|_p = \sum_{j=1}^{n} \frac{\partial(y^j \circ x^{-1})}{\partial x^i}\bigg|_{x(p)} \cdot \frac{\partial}{\partial y^j}\bigg|_p \tag{1.7}$$

If we use the Einstein summation convention, meaning that when an index appears twice in an expression, once as an upper index and once as a lower index, then summation over this index is understood, then the above equation could be written as

$$\frac{\partial}{\partial x^i}\bigg|_p = \frac{\partial(y^j \circ x^{-1})}{\partial x^i}\bigg|_{x(p)} \cdot \frac{\partial}{\partial y^j}\bigg|_p$$

or even shorter as

$$\frac{\partial}{\partial x^i} = \frac{\partial y^j}{\partial x^i} \cdot \frac{\partial}{\partial y^j}.$$

6.1 Differential of a map in coordinates

Recall that we defined the differential of a map f in terms of using the notion of tangent vectors via curves as well as using derivations. We would like to do some explicit computations and would like to see how the term df for some $f : M \to N$ looks like in coordinates. Let's try to understand this in the spacial case when $M = \mathbb{R}^n$ and $N = \mathbb{R}^m$. Let $f : U \subset \mathbb{R}^n \to V \subset \mathbb{R}^m$ be a smooth map. For any $p \in U$, we expect df to be a $m \times n$ matrix. We determine the matrix explicitly, in terms of the standard coordinate bases. Using $(x^1, \dots, x^n)$ to denote the coordinates in the domain and $(y^1, \dots, y^m)$ to denote those in the codomain, we use the chain rule to compute the action of df_p on a typical basis vector as follows:

$$\begin{aligned}
df_p\left(\frac{\partial}{\partial x^i}\bigg|_p\right) h &= \frac{\partial}{\partial x^i}\bigg|_p (h \circ f) \\
&= \frac{\partial h}{\partial y^j}(f(p)) \, \frac{\partial f^j}{\partial x^i}(p) \\
&= \left(\frac{\partial f^j}{\partial x^i}(p) \, \frac{\partial}{\partial y^j}\bigg|_{f(p)}\right) h.
\end{aligned}$$

Thus,

$$df_p\left(\frac{\partial}{\partial x^i}\bigg|_p\right) = \frac{\partial f^j}{\partial x^i}(p)\ \frac{\partial}{\partial y^j}\bigg|_{f(p)}. \tag{6.1}$$

This says that the matrix of the linear map df_p in terms of the coordinate bases is

$$\begin{pmatrix} \frac{\partial f^1}{\partial x^1}(p) & \cdots & \frac{\partial f^1}{\partial x^n}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial f^m}{\partial x^1}(p) & \cdots & \frac{\partial f^m}{\partial x^n}(p) \end{pmatrix},$$

which is precisely the Jacobian matrix of f at p. Therefore, in this case,

$$df_p : T_p\mathbb{R}^n \to T_{f(p)}\mathbb{R}^m$$

corresponds to the total derivative

$$Df(p) : \mathbb{R}^n \to \mathbb{R}^m.$$

Now consider the more general case of a smooth map $F : M \to N$ between smooth manifolds. Choosing smooth coordinate charts (U, φ) for M containing p and (V, ψ) for N containing $F(p)$, we obtain the coordinate representation

$$\hat{F} = \psi \circ F \circ \varphi^{-1} : \varphi(U \cap F^{-1}(V)) \to \psi(V).$$

Let $\hat{p} = \varphi(p)$ denote the coordinate representation of p. By the computation above, $d\hat{F}_{\hat{p}}$ is represented with respect to the standard coordinate bases by the Jacobian matrix of $\hat{F}$ at $\hat{p}$. Using the fact that $F \circ \varphi^{-1} = \psi^{-1} \circ \hat{F}$, we compute

$$dF_p\left(\frac{\partial}{\partial x^i}\bigg|_p\right) = dF_p\left(d(\varphi^{-1})_{\hat{p}}\left(\frac{\partial}{\partial x^i}\bigg|_{\hat{p}}\right)\right) = d(\psi^{-1})_{\hat{F}(\hat{p})}\left(d\hat{F}_{\hat{p}}\left(\frac{\partial}{\partial x^i}\bigg|_{\hat{p}}\right)\right).$$

latex

$$\begin{aligned} dF_p\left(\frac{\partial}{\partial x^i}\bigg|_p\right) &= d(\psi^{-1})_{\hat{F}(\hat{p})}\left(\frac{\partial \hat{F}^j}{\partial x^i}(\hat{p})\ \frac{\partial}{\partial y^j}\bigg|_{\hat{F}(\hat{p})}\right) \\ &= \frac{\partial \hat{F}^j}{\partial x^i}(\hat{p})\ \frac{\partial}{\partial y^j}\bigg|_{F(p)}. \end{aligned} \tag{3.10}$$

Thus, dF_p is represented in coordinate bases by the Jacobian matrix of (the coordinate representative of) F. Sometimes, the differential of a map is also called the **pushforward of F** and is denoted as F_*.

7. Vector fields

Next we want to introduce vector fields. Vector fields are maps which associate to each point of a manifold a tangent vector in the corresponding tangent space. We have the following:

Definition 7.1. A map $\xi : M \to TM$ is called a **vector field** on M, if for every $p \in M$ we have

$$\pi(\xi(p)) = p. \tag{7.1}$$

We can use the coordinate representation of a chart to write the local coordinate description of any vector field. Let $\phi : U \to V$ be a chart of M. A vector field ξ on U is characterized by the functions which are coefficients of the tangent vector $\xi(p),\ p \in M$. More precisely, we have

$$\xi^1, \ldots, \xi^n : V \to \mathbb{R}$$

for which
$$\xi(p) = \sum_{i=1}^{n} \xi^i(\phi(p)) \left.\frac{\partial}{\partial x^i}\right|_p .$$
For the chart ϕ of M we consider the corresponding chart Φ_ϕ on TM. Using the notation as before, we see that ξ corresponds to the map
$$v \mapsto (v, \xi^1(v), \ldots, \xi^n(v)), \ \ v \in V = \phi(U) \subset \mathbb{R}^n.$$
Thus ξ is C^k on U if and only if the coefficient functions $\xi^1, \ldots, \xi^n$ are C^k on V.

8. Submanifolds

As with any mathematical structure, once we have seen and understood the theory of manifolds, it makes sense to ask whether subsets of manifolds also admit a smooth structure. We will see that the notion of *immersions, submersions and embeddings* are the key properties which allow us to study special subsets of a manifold which are manifolds in their own right.

8.1 Submersions, immersions and embeddings

Definition 8.1. Let M and N are smooth manifolds and let $F\colon M \to N$ be a smooth point. For $p \in M$, we define the **rank of F at p** to be the rank of the linear map

$$dF_p\colon T_pM \to T_{F(p)}N.$$

In other words, it is the rank of the Jacobian matrix of F in any smooth chart, or the dimension of $\operatorname{Im} dF_p \subseteq T_{F(p)}N$.
If F has the same rank r at every point, we say that it has **constant rank**, and write $\operatorname{rank} F = r$.

We know by the rank-nullity theorem that rank of $F \leq \min\{\dim M, \dim N\}$. If the rank of dF_p is equal to this upper bound, we say that F **has full rank at** p, and if F has full rank everywhere, we say F **has full rank.**

Definition 8.2. A smooth map $F\colon M \to N$ is called a **smooth submersion** if its differential is surjective at each point (or equivalently, if $\operatorname{rank} F = \dim N$).

It is called a **smooth immersion** if its differential is injective at each point (equivalently, $\operatorname{rank} F = \dim M$).

Proposition 8.3. *Suppose $F\colon M \to N$ is a smooth map and $p \in M$. If dF_p is surjective, then p has a neighborhood U such that $F|_U$ is a submersion. If dF_p is injective, then p has a neighborhood U such that $F|_U$ is an immersion.*

Proof. Choose any smooth coordinates for M near p and for N near $F(p)$, either hypothesis means that Jacobian matrix of F in coordinates has full rank at p. The result now follows from the fact that the set of $m \times n$ matrices of full rank is an open subset of $\mathrm{M}(m \times n, \mathbb{R})$[1], so by continuity, the Jacobian of F has full rank in some neighbourhood of p. $\square$

Example 8.4 (Submersions and Immersions)**.**

(a) Suppose $M_1, \dots, M_k$ are smooth manifolds. Then each of the projection maps $\pi_i\colon M_1 \times \cdots \times M_k \to M_i$ is a smooth submersion. In particular, the projection $\pi\colon \mathbb{R}^{n+k} \to \mathbb{R}^n$ onto the first n coordinates is a smooth submersion.

(b) If $\gamma\colon J \to M$ is a smooth curve in a smooth manifold M with or without boundary, then γ is a smooth immersion if and only if $\gamma'(t) \neq 0$ for all $t \in J$.

(c) Exercise: If M is a smooth manifold and its tangent bundle is denoted by TM prove that the projection $\pi\colon TM \to M$ is a smooth submersion.

[1]Let f be the function taking $A \in M(m \times n)$ to the $f(A) = \Sigma_B\,|\det(B)|$ where the sum is running over all $m \times m$ (assuming $m < n$) sub-matrices of A. Clearly, f is continuous so $f^{-1}(\mathbb{R} \setminus \{0\})$ is an open set is $M(m \times n)$

(d) The smooth map $X\colon \mathbb{R}^2 \to \mathbb{R}^3$ given by

$$X(u,v) = \big((2+\cos 2\pi u)\cos 2\pi v,\ (2+\cos 2\pi u)\sin 2\pi v,\ \sin 2\pi u\big)$$

is a smooth immersion of $\mathbb{R}^2$ into $\mathbb{R}^3$ whose image is the doughnut-shaped surface obtained by revolving the circle $(y-2)^2+z^2=1$ in the (y,z)-plane about the z-axis.

Exercise 8.5. Show that a composition of smooth submersions is a smooth submersion, and a composition of smooth immersions is a smooth immersion. Give a counterexample to show that a composition of maps of constant rank need not have constant rank.

We already saw the definition of a smooth local diffeomorphism. We also saw that if $f : M \to N$ is a local diffeomorphism then the linear map $df_p : T_pM \to T_{f(p)}N$ is an isomorphism. The converse of this statement is also true and is known as the **inverse function theorem.**

Theorem 8.6 (Inverse Function Theorem for Manifolds). *Suppose M and N are smooth manifolds, and $f\colon M \to N$ is a smooth map. If $p \in M$ is a point such that df_p is an isomorphism, then there are neighbourhoods U_0 of p and V_0 of $f(p)$ such that $f|_{U_0}\colon U_0 \to V_0$ is a diffeomorphism.*

Proof. The fact that df_p is bijective implies that M and N have the same dimension, say n. Choose smooth charts (U,φ) centered at p and (V,ψ) centered at $f(p)$, with $f(U) \subseteq V$. Then $\widehat{f} = \psi \circ f \circ \varphi^{-1}$ is a smooth map from the open subset $\widehat{U} = \varphi(U) \subseteq \mathbb{R}^n$ into $\widehat{V} = \psi(V) \subseteq \mathbb{R}^n$, with $\widehat{f}(p) = 0$. Because φ and ψ are diffeomorphisms, the differential $d\widehat{f}_0 = d\psi_{f(p)} \circ df_p \circ d(\varphi^{-1})_0$ is nonsingular. The inverse function theorem for Euclidean spaces implies that there are connected open subsets $\widehat{U}_0 \subseteq \widehat{U}$ and $\widehat{V}_0 \subseteq \widehat{V}$ containing 0 such that $\widehat{f}$ restricts to a diffeomorphism from $\widehat{U}_0$ to $\widehat{V}_0$. Then $U_0 = \varphi^{-1}(\widehat{U}_0)$ and $V_0 = \psi^{-1}(\widehat{V}_0)$ are connected neighborhoods of p and $f(p)$, respectively, and it follows by composition that $f|_{U_0}$ is a diffeomorphism from U_0 to V_0. □

The next result shows the relationship between immersions, submersions and local diffeomorphisms.

Proposition 8.7. *Suppose M and N are smooth manifolds and let $f\colon M \to N$ be a map.*

1. *f is a local diffeomorphism if and only if it is both a smooth immersion and a smooth submersion.*
2. *If $\dim M = \dim N$ and f is either a smooth immersion or a smooth submersion, then it is a local diffeomorphism.*

Proof. Suppose first that f is a local diffeomorphism. Given $p \in M$, there is a neighbourhood U of p such that f is a diffeomorphism from U to $f(U)$ and $df_p\colon T_pM \to T_{f(p)}N$ is an isomorphism. Thus $\operatorname{rank} f = \dim M = \dim N$, so f is both a smooth immersion and a smooth submersion. Conversely, if f is both a smooth immersion and a smooth submersion, then df_p is an isomorphism at each $p \in M$, and the inverse function theorem Theorem refthm:IFT shows that p has a neighborhood on which f restricts to a diffeomorphism onto its image. This proves 1.

To prove 2., note that if M and N have the same dimension, then either injectivity or surjectivity of df_p implies bijectivity, so f is a smooth submersion if and only if it is a smooth immersion, and thus 2 follows from 1. □

8.2 The Rank Theorem

We see example of a submersion and immersion.

Example 8.8. Canonical Projection $\pi : \mathbb{R}^m \to \mathbb{R}^n$ (where $m > n$)

$$\pi(x^1, x^2, \ldots, x^m) = (x^1, x^2, \ldots, x^n)$$

is a submersion. The differential at any point p is the linear map:

$$d\pi_p : \mathbb{R}^m \to \mathbb{R}^n$$

represented by the $n \times m$ matrix

$$d\pi_p = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots & & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \end{bmatrix}$$

This matrix has rank n everywhere $\implies d\pi_p$ is surjective $\implies \pi$ is a submersion. Also notice that each fiber $\pi^{-1}(c) \cong \mathbb{R}^{m-n}$ is an affine subspace.

Example 8.9. The Canonical Inclusion $\iota : \mathbb{R}^m \to \mathbb{R}^n$ (where $m < n$)

$$\iota(x^1, \ldots, x^m) = (x^1, \ldots, x^m, 0, \ldots, 0)$$

is an immersion. The differential at any point p is:

$$d\iota_p = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \in M_{n\times m}(\mathbb{R})$$

This is an $n \times m$ matrix of rank m everywhere $\implies d\iota_p$ is injective.

The most important fact about constant-rank maps is the following consequence of the inverse function theorem, which says that a constant-rank smooth map can be placed locally into a particularly simple canonical form by a change of coordinates and so in a way the examples provided above of submersions and immersions are the model examples.

We first recall the following canonical form result for surjective linear maps.

Theorem 8.10 (Canonical Form for a Linear Map)**.** *Suppose V and W are finite-dimensional vector spaces, and $T : V \to W$ is a linear map of rank r. Then there are bases for V and W with respect to which T has the following matrix representation*

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix},$$

where I_r is the $r \times r$ identity matrix.

Theorem 8.11 (Rank Theorem). *Suppose M and N are smooth manifolds of dimensions m and n, respectively, and $F\colon M \to N$* ***is a smooth map with constant rank*** *r. For each $p \in M$ there exist smooth charts $(U, \varphi) \ni p$ for M and $(V, \psi) \ni F(p)$ for N such that $F(U) \subseteq V$, in which F* ***has a coordinate representation of the form***

$$\widehat{F}\big(x^1, \ldots, x^r, x^{r+1}, \ldots, x^m\big) = \big(x^1, \ldots, x^r, 0, \ldots, 0\big). \tag{8.1}$$

If F is a smooth submersion then the representation becomes

$$\widehat{F}\big(x^1, \ldots, x^n, x^{n+1}, \ldots, x^m\big) = \big(x^1, \ldots, x^n\big), \tag{8.2}$$

and if F is a smooth immersion, it is

$$\widehat{F}\big(x^1, \ldots, x^m\big) = \big(x^1, \ldots, x^m, 0, \ldots, 0\big). \tag{8.3}$$

Proof. Because the theorem is local, after choosing smooth coordinates we can replace M and N by open subsets $U \subseteq \mathbb{R}^m$ and $V \subseteq \mathbb{R}^n$. Since $DF(p)$ has rank r hence its matrix representation has some $r \times r$ submatrix with nonzero determinant. By reordering the coordinates, w.l.o.g., we assume that it is the upper left submatrix, $(\partial F^i/\partial x^j)$ for $i, j = 1, \dots, r$. We write the standard coordinates on $\mathbb{R}^m$ and $\mathbb{R}^n$ as $(x, y) = (x^1, \dots, x^r, y^1, \dots, y^{m-r})$ in $\mathbb{R}^m$ and $(v, w) = (v^1, \dots, v^r, w^1, \dots, w^{n-r})$ in $\mathbb{R}^n$. By initial translations of the coordinates, we may assume without loss of generality that $p = (0, 0)$ and $F(p) = (0, 0)$.

If we write $F(x, y) = \big(Q(x, y), R(x, y)\big)$ for some smooth maps $Q: U \to \mathbb{R}^r$ and $R: U \to \mathbb{R}^{n-r}$, then our hypothesis is that $(\partial Q^i/\partial x^j)$ is nonsingular at $(0, 0)$.

Define $\varphi: U \to \mathbb{R}^m$ by $\varphi(x, y) = \big(Q(x, y), y\big)$. Its total derivative at $(0, 0)$ is

$$D\varphi(0,0) = \begin{pmatrix} \dfrac{\partial Q^i}{\partial x^j}(0,0) & \dfrac{\partial Q^i}{\partial y^j}(0,0) \\ 0 & \delta^i_j \end{pmatrix},$$

where δ^i_j is the **Kronecker delta**, which is defined by

$$\delta^i_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

The matrix $D\varphi(0, 0)$ is nonsingular because one block $\frac{\partial Q^i}{\partial x^j}(0, 0)$ is non-singular at $(0, 0)$ and the second diagonal block is just the identity matrix. Therefore, by the inverse function theorem Theorem 8.6, there are connected neighbourhoods U_0 of $(0, 0)$ and $\widetilde{U}_0$ of $\varphi(0, 0) = (0, 0)$ such that $\varphi: U_0 \to \widetilde{U}_0$ is a diffeomorphism. By shrinking U_0 and $\widetilde{U}_0$ if necessary, we may assume that $\widetilde{U}_0$ is an open cube.

Writing the inverse map as $\varphi^{-1}(x, y) = \big(A(x, y), B(x, y)\big)$ for some smooth functions $A: \widetilde{U}_0 \to \mathbb{R}^r$ and $B: \widetilde{U}_0 \to \mathbb{R}^{m-r}$, we compute

$$(x, y) = \varphi\big(A(x, y), B(x, y)\big) = \big(Q\big(A(x, y), B(x, y)\big), B(x, y)\big). \tag{8.4}$$

Comparing y components shows that $B(x, y) = y$, and therefore φ^{-1} has the form

$$\varphi^{-1}(x, y) = \big(A(x, y), y\big).$$

On the other hand, $\varphi \circ \varphi^{-1} = \mathrm{Id}$ implies $Q\big(A(x, y), y\big) = x$, and therefore $F \circ \varphi^{-1}$ has the form

$$F \circ \varphi^{-1}(x, y) = \big(x, \widetilde{R}(x, y)\big),$$

where $\widetilde{R}: \widetilde{U}_0 \to \mathbb{R}^{n-r}$ is defined by $\widetilde{R}(x, y) = R\big(A(x, y), y\big)$. The Jacobian matrix of this composite map at an arbitrary point $(x, y) \in \widetilde{U}_0$ is

$$D(F \circ \varphi^{-1})(x, y) = \begin{pmatrix} \delta^i_j & 0 \\ \dfrac{\partial \widetilde{R}^i}{\partial x^j}(x, y) & \dfrac{\partial \widetilde{R}^i}{\partial y^j}(x, y) \end{pmatrix}.$$

Since composing with a diffeomorphism does not change the rank of a map, this matrix has rank r everywhere in $\widetilde{U}_0$ because F was assumed to have rank r and φ^{-1} is a diffeomorphism. The first r columns are obviously linearly independent, so the rank can be r only if the derivatives $\partial \widetilde{R}^i/\partial y^j$ vanish identically on $\widetilde{U}_0$, which implies that $\widetilde{R}$ is actually independent of $(y^1, \dots, y^{m-r})$. Thus, if we let $S(x) = \widetilde{R}(x, 0)$, then we have

$$F \circ \varphi^{-1}(x, y) = \big(x, S(x)\big). \tag{8.5}$$

We complete the proof by defining an appropriate smooth chart in some neighborhood of $(0, 0) \in V$. Let $V_0 \subseteq V$ be the open subset defined by $V_0 = \{(v, w) \in V : (v, 0) \in \widetilde{U}_0\}$. Then V_0 is a neighborhood of $(0, 0)$. Because $\widetilde{U}_0$ is a cube and $F \circ \varphi^{-1}$ has the form (8.5), it follows that $F \circ \varphi^{-1}(\widetilde{U}_0) \subseteq V_0$, and therefore $F(U_0) \subseteq V_0$. Define $\psi: V_0 \to \mathbb{R}^n$ by $\psi(v, w) = \big(v, w - S(v)\big)$. This is a diffeomorphism onto its image, because its inverse is given explicitly by $\psi^{-1}(s,t) = \big(s, t + S(s)\big)$; thus (V_0, ψ) is a smooth chart. It follows from (8.5) that

$$\psi \circ F \circ \varphi^{-1}(x,y) = \psi\big(x, S(x)\big) = \big(x, S(x) - S(x)\big) = (x, 0),$$

which was to be proved. $\square$

The next corollary can be viewed as a more invariant statement of the rank theorem. It says that constant-rank maps are precisely the ones whose local behaviour is the same as that of their differentials.

Corollary 8.12. *Let M and N be smooth manifolds, let $F : M \to N$ be a smooth map, and suppose M is connected. Then the following are equivalent:*

(a) For each $p \in M$ there exist smooth charts containing p and $F(p)$ in which the coordinate representation of F is linear.

(b) F has constant rank.

Proof. First suppose F has a linear coordinate representation in a neighbourhood of each point. Since every linear map has constant rank, it follows that the rank of F is constant in a neighbourhood of each point, and thus by connectedness it is constant on all of M. Conversely, if F has constant rank, the rank theorem shows that it has the linear coordinate representation (8.1) in a neighbourhood of each point. $\square$

We now use the rank theorem, which is purely a local statement to say something about the global nature of smooth maps based on their rank. Before we do that, we recall the following topological notions.

Definition 8.13. Let X be a topological space.

(a) A subset $A \subseteq X$ is called **nowhere dense** if its closure has empty interior:

$$\operatorname{int}(\overline{A}) = \emptyset.$$

(b) A subset $A \subseteq X$ is called **meagre** (or **of the first category**) if it can be written as a countable union of nowhere dense sets:

$$A = \bigcup_{n=1}^{\infty} A_n, \qquad \operatorname{int}(\overline{A_n}) = \emptyset \quad \forall n \in \mathbb{N}.$$

(c) A subset $A \subseteq X$ is called **comeagre** (or **residual**, or **of the second category**) if its complement $X \setminus A$ is meagre.

Theorem 8.14 (Baire Category Theorem). *The following two classes of topological spaces are **Baire spaces**, meaning that every countable intersection of open dense subsets is dense:*

*(i) Every **complete metric space** (X, d) is a Baire space.*

*(ii) Every **locally compact Hausdorff space** X is a Baire space.*

Another may wo reformulate this is the following. Let X be either a complete metric space or a locally compact Hausdorff space. Then the following equivalent statements hold:

*(a) **(Open dense sets)** If $\{U_n\}_{n=1}^{\infty}$ is a countable collection of open dense subsets of X, then*

$$\bigcap_{n=1}^{\infty} U_n \neq \emptyset.$$

In fact, $\bigcap_{n=1}^{\infty} U_n$ is dense in X.

*(b) **(Nowhere dense sets)** If $\{A_n\}_{n=1}^{\infty}$ is a countable collection of nowhere dense subsets of X, then*

$$\operatorname{int}\left(\bigcup_{n=1}^{\infty} A_n\right) = \emptyset.$$

That is, X cannot be written as a countable union of nowhere dense sets.

*(c) **(Meagre sets)** No nonempty open subset of X is meagre. Equivalently, X itself is not meagre:*

$$X \neq \bigcup_{n=1}^{\infty} A_n \quad \text{whenever each } A_n \text{ is nowhere dense in } X.$$

*(d) **(Residual sets)** Every comeagre (residual) subset of X is dense in X.*

We now state and prove the Global Rank theorem.

Theorem 8.15 (Global Rank Theorem). *Let M and N be smooth manifolds, and suppose $F : M \to N$ is a smooth map of constant rank.*

(a) If F is surjective, then it is a smooth submersion.

(b) If F is injective, then it is a smooth immersion.

(c) If F is bijective, then it is a diffeomorphism.

Proof. Let $m = \dim M$, $n = \dim N$, and suppose F has constant rank r. To prove (a), assume that F is not a smooth submersion, which means that $r < n$. By the rank theorem, Theorem 8.11, for each $p \in M$ there are smooth charts (U, φ) for M centered at p and (V, ψ) for N centered at $F(p)$ such that $F(U) \subseteq V$ and the coordinate representation of F is given by (8.1).

Shrinking U if necessary, we may assume that it is a regular coordinate ball and $F(\bar{U}) \subseteq V$, where $\bar{U}$ is the closed coordinate ball and is the closure of the open set U. This implies that $F(\bar{U})$ is a compact subset of the set $\{y \in V : y^{r+1} = \cdots = y^n = 0\}$, so it is closed in N and contains no open subset of N; hence it is nowhere dense in N. Since every open cover of a manifold has a countable subcover, we can choose countably many such charts $\{(U_i, \varphi_i)\}$ covering M, with corresponding charts $\{(V_i, \psi_i)\}$ covering $F(M)$. Because $F(M)$ is equal to the countable union of the nowhere dense sets $F(\bar{U}_i)$, it follows from the Baire category theorem, Theorem 8.14, that $F(M)$ has empty interior in N, which means F cannot be surjective.

To prove (b), assume that F is not a smooth immersion, so that $r < m$. By the rank theorem, for each $p \in M$ we can choose charts on neighbourhoods of p and $F(p)$ in which F has the coordinate representation (8.1). It follows that $F(0, \dots, 0, \varepsilon) = F(0, \dots, 0, 0)$ for any sufficiently small ε, so F is not injective.

Finally, (c) follows from (a) and (b), because a bijective smooth map of constant rank is a smooth submersion by part (a) and a smooth immersion by part (b); so Proposition 8.7 implies that F is a local diffeomorphism, and because it is bijective, it is a diffeomorphism. □

8.3 Embeddings

One special kind of immersion is particularly important.

Definition 8.16. If M and N are smooth manifolds, a ***smooth embedding of*** M ***into*** N is a smooth immersion $F : M \to N$ that is also a topological embedding, i.e., a homeomorphism onto its image $F(M) \subseteq N$ in the subspace topology.

Example 8.17 (Smooth Embeddings)**.**

1. If M is a smooth manifold with and $U \subseteq M$ is an open submanifold, the inclusion map $U \hookrightarrow M$ is a smooth embedding.

2. If $M_1, \dots, M_k$ are smooth manifolds and $p_i \in M_i$ are arbitrarily chosen points, each of the maps $\iota_j : M_j \to M_1 \times \cdots \times M_k$ given by

$$\iota_j(q) = (p_1, \dots, p_{j-1}, q, p_{j+1}, \dots, p_k)$$

is a smooth embedding. In particular, the inclusion map $\mathbb{R}^n \hookrightarrow \mathbb{R}^{n+k}$ given by sending $(x^1, \dots, x^n)$ to $(x^1, \dots, x^n, 0, \dots, 0)$ is a smooth embedding.

Let us see some *non-examples* of smooth embeddings. Below we give some examples where one or the other criteria for being a smooth embedding fails.

Example 8.18 (A Smooth Topological Embedding)**.** The map $\gamma : \mathbb{R} \to \mathbb{R}^2$ given by $\gamma(t) = (t^3, 0)$ is a smooth map and a topological embedding, but it is not a smooth embedding. The map γ is smooth and is a topological embedding as

$$\gamma(\mathbb{R}) = \{(x, 0) : x \in \mathbb{R}\},$$

which is the x-axis in $\mathbb{R}^2$. The inverse map from the image $\gamma(\mathbb{R})$ to $\mathbb{R}$:

$$\gamma^{-1}(x, 0) = \sqrt[3]{x}.$$

Since $x \mapsto \sqrt[3]{x}$ is continuous, γ^{-1} is continuous. Therefore γ is a homeomorphism of $\mathbb{R}$ onto its image, so it is a topological embedding.

But γ is not a smooth embedding as

$$\gamma'(t) = (3t^2, 0).$$

and hence at $t = 0$,

$$\gamma'(0) = (0, 0).$$

Thus the differential

$$d\gamma_0 : T_0\mathbb{R} \to T_{(0,0)}\mathbb{R}^2$$

is the zero map, so it is not injective.

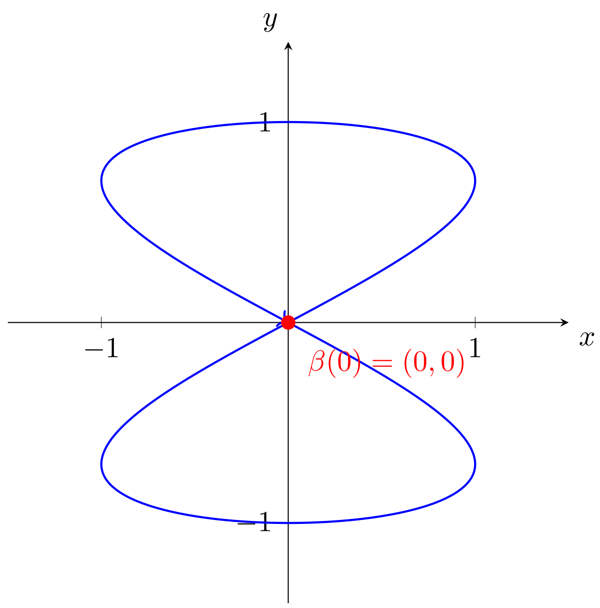
Example 8.19 (The Figure-Eight Curve)**.** Consider the curve $\beta : (-\pi, \pi) \to \mathbb{R}^2$ defined by

$$\beta(t) = (\sin 2t,\ \sin t).$$

Its image looks like a figure-eight in the plane, see the figure below. We show that β is *not* a smooth embedding. We have

$$\beta'(t) = (2\cos 2t, \cos t),$$

and hence$\beta'(t) \neq 0$ for all $t \in (-\pi, \pi)$ as, at $t = \frac{\pi}{2}$, $\beta'\left(\frac{\pi}{2}\right) = \left(2\cos\pi, \cos\frac{\pi}{2}\right) = (-2, 0) \neq 0$. and at $t = 0$, $\beta'(0) = (2, 1) \neq 0$.



We check that β is *not* a homeomorphism onto its image. The inverse $\beta^{-1} : \beta((-\pi, \pi)) \to (-\pi, \pi)$ is **not continuous**. First observe that

$$\beta(0) = (\sin 0, \sin 0) = (0, 0).$$

Now define the sequence

$$t_n = \pi - \frac{1}{n} \in (-\pi, \pi), \qquad n \geq 2.$$

Then $t_n \to \pi$ as $n \to \infty$, and

$$\begin{aligned}\sin 2t_n &= \sin\bigl(2\pi - \tfrac{2}{n}\bigr) = -\sin\tfrac{2}{n} \to 0,\\ \sin t_n &= \sin\bigl(\pi - \tfrac{1}{n}\bigr) = \sin\tfrac{1}{n} \to 0.\end{aligned}$$

Therefore

$$\beta(t_n) = \bigl(-\sin\tfrac{2}{n},\, \sin\tfrac{1}{n}\bigr) \;\longrightarrow\; (0,0) = \beta(0) \qquad \text{in } \mathbb{R}^2.$$

So $\beta(t_n) \to \beta(0)$ in the subspace topology on $\beta((-\pi,\pi))$.

If β^{-1} were continuous, we would need

$$\beta^{-1}(\beta(t_n)) = t_n \;\longrightarrow\; \beta^{-1}(0,0) = 0.$$

But $t_n \to \pi \neq 0$. Therefore β^{-1} is not continuous at $(0,0)$, so β is not a homeomorphism onto its image.

Example 8.20. Let $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$ be the 2-torus, and let α be any irrational number. The map $\gamma : \mathbb{R} \to \mathbb{T}^2$ given by

$$\gamma(t) = \left(e^{2\pi i t},\; e^{2\pi i \alpha t}\right),$$

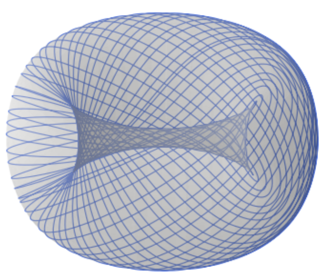
is a smooth immersion because $\gamma'(t)$ never vanishes. It is also injective, because $\gamma(t_1) = \gamma(t_2)$ implies that both $t_1 - t_2$ and $\alpha t_1 - \alpha t_2$ are integers, which is impossible unless $t_1 = t_2$. However, γ is *not* a homeomorphism onto its image. To see this, let $\varepsilon > 0$. Then we can find $n, m \in bZ$ such that $|\alpha n - m| < \varepsilon$. Thus,

$$|\gamma(n) - \gamma(0)| = |(e^{2\pi i n}, e^{2\pi i \alpha n}) - (1,1)| = |(1, e^{2\pi i \alpha n}) - (1,1)| \leq |e^{2\pi i \alpha n} - e^{2\pi i m}| \leq |2\pi(\alpha n - m)| \leq 2\pi\varepsilon.$$

Thus, $\gamma(0) \in \overline{\gamma(\mathbb{Z})}$ but $\mathbb{Z}$ has no limit point in $\mathbb{R}$.

