Differential Manifolds

Yuchen Bi

April 2026

Contents

Chapter 1

Topological Manifolds

These notes begin with a brief review of topology and the basic structural properties of topological manifolds that will be used throughout the text.

We start by recalling a few standard notions from topology.

1.1 Topological preliminaries

1.1.1 Topological spaces

We start with the notion of a topological space.

Definition 1.1 (Topological space)**.** A *topological space* is a pair $(X, \mathcal{O}_X)$, where X is a set , and $\mathcal{O}_X \subset 2^X$ is a set of subsets of X (equivalently, subset of power set of X) satisfies:

- $\emptyset, X \in \mathcal{O}_X$.
- $A, B \in \mathcal{O}_X \implies A \cap B \in \mathcal{O}_X$.
- $A_i \in \mathcal{O}_X, i \in I \implies \bigcup_{i \in I} A_i \in \mathcal{O}_X$.

We say A is an *open set* of X if and only if $A \in \mathcal{O}_X$.

Topology defines open sets to be stable under finite intersections and arbitrary unions. It is natural to ask if this stability extends to infinite intersections. To build intuition, consider the real line. The sets $A_n = (-1/n, 1/n)$ are all open intervals. What is $\bigcap_{n=1}^{\infty} A_n$? It is the single point 0, which is not an open set. This shows that openness is not preserved under infinite intersections.

Example 1.2. Given a set X, there are two natural topologies that can be defined on it:

Trivial topology $\mathcal{O}_X := \{\emptyset, X\}$.

Discrete topology $\mathcal{O}_X := 2^X$.

We now introduce two fundamental concepts related to open sets.

1. A subset $A \subset X$ is said to be *closed* if and only if its complement $X - A$ is open.

2. A subset $U \subset X$ is called a *neighborhood* of a point $p \in X$ if there exists an open set A such that $p \in A \subset U$.

The definition of a topological space is highly general. It is built on only three axioms, making it applicable in a wide range of contexts. As a result, there exist certain highly unconventional topologies that defy geometric intuition. A notable example of such a topology can be found in Furstenberg's proof of the infinitude of primes.

In these notes, we focus on well-behaved topological spaces. Arguably the most important class is that arising from metric spaces.

Example 1.3. Let (X, d) be a metric space. That is, X is a set, and $d : X \times X \to \mathbb{R}^+$ is a non-negative function satisfying:

positivity $d(x, y) = 0 \iff x = y$.

symmetry $d(x, y) = d(y, x)$.

triangle inequality $d(x, z) \leq d(x, y) + d(y, z)$.

Then we define

$$\mathcal{O}_X := \{A \subset X : \forall p \in A, \exists r > 0 \text{ such that } B(p, r) \subset A\}.$$

It is straightforward to verify $(X, \mathcal{O}_X)$ is a topological space.

Thus, $\mathbb{R}^n$ carries a natural topology induced by the Euclidean metric. When referring to $\mathbb{R}^n$ as a topological space, we implicitly consider this metric-induced topology, and the notation $\mathcal{O}_{\mathbb{R}^n}$ is often omitted for simplicity.

It is a good point to review another concept, which is called a *basis* for a topology.

Definition 1.4. Assume $(X, \mathcal{O}_X)$ is a topological space. Then $\mathcal{B} \subset 2^X$ is called a *basis* for this topological space if

$$\begin{aligned} A \in \mathcal{O}_X &\iff A \text{ is a union of elements in } \mathcal{B}. \\ &\iff \forall p \in A, \exists B \in \mathcal{B} \text{ such that } p \in B \subset A. \end{aligned}$$

Example 1.5.

1. $\mathcal{B} = \{B(x, r) \subset \mathbb{R}^n : x \in \mathbb{R}^n, r > 0\}$ is a basis for $\mathbb{R}^n$.
2. $\mathcal{B} = \{B(x, r) \subset \mathbb{R}^n : x \in \mathbb{Q}^n, r \in \mathbb{Q}, r > 0\}$ is also a basis for $\mathbb{R}^n$.

The previous example shows that $\mathbb{R}^n$ has a countable basis. This gives us a way to say that $\mathbb{R}^n$ is not too large. In general, we have the following definition.

Definition 1.6 (Second countable)**.** If the topological space $(X, \mathcal{O}_X)$ has a countable basis, we say X is *second countable.*

1.1.2 Continuous maps

Let us now discuss maps between topological spaces. It is convenient to simplify our notation at this stage. We will often write X to represent the entire topological space $(X, \mathcal{O}_X)$, implicitly assuming that a topology $\mathcal{O}_X$ (the family of open sets) has been assigned. This abbreviation is harmless, as the relevant topology will always be clear from the context.

Now, let $X = (X, \mathcal{O}_X)$, $Y = (Y, \mathcal{O}_Y)$ be topological spaces.A central idea in topology is that specifying a topology on a space allows us to define what continuity means. In fact, this is one of the primary motivations behind the very definition of a topological space.

Definition 1.7. A map $\varphi : X \to Y$ is said to be *continuous* if for any open set $B \subset Y$, the preimage $f^{-1}(B) \subset X$ is open.

It is essential to note that continuity is defined in terms of the preimage of open sets being open—not the image. For instance, a constant map $\mathbb{R} \to \mathbb{R}$ is continuous, even though the image of any non-empty open set is a single point, which is not open in $\mathbb{R}$.

Example 1.8. Let $X = Y = \mathbb{R}$, equipped with the standard topology. In this case, the topological definition of continuity given above is equivalent to the classical $\varepsilon - \delta$ definition from analysis.

The following defines when two topological spaces are considered "the same."

Definition 1.9. A continuous map $\varphi : X \to Y$ is called *homeomorphism* if φ is bijective, and both φ, φ^{-1} are continuous.

An important warning is necessary here: a continuous bijective map $\varphi \colon X \to Y$ is not necessarily a homeomorphism. In other words, the continuity of the inverse is not automatic. This phenomenon does occur, though in relatively special situations. One elementary example is given in Item 5 of Fact 1.20. A deeper example is provided by Brouwer's theorem on invariance of domain.

1.1.3 Subspace topology

Let X be a topological space and $Y \subset X$ a subset. We can naturally equip Y with a topology, called the *subspace topology*, by declaring the open sets in Y to be exactly those of the form

$$\mathcal{O}_Y = \{A \cap Y : A \in \mathcal{O}_X\}.$$

Whenever we refer to a map defined on Y—even without explicitly specifying its topology—we always assume that Y is endowed with this subspace topology. In particular, continuity of such a map is understood with respect to this induced topology.

It is important to understand (or at least feel comfortable with) the following two facts:

Fact 1.10. *If $f : X \to Z$ is continuous, then its restriction $f|_Y : Y \to Z$ is also continuous with respect to the subspace topology on Y.*

Fact 1.11. *The subspace topology is the coarsest topology on Y (i.e., the one with the fewest open sets) for which the inclusion map $Y \hookrightarrow X$ is continuous.*

Example 1.12. Consider the inclusion map $\iota : \mathbb{R} \to \mathbb{R}^2$ defined by $x \mapsto (x, 0)$, which embeds the real line as the x-axis in the plane. Then the subspace topology on $\mathbb{R}$ induced by this inclusion coincides with the standard topology on $\mathbb{R}$.

1.2 Topological manifolds

A topic of particular interest in this course concerns topological spaces that, in some sense, "resemble" $\mathbb{R}^n$. Indeed, $\mathbb{R}^n$ serves as a key model throughout our study. In this context, it is worth mentioning a theorem that, while important to know, is not central to our development—the so-called *topological invariance of dimension*.

Theorem 1.13 (Topological invariance of dimension)**.** *Let $\varphi\colon U \to V$ be a homeomorphism between nonempty open sets $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$. Then $m = n$.*

This tells us that, on a topological level, $\mathbb{R}^m$ and $\mathbb{R}^n$ are fundamentally different when $m \neq n$. While this may align with our intuition, the proof is surprisingly deep and lies beyond the scope of this course—it typically makes use of tools such as homology groups.

Let us now return to the idea of spaces that look like $\mathbb{R}^n$.

Definition 1.14 (Locally Euclidean)**.** A topological space X is called *locally Euclidean* at a point $p \in X$ of dimension n if there exists an open neighborhood $p \in U \subset X$ that is homeomorphic to some open subset of $\mathbb{R}^n$. That is, there exist an open set $\widetilde{U} \subset \mathbb{R}^n$ and a homeomorphism $\varphi\colon U \to \widetilde{U}$.

This means that, locally, the topology around p is indistinguishable from that of a Euclidean space. This is a very special property, and one we will encounter frequently. As a useful exercise, we note the following:

Exercise 1.15. One may always take $\widetilde{U} = B(0,1) \subset \mathbb{R}^n$ in the above definition. (Hint: use a dilation to identify $\mathbb{R}^n$ with the unit ball, and restrict the original homeomorphism accordingly.)

You might wonder why this property is so interesting. Consider, for example, the universe—or the surface of the Earth—as a topological space. We may not know its global structure (indeed, ancient people did not know the Earth was a sphere until they could observe it from outside), but we do know that locally, it appears Euclidean: our universe looks like $\mathbb{R}^3$ in small regions, and the Earth's surface looks like $\mathbb{R}^2$. Thus, while the local topology is trivial, the global structure can be far more interesting.

Here is a lemma whose proof is worth understanding, as the underlying argument will be used repeatedly.

Lemma 1.16. *The dimension n in the definition of a locally Euclidean space is uniquely determined at each point p.*

Proof. Suppose X is locally Euclidean at p with two possible dimensions n_1 and n_2. Then there exist open neighborhoods $p \in U_i \subset X$ and homeomorphisms $\varphi_i\colon U_i \to \widetilde{U}_i \subset \mathbb{R}^{n_i}$ for $i = 1, 2$. Since $p \in U_1 \cap U_2$, the intersection is nonempty. Now consider the composition

$$\varphi_2 \circ \varphi_1^{-1}\colon \varphi_1(U_1 \cap U_2) \subset \mathbb{R}^{n_1} \to \varphi_2(U_1 \cap U_2) \subset \mathbb{R}^{n_2},$$

which is a homeomorphism between nonempty open subsets of $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$. By the topological invariance of dimension, it follows that $n_1 = n_2$. $\square$

To further develop the language of manifolds, we introduce several key topological properties. The first is the following:

Definition 1.17. A topological space X is called *Hausdorff* if for any two distinct points $p, q \in X$, there exist open sets $U, V \subset X$ such that $p \in U$, $q \in V$, and $U \cap V = \emptyset$.

Example 1.18.
1. Any metric space is Hausdorff in its induced topology.
2. If X is *separated by continuous functions* in the sense that for every pair $p \neq q$ there exists a continuous $f : X \to \mathbb{R}$ with $f(p) \neq f(q)$, then X is Hausdorff.

Next, we recall the important notion of compactness.

Definition 1.19. A subset $K \subset X$ is called *compact* if every open cover $\{A_i\}_{i \in I}$ of K-that is, $K \subset \bigcup_{i \in I} A_i$-admits a finite subcover $\{A_i\}_{i \in I'}$ satisfying

$$K \subset \bigcup_{i \in I'} A_i, \quad I' \subset I, \quad |I'| < \infty.$$

If X itself is compact, we call X a *compact space.*

Although compactness is not always included in the definition of manifolds, it interacts fruitfully with the Hausdorff property. Moreover, Euclidean space $\mathbb{R}^n$ is locally compact—meaning that every point has a compact neighborhood inside any prescribed neighborhood—and we will later see that all manifolds inherit this property.

The following collection of basic topological facts will be useful throughout this course.

Fact 1.20.
1. *If X is Hausdorff and $Y \subset X$, then Y is Hausdorff under the subspace topology.*
2. *In a Hausdorff space X, every compact subset $K \subset X$ is closed.*
3. *If X is compact and $K \subset X$ is closed, then K is compact.*
4. *The continuous image of a compact set is compact: if $\varphi : X \to Y$ is continuous and $K \subset X$ is compact, then $\varphi(K)$ is compact.*
5. *If X is compact, Y is Hausdorff, and $\varphi : X \to Y$ is a continuous bijection, then φ is a homeomorphism.*

These facts complete our review of the necessary topological background. We now turn to the central object of interest in this course: manifolds.

Definition 1.21. A topological space M is called an *n-dimensional topological manifold* if it satisfies the following three conditions:

1. M is locally Euclidean of dimension n,
2. M is Hausdorff,
3. M is second countable.

The requirement that M be locally Euclidean is fairly intuitive—without it, the space could exhibit pathological local behavior. It is natural to ask, however, why we also impose the Hausdorff and second countable properties. To illustrate their importance, we now examine what can go wrong if either of these conditions is omitted.

Non-example 1.22 (Dropping the Hausdorff condition)**.** *Define* $X := (\mathbb{R} \times \{0, 1\})/\sim$ *as the quotient of two real lines, where* $(x, 0) \sim (x, 1)$ *for all* $x < 0$*. A subset* $A \subset X$ *is open if and only if its preimage under the quotient map* π *is open in* $\mathbb{R} \times \{0, 1\}$*. One may verify that the sets* $\{[(x, 0)] : x \in \mathbb{R}\} \cong \mathbb{R}$ *and* $\{[(x, 1)] : x \in \mathbb{R}\} \cong \mathbb{R}$ *are open in* X*. Thus,* X *is locally Euclidean. However,* X *is not Hausdorff: the two distinct points* $[(0, 0)]$ *and* $[(0, 1)]$ *do not admit disjoint neighborhoods.*

In essence, the Hausdorff condition prevents such pathological "branching" behavior, which is undesirable in a well-behaved geometric object.

Let us now examine the second countable condition.

Non-example 1.23 (Dropping the second countable condition)**.** *Define* $X := \mathbb{R}^2$ *with the topology given by*

$$\mathcal{O}_X = \{U \times \{y\} : U \in \mathcal{O}_{\mathbb{R}},\ y \in \mathbb{R}\}.$$

Then $(X, \mathcal{O}_X)$ *is locally Euclidean of dimension* 1 *at every point. However, this space fails to be second countable and has uncountably many connected components.*

Here, the second countable condition serves to exclude spaces with too many components-in particular, uncountably many.

A more sophisticated example is the long line, which also violates second countability. Its construction is somewhat involved and relies on set-theoretic concepts such as well-ordered sets. Interested readers may refer to the linked Wikipedia entry for further details.

Now we give some examples of manifolds.

Example 1.24. If M is a 0-dimensional topological manifold, then M is a finite or countable set equipped with the discrete topology.

Example 1.25. If M^n is a topological manifold and $M' \subset M^n$ is an open subset, then M' is also an n-dimensional topological manifold. For instance, any open subset of $\mathbb{R}^n$ is a manifold.

Example 1.26.

$$S^1 := \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}.$$

1.3 Connectivity

Let us recall some basic terminology regarding connectedness. There are essentially two fundamental ways to characterize when a topological space is connected.

Definition via Clopen Sets

The more basic definition states that a topological space X is *connected* if the only subsets that are both open and closed (clopen) are $\emptyset$ and X itself.

Why is this definition meaningful? Suppose there exists a non-trivial clopen subset $A \subset X$ with $A \neq \emptyset$ and $A \neq X$. Then its complement $X \setminus A$ is also clopen. Consequently, the topology of X decomposes into disjoint unions of open sets from A and $X \setminus A$, indicating that X can be separated into two independent components. Thus, connectedness precisely prohibits such a separation.

Path Connectedness

In practice, a more operational notion is the stronger concept of *path connectedness*. A space X is path connected if for any two points $p, q \in X$, there exists a continuous path $\gamma : [0,1] \to X$ such that $\gamma(0) = p$ and $\gamma(1) = q$.

It is straightforward to verify that path connectedness implies connectedness. However, the converse is generally false. A classic counterexample is the topologist's sine curve. Notably, for topological manifolds, these two notions coincide.

Theorem 1.27. *Let M^n be a topological manifold. Then*

$$M \text{ is connected } \iff M \text{ is path connected.}$$

Proof. ($\Longleftarrow$) This direction is straightforward and will be left as an exercise.

($\Longrightarrow$) We now prove the non-trivial direction. Assume M is connected and we want to show it is path connected. To understand the proof strategy, it is helpful to first consider the case when M is an open subset of $\mathbb{R}^n$, a proof that may be familiar from analysis or topology courses.

Case: $M \subset \mathbb{R}^n$ open. Fix a base point $p \in M$. Define the set

$$U = \{q \in M : \exists \text{ a continuous path from } p \text{ to } q \text{ within } M\}.$$

Our goal is to show that $U = M$.

First, we show U is open. Take any $q \in U$. Since M is open, there exists an open ball $B_\varepsilon(q) \subset M$. For any $r \in B_\varepsilon(q)$, we can connect q to r by a straight line segment (which is continuous) contained entirely in $B_\varepsilon(q) \subset M$. By composing this with the path from p to q, we obtain a path from p to r. Hence, $B_\varepsilon(q) \subset U$, proving U is open.

Next, we show U is closed in M. Suppose $q \in M \setminus U$. Again, by openness of M, there exists an open ball $B_\varepsilon(q) \subset M$. We claim $B_\varepsilon(q) \subset M \setminus U$. If not, there would exist some $r \in B_\varepsilon(q) \cap U$. But then we could connect p to r (since $r \in U$) and then connect r to q by a straight line segment within $B_\varepsilon(q)$, contradicting $q \notin U$. Thus, $M \setminus U$ is open, so U is closed.

Since M is connected and U is non-empty (as $p \in U$), we conclude $U = M$.

General case: M is a topological manifold. The proof follows the same strategy. Fix $p \in M$ and define U as above.

To show U is open: For $q \in U$, take a coordinate chart (V, φ) with $q \in V$ and $\varphi(V)$ homeomorphic to an open subset of $\mathbb{R}^n$. Since $\varphi(V)$ contains an open ball around $\varphi(q)$, we can find a neighborhood $W \subset V$ of q that is homeomorphic to an open ball. Any point in W can be connected to q by a path in W (via the straight line in coordinates), hence can be connected to p by extending the existing path. Thus, $W \subset U$.

To show U is closed: Suppose $q \in M \setminus U$. Take a coordinate chart (V, φ) around q with $\varphi(V)$ homeomorphic to an open ball. If there were some $r \in V \cap U$, we could connect p to r and then r to q within V (using the coordinate representation), contradicting $q \notin U$. Hence, $V \subset M \setminus U$, so $M \setminus U$ is open.

By connectedness of M and non-emptiness of U, we conclude $U = M$. □

1.4 Local compactness and paracompactness

We now turn to slightly more technical aspects of topological manifolds. We will see how the second countability condition translates into more practical tools for working with manifolds.

1.4.1 Local Compactness

Proposition 1.28. *Let M be a topological manifold. Then for every point $p \in M$ and every open neighborhood U of p, there exists a compact neighborhood K of p such that $K \subset U$.*

Proof. Since M is a topological manifold, it is locally Euclidean. Hence, there exists a coordinate chart (V, φ) such that:

- $p \in V \subset U$,
- $\varphi(V) \subset \mathbb{R}^n$ is open.

Choose an open ball $B_r(\varphi(p)) \subset \varphi(V)$ centered at $\varphi(p)$. Now, consider the closed ball $\overline{B}_{r/2}(\varphi(p)) \subset B_r(\varphi(p))$ and define:

$$K = \varphi^{-1}\left(\overline{B}_{r/2}(\varphi(p))\right).$$

We verify the required properties:

1. K **is compact**: The closed ball $\overline{B}_{r/2}(\varphi(p))$ is compact in $\mathbb{R}^n$. Since φ is a homeomorphism onto its image, K is compact in M.
2. K **is a neighborhood of** p: The open ball $B_{r/2}(\varphi(p))$ is contained in $\overline{B}_{r/2}(\varphi(p))$, so

$$\varphi^{-1}(B_{r/2}(\varphi(p))) \subset K.$$

   This set is open in M and contains p, so K is a neighborhood of p.
3. $K \subset U$: Since $K \subset V \subset U$, the result follows.

Thus, K is a compact neighborhood of p contained in U. □

This property shows that M is locally compact according to the definition:

Definition 1.29. A topological space X is *locally compact* if for every point $p \in X$ and every open neighborhood U of p, there exists a compact neighborhood K of p such that $p \in K \subset U$.

1.4.2 Exhaustion by Compact Sets

A crucial consequence of second countability and local compactness is the existence of exhaustion by compact sets:

Definition 1.30. A topological space X admits an *exhaustion by compact sets* if there exists a sequence $\{K_n\}_{n=1}^{\infty}$ of compact subsets such that:

1. $K_n \subset \text{int}(K_{n+1})$ for all $n \geq 1$,
2. $\bigcup_{n=1}^{\infty} K_n = X$.

Remark 1.31. For any exhaustion $\{K_i\}$ of X, we have:

$$X = \bigcup_{i=1}^{\infty} \text{Int}(K_i).$$

This has a useful consequence: any compact subset $K' \subset X$ is eventually contained in the interiors of the K_i's. Specifically, since $\{\mathrm{Int}(K_i)\}_{i=1}^{\infty}$ is an open cover of K' and K' is compact, there exists a finite subcover. Due to the nesting property $K_i \subset \mathrm{Int}(K_{i+1})$, this implies:

$$K' \subset \mathrm{Int}(K_i) \quad \text{for all } i \geq i_0(K').$$

In other words, any compact subset of X is eventually "absorbed" by the interiors of the exhaustion sets.

1.4.3 Construction of Compact Exhaustion

Proposition 1.32. *Let X be a second countable, locally compact, Hausdorff topological space. Then X admits an exhaustion by compact subsets.*

This result highlights one of the key motivations for assuming second countability in the definition of topological manifolds.

Proof. We construct the exhaustion explicitly. Let $\mathcal{B}$ be a countable basis for the topology of X. Define the subcollection:

$$\mathcal{B}' = \{B \in \mathcal{B} : \overline{B} \text{ is compact}\}.$$

By local compactness and Hausdorff condition, $\mathcal{B}'$ remains a basis for X. Actually, Let $p \in X$ be an arbitrary point and $A \in \mathcal{O}_X$ be any open neighborhood of p. By local compactness, there exists an open set A' and a compact set K such that $p \in A' \subset K \subset A$.

Since $\mathcal{B}$ is a basis, there exists some $B \in \mathcal{B}$ with $p \in B \subset A'$. Now, using the Hausdorff property, we observe that K is closed in X (as compact subsets of Hausdorff spaces are closed). Therefore, we have:

$$\overline{B} \subset \overline{A'} \subset K.$$

Since K is compact and $\overline{B}$ is a closed subset of K, it follows that $\overline{B}$ is compact. This shows that $B \in \mathcal{B}'$, and we have found $B \in \mathcal{B}'$ with $p \in B \subset A$, proving that $\mathcal{B}'$ is indeed a basis for X.

Enumerate the elements of $\mathcal{B}'$ as $U_1, U_2, U_3, \ldots$

We now construct the exhaustion recursively:

Base case: Let $K_1 = \overline{U_1}$.

Inductive step: Suppose K_n has been constructed. Since K_n is compact and $\{U_i\}$ is an open cover of X, there exists $m_n > n$ such that:

$$K_n \subset U_1 \cup U_2 \cup \cdots \cup U_{m_n}.$$

Define:

$$K_{n+1} = \overline{U_1} \cup \overline{U_2} \cup \cdots \cup \overline{U_{m_n}}.$$

By construction, each K_n is compact (finite union of compact sets), $K_n \subset \mathrm{int}(K_{n+1})$ (since K_n is contained in the union of the U_i's, which are open subsets of K_{n+1}), and $\bigcup_{n=1}^{\infty} K_n = X$ (since $\{U_i\}$ covers X). □

1.4.4 Paracompactness

We now introduce a fundamental topological property that will play a crucial role in our study of manifolds.

1.4.5 Basic Definitions

Definition 1.33. Let X be a topological space.

1. An *open cover* of X is a collection $\mathcal{U} \subset \mathcal{O}_X$ of open sets such that $X = \bigcup_{U \in \mathcal{U}} U$.
2. An open cover $\mathcal{U}$ is called *locally finite* if every point $p \in X$ has a neighborhood W that intersects only finitely many $U \in \mathcal{U}$.
3. An open cover $\mathcal{V}$ is a *refinement* of an open cover $\mathcal{U}$ if for every $V \in \mathcal{V}$, there exists $U \in \mathcal{U}$ such that $V \subset U$.
4. X is *paracompact* if every open cover has a locally finite refinement.

Remark 1.34. Paracompactness serves to "tame" potentially complicated open covers. In the context of manifolds, we frequently work with collections of coordinate charts covering the space. Although there may be infinitely many charts, paracompactness guarantees the existence of a refinement where only finitely many charts interact at any given point. This localization property makes many local arguments feasible and is essential for various global constructions.

1.4.6 Paracompactness of Topological Manifolds

Theorem 1.35. *Every topological manifold is paracompact.*

Remark 1.36. The primary application of paracompactness in differential geometry is the construction of *partitions of unity*. These are families of smooth functions that sum to 1 everywhere while maintaining local finiteness (only finitely many are nonzero in any compact neighborhood). Partitions of unity enable us to glue locally defined objects into global ones. We will explore this in detail in subsequent lectures.

Proof. The proof proceeds by constructing a locally finite refinement using the compact exhaustion of M. Let $\{K_i\}_{i=1}^{\infty}$ be an exhaustion by compact subsets with $K_i \subset \text{int}(K_{i+1})$ for all i.

Define the compact annuli:

$$N_i = K_i \setminus \text{int}(K_{i-1}), \quad \text{with } K_0 = \emptyset.$$

These sets cover M since $\bigcup_{i=1}^{\infty} K_i = M$ and $K_i \subset \text{int}(K_{i+1})$.

Now define open neighborhoods for these annuli:

$$\widetilde{N}_i = \text{int}(K_{i+1}) \setminus K_{i-2},$$

where we take $K_{-1} = K_0 = \emptyset$. Note that $N_i \subset \widetilde{N}_i$ for each i.

Let $\mathcal{U}$ be an arbitrary open cover of M. For each i, the compact set N_i can be covered by finitely many elements of $\mathcal{U}$. Let $\mathcal{U}_i$ be such a finite subcover for N_i.

Define the refinement:

$$\mathcal{V}_i = \{\widetilde{N}_i \cap U : U \in \mathcal{U}_i\},$$

and let $\mathcal{V} = \bigcup_{i=1}^{\infty} \mathcal{V}_i$.

We now verify that $\mathcal{V}$ is a locally finite refinement of $\mathcal{U}$:

1. **Refinement:** Each $V \in \mathcal{V}$ is of the form $\widetilde{N}_i \cap U$ for some $U \in \mathcal{U}_i \subset \mathcal{U}$, so $V \subset U$.

2. **Locally finite:** For any $p \in M$, there exists a minimal i such that $p \in K_i$. Consider the neighborhood $W = \operatorname{int}(K_{i+1}) \setminus K_{i-2}$. Then:

   - If $j > i+2$, then $W \cap \widetilde{N}_j \subset K_{i+1} \cap (M \setminus K_j) = \emptyset$ since $j > i+2$ implies $K_{i+1} \subset K_{j-2}$.
   - If $j < i-2$, then $W \cap \widetilde{N}_j \subset (M \setminus K_{i-2}) \cap K_{j+1} = \emptyset$ since $j < i-2$ implies $K_{j+1} \subset K_{i-2}$.

   Thus W intersects only $\widetilde{N}_{i-2}, \widetilde{N}_{i-1}, \widetilde{N}_i, \widetilde{N}_{i+1}, \widetilde{N}_{i+2}$ (with appropriate adjustments for boundary cases). Since each $\mathcal{V}_j$ is finite, W intersects only finitely many elements of $\mathcal{V}$.

Therefore, $\mathcal{V}$ is a locally finite refinement of $\mathcal{U}$, proving that M is paracompact. □

1.5 Classification of connected one-dimensional manifolds

We conclude this chapter with the classification of connected one-dimensional topological manifolds. The key point is that a one-dimensional manifold is locally an open interval, and global behavior is controlled by how such intervals can degenerate at their endpoints.

Lemma 1.37 (Interval extension lemma)**.** *Let $I, J \subset \mathbb{R}$ be bounded open intervals, and let $K \subset I$ and $L \subset J$ be closed subintervals. Then every homeomorphism $f_0\colon K \to L$ extends to a homeomorphism $f\colon I \to J$. In particular, there exists a homeomorphism $f\colon I \to J$ such that $f(K) = L$.*

Proof. Write

$$I = (a, b), \qquad J = (c, d), \qquad K = [\alpha, \beta], \qquad L = [\gamma, \delta].$$

Fix a homeomorphism $f_0\colon K \to L$. Define $f\colon I \to J$ by

$$f(x) = \begin{cases} c + (\gamma - c)\dfrac{x-a}{\alpha - a}, & x \in (a, \alpha], \\ f_0(x), & x \in [\alpha, \beta], \\ \delta + (d-\delta)\dfrac{x-\beta}{b-\beta}, & x \in [\beta, b). \end{cases}$$

This map is continuous and strictly monotone on each of the three pieces, and the endpoint values match at α and β. Hence f is a continuous bijection from the interval I onto the interval J, so it is a homeomorphism. By construction, $f|_K = f_0$. □

Proposition 1.38. *Let M be a connected 1-dimensional topological manifold. Suppose that M admits a compact exhaustion*

$$K_1 \subset K_2 \subset K_3 \subset \cdots, \qquad K_i \subset \operatorname{int}(K_{i+1}), \qquad \bigcup_{i=1}^{\infty} K_i = M,$$

in which each K_i is homeomorphic to $[0,1]$. Then M is homeomorphic to $\mathbb{R}$.

Proof. We construct homeomorphisms

$$F_i \colon K_i \to [-i, i]$$

inductively so that $F_{i+1}|_{K_i} = F_i$ for every i.

Choose any homeomorphism $F_1 \colon K_1 \to [-1,1]$. Assume that F_i has been constructed. Since $K_i \subset \operatorname{int}(K_{i+1})$ and $K_{i+1} \cong [0,1]$, the complement $K_{i+1} \setminus \operatorname{int}(K_i)$ consists of two nonempty boundary arcs. Equivalently, if we choose a homeomorphism

$$\theta_i \colon K_{i+1} \to [0,1],$$

then $\theta_i(K_i)$ is a closed subinterval contained in $(0,1)$. By Lemma 1.37, there exists a homeomorphism

$$\lambda_i \colon [0,1] \to [-i-1, i+1]$$

such that

$$\lambda_i|_{\theta_i(K_i)} = F_i \circ \theta_i^{-1}|_{\theta_i(K_i)}.$$

Set

$$F_{i+1} := \lambda_i \circ \theta_i.$$

Then F_{i+1} is a homeomorphism $K_{i+1} \to [-i-1, i+1]$ extending F_i.

The maps F_i therefore glue to a well-defined map

$$F \colon M \to \mathbb{R}$$

with $F|_{K_i} = F_i$ for every i. It is bijective because the intervals $[-i, i]$ exhaust $\mathbb{R}$. It remains to show that F is a homeomorphism.

The map F is continuous because each restriction $F|_{K_i}$ is continuous and the interiors of the K_i cover M. Likewise, F^{-1} is continuous because its restriction to each compact interval $[-i, i]$ is F_i^{-1}. Hence F is a homeomorphism. □

Proposition 1.39. *Let M be a 1-dimensional topological manifold, and let*

$$U_1 \subset U_2 \subset U_3 \subset \cdots$$

be an increasing sequence of open subsets of M, each homeomorphic to $(0,1)$. Then

$$U := \bigcup_{i=1}^{\infty} U_i$$

is also homeomorphic to $(0,1)$.

Proof. Since U is an open subset of a second countable Hausdorff manifold, it is itself a second countable Hausdorff locally compact space. Hence U admits a compact exhaustion

$$C_1 \subset C_2 \subset C_3 \subset \cdots, \qquad C_i \subset \operatorname{int}(C_{i+1}), \qquad \bigcup_{i=1}^{\infty} C_i = U.$$

By passing to a subsequence of the U_i, we may assume that $C_i \subset U_i$ for every i.

We now construct compact intervals

$$K_1 \subset K_2 \subset K_3 \subset \cdots$$

with

$$C_i \subset K_i \subset U_i, \qquad K_i \subset \operatorname{int}(K_{i+1}), \qquad K_i \cong [0,1].$$

Assume that K_{i-1} has been chosen. Because $C_i \cup K_{i-1}$ is a compact subset of U_i and $U_i \cong (0,1)$, choose a homeomorphism

$$\phi_i \colon U_i \to (0,1).$$

Then $\phi_i(C_i \cup K_{i-1})$ is a compact subset of $(0,1)$, hence contained in some closed interval $[a_i, b_i] \subset (0,1)$. Define

$$K_i := \phi_i^{-1}([a_i, b_i]).$$

Then $K_i \cong [0,1]$, it contains $C_i \cup K_{i-1}$, and because K_{i-1} is compact and K_i has nonempty interior in the interval U_i, we may enlarge $[a_i, b_i]$ slightly if necessary so that actually $K_{i-1} \subset \operatorname{int}(K_i)$.

Thus the K_i form a compact exhaustion of U by copies of $[0,1]$. Proposition 1.38 now implies that $U \cong \mathbb{R}$. Since $\mathbb{R} \cong (0,1)$, the proof is complete. □

Corollary 1.40. *Let M be a 1-dimensional topological manifold, and let $I \subset M$ be an open subset homeomorphic to $(0,1)$. Then there exists a maximal open subset $U \subset M$ such that $I \subset U$ and $U \cong (0,1)$.*

Proof. Consider the collection of all open subsets of M that contain I and are homeomorphic to $(0,1)$, partially ordered by inclusion. If $\{U_j\}_{j \in J}$ is a totally ordered subcollection, second countability of M allows us to choose an increasing countable cofinal subchain

$$U_{j_1} \subset U_{j_2} \subset \cdots,$$

and Proposition 1.39 shows that the union is again homeomorphic to $(0,1)$. Thus every chain has an upper bound, and Zorn's lemma yields a maximal element. □

Lemma 1.41 (Local containment lemma)**.** *Let M be a one-dimensional topological manifold, and let $f \colon (0,1) \to M$ be a continuous injective map. If there exists a strictly increasing sequence $t_i \to 1$ such that $f(t_i) \to p \in M$, then there exist a coordinate chart (V, ψ) around p with $\psi(V) = (-2,2)$ and an integer i_0 such that for every $i \geq i_0$,*

$$f([t_i, t_{i+1}]) \subset V \qquad \textit{and} \qquad \psi\big(f([t_i, t_{i+1}])\big) \subset (-1,1).$$

Proof. Because $f(t_i) \to p$, we may choose a chart (V, ψ) around p with $\psi(V) = (-2, 2)$ such that for all sufficiently large i we have $f(t_i) \in V$ and $\psi(f(t_i)) \in (-1, 1)$. Fix N so that this holds for every $i \geq N$.

If for every $i \geq N$ one already has

$$f([t_i, t_{i+1}]) \subset V \qquad \text{and} \qquad \psi\big(f([t_i, t_{i+1}])\big) \subset (-1, 1),$$

then there is nothing to prove. Suppose instead that $i \geq N$ is an index for which this fails. Set

$$S_i := \big\{s \in [t_i, t_{i+1}] : f([t_i, s]) \subset V \text{ and } \psi\big(f([t_i, s])\big) \subset (-1, 1)\big\}$$

and let $\tau_i := \sup S_i$. By continuity, one has

$$f([t_i, \tau_i]) \subset V \qquad \text{and} \qquad \psi\big(f([t_i, \tau_i])\big) \subset [-1, 1].$$

If $\psi(f(\tau_i)) \in (-1, 1)$, then the defining property of S_i would persist on a slightly larger interval, contradicting the definition of τ_i. Hence necessarily

$$\psi(f(\tau_i)) \in \{-1, 1\}.$$

Because f is injective, each of the two values -1 and 1 can occur in this way at most once. Therefore there are at most two indices $i \geq N$ for which the desired containment fails. Choosing i_0 larger than all of them proves the lemma. □

Lemma 1.42. *Let M be a one-dimensional topological manifold, and let $I \subset M$ be an open subset homeomorphic to $(0, 1)$. Then the closure $\overline{I}$ of I in M is homeomorphic to exactly one of the following spaces:*

$$(0, 1), \qquad [0, 1), \qquad [0, 1], \qquad S^1.$$

Proof. Choose a homeomorphism $f\colon (0, 1) \to I$. We analyze the possible limits of $f(t)$ as $t \to 0^+$ and $t \to 1^-$.

Suppose first that there exists a sequence $t_i \to 1^-$ such that $f(t_i) \to p \in M$. Choose a chart (V, ψ) around p with $\psi(V) = (-2, 2)$ and $\psi(p) = 0$. By Lemma 1.41, after passing to large i all image segments $f([t_i, t_{i+1}])$ lie in V and their coordinate images lie in $(-1, 1)$. Therefore, for some i_0, the map

$$g := \psi \circ f\colon [t_{i_0}, 1) \to (-1, 1)$$

is continuous and injective. A continuous injective real-valued function on an interval is strictly monotone, so the limit

$$L := \lim_{t \to 1^-} g(t)$$

exists in $[-1, 1]$. Setting $f(1) := \psi^{-1}(L)$ gives a continuous extension of f to the endpoint 1. By the Hausdorff property, this endpoint is uniquely determined whenever it exists. The same argument applies at 0.

Thus each endpoint of $(0, 1)$ either has no limit in M or contributes a single boundary point to the closure. If neither endpoint contributes a limit, then $\overline{I} \cong (0, 1)$. If exactly one endpoint contributes a limit, then $\overline{I} \cong [0, 1)$. If both contribute distinct limits, then $\overline{I} \cong [0, 1]$. If both contribute and the two limits agree, then the two ends are glued to one point, so $\overline{I} \cong S^1$. These are the only possibilities. □

Theorem 1.43 (Classification of connected 1-manifolds)**.** *Every connected one-dimensional topological manifold is homeomorphic either to* $\mathbb{R}$ *or to* S^1*.*

Proof. Let M be a connected one-dimensional topological manifold. Choose a nonempty open subset $I \subset M$ homeomorphic to $(0,1)$, and let $U \subset M$ be a maximal open subset containing I such that $U \cong (0,1)$, whose existence is guaranteed by Corollary 1.40. By Lemma 1.42, the closure $\overline{U}$ is homeomorphic to one of

$$(0,1), \qquad [0,1), \qquad [0,1], \qquad S^1.$$

We first rule out the half-open and closed interval cases. Suppose $\overline{U} \cong [0,1)$ or $\overline{U} \cong [0,1]$, and let $p \in \overline{U} \setminus U$ be an endpoint. Choose a chart (V,ψ) with $\psi(V) = (-2,2)$ and $\psi(p) = 0$. As in the proof of Lemma 1.42, after restricting to points of U sufficiently near p, the coordinate expression of $U \cap V$ is a one-sided interval, say $(0,\varepsilon)$ after possibly replacing ψ by $-\psi$. Then

$$W := \psi^{-1}((-\eta,\varepsilon))$$

for small $\eta > 0$ is an open interval neighborhood of p, and $U \cup W$ is an open subset strictly containing U that is still homeomorphic to $(0,1)$. This contradicts the maximality of U. Hence $\overline{U}$ cannot be homeomorphic to $[0,1)$ or $[0,1]$.

If $\overline{U} \cong S^1$, then $\overline{U}$ is compact and has no boundary points. For each $q \in \overline{U}$, a small coordinate neighborhood in M meets $\overline{U}$ in an open interval, hence lies entirely in $\overline{U}$. Thus $\overline{U}$ is open in M. It is also closed by definition, so connectedness of M implies $M = \overline{U} \cong S^1$.

The only remaining possibility is $\overline{U} \cong (0,1)$. In this case $\overline{U} = U$, so U is both open and closed in M. Since M is connected and $U \neq \emptyset$, we obtain $M = U$. Finally, $(0,1)$ is homeomorphic to $\mathbb{R}$, so $M \cong \mathbb{R}$. □

Remark 1.44. There is another useful route to the noncompact case. Using the compact exhaustion result proved earlier in this chapter, one can show that a noncompact connected one-dimensional manifold admits an exhaustion by compact intervals. The interval-extension result then allows one to glue these intervals inductively into a global homeomorphism with $\mathbb{R}$.

Chapter 2

Smooth Manifolds and Smooth Functions

In this chapter we pass from topological manifolds to smooth manifolds. We first introduce smooth atlases and smooth structures, then discuss basic examples and constructions, and finally develop the existence of cutoff functions, partitions of unity, and smooth approximation.

2.1 Smooth manifolds

We now equip a topological manifold with the additional structure that allows us to do calculus on it. This structure is encoded by smooth charts and smooth transition maps. We begin by recalling the notion of a smooth map between open subsets of Euclidean spaces.

Definition 2.1. Let $U \subset \mathbb{R}^m$ be an open set and $f : U \to \mathbb{R}^n$ be a function. We say that f is *smooth* (or C^∞) on U if:

1. f is continuous on U
2. All partial derivatives of f of all orders exist and are continuous on U

That is, for every multi-index $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{Z}^m_{\geq 0}$, the partial derivative

$$\partial^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_m^{\alpha_m}}$$

exists and is continuous on U, where $|\alpha| = \alpha_1 + \cdots + \alpha_m$.

Remark 2.2. This definition can be equivalently stated as:

- f is C^k for every $k \geq 0$
- f is infinitely differentiable (C^∞)
- The function f and all its derivatives vary continuously throughout U.

Let M^n be an n-dimensional topological manifold. Recall that a manifold is a locally Euclidean, Hausdorff, and second-countable topological space. For our present purposes, the most important feature is the *local Euclidean* structure. To make this precise, we introduce charts, which provide local coordinate descriptions of the manifold.

Definition 2.3. A *chart* on M is a pair (U, ϕ) where:

1. $U \subset M$ is an open set.
2. $\phi : U \to \phi(U) \subset \mathbb{R}^n$ is a homeomorphism onto an open subset of $\mathbb{R}^n$.

We can now compare how different charts relate to one another. The resulting transition maps will determine what it means for an atlas to be smooth.

Definition 2.4. Let (U, ϕ) and (V, ψ) be two charts on M. Let $W = U \cap V$ be their intersection. The maps

$$\psi \circ \phi^{-1}|_{\phi(W)} : \phi(W) \to \psi(W) \quad \text{and} \quad \phi \circ \psi^{-1}|_{\psi(W)} : \psi(W) \to \phi(W)$$

are called the *transition maps* between the charts. The two charts are called *smoothly compatible* if both transition maps are smooth (i.e., C^∞ maps between open subsets of $\mathbb{R}^n$).

Example 2.5 (Polar Coordinates)**.** Let $M = \mathbb{R}^2$ be the Euclidean plane. Consider two charts:

- The *Cartesian chart*: (U, ϕ) where $U = \mathbb{R}^2$ and $\phi = \mathrm{id}_{\mathbb{R}^2}$ is the identity map.
- The *polar chart*: (V, ψ) where $V = \mathbb{R}^2 \setminus \mathbb{R}^{\geq 0} \times \{0\}$ (the plane with the non-negative x-axis removed), and $\psi : V \to \mathbb{R}^2$ is defined by

$$\psi(x, y) = \left(\sqrt{x^2 + y^2}, \arctan\left(\frac{y}{x}\right)\right).$$

Denoting $r = \sqrt{x^2 + y^2}$ and $\theta = \arctan(y/x)$, we obtain the classical polar coordinates. The transition map from polar to Cartesian coordinates is given by:

$$\phi \circ \psi^{-1} : \psi(V) \to \mathbb{R}^2, \quad (r, \theta) \mapsto (r \cos\theta, r \sin\theta).$$

This is the familiar coordinate transformation between Cartesian and polar coordinates. Note that both transition maps $\psi \circ \phi^{-1}$ and $\phi \circ \psi^{-1}$ are smooth on their respective domains, making these charts smoothly compatible. The exclusion of the non-negative x-axis in the polar chart is necessary to ensure that ψ is a homeomorphism, as the angular coordinate θ requires a consistent branch choice.

Now we attempt to give the precise definition of a differentiable manifold. Intuitively, we need a "complete" family of charts whose domains cover the entire manifold, and the transition maps between any two charts should be smooth. This is analogous to how an atlas of maps is used in geography to describe the Earth.

Definition 2.6. An *atlas* $\mathcal{A}$ on M is a collection of charts on M such that:

1. The charts cover M: $\bigcup_{(U,\phi)\in\mathcal{A}} U = M$
2. Any two charts in $\mathcal{A}$ are mutually compatible.

Example 2.7 (Standard Charts on S^1)**.** The circle $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ can be endowed with a smooth structure using different atlases.

Atlas 1: Four Charts via Projections
Consider the following four open subsets of S^1:

- $U_1 = \{(x, y) \in S^1 : y > 0\}$ (upper semicircle)
- $U_2 = \{(x, y) \in S^1 : y < 0\}$ (lower semicircle)
- $U_3 = \{(x, y) \in S^1 : x > 0\}$ (right semicircle)
- $U_4 = \{(x, y) \in S^1 : x < 0\}$ (left semicircle)

Define the chart maps as projections onto coordinates:

- $\phi_1 : U_1 \to \mathbb{R}, \quad (x, y) \mapsto x$ (projection to x-axis)
- $\phi_2 : U_2 \to \mathbb{R}, \quad (x, y) \mapsto x$
- $\phi_3 : U_3 \to \mathbb{R}, \quad (x, y) \mapsto y$
- $\phi_4 : U_4 \to \mathbb{R}, \quad (x, y) \mapsto y$

Each ϕ_i is a homeomorphism onto $(-1, 1) \subset \mathbb{R}$. The transition maps between these charts are smooth. For example, for $U_1 \cap U_3 = \{(x, y) \in S^1 : x > 0, y > 0\}$, we have:

$$\phi_3 \circ \phi_1^{-1}(x) = \phi_3(x, \sqrt{1 - x^2}) = \sqrt{1 - x^2}, \quad x \in (0, 1)$$

which is smooth on $(0, 1)$. Similar calculations show all transition maps are smooth.

Atlas 2: Two Charts via Stereographic Projection

The stereographic projection provides a more elegant atlas with only two charts. Let:

- $V_1 = S^1 \setminus \{(0, 1)\}$ (circle without north pole)
- $V_2 = S^1 \setminus \{(0, -1)\}$ (circle without south pole)

Define the stereographic projections:

- $\psi_1 : V_1 \to \mathbb{R}$ (projection from north pole $(0, 1)$):
$$\psi_1(x, y) = \frac{x}{1 - y}$$
- $\psi_2 : V_2 \to \mathbb{R}$ (projection from south pole $(0, -1)$):
$$\psi_2(x, y) = \frac{x}{1 + y}$$

Both ψ_1 and ψ_2 are homeomorphisms onto $\mathbb{R}$. The transition map on $V_1 \cap V_2 = S^1 \setminus \{(0, 1), (0, -1)\}$ is:

$$\psi_2 \circ \psi_1^{-1}(t) = \psi_2\left(\frac{2t}{t^2 + 1}, \frac{t^2 - 1}{t^2 + 1}\right) = \frac{1}{t}, \quad t \in \mathbb{R} \setminus \{0\}$$

which is smooth on $\mathbb{R} \setminus \{0\}$. Similarly, $\psi_1 \circ \psi_2^{-1}(t) = 1/t$ is smooth.

Both atlases define the same smooth structure on S^1, demonstrating that different collections of charts can yield equivalent differentiable manifolds.

As the examples illustrate, a manifold can admit many different atlases. Analogous to the variety of physical atlases one might find for the Earth, we require a method to determine when two such atlases describe the same underlying smooth structure.

Definition 2.8. Two atlases $\mathcal{A}$ and $\mathcal{A}'$ on M are *compatible* if any of the following equivalent conditions holds:

1. $\mathcal{A} \cup \mathcal{A}'$ is an atlas.
2. Every chart in $\mathcal{A}$ is compatible with every chart in $\mathcal{A}'$.

Example 2.9 (Compatible Atlases on S^1)**.** The two atlases on S^1 described above—the four-chart atlas via projections and the two-chart atlas via stereographic projection—are in fact *compatible*. This means that they both belong to the same equivalence class of atlases and thus define the *same* smooth structure on the circle.

While one could verify this directly by checking the smoothness of all transition maps between charts of the first atlas and charts of the second, such a verification, though straightforward, is computationally tedious.

We will later establish this fact in a more conceptual and elegant way. Once we develop the theory of immersions and their relationship with differential structures, the compatibility of these atlases will follow naturally as a consequence of a more general principle. This approach highlights the power of the categorical viewpoint in differential geometry, where the fundamental properties of a manifold can be understood through its smooth maps to and from other manifolds.

Remark 2.10. Compatibility of atlases is an equivalence relation. Indeed, the reflexive and symmetric properties are obvious. To check transitivity, let $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$ be atlases such that $\mathcal{A}_1$ is compatible with $\mathcal{A}_2$, and $\mathcal{A}_2$ is compatible with $\mathcal{A}_3$. Take any charts $(U_1, \phi_1) \in \mathcal{A}_1$ and $(U_3, \phi_3) \in \mathcal{A}_3$. We must show that c_1 and c_3 are compatible. Let $V = U_1 \cap U_3$. If $V = \emptyset$, then c_1 and c_3 are trivially compatible. Suppose $V \neq \emptyset$. It suffices to check that $\phi_3 \circ \phi_1^{-1}$ is smooth on $\phi_1(V)$. We verify this by checking differentiability at $\phi_1(x)$ for each $x \in V$. Choose $(U_2, \phi_2) \in \mathcal{A}_2$ such that $x \in U_2$. Then:

$$\begin{aligned}
&\phi_2 \circ \phi_1^{-1} : \phi_1(U_1 \cap U_2) \to \phi_2(U_1 \cap U_2) \text{ is smooth at } \phi_1(x), \\
&\phi_3 \circ \phi_2^{-1} : \phi_2(U_2 \cap U_3) \to \phi_3(U_2 \cap U_3) \text{ is smooth at } \phi_2(x).
\end{aligned}$$

Hence, the composition $\phi_3 \circ \phi_1^{-1} = (\phi_3 \circ \phi_2^{-1}) \circ (\phi_2 \circ \phi_1^{-1})$ is smooth at $\phi_1(x)$ as desired.

This gives us an appropriate way to say when two atlases are essentially the same. We can now define:

A *differential manifold structure* on M is an equivalence class of compatible atlases on M.

An alternative definition can be formulated as follows. We say that an atlas $\mathcal{A}$ on M is *full* (or *maximal*) if whenever (U, ϕ) is a chart on M that is compatible with every chart $(V, \psi) \in \mathcal{A}$, then (U, ϕ) necessarily belongs to $\mathcal{A}$. It is straightforward to verify that each equivalence class of atlases on M contains exactly one full atlas.

We may therefore equivalently define:

A *differential manifold structure* on M is the choice of a full atlas on M.

To establish the equivalence of this definition with the previous one, we must show that for any atlas $\mathcal{A}$ on M, there exists a unique full atlas $\overline{\mathcal{A}}$ that is compatible with $\mathcal{A}$.

Theorem 2.11. *Let M be a topological manifold and $\mathcal{A}$ a smooth atlas on M. Then there exists a unique full atlas $\overline{\mathcal{A}}$ compatible with $\mathcal{A}$.*

Proof. The uniqueness is immediate. For existence, define $\overline{\mathcal{A}}$ to be the collection of all charts on M that are compatible with every chart in $\mathcal{A}$:

$$\overline{\mathcal{A}} := \{(U, \phi) : (U, \phi) \text{ is a chart on } M \text{ and is compatible with all } (V, \psi) \in \mathcal{A}\}.$$

We must verify that any two charts $(U_1, \phi_1), (U_2, \phi_2) \in \overline{\mathcal{A}}$ are smoothly compatible. Let $q \in \phi_1(U_1 \cap U_2)$ be arbitrary, and choose $p \in U_1 \cap U_2$ such that $\phi_1(p) = q$. Since $\mathcal{A}$ covers M, there exists a chart $(V, \psi) \in \mathcal{A}$ with $p \in V$.

Consider the transition map $\phi_2 \circ \phi_1^{-1}$ restricted to $\phi_1(U_1 \cap U_2 \cap V)$. This can be expressed as the composition:

$$\phi_2 \circ \phi_1^{-1}|_{\phi_1(U_1 \cap U_2 \cap V)} = \left(\phi_2 \circ \psi^{-1}|_{\psi(U_2 \cap V)}\right) \circ \left(\psi \circ \phi_1^{-1}|_{\phi_1(U_1 \cap V)}\right).$$

By the definition of $\overline{\mathcal{A}}$, both $\phi_2 \circ \psi^{-1}$ and $\psi \circ \phi_1^{-1}$ are smooth on their respective domains. Therefore, their composition $\phi_2 \circ \phi_1^{-1}$ is smooth in a neighborhood of q. Since q was arbitrary, the charts (U_1, ϕ_1) and (U_2, ϕ_2) are smoothly compatible.

Finally, $\overline{\mathcal{A}}$ is maximal by construction, which completes the proof of existence. □

Henceforth, we will assume that M is a topological space equipped with a fixed differential manifold structure. We denote by $\mathcal{A}(M)$ the full atlas corresponding to this structure. By a *chart* (U, ϕ) on M, we will always mean a chart belonging to $\mathcal{A}(M)$; similarly, an *atlas* $\mathcal{A}$ of M will always refer to a collection of charts that forms a subset of $\mathcal{A}(M)$ and covers M.

2.2 Examples and constructions

We next record several basic examples and constructions of smooth manifolds.

Example 2.12 (Euclidean Space)**.** The Euclidean space $\mathbb{R}^n$ carries a natural differential structure. The *standard smooth structure* is given by the maximal atlas containing the identity chart $(\mathbb{R}^n, \mathrm{id}_{\mathbb{R}^n})$. This atlas consists of all charts that are smoothly compatible with the identity map, forming what we call the *standard differentiable structure* on $\mathbb{R}^n$.

Example 2.13 (Open Submanifolds)**.** If M is a smooth manifold with maximal atlas $\mathcal{A}(M)$ and $U \subset M$ is an open subset, then U naturally inherits a smooth manifold structure. The *induced maximal atlas* on U is given by:

$$\mathcal{A}(U) = \{(V, \phi) : (V, \phi) \in \mathcal{A}(M) \text{ and } V \subset U\}.$$

That is, we simply restrict the charts of M to the open subset U. One readily verifies that this collection forms a maximal atlas on U, making it a smooth submanifold of M.

Example 2.14 (Distinct but Diffeomorphic Structures on $\mathbb{R}$)**.** Consider the topological manifold $\mathbb{R}$. The *standard differential structure* is given by the maximal atlas containing the identity chart $(\mathbb{R}, \mathrm{id}_{\mathbb{R}})$. However, we can define a *different* differential structure on the same underlying topological space using the chart $(\mathbb{R}, \phi)$ where $\phi(x) = x^3$.

These two charts are not smoothly compatible: the transition map $\mathrm{id}_{\mathbb{R}} \circ \phi^{-1}(t) = t^{1/3}$ is not differentiable at $t = 0$. Therefore, the maximal atlases generated by these charts are distinct, giving $\mathbb{R}$ (as a topological manifold) at least two different smooth structures.

This example demonstrates that *equality* of differential structures (i.e., requiring the maximal atlases to be identical) is too restrictive a notion of equivalence. A more appropriate concept, which we will introduce later, is that of *diffeomorphism* - two manifolds are considered equivalent if there exists a smooth bijection between them with smooth inverse, even if their maximal atlases differ.

There is also a converse point of view: instead of starting with an already defined topological manifold, one may begin with a family of compatible local charts and use them to construct the manifold itself. This is formalized in the following construction lemma.

Lemma 2.15 (Manifold Construction Lemma)**.** *Let X be a set. Suppose we are given a collection $\{U_\alpha\}_{\alpha\in I}$ of subsets of X and a collection of bijections $\{\phi_\alpha : U_\alpha \to \phi_\alpha(U_\alpha)\}_{\alpha\in I}$ where each $\phi_\alpha(U_\alpha)$ is an open subset of $\mathbb{R}^n$, such that:*

1. *For all $\alpha, \beta \in I$, the sets $\phi_\alpha(U_\alpha \cap U_\beta)$ and $\phi_\beta(U_\alpha \cap U_\beta)$ are open in $\mathbb{R}^n$.*
2. *The transition maps $\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$ are smooth.*
3. *The collection $\{U_\alpha\}$ covers X.*
4. ***(Hausdorff Separation)*** *For any two distinct points $p, q \in X$, there exist indices $\alpha, \beta \in I$ with $p \in U_\alpha$, $q \in U_\beta$, and open subsets $V_\alpha \subset U_\alpha$, $V_\beta \subset U_\beta$ such that:*
   - *$p \in V_\alpha$ and $q \in V_\beta$*
   - *$V_\alpha \cap V_\beta = \emptyset$*
   - *$\phi_\alpha(V_\alpha)$ and $\phi_\beta(V_\beta)$ are open in $\mathbb{R}^n$*
5. ***(Second Countability)*** *There exists a countable subcollection $\{U_{\alpha_i}\}_{i=1}^{\infty}$ that still covers X.*

Then there exists a unique topology and smooth structure on X making it a smooth n-dimensional manifold for which the (U_α, ϕ_α) form a smooth atlas.

Proof. Define a topology on X by declaring $O \subseteq X$ open if and only if $\phi_\alpha(O \cap U_\alpha)$ is open in $\mathbb{R}^n$ for all $\alpha \in I$. This is the unique topology making all ϕ_α homeomorphisms.

The given conditions ensure:

- Local Euclideanness: Each (U_α, ϕ_α) is a homeomorphism onto an open subset of $\mathbb{R}^n$.
- Hausdorff: Condition (4) provides separation via charts.
- Second countability: Condition (5) gives a countable cover, and $\mathbb{R}^n$ has countable bases.
- Smooth structure: Conditions (1)-(3) ensure the charts form a compatible atlas.

Uniqueness follows because any topology making the ϕ_α homeomorphisms must coincide with our definition. □

This lemma tells us that we don't need to start with a topology on X; the topology is *induced* by the charts.

It is instructive to compare this with another gluing construction:

Theorem 2.16 (Patching Smooth Structures)**.** *Let M be a topological manifold that can be written as a union of subsets $M = \bigcup_\alpha U_\alpha$. Suppose each U_α is endowed with a smooth manifold structure such that for every α, β, the smooth structures induced on $U_\alpha \cap U_\beta$ from U_α and U_β coincide. Then there exists a unique smooth structure on M that restricts to the given smooth structure on each U_α.*

Proof. For each α, let $\mathcal{A}_\alpha$ be the maximal smooth atlas on U_α defining its smooth structure. Define an atlas on M by taking the union:

$$\mathcal{A} = \bigcup_\alpha \mathcal{A}_\alpha.$$

Then:

- $\mathcal{A}$ covers M since the U_α cover M and each $\mathcal{A}_\alpha$ covers U_α.
- Any two charts in $\mathcal{A}$ are compatible: if $(V, \phi) \in \mathcal{A}_\alpha$ and $(W, \psi) \in \mathcal{A}_\beta$, then the transition map $\psi \circ \phi^{-1}$ is smooth because the smooth structures on $U_\alpha \cap U_\beta$ from $\mathcal{A}_\alpha$ and $\mathcal{A}_\beta$ coincide.

The maximal atlas containing $\mathcal{A}$ gives the desired smooth structure on M. Uniqueness follows since any such structure must contain all $\mathcal{A}_\alpha$. □

2.3 Grassmann Manifolds

We now give a more substantial discussion of the Grassmann manifold $G(k, n)$, the space of all k-dimensional linear subspaces of $\mathbb{R}^n$. We first construct it using the Manifold Construction Lemma and then record several basic geometric properties that will be useful later.

2.3.1 Coordinate Charts

There is a natural identification between k-dimensional subspaces of $\mathbb{R}^n$ and equivalence classes of full-rank $n \times k$ matrices, where two matrices are equivalent if their column spaces coincide. More concretely, let

$$\mathcal{M} = \{ X \in \mathrm{Mat}(n \times k, \mathbb{R}) : \mathrm{rank}(X) = k \}.$$

For $X \in \mathcal{M}$, we write

$$\mathrm{col}(X) := \left\{ Xu : u \in \mathbb{R}^k \right\} = \mathrm{span}\{ X_1, \dots, X_k \} \subset \mathbb{R}^n,$$

where $X_1, \dots, X_k$ denote the columns of X. Thus $\mathrm{col}(X)$ is the k-dimensional subspace spanned by the columns of X.

We declare $X \sim Y$ if and only if $\mathrm{col}(X) = \mathrm{col}(Y)$. Then the quotient set $\mathcal{M}/\sim$ is in bijection with $G(k, n)$ via

$$[X] \longmapsto \mathrm{col}(X).$$

Under this identification, each k-dimensional subspace $W \subset \mathbb{R}^n$ is represented by any full-rank matrix X whose columns form a basis of W.

For each k-element subset $I \subset \{1, \ldots, n\}$, define

$$U_I = \{W \in G(k,n) : \text{the projection } \pi_I : W \to \mathbb{R}^I \cong \mathbb{R}^k \text{ is an isomorphism}\}.$$

Equivalently, if we represent W as the column space of an $n \times k$ matrix X of full rank, then

$$U_I = \{[X] : \det(X_I) \neq 0\},$$

where X_I is the $k \times k$ submatrix of X with rows indexed by I.

Define the chart map $\phi_I : U_I \to \mathrm{Mat}((n-k) \times k, \mathbb{R}) \cong \mathbb{R}^{k(n-k)}$ by

$$\phi_I([X]) = X_{I^c} X_I^{-1},$$

where I^c is the complement of I. Note that if $X \sim Y$, then $Y = XG$ for some invertible $k \times k$ matrix G, so

$$Y_{I^c} Y_I^{-1} = (X_{I^c} G)(X_I G)^{-1} = X_{I^c} X_I^{-1},$$

and hence ϕ_I is well defined on equivalence classes.

Remark 2.17. For fixed X_I, the map $X_{I^c} \mapsto X_{I^c} X_I^{-1}$ is a linear isomorphism $\mathrm{Mat}((n-k) \times k, \mathbb{R}) \to \mathrm{Mat}((n-k) \times k, \mathbb{R})$. Thus ϕ_I sends open sets in U_I to open sets in $\mathbb{R}^{k(n-k)}$.

Proposition 2.18. *The collection $\{(U_I, \phi_I)\}$ endows $G(k,n)$ (equivalently, the quotient $\mathcal{M}/\sim$) with a natural structure of smooth manifold of dimension $k(n-k)$.*

Proof. We verify each condition of Manifold Construction Lemma:

Covering: For any k-dimensional subspace W, there exists some I such that the projection $\pi_I : W \to \mathbb{R}^I$ is an isomorphism. Thus the U_I cover $G(k,n)$.

Openness: For any I, J, the sets $\phi_I(U_I \cap U_J)$ and $\phi_J(U_I \cap U_J)$ are open in $\mathbb{R}^{k(n-k)}$, as they are defined by the non-vanishing of certain determinants.

