Manifolds

Bruno Martelli

Dipartimento di Matematica, Largo Pontecorvo 5, 56127 Pisa, Italy
E-mail address: `martelli at dm dot unipi dot it`

version: February 25, 2019

Contents

Introduction

CHAPTER 1

Preliminaries

We state here some basic notions of topology and analysis that we will use in this book. The proofs of some theorems are omitted and can be found in many excellent sources.

1.1. General topology

1.1.1. Topological spaces. A *topological space* is a pair (X, τ) where X is a set and τ is a collection of subsets of X called *open subsets*, satisfying the following axioms:

- $\varnothing$ and X are open subsets;
- the arbitrary union of open subsets is an open subset;
- the finite intersection of open subsets is an open subset.

The complement $X \setminus U$ of an open subset $U \in \tau$ is called *closed*. When we denote a topological space, we often write X instead of (X, τ) for simplicity.

A *neighbourhood* of a point $x \in X$ is any subset $N \subset X$ containing an open set U that contains x, that is $x \in U \subset N \subset X$.

1.1.2. Examples. There are many ways to construct topological spaces and we summarise them here very briefly.

Metric spaces. Every metric space (X, d) is also naturally a topological space: by definition, a subset $U \subset X$ is open $\iff$ for every $x_0 \in U$ there is an $r > 0$ such that the open ball

$$B(x_0, r) = \{x \in X \mid d(x, x_0) < r\}$$

is entirely contained in U.

In particular $\mathbb{R}^n$ is a topological space, whose topology is induced by the euclidean distance between points.

Product topology. The cartesian product $X = \prod_{i \in I} X_i$ of two or more topological spaces is a topological space: by definition, a subset $U \subset X$ is open $\iff$ it is a (possibly infinite) union of products $\prod_{i \in I} U_i$ of open subsets $U_i \subset X_i$, where $U_i \neq X_i$ only for finitely many i.

This is the coarsest topology (that is, the topology with the fewest open sets) on X such that the projections $X \to X_i$ are all continuous.

Subspace topology. Every subset $S \subset X$ of a topological space X is also naturally a topological space: by definition a subset $U \subset S$ is open $\iff$ there is an open subset $V \subset X$ such that $U = V \cap S$.

This is the coarsest topology on S such that the inclusion $i : S \hookrightarrow X$ is continuous.

In particular every subset $S \subset \mathbb{R}^n$ is naturally a topological space.

Quotient topology. Let $f : X \to Y$ be a surjective map. A topology on X induces one on Y as follows: by definition a set $U \subset Y$ is open $\iff$ its counterimage $f^{-1}(U)$ is open in X.

This is the finest topology (that is, the one with the most open subsets) on Y such that the map $f : X \to Y$ is continuous.

A typical situation is when Y is the quotient space $Y = X/_\sim$ for some equivalence relation $\sim$ on X, and $X \to Y$ is the induced projection.

1.1.3. Continuous maps. A map $f : X \to Y$ between topological spaces is *continuous* if the inverse of every open subset of Y is an open subset of X. The map f is a *homeomorphism* if it has an inverse $f^{-1} : Y \to X$ which is also continuous.

Two topological spaces X and Y are *homeomorphic* if there is a homeomorphism $f : X \to Y$ relating them. Being homeomorphic is clearly an equivalence relation.

1.1.4. Reasonable assumptions. A topological space can be very wild, but most of the spaces encountered in this book will satisfy some reasonable assumptions, that we now list.

Hausdorff. A topological space X is *Hausdorff* if every two distinct points $x, y \in X$ have disjoint open neighbourhoods U_x and U_y, that is $U_x \cap U_y = \varnothing$.

The euclidean space $\mathbb{R}^n$ is Hausdorff. Products and subspaces of Hausdorff spaces are also Hausdorff.

Second-countable. A *base* for a topological space X is a set of open subsets $\{U_i\}$ such that every open set is an arbitrary union of these. A topological space X is *second-countable* if it has a countable base.

The euclidean space $\mathbb{R}^n$ is second-countable. Countable products and subspaces of second-countable spaces are also second-countable.

Connected. A topological space X is *connected* if it is not the disjoint union $X = X_1 \sqcup X_2$ of two non-empty open subsets X_1, X_2. Every topological space X is partitioned canonically into maximal connected subsets, called *connected components*. Given this canonical decomposition, it is typically harmless to restrict our attention to connected spaces.

A slightly stronger notion is that of path-connectedness. A space X is *path-connected* if for every $x, y \in X$ there is a *path* connecting them, that is a continuous map $\alpha\colon [0,1] \to X$ with $\alpha(0) = x$ and $\alpha(1) = y$. Every path-connected space is connected. The converse is also true if one assumes the reasonable assumption that the topological space we are considering is *locally path-connected*, that is every point has a path-connected neighbourhood.

The Euclidean space $\mathbb{R}^n$ is path-connected. Products and quotients of (path-)connected spaces are (path-)connected.

Locally compact. A topological space X is *locally compact* if every point $x \in X$ has a compact neighbourhood. The euclidean space $\mathbb{R}^n$ is locally compact.

1.1.5. Reasonable consequences. The reasonable assumptions listed in the previous section have some nice and reasonable consequences.

Countable base with compact closure. We first note the following.

Proposition 1.1.1. *If a topological space X is Hausdorff and locally compact, every $x \in X$ has an open neighbourhood $U(x)$ with compact closure.*

Proof. Every $x \in X$ has a compact neighbourhood $V(x)$, that is closed since X is Hausdorff. The neighbourhood $V(x)$ contains an open neighbourhood $U(x)$ of x, whose closure is contained in $V(x)$ and hence compact. $\square$

Proposition 1.1.2. *Every locally compact second-countable Hausdorff space X has a countable base made of open sets with compact closure.*

Proof. Let $\{U_i\}$ be a countable base. For every open set $U \subset X$ and $x \in U$, there is an open neighbourhood $U(x) \subset U$ of x with compact closure, which contains a U_i that contains x. Therefore the U_i with compact closure suffice as a base for X. $\square$

Exhaustion by compact sets. Let X be a topological space. An *exhaustion by compact subsets* is a countable family $K_1, K_2, \ldots$ of compact subsets such that $K_i \subset \operatorname{int}(K_{i+1})$ for all i and $\cup_i K_i = X$.

The standard example is the exhaustion of $\mathbb{R}^n$ by closed balls

$$K_i = \overline{B(0,i)} = \left\{ x \in \mathbb{R}^n \;\middle|\; \|x\| \leq i \right\}.$$

Proposition 1.1.3. *Every locally compact second-countable Hausdorff space X has an exhaustion by compact subsets.*

Proof. The space X has a countable base $U_1, U_2, \ldots$ of open sets with compact closures. Define $K_1 = \overline{U_1}$ and

$$K_{i+1} = \overline{U_1} \cup \ldots \cup \overline{U_k}$$

where k is the smallest natural number such that $K_i \subset \cup_{j=1}^k U_j$. $\square$

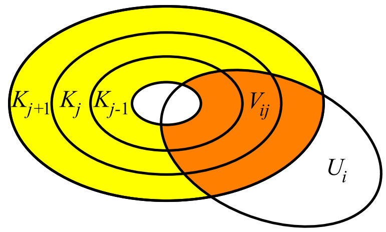

Figure 1.1. A locally compact second-countable Hausdorff space is paracompact: how to construct a locally finite refinement using an exhaustion by compact subsets.

Paracompactness. An *open cover* for a topological space X is a set $\{U_i\}$ of open sets whose union is the whole of X. An open cover $\{U_i\}$ is *locally finite* if every point in X has a neighbourhood that intersects only finitely many U_i. A *refinement* of an open cover $\{U_i\}$ is another open cover $\{V_j\}$ such that every V_j is contained in some U_i.

Definition 1.1.4. A topological space X is *paracompact* if every open cover $\{U_i\}$ has a locally finite refinement $\{V_j\}$.

Of course a compact space is paracompact, but the class of paracompact spaces is much larger.

Proposition 1.1.5. *Every locally compact second-countable Hausdorff space X is paracompact.*

Proof. Let $\{U_i\}$ be an open covering: we now prove that there is a locally finite refinement. We know that X has an exhaustion by compact subsets $\{K_j\}$, and we set $K_0 = K_{-1} = \varnothing$. For every i, j we define $V_{ij} = \big(\mathrm{int}(K_{j+1}) \setminus K_{j-2}\big) \cap U_i$ as in Figure 1.1. The family $\{V_{ij}\}$ is an open cover and a refinement of $\{U_i\}$, but it may not be locally finite.

For every fixed $j = 1, 2, \ldots$ only finitely many V_{ij} suffice to cover the compact set $K_j \setminus \mathrm{int}(K_{j-1})$, so we remove all the others. The resulting refinement $\{V_{ij}\}$ is now locally finite. $\square$

In particular the Euclidean space $\mathbb{R}^n$ is paracompact, and more generally every subspace $X \subset \mathbb{R}^n$ is paracompact. The reason for being interested in paracompactness may probably sound obscure at this point, and it will be unveiled in the next chapters.

1.1.6. Topological manifolds. Recall that the *open unit ball* in $\mathbb{R}^n$ is

$$B^n = \big\{ x \in \mathbb{R}^n \ \big|\ \|x\| < 1 \big\}.$$

A *topological manifold* of dimension n is a reasonable topological space locally modelled on B^n.

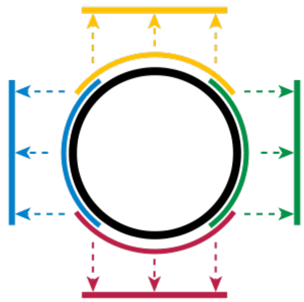

Figure 1.2. A topological manifold is covered by open subsets, each homeomorphic to B^n. Here the manifold is a circle, and is covered by four open arcs, each homeomorphic to the open interval B^1.

Definition 1.1.6. A *topological manifold* of dimension n (shortly, a *topological n-manifold*) is a Hausdorff second-countable topological space M such that every point x has an open neighbourhood U_x homeomorphic to B^n.

In other words, a Hausdorff second-countable topological space M is a manifold $\iff$ it has an open covering $\{U_i\}$ such that each U_i is homeomorphic to B^n. A schematic picture in Figure 1.2 shows that the circle is a topological 1-manifold: a more rigorous proof will be given in the next chapters.

Example 1.1.7. Every open subset of $\mathbb{R}^n$ is a topological n-manifold. In general, any open subset of a topological n-manifold is a topological n-manifold.

1.1.7. Pathologies. The two reasonability hypothesis in Definition 1.1.6 are there only to discard some spaces that are usually considered as pathological. Here are two examples. The impressionable reader may skip this section.

Exercise 1.1.8 (The double point). Consider two parallel lines $Y = \{y = \pm 1\} \subset \mathbb{R}^2$ and their quotient $X = Y/_\sim$ where $(x, y) \sim (x', y') \iff x = x'$ and ($y = y'$ or $x \neq 0$). Prove that every point in X has an open neighbourhood homeomorphic to B^1, but X is not Hausdorff.

The following is particularly crazy.

Exercise 1.1.9 (The long ray). Let α be an ordinal, and consider $X = \alpha \times [0, 1)$ with the lexicographic order. Remove from X the first element $(0, 0)$, and give X the *order topology*, having the intervals $(a, b) = \{a < x < b\}$ as a base. If α is countable, then X is homeomorphic to $\mathbb{R}$. If $\alpha = \omega_1$ is the first non countable ordinal, then X is the *long ray*: every point in X has an open neighbourhood homeomorphic to B^1, but X is not separable (it contains no countable dense subset) and hence is not second-countable. However, the long ray X is path-connected!

1.1.8. Homotopy. Let X and Y be two topological spaces. A *homotopy* between two continuous maps $f, g\colon X \to Y$ is another continuous map $F\colon X \times$

$[0,1] \to Y$ such that $F(\cdot, 0) = f$ and $F(\cdot, 1) = g$. Two maps f and g are *homotopic* if there is a homotopy between them, and we may write $f \sim g$.

Two topological spaces X and Y are *homotopically equivalent* if there are two continuous maps $f: X \to Y$ and $g: Y \to X$ such that $f \circ g \sim \mathrm{id}_Y$ and $g \circ f \sim \mathrm{id}_X$.

Two homeomorphic spaces are homotopically equivalent, but the converse may not hold. For instance, the euclidean space $\mathbb{R}^n$ is homotopically equivalent to a point for every n. A topological space that is homotopically equivalent to a point is called *contractible*.

1.2. Algebraic topology

1.2.1. Fundamental group. Let X be a topological space and $x_0 \in X$ a base point. The *fundamental group* of the pair (X, x_0) is a group

$$\pi_1(X, x_0)$$

defined by taking all *loops*, that is all paths starting and ending at x_0, considered up to homotopies with fixed endpoints. Loops may be concatenated, and this operation gives a group structure to $\pi_1(X, x_0)$.

If x_1 is another base point, every arc from x_0 to x_1 defines an isomorphism between $\pi_1(X, x_0)$ and $\pi_1(X, x_1)$. Therefore if X is path-connected the fundamental group is base point independent, at least up to isomorphisms, and we write it as $\pi_1(X)$. If $\pi_1(X)$ is trivial we say that X is *simply connected*.

Every continuous map $f: X \to Y$ between topological spaces induces a homomorphism

$$f_*: \pi_1(X, x_0) \longrightarrow \pi_1\big(Y, f(x_0)\big).$$

The transformation from f to f_* is a *functor* from the category of pointed topological spaces to that of groups. This means that $(f \circ g)_* = f_* \circ g_*$ and $(\mathrm{id}_X)_* = \mathrm{id}_{\pi_1(X, x_0)}$. It implies in particular that homeomorphic spaces have isomorphic fundamental groups.

Exercise 1.2.1. Every topological connected manifold M has a countable fundamental group.

Hint. Since M is second countable, we may find an open covering of M that consists of countably many open sets homeomorphic to open balls called *islands*. Every pair of such sets intersect in an open set that has at most countably many connected components called *bridges*. Every loop in $\pi_1(M, x_0)$ may be determined by a (non unique!) finite sequence of symbols saying which islands and bridges it crosses. There are only countably many sequences. $\square$

Two maps $f, g: (X, x_0) \to (Y, y_0)$ that are homotopic, via a homotopy that sends x_0 to y_0 at each time, induce the same homomorphisms $f_* =$

g_* on fundamental groups. This implies that homotopically equivalent path-connected spaces have isomorphic fundamental groups, so in particular every contractible topological space is simply connected.

There are simply connected manifolds that are not contractible, as we will discover in the next chapters.

1.2.2. Coverings. Let $\tilde{X}$ and X be two path-connected topological spaces. A continuous surjective map $p\colon \tilde{X} \to X$ is a *covering map* if every $x \in X$ has an open neighbourhood U such that

$$p^{-1}(U) = \bigsqcup_{i \in I} U_i$$

where U_i is open and $p|_{U_i}\colon U_i \to U$ is a homeomorphism for all $i \in I$.

A *local homeomorphism* is a continuous map $f\colon X \to Y$ where every $x \in X$ has an open neighbourhood U such that $f(U)$ is open and $f|_U\colon U \to f(U)$ is a homeomorphism. A covering map is always a local homeomorphism, but the converse may not hold.

The *degree* of a covering $p\colon \tilde{X} \to X$ is the cardinality of a fibre $p^{-1}(x)$ of a point x, a number which does not depend on x.

Two coverings $p\colon \tilde{X} \to X$ and $p'\colon \tilde{X}' \to X$ of the same space X are *isomorphic* if there is a homeomorphism $f\colon \tilde{X} \to \tilde{X}'$ such that $p = p' \circ f$.

1.2.3. Coverings and fundamental group. One of the most beautiful aspects of algebraic topology is the exceptionally strong connection between fundamental groups and covering maps.

Let $p\colon \tilde{X} \to X$ be a covering map. We fix a basepoint $x_0 \in X$ and a lift $\tilde{x}_0 \in p^{-1}(x_0)$ in the fibre of x_0. The induced homomorphism

$$p_*\colon \pi_1(\tilde{X}, \tilde{x}_0) \longrightarrow \pi_1(X, x_0)$$

is always injective. If we modify $\tilde{x}_0$ in the fibre of x_0, the image subgroup $\operatorname{Im} p_*$ changes only by a conjugation inside $\pi_1(X, x_0)$. The degree of p equals the index of $\operatorname{Im} p_*$ in $\pi_1(X, x_0)$.

A topological space Y is *locally contractible* if every point $y \in Y$ has a contractible neighbourhood. This is again a very reasonable assumption: every topological space considered in this book will be of this kind.

We now consider a connected and locally contractible topological space X and fix a base-point $x_0 \in X$.

Theorem 1.2.2. *By sending p to $\operatorname{Im} p_*$ we get a bijective correspondence*

$$\left\{ \begin{array}{c} \text{coverings } p\colon \tilde{X} \to X \\ \text{up to isomorphism} \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \text{subgroups of } \pi_1(X, x_0) \\ \text{up to conjugacy} \end{array} \right\}$$

The covering corresponding to the trivial subgroup is called the *universal covering*. In other words, a covering $\tilde{X} \to X$ is *universal* if $\tilde{X}$ is simply connected, and we have just discovered that this covering is unique up to isomorphism.

Exercise 1.2.3. Let $p\colon \tilde{X} \to X$ be a covering map. If X is a topological manifold, then $\tilde{X}$ also is.

Hint. To lift the second countability from X to $\tilde{X}$, use that $\pi_1(X)$ is countable by Exercise 1.2.1 and hence p has countable degree. □

1.2.4. Deck transformations. Let $p\colon \tilde{X} \to X$ be a covering map. A *deck transformation* or *automorphism* for p is a homeomorphism $f\colon \tilde{X} \to \tilde{X}$ such that $p \circ f = p$. The deck transformations form a group $\mathrm{Aut}(p)$ called the *deck transformation group* of p.

If $\mathrm{Im}\, p_*$ is a normal subgroup, the covering map is called *regular*. For instance, the universal cover is regular. Regular covering maps behave nicely in many aspects: for instance we have a natural isomorphism

$$\mathrm{Aut}(p) \cong \pi_1(X)/_{\pi_1(\tilde{X})}.$$

To be more specific, we need to recall some basic notions on group actions.

1.2.5. Group actions. An *action* of a group G on a set X is a group homomorphism

$$\rho\colon G \longrightarrow S(X)$$

where $S(X)$ is the group of all the bijections $X \to X$. We denote $\rho(g)$ simply by g, and hence write $g(x)$ instead of $\rho(g)(x)$. We note that

$$g(h(x)) = (gh)(x), \qquad e(x) = x$$

for every $g, h \in G$ and $x \in X$. In particular if $g(x) = y$ then $g^{-1}(y) = x$.