Of course, if $\alpha \in \mathbb{Q}$ then γ is not even injective.

Dense curve on $\mathbb{T}^2$: $\gamma(t) = (e^{2\pi i t}, e^{2\pi i\sqrt{2}\,t})$



8.4 Submanifolds

As the name suggests, submanifolds should be part of a manifold which should be manifolds in their own right. This is the nottion we want to make precise in this section. We will give two ways to define a submanifold. First, let us recall from linear algebra that a linear subspace of a vector space, for instance

$$\mathbb{R}^\ell \times \{0\} = \{(x^1, \ldots, x^\ell, 0, \ldots, 0) \in \mathbb{R}^n \mid (x^1, \ldots, x^\ell) \in \mathbb{R}^\ell\} \subset \mathbb{R}^n. \tag{8.6}$$

is definitely a manifold in its own right and in fact, this is the typical example of a submanifold. Indeed, any ℓ-dimensional subspace of an n-dimensional vector space looks like (8.6) after a suitable linear change of coordinates. The notion of a smooth submanifold generalizes this idea. We first start with the following definition.

Definition 8.21 (Slice charts)**.** A chart $(\mathcal{U}, x)$ on an n-manifold M is called an **ℓ-dimensional slice chart** for a subset $L \subset M$ if

$$L \cap \mathcal{U} = x^{-1}(\mathbb{R}^{\ell} \times \{0\}),$$

i.e. the points in $\mathcal{U}$ belong to L if and only if their coordinates $x^{\ell+1}, \ldots, x^n$ vanish.

Definition 8.22. Suppose M is a smooth n-manifold. A subset $L \subset M$ is called an **ℓ-dimensional smooth submanifold** of M if M admits a collection of smooth slice charts for L whose domains cover every point of L.

Example 8.23. If we use the polar coordinates for describing the smooth structure on $S^1 \subset \mathbb{R}^2$ with the two charts (U, θ) and (V, ϕ), where $\theta(U) = (0, 2\pi)$ and $\phi(V) = (-\pi, \pi)$., then these two coordinates were defined on open subsets of S^1, but they also have natural extensions to open subsets of $\mathbb{R}^2$, namely

$$\mathcal{U}' := \{tv \in \mathbb{R}^2 \mid v \in U,\ t > 0\}, \qquad \mathcal{V}' := \{tv \in \mathbb{R}^2 \mid v \in V,\ t > 0\}.$$

Recall that a **1-dimensional slice chart** for $S^1 \subset \mathbb{R}^2$ is a chart $(\mathcal{U}', (x^1, x^2))$ on $\mathbb{R}^2$ such that a point of $\mathcal{U}'$ lies on S^1 if and only if $x^2 = 0$.

Define the radial and angular coordinates on $\mathbb{R}^2 \setminus \{0\}$ as usual, and set:

$$\rho := r - 1 \quad \text{so that} \quad S^1 = \{r = 1\} = \{\rho = 0\}.$$

Define

$$\mathcal{U}' = \mathbb{R}^2 \setminus \big(\{0\} \cup \{(x, 0) \mid x > 0\}\big).$$

which is $\mathbb{R}^2$ with the origin and the positive x-axis removed. The coordinate map is:

$$(\theta,\ \rho) : \mathcal{U}' \to (0, 2\pi) \times (-1, \infty),$$

where $\theta \in (0, 2\pi)$ is the polar angle and $\rho = \sqrt{x^2 + y^2} - 1$. Then:

$$S^1 \cap \mathcal{U}' = (\theta, \rho)^{-1}(\mathbb{R} \times \{0\}),$$

since $\rho = 0 \iff r = 1$. Similarly,

$$\mathcal{V}' = \mathbb{R}^2 \setminus \big(\{0\} \cup \{(x, 0) \mid x < 0\}\big).$$

which is $\mathbb{R}^2$ with the origin and the **negative** x-axis removed. The coordinate map is

$$(\phi,\ \rho) : \mathcal{V}' \to (-\pi, \pi) \times (-1, \infty),$$

where $\phi \in (-\pi, \pi)$ is the polar angle and $\rho = \sqrt{x^2 + y^2} - 1$. Then:

$$S^1 \cap \mathcal{V}' = (\phi, \rho)^{-1}(\mathbb{R} \times \{0\}).$$

So $S^1 \subset \mathcal{U}' \cup \mathcal{V}'$, and the two slice charts cover every point of S^1, confirming that S^1 is a **smooth 1-dimensional submanifold** of $\mathbb{R}^2$.

Let us now prove that a submanifold of a manifold is also a manifold in its own right.

Proposition 8.24. *If L is an ℓ-dimensional C^k-submanifold of an n-dimensional C^k-manifold M, then L inherits naturally from M the structure of an ℓ-dimensional C^k-manifold such that the inclusion map $i : L \hookrightarrow M$ is of class C^k. Moreover, for each $p \in L$, the tangent space T_pL is naturally an ℓ-dimensional linear subspace of T_pM.*

Proof. We associate to every slice chart $(\mathcal{U}, x)$ for $L \subset M$ a chart of the form $(\mathcal{U} \cap L, x_L)$ on L, where we use the coordinate projection

$$\pi_\ell(x^1, \dots, x^n) := (x^1, \dots, x^\ell)$$

to define

$$x_L = \pi_\ell \circ x|_{\mathcal{U} \cap L} : \mathcal{U} \cap L \to \mathbb{R}^\ell.$$

Since L is a submanifold, L can be covered by slice charts and hence the collection of all charts of this form defines an atlas on L. We now check the compatibility conditions. Given two such charts $(\mathcal{U} \cap L, x_L)$ and $(\mathcal{V} \cap L, y_L)$ derived from two C^k-compatible slice charts $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$, the transition map $y \circ x^{-1}$ preserves the subspace $\mathbb{R}^\ell \times \{0\} \subset \mathbb{R}^n$, and its restriction to the intersection of its domain with this subspace is the transition map $y_L \circ x_L^{-1}$, which is therefore of class C^k as $y \circ x^{-1}$ is of class C^k. Similarly, second countability and Hausdorff property for L follows from that of M and thus L is a C^k-manifold. The inclusion map $i : L \hookrightarrow M$, with respect to any slice chart $(\mathcal{U}, x)$ and the associated chart $(\mathcal{U} \cap L, x_L)$ on L, looks like

$$(x^1, \dots, x^\ell) \mapsto (x^1, \dots, x^\ell, 0, \dots, 0),$$

which is clearly smooth, thus the inclusion is of class C^k.[2]

For each $p \in L$, the tangent map $di_p : T_pL \to T_pM$ is simply the canonical inclusion $T_pL \hookrightarrow T_pM$ defined by regarding each path in L as a path in M. Since its image is a linear subspace, it gives a canonical isomorphism of T_pL to a linear subspace of T_pM. □

Thus a submanifold $L \subset M$ is a manifold and from now on, we will assume that L is endowed with the differentiable structure described in Proposition 8.24.

Exercise 8.25. Suppose M and N are both C^k-manifolds and $f : M \to N$ is a map of class C^k. Prove:

(a) For any C^k-submanifold $L \subset M$, the restriction $f|_L : L \to N$ is also a map of class C^k.

(b) If $L \subset N$ is a C^k-submanifold such that $f(M) \subset L$, then the resulting map $f : M \to L$ is also of class C^k.

8.5 Embeddings and level sets

We nowtry to understand the relation between embeddings and submanifolds which will also allow us to produce many more examples of submanifolds. The main result in this direction is the following.

Theorem 8.26. *If $f : M \to N$ is an embedding, then its image $f(M)$ is a smooth submanifold of N.*

[2] Recall that if both L and M are manifolds of class C^k but $k < \infty$, then it does not make sense to say that the inclusion $L \hookrightarrow M$ is smooth, even though it looks smooth in the particular local coordinates we chose. The point is that one could also choose different coordinates in which it would still appear to be a map of class C^k, but not necessarily C^∞.

Proof. We want to produce slice charts for $f(M) \subset N$. Suppose $q \in f(M)$. Since f is injective, there is a unique point $p \in M$ such that $f(p) = q$, and since f is an immersion, the Rank Theorem 8.11 gives us charts $(\mathcal{U}, x)$ on M and $(\mathcal{V}, y)$ on N with $x(p) = 0$ and $y(q) = 0$ such that $y \circ f \circ x^{-1}$ takes the form $(x^1, \dots, x^m) \mapsto (x^1, \dots, x^m, 0, \dots, 0)$. Since the inverse $f(M) \to M$ is also continuous, we are free to assume after possibly shrinking $\mathcal{V} \subset N$ to a smaller neighborhood of q that

$$f^{-1}(\mathcal{V} \cap f(M)) \subset \mathcal{U},$$

or in other words, $\mathcal{V} \cap f(M) = f(\mathcal{U})$. This proves that $(\mathcal{V}, y)$ is a slice chart for the subset $f(M)$. $\square$

In fact, we can give an alternative definition of the notion of a submanifold as a corollary of the previous theorem.

Corollary 8.27. *A subset $L \subset M$ of a smooth manifold M is a smooth submanifold if and only if it admits a smooth structure for which the inclusion map $L \hookrightarrow M$ is a smooth embedding.* $\square$

We now see the relations between submersions and submanifolds. We make the following

Definition 8.28. Let $f : M \to N$ be a smooth map. A point $p \in M$ is called a **regular point** of f if f is a submersion at p, and a **critical point** otherwise. A point $q \in N$ is a **critical value** of f if $q = f(p)$ for some critical point p, and q is otherwise called a **regular value** of f.

Theorem 8.29. *For any smooth map $f : M^m \to N^n$ with regular value $q \in N$, $L := f^{-1}(q) \subset M$ is a smooth submanifold with $\dim L = \dim M - \dim N$, and the tangent space at any point $p \in L$ is $T_pL = \ker df_p \subset T_pM$.*

Proof. We would like to find slice charts for L. For each $p \in L = f^{-1}(q)$, f is by assumption a submersion at p, so Theorem 8.11 provides charts x near p and y near q such that $x(p)$ and $y(q)$ are both the origin in their respective Euclidean spaces and $y \circ f \circ x^{-1}$ becomes the map $(x^1, \dots, x^m) \mapsto (x^1, \dots, x^n)$. The zero-set of this map is a neighborhood of p in $f^{-1}(q)$ as seen in the x-coordinates, thus x is a slice chart. We now look at the tangent space and prove that $T_pL = \ker df_p$. Notice first that for any path γ in L through p, $f \circ \gamma$ is a constant path at $q \in N$, and hence $df_p([\gamma]) = 0 \in T_qN$, proving $T_pL \subset \ker df_p$. Now q is a regular value and hence $df_p : T_pM \to T_qN$ is surjective. This implies

$$\dim T_pL = \dim L = \dim M - \dim N = \dim T_pM - \dim T_qN = \dim \ker T_pf,$$

by the rank-nullity theorem. $\square$

Definition 8.30. Submanifolds of the form $f^{-1}(q) \subset M$ for regular values $q \in N$ are called **regular level sets** of f.

Thus, a submersion $f : M \to N$ is distinguished by the property that *all* of its level sets are regular, and are thus smooth submanifolds.

In fact, the proof of Theorem 8.29 also proves the following result on maps of constant rank.

Theorem 8.31 (Constant-Rank Level set theorem)**.** *Let $f : M^m \to N^n$ be a smooth map between smooth manifolds such that f is of constant rank r. Then the level sets of f, i.e., $f^{-1}(q) \subset M$, $q \in N$ are smooth submanifolds of M of dimension $m - r$.*

We now have a *very* easy way of proving that simple examples like the unit spheres $S^n \subset \mathbb{R}^{n+1}$ really are smooth submanifolds.

Example 8.32. Define $f : \mathbb{R}^{n+1} \to \mathbb{R}$ in terms of the standard Euclidean inner product by $f(x) = |x|^2 = \langle x, x \rangle$. This is a smooth map, with differential at any point $x \in \mathbb{R}^{n+1}$ given by $d_x f(v) = 2\langle x, v \rangle$, so it is a submersion everywhere except at the origin. This makes $S^n = f^{-1}(1)$ into a smooth submanifold of dimension $(n+1) - 1 = n$, so in particular, S^n inherits a natural smooth structure for which the inclusion $S^n \hookrightarrow \mathbb{R}^{n+1}$ is a smooth embedding. The kernel of $d_x f$ at a point $x \in S^n$ is the orthogonal complement of x, hence

$$T_x S^n = x^\perp \subset \mathbb{R}^{n+1}.$$

which is something you prove explicitly in Problem 3 in Problem Set 3.

Example 8.33. The smooth map $f : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto xy$ has only one critical point, at $(x, y) = (0, 0)$, thus $f^{-1}(t)$ is a smooth submanifold (a hyperbola) for every $t \neq 0$, and so is $f^{-1}(0) \setminus \{(0, 0)\}$, but $f^{-1}(0)$ fails to be a submanifold at the origin.

9. Vector Fields

9.1 Smooth vector fields

Definition 9.1. Let M be a smooth manifold. A ***vector field on*** M is a continuous map $X : M \to TM$, usually written $p \mapsto X_p$, with the property that

$$\pi \circ X = \mathrm{id}_M, \qquad \pi : TM \to M \text{ the projection map} \tag{9.1}$$

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

One can visualize a vector field on M in the same way as to visualize vector fields in Euclidean space: as an arrow attached to each point of M, chosen to be tangent to M and to vary continuously from point to point.

Since we already saw that if M is a smooth manifold then the tangent bundle TM inherits a smooth structure from M in Theorem 5.8, we say that a vector field is a ***smooth vector fields*** if it is smooth as map from M to TM.

Just as for functions, if X is a vector field on M, the ***support of*** X is defined to be the closure of the set $\{p \in M : X_p \neq 0\}$. A vector field is said to be ***compactly supported*** if its support is a compact set.

Suppose M is a smooth n-manifold. If $X : M \to TM$ is a vector field and $(U, (x^i))$ is any smooth coordinate chart for M, we can write the value of X at any point $p \in U$ in terms of the coordinate basis vectors:

$$X_p = X^i(p) \left.\frac{\partial}{\partial x^i}\right|_p . \tag{9.2}$$

This defines n functions $X^i : U \to \mathbb{R}$, called the ***component functions of*** X in the given chart. We can give the smoothness of vector fields in terms of the smoothness of the component functions.

Proposition 9.2. *Let M be a smooth manifold, and let $X : M \to TM$ be a vector field. If $(U, (x^i))$ is any smooth coordinate chart on M, then the restriction of X to U is smooth if and only if its component functions with respect to this chart are smooth.*

Proof. Let (x^i, v^i) be the natural coordinates on $\pi^{-1}(U) \subseteq TM$ associated with the chart $(U, (x^i))$. By definition of natural coordinates, the coordinate representation of $X : M \to TM$ on U is

$$\widehat{X}(x) = (x^1, \ldots, x^n, X^1(x), \ldots, X^n(x)),$$

where X^i is the i-th component function of X in x^i-coordinates. It follows immediately that smoothness of X in U is equivalent to smoothness of its component functions. $\square$

Let's see some examples of smooth vector fields.

Example 9.3. (Coordinate Vector Fields). If $(U, (x^i))$ is any smooth chart on M, the assignment

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

determines a vector field on U, called the ***i-th coordinate vector field*** and denoted by $\partial/\partial x^i$. It is smooth because its component functions are constants.

Example 9.4. (The Euler Vector Field). The vector field V on $\mathbb{R}^n$ whose value at $x \in \mathbb{R}^n$ is

$$V_x = x^1 \left.\frac{\partial}{\partial x^1}\right|_x + \cdots + x^n \left.\frac{\partial}{\partial x^n}\right|_x$$

is smooth because its coordinate functions are linear. It vanishes at the origin, and points radially outward everywhere else. It is called the **Euler vector field**.

If $U \subseteq M$ is open, the fact that T_pU is naturally identified with T_pM for each $p \in U$ allows us to identify TU with the open subset $\pi^{-1}(U) \subseteq TM$. Therefore, a vector field on U can be thought of either as a map from U to TU or as a map from U to TM, whichever is more convenient. If X is a vector field on M, its restriction $X|_U$ is a vector field on U, which is smooth if X is.

If M is a smooth manifold we would use

Notation: the notation $\mathfrak{X}(M)$ or $\Gamma(TM)$ to denote the set of all smooth vector fields on M.

The set $\mathfrak{X}(M)$ is a $\mathbb{R}$-vector space under point wise addition and scalar multiplication:

$$(aX + bY)_p = aX_p + bY_p.$$

The zero element of this vector space is the zero vector field, whose value at each $p \in M$ is $0 \in T_pM$. In addition, smooth vector fields can be multiplied by smooth real-valued functions: if $f \in C^\infty(M)$ and $X \in \mathfrak{X}(M)$, we define $fX : M \to TM$ by

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

The next proposition shows that these operations indeed yield smooth vector fields.

Proposition 9.5. *Let M be a smooth manifold.*

(1) If X and Y are smooth vector fields on M and $f, g \in C^\infty(M)$, then $fX + gY$ is a smooth vector field.

(2) $\mathfrak{X}(M)$ is a module over the ring $C^\infty(M)$.

Proof. Exercise. □

9.2 Local and Global Frames

Coordinate vector fields in a smooth chart provide a convenient way of representing vector fields, because their values form a basis for the tangent space at each point. However, they are not the only choices.

Suppose M is a smooth n-manifold. An ordered k-tuple $(X_1, \ldots, X_k)$ of vector fields defined on some subset $A \subseteq M$ is said to be **linearly independent** if $(X_1|_p, \ldots, X_k|_p)$ is a linearly independent k-tuple in T_pM for each $p \in A$, and is said to ***span the tangent bundle*** if the k-tuple $(X_1|_p, \ldots, X_k|_p)$ spans T_pM at each $p \in A$.

Definition 9.6. A **local frame for** M is an ordered n-tuple of vector fields $(E_1, \ldots, E_n)$ defined on an open subset $U \subseteq M$ that is linearly independent and spans the tangent bundle. In other words, the vectors $(E_1|_p, \ldots, E_n|_p)$ form a basis for T_pM at each $p \in U$.

It is called a **global frame** if $U = M$, and a **smooth frame** if each of the vector fields E_i is smooth.

Example 9.7. (Local and Global Frames).

(a) The standard coordinate vector fields form a smooth global frame for $\mathbb{R}^n$.

(b) If $(U, (x^i))$ is any smooth coordinate chart for a smooth manifold M, then the coordinate vector fields form a smooth local frame $(\partial/\partial x^i)$ on U, called a **coordinate frame**. Every point of M is in the domain of such a local frame.

(c) The vector field $d/d\theta$ on S^1 where θ is the angle coordinate is a smooth global frame for the circle.

(d) The n-tuple of vector fields $(\partial/\partial\theta^1, \ldots, \partial/\partial\theta^n)$ on the n-torus $T^n = S^1 \times \cdots \times S^1$ is a smooth global frame for $\mathbb{T}^n$.

For subsets of $\mathbb{R}^n$, there is a special type of frame that is often more useful for geometric problems than arbitrary frames. A k-tuple of vector fields $(E_1, \ldots, E_k)$ defined on some subset $A \subseteq \mathbb{R}^n$ is said to be **orthonormal** if for each $p \in A$, the vectors $(E_1|_p, \ldots, E_k|_p)$ are orthonormal with respect to the Euclidean dot product (where we identify $T_p\mathbb{R}^n$ with $\mathbb{R}^n$ in the usual way). A (local or global) frame consisting of orthonormal vector fields is called an **orthonormal frame**.

9.3 Vector fields and derivations

Recall that we an view a tangent vector as a derivation. Similarly, vector fields define operators on the space of smooth real-valued functions $C^\infty(M)$ on a manifold M. If $X \in \mathfrak{X}(M)$ and f is a smooth real-valued function defined on an open subset $U \subseteq M$, we obtain a new function $Xf : U \to \mathbb{R}$, defined by

$$(Xf)(p) = X_p f,$$

where the tangent vector X_p is acting as a derivation on the function f. Since the action of a tangent vector on a function is determined by the values of the function in an arbitrarily small neighbourhood, it follows that Xf is locally determined. In particular, for any open subset $V \subseteq U$,

$$(Xf)|_V = X(f|_V). \tag{9.3}$$

This construction yields another useful smoothness criterion for vector fields.

Proposition 9.8. *Let M be a smooth manifold and let $X : M \to TM$ be a vector field. The following are equivalent:*

(a) X is smooth.

(b) For every $f \in C^\infty(M)$, the function Xf is smooth on M.

(c) For every open subset $U \subseteq M$ and every $f \in C^\infty(U)$, the function Xf is smooth on U.

Proof. We will prove that (a) ⇒ (b) ⇒ (c) ⇒ (a).

To prove (a) ⇒ (b), assume X is smooth, and let $f \in C^\infty(M)$. For any $p \in M$, we can choose smooth coordinates (x^i) on a neighbourhood U of p. Then for $x \in U$, we can write

$$Xf(x) = \left(X^i(x) \left.\frac{\partial}{\partial x^i}\right|_x \right) f = X^i(x) \frac{\partial f}{\partial x^i}(x).$$

Since the component functions X^i are smooth on U, it follows that Xf is smooth in U. Since the same is true in a neighbourhood of each point, Xf is smooth on M.

To prove (b) ⇒ (c), suppose $U \subseteq M$ is open and $f \in C^\infty(U)$. For any $p \in U$, let ψ be a smooth bump function that is equal to 1 in a neighbourhood of p and supported in U, and define $\widetilde{f} = \psi f$, extended to be zero on $M \setminus \operatorname{supp} \psi$. Then $X\widetilde{f}$ is smooth by assumption, and is equal to Xf in a neighbourhood of p. This shows that Xf is smooth in a neighbourhood of each point of U.

Finally, to prove (c) ⇒ (a), suppose Xf is smooth whenever f is smooth on an open subset of M. If (x^i) are any smooth local coordinates on $U \subseteq M$, we can think of each coordinate x^i as a smooth function on U. Applying X to one of these functions, we obtain

$$Xx^i = X^j \frac{\partial}{\partial x^j}(x^i) = X^i.$$

Because Xx^i is smooth by assumption, it follows that the component functions of X are smooth, so X is smooth. □

Thus, a smooth vector field $X \in \mathfrak{X}(M)$ defines a map from

$$C^\infty(M) \to C^\infty(M), \quad f \mapsto Xf.$$

This map is clearly linear over $\mathbb{R}$. Moreover, the Leibniz rule for tangent vectors as derivations implies the product rule for vector fields:

$$X(fg) = f\,Xg + g\,Xf, \tag{9.4}$$

as you can easily check by evaluating both sides at an arbitrary point $p \in M$.

A map $X : C^\infty(M) \to C^\infty(M)$ is called a **derivation** (as distinct from a *derivation at* p) if it is linear over $\mathbb{R}$ and satisfies (9.4) for all $f, g \in C^\infty(M)$. Thus, any vector field is a derivation on $C^\infty(M)$.

The function Xf is sometimes also denoted by $\mathcal{L}_X f = Xf$ and is called the **Lie derivative of f in the direction of X.**

The next proposition shows that derivations of $C^\infty(M)$ can be identified with smooth vector fields.

Lemma 9.9. *Let M be a smooth manifold. A map $D : C^\infty(M) \to C^\infty(M)$ is a derivation if and only if it is of the form $Df = Xf = \mathcal{L}_X f$ for some smooth vector field $X \in \mathfrak{X}(M)$.*

Proof. We just showed above that every smooth vector field induces a derivation. Conversely, suppose $D : C^\infty(M) \to C^\infty(M)$ is a derivation. We need to come up with a vector field X such that $Df = Xf$ for all f. From the discussion above, it is clear that if there is such a vector field, its value at $p \in M$ must be the derivation at p whose action on any smooth real-valued function f is given by

$$X_p f = Df(p).$$

The linearity of D guarantees that this expression depends linearly on f, and the fact that D is a derivation yields the Leibniz rule for tangent vectors. Thus, the map $X_p : C^\infty(M) \to \mathbb{R}$ so defined is indeed a tangent vector, that is, a derivation of $C^\infty(M)$ at p. This defines X as a vector field. Because $Xf = Df$ is smooth whenever $f \in C^\infty(M)$, this vector field is smooth by Proposition 9.8. $\square$

Thus,

smooth vector fields on $M \leftrightarrow$ derivations of $C^\infty(M)$.

9.4 Vector fields, smooth maps, pushforwards and pullbacks

Recall that if $F : M \to N$ is a smooth map and X is a vector field on M, then for each point $p \in M$, we obtain a vector $dF_p(X_p) \in T_{F(p)}N$ by applying the differential of F to X_p. This begs the question: Can this process when applied globally, gives a vector field on N instead of just tangent vectors which depend on the point? The answer is **no**, this does not in general define a *vector field* on N. For example, if F is not surjective, there is no way to decide what vector to assign to a point $q \in N \setminus F(M)$. If F is not injective, then for some points of N there may be several different vectors obtained by applying dF to X at different points of M.

Suppose $F : M \to N$ is smooth and X is a vector field on M, and suppose there happens to be a vector field Y on N with the property that for each $p \in M$, $dF_p(X_p) = Y_{F(p)}$. In this case, we say the vector fields X and Y are **F-related** . The next proposition shows how F-related vector fields act on smooth functions.

Proposition 9.10. *Suppose $F : M \to N$ is a smooth map between manifolds, $X \in \mathfrak{X}(M)$, and $Y \in \mathfrak{X}(N)$. Then X and Y are F-related if and only if for every smooth real-valued function f defined on an open subset of N,*

$$X(f \circ F) = (Yf) \circ F. \tag{9.5}$$

Proof. For any $p \in M$ and any smooth real-valued f defined in a neighbourhood of $F(p)$,

$$X(f \circ F)(p) = X_p(f \circ F) = dF_p(X_p)f,$$

while

$$(Yf) \circ F(p) = (Yf)(F(p)) = Y_{F(p)}f.$$

Thus, (9.5) is true for all f if and only if $dF_p(X_p) = Y_{F(p)}$ for all p, i.e., if and only if X and Y are F-related. □

Example 9.11. Let $F : \mathbb{R} \to \mathbb{R}^2$ be the smooth map $F(t) = (\cos t, \sin t)$. Then $\frac{d}{dt} \in \mathfrak{X}(\mathbb{R})$ is F-related to the vector field $Y \in \mathfrak{X}(\mathbb{R}^2)$ defined by

$$Y = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}.$$

We show this in two different ways.

We show that for every $t \in \mathbb{R}$,

$$dF_t\left(\left.\frac{d}{dt}\right|_t\right) = Y_{F(t)}.$$

Recall that for a smooth map $F : \mathbb{R} \to \mathbb{R}^2$ written in coordinates as $F(t) = (F^1(t), F^2(t))$, the differential acts by

$$dF_t\left(\left.\frac{d}{dt}\right|_t\right) = \frac{dF^1}{dt}(t)\,\left.\frac{\partial}{\partial x}\right|_{F(t)} + \frac{dF^2}{dt}(t)\,\left.\frac{\partial}{\partial y}\right|_{F(t)}.$$

So

$$dF_t\left(\left.\frac{d}{dt}\right|_t\right) = -\sin t\,\left.\frac{\partial}{\partial x}\right|_{F(t)} + \cos t\,\left.\frac{\partial}{\partial y}\right|_{F(t)}.$$

By the definition of Y,

$$Y_{(x,y)} = -y\,\left.\frac{\partial}{\partial x}\right|_{(x,y)} + x\,\left.\frac{\partial}{\partial y}\right|_{(x,y)}.$$

Evaluating at $(x, y) = F(t) = (\cos t, \sin t)$,

$$Y_{F(t)} = -\sin t\,\left.\frac{\partial}{\partial x}\right|_{F(t)} + \cos t\,\left.\frac{\partial}{\partial y}\right|_{F(t)},$$

and hence

$$dF_t\left(\left.\frac{d}{dt}\right|_t\right) = Y_{F(t)} \qquad \text{for all } t \in \mathbb{R}.$$

Hence $\dfrac{d}{dt}$ is F-related to Y.

We can also use (9.5) to show the same. Let $f \in C^\infty(\mathbb{R}^2)$ be arbitrary. Using the chain rule,

$$\frac{d}{dt}(f \circ F)(t) = \frac{d}{dt}\,f(\cos t,\, \sin t) = \frac{\partial f}{\partial x}(F(t)) \cdot \frac{d}{dt}(\cos t) + \frac{\partial f}{\partial y}(F(t)) \cdot \frac{d}{dt}(\sin t),$$

which gives

$$\frac{d}{dt}(f \circ F)(t) = -\sin t \cdot \frac{\partial f}{\partial x}(F(t)) + \cos t \cdot \frac{\partial f}{\partial y}(F(t)).$$

By the definition of Y,

$$(Yf)(x, y) = -y\,\frac{\partial f}{\partial x}(x, y) + x\,\frac{\partial f}{\partial y}(x, y).$$

Evaluating at $F(t) = (\cos t, \sin t)$,

$$(Yf)(F(t)) = -\sin t \cdot \frac{\partial f}{\partial x}(F(t)) + \cos t \cdot \frac{\partial f}{\partial y}(F(t)),$$

and hence

$$\frac{d}{dt}(f \circ F)(t) = (Yf)(F(t)) \qquad \text{for all } t \in \mathbb{R} \text{ and all } f \in C^\infty(\mathbb{R}^2).$$

Therefore $\dfrac{d}{dt}$ is F-related to Y.

It is important to remember that for a given smooth map $F : M \to N$ and vector field $X \in \mathfrak{X}(M)$, there may not be *any* vector field on N that is F-related to X. There is one special case, however, in which there is always such a vector field, as the next lemma shows.

Lemma 9.12. *Suppose M and N are smooth manifolds and $F : M \to N$ is a diffeomorphism. For every $X \in \mathfrak{X}(M)$, there is a unique smooth vector field on N that is F-related to X.*

Proof. For $Y \in \mathfrak{X}(N)$ to be F-related to X means that $dF_p(X_p) = Y_{F(p)}$ for every $p \in M$. If F is a diffeomorphism, we *define* Y by

$$Y_q = dF_{F^{-1}(q)}\left(X_{F^{-1}(q)}\right).$$

It is clear that Y, so defined, is the unique vector field that is F-related to X. Note that $Y : N \to TN$ is the composition of the following smooth maps:

$$N \xrightarrow{F^{-1}} M \xrightarrow{X} TM \xrightarrow{dF} TN.$$

It follows that Y is smooth. □

We can now make the following definition.

Definition 9.13 (Pushforwards and Pullbacks). Let $F : M \to N$ be a diffeomorphism between smooth manifolds M and N. We denote the unique vector field that is F-related to X by F_*X, and call it the **pushforward of X by F**. It is explicitly described by

$$(F_*X)_q = dF_{F^{-1}(q)}\left(X_{F^{-1}(q)}\right). \tag{9.6}$$

The **pullback of a vector field** $Y \in \mathfrak{X}(N)$ is denoted by F^*Y and is defined as

$$F^*Y = dF^{-1} \circ Y \circ F \in \mathfrak{X}(M). \tag{9.7}$$

As long as the inverse map F^{-1} can be computed explicitly, the pushforward of a vector field can be computed directly from the formula in (9.6).

Let us compute the pushforward of a vector field explicitly.

Let M and N be the following open submanifolds of $\mathbb{R}^2$:

$$M = \{(x, y) : y > 0 \text{ and } x + y > 0\},$$

$$N = \{(u, v) : u > 0 \text{ and } v > 0\},$$

and define $F : M \to N$ by $F(x, y) = (x + y, x/y + 1)$. Then F is a diffeomorphism because its inverse is easily computed: just solve $(u, v) = (x + y, x/y + 1)$ for x and y to obtain the formula $(x, y) = F^{-1}(u, v) = (u - u/v, u/v)$. Let us compute the pushforward F_*X, where X is the following smooth vector field on M:

$$X_{(x,y)} = y^2 \left.\frac{\partial}{\partial x}\right|_{(x,y)}.$$

The differential of F at a point $(x, y) \in M$ is represented by its Jacobian matrix,

$$DF(x, y) = \begin{pmatrix} 1 & 1 \\ \frac{1}{y} & -\frac{x}{y^2} \end{pmatrix},$$

and thus $dF_{F^{-1}(u,v)}$ is represented by the matrix

$$DF\left(u - \frac{u}{v}, \frac{u}{v}\right) = \begin{pmatrix} 1 & 1 \\ \frac{v}{u} & \frac{v - v^2}{u} \end{pmatrix}.$$

For any $(u, v) \in N$,

$$X_{F^{-1}(u,v)} = \frac{u^2}{v^2} \frac{\partial}{\partial x}\bigg|_{F^{-1}(u,v)}.$$

Therefore, applying (9.6) with $p = (u, v)$ yields the formula for F_*X:

$$(F_*X)_{(u,v)} = \frac{u^2}{v^2} \frac{\partial}{\partial u}\bigg|_{(u,v)} + \frac{u}{v} \frac{\partial}{\partial v}\bigg|_{(u,v)}.$$

9.5 Vector Fields and Submanifolds

If $S \subseteq M$ is an embedded submanifold, a vector field X on M does not necessarily restrict to a vector field on S, because X_p may not lie in the subspace $T_pS \subseteq T_pM$ at a point $p \in S$. Given a point $p \in S$, a vector field X on M is said to be **tangent to S at p** if $X_p \in T_pS \subseteq T_pM$. It is **tangent to S** if it is tangent to S at every point of S.