Smooth transitions: For $[X] \in U_I \cap U_J$, the transition map is:

$$\phi_J \circ \phi_I^{-1}(A) = \phi_J\left(\left[\begin{pmatrix} I_k \\ A \end{pmatrix}\right]\right)$$

To compute this, we find $g \in \mathrm{GL}(k)$ such that:

$$\begin{pmatrix} I_k \\ A \end{pmatrix} g = \begin{pmatrix} B \\ C \end{pmatrix}$$

where B corresponds to rows J and C to rows J^c. Then $\phi_J([X]) = CB^{-1}$. This is a rational function in the entries of A, hence smooth.

Hausdorff separation: Let $[X], [Y] \in G(k,n)$ be distinct. Consider the matrix $\begin{pmatrix} X & Y \end{pmatrix}$ which has rank at least $k+1$ since $[X] \neq [Y]$.

By the lower semicontinuity of matrix rank, there exists $\varepsilon > 0$ such that any matrix within ε of $\begin{pmatrix} X & Y \end{pmatrix}$ has rank at least $k+1$.

Let V_X be the $\varepsilon/2$-neighborhood of X in the space of $n \times k$ matrices of rank k, and V_Y the $\varepsilon/2$-neighborhood of Y. Choose ε small enough so that $V_X \subset U_I$ and $V_Y \subset U_J$, where $X \in U_I$ and $Y \in U_J$.

Then $[V_X]$ and $[V_Y]$ are the required open subsets with $[X] \in [V_X] \subset U_I$, $[Y] \in [V_Y] \subset U_J$, $[V_X] \cap [V_Y] = \emptyset$, and $\phi_I([V_X])$, $\phi_J([V_Y])$ are open in $\mathbb{R}^{k(n-k)}$.

Second countability: There are only $\binom{n}{k}$ charts, which is finite, so the atlas is countable. □

Proposition 2.19. *The Grassmann manifold $G(k,n)$ is compact.*

Proof. Consider the set of orthonormal k–frames in $\mathbb{R}^n$,

$$V_k(\mathbb{R}^n) := \{(v_1, \dots, v_k) \in (\mathbb{R}^n)^k : \langle v_i, v_j \rangle = \delta_{ij} \text{ for all } i, j\}.$$

Equivalently, writing $v_1, \dots, v_k$ as the columns of an $n \times k$ matrix Q, this is

$$V_k(\mathbb{R}^n) = \{Q \in \mathbb{R}^{n\times k} : Q^T Q = I_k\}.$$

The conditions $Q^T Q = I_k$ are given by polynomial equations in the entries of Q, so $V_k(\mathbb{R}^n)$ is a closed subset of $\mathbb{R}^{nk}$. Moreover, each column of Q has length 1, so $V_k(\mathbb{R}^n)$ is bounded. By the Heine–Borel theorem, $V_k(\mathbb{R}^n)$ is compact.

Define a map

$$\pi : V_k(\mathbb{R}^n) \to G(k,n), \qquad \pi(v_1, \dots, v_k) := \operatorname{span}\{v_1, \dots, v_k\}.$$

This map is continuous, and it is surjective because every k–dimensional subspace $W \subset \mathbb{R}^n$ admits an orthonormal basis obtained by applying the Gram–Schmidt process to any basis of W.

Since $V_k(\mathbb{R}^n)$ is compact and π is continuous, the image $\pi(V_k(\mathbb{R}^n))$ is compact. But π is surjective, so $\pi(V_k(\mathbb{R}^n)) = G(k,n)$. Hence $G(k,n)$ is compact. □

We note that the quotient map

$$\pi : \mathcal{M} \longrightarrow \mathcal{M}/\!\sim\, \cong G(k,n)$$

is a submersion. This follows from the existence of smooth local right inverses of π.

Fix $[X] \in \mathcal{M}/\!\sim$, and choose the unique representative $X \in \mathcal{M}$ whose upper $k \times k$ block is the identity matrix:

$$X = \begin{pmatrix} I_k \\ B \end{pmatrix}, \qquad B \in \operatorname{Mat}((n-k) \times k, \mathbb{R}).$$

Such a representative exists whenever $[X] \in U_I$ for some I of size k, simply by reordering rows so that $I = \{1, \dots, k\}$.

Define a map

$$s : \phi_I(U_I) \subset \mathbb{R}^{k(n-k)} \longrightarrow \mathcal{M}, \qquad s(B) = \begin{pmatrix} I_k \\ B \end{pmatrix}.$$

By construction, $\pi \circ s = \phi_I^{-1}$ on $\phi_I(U_I)$. In particular, s is a smooth *local right inverse* of π:

$$\pi \circ s = \mathrm{id} \quad \text{locally.}$$

It now follows from the standard criterion for submersions that π is a submersion.

We now describe the subset of $G(k,n)$ consisting of k–planes containing a fixed nonzero vector $v \in \mathbb{R}^n$ in terms of the matrix model $\mathcal{M}/\sim$ introduced above.

Lemma 2.20. *Let $v \in \mathbb{R}^n$ be a nonzero vector, and define*

$$\mathcal{G}_v^{(k)} := \{\, W \in G(k,n) : v \in W \,\}.$$

Then $\mathcal{G}_v^{(k)}$ is precisely the image, under the quotient map

$$\pi : \mathcal{M} \to \mathcal{M}/\sim \cong G(k,n), \qquad X \mapsto [X],$$

of the subset of $\mathcal{M}$ consisting of matrices which, after a suitable choice of coordinates sending v to e_1, have the normalized block form

$$X = \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}, \qquad \operatorname{rank}(A) = k-1.$$

Proof. As before, let

$$\mathcal{M} = \{X \in \operatorname{Mat}(n \times k, \mathbb{R}) : \operatorname{rank} X = k\}$$

and declare $X \sim Y$ if $\operatorname{col}(X) = \operatorname{col}(Y)$. The quotient $\mathcal{M}/\sim$ is identified with $G(k,n)$ via $[X] \mapsto \operatorname{col}(X)$.

Since our construction of $G(k,n)$ is natural with respect to linear isomorphisms of $\mathbb{R}^n$, we may first reduce to the case $v = e_1 = (1,0,\dots,0)$. Indeed, if $T : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear map with $T(e_1) = v$, then T induces a bijection on $\mathcal{M}$ (by left multiplication) which respects the equivalence relation $\sim$ and hence induces a bijection $G(k,n) \to G(k,n)$ carrying $\mathcal{G}_{e_1}^{(k)}$ onto $\mathcal{G}_v^{(k)}$. Thus it suffices to describe $\mathcal{G}_{e_1}^{(k)}$.

Write $\mathbb{R}^n = \Re_1 \oplus \mathbb{R}^{n-1}$, where $\mathbb{R}^{n-1}$ is spanned by $e_2, \dots, e_n$. We are interested in k–dimensional subspaces $W \subset \mathbb{R}^n$ containing e_1. Consider the subset $\mathcal{S} \subset \mathcal{M}$ consisting of all matrices of the block form

$$X = \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix},$$

where we have split the rows as $1 + (n-1)$ and the columns as $1 + (k-1)$, the zero blocks have the appropriate sizes, and $A \in \operatorname{Mat}((n-1) \times (k-1), \mathbb{R})$ has rank $k-1$. Equivalently, X has the form

$$X = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix} \in \mathbb{R}^{n \times k}.$$

We claim that $\pi(\mathcal{S}) = \mathcal{G}_{e_1}^{(k)}$.

Let $\mathcal{S}$ be the set of matrices in block form

$$X = \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}, \qquad \operatorname{rank}(A) = k-1.$$

If $X \in \mathcal{S}$, its first column is e_1 and its remaining columns lie in $\{0\} \oplus \mathbb{R}^{n-1}$. Hence the column space of X is a k–plane containing e_1, so $\pi(\mathcal{S}) \subset \mathcal{G}_{e_1}^{(k)}$.

Conversely, if $W \in \mathcal{G}_{e_1}^{(k)}$, then $e_1 \in W$. Since W has dimension k, we may extend e_1 to a basis

$$\{e_1, u_2, \dots, u_k\} \subset W.$$

Replacing u_j by $u_j - (u_j)_1 e_1$ if necessary, we may assume each u_j has first coordinate 0. Writing

$$u_j = (0, w_j), \qquad w_j \in \mathbb{R}^{n-1},$$

the matrix

$$X = \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}, \qquad A = [\, w_2 \ \cdots \ w_k \,],$$

has column space W, and A automatically has rank $k-1$. Thus $X \in \mathcal{S}$ and $\pi(X) = W$.

Therefore $\pi(\mathcal{S}) = \mathcal{G}_{e_1}^{(k)}$. For a general nonzero vector v, conjugate by a linear isomorphism sending e_1 to v.

□

Lemma 2.21. *Let $v \in \mathbb{R}^n$ be a nonzero vector and define*

$$\mathcal{G}_v^{(k)} := \{\, W \in G(k,n) : v \in W \,\}.$$

Then, after a linear change of coordinates sending v to e_1, the following holds: for every k–element subset $I \subset \{1, \dots, n\}$ with $1 \in I$, the image

$$\phi_I(\mathcal{G}_{e_1}^{(k)} \cap U_I) \subset \mathrm{Mat}((n-k) \times k, \mathbb{R}) \cong \mathbb{R}^{k(n-k)}$$

is a linear subspace.

Proof. Since the Grassmann manifold construction is natural with respect to linear isomorphisms of $\mathbb{R}^n$, we may assume without loss of generality that $v = e_1$.

Fix an index set $I \subset \{1, \dots, n\}$ with $|I| = k$ and $1 \in I$, and write I^c for its complement. Under the identification

$$\mathbb{R}^n \cong \mathbb{R}^I \oplus \mathbb{R}^{I^c} \cong \mathbb{R}^k \oplus \mathbb{R}^{n-k},$$

every k–plane $W \in U_I$ can be uniquely written as the graph of a linear map

$$A : \mathbb{R}^k \to \mathbb{R}^{n-k}, \qquad W = \{(x, Ax) : x \in \mathbb{R}^k\}.$$

In the chart (U_I, ϕ_I) we have $\phi_I(W) = A$, viewed as an element of $\mathrm{Mat}((n-k) \times k, \mathbb{R})$.

Because $1 \in I$, the vector e_1 lies in the "horizontal" component $\mathbb{R}^I \cong \mathbb{R}^k$, so we can write

$$e_1 = (x_*, 0), \qquad x_* \in \mathbb{R}^k.$$

We may choose a basis of $\mathbb{R}^I$ so that $x_* = e_1$ under the identification $\mathbb{R}^I \cong \mathbb{R}^k$. In this basis, a matrix $A \in \mathrm{Mat}((n-k) \times k, \mathbb{R})$ has columns

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_k \end{bmatrix}, \qquad a_j \in \mathbb{R}^{n-k}.$$

We now compute the condition $e_1 \in W$:

$$e_1 \in W \iff \exists\, x \in \mathbb{R}^k : \ (x, Ax) = (e_1, 0) \iff x = e_1 \text{ and } Ae_1 = 0.$$

In terms of the columns of A, this means

$$Ae_1 = a_1 = 0.$$

Thus

$$W \in \mathcal{G}^{(k)}_{e_1} \cap U_I \qquad \Longleftrightarrow \qquad A = \phi_I(W) \text{ satisfies } a_1 = 0.$$

Equivalently,

$$\phi_I(\mathcal{G}^{(k)}_{e_1} \cap U_I) = \big\{ A \in \mathrm{Mat}((n-k) \times k, \mathbb{R}) : \text{the first column of } A \text{ is zero} \big\},$$

which is a linear subspace of $\mathrm{Mat}((n-k) \times k, \mathbb{R})$.

For a general nonzero vector v, apply a linear isomorphism sending e_1 to v. □

Lemma 2.22. *Let $v \in \mathbb{R}^n$ be a nonzero vector and define*

$$\mathcal{G}^{(k)}_v := \{ W \in G(k,n) : v \in W \}.$$

Then, after a linear change of coordinates sending v to e_1, the following holds:

(i) For every k–element subset $I \subset \{1, \dots, n\}$ with $1 \in I$, the image

$$\phi_I(\mathcal{G}^{(k)}_{e_1} \cap U_I) \subset \mathrm{Mat}((n-k) \times k, \mathbb{R}) \cong \mathbb{R}^{k(n-k)}$$

is a linear subspace.

(ii) The subset $\mathcal{G}^{(k)}_{e_1}$ is an embedded submanifold of $G(k,n)$, naturally diffeomorphic to $G(k-1, n-1)$.

Consequently, for any nonzero v, the subset $\mathcal{G}^{(k)}_v$ is an embedded submanifold diffeomorphic to $G(k-1, n-1)$.

Proof. Since the Grassmann manifold construction is natural with respect to linear isomorphisms of $\mathbb{R}^n$, we may assume without loss of generality that $v = e_1$.

Step 1: Local description as a linear subspace in coordinates. Fix an index set $I \subset \{1, \dots, n\}$ with $|I| = k$ and $1 \in I$, and write I^c for its complement. Under the identification

$$\mathbb{R}^n \cong \mathbb{R}^I \oplus \mathbb{R}^{I^c} \cong \mathbb{R}^k \oplus \mathbb{R}^{n-k},$$

every k–plane $W \in U_I$ can be uniquely written as the graph of a linear map

$$A : \mathbb{R}^k \to \mathbb{R}^{n-k}, \qquad W = \{(x, Ax) : x \in \mathbb{R}^k\}.$$

In the chart (U_I, ϕ_I) we have $\phi_I(W) = A$, viewed as an element of $\mathrm{Mat}((n-k) \times k, \mathbb{R})$.

Because $1 \in I$, the vector e_1 lies in the "horizontal" component $\mathbb{R}^I \cong \mathbb{R}^k$, so we can write

$$e_1 = (x_*, 0), \qquad x_* \in \mathbb{R}^k.$$

We may choose a basis of $\mathbb{R}^I$ so that $x_* = e_1$ under the identification $\mathbb{R}^I \cong \mathbb{R}^k$. In this basis, a matrix $A \in \mathrm{Mat}((n-k) \times k, \mathbb{R})$ has columns

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_k \end{bmatrix}, \qquad a_j \in \mathbb{R}^{n-k}.$$

We now compute the condition $e_1 \in W$:

$$e_1 \in W \iff \exists\, x \in \mathbb{R}^k : \ (x, Ax) = (e_1, 0) \iff x = e_1 \text{ and } Ae_1 = 0.$$

In terms of the columns of A, this means

$$Ae_1 = a_1 = 0.$$

Thus

$$W \in \mathcal{G}_{e_1}^{(k)} \cap U_I \qquad \Longleftrightarrow \qquad A = \phi_I(W) \text{ satisfies } a_1 = 0.$$

Equivalently,

$$\phi_I(\mathcal{G}_{e_1}^{(k)} \cap U_I) = \big\{A \in \mathrm{Mat}((n-k) \times k, \mathbb{R}) : \text{the first column of } A \text{ is zero}\big\},$$

which is a linear subspace of $\mathrm{Mat}((n-k) \times k, \mathbb{R})$. This proves (i), and shows that $\mathcal{G}_{e_1}^{(k)}$ is locally cut out by linear equations in each Grassmann chart, hence is an embedded submanifold.

Step 2: The induced charts coincide with those of $G(k-1, n-1)$.

Fix $I \subset \{1, \dots, n\}$ with $|I| = k$ and $1 \in I$, and write

$$I = \{1\} \cup I', \qquad |I'| = k - 1.$$

On $G(k, n)$ the chart ϕ_I identifies $W \in U_I$ with a matrix

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_k \end{pmatrix} \in \mathrm{Mat}((n-k) \times k, \mathbb{R}).$$

By Step 1, the submanifold $\mathcal{G}_{e_1}^{(k)}$ is described inside U_I by the linear condition $a_1 = 0$, hence its induced coordinate domain is

$$\{A = [\,0 \ \ B\,] : B \in \mathrm{Mat}((n-k) \times (k-1), \mathbb{R})\}.$$

Define

$$\tilde{\phi}_I : \mathcal{G}_{e_1}^{(k)} \cap U_I \ \longrightarrow \ \mathrm{Mat}((n-k) \times (k-1), \mathbb{R}), \qquad \tilde{\phi}_I(W) := B,$$

i.e. $\tilde{\phi}_I$ is obtained from ϕ_I by deleting the (zero) first column.

But the chart of $G(k-1, n-1)$ on the index set I' is defined by the same formula: for $U \in G(k-1, n-1)$,

$$\psi_{I'}(U) = B.$$

Therefore the induced chart $(\mathcal{G}_{e_1}^{(k)} \cap U_I, \tilde{\phi}_I)$ *coincides* with the chart $(V_{I'}, \psi_{I'})$ of $G(k-1, n-1)$ under the identification

$$W \ \longleftrightarrow \ F(W) := W \cap \mathbb{R}^{n-1}.$$

Finally, for two charts I, J with $1 \in I, J$, the transition map

$$\tilde{\phi}_J \circ \tilde{\phi}_I^{-1}$$

is obtained by restricting $\phi_J \circ \phi_I^{-1}$ to the hyperplane $a_1 = 0$, which is exactly the transition map for $G(k-1, n-1)$ written in the charts indexed by I', J'. Hence the induced atlas is identical to the Grassmann atlas of $G(k-1, n-1)$.

We conclude that

$$\mathcal{G}_{e_1}^{(k)} \cong G(k-1, n-1)$$

as smooth manifolds. Together with the initial linear change of coordinates, this yields $\mathcal{G}_v^{(k)} \cong G(k-1, n-1)$ for every nonzero v. □

More generally, we can consider k–planes containing a fixed subspace $P \subset \mathbb{R}^n$.

Definition 2.23. Let $P \subset \mathbb{R}^n$ be a linear subspace of dimension r with $1 \le r \le k < n$. We define

$$\mathcal{G}_P^{(k)} := \{\, W \in G(k,n) : P \subset W \,\}.$$

We now show that $\mathcal{G}_P^{(k)}$ is a Grassmann submanifold, and that the general case can be reduced inductively to the case of containing a single vector.

Proposition 2.24. *Let $P \subset \mathbb{R}^n$ be a linear subspace of dimension r with $1 \le r \le k$. Then $\mathcal{G}_P^{(k)}$ is an embedded submanifold of $G(k,n)$, and there is a natural diffeomorphism*

$$\mathcal{G}_P^{(k)} \cong G(k-r, n-r).$$

In particular,

$$\dim \mathcal{G}_P^{(k)} = (k-r)(n-k).$$

Proof. Step 1: *Reduction to the standard position.* Choose a linear isomorphism $T : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$T(P) = \operatorname{span}(e_1, \dots, e_r).$$

Then T induces a diffeomorphism

$$T_* : G(k,n) \to G(k,n), \qquad W \mapsto T(W),$$

which carries $\mathcal{G}_P^{(k)}$ onto

$$\mathcal{G}_{P_0}^{(k)}, \qquad P_0 := \operatorname{span}(e_1, \dots, e_r).$$

Thus it suffices to treat the case $P = P_0$.

Step 2: *Inductive reduction to the vector case.* Let

$$P_0 = \operatorname{span}(e_1, \dots, e_r)$$

and define

$$P_j := \operatorname{span}(e_1, \dots, e_j), \qquad \mathcal{G}_{P_j}^{(k)} := \{\, W \in G(k,n) : P_j \subset W \,\}.$$

Then we have the descending chain

$$G(k,n) \supset \mathcal{G}_{P_1}^{(k)} \supset \mathcal{G}_{P_2}^{(k)} \supset \cdots \supset \mathcal{G}_{P_r}^{(k)} = \mathcal{G}_{P_0}^{(k)}.$$

By the previously established "vector case", the subset $\mathcal{G}_{P_1}^{(k)} = \mathcal{G}_{e_1}^{(k)}$ is an embedded submanifold of $G(k,n)$, and in every Grassmann chart (U_I, ϕ_I) with $1 \in I$ its image is a linear subspace defined by the vanishing of the first column of the coordinate matrix. In particular, $\mathcal{G}_{P_1}^{(k)}$ is diffeomorphic to $G(k-1, n-1)$.

Now consider the next condition

$$P_2 = \operatorname{span}(e_1, e_2) \subset W.$$

Inside the submanifold $\mathcal{G}^{(k)}_{P_1}$, this condition again becomes a *linear* condition in the same coordinate charts: after the first column has been set to zero, the requirement that $e_2 \in W$ forces the second column of the local matrix to vanish. Thus

$$\mathcal{G}^{(k)}_{P_2} \subset \mathcal{G}^{(k)}_{P_1}$$

is obtained by imposing one more linear condition in each chart, and hence is again an embedded submanifold. Moreover, the resulting coordinate space has dimension $k-2$ columns and $(n-2)$ rows, so $\mathcal{G}^{(k)}_{P_2}$ is diffeomorphic to $G(k-2, n-2)$.

Iterating this process, each step imposes one additional condition $e_{j+1} \in W$, which in the adapted Grassmann charts corresponds to setting the $(j+1)$-st column of the matrix to zero. Therefore each inclusion

$$\mathcal{G}^{(k)}_{P_{j+1}} \subset \mathcal{G}^{(k)}_{P_j}$$

is an embedded submanifold obtained by a linear coordinate condition, and

$$\mathcal{G}^{(k)}_{P_j} \cong G(k-j, n-j).$$

After r steps we obtain

$$\mathcal{G}^{(k)}_{P_0} \cong G(k-r, n-r),$$

which completes the reduction.

□

2.4 Smooth maps and diffeomorphisms

Let M^m and N^n be two differential manifolds. A map $f : M \to N$ is said to be a smooth map if:

1. f is continuous.
2. f is "locally given by smooth functions", that is, there exists atlases $\mathcal{A}$ of M and $\mathcal{B}$ of N such that if $(U, \phi) \in \mathcal{A}$ and $(V, \psi) \in \mathcal{B}$, then setting $W = U \cap f^{-1}(V)$, the composite

   $$\phi(W) \xrightarrow{\phi^{-1}} W \xrightarrow{f} V \xrightarrow{\psi} \psi(V)$$

   is smooth.

Remark 2.25. 1. We describe condition 2 by saying that f is "locally given by smooth functions" since, in coordinates, composites of the form $\psi \circ f \circ \phi^{-1}$ may be written as n-tuples of smooth function of m variables.

2. Condition 2 is independent of the choice of atlases $\mathcal{A}$ and $\mathcal{B}$, as is seen by an argument similar to the one showing that compatibility of atlases in an equivalence relation.

The following formal properties of smooth maps are easily verified:

1. The composition of smooth maps is a smooth map.
2. The identity map on a manifold is a smooth map.

Definition 2.26. Let $f\colon M \to N$ be a smooth map. We say that f is a *diffeomorphism* if there exists a smooth map $g\colon N \to M$ such that

$$g \circ f = \mathrm{id}_M \quad \text{and} \quad f \circ g = \mathrm{id}_N\,.$$

In this case, we say that M and N are *diffeomorphic*.

Remark 2.27. Bijective smooth maps is not necessary a diffeomorphism, as the example $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$ shows.

After defining differentiable manifolds and smooth maps between them, we can employ the perspective of categories and functors to describe differential structures. While this abstract language may seem simple and mundane, it will greatly simplify our subsequent discussions on the relationship between immersions and differential structures. Moreover, the categorical viewpoint is widely used in mathematical branches such as algebraic topology and algebraic geometry, making familiarity with this language beneficial. Roughly speaking, the underlying philosophy is that an object is uniquely determined by its relationships with other objects.

Theorem 2.28. *Let M be a topological manifold, let $\mathcal{A}$ and $\mathcal{B}$ be full atlases on M, and let $M_{\mathcal{A}}$ (resp. $M_{\mathcal{B}}$) denote the smooth manifold whose underlying space is M that is determined by $\mathcal{A}$ (resp. $\mathcal{B}$). Then the following are equivalent:*

1. *$M_{\mathcal{A}} = M_{\mathcal{B}}$, that is, $\mathcal{A} = \mathcal{B}$.*
2. *The identity map $\mathrm{id}_M : M_{\mathcal{A}} \to M_{\mathcal{B}}$ is a diffeomorphism.*
3. *For all manifolds N, we have the equality $C^\infty(M_{\mathcal{A}}, N) = C^\infty(M_{\mathcal{B}}, N)$.*
4. *For all manifolds N, we have the equality $C^\infty(N, M_{\mathcal{A}}) = C^\infty(N, M_{\mathcal{B}})$.*

Proof. Although this theorem can be viewed as a special case of the general categorical principle that an object representing a functor is determined up to unique isomorphism, we provide a concrete proof in this context.

The implications (1) $\Rightarrow$ (2) and (2) $\Leftrightarrow$ (3) $\Leftrightarrow$ (4)are straightforward.

We now prove (2) $\Rightarrow$ (1): We need to show that any chart $(U, \phi) \in \mathcal{A}$ is compatible with any chart $(V, \psi) \in \mathcal{B}$. If $U \cap V = \emptyset$, the compatibility is trivial. Assume $U \cap V \neq \emptyset$ and take any $p \in U \cap V$. Since id_M is a diffeomorphism, by definition there exist charts $(U', \phi') \in \mathcal{A}$ and $(V', \psi') \in \mathcal{B}$ such that $p \in U' \cap V'$ and the transition map

$$\psi' \circ \phi'^{-1}\big|_{\phi'(U'\cap V')} = \psi' \circ \mathrm{id}_M \circ \phi'^{-1}\big|_{\phi'(U'\cap V')} : \phi'(U' \cap V') \to \psi'(U' \cap V')$$

is a diffeomorphism. Since (U, ϕ) and (U', ϕ') belong to the same atlas $\mathcal{A}$, and (V, ψ) and (V', ψ') belong to $\mathcal{B}$, the transition maps

$$\phi' \circ \phi^{-1}\big|_{\phi(U\cap U')} : \phi(U \cap U') \to \phi'(U \cap U')$$

and

$$\psi \circ \psi'^{-1}\big|_{\psi'(V\cap V')} : \psi'(V \cap V') \to \psi(V \cap V')$$

are diffeomorphisms. Therefore, their composition

$$\psi \circ \phi^{-1}\big|_{\phi(U\cap V\cap U'\cap V')} : \phi(U \cap V \cap U' \cap V') \to \psi(U \cap V \cap U' \cap V')$$

is a diffeomorphism. Since p was arbitrary, (U, ϕ) and (V, ψ) are compatible. As $\mathcal{A}$ and $\mathcal{B}$ are both maximal atlases, this implies $\mathcal{A} = \mathcal{B}$. $\square$

This theorem demonstrates that the smooth structure of a manifold is completely determined by either:

- The collection of smooth maps from the manifold to arbitrary test manifolds, or
- The collection of smooth maps from arbitrary test manifolds to the manifold.

It has been (and still remains) a celebrated problem to compare the category of topological manifolds with the category of smooth manifolds. Here are some historical landmarks:

- H. Whitney (1936): For $k \geq 1$, every C^k structure is C^k equivalent to a smooth structure.
- E. Moise (1952): For $n < 4$, every topological manifold carries a unique smooth structure up to diffeomorphism.
- J. Milnor (1956): Exotic smooth structures exist. In particular, the topological manifold S^7 carries smooth structures that are not diffeomorphic to the standard one.
- M. Kervaire (1960): There exists a compact 10-dimensional topological manifold that carries no smooth structure at all.
- M. Freedman / S. Donaldson (1982): $\mathbb{R}^4$ carries exotic smooth structures.

Comment on spherical case: For spheres S^n with $n \neq 4$, the classification of smooth structures is intimately related to the homotopy groups of spheres. In fact, the number of distinct smooth structures on S^n for $n \neq 4$ is determined by a certain invariant connected to these homotopy groups. Currently, this classification is known for dimensions up to about 90, with the number of exotic spheres growing rapidly in higher dimensions. However, the case of S^4 remains one of the most famous open problems in differential topology–it is still unknown whether S^4 admits any exotic smooth structures (the smooth Poincaré conjecture in dimension 4).

In this course, we will provide a complete classification of smooth structures on 1-dimensional manifolds. While this result is relatively simple, it is by no means trivial and illustrates the fundamental ideas of differential topology. If time permits, we will present two different approaches to this classification: one using vector fields and their flows, and another using the technique of *handle straightening*–both methods offering distinct geometric insights into the structure of 1-manifolds.

The classification of 2-dimensional manifolds will be covered in subsequent courses on differential topology, where more sophisticated tools like Morse theory and handle decompositions are developed. As for the fascinating and complex world of 3-dimensional manifolds–including the celebrated Poincaré conjecture proved by Perelman–this typically becomes the subject of specialized research, requiring deep techniques from geometric analysis and low-dimensional topology.

2.5 Cutoff Functions

All constructions in this section rely on the existence of smooth functions that are positive on specified regions of a manifold and identically zero elsewhere. We begin by constructing a fundamental smooth function on the real line that transitions from zero to positive.

Lemma 2.29 (Smooth Transition Function). *The function $f : \mathbb{R} \to \mathbb{R}$ defined by*

$$f(t) = \begin{cases} e^{-1/t} & t > 0 \\ 0 & t \leq 0 \end{cases}$$

is smooth.

Proof. We proceed by induction on n to show that for $t > 0$,

$$f^{(n)}(t) = \frac{d^n f}{dt^n} = P_n(t^{-1})e^{-1/t}$$

where P_n is a polynomial. The base case $n = 0$ holds trivially. Assuming the formula holds for n, we differentiate to obtain:

$$f^{(n+1)}(t) = \left[P_n'(t^{-1})(-t^{-2}) + P_n(t^{-1})t^{-2}\right] e^{-1/t} = P_{n+1}(t^{-1})e^{-1/t}.$$

To prove smoothness at $t = 0$, we note that $\lim_{t \to 0^+} t^{-1} f^{(n)}(t) = 0$ for every $n \geq 0$. This ensures that all derivatives of f exist and are continuous at $t = 0$, and in fact $f^{(n)}(0) = 0$. □

Lemma 2.30 (One-Dimensional Cutoff Function). *Given real numbers $r_1 < r_2$, there exists a smooth function $h : \mathbb{R} \to \mathbb{R}$ such that:*

- $h(t) \equiv 1$ *for* $t \leq r_1$
- $0 < h(t) < 1$ *for* $r_1 < t < r_2$
- $h(t) \equiv 0$ *for* $t \geq r_2$

Proof. Let f be the function from the previous lemma, and define

$$h(t) = \frac{f(r_2 - t)}{f(r_2 - t) + f(t - r_1)}.$$

The denominator is always positive since for any $t \in \mathbb{R}$, at least one of $r_2 - t > 0$ or $t - r_1 > 0$ holds. The function h inherits smoothness from f, and direct verification shows it satisfies all the stated properties. □

A function with these properties is called a *cutoff function.*

Lemma 2.31 (Smooth Bump Function). *Given positive real numbers $r_1 < r_2$, there exists a smooth function $H : \mathbb{R}^n \to \mathbb{R}$ such that:*

- $H \equiv 1$ *on* $\overline{B(0, r_1)}$
- $0 < H(x) < 1$ *for* $x \in B(0, r_2) \setminus \overline{B(0, r_1)}$
- $H \equiv 0$ *on* $\mathbb{R}^n \setminus B(0, r_2)$

Proof. Define $H(x) = h(|x|)$, where h is the one-dimensional cutoff function from the previous lemma with the same radii $r_1 < r_2$. The function H is smooth on $\mathbb{R}^n \setminus \{0\}$ as a composition of smooth functions. At the origin, H is identically 1 in a neighborhood, hence smooth there as well. □

2.6 Partitions of unity

It is often convenient to express smooth functions as sums of locally supported functions. This leads us to the concept of a partition of unity, whose existence fundamentally relies on the assumptions that M is second-countable and Hausdorff. Recall that the *support* $\mathrm{supp}(\eta) \subset M$ of a continuous function $\eta : M \to \mathbb{R}$ is defined as the closure of the open subset $\{x : \eta(x) \neq 0\}$.

Definition 2.32. Let $\mathcal{W} = \{W_i\}_{i \in I}$ be an open cover of M. A *partition of unity subordinate to* $\mathcal{W}$ is a collection of smooth functions $\eta_i : M \to [0, 1]$ satisfying:

1. $\mathrm{supp}(\eta_i) \subset W_i$ for each $i \in I$,
2. The family $\{\mathrm{supp}(\eta_i) : i \in I\}$ is locally finite, meaning every point $p \in M$ has a neighborhood intersecting only finitely many $\mathrm{supp}(\eta_i)$,
3. $\sum_{i \in I} \eta_i(p) = 1$ for all $p \in M$.

Theorem 2.33 (Existence of Partitions of Unity)**.** *Every open cover* $\mathcal{W} = \{W_i\}_{i \in I}$ *of* M *admits a subordinate partition of unity.*

We now prove the existence of partitions of unity. We begin by recalling the existence of an exhaustion by compact sets.

Proposition 2.34 (Exhaustion by Compact Sets)**.** *Every topological manifold* M *admits an* exhaustion by compact sets, *i.e., there exists a sequence* $\{K_n\}_{n=1}^{\infty}$ *of compact subsets such that:*

1. $K_n \subset \mathrm{int}(K_{n+1})$ *for all* $n \geq 1$,
2. $\bigcup_{n=1}^{\infty} K_n = M$.

Recall also that if $\mathcal{U}$ is an open cover of M, an open cover $\mathcal{V}$ is called a *refinement* of $\mathcal{U}$ if every $V \in \mathcal{V}$ is contained in some $U \in \mathcal{U}$. The previous proposition implies that M is *paracompact*, meaning that every open cover admits a locally finite refinement. We first establish a slightly weaker form of Theorem 2.33.

Proposition 2.35. *Every open cover* $\mathcal{W} = \{W_i\}_{i \in I}$ *of* M *has a refinement that admits a subordinate partition of unity.*

Proof. Let $K_1 \subset \mathrm{int}(K_2) \subset K_2 \subset \mathrm{int}(K_3) \subset \cdots$ be an exhaustion of M by compact sets as established previously, and let $\mathcal{W} = \{W_i\}_{i \in I}$ be an open cover. For any point $p \in M$, there exists a unique n such that $p \in K_n \setminus K_{n-1}$ (where we take $K_0 = \emptyset$). Then $\mathrm{int}(K_{n+1}) \setminus K_{n-1}$ is an open neighborhood of p. We choose a chart (U_α, ϕ_α) around p such that $\phi_\alpha(U_\alpha) = B(z, r) \subset \mathbb{R}^n$ and

$$U_\alpha \subset W_i \cap (\mathrm{int}(K_{n+1}) \setminus K_{n-1})$$

for some $i \in I$.

Varying p over M (and thus over all $n \geq 1$), the collection of open sets $\phi_\alpha^{-1}(B(z, r/3))$ covers M. In particular, for each n, the compact set $K_n \setminus \mathrm{int}(K_{n-1})$ is covered by finitely many such sets, say $\phi_{\alpha_i^n}^{-1}(B(z_i^n, r_i^n/3))$ for $1 \leq i \leq j_n$. Taking the union over n of these finite families, we obtain a countable open cover of M, since

$$\bigcup_{n \geq 1} (K_n \setminus \mathrm{int}(K_{n-1})) \supset \bigcup_{n \geq 1} (K_n \setminus K_{n-1}) = M.$$

By construction, each $\phi_{\alpha_i^n}^{-1}(B(z_i^n, r_i^n/3))$ is contained in some W_i, so this cover is a refinement of $\mathcal{W}$.

We now verify local finiteness. Let $p \in M$. Then p lies in some $K_n \setminus K_{n-1}$, and hence in the open set $\mathrm{int}(K_{n+1}) \setminus K_{n-1}$. By construction, p is contained in one of the sets $U_{\alpha_i^n} = \phi_{\alpha_i^n}^{-1}(B(z_i^n, r_i^n))$, which is contained in $\mathrm{int}(K_{n+1}) \setminus K_{n-1}$. Note that if $U_{\alpha_i^n}$ intersects $U_{\alpha_j^m}$, then we must have

$$(\mathrm{int}(K_{n+1}) \setminus K_{n-1}) \cap (\mathrm{int}(K_{m+1}) \setminus K_{m-1}) \neq \emptyset,$$

which implies that m can only be $n-1$, n, or $n+1$. Therefore, each $U_{\alpha_i^n}$ can intersect only finitely many $U_{\alpha_j^m}$ (specifically, at most $j_{n-1} + j_n + j_{n+1}$ of them). This establishes local finiteness, proving that M is paracompact.

Claim 2.36. *For each i and n, there exists a smooth function $\tilde{\rho}_i^n : B(z_i^n, r_i^n) \to [0,1]$ that vanishes outside $B(z_i^n, r_i^n/2)$ and is identically 1 on $B(z_i^n, r_i^n/3)$.*

This follows directly from the existence of smooth bump functions (Lemma 2.31) by appropriate scaling and translation.

We now define smooth functions $\tilde{\eta}_i^n : M \to [0,1]$ by

$$\tilde{\eta}_i^n(p) = \begin{cases} \tilde{\rho}_i^n(\phi_{\alpha_i^n}(p)) & \text{if } p \in \phi_{\alpha_i^n}^{-1}(B(z_i^n, r_i^n)), \\ 0 & \text{otherwise.} \end{cases}$$

Since the collection $\{\phi_{\alpha_i^n}^{-1}(B(z_i^n, r_i^n/3))\}$ covers M and the collection $\{\phi_{\alpha_i^n}^{-1}(B(z_i^n, r_i^n))\}$ is locally finite, the sum

$$p \mapsto \sum_{i,n} \tilde{\eta}_i^n(p)$$

is locally a finite sum of non-zero terms and hence defines a smooth function $M \to \mathbb{R}_{>0}$. We then define

$$\eta_i^n(p) = \frac{\tilde{\eta}_i^n(p)}{\sum_{i,n} \tilde{\eta}_i^n(p)}.$$

This yields the desired partition of unity subordinate to the refinement $\{U_{\alpha_i^n}\}$ of $\mathcal{W}$. □

Proof of Theorem 2.33. By the previous proposition we can find a refinement $\mathcal{W}' = \{W_j'\}_{j\in J}$ of $\mathcal{W} = \{W_i\}_{i\in I}$ and partition of unity $\{\eta_j' : M \to [0,1]\}$ subordinate to it.

For $j \in J$, fix a W_i such that $W_j' \subset W_i$. This gives a function $\lambda : J \to I$. We claim that

$$\eta_i := \sum_{j\in\lambda^{-1}(i)} \eta_j'$$

gives the desired partition of unity. This is a locally finite sum and hence a smooth function and the sum of η_i is 1 everywhere. Also since that $\mathrm{supp}(\eta_j') \subset W_j'$ and hence is also contained in W_i. Now observe that

$$\mathrm{supp}(\eta_i) = \overline{\eta_i^{-1}((0,1])} = \overline{\bigcup_{j\in\lambda^{-1}(i)} (\eta_j')^{-1}((0,1])}.$$

In general, the closure of the union of subsets is larger than the union of closures of these subset. But Note that the latter is a closure of a locally finite union of open subsets. The local finiteness ensures that this is the union of the closures, by directly verify the definition. So we conclude that

$$\operatorname{supp}(\eta_i) = \bigcup_{j \in \lambda^{-1}(i)} \operatorname{supp}(\eta'_j) \subset W_i.$$

□

2.7 Applications of Partitions of Unity

Proposition 2.37 (Smooth Urysohn Lemma). *Let M be a smooth manifold, and let $A, B \subset M$ be disjoint closed subsets. Then there exists a smooth function $\lambda : M \to [0,1]$ such that $\lambda|_A \equiv 0$ and $\lambda|_B \equiv 1$.*

Proof. Consider the open cover of M given by $U_1 = M \setminus A$ and $U_2 = M \setminus B$. Let $\{\eta_1, \eta_2\}$ be a smooth partition of unity subordinate to this cover. Since $\operatorname{supp}(\eta_1) \subset U_1$, we have $\eta_1|_A \equiv 0$. Similarly, as $\operatorname{supp}(\eta_2) \subset U_2$, we have $\eta_2|_B \equiv 0$. Now define $\lambda = \eta_1$. Then $\lambda|_A = 0$, and on B we have $\lambda = \eta_1 = 1 - \eta_2 \equiv 1$, which completes the proof. □

The Smooth Urysohn Lemma has an equivalent formulation that is often more convenient for applications:

Proposition 2.38 (Smooth Bump Function). *Let M be a smooth manifold, $A \subset M$ a closed subset, and $U \subset M$ an open subset containing A. Then there exists a smooth function $\psi : M \to [0,1]$ such that:*

- $\psi|_A \equiv 1$
- $\operatorname{supp}(\psi) \subset U$

Proof. Apply the Smooth Urysohn Lemma to the disjoint closed sets A and $M \setminus U$. This yields a smooth function $\lambda : M \to [0,1]$ with $\lambda|_A \equiv 0$ and $\lambda|_{M \setminus U} \equiv 1$. Define $\psi = 1 - \lambda$. Then $\psi|_A \equiv 1$, and since $\lambda \equiv 1$ on $M \setminus U$, we have $\psi \equiv 0$ on $M \setminus U$, hence $\operatorname{supp}(\psi) \subset U$. □

Remark 2.39. These two propositions are equivalent: each can be derived from the other. The Smooth Bump Function version is particularly useful for constructing local extensions and cutoff functions, while the original Urysohn formulation provides a clear separation property for disjoint closed sets.

Our second application concerns the extension of smooth functions from closed subsets. Let M and N be smooth manifolds, and $A \subset M$ an arbitrary subset. We say that a map $F : A \to N$ is *smooth on A* if for every point $p \in A$, there exists an open neighborhood $W \subset M$ containing p and a smooth map $\widetilde{F} : W \to N$ that agrees with F on $W \cap A$.

Lemma 2.40 (Extension Lemma for Smooth Functions). *Let M be a smooth manifold, $A \subset M$ a closed subset, and $f : A \to \mathbb{R}^k$ a smooth function. For any open subset $U \subset M$ containing A, there exists a smooth function $\tilde{f} : M \to \mathbb{R}^k$ such that:*

- $\tilde{f}|_A = f$

- $\operatorname{supp}(\tilde{f}) \subset U$

Proof. For each point $p \in A$, choose a neighborhood $W_p \subset U$ and a smooth function $\tilde{f}_p : W_p \to \mathbb{R}^k$ that agrees with f on $W_p \cap A$. The collection $\{W_p : p \in A\} \cup \{M \setminus A\}$ forms an open cover of M. Let $\{\eta_p : p \in A\} \cup \{\eta_0\}$ be a smooth partition of unity subordinate to this cover, with $\operatorname{supp}(\eta_p) \subset W_p$ and $\operatorname{supp}(\eta_0) \subset M \setminus A$.

Define the global function $\tilde{f} : M \to \mathbb{R}^k$ by

$$\tilde{f}(x) = \sum_{p \in A} \eta_p(x) \tilde{f}_p(x).$$

This sum is well-defined since the partition of unity is locally finite. The function $\tilde{f}$ is smooth, agrees with f on A, and has support contained in U. □

Remark 2.41. The classical Whitney Extension Theorem provides a more refined result, requiring only compatibility conditions on the partial derivatives rather than assuming smoothness in the sense defined above.

For example, consider a function defined on the union of linear subspaces of Euclidean space in general position. If the restriction of the function to each subspace is smooth, the Whitney Extension Theorem guarantees the existence of a smooth extension to the whole Euclidean space. This situation frequently arises in geometry or analysis and cannot be handled by the Extension Lemma above, which requires the function to already have local smooth extensions near every point of the closed set.

Next, we use partitions of unity to construct a special type of smooth function. If M is a topological space, an *exhaustion function for M* is a continuous function $f : M \to \mathbb{R}$ such that for every $c \in \mathbb{R}$, the *sublevel set* $f^{-1}((-\infty, c])$ is compact. The terminology arises from the fact that as n ranges over the positive integers, the sublevel sets $f^{-1}((-\infty, n])$ form an exhaustion of M by compact sets.

Proposition 2.42 (Existence of Smooth Exhaustion Functions)**.** *Every smooth manifold admits a smooth positive exhaustion function.*

Proof. Let $\{V_j\}_{j=1}^{\infty}$ be a countable open cover of M by precompact open subsets, and let $\{\eta_j\}$ be a smooth partition of unity subordinate to this cover. Define $f \in C^\infty(M)$ by

$$f(p) = \sum_{j=1}^{\infty} j \eta_j(p).$$

To verify that f is an exhaustion function, fix $c \in \mathbb{R}$ and choose a positive integer $N > c$. If $p \notin \bigcup_{j=1}^{N} \overline{V}_j$, then $\eta_j(p) = 0$ for $1 \leq j \leq N$, and thus

$$f(p) = \sum_{j=N+1}^{\infty} j \eta_j(p) \geq \sum_{j=N+1}^{\infty} N \eta_j(p) = N \sum_{j=N+1}^{\infty} \eta_j(p).$$

Since $\{\eta_j\}$ is a partition of unity and $\eta_j(p) = 0$ for $j \leq N$, we have $\sum_{j=N+1}^{\infty} \eta_j(p) = 1$, which implies $f(p) \geq N > c$. Equivalently, if $f(p) \leq c$, then $p \in \bigcup_{j=1}^{N} \overline{V}_j$. Therefore, $f^{-1}((-\infty, c])$ is a closed subset of the compact set $\bigcup_{j=1}^{N} \overline{V}_j$ and is consequently compact. □

As a final application of partitions of unity, we show that every closed subset of a smooth manifold can be realized as the zero set of a smooth function.

Theorem 2.43 (Closed Sets as Zero Sets of Smooth Functions)**.** *Let M be a smooth manifold. For any closed subset $K \subset M$, there exists a smooth nonnegative function $f : M \to \mathbb{R}$ such that $f^{-1}(0) = K$.*

Proof. Since M is a smooth manifold, it is paracompact. Consider the open set $U = M \setminus K$. We construct a locally finite open cover $\{V_j\}_{j\in J}$ of U such that each V_j is precompact and $\overline{V_j} \subset U$.

For each $j \in J$, we construct a smooth nonnegative function $\psi_j : M \to \mathbb{R}$ with the following properties:

- $\psi_j > 0$ on V_j
- $\operatorname{supp}(\psi_j) \subset U$

Such functions can be constructed using bump functions supported in coordinate charts.

Now, let $\{\eta_j\}_{j\in J}$ be a smooth partition of unity subordinate to the cover $\{V_j\}_{j\in J}$. Define the function $f : M \to \mathbb{R}$ by

$$f(p) = \sum_{j\in J} \eta_j(p)\psi_j(p).$$

We verify that f has the desired properties:

If $p \in K$, then $p \notin U$ and thus $p \notin \operatorname{supp}(\eta_j)$ for all j, so $f(p) = 0$. Conversely, if $p \notin K$, then $p \in U$ and there exists some j with $\eta_j(p) > 0$ and $\psi_j(p) > 0$, so $f(p) > 0$. Therefore, $f^{-1}(0) = K$.

□

2.8 Density of Smooth Functions

We now establish a fundamental approximation theorem: smooth functions are dense in the space of continuous functions on manifolds with respect to uniform convergence. This result underpins many constructions in geometric analysis, allowing us to approximate continuous geometric structures by smooth ones.

2.8.1 Convolution in Euclidean Space

The key tool for our approximation is convolution, which provides a method to smooth out functions while preserving their essential features.

Definition 2.44 (Convolution)**.** Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be measurable functions. The *convolution* of f and g is defined as

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x - y)dy,$$

whenever this integral exists.

Convolution provides a powerful smoothing technique when we take g to be a bump function. This construction exemplifies the utility of the bump functions we developed earlier.

Theorem 2.45 (Convolution Approximation in $\mathbb{R}^n$)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function with compact support. For any $\varepsilon > 0$, there exists a smooth function $g : \mathbb{R}^n \to \mathbb{R}$ with compact support such that:*

$$\sup_{x \in \mathbb{R}^n} |f(x) - g(x)| < \varepsilon.$$

Moreover, if f is supported in a compact set K, then g can be chosen with support contained in any given open neighborhood of K.

Proof. Let $\rho : \mathbb{R}^n \to \mathbb{R}$ be a smooth bump function with $\rho \geq 0$, $\operatorname{supp}(\rho) \subset B(0,1)$, and $\int_{\mathbb{R}^n} \rho(x)dx = 1$. For $\delta > 0$, define the mollifier $\rho_\delta(x) = \delta^{-n}\rho(x/\delta)$ and consider the convolution

$$f_\delta(x) = (f * \rho_\delta)(x) = \int_{\mathbb{R}^n} f(y)\rho_\delta(x-y)dy.$$

This function is smooth (by differentiation under the integral sign) and has support contained in the δ-neighborhood of $\operatorname{supp}(f)$.

The key observation is that convolution averages the values of f near each point. Since f is uniformly continuous (by compact support), for any $\varepsilon > 0$ we can choose $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$. Then

$$|f(x) - f_\delta(x)| = \left| \int_{\mathbb{R}^n} [f(x) - f(y)]\rho_\delta(x-y)dy \right| \leq \int_{\mathbb{R}^n} |f(x) - f(y)|\rho_\delta(x-y)dy < \varepsilon,$$

establishing uniform convergence. □

2.8.2 Global Approximation on Manifolds

We now extend this result to smooth manifolds using partitions of unity. The strategy is to work locally in coordinate charts and then glue the approximations together.

Theorem 2.46 (Global Uniform Approximation on Manifolds)**.** *Let M be a smooth manifold and $f : M \to \mathbb{R}$ a continuous function. For any $\varepsilon > 0$, there exists a smooth function $g : M \to \mathbb{R}$ such that:*

$$\sup_{x \in M} |f(x) - g(x)| < \varepsilon.$$

That is, $C^\infty(M)$ is dense in $C^0(M)$ with respect to uniform convergence.

Proof. Since M is a smooth manifold, it is paracompact and second countable. Let $\{U_i\}_{i=1}^\infty$ be a locally finite open cover of M by precompact coordinate charts, with each $\phi_i : U_i \to \mathbb{R}^n$ a diffeomorphism onto its image. Let $\{\eta_i\}_{i=1}^\infty$ be a smooth partition of unity subordinate to this cover, with $\operatorname{supp}(\eta_i) \subset U_i$.

The precompactness condition ensures that each $\overline{U_i}$ is compact, which will be crucial for our local approximations.

For each i, define $f_i = \eta_i f$. This function is continuous and supported in U_i. Consider the pushforward $\tilde{f}_i = f_i \circ \phi_i^{-1}$ defined on $\phi_i(U_i) \subset \mathbb{R}^n$. Since U_i is precompact, $\tilde{f}_i$ has compact support in $\phi_i(U_i)$.

By the Euclidean approximation theorem, there exists a smooth function $\tilde{g}_i : \phi_i(U_i) \to \mathbb{R}$ with compact support such that

$$\sup_{y \in \phi_i(U_i)} |\tilde{f}_i(y) - \tilde{g}_i(y)| < \frac{\varepsilon}{2^i}.$$

Define $g_i : U_i \to \mathbb{R}$ by $g_i = \tilde{g}_i \circ \phi_i$, and extend it to all of M by setting it to zero outside U_i.

Now define the global approximation by

$$g(x) = \sum_{i=1}^{\infty} g_i(x).$$

This sum is well-defined and smooth because the cover is locally finite–each point has a neighborhood intersecting only finitely many U_i, so the sum reduces to a finite sum locally.

For the error estimate, observe that for any $x \in M$:

$$\begin{aligned} |f(x) - g(x)| &= \left| \sum_{i=1}^{\infty} \eta_i(x) f(x) - \sum_{i=1}^{\infty} g_i(x) \right| \\ &\leq \sum_{i=1}^{\infty} |\eta_i(x) f(x) - g_i(x)| \,. \end{aligned}$$

Note that

$$|\eta_i(x) f(x) - g_i(x)| = |\tilde{f}_i(\phi_i(x)) - \tilde{g}_i(\phi_i(x))| < \frac{\varepsilon}{2^i}.$$

Therefore,

$$|f(x) - g(x)| < \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \varepsilon,$$

which holds uniformly on M. □

Remark 2.47. Even in the familiar setting of Euclidean space $\mathbb{R}^n$, convolution alone is insufficient for global uniform approximation on non-compact domains. While convolution provides excellent local approximations, the partition of unity technique remains essential for gluing these into a global approximation with uniform control.

Indeed, convolution is not strictly necessary for this density theorem. An elementary alternative exists: take a locally finite open cover $\{U_i\}$ of $\mathrm{supp}(f)$ with each U_i sufficiently small, and construct

$$g(x) = \sum_i f(x_i) \eta_i(x)$$

using a subordinate partition of unity $\{\eta_i\}$ and points $x_i \in U_i$. This yields uniform approximation $\sup |f(x) - g(x)| < \varepsilon$.

However, convolution provides a more robust approach. Crucially, when f is C^k, the convolution approximation $f_\delta = f * \rho_\delta$ converges to f in the C^k topology.

2.9 Uniqueness of the smooth structure on $\mathbb{R}$ in dimension one

We end this chapter with a useful one-dimensional rigidity statement: although one can write down many atlases on the underlying topological space $\mathbb{R}$, every smooth structure on a one-dimensional manifold homeomorphic to $\mathbb{R}$ is in fact diffeomorphic to the standard one.

The proof proceeds by a local straightening argument on intervals, followed by a global patching construction.

Lemma 2.48 (Subdivision lemma for open covers of $\mathbb{R}$)**.** *Let $\mathcal{U}$ be an open cover of $\mathbb{R}$. Then there exists a strictly increasing sequence $\{t_j\}_{j\in\mathbb{Z}}$ with $t_j \to \pm\infty$ as $j \to \pm\infty$ such that for every $j \in \mathbb{Z}$ there is some $U \in \mathcal{U}$ with*

$$[t_j, t_{j+2}] \subset U.$$

Proof. For each integer i, let $I_i = [i, i+2]$. Since I_i is compact, the restricted cover $\mathcal{U}|_{I_i}$ has a Lebesgue number $\delta_i > 0$. Choose an integer $n_i \geq 1$ such that $1/n_i < \delta_i$, and set $N_i = 2n_i$.

Now subdivide $[i, i+1]$ into the points

$$i + \frac{k}{N_i}, \qquad k = 0, 1, \ldots, N_i.$$

Let $\{t_j\}_{j\in\mathbb{Z}}$ be the union of all these subdivision points, arranged in increasing order.

Take any consecutive triple $t_j < t_{j+1} < t_{j+2}$, and choose $i \in \mathbb{Z}$ such that $t_j \in [i, i+1]$. Then

$$t_{j+2} \leq t_j + \frac{2}{N_i} \leq i + 1 + \frac{1}{n_i} < i + 2,$$

so $[t_j, t_{j+2}] \subset I_i$. Moreover,

$$\operatorname{diam}[t_j, t_{j+2}] \leq \frac{2}{N_i} = \frac{1}{n_i} < \delta_i.$$

Hence $[t_j, t_{j+2}]$ lies in some member of the cover $\mathcal{U}$. □

Proposition 2.49 (Straightening a 0-handle in one dimension)**.** *Let $a < b < c$ be real numbers, let $U \subset \mathbb{R}$ be an open neighborhood of $[a, c]$, and let $f : U \to \mathbb{R}$ be a continuous injection. Then there exists a continuous injection $g : U \to \mathbb{R}$ and an open interval J containing b with $\overline{J} \subset (a, c)$ such that*

(i) $g = f$ on $U \setminus J$,

(ii) g is a diffeomorphism on J.

Proof. Since f is a continuous injection on an interval, it is strictly monotone. Replacing f by $-f$ if necessary, we may assume that f is strictly increasing.

Choose numbers $a < \alpha < b < \beta < c$ with $[\alpha, \beta] \subset U$, and set $J = (\alpha, \beta)$. Let $A = f(\alpha)$ and $B = f(\beta)$. Since f is strictly increasing, we have $A < B$.

Choose any orientation-preserving diffeomorphism

$$\lambda : (\alpha, \beta) \to (A, B)$$

that extends continuously to $[\alpha, \beta]$ with $\lambda(\alpha) = A$ and $\lambda(\beta) = B$. For instance, one may compose the affine homeomorphism $(\alpha, \beta) \to (0, 1)$ with a fixed smooth increasing diffeomorphism $(0, 1) \to (0, 1)$ whose continuous extension to $[0, 1]$ fixes the endpoints, and then compose with the affine map $(0, 1) \to (A, B)$.

Define

$$g(x) = \begin{cases} f(x), & x \in U \setminus J, \\ \lambda(x), & x \in J. \end{cases}$$

Because λ matches f at the endpoints of J, the map g is continuous. It is strictly increasing on each of the intervals $(-\infty, \alpha]$, $[\alpha, \beta]$, and $[\beta, \infty)$, and these pieces fit together monotonically because

$$g(\alpha) = f(\alpha) = A, \qquad \lambda(J) = (A, B), \qquad g(\beta) = f(\beta) = B.$$

Hence g is strictly increasing on U, and therefore injective. By construction, $g = f$ on $U \setminus J$, while $g|_J = \lambda$ is a diffeomorphism. □

Proposition 2.50 (Straightening a 1-handle in one dimension)**.** *Let $a < b < c$ be real numbers, let $U \subset \mathbb{R}$ be an open neighborhood of $[a, c]$, and let $f : U \to \mathbb{R}$ be a continuous injection. Assume that there exist open neighborhoods U_a of a and U_c of c such that f is a diffeomorphism on U_a and on U_c. Then there exists a smooth injection $g : U \to \mathbb{R}$ and neighborhoods $V_a \subset U_a$ of a and $V_c \subset U_c$ of c such that*

(i) $g = f$ on $V_a \cup V_c$,

(ii) g is a diffeomorphism on a neighborhood of $[a, c]$.

Proof. Since f is a continuous injection on an interval, it is strictly monotone. Replacing f by $-f$ if necessary, we may assume that f is strictly increasing.

Choose relatively compact open intervals $V_a \Subset U_a$ and $V_c \Subset U_c$ containing a and c, with disjoint closures. Since f is smooth with positive derivative on neighborhoods of $\overline{V_a}$ and $\overline{V_c}$, there exists $m_0 > 0$ such that

$$f'(x) \geq m_0 \qquad \text{for all } x \in \overline{V_a} \cup \overline{V_c}.$$

Let

$$K = [a, c] \setminus (V_a \cup V_c),$$

and choose an open set $\widetilde{U}$ with $K \subset \widetilde{U} \Subset U$ such that $\{V_a, V_c, \widetilde{U}\}$ covers $[a, c]$.

Let $f_\varepsilon = f * \rho_\varepsilon$ be a mollification of f on $\widetilde{U}$. For sufficiently small $\varepsilon > 0$, the function f_ε is smooth and strictly increasing on $\widetilde{U}$, and in addition

$$f_\varepsilon'(x) \geq \frac{m_0}{2} \qquad \text{on } \overline{V_a} \cup \overline{V_c},$$

while

$$\sup_{\widetilde{U}} |f_\varepsilon - f| \leq \zeta$$

for a parameter $\zeta > 0$ to be chosen later. Define

$$h(x) = f_\varepsilon(x) + \delta x,$$

where $\delta > 0$ is small. Then h is smooth on $\widetilde{U}$ and

$$h'(x) \geq \delta \qquad \text{on } \widetilde{U},$$

while on $\overline{V_a} \cup \overline{V_c}$ we have

$$h'(x) \geq \frac{m_0}{2} + \delta.$$

Let $\{\eta_a, \eta_c, \widetilde{\eta}\}$ be a smooth partition of unity subordinate to $\{V_a, V_c, \widetilde{U}\}$, and define

$$g(x) = \eta_a(x)f(x) + \eta_c(x)f(x) + \widetilde{\eta}(x)h(x).$$

Then g is smooth on a neighborhood of $[a, c]$, and $g = f$ on $V_a \cup V_c$ because there $\widetilde{\eta} = 0$.

It remains to show that $g' > 0$ near $[a, c]$. On the set where $\eta_a = 1$ or $\eta_c = 1$, we have $g = f$, so $g' \geq m_0$. On the set where $\widetilde{\eta} = 1$, we have $g = h$, so $g' \geq \delta$. On the overlap region,

$$g'(x) = (\eta_a + \eta_c)f'(x) + \widetilde{\eta}h'(x) + \eta_a'(f - h) + \eta_c'(f - h).$$

The first two terms are bounded below by

$$m_1 := \min\left\{m_0, \frac{m_0}{2} + \delta\right\} = \frac{m_0}{2} + \delta.$$

Let

$$R = \sup_{\operatorname{supp}(\eta_a')\cup\operatorname{supp}(\eta_c')} |x|, \qquad M_1 = \sup(|\eta_a'| + |\eta_c'|).$$

Since $h - f = (f_\varepsilon - f) + \delta x$, we obtain

$$|\eta_a'(f - h) + \eta_c'(f - h)| \leq M_1(\zeta + \delta R).$$

Choose $\delta > 0$ such that $\delta \leq m_0/4$ and $M_1\delta R \leq m_0/8$, and then choose $\zeta > 0$ such that $M_1\zeta \leq m_0/8$. Then

$$M_1(\zeta + \delta R) \leq \frac{m_0}{4} \leq \frac{m_1}{2},$$

so $g'(x) \geq m_1/2 > 0$ on the overlap region as well.

Thus $g' > 0$ on a neighborhood of $[a, c]$, and therefore g is a diffeomorphism on some open neighborhood of $[a, c]$. Since g is strictly increasing near $[a, c]$ and agrees with the injective map f near the two ends, after shrinking U if necessary we may view g as a smooth injection on U. □

Theorem 2.51 (Uniqueness of the smooth structure on $\mathbb{R}$ in dimension one)**.** *Let M be a one-dimensional smooth manifold that is homeomorphic to $\mathbb{R}$. Then M is diffeomorphic to $\mathbb{R}$ with its standard smooth structure.*

Proof. Choose a homeomorphism

$$f : \mathbb{R} \to M.$$

Let $\{(U_\alpha, \varphi_\alpha)\}$ be a smooth atlas on M. Applying the subdivision lemma to the open cover $\{f^{-1}(U_\alpha)\}$ of $\mathbb{R}$, we obtain a strictly increasing sequence $\{t_j\}_{j\in\mathbb{Z}}$ such that for every j there exists $\alpha(j)$ with

$$[t_j, t_{j+2}] \subset f^{-1}(U_{\alpha(j)}).$$

Equivalently,

$$f([t_j, t_{j+2}]) \subset U_{\alpha(j)}.$$

We now modify f in two stages.

Step 1: local smoothing near the odd subdivision points. For each $k \in \mathbb{Z}$, the map

$$f_{2k} := \varphi_{\alpha(2k)} \circ f$$

is a continuous injection on a neighborhood of $[t_{2k}, t_{2k+2}]$. Applying Proposition 2.49 to f_{2k} on the interval $[t_{2k}, t_{2k+2}]$, we may modify f_{2k} inside (t_{2k}, t_{2k+2}) to obtain a continuous injection which is smooth and locally a diffeomorphism near t_{2k+1}, while agreeing with f_{2k} near the two endpoints. Pulling back via $\varphi_{\alpha(2k)}^{-1}$ and performing this construction for all k, we obtain a global homeomorphism

$$g : \mathbb{R} \to M$$

that is a diffeomorphism on a neighborhood of each odd subdivision point t_{2k+1}.