The *stabiliser* of a point $x \in X$ is the subgroup $G_x < G$ consisting of all the elements g such that $g(x) = x$. The *orbit* of a point $x \in X$ is the subset

$$O(x) = \big\{ g(x) \bigm| g \in G \big\} \subset X.$$

Exercise 1.2.4. We have $x \in O(x)$. Two orbits $O(x)$ and $O(y)$ either coincide or are disjoint. They coincide $\Longleftrightarrow \exists g \in G$ such that $g(x) = y$.

Therefore the set X is partitioned into orbits. The action is:

- *transitive* if for every $x, y \in X$ there is a $g \in G$ such that $g(x) = y$;
- *faithful* if ρ is injective;
- *free* if the stabiliser of every point is trivial, that is $g(x) \neq x$ for every $x \in X$ and every non-trivial $g \in G$.

Exercise 1.2.5. The stabilisers G_x and G_y of two points x, y lying in the same orbit are conjugate subgroups of G.

Exercise 1.2.6. There is a natural bijection between the left cosets of G_x in G and the elements of $O(x)$. In particular the cardinality of $O(x)$ equals the index $[G : G_x]$ of G_x in G.

The space of all the orbits is denoted by $X/_G$. We have a natural projection $\pi\colon X \to X/_G$.

1.2.6. Topological actions. If X is a topological space, a *topological action* of a group G on X is a homomorphism

$$G \longrightarrow \mathrm{Homeo}(X)$$

where $\mathrm{Homeo}(X)$ is the group of all the self-homeomorphisms of X. We have a natural projection $\pi\colon X \to X/_G$ and we equip the quotient set $X/_G$ with the quotient topology. The action is:

- *properly discontinuous* if any two points $x, y \in X$ have neighbourhoods U_x and U_y such that the set
$$\big\{ g \in G \;\big|\; g(U_x) \cap U_y \neq \varnothing \big\}$$
is finite.

Example 1.2.7. The action of a finite group G is always properly discontinuous.

This definition is relevant mainly because of the following remarkable fact.

Proposition 1.2.8. *Let G act on a Hausdorff path-connected space X. The following are equivalent:*

(1) *G acts freely and properly discontinuously;*
(2) *the quotient $X/_G$ is Hausdorff and $X \to X/_G$ is a regular covering.*

Every regular covering between Hausdorff path-connected spaces arises in this way.

Concerning the last sentence: if $\tilde{X} \to X$ is a regular covering, the deck transformation group G acts transitively on each fibre, and we get $X = \tilde{X}/_G$. This does not hold for non-regular coverings.

We have here a formidable and universal tool to construct plenty of regular coverings and of topological spaces: it suffices to have X and a group G acting freely and properly discontinously on it.

Since every universal cover is regular, we also get the following.

Corollary 1.2.9. *Every path-connected locally contractible Hausdorff topological space X is the quotient $\tilde{X}/_G$ of its universal cover by the action of some group G acting freely and properly discontinuously.*

Note that the group G is isomorphic to $\pi_1(X)$. There are plenty of examples of this phenomenon, but in this introductory chapter we limit ourselves to a very basic one. More will come later.

Example 1.2.10. Let $G = \mathbb{Z}$ act on $X = \mathbb{R}$ as translations, that is $g(v) = v + g$. The action is free and properly discontinuous; hence we get a covering $\mathbb{R} \to \mathbb{R}/_{\mathbb{Z}}$. The quotient $\mathbb{R}/_{\mathbb{Z}}$ is in fact homeomorphic to S^1 (exercise).

In principle, one could now think of classifying all the (locally contractible, path-connected, Hausdorff) topological spaces by looking only at the simply connected ones and then studying the groups acting freely and properly discontinuously on them. It is of course impossible to carry on this too ambitious strategy in this wide generality, but the task becomes more reasonable if one restricts the attention to spaces of some particular kind like – as we will see – the riemannian manifolds having constant curvature.

Recall that a continuous map $f: X \to Y$ is *proper* if $f^{-1}(K)$ is compact for every compact $K \subset Y$.

Exercise 1.2.11. Let a group G act on a locally compact space X. Assign to G the discrete topology. The following are equivalent:

- the action is properly discontinuous;
- for every compact $K \subset X$, the set $\{g \mid g(K) \cap K \neq \emptyset\}$ is finite;
- the map $G \times X \to X \times X$ that sends (g, x) to $(g(x), x)$ is proper.

1.3. Multivariable analysis

1.3.1. Smooth maps. A map $f: U \to V$ between two open sets $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ is C^∞ or *smooth* if it has partial derivatives of any order. All the maps considered in this book will be smooth.

In particular, for every $p \in U$ we have a *differential*

$$df_p: \mathbb{R}^n \longmapsto \mathbb{R}^m$$

which is the linear map that best approximates f near p, that is we get

$$f(x) = f(p) + df_p(x-p) + o\big(\|x-p\|\big).$$

If we see df_p as a $m \times n$ matrix, it is called the *Jacobian* and we get

$$df_p = \left(\frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_n}\right) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

A fundamental property of differentials is the *chain rule*: if we are given two smooth functions

$$U \xrightarrow{f} V \xrightarrow{g} W$$

then for every $p \in U$ we have

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

1.3.2. Taylor theorem. A *multi-index* is a vector $\alpha = (\alpha_1, \ldots, \alpha_n)$ of non-negative integers $\alpha_i \geq 0$. We set

$$|\alpha| = \alpha_1 + \ldots + \alpha_n, \qquad \alpha! = \alpha_1! \cdots \alpha_n!, \qquad x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}.$$

Let $f: U \to \mathbb{R}$ be a smooth map defined on some open set $U \subset \mathbb{R}^n$. For every multi-index α we define the corresponding combination of partial derivatives:

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}.$$

We recall Taylor's Theorem:

Theorem 1.3.1. *Let $f: U \to \mathbb{R}$ be a smooth map defined on some open convex set $U \subset \mathbb{R}^n$. For every point $x_0 \in U$ and integer $k \geq 0$ we have*

$$f(x) = \sum_{|\alpha| \leq k} \frac{D^\alpha f(x_0)}{\alpha!}(x - x_0)^\alpha + \sum_{|\alpha| = k+1} h_\alpha(x)(x - x_0)^\alpha$$

where $h_\alpha: U \to \mathbb{R}$ is a smooth map that depends on α.

1.3.3. Diffeomorphisms. A smooth map $f: U \to V$ between two open sets $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ is a *diffeomorphism* if it is invertible and its inverse $f^{-1}: V \to U$ is also smooth.

Proposition 1.3.2. *If f is a diffeomorphism, then df_p is invertible for every $p \in U$. In particular we get $n = m$.*

Proof. The chain rule gives

$$\mathrm{id}_{\mathbb{R}^n} = d(\mathrm{id}_U)_p = d(f^{-1} \circ f)_p = df^{-1}_{f(p)} \circ df_p,$$

$$\mathrm{id}_{\mathbb{R}^m} = d(\mathrm{id}_V)_{f(p)} = d(f \circ f^{-1})_{f(p)} = df_p \circ df^{-1}_{f(p)}.$$

Therefore the linear map df_p is invertible. □

We now show that a weak converse of this statement holds.

1.3.4. Local diffeomorphisms. We say that a smooth map $f: U \to V$ is a *local diffeomorphism* at a point $p \in U$ if there is an open neighbourhood $U' \subset U$ of p such that $f(U')$ is open and $f|_{U'}: U' \to f(U')$ is a diffeomorphism.

Here is an important theorem, that we will use frequently.

Theorem 1.3.3 (Inverse Function Theorem). *A smooth map $f: U \to V$ is a local diffeomorphism at $p \in U \iff$ its differential df_p is invertible.*

We say that a smooth map $f: U \to V$ is a *local diffeomorphism* if it is so at every point $p \in U$. A diffeomorphism is always a local diffeomorphism, but the converse does not hold as the following example shows.

Example 1.3.4. The smooth map $f: \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e^x \cos y \\ e^x \sin y \end{pmatrix}$$

has Jacobian

$$df_{(x,y)} = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}$$

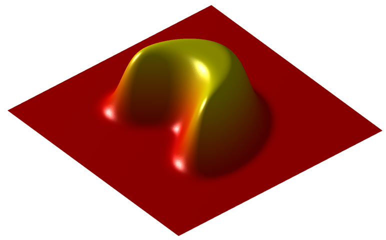

Figure 1.3. A smooth bump function $f\colon \mathbb{R}^2 \to \mathbb{R}$.

with determinant e^{2x} and hence everywhere invertible. By the Inverse Function Theorem, the map f is a local diffeomorphism. The map f is however not injective, hence it is not a diffeomorphism.

1.3.5. Bump functions. A *smooth bump function* is a smooth function $\rho\colon \mathbb{R}^n \to \mathbb{R}$ that has compact support (the *support* is the closure of the set of points $x \in \mathbb{R}^n$ where $\rho(x) \neq 0$). See Figure 1.3.

The existence of bump functions is a peculiar feature of the smooth environment that has many important consequences in differential topology. The main tool is the smooth function

$$h(t) = \begin{cases} e^{-\frac{1}{t}} & \text{if } t \geq 0, \\ 0 & \text{if } t \leq 0. \end{cases}$$

We may use it to build a bump function $\rho\colon \mathbb{R}^n \to \mathbb{R}$ as follows:

$$\rho(x) = h\big(1 - \|x\|^2\big).$$

The support of ρ is the closed unit disc $\|x\| \leq 1$, and it has a unique maximum at the origin $x = 0$.

Note that a bump function is never analytic (unless it is constantly zero). Sometimes it is useful to have a bump function that looks like a *plateau*, for instance consider $\eta\colon \mathbb{R}^n \to \mathbb{R}$ defined as follows:

$$\eta(x) = \frac{h\big(1 - \|x\|^2\big)}{h\big(1 - \|x\|^2\big) + h\big(\|x\|^2 - \frac{1}{4}\big)}.$$

Here $\eta(x) = 1$ for all $\|x\| \leq \frac{1}{2}$ and $\eta(x) = 0$ for all $\|x\| \geq 1$, while $\eta(x) \in (0, 1)$ for all $\frac{1}{2} < \|x\| < 1$.

1.3.6. Transition function. Another important smooth non-analytic functions is the *transition function* $\Psi\colon \mathbb{R} \to \mathbb{R}$ defined as

$$\Psi(x) = \frac{h(x)}{h(x) + h(1 - x)}$$

where $h(x)$ is the function defined above. The function Ψ is smooth and non-decreasing, and we have $\Psi(x) = 0$ for all $x \leq 0$ and $\Psi(x) = 1$ for all $x \geq 1$. See Figure 1.4.

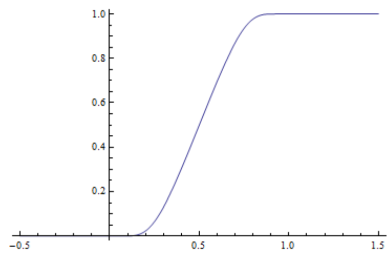

Figure 1.4. A smooth transition function Ψ.

1.3.7. Cauchy–Lipschitz theorem. The Cauchy–Lipschitz Theorem certifies the existence and uniqueness of solutions of a system of first-order differential equations, and also the smooth dependence on its initial values, when appropriate hypothesis are satisfied. Here is the version that we will use here.

Theorem 1.3.5. *Given a smooth function* $g\colon \mathbb{R}^n \to \mathbb{R}^n$, *there is a number* $\varepsilon > 0$ *and a unique smooth map*

$$f\colon B^n \times (-\varepsilon, \varepsilon) \to \mathbb{R}^n$$

such that

$$\begin{aligned} f(x,0) &= x, \\ \frac{\partial f}{\partial t}(x,t) &= g\big(f(x,t)\big). \end{aligned}$$

Uniqueness here means that if we get another ε' and another f' then $f(x,t) = f'(x,t)$ for all $x \in B^n$ and $|t| < \min\{\varepsilon, \varepsilon'\}$.

1.3.8. Integration. A *Borel set* $V \subset \mathbb{R}^n$ is any subset constructed from the open and closed sets by countable unions and intersections.

If $V \subset \mathbb{R}^n$ is a Borel set and $f\colon V \to \mathbb{R}$ is a non-negative measurable function, we may consider its *Lebesgue integral*

$$\int_V f.$$

If $\varphi\colon U \to V$ is a diffeomorphism between two open subsets of $\mathbb{R}^n$, then we get the following *changes of variables* formula

$$\int_{V'} f = \int_{U'} |\det d\varphi| f \circ \varphi$$

for any Borel subsets $U' \subset U$ and $V' = \varphi(U')$.

Remark 1.3.6. A diffeomorphism of course does not preserve the measure of Borel sets, but it sends zero-measure sets to zero-measure sets.

1.3.9. The Sard Lemma. Let $f: U \to \mathbb{R}^n$ be a smooth map defined on some open subset $U \subset \mathbb{R}^m$. We say that a point $p \in U$ is *regular* if the differential df_p is surjective, and *singular* otherwise. A value $q \in \mathbb{R}^n$ is a *regular value* if all its counterimages $p \in f^{-1}(q)$ are regular points, and *singular* otherwise.

Here is an important fact on smooth maps.

Lemma 1.3.7 (Sard's Lemma). *The singular values of f form a zero-measure subset of $\mathbb{R}^n$.*

Corollary 1.3.8. *If $m < n$, the image of f is a zero-measure subset.*

Recall that a *Peano curve* is a continuous surjection $\mathbb{R} \to \mathbb{R}^2$. Maps of this kind are forbidden in the smooth world.

1.3.10. Complex analysis. Let $U, V \subset \mathbb{C}$ be open subsets. Recall that a function $f: U \to V$ is *holomorphic* if for every $z_0 \in U$ the limit

$$f'(z_0) = \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

exists. The limit $f'(z_0)$ is a complex number called the *complex derivative* of f at z_0.

Quite surprisingly, a homolorphic function satisfies a wealth of very good properties: if we identify $\mathbb{C}$ with $\mathbb{R}^2$ in the usual way, we may interpret f as a function between open sets of $\mathbb{R}^2$, and it turns out that f is smooth (and even analytic) and its Jacobian at z_0 is such that

$$\det(df_{z_0}) = |f'(z_0)|^2.$$

It is indeed a remarkable fact that the presence of the complex derivative alone guarantees the existence of partial derivatives of any order.

1.4. Projective geometry

1.4.1. Projective spaces. Let $\mathbb{K}$ be any field: we will be essentially interested in the cases $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Let V be a finite-dimensional vector space on $\mathbb{K}$. The *projective space* of V is

$$\mathbb{P}(V) = \big(V \setminus \{0\}\big)/_{\sim}$$

where $v \sim w \iff v = \lambda w$ for some $\lambda \neq 0$. In particular we write

$$\mathbb{KP}^n = \mathbb{P}(\mathbb{K}^{n+1}).$$

Every non-zero vector $v = (x_0, \ldots, x_n) \in \mathbb{K}^{n+1}$ determines a point in $\mathbb{KP}^n$ which we denote as

$$[x_0, \ldots, x_n].$$

These are the *homogeneous coordinates* of the point. Of course not all the x_i are zero, and $[x_0, \ldots, x_n] = [\lambda x_0, \ldots, \lambda x_n]$ for all $\lambda \neq 0$.

1.4.2. Topology. When $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, the space $\mathbb{KP}^n$ inherits the quotient topology from $\mathbb{K}^{n+1}$ and is a Hausdorff compact connected topological space. A convenient way to see this is to consider the projections

$$\pi\colon S^n \longrightarrow \mathbb{RP}^n, \qquad \pi\colon S^{2n+1} \longrightarrow \mathbb{CP}^n$$

obtained by restricting the projections from $\mathbb{R}^n \setminus \{0\}$ and $\mathbb{C}^n \setminus \{0\}$. Note that

$$S^{2n+1} = \left\{ z \in \mathbb{C}^{n+1} \;\middle|\; |z_0|^2 + \ldots + |z_n|^2 = 1 \right\}.$$

Exercise 1.4.1. Show that the projections are surjective and deduce that the projective spaces are connected and compact.

Exercise 1.4.2. We have the following homeomorphisms

$$\mathbb{RP}^1 \cong S^1, \qquad \mathbb{CP}^1 \cong S^2.$$

The fundamental group of $\mathbb{RP}^n$ is $\mathbb{Z}$ when $n = 1$ and $\mathbb{Z}/_{2\mathbb{Z}}$ when $n > 1$. On the other hand the complex projective space $\mathbb{CP}^n$ is simply connected for every n.

CHAPTER 2

Tensors

2.1. Multilinear algebra

2.1.1. The dual space. In this book we will be concerned mostly with real finite-dimensional vector spaces. Given two such spaces V, W of dimension m, n, we denote by $\mathrm{Hom}(V, W)$ the set of all the linear maps $V \to W$. The set $\mathrm{Hom}(V, W)$ is itself naturally a vector space of dimension mn.

A space that will be quite relevant here is the *dual space* $V^* = \mathrm{Hom}(V, \mathbb{R})$, that consists of all the linear functionals $V \to \mathbb{R}$, also called *covectors*. The spaces V and V^* have the same dimension, but there is no canonical way to choose an isomorphism $V \to V^*$ between them: this fact will have important consequences in this book.

A basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ for V induces a *dual basis* $\mathcal{B}^* = \{\mathbf{v}^1, \ldots, \mathbf{v}^n\}$ for V^* by requiring that $\mathbf{v}^i(\mathbf{v}_j) = \delta_{ij}$. (Recall that the *Kronecker delta* δ_{ij} equals 1 if $i = j$ and 0 otherwise.) We can construct an isomorphism $V \to V^*$ by sending $\mathbf{v}_i$ to $\mathbf{v}^i$, but it heavily depends on the chosen basis $\mathcal{B}$.

On the other hand, a canonical isomorphism $V \to V^{**}$ exists between V and its *bidual space* $V^{**} = (V^*)^*$. The isomorphism is the following:

$$\mathbf{v} \longmapsto \big(\mathbf{v}^* \longmapsto \mathbf{v}^*(\mathbf{v})\big).$$

For that reason, the bidual space V^{**} will play no role here and will always be identified with V. In fact, it is useful to think of V and V^* as related by a bilinear pairing

$$V \times V^* \longrightarrow \mathbb{R}$$

that sends $(\mathbf{v}, \mathbf{v}^*)$ to $\mathbf{v}^*(\mathbf{v})$. Not only the vectors in V^* act on V, but also the vectors in V act on V^*.