Suppose $S \subseteq M$ is an embedded submanifold and Y is a smooth vector field on M. If there is a vector field $X \in \mathfrak{X}(S)$ that is ι-related to Y, where $\iota : S \hookrightarrow M$ is the inclusion map, then clearly Y is tangent to S, because $Y_p = d\iota_p(X_p)$ is in the image of $d\iota_p$ for each $p \in S$. The next proposition shows that the converse is true.

Proposition 9.14 (Restricting Vector Fields to Submanifolds). *Let M be a smooth manifold, let $S \subseteq M$ be an embedded submanifold and let $\iota : S \hookrightarrow M$ denote the inclusion map. If $Y \in \mathfrak{X}(M)$ is tangent to S, then there is a unique smooth vector field on S, denoted by $Y|_S$, that is ι-related to Y.*

Proof. The fact that Y is tangent to S means by definition that Y_p is in the image of $d\iota_p$ for each p. Thus, for each p there is a vector $X_p \in T_pS$ such that $Y_p = d\iota_p(X_p)$. Since $d\iota_p$ is injective, X_p is unique, so this defines X as a vector field on S. If we can show that X is smooth, it is the unique vector field that is ι-related to Y. It suffices to show that it is smooth in a neighbourhood of each point.

Let p be any point in S. Since S is an embedded submanifold, let $(U, (x^i))$ be a slice chart in M centred at p, so that $S \cap U$ is the subset where $x^{k+1} = \cdots = x^n = 0$, and $(x^1, \ldots, x^k)$ form local coordinates for S in U. If $Y = Y^1 \frac{\partial}{\partial x^1} + \cdots + Y^n \frac{\partial}{\partial x^n}$ in these coordinates, it follows from our construction that X has the coordinate representation $Y^1 \frac{\partial}{\partial x^1} + \cdots + Y^k \frac{\partial}{\partial x^k}$, which is clearly smooth on U. □

9.6 Lie Brackets

In this section we introduce an important way of combining two smooth vector fields to obtain another vector field.

Let X and Y be smooth vector fields on a smooth manifold M. Given a smooth function $f : M \to \mathbb{R}$, we can apply X to f and obtain another smooth function Xf. In turn, we can apply Y to this function, and obtain yet another smooth function $YXf = Y(Xf)$. The operation $f \mapsto YXf$, however, does not in general satisfy the product rule and thus cannot be a vector field, as the following example shows.

Example 9.15. Define vector fields $X = \frac{\partial}{\partial x}$ and $Y = x\frac{\partial}{\partial y}$ on $\mathbb{R}^2$, and let $f(x, y) = x$, $g(x, y) = y$. Then direct computation shows that $XY(fg) = 2x$, while $fXYg + gXYf = x$, so XY is not a derivation of $C^\infty(\mathbb{R}^2)$.

We can also apply the same two vector fields in the opposite order, obtaining a (usually different) function XYf. Applying both of these operators to f and subtracting, we obtain an operator

$$[X, Y] : C^\infty(M) \to C^\infty(M),$$

called the **Lie bracket of X and Y**, defined by

$$[X, Y]f = XYf - YXf.$$

The key fact is that this operator *is* a vector field.

Lemma 9.16. *The Lie bracket of any pair of smooth vector fields is a smooth vector field.*

Proof. It suffices to show that $[X, Y]$ is a derivation of $C^\infty(M)$. For arbitrary $f, g \in C^\infty(M)$, we compute

$$\begin{aligned}
[X, Y](fg) &= X(Y(fg)) - Y(X(fg)) \\
&= X(fYg + gYf) - Y(fXg + gXf) \\
&= XfYg + fXYg + XgYf + gXYf \\
&\quad - YfXg - fYXg - YgXf - gYXf \\
&= fXYg + gXYf - fYXg - gYXf \\
&= f[X, Y]g + g[X, Y]f.
\end{aligned}$$

□

The value of the vector field $[X, Y]$ at a point $p \in M$ is the derivation at p given by the formula

$$[X, Y]_p f = X_p(Yf) - Y_p(Xf). \tag{9.8}$$

However, this formula is of limited usefulness for computations, because it requires one to compute terms involving second derivatives of f that will always cancel each other out. The next proposition gives an extremely useful coordinate formula for the Lie bracket, in which the cancellations have already been accounted for.

Proposition 9.17 (Coordinate Formula for the Lie Bracket)**.** *Let X, Y be smooth vector fields on a smooth manifold M and let $X = X^i \frac{\partial}{\partial x^i}$ and $Y = Y^j \frac{\partial}{\partial x^j}$ be the coordinate expressions for X and Y in terms of some smooth local coordinates (x^i) for M. Then $[X, Y]$ has the following coordinate expression:*

$$[X, Y] = \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \right) \frac{\partial}{\partial x^j}, \tag{9.9}$$

or more concisely,

$$[X, Y] = (XY^j - YX^j) \frac{\partial}{\partial x^j}. \tag{9.10}$$

Proof. Because we know already that $[X, Y]$ is a smooth vector field, its action on a function is determined locally: $([X, Y]f)|_U = [X, Y](f|_U)$. Thus it suffices to compute in a single smooth chart, where we have

$$\begin{aligned}
[X, Y]f &= X^i \frac{\partial}{\partial x^i}\left(Y^j \frac{\partial f}{\partial x^j} \right) - Y^j \frac{\partial}{\partial x^j}\left(X^i \frac{\partial f}{\partial x^i} \right) \\
&= X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial f}{\partial x^j} + X^i Y^j \frac{\partial^2 f}{\partial x^i \partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \frac{\partial f}{\partial x^i} - Y^j X^i \frac{\partial^2 f}{\partial x^j \partial x^i} \\
&= X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial f}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \frac{\partial f}{\partial x^i},
\end{aligned}$$

where in the last step we have used the fact that mixed partial derivatives of a smooth function can be taken in any order. Interchanging the roles of the dummy indices i and j in the second term, we obtain (9.9). □

One trivial application of (9.9) is to compute the Lie brackets of the coordinate vector fields $(\partial/\partial x^i)$ in any smooth chart: because the component functions of the coordinate vector fields are all constants, it follows that

$$\left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right] \equiv 0 \quad \text{for all } i \text{ and } j. \tag{9.11}$$

Notice that this also follows from the definition of the Lie bracket in (9.8), and is essentially a restatement of the fact that mixed partial derivatives of smooth functions commute. Here is a slightly less trivial computation.

Example 9.18. Define smooth vector fields $X, Y \in \mathfrak{X}(\mathbb{R}^3)$ by

$$X = x\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + x(y+1)\frac{\partial}{\partial z},$$
$$Y = \frac{\partial}{\partial x} + y\frac{\partial}{\partial z}.$$

Then (9.10) gives us

$$\begin{aligned}[X,Y] &= X(1)\frac{\partial}{\partial x} + X(y)\frac{\partial}{\partial z} - Y(x)\frac{\partial}{\partial x} - Y(1)\frac{\partial}{\partial y} - Y(x(y+1))\frac{\partial}{\partial z} \\ &= 0\frac{\partial}{\partial x} + 1\frac{\partial}{\partial z} - 1\frac{\partial}{\partial x} - 0\frac{\partial}{\partial y} - (y+1)\frac{\partial}{\partial z} \\ &= -\frac{\partial}{\partial x} - y\frac{\partial}{\partial z},\end{aligned}$$

which is indeed a vector field.

Let us now see some basic properties of the Lie bracket.

Proposition 9.19 (Properties of the Lie Bracket)**.** *The Lie bracket satisfies the following identities for all $X, Y, Z \in \mathfrak{X}(M)$:*

(a) BILINEARITY: *For $a, b \in \mathbb{R}$,*

$$[aX + bY, Z] = a[X,Z] + b[Y,Z],$$
$$[Z, aX + bY] = a[Z,X] + b[Z,Y].$$

(b) ANTISYMMETRY:

$$[X,Y] = -[Y,X].$$

(c) JACOBI IDENTITY:

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0.$$

(d) For $f, g \in C^\infty(M)$,

$$[fX, gY] = fg[X,Y] + (fXg)Y - (gYf)X. \tag{9.12}$$

Proof. You can check bilinearity and antisymmetry yourself. The proof of the Jacobi identity is just a computation:

$$\begin{aligned}&[X,[Y,Z]]f + [Y,[Z,X]]f + [Z,[X,Y]]f \\ &\quad = X[Y,Z]f - [Y,Z]Xf + Y[Z,X]f \\ &\qquad - [Z,X]Yf + Z[X,Y]f - [X,Y]Zf \\ &\quad = XYZf - XZYf - YZXf + ZYXf + YZXf - YXZf \\ &\qquad - ZXYf + XZYf + ZXYf - ZYXf - XYZf + YXZf.\end{aligned}$$

In this last expression all the terms cancel in pairs. Part (d) is again a direct computation from the definition of the Lie bracket. Let $h \in C^\infty(M)$. We have

$$\begin{aligned}[fX, gY]h &= (fX)(gY)h - (gY)(fX)h \\ &= (fX)(gYh) - (gY)(fXh) \\ &= fg(XYh) + f(Xg)(Yh) - fg(YXh) - g(Yf)(Xh) \\ &= fg[X,Y]h + (fXg)Yh - (gYf)Xh.\end{aligned}$$

□

The significance of part (d) of this proposition might not be evident at this point, but it will become clearer in the future lectures when we will talk about the *Lie derivative of a vector field in the direction of another vector field* where it will be shown to be equal to the Lie bracket and hence the Lie derivative will indeed be a derivation using part (d).

The next result relates the Lie bracket of F-related vector fields.

Proposition 9.20 (Naturality of the Lie Bracket)**.** *Let $F : M \to N$ be a smooth map between manifolds and let $X_1, X_2 \in \mathfrak{X}(M)$ and $Y_1, Y_2 \in \mathfrak{X}(N)$ be vector fields such that X_i is F-related to Y_i for $i = 1, 2$. Then $[X_1, X_2]$ is F-related to $[Y_1, Y_2]$.*

Proof. Using (9.5) and the fact that X_i and Y_i are F-related,

$$X_1X_2(f \circ F) = X_1\big((Y_2f) \circ F\big) = (Y_1Y_2f) \circ F.$$

Similarly,

$$X_2X_1(f \circ F) = (Y_2Y_1f) \circ F.$$

Therefore,

$$\begin{aligned}
[X_1, X_2](f \circ F) &= X_1X_2(f \circ F) - X_2X_1(f \circ F) \\
&= (Y_1Y_2f) \circ F - (Y_2Y_1f) \circ F \\
&= \big([Y_1, Y_2]f\big) \circ F.
\end{aligned}$$

□

When applied in special cases, this result has the following important corollaries. First we consider the case in which the map is a diffeomorphism.

Corollary 9.21. *We have the following.*

1. *(Pushforwards of Lie Brackets) Suppose $F : M \to N$ is a diffeomorphism and $X_1, X_2 \in \mathfrak{X}(M)$. Then $F_*[X_1, X_2] = [F_*X_1, F_*X_2]$.*
2. *(Brackets of Vector Fields Tangent to Submanifolds) Let M be a smooth manifold and let S be an embedded submanifold of M. If Y_1 and Y_2 are smooth vector fields on M that are tangent to S, then $[Y_1, Y_2]$ is also tangent to S.*

Proof.
1. This is just the special case of Proposition 9.20 in which F is a diffeomorphism and $Y_i = F_*X_i$.
2. We know there exist smooth vector fields X_1 and X_2 on S such that X_i is ι-related to Y_i for $i = 1, 2$, where $\iota : S \to M$ is the inclusion. By Proposition 9.20, $[X_1, X_2]$ is ι-related to $[Y_1, Y_2]$, which is therefore tangent to S.

□

10. Integral curves and flows of vector fields

10.1 Integral Curves

Suppose M is a smooth manifold. If $\gamma\colon J \to M$ is a smooth curve, then for each $t \in J$, the velocity vector $\gamma'(t)$ is a vector in $T_{\gamma(t)}M$. What about the reverse process? That is, given $V \in \mathfrak{X}(M)$, can we come up with a curve γ such that the tangent vector is the corresponding image of the vector field?

Definition 10.1. Let V is a vector field on M. An **integral curve** of V is a differentiable curve $\gamma\colon J \to M$ whose velocity at each point is equal to the value of V at that point:

$$\gamma'(t) = V_{\gamma(t)} \quad \text{for all } t \in J. \tag{10.1}$$

If $0 \in J$, the point $\gamma(0)$ is called the **starting point of γ**.

Before dwelling further in the theory, let us see some simple examples.

Example 10.2. Let (x, y) be standard coordinates on $\mathbb{R}^2$, and let $V = \frac{\partial}{\partial x} \in \mathfrak{X}(\mathbb{R}^2)$. Then

$$\gamma(t) = (a + t, b), \quad a, b \text{ constants}$$

is an integral curve for V. Things to notice is that given any $(p, q) \in \mathbb{R}^2$ $\exists!$ integral curve of V starting from (p, q) and the images of two integral curves are either identical or disjoint. It is easy to check that the integral curves of V are precisely the straight lines parallel to the x-axis.

Example 10.3. Let $W = x\frac{\partial}{\partial y} - y\,\partial\partial x \in \mathfrak{X}(\mathbb{R}^2)$. If $\gamma\colon \mathbb{R} \to \mathbb{R}^2$ is a smooth curve, written in standard coordinates as $\gamma(t) = (x(t), y(t))$, then the condition $\gamma'(t) = W_{\gamma(t)}$ for γ to be an integral curve translates to

$$x'(t)\left.\frac{\partial}{\partial x}\right|_{\gamma(t)} + y'(t)\left.\frac{\partial}{\partial y}\right|_{\gamma(t)} = x(t)\left.\frac{\partial}{\partial y}\right|_{\gamma(t)} - y(t)\left.\frac{\partial}{\partial x}\right|_{\gamma(t)}.$$

Comparing the components of these vectors, we see that this is equivalent to the system of ordinary differential equations

$$\begin{aligned} x'(t) &= -y(t), \\ y'(t) &= x(t). \end{aligned}$$

These equations have the solutions

$$x(t) = a\cos t - b\sin t, \qquad y(t) = a\sin t + b\cos t,$$

for arbitrary constants a and b, and thus each curve of the form $\gamma(t) = (a\cos t - b\sin t, a\sin t + b\cos t)$ is an integral curve of W starting at (a, b). When $(a, b) = (0, 0)$, this is the constant curve $\gamma(t) \equiv (0, 0)$; otherwise, it is a circle traversed counterclockwise. Thus, given any $(a, b) \in \mathbb{R}^2$, we see once again that there is a unique integral curve starting at each point $(a, b) \in \mathbb{R}^2$, and the images of the various integral curves are either identical or disjoint.

From this example we see that given $V \in \mathfrak{X}(M)$, finding integral curve of V boils down to solving a system of ordinary differential equations in a smooth chart. More precisely, suppose V is a smooth vector field on M and $\gamma\colon J \to M$ is a smooth curve. On a smooth coordinate domain $U \subseteq M$, we can write γ in local coordinates as $\gamma(t) = (\gamma^1(t), \ldots, \gamma^n(t))$. Then (10.1) for γ to be an integral curve of V can be written

$$\dot{\gamma}^i(t)\left.\frac{\partial}{\partial x^i}\right|_{\gamma(t)} = V^i(\gamma(t))\left.\frac{\partial}{\partial x^i}\right|_{\gamma(t)},$$

which reduces to the following autonomous system of ordinary differential equations (ODEs):

$$\begin{aligned}\dot\gamma^1(t) &= V^1(\gamma^1(t),\dots,\gamma^n(t)),\\ &\vdots \\ \dot\gamma^n(t) &= V^n(\gamma^1(t),\dots,\gamma^n(t)).\end{aligned} \tag{10.2}$$

This means that to guarantee the existence of integral curves, we would need to solve an ordinary differential equation (ODE). Thus, we first recall the following fundamental result on ODEs.

Theorem 10.4 (Fundamental theorem for ODEs)**.** *Suppose $U \subset \mathbb{R}^n$ is open and $V : U \to \mathbb{R}^n$ is a smooth vector-valued function. Consider the system of ODE*

$$\begin{aligned}\dot\gamma^i(t) &= V^i(\gamma^1(t),\dots,\gamma^n(t)),\ i = 1,\dots,n,\\ \gamma^i(t_0) &= c^i,\ i = 1,\dots,n.\end{aligned}$$

with $c = (c^1,\dots,c^n) \in U$. Then

1. *(existence) $\forall t_0 \in \mathbb{R}$ and $x_0 \in U$ there exists an open interval $J_0 \ni t_0$ and an open subset $U_0 \subset U$ with $x_0 \in U_0$ such that for all $c \in U_0$ there exists a C^1-curve $\gamma : J_0 \to U$ that solves the ODE.*
2. *(Uniqueness) Any two solutions γ_1 and γ_2 of the above ODE are same on their common domain.*
3. *The map $\theta : J_0 \times U_0 \to U$ given by $\theta(t, x) = \gamma(t)$ where the γ is the unique solution above and $\gamma(t_0) = x$, is a smooth map.*

Using this theorem we have the following simple result.

Proposition 10.5. *Let V be a smooth vector field on a smooth manifold M. For each point $p \in M$, there exist $\varepsilon > 0$ and a smooth curve $\gamma\colon (-\varepsilon,\varepsilon) \to M$ that is an integral curve of V starting at p.* □

We notice some elementary properties of integral curves.

- If $a \in \mathbb{R}$ and $\gamma\colon J \to M$ is an integral curve of $V \in \mathfrak{X}(M)$, then the curve $\widetilde\gamma\colon \widetilde J \to M$ defined by $\widetilde\gamma(t) = \gamma(at)$ is an integral curve of the vector field aV, where $\widetilde J = \{t : at \in J\}$.

  Proof. We can examine the action of $\widetilde\gamma'(t)$ on a smooth real-valued function f defined in a neighbourhood of a point $\widetilde\gamma(t_0)$. By the chain rule and the fact that γ is an integral curve of V,

  $$\begin{aligned}\widetilde\gamma'(t_0)f &= \left.\frac{d}{dt}\right|_{t=t_0}(f\circ\widetilde\gamma)(t) = \left.\frac{d}{dt}\right|_{t=t_0}(f\circ\gamma)(at)\\ &= a(f\circ\gamma)'(at_0) = a\gamma'(at_0)f = aV_{\widetilde\gamma(t_0)}f. \qquad \square\end{aligned}$$

- Let V, M, J, and γ be as before. For any $b \in \mathbb{R}$, the curve $\widehat\gamma\colon \widehat J \to M$ defined by $\widehat\gamma(t) = \gamma(t+b)$ is also an integral curve of V, where $\widehat J = \{t : t + b \in J\}$.

- Suppose M and N are smooth manifolds and $F\colon M \to N$ is a smooth map. Then $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$ are F-related if and only if F takes integral curves of X to integral curves of Y, meaning that for each integral curve γ of X, $F\circ\gamma$ is an integral curve of Y.

  Proof. Suppose first that X and Y are F-related, and $\gamma\colon J \to M$ is an integral curve of X. If we define $\sigma\colon J \to N$ by $\sigma = F\circ\gamma$, then

  $$\sigma'(t) = (F\circ\gamma)'(t) = dF_{\gamma(t)}(\gamma'(t)) = dF_{\gamma(t)}(X_{\gamma(t)}) = Y_{F(\gamma(t))} = Y_{\sigma(t)},$$

  so σ is an integral curve of Y.

Conversely, suppose F takes integral curves of X to integral curves of Y. Given $p \in M$, let $\gamma\colon (-\varepsilon, \varepsilon) \to M$ be an integral curve of X starting at p. Since $F \circ \gamma$ is an integral curve of Y starting at $F(p)$, we have

$$Y_{F(p)} = (F \circ \gamma)'(0) = dF_p(\gamma'(0)) = dF_p(X_p),$$

which shows that X and Y are F-related. $\square$

10.2 Flows of vector fields

Let M be a smooth manifold and $V \in \mathfrak{X}(M)$, and suppose that for each point $p \in M$, V has a unique integral curve starting at p and defined for all $t \in \mathbb{R}$, which we denote by $\theta^{(p)} : \mathbb{R} \to M$. (We don't know whether this is always possible or not.) For each $t \in \mathbb{R}$, we can define a map $\theta_t : M \to M$ by sending each $p \in M$ to the point obtained by following for time t the integral curve starting at p:

$$\theta_t(p) = \theta^{(p)}(t).$$

Each map θ_t "slides" the manifold along the integral curves for time t. Then $t \mapsto \theta^{(p)}(t+s)$ is an integral curve of V starting at $q = \theta^{(p)}(s)$ and since we are assuming uniqueness of integral curves, $\theta^{(q)}(t) = \theta^{(p)}(t+s)$. When we translate this into a statement about the maps θ_t, it becomes

$$\theta_t \circ \theta_s(p) = \theta_{t+s}(p).$$

Together with the equation $\theta_0(p) = \theta^{(p)}(0) = p$, which holds by definition, this implies that the map

$\theta : \mathbb{R} \times M \to M$ is an action of the additive group $\mathbb{R}$ on M.

Definition 10.6. A **global flow** on M which is also sometimes called a **one-parameter group action** is a continuous left $\mathbb{R}$-action on M; that is, a continuous map $\theta : \mathbb{R} \times M \to M$ satisfying the following properties for all $s, t \in \mathbb{R}$ and $p \in M$:

$$\theta(t, \theta(s,p)) = \theta(t+s, p), \qquad \theta(0,p) = p. \tag{10.3}$$

Given a global flow θ on M, we have the following two collections of maps:

- For each $t \in \mathbb{R}$, define a continuous map $\theta_t : M \to M$ by

  $$\theta_t(p) = \theta(t,p).$$

  which satisfy

  $$\theta_t \circ \theta_s = \theta_{t+s}, \qquad \theta_0 = \mathrm{Id}_M. \tag{10.4}$$

  As is the case for any continuous group action, each map $\theta_t : M \to M$ is a homeomorphism, and if the flow is smooth, θ_t is a diffeomorphism.

- For each $p \in M$, define a curve $\theta^{(p)} : \mathbb{R} \to M$ by

  $$\theta^{(p)}(t) = \theta(t,p).$$

  The image of this curve is **the orbit of p under the group action.**

The next proposition shows that every smooth global flow is derived from the integral curves of some smooth vector field in precisely the way we described above. Before stating the result let us make the following

Definition 10.7. If $\theta : \mathbb{R} \times M \to M$ is a smooth global flow, for each $p \in M$ we define a tangent vector $V_p \in T_pM$ by

$$V_p = \theta^{(p)\prime}(0). \tag{10.5}$$

The assignment $p \mapsto V_p$ is a vector field on M, which is called the **infinitesimal generator of θ.**

Lemma 10.8. *Let $\theta : \mathbb{R} \times M \to M$ be a smooth global flow on a smooth manifold M. The infinitesimal generator V of θ is a smooth vector field on M, and each curve $\theta^{(p)}$ is an integral curve of V.*

Proof. To show that V is smooth we show that Vf is smooth for every $f \in C^\infty(U)$, $U \subseteq M$. For any such f and any $p \in U$, just note that

$$Vf(p) = V_p f = \theta^{(p)\prime}(0)f = \left.\frac{d}{dt}\right|_{t=0} f(\theta^{(p)}(t)) = \left.\frac{\partial}{\partial t}\right|_{(0,p)} f(\theta(t,p)).$$

Because $f(\theta(t,p))$ is a smooth function of (t,p) by composition, so is its partial derivative with respect to t. Thus, $Vf(p)$ depends smoothly on p, so V is smooth.

Next we need to show that $\theta^{(p)}$ is an integral curve of V, which means that $\theta^{(p)\prime}(t) = V_{\theta^{(p)}(t)}$ for all $p \in M$ and all $t \in \mathbb{R}$. Let $t_0 \in \mathbb{R}$ be arbitrary, and set $q = \theta^{(p)}(t_0) = \theta_{t_0}(p)$, so what we have to show is $\theta^{(p)\prime}(t_0) = V_q$. By the group law, for all t,

$$\theta^{(q)}(t) = \theta_t(q) = \theta_t(\theta_{t_0}(p)) = \theta_{t+t_0}(p) = \theta^{(p)}(t+t_0).$$

Therefore, for any smooth real-valued function f defined in a neighbourhood of q,

$$\begin{aligned} V_q f = \theta^{(q)\prime}(0)f &= \left.\frac{d}{dt}\right|_{t=0} f(\theta^{(q)}(t)) = \left.\frac{d}{dt}\right|_{t=0} f(\theta^{(p)}(t+t_0)) \\ &= \theta^{(p)\prime}(t_0)f, \end{aligned}$$

which was to be shown. □

Example 10.9 ((Global Flows).). The two vector fields on $\mathbb{R}^2$ described in Examples 10.2 and 10.3 both had integral curves defined for all $t \in \mathbb{R}$, so they generate global flows. Using the results of that example, we can write down the flows explicitly.

(a) The flow of $V = \partial/\partial x$ in $\mathbb{R}^2$ is the map $\tau\colon \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\tau_t(x,y) = (x+t, y).$$

For each nonzero $t \in \mathbb{R}$, τ_t translates the plane to the right ($t > 0$) or left ($t < 0$) by a distance $|t|$.

(b) The flow of $W = x\,\partial/\partial y - y\,\partial/\partial x$ is the map $\theta\colon \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\theta_t(x,y) = (x\cos t - y\sin t, x\sin t + y\cos t).$$

For each $t \in \mathbb{R}$, θ_t rotates the plane through an angle t about the origin.

Thus,

infinitesimal generators of global flows are smooth vector fields

A natural question arises: Is the converse of the above statement true, i.e., given a smooth $V \in \mathfrak{X}(M)$, must it be an infinitesimal generator of some global flow? Let's see some examples with smooth vector fields whose integral curves are not defined for all $t \in \mathbb{R}$.

Example 10.10. Let $M = \mathbb{R}^2 \setminus \{0\}$ with standard coordinates (x,y), and let V be the vector field $\frac{\partial}{\partial x}$ on M. The unique integral curve of V starting at $(-1,0) \in M$ is $\gamma(t) = (t-1, 0)$. However, in this case, γ cannot be extended continuously past $t = 1$.

Example 10.11. Let M be all of $\mathbb{R}^2$ and let $W = x^2\frac{\partial}{\partial x}$. The unique integral curve of W starting at $(1,0)$ is

$$\gamma(t) = \left(\frac{1}{1-t}, 0\right).$$

This curve also cannot be extended past $t = 1$, because its x-coordinate is unbounded as $t \nearrow 1$.

Thus, not every smooth vector field V gives rise to global flows by being their infinitesimal generator.

Definition 10.12. If M is a manifold, a **flow domain** for M is an open subset $\mathcal{D} \subseteq \mathbb{R} \times M$ with the property that for each $p \in M$, the set

$$\mathcal{D}^{(p)} = \{t \in \mathbb{R} : (t, p) \in \mathcal{D}\}$$

is an open interval containing 0.

A **flow** on M is a continuous map $\theta \colon \mathcal{D} \to M$, where $\mathcal{D} \subseteq \mathbb{R} \times M$ is a flow domain, that satisfies the following group laws: for all $p \in M$,

$$\theta(0, p) = p, \tag{10.6}$$

and for all $s \in \mathcal{D}^{(p)}$ and $t \in \mathcal{D}^{(\theta(s,p))}$ such that $s + t \in \mathcal{D}^{(p)}$,

$$\theta(t, \theta(s, p)) = \theta(t + s, p). \tag{10.7}$$

The flow θ is also sometimes called a **local flow** or a **local one-parameter group action**.

If θ is smooth, the ***infinitesimal generator of θ*** is defined by $V_p = \theta^{(p)\prime}(0)$.

Just like Lemma 10.8, we have the following result.

Lemma 10.13. *If $\theta \colon \mathcal{D} \to M$ is a smooth flow, then the infinitesimal generator V of θ is a smooth vector field, and each curve $\theta^{(p)}$ is an integral curve of V.*

We make the following definition.

Definition 10.14. A **maximal integral curve** is one that cannot be extended to an integral curve on any larger open interval and a **maximal flow** is a flow that admits no extension to a flow on a larger flow domain.

The next theorem is the main result of about flows of vector fields.

Theorem 10.15 (Fundamental Theorem on Flows). *Let V be a smooth vector field on a smooth manifold M. There is a unique smooth maximal flow $\theta \colon \mathcal{D} \to M$ whose infinitesimal generator is V. This flow has the following properties:*

(a) For each $p \in M$, the curve $\theta^{(p)} : \mathcal{D}^{(p)} \to M$ is the unique maximal integral curve of V starting at p.

(b) If $s \in \mathcal{D}^{(p)}$, then $\mathcal{D}^{(\theta(s,p))}$ is the interval $\mathcal{D}^{(p)} - s = \{t - s : t \in \mathcal{D}^{(p)}\}$.

(c) For each $t \in \mathbb{R}$, the set

$$M_t = \{p \in M \mid (t, p) \in \mathcal{D}\}$$

is open in M, and $\theta_t \colon M_t \to M_{-t}$ is a diffeomorphism with inverse θ_{-t}.

(d) For each $(t, p) \in \mathcal{D}$, $(\theta_t)_* V(p) = V(\theta(p))$.

Proof. The proof is a bit long so we proceed in stages. We first prove uniqueness in part (a). We know from Proposition 10.5 that given the smooth vector field V, there exists an integral curve starting at each point $p \in M$ for atleast some short time. Suppose $\gamma, \widetilde{\gamma} \colon J \to M$ are two integral curves of V defined on the same open interval J such that $\gamma(t_0) = \widetilde{\gamma}(t_0)$ for some $t_0 \in J$. We want to prove that in this case when $\gamma, \widetilde{\gamma}$ agree at one point, they must agree on every point of their common domain. Let

$$\mathcal{S} = \{t \in J \mid \gamma(t) = \widetilde{\gamma}(t)\}.$$

Clearly, $\mathcal{S} \neq \emptyset$, because $t_0 \in \mathcal{S}$ by hypothesis, and $\mathcal{S}$ is closed in J by continuity as if (t_n) is a sequence in $\mathcal{S}$ which converges to t, we can use continuity of $\gamma, \widetilde{\gamma}$ to get $\gamma(t_n) \to \gamma(t) = \widetilde{\gamma}(t_n) = \widetilde{\gamma}(t)$, thus proving that $t \in \mathcal{S}$. We now want to prove that $\mathcal{S}$ is also open and then use connectedness of J to prove that $J = \mathcal{S}$. Suppose $t_1 \in \mathcal{S}$. Then in a smooth coordinate neighbourhood around the point $p = \gamma(t_1)$, γ and $\widetilde{\gamma}$ are both solutions to same ODE with the same initial condition $\gamma(t_1) = \widetilde{\gamma}(t_1) = p$. By the uniqueness part of Theorem 10.4, $\gamma \equiv \widetilde{\gamma}$ on an interval containing t_1, which implies that $\mathcal{S}$ is open in J. Since J is connected, $\mathcal{S} = J$, which implies that $\gamma = \widetilde{\gamma}$ on all of J. Thus, any two integral curves that agree at one point agree on their common domain. Thus, if we manage to define a flow θ and then the coorresponding $\theta^{(p)}$ then it must be the unique integral curve starting at p.

We now prove maximality of the integral curve.Let $p \in M$ and let $\mathcal{D}^{(p)}$ be the union of all open intervals $J \subseteq \mathbb{R}$ containing 0 on which an integral curve starting at p is defined. Since we want to prove maximality of the domain, a natural thing would be to take the union and that is exactly what we are doing above. Define $\theta^{(p)} \colon \mathcal{D}^{(p)} \to M$ by letting $\theta^{(p)}(t) = \gamma(t)$, where γ is any integral curve starting at p and defined on an open interval containing 0 and t. Given any point t, since all such integral curves agree at t as explained above, $\theta^{(p)}$ is well defined,, i.e., it's image is independent of the integral curve chosen and is obviously the unique maximal integral curve starting at p as we have taken the union of all possible intervals.