Step 2: smoothing the connecting intervals. For each $k \in \mathbb{Z}$, consider the interval $[t_{2k-1}, t_{2k+1}]$. The map

$$g_{2k-1} := \varphi_{\alpha(2k-1)} \circ g$$

is a continuous injection on a neighborhood of this interval, and by construction it is already a diffeomorphism near the two endpoints t_{2k-1} and t_{2k+1}. Applying Proposition 2.50 to each such interval, and gluing the resulting modifications, we obtain a homeomorphism

$$h : \mathbb{R} \to M$$

that is smooth and locally a diffeomorphism on a neighborhood of every point of $\mathbb{R}$.

Since the neighborhoods arising from the intervals $[t_{2k-1}, t_{2k+1}]$ cover $\mathbb{R}$, the map h is a smooth bijection whose local coordinate expressions have nonvanishing derivative everywhere. Hence h is a local diffeomorphism. A bijective local diffeomorphism has smooth inverse, so h is a global diffeomorphism. □

Remark 2.52. The argument shows that in dimension one the only smooth structure on the underlying topological manifold $\mathbb{R}$ is the standard one, up to diffeomorphism. The key point is that every topological coordinate change can be smoothed on overlapping intervals without changing it near the overlap endpoints.

Chapter 3

Tangent Spaces and the Tangent Bundle

We now pass from the global theory of smooth manifolds to their infinitesimal geometry. The basic object at a point $p \in M$ is the tangent space T_pM, whose elements encode first-order directional information. We begin with the definition by derivations, explain its local nature and coordinate description, compare it with the geometric definition by curves, and finally assemble all tangent spaces into the tangent bundle.

3.1 Tangent vectors as derivations

The Euclidean model suggests the right definition. If $a \in \mathbb{R}^n$ and $p \in \mathbb{R}^n$, then the directional derivative

$$D_a f(p) = \nabla f(p) \cdot a$$

is a linear operator on smooth functions, and it satisfies the Leibniz rule

$$D_a(fg)(p) = f(p)D_a g(p) + g(p)D_a f(p).$$

On a manifold we take this property as the definition.

Definition 3.1. Let M be a smooth manifold and $p \in M$. A *derivation at* p is a linear map

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

such that

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

The set of all derivations at p is called the *tangent space* of M at p and is denoted by T_pM. Its elements are called *tangent vectors*.

Lemma 3.2. *Let $v \in T_pM$ and $f, g \in C^\infty(M)$.*

1. *If f is constant, then $v(f) = 0$.*
2. *If $f(p) = g(p) = 0$, then $v(fg) = 0$.*

Proof. For (1), applying the Leibniz rule to the constant function 1 gives

$$v(1) = v(1 \cdot 1) = 2v(1),$$

so $v(1) = 0$. Hence $v(c) = cv(1) = 0$ for every constant $c \in \mathbb{R}$.

For (2), the Leibniz rule gives

$$v(fg) = f(p)v(g) + g(p)v(f) = 0.$$

□

Example 3.3 (The tangent space of $\mathbb{R}^n$)**.** Fix $p \in \mathbb{R}^n$. For each $a = (a^1, \dots, a^n) \in \mathbb{R}^n$, define

$$D_a(f) = \sum_{i=1}^{n} a^i \frac{\partial f}{\partial x^i}(p).$$

Then $D_a \in T_p\mathbb{R}^n$.

Conversely, let $X \in T_p\mathbb{R}^n$. Set

$$a^i = X(x^i), \qquad a = (a^1, \dots, a^n),$$

where $x^1, \dots, x^n$ are the standard coordinate functions on $\mathbb{R}^n$. For any $f \in C^\infty(\mathbb{R}^n)$ we may write

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

for smooth functions g_i with $g_i(p) = \frac{\partial f}{\partial x^i}(p)$. Since X kills constants and $X\big((x^i - p^i)(x^j - p^j)\big) = 0$ by Theorem 3.2, it follows that

$$X(f) = \sum_{i=1}^{n} g_i(p) X(x^i) = \sum_{i=1}^{n} a^i \frac{\partial f}{\partial x^i}(p) = D_a(f).$$

Thus the map $a \mapsto D_a$ is an isomorphism $\mathbb{R}^n \cong T_p\mathbb{R}^n$. In particular, the standard basis vectors correspond to the coordinate derivations

$$e_i \longleftrightarrow \left.\frac{\partial}{\partial x^i}\right|_p.$$

Definition 3.4. Let $F : M \to N$ be a smooth map and let $p \in M$. The *differential* of F at p is the map

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

defined by

$$dF_p(v)(f) = v(f \circ F) \qquad \text{for } v \in T_pM,\ f \in C^\infty(N).$$

The map $dF_p(v)$ is again a derivation: for $f, g \in C^\infty(N)$,

$$\begin{aligned} dF_p(v)(fg) &= v\big((fg) \circ F\big) = v\big((f \circ F)(g \circ F)\big) \\ &= f(F(p))v(g \circ F) + g(F(p))v(f \circ F) \\ &= f(F(p))dF_p(v)(g) + g(F(p))dF_p(v)(f). \end{aligned}$$

Proposition 3.5 (Basic properties of the differential)**.** *Let $F : M \to N$ and $G : N \to P$ be smooth maps, and let $p \in M$.*

1. *The map $dF_p : T_pM \to T_{F(p)}N$ is linear.*
2. *$d(G \circ F)_p = dG_{F(p)} \circ dF_p$.*
3. *$d(\mathrm{id}_M)_p = \mathrm{id}_{T_pM}$.*
4. *If F is a diffeomorphism, then dF_p is an isomorphism and*

$$(dF_p)^{-1} = d(F^{-1})_{F(p)}.$$

Proof. Linearity is immediate from the definition. For the chain rule, if $v \in T_pM$ and $f \in C^\infty(P)$, then

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

The identity statement is immediate, and (4) follows by applying (2) and (3) to the identities

$$F^{-1} \circ F = \mathrm{id}_M, \qquad F \circ F^{-1} = \mathrm{id}_N .$$

□

3.2 The local nature of tangent vectors

Tangent vectors are first-order objects, so they should depend only on the germ of a function at the base point.

Proposition 3.6 (Locality)**.** *If $f, g \in C^\infty(M)$ agree on a neighborhood of $p \in M$, then*

$$v(f) = v(g)$$

for every $v \in T_pM$.

Proof. Let $h = f - g$, so h vanishes on some neighborhood U of p. Choose a bump function $\eta \in C^\infty(M)$ such that $\eta(p) = 1$ and $\mathrm{supp}(\eta) \subset U$. Then $\eta h \equiv 0$, hence

$$0 = v(\eta h) = \eta(p)v(h) + h(p)v(\eta) = v(h).$$

Thus $v(f) = v(g)$. □

A first consequence is that tangent spaces do not change when we restrict to an open neighborhood.

Theorem 3.7. *Let $U \subset M$ be open and let $p \in U$. Then the inclusion map $\iota : U \hookrightarrow M$ induces a natural isomorphism*

$$d\iota_p : T_pU \xrightarrow{\cong} T_pM.$$

Proof. Define $\Phi : T_pU \to T_pM$ by

$$\Phi(v)(f) = v(f|_U), \qquad f \in C^\infty(M).$$

This is clearly a derivation, hence $\Phi(v) \in T_pM$.

If $\Phi(v) = 0$ and $g \in C^\infty(U)$, choose a smooth extension $\tilde{g} \in C^\infty(M)$ that agrees with g near p. Then, by locality,

$$v(g) = v(\tilde{g}|_U) = \Phi(v)(\tilde{g}) = 0,$$

so $v = 0$. Thus Φ is injective.

Conversely, if $w \in T_pM$ and $g \in C^\infty(U)$, choose a smooth extension $\tilde{g} \in C^\infty(M)$ near p and define

$$v(g) = w(\tilde{g}).$$

By Theorem 3.6, this does not depend on the chosen extension. One checks immediately that $v \in T_pU$ and that $\Phi(v) = w$. Therefore Φ is an isomorphism, and it is precisely $d\iota_p$. □

Remark 3.8. Since $T_q\mathbb{R}^n \cong \mathbb{R}^n$ for each $q \in \mathbb{R}^n$ by Theorem 3.3, Theorem 3.7 implies that if M is an n-dimensional manifold, then every tangent space T_pM is an n-dimensional real vector space.

In practice, once $U \subset M$ is open we usually identify T_pU with T_pM through this theorem without further comment.

3.3 Coordinates on the tangent space

Let (U, φ) be a chart around $p \in M$, where

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

Using Theorem 3.7, we identify T_pU with T_pM and regard

$$d\varphi_p : T_pM \to T_{\varphi(p)}\mathbb{R}^n$$

as an isomorphism.

Definition 3.9. The *coordinate vectors* associated with the chart (U, φ) are defined by

$$\frac{\partial}{\partial x^i}\bigg|_p = (d\varphi_p)^{-1}\left(\frac{\partial}{\partial x^i}\bigg|_{\varphi(p)}\right), \qquad i = 1, \dots, n.$$

Equivalently, for every $f \in C^\infty(M)$,

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

Proposition 3.10. *For every chart $(U, \varphi = (x^1, \dots, x^n))$ around p, the vectors*

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

form a basis of T_pM. Moreover, every $v \in T_pM$ has the unique expression

$$v = \sum_{i=1}^n v(x^i) \frac{\partial}{\partial x^i}\bigg|_p.$$

Proof. The first statement follows because $d\varphi_p$ is an isomorphism and the standard coordinate vectors form a basis of $T_{\varphi(p)}\mathbb{R}^n$. For the expansion formula, write $d\varphi_p(v) = \sum_i a^i \frac{\partial}{\partial x^i}\big|_{\varphi(p)}$ and apply $(d\varphi_p)^{-1}$. The coefficients are determined by testing against the coordinate functions, so necessarily $a^i = v(x^i)$. □

Proposition 3.11 (Change of coordinates)**.** *Let* $(U, x^1, \dots, x^n)$ *and* $(V, y^1, \dots, y^n)$ *be two coordinate charts with* $p \in U \cap V$*. Then*

$$\frac{\partial}{\partial y^j}\bigg|_p = \sum_{i=1}^n \frac{\partial x^i}{\partial y^j}(p) \frac{\partial}{\partial x^i}\bigg|_p.$$

Equivalently, if

$$v = \sum_i v^i \frac{\partial}{\partial x^i}\bigg|_p = \sum_j w^j \frac{\partial}{\partial y^j}\bigg|_p,$$

then the components satisfy

$$w^j = \sum_i \frac{\partial y^j}{\partial x^i}(p) v^i.$$

Proof. For any $f \in C^\infty(M)$, the chain rule gives

$$\begin{aligned}
\frac{\partial}{\partial y^j}\bigg|_p (f) &= \frac{\partial (f \circ \psi^{-1})}{\partial y^j}(\psi(p)) \\
&= \sum_{i=1}^n \frac{\partial x^i}{\partial y^j}(p) \frac{\partial (f \circ \varphi^{-1})}{\partial x^i}(\varphi(p)).
\end{aligned}$$

This is exactly the claimed basis-change formula. The transformation rule for components follows by equating the two expansions of v and comparing coefficients. □

Remark 3.12. The basis vectors transform by the inverse Jacobian matrix, whereas the components of a tangent vector transform by the Jacobian matrix. This is why tangent vectors are often called *contravariant* objects.

3.4 Tangent vectors as velocities of curves

There is a second, more geometric, description of tangent vectors: they are velocities of curves through the point.

Definition 3.13. Let $p \in M$. Consider smooth curves $\gamma : (-\varepsilon, \varepsilon) \to M$ with $\gamma(0) = p$. We say that two such curves γ_1 and γ_2 are equivalent if, for one (and hence every) chart (U, φ) around p,

$$(\varphi \circ \gamma_1)'(0) = (\varphi \circ \gamma_2)'(0).$$

The set of equivalence classes is denoted by $T_p^{\text{curve}} M$.

Remark 3.14. The chart-independence follows from the ordinary chain rule in Euclidean space. Moreover, using coordinates one may transport addition and scalar multiplication from $\mathbb{R}^n$ to $T_p^{\text{curve}} M$, making it into a vector space.

Theorem 3.15 (Equivalence of the two definitions)**.** *For every $p \in M$, there is a natural isomorphism*

$$T_p^{\text{curve}} M \cong T_p M.$$

Proof. Given a curve class $[\gamma] \in T_p^{\text{curve}} M$, define

$$\Phi([\gamma])(f) = (f \circ \gamma)'(0), \qquad f \in C^\infty(M).$$

This is well defined and satisfies the Leibniz rule, so $\Phi([\gamma]) \in T_p M$.

Conversely, let $v \in T_p M$. Choose a chart $(U, \varphi = (x^1, \dots, x^n))$ with $\varphi(p) = 0$, and write

$$v = \sum_{i=1}^n v(x^i) \frac{\partial}{\partial x^i}\Big|_p.$$

Define a curve by

$$\gamma_v(t) = \varphi^{-1}(t\, v(x^1), \dots, t\, v(x^n))$$

for $|t|$ sufficiently small. Then $\gamma_v(0) = p$, and its coordinate velocity is precisely $(v(x^1), \dots, v(x^n))$. Hence $\Psi(v) := [\gamma_v]$ defines a map $\Psi : T_p M \to T_p^{\text{curve}} M$.

To check that Φ and Ψ are inverse, first note that for $v \in T_p M$ and any coordinate function x^i,

$$\Phi(\Psi(v))(x^i) = (x^i \circ \gamma_v)'(0) = v(x^i).$$

By Theorem 3.10, this implies $\Phi(\Psi(v)) = v$. Conversely, if $[\gamma] \in T_p^{\text{curve}} M$, then the coordinate velocity of γ equals the coordinate velocity of the curve constructed from the derivation $\Phi([\gamma])$, so $\Psi(\Phi([\gamma])) = [\gamma]$. □

3.5 The tangent bundle

The tangent spaces of a manifold fit together into a new manifold.

Definition 3.16. The *tangent bundle* of a smooth manifold M is the disjoint union

$$TM = \bigsqcup_{p \in M} T_p M = \{(p, v) : p \in M,\ v \in T_p M\}.$$

The natural projection

$$\pi : TM \to M$$

is given by $\pi(p, v) = p$.

If $(U, x^1, \dots, x^n)$ is a coordinate chart on M, every vector $v \in T_p M$ can be written uniquely as

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

This gives coordinates on $\pi^{-1}(U)$ by

$$\widetilde{\varphi}(p, v) = \big(x^1(p), \dots, x^n(p), v^1, \dots, v^n\big) \in \mathbb{R}^{2n}.$$

Theorem 3.17. *Let M be a smooth n-manifold. The above charts endow TM with a natural structure of a smooth $2n$-dimensional manifold.*

Proof. Each map $\widetilde{\varphi}$ is a bijection from $\pi^{-1}(U)$ onto the open set $\varphi(U) \times \mathbb{R}^n \subset \mathbb{R}^{2n}$. On an overlap of two coordinate charts (U, x) and (V, y), the induced transition map is

$$(x^1, \dots, x^n, v^1, \dots, v^n) \mapsto \Big(y^1(x), \dots, y^n(x), \sum_{j=1}^n \frac{\partial y^1}{\partial x^j}(x) v^j, \dots, \sum_{j=1}^n \frac{\partial y^n}{\partial x^j}(x) v^j\Big).$$

By smoothness of the coordinate transition map and its Jacobian entries, this transition map is smooth. Therefore these charts define a smooth atlas on TM. □

Remark 3.18. The tangent bundle is more than just a manifold: each fiber T_pM is a vector space, and in local coordinates the transition maps are linear in the fiber variables. This is the prototype of a *vector bundle*, a notion that will be developed systematically later.

Chapter 4

Submanifolds and Transversality

We now turn from tangent spaces to the local structure of smooth maps. The inverse function theorem gives the basic local normal form for maps with invertible differential. From it one obtains the local models for immersions and submersions, the theory of submanifolds, and the basic transversality constructions. The chapter culminates in the Whitney embedding theorem, tubular neighborhoods, Whitney approximation, and the homotopical form of transversality.

4.1 Parameterized Contraction Principle and the Inverse Function Theorem

Theorem 4.1 (Parameterized Contraction Mapping Principle)**.** *Let X be a complete metric space and Y a metric space. Let $\Phi : X \times Y \to X$ be a continuous map such that for some $0 \leq \rho < 1$ and all $y \in Y$:*

$$d(\Phi(x_1, y), \Phi(x_2, y)) \leq \rho d(x_1, x_2) \quad \text{for all } x_1, x_2 \in X.$$

Then for each $y \in Y$ there exists a unique $x \in X$ with $\Phi(x, y) = x$. If we denote this fixed point by $x = S(y)$, then the map $S : Y \to X$ is continuous.

Proof. Fix $x_0 \in X$ and define the sequence $\{x_n(y)\}$ by $x_0(y) = x_0$, $x_{n+1}(y) = \Phi(x_n(y), y)$.

By the contraction property, $d(x_{n+1}(y), x_n(y)) \leq \rho^n d(x_1(y), x_0(y))$. For $m > n$, the triangle inequality gives:

$$d(x_m(y), x_n(y)) \leq \sum_{k=n}^{m-1} \rho^k d(x_1(y), x_0(y)) \leq \frac{\rho^n}{1-\rho} d(x_1(y), x_0(y))$$

Since $\rho < 1$, $\{x_n(y)\}$ is Cauchy and converges to some $x(y) \in X$, which satisfies $\Phi(x(y), y) = x(y)$ by continuity. Uniqueness follows from the contraction property.

For continuity of $S(y) = x(y)$, given $\varepsilon > 0$, choose N so that $d(x_N(y), S(y)) < \varepsilon/3$ for all y. By continuity of Φ, $x_N(y)$ is continuous in y, so there exists $\delta > 0$ such that $d(y, y_0) < \delta$ implies $d(x_N(y), x_N(y_0)) < \varepsilon/3$. Then:

$$d(S(y), S(y_0)) \leq d(S(y), x_N(y)) + d(x_N(y), x_N(y_0)) + d(x_N(y_0), S(y_0)) < \varepsilon$$

□

Let $p \in M$ and let $x^1, \ldots, x^m$ be differentiable functions on a neighborhood U of p. Let $\varphi(q) = (x^1(q), \ldots, x^m(q))$ for $q \in U$. We say that $\{x^i\}_{1 \le i \le m}$ defines a coordinate system at p if there exists an open neighborhood U' of p, contained in U, such that $(U', \varphi|_{U'}, m)$ is a chart on M.

Theorem 4.2. *The following are equivalent:*

1. $\{x^i\}$ *defines a coordinate system at* p.
2. $\frac{\partial}{\partial x^i}$ *form a basis of* T_pM.

Theorem 4.2 is a consequence of the following more general theorem:

Theorem 4.3. *Let M and N be manifolds, $p \in M$ and $q \in N$, and let $F : M \to N$ be a smooth map such that $F(p) = q$. Then the following are equivalent:*

1. F *is a local diffeomorphism at* p.
2. $d_pF : T_pM \to T_qN$ *is an isomorphism.*

Theorem 4.4 (Inverse Function Theorem on Euclidean Space)**.** *Let $P : U \subset \mathbb{R}^m \to \mathbb{R}^m$ be a C^∞ map. Suppose that at some point $a \in U$, the derivative $DP(a) : \mathbb{R}^m \to \mathbb{R}^m$ is an invertible linear map. Then there exist neighborhoods $\tilde{U}$ of a and $\tilde{V}$ of $b = P(a)$ such that P restricts to a C^∞-diffeomorphism from $\tilde{U}$ onto $\tilde{V}$.*

Proof. Assume $a = 0$, $P(a) = 0$, and by composing with $[DP(0)]^{-1}$, assume $DP(0) = I$.

Define $\Phi(x, y) = x - P(x) + y$. Then $P(x) = y$ if and only if $\Phi(x, y) = x$.

Since $DP(0) = I$ and P is C^1, choose $\varepsilon > 0$ such that for $\|x_1\|, \|x_2\| \le \varepsilon$:

$$\|P(x_1) - P(x_2) - (x_1 - x_2)\| \le \tfrac{1}{2}\|x_1 - x_2\|$$

Then $\|\Phi(x_1, y) - \Phi(x_2, y)\| \le \frac{1}{2}\|x_1 - x_2\|$, so $\Phi(\cdot, y)$ is a contraction.

Let $X = \{x : \|x\| \le \varepsilon\}$, $Y = \{y : \|y\| \le \varepsilon/2\}$. For $x \in X$, $y \in Y$:

$$\|\Phi(x, y)\| \le \|x - P(x)\| + \|y\| \le \tfrac{1}{2}\|x\| + \|y\| \le \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon$$

so $\Phi(X \times Y) \subset X$.

By the contraction mapping theorem, for each $y \in Y$ there exists a unique $x \in X$ with $\Phi(x, y) = x$, i.e., $P(x) = y$. Denote this by $x = S(y)$. Then $S : Y \to X$ is continuous and $P \circ S = \mathrm{id}_Y$.

To show S is C^1:

Let $y_0 \in Y$ and $x_0 = S(y_0)$. Since $DP(x_0)$ is invertible (by continuity of DP and $DP(0) = I$), consider the difference quotient:

$$\frac{S(y) - S(y_0) - [DP(x_0)]^{-1}(y - y_0)}{\|y - y_0\|}$$

Using $P(S(y)) = y$ and $P(S(y_0)) = y_0$, we have:

$$\begin{aligned} &S(y) - S(y_0) - [DP(x_0)]^{-1}(y - y_0) \\ &= S(y) - S(y_0) - [DP(x_0)]^{-1}(P(S(y)) - P(S(y_0))) \end{aligned}$$

$$= [DP(x_0)]^{-1}\left(DP(x_0)(S(y) - S(y_0)) - (P(S(y)) - P(S(y_0)))\right)$$

By the differentiability of P at x_0:

$$P(S(y)) - P(S(y_0)) = DP(x_0)(S(y) - S(y_0)) + o(\|S(y) - S(y_0)\|)$$

Thus:

$$\frac{\|S(y) - S(y_0) - [DP(x_0)]^{-1}(y - y_0)\|}{\|y - y_0\|} \leq \|[DP(x_0)]^{-1}\| \cdot \frac{o(\|S(y) - S(y_0)\|)}{\|y - y_0\|}$$

Since S is continuous and $DP(x_0)$ is invertible, $\|S(y) - S(y_0)\|/\|y - y_0\|$ is bounded. Therefore the right-hand side tends to 0 as $y \to y_0$, proving that S is differentiable at y_0 with $DS(y_0) = [DP(x_0)]^{-1}$.

The continuity of DS follows from the continuity of DP and S.

Higher regularity: The C^k case for $k \geq 2$ follows by induction. We already have the derivative formula $DS(y) = [DP(S(y))]^{-1}$. If P is C^k, then DP is C^{k-1}, and since matrix inversion is smooth, the composition $[DP(S(y))]^{-1}$ is C^{k-1} by the chain rule and the induction hypothesis that S is C^{k-1}. Thus DS is C^{k-1}, meaning S is C^k. □

Remark 4.5. The contraction mapping principle is not strictly necessary here, as the inverse mapping is clearly Lipschitz continuous. Nevertheless, it is a convenient tool in other contexts, such as proving continuous dependence on initial conditions for ordinary differential equations.

4.2 Immersions, submersions, and subimmersions

Let M and N be smooth manifolds, $p \in M$ and $q \in N$, and let $f : M \to N$ be a smooth map such that $f(p) = q$. Let $m = \dim M$ and $n = \dim N$.

Definition 4.6. Let $\tilde{M}$ and $\tilde{N}$ be smooth manifolds, $\tilde{p} \in \tilde{M}$ and $\tilde{q} \in \tilde{N}$, and let $\tilde{f} : \tilde{M} \to \tilde{N}$ be a smooth map such that $\tilde{f}(\tilde{p}) = \tilde{q}$. Then (M, N, p, q, f) *looks locally like* $(\tilde{M}, \tilde{N}, \tilde{p}, \tilde{q}, \tilde{f})$ if there exist open neighborhoods U of p, V of q, $\tilde{U}$ of $\tilde{p}$, $\tilde{V}$ of $\tilde{q}$ and diffeomorphisms $g : U \to \tilde{U}$ and $h : V \to \tilde{V}$ such that:

1. $f(U) \subset V$ and $\tilde{f}(\tilde{U}) \subset \tilde{V}$,
2. $g(p) = \tilde{p}$ and $h(q) = \tilde{q}$,
3. The following diagram commutes:

$$\begin{array}{ccc} U & \xrightarrow{f} & V \\ {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle h} \\ \tilde{U} & \xrightarrow{\tilde{f}} & \tilde{V}. \end{array}$$

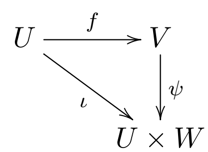
Remark 4.7. We shall apply this definition mainly when $\bar{M}$ is a vector space $\mathbb{R}^m$, $\bar{N}$ is a vector space $\mathbb{R}^n$ and $\bar{f}$ is a linear map. In this case, we will take $\bar{p} = 0$, $\bar{q} = 0$ without explicit mention.

4.2.1 Immersions

Notation 4.8. We write d_pf for the differential (or tangent map) of a smooth map f at a point p. For the induced smooth map $TMoTN$ between tangent bundles, we use the notation df.

Theorem 4.9. *The following are equivalent:*

1. d_pf *is injective.*
2. *There exist open neighborhoods* U *of* p, V *of* q, *and* W *of* 0 *(in* $\mathbb{R}^{n-m}$*) and a diffeomorphism* $\psi : V \to U \times W$ *such that:*
   (a) $f(U) \subset V$,
   (b) *If* ι *denotes the inclusion* $U \to U \times \{0\} \subset U \times W$, *then the following diagram commutes:*



3. (M, N, p, q, f) *looks locally like a linear injection* $\bar{f} : \mathbb{R}^m \to \mathbb{R}^n$.
4. *There exist local coordinates* $\{x^i\}$ *at* p *and* $\{y^j\}$ *at* q *such that* $x^i = y^i \circ f$ *for* $1 \leq i \leq m$ *and* $0 = y^j \circ f$ *for* $m + 1 \leq j \leq n$.
5. *There exist open neighborhoods* U *of* p *and* V *of* q, *and a smooth map* $\sigma : V \to U$ *such that* $f(U) \subset V$ *and* $\sigma \circ f = \mathrm{id}_U$.

Proof. The implications (2) ⇒ (3) ⇒ (4) ⇒ (5) ⇒ (1) are elementary.

We show (1) ⇒ (2). Since the question is local, we may assume that the following conditions are satisfied:

a. N is an open subset of $\mathbb{R}^n$,

b. $f(p) = 0$ and $\operatorname{Im} d_pf = \mathbb{R}^m \times \{0\} \subset \mathbb{R}^m \times \mathbb{R}^{n-m} = \mathbb{R}^n$.

Let W be $\{0\} \times \mathbb{R}^{n-m} \subset \mathbb{R}^n$. Define $f' : M \times W \to N$ by $f'(p, w) = f(p) + w$. Then by the inverse function theorem, f' is a local diffeomorphism at $(p, 0)$. Hence, by shrinking M, N and W, we may assume that f' is a diffeomorphism. Then the inverse ψ of f' satisfies the condition of (2).

Let us verify the commutativity $\psi \circ f = \iota$ in detail. By the definition of f' and the fact that f' is a diffeomorphism between $U \times W$ and V, we know that $f : U \to V$ is an injection. If $y \in f(U)$, then $y \in f'(U \times \{0\})$ and $f'^{-1}(y) = (f^{-1}(y), 0)$ by the definition of f'. Therefore, $\psi(f(x)) = (x, 0)$ for all $x \in U$, which means $\psi \circ f = \iota$. □

Definition 4.10. A smooth map f satisfying the equivalent conditions of the preceding theorem at p is called an *immersion at* p. A smooth map f which is an immersion at all $p \in M$ is called an *immersion*.

Example 4.11 (Inclusion into product manifold)**.** Let M and N be smooth manifolds, and fix a point $q_0 \in N$. Consider the inclusion map $\iota : M \to M \times N$ defined by

$$\iota(p) = (p, q_0).$$

This map is an immersion. To see this, consider the projection map $\pi : M \times N \to M$ defined by $\pi(p, q) = p$. Then we have

$$\pi \circ \iota = \mathrm{id}_M .$$

This means that π is a smooth left inverse for ι on the entire manifold M. By condition (5) of the theorem, with $U = M$ and $V = M \times N$, it follows immediately that ι is an immersion at every point $p \in M$.

Example 4.12 (Dense immersion of $\mathbb{R}^1$ into T^2)**.** Consider the 2-torus $T^2 = S^1 \times S^1$ and let α be a real number. Define the map $f : \mathbb{R}^1 \to T^2$ by

$$f(t) = (e^{2\pi i t}, e^{2\pi i \alpha t}).$$

This map is an immersion. To see this, consider the natural angular coordinates on T^2. Let (θ, ϕ) be coordinates on the universal cover $\mathbb{R}^2$ of T^2, with the identification $(\theta, \phi) \sim (\theta + m, \phi + n)$ for $m, n \in \mathbb{Z}$.

In these coordinates, the map f becomes

$$f(t) = (t, \alpha t) \mod \mathbb{Z}^2.$$

Now, at any point $t_0 \in \mathbb{R}$, we can choose a neighborhood U of t_0 small enough so that the projection $\mathbb{R}^2 \to T^2$ restricts to a diffeomorphism on $(t_0 - \varepsilon, t_0 + \varepsilon) \times (\alpha t_0 - \varepsilon, \alpha t_0 + \varepsilon)$ for some $\varepsilon > 0$. In this local coordinate chart, the map is simply

$$f(t) = (t, \alpha t),$$

and its derivative is

$$T_{t_0} f = \begin{pmatrix} 1 \\ \alpha \end{pmatrix},$$

which has full rank 1. Therefore, f is an immersion.

Note that when α is irrational, the image $f(\mathbb{R}^1)$ is dense in T^2. This follows from the fact that the irrational rotation on the circle is minimal (every orbit is dense). More precisely, for any open set $U \subset T^2$, there exists some $t \in \mathbb{R}$ such that $f(t) \in U$.

4.3 Submersions

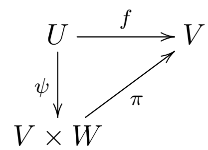
Theorem 4.13. *The following are equivalent:*

1. $d_p f$ *is surjective.*

2. *There exist open neighborhoods* U *of* p*,* V *of* q *and* W *of* 0 *(in* $\mathbb{R}^{m-n}$*) and a diffeomorphism* $\psi : U \to V \times W$ *such that:*

(a) $f(U) = V$,

(b) If π *denotes the projection* $V \times W \to V$, *then the following diagram commutes:*



3. (M, N, p, q, f) *looks locally like a linear surjection* $\bar{f} : \mathbb{R}^m \to \mathbb{R}^n$.

4. *There exist local coordinates* $\{x^i\}$ *at* p *and* $\{y^i\}$ *at* q *such that* $x^i = y^i \circ f$ *for* $1 \leq i \leq n$.

5. *There exist open neighborhoods* U *of* p *and* V *of* q *and a smooth map* $\sigma : V \to U$ *such that* $f(U) \subset V$ *and* $f \circ \sigma = \mathrm{id}_V$.

Proof. The proof is similar to the proof of the corresponding theorem on immersions.

We only need to show (1) ⇒ (2). Since the question is local, we may assume that the following conditions are satisfied:

a. M is an open subset of $\mathbb{R}^m$ and N is an open subset of $\mathbb{R}^n$,

b. $p = 0$ and $\operatorname{Ker} d_p f = \{0\} \times \mathbb{R}^{m-n} \subset \mathbb{R}^n \times \mathbb{R}^{m-n} = \mathbb{R}^m$.

Let W be $\{0\} \times \mathbb{R}^{m-n} \subset \mathbb{R}^m$. Define $\psi : M \to N \times W$ by $\psi(x) = (f(x), (x^{n+1}, \ldots, x^m))$. Then by the inverse function theorem, ψ is a local diffeomorphism at $p = 0$. Hence, by shrinking M, N and W, we may assume that ψ is a diffeomorphism. □

Definition 4.14. A smooth map f satisfying the equivalent conditions of the preceding theorem at p is called a *submersion at* p. A smooth map f which is a submersion at all $p \in M$ is called a *submersion*.

Example 4.15 (Projection from product manifold)**.** Let M and N be smooth manifolds. Consider the projection map $\pi : M \times N \to M$ defined by $\pi(p, q) = p$.

This map is a submersion. To see this, note that for any point $(p, q) \in M \times N$, the derivative $T_{(p,q)}\pi$ is surjective. Indeed, we can use condition (5) of the theorem by taking the section $\sigma : M \to M \times N$ defined by $\sigma(p) = (p, q_0)$ for any fixed $q_0 \in N$. Then $\pi \circ \sigma = \mathrm{id}_M$, so π is a submersion.

Example 4.16 (Projection to projective space)**.** Consider the map $\pi : \mathbb{R}^n \setminus \{0\} \to \mathbb{RP}^{n-1}$ defined by $\pi(x) = [x]$, where $[x]$ denotes the line through the origin containing x.

This map is a submersion. To see this, consider the coordinate chart $U_1 = \{[x_1 : \cdots : x_n] \mid x_1 \neq 0\}$ on $\mathbb{RP}^{n-1}$. On U_1, we have coordinates $(y_2, \ldots, y_n)$ with $y_j = x_j/x_1$ for $j = 2, \ldots, n$.

Define a smooth section $\sigma : U_1 \to \mathbb{R}^n \setminus \{0\}$ by

$$\sigma(y_2, \ldots, y_n) = (1, y_2, \ldots, y_n).$$

Then we have

$$\pi \circ \sigma(y_2, \ldots, y_n) = [1 : y_2 : \cdots : y_n] = (y_2, \ldots, y_n),$$

where the last equality holds because in the coordinate chart U_1, the point $[1 : y_2 : \cdots : y_n]$ corresponds exactly to $(y_2, \ldots, y_n)$.

Therefore, $\pi \circ \sigma = \mathrm{id}_{U_1}$, and by condition (5) of the theorem, π is a submersion on $\pi^{-1}(U_1)$. Since $\mathbb{RP}^{n-1}$ is covered by similar coordinate charts $U_i = \{[x_1 : \cdots : x_n] \mid x_i \neq 0\}$, and on each such chart we can define an analogous section, it follows that π is a submersion on all of $\mathbb{R}^n \setminus \{0\}$.

More generally, if we consider the space of full-rank $m \times n$ matrices ($m \geq n$) and the projection to the Grassmannian $\mathrm{Gr}(n, m)$ which sends a matrix to its column space, this map is also a submersion, and a similar local section argument can be used to prove it.

4.3.1 Remarks

1. Sometimes the phrase "f has maximal rank" (meaning $d_p f$ is injective if $m \leq n$ and $d_p f$ is surjective if $m \geq n$) is used to include both concepts.

2. An *embedding* is a smooth map f such that:

   (a) f is an immersion,

   (b) $f : M \to f(M)$ is a homeomorphism.

4.4 Subimmersions

Definition 4.17. f is a *subimmersion at p* if the following equivalent conditions are satisfied:

1. f looks locally like a composition $\bar{M} \xrightarrow{s} \bar{Z} \xrightarrow{\iota} \bar{N}$ where s is a submersion and ι is an immersion.

2. f looks locally like a linear map $\bar{f} : \mathbb{R}^m \to \mathbb{R}^n$.

A smooth map f which is a subimmersion at all $p \in M$ is called a *subimmersion*.

Remark 4.18. 1. The set of points $p \in M$ where a smooth map $f : M \to N$ is an immersion (resp. a submersion, a subimmersion) is *open* in M.

2. The composition of two immersions (resp. submersions) is an immersion (resp. a submersion). The analogous statement for subimmersions is false.

3. In the definition of subimmersion, the order (submersion then immersion) ensures stable rank behavior. For a linear map $\bar{f} : \mathbb{R}^m \to \mathbb{R}^n$ factoring as $\mathbb{R}^m \xrightarrow{s} \mathbb{R}^r \xrightarrow{\iota} \mathbb{R}^n$ with s surjective and ι injective, the rank is r. Reversing the order (immersion then submersion) may yield different ranks under local coordinates, as seen in the example $x \mapsto (x, 0)$ composed with different projections.

Theorem 4.19. *The following are equivalent:*

1. *f is a subimmersion at p.*

2. $\mathrm{rank}\, T_{p'} f$ *is constant for $p' \in U$ and U some neighborhood of p.*

Proof. (1) ⇒ (2): Clear.

(2) ⇒ (1): Let $r = \dim \operatorname{Im} d_p f$. Then, since the question is local, we may assume the following conditions are satisfied:

a. $N = V_1 \times V_2$ is open in $\mathbb{R}^r \times \mathbb{R}^{n-r}$,

b. $f(p) = 0$ and $\operatorname{Im} d_p f = \mathbb{R}^r \times \{0\}$.

Let $\pi : \mathbb{R}^r \times \mathbb{R}^{n-r} \to \mathbb{R}^r$ be the projection on the first factor. Then $\pi \circ f$ is a submersion. Hence we may further assume that:

a. $M = V_1 \times U_2$ is open in $\mathbb{R}^r \times \mathbb{R}^{m-r}$,

b. $\pi \circ f : V_1 \times U_2 \to V_1$ is the projection on the first factor.

The map f then has the following form:

$$f(x_1, x_2) = (x_1, \psi(x_1, x_2))$$

Finally, since $T_{p'} f$ has locally constant rank, we may assume that the rank of $T_{p'} f$ is in fact constant on $V_1 \times U_2$ (rank $= r$).

We claim that ψ must be independent of x_2 in a neighborhood of zero. Indeed, $D_2\psi(x_1, x_2) = 0$ since otherwise f would have rank greater than r at (x_1, x_2). Our claim is therefore a consequence of the following lemma:

Lemma 4.20. *Let $f : U \times V \to \mathbb{R}$ be a smooth function such that $D_2 f$ is identically 0. Then f is locally independent of the V coordinate.*

We conclude the proof of the theorem by noting that f may now be written as $V_1 \times U_2 \to V_1 \to V_1 \times V_2$ where the first map is pr_1 and the second is $\mathrm{id}_{V_1} \times \psi$. The first map is a submersion and the second is an immersion. □

Remark 4.21. In differential geometry, the concept of subimmersion naturally arises in several important contexts. In this course, we will encounter one particularly fundamental example:

Vector bundle theory: Consider a vector bundle homomorphism $\Phi : E \to F$ between two vector bundles over the same base manifold M. If the rank of the linear map $\Phi_p : E_p \to F_p$ induced on each fiber is constant, then the bundle homomorphism Φ (when viewed as a map between smooth manifolds) is a subimmersion.

This example is of fundamental importance. The subimmersion property ensures that we can define the **kernel bundle** $\ker(\Phi)$, the **image bundle** $\operatorname{image}(\Phi)$, and the **cokernel bundle** $\operatorname{coker}(\Phi)$. These constructions form the foundation of linear algebra over vector bundles and play crucial roles in many areas of differential geometry.

Another context (which we may mention if time permits) is **Lie group theory**. For Lie groups (manifolds with group structure), every continuous Lie group homomorphism $\phi : G \to H$ is automatically smooth and a subimmersion. While Lie groups form an important class of examples, their theory is rich enough to warrant a separate course.

In this course, the primary (and likely only) concrete examples of subimmersions we will encounter in detail will be constant-rank vector bundle homomorphisms.

4.5 Submanifold

Suppose M is a subspace of N (with the induced topology) and let

$$\iota : M \to N$$

be the inclusion map. We say M is **locally a submanifold of N at** $p \in M$ if either of the following equivalent conditions holds:

1. There exists an open neighborhood U of p in M, a chart (V, ψ) of N about p, and a linear subspace $E \subset \mathbb{R}^n$ such that $U \subset V$ and $\psi(U) = E \cap \psi(V)$.

2. There exist local coordinates $x^1, \dots, x^n$ defined near p in N and an integer $0 \le k \le n$ such that M is locally given by $x^1 = \cdots = x^k = 0$.

If M is locally a submanifold of N at every $p \in M$, we say M is a **submanifold of** N. We now justify this terminology.

Theorem 4.22. *The following are equivalent:*

1. *There exists a smooth structure on M such that ι is an immersion.*

2. *M is a submanifold of N.*

Proof. ($\Rightarrow$) This follows from the local coordinate characterization (item 4) of immersions.

($\Leftarrow$) Choose an open cover $\{U_i\}_{i\in I}$ of M such that for each $i \in I$, there exists a chart (V_i, ψ_i) of N and a linear subspace $E_i \subset \mathbb{R}^n$ satisfying $U_i \subset V_i$ and $\psi_i(U_i) = E_i \cap \psi_i(V_i)$.

Each U_i inherits a smooth structure making $\iota|_{U_i}$ an immersion. On overlaps $U_i \cap U_j$, these structures agree: if we take coordinates from (V_i, ψ_i) and (V_j, ψ_j), the transition map $\psi_j \circ \psi_i^{-1}$ is smooth on N and preserves the subspace structure defining M, hence $\{U_i, \varphi_i = \psi_i|_{U_i}\}_{i\in I}$ define a global smooth structure on M for which ι is an immersion.

Alternatively, once we prove the uniqueness theorem later, we could argue that both smooth structures make $\iota|_{U_i\cap U_j}$ an immersion, hence they must coincide by uniqueness. □

Now we address the uniqueness of this smooth manifold structure:

Theorem 4.23. *Let M be a topological space, N a smooth manifold, and $f : M \to N$ a continuous map. If there exists a smooth structure on M making f an immersion, then this smooth structure is unique.*

Proof. Suppose $\mathcal{A}$ is a smooth structure on M such that $f : (M, \mathcal{A}) \to N$ is an immersion. We claim that for any manifold P, a map $g : P \to (M, \mathcal{A})$ is smooth if and only if $f \circ g : P \to N$ is smooth.

The "only if" direction is clear. For the "if" direction, we work locally. Let $r \in P$, $p = g(r)$, and $q = f(p)$. Since f is an immersion, there exist neighborhoods U of p in M and V of q in N, and a smooth map $h : V \to U$ such that $h \circ f|_U = \mathrm{id}_U$.

By continuity of f, we can find a neighborhood W of r such that $g(W) \subset U$. Then on W we have:

$$g|_W = \mathrm{id}_U \circ g|_W = h \circ f \circ g|_W.$$

Since $f \circ g$ is smooth by assumption, $g|_W = h \circ (f \circ g|_W)$ is smooth. As smoothness is local, g is smooth everywhere.

Now, if $\mathcal{A}'$ is another smooth structure on M making f an immersion, then for any manifold P and map $g : P \to M$:

$$g \in C^\infty(P, (M, \mathcal{A}')) \iff f \circ g \in C^\infty(P, N) \iff g \in C^\infty(P, (M, \mathcal{A})).$$

By the characterization of smooth structures via smooth maps, this implies $\mathcal{A} = \mathcal{A}'$. □

Example 4.24. Recall the two standard smooth structures on S^1: one defined by four charts via coordinate projections, and another by two charts via stereographic projection. Since both atlases realize S^1 as a submanifold of $\mathbb{R}^2$, the theorem shows that they are compatible.

Example 4.25. Let M and N be topological manifolds with $f : M \to N$ a homeomorphism. If N has a smooth structure, then there exists a unique smooth structure on M making f a diffeomorphism.

Proof. For existence, pull back the smooth structure from N via f: if $\{(V_\alpha, \psi_\alpha)\}$ is a smooth atlas for N, then $\{(f^{-1}(V_\alpha), \psi_\alpha \circ f)\}$ is a smooth atlas for M.

For uniqueness, if $\mathcal{A}_1$ and $\mathcal{A}_2$ both make f a diffeomorphism, then the identity map $\mathrm{id}_M = f^{-1} \circ f$ is a diffeomorphism between $(M, \mathcal{A}_1)$ and $(M, \mathcal{A}_2)$, so $\mathcal{A}_1 = \mathcal{A}_2$. Alternatively, this follows from the uniqueness theorem for immersions. □

The uniqueness theorem also has a submersion counterpart.

Theorem 4.26. *Let M be a smooth manifold, N a topological space, and $f : M \to N$ a surjective map (not necessarily continuous a priori). If there exists a smooth structure on N such that f is a submersion, then this smooth structure is unique.*

Proof. Suppose $\mathcal{A}$ is a smooth structure on N such that $f : M \to N_{\mathcal{A}}$ is a submersion. We claim that for any manifold P, a map $g : N_{\mathcal{A}} \to P$ is smooth if and only if $g \circ f : M \to P$ is smooth.

The "only if" direction is clear. For the "if" direction, we work locally. Let $q \in N$ and choose $p \in M$ with $f(p) = q$ by the surjectivity of f. Since f is a submersion, , there exist neighborhoods U of p in M and V of q in N, and a smooth map $h : V \to U$ such that $f \circ h|_V = \mathrm{id}_V$.

The map $g \circ f$ being smooth implies that $g \circ \mathrm{id}_V = g \circ f \circ h|_V$ is smooth. hence g is smooth in a neighborhood of q. As smoothness is local, g is smooth everywhere.

Now, if $\mathcal{A}'$ is another smooth structure on N making f a submersion, then for any manifold P and map $g : N \to P$:

$$g \in C^\infty(N_{\mathcal{A}'}, P) \iff g \circ f \in C^\infty(M, P) \iff g \in C^\infty(N_{\mathcal{A}}, P).$$

By the same characterization of smooth structures, this again implies $\mathcal{A} = \mathcal{A}'$. □

Remark 4.27. From the proof of this theorem, we see that a surjective submersion plays a role analogous to a quotient map in topology. To verify that a map from a quotient manifold to another manifold is smooth, it suffices to check that its composition with the submersion is smooth.

A direct example is provided by homogeneous polynomials on Euclidean space. Consider the projection $\pi : \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{RP}^n$ onto projective space. If $F : \mathbb{R}^{n+1} \to \mathbb{R}$ is a homogeneous polynomial, then it naturally induces a map $f : \mathbb{RP}^n \to \mathbb{R}$ defined by $f([x]) = F(x)$. To verify that f is smooth, we only need to check that $f \circ \pi = F|_{\mathbb{R}^{n+1}\setminus\{0\}}$ is smooth, which is true since polynomials are smooth.

4.6 Embedding

Recall that a subset M of a smooth manifold N is a *smooth submanifold at p* if there exist local coordinates $x^1, \ldots, x^n$ defined near p in N and an integer $0 \leq k \leq n$ such that M is locally given by $x^1 = \cdots = x^k = 0$. If M is locally a smooth submanifold of N at every point $x \in M$, we say M is a *(smooth) submanifold* of N. This definition captures the "regularity" of the subset as it sits inside N.

The preceding discussion justifies the term "submanifold": there is a unique smooth structure on M compatible with the subspace topology for which the inclusion map $\iota : M \to N$ is an immersion.

Now consider a smooth immersion $f : M \to N$ such that $f : M \to f(M)$ is a homeomorphism, where $f(M)$ carries the subspace topology inherited from N. In this case, $f(M)$ inherits a smooth submanifold structure from M, and we call such an immersion an *embedding*.

To see this explicitly, take any $q \in f(M) \subset N$ and choose $p \in M$ with $f(p) = q$. Since f is an immersion, there exist coordinate charts $\varphi = (x^1, \ldots, x^m) : U \to \mathbb{R}^m$ around p and $\psi = (y^1, \ldots, y^n) : V \to \mathbb{R}^n$ around q with $f(U) \subset V$, such that:

$$y^i \circ f = x^i \quad \text{for } 1 \leq i \leq m,$$

and

$$y^i \circ f = 0 \quad \text{for } m+1 \leq i \leq n.$$

Moreover, since f is a homeomorphism onto its image, there exists an open neighborhood $W \subset N$ of q such that $W \cap f(M) = W \cap f(U)$. Therefore, within $W \cap V$, the set $f(M)$ is precisely given by $y^{m+1} = \cdots = y^n = 0$, confirming that $f(M)$ is a smooth submanifold of N.

4.7 Preimage Construction

Recall condition (4) in the characterization of submersions: a smooth map $f : M \to N$ is a submersion at p if there exist local coordinates $\{x^i\}$ around p and $\{y^i\}$ around $f(p)$ such that $x^i = y^i \circ f$ for $1 \leq i \leq n$ (where $n = \dim N$).

This local form immediately implies that the fiber $f^{-1}(q)$ is locally given by $x^1 = \cdots = x^n = 0$ near p, hence is a smooth submanifold of dimension $m - n$ at p. This observation motivates the following definition.

Definition 4.28. Let $f : M \to N$ be a smooth map. A point $q \in N$ is called a *regular value* of f if f is a submersion at every point $x \in f^{-1}(q)$.

The discussion above leads directly to the following fundamental result:

Theorem 4.29 (Preimage Theorem). *If $f : M \to N$ is a smooth map and $q \in N$ is a regular value, then $f^{-1}(q)$ is a smooth submanifold of M of dimension* $\dim M - \dim N$. *Moreover, for every $p \in f^{-1}(q)$, the tangent space satisfies*

$$T_p(f^{-1}(q)) = \ker(d_p f : T_p M \to T_q N).$$

It is often convenient to think not in terms of the absolute dimension of $f^{-1}(q)$, but rather its *codimension*—the amount by which its dimension is less than that of M. In the theorem above, $f^{-1}(q)$ has codimension $n = \dim N$.

Example 4.30 (The Sphere). The $(n-1)$-sphere $S^{n-1} \subset \mathbb{R}^n$ can be realized as $f^{-1}(1)$, where $f : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$f(x_1, \ldots, x_n) = x_1^2 + \cdots + x_n^2.$$

This map is smooth, and its derivative at any point is $[2x_1, \ldots, 2x_n]$. Since this is nonzero for all $(x_1, \ldots, x_n) \neq (0, \ldots, 0)$, every nonzero real number is a regular value of f. In particular, 1 is a regular value, so S^{n-1} is a smooth manifold of dimension $n-1$ (or codimension 1).

Example 4.31 (Orthogonal Group). The orthogonal group $O(n)$, consisting of $n \times n$ orthogonal matrices, can be described as $f^{-1}(I)$, where $f : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{Sym}(n)$ is defined by $f(A) = A^T A$. Here $\mathrm{Sym}(n)$ denotes the space of symmetric $n \times n$ matrices. One can verify that I is a regular value of f, hence $O(n)$ is a smooth submanifold of $\mathrm{GL}(n, \mathbb{R})$ of dimension $\frac{1}{2}n(n-1)$.

Definition 4.32. Let $f : M \to N$ be a smooth map. A point $q \in N$ is called a *critical value* of f if it is not a regular value.

Example 4.33. If $\dim M = m < n = \dim N$ and $f : M \to N$ is a smooth map, then every point in the image $f(M)$ is a critical value of f. This is because at any $p \in M$, the linear map $d_p f : T_p M \to T_{f(p)} N$ cannot be surjective for dimensional reasons.

The preimage construction provides a powerful method for producing new manifolds from old ones, often avoiding the need for explicit coordinate charts. The sphere and orthogonal group examples demonstrate how naturally occurring geometric objects can be recognized as smooth manifolds through this approach.

4.8 Transversality

We now address the natural question: when is the intersection of two submanifolds again a submanifold? In general, the intersection of two submanifolds can be quite pathological. For instance, when discussing partitions of unity, we showed that any closed subset $K \subset M$ of a manifold M can arise as the zero set of a non-negative smooth function $f \in C^\infty(M)$.

Let us consider the graph of this function:

$$\Gamma_f := \{(x, f(x)) \in M \times \mathbb{R} : x \in M\},$$

which is a smooth submanifold of $M \times \mathbb{R}$. On the other hand, we have the zero section:

$$\Gamma_0 := M \times \{0\},$$

which is also a smooth submanifold. Their intersection is

$$\Gamma_f \cap \Gamma_0 = \{(x, 0) \in M \times \mathbb{R} : f(x) = 0\} = K \times \{0\},$$

which, in general, may have no manifold structure at all. This shows that submanifold intersections need not themselves be submanifolds.

We will now introduce a sufficient condition that ensures the intersection of two submanifolds is again a submanifold. This condition is called *transversality.* The intuitive idea is that if two submanifolds intersect "as little as possible," their intersection will behave nicely and inherit a manifold structure.

Before giving the definition, we recall a basic lemma from linear algebra:

Lemma 4.34. *Let V be a finite-dimensional vector space and $V_1, V_2 \subset V$ be subspaces. Then the sequence*

$$0 \to V_1 \cap V_2 \xrightarrow{i} V \xrightarrow{j} V/V_1 \oplus V/V_2 \xrightarrow{k} V/(V_1 + V_2) \to 0$$

is exact, where:

- *$i : V_1 \cap V_2 \to V$ is the natural inclusion*
- *$j : V \to V/V_1 \oplus V/V_2$ is given by $j(v) = (v + V_1, v + V_2)$*
- *$k : V/V_1 \oplus V/V_2 \to V/(V_1 + V_2)$ is given by $k(v_1 + V_1, v_2 + V_2) = (v_1 - v_2) + (V_1 + V_2)$*

Proof of Lemma. We verify exactness at each term:

1. At $V_1 \cap V_2$: The map i is injective, so $\ker i = 0$.
2. At V: We have $\ker j = \{v \in V : v + V_1 = 0 \text{ and } v + V_2 = 0\} = V_1 \cap V_2 = \operatorname{im} i$.
3. At $V/V_1 \oplus V/V_2$:
   - $\operatorname{im} j \subset \ker k$: For any $v \in V$, $k(j(v)) = k(v + V_1, v + V_2) = (v - v) + (V_1 + V_2) = 0$.
   - $\ker k \subset \operatorname{im} j$: Suppose $k(v_1 + V_1, v_2 + V_2) = 0$. Then $v_1 - v_2 \in V_1 + V_2$, so we can write $v_1 - v_2 = w_1 + w_2$ with $w_i \in V_i$. Let $v = v_1 - w_1 = v_2 + w_2$. Then $j(v) = (v + V_1, v + V_2) = (v_1 + V_1, v_2 + V_2)$.
4. At $V/(V_1 + V_2)$: The map k is surjective since for any $w + (V_1 + V_2)$, we have $k(w + V_1, 0 + V_2) = w + (V_1 + V_2)$.

□

Theorem 4.35. *Let M be a manifold, N_1 and N_2 be submanifolds of M, and $p \in N_1 \cap N_2$. If $T_pM = T_pN_1 + T_pN_2$, then there exists a chart (U, φ) at p of M such that*

$$\varphi(U) = V_1 \times V_2 \times W$$
$$\varphi(U \cap N_1) = \{0\} \times V_2 \times W$$
$$\varphi(U \cap N_2) = V_1 \times \{0\} \times W.$$

Equivalently, there exists a coordinate system $x^1, \ldots, x^n$ at p and integers $r_1, r_2 \geq 0$ with $r_1 + r_2 \leq n$ such that:

$$N_1 \text{ is given by } x^1 = \ldots = x^{r_1} = 0 \text{ in a neighborhood of } p,$$
$$N_2 \text{ is given by } x^{r_1+1} = \ldots = x^{r_1+r_2} = 0 \text{ in a neighborhood of } p.$$

Proof. Since N_1 and N_2 are submanifolds of M, we can find submersions

$$f_1 : M \to \mathbb{R}^{r_1}, \quad f_2 : M \to \mathbb{R}^{r_2}$$

such that $N_i = f_i^{-1}(0)$ for $i = 1, 2$. Let $(x^1, \ldots, x^{r_1})$ and $(x^{r_1+1}, \ldots, x^{r_1+r_2})$ be the coordinate components of f_1 and f_2, respectively.

Apply Lemma 4.34 to $V = T_pM$, $V_1 = T_pN_1$, $V_2 = T_pN_2$:

$$0 \to T_p(N_1 \cap N_2) \xrightarrow{i} T_pM \xrightarrow{j} T_pM/T_pN_1 \oplus T_pM/T_pN_2 \xrightarrow{k} T_pM/(T_pN_1 + T_pN_2) \to 0.$$

The condition $T_pM = T_pN_1 + T_pN_2$ implies $T_pM/(T_pN_1 + T_pN_2) = 0$, so the sequence becomes short exact:

$$0 \to T_p(N_1 \cap N_2) \xrightarrow{i} T_pM \xrightarrow{j} T_pM/T_pN_1 \oplus T_pM/T_pN_2 \to 0.$$

In particular, j is surjective.

Now, the differential $d_p(f_1, f_2) : T_pM \to T_0\mathbb{R}^{r_1} \oplus T_0\mathbb{R}^{r_2}$ factors through j and the isomorphisms

$$T_pM/T_pN_1 \oplus T_pM/T_pN_2 \cong T_0\mathbb{R}^{r_1} \oplus T_0\mathbb{R}^{r_2}$$

induced by the submersions f_i. More precisely, we have a factorization:

$$d_p(f_1, f_2) = \Phi \circ j,$$

where $\Phi : T_pM/T_pN_1 \oplus T_pM/T_pN_2 \to T_0\mathbb{R}^{r_1} \oplus T_0\mathbb{R}^{r_2}$ is an isomorphism.

Since j is surjective and Φ is an isomorphism, it follows that $d_p(f_1, f_2)$ is also surjective. Therefore, $(f_1, f_2) : M \to \mathbb{R}^{r_1} \times \mathbb{R}^{r_2}$ is a submersion at p.

By the submersion theorem, there exists a coordinate system $(x^1, \ldots, x^n)$ around p such that (f_1, f_2) corresponds to the projection onto the first $r_1 + r_2$ coordinates. In these coordinates:

- $N_1 = f_1^{-1}(0)$ is given by $x^1 = \cdots = x^{r_1} = 0$
- $N_2 = f_2^{-1}(0)$ is given by $x^{r_1+1} = \cdots = x^{r_1+r_2} = 0$

This completes the proof. □

If N_1 and N_2 satisfy the condition of the preceding theorem at p, we say that N_1 and N_2 are **transversal at** p.

Corollary 4.36. *Suppose N_1 and N_2 are transversal at p. Then:*

1. *N_1 and N_2 are transversal in a neighborhood of p.*
2. *$N_1 \cap N_2$ is locally a submanifold of M at p.*
3. *$T_p(N_1 \cap N_2) = T_pN_1 \cap T_pN_2$.*

Proof. The first statement follows from the continuity of the transversality condition. The second and third statements are immediate consequences of the coordinate representation in the preceding theorem. □

4.9 Fibre Products, Pullbacks, and Cartesian Squares

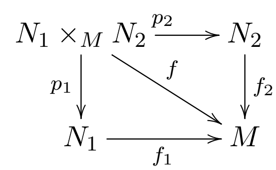
We now present a unified framework that generalizes both the preimage construction and the transversal intersection of submanifolds: the concept of *fibre products* or *pullbacks*.

4.9.1 Definition and Basic Properties

Consider a pair of smooth maps $f_i : N_i \to M$, $i = 1, 2$. Their **fibre product** (also called **pullback**) is defined as:

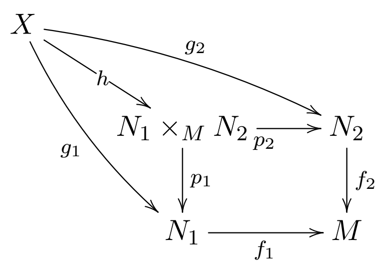
$$N_1 \times_M N_2 = \{(y_1, y_2) \in N_1 \times N_2 : f_1(y_1) = f_2(y_2)\}.$$

Let $p_i : N_1 \times_M N_2 \to N_i$ be the natural projections, and define $f = f_1 \circ p_1 = f_2 \circ p_2$. This yields a commutative diagram:



In category theory, such a square is called a **cartesian square** or **pullback square** because it satisfies the following universal property:

> For any manifold X with maps $g_1 : X \to N_1$ and $g_2 : X \to N_2$ such that $f_1 \circ g_1 = f_2 \circ g_2$, there exists a unique map $h : X \to N_1 \times_M N_2$ making the following diagram commute:



4.9.2 Transversality and Smoothness

In the smooth category, the fibre product $N_1 \times_M N_2$ may not be a manifold in general. To ensure smoothness, we need a transversality condition.

Let $(y_1, y_2) \in N_1 \times_M N_2$ and let $p = f(y_1, y_2)$. We say that f_1 and f_2 are **transverse at** (y_1, y_2) if

$$T_pM = \mathrm{Im}(T_{y_1} f_1) + \mathrm{Im}(T_{y_2} f_2).$$

we write $f_1 \pitchfork_p f_2$ indicating that the transversality condition holds at the point p, that is , f_1 and $_2$ are transverse at $\forall (y_1, y_2) \in f^{-1}(p)$. A commutative square of the above form is called **transversal cartesian at** p if this transversality condition holds.

As usual, if this condition holds for every $(y_1, y_2) \in N_1 \times_M \times N_2$, we simply say that f_1 and f_2 are **transverse** and write

$$f \pitchfork g.$$

Theorem 4.37 (Smoothness of Transverse Fibre Products)**.** *Suppose f_1 and f_2 are transverse at (y_1, y_2). Then:*

1. *f_1 and f_2 are transverse at all points in a neighborhood of (y_1, y_2) in $N_1 \times_M N_2$.*
2. *$N_1 \times_M N_2$ is locally a submanifold of $N_1 \times N_2$ at (y_1, y_2).*
3. *The tangent space is given by the fibre product:*

$$T_{(y_1,y_2)}(N_1 \times_M N_2) = T_{y_1}N_1 \times_{T_pM} T_{y_2}N_2,$$
$$T_{(y_1,y_2)}(N_1 \times_M N_2) = \{(v_1, v_2) \in T_{y_1}N_1 \times T_{y_2}N_2 : T_{y_1}f_1(v_1) = T_{y_2}f_2(v_2)\}.$$

Proof. Set $N = N_1 \times N_2$ and $P = N_1 \times_M N_2$. Define maps $g_i : N \to N \times M$ by:

$$g_1(y_1, y_2) = ((y_1, y_2), f_1(y_1)), \quad g_2(y_1, y_2) = ((y_1, y_2), f_2(y_2)).$$

Let $g = g_i|_P$. Then:

1. $g_1(N)$ and $g_2(N)$ are transverse at $g(y_1, y_2)$.
2. $g(P) = g_1(N) \cap g_2(N)$.

The result follows by applying the transverse intersection theorem to the submanifolds $g_1(N)$ and $g_2(N)$ of $N \times M$. □

4.9.3 Special Cases and Examples

The fibre product construction unifies several important concepts:

Example 4.38 (Preimage as Fibre Product)**.** Given $f : N \to M$ and a point $q \in M$, consider f and the inclusion $i : \{q\} \hookrightarrow M$. Their fibre product is:

$$N \times_M \{q\} = \{(y, q) \in N \times \{q\} : f(y) = q\} \cong f^{-1}(q).$$

The transversality condition becomes the requirement that q is a regular value of f.