Every linear map $L\colon V \to W$ induces an *adjoint* linear map $L^*\colon W^* \to V^*$ that sends f to $f \circ L$. Of course we get $L^{**} = L$.

2.1.2. Multilinear maps. Given some vector spaces $V_1, \ldots, V_k, W$, a map

$$F\colon V_1 \times \cdots \times V_k \longrightarrow W$$

is *multilinear* if it is linear on each component.

Let $\mathcal{B}_i = \{\mathbf{v}_{i,1}, \ldots, \mathbf{v}_{i,m_i}\}$ be a basis of V_i and $\mathcal{C} = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ a basis of W. The *coefficients* of F with respect to these basis are the numbers

$$F^j_{j_1,\ldots,j_k}$$

with $1 \leq j_i \leq m_i$ and $1 \leq j \leq n$ such that

$$F(\mathbf{v}_{1,j_1}, \ldots, \mathbf{v}_{k,j_k}) = \sum_{j=1}^{n} F^{j}_{j_1,\ldots,j_k} \mathbf{w}_j.$$

Exercise 2.1.1. Every multilinear F is determined by its coefficients, and every choice of coefficients determines a multilinear F.

We denote by $\mathrm{Mult}(V_1, \ldots, V_k; W)$ the space of all the multilinear maps $V_1 \times \cdots \times V_k \to W$. This is naturally a vector space.

Corollary 2.1.2. *We have*

$$\dim \mathrm{Mult}(V_1, \ldots, V_k; W) = \dim V_1 \cdots \dim V_k \dim W.$$

When $W = \mathbb{R}$ we omit it from the notation and write $\mathrm{Mult}(V_1, \ldots, V_k)$. In that case of course we have

$$\dim \mathrm{Mult}(V_1, \ldots, V_k) = \dim V_1 \cdots \dim V_k.$$

In fact, every space $\mathrm{Mult}(V_1, \ldots, V_k; W)$ may be transformed canonically into a similar one where the target vector space is $\mathbb{R}$, thanks to the following:

Exercise 2.1.3. There is a canonical isomorphism

$$\mathrm{Mult}(V_1, \ldots, V_k; W) \longrightarrow \mathrm{Mult}(V_1, \ldots, V_k, W^*)$$

defined by sending $F \in \mathrm{Mult}(V_1, \ldots, V_k; W)$ to the map

$$(\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{w}^*) \longmapsto \mathbf{w}^*\big(F(\mathbf{v}_1, \ldots, \mathbf{v}_k)\big).$$

Hint. The spaces have the same dimension and the map is injective. □

2.1.3. Sum and product of spaces. We now introduce a couple of operations $\oplus$ and $\otimes$ on vector spaces. Let $V_1, \ldots, V_k$ be some real finite-dimensional vector spaces.

Sum. The *sum* $V_1 \oplus \cdots \oplus V_k$ is just the cartesian product with componentwise vector space operations. That is:

$$V_1 \oplus \cdots \oplus V_k = \big\{(\mathbf{v}_1, \ldots, \mathbf{v}_k) \bigm| \mathbf{v}_1 \in V_1, \ldots, \mathbf{v}_k \in V_k\big\}$$

and the vector space operations are

$$(\mathbf{v}_1, \ldots, \mathbf{v}_k) + (\mathbf{w}_1, \ldots, \mathbf{w}_k) = (\mathbf{v}_1 + \mathbf{w}_1, \ldots, \mathbf{v}_k + \mathbf{w}_k),$$
$$\lambda(\mathbf{v}_1, \ldots, \mathbf{v}_k) = (\lambda\mathbf{v}_1, \ldots, \lambda\mathbf{v}_k).$$

Let $\mathcal{B}_i = \{\mathbf{v}_{i,1}, \ldots, \mathbf{v}_{i,m_i}\}$ be a basis of V_i, for all $i = 1, \ldots, k$.

Exercise 2.1.4. A basis for $V_1 \oplus \cdots \oplus V_k$ is

$$\big\{(\mathbf{v}_{1,j_1}, \mathbf{0}, \ldots, \mathbf{0}), \ldots, (\mathbf{0}, \ldots, \mathbf{0}, \mathbf{v}_{i,j_i}, \mathbf{0}, \ldots, \mathbf{0}), \ldots, (\mathbf{0}, \ldots, \mathbf{0}, \mathbf{v}_{k,j_k})\big\}$$

where $1 \leq j_i \leq m_i$ varies for each $i = 1, \ldots, k$.

We deduce that

$$\dim(V_1 \oplus \cdots \oplus V_k) = \dim V_1 + \ldots + \dim V_k.$$

Tensor product. The *tensor product* $V_1 \otimes \cdots \otimes V_k$ is defined (a bit more obscurely...) as the space of all the multilinear maps $V_1^* \times \cdots \times V_k^* \to \mathbb{R}$, *i.e.*

$$V_1 \otimes \cdots \otimes V_k = \operatorname{Mult}(V_1^*, \ldots, V_k^*).$$

We already know that

$$\dim(V_1 \otimes \cdots \otimes V_k) = \dim V_1 \cdots \dim V_k.$$

Any k vectors $\mathbf{v}_1 \in V_1, \ldots, \mathbf{v}_k \in V_k$ determine an element

$$\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_k \in V_1 \otimes \cdots \otimes V_k$$

which is by definition the multilinear map

$$(\mathbf{v}_1^*, \ldots, \mathbf{v}_k^*) \longmapsto \mathbf{v}_1^*(\mathbf{v}_1) \cdots \mathbf{v}_k^*(\mathbf{v}_k).$$

As opposite to the sum operation, it is important to note that *not* all the elements of $V_1 \otimes \cdots \otimes V_k$ are of the form $\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_k$. The elements of this type (sometimes called *pure* or *simple*) can however generate the space, as the next proposition shows. Let $\mathcal{B}_i = \{\mathbf{v}_{i,1}, \ldots, \mathbf{v}_{i,m_i}\}$ be a basis of V_i for all $1 \leq i \leq k$.

Proposition 2.1.5. *A basis for the tensor product $V_1 \otimes \cdots \otimes V_k$ is*

$$\{\mathbf{v}_{1,j_1} \otimes \cdots \otimes \mathbf{v}_{k,j_k}\}$$

where $1 \leq j_i \leq m_i$ varies for each $i = 1, \ldots, k$.

Proof. The number of elements is precisely the dimension $\dim V_1 \cdots \dim V_k$ of the space, hence we only need to show that they are independent. Let $\mathcal{B}^i = \{\mathbf{v}^{i,1}, \ldots, \mathbf{v}^{i,m_i}\}$ be the dual basis of $\mathcal{B}_i$. Suppose that

$$\sum_J \lambda_J \mathbf{v}_{1,j_1} \otimes \cdots \otimes \mathbf{v}_{k,j_k} = 0$$

where $J = (j_1, \ldots, j_k)$. By applying both members of the equation to the element $(\mathbf{v}^{1,j_1}, \ldots, \mathbf{v}^{k,j_k})$ we get $\lambda_J = 0$ where $J = (j_1, \ldots, j_k)$, and this for every multi-index J. $\square$

Example 2.1.6. A basis for $\mathbb{R}^2 \otimes \mathbb{R}^2$ is given by the elements

$$\begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix}, \quad \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix}, \quad \begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix}, \quad \begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix}.$$

Exercise 2.1.7. The following relations hold in $V \otimes W$:

$$(\mathbf{v} + \mathbf{v}') \otimes \mathbf{w} = \mathbf{v} \otimes \mathbf{w} + \mathbf{v}' \otimes \mathbf{w}, \qquad \mathbf{v} \otimes (\mathbf{w} + \mathbf{w}') = \mathbf{v} \otimes \mathbf{w} + \mathbf{v} \otimes \mathbf{w}',$$

$$\lambda(\mathbf{v} \otimes \mathbf{w}) = (\lambda \mathbf{v}) \otimes \mathbf{w} = \mathbf{v} \otimes (\lambda \mathbf{w}),$$

$$\mathbf{v} \otimes \mathbf{w} = \mathbf{0} \iff \mathbf{v} = \mathbf{0} \text{ or } \mathbf{w} = \mathbf{0}.$$

Exercise 2.1.8. Let $\mathbf{v}, \mathbf{v}' \in V$ and $\mathbf{w}, \mathbf{w}' \in W$ be non-zero vectors. If $\mathbf{v}$ and $\mathbf{v}'$ are independent, then $\mathbf{v} \otimes \mathbf{w}$ and $\mathbf{v}' \otimes \mathbf{w}'$ also are.

Exercise 2.1.9. Let $\mathbf{v}, \mathbf{v}' \in V$ and $\mathbf{w}, \mathbf{w}' \in W$ be two pairs of independent vectors. Show that

$$\mathbf{v} \otimes \mathbf{w} + \mathbf{v}' \otimes \mathbf{w}' \in V \otimes W$$

is not a pure element.

2.1.4. Canonical isomorphisms. We now introduce some canonical isomorphisms, that may look quite abstract at a first sight, but that will help us a lot to simplify many situations: two spaces that are canonically isomorphic may be harmlessly considered as the same space.

We start with the following easy:

Proposition 2.1.10. *The map* $\mathbf{v} \mapsto \mathbf{v} \otimes \mathbf{1}$ *defines a canonical isomorphism*

$$V \longrightarrow V \otimes \mathbb{R}.$$

Proof. The spaces have the same dimension and the map is linear and injective by Exercise 2.1.7. □

Let $V_1, \ldots, V_k, Z$ be any vector spaces.

Proposition 2.1.11. *There is a canonical isomorphism*

$$\mathrm{Mult}(V_1, \ldots, V_k; Z) \longrightarrow \mathrm{Hom}(V_1 \otimes \cdots \otimes V_k, Z)$$

defined by sending $F \in \mathrm{Mult}(V_1, \ldots, V_k; Z)$ *to the unique homomorphism* F' *that satisfies the relation*

$$F'(\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_k) = F(\mathbf{v}_1, \ldots, \mathbf{v}_k).$$

for every $\mathbf{v}_1 \in V_1, \ldots, \mathbf{v}_k \in V_k$.

Proof. It is easier to define the inverse map: every homomorphism $F' \in \mathrm{Hom}(V_1 \otimes \cdots \otimes V_k, Z)$ gives rise to an element $F \in \mathrm{Mult}(V_1, \ldots, V_k; Z)$ just by setting $F(\mathbf{v}_1, \ldots, \mathbf{v}_k) = F'(\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_k)$. This gives rise to a linear map

$$\mathrm{Hom}(V_1 \otimes \cdots \otimes V_k, Z) \longrightarrow \mathrm{Mult}(V_1, \ldots, V_k; Z)$$

between spaces of the same dimension. The map is injective (exercise: use Proposition 2.1.5), hence it is an isomorphism. □

This canonical isomorphism is called the *universal property* of $\otimes$ and one can also show that it characterises the tensor product uniquely. This is typically stated by drawing a commutative diagram like this:

(1)
$$\begin{tikzcd} V_1 \times \cdots \times V_k \arrow[r] \arrow[dr, "F"'] & V_1 \otimes \cdots \otimes V_k \arrow[d, dotted, "F'"] \\ & Z \end{tikzcd}$$

The universal property is very useful to construct maps. For instance, we may use it to construct more canonical isomorphisms:

Proposition 2.1.12. *There are canonical isomorphisms*

$$V \oplus W \cong W \oplus V, \qquad (V \oplus W) \oplus Z \cong V \oplus W \oplus Z \cong V \oplus (W \oplus Z),$$

$$V \otimes W \cong W \otimes V, \qquad (V \otimes W) \otimes Z \cong V \otimes W \otimes Z \cong V \otimes (W \otimes Z),$$

$$V \otimes (W \oplus Z) \cong (V \otimes W) \oplus (V \otimes Z),$$

$$(V_1 \oplus \cdots \oplus V_k)^* \cong V_1^* \oplus \cdots \oplus V_k^*, \qquad (V_1 \otimes \cdots \otimes V_k)^* \cong V_1^* \otimes \cdots \otimes V_k^*.$$

Proof. The isomorphisms in the first line are

$$(\mathbf{v}, \mathbf{w}) \mapsto (\mathbf{w}, \mathbf{v}), \qquad (\mathbf{v}, \mathbf{w}, \mathbf{z}) \mapsto \big((\mathbf{v}, \mathbf{w}), \mathbf{z}\big), \qquad (\mathbf{v}, \mathbf{w}, \mathbf{z}) \mapsto \big(\mathbf{v}, (\mathbf{w}, \mathbf{z})\big).$$

Those in the second line are uniquely determined by the conditions

$$\mathbf{v} \otimes \mathbf{w} \mapsto \mathbf{w} \otimes \mathbf{v}, \qquad \mathbf{v} \otimes \mathbf{w} \otimes \mathbf{z} \mapsto (\mathbf{v} \otimes \mathbf{w}) \otimes \mathbf{z}, \qquad \mathbf{v} \otimes \mathbf{w} \otimes \mathbf{z} \mapsto \mathbf{v} \otimes (\mathbf{w} \otimes \mathbf{z})$$

thanks to the universal property of the tensor products. Analogously the isomorphism of the third line is determined by

$$\mathbf{v} \otimes (\mathbf{w}, \mathbf{z}) \mapsto (\mathbf{v} \otimes \mathbf{w}, \mathbf{v} \otimes \mathbf{z}).$$

Concerning the last line, the first isomorphism is straightforward. For the second, we have

$$(V_1 \otimes \cdots \otimes V_k)^* = \mathrm{Hom}(V_1 \otimes \cdots \otimes V_k, \mathbb{R}) = \mathrm{Mult}(V_1, \ldots, V_k) = V_1^* \otimes \cdots \otimes V_k^*.$$

More concretely, every element $\mathbf{v}^1 \otimes \cdots \otimes \mathbf{v}^k \in V_1^* \otimes \cdots \otimes V_k^*$ is naturally an element of $(V_1 \otimes \cdots \otimes V_k)^*$ as follows:

$$(\mathbf{v}^1 \otimes \cdots \otimes \mathbf{v}^k)(\mathbf{w}_1 \otimes \cdots \otimes \mathbf{w}_k) = \mathbf{v}^1(\mathbf{w}_1) \cdots \mathbf{v}^k(\mathbf{w}_k).$$

The proof is complete. □

There are yet more canonical isomorphisms to discover! The following is a consequence of Exercise 2.1.3 and is particularly useful.

Corollary 2.1.13. *There is a canonical isomorphism*

$$\mathrm{Hom}(V, W) \cong V^* \otimes W.$$

In particular we have $\mathrm{End}(V) \cong V^* \otimes V = \mathrm{Mult}(V, V^*)$. In this canonical isomorphism, the identity endomorphism id_V corresponds to the bilinear map $V \times V^* \to \mathbb{R}$ that sends $(\mathbf{v}, \mathbf{v}^*)$ to $\mathbf{v}^*(\mathbf{v})$.

Exercise 2.1.14. Given $\mathbf{v}^* \in V^*$ and $\mathbf{w} \in W$, the element $\mathbf{v}^* \otimes \mathbf{w}$ corresponds via the canonical isomorphism $\mathrm{Hom}(V, W) \cong V^* \otimes W$ to the homomorphism $\mathbf{v} \mapsto \mathbf{v}^*(\mathbf{v})\mathbf{w}$. Deduce that the pure elements in $V^* \otimes W$ correspond precisely to the homomorphisms $V \to W$ of rank ≤ 1.

2.1.5. The Segre embedding. We briefly show a geometric application of the algebra introduced in this section. Let U, V be vector spaces. The natural map $U \times V \to U \otimes V$ induces an injective map on projective spaces

$$\mathbb{P}(U) \times \mathbb{P}(V) \hookrightarrow \mathbb{P}(U \otimes V)$$

called the *Segre embedding*. The map is injective thanks to Exercise 2.1.8.

We have just discovered a simple method for embedding a product of projective spaces in a bigger projective space. If $U = \mathbb{R}^{m+1}$ and $V = \mathbb{R}^{n+1}$ we have an isomorphism $U \otimes V \cong \mathbb{R}^{(m+1)(n+1)}$ and we get an embedding

$$\mathbb{RP}^m \times \mathbb{RP}^n \hookrightarrow \mathbb{RP}^{mn+m+n}.$$

Example 2.1.15. When $m = n = 1$ we get $\mathbb{RP}^1 \times \mathbb{RP}^1 \hookrightarrow \mathbb{RP}^3$. Note that $\mathbb{RP}^1 \times \mathbb{RP}^1$ is topologically a torus. The Segre map is

$$\big([x_0, x_1], [y_0, y_1]\big) \longmapsto \left[\begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \otimes \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}\right]$$

and the right member equals

$$\left[x_0y_0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_0y_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} + x_1y_0 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_1y_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right].$$

In coordinates with respect to the canonical basis the Segre embedding is

$$\big([x_0, x_1], [y_0, y_1]\big) \longmapsto [x_0y_0, x_0y_1, x_1y_0, x_1y_1].$$

It is now an exercise to show that the image is precisely the quadric $z_0z_3 = z_1z_2$ in $\mathbb{RP}^3$. We recover the well-known fact that such a quadric is a torus.

2.1.6. Infinite-dimensional spaces. In very few points in this book we will be concerned with infinite dimensional real vector spaces. We summarise briefly how to extend some of the operations introduced above to an infinite-dimensional context.

The *dual* V^* of a vector space V is always the space of all functionals $V \to \mathbb{R}$. There is a canonical injective map $V \hookrightarrow V^{**}$ which is surjective if and only if V has finite dimension.

Let $V_1, V_2, \ldots$ be vector spaces. The *direct product* and the *direct sum*

$$\prod_i V_i, \qquad \bigoplus_i V_i$$

are respectively the space of all sequences $(v_1, v_2, \ldots)$ with $v_i \in V_i$, and the subspace consisting of sequences with only finitely many non-zero elements. In the latter case, when the spaces V_i are clearly distinct, one may write every sequence simply as a sum

$$v_{i_1} + \ldots + v_{i_h}$$

of the non-zero elements in the sequence. There is a canonical isomorphism

$$(\oplus_i V_i)^* = \textstyle\prod_i V_i^*.$$

The *tensor product* $V \otimes W$ of two vector spaces of arbitrary dimension may be defined as the unique vector space that satisfies the universal property (1). Uniqueness is easy to prove, but existence is more involved: the space $\mathrm{Mult}(V^*, W^*)$ does not work here, it is too big because $V \neq V^{**}$. Instead we may define $V \otimes W$ as a quotient

$$V \otimes W = F(V \times W)/_{\sim}$$

where $F(S)$ is the *free vector space* generated by the set S, that is the abstract vector space with basis S, and $\sim$ is the equivalence relation generated by equivalences of this type:

$$\begin{aligned}(v_1, w) + (v_2, w) &\sim (v_1 + v_2, w),\\ (v, w_1) + (v, w_2) &\sim (v, w_1 + w_2),\\ (\lambda v, w) \sim \lambda(v, w) &\sim (v, \lambda w).\end{aligned}$$

The equivalence class of (v, w) is indicated as $v \otimes w$. More concretely, if $\{v_i\}$ and $\{w_j\}$ are basis of V and W, then $\{v_i \otimes w_j\}$ is a basis of $V \otimes W$, and this is the most important thing to keep in mind.