Now let

$$\mathcal{D} = \{(t,p) \in \mathbb{R} \times M : t \in \mathcal{D}^{(p)}\},$$

and define $\theta \colon \mathcal{D} \to M$ by $\theta(t,p) = \theta^{(p)}(t)$. By definition, for each $p \in M$, $\theta^{(p)}$ is the unique maximal integral curve of V starting at p. To verify that θ is actually a flow, we will need to verify group laws and show that $\mathcal{D}$ is a flow domain. Fix any $p \in M$ and $s \in \mathcal{D}^{(p)}$, and write $q = \theta(s,p) = \theta^{(p)}(s)$. The curve $\gamma \colon \mathcal{D}^{(p)} - s \to M$ defined by $\gamma(t) = \theta^{(p)}(t+s)$ starts at q, and the translation property of integral curves shows that γ is an integral curve of V. By uniqueness of ODE solutions, γ agrees with $\theta^{(q)}$ on their common domain, which is equivalent to the second group law (10.7), and the first group law (10.6) is a by-product of the definition. Thus, θ is actually a flow.

By maximality of $\theta^{(q)}$, the domain of γ cannot be larger than $\mathcal{D}^{(q)}$, which means that $\mathcal{D}^{(p)} - s \subseteq \mathcal{D}^{(q)}$. Since $0 \in \mathcal{D}^{(p)}$, this implies that $-s \in \mathcal{D}^{(q)}$, and the group law implies that $\theta^{(q)}(-s) = \theta^{(p)}(s-s) = \theta^{(p)}(0) = p$. Applying the same argument with $(-s,q)$ in place of (s,p), we find that $\mathcal{D}^{(q)} + s \subseteq \mathcal{D}^{(p)}$, which is the same as $\mathcal{D}^{(q)} \subseteq \mathcal{D}^{(p)} - s$. This proves part (b).

We continue the proof of showing that the flow θ obtained above indeed is a smooth flow. That is, we show that $\mathcal{D}$ is open in $\mathbb{R} \times M$ (so it is a flow domain), and that $\theta \colon \mathcal{D} \to M$ is smooth. Define a subset $W \subseteq \mathcal{D}$ as the set of all $(t,p) \in \mathcal{D}$ such that θ is defined and smooth on a product neighbourhood of (t,p) of the form $J \times U \subseteq \mathcal{D}$, where $J \subseteq \mathbb{R}$ is an open interval containing 0 and $U \subseteq M$ is a neighbourhood of p. Then W is open in $\mathbb{R} \times M$, and the restriction of θ to W is smooth because of the smoothness of each integral curve $\theta^{(p)}$ which follows from the smoothness of the vector field V, so it suffices to show that $W = \mathcal{D}$. Suppose this is not the case. Then there exists some point $(\tau, p_0) \in \mathcal{D} \setminus W$. Without loss of generality, assume $\tau > 0$.

Let $t_0 = \inf\{t \in \mathbb{R}_+ : (t,p_0) \notin W\}$ (note that this is different from what we saw in the lectures because we added $\mathbb{R}_+$ in the definition, you can also take $\sup$ instead of $\inf$ and proceed similarly. The main idea remains unchanged.). By the ODE theorem, θ is defined and smooth in some product neighbourhood of $(0,p_0)$, so $t_0 > 0$. Since $t_0 \leq \tau$ and $\mathcal{D}^{(p_0)}$ is an open interval containing 0 and τ, it follows that $t_0 \in \mathcal{D}^{(p_0)}$. Let $q_0 = \theta^{(p_0)}(t_0)$. By the ODE theorem again, there exist $\varepsilon > 0$ and a neighbourhood U_0 of q_0 such that $(-\varepsilon, \varepsilon) \times U_0 \subseteq W$. We will use the group law to show that θ extends smoothly to a neighborhood of (t_0, p_0), which contradicts our choice of t_0. We will do this by using the translation property of integral curves.

Choose some $t_1 < t_0$ such that $t_1 + \varepsilon > t_0$ and $\theta^{(p_0)}(t_1) \in U_0$. We can do this because of the neighbourhood of smoothness being $(-\varepsilon, \varepsilon) \times U_0$ for the point q_0. Since $t_1 < t_0$, we have $(t_1, p_0) \in W$, as t_0 is the infimum, and so there is a product neighborhood $(t_1 - \delta, t_1 + \delta) \times U_1 \subseteq W$. By definition of W, this implies that θ is defined and smooth on $[0, t_1 + \delta) \times U_1$. Because $\theta(t_1, p_0) \in U_0$, we can choose U_1 small enough that θ

maps $\{t_1\} \times U_1$ into U_0. Define $\widetilde{\theta}\colon [0, t_1+\varepsilon) \times U_1 \to M$ by

$$\widetilde{\theta}(t,p) = \begin{cases} \theta_t(p), & p \in U_1,\ 0 \leq t < t_1, \\ \theta_{t-t_1} \circ \theta_{t_1}(p), & p \in U_1,\ t_1 - \varepsilon < t < t_1 + \varepsilon. \end{cases}$$

The group law for θ guarantees that these definitions agree where they overlap, and our choices of U_1, t_1, and ε ensure that this defines a smooth map. By the translation property, each map $t \mapsto \widetilde{\theta}(t,p)$ is an integral curve of V, so $\widetilde{\theta}$ is a smooth extension of θ to a neighbourhood of (t_0, p_0), contradicting our choice of t_0. This completes the proof that $W = \mathcal{D}$ and completes the proof of the statement that θ is a smooth flow and $\mathcal{D}$ as defined above is a flow domain.

We now prove (c). The fact that M_t is open is an immediate consequence of the fact that $\mathcal{D}$ is open as each slice at any fixed time t is precisely M_t. We now prove that θ_t indeed maps M_t to M_{-t}. From part (b) we deduce

$$\begin{aligned} p \in M_t \Rightarrow t \in \mathcal{D}^{(p)} &\Rightarrow \mathcal{D}^{(\theta_t(p))} = \mathcal{D}^{(p)} - t \\ &\Rightarrow -t \in \mathcal{D}^{(\theta_t(p))} \Rightarrow \theta_t(p) \in M_{-t}, \end{aligned}$$

which shows that θ_t maps M_t to M_{-t}. Moreover, the group laws then show that $\theta_{-t} \circ \theta_t$ is equal to the identity on M_t. Reversing the roles of t and $-t$ shows that $\theta_t \circ \theta_{-t}$ is the identity on M_{-t}, which proves that θ is a diffeomorphism.

Finally, we prove part (d). Let $\theta(t_0, p) = q$ so we would like to prove that $(\theta_{t_0})_*(V(p)) = V(\theta_{t_0}(p)) = V(q)$. Let $f \in C^\infty(M)$. Then

$$\begin{aligned} ((\theta_{t_0})_*(V(p)))f = V_p(f \circ \theta_{t_0}) &= \left.\frac{d}{dt}\right|_{t=0} f \circ \theta_{t_0} \circ \theta^p(t) \\ &= \left.\frac{d}{dt}\right|_{t=0} f(\theta_{t_0+t}(p)) \qquad \text{(using the group law)} \\ &= \left.\frac{d}{dt}\right|_{t=0} f(\theta^q(t)) \\ &= \left(\left.\frac{d}{dt}\right|_{t=0} \theta^q(t) \right) f = V(q)f \qquad \text{(as } \theta^q \text{ is an integral curve for } V\text{),} \end{aligned}$$

thus completeing the proof of the theorem. ☐

Definition 10.16. The flow θ in Theorem 10.15 is called the ***flow generated by*** $\boldsymbol{V}$, or just the ***flow of*** $\boldsymbol{V}$.

Heuristically, the term “infinitesimal generator” comes from fact that in a smooth chart, a good approximation to an integral curve can be obtained by composing many small straight-lines, with the direction and length of each motion determined by the value of the vector field at the point arrived at in the previous step. Intuitively, one can think of a flow as a sequence of infinitely many infinitesimally small linear steps and hence V is “generating" the flow θ.

Proposition 10.17 (Naturality of Flows)**.** *Suppose M and N are smooth manifolds, $F\colon M \to N$ is a smooth map, $X \in \mathfrak{X}(M)$, and $Y \in \mathfrak{X}(N)$. Let θ be the flow of X and η the flow of Y. If X and Y are F-related, then for each $t \in \mathbb{R}$, $F(M_t) \subseteq N_t$ and $\eta_t \circ F = F \circ \theta_t$ on M_t:*

$$\begin{array}{ccc} M_t & \xrightarrow{F} & N_t \\ \theta_t \downarrow & & \downarrow \eta_t \\ M_{-t} & \xrightarrow{F} & N_{-t}. \end{array}$$

Proof. We know from the property of integral curves that for any $p \in M$, the curve $F \circ \theta^{(p)}$ is an integral curve of Y starting at $F \circ \theta^{(p)}(0) = F(p)$. Thus, by uniqueness of integral curves, the maximal integral curve $\eta^{(F(p))}$ must be defined at least on the interval $\mathcal{D}^{(p)}$, and $F \circ \theta^{(p)} = \eta^{(F(p))}$ on that interval. This means that

$$p \in M_t \Rightarrow t \in \mathcal{D}^{(p)} \Rightarrow t \in \mathcal{D}^{(F(p))} \Rightarrow F(p) \in N_t,$$

which is equivalent to $F(M_t) \subseteq N_t$, and

$$F(\theta^{(p)}(t)) = \eta^{(F(p))}(t) \quad \text{for all } t \in \mathcal{D}^{(p)},$$

which is equivalent to $\eta_t \circ F(p) = F \circ \theta_t(p)$ for all $p \in M_t$. $\square$

As a corollary, we have that if $F : M \to N$ is a diffeomorphism and θ is the flow of a vector field $X \in \mathfrak{X}(M)$ then the flow of $F_* X$ is $\eta_t = F \circ \theta_t \circ F^{-1}$, with domain $N_t = F(M_t)$ for each $t \in \mathbb{R}$.

10.3 Complete Vector Fields

Definition 10.18 (Complete vector fields)**.** We say that a smooth vector field is ***complete*** if it generates a global flow, or equivalently if each of its maximal integral curves is defined for all $t \in \mathbb{R}$.

It is not always easy to determine by looking at a vector field whether it is complete or not. If you can solve the ODE explicitly to find all of the integral curves, and they all exist for all time, then the vector field is complete. On the other hand, if you can find a single integral curve that cannot be extended to all of $\mathbb{R}$, as we did for the vector fields of examples above, then it is not complete. However, it is often impossible to solve the ODE explicitly, so it is useful to have some general criteria for determining when a vector field is complete.

We will show below that all compactly supported smooth vector fields, and therefore all smooth vector fields on a compact manifold, are complete. The proof will be based on the following lemma.

Lemma 10.19. *Let V be a smooth vector field on a smooth manifold M, and let θ be its flow. Suppose there is a positive number ε such that for every $p \in M$, the domain of $\theta^{(p)}$ contains $(-\varepsilon, \varepsilon)$. Then V is complete.*

Remark 10.20. What is this lemma actually saying? It is saying that if an integral curve of a vector field stops existing then the vector field must be "blowing up" in finite time. In other words, if an integral curve of a vector field can survive for even a short uniform time for every point in the manifold then it must exist for all time. Intuitively, the hypothesis that $\theta^{(p)}$ contains $(-\varepsilon, \varepsilon)$ for all p should explain that you can keep on applying the translation lemma and extending the existence time for the integral curve.

Proof. We prove it by contradiction. The basic idea is to use the translation property of integral curves. Suppose for the sake of contradiction that for some $p \in M$, the domain $\mathcal{D}^{(p)}$ of $\theta^{(p)}$ is bounded above. (A similar proof works if it is bounded below.) Let $b = \sup \mathcal{D}^{(p)}$, let t_0 be a positive number such that $b - \varepsilon < t_0 < b$, and let $q = \theta^{(p)}(t_0)$. The hypothesis implies that $\theta^{(q)}(t)$ is defined at least for $t \in (-\varepsilon, \varepsilon)$. Define a curve $\gamma \colon (-\varepsilon, t_0 + \varepsilon) \to M$ by

$$\gamma(t) = \begin{cases} \theta^{(p)}(t), & -\varepsilon < t < b, \\ \theta^{(q)}(t - t_0), & t_0 - \varepsilon < t < t_0 + \varepsilon. \end{cases}$$

These two definitions agree where they overlap, because $\theta^{(q)}(t - t_0) = \theta_{t-t_0}(q) = \theta_{t-t_0} \circ \theta_{t_0}(p) = \theta_t(p) = \theta^{(p)}(t)$ by the group law for θ. By the translation property, γ is an integral curve starting at p. Since $t_0 + \varepsilon > b$, this means that γ starting from p is existing beyond b, which is a contradiction to the definition of b. Thus, γ exists for all $t \in \mathbb{R}$ and hence V is a complete vector field. $\square$

Theorem 10.21. *Every compactly supported smooth vector field on a smooth manifold is complete. Thus, on a compact smooth manifold, every smooth vector field is complete.*

Proof. Suppose V is a compactly supported vector field on a smooth manifold M, and let $K = \operatorname{supp} V$. For each $p \in K$, there is a neighbourhood U_p of p and a positive number ε_p such that the flow of V is defined at least on $(-\varepsilon_p, \varepsilon_p) \times U_p$. By compactness, finitely many such sets $U_{p_1}, \dots, U_{p_k}$ cover K. With $\varepsilon = \min\{\varepsilon_{p_1}, \dots, \varepsilon_{p_k}\}$, it follows that every maximal integral curve starting in K is defined at least on $(-\varepsilon, \varepsilon)$. Since $V \equiv 0$ outside of K, every integral curve starting in $M \smallsetminus K$ is constant and thus can be defined on all of $\mathbb{R}$. Thus the hypotheses of Lemma 10.19 are satisfied, so V is complete. $\square$

11. Lie Derivatives

We know how to differentiate real-valued functions in the direction of a vector field on a manifold. Indeed, a tangent vector $v \in T_pM$ is by definition an operator that acts on a smooth function f to give a number vf that we interpret as a directional derivative of f at p. How can we make sense of directional derivatives of a vector field in the direction of a given vector? Let us start from the familiar case of Euclidean spaces. We can define the directional derivative of a smooth vector field W in the direction of a vector $v \in T_p\mathbb{R}^n$. It is the vector

$$D_vW(p) = \left.\frac{d}{dt}\right|_{t=0} W_{p+tv} = \lim_{t\to 0} \frac{W_{p+tv} - W_p}{t}. \tag{11.1}$$

An easy calculation shows that $D_vW(p)$ can be evaluated by applying D_v to each component of W separately:

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

Unfortunately, this definition is heavily dependent upon the fact that $\mathbb{R}^n$ is a vector space, so that the tangent vectors W_{p+tv} and W_p can both be viewed as elements of $\mathbb{R}^n$. If we want to mimic (11.1) on a manifold, we can try to replace $p + tv$ by a curve $\gamma(t)$ that starts at p and whose initial velocity is v. But even with this substitution, the difference quotient still makes no sense because $W_{\gamma(t)}$ and $W_{\gamma(0)}$ are elements of the two different vector spaces $T_{\gamma(t)}M$ and $T_{\gamma(0)}M$, respectively. It was not a problem in the Euclidean space because there is a canonical identification of each tangent space with $\mathbb{R}^n$ itself; but on a manifold there is no such identification. Thus there is no coordinate-independent way to make sense of the directional derivative of W in the direction of a vector v.

We fix this problem by replacing the vector $v \in T_pM$ with a *vector field* $V \in \mathfrak{X}(M)$, so we can use the flow of V to push values of W back to p and then differentiate.

Definition 11.1. Suppose M is a smooth manifold, V is a smooth vector field on M, and θ is the flow of V. For any smooth vector field W on M, define a vector field on M, denoted by $\mathcal{L}_VW$ and called the ***Lie derivative of W with respect to V***, by

$$\begin{aligned}(\mathcal{L}_VW)_p &= \left.\frac{d}{dt}\right|_{t=0} d(\theta_{-t})_{\theta_t(p)}\big(W_{\theta_t(p)}\big) \\ &= \lim_{t\to 0} \frac{d(\theta_{-t})_{\theta_t(p)}\big(W_{\theta_t(p)}\big) - W_p}{t},\end{aligned} \tag{11.2}$$

provided the derivative exists.

For small $t \neq 0$, at least the difference quotient makes sense: θ_t is defined in a neighbourhood of p, and θ_{-t} is the inverse of θ_t, so both $d(\theta_{-t})_{\theta_t(p)}\big(W_{\theta_t(p)}\big)$ and W_p are elements of T_pM.

Lemma 11.2. *Suppose M is a smooth manifold and $V, W \in \mathfrak{X}(M)$. Then $(\mathcal{L}_VW)_p$ exists for every $p \in M$, and $\mathcal{L}_VW$ is a smooth vector field.*

Proof. We check that the component functions of the vector field $\mathcal{L}_VW$ are smooth in any coordinate chart. Let θ be the flow of V and for an arbitrary $p \in M$, let $(U, (x^i))$ be a smooth chart containing p. Choose an open interval J_0 containing 0 and an open subset $U_0 \subseteq U$ containing p such that θ maps $J_0 \times U_0$ into U. We can always do that using continuity of the flow θ. For $(t, x) \in J_0 \times U_0$, write the component functions of θ as $\big(\theta^1(t, x), \dots, \theta^n(t, x)\big)$. Then for any $(t, x) \in J_0 \times U_0$, the matrix of $d(\theta_{-t})_{\theta_t(x)} : T_{\theta_t(x)}M \to T_xM$ is

$$\left[\frac{\partial \theta^i}{\partial x^j}(-t, \theta(t, x))\right]_{ij},$$

and hence,

$$d(\theta_{-t})_{\theta_t(x)}\big(W_{\theta_t(x)}\big) = \frac{\partial \theta^i}{\partial x^j}(-t, \theta(t,x))\, W^j(\theta(t,x))\, \frac{\partial}{\partial x^i}\bigg|_x .$$

Since θ^i and W^j are smooth functions, the coefficient of $\frac{\partial}{\partial x^i}|_x$ depends smoothly on (t,x). It follows that $(\mathcal{L}_V W)_x$, which is obtained by taking the derivative of this expression with respect to t and setting $t = 0$, exists for each $x \in U_0$ and depends smoothly on x. □

While equation 11.2 explains exactly how to make sense of the derivative of a vector field in the direction of another, it is not very useful for explicit computations, because typically the flow is difficult or impossible to write down explicitly. Fortunately, there is a simple formula for computing the Lie derivative without explicitly finding the flow.

Theorem 11.3. *If M is a smooth manifold and $X, Y \in \mathfrak{X}(M)$, then $\mathcal{L}_X Y = [X, Y]$.*

Proof. Let θ be the flow of X and let $p \in M$. We first note that (11.2) is equivalent to the expression

$$(\mathcal{L}_X Y)_p = \frac{d}{dt}(\theta_t^* Y)_p|_{t=0}, \quad \text{where } (\theta_t^* Y)_p = (\theta_t^{-1})_{*\theta_t(p)} Y_{\theta_t(p)}.$$

Recall also that for any $f \in C^\infty(M)$, $(Xf)_p = \frac{d}{dt}(\theta_t^* f)_p|_{t=0}$ where $\theta_t^* f = f \circ \theta_t$ and similarly we have $(\theta_t^* Y f)_p = (\theta_t^* Y)_p(\theta_t^* f)$. We now use all this and compute. We have

$$(X(Yf))_p = \lim_{t\to 0} \frac{(\theta_t^* Y f)_p - (Yf)_p}{t} = \lim_{t\to 0} \frac{(\theta_t^* Y)_p(\theta_t^* f) - (Yf)_p}{t}.$$

We add and subtract $(\theta_t^* Y)_p f$ to get

$$\begin{aligned}(X(Yf))_p &= \lim_{t\to 0} \frac{(\theta_t^* Y)_p(\theta_t^* f) - (\theta_t^* Y)_p f + (\theta_t^* Y)_p f - Y_p f}{t} \\ &= \lim_{t\to 0} (\theta_t^* Y)_p \frac{(\theta_t^* f) - f}{t} + \lim_{t\to 0} \frac{(\theta_t^* Y)_p - Y_p}{t} f = Y_p X f + (\mathcal{L}_X Y)_p.\end{aligned}$$

Thus, we get that $XY = YX + \mathcal{L}_X Y$.

□

We now have a geometric interpretation of the Lie bracket of two vector fields: it is the directional derivative of the second vector field along the flow of the first. We immediately get the following.

Proposition 11.4. Suppose M is a smooth manifold and $V, W, X \in \mathfrak{X}(M)$.

1. *$\mathcal{L}_V W = -\mathcal{L}_W V$.*
2. *$\mathcal{L}_V [W, X] = [\mathcal{L}_V W, X] + [W, \mathcal{L}_V X]$.*
3. *$\mathcal{L}_{[V,W]} X = \mathcal{L}_V \mathcal{L}_W X - \mathcal{L}_W \mathcal{L}_V X$.*
4. *If $g \in C^\infty(M)$, then $\mathcal{L}_V(gW) = (Vg)W + g\mathcal{L}_V W$.*
5. *If $F : M \to N$ is a diffeomorphism, then $F_*(\mathcal{L}_V X) = \mathcal{L}_{F_* V} F_* X$.*

Notice that Proposition 11.4 (d) is immediate from Proposition 9.19 (d) and it is of this form because the Lie bracket $[fV, gW]$ can be thought of as the Lie derivative $\mathcal{L}_{fV}(gW)$, it satisfies a product rule in g and W; and because it can also be thought of as $-\mathcal{L}_{gW}(fV)$, it satisfies a product rule in f and V as well.

If V and W are vector fields on M and θ is the flow of V, the Lie derivative $(\mathcal{L}_V W)_p$, by definition, expresses the t-derivative of the time-dependent vector $d(\theta_{-t})_{\theta_t(p)}\big(W_{\theta_t(p)}\big) \in T_p M$ at $t = 0$. The next proposition shows how it can also be used to compute the derivative of this expression at other times.

Proposition 11.5. Suppose M is a smooth manifold and $V, W \in \mathfrak{X}(M)$. Let θ be the flow of V. For any (t_0, p) in the domain of θ,

$$\left.\frac{d}{dt}\right|_{t=t_0} d(\theta_{-t})_{\theta_t(p)}\big(W_{\theta_t(p)}\big) = d(\theta_{-t_0})\big((\mathcal{L}_V W)_{\theta_{t_0}(p)}\big). \tag{11.3}$$

Proof. Let $p \in M$ be arbitrary, let $\mathcal{D}^{(p)} \subseteq \mathbb{R}$ denote the domain of the integral curve $\theta^{(p)}$, and consider the map $X : \mathcal{D}^{(p)} \to T_pM$ given by $X(t) = d(\theta_{-t})_{\theta_t(p)}\big(W_{\theta_t(p)}\big)$. The same idea as in the proof of Lemma 11.2 shows that X is a smooth curve in the vector space T_pM. Making the change of variables $t = t_0 + s$, we obtain

$$\begin{aligned} X'(t_0) &= \left.\frac{d}{ds}\right|_{s=0} X(t_0 + s) = \left.\frac{d}{ds}\right|_{s=0} d(\theta_{-t_0-s})\big(W_{\theta_{s+t_0}(p)}\big) \\ &= \left.\frac{d}{ds}\right|_{s=0} d(\theta_{-t_0}) \circ d(\theta_{-s})\big(W_{\theta_s(\theta_{t_0}(p))}\big) \\ &= d(\theta_{-t_0})\left(\left.\frac{d}{ds}\right|_{s=0} d(\theta_{-s})\big(W_{\theta_s(\theta_{t_0}(p))}\big)\right). \end{aligned}$$

The last equality follows because $d(\theta_{-t_0}) : T_{\theta_{t_0}(p)}M \to T_pM$ is a linear map that is independent of s. By definition of the Lie derivative, this last expression is equal to the right-hand side of (11.3). $\square$

11.1 Commuting Vector Fields

Definition 11.6. Let M be a smooth manifold and $V, W \in \mathfrak{X}(M)$. We say that V ***and*** W ***commute*** if $VWf = WVf$ for every smooth function f, or equivalently if $[V, W] \equiv 0$.

If θ is a smooth flow, a vector field W is said to be ***invariant under*** θ if W is θ_t-related to itself for each t or equivalently that $d(\theta_t)_p(W_p) = W_{\theta_t(p)}$ for all (t, p) in the domain of θ.

The next proposition shows that these two concepts are intimately related.

Proposition 11.7. *For smooth vector fields V and W on a smooth manifold M, the following are equivalent:*

(a) V and W commute.

(b) W is invariant under the flow of V.

(c) V is invariant under the flow of W.

As a result, every smooth vector field is invariant under its own flow.

Proof. Suppose $V, W \in \mathfrak{X}(M)$, and let θ denote the flow of V. We prove (b) $\implies$ (a). If (b) holds, then $W_{\theta_t(p)} = d(\theta_t)_p(W_p)$ whenever (t, p) is in the domain of θ. Applying $d(\theta_{-t})_{\theta_t(p)}$ to both sides, we conclude that $d(\theta_{-t})_{\theta_t(p)}\big(W_{\theta_t(p)}\big) = W_p$, which by the definition of the Lie derivative or (11.2) implies $[V, W] = \mathcal{L}_V W = 0$. The same argument shows that (c) implies (a).

To prove that (a) implies (b), assume that $[V, W] = \mathcal{L}_V W = 0$. Let $p \in M$ be arbitrary, and let $X(t) = d(\theta_{-t})_{\theta_t(p)}\big(W_{\theta_t(p)}\big)$ for $t \in \mathcal{D}^{(p)}$. Proposition 11.5 shows that $X'(t) \equiv 0$. Since $X(0) = W_p$, this implies that $X(t) = W_p$ for all $t \in \mathcal{D}^{(p)}$, and applying $d(\theta_t)_p$ to both sides yields the identity that says W is invariant under θ. Similarly (a) implies (c).

$\square$

We now relate commuting vector fields in terms of the relationship between their respective flows. More precisely, the next theorem will prove that two vector fields commute if and only if their flows commute. But what does two flows commuting with each other means?

Suppose V and W are smooth vector fields on M, and let θ and ψ be their respective flows. If V and W are complete vector fields and hence θ and ψ are global flows, then their flows commute means

$$\theta_t \circ \psi_s = \psi_s \circ \theta_t \qquad \text{for all } s, t \in \mathbb{R}.$$

However, if either V or W is not complete, the most we can hope for is that this equation holds for all s and t such that both sides are defined.

Unfortunately, even when the vector fields commute, their flows might not commute in this naive sense, because there are examples of commuting vector fields V and W and particular choices of t, s, and p for which both $\theta_t \circ \psi_s(p)$ and $\psi_s \circ \theta_t(p)$ are defined, but they are not equal.

Exercise 11.8. Consider $\mathbb{R}^3$ and consider the vector fields

$$V = \frac{\partial}{\partial x} - \frac{y}{x^2+y^2}\frac{\partial}{\partial z}, \quad W = \frac{\partial}{\partial y} + \frac{x}{x^2+y^2}\frac{\partial}{\partial z}.$$

Let θ and ψ be the flow of V and W respectively. Show that V and W commute but there exist $p \in \mathbb{R}^3$ and $s, t \in \mathbb{R}$ such that $\theta_t \circ \psi_s(p)$ and $\psi_s \circ \theta_t(p)$ both exist but are not equal. (**Hint:** for instance, $p = (1, -1, 0)$ and $s = 2, t = -2$ does the job. Maybe write the equation for the integral curves.)

Here is the problem: if $\theta_t \circ \psi_s(p)$ is defined for $t = t_0$ and $s = s_0$, then by the properties of flow domains, it must be defined for all t in some open interval containing 0 and t_0, but the analogous statement need not be true of s - there might be values of s between 0 and s_0 for which the integral curve of V starting at $\psi_s(p)$ does not extend all the way to $t = t_0$.

Thus we make the following definition.

Definition 11.9. If θ and ψ are flows on M, we say that θ ***and*** ψ ***commute*** if the following condition holds for every $p \in M$: whenever J and K are open intervals containing 0 such that one of the expressions $\theta_t \circ \psi_s(p)$ or $\psi_s \circ \theta_t(p)$ is defined for all $(s,t) \in J \times K$, both are defined and they are equal. For global flows, this is the same as saying that $\theta_t \circ \psi_s = \psi_s \circ \theta_t$ for all s and t.

Here is the main result for commuting flows.

Theorem 11.10. *Smooth vector fields commute if and only if their flows commute.*

Proof. Let V and W be smooth vector fields on a smooth manifold M, and let θ and ψ denote their respective flows. Assume first that V and W commute. Suppose that $p \in M$, and J and K are open intervals containing 0 such that $\psi_s \circ \theta_t(p)$ is defined for all $(s,t) \in J \times K$. By Proposition 11.7, the hypothesis implies that V is invariant under ψ. Fix any $s \in J$, and consider the curve $\gamma : K \to M$ defined by $\gamma(t) = \psi_s \circ \theta_t(p) = \psi_s\big(\theta^{(p)}(t)\big)$. This curve satisfies $\gamma(0) = \psi_s(p)$, and its velocity at $t \in K$ is

$$\gamma'(t) = \frac{d}{dt}\big(\psi_s(\theta^{(p)}(t))\big) = d(\psi_s)\big(\theta^{(p)\prime}(t)\big) = d(\psi_s)(V_{\theta^{(p)}(t)}) = V_{\gamma(t)}.$$

Thus, γ is an integral curve of V starting at $\psi_s(p)$. By uniqueness, therefore,

$$\gamma(t) = \theta^{\psi_s(p)}(t) = \theta_t\big(\psi_s(p)\big).$$

This proves that θ and ψ commute.

Conversely, assume that the flows commute, and let $p \in M$. If $\varepsilon > 0$ is chosen small enough that $\psi_s \circ \theta_t(p)$ is defined whenever $|s| < \varepsilon$ and $|t| < \varepsilon$, then the hypothesis guarantees that $\psi_s \circ \theta_t(p) = \theta_t \circ \psi_s(p)$ for all such s and t. This can be rewritten in the form

$$\psi^{\theta_t(p)}(s) = \theta_t\big(\psi^{(p)}(s)\big).$$

Differentiating this relation with respect to s, we get

$$W_{\theta_t(p)} = \frac{d}{ds}\bigg|_{s=0} \psi^{\theta_t(p)}(s) = \frac{d}{ds}\bigg|_{s=0} \theta_t\big(\psi^{(p)}(s)\big) = d(\theta_t)_p(W_p).$$

Applying $d(\theta_{-t})_{\theta_t(p)}$ to both sides, we conclude

$$d(\theta_{-t})_{\theta_t(p)}\big(W_{\theta_t(p)}\big) = W_p.$$

Differentiating with respect to t and applying the definition of the Lie derivative shows that $(\mathcal{L}_V W)_p = 0$. $\square$

12. Cotangent bundle and 1-forms

Recall that if V is a finite-dimensional $\mathbb{R}$-vector space then the space of all linear functionals, i.e., linear maps $\omega : V \to \mathbb{R}$ is denoted by V^* and is called the dual space to V. Elements of V^* are called **covectors** on V. In fact, we immediately get a basis of V^* once we choose a basis of V as given in the next proposition.

Proposition 12.1. *Let V be a finite-dimensional vector space. Given any basis $(e_1, \ldots, e_n)$ for V, let $\varepsilon^1, \ldots, \varepsilon^n \in V^*$ be the covectors defined by*

$$\varepsilon^i(e_j) = \delta^i_j,$$

*where δ^i_j is the Kronecker delta symbol. Then $(\varepsilon^1, \ldots, \varepsilon^n)$ is a basis for V^*, called the **dual basis** to (e_j) and hence $\dim V^* = \dim V$.*

So, for instance, for the standard basis $(e_1, \ldots, e_n)$ of $\mathbb{R}^n$, the dual basis is denoted by $\left(e^1, \ldots, e^n\right)$, and is called the ***standard dual basis***. Notice that we are using upper indices for denoting the covectors. These basis covectors are the linear functionals on $\mathbb{R}^n$ given by

$$e^i(v) = e^i\left(v^1, \ldots, v^n\right) = v^i.$$

In matrix notation, a linear map from $\mathbb{R}^n$ to $\mathbb{R}$ is represented by a $1 \times n$ matrix, that is, it is a ***row matrix***. The basis covectors can therefore also be thought of as the linear functionals represented by the row matrices

$$e^1 = (1\ 0\ \ldots\ 0), \quad e^2 = (0\ 1\ 0\ \ldots\ 0), \quad \ldots, \quad e^n = (0\ \ldots\ 0\ 1).$$

In general, if (e_j) is a basis for V and (ε^i) is its dual basis, then for any vector $v = v^j e_j \in V$, we have

$$\varepsilon^i(v) = v^j\, \varepsilon^i(e_j) = v^j\, \delta^i_j = v^i.$$

We can express an arbitrary covector $\omega \in V^*$ in terms of the dual basis as

$$\omega = \omega_i\, \varepsilon^i, \tag{12.1}$$

where the components are determined by $\omega_i = \omega(e_i)$ and similarly, the action of ω on a vector $v = v^j e_j$ is

$$\omega(v) = \omega_i\, v^i. \tag{12.2}$$

Notation: We always write basis covectors with upper indices, and components of a covector with lower indices.