Example 4.39 (Intersection as Fibre Product)**.** For submanifolds $N_1, N_2 \subset M$ with inclusions $\iota_i : N_i \hookrightarrow M$, their fibre product is:

$$N_1 \times_M N_2 = \{(y_1, y_2) \in N_1 \times N_2 : \iota_1(y_1) = \iota_2(y_2)\} \cong N_1 \cap N_2.$$

The transversality condition is exactly

$$T_pM = T_pN_1 + T_pN_2, \qquad p \in N_1 \cap N_2.$$

When this condition holds, we say that N_1 and N_2 intersect transversely and write

$$N_1 \pitchfork N_2.$$

Example 4.40 (Product as Fibre Product)**.** When M is a single point, the fibre product reduces to the ordinary product:

$$N_1 \times_{\{*\}} N_2 \cong N_1 \times N_2.$$

Example 4.41 (Graph of a Map)**.** Given $f : N_1 \to N_2$, consider the maps $f : N_1 \to N_2$ and $\mathrm{id} : N_2 \to N_2$. Their fibre product is the graph of f:

$$N_1 \times_{N_2} N_2 = \{(y_1, y_2) \in N_1 \times N_2 : f(y_1) = y_2\} =: \Gamma_f.$$

4.10 The Morse-Sard Theorem

The usefulness of the preimage construction and of the transversality condition lies in the fact that the hypotheses are, in a very strong sense, "typically" satisfied. The Morse-Sard theorem guarantees that the set of non-regular values has measure zero, so for a fixed smooth map, almost every value is regular.

Theorem 4.42 (Morse–Sard)**.** *Let $f \in C^\infty(U, \mathbb{R}^n)$, where $U \subset \mathbb{R}^m$ is open. Then the set of critical values of f has Lebesgue measure 0 in $\mathbb{R}^n$.*

Proof. The proof proceeds by induction on m. Assume the theorem is already known for dimension $m-1$ whenever $m > 1$.

For $j \geq 1$, define

$$C_j = \{x \in U : f'(x) = 0,\ f''(x) = 0,\ \dots, f^{(j)}(x) = 0\}.$$

We first show:

$$\mathcal{L}^n(f(C_j)) = 0 \qquad \text{whenever } (j+1)n > m. \tag{4.1}$$

It suffices to prove $\mathcal{L}^n(f(K \cap C_j)) = 0$ for a compact cube $K \subset U$ of side length ℓ. Divide K into k^m subcubes of side $\varepsilon = \ell/k$. Let $I_1, \dots, I_N$ be the subcubes intersecting C_j, and choose $x_t \in I_t \cap C_j$.

Taylor's theorem and $x_t \in C_j$ give

$$|f(x) - f(x_t)| \leq A|x - x_t|^{j+1} \leq A\varepsilon^{j+1}, \qquad x \in I_t.$$

If $(j+1)n > m$, then

$$\mathcal{L}^n(f(I_t)) \leq A^n \varepsilon^{(j+1)n} \leq A^n \varepsilon^{m+1} = A^n \varepsilon\, \mathcal{L}^m(I_t).$$

Summing over t,

$$\mathcal{L}^n(f(K \cap C_j)) \leq \sum_{t=1}^{N} \mathcal{L}^n(f(I_t)) \leq A^n \varepsilon \sum_{t=1}^{N} \mathcal{L}^m(I_t) \leq A^n \ell^m \varepsilon.$$

Letting $\varepsilon \to 0$ proves (4.1).

Next, note that the set

$$E_j := C_j \setminus C_{j+1}$$

is contained in a smooth submanifold of codimension 1 near each of its points. Indeed, at $x_0 \in E_j$ there exists a component g of $f^{(j)}$ with $dg(x_0) \neq 0$. Thus, E_j lies locally in the level set $g^{-1}(g(x_0))$, which is a smooth $(m-1)$-dimensional submanifold S.

If f has a critical point on S, then this point is also a critical point of $f|_S$. By the induction hypothesis (applied to the dimension $m-1$ domain), the set of critical values of $f|_S$ has measure zero. Since E_j is covered by countably many such neighborhoods, we obtain

$$\mathcal{L}^n(f(C_j \setminus C_{j+1})) = 0.$$

Finally, we handle the set $C \setminus C_1$, where C is the full critical set. Since C is invariant under precomposition with local diffeomorphisms, we may work in coordinates. At a point of $C \setminus C_1$ where $\frac{\partial f_1}{\partial x_1} \neq 0$, let ψ be the inverse of the map

$$x \mapsto (f_1(x), x_2, \ldots, x_m).$$

Then

$$f \circ \psi(y) = (y_1, g(y)), \qquad g : \mathbb{R}^m \to \mathbb{R}^{n-1}.$$

A point $y = (y_1, y')$ is critical for $f \circ \psi$ iff y' is a critical point of the map $y' \mapsto g(y_1, y')$. For each fixed y_1, the set of critical values of $f \circ \psi$ in the slice $\{y_1\} \times \mathbb{R}^{n-1}$ has measure zero. Since the critical set in a compact K is compact, the image under f is compact and hence measurable. By Fubini's theorem, for each compact K

$$\mathcal{L}^n(f(C \setminus C_1) \cap f(K)) = 0.$$

Thus $f(C \setminus C_1)$ has measure zero, completing the proof. □

Appendix: categorical properties of Cartesian squares

Proposition 4.43. *The following properties hold for transversal cartesian squares:*

Consider a commutative diagram of smooth manifolds:

$$\begin{array}{ccccc} A & \xrightarrow{f} & B & \xrightarrow{h} & C \\ {\scriptstyle g}\downarrow & & {\scriptstyle k}\downarrow & & \downarrow{\scriptstyle l} \\ D & \xrightarrow[m]{} & E & \xrightarrow[n]{} & F \end{array}$$

Then:

1. *If both squares are transversal cartesian, then the outer rectangle is transversal cartesian.*
2. *If the right square and the outer rectangle are transversal cartesian, then the left square is transversal cartesian.*

Proof. By the Pasting Lemma for pullbacks below, the corresponding statements for the underlying commutative squares as pullbacks in the category of smooth manifolds hold. It remains to verify the transversality conditions.

(1) Let $a \in A$, $b = f(a)$, $c = h(b)$, $d = g(a)$, $e = m(d) = k(b)$, $f = n(e) = l(c)$. Given the transversality conditions:

$$T_eE = \mathrm{Im}(T_bk) + \mathrm{Im}(T_dm)$$
$$T_fF = \mathrm{Im}(T_cl) + \mathrm{Im}(T_en)$$

For any $v \in T_fF$, write $v = T_cl(u) + T_en(w)$ with $u \in T_cC$, $w \in T_eE$. Then write $w = T_bk(x) + T_dm(y)$ with $x \in T_bB$, $y \in T_dD$. Thus $v = T_cl(u) + T_en(T_bk(x)) + T_en(T_dm(y)) = T_b(l \circ h)(u' + x) + T_d(n \circ m)(y)$ for some $u' \in T_bB$, showing $v \in \mathrm{Im}(T_a(l \circ h \circ f)) + \mathrm{Im}(T_d(n \circ m))$.

(2) Let $e = k(b) = m(d)$, $w \in T_eE$. Then $T_en(w) \in T_fF$. By outer rectangle transversality: $T_en(w) = T_c(l)(u) + T_d(n \circ m)(v)$ for some $u \in T_cC$, $v \in T_dD$. Thus $T_en(w) = T_c(l)(u) + T_en(T_dm(v))$. Since the right square is transversal cartesian, there is $x \in T_bB$ such that $T_bk(x) = w - T_dm(v)$ and $T_bh(x) = u$, giving $w = T_bk(x) + T_dm(v)$. □

Note: The properties in Proposition 4.43 are standard for cartesian squares (pullbacks) in category theory. However, for transversal cartesian squares we must additionally verify that the transversality condition is preserved under composition and decomposition of squares. The above proofs establish precisely that the transversality condition holds in these situations.

Lemma 4.44 (Pasting Lemma for Cartesian Squares)**.** *In any category with pullbacks, consider a commutative diagram:*

$$\begin{array}{ccccc} A & \xrightarrow{f} & B & \xrightarrow{h} & C \\ {\scriptstyle g}\downarrow & & {\scriptstyle k}\downarrow & & \downarrow{\scriptstyle l} \\ D & \xrightarrow[m]{} & E & \xrightarrow[n]{} & F \end{array}$$

Then:

1. *If both squares are cartesian, then the outer rectangle is cartesian.*

2. *If the right square and the outer rectangle are cartesian, then the left square is cartesian.*

Proof. We prove these statements using the universal property of pullbacks.

(1) Assume both inner squares are cartesian. We need to show the outer rectangle is cartesian, i.e., that A is the pullback of $D \to F$ and $C \to F$.

Let X be any object with maps $\alpha : X \to D$ and $\gamma : X \to C$ such that $n \circ m \circ \alpha = l \circ \gamma$. Since the right square is cartesian, there exists a unique map $\beta : X \to B$ such that $k \circ \beta = m \circ \alpha$ and $h \circ \beta = \gamma$.

Now, since the left square is cartesian, there exists a unique map $\delta : X \to A$ such that $g \circ \delta = \alpha$ and $f \circ \delta = \beta$.

This δ satisfies $g \circ \delta = \alpha$ and $(h \circ f) \circ \delta = \gamma$, showing that A is indeed the pullback.

(2) Assume the right square and outer rectangle are cartesian. We need to show the left square is cartesian.

Let X be any object with maps $\alpha : X \to D$ and $\beta : X \to B$ such that $m \circ \alpha = k \circ \beta$.

Consider the map $l \circ h \circ \beta : X \to F$. Note that $n \circ m \circ \alpha = n \circ k \circ \beta = l \circ h \circ \beta$, where the last equality follows from commutativity of the right square.

Since the outer rectangle is cartesian, there exists a unique map $\delta : X \to A$ such that $g \circ \delta = \alpha$ and $(h \circ f) \circ \delta = h \circ \beta$.

Now, both $f \circ \delta$ and β are maps from X to B that satisfy:

- $k \circ (f \circ \delta) = k \circ f \circ \delta = m \circ g \circ \delta = m \circ \alpha = k \circ \beta$

- $h \circ (f \circ \delta) = h \circ f \circ \delta = h \circ \beta$

Since the right square is cartesian, the map $(k, h) : B \to E \times_F C$ is injective on the level required by the universal property (equivalently, it is the pullback of l along n). Therefore $f \circ \delta = \beta$.

Thus δ satisfies $g \circ \delta = \alpha$ and $f \circ \delta = \beta$, showing that the left square is cartesian. □

4.11 Whitney embedding

Before turning to the Whitney Embedding Theorem, let us record a few remarks about Sard's theorem in the manifold setting. Last time we proved the Sard theorem for smooth maps defined on open subsets of Euclidean space. Since all our work from now on takes place on smooth manifolds, it is useful to clarify what "measure zero" and "almost everywhere" mean in that context.

Remark 4.45 (Measure zero sets on manifolds)**.** Although we have not formally introduced a measure on smooth manifolds, the concept of a measure-zero subset makes perfect sense. A subset $A \subset M$ is said to have measure zero if, for every coordinate chart (U, φ), the set

$$\varphi(A \cap U) \subset \mathbb{R}^m$$

has Lebesgue measure zero. Since diffeomorphisms preserve measure-zero sets, this definition is coordinate-independent. Thus it is meaningful to speak of properties holding "almost everywhere" on a manifold.

Remark 4.46 (Measure zero vs. meagre sets)**.** Measure-zero sets provide one notion of "smallness" of a subset. A different notion, coming from topology rather than measure theory, is that of a *first category* or *meagre* set: a subset is meagre if it is a countable union of nowhere dense sets (also known as a set of the first Baire category). Critical value sets in the Morse–Sard theorem are unions of compact measure-zero sets, hence are nowhere dense and therefore meagre. This notion extends naturally to smooth manifolds.

In infinite-dimensional settings (for example, Banach or Hilbert manifolds), Lebesgue measure is no longer available. The corresponding version of Sard's theorem is the *Sard–Smale theorem*, which states that the set of regular values of a smooth Fredholm map is residual (comeagre).

We continue the Whitney Embedding Theorem. This is a fundamental and beautiful result: although we have defined manifolds abstractly, the manifolds we visualize in our minds are almost always realized as submanifolds of Euclidean space. Whitney's theorem guarantees that our abstract definition does not produce objects that deviate significantly from this intuition.

Recall that in the proof of the partition of unity theorem, we constructed a locally finite collection $\{W_i\}$, where each W_i is contained in the domain U_i of some coordinate chart (U_i, φ_i). For each i we also produced a compactly supported smooth function η_i, which is identically 1 on a neighborhood of W_i and whose support lies inside U_i. If you like Lee's terminology, the sets W_i are called *regular coordinate balls.*

We begin with the following lemma, which will be the starting point for the construction of embeddings.

Lemma 4.47. *Let $K \subset M$ be a compact subset of an m-dimensional manifold M, and let U be a neighborhood of K. Then there exists a natural number n and a map*

$$g \in C_c^\infty(M, \mathbb{R}^n)$$

such that g is an injective immersion on a neighborhood of K, and $g \equiv 0$ outside U.

Proof. Choose finitely many sets W_j, $j = 1, \dots, k$, from the locally finite family constructed above so that they cover K. For each j, define

$$g_j = (\eta_j \varphi_j, \eta_j) \colon M \to \mathbb{R}^{m+1}.$$

Each g_j is an injective immersion on a neighborhood of $\overline{W}_j$. Therefore the direct sum

$$g = g_1 \oplus \cdots \oplus g_k$$

is an injective immersion on some neighborhood of K. Here

$$g = g_1 \oplus \cdots \oplus g_k$$

means the map whose value at x is the concatenation of the vectors $g_1(x), \dots, g_k(x)$. Thus each $g_j : M \to \mathbb{R}^{m+1}$ yields

$$g : M \to \mathbb{R}^{k(m+1)}, \qquad g(x) = (g_1(x), \dots, g_k(x)).$$

Finally, multiply g by a bump function which is identically 1 near K and vanishes outside U. This yields a compactly supported smooth map into $\mathbb{R}^n$ that is an injective immersion near K, as required. □

The weakness of this argument is that it gives no control over the target dimension n. When M is compact, this is harmless: we only need a single embedding into some $\mathbb{R}^n$. However, when M is noncompact, one typically exhausts M by larger and larger compact subsets, and the above construction forces us to choose larger values of n at each stage. If we proceed naïvely, we are eventually led to embeddings into an infinite-dimensional Euclidean space–an outcome that is neither geometric nor desirable.

We now turn to the question of *dimension reduction*, which overcomes this issue and leads to the finite-dimensional Whitney Embedding Theorem.

We first introduce a useful family of linear projections. For $a = (a_1, \dots, a_{n-1}) \in \mathbb{R}^{n-1}$ we denote by

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

the projection along the vector $(a, 1)$ onto the hyperplane $\{x_n = 0\}$, i.e.

$$\pi_a(x^1, \dots, x^{n-1}, x^n) = \big(x^1 - a_1 x^n, \dots, x^{n-1} - a_{n-1} x^n\big).$$

We now prove that, for a generic choice of a, such a projection preserves immersions on a fixed compact set.

Lemma 4.48. *Let M be an m-dimensional manifold, $K \subset M$ a compact set, and $f \in C^\infty(M, \mathbb{R}^n)$ an immersion on a neighborhood of K. Assume $n > 2m$. Then there is a closed measure-zero set $E \subset \mathbb{R}^{n-1}$ such that, for all $a \in \mathbb{R}^{n-1} \setminus E$, the map $\pi_a \circ f$ is an immersion on K.*

Proof. Since measure zero is preserved under countable unions and this is a local statement, we may assume that K is contained in a single coordinate chart and, furthermore, that M is an open subset of $\mathbb{R}^m$.

Let $E \subset \mathbb{R}^{n-1}$ be the set of all parameters a for which $\pi_a \circ f$ fails to be an immersion at some point of K. Concretely, $a \in E$ if and only if there exist $x \in K$ and $\lambda \in \mathbb{R}^m$ with $|\lambda| = 1$ such that

$$\sum_{k=1}^{m} \lambda_k \Big(\frac{\partial f_j}{\partial x_k}(x) - a_j \frac{\partial f_n}{\partial x_k}(x) \Big) = 0, \qquad j = 1, \dots, n-1.$$

These equations define a closed subset of

$$K \times S^{m-1} \times \mathbb{R}^{n-1},$$

and the projection to the a-factor is proper (since K is compact), hence E is closed.

Set

$$\mu = \sum_{k=1}^{m} \lambda_k \frac{\partial f_n}{\partial x_k}(x),$$

so that the above equations can be rewritten as

$$\sum_{k=1}^{m} \lambda_k \frac{\partial f_j}{\partial x_k}(x) = \mu\, a_j, \qquad j = 1, \dots, n-1,$$

and, with $a_n := 1$,

$$\sum_{k=1}^{m} \lambda_k \frac{\partial f_j}{\partial x_k}(x) = \mu\, a_j, \qquad j = 1, \dots, n.$$

This means that the vector $(a, 1) \in \mathbb{R}^n$ is tangent to the immersed submanifold $f(M)$ at the point $f(x)$.

Since f is an immersion, the vector

$$v = \sum_{k=1}^{m} \lambda_k \frac{\partial f}{\partial x_k}(x)$$

is nonzero, hence $\mu \neq 0$ and $(a, 1)$ lies in the range of the smooth map

$$F : \mathbb{R}^m \times K \to \mathbb{R}^n, \qquad F(\lambda, x) = \sum_{k=1}^{m} \lambda_k \frac{\partial f}{\partial x_k}(x).$$

The domain of F has dimension $2m < n$, so by a simple special case of the Morse-Sard theorem the image $F(\mathbb{R}^m \times K)$ has measure zero in $\mathbb{R}^n$. For each fixed $\mu \neq 0$, the intersection of this image with the affine hyperplane

$$H_\mu = \{(a, \mu) \in \mathbb{R}^n : a \in \mathbb{R}^{n-1}\}$$

also has measure zero (by Fubini's theorem and homogeneity). Projecting $H_\mu \cap \operatorname{im} F$ onto the first $n-1$ coordinates yields a measure-zero subset of $\mathbb{R}^{n-1}$.

Since, for $a \in E$, the direction $(a, 1)$ lies in $\operatorname{im} F$ with some $\mu \neq 0$, we conclude that E is a closed measure-zero subset of $\mathbb{R}^{n-1}$. □

We next analyze when injectivity is preserved under projection.

Lemma 4.49. *Let M be an m-dimensional manifold, $K \subset M$ a compact set, and $f \in C^\infty(M, \mathbb{R}^n)$ an injective immersion on a neighborhood of K. Assume $n > 2m+1$. Then there exists a closed measure-zero set $F \subset \mathbb{R}^{n-1}$ such that, for all $a \in \mathbb{R}^{n-1} \setminus F$, the map $\pi_a \circ f$ is an injective immersion on a neighborhood of K.*

Proof. By Lemma 4.48, there is a closed measure-zero set $E \subset \mathbb{R}^{n-1}$ such that $\pi_a \circ f$ is an immersion near K for all $a \notin E$. It remains to rule out failure of injectivity.

Let $E' \subset \mathbb{R}^{n-1}$ be the set of parameters a such that $\pi_a \circ f$ is not injective on K. First observe that $E \cup E'$ is closed. Indeed, suppose $a_j \in E'$ and $a_j \to a$. For each j there exist $x'_j, x''_j \in K$, $x'_j \neq x''_j$, with

$$\pi_{a_j} f(x'_j) = \pi_{a_j} f(x''_j).$$

Passing to a subsequence, we may assume $x'_j \to x'$ and $x''_j \to x''$ for some $x', x'' \in K$. If $a \notin E$, then for j sufficiently large, $\pi_{a_j} \circ f$ is an injective immersion on a fixed neighborhood of x', so in particular $x' \neq x''$ and

$$\pi_a f(x') = \pi_a f(x''),$$

showing that $a \in E'$. Thus $E \cup E'$ is closed.

Now describe E' more explicitly. The condition $a \in E'$ means that there exist $x', x'' \in K$, $x' \neq x''$, such that

$$f_j(x') - a_j f_n(x') = f_j(x'') - a_j f_n(x''), \qquad j = 1, \dots, n-1.$$

Setting $a_n = 1$ and

$$\mu = f_n(x') - f_n(x''),$$

we can rewrite this as

$$f(x') - f(x'') = \mu\,(a_1, \dots, a_{n-1}, 1).$$

Since f is injective on K, we have $f(x') \neq f(x'')$ and therefore $\mu \neq 0$. Thus the vector $(a, 1)$ lies in the range of the smooth map

$$G : \mathbb{R} \times K \times K \to \mathbb{R}^n, \qquad G(t, x', x'') = t\,\big(f(x') - f(x'')\big).$$

The domain of G has dimension $1 + 2m < n$ by assumption, so the image $G(\mathbb{R} \times K \times K)$ has measure zero in $\mathbb{R}^n$. As before, by homogeneity and Fubini's theorem, its intersection with each hyperplane $\{(a, \mu) : a \in \mathbb{R}^{n-1}\}$, $\mu \neq 0$, has measure zero, and hence the corresponding sets of a in $\mathbb{R}^{n-1}$ also have measure zero.

It follows that E' is a measure-zero subset of $\mathbb{R}^{n-1}$, and hence so is $F := E \cup E'$, which is closed. For $a \notin F$, the map $\pi_a \circ f$ is both an immersion and injective on a neighborhood of K, as claimed. □

In these two lemmas we have successively excluded two types of "bad" projection directions: in Lemma 4.48 we avoided projecting along directions tangent to $f(M)$, and in Lemma 4.49 we further avoided directions parallel to chords joining distinct points of $f(K)$. The sets of forbidden directions are controlled by $2m$ and $2m + 1$ parameters, respectively, which explains the dimension assumptions $n > 2m$ and $n > 2m + 1$.

Theorem 4.50 (Approximation by proper embeddings)**.** *Let M be an m-dimensional manifold and let*

$$f \in C^\infty(M, \mathbb{R}^n)$$

be a proper map, where $n \geq 2m+1$. Then for every positive continuous function $\varepsilon : M \to (0, \infty)$ there exists a proper embedding

$$g \in C^\infty(M, \mathbb{R}^n)$$

such that

$$|g(x) - f(x)| \leq \varepsilon(x), \qquad x \in M.$$

Proof. First replace ε by

$$\varepsilon'(x) := \min\{1, \varepsilon(x)\},$$

and relabel $\varepsilon' = \varepsilon$. Then the function

$$M \to \mathbb{R}, \qquad x \mapsto |f(x)| - \varepsilon(x)$$

is proper: indeed, for any $c \in \mathbb{R}$ we have

$$\{x : |f(x)| - \varepsilon(x) \leq c\} \subset \{x : |f(x)| \leq c + 1\},$$

and the latter set is compact because f is proper and $\varepsilon \leq 1$. Consequently, if g satisfies $|g - f| \leq \varepsilon$ on M, then

$$|g(x)| \geq |f(x)| - |g(x) - f(x)| \geq |f(x)| - \varepsilon(x),$$

so the properness of $x \mapsto |f(x)| - \varepsilon(x)$ implies that g is also proper. Thus it suffices to construct an embedding g with $|g - f| \leq \varepsilon$.

Let

$$K_1 \subset K_2 \subset \cdots \subset M$$

be a compact exhaustion of M. We construct inductively a sequence

$$g_j \in C^\infty(M, \mathbb{R}^n), \qquad j = 0, 1, 2, \ldots,$$

such that:

1. $g_0 = f$;
2. for all $j \geq 1$,
$$|g_j(x) - g_{j-1}(x)| \leq \frac{\varepsilon(x)}{2^j}, \qquad x \in M;$$
3. g_j is an injective immersion on a neighborhood of K_j;
4. $g_j = g_{j-1}$ on K_{j-1}.

Assume g_{j-1} has been constructed for some $j \geq 1$. By Lemma 4.47, there exist an integer ℓ and a map

$$h \in C_c^\infty(M, \mathbb{R}^\ell)$$

such that $g_{j-1} \oplus h$ is an injective immersion on a neighborhood of K_j. Moreover, we may choose h so that $h \equiv 0$ near K_{j-1}: for instance, apply Lemma 4.47 with

$$K = K_j \setminus V_{j-1}, \qquad U = M \setminus K_{j-1},$$

where V_{j-1} is a neighborhood of K_{j-1} on which g_{j-1} is already an injective immersion.

Now repeatedly apply Lemma 4.49 to the map

$$M \to \mathbb{R}^{n+\ell}, \qquad x \mapsto (g_{j-1}(x), h(x)),$$

to obtain a linear map $T : \mathbb{R}^\ell \to \mathbb{R}^n$ with arbitrarily small operator norm such that

$$g_j := g_{j-1} + T \circ h$$

is an injective immersion on a neighborhood of K_j. Since $h \equiv 0$ near K_{j-1}, we have $g_j = g_{j-1}$ on K_{j-1}. By choosing $\|T\|$ sufficiently small, we can also ensure that

$$|g_j(x) - g_{j-1}(x)| \leq \frac{\varepsilon(x)}{2^j}, \qquad x \in M.$$

This completes the inductive step.

By construction,

$$\sum_{j=1}^{\infty} |g_j(x) - g_{j-1}(x)| \leq \sum_{j=1}^{\infty} \frac{\varepsilon(x)}{2^j} = \varepsilon(x),$$

so the series converges pointwise, and in fact stabilizes on each compact set: for any fixed K_N, all maps g_j with $j \geq N$ coincide on K_N because of property (4). We can therefore define

$$g(x) := \lim_{j \to \infty} g_j(x),$$

and obtain a smooth map $g \in C^\infty(M, \mathbb{R}^n)$ with

$$|g(x) - f(x)| \leq \varepsilon(x), \qquad x \in M.$$

Moreover, on each K_j we have $g = g_j$, and g_j is an injective immersion near K_j, hence g is an injective immersion on all of M.

Finally, since g is a proper injective immersion, the following lemma shows that g is an embedding. □

Lemma 4.51. *Let X and Y be Hausdorff spaces, with Y locally compact, and let $f : X \to Y$ be proper (i.e. $f^{-1}(K)$ is compact for every compact $K \subset Y$). Then f is a closed map: the image of every closed subset of X is closed in Y.*

Proof. Let $C \subset X$ be closed, and let $y \in \overline{f(C)}$. We must show $y \in f(C)$.

Since Y is locally compact and Hausdorff, there exists a compact neighborhood $K \subset Y$ of y. Then

$$f(C) \cap K = f\big(C \cap f^{-1}(K)\big).$$

The set $C \cap f^{-1}(K)$ is closed in the compact set $f^{-1}(K)$, hence compact. Thus $f(C) \cap K$ is the continuous image of a compact set and therefore compact, in particular closed in K.

Because $y \in \overline{f(C)}$, we also have $y \in \overline{f(C) \cap K}$. But $f(C) \cap K$ is closed in K and $y \in K$, so $y \in f(C) \cap K$. Hence $y \in f(C)$.

We have shown that every point in $\overline{f(C)}$ lies in $f(C)$, so $f(C)$ is closed in Y. Thus f is a closed map. □

Finally, let us explain why proper maps are abundant, so the approximation theorem above always applies. When we constructed partitions of unity, we also proved the existence of an *exhaustion function* on any smooth manifold M. Recall that an exhaustion function is a continuous map

$$\varphi : M \to \mathbb{R}$$

such that every sublevel set

$$\varphi^{-1}((-\infty, c])$$

is compact for all $c \in \mathbb{R}$. In particular, every exhaustion function is a *proper map.*

Therefore, given any $n \geq 1$, we may consider the map

$$x \longmapsto (\varphi(x), 0, \ldots, 0) \in \mathbb{R}^n,$$

which is again proper. Thus proper maps $M \to \mathbb{R}^n$ always exist, and Theorem 4.50 shows that we can approximate any such proper map by a proper embedding arbitrarily well. This guarantees the existence of smooth embeddings of M into $\mathbb{R}^n$ for all sufficiently large n.

We now describe the normal bundle of an embedded submanifold and show that it is an embedded submanifold of the tangent bundle of $\mathbb{R}^n$.

Let $M \subset \mathbb{R}^n$ be an embedded m-dimensional submanifold. For each $x \in \mathbb{R}^n$, the tangent space $T_x\mathbb{R}^n$ is canonically identified with $\mathbb{R}^n$, and hence inherits the standard Euclidean inner product $\langle \cdot, \cdot \rangle$.

For $x \in M$, define the normal space

$$N_xM \;:=\; \{v \in T_x\mathbb{R}^n : \langle v, w \rangle = 0 \text{ for all } w \in T_xM\}.$$

The *normal bundle* of M is then defined as

$$NM \;:=\; \{(x, v) \in T\mathbb{R}^n \cong \mathbb{R}^n \times \mathbb{R}^n \mid x \in M,\ v \in N_xM\}.$$

There is a natural projection

$$\pi_{NM} : NM \longrightarrow M, \qquad \pi_{NM}(x, v) = x,$$

which is just the restriction of the standard bundle projection $\pi : T\mathbb{R}^n \to \mathbb{R}^n$.

Proposition 4.52. *The normal bundle NM is an embedded submanifold of $T\mathbb{R}^n$.*

Proof. Let $p_0 \in M$ and choose a coordinate chart

$$(U, \varphi = (x^1, \ldots, x^n))$$

on $\mathbb{R}^n$ such that $p_0 \in U$ and

$$M \cap U = \{p \in U : x^{m+1}(p) = \cdots = x^n(p) = 0\}.$$

On U the coordinate vector fields

$$\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n},$$

form a basis of $T_p\mathbb{R}^n$ for each $p \in U$.

Let $(u^1, \dots, u^n)$ be the standard coordinates on $\mathbb{R}^n$. Then on U we can write

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

Define a smooth map

$$\Phi : U \times \mathbb{R}^n \to \varphi(U) \times \mathbb{R}^n$$

by

$$\Phi(p, v) = \big(x^1(p), \dots, x^n(p), \langle v, \tfrac{\partial}{\partial x^1} \rangle, \dots, \langle v, \tfrac{\partial}{\partial x^n} \rangle\big).$$

In product coordinates (p, v) the differential has block form

$$D\Phi_{(p,v)} = \begin{pmatrix} A & 0 \\ * & A^{-1} \end{pmatrix},$$

where $A = (\partial x^i/\partial u^j)_{i,j}$ is the Jacobian of $u \mapsto x(u)$. Since φ is a local diffeomorphism, A is invertible, hence $D\Phi_{(p,v)}$ is invertible. So Φ is a local diffeomorphism.

We claim that Φ is injective. If $\Phi(p, v) = \Phi(p', v')$, then

$$(x^1(p), \dots, x^n(p)) = (x^1(p'), \dots, x^n(p')),$$

so $p = p'$ by injectivity of φ. For this p,

$$\langle v, \tfrac{\partial}{\partial x^i} \rangle = \langle v', \tfrac{\partial}{\partial x^i} \rangle, \quad i = 1, \dots, n,$$

so $v = v'$ because the $\partial/\partial x^i$ form a basis of $T_p\mathbb{R}^n$. Thus Φ is injective and defines a smooth coordinate chart on $U \times \mathbb{R}^n$.

Now $(p, v) \in NM$ iff:

$$x^{m+1}(p) = \cdots = x^n(p) = 0, \qquad \langle v, \tfrac{\partial}{\partial x^i} \rangle = 0 \ (i = 1, \dots, m).$$

Hence, writing

$$\Phi(p, v) = (x^1, \dots, x^n, w^1, \dots, w^n), \qquad w^i = \langle v, \tfrac{\partial}{\partial x^i} \rangle,$$

we have

$$(p, v) \in NM \iff x^{m+1} = \cdots = x^n = 0, \quad w^1 = \cdots = w^m = 0.$$

Thus

$$\Phi\big(NM \cap (U \times \mathbb{R}^n)\big) = \{(x, w) : x^{m+1} = \cdots = x^n = 0, \ w^1 = \dots w^m = 0\},$$

a linear subspace of $\varphi(U) \times \mathbb{R}^n$, hence an embedded submanifold.

Therefore NM is an embedded submanifold of $T\mathbb{R}^n$. □

4.12 Tubular neighborhoods

Let $M \subset \mathbb{R}^n$ be an embedded m–dimensional submanifold. Its normal bundle

$$NM = \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^n : x \in M,\ v \perp T_xM\}$$

is an embedded submanifold of $T\mathbb{R}^n \cong \mathbb{R}^n \times \mathbb{R}^n$. Define the (normal) exponential map

$$E : NM \to \mathbb{R}^n, \qquad E(x, v) = x + v.$$

This is smooth because it is the restriction of the addition map on $\mathbb{R}^n \times \mathbb{R}^n$.

Definition 4.53. A *tubular neighborhood* of M is an open set $U \subset \mathbb{R}^n$ for which there exists an open subset

$$V = \{(x, v) \in NM : |v| < \delta(x)\}$$

with $\delta : M \to (0, \infty)$ continuous, such that $E|_V : V \to U$ is a diffeomorphism.

Theorem 4.54 (Tubular Neighborhood Theorem)**.** *Every embedded submanifold $M \subset \mathbb{R}^n$ admits a tubular neighborhood.*

Proof. Let $M_0 = \{(x, 0) : x \in M\}$ be the zero section of NM. We first show that E is a local diffeomorphism near M_0.

Fix $x \in M$. On M_0, the restriction

$$E|_{M_0} : M_0 \to M, \qquad (x, 0) \mapsto x$$

is a diffeomorphism, so its differential

$$dE_{(x,0)} : T_{(x,0)}M_0 \to T_xM$$

is an isomorphism. On the other hand, the restriction of E to the fiber N_xM is the affine map

$$N_xM \to \mathbb{R}^n, \qquad w \mapsto x + w,$$

whose differential at $w = 0$ is the identity on N_xM. Thus

$$dE_{(x,0)} : T_{(x,0)}(N_xM) \to N_xM$$

is also an isomorphism.

Since E restricts to a diffeomorphism $M_0 \to M$ along the zero section $M_0 = \{(x, 0) : x \in M\}$, its differential

$$dE_{(x,0)} : T_{(x,0)}M_0 \longrightarrow T_xM$$

is an isomorphism. Similarly, the restriction of E to the fiber N_xM is the affine map $w \mapsto x+w$, whose differential at 0 is the identity on N_xM. Hence

$$dE_{(x,0)} : T_{(x,0)}(N_xM) \longrightarrow N_xM$$

is also an isomorphism.

Now set

$$V_1 := T_{(x,0)}M_0, \qquad V_2 := T_{(x,0)}(N_xM) \subset T_{(x,0)}NM.$$

Since M_0 and the fiber directions meet transversely in NM at $(x, 0)$, we have

$$V_1 \cap V_2 = \{0\}, \qquad \dim V_1 = \dim T_x M = m, \qquad \dim V_2 = \dim N_x M = n - m.$$

Because $\dim T_{(x,0)} NM = n = \dim V_1 + \dim V_2$, it follows that

$$T_{(x,0)} NM = V_1 \oplus V_2.$$

Choose a basis of $T_{(x,0)} NM$ consisting of a basis of V_1 followed by a basis of V_2, and choose a basis of $T_x \mathbb{R}^n$ consisting of a basis of $T_x M$ followed by a basis of $N_x M$. With respect to these adapted bases, the matrix of

$$dE_{(x,0)} : T_{(x,0)} NM \longrightarrow T_x \mathbb{R}^n$$

has block diagonal form

$$dE_{(x,0)} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix},$$

where

$$A = dE_{(x,0)}|_{V_1} : V_1 \to T_x M, \qquad B = dE_{(x,0)}|_{V_2} : V_2 \to N_x M.$$

Both A and B are linear isomorphisms by the discussion above, hence the entire block diagonal matrix is invertible. Thus $dE_{(x,0)}$ is a linear isomorphism.

By the inverse function theorem, E is a local diffeomorphism at $(x, 0)$.

Since NM is a vector bundle, we have a direct sum decomposition

$$T_{(x,0)} NM = T_{(x,0)} M_0 \oplus T_{(x,0)} (N_x M),$$

and similarly

$$T_x \mathbb{R}^n = T_x M \oplus N_x M.$$

By the preceding discussion, $dE_{(x,0)}$ maps each summand isomorphically onto the corresponding summand, hence is an isomorphism $T_{(x,0)} NM \to T_x \mathbb{R}^n$. By the inverse function theorem, E is a local diffeomorphism at $(x, 0)$.

Therefore, for each $x \in M$ there exists $\delta > 0$ such that

$$E : V_\delta(x) \to E\big(V_\delta(x)\big)$$

is a diffeomorphism, where we may take

$$V_\delta(x) := \{(x', v') \in NM : |x - x'| < \delta,\ |v'| < \delta\}.$$

For each $x \in M$, define

$$\rho(x) := \sup\big\{\delta \in (0, 1] : E|_{V_\delta(x)} \text{ is a diffeomorphism onto its image}\big\}.$$

The local argument above shows $\rho(x) > 0$ for all x. To see that ρ is 1–Lipschitz, fix $x, x' \in M$ and assume $|x - x'| < \rho(x)$. Set $\delta := \rho(x) - |x - x'| > 0$. Then $V_\delta(x') \subset V_{\rho(x)}(x)$, so E is a diffeomorphism on $V_\delta(x')$, hence $\rho(x') \geq \delta$ and

$$\rho(x) - \rho(x') \leq |x - x'|.$$

If $|x - x'| \geq \rho(x)$ this inequality is trivial. Exchanging x and x' gives $\rho(x') - \rho(x) \leq |x - x'|$, so

$$|\rho(x) - \rho(x')| \leq |x - x'|.$$

Now set

$$V := \{(x, v) \in NM : |v| < \tfrac{1}{2}\rho(x)\}.$$

We claim that E is injective on V. Suppose $(x, v), (x', v') \in V$ satisfy $E(x, v) = E(x', v')$, i.e. $x + v = x' + v'$. Then

$$|x - x'| = |v - v'| \leq |v| + |v'| < \tfrac{1}{2}\rho(x) + \tfrac{1}{2}\rho(x').$$

Without loss of generality assume $\rho(x') \leq \rho(x)$. Then

$$|x - x'| < \rho(x),$$

so both (x, v) and (x', v') lie in $V_{\rho(x)}(x)$. By definition of $\rho(x)$, E is injective on $V_{\rho(x)}(x)$, hence $(x, v) = (x', v')$. Thus E is injective on V.

Finally, let $U := E(V) \subset \mathbb{R}^n$. Since E is a local diffeomorphism and V is open, U is open and

$$E|_V : V \to U$$

is a bijective local diffeomorphism, hence a global diffeomorphism. By construction, V is of the form $\{(x, v) : |v| < \delta(x)\}$ with $\delta(x) = \frac{1}{2}\rho(x)$, so U is a tubular neighborhood of M. □

Proposition 4.55 (Existence of a smooth retraction)**.** *If U is any tubular neighborhood of M, then there exists a smooth map*

$$r : U \to M$$

which is a retraction ($r|_M = \mathrm{id}_M$) and a smooth submersion.

Proof. Write $U = E(V)$ with

$$V = \{(x, v) \in NM : |v| < \delta(x)\},$$

and $E|_V : V \to U$ a diffeomorphism. Let $\pi_{NM} : NM \to M$ be the bundle projection $(x, v) \mapsto x$. Define

$$r := \pi_{NM} \circ (E|_V)^{-1} : U \to M.$$

Then r is smooth and $r(x) = x$ for all $x \in M$, so r is a retraction. Moreover, π_{NM} is a submersion, and $(E|_V)^{-1}$ is a diffeomorphism, so r is a submersion as well. □

Proposition 4.56 (Existence of a smooth retraction)**.** *If U is any tubular neighborhood of M, then there exists a smooth map*

$$r : U \to M$$

which is a retraction ($r|_M = \mathrm{id}_M$) and a smooth submersion.

Proof. Write $U = E(V)$ with

$$V = \{(x, v) \in NM : |v| < \delta(x)\},$$

and $E|_V : V \to U$ a diffeomorphism. Let $\pi_{NM} : NM \to M$ be the bundle projection $(x, v) \mapsto x$. Define

$$r := \pi_{NM} \circ (E|_V)^{-1} : U \to M.$$

Then r is smooth and $r(x) = x$ for all $x \in M$, so r is a retraction. Moreover, π_{NM} is a submersion, and $(E|_V)^{-1}$ is a diffeomorphism, so r is a submersion as well. □

4.13 Whitney approximation theorem

Earlier we proved the following approximation result for real-valued functions: if M is a smooth manifold, $f : M \to \mathbb{R}$ is continuous, and $\varepsilon > 0$, then there exists a smooth function $g : M \to \mathbb{R}$ such that $|f(x) - g(x)| < \varepsilon$ for all $x \in M$. The same argument, using partitions of unity and convolution in coordinate charts, shows more generally:

- For a continuous map $f : M \to \mathbb{R}^n$ and a positive continuous function $\varepsilon : M \to (0, \infty)$, there exists a smooth $\tilde{f} : M \to \mathbb{R}^n$ with

$$\|\tilde{f}(x) - f(x)\| < \varepsilon(x) \quad \text{for all } x \in M.$$

- If $A \subset M$ is closed and f is already smooth on a neighborhood of A, then $\tilde{f}$ can be chosen to agree with f on a (possibly smaller) neighborhood of A.

The vector space structure of $\mathbb{R}^n$ is crucial in that proof, since we multiply by cut-off functions and sum.

Using the Whitney embedding theorem and the tubular neighborhood theorem, we can now generalize this approximation result to maps with values in an arbitrary manifold.

Theorem 4.57 (Whitney Approximation Theorem)**.** *Let M and N be smooth manifolds and let $f : M \to N$ be a continuous map. Then f is homotopic to a smooth map $g : M \to N$.*

Moreover, if $A \subset M$ is a closed subset and f is smooth on a neighborhood of A, then g can be chosen so that $g = f$ on A, and the homotopy from f to g is stationary on A (i.e. relative to A).

Proof. By the Whitney embedding theorem, we may assume that N is a properly embedded submanifold of some Euclidean space $\mathbb{R}^n$; that is, we view N as a closed embedded submanifold $N \subset \mathbb{R}^n$.

By the tubular neighborhood theorem, there exists an open neighborhood $U \subset \mathbb{R}^n$ of N and a smooth map

$$r : U \to N$$

which is a retraction, i.e. $r(x) = x$ for all $x \in N$.

For each $x \in N$, define

$$\delta(x) := \sup\{\varepsilon \in (0, 1] : B(x, \varepsilon) \subset U\},$$

where $B(x, \varepsilon)$ is the open Euclidean ball of radius ε centered at x. A triangle-inequality argument exactly like the one used in the proof of the tubular neighborhood theorem shows that

$$|\delta(x) - \delta(x')| \leq \|x - x'\| \quad \text{for all } x, x' \in N,$$

so $\delta : N \to (0, 1]$ is 1–Lipschitz in particular continuous.

Now consider the original continuous map $f : M \to N \subset \mathbb{R}^n$. Set

$$\tilde{\delta} := \delta \circ f : M \to (0, 1].$$

By the approximation result for $\mathbb{R}^n$–valued maps, there exists a smooth map $\tilde{f} : M \to \mathbb{R}^n$ such that

$$\|\tilde{f}(p) - f(p)\| < \tilde{\delta}(p) = \delta(f(p)) \quad \text{for all } p \in M,$$

and such that $\tilde{f} = f$ on a neighborhood of A (since f is smooth near A by assumption).

For each $p \in M$, the inequality $\|\tilde{f}(p) - f(p)\| < \delta(f(p))$ means precisely that $\tilde{f}(p)$ lies in the ball $B(f(p), \delta(f(p)))$, and by definition of δ this ball is contained in U. Since U is convex along the line segment between $f(p)$ and $\tilde{f}(p)$, we have

$$(1-t)f(p) + t\tilde{f}(p) \in U \quad \text{for all } t \in [0,1].$$

Define a homotopy $H : M \times [0,1] \to N$ by

$$H(p,t) := r\big((1-t)f(p) + t\tilde{f}(p)\big).$$

This is well-defined because the point inside $r(\,\cdot\,)$ lies in U for all t. The map H is continuous (and smooth in p for each fixed $t > 0$) as a composition of continuous maps. We have

$$H(p,0) = r(f(p)) = f(p), \qquad H(p,1) = r(\tilde{f}(p)) =: g(p).$$

Thus H is a homotopy from f to $g := r \circ \tilde{f}$.

The map g is smooth, being the composition of the smooth maps $r : U \to N$ and $\tilde{f} : M \to \mathbb{R}^n$, and it takes values in N. Moreover, on a neighborhood of A we have $\tilde{f} = f$ and $f(p) \in N$, so $r(f(p)) = f(p)$; hence

$$H(p,t) = r\big((1-t)f(p) + t\tilde{f}(p)\big) = r(f(p)) = f(p)$$

for all p near A and all $t \in [0,1]$. In particular, $g(p) = f(p)$ for all $p \in A$, and the homotopy is fixed on A.

This proves the theorem. □

4.14 Smooth homotopy and topological homotopy

Let M and N be smooth manifolds. A homotopy $H : M \times I \to N$ is called a *smooth homotopy* if H is a smooth map, in the sense that it extends to a smooth map on some open neighborhood of $M \times I$ in $M \times \mathbb{R}$. Two smooth maps $f, g : M \to N$ are said to be *smoothly homotopic* if there exists a smooth homotopy from f to g.

We now prove that smooth homotopy is equivalent to ordinary homotopy for smooth maps.

Theorem 4.58. *Let M and N be smooth manifolds, and let $f, g : M \to N$ be smooth maps. If f and g are homotopic as continuous maps, then they are smoothly homotopic.*

Proof. Let $H : M \times I \to N$ be a (continuous) homotopy from f to g. Define an extended map

$$\overline{H} : M \times \mathbb{R} \to N$$

by

$$\overline{H}(x,t) = \begin{cases} f(x), & t \le 0, \\ H(x,t), & 0 \le t \le 1, \\ g(x), & t \ge 1. \end{cases}$$

Then $\overline{H}$ is continuous, and for every $\delta > 0$ it is smooth on a neighborhood of

$$M \times (-\infty, -\delta] \;\cup\; M \times [1+\delta, \infty),$$

since there it is equal to the smooth maps f and g.

Apply the Whitney approximation theorem to the continuous map $\overline{H} : M \times \mathbb{R} \to N$. Fix $\delta > 0$ and let $A := M \times ((-\infty, -\delta] \cup [1+\delta, \infty))$. Then $\overline{H}$ is smooth on a neighborhood of A, so Whitney's theorem yields a smooth map

$$\widetilde{H} : M \times \mathbb{R} \to N$$

such that $\widetilde{H} = \overline{H}$ on A.

In particular,

$$\widetilde{H}(x, -\delta) = \overline{H}(x, -\delta) = f(x), \qquad \widetilde{H}(x, 1+\delta) = \overline{H}(x, 1+\delta) = g(x).$$

Restrict $\widetilde{H}$ to the compact interval $[-\delta, 1+\delta]$ and reparameterize it linearly to obtain a smooth homotopy

$$K : M \times I \to N$$

from f to g.

Thus f and g are smoothly homotopic. □

4.15 The transversality homotopy theorem

Theorem 4.59 (Parametric Transversality Theorem)**.** *Suppose N, M, X, and S are smooth manifolds. If the smooth map $f : N \times S \to M$ is transverse to the smooth map $g : X \to M$, then for almost every $s \in S$, the map*

$$f_s : N \to M, \qquad f_s(n) := f(n, s),$$

is transverse to g.

Proof. Since f is transverse to g, the fibre product

$$W := (N \times S) \times_M X = \{(n, s, x) \in N \times S \times X : f(n, s) = g(x)\}$$

is a smooth manifold. Let

$$\pi_1 : W \to N \times S, \qquad \pi_2 : W \to X$$

be the projection maps, so that $f \circ \pi_1 = g \circ \pi_2$.

Let

$$p_1 : N \times S \to N, \qquad p_2 : N \times S \to S$$

be the standard projections. Define

$$h := p_2 \circ \pi_1\big|_W : W \to S.$$

By Sard's theorem, for almost every $s \in S$, the point s is a regular value of h.

Fix such an $s \in S$. We claim that f_s is transverse to g.

Let $(n, x) \in N \times X$ satisfy $f_s(n) = g(x) =: m$, and write $w = (n, s, x) \in W$. Since s is a regular value of h, the map h is transverse to the submanifold $S \supset \{s\}$, which means that

$$d(\pi_1)_w(T_w W) + T_{(n,s)}(N \times \{s\}) = T_{(n,s)}(N \times S). \tag{4.2}$$

On the other hand, the assumption that f is transverse to g means that

$$df_{(n,s)}\big(T_{(n,s)}(N \times S)\big) \; + \; dg_x(T_x X) = T_m M.$$

Using $f \circ \pi_1 = g \circ \pi_2$ on W, we get

$$df_{(n,s)} \circ d(\pi_1)_w(T_w W) = dg_x \circ d(\pi_2)_w(T_w W).$$

Now apply $df_{(n,s)}$ to (4.2),

$$dg_x \circ d(\pi_2)_w(T_w W) \; + \; df_{(n,s)}(T_n N) \; + \; dg_x(T_x X) = T_m M.$$

Since $dg_x \circ d(\pi_2)_w(T_w W) \subset dg_x(T_x X)$, the previous equation simplifies to

$$df_{(n,s)}(T_n N) + dg_x(T_x X) = T_m M.$$

But $df_{(n,s)}|_{T_n N} = d(f_s)_n$, hence

$$d(f_s)_n(T_n N) + dg_x(T_x X) = T_m M,$$

which is exactly the transversality condition for f_s and g at (n, x).

Therefore f_s is transverse to g for almost every $s \in S$. □

Theorem 4.60 (Transversality Homotopy Theorem)**.** *Let N, M, and X be smooth manifolds, and let $g : X \to M$ be a smooth map. Then for any smooth map $f : N \to M$ there exists a smooth map $h : N \to M$ such that h is homotopic to f and h is transverse to g.*

Proof. By the Whitney embedding theorem, we may regard M as a properly embedded submanifold of some Euclidean space $\mathbb{R}^k$. Let $U \subset \mathbb{R}^k$ be a tubular neighborhood of M, and let

$$r : U \to M$$

be the corresponding smooth retraction, which is also a smooth submersion.

As in the proof of the Whitney approximation theorem, for each $x \in M$ define

$$\delta(x) := \sup\{\varepsilon \in (0, 1] : B(x, \varepsilon) \subset U\},$$

so that $\delta : M \to (0, 1]$ is continuous. Choose a smooth function

$$\varepsilon : N \to (0, \infty)$$

such that

$$0 < \varepsilon(p) < \delta(f(p)) \quad \text{for all } p \in N.$$

Let $S = B^k \subset \mathbb{R}^k$ be the open unit ball, and define

$$F : N \times S \to M, \qquad F(p, s) := r\big(f(p) + \varepsilon(p)s\big).$$

This is well-defined: for each $(p, s) \in N \times S$ we have

$$\| \, \varepsilon(p)s \, \| < \varepsilon(p) < \delta(f(p)),$$

so $f(p)+\varepsilon(p)s \in B(f(p),\delta(f(p))) \subset U$, and hence $r(f(p)+\varepsilon(p)s)$ is defined. Clearly F is smooth, and

$$F(p,0) = r(f(p)) = f(p),$$

so $F_0 = f$, where $F_s(p) := F(p,s)$.

We claim that F is transverse to $g : X \to M$. In fact, we show that F is a submersion. Fix $(p,s) \in N \times S$ and set $y := f(p)+\varepsilon(p)s \in U$, $m := F(p,s) = r(y) \in M$. The differential of F at (p,s) is

$$dF_{(p,s)} = dr_y \circ d\Phi_{(p,s)},$$

where

$$\Phi : N \times S \to U, \qquad \Phi(p,s) := f(p)+\varepsilon(p)s.$$

The partial derivative of Φ in the s–direction is the linear map

$$d\Phi_{(p,s)}\big|_{\{0\}\times T_sS} : T_sS \cong \mathbb{R}^k \longrightarrow T_y\mathbb{R}^k \cong \mathbb{R}^k, \quad v \mapsto \varepsilon(p)v,$$

which is an isomorphism. Since $dr_y : T_yU \to T_mM$ is surjective (because r is a submersion), it follows that the composition

$$T_sS \xrightarrow{d\Phi_{(p,s)}} T_yU \xrightarrow{dr_y} T_mM$$

is also surjective. Thus $dF_{(p,s)}$ is surjective, so F is a smooth submersion and therefore transverse to g:

$$dF_{(p,s)}\big(T_{(p,s)}(N \times S)\big) = T_mM \supset dg_x(T_xX)$$

whenever $F(p,s) = g(x) = m$.

Now apply the Parametric Transversality Theorem to the maps

$$F : N \times S \to M, \qquad g : X \to M.$$

Since F is transverse to g, the theorem implies that for almost every $s \in S$, the map

$$F_s : N \to M, \qquad F_s(p) := F(p,s),$$

is transverse to g.

Choose such an $s_0 \in S$. Then $h := F_{s_0}$ is smooth and transverse to g. Moreover, f and h are smoothly homotopic via the homotopy

$$H : N \times I \to M, \qquad H(p,t) := F(p,ts_0).$$

Indeed, $H(p,0) = F(p,0) = f(p)$ and $H(p,1) = F(p,s_0) = h(p)$.

Thus we have found a smooth map h homotopic to f and transverse to g, as required. $\square$

Chapter 5

Vector Fields, Flows, and Lie Brackets

We now develop the basic calculus of vector fields on a smooth manifold. We begin with vector fields as sections of the tangent bundle and as derivations of the algebra $C^\infty(M)$. We then study vector fields related by smooth maps, the Lie bracket, integral curves and flows, and finally the rectification of vector fields and the local structure of several commuting flows.

5.1 Vector fields and derivations

Let M be a smooth manifold. Formally, a (smooth) vector field on M is a section of the tangent bundle $\pi : TM \to M$, that is, a smooth map

$$X : M \to TM$$

such that $\pi \circ X = \mathrm{id}_M$. Following Lee's notation, we usually write $X_p \in T_pM$ for the tangent vector at p given by $X(p)$, so a vector field assigns to each point $p \in M$ a tangent vector $X_p \in T_pM$.

As usual, we restrict attention to C^∞ objects: we say that a vector field X is smooth if the section $X : M \to TM$ is a smooth map of manifolds.

Just as we have local immersions and local submersions, we can also define *local vector fields*: If $U \subset M$ is an open set, a vector field on U is a smooth section of the restricted bundle

$$\pi|_{TU} : TU \to U,$$

that is, a smooth map $X : U \to TU$ with $\pi \circ X = \mathrm{id}_U$. In other words, a local vector field is simply a vector field defined on an open subset of M.

The *support* of a vector field X is defined to be the closure of the set of points where X does not vanish:

$$\operatorname{supp} X := \overline{\{p \in M : X_p \neq 0\}}.$$

We say that X is *compactly supported* if its support is a compact subset of M.

At first sight these definitions may look somewhat formal. To better understand them, let us look at the local coordinate description. Recall that for a chart

$$(U, \varphi) = (U, (x^1, \dots, x^m))$$

on M, we can define an induced chart on TM by

$$(\pi^{-1}(U), \tilde{\varphi}),$$

where

$$\tilde{\varphi}(p, v) = \big(x^1(p), \dots, x^m(p), v^1, \dots, v^m\big) \in \mathbb{R}^{2m},$$

and we write the tangent vector $v \in T_pM$ in the form

$$v = \sum_{i=1}^{m} v^i \frac{\partial}{\partial x^i}\Big|_p.$$

In this local picture, a section $X : U \to TU$ can be written in coordinates as

$$\tilde{\varphi} \circ X \circ \varphi^{-1} : \varphi(U) \to \varphi(U) \times \mathbb{R}^m, \quad (x^1, \dots, x^m) \mapsto \big(x^1, \dots, x^m, X^1(x), \dots, X^m(x)\big),$$

where $X^1, \dots, X^m$ are smooth functions on $\varphi(U)$ (equivalently, on U) giving the components of X in this chart.

Using our previous coordinate vector convention, we write

$$X_p = X^i(p) \frac{\partial}{\partial x^i}\Big|_p, \quad X^i(p) := X^i(\varphi(p)).$$

We call the functions $X^1, \dots, X^m$ the *component functions* of X in the chart (U, φ).

Thus, once we equip TM with the atlas induced from that of M, a smooth vector field is completely described by its component functions in each chart, subject to the usual compatibility conditions on overlaps. More concretely, suppose

$$(U, \varphi = (x^1, \dots, x^m)), \qquad (V, \psi = (y^1, \dots, y^m))$$

are two overlapping coordinate charts, and the corresponding component functions of a vector field X are $\{X^i\}$ in the x-coordinates and $\{Y^j\}$ in the y-coordinates. Then on $U \cap V$ we have

$$X = X^i \frac{\partial}{\partial x^i} = Y^j \frac{\partial}{\partial y^j},$$

so the components are related by the usual change-of-coordinates formula

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

Thus, a smooth vector field on M may equivalently be described by a collection of smooth component functions in each chart whose domains form a cover of M, subject to the familiar coordinate transformation law above. This gives a complete coordinate description of smooth vector fields on a manifold.

Let us look at two simple examples. Given a coordinate chart $(U, \varphi = (x^1, \dots, x^m))$ on M, we obtain the *coordinate vector fields*

$$\frac{\partial}{\partial x^i} : U \longrightarrow TU,$$

which are local vector fields on U.

As a second example, recall from the coordinate description of vector fields that on Euclidean space $\mathbb{R}^m$ our definition reduces to the familiar notion from multivariable calculus. Thus all the vector fields one encounters in basic analysis are special cases of the present definition. In particular, $\mathbb{R}^m$ carries the *position vector field* (called the Euler vector field in Lee), given in coordinates by

$$V_x = x^1 \frac{\partial}{\partial x^1} + \cdots + x^m \frac{\partial}{\partial x^m}.$$

Before continuing, we record a convention that we will use frequently. Throughout these notes we adopt the *Einstein summation convention*: whenever an index appears once upstairs and once downstairs, we implicitly sum over that index. For example,

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

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

Next we introduce a slight generalization of the notion of a vector field. Let $f : M \to N$ be a smooth map between smooth manifolds. A smooth map

$$X : M \longrightarrow TN \quad \text{satisfying} \quad \pi \circ X = f$$

is called a *vector field along* f. In the special case where M is a submanifold of N and f is the inclusion $\iota : M \hookrightarrow N$, we call such an X simply a vector field along M. As an extreme example, if M consists of a single point $q \in N$, then a vector field along the inclusion $M \hookrightarrow N$ is just a tangent vector in T_qN.

We now prove a basic fact.

> *Given a smooth manifold M, a point $p \in M$, and a tangent vector $v \in T_pM$, there exists a smooth vector field X on M such that $X_p = v$.*

Proof. Choose a coordinate chart $(U, \varphi = (x^1, \dots, x^m))$ around p. In these coordinates, we may write

$$v = v^i \frac{\partial}{\partial x^i}\Big|_p$$

for suitable real numbers v^i. Since the coordinate vector fields $\frac{\partial}{\partial x^i}$ are defined on all of U, choose a bump function $\eta : M \to \mathbb{R}$ supported in U such that $\eta(p) = 1$. Define a vector field

$$X = \eta\, v^i \frac{\partial}{\partial x^i} \quad \text{on } U, \qquad X = 0 \quad \text{on } M \setminus U.$$

This yields a smooth vector field on M satisfying $X_p = v$. □

Let $\mathfrak{X}(M)$ denote the set of all smooth vector fields on M. This is a vector space over $\mathbb{R}$ with pointwise addition and scalar multiplication:

$$(aX + bY)_p = a\, X_p + b\, Y_p.$$

More importantly, $\mathfrak{X}(M)$ is a module over the ring $C^\infty(M)$ of smooth real-valued functions on M. For $f \in C^\infty(M)$ and $X \in \mathfrak{X}(M)$, define

$$(fX)_p := f(p)\, X_p.$$

Using the local coordinate description of vector fields, it follows immediately that fX is again a smooth vector field on M, so $fX \in \mathfrak{X}(M)$. Thus $\mathfrak{X}(M)$ naturally carries the structure of a $C^\infty(M)$-module.

We now adopt another point of view on vector fields. Recall that a tangent vector $v \in T_pM$ can be defined as an $\mathbb{R}$-linear functional

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

satisfying the Leibniz rule

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

If $X \in \mathfrak{X}(M)$ is a smooth vector field, then for each $p \in M$ the value X_p is a tangent vector at p, hence a derivation at p in the above sense. We may therefore define, for every $f \in C^\infty(M)$, a new function $Xf : M \to \mathbb{R}$ by

$$(Xf)(p) := X_p(f).$$

In local coordinates we have

$$X = X^i \frac{\partial}{\partial x^i} \qquad \Rightarrow \qquad Xf = X^i \frac{\partial f}{\partial x^i},$$

so Xf is again a smooth function. Thus every smooth vector field $X \in \mathfrak{X}(M)$ gives rise to a map

$$X : C^\infty(M) \longrightarrow C^\infty(M), \qquad f \longmapsto Xf.$$

Proposition 5.1. *For any smooth vector field $X \in \mathfrak{X}(M)$, the associated map $X : C^\infty(M) \to C^\infty(M)$ satisfies:*

1. *X is $\mathbb{R}$-linear:*

$$X(af + bg) = a\,Xf + b\,Xg \qquad \textit{for all } a, b \in \mathbb{R},\ f, g \in C^\infty(M).$$

2. *X satisfies the Leibniz rule:*

$$X(fg) = f\,Xg + g\,Xf \qquad \textit{for all } f, g \in C^\infty(M).$$

Proof. Fix $p \in M$. The map $X_p : C^\infty(M) \to \mathbb{R}$ is a tangent vector at p, hence by definition an $\mathbb{R}$–linear map satisfying the Leibniz rule at p. Thus

$$(X(af + bg))(p) = X_p(af + bg) = a\,X_p(f) + b\,X_p(g) = (a\,Xf + b\,Xg)(p),$$

which proves the linearity of X pointwise, and hence globally.

Similarly,

$$(X(fg))(p) = X_p(fg) = f(p)\,X_p(g) + g(p)\,X_p(f) = (f\,Xg + g\,Xf)(p)$$

for all $p \in M$, which shows the Leibniz rule holds as an identity of functions. □

Remark 5.2. In algebra, a map $D : A \to A$ on an algebra A is called a *derivation* if it is linear and satisfies the Leibniz rule

$$D(ab) = a\,D(b) + b\,D(a).$$

In our setting, if the target is $\mathbb{R}$ and the Leibniz rule involves evaluation at a fixed point p (as above), then derivations $C^\infty(M) \to \mathbb{R}$ are precisely tangent vectors in T_pM. If the target is $C^\infty(M)$ and we have the global Leibniz rule

$$D(fg) = f\,Dg + g\,Df,$$

then such derivations $C^\infty(M) \to C^\infty(M)$ will turn out to be exactly smooth vector fields on M.

We now show that this derivation property can in fact be taken as the *definition* of a vector field.

Proposition 5.3. *A map $D : C^\infty(M) \to C^\infty(M)$ is a derivation (i.e. $\mathbb{R}$-linear and satisfying $D(fg) = f\,Dg + g\,Df$ for all f, g) if and only if there exists a smooth vector field $X \in \mathfrak{X}(M)$ such that*

$$Df = Xf \quad \textit{for all } f \in C^\infty(M).$$

In other words, derivations of $C^\infty(M)$ are in one-to-one correspondence with smooth vector fields on M.

Proof. The "if" direction has already been proved: given a smooth vector field X, the map $f \mapsto Xf$ is a derivation by the previous proposition.