The tensor product is distributive with respect to direct sum, that is there are canonical isomorphisms

$$V \otimes \big(\oplus_i W_i \big) \cong \oplus_i (V \otimes W_i)$$

but the tensor product is *not* distributive with respect to the direct product in general! We need $\dim V < \infty$ for that:

Exercise 2.1.16. If V has finite dimension, there is a canonical isomorphism

$$V \otimes \big(\textstyle\prod_i W_i\big) \cong \prod_i (V \otimes W_i).$$

Dimostrare?

2.2. Tensors

We have defined the operations $\oplus, \otimes, *$ in full generality, and we now apply them to a single finite-dimensional real vector space V.

2.2.1. Definition. Let V be a real vector space of dimension n and $h, k \geq 0$ some integers. A *tensor* of type (h, k) is an element T of the vector space

$$\mathcal{T}_h^k(V) = \underbrace{V \otimes \cdots \otimes V}_{h} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{k}.$$

In other words T is a multilinear map

$$T\colon \underbrace{V^* \times \cdots \times V^*}_{h} \times \underbrace{V \times \cdots \times V}_{k} \longrightarrow \mathbb{R}.$$

This elegant definition gathers many well-known notions in a single word:

- a tensor of type $(0, 0)$ is by convention an element of $\mathbb{R}$, a *scalar*;
- a tensor of type $(1, 0)$ is an element of V, a *vector*;
- a tensor of type $(0, 1)$ is an element of V^*, a *covector*;

- a tensor of type $(0,2)$ is a *bilinear form* $V \times V \to \mathbb{R}$;
- a tensor of type $(1,1)$ is an element of $V \otimes V^*$ and hence may be interpreted as an *endomorphism* $V \to V$, by Corollary 2.1.13;

More generally, every tensor T of type (h,k) may be interpreted as a multilinear map

$$T' \colon \underbrace{V \times \cdots \times V}_{k} \longrightarrow \underbrace{V \otimes \cdots \otimes V}_{h}$$

by writing

$$T'(\mathbf{v}_1, \ldots, \mathbf{v}_k)(\mathbf{v}_1^*, \ldots, \mathbf{v}_h^*) = T(\mathbf{v}_1^*, \ldots, \mathbf{v}_h^*, \mathbf{v}_1, \ldots, \mathbf{v}_k).$$

In particular a tensor of type $(1,k)$ can be interpreted as a multilinear map

$$T \colon \underbrace{V \times \cdots \times V}_{k} \longrightarrow V.$$

Example 2.2.1. The *euclidean scalar product* in $\mathbb{R}^n$ is defined as

$$(x_1, \ldots, x_n) \cdot (x_1', \ldots, x_n') = x_1 x_1' + \ldots + x_n x_n'.$$

It is a bilinear map $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and hence a tensor of type $(0,2)$.

Example 2.2.2. The *cross product* in $\mathbb{R}^3$ is defined as

$$(x, y, z) \wedge (x', y', z') = (yz' - zy', zx' - xz', xy' - yx').$$

It is a bilinear map $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ and hence a tensor of type $(1,2)$.

Example 2.2.3. The *determinant* may be interpreted as a multilinear map

$$\underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n} \longrightarrow \mathbb{R}$$

that sends $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ to $\det(\mathbf{v}_1 \cdots \mathbf{v}_n)$. As such, it is a tensor of type $(0,n)$.

2.2.2. Coordinates. Every abstract and ethereal object in linear algebra transforms into a more reassuring multidimensional array of numbers, called *coordinates*, as soon as we choose a basis.

Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis of V, and $\mathcal{B}^* = \{\mathbf{v}^1, \ldots, \mathbf{v}^n\}$ be the dual basis of V^*. A basis of the tensor space $\mathcal{T}_h^k(V)$ consists of all the vectors

$$\mathbf{v}_{i_1} \otimes \cdots \otimes \mathbf{v}_{i_h} \otimes \mathbf{v}^{j_1} \otimes \cdots \otimes \mathbf{v}^{j_k}$$

where $1 \leq i_1, \ldots, i_h, j_1, \ldots, j_k \leq n$. Overall, this basis consists of n^{h+k} vectors. Every tensor T of type (h,k) can be written uniquely as

$$T = T^{i_1, \ldots, i_h}_{j_1, \ldots, j_k} \mathbf{v}_{i_1} \otimes \cdots \otimes \mathbf{v}_{i_h} \otimes \mathbf{v}^{j_1} \otimes \cdots \otimes \mathbf{v}^{j_k}. \tag{2}$$

We are using here the *Einstein summation convention*: every index that is repeated at least twice should be summed over the values of the index. Therefore in (2) we sum over all the indices $i_1, \ldots, i_h, j_1, \ldots, j_k$. The following proposition shows how to compute the coordinates of T directly.

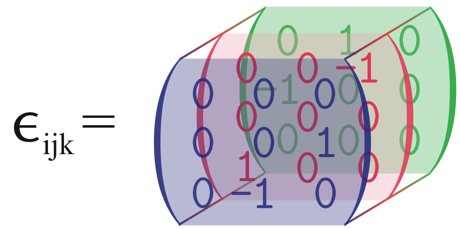

Figure 2.1. The coordinates of the cross product tensor with respect to the canonical basis of $\mathbb{R}^3$ (or any positive orthonormal basis) form the *Levi-Civita symbol* ϵ_{ijk}.

Proposition 2.2.4. *The coordinates of T are*

$$T^{i_1,\ldots,i_h}_{j_1,\ldots,j_k} = T\big(\mathbf{v}^{i_1},\ldots,\mathbf{v}^{i_h},\mathbf{v}_{j_1},\ldots,\mathbf{v}_{j_k}\big).$$

Proof. Apply both members of (2) to $\big(\mathbf{v}^{i_1},\ldots,\mathbf{v}^{i_h},\mathbf{v}_{j_1},\ldots,\mathbf{v}_{j_k}\big)$. $\square$

Example 2.2.5. The coordinates of the Euclidean scalar product g on $\mathbb{R}^n$ with respect to an orthonormal basis are $g_{ij} = \delta_{ij}$.

Example 2.2.6. The coordinates of $\mathrm{id} \in \mathrm{Hom}(V,V) = V \otimes V^*$ with respect to *any* basis are $\mathrm{id}^i_j = \delta^i_j$. This is again the Kronecker delta, written as δ^i_j for convenience.

Exercise 2.2.7. The coordinates of the cross product tensor in $\mathbb{R}^3$ with respect to any positive orthonormal basis are

$$T^i_{jk} = \epsilon_{ijk} = \begin{cases} +1 & \text{if } (i,j,k) \text{ is } (1,2,3), (2,3,1), \text{ or } (3,1,2), \\ -1 & \text{if } (i,j,k) \text{ is } (3,2,1), (1,3,2), \text{ or } (2,1,3), \\ \ \ 0 & \text{if } i=j, \text{ or } j=k, \text{ or } k=i. \end{cases}$$

The three-dimensional array ϵ_{ijk} is called the *Levi-Civita symbol* and is shown in Figure 2.1.

Exercise 2.2.8. The determinant in $\mathbb{R}^3$ may be interpreted as a tensor of type $(0,3)$. Show that its coordinates with respect to any positive orthonormal basis are also ϵ_{ijk}.

2.2.3. Coordinates manipulation.

The coordinates and the Einstein convention are powerful tools that enable us to describe complicated tensor manipulations in a very concise way, and the reader should familiarise with them. We start by exhibiting some simple examples. We fix a basis $\mathcal{B} = \{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ for V and consider coordinates with respect to this basis. We write the coordinates of a generic vector $\mathbf{v}$ as v^i, that is we have

$$\mathbf{v} = v^i \mathbf{v}_i.$$

If $\mathbf{v} \in V$ is a vector and $T: V \to V$ is an endomorphism, that is $T \in \mathcal{T}_1^1(V)$, we may write $\mathbf{w} = T(\mathbf{v})$ directly in coordinates as follows:

$$w^j = T_i^j v^i$$

where v^i, w^j, T_i^j are the coordinates of $\mathbf{v}, \mathbf{w}, T$. The trace of T is simply

$$T_i^i.$$

If $\mathbf{v}, \mathbf{w} \in V$ are vectors and $g: V \times V \to \mathbb{R}$ is a bilinear form, that is $g \in \mathcal{T}_0^2(V)$, it has coordinates g_{ij} and we may write the scalar $g(\mathbf{v}, \mathbf{w})$ as follows:

$$v^i g_{ij} w^j.$$

The expressions $w^j = T_i^j v^i$ and $v^i g_{ij} w^j$ are just the usual products matrix-times-vector(s) that describe endomorphisms and bilinear forms in coordinates: we are only rewriting them using the Einstein convention.

Let T be the tensor of type $(1,2)$ that describes the cross product in $\mathbb{R}^3$. The equality $\mathbf{z} = \mathbf{v} \wedge \mathbf{w}$ can be written in coordinates as

$$z^i = T_{jk}^i v^j w^k.$$

Note that in all the cases described so far the Einstein convention is applied to pairs of indices, one being a superscript and the other a subscript. This is in fact a more general phenomenon.

Example 2.2.9. We prove the well-known equalities

$$(\mathbf{v} \wedge \mathbf{w}) \cdot \mathbf{z} = \mathbf{v} \cdot (\mathbf{w} \wedge \mathbf{z}) = \det(\mathbf{v}\,\mathbf{w}\,\mathbf{z})$$

using coordinates. The three members may be written as

$$v^j T_{jk}^i w^k g_{il} z^l, \qquad v^l g_{li} w^j T_{jk}^i z^k, \qquad \det_{ijk} v^i w^j z^k.$$

Now we take an orthonormal basis $\mathcal{B}$, so that $g_{ij} = \delta_{ij}$ and $T_{jk}^i = \epsilon_{ijk} = \det_{ijk}$. The three members simplify as

$$\epsilon_{ijk} v^j w^k z^i, \qquad \epsilon_{ijk} v^i w^j z^k, \qquad \epsilon_{ijk} v^i w^j z^k$$

and they represent the same number thanks to the symmetries of ϵ.

2.2.4. Change of basis. If $\mathcal{C} = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ is another basis of V then

$$\mathbf{w}_j = A_j^i \mathbf{v}_i, \qquad \mathbf{v}_j = B_j^i \mathbf{w}_i$$

for some matrices A and $B = A^{-1}$. Here A_j^i is the entry at the i-th row and the j-th column of A, and we use the Einstein convention: we sum along the repeated index i. The relation $B = A^{-1}$ may be written as

$$A_k^i B_j^k = \delta_j^i = B_k^i A_j^k$$

where δ_j^i is the Kronecker delta.

Proposition 2.2.10. *The dual bases change as follows:*

$$\mathbf{w}^i = B_j^i \mathbf{v}^j, \qquad \mathbf{v}^i = A_j^i \mathbf{w}^j.$$

Proof. We check that the proposed $\mathbf{w}^i$ form the dual basis of $\mathbf{w}_i$:

$$\mathbf{w}^i(\mathbf{w}_j) = (B^i_k \mathbf{v}^k)(A^l_j \mathbf{v}_l) = B^i_k A^l_j \mathbf{v}^k(\mathbf{v}_l) = B^i_k A^l_j \delta^k_l = B^i_k A^k_j = \delta^i_j.$$

It is a useful exercise to fully understand each of the previous equalities! In the fourth one we removed the Kronecker delta and set $k = l$. $\square$

Let T be a tensor as in (2). We now want to determine the coordinates $\hat{T}^{i_1,\ldots,i_h}_{j_1,\ldots,j_k}$ of T in the new basis $\mathcal{C}$, in terms of the coordinates $T^{i_1,\ldots,i_h}_{j_1,\ldots,j_k}$ in the old basis $\mathcal{B}$ and of the matrices A and B.

Proposition 2.2.11. *We have*

$$\hat{T}^{i_1\ldots i_h}_{j_1\ldots j_k} = B^{i_1}_{l_1} \cdots B^{i_h}_{l_h} A^{m_1}_{j_1} \cdots A^{m_k}_{j_k} T^{l_1\ldots l_h}_{m_1\ldots m_k} \tag{3}$$

This complicated equation may be memorised by noting that we need one A for every lower index of T, and one B for every upper index.

Proof. By Proposition 2.2.4 we have

$$\begin{aligned}
\hat{T}^{i_1,\ldots,i_h}_{j_1,\ldots,j_k} &= T\big(\mathbf{w}^{i_1}, \ldots, \mathbf{w}^{i_h}, \mathbf{w}_{j_1}, \ldots, \mathbf{w}_{j_k}\big) \\
&= T\big(B^{i_1}_{l_1}\mathbf{v}^{l_1}, \ldots, B^{i_h}_{l_h}\mathbf{v}^{l_h}, A^{m_1}_{j_1}\mathbf{v}_{m_1}, \ldots, A^{m_k}_{j_k}\mathbf{v}_{m_k}\big) \\
&= B^{i_1}_{l_1} \cdots B^{i_h}_{l_h} A^{m_1}_{j_1} \cdots A^{m_k}_{j_k} T\big(\mathbf{v}^{l_1}, \ldots, \mathbf{v}^{l_h}, \mathbf{v}_{m_1}, \ldots, \mathbf{v}_{m_k}\big) \\
&= B^{i_1}_{l_1} \cdots B^{i_h}_{l_h} A^{m_1}_{j_1} \cdots A^{m_k}_{j_k} T^{l_1\ldots l_h}_{m_1\ldots m_k}.
\end{aligned}$$

The proof is complete. $\square$

The reader should appreciate the generality of the formula (3): it describes in a single equality the coordinate changes of vectors, covectors, endomorphisms, bilinear forms, the cross product in $\mathbb{R}^3$, the determinant, and some more complicate tensors that we will encounter in this book. We write some of them:

$$\hat{v}^i = B^i_l v^l, \quad \hat{v}_j = A^m_j v_m, \quad \hat{T}^i_j = B^i_l A^m_j T^l_m, \quad \hat{g}_{ij} = A^m_i A^n_j g_{mn}.$$

The formula (3) contains many indices and may look complicated at a first glance, but in fact it only says that the lower indices $j_1, \ldots, j_k$ change through the matrix A, while the upper indices $i_1, \ldots, i_h$ change via the inverse matrix $B = A^{-1}$. For that reason, the lower and upper indices are also called respectively *covariant* and *contravariant*.

Remark 2.2.12. In some physics and engineering text books, the formula (3) is used as a *definition* of tensor: a tensor is simply a multi-dimensional array, that changes as prescribed by the formula if one modifies the basis of the vector space.

We now introduce some operations with tensors.

2.2.5. Tensor product. It follows from the definitions that

$$\mathcal{T}_h^k(V) \otimes \mathcal{T}_m^n(V) = \mathcal{T}_{h+m}^{k+n}(V).$$

In particular, given two tensors $S \in \mathcal{T}_h^k(V)$ and $T \in \mathcal{T}_m^n(V)$, their *product* $S \otimes T$ is an element of $\mathcal{T}_{h+m}^{k+n}(V)$. In coordinates with respect to some basis $\mathcal{B}$, it may be written as

$$(S \otimes T)_{j_1 \ldots j_k j_{k+1} \ldots j_{k+n}}^{i_1 \ldots i_h i_{h+1} \ldots i_{h+m}} = S_{j_1 \ldots j_k}^{i_1 \ldots i_h} T_{j_{k+1} \ldots j_{k+n}}^{i_{h+1} \ldots i_{h+m}}.$$

2.2.6. The tensor algebra. The *tensor algebra* of V is

$$\mathcal{T}(V) = \bigoplus_{h,k \geq 0} \mathcal{T}_h^k(V).$$

The product $\otimes$ is defined on every pair of tensors, and it extends distributively on the whole of $\mathcal{T}(V)$. With this operation $\mathcal{T}(V)$ is an associative algebra and an infinite-dimensional vector space (if V is not trivial). Recall that

$$\mathcal{T}_0^0(V) = \mathbb{R}, \qquad \mathcal{T}_1^0(V) = V, \qquad \mathcal{T}_0^1(V) = V^*.$$

Exercise 2.2.13. If $\dim V \geq 2$ the algebra is not commutative: if $\mathbf{v}, \mathbf{w} \in V$ are independent vectors, then $\mathbf{v} \otimes \mathbf{w} \neq \mathbf{w} \otimes \mathbf{v}$.

Hint. Extend them to a basis $\mathbf{v}_1 = \mathbf{v}, \mathbf{v}_2 = \mathbf{w}, \mathbf{v}_3, \ldots, \mathbf{v}_n$, consider the dual basis $\mathbf{v}^1, \ldots, \mathbf{v}^n$ and determine the value of $\mathbf{v} \otimes \mathbf{w}$ and $\mathbf{w} \otimes \mathbf{v}$ on $(\mathbf{v}^1, \mathbf{v}^2)$. $\square$

We denote for simplicity

$$\mathcal{T}_h(V) = \mathcal{T}_h^0(V), \qquad \mathcal{T}^k(V) = \mathcal{T}_0^k(V).$$

The vector spaces

$$\mathcal{T}_*(V) = \bigoplus_{h \geq 0} \mathcal{T}_h(V), \qquad \mathcal{T}^*(V) = \bigoplus_{k \geq 0} \mathcal{T}^k(V)$$

are both subalgebras of $\mathcal{T}(V)$ and are called the *covariant* and *contravariant tensor algebras*, respectively.

Exercise 2.2.14. The algebras $\mathcal{T}_*(\mathbb{R})$ and $\mathbb{R}[x]$ are isomorphic.

Remark 2.2.15. Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis of V. The elements $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathcal{T}_1(V)$ generate $\mathcal{T}_*(V)$ as a *free algebra*. This means that every element of $\mathcal{T}_*(V)$ may be written as a polynomial in the variables $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in a unique way up to permuting its addenda. Note that $\otimes$ is not commutative, hence the ordering in each monomial is important. As an example:

$$3 + \mathbf{v}_1 - 7\mathbf{v}_2 + \mathbf{v}_1 \otimes \mathbf{v}_2 - 3\mathbf{v}_2 \otimes \mathbf{v}_1.$$

2.2.7. Contractions. We now introduce a general important operation on tensors called *contraction* that generalises the trace of endomorphisms.