Suppose V and W are vector spaces and $L\colon V \to W$ is a linear map. We define a linear map $L^*\colon W^* \to V^*$, called the ***dual map*** or ***transpose of L***, by

$$(L^*\omega)(v) = \omega(Lv) \quad \text{for } \omega \in W^*,\ v \in V.$$

The dual maps satisfies the following properties.

(1) $(A \circ B)^* = B^* \circ A^*$.

(2) $(\mathrm{Id}_V)^*\colon V^* \to V^*$ *is the identity map of* V^*.

For a vector space V, we also have a ***second dual space*** $V^{**} = (V^*)^*$. For each vector space V there is a natural, basis-independent map $\xi\colon V \to V^{**}$, defined as follows. For each vector $v \in V$, define a linear functional $\xi(v)\colon V^* \to \mathbb{R}$ by

$$\xi(v)(\omega) = \omega(v) \quad \text{for } \omega \in V^*.$$

Proposition 12.2. *For any finite-dimensional vector space V, the map $\xi\colon V \to V^{**}$ is an isomorphism.*

Proof. Because $\dim V = \dim V^{**}$, it suffices to verify that ξ is injective. Suppose $v \in V$ is not zero. Extend v to a basis $(v = e_1, \dots, e_n)$ for V, and let $(\varepsilon^1, \dots, \varepsilon^n)$ denote the dual basis for V^*. Then $\xi(v) \neq 0$ because

$$\xi(v)\left(\varepsilon^1\right) = \varepsilon^1(v) = \varepsilon^1(e_1) = 1.$$

□

The preceding proposition shows that when V is finite-dimensional, we can unambiguously identify V^{**} with V itself, because the map ξ is canonically defined, without reference to any basis. It is important to observe that although V^* is also isomorphic to V (for the simple reason that any two finite-dimensional vector spaces of the same dimension are isomorphic), there is no *canonical* isomorphism $V \cong V^*$.

12.1 Tangent Covectors on Manifolds

Now let M be a smooth manifold. For each $p \in M$, we define the ***cotangent space at*** p, denoted by T_p^*M, to be the dual space to T_pM

$$T_p^*M = (T_pM)^*.$$

Elements of T_p^*M are called ***tangent covectors at*** p, or just ***covectors at*** p.

Let (U, x^i) be a coordinate chart. Then the coordinate basis $\left(\frac{\partial}{\partial x^i}\big|_p\right)$ gives rise to a dual basis for T_p^*M, which we denote by $\left(\lambda^i\big|_p\right)$. Any covector $\omega \in T_p^*M$ can thus be written uniquely as $\omega = \omega_i\, \lambda^i\big|_p$, where

$$\omega_i = \omega\left(\frac{\partial}{\partial x^i}\bigg|_p\right).$$

Let us try to understand how the covectors look like under change of coordinate system. Suppose that $(\tilde{x}^j)$ is another set of smooth coordinates whose domain contains p, and let $\left(\tilde{\lambda}^j\big|_p\right)$ denote the basis for T_p^*M dual to $\left(\frac{\partial}{\partial \tilde{x}^j}\big|_p\right)$. We recall that the coordinate vector fields transform as follows:

$$\frac{\partial}{\partial x^i}\bigg|_p = \frac{\partial \tilde{x}^j}{\partial x^i}(p)\ \frac{\partial}{\partial \tilde{x}^j}\bigg|_p. \tag{12.3}$$

Writing ω in both systems as $\omega = \omega_i\, \lambda^i\big|_p = \tilde{\omega}_j\, \tilde{\lambda}^j\big|_p$, we can use (12.3) to compute the components ω_i in terms of $\tilde{\omega}_j$:

$$\omega_i = \omega\left(\frac{\partial}{\partial x^i}\bigg|_p\right) = \omega\left(\frac{\partial \tilde{x}^j}{\partial x^i}(p)\ \frac{\partial}{\partial \tilde{x}^j}\bigg|_p\right) = \frac{\partial \tilde{x}^j}{\partial x^i}(p)\tilde{\omega}_j. \tag{12.4}$$

Thus, the components of a vector, i.e., the n-tuples $(v^1, \dots, v^n)$ and $(\tilde{v}^1, \dots, \tilde{v}^n)$ assigned to two different coordinate systems (x^i) and $(\tilde{x}^j)$ were related by the transformation law

$$\tilde{v}^j = \frac{\partial \tilde{x}^j}{\partial x^i}(p)v^i,$$

and a covector transforms, according to the following slightly different rule:

$$\omega_i = \frac{\partial \tilde{x}^j}{\partial x^i}(p)\tilde{\omega}_j. \tag{11.7}$$

12.2 Covector Fields or 1-forms

For any smooth manifold M the disjoint union

$$T^*M = \coprod_{p \in M} T_p^*M$$

is called the ***cotangent bundle of*** M. It has a natural projection map $\pi\colon T^*M \to M$ sending $\omega \in T_p^*M$ to $p \in M$. As above, given any smooth local coordinates (x^i) on an open subset $U \subseteq M$, for each $p \in U$ we denote the basis for T_p^*M dual to $\left(\frac{\partial}{\partial x^i}\big|_p\right)$ by $\left(\lambda^i\big|_p\right)$. This defines n maps $\lambda^1, \dots, \lambda^n\colon U \to T^*M$, called ***coordinate covector fields***. Just like the case for tangent bundle Theorem 5.8, the next result states that the cotangent bundle is also a $2n$-dimensional manifold with a smooth structure inherited from M.

Theorem 12.3. *Let M be a smooth n-manifold. The cotangent bundle T^*M has a unique topology and smooth structure making it into a smooth $2n$-dimensional manifold.*

Proof. We leave the proof as an exercise. □

As in the case of the tangent bundle, smooth local coordinates for M yield smooth local coordinates for its cotangent bundle. If (x^i) are smooth coordinates on an open subset $U \subseteq M$,then the map from $\pi^{-1}(U)$ to $\mathbb{R}^{2n}$ given by

$$\xi_i \, \lambda^i\big|_p \mapsto \left(x^1(p), \dots, x^n(p), \xi_1, \dots, \xi_n\right)$$

is a smooth coordinate chart for T^*M. We call (x^i, ξ_i) the ***natural coordinates for*** T^*M associated with (x^i).

Again, as was the case with vector fields, we have the following definition.

Definition 12.4. A **covector field** or a **(differential) 1-form** is a continuous map $\omega : M \to T^*M$ such that

$$\pi \circ \omega = \mathrm{id}_M.$$

Thus, for every $p \in M$, $\omega_p \in T_p^*M$.

In any smooth local coordinates on an open subset $U \subseteq M$, a covector field ω can be written in terms of the coordinate covector fields (λ^i) as $\omega = \omega_i \, \lambda^i$ for n functions $\omega_i\colon U \to \mathbb{R}$ called the ***component functions of ω***. They are characterized by

$$\omega_i(p) = \omega_p\left(\frac{\partial}{\partial x^i}\bigg|_p\right).$$

If ω is a covector field and X is a vector field on M, then we can form a function $\omega(X)\colon M \to \mathbb{R}$ by

$$\omega(X)(p) = \omega_p(X_p), \quad p \in M.$$

If we write $\omega = \omega_i \, \lambda^i$ and $X = X^j \frac{\partial}{\partial x^j}$ in terms of local coordinates, then $\omega(X)$ has the local coordinate representation $\omega(X) = \omega_i \, X^i$.

Just as in the case of vector fields, there are several ways to check for smoothness of a covector field.

Proposition 12.5. *Let M be a smooth manifold and let $\omega\colon M \to T^*M$ be a covector field. The following are equivalent:*

1. *ω is smooth.*
2. *In every smooth coordinate chart, the component functions of ω are smooth.*
3. *For every smooth vector field $X \in \mathfrak{X}(M)$, the function $\omega(X)$ is smooth on M.*
4. *For every open subset $U \subseteq M$ and every smooth vector field X on U, the function $\omega(X)\colon U \to \mathbb{R}$ is smooth on U.*

We denote the real vector space of all smooth covector fields on M by $\mathfrak{X}^*M$ or $\Omega^1(M)$. For $f \in C^\infty(M)$, the covector field $f\omega$ is defined as

$$(f\omega)_p = f(p)\omega_p. \tag{12.5}$$

12.3 The Differential of a Function

Recall that the gradient of a smooth real-valued function f on $\mathbb{R}^n$ is defined as the vector field whose components are the partial derivatives of f

$$\operatorname{grad} f = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^i}.$$

Although the partial derivatives of a smooth function cannot be interpreted in a coordinate-independent way as the components of a vector field, it turns out that they can be interpreted as the components of a covector field. This is the most important application of covector fields.

Definition 12.6. Let f be a smooth real-valued function on a smooth manifold M We define a covector field df, called the ***differential of f***, by

$$df_p(v) = vf \quad \text{for } v \in T_pM.$$

Proposition 12.7. *The differential of a smooth function is a smooth covector field.*

Proof. Clearly at each point $p \in M$, $df_p(v)$ depends linearly on v, so that df_p is indeed a covector at p. To see that df is smooth, we use Proposition 12.5 (3): for any smooth vector field X on M, the function $df(X)$ is smooth because it is equal to Xf. □

Let us compute df in coordinates or in other words, let's compute its coordinate representation. Let (x^i) be smooth coordinates on an open subset $U \subseteq M$, and let (λ^i) be the corresponding coordinate dual basis on U. Write df in coordinates as $df_p = A_i(p)\, \lambda^i\big|_p$ for some functions $A_i \colon U \to \mathbb{R}$; then the definition of df implies

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

This yields the following formula for the coordinate representation of df:

$$df_p = \frac{\partial f}{\partial x^i}(p)\, \lambda^i\Big|_p. \tag{12.6}$$

Thus, the component functions of df in any smooth coordinate chart are the partial derivatives of f with respect to those coordinates. As a result, the differential of f in coordinates is analogous to the gradient of f in local coordinates but df makes sense irrespective of any coordinate system.

Applying (12.6) to the special case in which f is one of the coordinate functions $x^j \colon U \to \mathbb{R}$, we obtain

$$dx^j\big|_p = \frac{\partial x^j}{\partial x^i}(p)\, \lambda^i\big|_p = \delta^j_i\, \lambda^i\big|_p = \lambda^j\big|_p.$$

In other words, the coordinate covector field λ^j is none other than the differential dx^j. Therefore, the formula (12.6) for df_p can be rewritten as

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

or as an equation between covector *fields* instead of covectors:

$$df = \frac{\partial f}{\partial x^i}\, dx^i. \tag{11.11}$$

From now on, we abandon the notation λ^i for the coordinate coframe, and use dx^i instead.

Example 12.8. If $f(x,y) = x^2 y \cos x$ on $\mathbb{R}^2$, then df is given by the formula

$$\begin{aligned} df &= \frac{\partial\left(x^2 y \cos x\right)}{\partial x}\, dx + \frac{\partial\left(x^2 y \cos x\right)}{\partial y}\, dy \\ &= \left(2xy\cos x - x^2 y \sin x\right) dx + x^2 \cos x\, dy. \end{aligned}$$

Recall that for a smooth real-valued function $f\colon M \to \mathbb{R}$, we now have two different definitions for the differential of f at a point $p \in M$. Previously, we defined df_p as a linear map from T_pM to $T_{f(p)}\mathbb{R}$.Now, we defined df_p as a covector at p, which is to say a linear map from T_pM to $\mathbb{R}$. These are really the same object, once we take into account the canonical identification between $\mathbb{R}$ and $T_{f(p)}\mathbb{R}$; one easy way to see this is to note that both are represented in coordinates by the row matrix whose components are the partial derivatives of f.

Proposition 12.9 (Properties of the Differential). *Let M be a smooth manifold and let $f, g \in C^\infty(M)$.*

(a) If a and b are constants, then $d(af + bg) = a\,df + b\,dg$.

(b) $d(fg) = f\,dg + g\,df$.

(c) $d(f/g) = (g\,df - f\,dg)/g^2$ on the set where $g \neq 0$.

(d) If f is constant, then $df = 0$.

Proposition 12.10 (Functions with Vanishing Differentials). *If f is a smooth real-valued function on a smooth manifold M then $df = 0$ if and only if f is constant on each component of M.*

Proof. It suffices to assume that M is connected and show that $df = 0$ if and only if f is constant. One direction is immediate: if f is constant, then $df = 0$. Conversely, suppose $df = 0$, let $p \in M$, and let $\mathcal{C} = \{q \in M : f(q) = f(p)\}$. If q is any point in $\mathcal{C}$, let U be a smooth coordinate ball centred at q. From (11.11) we see that $\frac{\partial f}{\partial x^i} \equiv 0$ in U for each i, so by elementary calculus f is constant on U. This shows that $\mathcal{C}$ is open, and since it is closed by continuity, it must be all of M. Thus, f is everywhere equal to the constant $f(p)$. □

12.4 Pullbacks of Covector Fields

As we have seen, a smooth map yields a linear map on tangent vectors called the differential. Dualizing this leads to a linear map on covectors going in the opposite direction.

Let $F\colon M \to N$ be a smooth map between smooth manifolds and let $p \in M$ be arbitrary. The differential $dF_p\colon T_pM \to T_{F(p)}N$ yields a dual linear map

$$dF_p^*\colon T_{F(p)}^*N \to T_p^*M,$$

called the **(pointwise) pullback by** F **at** p, or the **cotangent map of** F. Unraveling the definitions, we see that dF_p^* is characterized by

$$dF_p^*(\omega)(v) = \omega\big(dF_p(v)\big), \quad \text{for } \omega \in T_{F(p)}^*N,\ v \in T_pM.$$

Here comes the important part. Recall that when we discussed vector fields, we made a point of noting that pushforwards of vector fields under smooth maps are defined only in the special cases of diffeomorphisms or Lie group homomorphisms (recall the notion of being F-related). The surprising thing about covectors is that covector fields *always* pull back to covector fields.

Definition 12.11. Given a smooth map $F\colon M \to N$ and a covector field ω on N, define a covector field $F^*\omega$ on M, called the ***pullback of*** ω ***by*** F, by

$$(F^*\omega)_p = dF_p^*\left(\omega_{F(p)}\right). \tag{12.7}$$

It acts on a vector $v \in T_pM$ by

$$(F^*\omega)_p(v) = \omega_{F(p)}\big(dF_p(v)\big).$$

In contrast to the vector field case, there is no ambiguity here about what point to pull back from: the value of $F^*\omega$ at p is the pullback of ω at $F(p)$.

Let us prove two important properties of the pullback.

Proposition 12.12. *Let $F\colon M \to N$ be a smooth map between smooth manifolds. Suppose u is a continuous real-valued function on N, and ω is a covector field on N. Then*

$$F^*(u\omega) = (u \circ F)F^*\omega. \tag{12.8}$$

If in addition u is smooth, then

$$F^*du = d(u \circ F). \tag{12.9}$$

Proof. To prove (12.8) we compute

$$\begin{aligned}
\big(F^*(u\omega)\big)_p &= dF_p^*\big((u\omega)_{F(p)}\big) \\
&= dF_p^*\big(u(F(p))\omega_{F(p)}\big) \\
&= u\big(F(p)\big)dF_p^*\big(\omega_{F(p)}\big) && \text{(by linearity of } dF_p^*) \\
&= u\big(F(p)\big)(F^*\omega)_p && \text{(by definition)} \\
&= \big((u \circ F)F^*\omega\big)_p.
\end{aligned}$$

For (12.9), we let $v \in T_pM$ be arbitrary, and compute

$$\begin{aligned}
(F^*du)_p(v) &= \big(dF_p^*\big(du_{F(p)}\big)\big)(v) && \text{(by (12.7))} \\
&= du_{F(p)}\big(dF_p(v)\big) && \text{(by definition of } dF_p^*) \\
&= dF_p(v)u && \text{(by definition of } du) \\
&= v(u \circ F) && \text{(by definition of } dF_p) \\
&= d(u \circ F)_p(v) && \text{(by definition of } d(u \circ F)).
\end{aligned}$$

□

Proposition 12.13. *Suppose $F\colon M \to N$ is a smooth map between smooth manifolds and let ω be a covector field on N. Then $F^*\omega$ is a covector field on M. If ω is smooth, then so is $F^*\omega$.*

Proof. Let $p \in M$ be arbitrary, and choose smooth coordinates (y^j) for N in a neighborhood V of $F(p)$. Let $U = F^{-1}(V)$, which is a neighborhood of p. Writing ω in coordinates as $\omega = \omega_j\, dy^j$ for continuous functions ω_j on V and using Proposition 12.12 twice, we have the following computation in U:

$$F^*\omega = F^*\left(\omega_j\, dy^j\right) = (\omega_j \circ F)F^*\, dy^j = (\omega_j \circ F)d\left(y^j \circ F\right). \tag{12.10}$$

This expression is continuous, and is smooth if ω is smooth. □

Example 12.14. Let $F\colon \mathbb{R}^3 \to \mathbb{R}^2$ be the map given by

$$(u, v) = F(x, y, z) = \left(x^2 y,\ y \sin z\right),$$

and let $\omega \in \mathfrak{X}^*\left(\mathbb{R}^2\right)$ be the covector field

$$\omega = u\, dv + v\, du.$$

According to (12.10), the pullback $F^*\omega$ is given by

$$\begin{aligned}
F^*\omega &= (u \circ F)d(v \circ F) + (v \circ F)d(u \circ F) \\
&= \left(x^2 y\right)\, d(y \sin z) + (y \sin z)d\left(x^2 y\right) \\
&= x^2 y(\sin z\, dy + y \cos z\, dz) + y \sin z\left(2xy\, dx + x^2\, dy\right) \\
&= 2xy^2 \sin z\, dx + 2x^2 y \sin z\, dy + x^2 y^2 \cos z\, dz.
\end{aligned}$$

Suppose M is a smooth manifold, $S \subseteq M$ is an immersed submanifold and $\iota\colon S \hookrightarrow M$ is the inclusion map. If ω is any smooth covector field on M, the pullback by ι yields a smooth covector field $\iota^*\omega$ on S. To see what this means, let $v \in T_pS$ be arbitrary, and compute

$$(\iota^*\omega)_p(v) = \omega_p\big(d\iota_p(v)\big) = \omega_p(v),$$

since $d\iota_p\colon T_pS \to T_pM$ is just the inclusion map, under our usual identification of T_pS with a subspace of T_pM. Thus, $\iota^*\omega$ is just the restriction of ω to vectors tangent to S. For this reason, $\iota^*\omega$ is often called the ***restriction of*** ω ***to*** S.

12.5 Lie Derivatives of 1-forms

The Lie derivative operation can be extended to covector fields. Suppose M is a smooth manifold, V is a smooth vector field on M, and θ is its flow. For any $p \in M$, if t is sufficiently close to zero, then θ_t is a diffeomorphism from a neighborhood of p to a neighbourhood of $\theta_t(p)$, so $d(\theta_t)_p^*$ pulls back covectors at $\theta_t(p)$ to ones at p by the formula

$$d(\theta_t)_p^*(\omega_{\theta_t(p)})(v) = \omega_{\theta_t(p)}(d(\theta_t)_p(v_1)).$$

Note that $d(\theta_t)_p^*(\omega_{\theta_t(p)})$ is just the value of the pullback covector field $\theta_t^*\omega$ at p.

Definition 12.15. Given a smooth covector field ω on M, we define the ***Lie derivative of ω with respect to V***, denoted by $\mathcal{L}_V\omega$, by

$$(\mathcal{L}_V\omega)_p = \frac{d}{dt}\bigg|_{t=0}(\theta_t^*\omega)_p = \lim_{t\to 0}\frac{d(\theta_t)_p^*(\omega_{\theta_t(p)}) - \omega_p}{t}, \tag{12.11}$$

provided the derivative exists.

Just as for the case of vector fields we have

Proposition 12.16. *With M, V, and ω as above, the derivative in (12.11) exists for every $p \in M$ and defines $\mathcal{L}_V\omega$ as a smooth covector field on M.*

Lemma 12.17. *Let M be a smooth manifold and let $V \in \mathfrak{X}(M)$. Suppose f is a smooth real-valued function on M, and ω is a smooth covector field on M. Then*

(a) $\mathcal{L}_V(f\omega) = (\mathcal{L}_V f)\omega + f\mathcal{L}_V\omega$.

(b) If X is a smooth vector field then,

$$\mathcal{L}_V(\omega(X)) = (\mathcal{L}_V\omega)(X) + \omega(\mathcal{L}_V X) \tag{12.12}$$

One consequence of this proposition is the following formula expressing the Lie derivative of any smooth covector field in terms of Lie brackets and ordinary directional derivatives of functions, which allows us to compute Lie derivatives without first determining the flow.

Corollary 12.18. *We have the following.*

1. *If V is a smooth vector field and ω is a smooth covector field, then for any smooth vector fields X*

$$(\mathcal{L}_V\omega)(X) = V(\omega(X)) - \omega([V, X]). \tag{12.13}$$

2. *If $f \in C^\infty(M)$, then $\mathcal{L}_V(df) = d(\mathcal{L}_V f)$.*

Proof. We just prove (2). Using (12.13), for any $X \in \mathfrak{X}(M)$ we compute

$$\begin{aligned}(\mathcal{L}_V df)(X) &= V(df(X)) - df([V, X]) = VXf - [V, X]f \\ &= VXf - (VXf - XVf) = XVf \\ &= d(Vf)(X) = d(\mathcal{L}_V f)(X).\end{aligned}$$

□

13. Tensors

After having studied the concept of vector fields and covector fields in detail and understanding the techniques to study them, we now start with the study of tensors which provide a uniform way to combine the study of tangent and cotangent bundles. We begin with tensors on a vector space, which are multilinear generalizations of covectors; a covector is the special case of a tensor of rank one.

Many geometric structures which we will see in the study of differential geometry are tensors: for instance, Riemannian metrics, almost complex structures, symplectic structures, differential forms etc.

13.1 Multilinear Algebra

Recall that covectors are real-valued linear functions on a vector space. A tensor is a generalization of this in the sense that tensors are real-valued *multilinear* functions of one or more variables.

Suppose $V_1, \dots, V_k$, and W are vector spaces. A map

$$F\colon V_1 \times \cdots \times V_k \to W$$

is said to be *multilinear* if it is linear as a function of each variable for each i,

$$F(v_1, \dots, av_i + bv_i', \dots, v_k) = aF(v_1, \dots, v_i, \dots, v_k) + bF(v_1, \dots, v_i', \dots, v_k).$$

In this way, a multilinear function of one variable is just a linear function, and a multilinear function of two variables is generally called *bilinear*. Let us write $\mathrm{L}(V_1, \dots, V_k; W)$ for the set of all multilinear maps from $V_1 \times \cdots \times V_k$ to W. It is a vector space under the usual operations of pointwise addition and scalar multiplication:

$$(F + F')(v_1, \dots, v_k) = F(v_1, \dots, v_k) + F'(v_1, \dots, v_k),$$
$$(aF)(v_1, \dots, v_k) = a(F(v_1, \dots, v_k)).$$

For example, dot product in $\mathbb{R}^n$, the cross product in $\mathbb{R}^3$, the determinant are all examples of multilinear maps. Let's see a different but rather important example.

Example 13.1 (Tensor product of covectors)**.** Suppose V is a vector space, and $\omega, \eta \in V^*$. Define a function $\omega \otimes \eta\colon V \times V \to \mathbb{R}$ by

$$\omega \otimes \eta(v_1, v_2) = \omega(v_1)\eta(v_2),$$

where the product on the right is just ordinary multiplication of real numbers. The linearity of ω and η guarantees that $\omega \otimes \eta$ is a bilinear function of v_1 and v_2. If (e^1, e^2) denotes the standard dual basis for $(\mathbb{R}^2)^*$, then $e^1 \otimes e^2\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ is the bilinear function

$$e^1 \otimes e^2((w, x), (y, z)) = wz.$$

We can generalize the last example to any arbitrary real-valued multilinear functions as follows: let $V_1, \dots, V_k, W_1, \dots, W_l$ be real vector spaces, and suppose $F \in \mathrm{L}(V_1, \dots, V_k; \mathbb{R})$ and $G \in \mathrm{L}(W_1, \dots, W_l; \mathbb{R})$. Define a function

$$F \otimes G\colon V_1 \times \cdots \times V_k \times W_1 \times \cdots \times W_l \to \mathbb{R}$$

by

$$F \otimes G(v_1, \dots, v_k, w_1, \dots, w_l) = F(v_1, \dots, v_k)G(w_1, \dots, w_l). \tag{12.1}$$

It follows from the multilinearity of F and G that $F \otimes G(v_1, \dots, v_k, w_1, \dots, w_l)$ depends linearly on each argument v_i or w_j separately, so $F \otimes G$ is an element of $\mathrm{L}(V_1, \dots, V_k, W_1, \dots, W_l; \mathbb{R})$, called the **tensor product of** F **and** G. One can show that

$$(F \otimes G) \otimes H = F \otimes (G \otimes H).$$

Thus, we can write tensor products of three or more multilinear functions unambiguously without parentheses. If $F_1, \dots, F_l$ are multilinear functions depending on $k_1, \dots, k_l$ variables, respectively, their tensor product $F_1 \otimes \cdots \otimes F_l$ is a multilinear function of $k = k_1 + \cdots + k_l$ variables, whose action on k vectors is given by inserting the first k_1 vectors into F_1, the next k_2 vectors into F_2, and so forth, and multiplying the results together.

If $\omega^j \in V_j^*$ for $j = 1, \dots, k$, then $\omega^1 \otimes \cdots \otimes \omega^k \in \mathrm{L}(V_1, \dots, V_k; \mathbb{R})$ is the multilinear function given by

$$\omega^1 \otimes \cdots \otimes \omega^k(v_1, \dots, v_k) = \omega^1(v_1) \cdots \omega^k(v_k). \tag{13.1}$$

The next proposition shows that a basis for any space of multilinear functions can be formed by taking all possible tensor products of basis covectors.

Proposition 13.2 (A Basis for the Space of Multilinear Functions). *Let $V_1, \dots, V_k$ be real vector spaces of dimensions $n_1, \dots, n_k$, respectively. For each $j \in \{1, \dots, k\}$, let $(E_1^{(j)}, \dots, E_{n_j}^{(j)})$ be a basis for V_j, and let $(\varepsilon_{(j)}^1, \dots, \varepsilon_{(j)}^{n_j})$ be the corresponding dual basis for V_j^*. Then the set*

$$\mathcal{B} = \left\{ \varepsilon_{(1)}^{i_1} \otimes \cdots \otimes \varepsilon_{(k)}^{i_k} : 1 \le i_1 \le n_1, \dots, 1 \le i_k \le n_k \right\}$$

is a basis for $\mathrm{L}(V_1, \dots, V_k; \mathbb{R})$, which therefore has dimension equal to $n_1 \cdots n_k$.

Proof. We need to show that $\mathcal{B}$ is linearly independent and spans $\mathrm{L}(V_1, \dots, V_k; \mathbb{R})$. Suppose $F \in \mathrm{L}(V_1, \dots, V_k; \mathbb{R})$ is arbitrary. For each ordered k-tuple $(i_1, \dots, i_k)$ of integers with $1 \le i_j \le n_j$, define a number $F_{i_1 \dots i_k}$ by

$$F_{i_1 \dots i_k} = F\left(E_{i_1}^{(1)}, \dots, E_{i_k}^{(k)}\right). \tag{13.2}$$

If we show that

$$F = F_{i_1 \dots i_k} \varepsilon_{(1)}^{i_1} \otimes \cdots \otimes \varepsilon_{(k)}^{i_k}$$

the it will follow that $\mathcal{B}$ spans $\mathrm{L}(V_1, \dots, V_k; \mathbb{R})$. For any k-tuple of vectors $(v_1, \dots, v_k) \in V_1 \times \cdots \times V_k$, write $v_1 = v_1^{i_1} E_{i_1}^{(1)}, \dots, v_k = v_k^{i_k} E_{i_k}^{(k)}$, and compute

$$\begin{aligned} F_{i_1 \dots i_k} \varepsilon_{(1)}^{i_1} \otimes \cdots \otimes \varepsilon_{(k)}^{i_k}(v_1, \dots, v_k) &= F_{i_1 \dots i_k} \varepsilon_{(1)}^{i_1}(v_1) \cdots \varepsilon_{(k)}^{i_k}(v_k) \\ &= F_{i_1 \dots i_k} v_1^{i_1} \cdots v_k^{i_k}, \end{aligned}$$

while $F(v_1, \dots, v_k)$ is equal to the same thing by multilinearity. This proves the claim.

To show that $\mathcal{B}$ is linearly independent, suppose some linear combination equals zero:

$$F_{i_1 \dots i_k} \varepsilon_{(1)}^{i_1} \otimes \cdots \otimes \varepsilon_{(k)}^{i_k} = 0.$$

Apply this to any ordered k-tuple of basis vectors, $(E_{j_1}^{(1)}, \dots, E_{j_k}^{(k)})$. By the same computation as above, this implies that each coefficient $F_{j_1 \dots j_k}$ is zero. Thus, the only linear combination of elements of $\mathcal{B}$ that sums to zero is the trivial one. □

This proof shows, by the way, that the components $F_{i_1 \dots i_k}$ of a multilinear function F in terms of the basis elements in $\mathcal{B}$ are given by (13.2). Thus, F is completely determined by its action on all possible sequences of basis vectors.

13.2 Covariant and Contravariant Tensors

Definition 13.3. Let V be a finite-dimensional real vector space. If k is a positive integer, a *covariant k-tensor on V* is an element of the k-fold tensor product $V^* \otimes \cdots \otimes V^*$, which can also be thoght of as a real-valued multilinear function of k elements of V:

$$\alpha \colon \underbrace{V \times \cdots \times V}_{k \text{ copies}} \longrightarrow \mathbb{R}.$$

The number k is called the *rank of α*.

Thus, a covriant tensor eats k-vectors and spits out a real number. A 0-tensor is, by convention, just a real number. We denote the vector space of all covariant k-tensors on V by the notation

$$T^k(V^*) = \underbrace{V^* \otimes \cdots \otimes V^*}_{k \text{ copies}}.$$

For instance, a **covector** is a covariant tensor of rank 1. An inner product is a covariant tensor of rank 2.

Definition 13.4. For any finite-dimensional real vector space V, we define the space of *contravariant tensors on V of rank k* to be the vector space

$$T^k(V) = \underbrace{V \otimes \cdots \otimes V}_{k \text{ copies}}.$$

Elements of $T^k(V)$ can also be thought of as

$$T^k(V) \cong \left\{ \text{multilinear functions } \alpha \colon \underbrace{V^* \times \cdots \times V^*}_{k \text{ copies}} \to \mathbb{R} \right\}.$$

Combining the previous two definitions we have the following

Definition 13.5. For any nonnegative integers k, l, we define the space of *mixed tensors on V of type (k, l)* as

$$T^{(k,l)}(V) = \underbrace{V \otimes \cdots \otimes V}_{k \text{ copies}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{l \text{ copies}}.$$

Thus, in our notation, for instance,

$$T^{(0,0)}(V) = T^0(V^*) = T^0(V) = \mathbb{R},$$

$$T^{(0,1)}(V) = T^1(V^*) = V^*,$$
$$T^{(1,0)}(V) = T^1(V) = V,$$
$$T^{(0,k)}(V) = T^k(V^*),$$
$$T^{(k,0)}(V) = T^k(V).$$

Remark 13.6. The notation $T^{(k,l)}(V)$ is not universal. Another common notation is $T^k_l(V)$. Another notation reverse the roles of k and l so one must be careful in reading another source.

A corollary of Proposition 13.2 is the following statement on basis of the space of tensors.

Corollary 13.7. *Let V be an n-dimensional real vector space with a basis (E_i) and let (ε^j) is the dual basis for V^*. Then the following sets constitute bases for the tensor spaces over V:*

$$\{\varepsilon^{i_1} \otimes \cdots \otimes \varepsilon^{i_k} : 1 \leq i_1, \ldots, i_k \leq n\} \qquad \text{for } T^k(V^*);$$

$$\{E_{i_1} \otimes \cdots \otimes E_{i_k} : 1 \leq i_1, \ldots, i_k \leq n\} \qquad \text{for } T^k(V);$$

$$\{E_{i_1} \otimes \cdots \otimes E_{i_k} \otimes \varepsilon^{j_1} \otimes \cdots \otimes \varepsilon^{j_l} : 1 \leq i_1, \ldots, i_k, j_1, \ldots, j_l \leq n\} \qquad \text{for } T^{(k,l)}(V).$$

Therefore,

$$\dim T^k(V^*) = \dim T^k(V) = n^k \quad \text{and} \quad \dim T^{(k,l)}(V) = n^{k+l}.$$

Thus, given a basis of V and the corresponding dual basis, we can explicitly write the expression of any element $A \in T^k(V^*)$. We have

$$A = A_{i_1 \ldots i_k} \varepsilon^{i_1} \otimes \cdots \otimes \varepsilon^{i_k},$$

where the n^k coefficients $A_{i_1 \ldots i_k}$ are determined by

$$A_{i_1 \ldots i_k} = A(E_{i_1}, \ldots, E_{i_k}).$$

13.3 Tensor Fields on Manifolds

Let M be a smooth manifold. We mimic our descriptions of tangent bundle and cotangent bundle to first define the **bundle of covariant k-tensors on M** by

$$T^k T^* M = \coprod_{p \in M} T^k(T_p^* M).$$

Analogously, we define the **bundle of contravariant k-tensors** by

$$T^k T M = \coprod_{p \in M} T^k(T_p M),$$

and the **bundle of mixed tensors of type (k, l)** by

$$T^{(k,l)} T M = \coprod_{p \in M} T^{(k,l)}(T_p M).$$

Immediately, we see, for instance that

$$T^{(0,0)} T M = T^0 T^* M = T^0 T M = M \times \mathbb{R},$$

$$T^{(0,1)} T M = T^1 T^* M = T^* M,$$
$$T^{(1,0)} T M = T^1 T M = T M,$$
$$T^{(0,k)} T M = T^k T^* M,$$
$$T^{(k,0)} T M = T^k T M.$$

Just like the case with the tangent and cotangent bundle, one can prove that the above space, also called **tensor bundle over M** are smooth manifolds with a smooth structure inherited from M and they also come equipped with a projection map $\pi : T^{(k,l)}(M) \to M$. Thus, we can also mimic the description of vector fields and covector fields to give the following definition.