For the converse, suppose $D : C^\infty(M) \to C^\infty(M)$ is a derivation. For each point $p \in M$, define

$$D_p : C^\infty(M) \to \mathbb{R}, \qquad D_p(f) := (Df)(p).$$

Each D_p is $\mathbb{R}$-linear and satisfies the Leibniz rule

$$D_p(fg) = f(p)\,D_p(g) + g(p)\,D_p(f),$$

so D_p is a derivation at p, hence a tangent vector in T_pM. Thus we obtain a map

$$X : M \longrightarrow TM, \qquad X_p := D_p.$$

Fix a coordinate chart $(U, \varphi = (x^1, \ldots, x^m))$ on M. By the coordinate description of tangent vectors from the coordinate description of tangent vectors proved in the previous chapter, for each $p \in U$ there exist unique real numbers $X^1(p), \ldots, X^m(p)$ such that

$$D_p(f) = X^i(p)\,\frac{\partial f}{\partial x^i}(p) \qquad \text{for all } f \in C^\infty(M).$$

In other words,

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

We claim that each coefficient function $X^i : U \to \mathbb{R}$ is smooth. Let $K \Subset U$ be a compact subset of U. Choose a bump function $\eta \in C^\infty(M)$ such that $\eta \equiv 1$ on K and $\operatorname{supp} \eta \subset U$. For each i define

$$f_i := \eta\, x^i.$$

Since f_i is smooth and has support in U, the function Df_i is smooth on M. For $p \in K$ we have $\eta(p) = 1$ and $d\eta_p = 0$, so

$$\frac{\partial f_i}{\partial x^j}(p) = \frac{\partial}{\partial x^j}(\eta x^i)(p) = \delta^i_j.$$

Hence, using the formula for D_p in coordinates,

$$(Df_i)(p) = D_p(f_i) = X^j(p)\, \frac{\partial f_i}{\partial x^j}(p) = X^j(p)\, \delta^i_j = X^i(p).$$

Thus on K we have

$$X^i = Df_i\big|_K,$$

and since Df_i is smooth, the restriction $X^i|_K$ is smooth. Because $K \Subset U$ was arbitrary, we conclude that each X^i is smooth on U.

Therefore, in every coordinate chart (U, x^i) we can write

$$X\big|_U = X^i\, \frac{\partial}{\partial x^i}$$

with smooth coefficient functions X^i, so X is a smooth vector field on M. Finally, by construction

$$(Df)(p) = D_p(f) = X_p(f) = (Xf)(p) \quad \text{for all } p \in M,\ f \in C^\infty(M),$$

so $Df = Xf$ as functions. This completes the proof. □

Proposition 5.4 (Locality of derivations). *Let $D : C^\infty(M) \to C^\infty(M)$ be a derivation. If $f, g \in C^\infty(M)$ agree on an open set $U \subset M$, then*

$$Df = Dg \quad \text{on } U.$$

Equivalently, $D(f - g)$ vanishes on U.

Proof. Let $h := f - g$. Then $h = 0$ on U. Choose a bump function $\eta \in C^\infty(M)$ such that $\eta > 0$ on U and $\eta = 0$ outside a slightly larger open set containing U. Then $\eta h = 0$, so by the Leibniz rule,

$$0 = D(\eta h) = \eta\, Dh + D(\eta)\, h.$$

On U we have $h = 0$, so the second term vanishes there, and hence $\eta\, Dh = 0$ on U. Since $\eta > 0$ on U, it follows that $Dh = 0$ on U, i.e. $Df = Dg$ on U. □

We record a closely related locality statement for derivations.

Proposition 5.5 (Restriction of derivations to open subsets)**.** *Let M be a smooth manifold and let $X : C^\infty(M) \to C^\infty(M)$ be a derivation. For every open set $U \subset M$ there exists a unique derivation*

$$X\big|_U : C^\infty(U) \to C^\infty(U)$$

such that for every $f \in C_c^\infty(U)$, if we identify f with its extension by 0 to a compactly supported function on M, then

$$X(f)\big|_U \;=\; X\big|_U(f).$$

In other words, $X\big|_U$ is characterized by the property that its action on compactly supported functions in U agrees with the action of X on their zero extensions to M. By abuse of notation, we will often simply write X instead of $X\big|_U$ when the open set is clear from context.

Proof. Existence: Fix an open set $U \subset M$. For $f \in C^\infty(U)$ and $p \in U$, choose a smooth function $\tilde{f} \in C^\infty(M)$ such that $\tilde{f} = f$ on some neighbourhood of p. Using the locality of X (proved earlier: if $h_1 = h_2$ near p, then $Xh_1(p) = Xh_2(p)$), we may define

$$(X\big|_U f)(p) := X(\tilde{f})(p),$$

and this is independent of the choice of $\tilde{f}$. Thus we obtain a well-defined map $X\big|_U : C^\infty(U) \to C^\infty(U)$.

Linearity and the Leibniz rule for $X\big|_U$ follow immediately from those of X, since near each point p the value $(X\big|_U f)(p)$ is computed using an extension $\tilde{f}$ and the derivation property of X. By construction, if $f \in C_c^\infty(U)$ and we identify f with its zero extension to M, then we may take $\tilde{f}$ to be exactly this extension, and hence $X\big|_U(f) = X(f)\big|_U$.

Uniqueness: Suppose $Y : C^\infty(U) \to C^\infty(U)$ is another derivation with the stated property, i.e. $Y(f) = X(f)\big|_U$ for all $f \in C_c^\infty(U)$ (via zero extension). Let $f \in C^\infty(U)$ and $p \in U$. Choose a bump function $\eta \in C_c^\infty(U)$ with $\eta(p) = 1$, and consider $\eta f \in C_c^\infty(U)$. Then

$$(X\big|_U f)(p) = (X\big|_U(\eta f))(p) = X(\eta f)(p),$$

while similarly

$$Y(f)(p) = Y(\eta f)(p) = X(\eta f)(p),$$

since ηf has compact support in U and both $X\big|_U$ and Y agree with X on such functions. Thus $(X\big|_U f)(p) = Y(f)(p)$ for all p, so $X\big|_U = Y$.

This proves existence and uniqueness of the restriction $X\big|_U$. □

Remark 5.6. If one is willing to work directly in local coordinates, the existence of the restricted derivation $X|_U$ is essentially immediate. Indeed, in a coordinate chart (U, x^i) we may write

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

so for any $f \in C^\infty(U)$ the expression

$$(X|_U)f := X^i \frac{\partial f}{\partial x^i}$$

is automatically a derivation on U, and clearly agrees with X on compactly supported functions once we identify a function in $C_c^\infty(U)$ with its extension by 0 to M. Thus the restriction property is completely transparent in coordinates.

The coordinate-free proof above is more formal but makes the locality of derivations conceptually clear: the construction of $X|_U$ depends only on the algebraic properties of D and not on the choice of coordinates. Moreover, in proving the coordinate formula for tangent vectors earlier, we have already used arguments of exactly this local algebraic nature.

5.2 Related vector fields and the Lie bracket

We now discuss the relationship between vector fields and smooth maps between manifolds. In our earlier discussion of tangent spaces, we defined the *differential* (or derivative) of a smooth map $f : M \to N$: for each $p \in M$ there is a linear map

$$d_p f : T_p M \longrightarrow T_{f(p)} N,$$

which pushes tangent vectors at p forward to tangent vectors at $f(p)$.

It is natural to ask whether f can push forward a vector field X on M to a vector field on N. In general, the answer is *no*. First, f need not be surjective, so even if we define $Y_{f(p)} := d_p f(X_p)$, this only gives us a vector field on $f(M) \subset N$, not on all of N. Even if f is surjective but not injective, we may have $f(p_1) = f(p_2) = q$ with $d_{p_1} f(X_{p_1}) \neq d_{p_2} f(X_{p_2})$, so there is no well-defined vector Y_q.

Even worse, the pushforward need not be smooth even when f is a smooth bijection. For example, consider $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$. Then f is a smooth bijection, but not a diffeomorphism (its inverse is not smooth at 0). If we try to push forward the constant vector field $\frac{\partial}{\partial x}$, we obtain

$$d_x f\Big(\frac{\partial}{\partial x}\Big) = 3x^2 \frac{\partial}{\partial y} = 3\, y^{2/3} \frac{\partial}{\partial y},$$

where $y = f(x) = x^3$. The coefficient $3y^{2/3}$ is not smooth at $y = 0$, so this is not a smooth vector field on $\mathbb{R}$.

Instead of trying to push forward arbitrary vector fields, we introduce a weaker notion.

Definition 5.7. Let $f : M \to N$ be a smooth map, and let $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$ be vector fields. We say that *X and Y are f-related* if

$$d_p f(X_p) = Y_{f(p)} \qquad \text{for all } p \in M.$$

Equivalently, X and Y are f-related if and only if for every $\varphi \in C^\infty(N)$ we have

$$(Y\varphi) \circ f \;=\; X(\varphi \circ f),$$

that is,

$$Y_{f(p)}(\varphi) \;=\; X_p(\varphi \circ f) \qquad \text{for all } p \in M \text{ and all } \varphi \in C^\infty(N).$$

In general there is no simple criterion to decide when two given vector fields are f-related. However, diffeomorphisms occupy a special place in the category of smooth manifolds: they are precisely the isomorphisms in this category, and hence should preserve all of the geometric structures we consider. In particular, a diffeomorphism always allows us to push forward vector fields.

Proposition 5.8. *Let $f : M \to N$ be a diffeomorphism. Then for every $X \in \mathfrak{X}(M)$ there exists a unique smooth vector field $Y \in \mathfrak{X}(N)$ such that X and Y are f-related.*

Proof. If Y is to be f-related to X, the defining condition forces

$$Y_q = d_{f^{-1}(q)}f\big(X_{f^{-1}(q)}\big) \qquad \text{for each } q \in N.$$

Thus there is at most one such Y, and the formula above shows how to define it:

$$Y : N \xrightarrow{f^{-1}} M \xrightarrow{X} TM \xrightarrow{df} TN.$$

Since f^{-1}, X, and df are smooth maps and composition of smooth maps is smooth, it follows that $Y : N \to TN$ is smooth. Thus Y is a smooth vector field on N and is f–related to X by construction. □

In this situation we write $Y = f_*X$ and call Y the *pushforward* of X by f.

Another useful setting where we can understand the relation between vector fields is the case of submanifolds. Let $S \subset M$ be an embedded submanifold. For each $p \in S$ we have a natural inclusion

$$T_pS \subset T_pM,$$

coming from the differential of the inclusion map $\iota : S \hookrightarrow M$. (We identify T_pS with its image $d_p\iota(T_pS) \subset T_pM$ and simply write $T_pS \subset T_pM$.)

Definition 5.9. Let $X \in \mathfrak{X}(M)$ be a smooth vector field on M. We say that X is *tangent to S at $p \in S$* if $X_p \in T_pS$. We say that X is *tangent to S* if $X_p \in T_pS$ for every $p \in S$.

The following characterization of tangency in terms of functions vanishing on S is very useful.

Proposition 5.10. *Let $S \subset M$ be an embedded submanifold and $X \in \mathfrak{X}(M)$. Then X is tangent to S if and only if*

$$(Xf)\big|_S = 0 \quad \textit{for every } f \in C^\infty(M) \textit{ with } f|_S \equiv 0.$$

Proof. First suppose X is tangent to S. Let $f \in C^\infty(M)$ vanish on S. Then $f \circ \iota \equiv 0$ on S, so for any $p \in S$ and $v \in T_pS$ we have

$$v(f \circ \iota) = 0.$$

In particular, taking $v = X_p \in T_pS$,

$$(Xf)(p) = X_pf = X_p(f \circ \iota) = 0.$$

Thus $(Xf)|_S \equiv 0$.

Conversely, suppose $X \in \mathfrak{X}(M)$ satisfies $(Xf)|_S \equiv 0$ whenever f vanishes on S. Fix $p \in S$ and consider the tangent vector $X_p \in T_pM$. We claim $X_p \in T_pS$. This is a local question, so we may choose a coordinate chart $(U, (x^1, \ldots, x^m))$ on M with $p \in U$ such that in these coordinates

$$S \cap U = \{x^1 = \cdots = x^k = 0\}$$

for some k. Write

$$X_p = a^i \frac{\partial}{\partial x^i}\Big|_p.$$

For each $1 \le j \le k$, the coordinate function x^j vanishes on $S \cap U$, so by assumption $(Xx^j)(p) = 0$. But

$$(Xx^j)(p) = X_p x^j = a^i \frac{\partial x^j}{\partial x^i}(p) = a^j.$$

Hence $a^j = 0$ for $1 \le j \le k$, so X_p has no components in the normal directions and therefore lies in T_pS. Since $p \in S$ was arbitrary, X is tangent to S. □

If $S \subset M$ is a submanifold, $Y \in \mathfrak{X}(M)$ is a vector field on M tangent to S, and $X \in \mathfrak{X}(S)$ is a vector field on S, we say that X and Y are ι-related, where $\iota : S \hookrightarrow M$ is the inclusion, if

$$d_p\iota(X_p) = Y_p \qquad \text{for all } p \in S.$$

In this case we will usually denote X by $Y|_S$.

Proposition 5.11. *Let $S \subset M$ be an embedded submanifold, and let $\iota : S \hookrightarrow M$ be the inclusion map. If $Y \in \mathfrak{X}(M)$ is tangent to S, then there exists a unique smooth vector field $X \in \mathfrak{X}(S)$, denoted $Y|_S$, which is ι-related to Y.*

Proof. Define $X_p \in T_pS$ for $p \in S$ by the requirement

$$d_p\iota(X_p) = Y_p.$$

This is possible and unique since $d_p\iota : T_pS \to T_pM$ is an injective linear map with image $T_pS \subset T_pM$ and $Y_p \in T_pS$ because Y is tangent to S. Thus we obtain a map

$$X : S \to TS, \qquad p \mapsto X_p.$$

We already know that the composition

$$S \xrightarrow{X} TS \hookrightarrow TM$$

agrees with the restriction of $Y : M \to TM$ to S, hence is smooth. Since TS is an embedded submanifold of TM and X takes values in TS, it follows by the corresponding result proved in the previous chapter that $X : S \to TS$ is smooth. Therefore X is a smooth vector field on S, and by construction it is ι-related to Y. Uniqueness is clear from the defining relation $d_p\iota(X_p) = Y_p$. □

We now introduce the Lie bracket of vector fields. Let X and Y be two smooth vector fields on a manifold M. Recall that we can equivalently view X and Y as derivations

$$X, Y : C^\infty(M) \longrightarrow C^\infty(M),$$

that is, as $\mathbb{R}$-linear maps satisfying the Leibniz rule $X(fg) = f\,Xg + g\,Xf$ and similarly for Y.

Definition (Lie bracket). The *Lie bracket* of X and Y is the map

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

defined by

$$[X, Y]f := X(Yf) - Y(Xf), \qquad f \in C^\infty(M).$$

We will see that $[X, Y]$ is again a derivation, hence comes from a unique smooth vector field on M, which we also denote by $[X, Y]$.

Lemma 5.12. *For smooth vector fields $X, Y \in \mathfrak{X}(M)$, the Lie bracket $[X, Y]$ defined above is a derivation of $C^\infty(M)$, i.e. it is $\mathbb{R}$-linear and satisfies the Leibniz rule*

$$[X, Y](fg) = f\,[X, Y]g + g\,[X, Y]f \quad \text{for all } f, g \in C^\infty(M).$$

Proof. Linearity in f is immediate from the definition and the linearity of X and Y. It remains to check the Leibniz rule. For $f, g \in C^\infty(M)$ we compute:

$$\begin{aligned}
[X, Y](fg) &= X\big(Y(fg)\big) - Y\big(X(fg)\big) \\
&= X\big(f\,Yg + g\,Yf\big) - Y\big(f\,Xg + g\,Xf\big) \\
&= X(f)\,Yg + f\,X(Yg) + X(g)\,Yf + g\,X(Yf) \\
&\quad - \big(Y(f)\,Xg + f\,Y(Xg) + Y(g)\,Xf + g\,Y(Xf)\big).
\end{aligned}$$

Rearranging terms, we group those multiplied by f and those multiplied by g:

$$\begin{aligned}
[X, Y](fg) &= f\big(X(Yg) - Y(Xg)\big) + g\big(X(Yf) - Y(Xf)\big) \\
&\quad + \big(X(f)\,Yg - Y(f)\,Xg\big) + \big(X(g)\,Yf - Y(g)\,Xf\big).
\end{aligned}$$

But the last two brackets cancel pairwise, since

$$X(f)\,Yg - Y(f)\,Xg = g\big(X(f)\,Y - Y(f)\,X\big)g = 0,$$

and similarly for the other term (more concretely, they recombine to 0 by symmetry). What matters is that the mixed terms arise in a symmetric way and cancel, leaving

$$[X, Y](fg) = f\,[X, Y]g + g\,[X, Y]f,$$

which is exactly the Leibniz rule for $[X, Y]$. Thus $[X, Y]$ is a derivation. □

Since derivations of $C^\infty(M)$ are in one-to-one correspondence with smooth vector fields on M, we may regard $[X, Y]$ as a smooth vector field, called the Lie bracket of X and Y.

The Lie bracket has a simple coordinate expression which is very useful for computations. Let $(U, (x^1, \ldots, x^m))$ be a coordinate chart on M, and write

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

on U (Einstein summation convention). Then we have the following formula.

Proposition 5.13. *In local coordinates as above,*

$$[X, Y] = \Big(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i}\Big)\,\frac{\partial}{\partial x^j}.$$

Proof. Let $f \in C^\infty(M)$, and compute on U:

$$Xf = X^i \frac{\partial f}{\partial x^i}, \qquad Yf = Y^i \frac{\partial f}{\partial x^i}.$$

Then

$$XYf = X\big(Y^i \partial_i f\big) = X^j \partial_j\big(Y^i\,\partial_i f\big)$$

$$
\begin{aligned}
&= X^j\big((\partial_j Y^i)\,\partial_i f + Y^i\,\partial_j\partial_i f\big),\\
YXf &= Y\big(X^i\partial_i f\big) = Y^j\partial_j\big(X^i\,\partial_i f\big)\\
&= Y^j\big((\partial_j X^i)\,\partial_i f + X^i\,\partial_j\partial_i f\big),
\end{aligned}
$$

where we write ∂_i for $\frac{\partial}{\partial x^i}$ and similarly for ∂_j. Subtracting,

$$
\begin{aligned}
[X,Y]f &= XYf - YXf\\
&= \big(X^j\,\partial_j Y^i - Y^j\,\partial_j X^i\big)\,\partial_i f + \big(X^jY^i - Y^jX^i\big)\,\partial_j\partial_i f.
\end{aligned}
$$

The second term vanishes because the second derivatives $\partial_j\partial_i f$ are symmetric in (i,j) while $X^jY^i - Y^jX^i$ is antisymmetric. Thus

$$[X,Y]f = \big(X^j\,\partial_j Y^i - Y^j\,\partial_j X^i\big)\,\partial_i f.$$

Since this holds for all f, we obtain

$$[X,Y] = \big(X^j\,\partial_j Y^i - Y^j\,\partial_j X^i\big)\,\frac{\partial}{\partial x^i},$$

as claimed (relabelling the dummy index if desired). □

In particular, for coordinate vector fields we have

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

Thus, if two vector fields have nonzero Lie bracket, then they cannot both be coordinate vector fields in any common coordinate system.

Example. On $\mathbb{R}^2$ with coordinates (x,y), let

$$X = \frac{\partial}{\partial x}, \qquad Y = \frac{\partial}{\partial x} + x\,\frac{\partial}{\partial y}.$$

Then

$$[X,Y] = \Big(X(Y^j) - Y(X^j)\Big)\,\frac{\partial}{\partial x^j} = \frac{\partial}{\partial y},$$

so X and Y do not commute.

The Lie bracket enjoys several fundamental algebraic properties.

Proposition 5.14 (Basic properties of the Lie bracket)**.** *For all $X,Y,Z \in \mathfrak{X}(M)$ and $f,g \in C^\infty(M)$, the Lie bracket satisfies:*

(1) Bilinearity:

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

for all $a,b \in \mathbb{R}$.

(2) Antisymmetry:

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

(3) Jacobi identity:

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

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

$$[fX, gY] = fg\,[X,Y] + (f\,Xg)\,Y - (g\,Yf)\,X.$$

Proof. (1) and (2) follow immediately from the definition $[X,Y]f = X(Yf) - Y(Xf)$ and the linearity of X, Y.

For (3), it is convenient to think of vector fields as linear operators on $C^\infty(M)$. Each $X \in \mathfrak{X}(M)$ acts on smooth functions by

$$C^\infty(M) \longrightarrow C^\infty(M), \qquad f \longmapsto Xf,$$

and our Lie bracket is just the commutator of these operators:

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

(Here $\circ$ denotes composition of operators, which is associative.)

We now compute the Jacobi expression using only linearity and associativity of composition. For any three such operators X, Y, Z, we have

$$\begin{aligned}
[X,[Y,Z]] &= X\big(YZ - ZY\big) - \big(YZ - ZY\big)X \\
&= XYZ - XZY - YZX + ZYX, \\
[Y,[Z,X]] &= Y\big(ZX - XZ\big) - \big(ZX - XZ\big)Y \\
&= YZX - YXZ - ZXY + XZY, \\
[Z,[X,Y]] &= Z\big(XY - YX\big) - \big(XY - YX\big)Z \\
&= ZXY - ZYX - XYZ + YXZ.
\end{aligned}$$

To see the cancellation clearly, it is useful to tabulate the coefficients of each word in X, Y, Z:

word	XYZ	XZY	YXZ	YZX	ZXY	ZYX
$[X,[Y,Z]]$	$+1$	-1	0	-1	0	$+1$
$[Y,[Z,X]]$	0	$+1$	-1	$+1$	-1	0
$[Z,[X,Y]]$	-1	0	$+1$	0	$+1$	-1
sum	0	0	0	0	0	0

Each word in X, Y, Z appears with total coefficient 0, so

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

which is the Jacobi identity.

For (4), let $h \in C^\infty(M)$ and compute:

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

Since this holds for all h, we obtain the stated formula. $\square$

Remark 5.15. When $f \equiv 1$ in (4), we obtain

$$[X, gY] = g[X, Y] + (Xg)\,Y,$$

which looks like a Leibniz rule in the Y-slot; for this reason one often thinks of $[X, Y]$ as the *Lie derivative of* Y *along* X. Note that this is quite different from the derivative of a function: to differentiate a vector field we need another vector field, whereas to differentiate a function we only need a single tangent vector. In particular, even if $X_p = 0$, it is not true in general that $[X, Y](p) = 0$, whereas for functions we have $(Xf)(p) = 0$ whenever $X_p = 0$.

One might ask whether there is a natural way to differentiate vector fields using only a tangent vector at a point, analogous to how we differentiate functions. This leads to the notion of a *connection*. There is no canonical choice of connection on a general smooth manifold, but once we add extra structure (a Riemannian metric, a Hermitian metric, or a sub-Riemannian structure), there are often canonical connections: the Levi–Civita connection in Riemannian geometry, the Chern connection in complex Hermitian geometry, and various canonical connections in sub-Riemannian geometry (for example, the Tanno connection in certain contact metric structures). By contrast, the Lie bracket depends only on the smooth structure and is completely canonical.

Finally, we record how the Lie bracket behaves under smooth maps.

Proposition 5.16. *Let* $f : M \to N$ *be a smooth map, and let* $X_1, X_2 \in \mathfrak{X}(M)$ *and* $Y_1, Y_2 \in \mathfrak{X}(N)$ *be smooth vector fields such that* X_i *is* f*-related to* Y_i *for* $i = 1, 2$*, i.e.*

$$(df)_p(X_i(p)) = Y_i(f(p)) \quad \text{for all } p \in M.$$

Then $[X_1, X_2]$ *is* f*-related to* $[Y_1, Y_2]$.

Proof. Recall that X being f-related to Y is equivalent to the identity

$$X(\phi \circ f) = (Y\phi) \circ f \quad \text{for all } \phi \in C^\infty(N).$$

Using this for X_i and Y_i, we compute, for any $\phi \in C^\infty(N)$,

$$\begin{aligned} [X_1, X_2](\phi \circ f) &= X_1\big(X_2(\phi \circ f)\big) - X_2\big(X_1(\phi \circ f)\big) \\ &= X_1\big((Y_2\phi) \circ f\big) - X_2\big((Y_1\phi) \circ f\big) \\ &= \big(Y_1(Y_2\phi)\big) \circ f - \big(Y_2(Y_1\phi)\big) \circ f \\ &= \big([Y_1, Y_2]\phi\big) \circ f. \end{aligned}$$

This shows that $[X_1, X_2]$ is f-related to $[Y_1, Y_2]$. □

Corollary 5.17. *Let* $F : M \to N$ *be a diffeomorphism, and let* $X_1, X_2 \in \mathfrak{X}(M)$*. Then*

$$F_*[X_1, X_2] \;=\; [F_*X_1,\, F_*X_2],$$

where F_*X *denotes the unique vector field on* N *that is* F*-related to* X*.*

Proof. By definition, F_*X_i is F-related to X_i for $i = 1, 2$. Applying the proposition with $f = F$, X_i as before, and $Y_i = F_*X_i$, we conclude that $[X_1, X_2]$ is F-related to $[F_*X_1, F_*X_2]$. By uniqueness of the F-related vector field, this means $F_*[X_1, X_2] = [F_*X_1, F_*X_2]$. □

Corollary 5.18. *Let M be a smooth manifold and $S \subset M$ an embedded submanifold with inclusion map $\iota : S \hookrightarrow M$. Suppose $X_1, X_2 \in \mathfrak{X}(M)$ are tangent to S. Then $[X_1, X_2]$ is also tangent to S, and*

$$[X_1, X_2]\big|_S \;=\; \big[\, X_1\big|_S,\, X_2\big|_S \,\big],$$

where $X_i\big|_S$ denotes the unique vector field on S that is ι-related to X_i.

Proof. Since X_i is tangent to S, there exists a unique vector field $Y_i := X_i\big|_S \in \mathfrak{X}(S)$ that is ι-related to X_i for $i = 1, 2$. Applying the proposition with $f = \iota$, we see that $[X_1, X_2]$ is ι-related to $[Y_1, Y_2] = [X_1|_S, X_2|_S]$. By the characterization of vector fields tangent to S via ι-related fields, this shows that $[X_1, X_2]$ is tangent to S and that its restriction to S is precisely $[X_1|_S, X_2|_S]$. □

5.3 Integral Curves and Flows

Let M be a smooth manifold and $V \in \mathfrak{X}(M)$ a smooth vector field.

Definition 5.19 (Integral curve / trajectory)**.** A smooth curve $\gamma : (a, b) \to M$ is called an *integral curve* (or *trajectory*) of V if for every $t \in (a, b)$ we have

$$\gamma'(t) := d\gamma_t\Big(\frac{\partial}{\partial t}\Big) = V_{\gamma(t)}.$$

You should already be somewhat familiar with this notion from ordinary differential equations. Let us make the definition more concrete in local coordinates.

Choose a local chart (U, φ) around a point of interest, with

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

On U we can write

$$V = V^i \,\frac{\partial}{\partial x^i}, \qquad V^1, \dots, V^m \in C^\infty(U).$$

If $\gamma : (a, b) \to U$ is a curve, we set

$$\gamma^i(t) := x^i(\gamma(t)).$$

Then

$$\gamma'(t) = d\gamma_t\Big(\frac{\partial}{\partial t}\Big) = (\gamma^i)'(t)\, \frac{\partial}{\partial x^i}\Big|_{\gamma(t)}.$$

On the other hand,

$$V_{\gamma(t)} = V^i(\gamma^1(t), \dots, \gamma^m(t))\, \frac{\partial}{\partial x^i}\Big|_{\gamma(t)}.$$

Thus the condition that γ is a trajectory of V is equivalent to the system of ordinary differential equations

$$(\gamma^i)'(t) = V^i\big(\gamma^1(t), \dots, \gamma^m(t)\big), \qquad i = 1, \dots, m.$$

In vector notation we can write

$$\vec{\gamma}'(t) = \vec{V}(\gamma(t)).$$

This is precisely the simplest type of first-order autonomous ODE system. Therefore, studying trajectories of a vector field on a manifold amounts to studying autonomous ODEs in local charts.

Note that our ODE system does not depend explicitly on t. As in the usual ODE theory, one useful fact that follows directly from the definition is the following "translation lemma".

Lemma 5.20 (Translation lemma)**.** *Let $\gamma : (a, b) \to M$ be a trajectory of V. Fix $t_0 \in \mathbb{R}$ such that $(a - t_0, b - t_0) \subset \mathbb{R}$. Define*

$$\tilde{\gamma} : (a - t_0, b - t_0) \to M, \qquad \tilde{\gamma}(t) := \gamma(t + t_0).$$

Then $\tilde{\gamma}$ is again a trajectory of V.

Proof. Let $\tau^{t_0} : (a - t_0, b - t_0) \to (a, b)$ denote the translation

$$\tau^{t_0}(t) := t + t_0.$$

Then $\tilde{\gamma} = \gamma \circ \tau^{t_0}$. We compute the derivative using the chain rule:

$$\tilde{\gamma}'(t) = d(\gamma \circ \tau^{t_0})_t\Big(\frac{\partial}{\partial t}\Big) = d\gamma_{\tau^{t_0}(t)}\Big(d\tau^{t_0}_t\Big(\frac{\partial}{\partial t}\Big)\Big).$$

Since τ^{t_0} is just a translation of $\mathbb{R}$, its differential is the identity:

$$d\tau^{t_0}_t\Big(\frac{\partial}{\partial t}\Big) = \frac{\partial}{\partial t}.$$

Hence

$$\tilde{\gamma}'(t) = d\gamma_{\tau^{t_0}(t)}\Big(\frac{\partial}{\partial t}\Big) = \gamma'(\tau^{t_0}(t)).$$

But γ is a trajectory of V, so

$$\gamma'(\tau^{t_0}(t)) = V_{\gamma(\tau^{t_0}(t))} = V_{\tilde{\gamma}(t)}.$$

Therefore $\tilde{\gamma}'(t) = V_{\tilde{\gamma}(t)}$ for all t, i.e. $\tilde{\gamma}$ is again a trajectory of V. □

To proceed further, we collect the ODE facts we need into the following lemma.

Lemma 5.21 (ODE-theory lemma)**.** *Let $V \in \mathfrak{X}(M)$ be a smooth vector field and $p \in M$. Then there exist open sets $U_0, U \subset M$ and a number $\varepsilon > 0$ such that*

$$p \in U_0 \subset U,$$

and a smooth map

$$\theta : (-\varepsilon, \varepsilon) \times U_0 \to U$$

with the following properties:

1. *For each $q \in U_0$, the curve*

$$\gamma_q : (-\varepsilon, \varepsilon) \to U, \qquad \gamma_q(t) := \theta(t, q)$$

   is a trajectory of V with $\gamma_q(0) = q$.

2. *If $\tilde{\gamma} : (a,b) \to U$ is any other trajectory of V with $0 \in (a,b)$ and $\tilde{\gamma}(0) = q \in U_0$, then*

$$\tilde{\gamma}(t) = \gamma_q(t) \quad \textit{for all } t \in (a,b) \cap (-\varepsilon, \varepsilon).$$

Sketch of proof. Choose a local chart (U, φ) around p. In coordinates, V becomes a smooth vector field on an open subset of $\mathbb{R}^m$ and the integral curves of V are solutions of a smooth autonomous ODE system

$$\vec{\gamma}'(t) = \vec{V}(\vec{\gamma}(t)).$$

By the standard existence, uniqueness, and smooth dependence on initial conditions for ODEs in $\mathbb{R}^m$, there exist:

- an open neighborhood $U_0 \subset U$ of p,
- a number $\varepsilon > 0$,
- and a smooth map

  $$\tilde{\theta} : (-\varepsilon, \varepsilon) \times \varphi(U_0) \to \varphi(U)$$

  such that for each $x \in \varphi(U_0)$, the curve $t \mapsto \tilde{\theta}(t, x)$ is the unique solution of the ODE with initial condition $\tilde{\theta}(0, x) = x$.

Transporting this construction back to M via the chart φ gives the desired map θ with properties (1) and (2). □

We now use Lemma 5.21 to derive some global consequences on the manifold.

Proposition 5.22 (Uniqueness of trajectories)**.** *Let $\gamma_i : (a_i, b_i) \to M$, $i = 1, 2$, be trajectories of V. Assume there exists $t_0 \in (a_1, b_1) \cap (a_2, b_2)$ such that*

$$\gamma_1(t_0) = \gamma_2(t_0).$$

Then

$$\gamma_1(t) = \gamma_2(t) \quad \textit{for all } t \in (a_1, b_1) \cap (a_2, b_2).$$

Proof. Set

$$I := (a_1, b_1) \cap (a_2, b_2), \qquad S := \{\, t \in I : \gamma_1(t) = \gamma_2(t) \,\}.$$

By assumption, $t_0 \in S$, so S is nonempty.

First, S is closed in I. Indeed, if (t_n) is a sequence in S with $t_n \to t \in I$, then by continuity of γ_1 and γ_2 we have

$$\gamma_1(t) = \lim_{n \to \infty} \gamma_1(t_n) = \lim_{n \to \infty} \gamma_2(t_n) = \gamma_2(t),$$

so $t \in S$.

Next, S is open in I. Let $t \in S$, and put $q := \gamma_1(t) = \gamma_2(t)$. Apply Lemma 5.21 at the point q and time t (using the translation lemma to shift the time origin if needed): there exists $\varepsilon > 0$ such that on the interval $(t - \varepsilon, t + \varepsilon)$ there is a unique trajectory of V passing through q at time t. Both γ_1 and γ_2 are such trajectories, hence they must coincide on $(t - \varepsilon, t + \varepsilon) \cap I$. Thus $(t - \varepsilon, t + \varepsilon) \cap I \subset S$, which shows that S is open in I.

We have shown that S is nonempty, open, and closed in the connected interval I, so $S = I$. Hence $\gamma_1 = \gamma_2$ on I. □

Definition 5.23 (Maximal trajectory)**.** A trajectory $\gamma : I \to M$ of V is called *maximal* if its domain $I \subset \mathbb{R}$ cannot be strictly enlarged while keeping γ a trajectory. That is, there is no trajectory $\tilde{\gamma} : J \to M$ with $I \subsetneq J$ and $\tilde{\gamma}|_I = \gamma$.

Proposition 5.24 (Existence and uniqueness of maximal trajectories)**.** *Let $t_0 \in \mathbb{R}$ and $p \in M$. Then there exists a unique maximal trajectory*

$$\gamma : I_{\max} \to M$$

of V such that $\gamma(t_0) = p$.

Proof. Consider the collection $\mathcal{C}$ of all trajectories

$$\gamma_\alpha : I_\alpha \to M$$

such that $t_0 \in I_\alpha$ and $\gamma_\alpha(t_0) = p$. By Proposition 5.22, any two such trajectories agree on the intersection of their domains: if $t \in I_\alpha \cap I_\beta$, then $\gamma_\alpha(t) = \gamma_\beta(t)$.

Define

$$I_{\max} := \bigcup_\alpha I_\alpha \subset \mathbb{R},$$

and define a map $\gamma : I_{\max} \to M$ by

$$\gamma(t) := \gamma_\alpha(t) \quad \text{for any } \alpha \text{ with } t \in I_\alpha.$$

This is well-defined because whenever $t \in I_\alpha \cap I_\beta$ we have $\gamma_\alpha(t) = \gamma_\beta(t)$.

By construction, γ is a trajectory of V (locally it coincides with some γ_α) and satisfies $\gamma(t_0) = p$. Moreover, if $\tilde{\gamma} : J \to M$ is any trajectory with $t_0 \in J$, $\tilde{\gamma}(t_0) = p$, and $I_{\max} \subsetneq J$, then $\tilde{\gamma}$ belongs to the family $\mathcal{C}$ and hence $J \subset I_{\max}$, a contradiction. Thus $I_{\max}$ cannot be strictly enlarged, so γ is maximal.

Uniqueness follows from Proposition 5.22: if $\gamma_1 : I_1 \to M$ and $\gamma_2 : I_2 \to M$ are maximal trajectories with $\gamma_1(t_0) = \gamma_2(t_0) = p$, then $\gamma_1 = \gamma_2$ on $I_1 \cap I_2$, and maximality forces $I_1 = I_2 = I_{\max}$ and $\gamma_1 = \gamma_2$. □

5.3.1 An Example and Completeness

Let us look at a simple example. Consider $M = \mathbb{R}$ and the vector field

$$V = x^2 \frac{\partial}{\partial x},$$

so the integral curves satisfy the ODE

$$\gamma'(t) = (\gamma(t))^2.$$

If we look for an integral curve with initial condition $\gamma(0) = 1$, we can solve this ODE explicitly. Separating variables gives

$$\frac{d\gamma}{\gamma^2} = dt,$$

so

$$-\frac{1}{\gamma(t)} = t + C.$$

Using $\gamma(0) = 1$ we find $C = -1$, hence

$$-\frac{1}{\gamma(t)} = t - 1 \qquad \Longrightarrow \qquad \gamma(t) = \frac{1}{1-t}.$$

This solution is defined as long as $1 - t \neq 0$, i.e. on the interval $(-\infty, 1)$. One checks that this is indeed the maximal interval of existence of the integral curve with $\gamma(0) = 1$.

This example shows that, in general, the maximal domain of an integral curve need not be all of $\mathbb{R}$; solutions may blow up in finite time.

In many applications, however, it is very convenient if the integral curves of a vector field exist for all times $t \in \mathbb{R}$. This motivates the following definition.

Definition 5.25 (Complete vector field)**.** Let $V \in \mathfrak{X}(M)$ be a smooth vector field. For each $p \in M$ let

$$\gamma_p : \mathcal{D}^{(p)} \to M$$

denote the maximal integral curve of V with $\gamma_p(0) = p$, where $\mathcal{D}^{(p)} \subset \mathbb{R}$ is its maximal domain of definition. We say that V is *complete* if

$$\mathcal{D}^{(p)} = \mathbb{R} \quad \text{for every } p \in M.$$

The next lemma gives a useful criterion for detecting incompleteness: if a maximal integral curve cannot be extended beyond some finite time, then it must "escape to infinity" in the sense that it eventually leaves every compact subset of M.

Lemma 5.26 (Escape lemma)**.** *Let $V \in \mathfrak{X}(M)$ and let $\gamma : (a,b) \to M$ be a maximal integral curve of V with $\gamma(0) = p$. Suppose that $b < \infty$. Then for any compact set $K \subset M$ and any $t_0 \in (a,b)$, we cannot have*

$$\gamma([t_0, b)) \subset K.$$

In other words, as $t \to b^-$ the curve $\gamma(t)$ leaves every compact subset of M.

Proof. Assume for contradiction that there exist a compact set $K \subset M$ and $t_0 \in (a,b)$ such that

$$\gamma([t_0, b)) \subset K.$$

For each $t \in [t_0, b)$ consider the set

$$K_t := \overline{\gamma([t,b))} \subset K.$$

Each K_t is nonempty, closed, and contained in the compact set K, hence compact. Moreover, if $t_1 < t_2$ then

$$\gamma([t_2, b)) \subset \gamma([t_1, b)) \quad \Longrightarrow \quad K_{t_2} \subset K_{t_1},$$

so the family $(K_t)_{t \in [t_0, b)}$ is nested and decreasing. By the finite intersection property for compact sets, the intersection

$$\bigcap_{t \in [t_0, b)} K_t$$

is nonempty. Let q be a point in this intersection.

We now apply Lemma 5.21 (the ODE-theory lemma) to the vector field V at the point q. There exist open sets $U_0 \subset U \subset M$ with $q \in U_0$ and a number $\varepsilon > 0$ such that for each $r \in U_0$ there is a unique integral curve

$$\gamma_r : (-\varepsilon, \varepsilon) \to U$$

of V with $\gamma_r(0) = r$.

Since $q \in K_t = \overline{\gamma([t, b))}$ for every $t \in [t_0, b)$, and $\gamma([t_0, b))$ is contained in K, we can find a sequence $t_n \in [t_0, b)$ with $t_n \to b$ and $\gamma(t_n) \to q$. For n large enough, we have $\gamma(t_n) \in U_0$ and also

$$b - t_n < \frac{\varepsilon}{2}.$$

Choose one such t_n and denote it by t_1. In particular, $\gamma(t_1) \in U_0$ and $b - t_1 < \varepsilon/2$.

Consider now the integral curve

$$\gamma_{\gamma(t_1)} : (-\varepsilon, \varepsilon) \to U$$

with initial condition $\gamma_{\gamma(t_1)}(0) = \gamma(t_1)$. By the translation lemma, the curve

$$\tilde{\gamma}(t) := \gamma_{\gamma(t_1)}(t - t_1)$$

is an integral curve of V defined on $(t_1 - \varepsilon, t_1 + \varepsilon)$ with $\tilde{\gamma}(t_1) = \gamma(t_1)$.

We claim that we can use γ and $\tilde{\gamma}$ to extend γ beyond b. Define a new curve

$$\Gamma : (a, b + \varepsilon/2) \to M$$

by

$$\Gamma(t) := \begin{cases} \gamma(t), & t < b, \\ \tilde{\gamma}(t), & t \in [t_1, b + \varepsilon/2). \end{cases}$$

First we check that Γ is well-defined. On the overlap (t_1, b) both γ and $\tilde{\gamma}$ are defined and are integral curves of V with the same value at $t = t_1$. By the uniqueness part of Lemma 5.21, they must coincide on this overlap. Hence the two definitions agree wherever they both apply, so Γ is a well-defined curve on $(a, b + \varepsilon/2)$.

Moreover, Γ is an integral curve of V on this larger interval: on (a, b) it agrees with γ, and on $(t_1, b + \varepsilon/2)$ it agrees with $\tilde{\gamma}$, both of which are integral curves of V. Thus Γ extends γ beyond b, contradicting the maximality of the domain (a, b).

This contradiction shows that our assumption was false. Hence for any compact $K \subset M$ and any $t_0 \in (a, b)$ we cannot have $\gamma([t_0, b)) \subset K$. ☐

As an immediate consequence, we obtain a useful sufficient condition for completeness.

Corollary 5.27 (Compact support implies completeness)**.** *Let $V \in \mathfrak{X}(M)$ be a smooth vector field with compact support* $\operatorname{supp}(V) \subset M$*. Then V is complete.*

Proof. Let $p \in M$, and let $\gamma_p : \mathcal{D}^{(p)} \to M$ be the maximal integral curve of V with $\gamma_p(0) = p$.

If $p \notin \operatorname{supp}(V)$, then there is a neighborhood U of p on which $V \equiv 0$. The constant curve

$$\gamma_p(t) \equiv p, \qquad t \in \mathbb{R},$$

is an integral curve of V with $\gamma_p(0) = p$, and by uniqueness it must coincide with the maximal integral curve. Thus in this case $\mathcal{D}^{(p)} = \mathbb{R}$.

Now suppose $p \in \mathrm{supp}(V)$. Let $\gamma_p : (a, b) \to M$ be its maximal domain. We show that $b = +\infty$; the argument for $a = -\infty$ is analogous.

Assume by contradiction that $b < \infty$. Observe that the image of γ_p for $t \in [0, b)$ is contained in K by uniqueness:

$$\gamma_p([0, b)) \subset K.$$

Applying Lemma 5.26 with this compact set K and $t_0 = 0$ contradicts the assumption that $b < \infty$. Hence we must have $b = +\infty$.

A symmetric argument applied to the vector field $-V$ shows that $a = -\infty$. Therefore $\mathcal{D}^{(p)} = \mathbb{R}$, and V is complete. □

5.3.2 Flows of Complete Vector Fields

Let $V \in \mathfrak{X}(M)$ be a complete smooth vector field. For each $p \in M$, let $\gamma_p : \mathbb{R} \to M$ denote the maximal integral curve of V satisfying $\gamma_p(0) = p$, which is defined on all of $\mathbb{R}$ by completeness.

We define the *flow* of V to be the map

$$\theta : \mathbb{R} \times M \to M, \qquad \theta(t, p) := \theta_t(p) := \gamma_p(t).$$

Thus the flow is a one-parameter family of maps

$$\{\theta_t : M \to M\}_{t \in \mathbb{R}}.$$

Proposition 5.28 (Flow properties)**.** *Let V be complete. Then:*

$$\theta_0 = \mathrm{id}_M, \qquad \theta_{t_1} \circ \theta_{t_2} = \theta_{t_1 + t_2} \quad \textit{for all } t_1, t_2 \in \mathbb{R}.$$

Proof. The identity $\theta_0 = \mathrm{id}_M$ follows immediately from the definition:

$$\theta_0(p) = \gamma_p(0) = p.$$

For the composition law, fix t_1, t_2 and $p \in M$. Consider the curve

$$t \mapsto \theta_{t_1}(\theta_{t_2}(p)).$$

This is again an integral curve of V (by the translation lemma), and at time $t = 0$ it takes the value $\theta_{t_1+t_2}(p)$. By uniqueness of integral curves, these two integral curves must coincide for all t, hence

$$\theta_{t_1}(\theta_{t_2}(p)) = \theta_{t_1+t_2}(p).$$

□

Thus $\{\theta_t\}$ is a smooth action of the additive group $\mathbb{R}$ on M. The last step is to prove smoothness of the flow.

Lemma 5.29. *The map $\theta : \mathbb{R} \times M \to M$ is smooth.*

Proof. Define

$$W := \{(t,p) \in \mathbb{R} \times M : \theta \text{ is smooth in a neighborhood of } (t,p)\}.$$

By definition, W is open. We claim that for each fixed $p \in M$, the set

$$W_p := \{\, t \in \mathbb{R} : (t,p) \in W \,\}$$

is closed in $\mathbb{R}$. Since $\mathbb{R}$ is connected, once we show that W_p is nonempty, open, and closed, we must have $W_p = \mathbb{R}$ for every p, hence θ is smooth everywhere.

First note that W_p is nonempty. Indeed, at $t = 0$ the ODE-theory lemma gives local smooth dependence on initial conditions near p, so $(0,p) \in W$.

Let $(t_0, p) \in W$. We wish to show that all t sufficiently close to t_0 also belong to W_p, and moreover that W_p is closed.

Fix such a point (t_0, p), and set

$$p_0 := \gamma_p(t_0) = \theta_{t_0}(p).$$

By the ODE-theory lemma, there exist open sets $U_0 \subset U$ with $p_0 \in U_0$ and some $\varepsilon > 0$ such that

$$(t,q) \mapsto \theta(t,q)$$

is smooth for $(t,q) \in (-\varepsilon, \varepsilon) \times U_0$. In particular, θ is smooth on a neighborhood of $(0, p_0)$.

Now suppose t_1 satisfies $|t_1 - t_0| \le \varepsilon/2$. Then $\gamma_p(t_1) = \theta_{t_1}(p)$ lies in U_0. Using the flow property $\theta_{t_1} = \theta_{t_1 - t_0} \circ \theta_{t_0}$ and the smoothness near both (t_0, p) and $(0, p_0)$, we can patch these expressions together exactly as in the proof of the escape lemma to conclude that θ is smooth in a neighborhood of (t_1, p).

This shows that W_p is closed.

□

We have now established that the flow of a complete vector field is a smooth one-parameter group of diffeomorphisms of M.

5.3.3 Further properties of flows of complete vector fields

Let $V \in \mathfrak{X}(M)$ be a complete smooth vector field. For each $p \in M$, let $\gamma_p : \mathbb{R} \to M$ denote the maximal integral curve of V satisfying $\gamma_p(0) = p$, which is defined on all of $\mathbb{R}$ by completeness.

We define the *flow* of V to be the map

$$\theta : \mathbb{R} \times M \to M, \qquad \theta(t,p) := \theta_t(p) := \gamma_p(t).$$

Thus the flow is a one-parameter family of maps

$$\{\theta_t : M \to M\}_{t \in \mathbb{R}}.$$

Proposition 5.30 (Flow properties)**.** *Let V be complete. Then:*

$$\theta_0 = \mathrm{id}_M, \qquad \theta_{t_1} \circ \theta_{t_2} = \theta_{t_1 + t_2} \quad \textit{for all } t_1, t_2 \in \mathbb{R}.$$

Proof. The identity $\theta_0 = \mathrm{id}_M$ follows immediately from the definition:

$$\theta_0(p) = \gamma_p(0) = p.$$

For the composition law, fix $t_1, t_2 \in \mathbb{R}$ and $p \in M$. Consider the curve

$$t \mapsto \theta_t(\theta_{t_2}(p)),$$

which is an integral curve of V. Consider also the curve

$$t \mapsto \theta_{t+t_2}(p).$$

This is again an integral curve of V (by the translation lemma). At time $t = 0$, both curves take the value $\theta_{t_2}(p)$. By uniqueness of integral curves, these two curves must coincide for all t. Hence,

$$\theta_{t_1}(\theta_{t_2}(p)) = \theta_{t_1+t_2}(p).$$

□

Thus $\{\theta_t\}$ is a smooth action of the additive group $\mathbb{R}$ on M.

Now we prove smoothness of the flow.

Lemma 5.31. *The map $\theta : \mathbb{R} \times M \to M$ is smooth.*

Proof. Define

$$W := \{(t, p) \in \mathbb{R} \times M : \theta \text{ is smooth in a neighborhood of } (t, p)\}.$$

By definition, W is open. For each fixed $p \in M$, define

$$W_p := \{\, t \in \mathbb{R} : (t, p) \in W \,\}.$$

We claim that W_p is closed in $\mathbb{R}$. Since $\mathbb{R}$ is connected, once we show that W_p is nonempty, open, and closed, it follows that $W_p = \mathbb{R}$ for every p, and hence θ is smooth everywhere.

The openness of W_p follows directly from the definition.

The set W_p is nonempty. Indeed, at $t = 0$ the ODE-theory lemma gives smooth dependence on initial conditions near p, so $(0, p) \in W$.

It remains to show that W_p is closed. Let t_0 be a limit point of W_p, and set

$$p_0 := \gamma_p(t_0) = \theta_{t_0}(p).$$

By the ODE-theory lemma, there exist open sets $U_0 \subset U$ with $p_0 \in U_0$ and some $\varepsilon > 0$ such that the map

$$(t, q) \mapsto \theta(t, q)$$

is smooth on $(-\varepsilon, \varepsilon) \times U_0$. In particular, θ is smooth in a neighborhood of $(0, p_0)$.

Now suppose t_1 satisfies $|t_1 - t_0| \le \varepsilon/2$ and that $\gamma_p(t_1) = \theta_{t_1}(p)$ lies in U_0. Recall the flow property in the form

$$\theta(t, r) = \theta\big(t - t_1, \theta_{t_1}(r)\big).$$

Since θ is smooth on $(-\varepsilon, \varepsilon) \times U_0$ and U_0 is a neighborhood of $\theta_{t_1}(p)$, continuity of θ_{t_1} implies that there exists a neighborhood V of p such that

$$\theta_{t_1}(V) \subset U_0.$$

It follows that $\theta(t, r)$ is smooth on

$$(t_1 - \varepsilon, t_1 + \varepsilon) \times V \supset \left(t_0 - \tfrac{\varepsilon}{2},\, t_0 + \tfrac{\varepsilon}{2}\right) \times V.$$

Hence θ is smooth in a neighborhood of (t_0, p), so $t_0 \in W_p$.

This shows that W_p is closed. □

We have now established that the flow of a complete vector field is a smooth one-parameter group of diffeomorphisms of M.

5.4 Lie derivatives and Lie brackets

Let $V \in \mathfrak{X}(M)$ be a complete vector field with flow

$$\theta : \mathbb{R} \times M \to M, \qquad \theta(t, p) = \theta_t(p).$$

Recall that each θ_t is a diffeomorphism of M, and that $\{\theta_t\}_{t \in \mathbb{R}}$ is a one-parameter group of diffeomorphisms.

Lie derivatives.

We first recall the push-forward of a vector field by a diffeomorphism.

Definition 5.32 (Push-forward of a vector field)**.** Let $\Phi : M \to M$ be a diffeomorphism and $X \in \mathfrak{X}(M)$. The *push-forward* of X by Φ is the vector field

$$\Phi_* X \in \mathfrak{X}(M), \qquad (\Phi_* X)_q := d\Phi_{\Phi^{-1}(q)}\big(X_{\Phi^{-1}(q)}\big), \quad q \in M.$$

In particular, for the flow θ_t of V we have a family of push-forwards $(\theta_t)_* : \mathfrak{X}(M) \to \mathfrak{X}(M)$.

Definition 5.33 (Lie derivative of functions and vector fields)**.** Let $V \in \mathfrak{X}(M)$ be complete with flow θ_t.

(a) For $f \in C^\infty(M)$, the *Lie derivative* of f along V is

$$(\mathcal{L}_V f)(p) := \frac{d}{dt}\bigg|_{t=0} f\big(\theta_t(p)\big), \qquad p \in M.$$

(b) For $X \in \mathfrak{X}(M)$, the *Lie derivative* of X along V is

$$(\mathcal{L}_V X)_p := \frac{d}{dt}\bigg|_{t=0} \big((\theta_{-t})_* X\big)_p, \qquad p \in M.$$

Thus for functions, $\mathcal{L}_V f = V(f)$ is just differentiation along the integral curves of V. For vector fields, the definition says that we transport X back along the flow of V and differentiate at $t = 0$.

Remark 5.34. At first sight, parts (a) and (b) of the definition look rather different: for functions we differentiate the pullback $f \circ \theta_t$, whereas for vector fields we differentiate the push-forward $(\theta_{-t})_* X$. Conceptually, however, they express the same idea: we transport the object along the flow of V and differentiate at $t = 0$. For functions we use the pullback

$$\theta_t^* f := f \circ \theta_t,$$

while for vector fields we use the push-forward

$$(\theta_{-t})_* X = d\theta_{-t} \circ X \circ \theta_{-t}^{-1},$$

which is the natural way to view X in the "moving frame" given by the inverse flow. In both cases,

$$\mathcal{L}_V(\,\cdot\,) = \frac{d}{dt}\bigg|_{t=0} (\text{transport by } \theta_t).$$

The Lie derivative satisfies a Leibniz rule with respect to multiplication by functions: for $f \in C^\infty(M)$ and $X \in \mathfrak{X}(M)$,

$$\mathcal{L}_V(fX) = f\,\mathcal{L}_V X + (\mathcal{L}_V f)\,X.$$

This follows directly from the definition. Indeed, fix $p \in M$ and $g \in C^\infty(M)$. Then

$$(\mathcal{L}_V(fX))_p(g) = \frac{d}{dt}\bigg|_{t=0} \big((\theta_{-t})_*(fX)\big)_p(g).$$

Using the definition of the push-forward,

$$\big((\theta_{-t})_*(fX)\big)_p(g) = (fX)_{\theta_t(p)}(g \circ \theta_{-t}) = f(\theta_t(p))\,X_{\theta_t(p)}(g \circ \theta_{-t}).$$

Set

$$A(t) := f(\theta_t(p)), \qquad B(t) := X_{\theta_t(p)}(g \circ \theta_{-t}).$$

Then

$$(\mathcal{L}_V(fX))_p(g) = \frac{d}{dt}\bigg|_{t=0} (A(t)B(t)) = A'(0)B(0) + A(0)B'(0).$$

By definition of $\mathcal{L}_V$ on functions,

$$A'(0) = (\mathcal{L}_V f)(p), \qquad B(0) = X_p(g),$$

and by the definition of $\mathcal{L}_V$ on vector fields,

$$B'(0) = \frac{d}{dt}\bigg|_{t=0} \big((\theta_{-t})_* X\big)_p(g) = (\mathcal{L}_V X)_p(g).$$

Hence

$$(\mathcal{L}_V(fX))_p(g) = (\mathcal{L}_V f)(p)\,X_p(g) + f(p)\,(\mathcal{L}_V X)_p(g) = \big((\mathcal{L}_V f)X + f\,\mathcal{L}_V X\big)_p(g).$$

Since this holds for all $g \in C^\infty(M)$, we obtain

$$\mathcal{L}_V(fX) = f\,\mathcal{L}_V X + (\mathcal{L}_V f)\,X.$$

Recall that for vector fields $V, X \in \mathfrak{X}(M)$ the *Lie bracket* $[V, X] \in \mathfrak{X}(M)$ is characterized by

$$[V, X](f) = V(Xf) - X(Vf) \qquad \text{for all } f \in C^\infty(M).$$

Lie brackets as Lie derivatives.

Recall that for vector fields $V, X \in \mathfrak{X}(M)$ the *Lie bracket* $[V, X] \in \mathfrak{X}(M)$ is characterized by

$$[V, X](f) = V(Xf) - X(Vf) \qquad \text{for all } f \in C^\infty(M).$$

Proposition 5.35. *For $V, X \in \mathfrak{X}(M)$ we have*

$$\mathcal{L}_V X = [V, X].$$

Proof. Let θ_t be the flow of V, and fix $p \in M$ and $f \in C^\infty(M)$. By definition of the Lie derivative of a vector field,

$$(\mathcal{L}_V X)_p = \left.\frac{d}{dt}\right|_{t=0} \big((\theta_{-t})_* X\big)_p,$$

so applied to the test function f we get

$$(\mathcal{L}_V X)_p(f) = \left.\frac{d}{dt}\right|_{t=0} \big((\theta_{-t})_* X\big)_p(f).$$

By the definition of push-forward,

$$\big((\theta_{-t})_* X\big)_p = d\theta_{-t}\big|_{\theta_t(p)}\big(X_{\theta_t(p)}\big),$$

and as a derivation acting on f this is

$$\big((\theta_{-t})_* X\big)_p(f) = X_{\theta_t(p)}\big(f \circ \theta_{-t}\big).$$

Thus

$$(\mathcal{L}_V X)_p(f) = \left.\frac{d}{dt}\right|_{t=0} X_{\theta_t(p)}\big(f \circ \theta_{-t}\big).$$

To differentiate this, it is convenient to separate the dependence on the base point and on the function. Define

$$g_t := f \circ \theta_{-t} \in C^\infty(M), \qquad h_t(q) := X_q(g_t),$$

so that

$$X_{\theta_t(p)}\big(f \circ \theta_{-t}\big) = h_t\big(\theta_t(p)\big).$$

Then by the chain rule,

$$\left.\frac{d}{dt}\right|_{t=0} h_t(\theta_t(p)) = \left.\frac{d}{dt}\right|_{t=0} h_t(p) \;+\; \left.\frac{d}{dt}\right|_{t=0} h_0(\theta_t(p)).$$

We now compute the two terms separately.

(1) *Variation of the function.* At a fixed point p we have

$$h_t(p) = X_p(g_t) = X_p(f \circ \theta_{-t}),$$

so

$$\left.\frac{d}{dt}\right|_{t=0} h_t(p) = X_p\left(\left.\frac{d}{dt}\right|_{t=0} (f \circ \theta_{-t}) \right).$$

But by the definition of the Lie derivative of f along V,

$$\left.\frac{d}{dt}\right|_{t=0} f(\theta_{-t}(q)) = -(\mathcal{L}_V f)(q) = -V(f)(q),$$

so in particular at $q = p$,

$$\left.\frac{d}{dt}\right|_{t=0} (f \circ \theta_{-t}) = -V(f).$$

Thus

$$\left.\frac{d}{dt}\right|_{t=0} h_t(p) = X_p\big(-V(f)\big) = -X(Vf)(p).$$

(2) *Variation of the base point.* Here $h_0(q) = X_q(g_0) = X_q(f)$, so

$$h_0(\theta_t(p)) = X_{\theta_t(p)}(f) = (Xf)(\theta_t(p)).$$

Therefore

$$\left.\frac{d}{dt}\right|_{t=0} h_0(\theta_t(p)) = \left.\frac{d}{dt}\right|_{t=0} (Xf)(\theta_t(p)) = (\mathcal{L}_V(Xf))(p) = V(Xf)(p).$$

Combining (1) and (2), we obtain

$$(\mathcal{L}_V X)_p(f) = V(Xf)(p) - X(Vf)(p).$$

Since this holds for every $f \in C^\infty(M)$, we have

$$(\mathcal{L}_V X)_p = [V, X]_p,$$

and therefore $\mathcal{L}_V X = [V, X]$ as vector fields. □

Remark 5.36. The identity $\mathcal{L}_V X = [V, X]$ can be rephrased as a Leibniz rule for the Lie derivative with respect to the natural pairing

$$\mathfrak{X}(M) \times C^\infty(M) \to C^\infty(M), \qquad (X, f) \mapsto Xf.$$

Indeed, the proof shows that for every $f \in C^\infty(M)$,

$$(\mathcal{L}_V X)(f) = V(Xf) - X(Vf),$$

so we can rewrite this as

$$\mathcal{L}_V(Xf) = V(Xf) = (\mathcal{L}_V X)(f) + X(\mathcal{L}_V f).$$

Thus $\mathcal{L}_V$ acts as a derivation with respect to this pairing:

$$\boxed{\mathcal{L}_V(Xf) = (\mathcal{L}_V X)(f) + X(\mathcal{L}_V f).}$$

In other words, Proposition 5.35 itself can be viewed as a Leibniz rule for the Lie derivative.

Remark 5.37 (Jacobi identity via Lie derivatives)**.** We can now give a conceptual proof of the Jacobi identity

$$[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]=0 \qquad \text{for all } X,Y,Z\in\mathfrak{X}(M).$$

Let $Z\in\mathfrak{X}(M)$ with flow θ_t. For any vector field $X\in\mathfrak{X}(M)$ we set

$$X_t := (\theta_t)_*X.$$

Recall that a basic fact about Lie brackets is that f-related vector fields have f-related brackets: if X_t and Y_t are θ_t-related to X and Y respectively, then

$$[X_t,Y_t]=(\theta_t)_*[X,Y].$$

Applying this to X_t and Y_t defined above, we obtain

$$[X_t,Y_t]=(\theta_t)_*[X,Y] \qquad \text{for all } t$$

(where both sides are defined). Differentiating at $t=0$ and using the definition of Lie derivative via the flow, we get

$$\mathcal{L}_Z[X,Y]=\left.\frac{d}{dt}\right|_{t=0}[X_t,Y_t]=[\mathcal{L}_ZX,Y]+[X,\mathcal{L}_ZY]\,.$$

Thus the Lie derivative along Z is a derivation of the Lie bracket:

$$\mathcal{L}_Z[X,Y]=[\mathcal{L}_ZX,Y]+[X,\mathcal{L}_ZY].$$

On the other hand, we have already shown that $\mathcal{L}_ZX=[Z,X]$ for all X. Substituting this into the derivation identity gives

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

Rewriting this as

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

yields precisely the Jacobi identity.

Lemma 5.38. *For any vector fields X,Y we have*

$$[\mathcal{L}_Y,\mathcal{L}_X]=\mathcal{L}_{[Y,X]}$$

as operators on smooth functions and on smooth vector fields.

Proof. On functions. Since $\mathcal{L}_Xf=X(f)$,

$$[\mathcal{L}_Y,\mathcal{L}_X]f=Y(Xf)-X(Yf)=[Y,X](f)=\mathcal{L}_{[Y,X]}f.$$

On vector fields. Using $\mathcal{L}_XZ=[X,Z]$,

$$[\mathcal{L}_Y,\mathcal{L}_X]Z=\mathcal{L}_Y([X,Z])-\mathcal{L}_X([Y,Z])=[Y,[X,Z]]-[X,[Y,Z]].$$

The Jacobi identity

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

is equivalent to

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

Thus

$$[\mathcal{L}_Y,\mathcal{L}_X]Z=[[Y,X],Z]=\mathcal{L}_{[Y,X]}Z.$$

□

5.5 Straightening of vector fields

We now use flows to study the local normal form of vector fields.