The trace is an operation that picks as an input an endomorphism, that is a $(1,1)$-tensor, and produces as an output a number, that is a $(0,0)$-tensor. More generally, a contraction is an operation that transforms a (h,k)-tensor into a $(h-1,k-1)$-tensor, and is defined for all $h,k \geq 1$. It depends on the choice of two integers $1 \leq a \leq h$ and $1 \leq b \leq k$ and results in a linear map

$$C \colon \mathcal{T}_h^k(V) \longrightarrow \mathcal{T}_{h-1}^{k-1}(V).$$

The contraction is defined as follows. Recall that

$$\mathcal{T}_h^k(V) = \underbrace{V \otimes \cdots \otimes V}_{h} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{k}.$$

The indices a and b indicate which factors V and V^* need to be "contracted". After a canonical isomorphism we may put these factors at the end and write

$$\mathcal{T}_h^k(V) = \underbrace{V \otimes \cdots \otimes V}_{h-1} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{k-1} \otimes V \otimes V^* = \mathcal{T}_{h-1}^{k-1}(V) \otimes V \otimes V^*.$$

The *contraction* is the linear map

$$C \colon \mathcal{T}_{h-1}^{k-1}(V) \otimes V \otimes V^* \longrightarrow \mathcal{T}_{h-1}^{k-1}(V)$$

determined by the condition

$$C(\mathbf{w} \otimes \mathbf{v} \otimes \mathbf{v}^*) = \mathbf{v}^*(\mathbf{v})\mathbf{w}.$$

Recall that C is well-defined because $(\mathbf{w}, \mathbf{v}, \mathbf{v}^*) \mapsto \mathbf{v}^*(\mathbf{v})\mathbf{w}$ is multilinear and hence the universal property applies.

Example 2.2.16. The contraction of a pure tensor is

$$C(\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_h \otimes \mathbf{v}^1 \otimes \cdots \otimes \mathbf{v}^k) =$$
$$\mathbf{v}^b(\mathbf{v}_a)\mathbf{v}_1 \otimes \cdots \otimes \widehat{\mathbf{v}_a} \otimes \cdots \otimes \mathbf{v}_h \otimes \mathbf{v}^1 \otimes \cdots \otimes \widehat{\mathbf{v}^b} \otimes \cdots \otimes \mathbf{v}^k$$

where $\hat{\mathbf{w}}$ indicates that the factor $\mathbf{w}$ is omitted.

2.2.8. In coordinates. The definition of a contraction may look abstruse, but we now see that everything is pretty simple in coordinates. Let $\mathcal{B}_i = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for V.

Proposition 2.2.17. *If T has coordinates $T_{j_1,\ldots,j_k}^{i_1,\ldots,i_h}$, then $C(T)$ has*

$$C(T)_{j_1,\ldots,j_{k-1}}^{i_1,\ldots,i_{h-1}} = T_{j_1,\ldots,l,\ldots,j_{k-1}}^{i_1,\ldots,l,\ldots,i_{h-1}}$$

where l is inserted at the positions a above and b below.

Proof. We write the coordinates of T as $T^{i_1,\dots,i,\dots,i_{h-1}}_{j_1,\dots,j,\dots,j_{k-1}}$ for convenience, where i and j occupy the places a and b. We have

$$\begin{aligned}
C(T) &= C\big(T^{i_1,\dots,i,\dots,i_{h-1}}_{j_1,\dots,j,\dots,j_{k-1}}\mathbf{v}_{i_1}\otimes\cdots\otimes\mathbf{v}_i\otimes\cdots\otimes\mathbf{v}_{i_{h-1}}\otimes\mathbf{v}^{j_1}\otimes\cdots\otimes\mathbf{v}^{j}\otimes\cdots\otimes\mathbf{v}^{j_{k-1}}\big)\\
&= T^{i_1,\dots,i,\dots,i_{h-1}}_{j_1,\dots,j,\dots,j_{k-1}}\delta^j_i\mathbf{v}_{i_1}\otimes\cdots\otimes\mathbf{v}_{i_{h-1}}\otimes\mathbf{v}^{j_1}\otimes\cdots\otimes\mathbf{v}^{j_{k-1}}\\
&= T^{i_1,\dots,l,\dots,i_{h-1}}_{j_1,\dots,l,\dots j_{k-1}}\mathbf{v}_{i_1}\otimes\cdots\otimes\mathbf{v}_{i_{h-1}}\otimes\mathbf{v}^{j_1}\otimes\cdots\otimes\mathbf{v}^{j_{k-1}}.
\end{aligned}$$

The proof is complete. □

This shows in particular that, as promised, the contraction of an endomorphism whose coordinates are T^i_j is indeed its trace T^i_i.

Contractions are handled very easily in coordinates. As an example, a tensor T of type $(1,2)$ has coordinates T^i_{jk} and can be contracted in two ways, producing two (typically distinct) covectors $\mathbf{v}$ and $\mathbf{v}'$ with coordinates

$$v_k = T^i_{ik}, \quad v'_j = T^i_{ji}.$$

It is important to remember that the coordinates depend on the choice of a basis $\mathcal{B}$, but the covectors $\mathbf{v}$ and $\mathbf{v}'$ obtained by contracting T do *not* depend on $\mathcal{B}$. Likewise, a tensor of type T^{ij}_{kl} has four types of contractions, producing four (possibly distinct) tensors of type (1,1), that is endomorphisms.

It is convenient to manipulate a tensor using its coordinates as we just did: remember however that we must always contract a covariant index together with a contravariant one! The "contraction" of two covariant (or contravariant) indices makes no sense because it is not basis-independent. This should not be surprising: the trace T^i_i of an endomorphism is basis-independent, but the trace g_{ii} of a bilinear form is notoriously not. Said with other words: there is a canonical homomorphism $V\otimes V^*\to\mathbb{R}$, but there is no canonical homomorphism $V\otimes V\to\mathbb{R}$.

Exercise 2.2.18. The tensor T that expresses the cross product in $\mathbb{R}^3$ has two contractions. Prove that they both give rise to the null covector.

Hint. This can be done by calculation, or abstractly: since T is invariant under orientation-preserving isometries, also its contractions are. □

Example 2.2.19. Let $T, \det, g$ be the tensors in $\mathbb{R}^3$ that represent the cross product, the determinant, and the Euclidean scalar product. They are of type $(1,2)$, $(0,3)$, and $(0,2)$ respectively. The tensor $T\otimes g$ is of type $(1,4)$ and may be written in coordinates as $T^k_{ij}g_{lm}$. It has four contractions $C(T\otimes g)$, that are all of type $(0,3)$. These are

$$T^k_{kj}g_{lm}, \quad T^k_{ik}g_{lm}, \quad T^k_{ij}g_{km}, \quad T^k_{ij}g_{lk}.$$

The first two are null by the previous exercise. The last two, expressed on a orthonormal basis, become ϵ_{ijm} and ϵ_{ijl}. Therefore for these two contractions we get $C(T\otimes g)=\det$.

Every time we sum over a pair of covariant and contravariant indices, we are doing a contraction. So for instance each of the operations

$$w^j = T^j_i v^i, \qquad v^i g_{ij} w^j$$

described in Section 2.2.3 may be interpreted as two-steps operations, where we first multiply some tensors and then we contract the result. Contractions are everywhere.

2.3. Scalar products

We now study vector spaces V equipped with a *scalar product* g. We investigate in particular the effects of g on the tensor algebra $\mathcal{T}(V)$. We start by recalling some basic facts on scalar products.

2.3.1. Definition. A *scalar product* on V is a symmetric bilinear form g that is *not degenerate*, that is

$$g(\mathbf{v}, \mathbf{w}) = 0 \ \forall \mathbf{v} \in V \Longleftrightarrow \mathbf{w} = 0.$$

Recall that the scalar product is

- *positive definite* if $g(\mathbf{v}, \mathbf{v}) > 0 \ \forall \mathbf{v} \neq 0$,
- *negative definite* if $g(\mathbf{v}, \mathbf{v}) < 0 \ \forall \mathbf{v} \neq 0$,
- *indefinite* in the other cases.

Every scalar product g has a *signature* (p, m) where p (respectively, m) is the maximum dimension of a subspace $W \subset V$ such that the restriction $g|_W$ is positive definite (respectively, negative definite). We have $p+m = n = \dim V$. The scalar product is positive definite (respectively, negative definite) $\Longleftrightarrow$ its signature is $(n, 0)$ (respectively, $(0, n)$).

A scalar product g is a tensor of type $(0, 2)$ and its coordinates with respect to some basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ are written as g_{ij}. The basis $\mathcal{B}$ is *orthonormal* if $g_{ij} = \pm\delta_{ij}$ for all i, j. In particular $g_{ii} = \pm 1$, and the sign $+1$ and -1 must occur p and m times as i varies. Every scalar product has an orthonormal basis.

We are mostly interested in positive definite scalar products, but indefinite scalar product also arise in some interesting contexts – notably in Einstein's general relativity.

2.3.2. Isometries. Let V and W be equipped with some scalar products g and h. A linear map $T\colon V \to V$ is an *isometry* if $g(\mathbf{u}, \mathbf{v}) = h\big(T(\mathbf{u}), T(\mathbf{v})\big)$ for all $\mathbf{u}, \mathbf{v} \in V$. This condition can be expressed in coordinates as

$$u^i g_{ij} v^j = u^i T^k_i h_{kl} T^l_j v^j$$

and since it must be verified for all $\mathbf{u}, \mathbf{v}$ we get

$$g_{ij} = T^k_i h_{kl} T^l_j.$$

2.3.3. The identification of V and V^*. Let V be equipped with a scalar product g. Our aim is now to show that g enriches the tensor algebra $\mathcal{T}(V)$ with some new interesting structures.

We first discover that g induces an isomorphism

$$V \longrightarrow V^*$$

that sends $\mathbf{v} \in V$ to the functional $\mathbf{v}^* \in V^*$ defined by $\mathbf{v}^*(\mathbf{w}) = g(\mathbf{v}, \mathbf{w})$. (This is an isomorphism because g is non-degenerate!) This is an important point: as we know, the spaces V and V^* are not canonically identified, but we can identify them once we have fixed a scalar product g.

Exercise 2.3.1. In coordinates, the isomorphism $V \to V^*$ sends a vector v^i to the covector

$$v_j = g_{ij} v^i.$$

The scalar product g induces a scalar product on V^*, that we lazily still name g, as follows:

$$g(\mathbf{v}^*, \mathbf{w}^*) = g(\mathbf{v}, \mathbf{w})$$

where $\mathbf{v}^*, \mathbf{w}^* \in V^*$ are the images of $\mathbf{v}, \mathbf{w} \in V$ along the isomorphism $V \to V^*$ defined above. The scalar product g on V^* is a tensor of type $(2, 0)$ and its coordinates are denoted by g^{ij}.

Proposition 2.3.2. *The matrix g^{ij} is the inverse of g_{ij}.*

Proof. Note that g_{ij} is invertible because g is non-degenerate. The equality defining g^{ij} may be rewritten in coordinates as

$$v^i g_{ik} g^{kl} g_{lj} w^j = v_k g^{kl} w_l = v^i g_{ij} w^j.$$

Since this holds for every $\mathbf{v}, \mathbf{w} \in V$ we get

$$g_{ik} g^{kl} g_{lj} = g_{ij}.$$

Read as a matrices multiplication, this is $GHG = G$ that implies $GH = HG = I$ because G is invertible and hence $H = G^{-1}$. The proof is complete. $\square$

Note that the proposition holds for every choice of a basis $\mathcal{B}$.

2.3.4. Raising and lowering indices. We may use the scalar product g on V to "raise" and "lower" the indices of any tensor at our pleasure. That is, the isomorphism $V \to V^*$ induces an isomorphism

$$\mathcal{T}_h^k(V) \longrightarrow \mathcal{T}_{h+k}(V)$$

for all $h, k \geq 0$. In coordinates, the isomorphism sends a tensor $T^{i_1,\ldots,i_h}_{j_1,\ldots,j_k}$ to

$$U^{i_1,\ldots,i_h,j_1,\ldots,j_k} = T^{i_1,\ldots,i_h}_{l_1,\ldots,l_k} g^{l_1 j_1} \cdots g^{l_k j_k}.$$

We can use g^{ij} to raise the indices of a tensor, and in the opposite direction we can use g_{ij} to lower them. This operation may be encoded efficiently and unambiguously by assigning different indices to distinct columns in the notation. So for instance we start with a tensor like

$$T_i{}^j{}_{kl}$$

and then we may raise or lower some indices to produce a new tensor that we may lazily indicate with the same letter; for instance we can move the indices i and j and get a new tensor

$$T^i{}_{j\ kl}.$$

If $g_{ij} = \delta_{ij}$, then $g^{ij} = \delta^{ij}$ and the coordinates of the two different tensors are just the same, that is $T_i{}^j{}_{kl} = T^i{}_{j\ kl}$ for every i, j, k, l. In general we have

$$T^i{}_{j\ kl} = T_{i'}{}^{j'}{}_{kl} g^{i'i} g_{j'j}.$$

2.3.5. Scalar product on the tensor spaces. A scalar product g on V induces a scalar product on each vector space $\mathcal{T}_h^k(V)$, still boringly denoted by g. This is done as follows: if $S, T \in \mathcal{T}_h^k(V)$, then $g(S, T)$ is the scalar

$$T^{i_1,\ldots,i_h}_{j_1,\ldots,j_k} g_{i_1 l_1} \cdots g_{i_h l_h} g^{j_1 m_1} \cdots g^{j_k m_k} S^{l_1,\ldots,l_h}_{m_1,\ldots,m_k}.$$

Note that this number is basis-independent: it is obtained by multiple contractions of a product of tensors.

Exercise 2.3.3. If $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthonormal basis of V, then

$$\{\mathbf{v}_{i_1} \otimes \cdots \otimes \mathbf{v}_{i_h} \otimes \mathbf{v}^{j_1} \otimes \cdots \otimes \mathbf{v}^{j_k}\}$$

is an orthonormal basis of $\mathcal{T}_h^k(V)$. If g is positive-definite on V then it is so also on $\mathcal{T}_h^k(V)$.

2.4. The symmetric and exterior algebras

Symmetric and antisymmetric matrices play an important role in linear algebra: both concepts can be generalised to tensors.

2.4.1. Symmetric and antisymmetric tensors. We now introduce two special types of contravariant tensors.

Definition 2.4.1. A tensor $T \in \mathcal{T}^k(V)$ is *symmetric* if

$$T(\mathbf{u}_1, \ldots, \mathbf{u}_k) = T\big(\mathbf{u}_{\sigma(1)}, \ldots, \mathbf{u}_{\sigma(k)}\big) \tag{4}$$

for every vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k \in V$ and every permutation $\sigma \in S_k$. On the other hand T is *antisymmetric* if

$$T(\mathbf{u}_1, \ldots, \mathbf{u}_k) = (-1)^{\mathrm{sgn}(\sigma)} T\big(\mathbf{u}_{\sigma(1)}, \ldots, \mathbf{u}_{\sigma(k)}\big)$$

for every vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k \in V$ and every permutation $\sigma \in S_k$.

Both conditions are very easily expressed in coordinates. As usual we fix any basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ on V and consider the coordinates of T with respect to $\mathcal{B}$.

Proposition 2.4.2. *A tensor $T \in \mathcal{T}^k(V)$ is*

- *symmetric* $\iff T_{i_1,\ldots,i_k} = T_{i_{\sigma(1)},\ldots,i_{\sigma(k)}} \quad \forall i_1,\ldots,i_k,\ \forall\sigma;$
- *antisymmetric* $\iff T_{i_1,\ldots,i_k} = (-1)^{\mathrm{sgn}(\sigma)} T_{i_{\sigma(1)},\ldots,i_{\sigma(k)}} \ \forall i_1,\ldots,i_k, \forall\sigma.$

Proof. We prove the first sentence, the second is analogous. Recall that

$$T_{i_1,\ldots,i_k} = T\big(\mathbf{v}_{i_1},\ldots,\mathbf{v}_{i_k}\big).$$

Therefore we must prove that (4) holds for all vectors $\iff$ it holds for the vectors in the basis $\mathcal{B}$. This is left as an exercise. □

For instance a tensor T_{ij} is symmetric if $T_{ij} = T_{ji}$ and antisymmetric if $T_{ij} = -T_{ji}$, for all $1 \leq i, j \leq n$.

Example 2.4.3. Every scalar product on V is a symmetric tensor $g \in \mathcal{T}^2(V)$. The determinant is an antisymmetric tensor $\det \in \mathcal{T}^n(\mathbb{R}^n)$.

Remark 2.4.4. If T is antisymmetric and the indices $i_1,\ldots,i_k$ are not all distinct, then $T_{i_1,\ldots,i_k} = 0$.

2.4.2. Symmetrisation and antisymmetrisation of tensors. If a tensor T is not (anti-)symmetric, we can transform it by brute force into an (anti-)symmetric one.

Let $T \in \mathcal{T}^k(V)$ be a contravariant tensor. The *symmetrisation* of T is the tensor $S(T) \in \mathcal{T}^k(V)$ defined by averaging T on permutations as follows:

$$S(T)(\mathbf{v}_1,\ldots,\mathbf{v}_k) = \frac{1}{k!}\sum_{\sigma\in S_k} T\big(\mathbf{v}_{\sigma(1)},\ldots,\mathbf{v}_{\sigma(k)}\big).$$

Analogously, the *antisymmetrisation* of T is the tensor

$$A(T)(\mathbf{v}_1,\ldots,\mathbf{v}_k) = \frac{1}{k!}\sum_{\sigma\in S_k} (-1)^{\mathrm{sgn}(\sigma)} T\big(\mathbf{v}_{\sigma(1)},\ldots,\mathbf{v}_{\sigma(k)}\big).$$

Exercise 2.4.5. The tensors $S(T)$ and $A(T)$ are indeed symmetric and antisymmetric, respectively. We have $S(T) = T \iff T$ is symmetric and $A(T) = T \iff T$ is antisymmetric.

In coordinates with respect to some basis we have

$$S(T)_{i_1,\ldots,i_k} = \frac{1}{k!}\sum_{\sigma\in S_k} T_{i_{\sigma(1)},\ldots,i_{\sigma(k)}},$$
$$A(T)_{i_1,\ldots,i_k} = \frac{1}{k!}\sum_{\sigma\in S_k} (-1)^{\mathrm{sgn}(\sigma)} T_{i_{\sigma(1)},\ldots,i_{\sigma(k)}}.$$

The members on the right can be written more concisely as

$$T_{(i_1,\ldots,i_k)}, \qquad T_{[i_1,\ldots,i_k]}.$$

The round or square brackets indicate that we symmetrise or antisymmetrise by summing along all permutations on the indices (and dividing by $k!$).