Definition 13.8. A (k,l)-tensor field α on a smooth manifold M is a continuous map

$$\alpha : M \to T^{(k,l)}(M)$$

such that

$$\pi \circ \alpha = id_M.$$

The tensor field α is smooth if it is a smooth map between smooth manifolds.

We see that contravariant 1-tensor fields are the same as vector fields, and covariant 1-tensor fields are covector fields, a 0-tensor field is the same as a continuous real-valued function.
Notation: We will denote the space of all (k,l)-tensor fields by $\Gamma(T^{(k,l)}TM)$. These are infinite-dimensional vector spaces over $\mathbb{R}$, and are modules over$C^\infty(M)$. In any smooth local coordinates (x^i), elements of these bundles can be written (using the summation convention) as

$$A = \begin{cases} A_{i_1\dots i_k}\, dx^{i_1} \otimes \cdots \otimes dx^{i_k}, & A \in \Gamma(T^k T^* M); \\ A^{i_1\dots i_k} \dfrac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \dfrac{\partial}{\partial x^{i_k}}, & A \in \Gamma(T^k TM); \\ A^{i_1\dots i_k}_{j_1\dots j_l} \dfrac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \dfrac{\partial}{\partial x^{i_k}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_l}, & A \in \Gamma(T^{(k,l)}TM). \end{cases}$$

The functions $A_{i_1\dots i_k}$, $A^{i_1\dots i_k}$, or $A^{i_1\dots i_k}_{j_1\dots j_l}$ are called the **component functions of** A in the chosen coordinates. For brevity, we will denote the space of all smooth covariant k-tensor fields by

$$\mathcal{T}^k(M) = \Gamma(T^k T^* M).$$

The next proposition is analogous to Proposition 9.8 which gave the smoothness criteria for vector fields or Proposition 12.5 for covector fields. We only state this for covariant tensors but the result is similar for other types of tensors.

Proposition 13.9. *Let M be a smooth manifold and let $A\colon M \to T^k T^* M$ be a tensor field. The following are equivalent.*

(1) A is smooth.

(2) In every smooth coordinate chart, the component functions of A are smooth.

(3) If $X_1, \dots, X_k \in \mathfrak{X}(M)$, then the function $A(X_1, \dots, X_k)\colon M \to \mathbb{R}$, defined by

$$A(X_1, \dots, X_k)(p) = A_p(X_1|_p, \dots, X_k|_p),$$

is smooth.

(4) Whenever $X_1, \dots, X_k$ are smooth vector fields defined on some open subset $U \subseteq M$, the function $A(X_1, \dots, X_k)$ is smooth on U.

Thus, we again relegated the smoothness for a tensor field to either looking at its component functions or cooking up a function out of A by feeding k vector fields to it and then looking at the smoothness of the resultant function. Similar to the case of vector fields, we have the following

Proposition 13.10. *Suppose M is a smooth manifold, $A \in \mathcal{T}^k(M)$, $B \in \mathcal{T}^l(M)$, and $f \in C^\infty(M)$. Then fA and $A \otimes B$ are also smooth tensor fields, whose components in any smooth local coordinate chart are*

$$(fA)_{i_1\dots i_k} = f A_{i_1\dots i_k},$$
$$(A \otimes B)_{i_1\dots i_{k+l}} = A_{i_1\dots i_k} B_{i_{k+1}\dots i_{k+l}}.$$

Now a question arises: if we see a map which takes k vector fields as its arguments and gives us a smooth function, how can we guarantee that it is indeed a smooth covariant tensor field of rank k? First of all, we notice that any rank k-covariant tensor A induces a map

$$\underbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}_{k \text{ copies}} \longrightarrow C^\infty(M).$$

which is multilinear over $\mathbb{R}$. But, in fact, more is true: it is *multilinear over* $C^\infty(M)$, which means that for $f, f' \in C^\infty(M)$ and $X_i, X_i' \in \mathfrak{X}(M)$, we have

$$A(X_1, \ldots, fX_i + f'X_i', \ldots, X_k) = fA(X_1, \ldots, X_i, \ldots, X_k) + f'A(X_1, \ldots, X_i', \ldots, X_k).$$

This property turns out to be characteristic of smooth tensor fields, as the next lemma shows, which we state without proof.

Lemma 13.11 (Tensor Characterization Lemma)**.** *A map*

$$\mathcal{A}\colon \underbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}_{k \text{ copies}} \longrightarrow C^\infty(M), \tag{13.3}$$

is induced by a smooth covariant k-tensor field as above if and only if it is multilinear over $C^\infty(M)$.

A **symmetric tensor field** on a manifold is a covariant tensor field whose value at each point is a symmetric tensor, i.e., the value of the tensor doesn't change if we change the position of the vectors in its argument. The symmetric product of two or more tensor fields is defined pointwise, just like the tensor product. Thus, for example, if A and B are smooth covector fields, their symmetric product is the smooth 2-tensor field AB, which is given by

$$AB = \frac{1}{2}(A \otimes B + B \otimes A).$$

Let us mention two more examples of tensors.

Example 13.12 (Riemannian metrics). Let M be a manifold. A Riemannian metric g on M associates to every point $p \in M$ an inner product g_p on T_pM and thus we get a bilinear map

$$g_p : T_p \times T_pM \to \mathbb{R}$$

which is also symmetric and positive-definite. A Riemannian metric is called smooth if the function $M \to \mathbb{R} : p \mapsto g(X(p), Y(p))$ is smooth for any pair of vector fields $X, Y \in \mathfrak{X}(M)$. Thus, g is an example of a $(0, 2)$-tensor which is symmetric.

Example 13.13 (Almost complex structure). Suppose V is a $2n$-dimensional real vector space. We can make it into an n-dimensional complex vector space if we have a map $J : V \to V$ such that $J^2 = -id$ and then defining $(a + ib)v = av + bJV$. Such a J is called a **complex structure** on V.

An **almost complex structure** on M is a map $J : \Gamma(TM) \to \Gamma(TM)$ such that it a complex structure on T_pM for all $p \in M$. Thus an almost complex structure is an example of a $(1, 1)$-tensor field.

13.4 Pullbacks of Tensor Fields

Just like covector fields, covariant tensor fields can be pulled back by a smooth map to yield tensor fields on the domain. This construction works only for covariant tensor fields, which is one reason why we focus most of our attention on the covariant case. The reason is the same as to why the pull-back of a covector field is always defined.

Definition 13.14. Suppose $F\colon M \to N$ is a smooth map. For any point $p \in M$ and any k-tensor $\alpha \in T^k(T^*_{F(p)}N)$, we define a tensor $dF^*_p(\alpha) \in T^k(T^*_pM)$, called the **pointwise pullback of** α **by** F **at** p, by

$$dF^*_p(\alpha)(v_1, \dots, v_k) = \alpha(dF_p(v_1), \dots, dF_p(v_k)), \qquad v_1, \dots, v_k \in T_pM. \tag{13.4}$$

If A is a covariant k-tensor field on N, we define a k-tensor field F^*A on M, called the **pullback of** A **by** F, by

$$(F^*A)_p = dF^*_p(A_{F(p)}),$$

which acts on vectors $v_1, \dots, v_k \in T_pM$ by

$$(F^*A)_p(v_1, \dots, v_k) = A_{F(p)}(dF_p(v_1), \dots, dF_p(v_k)). \tag{13.5}$$

In precisely, the same way as we proved the properties about pullback of vector fields and covector fields, we get the following proposition which collects together some basic properties of pullback of tensor fields.

Proposition 13.15. Suppose $F\colon M \to N$ and $G\colon N \to P$ are smooth maps, A and B are covariant tensor fields on N, C a tensor field on P and $f \in C^\infty(N)$.

(a) $F^*(fA) = (f \circ F)F^*A$.

(b) $F^*(A \otimes B) = F^*A \otimes F^*B$.

(c) $F^*(A + B) = F^*A + F^*B$.

(d) F^*B is a (continuous) tensor field, and is smooth if B is smooth.

(e) $(G \circ F)^*C = F^*(G^*C)$.

(f) $(\mathrm{Id}_N)^*B = B$.

This immediately leads to the following corollary which provides local coordinates description of the pullback of a tensor field and should be compared with equation (12.10) for covector fields.

Corollary 13.16. *Let $F\colon M \to N$ be smooth, and let B be a covariant k-tensor field on N. If $p \in M$ and (y^i) are smooth coordinates for N on a neighborhood of $F(p)$, then F^*B has the following expression in a neighborhood of p:*

$$F^*\left(B_{i_1 \dots i_k}\, dy^{i_1} \otimes \cdots \otimes dy^{i_k}\right) = (B_{i_1 \dots i_k} \circ F)\, d(y^{i_1} \circ F) \otimes \cdots \otimes d(y^{i_k} \circ F). \tag{13.6}$$

In general, there is neither a pushforward nor a pullback operation for mixed tensor fields. However, in the special case of a diffeomorphism, tensor fields of any variance can be pushed forward and pulled back. We just state the result and important properties below.

Proposition 13.17. *Let $F : M \to N$ be a diffeomorphism between smooth manifold, Then*

$$F_* : \Gamma(T^{(k,l)}TM) \to \Gamma(T^{(k,l)}TN), \quad F^* : \Gamma(T^{(k,l)}TN) \to \Gamma(T^{(k,l)}TM),$$

generalize the usual pushforwards and pullbacks of vector fields and covector fields respectively such that

1. $F_* = (F^*)^{-1}$.
2. $F^*(A \otimes B) = F^*(A) \otimes F^*(B)$.
3. $(F \circ G)_* = F_* \circ G_*$ *and* $(F \circ G)^* = G^* \circ F^*$.
4. *For any covariant k-tensor A on N,* $F^*(A(X_1, \ldots, X_k)) = F^*A(F_*^{-1}(X_1), \ldots, F_*^{-1}(X_k))$.

13.5 Lie Derivatives of Tensor Fields

Once we know what pullback of tensor fields are, we can mimic the idea about taking Lie derivatives of covector fields in the direction of a vector field and extend it to define the Lie derivative of an arbitrary tensor field in the direction of a vector field.

Suppose M is a smooth manifold, V is a smooth vector field on M, and θ is its flow. For any $p \in M$, if t is sufficiently close to zero, then θ_t is a diffeomorphism from a neighborhood of p to a neighbourhood of $\theta_t(p)$, so $d(\theta_t)_p^*$ pulls back tensors at $\theta_t(p)$ to ones at p by the formula

$$d(\theta_t)_p^*(A_{\theta_t(p)})(v_1, \ldots, v_k) = A_{\theta_t(p)}(d(\theta_t)_p(v_1), \ldots, d(\theta_t)_p(v_k)).$$

where $d(\theta_t)_p^*(A_{\theta_t(p)})$ is just the value of the pullback tensor field $\theta_t^* A$ at p.

Definition 13.18. Given a smooth covariant tensor field A on M, we define the **Lie derivative of A with respect to V**, denoted by $\mathcal{L}_V A$, by

$$(\mathcal{L}_V A)_p = \frac{d}{dt}\bigg|_{t=0} (\theta_t^* A)_p = \lim_{t \to 0} \frac{d(\theta_t)_p^*(A_{\theta_t(p)}) - A_p}{t}, \tag{13.7}$$

provided the derivative exists. The quantity $\mathcal{L}_V A$ is a smooth tensor field on M and is a tensor of the same type as A.

Proposition 13.19. *Let M be a smooth manifold and let $V \in \mathfrak{X}(M)$. Suppose $f \in C^\infty(M)$, and A, B are smooth covariant tensor fields on M. Then*

(a) $\mathcal{L}_V(fA) = (\mathcal{L}_V f)A + f\mathcal{L}_V A$.

(b) $\mathcal{L}_V(A \otimes B) = (\mathcal{L}_V A) \otimes B + A \otimes \mathcal{L}_V B$.

(c) If $X_1, \ldots, X_k$ are smooth vector fields and A is a smooth k-tensor field,

$$\mathcal{L}_V(A(X_1, \ldots, X_k)) = (\mathcal{L}_V A)(X_1, \ldots, X_k) + A(\mathcal{L}_V X_1, \ldots, X_k) + \cdots + A(X_1, \ldots, \mathcal{L}_V X_k). \tag{13.8}$$

(d) *We also have*

$$\begin{aligned}(\mathcal{L}_V A)(X_1, \ldots, X_k) = V(A(X_1, \ldots, X_k)) - A([V, X_1], X_2, \ldots, X_k) - \cdots \\ \cdots - A(X_1, \ldots, X_{k-1}, [V, X_k]).\end{aligned} \tag{13.9}$$

Proof. Erercise, which can be solved in the same way as that for vector fields. □

Let us see an example of how we can compute the Lie derivative of a $(0,2)$-tensor in local coordinates.

Example 13.20. Suppose A is an arbitrary smooth covariant 2-tensor field, and V is a smooth vector field. We compute the Lie derivative $\mathcal{L}_V A$ in smooth local coordinates (x^i). First, we observe that $\mathcal{L}_V dx^i = d(\mathcal{L}_V x^i) = d(Vx^i) = dV^i$. Therefore,

$$\begin{aligned}\mathcal{L}_V A &= \mathcal{L}_V (A_{ij} dx^i \otimes dx^j) \\ &= \mathcal{L}_V (A_{ij}) dx^i \otimes dx^j + A_{ij}(\mathcal{L}_V dx^i) \otimes dx^j + A_{ij} dx^i \otimes (\mathcal{L}_V dx^j) \\ &= VA_{ij}\, dx^i \otimes dx^j + A_{ij}\, dV^i \otimes dx^j + A_{ij}\, dx^i \otimes dV^j \\ &= \left(VA_{ij} + A_{kj}\frac{\partial V^k}{\partial x^i} + A_{ik}\frac{\partial V^k}{\partial x^j}\right) dx^i \otimes dx^j.\end{aligned}$$

Thus, once we know the expression of the vector field V in local coordinates and the expression of the tensor A in local coordinates, we can explicitly write down the expression for $\mathcal{L}_V A$ in local coordinates in terms of the derivatives of component functions of V and A. Similar formulas hold for higher rank tensors.

Recall that the Lie derivative of a vector field W with respect to V is zero if and only if W is invariant under the flow of V which was the content of Proposition 11.7. We can say a similar thing for tensors. If A is a smooth tensor field on M and θ is a flow on M, we say that A **is invariant under** θ if for each t, the map θ_t pulls A back to itself wherever it is defined, that is,

$$d(\theta_t)_p^*(A_{\theta_t(p)}) = A_p \tag{12.11}$$

for all (t,p) in the domain of θ. If θ is a global flow, this is equivalent to $\theta_t^* A = A$ for all $t \in \mathbb{R}$.

The following result is an analog of Proposition 11.5 which shows how the Lie derivative can be used to compute time derivatives at times other than $t = 0$.

Proposition 13.21. *Suppose M is a smooth manifold $V \in \mathfrak{X}(M)$. Let θ be the flow of V. For any smooth covariant tensor field A and any (t_0, p) in the domain of θ,*

$$\left.\frac{d}{dt}\right|_{t=t_0} (\theta_t^* A)_p = (\theta_{t_0}^*(\mathcal{L}_V A))_p. \tag{13.10}$$

Proof. Use the definition of the pullback of tensors to notice that (13.10) is equivalent to proving that have to prove

$$\left.\frac{d}{dt}\right|_{t=t_0} d(\theta_t)_p^*(A_{\theta_t(p)}) = d(\theta_{t_0})_p^*((\mathcal{L}_V A)_{\theta_{t_0}(p)}).$$

We mimic the proof of Proposition 11.5 and do a change of variables $t = s + t_0$ to get

$$\begin{aligned}\left.\frac{d}{dt}\right|_{t=t_0} d(\theta_t)_p^*(A_{\theta_t(p)}) &= \left.\frac{d}{ds}\right|_{s=0} d(\theta_{s+t_0})_p^*(A_{\theta_{s+t_0}(p)}) \\ &= \left.\frac{d}{ds}\right|_{s=0} d(\theta_{t_0})_p^* d(\theta_s)_{\theta_{t_0}(p)}^*(A_{\theta_s(\theta_{t_0}(p))}) \\ &= d(\theta_{t_0})_p^* \left.\frac{d}{ds}\right|_{s=0} d(\theta_s)_{\theta_{t_0}(p)}^*(A_{\theta_s(\theta_{t_0}(p))}) \\ &= d(\theta_{t_0})_p^*((\mathcal{L}_V A)_{\theta_{t_0}(p)}).\end{aligned}$$

□

Thus, we immediately get the analog of Proposition 11.7 for tensors.

Theorem 13.22. *Let M be a smooth manifold and let $V \in \mathfrak{X}(M)$. A smooth covariant tensor field A is invariant under the flow of V if and only if $\mathcal{L}_V A = 0$.* □

14. Differential forms

We learned about tensors in the previous chapter. Among all tensors, there are specail types of tensors called symmetric and skew-symmetric tensors. In this chapter, we will learn about the latter. But first, we define both symmetric and skew-symmetric tensors.

Definition 14.1. A rank k-covariant tensor α on a finite-dimensional real vector space V is called **symmetric** if

$$\alpha(v_1, ..., v_i, ..., v_j, ..., v_k) = \alpha(v_1, ..., v_j, ..., v_i, ..., v_k) \tag{14.1}$$

for all $1 \leq i < j \leq k$.

In fact, (14.1) implies that the value of α is unchanged when the vectors $v_1, ..., v_k$ are rearranged in any order. In terms of the component functions, this means that $\alpha_{i_1 \ldots i_k}$ are unchanged by any permutation of indices.

The set of symmetric covariant k-tensors $\Sigma^k(V^*)$ is a linear subspace of the space $T^k(V^*)$ of all covariant k-tensors on V. Given any tensor α we can make it into a symmetric tensor as follows. Let S_k denote the symmetric group on k elements. Given a k-tensor α and a permutation $\sigma \in S_k$, we define a new k-tensor ${}^\sigma\alpha$ by

$${}^\sigma\alpha(v_1, \ldots, v_k) = \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)}).$$

Note that ${}^\tau({}^\sigma\alpha) = {}^{\tau\sigma}\alpha$, where $\tau\sigma$ represents the composition of τ and σ, that is, $\tau\sigma(i) = \tau(\sigma(i))$.

We define a projection $\mathrm{Sym}\colon T^k(V^*) \to \Sigma^k(V^*)$ called **symmetrization** by

$$\mathrm{Sym}\,\alpha = \frac{1}{k!}\sum_{\sigma\in S_k} {}^\sigma\alpha = \frac{1}{k!}\sum_{\sigma\in S_k} \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)}).$$

Clearly $\mathrm{Sym}\,\alpha$ is a symmetric tensor.

If α and β are symmetric tensors on V, then $\alpha \otimes \beta$ is not symmetric in general. For example, consider α, β as covectors. Individually, both are symmetric, but

$$\alpha \otimes \beta(v_1, v_2) = \alpha(v_1)\beta(v_2) \neq \alpha \otimes \beta(v_2, v_1).$$

But we can define the symmetrization of $\alpha \otimes \beta$ as above. If $\alpha \in \Sigma^k(V^*)$ and $\beta \in \Sigma^l(V^*)$, we define their ***symmetric product*** to be the $(k+l)$-tensor $\alpha\beta$ by

$$\alpha\beta = \mathrm{Sym}(\alpha \otimes \beta).$$

which acts on vectors $v_1, \ldots, v_{k+l}$ is given by

$$\alpha\beta(v_1, \ldots, v_{k+l}) = \frac{1}{(k+l)!}\sum_{\sigma\in S_{k+l}} \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)})\beta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)}).$$

Thus, for example, if α and β are covectors, then

$$\alpha\beta = \frac{1}{2}(\alpha \otimes \beta + \beta \otimes \alpha).$$

The same analysis on each tangent space gives the notion of symmetric tensors on manifolds. We already saw an example before: Riemannian metric on a manifold. We will learn more about this later.

14.1 Alternating Tensors

Let us first discuss various notions related to alternating tensors on a vector space and then we will just translate all the knowledge to manifolds. We have the following definition.

Definition 14.2. A covariant k-tensor α on V is said to be **alternating** or **antisymmetric** if it changes sign whenever two of its arguments are interchanged, that is, for all vectors $v_1, \ldots, v_k \in V$ and every pair of distinct indices i, j we have satisfies

$$\alpha(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -\alpha(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k). \tag{14.2}$$

Alternating covariant k-tensors are also called **exterior forms**.

Recall that for any permutation $\sigma \in S_k$, the **sign** of σ, denoted by $\operatorname{sgn} \sigma$, is equal to $+1$ if σ is even and -1 if σ is odd. Thus, (14.2) implies that if α is alternating then for any vectors $v_1, \ldots, v_k$ and any permutation $\sigma \in S_k$,

$$\alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = (\operatorname{sgn} \sigma)\alpha(v_1, \ldots, v_k),$$

and the components $\alpha_{i_1 \ldots i_k}$ of α change sign whenever two indices are interchanged.
Notation. The vector space of all alternating k-tensors on V is denoted by $\Lambda^k(V^*)$. Clearly all 0-tensors and 1-tensors are alternating.

Proposition 14.3. Let α be a covariant k-tensor on a finite-dimensional vector space V. The following are equivalent:

(a) α *is alternating.*

(b) $\alpha(v_1, \ldots, v_k) = 0$ *whenever the* k*-tuple* $(v_1, \ldots, v_k)$ *is linearly dependent.*

(c) α *gives the value zero whenever two of its arguments are equal:*

$$\alpha(v_1, \ldots, w, \ldots, w, \ldots, v_k) = 0.$$

Proof. The implications (a) $\Rightarrow$ (c) and (b) $\Rightarrow$ (c) are immediate. We complete the proof by showing that (c) implies both (a) and (b).

Assume that α satisfies (c). For any vectors $v_1, \ldots, v_k$, the hypothesis implies

$$\begin{aligned} 0 &= \alpha(v_1, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_k) \\ &= \alpha(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_k) + \alpha(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) \\ &\quad + \alpha(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k) + \alpha(v_1, \ldots, v_j, \ldots, v_j, \ldots, v_k) \\ &= \alpha(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) + \alpha(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k). \end{aligned}$$

Thus α is alternating. On the other hand, if $(v_1, \ldots, v_k)$ is a linearly dependent k-tuple, then one of the v_i's can be written as a linear combination of the others. For simplicity, let us assume that

$$v_k = \sum_{j=1}^{k-1} a^j v_j.$$

Then multilinearity of α implies

$$\alpha(v_1, \ldots, v_k) = \sum_{j=1}^{k-1} a^j \alpha(v_1, \ldots, v_{k-1}, v_j).$$

In each of these terms, α has two identical arguments, so every term is zero. $\square$

We now define the analogue of the symmetrization operator for obtaining alternating tensors.

We define the projection $\mathrm{Alt}\colon T^k(V^*) \to \Lambda^k(V^*)$, called **alternation** or **skew-symmetrization**, as:

$$\mathrm{Alt}\,\alpha = \frac{1}{k!}\sum_{\sigma\in S_k}(\operatorname{sgn}\sigma)({}^\sigma\alpha),$$

where S_k is the symmetric group on k elements. More explicitly, this means

$$(\mathrm{Alt}\,\alpha)(v_1,\dots,v_k) = \frac{1}{k!}\sum_{\sigma\in S_k}(\operatorname{sgn}\sigma)\alpha(v_{\sigma(1)},\dots,v_{\sigma(k)}).$$

Thus, for instance, if β is a 2-tensor, then

$$(\mathrm{Alt}\,\beta)(v,w) = \frac{1}{2}(\beta(v,w)-\beta(w,v)).$$

For a 3-tensor γ,

$$(\mathrm{Alt}\,\gamma)(v_1,v_2,v_3) = \frac{1}{6}\big(\gamma(v_1,v_2,v_3)+\gamma(v_2,v_3,v_1)+\gamma(v_3,v_1,v_2)-\gamma(v_2,v_1,v_3)-\gamma(v_1,v_3,v_2)-\gamma(v_3,v_2,v_1)\big).$$

We now try to produce an explicit basis of the space $\Lambda^k(V*)$ by introducing the concept of elementary exterior forms.

Notation. For $k\in\mathbb{Z}$, an ordered k-tuple $I=(i_1,\dots,i_k)$ of positive integers is called a **multi-index of length** k. If I is such a multi-index and $\sigma\in S_k$, we write I_σ for the multi-index $I_\sigma=(i_{\sigma(1)},\dots,i_{\sigma(k)})$. Note that $I_{\sigma\tau}=(I_\sigma)_\tau$ for $\sigma,\tau\in S_k$.

Let V be an n-dimensional vector space, and suppose $(\varepsilon^1,\dots,\varepsilon^n)$ is any basis for V^*. We now define a collection of k-exterior forms on V that generalize the determinant function on $\mathbb{R}^n$. For each multi-index $I=(i_1,\dots,i_k)$ of length k such that $1\le i_1,\dots,i_k\le n$, define a covariant k-tensor $\varepsilon^I=\varepsilon^{i_1\dots i_k}$ by

$$\varepsilon^I(v_1,\dots,v_k)=\det\begin{pmatrix}\varepsilon^{i_1}(v_1) & \cdots & \varepsilon^{i_1}(v_k)\\ \vdots & \ddots & \vdots\\ \varepsilon^{i_k}(v_1) & \cdots & \varepsilon^{i_k}(v_k)\end{pmatrix}=\det\begin{pmatrix}v_1^{i_1} & \cdots & v_k^{i_1}\\ \vdots & \ddots & \vdots\\ v_1^{i_k} & \cdots & v_k^{i_k}\end{pmatrix}. \tag{14.3}$$

Since the determinant changes sign whenever two columns are interchanged, it is clear that ε^I is an alternating k-tensor. We call ε^I an **elementary alternating tensor**.

Example 14.5. In terms of the standard dual basis (e^1,e^2,e^3) for $(\mathbb{R}^3)^*$, we have

$$e^{13}(v,w)=v^1w^3-w^1v^3.$$

Proposition 14.4 (Properties of Elementary k-exterior forms)**.** Let (E_i) be a basis for V, let (ε^i) be the dual basis for V^*, and let ε^I be as defined above.

1. If I has a repeated index, then $\varepsilon^I=0$.
2. If $J=I_\sigma$ for some $\sigma\in S_k$, then $\varepsilon^I=(\operatorname{sgn}\sigma)\varepsilon^J$.
3. The result of evaluating ε^I on a sequence of basis vectors is

$$\varepsilon^I(E_{j_1},\dots,E_{j_k})=\det\begin{pmatrix}\delta^{i_1}_{j_1} & \cdots & \delta^{i_1}_{j_k}\\ \vdots & \ddots & \vdots\\ \delta^{i_k}_{j_1} & \cdots & \delta^{i_k}_{j_k}\end{pmatrix}.$$

Proof. If I has a repeated index, then for any vectors $v_1, \dots, v_k$, the determinant in (14.3) has two identical rows and thus is equal to zero, which proves (a). On the other hand, if J is obtained from I by interchanging two indices, then the corresponding determinants have opposite signs; this implies (b). Finally, (c) follows immediately from the definition of ε^I. $\square$

We can now use the elementary k-exterior forms to write a basis for $\Lambda^k(V^*)$. A multi-index $I = (i_1, \dots, i_k)$ is said to be **increasing** if $i_1 < \dots < i_k$.

Proposition 14.5 (A basis for $\Lambda^k(V^*)$)**.** Let V be an n-dimensional vector space. If (ε^i) is any basis for V^*, then for each positive integer $k \leq n$, the collection of k-exterior forms

$$\mathcal{E} = \{\varepsilon^I : I \text{ is an increasing multi-index of length } k\}$$

is a basis for $\Lambda^k(V^*)$. Therefore,

$$\dim \Lambda^k(V^*) = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

If $k > n$, then $\dim \Lambda^k(V^*) = 0$.

Proof. Note that $\Lambda^k(V^*)$ is the trivial vector space when $k > n$ follows immediately from Proposition 13.3 (b), since every k-tuple of vectors is linearly dependent in that case. For the case $k \leq n$, we need to show that the set $\mathcal{E}$ spans $\Lambda^k(V^*)$ and is linearly independent. Let (E_i) be the basis for V dual to (ε^i).

To show that $\mathcal{E}$ spans $\Lambda^k(V^*)$, let $\alpha \in \Lambda^k(V^*)$ be arbitrary. For each multi-index $I = (i_1, \dots, i_k)$ (not necessarily increasing), define a real number α_I by

$$\alpha_I = \alpha(E_{i_1}, \dots, E_{i_k}).$$

Since α is alternating hence $\alpha_I = 0$ if I contains a repeated index, and $\alpha_J = (\operatorname{sgn}\sigma)\alpha_I$ if $J = I_\sigma$ for $\sigma \in S_k$. Thus, for any multi-index J, we have,

$$\sum_{I-\text{increasing}} \alpha_I \varepsilon^I(E_{j_1}, \dots, E_{j_k}) = \alpha_J = \alpha(E_{j_1}, \dots, E_{j_k}).$$

Thus $\sum_{I-\text{increasing}} \alpha_I \varepsilon^I = \alpha$, so $\mathcal{E}$ spans $\Lambda^k(V^*)$. The proof of linear independence is an exercise. $\square$

14.2 The Wedge Product

We now see an operation which allows us to combine two or more exterior forms to again give an exterior form. We still work on a real finite dimensional vector space V. As we see below, the operation of wedge product is just the alternation of tensor product with appropriately chosen normalization constant.

Definition 14.6. Given $\omega \in \Lambda^k(V^*)$ and $\eta \in \Lambda^l(V^*)$, we define their **wedge product** or **exterior product** to be the following $(k+l)$-exterior form:

$$\omega \wedge \eta = \frac{(k+l)!}{k!l!} \operatorname{Alt}(\omega \otimes \eta). \tag{14.4}$$

One of the reasons we choose this normalization constant in (14.4) is because of the following relation for elementary alternating tensors. One can check that for any multi-indices $I = (i_1, \dots, i_k)$ and $J = (j_1, \dots, j_l)$,

$$\varepsilon^I \wedge \varepsilon^J = \varepsilon^{IJ},$$

where $IJ = (i_1, \dots, i_k, j_1, \dots, j_l)$ is obtained by concatenating I and J.

We now see some important properties of the wedge product.

Proposition 14.7. Suppose $\omega, \omega', \eta, \eta'$, and ξ are exterior forms on a finite-dimensional vector space V.

1. ***Bilinearity:*** For $a, a' \in \mathbb{R}$,
$$(a\omega + a'\omega') \wedge \eta = a(\omega \wedge \eta) + a'(\omega' \wedge \eta),$$
$$\eta \wedge (a\omega + a'\omega') = a(\eta \wedge \omega) + a'(\eta \wedge \omega').$$
2. ***Associativity:***
$$\omega \wedge (\eta \wedge \xi) = (\omega \wedge \eta) \wedge \xi.$$
3. ***Anticommutativity:*** For $\omega \in \Lambda^k(V^*)$ and $\eta \in \Lambda^l(V^*)$,
$$\omega \wedge \eta = (-1)^{kl} \eta \wedge \omega. \tag{14.5}$$
4. If (ε^i) is any basis for V^* and $I = (i_1, \ldots, i_k)$ is any multi-index, then
$$\varepsilon^{i_1} \wedge \cdots \wedge \varepsilon^{i_k} = \varepsilon^I. \tag{14.6}$$
5. For any covectors $\omega^1, \ldots, \omega^k$ and vectors $v_1, \ldots, v_k$,
$$\omega^1 \wedge \cdots \wedge \omega^k(v_1, \ldots, v_k) = \det(\omega^j(v_i)). \tag{14.7}$$

Proof. Since the wedge product is the alternation of tensor product and the tensor product is a bilinear operation, we see bilinearity of the wedge product. We prove 2. and 3. on elementary alternating tensors from which the general case will follows. Note that

$$(\varepsilon^I \wedge \varepsilon^J) \wedge \varepsilon^K = \varepsilon^{IJ} \wedge \varepsilon^K = \varepsilon^{IJK} = \varepsilon^I \wedge \varepsilon^{JK} = \varepsilon^I \wedge (\varepsilon^J \wedge \varepsilon^K),$$

and hence the associativity of wedge product follows. For part 3., we again observe that from the property of being an exterior form, we get

$$\varepsilon^I \wedge \varepsilon^J = \varepsilon^{IJ} = (\operatorname{sgn} \sigma)\varepsilon^{JI} = (\operatorname{sgn} \sigma)\varepsilon^J \wedge \varepsilon^I,$$

where σ is the permutation that sends IJ to JI. Now the main point is that once you write both I and J as product of 2-cycles, moving J to front and I to back requires moving all the indices of J to front and hence $\operatorname{sgn} \tau = (-1)^{kl}$, because τ. This proves anti-commutativity for alternating tensors and then the general case follows from bilinearity.