Definition 5.39. Let $V \in \mathfrak{X}(M)$. A point $p \in M$ is called a *singular point* of V if $V_p = 0$. If $V_p \neq 0$, we call p a *regular point* of V.

The following theorem shows that near a regular point, a vector field can be "straightened" to a constant vector field in suitable coordinates.

Theorem 5.40 (Rectification (straightening) of a single vector field)**.** *Let $V \in \mathfrak{X}(M)$ and let $p \in M$ be a regular point of V. Then there exists a coordinate chart $(U, x^1, \dots, x^m)$ with $p \in U$ such that*

$$V = \frac{\partial}{\partial x^1} \quad \text{on } U.$$

Moreover, suppose $S \subset M$ is a hypersurface through p with $V_p \notin T_pS$. Then the chart can be chosen so that

$$S \cap U = \{x^1 = 0\}.$$

Proof. We may assume that near p there is a chart $(W, y^1, \dots, y^m)$ with $p \in W$ such that

$$S \cap W = \{y^1 = 0\}.$$

Let θ_t be the local flow of V (defined for $|t|$ sufficiently small and y near p). Consider the map

$$\Phi : (-\varepsilon, \varepsilon) \times (S \cap W) \to M, \qquad \Phi(t, q) := \theta_t(q).$$

For $\varepsilon > 0$ sufficiently small, Φ is well-defined and smooth.

We first note that for every (t, q) in the domain of Φ we have

$$d\Phi_{(t,q)}\Big(\frac{\partial}{\partial t}\Big) = V_{\Phi(t,q)}.$$

Indeed, fix (t_0, q_0) and consider the curve

$$\gamma(s) := \Phi(t_0 + s, q_0) = \theta_{t_0+s}(q_0).$$

By definition of the flow, γ is an integral curve of V, so

$$\gamma'(0) = V_{\gamma(0)} = V_{\theta_{t_0}(q_0)} = V_{\Phi(t_0,q_0)}.$$

On the other hand, by the chain rule in the product $(-\varepsilon, \varepsilon) \times (S \cap W)$,

$$\gamma'(0) = d\Phi_{(t_0,q_0)}\Big(\frac{\partial}{\partial t}\Big).$$

Comparing these two expressions gives the desired identity.

In particular, at $(t, q) = (0, p)$ we have

$$d\Phi_{(0,p)}\Big(\frac{\partial}{\partial t}\Big) = V_p, \qquad d\Phi_{(0,p)}(w) = w \quad \text{for } w \in T_pS,$$

since $\Phi(0, q) = q$ for $q \in S$. By construction, V_p is transverse to T_pS, so $d\Phi_{(0,p)}$ is an isomorphism

$$d\Phi_{(0,p)} : T_{(0,p)}\big((-\varepsilon, \varepsilon) \times S\big) \to T_pM.$$

Thus, by the inverse function theorem, Φ is a diffeomorphism from a neighborhood of $(0, p)$ in $(-\varepsilon, \varepsilon) \times S$ onto a neighborhood U of p in M.

We now use (t, z) as coordinates, where t is the coordinate on $(-\varepsilon, \varepsilon)$ and $z = (z^2, \ldots, z^m)$ are local coordinates on S near p (obtained from restricting $(y^2, \ldots, y^m)$ to S). Define

$$x^1 := t, \qquad x^j := z^j, \quad j = 2, \ldots, m,$$

pulled back to U via Φ^{-1}.

Let $q \in U$ and write $(t, z) := \Phi^{-1}(q)$. Then the curve

$$s \longmapsto (t + s, z)$$

in $(-\varepsilon, \varepsilon) \times S$ is mapped by Φ to

$$s \longmapsto \Phi(t + s, z) = \theta_{t+s}(q_0),$$

where $q_0 \in S$ is fixed. As above, this is an integral curve of V, so its derivative at $s = 0$ is V_q. On the other hand, its derivative at $s = 0$ in the $(x^1, \ldots, x^m)$–coordinates is just $\partial/\partial x^1|_q$. Thus

$$V_q = \frac{\partial}{\partial x^1}\bigg|_q$$

for all $q \in U$, i.e. $V = \partial/\partial x^1$ on U. Moreover, S corresponds to $\{t = 0\}$, hence $S \cap U = \{x^1 = 0\}$, as required. □

5.6 Several commuting vector fields

We now generalize the rectification theorem to several commuting vector fields.

Lemma 5.41. *Let $X, Y \in \mathfrak{X}(M)$ be vector fields with local flows θ_t and η_s respectively. If $[X, Y] = 0$, then for all sufficiently small t, s for which both sides are defined, we have*

$$\theta_t \circ \eta_s = \eta_s \circ \theta_t.$$

Proof. We proceed in two steps.

Step 1: the flow of Y preserves X.

Define

$$X_s := (\eta_{-s})_* X.$$

By the general identity for Lie derivatives,

$$\mathcal{L}_Y X = \frac{d}{ds}\bigg|_0 (\eta_{-s})_* X = [Y, X].$$

A standard argument using the group property $\eta_{s+h} = \eta_h \circ \eta_s$ gives the differential equation

$$\frac{d}{ds} X_s = -(\eta_{-s})_*[Y, X].$$

(Indeed, differentiating $X_{s+h} = (\eta_{-(s+h)})_* X = (\eta_{-h})_* X_s$ at $h = 0$ yields this.)

If $[X, Y] = 0$, then $[Y, X] = 0$ and therefore

$$\frac{d}{ds} X_s = 0.$$

Thus X_s is constant in s, and hence

$$(\eta_{-s})_* X = X_s = X_0 = X.$$

Equivalently,

$$(\eta_s)_* X = X \qquad \text{for all sufficiently small } s.$$

Step 2: commuting of the flows.

Fix $p \in M$ and s such that $\eta_s(p)$ is defined. Consider the curve

$$\gamma(t) := \eta_{-s}(\theta_t(\eta_s(p))).$$

We compute its derivative:

$$\gamma'(t) = d\eta_{-s}\big|_{\theta_t(\eta_s(p))}\big(X_{\theta_t(\eta_s(p))}\big) = (\eta_{-s})_* X\big|_{\gamma(t)}.$$

By Step 1, $(\eta_{-s})_* X = X$, hence

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

Thus γ is an integral curve of X.

On the other hand,

$$t \mapsto \theta_t(p)$$

is also an integral curve of X, and they satisfy the same initial condition:

$$\gamma(0) = \eta_{-s}(\theta_0(\eta_s(p))) = \eta_{-s}(\eta_s(p)) = p = \theta_0(p).$$

By uniqueness of integral curves,

$$\gamma(t) = \theta_t(p)$$

for all small t. Expanding the definition of γ gives

$$\eta_{-s}(\theta_t(\eta_s(p))) = \theta_t(p),$$

or equivalently

$$\theta_t(\eta_s(p)) = \eta_s(\theta_t(p)).$$

As p and s were arbitrary, the flows commute:

$$\theta_t \circ \eta_s = \eta_s \circ \theta_t.$$

□

We can now state and prove the general rectification theorem.

Theorem 5.42 (Rectification of several commuting vector fields)**.** *Let $V_1, \dots, V_k \in \mathfrak{X}(M)$ be smooth vector fields defined near a point $p \in M$ such that:*

1. $[V_i, V_j] = 0$ *for all* $1 \le i, j \le k$ *(the* V_i *pairwise commute),*
2. $V_1(p), \ldots, V_k(p)$ *are linearly independent in* T_pM*.*

Then there exists a coordinate chart $(U, x^1, \ldots, x^m)$ *with* $p \in U$ *such that*

$$V_i = \frac{\partial}{\partial x^i} \quad \text{on } U, \qquad i = 1, \ldots, k.$$

Moreover, suppose $S \subset M$ *is a submanifold of codimension* k *through* p *such that*

$$T_pM = T_pS \oplus \operatorname{span}\{V_1(p), \ldots, V_k(p)\}.$$

Then the chart can be chosen so that

$$S \cap U = \{x^1 = \cdots = x^k = 0\}.$$

Proof. By the submanifold theorem, we may choose an initial chart $(W, y^1, \ldots, y^m)$ around p such that

$$S \cap W = \{y^1 = \cdots = y^k = 0\},$$

and such that the vectors $\partial/\partial y^1, \ldots, \partial/\partial y^k$ at p span the same subspace as $V_1(p), \ldots, V_k(p)$. After a linear change of the coordinates $y^1, \ldots, y^k$, we may assume

$$V_i(p) = \left.\frac{\partial}{\partial y^i}\right|_p, \qquad i = 1, \ldots, k.$$

Let θ^i_t denote the local flow of V_i, defined for $|t|$ small in a neighborhood of p. For $x = (x^1, \ldots, x^m)$ close to $0 \in \mathbb{R}^m$, consider the point

$$q(x) := (0, \ldots, 0, x^{k+1}, \ldots, x^m) \in S$$

in y-coordinates, and define

$$\Phi(x^1, \ldots, x^m) := \theta^1_{x^1} \circ \theta^2_{x^2} \circ \cdots \circ \theta^k_{x^k}\big(q(x)\big).$$

By Lemma 5.41 the flows θ^i_t commute, so the order of composition is irrelevant (for x sufficiently close to 0).

We now use $x = (x^1, \ldots, x^m)$ as coordinates on the image of Φ, so a point of M near p has x-coordinates precisely when it is of the form $\Phi(x)$. In order to identify V_i in these coordinates, fix x and an index $i \in \{1, \ldots, k\}$ and consider the curve

$$\gamma_i(t) := \Phi(x^1, \ldots, x^{i-1}, x^i + t, x^{i+1}, \ldots, x^m).$$

By definition of the coordinate vector field,

$$\gamma_i'(0) = d\Phi_x\Big(\frac{\partial}{\partial x^i}\Big).$$

On the other hand, using the definition of Φ and commutativity of the flows, we can rewrite γ_i as

$$\gamma_i(t) = \theta^1_{x^1} \circ \cdots \circ \theta^{i-1}_{x^{i-1}} \circ \theta^i_{x^i+t} \circ \theta^{i+1}_{x^{i+1}} \circ \cdots \circ \theta^k_{x^k}\big(q(x)\big)$$

$$
\begin{aligned}
&= \theta^i_t\Big(\theta^1_{x^1} \circ \cdots \circ \theta^k_{x^k}(q(x))\Big) \\
&= \theta^i_t\big(\Phi(x)\big).
\end{aligned}
$$

Thus γ_i is precisely the integral curve of V_i through the point $\Phi(x)$, and so

$$\gamma_i'(0) = V_i(\Phi(x)).$$

Comparing with the expression above, we obtain

$$d\Phi_x\Big(\frac{\partial}{\partial x^i}\Big) = V_i(\Phi(x)), \qquad x \text{ near } 0, \quad i = 1, \dots, k. \tag{$*$}$$

Next, we examine $d\Phi_0$. For $i > k$, varying x^i moves $q(x)$ tangentially along S, and at $x = 0$ we have

$$\Phi(0, \dots, 0, x^{k+1}, \dots, x^m) = q(x), \quad q(0) = p.$$

Hence

$$d\Phi_0\Big(\frac{\partial}{\partial x^i}\Big) = \frac{\partial}{\partial y^i}\bigg|_p, \qquad i = k+1, \dots, m.$$

For $1 \le i \le k$, applying $(*)$ at $x = 0$ gives

$$d\Phi_0\Big(\frac{\partial}{\partial x^i}\Big) = V_i(\Phi(0)) = V_i(p), \qquad i = 1, \dots, k.$$

By construction, the vectors

$$V_1(p), \dots, V_k(p), \quad \frac{\partial}{\partial y^{k+1}}\bigg|_p, \dots, \frac{\partial}{\partial y^m}\bigg|_p$$

form a basis of T_pM, so $d\Phi_0 : T_0\mathbb{R}^m \to T_pM$ is an isomorphism. The inverse function theorem then implies that there exists a neighborhood U_0 of $0 \in \mathbb{R}^m$ such that

$$\Phi : U_0 \to U := \Phi(U_0)$$

is a diffeomorphism onto an open neighborhood U of p in M.

Now regard $x = (x^1, \dots, x^m)$ as coordinates on U via Φ^{-1}. The relation $(*)$ holds for all $x \in U_0$, so in these coordinates $d\Phi_x(\partial/\partial x^i) = V_i(\Phi(x))$ simply says that on U we have

$$V_i = \frac{\partial}{\partial x^i}, \qquad i = 1, \dots, k.$$

Finally, note that if $x^1 = \cdots = x^k = 0$, then

$$\Phi(0, \dots, 0, x^{k+1}, \dots, x^m) = q(x) \in S,$$

so $\{x^1 = \cdots = x^k = 0\} \subset S \cap U$. Conversely, if $\Phi(x) \in S$, then, by construction, the unique point of S on the joint orbit of the commuting flows through $q(x)$ is $q(x)$ itself, which forces $x^1 = \cdots = x^k = 0$. Hence

$$S \cap U = \{x^1 = \cdots = x^k = 0\},$$

and the theorem is proved. $\square$

We now compute the first nontrivial term in the commutator of the flows of X and Y. Fix $p \in M$ and define

$$\Phi_{t,s}(p) := \theta_{-t}\big(\eta_{-s}(\theta_t(\eta_s(p)))\big),$$

where θ_t and η_s are the flows of X and Y. For any $f \in C^\infty(M)$ set

$$\varphi(t,s) := f\big(\Phi_{t,s}(p)\big).$$

We compute the partial derivatives of φ at $(0,0)$. Using pullback notation,

$$\varphi(t,s) = \big(\eta_s^*\theta_t^*\eta_{-s}^*\theta_{-t}^* f\big)(p).$$

Recall the standard first-order expansions

$$\theta_t^* = \mathrm{id} + t\mathcal{L}_X + O(t^2), \qquad \eta_s^* = \mathrm{id} + s\mathcal{L}_Y + O(s^2),$$

and similarly

$$\theta_{-t}^* = \mathrm{id} - t\mathcal{L}_X + O(t^2), \qquad \eta_{-s}^* = \mathrm{id} - s\mathcal{L}_Y + O(s^2).$$

Since $\Phi_{t,0}(p) = \Phi_{0,s}(p) = p$, it follows immediately that

$$\partial_t\varphi(0,0) = \partial_s\varphi(0,0) = \partial_{tt}\varphi(0,0) = \partial_{ss}\varphi(0,0) = 0.$$

We compute the mixed derivative. Set

$$G(t,s) := \eta_s^*\theta_t^*\eta_{-s}^*\theta_{-t}^*.$$

Then $\varphi(t,s) = (G(t,s)f)(p)$, so

$$\partial_{ts}\varphi(0,0) = (\partial_{ts}G)(0,0)f(p).$$

First differentiate in t at $t = 0$:

$$\partial_t G(t,s)\big|_{t=0} = \eta_s^*(\mathcal{L}_X)\eta_{-s}^* - \mathcal{L}_X.$$

Differentiating this in s at $s = 0$ yields

$$\partial_{ts}G(0,0) = \mathcal{L}_Y\mathcal{L}_X - \mathcal{L}_X\mathcal{L}_Y = [\mathcal{L}_Y, \mathcal{L}_X] = \mathcal{L}_{[Y,X]} = -\mathcal{L}_{[X,Y]}.$$

Hence

$$\partial_{ts}\varphi(0,0) = -[X,Y]f(p).$$

The two-variable Taylor expansion therefore gives

$$\varphi(t,s) = f(p) - ts\,[X,Y]f(p) + O\big(|ts|(|t| + |s|)\big).$$

Since this holds for every smooth function f, in local coordinates around p we obtain the vector expansion

$$\Phi_{t,s}(p) = p - ts\,[X,Y]_p + O\big(|ts|(|t| + |s|)\big).$$

This shows that the leading term in the failure of the flows of X and Y to commute is exactly the Lie bracket $[X,Y]$.

Alternative proof of Lemma 5.41. We use the commutator expansion of the flows. Fix $p \in M$ and consider

$$\Phi_{t,s}(p) := \theta_{-t}\big(\eta_{-s}(\theta_t(\eta_s(p)))\big).$$

For any $f \in C^\infty(M)$ set

$$\varphi(t,s) := f\big(\Phi_{t,s}(p)\big).$$

As computed above, using pullbacks and the definition of the Lie derivative,

$$\partial_{ts}\varphi(0,0) = -[X,Y]f(p),$$

while

$$\partial_t\varphi(0,0) = \partial_s\varphi(0,0) = \partial_{tt}\varphi(0,0) = \partial_{ss}\varphi(0,0) = 0.$$

Thus the two–variable Taylor expansion at $(0,0)$ yields

$$\varphi(t,s) = f(p) - ts\,[X,Y]f(p) + R_f(t,s),$$

where the remainder satisfies

$$R_f(t,s) = O\big(|ts|(|t|+|s|)\big) \quad \text{as } (t,s) \to (0,0).$$

If $[X,Y] = 0$, the mixed term vanishes and we obtain

$$\varphi(t,s) = f(p) + R_f(t,s), \qquad R_f(t,s) = O\big(|ts|(|t|+|s|)\big).$$

In a coordinate chart around p, this is equivalent to

$$\Phi_{t,s}(p) = p + E(t,s), \qquad E(t,s) = O\big(|ts|(|t|+|s|)\big).$$

Moreover, by smoothness, this estimate is uniform for p in a fixed compact set and for $|t|, |s|$ sufficiently small.

Now fix small t, s and subdivide the rectangle $[0,t] \times [0,s]$ into an $N \times N$ grid of smaller rectangles of size

$$\Delta t = \frac{t}{N}, \qquad \Delta s = \frac{s}{N}.$$

One checks that $\Phi_{t,s}$ can be written as a composition of N^2 "elementary commutators" of the form

$$q \longmapsto \Phi_{\Delta t,\Delta s}(q) = \theta_{-\Delta t}\big(\eta_{-\Delta s}(\theta_{\Delta t}(\eta_{\Delta s}(q)))\big),$$

based at various points q along the grid.

For each such small rectangle and each base point q in a fixed compact neighborhood of p, the above Taylor expansion gives

$$\Phi_{\Delta t,\Delta s}(q) = q + O\big(|\Delta t\,\Delta s|(|\Delta t| + |\Delta s|)\big),$$

with a constant independent of q (for N large enough so that all points stay in the chosen neighborhood). Since $|\Delta t| = |t|/N$ and $|\Delta s| = |s|/N$, we obtain

$$\big|\Phi_{\Delta t,\Delta s}(q) - q\big| \;\le\; C\,\frac{|ts|}{N^2}\Big(\frac{|t|}{N} + \frac{|s|}{N}\Big) \;=\; C\,\frac{|ts|(|t|+|s|)}{N^3},$$

for some constant C.

Composing over the N^2 small rectangles, the total deviation of $\Phi_{t,s}(p)$ from p is bounded by

$$N^2 \cdot C \, \frac{|ts|(|t| + |s|)}{N^3} = C \, \frac{|ts|(|t| + |s|)}{N}.$$

Letting $N \to \infty$ shows that

$$\Phi_{t,s}(p) = p$$

for all sufficiently small t, s. Since p was arbitrary, we conclude that $\Phi_{t,s} = \mathrm{id}$ near $(0, 0)$, which is equivalent to

$$\theta_t \circ \eta_s = \eta_s \circ \theta_t$$

for all small t, s. This recovers Lemma 5.41 by a discrete argument. □

Chapter 6

Vector Bundles, Tensor Fields, Differential Forms, and Stokes' Theorem

We now turn from vector fields to the general language of vector bundles and tensor calculus on manifolds. After introducing vector bundles, sections, frames, and bundle homomorphisms, we develop the standard fiberwise constructions such as dual bundles, tensor products, and exterior powers. This leads naturally to tensor bundles, Lie derivatives of tensor fields, differential forms, the exterior derivative, and finally integration on oriented manifolds and Stokes' theorem together with several applications.

6.1 Vector Bundles

We introduce *vector bundles* in the smooth category. (One can develop the theory equally well in the continuous category using purely topological language; here we will work throughout with smooth manifolds and smooth maps.)

Definition 6.1 (Smooth real vector bundle)**.** Let M be a smooth manifold. A *(real) vector bundle of rank k over M* consists of a smooth manifold E (the *total space*) together with a smooth surjective submersion

$$\pi : E \to M$$

such that:

1. For each $p \in M$, the fibre

   $$E_p := \pi^{-1}(p)$$

   is endowed with the structure of a k-dimensional real vector space.

2. (*Local triviality*) For each $p \in M$ there exist an open neighborhood $U \subset M$ of p and a diffeomorphism

   $$\Phi : \pi^{-1}(U) \longrightarrow U \times \mathbb{R}^k,$$

   called a *local trivialization* of E over U, such that:

(a) $\mathrm{pr}_1 \circ \Phi = \pi$ on $\pi^{-1}(U)$, where $\mathrm{pr}_1 : U \times \mathbb{R}^k \to U$ is the projection;

(b) for each $q \in U$, the restriction

$$\Phi|_{E_q} : E_q \longrightarrow \{q\} \times \mathbb{R}^k \simeq \mathbb{R}^k$$

is a linear isomorphism of real vector spaces.

A rank-1 vector bundle is often called a *(real) line bundle*. We sometimes refer to E itself as *the vector bundle* and to π as the *bundle projection*.

Definition 6.2 (Trivial bundle)**.** A rank-k vector bundle $\pi : E \to M$ is called *trivial* if it is (bundle-)diffeomorphic to the product bundle $M \times \mathbb{R}^k$, i.e. if there exists a diffeomorphism $\Psi : E \to M \times \mathbb{R}^k$ such that $\mathrm{pr}_1 \circ \Psi = \pi$ and $\Psi|_{E_p} : E_p \to \{p\} \times \mathbb{R}^k$ is linear for every $p \in M$.

Examples

Example 6.3 (Product bundle)**.** Let M be a smooth manifold. The projection

$$\pi : M \times \mathbb{R}^k \to M, \qquad \pi(p, v) = p$$

defines a rank-k vector bundle, called the *product bundle*. It is trivial by definition.

Example 6.4 (The Möbius line bundle over S^1)**.** Define an action of $\mathbb{Z}$ on $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ by

$$n \cdot (x, y) := (x + n,\, (-1)^n y), \qquad n \in \mathbb{Z}.$$

Let $E := \mathbb{R}^2/\mathbb{Z}$ be the quotient and $q : \mathbb{R}^2 \to E$ the quotient map. Let $S^1 := \mathbb{R}/\mathbb{Z}$ and denote by $[x] \in S^1$ the class of $x \in \mathbb{R}$. Define

$$\pi : E \to S^1, \qquad \pi\big(q(x, y)\big) = [x].$$

Then $\pi : E \to S^1$ is a rank-1 (real) vector bundle: each fibre is naturally identified with $\mathbb{R}$ via $y \mapsto q(x, y)$, and local trivializations exist (e.g. over any open arc in S^1 admitting a smooth lift to $\mathbb{R}$). The total space E is diffeomorphic to the Möbius strip, so this is a nontrivial line bundle over S^1.

Example 6.5 (Tangent bundle)**.** Let M be a smooth m-manifold. The tangent bundle

$$\pi : TM \to M, \qquad \pi(v_p) = p$$

is a rank-m vector bundle. Indeed, for any chart (U, φ) with $\varphi : U \to \varphi(U) \subset \mathbb{R}^m$, the induced map

$$T\varphi : \pi^{-1}(U) = TU \longrightarrow \varphi(U) \times \mathbb{R}^m, \qquad v_p \longmapsto \big(\varphi(p),\, d\varphi_p(v_p)\big)$$

is a local trivialization.

6.1.1 Sections and Frames

Let $\pi : E \to M$ be a (smooth real) vector bundle of rank k over a smooth manifold M.

Definition 6.6 (Sections)**.** A *(smooth) section* of E is a smooth map $\sigma : M \to E$ such that

$$\pi \circ \sigma = \mathrm{id}_M .$$

Equivalently, $\sigma(p) \in E_p$ for every $p \in M$. We denote the set of all (global) smooth sections by

$$\Gamma(E) := \Gamma(M, E).$$

More generally, if $U \subset M$ is open, a *local section over* U is a smooth map $\sigma : U \to E$ satisfying

$$\pi \circ \sigma = \mathrm{id}_U .$$

The set of smooth sections over U is denoted by $\Gamma(U, E)$. When it is useful to emphasize the distinction, we refer to elements of $\Gamma(E)$ as *global sections*.

Definition 6.7 (Support of a section)**.** For $\sigma \in \Gamma(E)$ we define its *support* by

$$\mathrm{supp}(\sigma) := \overline{\{\, p \in M : \sigma(p) \neq 0 \in E_p \,\}}.$$

(And similarly for $\sigma \in \Gamma(U, E)$, taking the closure in U.)

Example 6.8 (Zero section)**.** Every vector bundle $\pi : E \to M$ has a canonical global section

$$\zeta : M \to E, \qquad \zeta(p) := 0 \in E_p,$$

called the *zero section* of E.

Example 6.9 (Vector fields)**.** Sections of the tangent bundle are precisely vector fields:

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

Example 6.10 (Product bundles)**.** If $E = M \times \mathbb{R}^k$ is the product bundle with projection $\pi(p, v) = p$, then there is a natural identification

$$\Gamma(E) \;\cong\; C^\infty(M, \mathbb{R}^k).$$

Indeed, given $f \in C^\infty(M, \mathbb{R}^k)$ we obtain a section

$$\tilde{f} : M \to M \times \mathbb{R}^k, \qquad \tilde{f}(p) := (p, f(p)),$$

and conversely any section $\sigma(p) = (p, f(p))$ determines such a map f.

The $C^\infty(M)$–module structure on $\Gamma(E)$

Just as for vector fields, $\Gamma(E)$ is naturally a module over the ring $C^\infty(M)$. Namely, for $a \in C^\infty(M)$ and $\sigma, \tau \in \Gamma(E)$ we define

$$(\sigma + \tau)(p) := \sigma(p) + \tau(p) \in E_p, \qquad (a\sigma)(p) := a(p)\,\sigma(p) \in E_p.$$

These operations are well-defined since each fibre E_p is a vector space, and they are smooth because, in any local trivialization, they are given by the usual pointwise operations on smooth $\mathbb{R}^k$-valued maps. With these operations, $\Gamma(E)$ is an abelian group under addition and satisfies the usual module axioms over $C^\infty(M)$.

6.1.2 Frames and trivializations

To perform concrete computations, it is often convenient to regard a vector bundle as a smoothly varying family of vector spaces. A key tool is the notion of a *local frame*.

Definition 6.11 (Linear independence, spanning, frames)**.** Let $U \subset M$ be open and let $\sigma_1, \dots, \sigma_k \in \Gamma(U, E)$.

1. The k-tuple $(\sigma_1, \dots, \sigma_k)$ is *pointwise linearly independent* if for each $p \in U$ the vectors
$$\sigma_1(p), \dots, \sigma_k(p) \in E_p$$
are linearly independent.

2. The k-tuple *spans E over U* if for each $p \in U$ the vectors $\sigma_1(p), \dots, \sigma_k(p)$ span the fibre E_p.

3. A *local frame* for E over U is an ordered k-tuple $(\sigma_1, \dots, \sigma_k)$ of local sections over U that is pointwise linearly independent and spans E over U. Equivalently,
$$(\sigma_1(p), \dots, \sigma_k(p))$$
is a basis of E_p for every $p \in U$.

4. If $U = M$, we call $(\sigma_1, \dots, \sigma_k)$ a *global frame*.

Example 6.12 (Standard frame on a product bundle)**.** If $E = M \times \mathbb{R}^k$, let $(e_1, \dots, e_k)$ be the standard basis of $\mathbb{R}^k$. Then
$$\tilde{e}_i : M \to M \times \mathbb{R}^k, \qquad \tilde{e}_i(p) := (p, e_i),$$
defines a global frame $(\tilde{e}_1, \dots, \tilde{e}_k)$ for E.

Frames and trivializations

In fact, local frames and local trivializations are equivalent data.

Lemma 6.13 (Trivializations yield frames)**.** *Let $\Phi : \pi^{-1}(U) \to U \times \mathbb{R}^k$ be a local trivialization and let $(e_1, \dots, e_k)$ be the standard basis of $\mathbb{R}^k$. Define*
$$\sigma_i : U \to E, \qquad \sigma_i(p) := \Phi^{-1}(p, e_i).$$
Then $(\sigma_1, \dots, \sigma_k)$ is a local frame for E over U, called the frame associated with Φ.

Lemma 6.14 (Frames yield trivializations)**.** *Conversely, let $(\sigma_1, \dots, \sigma_k)$ be a local frame for E over U. Define*
$$\Psi : U \times \mathbb{R}^k \longrightarrow \pi^{-1}(U), \qquad \Psi\big(p, (v^1, \dots, v^k)\big) := \sum_{i=1}^k v^i \, \sigma_i(p).$$
Then Ψ is a diffeomorphism, and its inverse
$$\Phi := \Psi^{-1} : \pi^{-1}(U) \to U \times \mathbb{R}^k$$
is a local trivialization of E over U.

Proof. Since $(\sigma_1(p), \ldots, \sigma_k(p))$ is a basis of E_p for each $p \in U$, the map Ψ is bijective: for fixed p it is exactly the linear isomorphism $\mathbb{R}^k \to E_p$ sending $(v^1, \ldots, v^k)$ to $\sum_i v^i \sigma_i(p)$.

To prove smoothness and smoothness of the inverse, it suffices to work locally. Fix $p \in U$ and choose an open neighborhood $V \subset U$ of p together with a local trivialization $\Theta : \pi^{-1}(V) \to V \times \mathbb{R}^k$. Write

$$\Theta \circ \sigma_i(q) = (q,\, s_i(q)), \qquad s_i \in C^\infty(V, \mathbb{R}^k).$$

In coordinates, the map $\Theta \circ \Psi : V \times \mathbb{R}^k \to V \times \mathbb{R}^k$ takes the form

$$(\Theta \circ \Psi)\big(q, v\big) = \Big(q,\, \sum_{i=1}^{k} v^i\, s_i(q)\Big), \qquad v = (v^1, \ldots, v^k).$$

This map is smooth. Moreover, for each $q \in V$ the vectors $s_1(q), \ldots, s_k(q) \in \mathbb{R}^k$ form a basis (because Θ restricts to linear isomorphisms on fibres), so the $k \times k$ matrix whose columns are $s_i(q)$ is invertible. Hence, for each fixed q, the map $v \mapsto \sum_i v^i s_i(q)$ is an isomorphism of $\mathbb{R}^k$ and its inverse depends smoothly on q (e.g. by Cramer's rule, or by smoothness of matrix inversion on $\mathrm{GL}(k, \mathbb{R})$). Therefore $(\Theta \circ \Psi)^{-1}$ is smooth, and hence Ψ^{-1} is smooth on $\pi^{-1}(V)$. Since p was arbitrary, Ψ is a diffeomorphism on $\pi^{-1}(U)$, and $\Phi = \Psi^{-1}$ is a local trivialization. □

6.1.3 Vector bundles via transition functions

When defining manifolds, there are two complementary points of view. One starts from a topological space and imposes regularity conditions, leading to the *abstract* definition of a smooth manifold. The other starts from an *atlas*: one glues open subsets of Euclidean space along smooth transition maps to obtain a manifold.

Vector bundles admit the same two viewpoints. Our definition so far has followed the first approach (a total space together with local trivializations). We now explain the second approach, in which a bundle is described by its *transition functions*.

Definition 6.15 (Transition functions and cocycle conditions)**.** Let $\pi : E \to M$ be a rank-k vector bundle. Choose an open cover $\{U_\alpha\}_{\alpha \in I}$ of M such that the restricted bundle $E|_{U_\alpha} := \pi^{-1}(U_\alpha) \to U_\alpha$ is trivial for every α. Fix local trivializations

$$\Phi_\alpha : \pi^{-1}(U_\alpha) \longrightarrow U_\alpha \times \mathbb{R}^k.$$

By Lemma 6.13 this is equivalent to choosing a local frame $(\sigma_1^\alpha, \ldots, \sigma_k^\alpha)$ over U_α.

On the overlap $U_{\alpha\beta} := U_\alpha \cap U_\beta$, the two trivializations differ by a smooth map with values in $\mathrm{GL}(k, \mathbb{R})$. Concretely, there is a unique map

$$\varphi_{\alpha\beta} : U_{\alpha\beta} \longrightarrow \mathrm{GL}(k, \mathbb{R})$$

such that, for each $p \in U_{\alpha\beta}$,

$$\sigma_i^\alpha(p) = \sum_{j=1}^{k} (\varphi_{\alpha\beta}(p))_i^{\,j}\, \sigma_j^\beta(p), \qquad i = 1, \ldots, k. \tag{6.1}$$

Equivalently, in terms of trivializations one has

$$\Phi_\alpha \circ \Phi_\beta^{-1}(p, v) = \big(p, \varphi_{\alpha\beta}(p)\, v\big), \qquad (p, v) \in U_{\alpha\beta} \times \mathbb{R}^k. \tag{6.2}$$

The family $\{\varphi_{\alpha\beta}\}$ satisfies the *cocycle conditions*:

$$\varphi_{\alpha\alpha}(p) = I_k \qquad \text{for all } p \in U_\alpha, \tag{6.3}$$

$$\varphi_{\alpha\beta}(p)\,\varphi_{\beta\gamma}(p) = \varphi_{\alpha\gamma}(p) \qquad \text{for all } p \in U_\alpha \cap U_\beta \cap U_\gamma. \tag{6.4}$$

In particular, setting $\gamma = \alpha$ in (6.4) yields

$$\varphi_{\alpha\beta} = \varphi_{\beta\alpha}^{-1} \quad \text{on } U_{\alpha\beta}. \tag{6.5}$$

The moral is that a vector bundle is completely encoded by its transition functions. Conversely, any collection of transition functions satisfying the cocycle conditions can be used to *construct* a vector bundle by gluing.

Lemma 6.16 (Vector bundle construction from transition functions)**.** *Let $\{U_\alpha\}_{\alpha\in I}$ be an open cover of M, and let*

$$\varphi_{\alpha\beta} \in C^\infty(U_\alpha \cap U_\beta, \mathrm{GL}(k,\mathbb{R}))$$

be a family of maps satisfying the cocycle conditions (6.3)–(6.4). Then there exists a rank-k vector bundle $\pi : E \to M$ and local trivializations $\Phi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^k$ such that the associated transition functions are precisely $\varphi_{\alpha\beta}$.

Proof. Define a set

$$\widetilde{E} := \bigsqcup_{\alpha\in I} (U_\alpha \times \mathbb{R}^k),$$

the disjoint union of copies of $U_\alpha \times \mathbb{R}^k$. We write elements as triples (p, v, α) with $p \in U_\alpha$ and $v \in \mathbb{R}^k$.

Define an equivalence relation $\sim$ on $\widetilde{E}$ by declaring that

$$(p, v, \alpha) \sim (p, w, \beta) \quad \Longleftrightarrow \quad p \in U_\alpha \cap U_\beta \;\text{ and }\; w = \varphi_{\beta\alpha}(p)\,v.$$

We check that this is indeed an equivalence relation. Reflexivity follows from $\varphi_{\alpha\alpha} = I_k$. Symmetry follows from $\varphi_{\beta\alpha} = \varphi_{\alpha\beta}^{-1}$. For transitivity, suppose $(p, v, \alpha) \sim (p, w, \beta)$ and $(p, w, \beta) \sim (p, z, \gamma)$. Then $w = \varphi_{\beta\alpha}(p)v$ and $z = \varphi_{\gamma\beta}(p)w$, hence

$$z = \varphi_{\gamma\beta}(p)\varphi_{\beta\alpha}(p)v = \varphi_{\gamma\alpha}(p)v$$

by the cocycle condition (6.4), so $(p, v, \alpha) \sim (p, z, \gamma)$.

Let

$$E := \widetilde{E}/\sim$$

and denote the equivalence class of (p, v, α) by $[p, v, \alpha]$. Define

$$\pi : E \to M, \qquad \pi\big([p, v, \alpha]\big) := p.$$

This is well-defined because $\sim$ only relates points with the same base point p.

For each α, define a map

$$\iota_\alpha : U_\alpha \times \mathbb{R}^k \to E, \qquad \iota_\alpha(p, v) := [p, v, \alpha].$$

These maps are injective, and on overlaps one has

$$\iota_\alpha(p, v) = \iota_\beta\big(p, \varphi_{\beta\alpha}(p)v\big), \qquad p \in U_\alpha \cap U_\beta.$$

We now give E the unique topology and smooth structure such that each ι_α is a smooth embedding and the images $\iota_\alpha(U_\alpha \times \mathbb{R}^k)$ form an open cover of E, with transition maps on overlaps given by

$$(p, v) \longmapsto \big(p, \varphi_{\beta\alpha}(p)v\big),$$

which are smooth by assumption. With this smooth structure, π is a smooth submersion.

Finally, define local trivializations by

$$\Phi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^k, \qquad \Phi_\alpha\big([p, v, \alpha]\big) := (p, v).$$

This is well-defined and a diffeomorphism onto $U_\alpha \times \mathbb{R}^k$, and the transition functions between Φ_α and Φ_β are exactly $\varphi_{\alpha\beta}$ by construction (cf. (6.2)). □

6.1.4 Bundle homomorphisms

We now consider maps between vector bundles, often called *bundle homomorphisms*. As usual, we restrict attention to maps that are compatible with the bundle structure.

Definition 6.17 (Bundle homomorphism)**.** Let $\pi : E \to M$ and $\pi' : F \to N$ be smooth vector bundles. A smooth map $f : E \to F$ is called a *bundle homomorphism* if there exists a smooth map $f_0 : M \to N$ such that

$$\pi' \circ f = f_0 \circ \pi,$$

and such that for every $p \in M$ the restricted map

$$f|_{E_p} : E_p \longrightarrow F_{f_0(p)}$$

is linear. We express the relationship between f and f_0 by saying that f *covers* f_0.

A bijective bundle homomorphism $f : E \to F$ whose inverse is also a bundle homomorphism is called a *bundle isomorphism*. A bundle homomorphism *over* M is a bundle homomorphism covering the identity map $\mathrm{id}_M : M \to M$.

If E and F are vector bundles over the same base M and $f : E \to F$ is a bundle homomorphism over M, then f induces a natural map on sections.

Definition 6.18 (Induced map on sections)**.** Let $f : E \to F$ be a bundle homomorphism over M. Define

$$\tilde{f} : \Gamma(E) \longrightarrow \Gamma(F), \qquad (\tilde{f}(\sigma))(p) := f\big(\sigma(p)\big).$$

Then $\tilde{f}$ is linear over $C^\infty(M)$, i.e.

$$\tilde{f}(\sigma_1 + \sigma_2) = \tilde{f}(\sigma_1) + \tilde{f}(\sigma_2), \qquad \tilde{f}(a\,\sigma) = a\,\tilde{f}(\sigma)$$

for all $\sigma, \sigma_1, \sigma_2 \in \Gamma(E)$ and $a \in C^\infty(M)$.

In fact, this construction can be reversed: $C^\infty(M)$–linear maps on sections arise precisely from bundle homomorphisms.

Lemma 6.19 (Characterization of bundle homomorphisms)**.** *Let $\pi : E \to M$ and $\pi' : F \to M$ be smooth vector bundles over the same base. Suppose*

$$T : \Gamma(E) \longrightarrow \Gamma(F)$$

is a map that is linear over $C^\infty(M)$. Then there exists a unique bundle homomorphism $f : E \to F$ over M such that

$$T(\sigma) = \tilde{f}(\sigma) = f \circ \sigma \qquad \textit{for all } \sigma \in \Gamma(E).$$

Proof. We proceed in several steps.

Step 1: Locality. We first show that T is local. Namely, if $\sigma_1, \sigma_2 \in \Gamma(E)$ satisfy $\sigma_1 = \sigma_2$ on an open set $U \subset M$, then $T(\sigma_1) = T(\sigma_2)$ on U.

Indeed, let $\rho \in C^\infty(M)$ be a bump function with $\rho \equiv 1$ on a neighborhood of a given point $p \in U$ and $\operatorname{supp} \rho \subset U$. Then $\rho(\sigma_1 - \sigma_2) = 0$, and by $C^\infty(M)$–linearity we obtain

$$\rho\big(T(\sigma_1) - T(\sigma_2)\big) = T\big(\rho(\sigma_1 - \sigma_2)\big) = 0.$$

Evaluating at p and using $\rho(p) = 1$ yields $T(\sigma_1)(p) = T(\sigma_2)(p)$, as claimed.

Step 2: Pointwise dependence. We now show that T is pointwise in the sense that $T(\sigma)(p)$ depends only on the value $\sigma(p) \in E_p$.

Let $\sigma_1, \sigma_2 \in \Gamma(E)$ satisfy $\sigma_1(p) = \sigma_2(p)$ for some $p \in M$, and set $\tau := \sigma_1 - \sigma_2$. Then $\tau(p) = 0 \in E_p$. Choose a coordinate neighborhood V of p together with a local frame $(\tilde{\sigma}_1, \dots, \tilde{\sigma}_k)$ of E over V. On V we can write

$$\tau = \sum_{i=1}^{k} u^i \, \tilde{\sigma}_i,$$

where $u^i \in C^\infty(V)$ satisfy $u^i(p) = 0$. Choose a bump function $\chi \in C^\infty(M)$ with $\operatorname{supp} \chi \subset V$ and $\chi \equiv 1$ near p, and set

$$\tilde{\tau} := \sum_{i=1}^{k} \chi \, u^i \, \tilde{\sigma}_i \in \Gamma(E).$$

Then $\tilde{\tau} = \tau$ near p, hence by locality $T(\tilde{\tau})(p) = T(\tau)(p)$. Using $C^\infty(M)$–linearity,

$$T(\tilde{\tau}) = \sum_{i=1}^{k} \chi \, u^i \, T(\tilde{\sigma}_i),$$

and evaluating at p gives $T(\tilde{\tau})(p) = 0$ since $\chi(p) = 1$ and $u^i(p) = 0$. Thus $T(\tau)(p) = 0$, which implies $T(\sigma_1)(p) = T(\sigma_2)(p)$.

Step 3: Definition of f. For $v \in E_p$, choose any section $\sigma \in \Gamma(E)$ with $\sigma(p) = v$, and define

$$f(v) := T(\sigma)(p) \in F_p.$$

By Step 2 this definition is independent of the choice of σ. By construction, f is linear on each fibre and satisfies $\pi' \circ f = \pi$.

Step 4: Smoothness and conclusion. We now show that f is smooth and complete the proof.

Work in local trivializations of E and F over an open set $V \subset M$, so that sections of $E|_V$ and $F|_V$ may be identified with smooth $\mathbb{R}^k$-valued functions on V. By construction we already know that

$$T(\sigma) = f \circ \sigma \qquad \text{for all } \sigma \in \Gamma(E).$$

In these local coordinates, this means that T acts on sections by pointwise multiplication with the matrix representing the fibrewise linear map $f|_{E_p} : E_p \to F_p$. Since T sends smooth sections to smooth sections and is $C^\infty(M)$–linear, the corresponding matrix-valued function depends smoothly on $p \in V$. Hence f is smooth in local trivializations, and therefore smooth globally.

The identity $T(\sigma) = f \circ \sigma$ holds by construction for all $\sigma \in \Gamma(E)$, and uniqueness of f is immediate from this relation.

□

6.2 From vector spaces to vector bundles: duals, sums, tensors, and exterior powers

Let M be a smooth manifold. A (real) smooth vector bundle $\pi : E \to M$ of rank r is locally trivial: for each $p \in M$ there exists an open set $U \ni p$ and a diffeomorphism (a trivialization)

$$\Psi : E|_U \longrightarrow U \times \mathbb{R}^r$$

that restricts on each fiber to a linear isomorphism $E_q \simeq \mathbb{R}^r$.

6.2.1 Local frames, transition functions, and the "fiberwise" principle

A trivialization Ψ determines a *local frame* $e_1, \dots, e_r \in \Gamma(E|_U)$ by declaring $\Psi(e_i(q)) = (q, e_i^{\mathrm{std}})$ for the standard basis (e_i^{std}) of $\mathbb{R}^r$. On overlaps $U \cap V$, two local frames $e = (e_1, \dots, e_r)$ and $\tilde{e} = (\tilde{e}_1, \dots, \tilde{e}_r)$ are related by a smooth map $g : U \cap V \to GL(r, \mathbb{R})$ such that

$$\tilde{e}_j = \sum_{i=1}^{r} g_{ij}\, e_i \qquad \text{on } U \cap V. \tag{6.6}$$

The maps g are the *transition functions* of the bundle.

Many constructions from linear algebra (direct sum, dual, tensor product, exterior power) can be performed fiberwise. The subtle point is to check that the resulting family of fibers assembles into a *smooth* vector bundle. In practice, one proves this by describing the transition functions of the new bundle in terms of those of the original one(s).

6.2.2 Direct sums

Definition 6.20 (Direct sum bundle)**.** Let $E \to M$ and $F \to M$ be vector bundles. Their *direct sum bundle* $E \oplus F \to M$ is defined fiberwise by

$$(E \oplus F)_p := E_p \oplus F_p.$$

Proposition 6.21 (Smooth structure on $E \oplus F$)**.** *The fiberwise direct sum $E \oplus F$ carries a natural smooth vector bundle structure. In local trivializations $E|_U \simeq U \times \mathbb{R}^r$ and $F|_U \simeq U \times \mathbb{R}^s$, one has*

$$(E \oplus F)|_U \simeq U \times (\mathbb{R}^r \oplus \mathbb{R}^s) \simeq U \times \mathbb{R}^{r+s}.$$

Moreover, if $g : U \cap V \to GL(r)$ and $h : U \cap V \to GL(s)$ are transition functions for E and F, then the transition function for $E \oplus F$ is

$$U \cap V \ni q \longmapsto \begin{pmatrix} g(q) & 0 \\ 0 & h(q) \end{pmatrix} \in GL(r+s).$$

Proof. Choose local frames $e_1, \dots, e_r$ for $E|_U$ and $f_1, \dots, f_s$ for $F|_U$. Then $(e_1, \dots, e_r, f_1, \dots, f_s)$ is a local frame for $(E \oplus F)|_U$. On overlaps, (6.6) gives block-diagonal transition matrices as stated, hence the usual cocycle conditions and smoothness follow. □

6.2.3 The dual bundle and $C^\infty(M)$–linear functionals

Definition 6.22 (Dual bundle)**.** Let $E \to M$ be a rank r vector bundle. Its *dual bundle* $E^* \to M$ is defined fiberwise by

$$(E^*)_p := (E_p)^* = \mathrm{Hom}_{\mathbb{R}}(E_p, \mathbb{R}).$$

Proposition 6.23 (Transition functions of E^*)**.** *If $g : U \cap V \to GL(r)$ is a transition function for E in local frames, then the transition function for E^* is $(g^{-1})^\mathsf{T}$.*

Proof. Let $e = (e_i)$ and $\tilde{e} = (\tilde{e}_i)$ be frames related by $\tilde{e} = e\, g$ (i.e. (6.6)). Let $e^* = (e^i)$ and $\tilde{e}^* = (\tilde{e}^i)$ be the dual coframes characterized by $e^i(e_j) = \delta^i_j$ and $\tilde{e}^i(\tilde{e}_j) = \delta^i_j$. Writing $\tilde{e}^i = \sum_j a_{ij} e^j$, the condition $\delta^i_k = \tilde{e}^i(\tilde{e}_k)$ becomes $\delta^i_k = \sum_j a_{ij} g_{jk}$, hence $A = (a_{ij})$ satisfies $Ag = I$, i.e. $A = g^{-1}$, which in column convention is $(g^{-1})^\mathsf{T}$. □

The dual bundle is particularly useful because its sections represent precisely the $C^\infty(M)$–linear functionals on $\Gamma(E)$.

Lemma 6.24. *Let $T : \Gamma(E) \to C^\infty(M)$ be $C^\infty(M)$–linear. If $s \in \Gamma(E)$ satisfies $s(p) = 0$ at some point $p \in M$, then $T(s)(p) = 0$.*

Proof. Choose a local frame $e_1, \dots, e_r$ on a neighborhood $U \ni p$ and write $s|_U = \sum_i f_i e_i$ with $f_i \in C^\infty(U)$. Since $s(p) = 0$, we have $f_i(p) = 0$ for all i. Extend f_i by 0 outside U to functions in C^∞ by multiplying with a cutoff supported in U (still denoted f_i), and extend e_i to global sections by multiplying with a cutoff supported in U (still denoted e_i). Then globally $s = \sum_i f_i e_i$ and by $C^\infty(M)$–linearity,

$$T(s) = \sum_i T(f_i e_i) = \sum_i f_i\, T(e_i).$$

Evaluating at p gives $T(s)(p) = \sum_i f_i(p) T(e_i)(p) = 0$. □

Theorem 6.25 (Dual sections as C^∞–linear functionals)**.** *The map*

$$\Gamma(E^*) \longrightarrow \mathrm{Hom}_{C^\infty(M)}\big(\Gamma(E), C^\infty(M)\big), \qquad \alpha \longmapsto \Big(s \mapsto \big[p \mapsto \alpha(p)(s(p))\big]\Big),$$

is an isomorphism of $C^\infty(M)$–modules.

Proof. Step 1: From $\Gamma(E^*)$ *to* $\mathrm{Hom}_{C^\infty(M)}(\Gamma(E), C^\infty(M))$. Given $\alpha \in \Gamma(E^*)$, define $T_\alpha(s)(p) := \alpha(p)(s(p))$. Smoothness of $T_\alpha(s)$ is checked in local trivializations, and $C^\infty(M)$–linearity is immediate: $T_\alpha(fs)(p) = \alpha(p)(f(p)s(p)) = f(p)\alpha(p)(s(p))$.

Step 2: From T *to a section* $\alpha \in \Gamma(E^*)$. Let $T : \Gamma(E) \to C^\infty(M)$ be $C^\infty(M)$–linear. For each $p \in M$, define a linear functional $\alpha(p) \in (E_p)^*$ by

$$\alpha(p)(v) := T(s)(p), \tag{6.7}$$

where $s \in \Gamma(E)$ is any section with $s(p) = v$.

We must show (6.7) is well-defined. If $s'(p) = v$ as well, then $(s - s')(p) = 0$, hence by Lemma 6.24 we have $T(s - s')(p) = 0$, so $T(s)(p) = T(s')(p)$. Thus $\alpha(p)$ is well-defined and clearly linear in v.

Step 3: Smoothness of α. Choose a local frame $e_1, \ldots, e_r$ on U. Let $e^1, \ldots, e^r$ be the dual coframe of $E^*|_U$. For $q \in U$ we claim

$$\alpha|_U = \sum_{i=1}^{r} a_i\, e^i \qquad \text{with} \qquad a_i := T(e_i)|_U \in C^\infty(U). \tag{6.8}$$

Indeed, for any $v = \sum_i v^i e_i(q) \in E_q$, choose the section $s = \sum_i v^i e_i$ (constant coefficients in the chosen frame). Then $s(q) = v$ and by C^∞–linearity,

$$\alpha(q)(v) = T(s)(q) = \sum_i v^i\, T(e_i)(q) = \sum_i a_i(q)\, v^i,$$

which exactly says (6.8). Since the coefficients a_i are smooth by Lemma 6.24, α is a smooth section of E^* on U. These local expressions agree on overlaps because the construction (6.7) is intrinsic.

Step 4: Inverses. By construction, $T = T_\alpha$. Conversely, starting from α and forming T_α, then reconstructing α from T_α recovers the original fiberwise pairing. Hence the two maps are inverse isomorphisms. □

6.2.4 Tensor products and multilinear maps

Definition 6.26 (Tensor product bundle)**.** For vector bundles $E \to M$ and $F \to M$, define $E \otimes F \to M$ fiberwise by

$$(E \otimes F)_p := E_p \otimes F_p.$$

Proposition 6.27 (Transition functions of $E \otimes F$)**.** *If* $g : U \cap V \to GL(r)$ *and* $h : U \cap V \to GL(s)$ *are transition functions for* E *and* F*, then* $E \otimes F$ *has transition function*

$$U \cap V \ni q \longmapsto g(q) \otimes h(q) \in GL(rs),$$

where $g \otimes h$ *denotes the induced linear map on* $\mathbb{R}^r \otimes \mathbb{R}^s$*.*

Proof. In local frames e_i for E and f_a for F, the tensors $e_i \otimes f_a$ form a frame for $E \otimes F$. On overlaps, replacing e_i and f_a by their transformed frames yields exactly the matrix representing $g \otimes h$. □

Next we identify tensor sections with $C^\infty(M)$–multilinear maps.

Theorem 6.28 (Tensor sections as $C^\infty(M)$–multilinear maps)**.** *Let $E_1, \dots, E_k, F \to M$ be vector bundles. There is a natural isomorphism*

$$\Gamma\big(E_1^* \otimes \cdots \otimes E_k^* \otimes F\big) \;\cong\; \mathrm{Mult}_{C^\infty}\big(\Gamma(E_1) \times \cdots \times \Gamma(E_k), \Gamma(F)\big),$$

where the right-hand side denotes $C^\infty(M)$–multilinear maps.

Proof. Step 1: From a section to a multilinear map. Let $A \in \Gamma(E_1^* \otimes \cdots \otimes E_k^* \otimes F)$. For sections $s_i \in \Gamma(E_i)$ define a section $T_A(s_1, \dots, s_k) \in \Gamma(F)$ by

$$\big(T_A(s_1, \dots, s_k)\big)(p) := A(p)\big(s_1(p), \dots, s_k(p)\big) \in F_p.$$

This is smooth (check in local trivializations) and $C^\infty(M)$–multilinear by fiberwise linearity of $A(p)$.

Step 2: From a multilinear map to a section. Let $T : \Gamma(E_1) \times \cdots \times \Gamma(E_k) \to \Gamma(F)$ be $C^\infty(M)$–multilinear. For $p \in M$ and vectors $v_i \in (E_i)_p$, choose sections $s_i \in \Gamma(E_i)$ with $s_i(p) = v_i$ and define

$$A(p)(v_1, \dots, v_k) := T(s_1, \dots, s_k)(p) \in F_p.$$

We must check this is well-defined. Suppose $s_i'(p) = v_i$ as well. Fix all slots except the i-th and consider the difference

$$T(s_1, \dots, s_i, \dots, s_k) - T(s_1, \dots, s_i', \dots, s_k) = T(s_1, \dots, s_i - s_i', \dots, s_k).$$

The section $s_i - s_i'$ vanishes at p, hence (in a local frame for E_i) it can be written as $\sum_j f_j e_j$ with $f_j(p) = 0$. By $C^\infty(M)$–multilinearity,

$$T(\dots, f_j e_j, \dots) = f_j\, T(\dots, e_j, \dots),$$

and evaluating at p gives $T(\dots, f_j e_j, \dots)(p) = 0$. Summing over j yields $T(\dots, s_i - s_i', \dots)(p) = 0$. Thus $A(p)$ is well-defined and multilinear.

Step 3: Smoothness. Choose local frames $e_a^{(i)}$ for E_i on U and a frame f_b for F on U. Then the values

$$T\big(e_{a_1}^{(1)}, \dots, e_{a_k}^{(k)}\big) \in \Gamma(F|_U)$$

can be expanded as $\sum_b c^b_{a_1 \cdots a_k} f_b$ with smooth coefficients $c^b_{a_1 \cdots a_k} \in C^\infty(U)$. These coefficients are exactly the local components of the tensor field $A|_U$ in the induced frame of $E_1^* \otimes \cdots \otimes E_k^* \otimes F$, hence A is a smooth section. Compatibility on overlaps follows from the intrinsic definition.

Step 4: Inverses. Both constructions are inverse to each other by inspection. □

6.2.5 Exterior powers as subbundles

Definition 6.29 (Alternation on tensor powers)**.** Let V be a vector space. Define the *alternation operator*

$$\mathrm{Alt} : (V^*)^{\otimes k} \to (V^*)^{\otimes k}, \qquad \mathrm{Alt}(\tau) := \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma)\, \tau \circ \sigma.$$

Here S_k denotes the symmetric group on $\{1, \dots, k\}$, and $\mathrm{sgn}(\sigma) \in \{+1, -1\}$ is the *sign* of the permutation σ. It can be defined in either of the following equivalent ways:

- $\operatorname{sgn}(\sigma) = +1$ if σ is an even permutation (a product of an even number of transpositions), and $\operatorname{sgn}(\sigma) = -1$ if σ is odd;
- $\operatorname{sgn}(\sigma) = (-1)^{N(\sigma)}$, where $N(\sigma)$ is the number of *inversions* of σ, i.e.

$$N(\sigma) := \#\{(i,j) \mid 1 \le i < j \le k,\ \sigma(i) > \sigma(j)\}.$$

For $\tau \in (V^*)^{\otimes k}$ and $\sigma \in S_k$, the tensor $\tau \circ \sigma$ is defined by permuting the arguments:

$$(\tau \circ \sigma)(v_1, \dots, v_k) := \tau\big(v_{\sigma(1)}, \dots, v_{\sigma(k)}\big), \qquad v_1, \dots, v_k \in V.$$

A k–linear form $\omega : V^k \to \mathbb{R}$ is called *alternating* if

$$\omega(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = \operatorname{sgn}(\sigma)\, \omega(v_1, \dots, v_k) \quad \text{for all } \sigma \in S_k.$$

Equivalently, ω is alternating if and only if

$$\omega(v_1, \dots, v_i, \dots, v_j, \dots, v_k) = -\omega(v_1, \dots, v_j, \dots, v_i, \dots, v_k)$$

for every transposition $(i\, j)$, or, equivalently again,

$$\omega(v_1, \dots, v_k) = 0 \quad \text{whenever } v_i = v_j \text{ for some } i \neq j.$$

The image of the alternation operator Alt is precisely the subspace of alternating k–linear forms, denoted $\Lambda^k V^*$.

Example 6.30 (Determinant as an alternating n–form and volume)**.** Let $\dim V = n$. Choose a basis $(e_1, \dots, e_n)$ of V and let $(e^1, \dots, e^n)$ be the dual basis of V^*. The tensor

$$\omega_0 := e^1 \wedge \cdots \wedge e^n \in \Lambda^n V^*$$

is characterized by $\omega_0(e_1, \dots, e_n) = 1$. For any $v_1, \dots, v_n \in V$, write $v_j = \sum_i a_{ij} e_i$ and set $A = (a_{ij}) \in \mathbb{R}^{n \times n}$. Then

$$\omega_0(v_1, \dots, v_n) = \det(A).$$

In other words, relative to the chosen basis, the determinant is precisely the alternating n–linear form ω_0 applied to $(v_1, \dots, v_n)$.

Geometrically, $|\det(A)|$ is the Euclidean n–dimensional volume of the parallelotope spanned by $v_1, \dots, v_n$. This is why top-degree alternating forms (and, on manifolds, top-degree differential forms) are the natural objects encoding volume and integration.

Proposition 6.31 (Exterior powers of a bundle)**.** *Let $E \to M$ be a rank r vector bundle. The fiberwise definition*

$$(\Lambda^k E^*)_p := \Lambda^k (E_p)^*$$

assembles into a smooth vector bundle $\Lambda^k E^* \to M$, and it is a smooth vector subbundle of $(E^*)^{\otimes k}$.

Proof. Work in a trivialization $E|_U \simeq U \times \mathbb{R}^r$, hence $(E^*)^{\otimes k}|_U \simeq U \times ((\mathbb{R}^r)^*)^{\otimes k}$. The alternation map Alt is a fixed linear projection on the model fiber $((\mathbb{R}^r)^*)^{\otimes k}$, hence defines a smooth bundle map

$$\mathrm{Alt} : (E^*)^{\otimes k}|_U \longrightarrow (E^*)^{\otimes k}|_U$$

by acting fiberwise. Its image is the trivial subbundle $U \times \Lambda^k(\mathbb{R}^r)^*$, so it is smooth on U.

On an overlap $U \cap V$, the transition function of E^* is $(g^{-1})^\mathsf{T}$ (Proposition 6.23), so the transition on $(E^*)^{\otimes k}$ is $((g^{-1})^\mathsf{T})^{\otimes k}$. This linear map preserves alternating tensors, hence restricts to a transition function on $\Lambda^k(\mathbb{R}^r)^*$. Therefore the local subbundles patch, and $\Lambda^k E^*$ is a well-defined smooth subbundle of $(E^*)^{\otimes k}$. □

Remark 6.32. In particular, sections of $\Lambda^k E^*$ can be viewed equivalently as alternating k–tensor fields, i.e. as sections of $(E^*)^{\otimes k}$ that are alternating in the sense of Definition 6.29.

Conceptually, this identification is entirely parallel to the descriptions given earlier: sections of E^* correspond to $C^\infty(M)$–linear functionals on $\Gamma(E)$, and sections of tensor bundles $E_1^* \otimes \cdots \otimes E_k^* \otimes F$ correspond to $C^\infty(M)$–multilinear maps on sections. The exterior power $\Lambda^k E^*$ simply singles out those multilinear maps that are alternating in the arguments.

For this reason, we will freely identify sections of $\Lambda^k E^*$ with alternating k–multilinear maps on $\Gamma(E)$, without giving a separate proof: the argument is a straightforward adaptation of Theorem 6.28, together with the observation that alternation is a purely fiberwise, algebraic condition.

6.2.6 Naturality of fiberwise constructions

The constructions above are also *functorial*: any bundle isomorphism induces canonical isomorphisms on the bundles obtained from it by direct sums, duals, tensor products, and exterior powers.

Proposition 6.33 (Induced maps on direct sums, duals, tensors, and exterior powers)**.** *Let $E \to M$ and $F \to M$ be smooth vector bundles, and let $\Phi : E \to F$ be a vector bundle isomorphism over Id_M (i.e. Φ is smooth, $\pi_F \circ \Phi = \pi_E$, and each $\Phi_p : E_p \to F_p$ is a linear isomorphism). Then:*

*(a) (*Direct sums*) For any bundle isomorphism $\Psi : E' \to F'$ over Id_M, there is an induced bundle isomorphism*

$$\Phi \oplus \Psi : E \oplus E' \longrightarrow F \oplus F', \qquad (\Phi \oplus \Psi)_p(v, w) := (\Phi_p v, \Psi_p w).$$

*(b) (*Dual bundle, contravariant*) There is an induced bundle isomorphism*

$$\Phi^* : F^* \longrightarrow E^*, \qquad (\Phi^*)_p(\alpha) := \alpha \circ \Phi_p \ \in \ (E_p)^*.$$

*(c) (*Tensor products*) For any bundle isomorphism $\Psi : E' \to F'$ over Id_M, there is an induced bundle isomorphism*

$$\Phi \otimes \Psi : E \otimes E' \longrightarrow F \otimes F', \qquad (\Phi \otimes \Psi)_p(v \otimes w) := \Phi_p v \otimes \Psi_p w.$$

*(d) (*Exterior powers*) For each $k \geq 0$, there is an induced bundle isomorphism*

$$\Lambda^k\Phi^* : \Lambda^k F^* \longrightarrow \Lambda^k E^*, \qquad (\Lambda^k\Phi^*)_p(\omega) := \omega \circ (\Phi_p, \ldots, \Phi_p).$$

Equivalently, viewing $\Lambda^k F^*$ as a subbundle of $(F^*)^{\otimes k}$, $\Lambda^k\Phi^*$ is the restriction of $(\Phi^*)^{\otimes k}$ to alternating tensors.