2.4.3. The symmetric and antisymmetric algebras. We now introduce two more algebras associated to V. For every $k \geq 0$ we denote by

$$S^k(V), \qquad \Lambda^k(V)$$

the vector subspace of $\mathcal{T}^k(V)$ consisting of all the symmetric or antisymmetric tensors, respectively. We now define

$$S^*(V) = \bigoplus_{k\geq 0} S^k(V), \qquad \Lambda^*(V) = \bigoplus_{k\geq 0} \Lambda^k(V).$$

These are both vector subspaces of the contravariant tensor algebra $\mathcal{T}^*(V)$. These are *not* subalgebras of $\mathcal{T}^*(V)$, because they are not closed under $\otimes$. Note that

$$S^1(V) = \Lambda^1(V) = \mathcal{T}^1(V) = V^*$$

but $S^2(V)$ and $\Lambda^2(V)$ are strictly smaller than $\mathcal{T}^2(V)$ if $\dim V \geq 2$, because of the following:

Exercise 2.4.6. If $\mathbf{v}^*, \mathbf{w}^* \in V^*$ are independent, then $\mathbf{v}^* \otimes \mathbf{w}^*$ is neither symmetric nor antisymmetric. Moreover

$$S(\mathbf{v}^* \otimes \mathbf{w}^*) = \frac{1}{2}(\mathbf{v}^* \otimes \mathbf{w}^* + \mathbf{w}^* \otimes \mathbf{v}^*), \quad A(\mathbf{v}^* \otimes \mathbf{w}^*) = \frac{1}{2}(\mathbf{v}^* \otimes \mathbf{w}^* - \mathbf{w}^* \otimes \mathbf{v}^*).$$

The spaces $S^*(V)$ and $\Lambda^*(V)$ are actually algebras, but with some products different from $\otimes$, that we now define. The *symmetrised product* of some contravariant tensors $T^1 \in \mathcal{T}^{k_1}(V), \ldots, T^m \in \mathcal{T}^{k_m}(V)$ is

$$T^1 \odot \cdots \odot T^m = \frac{(k_1 + \ldots + k_m)!}{k_1! \cdots k_m!} S(T^1 \otimes \cdots \otimes T^m)$$

while their *antisymmetrised product* is

$$T^1 \wedge \ldots \wedge T^m = \frac{(k_1 + \ldots + k_m)!}{k_1! \cdots k_m!} A(T^1 \otimes \cdots \otimes T^m).$$

For instance if $\mathbf{v}^*, \mathbf{w}^* \in V^*$ then

$$\mathbf{v}^* \odot \mathbf{w}^* = \mathbf{v}^* \otimes \mathbf{w}^* + \mathbf{w}^* \otimes \mathbf{v}^*, \qquad \mathbf{v}^* \wedge \mathbf{w}^* = \mathbf{v}^* \otimes \mathbf{w}^* - \mathbf{w}^* \otimes \mathbf{v}^*.$$

Note that

$$\mathbf{v}^* \odot \mathbf{w}^* = \mathbf{w}^* \odot \mathbf{v}^*, \qquad \mathbf{v}^* \wedge \mathbf{w}^* = -\mathbf{w}^* \wedge \mathbf{v}^*.$$

More generally, if $\mathbf{v}^1, \ldots, \mathbf{v}^m \in V^*$ then

$$\mathbf{v}^1 \odot \cdots \odot \mathbf{v}^m = \sum_{\sigma \in S_m} \mathbf{v}^{\sigma(1)} \otimes \cdots \otimes \mathbf{v}^{\sigma(m)},$$

$$\mathbf{v}^1 \wedge \cdots \wedge \mathbf{v}^m = \sum_{\sigma \in S_m} (-1)^{\mathrm{sgn}(\sigma)} \mathbf{v}^{\sigma(1)} \otimes \cdots \otimes \mathbf{v}^{\sigma(m)}.$$

Using coordinates with respect to some basis $\mathcal{B}$ of V we can write

$$(T \odot U)_{i_1,\ldots,i_{p+q}} = \frac{(p+q)!}{p!q!} T_{(i_1,\ldots,i_p} U_{i_{p+1},\ldots,i_{p+q})},$$
$$(T \wedge U)_{i_1,\ldots,i_{p+q}} = \frac{(p+q)!}{p!q!} T_{[i_1,\ldots,i_p} U_{i_{p+1},\ldots,i_{p+q}]}.$$

Proposition 2.4.7. *The vector spaces $S^*(V)$ and $\Lambda^*(V)$ form two associative algebras with the products $\odot$ and $\wedge$ respectively.*

Proof. Everything is immediate except associativity. We prove it for Λ, the other is analogous. Pick $S \in \Lambda^p$, $T \in \Lambda^q$, and $U \in \Lambda^r$. In coordinates

$$\begin{aligned}
\big((S \wedge T) \wedge U\big)_{i_1,\ldots,i_{p+q+r}} &= \frac{1}{(p+q)!r!}(S \wedge T)_{[i_1,\ldots,i_{p+q}} U_{i_{p+q+1},\ldots,i_{p+q+r}]} \\
&= \frac{1}{(p+q)!p!q!r!} S_{[[i_1,\ldots,i_p} T_{i_{p+1},\ldots,i_{p+q}]} U_{i_{p+q+1},\ldots,i_{p+q+r}]} \\
&= \frac{1}{p!q!r!} S_{[i_1,\ldots,i_p} T_{i_{p+1},\ldots,i_{p+q}} U_{i_{p+q+1},\ldots,i_{p+q+r}]} \\
&= (S \wedge T \wedge U)_{i_1,\ldots,i_{p+q+r}}.
\end{aligned}$$

The third equality follows from the fact that the same permutation in the symmetric group S_{p+q+r} is obtained $(p+q)!$ times. Analogously we can prove that $S \wedge (T \wedge U) = S \wedge T \wedge U$. The proof is complete. □

The two algebras $S^*(V)$ and $\Lambda^*(V)$ are called the *contravariant symmetric algebra* and the *contravariant exterior algebra*. The products $\otimes$ and $\wedge$ are called the *symmetric* and *exterior product*.

2.4.4. Dimensions. We now construct some standard basis for $S^k(V)$ and $\Lambda^k(V)$ and calculate their dimensions. Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for V and $\mathcal{B}^* = \{\mathbf{v}^1, \ldots, \mathbf{v}^n\}$ the dual basis of V^*.

Proposition 2.4.8. *A basis for $S^k(V)$ is*

$$\left\{\mathbf{v}^{i_1} \odot \cdots \odot \mathbf{v}^{i_k}\right\}$$

where $1 \leq i_1 \leq \ldots \leq i_k \leq n$ vary. A basis for $\Lambda^k(V)$ is

$$\left\{\mathbf{v}^{i_1} \wedge \cdots \wedge \mathbf{v}^{i_k}\right\}$$

where $1 \leq i_1 < \ldots < i_k \leq n$ vary.

Proof. This is a consequence of Propositions 2.4.2 and Remark 2.4.4. □

Example 2.4.9. The following is a basis for $S^2(\mathbb{R}^2)$:

$$\mathbf{e}^1 \odot \mathbf{e}^1, \quad \mathbf{e}^1 \odot \mathbf{e}^2, \quad \mathbf{e}^2 \odot \mathbf{e}^2.$$

The following is a basis for $\Lambda^2(\mathbb{R}^3)$:

$$\mathbf{e}^1 \wedge \mathbf{e}^2, \quad \mathbf{e}^1 \wedge \mathbf{e}^3, \quad \mathbf{e}^2 \wedge \mathbf{e}^3.$$

Corollary 2.4.10. *We have*

$$\dim S^k(V) = \binom{n+k-1}{k},$$

$$\dim \Lambda^k(V) = \begin{cases} \binom{n}{k} & \text{if } k \leq n, \\ 0 & \text{if } k > n. \end{cases}$$

Corollary 2.4.11. *The algebra $S^*(V)$ is commutative, while $\Lambda^*(V)$ is* anti-commutative, *that is*

$$T \wedge U = (-1)^{pq} U \wedge T$$

for every $T \in \Lambda^p(V)$ and $U \in \Lambda^q(V)$.

Proof. We prove anticommutativity. It suffices to prove this when T, U are basis elements, that is we must show that

$$\mathbf{v}^{i_1} \wedge \ldots \wedge \mathbf{v}^{i_p} \wedge \mathbf{v}^{j_1} \wedge \ldots \wedge \mathbf{v}^{j_q} = (-1)^{pq} \mathbf{v}^{j_1} \wedge \ldots \wedge \mathbf{v}^{j_q} \wedge \mathbf{v}^{i_1} \wedge \ldots \wedge \mathbf{v}^{i_p}.$$

This equality follows from applying pq times the simple equality

$$\mathbf{v}^* \wedge \mathbf{w}^* = -\mathbf{w}^* \wedge \mathbf{v}^*.$$

The proof is complete. □

Corollary 2.4.12. *If $T \in \Lambda^k(V)$ with odd k then $T \wedge T = 0$.*

Corollary 2.4.13. *We have* $\dim S^*(V) = \infty$ *and* $\dim \Lambda^*(V) = 2^n$.

Exercise 2.4.14. The algebras $S^*(V)$ and $\mathbb{R}[x_1, \ldots, x_n]$ are isomorphic.

2.4.5. The determinant line. One of the most important aspect of the theory, that will have important applications in the next chapters, is the following – apparently innocuous – fact:

$$\dim \Lambda^n(V) = 1.$$

The space $\Lambda^n(V)$ is called the *determinant line*. If $\mathbf{v}^1, \ldots, \mathbf{v}^n$ is a basis of V^*, then a generator for $\Lambda^n(V)$ is the tensor

$$\mathbf{v}^1 \wedge \ldots \wedge \mathbf{v}^n.$$

In fact, we already know that there is only one alternating n-linear form in V up to rescaling – this is exactly where the determinant comes from. When $V = \mathbb{R}^n$, we get

$$\det = \mathbf{e}^1 \wedge \ldots \wedge \mathbf{e}^n$$

where $\mathbf{e}^1, \ldots, \mathbf{e}^n$ is the canonical basis of $(\mathbb{R}^n)^* = \mathbb{R}^n$. Note however that $\Lambda^n(V)$ has no canonical generator unless we make some choice, like for instance a basis of V.

Let now $\mathbf{v}^1, \ldots, \mathbf{v}^n$ and $\mathbf{w}^1, \ldots, \mathbf{w}^n$ be two basis of V^*, and let A the change of basis matrix, so that $\mathbf{v}^i = A^i_j \mathbf{w}^j$.

Proposition 2.4.15. *The following equality holds:*

$$\mathbf{v}^1 \wedge \ldots \wedge \mathbf{v}^n = \det A \cdot \mathbf{w}^1 \wedge \ldots \wedge \mathbf{w}^n.$$

Proof. We have

$$\begin{aligned}
\mathbf{v}^1 \wedge \ldots \wedge \mathbf{v}^n &= A^1_{j_1} \cdots A^n_{j_n} \mathbf{w}^{j_1} \wedge \ldots \wedge \mathbf{w}^{j_n} \\
&= \sum_{\sigma \in S_n} A^1_{\sigma(1)} \cdots A^n_{\sigma(n)} \mathbf{w}^{\sigma(1)} \wedge \ldots \wedge \mathbf{w}^{\sigma(n)} \\
&= \sum_{\sigma \in S_n} (-1)^{\mathrm{sgn}(\sigma)} A^1_{\sigma(1)} \cdots A^n_{\sigma(n)} \mathbf{w}^1 \wedge \ldots \wedge \mathbf{w}^n \\
&= \det A \cdot \mathbf{w}^1 \wedge \ldots \wedge \mathbf{w}^n.
\end{aligned}$$

The proof is complete. □

We have discovered here another important fact: the equality looks like the formula in the change of variables in multiple integrals, see Section 1.3.8. This will allow us to connect alternating tensors with integration and volume.

2.4.6. Totally decomposable antisymmetric tensors. An antisymmetric tensor $T \in \Lambda^k(V)$ is *totally decomposable* if it may be written as

$$T = \mathbf{w}^1 \wedge \ldots \wedge \mathbf{w}^k$$

for some covectors $\mathbf{w}^1, \ldots, \mathbf{w}^k \in V^*$. This notion is similar to that of a pure tensor, only with the product $\wedge$ instead of $\otimes$.

Proposition 2.4.16. *The element* $T = \mathbf{w}^1 \wedge \ldots \wedge \mathbf{w}^k$ *is non-zero* $\Longleftrightarrow$ *the covectors* $\mathbf{w}^1, \ldots, \mathbf{w}^k$ *are linearly independent.*

Proof. If $\mathbf{w}^1 = \lambda_i \mathbf{w}^i$, then T is a combination of totally decomposable elements where the same covector $\mathbf{w}^i$ appears twice, and $\mathbf{w}^i \wedge \mathbf{w}^i = 0$.

Conversely, if they are independent they can be completed to a basis $\mathbf{w}^1, \ldots, \mathbf{w}^n$ of V and we know that $\mathbf{w}^1 \wedge \ldots \wedge \mathbf{w}^n \neq 0$, hence $T \neq 0$. □

Not all the antisymmetric tensors are totally decomposable:

Exercise 2.4.17. If $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4 \in V^*$ are linearly independent then

$$\mathbf{v}_1 \wedge \mathbf{v}_2 + \mathbf{v}_3 \wedge \mathbf{v}_4$$

is not totally decomposable.

Hint. If $\mathbf{w}$ is totally decomposable, then $\mathbf{w} \wedge \mathbf{w} = 0$. □

2.4.7. Covariant versions. We have established the theory of symmetric and antisymmetric contravariant tensors, but actually everything we said also holds *verbatim* for the *covariant* tensors: we can therefore denote by

$$S_k(V), \qquad \Lambda_k(V)$$

the subspaces of $\mathcal{T}_k(V)$ consisting of all the symmetric or antisymmetric tensors, and define

$$S_*(V) = \bigoplus_{k\geq 0} S_k(V), \qquad \Lambda_*(V) = \bigoplus_{k\geq 0} \Lambda_k(V).$$

These form two algebras, called the *covariant symmetric algebra* and *covariant exterior algebra*.

2.4.8. Linear maps. Every linear map $L\colon V \to W$ between vector spaces induces some algebra homomorphisms

$$\begin{aligned} L_*\colon \mathcal{T}_*(V) &\longrightarrow \mathcal{T}_*(W), \qquad & L^*\colon \mathcal{T}^*(W) &\longrightarrow \mathcal{T}^*(V), \\ L_*\colon S_*(V) &\longrightarrow S_*(W), \qquad & L^*\colon S^*(W) &\longrightarrow S^*(V), \\ L_*\colon \Lambda_*(V) &\longrightarrow \Lambda_*(W), \qquad & L^*\colon \Lambda^*(W) &\longrightarrow \Lambda^*(V). \end{aligned}$$

The passing from L to L_* or L^* is functorial, that is

$$(L' \circ L)_* = L'_* \circ L_*, \qquad \mathrm{id}_* = \mathrm{id},$$

$$(L' \circ L)^* = L^* \circ (L')^*, \qquad \mathrm{id}^* = \mathrm{id}.$$

From this we deduce that if L is an isomorphism then L_* is an isomorphism. More than that:

- if L is injective then L_* is injective and L^* is surjective,
- if L is surjective then L_* is surjective and L^* injective.

This holds because if L is injective (surjective) there is a linear map $L'\colon W \to V$ such that $L'\circ L = \mathrm{id}_V$ ($L\circ L' = \mathrm{id}_W$), as one proves with standard linear algebra techniques.

Remark 2.4.18. The terms *covariance* and its opposite *contravariance* are used for similar objects in two quite different contexts, and this is a permanent source of confusion. In general, a mathematical entity is "covariant" if it changes "in the same way" as some other preferred entity when some modification is made. But which modifications are we considering here?

Physicists are interested in changes of frame, that is of basis, and they note that if we change a basis with a matrix A, then the coordinates of a vector change with $B = A^{-1}$, that is *contravariantly*. On the other hand, mathematicians are mostly interested in functoriality, and note that a map $L\colon V \to W$ induces maps $L_*\colon \mathcal{T}_*(V) \to \mathcal{T}_*(W)$ and $L^*\colon \mathcal{T}^*(W) \to \mathcal{T}^*(V)$ on tensors, and they call *contravariant* the second ones because arrows are reversed.

The reader can ignore all these matters – in fact, these issues start to annoy you only when you decide to write a textbook, and you are forced to choose a notation that is both reasonable and consistent.

2.5. Grassmannians

After many pages of algebra, we now would like to see some geometric applications of the structures that we have just introduced. Here is one.

2.5.1. Definition. Let V be a real vector space of dimension n. Remember that the projective space $\mathbb{P}(V)$ is the set of all the vector lines in V. More generally, fix $0 < k < n = \dim V$.

Definition 2.5.1. The *Grassmannian* $\mathrm{Gr}_k(V)$ is the set consisting of all the k-dimensional vector subspaces $W \subset V$.

Recall that every $W \subset V$ determines a dual subspace $W^* \subset V^*$ consisting of all the functionals that vanish on W. We have $\dim W^* = n - \dim W$. Therefore the sets $\mathrm{Gr}_k(V)$ and $\mathrm{Gr}_{n-k}(V^*)$ may be identified canonically. In particular we get

$$\mathrm{Gr}_1(V) = \mathbb{P}(V), \qquad \mathrm{Gr}_{n-1}(V) = \mathbb{P}(V^*).$$

The simplest new interesting set to investigate is the Grassmannian $\mathrm{Gr}_2(\mathbb{R}^4)$ of vector planes in $\mathbb{R}^4$. How can we study such an object?

2.5.2. The Plücker embedding. A generic Grassmannian is not a projective space in any sense, but we now show that it can be embedded in some (bigger) projective space. We do this using the exterior algebra.

For every k-dimensional subspace $W \subset V$ of V we have the inclusion map $L \colon W \to V$ which induces an injective linear map

$$\Lambda_k(W) \longrightarrow \Lambda_k(V).$$

Since $\dim \Lambda_k(W) = 1$, the image of this map is a line in $\Lambda_k(V)$ that depends only on W. By sending W to this line we get a map

$$\mathrm{Gr}_k(V) \longrightarrow \mathbb{P}\big(\Lambda_k(V)\big)$$

called the *Plücker embedding*. Concretely, the map sends $W \subset V$ to

$$[\mathbf{w}_1 \wedge \ldots \wedge \mathbf{w}_k]$$

where $\mathbf{w}_1, \ldots, \mathbf{w}_k$ is any basis of W.