Part (4) is an immediate consequence of the fact that $\varepsilon^I \wedge \varepsilon^J = \varepsilon^{IJ}$. Part 5. is left as an exercise. □

Notation. We will write $\varepsilon^I = \varepsilon^{i_1} \wedge \cdots \wedge \varepsilon^{i_k}$.

Definition 14.8. For any n-dimensional vector space V, we define a vector space $\Lambda(V^*)$, called the ***exterior algebra*** (or ***Grassmann algebra***) *of* V. by

$$\Lambda(V^*) = \bigoplus_{k=0}^{n} \Lambda^k(V^*).$$

Note that $\dim \Lambda(V^*) = \binom{n}{0} + \binom{n}{1} + \cdots \binom{n}{n-1} + \binom{n}{n} = 2^n$.
The wedge product is an operation on $\Lambda(V^*)$ and it makes $\Lambda(V^*)$ into an associative algebra but not commutative.

Remark 14.9. An algebra A is said to be **graded** if it has a direct sum decomposition

$$A = \bigoplus_{k \in \mathbb{Z}} A^k$$

such that the product satisfies

$$(A^k)(A^l) \subseteq A^{k+l}$$

for each k and l. A graded algebra is ***anticommutative*** if the product satisfies

$$ab = (-1)^{kl} ba$$

for $a \in A^k$, $b \in A^l$. Thus, $\Lambda(V^*)$ is an anticommutative graded algebra.

We say that a k-form ω is **decomposable** is it can be written as a wedge product of k 1-forms, $\omega = \omega_1 \wedge \cdots \wedge \omega_k$ with ω_i beign 1-forms. Not every k-form is decomposable but every k-form can be written a linear combination of decomposable forms.

14.3 Interior Product

We now briefly discuss the notion of interior product.

Definition 14.10. Let V be a finite-dimensional vector space. For each $v \in V$, we define a linear map

$$i_v \colon \Lambda^k(V^*) \to \Lambda^{k-1}(V^*),$$

called **interior product by** v, as

$$i_v \omega(w_1, \ldots, w_{k-1}) = \omega(v, w_1, \ldots, w_{k-1}). \tag{14.8}$$

That is, $i_v\omega$ is obtained from ω by inserting v into the first slot. We will also use the notation

$$v \lrcorner \omega = i_v \omega.$$

and read it as "v hook ω."

Let us see some elementary properties of the interior product.

Proposition 14.11. Let V be a finite-dimensional vector space and $v \in V$.

1. $i_v \circ i_v = 0$.
2. *If* $\omega \in \Lambda^k(V^*)$ *and* $\eta \in \Lambda^l(V^*)$,

$$v \lrcorner (\omega \wedge \eta) = (v \lrcorner \omega) \wedge \eta + (-1)^k \omega \wedge (v \lrcorner \eta). \tag{14.9}$$

Proof. If ω is a function or a covector then taking interior product twice gives 0. On exterior forms of rank k for $k \geq 2$, notice that $i_v \circ i_v$ means we insert the same vector in the first two slots and since ω is alternating, we get part 1.

We now prove part 2. and note that it suffices to consider the case in which both ω and η are decomposable, since every exterior form can be written as a linear combination of decomposable ones.

Let $v_1 = v \in V$ and pick an arbitrary $(k-1)$-tuple of vectors $(v_2, \ldots, v_k)$. Let $\omega^1, \ldots, \omega^k$ be 1-forms. We first prove that

$$(\omega^1 \wedge \cdots \wedge \omega^k)(v_1, \ldots, v_k) = \sum_{i=1}^{k} (-1)^{i-1} \omega^i(v_1)(\omega^1 \wedge \cdots \wedge \widehat{\omega^i} \wedge \cdots \wedge \omega^k)(v_2, \ldots, v_k), \tag{14.10}$$

where the hat indicates that ω^i is omitted. To see why (14.10) is true, notice that the LHS is the determinant of the matrix $[A]_{ij}$ whose (i,j)-entry is $\omega^i(v_j)$. To analyze the right-hand side, let v_j^i denote the $(k-1)\times(k-1)$ submatrix of $[A]$ obtained by deleting the ith row and jth column. Then the right-hand side of (14.10) is

$$\sum_{i=1}^{k}(-1)^{i-1}\omega^i(v_1)\det v_1^i.$$

This is just the expansion of $\det[A]$ by minors along the first column, and therefore is equal to $\det[A]$. Thus, (14.10) holds. Thus, we get that

$$v\lrcorner(\omega^1\wedge\cdots\wedge\omega^k)=\sum_{i=1}^{k}(-1)^{i-1}\omega^i(v)\,\omega^1\wedge\cdots\wedge\widehat{\omega^i}\wedge\cdots\wedge\omega^k,$$

which proves part 2 for decomposable forms and hence for all exterior forms. $\square$

14.4 Differential forms on manifolds

After having understood exterior forms on vector spaces and various operations associated to a differential form, we now take $V = T_pM$ where M is a smooth manifold and $p \in M$ and mimic the previous sections. This gives the notion of the space of differential forms on a manifold as well as differential forms on a manifolds.

We denote the subset of T^kT^*M consisting of alternating tensors by Λ^kT^*M:

$$\Lambda^kT^*M = \coprod_{p\in M} \Lambda^k(T_p^*M).$$

As before, for every $p \in M$, $\Lambda_p^kT^*M$ is a $\binom{n}{k}$-dimensional vector space.

Definition 14.12. A **differential k-form** is a tensor field whose value at each point is an alternating tensor. The integer k is called the **degree** of the form and we denote the vector space of smooth k-forms by

$$\Omega^k(M) = \Gamma(\Lambda^kT^*M).$$

The wedge product of two differential forms is defined pointwise: $(\omega \wedge \eta)_p = \omega_p \wedge \eta_p$. Thus, the wedge product of a k-form with an l-form is a $(k+l)$-form. If f is a 0-form and η is a k-form, we interpret the wedge product $f \wedge \eta$ to mean the ordinary product $f\eta$. If we define

$$\Omega^*(M) = \bigoplus_{k=0}^{n} \Omega^k(M), \tag{14.11}$$

then the wedge product turns $\Omega^*(M)$ into an associative, anticommutative graded algebra.

In any smooth chart, a k-form ω can be written locally as

$$\omega = \sum_I \omega_I \, dx^{i_1} \wedge \cdots \wedge dx^{i_k},$$

where the coefficients ω_I are continuous functions defined on the coordinate domain. Note that

$$dx^{i_1} \wedge \cdots \wedge dx^{i_k} \left(\frac{\partial}{\partial x^{j_1}}, \cdots, \frac{\partial}{\partial x^{j_k}} \right) = \det \begin{pmatrix} \delta^{i_1}_{j_1} & \cdots & \delta^{i_1}_{j_k} \\ \vdots & \ddots & \vdots \\ \delta^{i_k}_{j_1} & \cdots & \delta^{i_k}_{j_k} \end{pmatrix}.$$

and the component functions ω_I of ω are determined by

$$\omega_I = \omega \left(\frac{\partial}{\partial x^{i_1}}, \cdots, \frac{\partial}{\partial x^{i_k}} \right).$$

We just state the expressions for pull-back of differential forms which are just special cases of pullback of tensor fields in § 13.4. Suppose $F\colon M \to N$ is smooth. Then

$$F^*(\omega \wedge \eta) = (F^*\omega) \wedge (F^*\eta), \tag{14.12}$$

and the local coordinate expressions is

$$F^* \left(\sum_{I-\text{increasing}} \omega_I \, dy^{i_1} \wedge \ldots \wedge dy^{i_k} \right) = \sum_{I-\text{increasing}} (\omega_I \circ F) \, d(y^{i_1} \circ F) \wedge \ldots \wedge d(y^{i_k} \circ F).$$

Let us see an explicit example.

Example 14.13. Define $F\colon \mathbb{R}^2 \to \mathbb{R}^3$ by $F(u,v) = (u, v, u^2 - v^2)$, and let ω be the 2-form $y\,dx \wedge dz + x\,dy \wedge dz$ on $\mathbb{R}^3$. The pullback $F^*\omega$ is computed as follows:

$$\begin{aligned} F^*(y\,dx \wedge dz + x\,dy \wedge dz) &= v\,du \wedge d(u^2 - v^2) + u\,dv \wedge d(u^2 - v^2) \\ &= v\,du \wedge (2u\,du - 2v\,dv) + u\,dv \wedge (2u\,du - 2v\,dv) \\ &= -2v^2\,du \wedge dv + 2u^2\,dv \wedge du = -2(u^2 + v^2)\,du \wedge dv, \end{aligned}$$

where we have used the fact $du \wedge du = dv \wedge dv = 0$ and $du \wedge dv = -dv \wedge du$ by anticommutativity.

Interior multiplication also extends naturally to vector fields and differential forms, simply by letting it act pointwise: if $X \in \mathfrak{X}(M)$ and $\omega \in \Omega^k(M)$, define a $(k-1)$-form $X \lrcorner\, \omega = i_X\omega$ by

$$(X \lrcorner\, \omega)_p = X_p \lrcorner\, \omega_p.$$

14.5 Exterior Derivatives

In this section we define a natural and probably the most important differential operator on smooth forms, called the **exterior derivative.** It is a generalization of the differential of a function.

For each smooth manifold M, we will show that there is a differential operator $d\colon \Omega^k(M) \to \Omega^{k+1}(M)$ satisfying $d(d\omega) = 0$ for all ω. We will do this in stages. First we will show below the definition of the operator d on $\mathbb{R}^n$ and then use that to prove the existence of an operator d on any smooth manifold which will satisfy the desired properties from $\mathbb{R}^n$. We will do so by giving an invariant formula for the exterior derivative d.

The definition of d on Euclidean space is as follows.

Definition 14.14. Let $\omega = \sum_J \omega_J\,dx^{j_1} \wedge \cdots \wedge dx^{j_k}$ be a smooth k-form on an open subset $U \subseteq \mathbb{R}^n$, we define its **exterior derivative** $d\omega \in \Omega^{k+1}$ by

$$d\left(\sum_J \omega_J\,dx^{j_1} \wedge \cdots \wedge dx^{j_k}\right) = \sum_J d\omega_J \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_k}, \tag{14.13}$$

where $d\omega_J$ is the differential of the function ω_J, and hence (14.13) can be written as

$$d\left(\sum_J \omega_J\,dx^{j_1} \wedge \ldots \wedge dx^{j_k}\right) = \sum_J \sum_i \frac{\partial \omega_J}{\partial x^i}\,dx^i \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_k}. \tag{14.14}$$

So for instance, if ω is a 1-form, then (14.14) implies

$$\begin{aligned} d(\omega_j\,dx^j) &= \sum_{i,j} \frac{\partial \omega_j}{\partial x^i}\,dx^i \wedge dx^j \\ &= \sum_{i<j} \frac{\partial \omega_j}{\partial x^i}\,dx^i \wedge dx^j + \sum_{i>j} \frac{\partial \omega_j}{\partial x^i}\,dx^i \wedge dx^j \\ &= \sum_{i<j} \left(\frac{\partial \omega_j}{\partial x^i} - \frac{\partial \omega_i}{\partial x^j}\right) dx^i \wedge dx^j \end{aligned}$$

after we interchange i and j in the second sum and use the fact $dx^j \wedge dx^i = -dx^i \wedge dx^j$. For a smooth 0-form f, (14.14) implies

$$df = \frac{\partial f}{\partial x^i}\,dx^i,$$

which is just the differential of f.

Let us see some properties of the exterior derivative on $\mathbb{R}^n$ before seeing its definition on a manifold.

Proposition 14.15. *The exterior derivative d satisfies the following.*

(a) *d is linear over $\mathbb{R}$.*

(b) *If $\omega \in \Omega^k(U)$ and $\eta \in \Omega^l(U)$, $U \subset \mathbb{R}^n$ open, then*

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta. \tag{14.15}$$

(c) *$d \circ d \equiv 0$. In particular, this tells us that $\operatorname{Im}(d_{k-1}) \subset \ker(d_k)$.*

(d) *d commutes with pullbacks: for $U \underset{open}{\subset} \mathbb{R}^n$, $V \underset{open}{\subset} \mathbb{R}^m$, $F \colon U \to V$ a smooth map, and $\omega \in \Omega^k(V)$, we have*

$$F^*(d\omega) = d(F^*\omega). \tag{14.16}$$

Proof. Linearity of d is an immediate consequence of the definition. To prove (b), by linearity it suffices to consider terms of the form $\omega = u\, dx^I \in \Omega^k(U)$ and $\eta = v\, dx^J \in \Omega^l(U)$ for smooth real-valued functions u and v. We compute

$$\begin{aligned}
d(\omega \wedge \eta) &= d((u\, dx^I) \wedge (v\, dx^J)) \\
&= d(uv\, dx^I \wedge dx^J) && \text{(linearity of } \wedge \text{ over } C^\infty(M)) \\
&= (v\, du + u\, dv) \wedge dx^I \wedge dx^J && \text{(product rule)} \\
&= (du \wedge dx^I) \wedge (v\, dx^J) + (-1)^k (u\, dx^I) \wedge (dv \wedge dx^J) && \text{(associativity of } \wedge) \\
&= d\omega \wedge \eta + (-1)^k \omega \wedge d\eta,
\end{aligned}$$

where the $(-1)^k$ comes from the fact that $dv \wedge dx^I = (-1)^k dx^I \wedge dv$ because dv is a 1-form and dx^I is a k-form.

Note that the second order derivatives commute for functions, hence $d \circ d = 0$ for 0-forms, i.e., for $f \in C^\infty(M)$.

We prove (c) first for functions which are 0-forms, which is just a real-valued function. In this case,

$$\begin{aligned}
d(du) = d\left(\frac{\partial u}{\partial x^j}\, dx^j\right) &= \frac{\partial^2 u}{\partial x^i \partial x^j}\, dx^i \wedge dx^j \\
&= \sum_{i<j} \left(\frac{\partial^2 u}{\partial x^i \partial x^j} - \frac{\partial^2 u}{\partial x^j \partial x^i}\right) dx^i \wedge dx^j = 0.
\end{aligned}$$

For the general case, we have

$$\begin{aligned}
d(d\omega) &= d\left(\sum_{J-\text{increasing}} d\omega_J \wedge dx^J\right) \\
&= \sum_{J-\text{increasing}} d(d\omega_J) \wedge dx^J - \sum_{J-\text{increasing}} d\omega_j \wedge d(dx^J) && \text{(using (b))} \\
&= \sum_{J-\text{increasing}} d(d\omega_J) \wedge dx^J + \sum_{J-\text{increasing}} \sum_{i=1}^{k} (-1)^i d\omega_J \wedge dx^{j_1} \wedge \ldots \wedge d(dx^{j_i}) \wedge \ldots \wedge dx^{j_k} = 0,
\end{aligned}$$

where we used the $k = 0$ case for the functions ω_J and each of the coordinate functions x^{j_i}.

Finally, to prove (d), again it suffices to consider $\omega = u\, dx^{i_1} \wedge \ldots \wedge dx^{i_k}$. For such a form, the left-hand side of (14.16) is

$$\begin{aligned}
F^*(d(u\, dx^{i_1} \wedge \ldots \wedge dx^{i_k})) &= F^*(du \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k}) \\
&= d(u \circ F) \wedge d(x^{i_1} \circ F) \wedge \ldots \wedge d(x^{i_k} \circ F),
\end{aligned}$$

and the right-hand side is

$$\begin{aligned} d(F^*(u\,dx^{i_1} \wedge \ldots \wedge dx^{i_k})) &= d\left((u \circ F)\, d(x^{i_1} \circ F) \wedge \ldots \wedge d(x^{i_k} \circ F)\right) \\ &= d(u \circ F) \wedge d(x^{i_1} \circ F) \wedge \ldots \wedge d(x^{i_k} \circ F), \end{aligned}$$

so they are equal. □

We now want to mimic the situation in $\mathbb{R}^n$ and define the exterior derivative on manifolds. Obviously, we would like the exterior derivative opeartor $d : \Omega^k(M) \to \Omega^{k+1}(M)$ to satisfy the properties in Proposition 14.15. It should also generalize the differential of functions. Unlike the wedge product, the interior product and the pull-back operations that we defined earlier, the exterior derivative is no longer a pointwise operation, but is a local operation (i.e. depends on the "nearby values")

Let $U \underset{\text{open}}{\subset} M$. For $f \in \Omega^0(U) = C^\infty(U)$ we have already seen that $df \in \Omega^1(U)$. So we get a linear map

$$d : \Omega^0(U) \to \Omega^1(U), \quad f \mapsto df.$$

Locally on each coordinate chart we have

$$df = \sum_i (\partial_i f) dx^i,$$

where we use the shorthand notation $\partial_i = \dfrac{\partial}{\partial x^i}$. We also have an "invariant definition" of $df \in \Omega^1(U)$, via

$$df(X) = Xf, \quad \forall X \in \Gamma(TU).$$

Now suppose ω is a k-form on M, so that locally

$$\omega = \sum_I \omega_{i_1,\ldots,i_k} dx^{i_1} \wedge \ldots \wedge dx^{i_k}.$$

We use (14.13) to give the following definition.

Definition 14.16. The exterior derivative of $\omega \in \Omega(U)$ is the $(k+1)$-form $d\omega$ given by the formula

$$\begin{aligned} d\omega &= \sum_I d\omega_{i_1,\ldots,i_k} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} \\ &= \sum_{I,i} \partial_i(\omega_{i_1,\ldots,i_k}) dx^i \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k}. \end{aligned} \tag{14.17}$$

Before we proceed, we need to clarify that $d\omega$ defined above is well-defined. In other words, the $(k+1)$-form $d\omega$ defined above should be independent of the choices of coordinate patches.

We can do it either by checking that (14.17) is unchanged if one use another coordinate chart or we can come up with an equivalent but coordinate-free definition (usually called the invariant formulation). We will take the second approach here which will prove the existence of the operator d on M which when restricted to a coordinate chart (U, x) will give (14.17).

Let's start with small k's to find out the invariant formula of $d\omega$.

- For $k = 0$, i.e. $\omega = f \in C^\infty(U)$, we can regard df as a $C^\infty(U)$-linear map

  $$df : \Gamma(TU) \to C^\infty(U)$$

  such that

  $$df(X) = Xf.$$

- For $k = 1$, i.e. $\omega \in \Omega^1(U)$, $d\omega \in \Omega^2(U)$ and so should have two vector fields as its argument and a function as its output. So
$$d\omega : \Gamma(TU) \times \Gamma(TU) \to C^\infty(U).$$
Let us use the formula (14.17) to see an invariant expression of $d\omega$. We write $\omega = \sum_i \omega_i dx^i$, $X = \sum_k X^k \partial_k$ and $Y = \sum_l Y^l \partial_l$. Then
$$\begin{aligned} d\omega(X,Y) &= \sum_{i,j,k,l} (\partial_j \omega_i) dx^j \wedge dx^i (X^k \partial_k, Y^l \partial_l) \\ &= \sum_{i,j} \big((\partial_j\omega_i) X^j Y^i - (\partial_j \omega_i) X^i Y^j \big) \\ &= \sum_{i,j} \big(X^j \partial_j (\omega_i Y^i) - \omega_i X^j \partial_j (Y^i) - Y^j \partial_j (\omega_i X^i) + \omega_i Y^j \partial_j (X^i) \big) \\ &= X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y]) \qquad \left(\text{using (9.9) } [X,Y] = \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \right) \frac{\partial}{\partial x^j} \right) \end{aligned}$$
So we arrive at
$$d\omega(X,Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y]).$$
- For $k = 2$, i.e. $\omega \in \Omega^2(U)$, by a tedious computation one can prove that as a $C^\infty(U)$-trilinear map
$$d\omega : \Gamma(TU) \times \Gamma(TU) \times \Gamma(TU) \to C^\infty(U),$$
we have
$$\begin{aligned} d\omega(X,Y,Z) =& X(\omega(Y,Z)) - Y(\omega(X,Z)) + Z(\omega(X,Y)) \\ & - \omega([X,Y],Z) + \omega([X,Z],Y) - \omega([Y,Z],X). \end{aligned}$$

So we are led to the following *the invariant formula* for $d\omega$

Theorem 14.17. *For any $\omega \in \Omega^k(U)$, the $(k+1)$-form $d\omega$, viewed as a $C^\infty(U)$-multilinear map*
$$d\omega : \underbrace{\Gamma(TU) \times \ldots \times \Gamma(TU)}_{k+1-times} \to C^\infty(U),$$
is given by the formula
$$d\omega(X_1, \ldots, X_{k+1}) := \sum_i (-1)^{i-1} X_i(\omega(X_1, \ldots, \widehat{X}_i, \ldots, X_{k+1})) \tag{14.18}$$
$$+ \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_1, \ldots, \widehat{X}_i, \ldots, \widehat{X}_j, \ldots, X_{k+1}). \tag{14.19}$$

Proof. Define $d\omega$ via the formula in (14.19). We need to show that

1. $d\omega$ is an alternating tensor, i.e. for any $m < n$,

$$d\omega(X_1, \dots, X_m, \dots, X_n, \dots, X_{k+1}) = -d\omega(X_1, \dots, X_n, \dots, X_m, \dots, X_{k+1}).$$

   This can be checked to be true from (14.19).

(2) $d\omega$ is multi-linear is $C^\infty(U)$-linear on U. Note that $d\omega$ is obviously $\mathbb{R}$-linear. We prove that for any $f \in C^\infty(U)$,

$$d\omega(fX_1, X_2, \dots, X_{k+1}) = f d\omega(X_1, \dots, X_{k+1}).$$

Let use compute.

$$\begin{aligned}
d\omega(fX_1, X_2, \dots, X_{k+1}) =& fX_1(\omega(X_2, \dots, X_{k+1})) \\
&+ \sum_{i>1} (-1)^{i-1} X_i(\omega(fX_1, \dots, \widehat{X}_i, \dots, X_{k+1})) \\
&+ \sum_{i>1} (-1)^{1+i} \omega([fX_1, X_i], X_2, \dots, \widehat{X}_i, \dots, X_{k+1}) \\
&+ \sum_{1<i<j} (-1)^{i+j} \omega([X_i, X_j], fX_1, \dots, \widehat{X}_i, \dots, \widehat{X}_j, \dots, X_{k+1}) \\
=& f d\omega(X_1, \dots, X_{k+1}) \\
&+ \sum_{i>1} (-1)^{i-1} (X_i f) \omega(X_1, \dots, \widehat{X}_i, \dots, X_{k+1}) \\
&- \sum_{i>1} (-1)^{1+i} (X_i f) \omega(X_1, \dots, \widehat{X}_i, \dots, X_{k+1}) \\
=& f d\omega(X_1, \dots, X_{k+1}),
\end{aligned}$$

where we used the property $([fX, gY] = fg[X, Y] + (fXg)Y - (gYf)X)$ in the second to last equality.

(3) We now check that $d\omega$ has the local expression (14.17). This is an exercise.

□

We get the same formulas for wedge product, interior product and pull-backs of differential forms on manifolds as we had them for forms on vector spaces.

The *wedge product* $\wedge : \Omega^k(U) \times \Omega^l(U) \to \Omega^{k+l}(U)$.
For example,

$$(dx^1 + 2dx^2) \wedge (dx^1 \wedge dx^2 - dx^2 \wedge dx^3 + 3dx^1 \wedge dx^3) = -7dx^1 \wedge dx^2 \wedge dx^3.$$

For any $V \in \Gamma(TM)$ we have the *interior product* $i_V : \Omega^k(M) \to \Omega^{k-1}(M)$ with

$$V \lrcorner \omega(X_1, \dots, X_{k-1}) = i_V \omega(X_1, \dots, X_{k-1}) = \omega(V, X_1, \dots, X_{k-1}).$$

For any smooth map $F : M \to N$, one has the *pull-back* $F^* : \Omega^k(N) \to \Omega^k(M)$. This is defined pointwise via the linear map $dF_p : T_pM \to T_{F(p)}N$. So if $\omega \in \Omega^k(N)$, then

$$(F^*\omega)_p(X_1, \dots, X_k) = \omega_{F(p)}(dF_p(X_1), \dots, dF_p(X_k)).$$

We list all the properties related to differential forms on manifolds.

Lemma 14.18. *Let M and N be a smooth manifolds. Suppose $\omega \in \Omega^k(M)$, $\eta \in \Omega^l(M)$, $X \in \Gamma(TM)$ and $F : M \to N$ be a smooth map. Then*

1. $\omega \wedge \eta = (-1)^{kl}\eta \wedge \omega$.
2. $F^*(\omega \wedge \eta) = F^*\omega \wedge F^*\eta$.
3. $i_X(\omega \wedge \eta) = (i_X\omega) \wedge \eta + (-1)^k\omega \wedge i_X\eta$.
4. $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k\omega \wedge d\eta$.
5. $i_X \circ i_X = 0$.
6. $d \circ d = 0$.
7. $F * \circ d = d \circ F^*$.

Proof. The proof is similar to the proofs we saw for the vector spaces. □

Definition 14.19. A differential form ω is called **closed** if $d\omega = 0$. A differential form $\omega \in \Omega^k(M)$ is called **exact** if $\omega = d\eta$ for some $\eta \in \Omega^{k-1}(M)$. Since $d \circ d = 0$, we get that

Every exact form is closed.

14.6 Lie Derivatives of Differential Forms

We derived formulas for computing Lie derivatives of smooth tensor fields in § 13.5, which apply equally well to differential forms. However, in the case of differential forms, the exterior derivative yields a much more powerful formula for computing Lie derivatives, which also has significant theoretical consequences.

So the concepts which we want to understand are how the Lie derivative behaves with:wedge products, interior products and the exterior derivative. We start with the wedge product and recall that the wedge product $\omega \wedge \eta$ is the skew-symmetrization of the tensor product, Moreover, Proposition 13.19 shows that $\mathcal{L}_V(A \otimes B) = \mathcal{L}_V A \otimes B + A \otimes \mathcal{L}_V B$. Thus, we get the following

Proposition 14.20. *Suppose M is a smooth manifold, $V \in \mathfrak{X}(M)$, and $\omega, \eta \in \Omega^*(M)$. Then*

$$\mathcal{L}_V(\omega \wedge \eta) = (\mathcal{L}_V\omega) \wedge \eta + \omega \wedge (\mathcal{L}_V\eta).$$

The next theorem is one of the most important equations about the derivatives of differential forms. It gives a remarkable formula for Lie derivatives of differential forms, which dates back to Élie Cartan, the French mathematician who invented the theory of differential forms.

Theorem 14.21 (Cartan's magic formula). *On a smooth manifold M, for any smooth vector field V and any smooth differential form ω,*

$$\mathcal{L}_V\omega = V \lrcorner (d\omega) + d(V \lrcorner \omega). \tag{14.20}$$

Proof. We prove that (14.20) holds for smooth k-forms by induction on k. We begin with a smooth 0-form f, in which case

$$V \lrcorner (df) + d(V \lrcorner f) = V \lrcorner df = df(V) = Vf = \mathcal{L}_V f,$$

which is (14.20).

Now let $k \geq 1$, and suppose (14.20) has been proved for forms of degree less than k. Let ω be an arbitrary smooth k-form, written in smooth local coordinates as

$$\omega = \sum_I \omega_I \, dx^{i_1} \wedge \ldots \wedge dx^{i_k}.$$

We want to use the induction hypothesis and the base case so we write $f = x^{i_1}$ and $\beta = \omega_I \, dx^{i_2} \wedge \ldots \wedge dx^{i_k}$ and observe that each term in the expression of ω can be written in the form $df \wedge \beta$, where f is a smooth function and β is a smooth $(k-1)$-form. We know from Corollary 12.18 that $\mathcal{L}_V df = d(\mathcal{L}_V f) = d(Vf)$. Thus,

$$\begin{aligned} \mathcal{L}_V(df \wedge \beta) &= (\mathcal{L}_V df) \wedge \beta + df \wedge (\mathcal{L}_V \beta) \\ &= d(Vf) \wedge \beta + df \wedge (V \lrcorner \, d\beta + d(V \lrcorner \, \beta)). \end{aligned}$$

On the other hand, noting that $V \lrcorner \, df = df(V) = Vf$, we compute

$$\begin{aligned} V \lrcorner \, d(df \wedge \beta) + d(V \lrcorner \, (df \wedge \beta)) &= V \lrcorner \, (-df \wedge d\beta) + d((Vf)\beta - df \wedge (V \lrcorner \, \beta)) \\ &= -(Vf)d\beta + df \wedge (V \lrcorner \, d\beta) + d(Vf) \wedge \beta + (Vf)d\beta + df \wedge d(V \lrcorner \, \beta), \end{aligned}$$

where we used the fact that $V \lrcorner (\omega \wedge \eta) = (V \lrcorner \omega) \wedge \eta + (-1)^k \omega \wedge (V \lrcorner \eta)$. Thus,

$$\mathcal{L}_V(df \wedge \beta) = V \lrcorner \, d(df \wedge \beta) + d(V \lrcorner \, (df \wedge \beta))$$

and (14.20) follows from induction. □

We get the following corollary.

Corollary 14.22 (Lie Derivative Commutes with d). *If V is a smooth vector field and ω is a smooth differential form, then*

$$\mathcal{L}_V(d\omega) = d(\mathcal{L}_V \omega).$$

Proof. This follows from Cartan's formula and the fact that $d \circ d = 0$:

$$\begin{aligned} \mathcal{L}_V d\omega &= V \lrcorner \, d(d\omega) + d(V \lrcorner \, d\omega) = d(V \lrcorner \, d\omega); \\ d\mathcal{L}_V \omega &= d(V \lrcorner \, d\omega) + d(d(V \lrcorner \, \omega)) = d(V \lrcorner \, d\omega). \end{aligned}$$ □

Exercise 14.23. Let M be a smooth manifold and $V \in \mathfrak{X}(M)$.

(a) Show that $\mathcal{L}_V \circ i_V = i_V \circ \mathcal{L}_V$.

(b) Show that $i_V \circ \mathcal{L}_V \omega = i_V(V \lrcorner \, d\omega)$ for any smooth form ω.

Using the invarinat formula (14.19) for $d\omega$ and Cartan's magic formula we get another formula for the Lie derivative of a form.

Proposition 14.24. *If ω is a smooth k-form and $X, Y_1, \ldots, Y_k$ are smooth vector fields on a smooth manifold M, then*

$$(\mathcal{L}_X \omega)(Y_1, \ldots, Y_k) = X(\omega(Y_1, \ldots, Y_k)) - \sum_{i=1}^{k} \omega(Y_1, \ldots, [X, Y_i], \ldots, Y_k). \tag{14.21}$$

15. Integration on manifolds

15.1 Partitions of unity

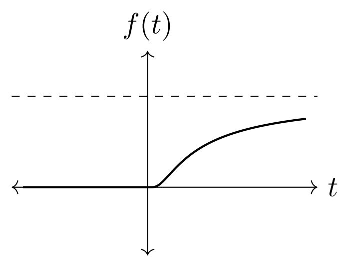
We start our discussions with one of the most important classes of smooth functions called **bump functions**. Recall that we used these functions in our proof of the Tensor characterization lemma. We start with the following fact from real analysis.

Proposition 15.1. *The function $f : \mathbb{R} \to \mathbb{R}$ defined by*

$$f(t) = \begin{cases} e^{-1/t}, & t > 0, \\ 0, & t \leq 0, \end{cases} \tag{15.1}$$

is smooth. $\square$

The function's graph look like the figure below.



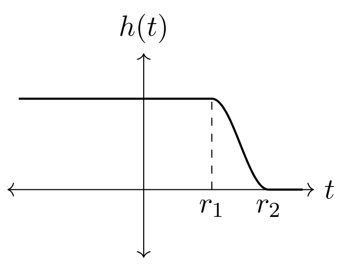
As a result, given any real numbers r_1 and r_2 such that $r_1 < r_2$, there exists a smooth function $h : \mathbb{R} \to \mathbb{R}$ such that

$$\begin{cases} h(t) \equiv 1, & t \leq r_1, \\ 0 < h(t) < 1, & r_1 < t < r_2, \\ h(t) \equiv 0, & t \geq r_2 \end{cases}$$

by defining

$$h(t) = \frac{f(r_2 - t)}{f(r_2 - t) + f(t - r_1)}.$$

with $f(t)$ as in (15.1). Such an h is called a **cutoff function** and it looks like the figure below.