Moreover, all these assignments are compatible with composition and identities: for bundle isomorphisms $\Phi : E \to F$ and $\Psi : F \to G$ over Id_M,

$$(\Psi \circ \Phi) \oplus (\Psi' \circ \Phi') = (\Psi \oplus \Psi') \circ (\Phi \oplus \Phi'),$$
$$(\Psi \circ \Phi)^* = \Phi^* \circ \Psi^*,$$
$$(\Psi \circ \Phi) \otimes (\Psi' \circ \Phi') = (\Psi \otimes \Psi') \circ (\Phi \otimes \Phi').$$

and similarly for $\Lambda^k(\,\cdot\,)^*$, and each construction sends Id *to* Id*.*

Proof. All maps are defined fiberwise, and by construction they cover Id_M and are linear isomorphisms on each fiber.

Smoothness is checked in local trivializations. For instance, in a local frame for $E|_U$ and $F|_U$, the map $\Phi|_U$ is represented by a smooth $GL(r)$–valued matrix $A : U \to GL(r)$. Then $\Phi \oplus \Psi$ is represented by a block matrix, Φ^* by $(A^{-1})^{\mathsf{T}}$, $\Phi \otimes \Psi$ by $A \otimes B$, and $\Lambda^k\Phi^*$ by the induced action on alternating tensors. Thus each induced map is smooth with smooth inverse.

Compatibility with composition and identities follows immediately from the fiberwise formulas. □

Remark 6.34. In categorical language: over a fixed base M, the assignments $E \mapsto E \oplus E'$, $E \mapsto E \otimes E'$, and $E \mapsto \Lambda^k E$ are covariant functors on the category of vector bundles over M, while $E \mapsto E^*$ is a contravariant functor.

6.2.7 Why these constructions matter

Isomorphism classes of vector bundles over M form a commutative monoid under direct sum. Its group completion leads to the *K–group* $K^0(M)$, an important invariant and a prototype of a generalized cohomology theory. We will not pursue this further here. In this course we focus mainly on the tangent bundle TM and its dual T^*M, together with the bundles $\Lambda^k T^*M$ of differential k–forms.

6.3 Tensor bundles and natural operations

Let M be a smooth manifold. We write $\mathfrak{X}(M)$ for the space of smooth vector fields and $C^\infty(M)$ for the space of smooth real-valued functions.

6.3.1 Tensor bundles

Recall that $TM \to M$ is the tangent bundle and $T^*M \to M$ is the cotangent bundle. For each $p \in M$, the fibers are the vector space T_pM and its dual T_p^*M. As a notational convention, if $(x^1, \ldots, x^n)$ is a local chart and

$$\left\{ \frac{\partial}{\partial x^i} \right\}_{i=1}^n$$

denotes the associated coordinate basis of T_pM, then we write $\{dx^i\}_{i=1}^n$ for the dual basis of T_p^*M, characterized by

$$dx^i\left(\frac{\partial}{\partial x^j}\right) = \delta^i_j.$$

Definition 6.35 (Tensor bundles)**.** For integers $r, s \geq 0$, the *tensor bundle of type* (r,s) is

$$T^r_sM := (TM)^{\otimes r} \otimes (T^*M)^{\otimes s}.$$

A *tensor field of type* (r,s) is a smooth section $T \in \Gamma(T^r_sM)$.

Concretely, a section $T \in \Gamma(T^r_sM)$ assigns to each $p \in M$ a multilinear map

$$T_p : \underbrace{T_p^*M \times \cdots \times T_p^*M}_{r \text{ times}} \times \underbrace{T_pM \times \cdots \times T_pM}_{s \text{ times}} \longrightarrow \mathbb{R},$$

or, equivalently (by currying and the natural identifications), a multilinear map with r vector inputs and s covector inputs. We will switch freely between these equivalent viewpoints; the bookkeeping is encoded by the type (r,s).

Basic examples.

- $T^0_0M = M \times \mathbb{R}$ and $\Gamma(T^0_0M) = C^\infty(M)$.
- $T^1_0M = TM$ and $\Gamma(T^1_0M) = \mathfrak{X}(M)$.
- $T^0_1M = T^*M$ and $\Gamma(T^0_1M) = \Omega^1(M)$ (smooth 1–forms).
- A *Riemannian metric* is a smooth section $g \in \Gamma(T^0_2M)$ such that each g_p is a positive definite symmetric bilinear form on T_pM.
- A $(1,1)$–*tensor field* can be viewed pointwise as a linear map $T_pM \to T_pM$; for instance, the differential of a smooth map or an almost complex structure J (when $J^2 = -\mathrm{Id}$).

Tensor products and contractions. If $S \in \Gamma(T^{r_1}_{s_1}M)$ and $T \in \Gamma(T^{r_2}_{s_2}M)$, their tensor product is the section

$$S \otimes T \in \Gamma\big(T^{r_1+r_2}_{s_1+s_2}M\big),$$

defined fiberwise by the usual tensor product of multilinear maps.

A *contraction* is obtained by pairing one TM–factor with one T^*M–factor using the canonical pairing $\langle \alpha, v\rangle = \alpha(v)$. For example, for $A \in \Gamma(T^1_1M)$ (an endomorphism field), the function $\mathrm{tr}(A)$ is a contraction:

$$\mathrm{tr}(A)(p) := \sum_{i=1}^n e^i\big(A(e_i)\big),$$

where $\{e_i\}$ is any basis of T_pM and $\{e^i\}$ its dual basis. The right-hand side is independent of the chosen basis because it is defined by the canonical pairing.

6.3.2 Pullback under diffeomorphisms

Let $\Phi : M \to M$ be a diffeomorphism. It induces the tangent map

$$\Phi_* : TM \to TM, \qquad \Phi_* : T_pM \to T_{\Phi(p)}M,$$

and the pullback on covectors

$$\Phi^* : T^*M \to T^*M, \qquad \Phi^* : T^*_{\Phi(p)}M \to T^*_pM, \quad (\Phi^*\beta)_p(v) := \beta_{\Phi(p)}(\Phi_*v).$$

By tensoring these maps fiberwise, we obtain a natural pullback on (r, s)–tensor bundles, still denoted by Φ^*:

$$\Phi^* : \Gamma(T^r_sM) \longrightarrow \Gamma(T^r_sM).$$

Concretely, if T is a (r, s)–tensor field and $v_i \in T_pM$, $\alpha_j \in T^*_pM$, then

$$(\Phi^*T)_p(v_1, \ldots, v_r, \alpha_1, \ldots, \alpha_s) := T_{\Phi(p)}\big(\Phi_*v_1, \ldots, \Phi_*v_r,\ (\Phi^{-1})^*\alpha_1, \ldots, (\Phi^{-1})^*\alpha_s\big). \tag{6.9}$$

Here the appearance of $(\Phi^{-1})^*$ on covectors ensures that the right-hand side depends only on data at p and uses the canonical pairing consistently.

Remark 6.36 (Compatibility with functions and vector fields)**.** For $f \in C^\infty(M)$, the pullback is $(\Phi^*f)(p) = f(\Phi(p))$. For a 1–form $\omega \in \Omega^1(M)$, $\Phi^*\omega$ is the usual pullback: $(\Phi^*\omega)_p(v) = \omega_{\Phi(p)}(\Phi_*v)$. For a vector field $Y \in \mathfrak{X}(M)$, the *pushforward* Φ_*Y is the vector field defined by $(\Phi_*Y)_{\Phi(p)} := \Phi_*Y_p$.

Two functorial properties will be used repeatedly.

Proposition 6.37 (Naturality of pullback)**.** *For any diffeomorphism Φ and any tensor fields S, T,*

(i) $\Phi^*(S \otimes T) = \Phi^*S \otimes \Phi^*T$.

(ii) If C is any fixed index contraction (pairing one TM–slot with one T^*M–slot), then $\Phi^*(C(T)) = C(\Phi^*T)$.

Proof. Both statements are fiberwise and reduce to linear algebra. By construction, Φ^* on T^r_sM is obtained by tensoring the linear maps $\Phi_* : T_pM \to T_{\Phi(p)}M$ and $(\Phi^{-1})^* : T^*_pM \to T^*_{\Phi(p)}M$ in the appropriate slots. Tensor products commute with composition, which gives (i). For (ii), note that contraction is defined using the canonical pairing $T^*_{\Phi(p)}M \times T_{\Phi(p)}M \to \mathbb{R}$, and the maps Φ_* and Φ^* preserve this pairing:

$$\big((\Phi^{-1})^*\alpha\big)(\Phi_*v) = \alpha(v) \qquad (\alpha \in T^*_pM,\ v \in T_pM).$$

Hence performing a contraction before or after pullback yields the same result. □

Remark 6.38 (Functoriality for compositions)**.** If $\Psi : M \to M$ is another diffeomorphism, then

$$(\Psi \circ \Phi)^* = \Phi^* \circ \Psi^*, \qquad (\mathrm{Id})^* = \mathrm{Id}.$$

This is immediate from the definitions and expresses that pullback is a contravariant functor.

6.4 Lie derivatives of tensor fields

Let $X \in \mathfrak{X}(M)$ be a smooth vector field, and let Φ_t denote its (local) flow, i.e. $\Phi_0 = \mathrm{Id}$ and

$$\frac{d}{dt}\Phi_t(p) = X_{\Phi_t(p)}$$

whenever the flow is defined. Pullback by Φ_t transports tensor fields along the flow, and differentiating at $t = 0$ produces the Lie derivative.

Definition 6.39 (Lie derivative on tensors)**.** For a tensor field $T \in \Gamma(T^r_s M)$, define

$$\mathcal{L}_X T := \left.\frac{d}{dt}\right|_{t=0} \Phi_t^* T.$$

Remark 6.40 (Lie derivative on functions)**.** If $f \in C^\infty(M)$, then $\Phi_t^* f = f \circ \Phi_t$ and

$$\mathcal{L}_X f = \left.\frac{d}{dt}\right|_{t=0} (f \circ \Phi_t) = df(X) = X(f),$$

the directional derivative of f along X.

6.4.1 Leibniz rules and contractions

The Lie derivative is a derivation with respect to all natural tensor operations built from tensor products and contractions.

Proposition 6.41 (Leibniz rule for tensor products)**.** *For any tensor fields S, T,*

$$\mathcal{L}_X(S \otimes T) = (\mathcal{L}_X S) \otimes T + S \otimes (\mathcal{L}_X T).$$

Proof. By Proposition 6.37(i),

$$\Phi_t^*(S \otimes T) = \Phi_t^* S \otimes \Phi_t^* T.$$

Differentiate at $t = 0$ and use the product rule for the tensor product in each fiber. □

Proposition 6.42 (Lie derivative commutes with contraction)**.** *For any fixed contraction map C,*

$$\mathcal{L}_X(C(T)) = C(\mathcal{L}_X T).$$

Proof. By Proposition 6.37(ii),

$$\Phi_t^*(C(T)) = C(\Phi_t^* T).$$

Differentiate at $t = 0$ and use that C is fiberwise linear, hence commutes with differentiation. □

Corollary 6.43 (Lie derivative respects all "natural" tensor expressions)**.** *Any tensor field obtained from given tensor fields by repeated tensor products and contractions satisfies the expected Leibniz rule under $\mathcal{L}_X$.*

Proof. Combine Proposition 6.41 and Proposition 6.42 and argue by induction on the number of operations used to build the expression. □

6.5 Differential forms: wedge and contraction

6.5.1 Forms as alternating covariant tensors

A *k–form* is a smooth section of the bundle $\Lambda^k T^*M$. We write $\Omega^k(M) := \Gamma(\Lambda^k T^*M)$ and $\Omega^0(M) = C^\infty(M)$.

6.5.2 Wedge product

The wedge product $\wedge : \Omega^p(M) \times \Omega^q(M) \to \Omega^{p+q}(M)$ is the alternation of the tensor product. In particular, alternation commutes with pullback, hence with Lie derivatives.

Proposition 6.44 (Lie derivative and wedge)**.** *For $\alpha \in \Omega^p(M)$ and $\beta \in \Omega^q(M)$,*

$$\mathcal{L}_X(\alpha \wedge \beta) = (\mathcal{L}_X \alpha) \wedge \beta + \alpha \wedge (\mathcal{L}_X \beta).$$

6.5.3 Interior product

Definition 6.45 (Interior product)**.** Let $X \in \mathfrak{X}(M)$. The *interior product* (contraction) is the map

$$\iota_X : \Omega^k(M) \to \Omega^{k-1}(M)$$

defined by

$$(\iota_X \omega)(Y_1, \ldots, Y_{k-1}) := \omega(X, Y_1, \ldots, Y_{k-1}). \tag{6.10}$$

Proposition 6.46 (Graded Leibniz rule for ι)**.** *If $\alpha \in \Omega^p(M)$ and $\beta \in \Omega^q(M)$, then*

$$\iota_X(\alpha \wedge \beta) = (\iota_X \alpha) \wedge \beta + (-1)^p \, \alpha \wedge (\iota_X \beta).$$

Proof. This is the standard alternating-multilinear check from the definition of $\wedge$ and (6.10). □

6.5.4 A useful commutator identity

Lemma 6.47 (Commutator $[\mathcal{L}, \iota]$)**.** *For all $X, Y \in \mathfrak{X}(M)$ and all $\omega \in \Omega^k(M)$,*

$$\mathcal{L}_X(\iota_Y \omega) = \iota_Y(\mathcal{L}_X \omega) + \iota_{[X,Y]}\omega. \tag{6.11}$$

Proof. Fix $Z_1, \ldots, Z_{k-1} \in \mathfrak{X}(M)$. We evaluate both sides of (6.11) on $(Z_1, \ldots, Z_{k-1})$.

Step 1: Expand the left-hand side using the Lie derivative formula on forms. Recall the standard formula: for any $\eta \in \Omega^m(M)$ and vector fields $V_1, \ldots, V_m$,

$$(\mathcal{L}_X \eta)(V_1, \ldots, V_m) = X\big(\eta(V_1, \ldots, V_m)\big) - \sum_{i=1}^{m} \eta(V_1, \ldots, [X, V_i], \ldots, V_m). \tag{6.12}$$

Apply this with $\eta = \iota_Y \omega \in \Omega^{k-1}(M)$ and $V_i = Z_i$. Using

$$(\iota_Y \omega)(Z_1, \ldots, Z_{k-1}) = \omega(Y, Z_1, \ldots, Z_{k-1}),$$

we obtain

$$\big(\mathcal{L}_X(\iota_Y\omega)\big)(Z_1,\dots,Z_{k-1}) = X\Big(\omega(Y,Z_1,\dots,Z_{k-1})\Big) - \sum_{j=1}^{k-1}\omega\big(Y,Z_1,\dots,[X,Z_j],\dots,Z_{k-1}\big). \tag{6.13}$$

Step 2: Expand the first term on the right-hand side. Again by (6.12),

$$\begin{aligned}\big(\iota_Y(\mathcal{L}_X\omega)\big)(Z_1,\dots,Z_{k-1}) &= (\mathcal{L}_X\omega)(Y,Z_1,\dots,Z_{k-1})\\ &= X\Big(\omega(Y,Z_1,\dots,Z_{k-1})\Big) - \omega\big([X,Y],Z_1,\dots,Z_{k-1}\big)\\ &\quad - \sum_{j=1}^{k-1}\omega\big(Y,Z_1,\dots,[X,Z_j],\dots,Z_{k-1}\big).\end{aligned} \tag{6.14}$$

Step 3: Compare the expansions. Subtracting (6.14) from (6.13), all terms cancel except the $[X,Y]$–term, and we get

$$\big(\mathcal{L}_X(\iota_Y\omega)\big)(Z_1,\dots,Z_{k-1}) - \big(\iota_Y(\mathcal{L}_X\omega)\big)(Z_1,\dots,Z_{k-1}) = \omega\big([X,Y],Z_1,\dots,Z_{k-1}\big).$$

The right-hand side is exactly

$$(\iota_{[X,Y]}\omega)(Z_1,\dots,Z_{k-1}) = \omega\big([X,Y],Z_1,\dots,Z_{k-1}\big).$$

Therefore,

$$\big(\mathcal{L}_X(\iota_Y\omega)\big)(Z_1,\dots,Z_{k-1}) = \big(\iota_Y(\mathcal{L}_X\omega)\big)(Z_1,\dots,Z_{k-1}) + \big(\iota_{[X,Y]}\omega\big)(Z_1,\dots,Z_{k-1}).$$

Since $Z_1,\dots,Z_{k-1}$ were arbitrary, this proves (6.11). □

6.6 The exterior derivative and Koszul's formula

6.6.1 A coordinate definition

We begin with a concrete definition of the exterior derivative in local coordinates. We will then prove that the resulting form is independent of the chosen chart (in particular it is compatible on overlaps), and hence defines a global operator. This local-to-global viewpoint naturally suggests that there should be an intrinsic coordinate-free formula, which is provided by Koszul's formula in the next subsection.

Definition 6.48 (Exterior derivative in a chart)**.** Let $\omega \in \Omega^k(M)$. Given a coordinate chart $(U;x^1,\dots,x^n)$, write

$$\omega|_U = \sum_{1\le i_1<\cdots<i_k\le n}\omega_{i_1\cdots i_k}\,dx^{i_1}\wedge\cdots\wedge dx^{i_k}, \qquad \omega_{i_1\cdots i_k}\in C^\infty(U).$$

We define $d\omega$ on U by

$$(d\omega)|_U := \sum_{1\le i_1<\cdots<i_k\le n} d\omega_{i_1\cdots i_k}\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_k},$$

where for $f\in C^\infty(U)$ we set $df := \sum_{j=1}^n \partial_j f\,dx^j$.

Lemma 6.49 (Graded Leibniz rule)**.** *Let* $\alpha \in \Omega^p(M)$ *and* $\beta \in \Omega^q(M)$*. Then*

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \, \alpha \wedge d\beta.$$

Proof. Fix a chart $(U; x^1, \ldots, x^n)$. It suffices to prove the identity on U. First assume $\alpha = f \, dx^{i_1} \wedge \cdots \wedge dx^{i_p}$ and $\beta = g \, dx^{j_1} \wedge \cdots \wedge dx^{j_q}$ with $f, g \in C^\infty(U)$. Then

$$\alpha \wedge \beta = fg \, dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q}.$$

By Definition 6.48,

$$d(\alpha \wedge \beta) = d(fg) \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q}.$$

Using $d(fg) = f \, dg + g \, df$, we get

$$d(\alpha \wedge \beta) = f \, dg \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q} + g \, df \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q}.$$

On the other hand,

$$d\alpha \wedge \beta = (df \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p}) \wedge (g \, dx^{j_1} \wedge \cdots \wedge dx^{j_q}) = g \, df \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q},$$

and

$$\begin{aligned} \alpha \wedge d\beta &= (f \, dx^{i_1} \wedge \cdots \wedge dx^{i_p}) \wedge (dg \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q}) \\ &= f \, dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dg \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q} \\ &= (-1)^p \, f \, dg \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q}, \end{aligned}$$

where the last equality uses that moving the 1–form dg past p one-forms $dx^{i_1}, \ldots, dx^{i_p}$ produces the sign $(-1)^p$. Combining the last three displays yields

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \, \alpha \wedge d\beta \qquad \text{on } U$$

for such decomposable α, β.

For general $\alpha \in \Omega^p(U)$ and $\beta \in \Omega^q(U)$, write

$$\alpha = \sum_I \alpha_I \, dx^I, \qquad \beta = \sum_J \beta_J \, dx^J,$$

and use bilinearity of $\wedge$ together with the definition of d in the chart, which is $\mathbb{R}$–linear and satisfies $d(\alpha_I \, dx^I) = d\alpha_I \wedge dx^I$. Applying the already proven decomposable case term-by-term gives the desired identity on U, hence on M. □

Lemma 6.50 (The local definition is well-defined)**.** *Let* $(U; x^1, \ldots, x^n)$ *and* $(U; y^1, \ldots, y^n)$ *be two coordinate charts on the same open set* U*. Let* d_x *(resp.* d_y*) denote the operator defined in Definition 6.48 when computed using the* x*–chart (resp. the* y*–chart). Then for every* $\omega \in \Omega^k(U)$ *we have*

$$d_x \omega = d_y \omega \qquad \text{on } U.$$

Equivalently, the form $(d\omega)|_U$ *defined by Definition 6.48 does not depend on the chosen coordinate chart.*

Proof. We compute everything using d_x. The key identity is

$$d_x(dy^a) = 0 \qquad \text{for each } a = 1, \dots, n, \tag{6.15}$$

where we view $y^a = y^a(x)$ as a smooth function of x on U.

Indeed, by the chain rule,

$$dy^a = \sum_{i=1}^{n} \frac{\partial y^a}{\partial x^i}(x)\, dx^i.$$

Applying d_x and using $d_x(dx^i) = 0$ gives

$$d_x(dy^a) = \sum_{i=1}^{n} d_x\left(\frac{\partial y^a}{\partial x^i}(x)\right) \wedge dx^i = \sum_{i,j=1}^{n} \frac{\partial^2 y^a}{\partial x^j \partial x^i}(x)\, dx^j \wedge dx^i = 0,$$

since mixed partials commute and $dx^j \wedge dx^i = -dx^i \wedge dx^j$. By the graded Leibniz rule, (6.15) implies

$$d_x\big(dy^{i_1} \wedge \cdots \wedge dy^{i_k}\big) = 0 \qquad \text{for all } 1 \le i_1 < \cdots < i_k \le n. \tag{6.16}$$

Now expand ω in the y–coordinates:

$$\omega = \sum_I \omega_I^{(y)}(y)\, dy^I, \qquad dy^I := dy^{i_1} \wedge \cdots \wedge dy^{i_k}.$$

When applying d_x, the coefficients must be viewed as functions of x via the composition $y = y(x)$, i.e. $\omega_I^{(y)}(y(x))$. Using the graded Leibniz rule and (6.16), we obtain

$$d_x\omega = \sum_I d_x\big(\omega_I^{(y)}(y(x))\big) \wedge dy^I.$$

Next, by the chain rule,

$$d_x\big(\omega_I^{(y)}(y(x))\big) = \sum_{j=1}^{n} \frac{\partial}{\partial x^j}\big(\omega_I^{(y)}(y(x))\big)\, dx^j = \sum_{j=1}^{n}\sum_{a=1}^{n} \frac{\partial \omega_I^{(y)}}{\partial y^a}(y(x))\, \frac{\partial y^a}{\partial x^j}(x)\, dx^j.$$

Recalling $dy^a = \sum_{j=1}^{n} \frac{\partial y^a}{\partial x^j}(x)\, dx^j$, this rewrites as

$$d_x\big(\omega_I^{(y)}(y(x))\big) = \sum_{a=1}^{n} \frac{\partial \omega_I^{(y)}}{\partial y^a}(y(x))\, dy^a.$$

Therefore

$$d_x\omega = \sum_I \left(\sum_{a=1}^{n} \frac{\partial \omega_I^{(y)}}{\partial y^a}(y)\, dy^a\right) \wedge dy^I = \sum_I d_y\big(\omega_I^{(y)}(y)\big) \wedge dy^I = d_y\omega,$$

which proves $d_x\omega = d_y\omega$ on U. □

6.6.2 An intrinsic formula

We have defined the exterior derivative $d\omega$ in local coordinates and proved that the resulting $(k+1)$–form is independent of the choice of chart. In particular, d is a globally well-defined operator

$$d : \Omega^k(M) \longrightarrow \Omega^{k+1}(M).$$

Once we know that d is intrinsic, it is natural to ask for a coordinate-free formula expressing $d\omega$ directly in terms of vector fields and Lie brackets. Such a formula exists; it is usually attributed to Koszul. We present it for 1–forms, where the computation is most transparent.

Koszul's formula for 1–forms.

Proposition 6.51 (Koszul's formula for 1–forms)**.** *Let $\alpha \in \Omega^1(M)$. Define a map $\beta : \mathfrak{X}(M) \times \mathfrak{X}(M) \to C^\infty(M)$ by*

$$\beta(X, Y) := X(\alpha(Y)) - Y(\alpha(X)) - \alpha([X, Y]). \tag{6.17}$$

Then β is $C^\infty(M)$–bilinear and alternating, hence determines a unique 2–form on M, denoted by $d\alpha$. Moreover, this $d\alpha$ agrees with the exterior derivative defined in local coordinates.

Proof. Alternation is immediate: swapping X and Y changes the sign since $[Y, X] = -[X, Y]$.

$C^\infty(M)$–bilinearity.

Let $f \in C^\infty(M)$. Using that α is $C^\infty(M)$–linear and the product rule, we compute

$$\begin{aligned}
\beta(fX, Y) &= (fX)(\alpha(Y)) - Y(\alpha(fX)) - \alpha([fX, Y]) \\
&= f\,X(\alpha(Y)) - Y\big(f\,\alpha(X)\big) - \alpha\big(f[X, Y] - (Yf)\,X\big) \\
&= f\,X(\alpha(Y)) - (Yf)\,\alpha(X) - f\,Y(\alpha(X)) - f\,\alpha([X, Y]) + (Yf)\,\alpha(X) \\
&= f\big(X(\alpha(Y)) - Y(\alpha(X)) - \alpha([X, Y])\big) \\
&= f\,\beta(X, Y),
\end{aligned}$$

where we used the Leibniz rule for Lie brackets

$$[fX, Y] = f[X, Y] - (Yf)\,X. \tag{6.18}$$

Agreement with the coordinate definition. Let $(U; x^1, \ldots, x^n)$ be a chart and write $\alpha = \sum_{i=1}^n a_i\,dx^i$ with $a_i \in C^\infty(U)$. The coordinate definition gives

$$(d\alpha)_{\text{coord}} := \sum_{i=1}^n da_i \wedge dx^i \in \Omega^2(U).$$

On the other hand, let $\beta \in \Omega^2(U)$ be the 2–form defined by the intrinsic formula (e.g. via $\beta(X, Y) = X(\alpha(Y)) - Y(\alpha(X)) - \alpha([X, Y])$).

Both $(d\alpha)_{\text{coord}}$ and β are 2–forms, hence are $C^\infty(U)$–bilinear in (X, Y). Therefore it suffices to check equality pointwise: for each $p \in U$ we only need to show

$$(d\alpha)_{\text{coord},p}(v, w) = \beta_p(v, w) \qquad \text{for all } v, w \in T_pU.$$

Fix $p \in U$ and $v, w \in T_pU$. Choose smooth vector fields $X, Y \in \mathfrak{X}(U)$ such that $X_p = v$, $Y_p = w$, and such that in the given coordinates their components are constant in a neighborhood of p. For instance, if $v = \sum_j v^j \partial_j|_p$ and $w = \sum_j w^j \partial_j|_p$, one may take

$$X := \sum_{j=1}^{n} v^j \, \partial_j, \qquad Y := \sum_{j=1}^{n} w^j \, \partial_j,$$

which satisfy $X_p = v$, $Y_p = w$ and $[X, Y] = 0$ everywhere.

Now compute at the point p. First, using $\alpha(Y) = \sum_i a_i \, dx^i(Y) = \sum_i a_i \, Y^i$ and that the components Y^i are constant, we get

$$X(\alpha(Y))_p = \sum_{i=1}^{n} X(a_i)_p \, Y^i = \sum_{i=1}^{n} X(a_i)_p \, dx^i(Y_p),$$

and similarly

$$Y(\alpha(X))_p = \sum_{i=1}^{n} Y(a_i)_p \, dx^i(X_p).$$

Since $[X, Y]_p = 0$, the intrinsic definition yields

$$\beta_p(v, w) = \beta_p(X_p, Y_p) = X(\alpha(Y))_p - Y(\alpha(X))_p.$$

On the other hand, evaluating the coordinate form gives

$$\begin{aligned}
(d\alpha)_{\text{coord},p}(v, w) &= \sum_{i=1}^{n} (da_i \wedge dx^i)_p(X_p, Y_p) \\
&= \sum_{i=1}^{n} \big(da_i(X)_p \, dx^i(Y)_p - da_i(Y)_p \, dx^i(X)_p\big) \\
&= \sum_{i=1}^{n} \big(X(a_i)_p \, dx^i(Y_p) - Y(a_i)_p \, dx^i(X_p)\big).
\end{aligned}$$

Comparing the last expression with the formula for $\beta_p(v, w)$ shows

$$(d\alpha)_{\text{coord},p}(v, w) = \beta_p(v, w).$$

Since p, v, w were arbitrary, $(d\alpha)_{\text{coord}} = \beta$ on U, and hence the two definitions agree globally. □

6.6.3 Pullbacks of covariant tensors and differential forms

We previously discussed pullbacks (and pushforwards) for diffeomorphisms. In fact, for covariant tensors—in particular for $(0, k)$–tensors and differential forms—the pullback construction works for *any* smooth map, not only for diffeomorphisms. By contrast, for vector fields (and more generally contravariant tensors), a smooth map $f : M \to N$ does *not* in general induce a well-defined pushforward to N: even though df_p sends T_pM to $T_{f(p)}N$, there is typically no canonical way to assign a single vector at a point $q \in N$ from the values over the whole fiber

$f^{-1}(q)$. This is why $(0,k)$–tensors have better functoriality with respect to smooth maps than vector fields do.

Pullback of $(0,k)$*–tensors.*

Let $f : M \to N$ be a smooth map. Recall that for each $p \in M$ we have a linear map on tangent spaces

$$df_p : T_pM \longrightarrow T_{f(p)}N.$$

Definition 6.52 (Pullback of covariant tensors)**.** Let T be a smooth $(0,k)$–tensor field on N, i.e. $T \in \Gamma((T^*N)^{\otimes k})$. We define its *pullback* $f^*T \in \Gamma((T^*M)^{\otimes k})$ by

$$(f^*T)_p(v_1, \dots, v_k) := T_{f(p)}\big(df_p(v_1), \dots, df_p(v_k)\big), \qquad p \in M,\ v_i \in T_pM. \tag{6.19}$$

Pullback of differential forms.

Since $\Omega^k(N)$ is a subbundle of $(T^*N)^{\otimes k}$, Definition 6.52 restricts to k–forms.

Definition 6.53 (Pullback of forms)**.** Let $\omega \in \Omega^k(N)$. We define $f^*\omega \in \Omega^k(M)$ by

$$(f^*\omega)_p(v_1, \dots, v_k) := \omega_{f(p)}\big(df_p(v_1), \dots, df_p(v_k)\big).$$

Lemma 6.54 (Basic algebraic properties)**.** *For* $\eta \in \Omega^k(N)$ *and* $\theta \in \Omega^\ell(N)$ *we have*

$$f^*(\eta \wedge \theta) = f^*\eta \wedge f^*\theta.$$

Moreover, for $g \in C^\infty(N)$ *we have* $f^*g = g \circ f$*, and for composable smooth maps* $M \xrightarrow{f} N \xrightarrow{g} P$*,*

$$(g \circ f)^* = f^* \circ g^*.$$

Proof. All statements follow immediately from the definition and multilinearity. For instance, for $v_1, \dots, v_{k+\ell} \in T_pM$ we compute

$$(f^*(\eta \wedge \theta))_p(v_1, \dots, v_{k+\ell}) = (\eta \wedge \theta)_{f(p)}(df_pv_1, \dots, df_pv_{k+\ell}),$$

which equals $(f^*\eta \wedge f^*\theta)_p(v_1, \dots, v_{k+\ell})$ by the standard definition of the wedge product. □

Compatibility with the exterior derivative.

The pullback is compatible with d.

Proposition 6.55 (Naturality of d)**.** *For every smooth map* $f : M \to N$ *and every* $\omega \in \Omega^k(N)$*,*

$$d(f^*\omega) = f^*(d\omega). \tag{6.20}$$

Proof. It suffices to check the identity locally. Fix $p \in M$ and choose coordinate charts $(U; x^1, \dots, x^m)$ around p and $(V; y^1, \dots, y^n)$ around $f(p)$ with $f(U) \subset V$. Write $f = (f^1, \dots, f^n)$ in these charts, i.e. $f^a = y^a \circ f$ is a smooth function on U.

Step 1: a coordinate formula for pullback. Write on V

$$\omega = \sum_I \omega_I(y)\, dy^I, \qquad dy^I := dy^{i_1} \wedge \cdots \wedge dy^{i_k}.$$

Then on U,

$$f^*\omega = \sum_I (\omega_I \circ f)\, d(f^{i_1}) \wedge \cdots \wedge d(f^{i_k}), \tag{6.21}$$

because $f^*(dy^a) = d(y^a \circ f) = d(f^a)$ and f^* is multiplicative with respect to $\wedge$ (Lemma 6.54).

Step 2: compute $d(f^*\omega)$. Using the local definition of d and the graded Leibniz rule,

$$d(f^*\omega) = \sum_I d(\omega_I \circ f) \wedge d(f^{i_1}) \wedge \cdots \wedge d(f^{i_k}) + \sum_I (\omega_I \circ f)\, d\big(d(f^{i_1}) \wedge \cdots \wedge d(f^{i_k})\big).$$

The second sum vanishes because $d(d(f^a)) = 0$ for each a (exactly as in Lemma 6.50), hence

$$d\big(d(f^{i_1}) \wedge \cdots \wedge d(f^{i_k})\big) = 0. \tag{6.22}$$

Therefore

$$d(f^*\omega) = \sum_I d(\omega_I \circ f) \wedge d(f^{i_1}) \wedge \cdots \wedge d(f^{i_k}). \tag{6.23}$$

Step 3: compute $f^*(d\omega)$. By the coordinate definition,

$$d\omega = \sum_I d\omega_I(y) \wedge dy^I.$$

Pulling back and using $f^*(dy^a) = d(f^a)$ gives

$$f^*(d\omega) = \sum_I f^*(d\omega_I) \wedge d(f^{i_1}) \wedge \cdots \wedge d(f^{i_k}). \tag{6.24}$$

It remains to note that for a function $\varphi \in C^\infty(V)$,

$$f^*(d\varphi) = d(\varphi \circ f), \tag{6.25}$$

which follows by writing $d\varphi = \sum_a (\partial_{y^a}\varphi)\, dy^a$ and pulling back:

$$f^*(d\varphi) = \sum_a (\partial_{y^a}\varphi) \circ f\, f^*(dy^a) = \sum_a (\partial_{y^a}\varphi) \circ f\, d(f^a) = d(\varphi \circ f).$$

Applying (6.25) to $\varphi = \omega_I$ shows that $f^*(d\omega_I) = d(\omega_I \circ f)$, so (6.24) becomes exactly (6.23). This proves (6.20) on U, hence on all of M. □

6.7 Integration of differential forms and orientability

In order to define integration of forms on a manifold, we first treat the Euclidean case. The outcome will show that the integral of a top–degree form is invariant precisely under *orientation–preserving* diffeomorphisms. This naturally leads to the notion of an oriented (i.e. orientable together with a choice of orientation) manifold, and then to the definition of the integral on M by a partition of unity.

6.7.1 Integration of n–forms in $\mathbb{R}^n$

Let $U \subset \mathbb{R}^n$ be open with standard coordinates $(x^1, \dots, x^n)$. Any $\omega \in \Omega^n(U)$ can be written uniquely as

$$\omega = f\, dx^1 \wedge \cdots \wedge dx^n, \qquad f \in C^\infty(U).$$

Lemma 6.56 ($\Lambda^n V^*$ is 1–dimensional)**.** *Let V be an n–dimensional real vector space. Then $\Lambda^n V^*$ is a one–dimensional vector space. More precisely, if $(e_1, \dots, e_n)$ is a basis of V and $(\varepsilon^1, \dots, \varepsilon^n)$ is the dual basis of V^*, then*

$$\varepsilon^1 \wedge \cdots \wedge \varepsilon^n \neq 0$$

and every $\eta \in \Lambda^n V^*$ can be written uniquely as

$$\eta = c\, \varepsilon^1 \wedge \cdots \wedge \varepsilon^n \qquad \textit{for a unique } c \in \mathbb{R}.$$

Proof. We first show $\varepsilon^1 \wedge \cdots \wedge \varepsilon^n \neq 0$.

We first show that $\varepsilon^1 \wedge \cdots \wedge \varepsilon^n \neq 0$. Recall that the wedge product of 1–forms is defined as the alternation of their tensor product:

$$\varepsilon^1 \wedge \cdots \wedge \varepsilon^n := \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, \varepsilon^{\sigma(1)} \otimes \cdots \otimes \varepsilon^{\sigma(n)}.$$

Evaluating this n–form on the basis vectors $(e_1, \dots, e_n)$ gives

$$\begin{aligned}(\varepsilon^1 \wedge \cdots \wedge \varepsilon^n)(e_1, \dots, e_n) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, \varepsilon^{\sigma(1)}(e_1) \cdots \varepsilon^{\sigma(n)}(e_n) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, \delta_{\sigma(1)1} \cdots \delta_{\sigma(n)n}.\end{aligned}$$

The product $\delta_{\sigma(1)1} \cdots \delta_{\sigma(n)n}$ is equal to 1 if and only if $\sigma = \mathrm{id}$, and vanishes otherwise. Hence the sum reduces to a single nonzero term, and we obtain

$$(\varepsilon^1 \wedge \cdots \wedge \varepsilon^n)(e_1, \dots, e_n) = 1.$$

In particular, $\varepsilon^1 \wedge \cdots \wedge \varepsilon^n$ does not vanish identically and therefore is a nonzero element of $\Lambda^n V^*$.

Now let $\eta \in \Lambda^n V^*$ be arbitrary. Since η is alternating and multilinear, for any $v_1, \dots, v_n \in V$ we can write each vector $v_j \in V$ in the basis $(e_1, \dots, e_n)$ as

$$v_j = \sum_{m=1}^n a_{mj}\, e_m, \qquad j = 1, \dots, n,$$

and set $A := (a_{mj})_{m,j=1}^n$. By multilinearity of η we expand

$$\begin{aligned}\eta(v_1, \dots, v_n) &= \eta\Big(\sum_{i_1=1}^n a_{i_1 1} e_{i_1},\ \dots,\ \sum_{i_n=1}^n a_{i_n n} e_{i_n}\Big) \\ &= \sum_{i_1, \dots, i_n = 1}^n a_{i_1 1} \cdots a_{i_n n}\, \eta(e_{i_1}, \dots, e_{i_n}).\end{aligned}$$

Since η is alternating, every term with $i_r = i_s$ for some $r \neq s$ vanishes. Hence only tuples $(i_1, \ldots, i_n)$ with pairwise distinct entries contribute, i.e. exactly those of the form

$$(i_1, \ldots, i_n) = (\sigma(1), \ldots, \sigma(n)) \qquad \text{for some } \sigma \in S_n.$$

Therefore

$$\eta(v_1, \ldots, v_n) = \sum_{\sigma \in S_n} a_{\sigma(1)1} \cdots a_{\sigma(n)n} \, \eta(e_{\sigma(1)}, \ldots, e_{\sigma(n)}). \tag{6.26}$$

Now use alternation once more:

$$\eta(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) = \operatorname{sgn}(\sigma) \, \eta(e_1, \ldots, e_n).$$

With $c := \eta(e_1, \ldots, e_n)$, (6.26) becomes

$$\eta(v_1, \ldots, v_n) = c \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \, a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

By the Leibniz formula for the determinant,

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \, a_{\sigma(1)1} \cdots a_{\sigma(n)n},$$

hence

$$\eta(v_1, \ldots, v_n) = \eta(e_1, \ldots, e_n) \, \det(a_{mj})_{m,j=1}^n.$$

In particular

$$(\varepsilon^1 \wedge \cdots \wedge \varepsilon^n)(v_1, \ldots, v_n) = \det(a_{mj})_{m,j=1}^n.$$

Hence for all $v_1, \ldots, v_n \in V$,

$$\eta(v_1, \ldots, v_n) = c \, (\varepsilon^1 \wedge \cdots \wedge \varepsilon^n)(v_1, \ldots, v_n),$$

which proves $\eta = c \, \varepsilon^1 \wedge \cdots \wedge \varepsilon^n$.

Finally, uniqueness: if $\eta = c \, \varepsilon^1 \wedge \cdots \wedge \varepsilon^n = d \, \varepsilon^1 \wedge \cdots \wedge \varepsilon^n$, evaluate both sides at $(e_1, \ldots, e_n)$ to get $c = d$. Thus $\Lambda^n V^*$ is spanned by one nonzero element, hence is 1–dimensional. □

If ω has compact support in U, we define

$$\int_U \omega \; := \; \int_U f(x) \, dx, \tag{6.27}$$

where the right-hand side is the usual Lebesgue (or Riemann) integral on $\mathbb{R}^n$.

6.7.2 Behavior under diffeomorphisms

Let $U, V \subset \mathbb{R}^n$ be open and let $\Phi : V \to U$ be a diffeomorphism. Write $(y^1, \ldots, y^n)$ for the standard coordinates on V. We first record the effect of pullback on the volume form.

Lemma 6.57. *For a diffeomorphism $\Phi : V \to U$ we have*

$$\Phi^*(dx^1 \wedge \cdots \wedge dx^n) = \det(D\Phi) \, dy^1 \wedge \cdots \wedge dy^n,$$

where $D\Phi$ denotes the Jacobian matrix of Φ in the coordinates y.

Proof. Write $\Phi = (\Phi^1, \dots, \Phi^n)$ in the x–coordinates on U, so each Φ^i is a smooth function on V. By definition of pullback on 1–forms,

$$\Phi^* dx^i = d(x^i \circ \Phi) = d\Phi^i = \sum_{j=1}^n \frac{\partial \Phi^i}{\partial y^j}\, dy^j.$$

Hence, using that Φ^* is an algebra homomorphism with respect to $\wedge$,

$$\Phi^*(dx^1 \wedge \cdots \wedge dx^n) = \Phi^* dx^1 \wedge \cdots \wedge \Phi^* dx^n = \left(\sum_{j_1} \frac{\partial \Phi^1}{\partial y^{j_1}} dy^{j_1}\right) \wedge \cdots \wedge \left(\sum_{j_n} \frac{\partial \Phi^n}{\partial y^{j_n}} dy^{j_n}\right).$$

Expanding by multilinearity, all terms with repeated indices vanish since $dy^j \wedge dy^j = 0$, so only permutations survive:

$$\Phi^*(dx^1 \wedge \cdots \wedge dx^n) = \sum_{\sigma \in S_n} \frac{\partial \Phi^1}{\partial y^{\sigma(1)}} \cdots \frac{\partial \Phi^n}{\partial y^{\sigma(n)}}\, dy^{\sigma(1)} \wedge \cdots \wedge dy^{\sigma(n)}.$$

Now reorder each wedge product to $(dy^1 \wedge \cdots \wedge dy^n)$:

$$dy^{\sigma(1)} \wedge \cdots \wedge dy^{\sigma(n)} = \operatorname{sgn}(\sigma)\, dy^1 \wedge \cdots \wedge dy^n.$$

Therefore

$$\Phi^*(dx^1 \wedge \cdots \wedge dx^n) = \left(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, \frac{\partial \Phi^1}{\partial y^{\sigma(1)}} \cdots \frac{\partial \Phi^n}{\partial y^{\sigma(n)}}\right) dy^1 \wedge \cdots \wedge dy^n.$$

By the Leibniz formula, the coefficient in parentheses is exactly $\det(D\Phi)$, where $D\Phi = (\partial \Phi^i / \partial y^j)_{1 \le i,j \le n}$. □

When Φ is a diffeomorphism, $\det(D\Phi)$ never vanishes and hence has a constant sign on each connected component of V.

Definition 6.58 (Orientation–preserving / reversing)**.** Let $\Phi : V \to U$ be a diffeomorphism between open subsets of $\mathbb{R}^n$. We call Φ *orientation–preserving* if $\det(D\Phi) > 0$ everywhere, and *orientation–reversing* if $\det(D\Phi) < 0$ everywhere.

Proposition 6.59 (Change of variables for n–forms)**.** *Let $\omega \in \Omega^n_c(U)$. If Φ is an orientation-preserving diffeomorphism, then*

$$\int_V \Phi^* \omega = \int_U \omega. \tag{6.28}$$

If $\omega = f\, dx^1 \wedge \cdots \wedge dx^n$, then

$$\Phi^* \omega = (f \circ \Phi)\, \det(D\Phi)\, dy^1 \wedge \cdots \wedge dy^n,$$

and (6.28) is the classical substitution formula.

Proof. By Lemma 6.57,

$$\Phi^* \omega = (f \circ \Phi)\, \Phi^*(dx^1 \wedge \cdots \wedge dx^n) = (f \circ \Phi)\, \det(D\Phi)\, dy^1 \wedge \cdots \wedge dy^n.$$

Integrating and applying the usual change–of–variables theorem yields (6.28). □

6.7.3 Orientability and orientations on manifolds

Let M be an n–dimensional smooth manifold.

Definition 6.60 (Oriented atlas and orientability)**.** An atlas $\{(U_\alpha, \varphi_\alpha)\}$ on M is called *oriented* if for every pair α, β with $U_\alpha \cap U_\beta \neq \emptyset$, the transition map

$$\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \longrightarrow \varphi_\beta(U_\alpha \cap U_\beta)$$

is orientation–preserving. We say that M is *orientable* if it admits an oriented atlas.

Exactly as for smooth structures, one can enlarge an oriented atlas to a maximal one.

Definition 6.61 (Orientation as a maximal oriented atlas)**.** Assume M is orientable. An *orientation* on M is a maximal oriented atlas, i.e. an oriented atlas which is not properly contained in any larger oriented atlas. An *oriented manifold* is an orientable manifold together with a choice of orientation.

Remark 6.62 (Two possible orientations)**.** If M is connected and orientable, then it has exactly two orientations. Indeed, once a maximal oriented atlas is fixed, replacing a single chart φ_α by a composition with an orientation–reversing linear map $\mathbb{R}^n \to \mathbb{R}^n$ produces the other maximal oriented atlas; there is no third possibility because any two oriented atlases must have transition maps with Jacobian of constant sign, hence either all overlaps preserve orientation or all overlaps reverse it.

Example 6.63.

- $\mathbb{R}^n$ is orientable (with the standard orientation).
- S^n is orientable (e.g. by stereographic coordinates).
- $\mathbb{RP}^n$ is orientable if and only if n is odd.
- Every complex manifold is orientable: the complex structure canonically induces an orientation on the underlying real tangent spaces.

6.7.4 Integration on an oriented manifold

From now on let M be an oriented n–manifold.

Definition 6.64 (Integral of a compactly supported top–degree form)**.** Let $\omega \in \Omega_c^n(M)$. Choose a (positively) oriented atlas $\{(U_\alpha, \varphi_\alpha)\}$ representing the given orientation, and a smooth partition of unity $\{\rho_\alpha\}$ subordinate to $\{U_\alpha\}$. We define

$$\int_M \omega \;:=\; \sum_\alpha \int_{\varphi_\alpha(U_\alpha)} (\varphi_\alpha^{-1})^*(\rho_\alpha\, \omega), \tag{6.29}$$

where each integral on the right is the Euclidean integral (6.27). The sum is finite because ω has compact support.

Proposition 6.65 (Well-definedness)**.** *The value $\int_M \omega$ in (6.29) depends only on the oriented smooth structure on M. In particular, it is independent of the choice of oriented atlas representing the given orientation and independent of the chosen subordinate partition of unity.*

Proof. We first show independence of the atlas (for a fixed partition argument). Let (U, φ) and (U, ψ) be two positively oriented charts on the same open set. Set $\Phi := \varphi \circ \psi^{-1}$, a diffeomorphism between open subsets of $\mathbb{R}^n$. Since both charts are positively oriented, the transition map Φ is orientation–preserving, i.e. $\det(D\Phi) > 0$. For any $\eta \in \Omega_c^n(U)$, Proposition 6.59 gives

$$\int_{\varphi(U)} (\varphi^{-1})^* \eta = \int_{\psi(U)} (\psi^{-1})^* \eta.$$

Applying this to $\eta = \rho\,\omega$ shows that each local term in (6.29) is independent of the choice of positively oriented coordinates.

Next, we show independence of the partition of unity. Let $\{\rho_\alpha\}$ and $\{\tilde{\rho}_\beta\}$ be two partitions of unity subordinate to the same oriented atlas. Using linearity of the integral and $\sum_\beta \tilde{\rho}_\beta = 1$ on $\operatorname{supp} \omega$, we obtain

$$\sum_\alpha \int_M \rho_\alpha\, \omega = \sum_\alpha \int_M \Big(\rho_\alpha \sum_\beta \tilde{\rho}_\beta\Big) \omega = \sum_{\alpha,\beta} \int_M (\rho_\alpha \tilde{\rho}_\beta)\, \omega.$$

By the already established coordinate independence, each term $\int_M (\rho_\alpha \tilde{\rho}_\beta)\, \omega$ may be computed in any positively oriented chart contained in $U_\alpha \cap U_\beta$. Reversing the roles of ρ and $\tilde{\rho}$ yields

$$\sum_{\alpha,\beta} \int_M (\rho_\alpha \tilde{\rho}_\beta)\, \omega = \sum_\beta \int_M \tilde{\rho}_\beta\, \omega,$$

hence the two definitions agree. □

6.8 Stokes' theorem

We now turn to the fundamental theorem relating exterior differentiation and integration.

6.8.1 Euclidean Stokes for compactly supported forms

Let $\mathbb{R}^n$ carry its standard orientation given by $dx^1 \wedge \cdots \wedge dx^n$.

Theorem 6.66 (Euclidean Stokes for compact support)**.** *For every $\omega \in \Omega_c^{n-1}(\mathbb{R}^n)$ one has*

$$\int_{\mathbb{R}^n} d\omega \;=\; 0.$$

Proof. Write ω in the standard basis of $(n-1)$–forms:

$$\omega = \sum_{j=1}^{n} a_j\, dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^n,$$

where the hat indicates that the corresponding factor is omitted, and $a_j \in C_c^\infty(\mathbb{R}^n)$. Since $dx^1, \ldots, dx^n$ are closed, we have

$$d\omega = \sum_{j=1}^{n} da_j \wedge \big(dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^n\big).$$

Only the dx^j–term in $da_j = \sum_i \frac{\partial a_j}{\partial x^i} dx^i$ survives in the wedge product, hence

$$d\omega = \sum_{j=1}^{n} \frac{\partial a_j}{\partial x^j}\, dx^j \wedge dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^n = \sum_{j=1}^{n} (-1)^{j-1} \frac{\partial a_j}{\partial x^j}\, dx^1 \wedge \cdots \wedge dx^n.$$

Choose a cube $Q = [-R, R]^n$ containing $\operatorname{supp}\omega$. Then $\operatorname{supp}(d\omega) \subset Q$ and therefore

$$\int_{\mathbb{R}^n} d\omega = \int_Q d\omega = \sum_{j=1}^{n} (-1)^{j-1} \int_Q \frac{\partial a_j}{\partial x^j}\, dx^1 \cdots dx^n.$$

Fix j. By Fubini, writing $x = (x', x^j)$ with $x' \in [-R, R]^{n-1}$,

$$\int_Q \frac{\partial a_j}{\partial x^j}\, dx = \int_{[-R,R]^{n-1}} \left(\int_{-R}^{R} \frac{\partial a_j}{\partial x^j}(x', t)\, dt \right) dx'.$$

By the Newton–Leibniz formula,

$$\int_{-R}^{R} \frac{\partial a_j}{\partial x^j}(x', t)\, dt = a_j(x', R) - a_j(x', -R) = 0,$$

because a_j has compact support contained in the interior of Q. Thus each term vanishes, and $\int_{\mathbb{R}^n} d\omega = 0$. □

6.8.2 Stokes on oriented manifolds without boundary

Theorem 6.67 (Stokes, no boundary)**.** *Let M be an oriented smooth n–manifold without boundary. Then for every $\omega \in \Omega_c^{n-1}(M)$ one has*

$$\int_M d\omega \;=\; 0.$$

Proof. Let $\{(U_\alpha, \varphi_\alpha)\}$ be a positively oriented atlas and $\{\rho_\alpha\}$ a partition of unity subordinate to $\{U_\alpha\}$. By Definition 6.64 and linearity,

$$\int_M d\omega = \sum_\alpha \int_M \rho_\alpha\, d\omega = \sum_\alpha \int_M d(\rho_\alpha\, \omega) - \sum_\alpha \int_M d\rho_\alpha \wedge \omega.$$

Since $\sum_\alpha \rho_\alpha = 1$ on $\operatorname{supp}\omega$, we have $\sum_\alpha d\rho_\alpha = d(\sum_\alpha \rho_\alpha) = 0$ on $\operatorname{supp}\omega$, hence $\sum_\alpha d\rho_\alpha \wedge \omega = 0$ and therefore

$$\int_M d\omega = \sum_\alpha \int_M d(\rho_\alpha\, \omega).$$

Each form $d(\rho_\alpha \omega)$ has compact support contained in U_α, so by the coordinate definition of the integral,

$$\int_M d(\rho_\alpha \omega) = \int_{\varphi_\alpha(U_\alpha)} (\varphi_\alpha^{-1})^* d(\rho_\alpha \omega) = \int_{\varphi_\alpha(U_\alpha)} d\big((\varphi_\alpha^{-1})^*(\rho_\alpha \omega)\big),$$

where we used that pullback commutes with d. The form $(\varphi_\alpha^{-1})^*(\rho_\alpha \omega)$ is a compactly supported $(n-1)$–form on an open subset of $\mathbb{R}^n$, so by Theorem 6.66 the last integral is 0. Summing over α yields $\int_M d\omega = 0$. □

Manifolds with boundary and the induced boundary orientation. Let M be a smooth n–manifold with boundary ∂M. A *boundary chart* is a pair (U, φ) where $U \subset M$ is open and

$$\varphi : U \longrightarrow V \subset \mathbb{H}^n := \{x \in \mathbb{R}^n : \ x^1 \leq 0\}$$

is a diffeomorphism onto an open subset V of the closed half–space, such that $\varphi(U \cap \partial M) = V \cap \{x^1 = 0\}$. An *oriented atlas* for M is an atlas consisting of interior charts to $\mathbb{R}^n$ and boundary charts to $\mathbb{H}^n$ whose transition maps are orientation–preserving on the interior.

an *orientation* on an n–manifold M is specified by a maximal oriented atlas $\{(U_\alpha, \varphi_\alpha)\}$

Definition 6.68 (Induced orientation on the boundary)**.** Let $p \in \partial M$. Choose a positively oriented boundary chart (U, φ) with $p \in U$, where

$$\varphi : U \longrightarrow V \subset \mathbb{H}^n = \{x^1 \leq 0\}, \qquad \varphi(U \cap \partial M) = V \cap \{x^1 = 0\}.$$

An ordered basis $(v_1, \ldots, v_{n-1})$ of $T_p(\partial M)$ is declared *positively oriented* if and only if

$$\big(D\varphi_p(v_1), \ldots, D\varphi_p(v_{n-1})\big)$$

is a positively oriented basis of $\mathbb{R}^{n-1}$ with respect to the standard orientation induced by the coordinates $(x^2, \ldots, x^n)$ on $\{x^1 = 0\}$.

Equivalently, $(v_1, \ldots, v_{n-1})$ is positively oriented in $T_p(\partial M)$ if and only if

$$\big(D\varphi_p(\partial/\partial x^1), \ D\varphi_p(v_1), \ldots, D\varphi_p(v_{n-1})\big)$$

is a positively oriented basis of $\mathbb{R}^n$.

With this induced orientation on $\partial\mathbb{H}^n = \{x^1 = 0\}$, the standard coordinates $(x^2, \ldots, x^n)$ are positively oriented. Consequently, for any compactly supported function η on $\partial\mathbb{H}^n$ we simply have

$$\int_{\partial\mathbb{H}^n} \eta \, dx^2 \wedge \cdots \wedge dx^n \ = \ \int_{\mathbb{R}^{n-1}} \eta,$$

where the right–hand side denotes the usual Euclidean integral on $\mathbb{R}^{n-1}$. In particular, no additional sign appears in the boundary integral under this convention.

6.8.3 Stokes on the half–space

Let $\iota : \partial\mathbb{H}^n \hookrightarrow \mathbb{H}^n$ denote the inclusion.

Theorem 6.69 (Stokes on $\mathbb{H}^n$)**.** *Let $\omega \in \Omega_c^{n-1}(\mathbb{H}^n)$ have compact support in $\mathbb{H}^n$ (which is automatically compact in $\mathbb{R}^n$). Then*

$$\int_{\mathbb{H}^n} d\omega \ = \ \int_{\partial\mathbb{H}^n} \iota^*\omega,$$

where $\partial\mathbb{H}^n$ is oriented by Definition 6.68.

Proof. Write

$$\omega = \sum_{j=1}^{n} a_j \, dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^n, \qquad a_j \in C_c^\infty(\mathbb{H}^n).$$

As in the Euclidean computation,

$$d\omega = \sum_{j=1}^{n} (-1)^{j-1} \frac{\partial a_j}{\partial x^j} \, dx^1 \wedge \cdots \wedge dx^n.$$

Choose a rectangular box

$$Q = [-R, 0] \times [-R, R]^{n-1} \subset \mathbb{H}^n$$

that contains $\operatorname{supp} \omega$. Then $\operatorname{supp}(d\omega) \subset Q$ and hence

$$\int_{\mathbb{H}^n} d\omega = \int_Q d\omega = \sum_{j=1}^{n} (-1)^{j-1} \int_Q \frac{\partial a_j}{\partial x^j} \, dx^1 \cdots dx^n.$$

For $j \geq 2$, the same Fubini–Newton–Leibniz argument as in the Euclidean case shows

$$\int_Q \frac{\partial a_j}{\partial x^j} \, dx = 0,$$

because the support is away from the lateral faces $x^j = \pm R$. For $j = 1$, Fubini gives

$$\int_Q \frac{\partial a_1}{\partial x^1} \, dx = \int_{[-R,R]^{n-1}} \left(\int_0^R \frac{\partial a_1}{\partial x^1}(t, x') \, dt \right) dx',$$

and Newton–Leibniz yields

$$\int_{-R}^{0} \frac{\partial a_1}{\partial x^1}(t, x') \, dt = a_1(0, x') - a_1(-R, x') = a_1(0, x'),$$

since a_1 vanishes near $x^1 = -R$. Therefore

$$\int_{\mathbb{H}^n} d\omega = \int_Q d\omega = -\int_{[-R,R]^{n-1}} a_1(0, x') \, dx^2 \cdots dx^n.$$

On the other hand, the pullback to the boundary $\{x^1 = 0\}$ kills every term containing dx^1, so only the $j = 1$ term survives:

$$\iota^* \omega = a_1(0, x') \, dx^2 \wedge \cdots \wedge dx^n.$$

With the induced boundary orientation, the boundary integral is

$$\int_{\partial \mathbb{H}^n} \iota^* \omega = \int_{\mathbb{R}^{n-1}} a_1(0, x') \left(dx^2 \cdots dx^n \right) = -\int_{\mathbb{R}^{n-1}} a_1(0, x') \, dx^2 \cdots dx^n,$$

and compact support allows us to restrict to $[-R, R]^{n-1}$. This agrees with the expression for $\int_{\mathbb{H}^n} d\omega$ above. □

6.8.4 Stokes' theorem on manifolds with boundary

Let $\iota : \partial M \hookrightarrow M$ denote the inclusion.

Theorem 6.70 (Stokes' theorem)**.** *Let M be an oriented smooth n–manifold with boundary, and orient ∂M by Definition 6.68. Then for every $\omega \in \Omega_c^{n-1}(M)$ one has*

$$\int_M d\omega \;=\; \int_{\partial M} \iota^*\omega.$$

Proof. Let $\{(U_\alpha, \varphi_\alpha)\}$ be an oriented atlas consisting of interior charts $\varphi_\alpha : U_\alpha \to V_\alpha \subset \mathbb{R}^n$ and boundary charts $\varphi_\alpha : U_\alpha \to V_\alpha \subset \mathbb{H}^n$, and let $\{\rho_\alpha\}$ be a partition of unity subordinate to $\{U_\alpha\}$.

As in the proof of Theorem 6.67,

$$\int_M d\omega = \sum_\alpha \int_M d(\rho_\alpha\omega),$$

because $\sum_\alpha d\rho_\alpha = 0$ on $\operatorname{supp}\omega$. Fix α. If U_α is an interior chart, then by pulling back to $\mathbb{R}^n$ and applying Theorem 6.66,

$$\int_M d(\rho_\alpha\omega) = 0.$$

If U_α is a boundary chart, then pulling back to $\mathbb{H}^n$ gives

$$\int_M d(\rho_\alpha\omega) = \int_{V_\alpha} d\big((\varphi_\alpha^{-1})^*(\rho_\alpha\omega)\big) = \int_{\partial V_\alpha} \iota^*\big((\varphi_\alpha^{-1})^*(\rho_\alpha\omega)\big),$$

by Theorem 6.69. The boundary ∂V_α corresponds precisely to $U_\alpha \cap \partial M$ and the boundary orientation is compatible with Definition 6.68 by construction of oriented boundary charts. Therefore this equals

$$\int_{U_\alpha\cap\partial M} \iota^*(\rho_\alpha\omega).$$

Summing over α, we obtain

$$\int_M d\omega = \sum_\alpha \int_{U_\alpha\cap\partial M} \iota^*(\rho_\alpha\omega) = \int_{\partial M} \iota^*\Big(\sum_\alpha \rho_\alpha\omega\Big) = \int_{\partial M} \iota^*\omega,$$

since $\sum_\alpha \rho_\alpha = 1$ on $\operatorname{supp}\omega$ and hence on $\operatorname{supp}(\iota^*\omega) \subset \partial M$. □

Remark 6.71 (Fundamental theorem of calculus)**.** Let $M = [a,b] \subset \mathbb{R}$ with its standard orientation. For $f \in C^\infty([a,b])$ consider the 0–form $\omega = f$. Then $d\omega = f'(x)\,dx$, and Stokes' theorem gives

$$\int_a^b f'(x)\,dx = f(b) - f(a),$$

where the boundary $\partial[a,b] = \{b\} - \{a\}$ carries the induced orientation.

6.9 Orientation and volume forms

We now record an equivalent description of orientations in terms of nowhere–vanishing top–degree forms, and then give several applications of Stokes' theorem.

6.9.1 Orientations are the same as volume forms

Definition 6.72 (Volume form)**.** Let M be a smooth n–manifold. A *volume form* on M is a smooth n–form $\omega \in \Omega^n(M)$ which is nowhere vanishing.

Theorem 6.73 (Orientation and volume forms)**.** *Let M be a smooth n–manifold. The following are equivalent:*

(i) M is orientable.

(ii) M admits a volume form.

Moreover, if M is connected, then choosing an orientation on M is equivalent to choosing a volume form ω, where two volume forms determine the same orientation if and only if they differ by multiplication with a positive smooth function.