Proposition 2.5.2. *The Plücker embedding is injective.*

Proof. Consider $W \neq W'$. Let $\mathbf{w}_1, \ldots, \mathbf{w}_k$ and $\mathbf{w}'_1, \ldots, \mathbf{w}'_k$ be any basis of W and W'. Pick any vector $\mathbf{w} \in W \setminus W'$. By Proposition 2.4.16 we have

$$\mathbf{w}_1 \wedge \ldots \wedge \mathbf{w}_k \wedge \mathbf{w} = 0, \qquad \mathbf{w}'_1 \wedge \ldots \wedge \mathbf{w}'_k \wedge \mathbf{w} \neq 0.$$

Therefore the tensors $\mathbf{w}_1 \wedge \ldots \wedge \mathbf{w}_k$ and $\mathbf{w}'_1 \wedge \ldots \wedge \mathbf{w}'_k$ cannot be proportional. $\square$

For instance, we get the Plücker embedding

$$\mathrm{Gr}_2(\mathbb{R}^4) \longhookrightarrow \mathbb{P}\big(\Lambda_2(\mathbb{R}^4)\big) \cong \mathbb{P}\Big(\mathbb{R}^{\binom{4}{2}}\Big) = \mathbb{RP}^5.$$

This embedding is clearly not surjective because of Exercise 2.4.17. We can consider the set $\mathrm{Gr}_k(\mathbb{R}^n)$ canonically embedded in $\mathbb{RP}^N$ with $N = \binom{n}{k} - 1$ and in particular we may assign it the subspace topology.

2.5.3. The Veronese embedding. Here is another geometric application. Fix $k > 0$ and consider the natural map $V \to S^k(V)$ defined as

$$\mathbf{v} \longmapsto \underbrace{\mathbf{v} \odot \cdots \odot \mathbf{v}}_{k}.$$

This map is not linear in general, however it is injective (exercise) and it also induces an injective map between projective spaces

$$\mathbb{P}(V) \hookrightarrow \mathbb{P}\big(S^k(V)\big)$$

called the *Veronese embedding*. This map is not a projective map in general.

Exercise 2.5.3. If $V = \mathbb{R}^{n+1}$ and we use the canonical basis, we get

$$\mathbb{P}^n \longhookrightarrow \mathbb{P}^N$$

where $N = \binom{n+k}{k} - 1$. The map sends $[x_0, \ldots, x_n]$ to $[x_0^k, x_0^{k-1}x_1, \ldots]$ where the square brackets contain all the possible degree-k monomials in the variables $x_0, \ldots, x_n$. For instance for $k = n = 2$ we get

$$\mathbb{P}^2 \longhookrightarrow \mathbb{P}^5$$

given by

$$[x, y, z] \longmapsto [x^2, y^2, z^2, xy, yz, zx].$$

For $n = 1$ we get

$$\mathbb{P}^1 \longhookrightarrow \mathbb{P}^k$$

given by

$$[x, y] \longmapsto [x^k, x^{k-1}y, \ldots, xy^{k-1}, y^k].$$

2.6. Orientation

We end this chapter with a short section, where we introduce and discuss the notion of orientation on a real vector space V.

2.6.1. Definition. Let us say that two basis of V are *cooriented* if the change of basis matrix relating them has positive determinant. Being cooriented is an equivalence relation on the set of all the basis in V, and one checks immediately that we get precisely two equivalence classes of basis.

Definition 2.6.1. An *orientation* on V is the choice of one of these two equivalence classes.

If V is oriented, the bases belonging to the preferred equivalence class are called *positive*, and the other *negative*. Of course V has two distinct orientations. The space $\mathbb{R}^n$ has a canonical orientation given by the canonical basis, but a space V may not have a canonical orientation in general.

Exercise 2.6.2. If $V = U \oplus W$, then an orientation on any two of the spaces U, V, W induces an orientation on the third, by requiring that, for every positive basis $u_1, \ldots, u_k$ of U and $w_1, \ldots, w_h$ of W, the basis $u_1, \ldots, u_k, w_1, \ldots, w_h$ of V is also positive.

2.6.2. Via the exterior algebra. We now study briefly the relation between the orientation on V and on some other tensor spaces.

Exercise 2.6.3. An orientation on V induces one on V^* and vice-versa, as follows: a basis on V is positive $\iff$ its dual basis on V^* is positive.

Proposition 2.4.15 in turn shows that an orientation on V^* induces one on $\Lambda^n(V)$ and vice-versa: a basis $\mathbf{v}^1, \ldots, \mathbf{v}^n$ is positive in $V^* \iff$ the element $\mathbf{v}^1 \wedge \ldots \wedge \mathbf{v}^n$ is a positive basis for the line $\Lambda^n(V)$.

Indeed we could define an orientation on V to be an orientation on the determinant line $\Lambda^n(V)$.

2.6.3. Scalar product. Finally, we note that if V is equipped with both an orientation and a positive-definite scalar product g, then we get for free a canonical generator T for the determinant line $\Lambda^n(V)$ by taking

$$T = \mathbf{v}^1 \wedge \ldots \wedge \mathbf{v}^n$$

where $\mathbf{v}^1, \ldots, \mathbf{v}^n$ is *any* positive orthonormal basis of V^*. The generator T does not depend on the basis, because any two such basis are related by an orthogonal matrix A with $\det A = 1$ and hence Proposition 2.4.15 applies. The element T is also determined by requiring that

$$T(\mathbf{v}_1, \ldots, \mathbf{v}_n) = 1$$

on every positive orthonormal basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of V.

CHAPTER 3

Smooth manifolds

3.1. Smooth manifolds

We introduce here the notion of *smooth manifold*, the main protagonist of the book.

3.1.1. Definition. The definition of topological manifold that we have proposed in Section 1.1.6 is simple but also very poor, and it is quite hard to employ it concretely: for instance, it is already non obvious to answer such a natural question as whether $\mathbb{R}^n$ and $\mathbb{R}^m$ are homeomorphic when $n \neq m$. To make life easier, we enrich the definition by adding a *smooth structure* that exploits the power of differential calculus.

Let M be a topological n-manifold. A *chart* is a homeomorphism $\varphi\colon U \to V$ from some open set $U \subset M$ onto an open set $V \subset \mathbb{R}^n$. The inverse map $\varphi^{-1}\colon V \to U$ is called a *parametrisation*. An *atlas* on M is a set $\{\varphi_i\}$ of charts $\varphi_i\colon U_i \to V_i$ that cover M, that is such that $\cup U_i = M$.

Let $\{\varphi_i\}$ be an atlas on M. Whenever $U_i \cap U_j \neq \varnothing$, we define a *transition map*

$$\varphi_{ij}\colon \varphi_i(U_i \cap U_j) \longrightarrow \varphi_j(U_i \cap U_j)$$

by setting $\varphi_{ij} = \varphi_j \circ \varphi_i^{-1}$. The reader should visualise this definition by looking at Figure 3.1. Note that both the domain and codomain of φ_{ij} are open sets of $\mathbb{R}^n$, and hence it makes perfectly sense to ask whether the transition functions φ_{ij} are smooth. We say that the atlas is *smooth* if all the transition functions φ_{ij} are smooth. Here is the most important definition of the book:

Definition 3.1.1. A *smooth n-manifold* is a topological n-manifold equipped with a smooth atlas.

To be more precise, we allow the same smooth manifold to be described by different atlases, as follows: we say that two smooth atlases $\{\varphi_i\}$ and $\{\varphi'_j\}$ are *compatible* if their union is again a smooth atlas; compatibility is an equivalent relation and we define a *smooth structure* on a topological manifold M to be an equivalence class of smooth atlases on M. The rigorous definition of a smooth manifold is a topological manifold M with a smooth structure on it.

Remark 3.1.2. The union of all the smooth atlases in M compatible with a given one is again a compatible smooth atlas, called a *maximal atlas*. The maximal atlas is uniquely determined by the smooth structure: hence one can

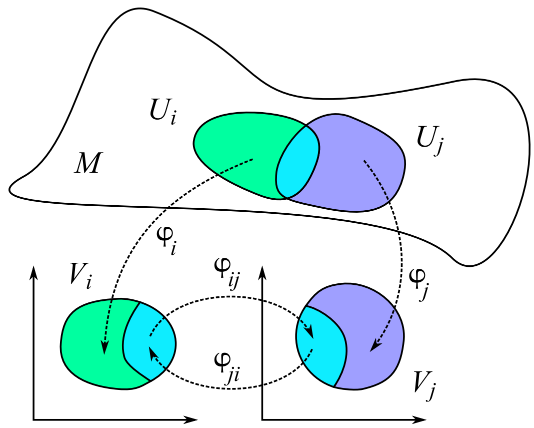

Figure 3.1. Two overlapping charts φ_i and φ_j induce a transition function $\varphi_{ij} = \varphi_j \circ \varphi_i^{-1}$.

also define a smooth manifold to be a topological manifold equipped with a maximal atlas, without using equivalence classes.

As a first example, every open subset $U \subset \mathbb{R}^n$ is naturally a smooth manifold, with an atlas that consists of a unique chart: the identity map $U \to U$.

The open subsets of $\mathbb{R}^n$ can be pretty complicated, but they are never compact. To construct some compact smooth manifolds we now build some atlases as in Figure 1.2.

3.1.2. Spheres. Recall that the *unit sphere* is

$$S^n = \left\{ x \in \mathbb{R}^{n+1} \;\middle|\; \|x\| = 1 \right\}.$$

This is the prototypical example of a compact smooth manifold. To build a smooth atlas on S^n, we may consider the hemispheres

$$U_i^+ = \left\{ x \in S^n \;\middle|\; x_i > 0 \right\}, \qquad U_i^- = \left\{ x \in S^n \;\middle|\; x_i < 0 \right\}$$

for $i = 1, \ldots, n+1$ and define a chart $\varphi_i^\pm \colon U_i^\pm \to B^n$ by forgetting x_i, that is

$$\varphi_i^\pm(x_1, \ldots, x_{n+1}) = (x_1, \ldots, \check{x}_i, \ldots, x_{n+1}).$$

Proposition 3.1.3. *These charts define a smooth atlas on S^n.*

Proof. The inverse $(\varphi_i^\pm)^{-1}$ is

$$(y_1, \ldots, y_n) \longmapsto \left(y_1, \ldots, y_{i-1}, \pm\sqrt{1 - y_1^2 - \ldots - y_n^2}, y_i, \ldots, y_n \right).$$

The transition functions are compositions $\varphi_i^\pm \circ (\varphi_j^\pm)^{-1}$ and are smooth. $\square$

We have equipped S^n with the structure of a smooth manifold. As we said, the same smooth structure on S^n can be built via a different atlas: for instance

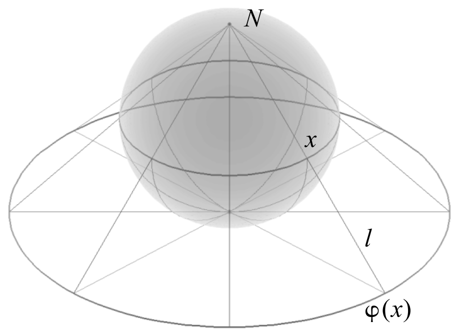

Figure 3.2. The stereographic projection sends a point $x \in S^n \setminus \{N\}$ to the point $\varphi(x)$ obtained by intersecting the line l containing N and x with the horizontal hyperplane $x_{n+1} = -1$.

we describe one now that contains only two charts. Consider the north pole $N = (0, \ldots, 0, 1)$ in S^n and the *stereographic projection* $\varphi_N\colon S^n \setminus \{N\} \to \mathbb{R}^n$,

$$\varphi_N(x_1, \ldots, x_{n+1}) = \frac{2}{1 - x_{n+1}}(x_1, \ldots, x_n).$$

The geometric interpretation of the stereographic projection is illustrated in Figure 3.2. The map φ_N is a homeomorphism, so in particular $S^n \setminus \{N\}$ is homeomorphic to $\mathbb{R}^n$. We can analogously define a stereographic projection φ_S via the south pole $S = (0, \ldots, 0, -1)$, and deduce that $S^n \setminus \{S\}$ is also homeomorphic to $\mathbb{R}^n$.

Exercise 3.1.4. The two charts $\{\varphi_S, \varphi_N\}$ form a smooth atlas for S^n, compatible with the one defined above.

The atlases $\{\varphi_i^\pm\}$ and $\{\varphi_S, \varphi_N\}$ define the same smooth structure on S^n.

Remark 3.1.5. The circle S^1 is quite special: we can identify $\mathbb{C}$ with $\mathbb{R}^2$ and write $S^1 = \{e^{i\theta} \mid \theta \in \mathbb{R}\}$. The universal covering $\mathbb{R} \to S^1$, $\theta \mapsto e^{i\theta}$ is of course not injective, but it furnishes an atlas of natural charts when restricted to the open segments (a, b) with $b - a < 2\pi$. The transition maps are translations.

3.1.3. Projective spaces. We now consider the real projective space $\mathbb{RP}^n$. Recall the every point in $\mathbb{RP}^n$ has some homogeneous coordinates $[x_0, \ldots, x_n]$.

For $i = 0, \ldots, n$ we set $U_i \subset \mathbb{RP}^n$ to be the open subset defined by the inequality $x_i \neq 0$. We define a chart $\varphi_i\colon U_i \to \mathbb{R}^n$ by setting

$$\varphi_i\big([x_0, \ldots, x_n]\big) = \left(\frac{x_0}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_n}{x_i}\right).$$

The inverse parametrisation $\varphi_i^{-1}\colon \mathbb{R}^n \to U_i$ may be written simply as

$$\varphi_i^{-1}(x_1, \ldots, x_n) = [x_1, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n].$$

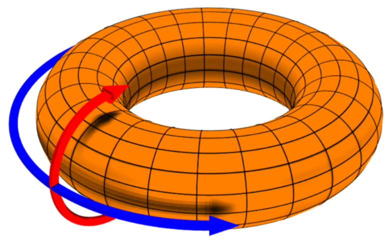

Figure 3.3. The torus $S^1 \times S^1$ embedded in $\mathbb{R}^3$. Every point $(e^{i\theta}, e^{i\varphi}) \in S^1 \times S^1$ of the torus may be interpreted on the figure as a point with (blue) longitude θ and (red) latitude φ. Note that the latitude and longitude behave very nicely on the torus, as opposite to the sphere where longitude is ambiguous at the poles. Cartographers would enjoy to live on a torus-shaped planet.

The open subsets $U_0, \ldots, U_n$ cover $\mathbb{RP}^n$ and the transition functions φ_{ij} are clearly smooth: hence the atlas $\{\varphi_i\}$ defines a smooth structure on $\mathbb{RP}^n$.

We have discovered that $\mathbb{RP}^n$ is naturally a smooth n-manifold. The same construction works for the *complex* projective space $\mathbb{CP}^n$ which is hence a smooth $2n$-manifold: it suffices to identify $\mathbb{C}^{n+1}$ with $\mathbb{R}^{2n+2}$ in the usual way.

Recall that $\mathbb{RP}^n$ and $\mathbb{CP}^n$ are connected and compact, see Exercise 1.4.1.

3.1.4. Products. The product $M \times N$ of two smooth manifolds M, N of dimension m, n is naturally a smooth $(m+n)$-manifold. Indeed, two smooth atlases $\{\varphi_i\}$, $\{\psi_j\}$ on M, N induce a smooth atlas $\{\varphi_i \times \psi_j\}$ on $M \times N$.

For instance the *torus* $S^1 \times S^1$ is a smooth manifold of dimension two. By the way, a 2-manifold is usually called a *surface*. The torus may be conveniently embedded in $\mathbb{R}^3$ as in Figure 3.3.

3.1.5. Alternative definition. We end this section with a slightly technical observation, that the reader may wish to skip. We note that it is not strictly necessary to priorly have a topology to define a smooth manifold structure: we can also proceed directly with atlases as follows.

Let X be any set. We define a *smooth atlas* on X to be a collection of subsets U_i covering X and of bijections $\varphi_i \colon U_i \to V_i$ onto open subsets of $\mathbb{R}^n$, such that $\varphi_i(U_i \cap U_j)$ is open for every i, j, and the transition maps $\varphi_{ij} = \varphi_j \circ \varphi_i^{-1}$ are smooth wherever they are defined.

Exercise 3.1.6. There is a unique topology on X such that every U_i is open and every $\varphi_i \colon U_i \to V_i$ is a homeomorphism. In this topology, a subset $U \subset X$ is open $\iff$ the sets $\varphi(U \cap U_i)$ are open for every i.

Therefore a smooth atlas on a set X defines a compatible topology. If this topology is Hausdorff and second-countable, this gives a smooth manifold structure on X.

3.2. Smooth maps

Every honest category of objects has its morphisms. We have defined the smooth manifolds, and we now introduce the right kind of maps between them.

We will henceforth use the following convention: if M is a given smooth manifold, we just call a *chart* on M any chart $\varphi\colon U \to V$ compatible with the smooth structure on M.

3.2.1. Definition. We say that a map $f\colon M \to N$ between two smooth manifolds is *smooth* if it is so when read along some charts. This means that for every $x \in M$ there are some charts $\varphi\colon U \to V$ and $\psi\colon W \to Z$ of M and N, with $x \in U$ and $f(U) \subset W$, such that the map

$$\psi \circ f \circ \varphi^{-1}\colon V \longrightarrow Z$$

is smooth. Note that the manifolds M and N may have different dimensions. It may be useful to visualise this definition via a commutative diagram:

$$\begin{tikzcd} U \arrow[r, "f"] \arrow[d, "\varphi"'] & W \arrow[d, "\psi"] \\ V \arrow[r, "F"'] & Z \end{tikzcd}$$

Here $F = \psi \circ f \circ \varphi^{-1}$ should be thought as "the map f read on charts".

Remark 3.2.1. If $f\colon M \to N$ is smooth then $\psi \circ f \circ \varphi^{-1}$ is also smooth for *any* charts φ and ψ as above. This is a typical situation: if something is smooth on some charts, it is so on all charts, because the transition functions are smooth and the composition of smooth maps is smooth.

A *curve* in M is a smooth map $\gamma\colon I \to M$ defined on some open interval $I \subset \mathbb{R}$, that may be bounded or unbounded. Curves play an important role in differential topology and geometry.

Exercise 3.2.2. The inclusion $S^n \hookrightarrow \mathbb{R}^{n+1}$ is a smooth map.