We can easily extend this construction to $\mathbb{R}^n$ as the following result shows.

Proposition 15.2. *Given any positive real numbers $r_1 < r_2$, there is a smooth function $H : \mathbb{R}^n \to \mathbb{R}$ such that*

$$\begin{cases} H(x) \equiv 1, & x \in \overline{B}_{r_1}(0) \\ 0 < H(x) < 1, & x \in B_{r_2}(0) \setminus \overline{B}_{r_1}(0) \\ H(x) \equiv 0, x \in \mathbb{R}^n \setminus B_{r_2}(0). \end{cases} \tag{15.2}$$

Proof. Set $H(x) = h(|x|)$, with h as defined above. Clearly, H is smooth on $\mathbb{R}^n \setminus \{0\}$, because it is a composition of smooth functions there. Since it is identically equal to 1 on $B_{r_1}(0)$, it is smooth there too. $\square$

The function H constructed in this proposition is an example of a **smooth bump function**, a smooth real-valued function that is equal to 1 on a specified set and is zero outside a specified neighbourhood of that set.

We can now define partitions of unity.

Definition 15.3. Let M be a manifold and let $\mathcal{U} = (U_\alpha)_{\alpha\in A}$ be an arbitrary open cover of M. A **partition of unity subordinate to** $\mathcal{U}$ is an indexed family $(\psi_\alpha)_{\alpha\in A}$ of continuous functions $\psi_\alpha : M \to \mathbb{R}$ with the following properties:

(i) $0 \leq \psi_\alpha(x) \leq 1$ for all $\alpha \in A$ and all $x \in M$.

(ii) $\operatorname{supp}\psi_\alpha \subseteq U_\alpha$ for each $\alpha \in A$.

(iii) The family of supports $(\operatorname{supp}\psi_\alpha)_{\alpha\in A}$ is locally finite, meaning that every point has a neighborhood that intersects $\operatorname{supp}\psi_\alpha$ for only finitely many values of α.

(iv) $\sum_{\alpha\in A}\psi_\alpha(x) = 1$ for all $x \in M$.

If M is a smooth manifold, a **smooth partition of unity** is one for which each of the functions ψ_α is smooth.

Because of the local finiteness condition (iii), the sum in (iv) actually has only finitely many nonzero terms in a neighbourhood of each point, so there is no issue of convergence.

Theorem 15.4. *Suppose M is a smooth manifold and $\mathcal{U} = (U_\alpha)_{\alpha\in A}$ is any indexed open cover of M. Then there exists a smooth partition of unity subordinate to $\mathcal{U}$.*

Proof. Each set U_α is a smooth manifold in its own right, and thus has a basis $\mathcal{B}_\alpha$ of coordinate balls and thus $\mathcal{B} = \bigcup_\alpha \mathcal{B}_\alpha$ is a basis for the topology of M. Since M is a manifold, the cover $\mathcal{U}$ has a countable, locally finite refinement $\{B_i\}$ consisting of elements of $\mathcal{B}$ and the cover $\{\overline{B}_i\}$ is also locally finite.

For each i, the fact that B_i is a coordinate ball in some U_α implies that there is a coordinate ball $B_i' \subseteq U_\alpha$ such that $B_i' \supseteq \overline{B}_i$, and a smooth coordinate map $\varphi_i : B_i' \to \mathbb{R}^n$ such that $\varphi_i(\overline{B}_i) = \overline{B}_{r_i}(0)$ and $\varphi_i(B_i') = B_{r_i'}(0)$ for some $r_i < r_i'$. For each i, define a function $f_i : M \to \mathbb{R}$ by

$$f_i = \begin{cases} H_i \circ \varphi_i & \text{on } B_i', \\ 0 & \text{on } M \setminus \overline{B}_i, \end{cases}$$

where $H_i : \mathbb{R}^n \to \mathbb{R}$ is a smooth function that is positive in $B_{r_i}(0)$ and zero elsewhere, as in Proposition 15.2. On the set $B_i' \setminus \overline{B}_i$ where the two definitions overlap, both definitions yield the zero function, so f_i is well defined and smooth, and $\operatorname{supp} f_i = \overline{B}_i$.

Define $f : M \to \mathbb{R}$ by

$$f(x) = \sum_i f_i(x).$$

Because of the local finiteness of the cover $\{\overline{B}_i\}$, this sum has only finitely many nonzero terms in a neighbourhood of each point and thus defines a smooth function. Because each f_i is non-negative everywhere and positive on B_i, and every point of M is in some B_i, it follows that $f(x) > 0$ everywhere on M. Thus, the functions $g_i : M \to \mathbb{R}$ defined by $g_i(x) = f_i(x)/f(x)$ are also smooth. It is immediate from the definition that $0 \leq g_i \leq 1$ and $\sum_i g_i \equiv 1$.

Finally, we need to re-index our functions so that they are indexed by the same set A as our open cover. Because the cover $\{B'_i\}$ is a refinement of $\mathcal{U}$, for each i we can choose some index $a(i) \in A$ such that $B'_i \subseteq U_{a(i)}$. For each $\alpha \in A$, define $\psi_\alpha : M \to \mathbb{R}$ by

$$\psi_\alpha = \sum_{i:a(i)=\alpha} g_i.$$

If there are no indices i for which $a(i) = \alpha$, then this sum should be interpreted as the zero function. Thus,

$$\operatorname{supp} \psi_\alpha = \overline{\bigcup_{i:a(i)=\alpha} B_i} = \bigcup_{i:a(i)=\alpha} \overline{B}_i \subseteq U_\alpha.$$

Each ψ_α is a smooth function that satisfies $0 \le \psi_\alpha \le 1$. Moreover, the family of supports $(\operatorname{supp} \psi_\alpha)_{\alpha \in A}$ is still locally finite, and $\sum_\alpha \psi_\alpha \equiv \sum_i g_i \equiv 1$, so this is the desired partition of unity. □

15.2 Top forms on manifolds

Let M^n be a smooth manifold. Since, $\Omega^k(M) = 0$ for $k > m$, we call any smooth n-form a **top form** on M. Let $p \in M$ and (φ, U) a local chart near p. Then $dx^1 \wedge \cdots \wedge dx^n$ is a top form on U. Note that for any $q \in U$, $(dx^1 \wedge \ldots \wedge dx^m)_q \neq 0$ since at any $q \in U$,

$$(dx^1 \wedge \cdots \wedge dx^n)_q(\partial_1, \ldots, \partial_n) = \det(dx^i(\partial_j))_{1 \le i,j \le n} = 1.$$

Moreover, since $\dim \Lambda^n T_p^* M = 1$, we see that for any top form ω on U and any $q \in U$, there is a real number λ_q such that

$$\omega_q = \lambda_q (dx^1 \wedge \ldots \wedge dx^m)_q.$$

Since ω is C^∞, the coefficient λ, as a function on U, is C^∞. So up to multiplication by functions, there is a "canonical top form" in the chart. Of course, this may not be true globally on M: For two top forms $\omega, \eta \in \Omega^n(M)$, it may happen that $\omega_p = 0, \omega_q \neq 0$ while $\eta_p \neq 0, \eta_q = 0$, and thus there exists no function $f \in C^\infty(M)$ with $\omega = f\eta$ or $\eta = f\omega$.

In particular, if we change coordinates from $(x_\alpha^1, \ldots, x_\alpha^n)$ to $(x_\beta^1, \ldots, x_\beta^n)$ on U, we get two top forms, $dx_\alpha^1 \wedge \ldots \wedge dx_\alpha^n$ and $dx_\beta^1 \wedge \ldots \wedge dx_\beta^n$. They should be related by a smooth function on U.

Proposition 15.5. *If $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ is a diffeomorphism and $y = \varphi(x)$, then*

$$\varphi^*(dy^1 \wedge \ldots \wedge dy^n) = \det(d\varphi_x)\, dx^1 \wedge \ldots \wedge dx^n$$

Proof. If we denote $\varphi = (\varphi^1, \ldots, \varphi^n)$, then $\varphi^* y^i = y^i \circ \varphi = \varphi^i$. So

$$\varphi^*(dy^1 \wedge \ldots \wedge dy^n) = d\varphi^1 \wedge \ldots \wedge d\varphi^n.$$

But since

$$d\varphi^1 \wedge \ldots \wedge d\varphi^m(\partial_1^x, \ldots, \partial_n^x) = \det(d\varphi_x),$$

we have

$$d\varphi^1 \wedge \ldots \wedge d\varphi^n = \det(d\varphi_x)\, dx^1 \wedge \ldots \wedge dx^n.$$

□

Now let $(\varphi_\alpha, U, V_\alpha)$ and $(\varphi_\beta, U, V_\beta)$ be two coordinate charts. Then the coordinate change map is $\varphi_{\alpha\beta} = \varphi_\beta \circ \varphi_\alpha^{-1}$, which maps $\varphi_\alpha(x)$ to $y = \varphi_\beta(x)$. So we get

$$(\varphi_{\alpha\beta})^*(\varphi_\beta^{-1})^* dx_\beta^1 \wedge \ldots \wedge dx_\beta^n = \det(d\varphi_{\alpha\beta})\, (\varphi_\alpha^{-1})^*\, dx_\alpha^1 \wedge \ldots \wedge dx_\alpha^n.$$

Since $(\varphi_{\alpha\beta})^*(\varphi_\beta^{-1})^* = (\varphi_\beta^{-1} \circ \varphi_{\alpha\beta})^* = \varphi_\alpha^{-1*}$, we get at

$$dx_\beta^1 \wedge \ldots \wedge dx_\beta^n = \det(d\varphi_{\alpha\beta})\, dx_\alpha^1 \wedge \ldots \wedge dx_\alpha^n.$$

15.3 Orientability

Let M be a smooth manifold of dimension m, and let $\omega \in \Omega^m(M)$ be a smooth m-form. We want to define the integral $\int_M \omega$. For simplicity let's suppose ω is supported on a chart (φ, U, V) with coordinates $\{x^1, \ldots, x^m\}$. Then we can write

$$\omega = f(\varphi(x))dx^1 \wedge \ldots \wedge dx^m,$$

where f is a smooth function on V. With the help of the Euclidean differential form $f(x)dx^1 \ldots dx^m$ on V, it is natural to define

$$\int_U \omega := \int_V f(x)dx^1 \ldots dx^m. \tag{15.3}$$

Then as usual one need to check that the integral in the right hand side is independent of the choice of coordinate charts.

So we let $(\varphi_\alpha, U, V_\alpha)$ and $(\varphi_\beta, U, V_\beta)$ be two coordinate systems on U, with transition map $\varphi_{\alpha\beta} = \varphi_\beta \circ \varphi_\alpha^{-1} : V_\alpha \to V_\beta$ which maps $x_\alpha = \varphi_\alpha(x)$ to $x_\beta = \varphi_\beta(x)$. Then

$$\omega = f_\beta(x_\beta)dx_\beta^1 \wedge \ldots \wedge dx_\beta^m = f_\beta(\varphi_{\alpha\beta}(x_\alpha)) \det(d\varphi_{\alpha\beta})dx_\alpha^1 \wedge \ldots \wedge dx_\alpha^m$$

So for the definition to be well-defined, we need

$$\int_{V_\beta} f_\beta(x_\beta)dx_\beta^1 \ldots dx_\beta^m = \int_{V_\alpha} f_\beta(\varphi_{\alpha\beta}(x_\alpha)) \det(d\varphi_{\alpha\beta})dx_\alpha^1 \ldots dx_\alpha^m.$$

Unfortunately this is not always true: In calculus we learned that for the integrals of multi-variable functions

$$\int f(x)dx^1 \ldots dx^m,$$

if $\varphi : V_1 \to V_2$ is a diffeomorphism, then we have the *change of variable formula*: $(y = \varphi(x))$

$$\int_{V_2} f(x)dy^1 \ldots dy^m = \int_{V_1} f(\varphi(x))|\det(d\varphi)(x)|dx^1 \ldots dx^m. \tag{15.4}$$

So in order to have

$$\int_{V_\beta} f_\beta(x_\beta)dx_\beta^1 \ldots dx_\beta^m = \int_{V_\alpha} f_\beta(\varphi_{\alpha\beta}(x_\alpha))\ \det(d\varphi_{\alpha\beta})(x_\alpha)\ dx_\alpha^1 \ldots dx_\alpha^m.$$

and thus, for the definition to be independent of the choice of charts, we need to assume

$$\det(d\varphi_{\alpha\beta}) > 0$$

for all charts. This motivates the following definition.

Definition 15.6. Let M be a smooth manifold of dimension n. Two charts $(\varphi_\alpha, U_\alpha, V_\alpha)$ and $(\varphi_\beta, U_\beta, V_\beta)$ are **orientation compatible** if the transition map $\varphi_{\alpha\beta} = \varphi_\beta \circ \varphi_\alpha^{-1}$ satisfies

$$\det(d\varphi_{\alpha\beta})_p > 0,$$

for all $p \in \varphi_\alpha(U_\alpha \cap U_\beta)$.
An **orientation** of M is an atlas $\mathcal{A} = \{(\varphi_\alpha, U_\alpha, V_\alpha) \mid \alpha \in A\}$ whose charts are pairwise orientation compatible.
We say M is **orientable** if it has an orientation.

Remark 15.7. Let U be a chart with coordinates $\{x^1, \ldots, x^m\}$. We use the notation $-U$ to represent the same coordinate chart U but with "twisted" coordinates $\{-x^1, x^2, \ldots, x^m\}$. Then $-U$ and U are not orientation compatible. Let $\tilde{U}$ be any other coordinate chart such that $\tilde{U} \cap U \neq \emptyset$ is *connected*. Then either

- $\tilde{U}$ and U are orientation compatible,

or

- $\tilde{U}$ and $-U$ are orientation compatible.

As a consequence, we immediately see

Corollary 15.8. *If M is connected and orientable, then M admits exactly two different orientations.*

Example 15.9. For the real projective space $\mathbb{RP}^n$, we have constructed an atlas consisting of $n+1$ charts. We have seen that $\mathbb{RP}^n$ is orientable for odd n. It turns out that $\mathbb{RP}^n$ is not orientable for even n.

15.4 Integrations of top forms on smooth manifolds.

Now assume M is a smooth orientable m-manifold and fix an orientation $\mathcal{A}$ on M. Let ω be any smooth m-form on M. To define $\int_M \omega$, we first assume that ω is supported in a coordinate chart (φ, U, V) which is orientation compatible with $\mathcal{A}$. In this case there is a function f supported in U such that

$$\omega = f(\varphi(x))dx^1 \wedge \ldots \wedge dx^m.$$

In this case we simply define

$$\int_U \omega := \int_V f(x)dx^1 \ldots dx^m, \tag{15.5}$$

where the right hand side is the Lebesgue integral on $V \subset \mathbb{R}^m$. We will assume f is integrable. In fact, in what follows the function f involved are compactly supported and hence integrable.

To integrate a general m-form $\omega \in \Omega^m(M)$, we take a locally finite cover $\{U_\alpha\}$ of M that are compatible with the orientation $\mathcal{A}$. Let $\{\psi_\alpha\}$ be a partition of unity subordinate to $\{U_\alpha\}$. Now since each ψ_α is supported in U_α, each $\psi_\alpha\omega$ is supported in U_α as well. We define

$$\int_M \omega := \sum_\alpha \int_{U_\alpha} \psi_\alpha \omega. \tag{15.6}$$

We say that ω is **integrable** if the right hand side converges absolutely for any such cover and any such partitions of unity. This is true, for example, if ω is compactly supported.

One need to check that the definition (15.6) is independent of the choices of orientation-compatible coordinate charts, and is independent of the choices of partition of unity.

Theorem 15.10. *Suppose ω is compactly supported, or more generally, ω is integrable. The expression (15.6) is independent of the choices of $\{U_\alpha\}$ and the choices of $\{\psi_\alpha\}$.*

Proof. We first show that (15.5) is well-defined. The argument is essentially the same as in the Euclidean case: if ω is supported in U, and if $\{x_\alpha^i\}$ and $\{x_\beta^j\}$ are two orientation-compatible coordinate systems on U, so that

$$\omega = f_\alpha dx_\alpha^1 \wedge \ldots \wedge dx_\alpha^m = f_\beta dx_\beta^1 \wedge \ldots \wedge dx_\beta^m,$$

then we want to prove

$$\int_{V_\alpha} f_\alpha dx_\alpha^1 \ldots dx_\alpha^m = \int_{V_\beta} f_\beta dx_\beta^1 \ldots dx_\beta^m.$$

This is true, because

$$dx_\beta^1 \wedge \ldots \wedge dx_\beta^m = \det(d\varphi_{\alpha\beta})dx_\alpha^1 \wedge \ldots \wedge dx_\alpha^m$$

implies $f_\alpha = \det(d\varphi_{\alpha\beta})f_\beta$. Since $\det(d\varphi_{\alpha\beta}) > 0$, the conclusion follows from the change of variable formula in $\mathbb{R}^n$.

To prove (15.6) is well-defined, we suppose $\{U_\alpha\}$ and $\{U_\beta\}$ are two locally finite cover of M consisting of orientation-compatible charts, and $\{\psi_\alpha\}$ and $\{\rho_\beta\}$ are partitions of unity subordinate to $\{U_\alpha\}$ and $\{U_\beta\}$ respectively. Then $\{U_\alpha \cap U_\beta\}$ is a new locally finite cover of M, and $\{\psi_\alpha\rho_\beta\}$ is a partition of unity subordinate to this new cover. It is enough to prove

$$\sum_\alpha \int_{U_\alpha} \psi_\alpha\omega = \sum_{\alpha,\beta} \int_{U_\alpha \cap U_\beta} \psi_\alpha\rho_\beta\omega.$$

This is true because for each fixed α,

$$\int_{U_\alpha} \psi_\alpha\omega = \int_{U_\alpha} \Big(\sum_\beta \rho_\beta\Big)\psi_\alpha\omega = \sum_\beta \int_{U_\alpha \cap U_\beta} \psi_\alpha\rho_\beta\omega.$$

□

15.5 Change of variable formula

Finally we extend the change of variable formula form $\mathbb{R}^m$ to manifolds.

Definition 15.11. Let M, N be orientable smooth n-manifolds, with orientations $\mathcal{A}$ and $\mathcal{B}$ respectively. A diffeomorphism $\varphi : M \to N$ is said to be **orientation-preserving** if for each $(\psi_\beta, X_\beta, Y_\beta) \in \mathcal{B}$, the chart $(\psi_\beta \circ \varphi, \varphi^{-1}(X_\beta), Y_\beta)$ on M is orientation compatible with $\mathcal{A}$.

Suppose M, N are connected. It is easy to see that a diffeomorphism $\varphi : M \to N$ is orientation-preserving if and only if there exists one chart $(\psi_\beta, X_\beta, Y_\beta) \in \mathcal{B}$, such that the chart $(\psi_\beta \circ \varphi, \varphi^{-1}(X_\beta), Y_\beta)$ on M is orientation compatible with $\mathcal{A}$. Similarly if there exists one chart $(\psi_\beta, X_\beta, Y_\beta) \in \mathcal{B}$, such that the chart $(\psi_\beta \circ \varphi, \varphi^{-1}(X_\beta), Y_\beta)$ on M is incompatible with $\mathcal{A}$, then for every chart $(\psi_\beta, X_\beta, Y_\beta) \in \mathcal{B}$, the chart $(\psi_\beta \circ \varphi, \varphi^{-1}(X_\beta), Y_\beta)$ on M is incompatible with $\mathcal{A}$. In this case we say φ is *orientation-reverting*.

Now we state:

Theorem 15.12 (The change of variable formula)**.** *Suppose M, N are n-dimensional orientable smooth manifolds, and $\varphi : M \to N$ is a diffeomorphism.*

1. *If φ is an orientation-preserving, then*
$$\int_M \varphi^*\omega = \int_N \omega.$$
2. *If φ is an orientation-reverting, then*
$$\int_M \varphi^*\omega = -\int_N \omega.$$

Proof. It is enough to prove this in local charts, in which case this is merely the change of variable formula in $\mathbb{R}^m$. ☐

Remark 15.13. If ω is a compactly supported k-form on M, where $k < m = \dim M$, then one cannot integrate ω over M. However, for any k-dimensional orientable submanifold $X \subset M$, one can define $\int_X \omega$ by setting it to be $\int_X \iota^*\omega$, where $\iota : X \hookrightarrow M$ is the inclusion map. By this way we get a "pairing" between k-forms on M and k-dimensional orientable submanifolds in M.

15.6 Volume form and volume measure

Next we show that orientability can be characterized via the existence of specific top forms:

Theorem 15.14. *An m-dimensional smooth manifold M is orientable if and only if M admits a nowhere vanishing smooth m-form μ.*

Proof. First let μ be a nowhere vanishing smooth m-form on M. Then on each local chart $(U, x^1, \ldots, x^m)$ (where U is always chosen to be connected), there is a smooth function $f \neq 0$ so that $\mu = f dx^1 \wedge \ldots \wedge dx^m$. It follows that
$$\mu(\partial_1, \ldots, \partial_m) = f \neq 0.$$
We can always take such a chart near each point so that $f > 0$, otherwise we can replace x^1 by $-x^1$. Now suppose $(U_\alpha, x^1_\alpha, \ldots, x^m_\alpha)$ and $(U_\beta, x^1_\beta, \ldots, x^m_\beta)$ be two such charts, so that on the intersection $U_\alpha \cap U_\beta$ one has
$$f dx^1_\alpha \wedge \ldots \wedge dx^m_\alpha = \mu = g dx^1_\beta \wedge \ldots \wedge dx^m_\beta = g \det(d\varphi_{\alpha\beta}) dx^1_\alpha \wedge \ldots \wedge dx^m_\alpha.$$
where $f, g > 0$. It follows that $\det(d\varphi_{\alpha\beta}) > 0$. So the atlas constructed by this way is an orientation.

Conversely, suppose $\mathcal{A}$ is an orientation. For each local chart U_α in $\mathcal{A}$, we let
$$\mu_\alpha = dx^1_\alpha \wedge \ldots \wedge dx^m_\alpha.$$

Pick a partition of unity $\{\psi_\alpha\}$ subordinate to the open cover $\{U_\alpha\}$. We claim that

$$\mu := \sum_\alpha \psi_\alpha \mu_\alpha,$$

is a nowhere vanishing smooth m-form on M. In fact, for each $p \in M$, there is a neighbourhood U of p so that the sum $\sum_\alpha \psi_\alpha \mu_\alpha$ is a finite sum $\sum_{i=1}^k \psi_i \mu_i$. It follows that near p,

$$\mu(\partial_1^1, \dots, \partial_m^1) = \sum_{i=1}^k (\det d\varphi_{1i})\, \psi_i > 0.$$

So $\mu \neq 0$ near p. $\square$

Definition 15.15. A nowhere vanishing smooth m-form μ on an m-dimensional smooth manifold M is called a *volume form*.

Thus, we can always integrate top forms and its multiples on manifolds admitting volume forms.

Remark 15.16. If M is orientable, and μ is a volume form, then the two orientations of M are represented by μ and $-\mu$ respectively. We denote the two orientations by $[\mu]$ and $[-\mu]$.

Remark 15.17. In particular, on any Lie group, one can define conceptions like left-invariant differential forms. Since any Lie group is orientable (Exercise), there exists left-invariant volume form on any Lie group G. The measures associated to left-invariant volume form on Lie groups are called **Haar measures**.

16. Stokes's theorem

16.1 Orientations on manifolds with boundary

We start with our discussion on manifold with boundary. It's almost the same concept as that of a manifold but now we allow points in the "boundary" as well which will be described as those points such that any neighbourhood looks like a part of "half-space". Let

$$\mathbb{R}^m_+ = \{(x^1, \dots, x^m) \mid x^m \geq 0\}.$$

We have the following.

Definition 16.1. A topological space M is called an m-dimensional **topological manifold with boundary** if it is Hausdorff, second-countable, and for any $p \in M$, there is a neighborhood U of p which is homeomorphic to either $\mathbb{R}^m$ or $\mathbb{R}^m_+$.
We define the boundary of M to be

$$\partial M = \{p \in M \mid p \text{ has no neighborhood in } M \text{ that is homeomorphic to } \mathbb{R}^m\}.$$

Thus, $p \in \partial M$ if and only if p has a neighborhood that is homeomorphic to $\mathbb{R}^m_+$ such that p gets mapped onto $\partial\mathbb{R}^m_+ = \{(x^1, \dots, x^m) \mid x^m = 0\}$. We call

$$\text{int}(M) = M \setminus \partial M$$

the interior of M.

Example 16.2. The closed ball

$$B^m(1) = \{x \in \mathbb{R}^m \mid |x| \leq 1\}$$

is a smooth manifold with boundary, and its boundary is $\partial B^m(1) = S^{m-1}$.

Lemma 16.3. *If M^m is a smooth manifold then the smooth structure on M induces a smooth structure on ∂M making it an $(m-1)$-dimensional smooth manifold.*

Proof. Let $p \in \partial M$ and let $(U, x^1, \dots, x^m)$ be a chart coming from the smooth structure on M which is homeomorphic to an open subset of $\mathbb{R}^m_+$. Then

$$U \cap \partial M = \{(x^1, \dots, x^m) \mid x^m = 0\},$$

and hence $(U \cap \partial M, x^1, \dots x^{m-1})$ is a chart on ∂M. □

Now let M be an orientable smooth manifold with boundary, and let $[\mu]$ be an orientation on M. This means that $\mu \in \Omega^m(M)$ is nowhere vanishing. We obtain an induced orientation on ∂M. The induced orientation on ∂M is defined as follows. Suppose locally near the boundary, we have

$$\mu = f dx^1 \wedge \cdots \wedge dx^m$$

with $f > 0$. Then, locally, we define the induced orientation on ∂M to be the one that is represented by the differential form

$$\eta = (-1)^m dx^1 \wedge \cdots \wedge dx^{m-1}. \tag{16.1}$$

By choosing a partitions of unity, we can glue these local $(m-1)$-forms on ∂M into a global $(m-1)$-form on ∂M which is nowhere vanishing. The nowhere vanishing property of η follows from the nowhere vanishing property of μ. Thus, from Theorem 15.14, we get a volume form that defines an orientation on ∂M, which is induced from an orientation on M.

We explain why η in (16.1) has that form. Note that in coordinate charts near boundary, M is defined by $x^m \geq 0$. So $-x^m$ is the "outward pointing normal" of M, and this boundary orientation is chosen so that

$$d(-x^m) \wedge \eta = d(-x^m) \wedge (-1)^m dx^1 \wedge \cdots \wedge dx^{m-1} = dx^1 \wedge \cdots \wedge dx^m.$$

16.2 Stokes's theorem

Now we can state and prove our main theorem:

Theorem 16.4 (Stokes's theorem). *Let M be a smooth oriented m-dimensional manifold with boundary ∂M with the induced orientation. For any $\omega \in \Omega^{m-1}(M)$ with compact support, we have*

$$\int_{\partial M} \iota_{\partial M}^* \omega = \int_M d\omega, \tag{16.2}$$

where $\iota_{\partial M} : \partial M \hookrightarrow M$ is the inclusion map.

Proof. The proof follows from the following three cases which we prove one by one.
(1) ω is supported in a chart U that is diffeomorphic to $\mathbb{R}^m$.
(2) ω is supported in a chart U that is diffeomorphic to $\mathbb{R}^m_+$.
(3) The general case.

Case (1): Since $\omega = 0$ on ∂M, we have $\int_{\partial M} \iota_{\partial M}^* \omega = 0$. To calculate $\int_M d\omega$, we denote

$$\omega = \sum_i (-1)^{i-1} f_i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^m,$$

where f_i's are compactly supported smooth functions. Then, using the formula for the exterior derivative, we have

$$d\omega = \sum_i \frac{\partial f_i}{\partial x^i} dx^1 \wedge \cdots \wedge dx^m.$$

So by definition,

$$\begin{aligned}
\int_M d\omega &= \int_{\mathbb{R}^m} \sum_i \frac{\partial f_i}{\partial x^i} dx^1 \cdots dx^m \\
&\underbrace{=}_{\text{Fubini's theorem}} \sum_i \int_{\mathbb{R}^{m-1}} \left(\int_{-\infty}^{\infty} \frac{\partial f_i}{\partial x^i} dx^i \right) dx^1 \cdots \widehat{dx^i} \cdots dx^m \\
&\underbrace{=}_{\text{FTC}} \sum_i \int_{\mathbb{R}^{m-1}} \left(\lim_{s \to \infty} f_i(s) - \lim_{s \to -\infty} f_i(s) \right) dx^1 \cdots \widehat{dx^i} \cdots dx^m \\
&= 0,
\end{aligned}$$

as f is compactly supported. Thus, (16.2) is true in Case 1.

Case (2): We have the same formula to calculate $\int_M d\omega$, except for the $i = m$ term, because in this case one of the limits is 0 instead of $-\infty$. We have

$$\int_M d\omega = \int_{\mathbb{R}^{m-1}} \left(\int_0^{\infty} \frac{\partial f_m}{\partial x^m} dx^m \right) dx^1 \cdots dx^{m-1} = - \int_{\mathbb{R}^{m-1}} f_m(x^1, \ldots, x^{m-1}, 0) dx^1 \cdots dx^{m-1}.$$

On the other hand, since $x^m = 0$ on ∂M, we see

$$\iota_{\partial M}^* \omega = (-1)^{m-1} f_m(x^1, \ldots, x^{m-1}, 0) dx^1 \wedge \cdots \wedge dx^{m-1} = -f_m(x^1, \ldots, x^{m-1}, 0) \cdot (-1)^m dx^1 \wedge \cdots \wedge dx^{m-1}.$$

So

$$\begin{aligned}
\int_{\partial M} \iota_{\partial M}^* \omega &= \int_{\mathbb{R}^{m-1}} (-f_m(x^1, \ldots, x^{m-1}, 0)) dx^1 \cdots dx^{m-1} \\
&= - \int_{\mathbb{R}^{m-1}} f_m(x^1, \ldots, x^{m-1}, 0) dx^1 \cdots dx^{m-1} \\
&= \int_M d\omega,
\end{aligned}$$

which is (16.2).

Case (3): In general, we cover the compact set supp(ω) by finitely many coordinate charts, and take a partition of unity $\{\psi_i\}$ subordinate to that cover. Then

$$\int_{\partial M} \iota^*_{\partial M}\omega = \sum_i \int_{\partial M} \iota^*_{\partial M}(\psi_i\omega) \underbrace{=}_{\text{Stokes's thm. in a chart}} \sum_i \int_M d(\psi_i\omega) = \sum_i \int_M d\psi_i \wedge \omega + \int_M d\omega = \int_M d\omega,$$

as

$$\sum_i \int_M d\psi_i \wedge \omega = \int_M d(\sum_i \psi_i) \wedge \omega = 0.$$

□

We see some simple consequences.

1. Let $M = [a, b]$ and $\omega = f \in C^\infty(M) = \Omega^0(M)$. Then

$$\int_M df = \int_a^b \frac{df}{dx} dx = \int_{\partial[a,b]} f = f(b) - f(a).$$

2. If $\partial M = \emptyset$ then for any compactly supported $m-1$-form ω we have $\int_M d\omega = \int_{\partial M} \omega = 0$. For instance, if $f : S^1 \to \mathbb{R}$ is a smooth function then $\int_{S^1} df = 0$.

3. **Green's theorem** Let D be a compact domain in $\mathbb{R}^2$ and $P, Q \in C^\infty(D)$. Then

$$\int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dxdy \right) = \int_{\partial D} Pdx + Qdy.$$

References

[Bär] Christian Bär. *Lecture notes on Differential Geometry*. URL: https://www.math.uni-potsdam.de/fileadmin/user_upload/Prof-Geometrie/Dokumente/Lehre/Lehrmaterialien/DiffGeo.pdf (cit. on p. 1).

[Bau] Helga Baum. *Differentialgeometrie*. URL: https://www.mathematik.hu-berlin.de/%7Ebaum/Skript/DG-05-06.15.pdf (cit. on p. 1).

[Ker60] Michel A. Kervaire. "A manifold which does not admit any differentiable structure". English. *Comment. Math. Helv.* 34 (1960), pp. 257–270. Zbl: 0145.20304. URL: https://eudml.org/doc/139196 (cit. on p. 12).

[Lee13] John M. Lee. *Introduction to smooth manifolds*. English. 2nd revised ed. Vol. 218. Grad. Texts Math. New York, NY: Springer, 2013. Zbl: 1258.53002 (cit. on p. 1).

[Tu11] Loring W. Tu. *An introduction to manifolds. 2nd revised ed.* English. 2nd ed. Universitext. New York, NY: Springer, 2011. Zbl: 1200.58001 (cit. on p. 1).

[Wal] Thomas Walpuski. *Lecture notes on Differential Geometry*. URL: https://walpu.ski/Teaching/WS2021/DifferentialGeometry/DifferentialGeometry.pdf (cit. on p. 1).

[Wen] Chris Wendl. *Differential Geometry 12*. URL: https://www.mathematik.hu-berlin.de/~wendl/Sommer2022/Diffgeo2/lecturenotes.pdf (cit. on p. 1).