Proof. $(i) \Rightarrow (ii)$. Assume M is orientable and fix a positively oriented atlas $\{(U_\alpha, \varphi_\alpha)\}$, where each $\varphi_\alpha : U_\alpha \to V_\alpha$ is a diffeomorphism onto an open set $V_\alpha \subset \mathbb{R}^n$. On each U_α define

$$\omega_\alpha := (\varphi_\alpha^{-1})^*(dx^1 \wedge \cdots \wedge dx^n),$$

which is a smooth nowhere–vanishing n–form on U_α.

If $U_\alpha \cap U_\beta \neq \emptyset$, let

$$\Phi_{\beta\alpha} := \varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$$

be the transition map. On $U_\alpha \cap U_\beta$ we have

$$\omega_\alpha = (\det(D\Phi_{\beta\alpha}) \circ \varphi_\alpha)\, \omega_\beta.$$

Since the atlas is positively oriented, $\det(D\Phi_{\beta\alpha}) > 0$ on overlaps, so the ratio between ω_α and ω_β is a positive smooth function on $U_\alpha \cap U_\beta$.

Choose a smooth partition of unity $\{\rho_\alpha\}$ subordinate to $\{U_\alpha\}$ and define the global n–form

$$\omega := \sum_\alpha \rho_\alpha\, \omega_\alpha.$$

At any point $p \in M$, at least one $\rho_\alpha(p) > 0$ and all nonzero $\omega_\alpha(p)$ determine the same orientation, so they cannot cancel. Hence $\omega(p) \neq 0$ for all $p \in M$, and ω is a volume form.

$(ii) \Rightarrow (i)$. Assume $\omega \in \Omega^n(M)$ is a volume form. Let (U, φ) be any chart, with coordinates $(x^1, \ldots, x^n)$ on $\varphi(U) \subset \mathbb{R}^n$. Since ω is a top–degree form, on U we can write uniquely

$$\omega|_U = f\, dx^1 \wedge \cdots \wedge dx^n$$

for some smooth function f on U. Because ω is nowhere vanishing, f is nowhere zero.

We declare the chart (U, φ) to be *positively oriented* if $f > 0$ on U. Now take two such positively oriented charts (U, φ) and (V, ψ) with $U \cap V \neq \emptyset$. On the overlap, we have

$$f_\varphi\, dx^1 \wedge \cdots \wedge dx^n = \omega = f_\psi\, dy^1 \wedge \cdots \wedge dy^n,$$

and we obtain

$$f_\psi = f_\varphi \cdot \det\big(D(\psi \circ \varphi^{-1})\big).$$

Since $f_\varphi > 0$ and $f_\psi > 0$, it follows that $\det(D(\psi \circ \varphi^{-1})) > 0$ on the overlap. Hence all transition maps between positively oriented charts are orientation–preserving. Therefore the collection of positively oriented charts forms an oriented atlas, and M is orientable.

Finally, if ω is a volume form and g is a smooth positive function, then $g\omega$ determines the same set of positively oriented charts, hence the same orientation. Conversely, if ω and $\tilde{\omega}$ determine the same orientation on a connected manifold, then in any positively oriented chart they are given by $\omega = f\, dx^1 \wedge \cdots \wedge dx^n$ and $\tilde{\omega} = \tilde{f}\, dx^1 \wedge \cdots \wedge dx^n$ with $f, \tilde{f} > 0$, so $\tilde{\omega} = (\tilde{f}/f)\,\omega$ with $\tilde{f}/f > 0$. ☐

Remark 6.74 (Positive total volume)**.** Let M be compact, connected and oriented, and let ω be a volume form compatible with the chosen orientation. Then

$$\int_M \omega > 0.$$

Indeed, in positively oriented coordinates $\omega = f\, dx^1 \wedge \cdots \wedge dx^n$ with $f > 0$, so the integral is locally positive and hence globally positive.

Remark 6.75. Let M be an oriented smooth n–manifold with boundary and let $\iota : \partial M \hookrightarrow M$ be the inclusion. If $\omega \in \Omega^n(M)$ is a volume form compatible with the chosen orientation on M, then $\iota^*(\iota_N \omega)$ is a volume form on ∂M compatible with the induced boundary orientation, where N denotes an outward normal vector field along ∂M.

6.9.2 Stokes and a "no retraction to the boundary" principle

The next lemma is a convenient formulation of the key idea behind the Brouwer fixed point theorem.

Lemma 6.76. *Let M be a compact oriented smooth n–manifold with nonempty boundary. There is no smooth map $F : M \to M$ such that*

$$F|_{\partial M} = \mathrm{id}_{\partial M}\,.$$

Proof. Let $\omega \in \Omega^{n-1}(\partial M)$ be a volume form on the boundary compatible with the induced orientation. Consider the $(n-1)$–form $\eta := F^*\omega$ on M.

If $\iota : \partial M \hookrightarrow M$ denotes the inclusion, then

$$d\eta = F^*(d\omega) = 0,$$

because $d\omega$ is an n–form on the $(n-1)$–dimensional manifold ∂M.

Then

$$\int_{\partial M} \omega = \int_{\partial M} \iota^*(F^*\omega) = \int_M d\eta = 0,$$

while $\int_{\partial M} \omega > 0$ by Remark 6.74. A contradiction. ☐

6.9.3 Brouwer fixed point theorem

Theorem 6.77 (Brouwer fixed point theorem)**.** *Every continuous map $f : B^n \to B^n$ has a fixed point.*

Proof. Assume that f has no fixed point. One then constructs a continuous map $r : B^n \to \partial B^n$ by sending x to the intersection point of ∂B^n with the ray starting at $f(x)$ and passing through x. This map restricts to the identity on ∂B^n. After smoothing (or using a smooth approximation argument), we obtain a smooth map $F : B^n \to B^n$ with $F|_{\partial B^n} = \mathrm{id}_{\partial B^n}$, contradicting Lemma 6.76. □

6.10 An application to vector fields on spheres

6.10.1 Even-dimensional spheres have no nowhere–vanishing tangent field

Theorem 6.78 (Hairy ball theorem for even spheres)**.** *If n is even, then there is no smooth nowhere–vanishing tangent vector field on S^n.*

Proof. Assume for contradiction that v is a smooth nowhere–vanishing tangent vector field on S^n. Normalize it so that $|v(x)| = 1$ for all $x \in S^n$.

Define, for $t \in [0, \pi]$, a smooth map $F_t : S^n \to S^n$ by

$$F_t(x) = \cos t\, x + \sin t\, v(x).$$

Since $v(x) \perp x$ and $|v(x)| = |x| = 1$, we have $|F_t(x)|^2 = \cos^2 t + \sin^2 t = 1$, so $F_t(x) \in S^n$ for all x, t. Moreover $F_0 = \mathrm{id}$ and $F_\pi(x) = -x$ is the antipodal map.

Embed $S^n \subset \mathbb{R}^{n+1}$ with standard coordinates $(x^1, \dots, x^{n+1})$, and let

$$\Omega := dx^1 \wedge \cdots \wedge dx^{n+1}$$

be the Euclidean volume form on $\mathbb{R}^{n+1}$. Define an n–form on $\mathbb{R}^{n+1} \setminus \{0\}$ by

$$\omega := \iota_R \Omega, \qquad R := \sum_{i=1}^{n+1} x^i \frac{\partial}{\partial x^i}$$

(the radial vector field). Restricting to S^n gives a smooth nowhere–vanishing n–form on S^n, hence a volume form:

$$\omega|_{S^n} \in \Omega^n(S^n).$$

In coordinates one can write

$$\omega = \sum_{i=1}^{n+1} (-1)^{i-1} x^i\, dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^{n+1}.$$

Let $A : S^n \to S^n$ be the antipodal map $A(x) = -x$. Then $A^* x^i = -x^i$ and $A^* dx^i = d(A^* x^i) = -dx^i$. Therefore each summand transforms as

$$A^*\Big(x^i\, dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^{n+1}\Big) = (-x^i) \cdot (-1)^n\, dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^{n+1}.$$

(The factor $(-1)^n$ comes from pulling back the wedge of n differentials dx^j with $j \neq i$.)

Hence

$$A^*\omega = (-1)^{n+1}\omega.$$

In particular, if n is even then $A^*\omega = -\omega$.

Consider the smooth map

$$F : S^n \times [0,\pi] \to S^n, \qquad F(x,t) = F_t(x).$$

Since $d\omega = 0$, we have $d(F^*\omega) = F^*(d\omega) = 0$. By Stokes' theorem on $S^n \times [0,\pi]$,

$$0 = \int_{S^n \times [0,\pi]} d(F^*\omega) = \int_{S^n} F_\pi^*\omega - \int_{S^n} F_0^*\omega.$$

Hence

$$\int_{S^n} F_\pi^*\omega = \int_{S^n} \omega.$$

Now F_π is the antipodal map $A(x) = -x$. When n is even,

$$A^*\omega = -\,\omega.$$

Consequently,

$$\int_{S^n} F_\pi^*\omega = \int_{S^n} A^*\omega = -\int_{S^n} \omega,$$

contradicting $\int_{S^n} \omega \neq 0$. This finishes the proof. □

Remark 6.79. The argument above is a special case of the following principle: if H_t is a smooth homotopy and ω is closed, then $\int H_t^*\omega$ is independent of t. A more general and systematic approach uses the Cartan magic formula

$$\mathcal{L}_X = d\,\iota_X + \iota_X\,d,$$

which we discuss below.

6.11 Classical vector calculus as consequences of Stokes

We briefly explain how the familiar integral identities from vector calculus are special cases of Stokes' theorem. In order to state this precisely we need to recall the Hodge star on Euclidean space and its relation to the standard volume form.

The Euclidean Hodge star in $\mathbb{R}^3$

Let (x^1, x^2, x^3) be the standard coordinates on $\mathbb{R}^3$ and let

$$dV := dx^1 \wedge dx^2 \wedge dx^3$$

be the standard volume form. We use the Euclidean inner product $\langle \cdot, \cdot \rangle$ and the associated musical isomorphism

$$\flat : \ \mathfrak{X}(\mathbb{R}^3) \to \Omega^1(\mathbb{R}^3), \qquad X = X^i \frac{\partial}{\partial x^i} \ \mapsto \ X^\flat := \sum_{i=1}^3 X^i\,dx^i.$$

The Hodge star operator

$$* : \Omega^k(\mathbb{R}^3) \to \Omega^{3-k}(\mathbb{R}^3)$$

is characterized by the requirement that for any $\alpha, \beta \in \Omega^k(\mathbb{R}^3)$,

$$\alpha \wedge *\beta = \langle \alpha, \beta \rangle \, dV, \tag{6.30}$$

where $\langle \alpha, \beta \rangle$ denotes the pointwise inner product on k–forms induced by the Euclidean metric. In particular, using that $\{dx^1, dx^2, dx^3\}$ is an orthonormal coframe, one checks from (6.30) that

$$*1 = dV, \qquad *dx^1 = dx^2 \wedge dx^3, \quad *dx^2 = dx^3 \wedge dx^1, \quad *dx^3 = dx^1 \wedge dx^2,$$

and

$$*(dx^2 \wedge dx^3) = dx^1, \quad *(dx^3 \wedge dx^1) = dx^2, \quad *(dx^1 \wedge dx^2) = dx^3, \quad *dV = 1.$$

In particular, for any vector field X we have

$$*X^\flat = X^1 \, dx^2 \wedge dx^3 + X^2 \, dx^3 \wedge dx^1 + X^3 \, dx^1 \wedge dx^2. \tag{6.31}$$

Proposition 6.80 (Divergence theorem and Stokes' theorem in $\mathbb{R}^3$)**.** *Let $\Omega \subset \mathbb{R}^3$ be a bounded domain with smooth boundary, and let ν be the outward unit normal along $\partial\Omega$.*

(i) (Gauss–Green / divergence theorem) For every smooth vector field X on a neighborhood of $\overline{\Omega}$,

$$\int_\Omega \operatorname{div} X \, dV = \int_{\partial\Omega} \langle X, \nu \rangle \, dS.$$

(ii) (Classical Stokes formula) Let $\Sigma \subset \mathbb{R}^3$ be a compact oriented smooth surface with (possibly empty) boundary $\partial\Sigma$. Let ν be the unit normal compatible with the chosen orientation of Σ, and let T be the positively oriented unit tangent along $\partial\Sigma$. Then for every smooth vector field X defined near Σ,

$$\int_\Sigma \langle \operatorname{curl} X, \nu \rangle \, dS = \int_{\partial\Sigma} \langle X, T \rangle \, ds.$$

Proof. (i) Consider the 2–form $\alpha := *X^\flat$ on $\overline{\Omega}$. By (6.31) we compute

$$d\alpha = d(*X^\flat) = (\partial_1 X^1 + \partial_2 X^2 + \partial_3 X^3) \, dx^1 \wedge dx^2 \wedge dx^3 = (\operatorname{div} X) \, dV.$$

Applying Stokes' theorem to α gives

$$\int_\Omega \operatorname{div} X \, dV = \int_\Omega d\alpha = \int_{\partial\Omega} \iota^*\alpha, \tag{6.32}$$

where $\iota : \partial\Omega \hookrightarrow \mathbb{R}^3$ is the inclusion and $\partial\Omega$ is oriented by the induced boundary orientation.

It remains to identify $\iota^*\alpha$ with $\langle X, \nu \rangle \, dS$. Fix $p \in \partial\Omega$ and choose an oriented orthonormal basis (e_1, e_2) of $T_p(\partial\Omega)$ such that (ν, e_1, e_2) is an oriented orthonormal basis of $\mathbb{R}^3$. By the defining property (6.30) with $k = 1$,

$$\iota^*\alpha(e_1, e_2) = (*X^\flat)(e_1, e_2) = dV(X, e_1, e_2),$$

because for a 1–form β one has $(*\beta)(u,v) = dV(\beta^\sharp, u, v)$, and here $\beta = X^\flat$. Since (ν, e_1, e_2) is oriented orthonormal, we have

$$dV(X, e_1, e_2) = \langle X, \nu\rangle\, dV(\nu, e_1, e_2) = \langle X, \nu\rangle.$$

On the other hand, the area form dS on $\partial\Omega$ is characterized by $dS(e_1, e_2) = 1$ for every oriented orthonormal basis (e_1, e_2) of $T_p(\partial\Omega)$. Therefore

$$\iota^*\alpha = \langle X, \nu\rangle\, dS,$$

and (6.32) becomes exactly the divergence theorem.

(ii) Let $\Sigma \subset \mathbb{R}^3$ be an oriented smooth surface and define the 1–form $\beta := X^\flat$ along a neighborhood of Σ. Stokes' theorem on Σ gives

$$\int_\Sigma d\beta = \int_{\partial\Sigma} \iota^*\beta, \tag{6.33}$$

where now $\iota : \partial\Sigma \hookrightarrow \Sigma$ is the inclusion.

We claim that

$$d\beta|_{T\Sigma} = \langle \operatorname{curl} X, \nu\rangle\, dS, \qquad \iota^*\beta = \langle X, T\rangle\, ds.$$

The second identity is immediate: along $\partial\Sigma$ the 1–form β evaluated on the positively oriented unit tangent T satisfies

$$\iota^*\beta(T) = \beta(T) = \langle X, T\rangle,$$

so $\iota^*\beta = \langle X, T\rangle\, ds$ by the characterization of arclength ds via $ds(T) = 1$.

For the first identity, fix $p \in \Sigma$ and choose an oriented orthonormal basis (e_1, e_2) of $T_p\Sigma$ such that (e_1, e_2, ν) is an oriented orthonormal basis of $\mathbb{R}^3$. Using the Euclidean identity

$$(\operatorname{curl} X)^\flat = *\, d(X^\flat)$$

(which can be verified by a direct coordinate computation in $\mathbb{R}^3$), we obtain

$$\langle \operatorname{curl} X, \nu\rangle = (\operatorname{curl} X)^\flat(\nu) = (*\, d\beta)(\nu).$$

By the defining property of the Hodge star for 2–forms (or equivalently by the explicit formulas above), we have

$$(*\, d\beta)(\nu) = d\beta(e_1, e_2).$$

Since $dS(e_1, e_2) = 1$, this means exactly that

$$d\beta|_{T\Sigma} = \langle \operatorname{curl} X, \nu\rangle\, dS.$$

Substituting these identifications into (6.33) yields

$$\int_\Sigma \langle \operatorname{curl} X, \nu\rangle\, dS = \int_{\partial\Sigma} \langle X, T\rangle\, ds,$$

which is the classical Stokes formula. ∎

Chapter 7

de Rham Cohomology and Degree

In this final chapter we introduce de Rham cohomology, compute a number of basic examples, and end with the degree of a smooth map between compact oriented manifolds. The central point is that the identity $d^2 = 0$ turns differential forms into a cochain complex, so global topology can be detected by closed forms modulo exact forms.

7.1 de Rham cohomology

We now turn to the last topic of this course: *de Rham cohomology.* The guiding idea is very simple:

- the exterior derivative satisfies $d^2 = 0$,
- therefore differential forms form a chain complex,
- and we can measure globally when a closed form fails to be exact.

This produces cohomology groups which are invariants of the manifold.

7.1.1 The identity $d^2 = 0$

Recall that the exterior derivative

$$d : \Omega^k(M) \to \Omega^{k+1}(M)$$

was defined locally in coordinates and then shown to be independent of the choice of charts. A basic property is that applying d twice gives zero.

Proposition 7.1. *For every smooth manifold M and every $k \geq 0$ one has*

$$d \circ d = 0 \qquad \text{on } \Omega^k(M).$$

Proof. Since the statement is local, it suffices to prove it in a coordinate chart on an open set $U \subset \mathbb{R}^n$ with coordinates $(x^1, \dots, x^n)$.

Write $\omega \in \Omega^k(U)$ as

$$\omega = \sum_I a_I \, dx^{i_1} \wedge \cdots \wedge dx^{i_k},$$

where $I = (i_1 < \cdots < i_k)$ runs over increasing multi–indices and $a_I \in C^\infty(U)$. Using $d(dx^j) = 0$ and the Leibniz rule for d, we obtain

$$d\omega = \sum_I da_I \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

Applying d once more,

$$d(d\omega) = \sum_I d(da_I) \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

Thus it suffices to show $d(da) = 0$ for a smooth function a. Write $da = \sum_{j=1}^n \frac{\partial a}{\partial x^j}\, dx^j$, then

$$d(da) = \sum_{j=1}^n d\Big(\frac{\partial a}{\partial x^j}\Big) \wedge dx^j = \sum_{j=1}^n \sum_{\ell=1}^n \frac{\partial^2 a}{\partial x^\ell \partial x^j}\, dx^\ell \wedge dx^j.$$

Split the sum into $\ell < j$ and $\ell > j$ and use $dx^\ell \wedge dx^j = -dx^j \wedge dx^\ell$. Since mixed second partial derivatives commute,

$$\frac{\partial^2 a}{\partial x^\ell \partial x^j} = \frac{\partial^2 a}{\partial x^j \partial x^\ell},$$

the terms cancel in pairs, hence $d(da) = 0$. Therefore $d(d\omega) = 0$ on U, and since d is defined chartwise this proves $d^2 = 0$ on M. □

7.1.2 Definition of de Rham cohomology and naturality

Because $d^2 = 0$, the spaces $\Omega^k(M)$ form a cochain complex

$$0 \longrightarrow \Omega^0(M) \xrightarrow{d} \Omega^1(M) \xrightarrow{d} \cdots \xrightarrow{d} \Omega^n(M) \longrightarrow 0.$$

This leads to the following definition.

Definition 7.2 (Closed and exact forms)**.** A form $\omega \in \Omega^k(M)$ is called *closed* if $d\omega = 0$. It is called *exact* if there exists $\eta \in \Omega^{k-1}(M)$ such that $\omega = d\eta$.

Definition 7.3 (de Rham cohomology)**.** The k–th de Rham cohomology group of M is

$$H^k_{\mathrm{dR}}(M) := \frac{\{\omega \in \Omega^k(M) : d\omega = 0\}}{\{d\eta : \eta \in \Omega^{k-1}(M)\}}.$$

We write $[\omega]$ for the class of a closed form ω.

Proposition 7.4 (Naturality)**.** *Let $f : M \to N$ be a smooth map. Then pullback induces a well–defined linear map*

$$f^* : H^k_{\mathrm{dR}}(N) \to H^k_{\mathrm{dR}}(M), \qquad [\omega] \mapsto [f^*\omega].$$

Proof. If $d\omega = 0$ then $d(f^*\omega) = f^*(d\omega) = 0$, so $f^*\omega$ is closed. If $\omega = d\eta$ then $f^*\omega = f^*(d\eta) = d(f^*\eta)$ is exact. Therefore the assignment $[\omega] \mapsto [f^*\omega]$ depends only on the cohomology class of ω and is well defined. □

7.1.3 Homotopy invariance

A fundamental property of de Rham cohomology is that it is invariant under smooth homotopies.

Theorem 7.5 (Homotopy invariance)**.** *Let $f_0, f_1 : M \to N$ be smooth maps and suppose there exists a smooth homotopy $F : M \times [0,1] \to N$ with $F(\cdot,0) = f_0$ and $F(\cdot,1) = f_1$. Then for every k,*

$$f_0^* = f_1^* : H^k_{\mathrm{dR}}(N) \to H^k_{\mathrm{dR}}(M).$$

Proof. Fix a closed form $\omega \in \Omega^k(N)$ and consider $F^*\omega \in \Omega^k(M \times [0,1])$. Write coordinates on $[0,1]$ by t. Any k–form on $M \times [0,1]$ can be written uniquely as

$$F^*\omega = \alpha(x,t) + dt \wedge \beta(x,t),$$

where for each t the form $\alpha(\cdot,t)$ is a k–form on M and $\beta(\cdot,t)$ is a $(k-1)$–form on M.

Compute $d(F^*\omega) = 0$ (since $d\omega = 0$ and pullback commutes with d). Using $d = d_M + dt \wedge \frac{\partial}{\partial t}$ on product manifolds, one obtains

$$0 = d(F^*\omega) = d_M\alpha + dt \wedge \Big(\frac{\partial \alpha}{\partial t} - d_M\beta\Big).$$

Hence

$$d_M\alpha(\cdot,t) = 0 \quad \text{and} \quad \frac{\partial \alpha}{\partial t} = d_M\beta.$$

Integrating the second identity in t yields

$$\alpha(\cdot,1) - \alpha(\cdot,0) = d_M\Big(\int_0^1 \beta(\cdot,t)\,dt\Big).$$

Finally, note that $\alpha(\cdot,0) = f_0^*\omega$ and $\alpha(\cdot,1) = f_1^*\omega$, because restricting F to $t = 0$ or $t = 1$ gives f_0 or f_1. Therefore

$$f_1^*\omega - f_0^*\omega = d\Big(\int_0^1 \beta(\cdot,t)\,dt\Big),$$

so $f_0^*\omega$ and $f_1^*\omega$ differ by an exact form and define the same class in $H^k_{\mathrm{dR}}(M)$. This proves $f_0^* = f_1^*$. □

Remark 7.6. There is a more conceptual proof using the Lie derivative and Cartan's magic formula

$$\mathcal{L}_X = d\,\iota_X + \iota_X\,d,$$

applied to the vector field $\partial/\partial t$ on $M \times [0,1]$. We will return to this viewpoint soon.

7.1.4 Poincaré lemma and the cohomology of $\mathbb{R}^n$

The Poincaré lemma explains why Euclidean space has no cohomology in positive degrees.

Theorem 7.7 (Poincaré lemma)**.** *Let $U \subset \mathbb{R}^n$ be star–shaped (in particular, contractible). Then every closed k–form on U is exact for $k \geq 1$. Equivalently,*

$$H^0_{\mathrm{dR}}(U) \cong \mathbb{R}, \qquad H^k_{\mathrm{dR}}(U) = 0 \ \ \textit{for } k \geq 1.$$

Remark 7.8. Use the straight–line contraction of U to a point.

7.1.5 The cohomology of S^1

We compute de Rham cohomology of the circle using its description $S^1 = \mathbb{R}/\mathbb{Z}$.

Proposition 7.9. *One has*

$$H^0_{\mathrm{dR}}(S^1) \cong \mathbb{R}, \qquad H^1_{\mathrm{dR}}(S^1) \cong \mathbb{R}, \qquad H^k_{\mathrm{dR}}(S^1) = 0 \ \ \textit{for } k \geq 2.$$

Proof. Since S^1 is connected, the only closed 0–forms are constant functions, hence $H^0_{\mathrm{dR}}(S^1) \cong \mathbb{R}$.

For H^1, view $S^1 = \mathbb{R}/\mathbb{Z}$ with coordinate x on $\mathbb{R}$. A smooth 1–form on S^1 corresponds to a $\mathbb{Z}$–periodic smooth function f on $\mathbb{R}$ via

$$\alpha = f(x)\,dx.$$

Every 1–form is automatically closed (since $d\alpha$ is a 2–form on a 1–manifold). The form α is exact if and only if there exists a smooth $\mathbb{Z}$–periodic function u with $du = u'(x)\,dx = \alpha$, i.e. $u' = f$.

If u is periodic then integrating $u' = f$ over one period gives

$$\int_0^1 f(x)\,dx = u(1) - u(0) = 0.$$

Conversely, if $\int_0^1 f(x)\,dx = 0$, define

$$u(x) := \int_0^x f(t)\,dt.$$

Then $u'(x) = f(x)$, and $u(x+1) - u(x) = \int_x^{x+1} f(t)\,dt = \int_0^1 f(t)\,dt = 0$, so u is periodic and $\alpha = du$ is exact.

Thus exact 1–forms are precisely those with $\int_{S^1} \alpha = 0$, and the map

$$H^1_{\mathrm{dR}}(S^1) \to \mathbb{R}, \qquad [\alpha] \mapsto \int_{S^1} \alpha$$

is a well–defined linear isomorphism. In particular, the class of dx is a generator and $H^1_{\mathrm{dR}}(S^1) \cong \mathbb{R}$. Finally, $H^k_{\mathrm{dR}}(S^1) = 0$ for $k \geq 2$ since $\Omega^k(S^1) = 0$ for $k \geq 2$. □

7.1.6 The cohomology of S^n

We now compute the cohomology groups of the sphere. First note that $\Omega^k(S^n) = 0$ for $k > n$, hence $H^k_{\mathrm{dR}}(S^n) = 0$ for $k > n$.

To treat the remaining degrees, we use the standard cover of S^n by two contractible open sets. Let U be the complement of the north pole and V the complement of the south pole. By stereographic projection, both U and V are diffeomorphic to $\mathbb{R}^n$, hence by the Poincaré lemma

$$H^k_{\mathrm{dR}}(U) = H^k_{\mathrm{dR}}(V) = 0 \quad \text{for } k \geq 1.$$

Moreover, $U \cap V$ is diffeomorphic to $S^{n-1} \times \mathbb{R}$ and hence homotopy equivalent to S^{n-1}.

Theorem 7.10. *For $n \geq 1$ one has*

$$H^0_{\mathrm{dR}}(S^n) \cong \mathbb{R}, \qquad H^n_{\mathrm{dR}}(S^n) \cong \mathbb{R}, \qquad H^k_{\mathrm{dR}}(S^n) = 0 \ \ \textit{for } 0 < k < n.$$

Proof. First, $H^0_{\mathrm{dR}}(S^n) \cong \mathbb{R}$ because S^n is connected.

Let $0 < k < n$ and let $\alpha \in \Omega^k(S^n)$ be closed. Since $H^k_{\mathrm{dR}}(U) = 0$ and $H^k_{\mathrm{dR}}(V) = 0$, there exist forms $u \in \Omega^{k-1}(U)$ and $v \in \Omega^{k-1}(V)$ such that

$$\alpha|_U = du, \qquad \alpha|_V = dv.$$

On the overlap $U \cap V$ we have $d(u - v) = 0$, so $u - v$ is a closed $(k-1)$–form on $U \cap V$. Since $U \cap V$ is homotopy equivalent to S^{n-1}, by induction on n we know that

$$H^{k-1}_{\mathrm{dR}}(U \cap V) = 0 \qquad \text{for } 1 \leq k-1 \leq n-2,$$

hence $u - v$ is exact: there exists $w \in \Omega^{k-2}(U \cap V)$ such that

$$u - v = dw \quad \text{on } U \cap V.$$

Choose a smooth cutoff function ρ on S^n such that $\rho \equiv 1$ on a neighborhood of the north pole and $\rho \equiv 0$ on a neighborhood of the south pole. In particular, ρ is supported in V and $1 - \rho$ is supported in U. Set

$$u' := u - d(\rho\, w) \ \text{ on } U, \qquad v' := v - d\big((\rho - 1)w\big) \ \text{ on } V.$$

Then $du' = du = \alpha|_U$ and $dv' = dv = \alpha|_V$, and one checks that u' and v' agree on $U \cap V$. Hence they glue to a global $(k-1)$–form θ on S^n with $d\theta = \alpha$. Thus α is exact and $H^k_{\mathrm{dR}}(S^n) = 0$ for $0 < k < n$.

Now consider $k = n$. The same argument shows that any closed n–form α is exact on U and on V, so $\alpha|_U = du$ and $\alpha|_V = dv$ for some $(n-1)$–forms u, v. On $U \cap V$ we have $d(u - v) = 0$, so $u - v$ is a closed $(n-1)$–form on $U \cap V \simeq S^{n-1} \times \mathbb{R}$. By induction, $H^{n-1}_{\mathrm{dR}}(U \cap V) \cong H^{n-1}_{\mathrm{dR}}(S^{n-1}) \cong \mathbb{R}$, so we can write

$$u - v = \lambda\, \eta + d\gamma \quad \text{on } U \cap V,$$

where $\lambda \in \mathbb{R}$, $\gamma \in \Omega^{n-2}(U \cap V)$, and η is a fixed closed $(n-1)$–form on $U \cap V$ whose cohomology class is nonzero. If $\lambda = 0$, then $u - v$ is exact and the previous gluing argument shows that α is exact on S^n. Therefore $H^n_{\mathrm{dR}}(S^n)$ has dimension at most 1.

Finally, $H^n_{\mathrm{dR}}(S^n)$ is nontrivial because S^n is orientable and hence admits a volume form ω. If ω were exact, $\omega = d\beta$, then Stokes' theorem on the closed manifold S^n would give

$$\int_{S^n} \omega = \int_{S^n} d\beta = 0,$$

contradicting $\int_{S^n} \omega \neq 0$. Hence $[\omega] \neq 0$ in $H^n_{\mathrm{dR}}(S^n)$, so $\dim H^n_{\mathrm{dR}}(S^n) = 1$ and $H^n_{\mathrm{dR}}(S^n) \cong \mathbb{R}$. □

Remark 7.11. The last step shows a general principle: if M is a compact oriented manifold without boundary, then the top–degree de Rham cohomology $H^n_{\mathrm{dR}}(M)$ is nontrivial, because a compatible volume form cannot be exact. Later we will see that $H^n_{\mathrm{dR}}(M)$ is in fact one–dimensional when M is also connected.

7.1.7 The case $M \times S^1$

In this subsection we compute the de Rham cohomology of $M \times S^1$ directly, without using any general Künneth theorem. The key inputs are:

- the decomposition of forms into a part involving dt and a part not involving dt,
- integration over the circle,
- and the homotopy formula from Theorem 7.5.

Write $S^1 = \mathbb{R}/\mathbb{Z}$ with coordinate $t \in [0, 1)$ and let

$$\theta := dt \in \Omega^1(S^1).$$

This is a globally defined closed 1–form on S^1.

Let $\pi_M : M \times S^1 \to M$ and $\pi_{S^1} : M \times S^1 \to S^1$ be the projections.

Decomposition of forms and a fiber integral.

Every k–form on $M \times S^1$ can be written uniquely as

$$\omega = \alpha(x, t) + dt \wedge \beta(x, t), \tag{7.1}$$

where for each fixed t the form $\alpha(\cdot, t)$ lies in $\Omega^k(M)$ and $\beta(\cdot, t)$ lies in $\Omega^{k-1}(M)$.

Define linear maps

$$A : \Omega^k(M \times S^1) \to \Omega^k(M), \qquad B : \Omega^k(M \times S^1) \to \Omega^{k-1}(M)$$

by

$$A(\omega) := \int_0^1 \alpha(\cdot, t)\, dt, \qquad B(\omega) := \int_0^1 \beta(\cdot, t)\, dt,$$

where α, β are as in (7.1). These are well defined because the decomposition is unique.

Lemma 7.12. *If ω is closed, then $A(\omega)$ is a closed k–form on M and $B(\omega)$ is a closed $(k-1)$–form on M. Moreover, if ω is exact, then $A(\omega)$ is exact and $B(\omega)$ is exact.*

Proof. Write $\omega = \alpha + dt \wedge \beta$. Using $d = d_M + dt \wedge \partial_t$ on the product and $d(dt) = 0$, we compute

$$d\omega = d_M\alpha + dt \wedge \big(\partial_t\alpha - d_M\beta\big).$$

If $d\omega = 0$, then $d_M\alpha = 0$ and $\partial_t\alpha = d_M\beta$. Integrating in t yields $d_M A(\omega) = A(d_M\alpha) = 0$ and

$$\alpha(\cdot, 1) - \alpha(\cdot, 0) = \int_0^1 \partial_t\alpha\, dt = \int_0^1 d_M\beta\, dt = d_M B(\omega).$$

Since $\alpha(\cdot, 1) = \alpha(\cdot, 0)$ (periodicity in t), we obtain $d_M B(\omega) = 0$ as well.

If $\omega = d\eta$, write $\eta = \mu + dt \wedge \nu$ with $\mu \in \Omega^{k-1}(M)$, $\nu \in \Omega^{k-2}(M)$ depending on t. Then

$$\omega = d\eta = d_M\mu + dt \wedge \big(\partial_t\mu - d_M\nu\big),$$

so $\alpha = d_M\mu$ and $\beta = \partial_t\mu - d_M\nu$. Thus

$$A(\omega) = \int_0^1 d_M\mu\, dt = d_M\Big(\int_0^1 \mu\, dt\Big)$$

is exact, and

$$B(\omega) = \int_0^1 (\partial_t \mu - d_M \nu)\, dt = \mu(\cdot, 1) - \mu(\cdot, 0) - d_M\Big(\int_0^1 \nu\, dt\Big) = -d_M\Big(\int_0^1 \nu\, dt\Big)$$

is exact. □

By the lemma, A and B descend to well–defined linear maps on cohomology:

$$A : H^k_{\mathrm{dR}}(M \times S^1) \to H^k_{\mathrm{dR}}(M), \qquad B : H^k_{\mathrm{dR}}(M \times S^1) \to H^{k-1}_{\mathrm{dR}}(M).$$

The main statement.

Theorem 7.13 (Künneth for $M \times S^1$)**.** *Let M be a smooth manifold and let $\theta = dt$ be the standard closed 1–form on $S^1 = \mathbb{R}/\mathbb{Z}$. For each $k \geq 0$ the map*

$$\Phi : \; H^k_{\mathrm{dR}}(M) \oplus H^{k-1}_{\mathrm{dR}}(M) \longrightarrow H^k_{\mathrm{dR}}(M \times S^1), \qquad ([u], [v]) \longmapsto [\pi_M^* u + \pi_M^* v \wedge \pi_{S^1}^* \theta]$$

is an isomorphism (with the convention $H^{-1}_{\mathrm{dR}}(M) = 0$).

Proof. Step 1: Φ is well defined. If u is replaced by $u + d\gamma$, then $\pi_M^*(u + d\gamma)$ differs from $\pi_M^* u$ by the exact form $d(\pi_M^* \gamma)$. If v is replaced by $v + d\delta$, then

$$\pi_M^*(v + d\delta) \wedge \pi_{S^1}^* \theta = \pi_M^* v \wedge \pi_{S^1}^* \theta + d(\pi_M^* \delta) \wedge \pi_{S^1}^* \theta = \pi_M^* v \wedge \pi_{S^1}^* \theta + d\big(\pi_M^* \delta \wedge \pi_{S^1}^* \theta\big),$$

since $d\theta = 0$. Hence the cohomology class of the image depends only on $[u]$ and $[v]$.

Step 2: Construct a candidate inverse. Given a cohomology class $[\omega] \in H^k_{\mathrm{dR}}(M \times S^1)$, choose a closed representative ω and write it as $\omega = \alpha + dt \wedge \beta$. Define

$$\Psi([\omega]) := \big([A(\omega)],\ [B(\omega)]\big) \in H^k_{\mathrm{dR}}(M) \oplus H^{k-1}_{\mathrm{dR}}(M).$$

This is well defined by Lemma 7.12.

Step 3: $\Psi \circ \Phi = \mathrm{id}$. Let u be a closed k–form on M and v a closed $(k-1)$–form on M. Then

$$\Phi([u], [v]) = [\pi_M^* u + \pi_M^* v \wedge dt],$$

so in the decomposition $\alpha = \pi_M^* u$ and $\beta = \pi_M^* v$ are independent of t. Hence $A = \int_0^1 \alpha\, dt = u$ and $B = \int_0^1 \beta\, dt = v$. Therefore $\Psi(\Phi([u], [v])) = ([u], [v])$.

Step 4: $\Phi \circ \Psi = \mathrm{id}$. Let ω be a closed k–form on $M \times S^1$ and write $\omega = \alpha + dt \wedge \beta$. Set

$$\bar{\alpha} := A(\omega) \in \Omega^k(M), \qquad \bar{\beta} := B(\omega) \in \Omega^{k-1}(M),$$

and consider the closed form

$$\omega_0 := \pi_M^* \bar{\alpha} + \pi_M^* \bar{\beta} \wedge dt.$$

We will show that $\omega - \omega_0$ is exact.

Define $H : M \times S^1 \times [0, 1] \to M \times S^1$ by

$$H(x, t, s) := (x,\, t + s),$$

where $t + s$ is taken modulo 1. Then $H(\cdot, \cdot, 0) = \mathrm{id}$ and $H(\cdot, \cdot, 1) = \mathrm{id}$, but the path $s \mapsto H(\cdot, \cdot, s)$ implements translation in the S^1–direction. Applying the homotopy formula from Theorem 7.5

to the closed form ω and the homotopy H, one obtains an explicit $(k-1)$–form $K(\omega)$ on $M \times S^1$ such that

$$\tau^*\omega - \omega = dK(\omega), \tag{7.2}$$

where $\tau(x,t) = (x, t+1)$ is the time–1 translation, hence the identity map. Averaging (7.2) over $s \in [0,1]$ yields that ω is cohomologous to its average in t: there exists $\eta \in \Omega^{k-1}(M \times S^1)$ such that

$$\omega - \int_0^1 (\mathrm{id} \times \tau_s)^* \omega \, ds = d\eta.$$

A direct computation of the average using the decomposition $\omega = \alpha + dt \wedge \beta$ shows that

$$\int_0^1 (\mathrm{id} \times \tau_s)^* \omega \, ds = \pi_M^* \bar{\alpha} + \pi_M^* \bar{\beta} \wedge dt = \omega_0.$$

Hence $\omega - \omega_0$ is exact, so $[\omega] = [\omega_0] = \Phi(\Psi([\omega]))$.

Combining Steps 3 and 4 shows that Φ and Ψ are inverse isomorphisms. □

Corollary 7.14. *Let $T^n = (S^1)^n = \mathbb{R}^n/\mathbb{Z}^n$ with periodic coordinates $(x^1, \ldots, x^n)$. Then for each $k = 0, 1, \ldots, n$ the de Rham cohomology group $H^k_{\mathrm{dR}}(T^n)$ is a real vector space of dimension $\binom{n}{k}$. More precisely, the cohomology classes*

$$\big[dx^{i_1} \wedge \cdots \wedge dx^{i_k}\big], \qquad 1 \le i_1 < \cdots < i_k \le n,$$

form a basis of $H^k_{\mathrm{dR}}(T^n)$. In particular,

$$H^0_{\mathrm{dR}}(T^n) \cong \mathbb{R}, \qquad H^1_{\mathrm{dR}}(T^n) \cong \mathbb{R}^n, \qquad H^n_{\mathrm{dR}}(T^n) \cong \mathbb{R},$$

and $H^k_{\mathrm{dR}}(T^n) = 0$ for $k > n$.

Proof. We argue by induction on n using Theorem 7.13.

For $n = 1$ this is exactly the computation of $H^*_{\mathrm{dR}}(S^1)$, with basis $[1]$ in degree 0 and $[dx^1]$ in degree 1.

Assume the statement holds for T^{n-1} and consider

$$T^n = T^{n-1} \times S^1$$

with coordinates $(x^1, \ldots, x^{n-1}, x^n)$. By Theorem 7.13, for each k there is an isomorphism

$$H^k_{\mathrm{dR}}(T^n) \cong H^k_{\mathrm{dR}}(T^{n-1}) \;\oplus\; H^{k-1}_{\mathrm{dR}}(T^{n-1}),$$

given explicitly by

$$([u],[v]) \longmapsto [\pi^* u + \pi^* v \wedge dx^n],$$

where $\pi : T^n \to T^{n-1}$ is the projection and we identify dx^n with the pullback of the standard 1–form on the last S^1 factor.

By the induction hypothesis, $H^k_{\mathrm{dR}}(T^{n-1})$ has basis $\big[dx^{i_1} \wedge \cdots \wedge dx^{i_k}\big]$ with $1 \le i_1 < \cdots < i_k \le n-1$, and $H^{k-1}_{\mathrm{dR}}(T^{n-1})$ has basis $\big[dx^{j_1} \wedge \cdots \wedge dx^{j_{k-1}}\big]$ with $1 \le j_1 < \cdots < j_{k-1} \le n-1$. Under the above isomorphism these basis elements map to the classes

$$\big[dx^{i_1} \wedge \cdots \wedge dx^{i_k}\big] \quad \text{and} \quad \big[dx^{j_1} \wedge \cdots \wedge dx^{j_{k-1}} \wedge dx^n\big],$$

which together are exactly the classes

$$\big[dx^{\ell_1} \wedge \cdots \wedge dx^{\ell_k}\big], \qquad 1 \le \ell_1 < \cdots < \ell_k \le n.$$

Hence they form a basis of $H^k_{\mathrm{dR}}(T^n)$. Counting them gives $\dim H^k_{\mathrm{dR}}(T^n) = \binom{n}{k}$. □

Remark 7.15. Theorem 7.13 is a special case of the general Künneth formula for de Rham cohomology, but the proof above uses only the homotopy formula and elementary manipulations with differential forms.

7.1.8 Top-degree cohomology of a compact connected manifold

We have computed the de Rham cohomology of the sphere and the torus. Next we prove a general result: for a compact connected *orientable* m–manifold, the top-degree de Rham cohomology is isomorphic to $\mathbb{R}$. The proof is elementary and reduces to a compactly supported construction in Euclidean space.

A divergence lemma on a cube.
Let $I^m = [0,1]^m \subset \mathbb{R}^m$.

Lemma 7.16. *Let $f \in C^\infty(\mathbb{R}^m)$ satisfy $\operatorname{supp}(f) \subset I^m = [0,1]^m$ and*

$$\int_{\mathbb{R}^m} f(x)\,dx = 0.$$

Then there exist smooth functions $f_1, \dots, f_m \in C^\infty(\mathbb{R}^m)$ such that $\operatorname{supp}(f_i) \subset I^m$ for each i, and

$$\sum_{i=1}^{m} \frac{\partial f_i}{\partial x^i} = f \qquad \text{on } \mathbb{R}^m.$$

Proof. We argue by induction on m.

Step 1: The case $m = 1$. Let $f \in C^\infty(\mathbb{R})$ be supported in $[0,1]$ and satisfy $\int_{\mathbb{R}} f(x)\,dx = 0$. Define

$$f_1(x) := \int_{-\infty}^{x} f(s)\,ds.$$

Then f_1 is smooth and $f_1'(x) = f(x)$. Moreover, for $x \le 0$ we have $f_1(x) = 0$, and for $x \ge 1$ we get

$$f_1(x) = \int_{-\infty}^{x} f(s)\,ds = \int_{-\infty}^{\infty} f(s)\,ds = 0,$$

so $\operatorname{supp}(f_1) \subset [0,1]$.

Step 2: Induction step. Assume the lemma holds in dimension $m-1$. Let $f \in C^\infty(\mathbb{R}^m)$ be supported in I^m and satisfy $\int_{\mathbb{R}^m} f\,dx = 0$.

Write $x = (x', x^m)$ with $x' \in \mathbb{R}^{m-1}$. Define

$$g(x') := \int_{\mathbb{R}} f(x', t)\,dt.$$

Then $g \in C^\infty(\mathbb{R}^{m-1})$ and $\operatorname{supp}(g) \subset I^{m-1}$. Moreover,

$$\int_{\mathbb{R}^{m-1}} g(x')\,dx' = \int_{\mathbb{R}^m} f(x)\,dx = 0.$$

By the induction hypothesis, there exist $h_1, \dots, h_{m-1} \in C^\infty(\mathbb{R}^{m-1})$, supported in I^{m-1}, such that

$$\sum_{i=1}^{m-1} \frac{\partial h_i}{\partial x^i} = g \qquad \text{on } \mathbb{R}^{m-1}.$$

Choose $\rho \in C^\infty(\mathbb{R})$ with $\operatorname{supp}(\rho) \subset [0,1]$ and

$$\int_{\mathbb{R}} \rho(t)\, dt = 1.$$

For $i = 1, \dots, m-1$ define

$$f_i(x', x^m) := h_i(x')\, \rho(x^m).$$

Then $\operatorname{supp}(f_i) \subset I^{m-1} \times [0,1] \subset I^m$, and

$$\sum_{i=1}^{m-1} \frac{\partial f_i}{\partial x^i}(x', x^m) = \Big(\sum_{i=1}^{m-1} \frac{\partial h_i}{\partial x^i}(x')\Big)\rho(x^m) = g(x')\, \rho(x^m).$$

Set

$$H(x', x^m) := f(x', x^m) - g(x')\, \rho(x^m).$$

Then $H \in C^\infty(\mathbb{R}^m)$, $\operatorname{supp}(H) \subset I^m$, and for every fixed x' we have

$$\int_{\mathbb{R}} H(x', t)\, dt = \int_{\mathbb{R}} f(x', t)\, dt - g(x') \int_{\mathbb{R}} \rho(t)\, dt = g(x') - g(x') = 0.$$

Define

$$f_m(x', x^m) := \int_{-\infty}^{x^m} H(x', t)\, dt.$$

Then f_m is smooth and satisfies

$$\frac{\partial f_m}{\partial x^m}(x', x^m) = H(x', x^m) = f(x', x^m) - g(x')\, \rho(x^m).$$

Moreover, since $\operatorname{supp}(H) \subset [0,1]$ in the x^m–variable, we have $f_m(x', x^m) = 0$ for $x^m \le 0$. For $x^m \ge 1$ we obtain

$$f_m(x', x^m) = \int_{-\infty}^{x^m} H(x', t)\, dt = \int_{-\infty}^{\infty} H(x', t)\, dt = 0,$$

so $\operatorname{supp}(f_m) \subset I^m$.

Finally,

$$\sum_{i=1}^{m} \frac{\partial f_i}{\partial x^i} = \Big(\sum_{i=1}^{m-1} \frac{\partial f_i}{\partial x^i}\Big) + \frac{\partial f_m}{\partial x^m} = g\,\rho + \big(f - g\,\rho\big) = f,$$

which completes the induction. □

A compactly supported primitive in top degree.

Corollary 7.17. *Let $\omega \in \Omega^m(\mathbb{R}^m)$ satisfy $\operatorname{supp}(\omega) \subset I^m$ and*

$$\int_{\mathbb{R}^m} \omega = 0.$$

Then there exists $\eta \in \Omega^{m-1}(\mathbb{R}^m)$ such that $\operatorname{supp}(\eta) \subset I^m$ and $d\eta = \omega$.

Proof. Write $\omega = f(x)\, dx^1 \wedge \cdots \wedge dx^m$ with $f \in C^\infty(\mathbb{R}^m)$. Then $\operatorname{supp}(f) \subset I^m$ and $\int_{\mathbb{R}^m} f\, dx = 0$. By Lemma 7.16 we can find $f_1, \ldots, f_m$ supported in I^m such that $\sum_i \partial_i f_i = f$.

Define

$$\eta := \sum_{i=1}^{m} (-1)^{i-1} f_i\, dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^m,$$

where the hat indicates that dx^i is omitted. Then η is smooth, supported in I^m, and a direct computation gives

$$d\eta = \Big(\sum_{i=1}^{m} \frac{\partial f_i}{\partial x^i}\Big)\, dx^1 \wedge \cdots \wedge dx^m = f\, dx^1 \wedge \cdots \wedge dx^m = \omega.$$

□

Moving the support of a top form along a chain of charts.

The next lemma says that, within a fixed cohomology class, we can move the support of a top-degree form from one coordinate neighborhood to another.

Lemma 7.18. *Let M^m be a connected smooth manifold without boundary, and let $U_0, U_\infty \subset M$ be coordinate neighborhoods. Assume in addition that both U_0 and U_∞ are diffeomorphic to the open cube $(0,1)^m$. Suppose $\omega_0 \in \Omega^m(M)$ satisfies $\operatorname{supp}(\omega_0) \subset U_0$. Then there exists $\eta \in \Omega^{m-1}(M)$ such that*

$$\operatorname{supp}(\omega_0 + d\eta) \subset U_\infty.$$

Proof. Choose coordinate neighborhoods $U_0, U_1, \ldots, U_N$ such that

$$U_{j-1} \cap U_j \neq \emptyset \quad \text{for all } j = 1, \ldots, N, \qquad U_N = U_\infty,$$

and each U_j is diffeomorphic to the open cube $(0,1)^m$. (This is possible since M is connected: first choose a chain of coordinate neighborhoods with consecutive overlaps, and then shrink each neighborhood so that it is still a coordinate neighborhood, still overlaps the next one, and its coordinate image is an open cube.)

We construct forms $\omega_j \in \Omega^m(M)$ inductively with

$$\operatorname{supp}(\omega_j) \subset U_{j-1} \cap U_j \quad \text{and} \quad \int_M \omega_j = \int_M \omega_{j-1}.$$

To do this, fix $j \geq 1$. Since $U_{j-1} \cap U_j$ is a nonempty coordinate neighborhood, we can choose a smooth bump function χ supported in $U_{j-1} \cap U_j$, not identically zero, and let Ω be the local coordinate volume form on that chart. Then $\chi\, \Omega$ is a compactly supported m–form with $\int_M \chi\, \Omega \neq 0$. Hence we may pick a constant $c_j \in \mathbb{R}$ so that

$$\omega_j := c_j\, \chi\, \Omega \quad \text{satisfies} \quad \int_M \omega_j = \int_M \omega_{j-1}.$$

Now $\omega_{j-1} - \omega_j$ has compact support contained in $U_{j-1} \cup U_j$ and satisfies $\int_M (\omega_{j-1} - \omega_j) = 0$.

Moreover, by construction $\operatorname{supp}(\omega_{j-1} - \omega_j)$ is contained in the coordinate neighborhood U_{j-1} (since ω_{j-1} is supported in $U_{j-2} \cap U_{j-1} \subset U_{j-1}$ and ω_j is supported in $U_{j-1} \cap U_j \subset U_{j-1}$). Thus we may identify U_{j-1} with an open set in $\mathbb{R}^m$. After multiplying by a cutoff function that equals 1 on $\mathrm{supp}(\omega_{j-1} - \omega_j)$ and has support in a cube contained in the chart, we can reduce to Corollary 7.17 in $\mathbb{R}^m$ and obtain $\eta_j \in \Omega^{m-1}(M)$ supported in U_{j-1} such that

$$d\eta_j = \omega_j - \omega_{j-1}.$$

Summing over $j = 1, \dots, N$ and setting $\eta := \sum_{j=1}^N \eta_j$, we get

$$\omega_0 + d\eta = \omega_N,$$

and $\mathrm{supp}(\omega_N) \subset U_{N-1} \cap U_N \subset U_\infty$. □

Localizing a top form to a fixed chart.

Proposition 7.19. *Let M^m be a compact connected smooth manifold without boundary, and let $U \subset M$ be a coordinate neighborhood. Then for every $\omega \in \Omega^m(M)$ there exists $\eta \in \Omega^{m-1}(M)$ such that*

$$\mathrm{supp}(\omega + d\eta) \subset U.$$

Proof. Choose finitely many coordinate neighborhoods $V_1, \dots, V_N$ covering M and a partition of unity $\{\rho_i\}_{i=1}^N$ subordinate to this cover. Then $\omega = \sum_{i=1}^N \rho_i\omega$, and each $\rho_i\omega$ has support contained in V_i.

Fix i. Apply Lemma 7.18 with $U_0 = V_i$ and $U_\infty = U$ to the form $\omega_0 := \rho_i\omega$. We obtain $\eta_i \in \Omega^{m-1}(M)$ such that $\mathrm{supp}(\rho_i\omega + d\eta_i) \subset U$. Summing over i and setting $\eta := \sum_{i=1}^N \eta_i$, we get

$$\omega + d\eta = \sum_{i=1}^N (\rho_i\omega + d\eta_i),$$

and the right-hand side is supported in U. □

Top-degree cohomology of a compact oriented manifold.

Assume now that M^m is compact, connected, and oriented. Then the integral $\int_M \omega$ is defined for $\omega \in \Omega^m(M)$ and vanishes on exact forms. Hence we obtain a linear map

$$I : \ H^m_{\mathrm{dR}}(M) \longrightarrow \mathbb{R}, \qquad I([\omega]) = \int_M \omega.$$

Theorem 7.20. *Let M^m be a compact connected oriented smooth manifold without boundary. Then I is an isomorphism. In particular,*

$$H^m_{\mathrm{dR}}(M) \cong \mathbb{R}.$$

Proof. Step 1: Surjectivity. Choose a volume form $\mu \in \Omega^m(M)$ with $\int_M \mu \neq 0$ (e.g. any positive volume form). Then $I([\mu]) = \int_M \mu$, so I is surjective.

Step 2: Injectivity. Let $[\omega] \in H^m_{\mathrm{dR}}(M)$ satisfy $I([\omega]) = \int_M \omega = 0$. Fix a coordinate neighborhood $U \subset M$. By Proposition 7.19 there exists $\eta_0 \in \Omega^{m-1}(M)$ such that

$$\mathrm{supp}(\omega + d\eta_0) \subset U.$$

Set $\omega_1 := \omega + d\eta_0$. Then $\int_M \omega_1 = \int_M \omega = 0$ and ω_1 is supported in a coordinate chart. Choose a smaller coordinate cube inside that chart containing $\mathrm{supp}(\omega_1)$. By Corollary 7.17 (transported to the chart) there exists $\eta_1 \in \Omega^{m-1}(M)$ supported in U such that $d\eta_1 = \omega_1$. Hence $\omega = d(\eta_1 - \eta_0)$ is exact, so $[\omega] = 0$.

Therefore I is injective, and hence an isomorphism. □

Remark 7.21 (Top-degree cohomology without orientability)**.** Proposition 7.19 does *not* assume that M is oriented. In particular, for any compact connected smooth manifold M^m (possibly nonorientable) and any fixed coordinate neighborhood $U \subset M$, we may define a linear map

$$J: \; H^m_{\mathrm{dR}}(M) \longrightarrow \mathbb{R}$$

as follows. Given $[\omega] \in H^m_{\mathrm{dR}}(M)$, choose $\eta \in \Omega^{m-1}(M)$ such that $\operatorname{supp}(\omega + d\eta) \subset U$ and set

$$J([\omega]) := \int_U (\omega + d\eta),$$

where the integral is taken by transporting $\omega + d\eta$ to an open subset of $\mathbb{R}^m$ via the chosen chart and using the standard integration on $\mathbb{R}^m$. (The value is independent of all choices: changing η alters $\omega + d\eta$ by an exact form supported in U, whose integral over U vanishes.)

Exactly as in the proof of Theorem 7.20, the map J is injective: if $J([\omega]) = 0$, then after localizing the support into a coordinate cube one concludes that ω is exact by Corollary 7.17.

If M is oriented, one checks that $J = I$ (hence J is an isomorphism). If M is *not* orientable, then $H^m_{\mathrm{dR}}(M) = 0$. Indeed, a nowhere vanishing m–form would determine an orientation, so on a nonorientable manifold every m–form must vanish somewhere; in particular, there is no volume form. Since J is injective and $J([\omega])$ can be computed locally in a chart, it follows that the only possible top-degree class is the zero class.

Degree of a map between compact oriented manifolds.

Let M^n and N^n be connected, compact, oriented smooth manifolds without boundary, and let $f : M \to N$ be a smooth map.

Definition 7.22 (Degree)**.** There exists a unique real number $\deg(f) \in \mathbb{R}$ such that for every $\omega \in \Omega^n(N)$,

$$\int_M f^*\omega \;=\; \deg(f) \int_N \omega. \tag{7.3}$$

This number is called the *degree* of f.

Existence and uniqueness. Choose $\theta \in \Omega^n(N)$ with $\int_N \theta = 1$ and set

$$\deg(f) := \int_M f^*\theta.$$

Given any $\omega \in \Omega^n(N)$, let $a := \int_N \omega$. By Theorem 7.20, we have $[\omega] = a[\theta]$ in $H^n_{\mathrm{dR}}(N)$, hence $\omega - a\theta = d\beta$ for some $\beta \in \Omega^{n-1}(N)$. Therefore

$$\int_M f^*\omega = \int_M f^*(a\theta + d\beta) = a \int_M f^*\theta + \int_M d(f^*\beta) = a \deg(f),$$

and since $a = \int_N \omega$, (7.3) follows.

Uniqueness is immediate: if $\int_M f^*\omega = k \int_N \omega$ holds for all ω, then taking $\omega = \theta$ gives $k = \int_M f^*\theta = \deg(f)$. □

Theorem 7.23. *Let $f : M^n \to N^n$ be smooth between connected compact oriented n–manifolds, and let $q \in N$ be a regular value of f. Then $f^{-1}(q)$ is a finite set and*

$$\deg(f) = \sum_{x \in f^{-1}(q)} \operatorname{sgn}(x),$$

where $\operatorname{sgn}(x) = +1$ *if* $df_x : T_xM \to T_qN$ *is orientation preserving, and* $\operatorname{sgn}(x) = -1$ *otherwise.*

Proof. If $f^{-1}(q) = \emptyset$, choose $\omega \in \Omega^n_c(N \setminus f(M))$. Then $f^*\omega = 0$, hence $\deg(f) \int_N \omega = \int_M f^*\omega = 0$. Picking ω with $\int_N \omega \neq 0$ forces $\deg(f) = 0$.

Assume now $f^{-1}(q) = \{x_1, \ldots, x_m\}$. Since q is a regular value, each x_i is a regular point. By the inverse function theorem, there exists a connected coordinate neighborhood $U \ni q$ and pairwise disjoint neighborhoods $V_i \ni x_i$ such that $f|_{V_i} : V_i \to U$ is a diffeomorphism for each i, and

$$f^{-1}(U) = \bigsqcup_{i=1}^{m} V_i.$$

Choose $\omega \in \Omega^n_c(U)$ with $\int_U \omega = 1$. Then

$$\deg(f) = \int_M f^*\omega = \sum_{i=1}^{m} \int_{V_i} f^*\omega.$$

If $f|_{V_i}$ preserves orientation, then by change of variables $\int_{V_i} f^*\omega = \int_U \omega = 1$; if it reverses orientation, then $\int_{V_i} f^*\omega = -\int_U \omega = -1$. Hence $\deg(f) = \sum_i \operatorname{sgn}(x_i)$. □

Remark 7.24 (Integrality and the mod 2 degree)**.** Theorem 7.23 shows in particular that $\deg(f) \in \mathbb{Z}$: for a regular value q it is a finite sum of ± 1. Moreover, reducing the identity

$$\deg(f) = \sum_{x \in f^{-1}(q)} \operatorname{sgn}(x)$$

modulo 2, we obtain

$$\deg(f) \equiv \# f^{-1}(q) \pmod 2,$$

which recovers the mod 2 degree defined earlier (counting the number of preimages of a regular value modulo 2) without any orientability assumptions.

Proposition 7.25 (Multiplicativity)**.** *If $f : M \to N$ and $g : N \to P$ are smooth maps between connected compact oriented n–manifolds, then*

$$\deg(g \circ f) = \deg(g) \deg(f).$$

Proof. For any $\omega \in \Omega^n(P)$ we have

$$\int_M (g \circ f)^*\omega = \int_M f^*(g^*\omega) = \deg(f) \int_N g^*\omega = \deg(f) \deg(g) \int_P \omega.$$

By uniqueness in Definition 7.22, this implies $\deg(g \circ f) = \deg(g) \deg(f)$. □

An application: the fundamental theorem of algebra.

Let $\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ be the extended complex plane. It is diffeomorphic to S^2 and hence a connected compact oriented smooth 2–manifold. Fix $k \geq 1$ and consider the polynomial map

$$f : \widehat{\mathbb{C}} \longrightarrow \widehat{\mathbb{C}}, \qquad f(z) = z^k + a_1 z^{k-1} + \cdots + a_k \quad (z \in \mathbb{C}), \qquad f(\infty) = \infty.$$

(Equivalently, in homogeneous coordinates $[Z : W] \in \mathbb{CP}^1$ one may write $f([Z : W]) = [P(Z, W) : W^k]$ where $P(Z, W) = Z^k + a_1 Z^{k-1} W + \cdots + a_k W^k$, from which smoothness at $\infty = [1 : 0]$ is immediate.)

Proposition 7.26. *The map $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ has degree $\deg(f) = k$.*

Proof. On $\mathbb{C} \cong \mathbb{R}^2$ we use the standard orientation determined by $dx \wedge dy$. Let $h : U \subset \mathbb{C} \to \mathbb{C}$ be holomorphic and let $p \in U$ with $h'(p) \neq 0$. Writing $h = u + iv$ in real and imaginary parts, the Cauchy–Riemann equations give at p

$$Dh_p = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \begin{pmatrix} u_x & u_y \\ -u_y & u_x \end{pmatrix}.$$

Hence

$$\det(Dh_p) = u_x^2 + u_y^2 = |h'(p)|^2 > 0,$$

so Dh_p is orientation preserving.

Consider first $f_0(z) = z^k$. Let $q \in \mathbb{C} \setminus \{0\}$ be a regular value of f_0. Then $f_0^{-1}(q)$ consists of the k distinct k–th roots of q. On $\mathbb{C}$ the map $z \mapsto z^k$ is holomorphic, hence orientation preserving at every point where $df_0 \neq 0$; for $q \neq 0$ all preimages satisfy $df_0 \neq 0$. Therefore Theorem 7.23 yields $\deg(f_0) = k$.

Now consider the homotopy of maps $f_t(z) = z^k + t(a_1 z^{k-1} + \cdots + a_k)$ on $\widehat{\mathbb{C}}$ (with $f_t(\infty) = \infty$). One shows that $\deg(f_t)$ is constant in t. Indeed, for any $\omega \in \Omega^n(N)$ the form $f_t^*\omega$ represents the cohomology class $f_t^*[\omega] \in H^n_{\mathrm{dR}}(M)$, and integration over M depends only on this cohomology class. Since de Rham cohomology is homotopy invariant, the map $f_t^* : H^n_{\mathrm{dR}}(N) \to H^n_{\mathrm{dR}}(M)$ does not depend on t. Consequently $\deg(f_t)$ is independent of t, and hence $\deg(f) = \deg(f_1) = \deg(f_0) = k$. □