The space of all the smooth maps $M \to N$ is usually denoted by $C^\infty(M, N)$. We will often encounter the space $C^\infty(M, \mathbb{R})$, written as $C^\infty(M)$ for short. We note that $C^\infty(M)$ is a real commutative algebra.

3.2.2. Diffeomorphisms. A smooth map $f\colon M \to N$ is a *diffeomorphism* if it is a homeomorphism and its inverse $f^{-1}\colon N \to M$ is also smooth.

Example 3.2.3. The map $f\colon B^n \to \mathbb{R}^n$ defined as

$$f(x) = \frac{x}{\sqrt{1 - \|x\|^2}}$$

is a diffeomorphism. Its inverse is

$$g(x) = \frac{x}{\sqrt{1 + \|x\|^2}}.$$

Two manifolds M, N are *diffeomorphic* if there is a diffeomorphism $f\colon M \to N$. Being diffeomorphic is clearly an equivalence relation. The open ball of radius $r > 0$ centred at $x_0 \in \mathbb{R}^n$ is by definition

$$B(x_0, r) = \left\{x \in \mathbb{R}^n \;\middle|\; \|x - x_0\| < r\right\}.$$

Exercise 3.2.4. Any two open balls in $\mathbb{R}^n$ are diffeomorphic.

As a consequence, every open ball in $\mathbb{R}^n$ is diffeomorphic to $\mathbb{R}^n$ itself.

Exercise 3.2.5. The antipodal map $\iota\colon S^n \to S^n$, $\iota(x) = -x$ is a diffeomorphism.

Example 3.2.6. The following diffeomorphisms hold:

$$\mathbb{RP}^1 \cong S^1, \qquad \mathbb{CP}^1 \cong S^2.$$

These are obtained as compositions

$$\mathbb{RP}^1 \longrightarrow \mathbb{R} \cup \{\infty\} \longrightarrow S^1$$

$$\mathbb{CP}^1 \longrightarrow \mathbb{C} \cup \{\infty\} \longrightarrow S^2$$

where the first map sends $[x_0, x_1]$ to x_1/x_0, and the second is the stereographic projection. All the maps are clearly 1-1 and we only need to check that the composition is smooth, and with smooth inverse. Everything is obvious except near the point $[0, 1]$. In the complex case, if we take the parametrisation $z \mapsto [z, 1]$, by calculating we find that the map is

$$[z, 1] \longmapsto \frac{1}{1 + 4|z|^2}\left(4\Re z, -4\Im z, 1 - 4|z|^2\right).$$

So it is smooth and has smooth inverse.

3.3. Partitions of unity

We now introduce a powerful tool that may look quite technical at a first reading, but which will have spectacular consequences in the next pages. The general idea is that smooth functions are flexible enough to be patched altogether: one can use bump functions (see Section 1.3.5) to extend smooth maps from local to global, or to approximate continuous maps with smooth maps.

3.3.1. Definition. Let M be a smooth manifold. We say that an atlas $\{\varphi_i\colon U_i \to V_i\}$ for M is *adequate* if the open sets $\{U_i\}$ form a locally finite covering of M, we have $V_i = \mathbb{R}^n$ for all i, and the open sets $\varphi_i^{-1}(B^n)$ also form a covering of M.

We already know that M is paracompact by Proposition 1.1.5, so every open covering has a locally finite refinement. We reprove here this fact in a stronger form.

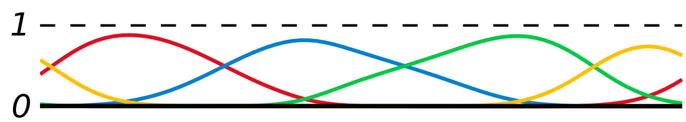

Figure 3.4. A partition of unity on S^1.

Proposition 3.3.1. *Let $\{U_i\}$ be an open covering of M. There is an adequate atlas $\{\varphi_k \colon W_k \to \mathbb{R}^n\}$ such that $\{W_k\}$ refines $\{U_i\}$.*

Proof. We readapt the proof of Proposition 1.1.5. We know that M has an exhaustion by compact subsets $\{K_j\}$, and we set $K_0 = K_{-1} = \varnothing$.

We construct the atlas inductively on $j = 1, 2\ldots$ For every $p \in K_j \setminus \mathrm{int}(K_{j-1})$ there is an open set U_i containing p. We fix a chart $\varphi_p \colon W_p \to \mathbb{R}^n$ with $p \in W_p \subset \big(\mathrm{int}(K_{j+1}) \setminus K_{j-2}\big) \cap U_i$.

The open sets $\varphi_p^{-1}(B^n)$ cover the compact set $K_j \setminus \mathrm{int}(K_{j-1})$ as p varies there, and finitely many of them suffice to cover it. By taking only these finitely many φ_p for every $j = 1, 2, \ldots$ we get an adequate covering. $\square$

Let $\{U_i\}$ be an open covering of M.

Definition 3.3.2. A *partition of unity* subordinate to the open covering $\{U_i\}$ is a family $\{\rho_i \colon M \to \mathbb{R}\}$ of smooth functions with values in $[0, 1]$, such that the following hold:

(1) the support of ρ_i is contained in U_i for all i,
(2) every $x \in M$ has a neighbourhood where all but finitely many of the ρ_i vanish, and $\sum_i \rho_i(x) = 1$.

See an example in Figure 3.4. What is important for us, is that partitions of unity exist.

Proposition 3.3.3. *For every open covering $\{U_i\}$ of M there is a partition of unity subordinate to $\{U_i\}$.*

Proof. Fix a smooth bump function $\lambda \colon \mathbb{R}^n \to \mathbb{R}$ with values in $[0, 1]$ such that $\lambda(x) = 1$ if $\|x\| \leq 1$ and $\lambda(x) = 0$ if $\|x\| \geq 2$, see Section 1.3.5.

Pick an adequate atlas $\big\{\varphi_k \colon W_k \to \mathbb{R}^n\big\}$ such that $\{W_k\}$ refines $\{U_i\}$. Define the function $\bar\rho_k \colon M \to \mathbb{R}$ as $\bar\rho_k(p) = \lambda(\varphi_k(p))$ if $p \in W_k$ and zero otherwise. The family $\{\bar\rho_k\}$ is almost a partition of unity subordinate to $\{W_k\}$, except that $\sum_j \bar\rho_j(p)$ may be any strictly positive number (note that it is not zero because the atlas is adequate). To fix this it suffices to set

$$\rho_k(p) = \frac{\bar\rho_k(p)}{\sum_j \bar\rho_j(p)}.$$

The family $\{\rho_k\}$ is a partition of unity subordinate to $\{W_k\}$. To get one $\{\eta_i\}$ subordinate to $\{U_i\}$ we fix a function $i(k)$ such that $W_k \subset U_{i(k)}$ for every k

and we define

$$\eta_i(p) = \sum_{i(k)=i} \rho_k(p).$$

The proof is complete. $\square$

3.3.2. Extension of smooth maps. We show an application of the partitions of unity. Let M and N be two smooth manifolds. The fact that we prove here is already interesting and non-trivial when M is $\mathbb{R}^m$ or some open set in it. We first need to define a notion of smooth map for arbitrary (not necessarily open) domains.

Definition 3.3.4. Let $S \subset M$ be any subset. A map $f\colon S \to N$ is *smooth* if it is locally the restriction of smooth functions. That is, for every $p \in S$ there are an open neighbourhood $U \subset M$ of p and a smooth map $g\colon U \to N$ such that $g|_{U\cap S} = f|_{U\cap S}$.

One may wonder whether the existence of local extensions implies that of a global one. This is true if the domain is closed and the codomain is $\mathbb{R}^n$.

Proposition 3.3.5. *If $S \subset M$ is a closed subset, every smooth map $f\colon S \to \mathbb{R}^n$ is the restriction of a smooth map $F\colon M \to \mathbb{R}^n$.*

Proof. By definition for every $p \in S$ there are an open neighbourhood $U(p)$ and a local extension $g_p\colon U(p) \to \mathbb{R}^n$ of f. Consider the open covering

$$\big\{U(p)\big\}_{p\in S} \cup \big\{M \setminus S\big\}$$

of M, and pick a partition of unity $\{\rho_p\} \cup \{\rho\}$ subordinate to it. For every $x \in M$ we define

$$F(x) = \sum \rho_p(x) g_p(x)$$

where the sum is taken over the finitely many $p \in M$ such that $\rho_p(x) \neq 0$. The function $F\colon M \to \mathbb{R}^n$ is locally a finite sum of smooth functions and is hence smooth. If $x \in S$ we have

$$F(x) = \sum \rho_p(x) g_p(x) = \sum \rho_p(x) f(x) = f(x) \sum \rho_p(x) = f(x).$$

Therefore $F\colon M \to \mathbb{R}^n$ is a smooth global extension of f. $\square$

Remark 3.3.6. Smooth (not even continuous) extensions cannot exist for every $S \subset M$ for obvious reasons. Take for instance $M = \mathbb{R}$ and $S = \mathbb{R}^* = \mathbb{R} \setminus \{0\}$ and $f\colon S \to \mathbb{R}$ with $f(x) = 1$ on $x > 0$ and $f(x) = 0$ on $x < 0$.

Remark 3.3.7. In the proof, the extension F vanishes outside $\cup_{p\in S} U(p)$. In the construction we may take the $U(p)$ to be arbitrarily small: hence we may require F to vanish outside of an arbitrary open neighbourhood of S.

3.3.3. Approximation of continuous maps. Here is another application of the partition of unity. Let M be a smooth manifold.

Proposition 3.3.8. *Let $f: M \to \mathbb{R}^n$ be a continuous map, whose restriction $f|_S$ to some (possibly empty) closed subset $S \subset M$ is smooth. For every continuous positive function $\varepsilon: M \to \mathbb{R}_{>0}$ there is a smooth map $g: M \to \mathbb{R}^n$ with $f(x) = g(x)$ for all $x \in S$ and $|f(x) - g(x)| < \varepsilon(x)$ for all $x \in M$.*

Proof. The map g is easily constructed locally: for every $p \in M$ there are an open neighbourhood $U(p) \subset M$ and a smooth map $g_p: U(p) \to \mathbb{R}^n$ such that $f(x) = g_p(x)$ for all $x \in U(p) \cap S$ and $|f(x) - g_p(x)| < \varepsilon(x)$ for all $x \in U(p)$. (This is proved as follows: if $p \in S$, let g_p be an extension of f, while if $p \notin S$ simply set $g_p(x) = f(p)$ constantly. The second condition is then achieved by restricting $U(p)$.)

We now paste the g_p to a global map by taking a partition of unity $\{\rho_p\}$ subordinated to $\big\{U(p)\big\}$ and defining

$$g(x) = \sum \rho_p(x) g_p(x).$$

The sum is taken over the finitely many $p \in M$. such that $\rho_p(x) \neq 0$. The map $g: M \to \mathbb{R}^n$ is smooth and $f(x) = g(x)$ for all $x \in S$. Moreover

$$\begin{aligned} |f(x) - g(x)| &= \Big| \sum \rho_p(x) f(x) - \sum \rho_p(x) g_p(x) \Big| \\ &\leq \sum \rho_p(x) \big| f(x) - g_p(x) \big| \leq \sum \rho_p(x) \varepsilon(x) = \varepsilon(x). \end{aligned}$$

The proof is complete. □

We have proved in particular that every continuous map $f: M \to \mathbb{R}^n$ may be approximated by smooth functions.

3.3.4. Smooth exhaustions. Here is another application. A *smooth exhaustion* on a manifold M is a smooth positive function $f: M \to \mathbb{R}_{>0}$ such that $f^{-1}[0, T]$ is compact for every T.

Proposition 3.3.9. *Every manifold M has a smooth exhaustion.*

Proof. Pick a locally finite covering $\big\{U_i\big\}$ where $\bar{U}_i$ is compact for every i, and a subordinated partition of unity ρ_i. The function

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

is easily seen to be a smooth exhaustion. □

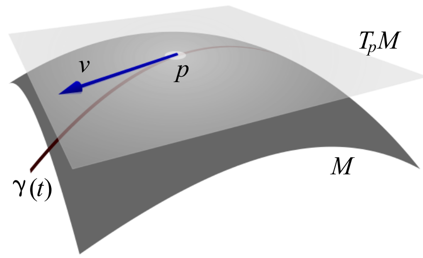

Figure 3.5. The tangent space T_pM is the set of all curves γ passing through p up to some equivalence relation.

3.4. Tangent space

Let M be a smooth n-manifold. We now define for every point $p \in M$ a n-dimensional real vector space T_pM called the *tangent space* of M at p.

Heuristically, the tangent space T_pM should generalise the intuitive notions of tangent line to a curve in $\mathbb{R}^2$ or $\mathbb{R}^3$, or of a tangent plane to a surface in $\mathbb{R}^3$, as in Figure 3.5. There is however a problem here in trying to formalise this idea: our manifold M is an abstract object and is not embedded in some bigger space like the surface in $\mathbb{R}^3$ depicted in the figure! For that reason we need to define T_pM intrinsically, using only the points that are contained *inside* M and not *outside* – since there is no outside at all. We do this by considering all the curves passing through p: as suggested in Figure 3.5, every such curve γ should define somehow a tangent vector $v \in T_pM$.

3.4.1. Definition via curves. Here is a rigorous definition of the tangent space T_pM at $p \in M$. We fix a point $p \in M$ and consider all the curves $\gamma\colon I \to M$ with $0 \in I$ and $\gamma(0) = p$. (The interval I may vary.) We want to define a notion of tangency of such curves at p. Let γ_1, γ_2 be two such curves.

If $M = \mathbb{R}^n$, the derivative $\gamma'(t)$ makes sense and we say as usual that γ_1 and γ_2 are *tangent* at p if $\gamma_1'(0) = \gamma_2'(0)$. On a more general M, we pick a chart $\varphi\colon U \to V$ and we say that γ_1 and γ_2 are *tangent* at p if the compositions $\varphi \circ \gamma_1$ and $\varphi \circ \gamma_2$ are tangent at $\varphi(p)$.[1]

This definition is chart-independent, that is it is not influenced by the choice of φ, because a transition map between two different charts transports tangent curves to tangent curves.

The tangency at p is an equivalence relation on the set of all curves $\gamma\colon I \to M$ with $\gamma(0) = p$. We are ready to define T_pM.

[1]To be precise, we may need to priorly restrict γ_1 and/or γ_2 to a smaller interval $I' \subset I$ in order for their images to lie in U.

Definition 3.4.1. The *tangent space* T_pM at $p \in M$ is the set of all curves $\gamma\colon I \to M$ with $0 \in I$ and $\gamma(0) = p$, considered up to tangency at p.

When $M = \mathbb{R}^n$, the space $T_p\mathbb{R}^n$ is naturally identified with $\mathbb{R}^n$ itself, by transforming every curve γ into its derivative $\gamma'(0)$. We will always write

$$T_p\mathbb{R}^n = \mathbb{R}^n.$$

This holds also for open subsets $M \subset \mathbb{R}^n$.

3.4.2. Definition via derivations. We now propose a more abstract and quite different definition of the tangent space at a point. It is always good to understand different equivalent definitions of the same mathematical object: the reader may choose the one she prefers, but we advise her to try to understand and remember both because, depending on the context, one definition may be more suitable than the other – for instance to prove theorems.

Let M be a smooth manifold and $p \in M$ be a point. A *derivation* v at p is an operation that assigns a number $v(f)$ to every smooth function $f\colon U \to \mathbb{R}$ defined in some open neighbourhood U of p, that fulfils the following requirements:

(1) if f and g agree on a neighbourhood of p, then $v(f) = v(g)$;
(2) v is linear, that is $v(\lambda f + \mu g) = \lambda v(f) + \mu v(g)$ for all numbers λ, μ;
(3) $v(fg) = v(f)g(p) + f(p)v(g)$.

In (2) and (3) we suppose that f and g are defined on the same open neighbourhood U. The term "derivation" is used here because the third requirement looks very much like the Leibnitz rule. Here is a fresh new definition of the tangent space at a point:

Definition 3.4.2. The tangent space T_pM is the set of all the derivations at p.

A linear combination $\lambda v + \lambda' v'$ of two derivations v, v' with $\lambda, \lambda' \in \mathbb{R}$ is again a derivation: therefore the tangent space T_pM has a natural structure of real vector space.

We study the model case $M = \mathbb{R}^n$. Here every vector $v \in \mathbb{R}^n$ determines the *directional derivative* ∂_v along v, defined as usual as

$$\partial_v f = \sum_{i=1}^n v^i \frac{\partial f}{\partial x_i},$$

which fulfils all the requirement (1-3) and is hence a derivation. Conversely:

Proposition 3.4.3. *If $M = \mathbb{R}^n$ every derivation is a directional derivative ∂_v along some vector $v \in \mathbb{R}^n$.*

Proof. We set $p = 0$ for simplicity. By the Taylor formula every smooth function f can be written near 0 as

$$f(x) = f(0) + \sum_i \frac{\partial f}{\partial x_i}(0)x_i + \sum_{i,j} h_{ij}(x)x_ix_j$$

for some smooth functions h_{ij}. If v is a derivation, by applying it to f we get

$$v(f) = f(0)v(1) + \sum_i \frac{\partial f}{\partial x_i}(0)v(x_i) + \sum_{i,j} v(h_{ij}x_ix_j).$$

The first and third term vanish because of the Leibnitz rule (exercise: use that $v(1) = v(1 \cdot 1)$). Therefore v is the partial derivative along the vector $(v(x_1), \ldots, v(x_n))$. $\square$

We have discovered that when $M = \mathbb{R}^n$ the tangent space T_pM is naturally identified with $\mathbb{R}^n$. This works also if $M \subset \mathbb{R}^n$ is an open subset.

We have shown in particular that the two definitions – via curves and via derivations – of T_pM are equivalent at least for the open subsets $M \subset \mathbb{R}^n$. On a general M, here is a direct way to pass from one definition to the other: for every curve $\gamma\colon I \to M$ with $\gamma(0) = p$, we may define a derivation v by setting

$$v(f) = (f \circ \gamma)'(0).$$

This gives indeed a 1-1 correspondence between curves up to tangency and derivations, as one can immediately deduce by taking one chart.

Summing up, we have two equivalent definitions: the one via curves may look more concrete, but derivations have the advantage of giving T_pM a natural structure of a n-dimensional real vector space.

It is important to note that T_pM is a vector space and *nothing more than that*: for instance there is no canonical norm or scalar product on T_pM, so it does not make any sense to talk about the *lengths* of tangent vectors – tangent vectors have no lengths. We are lucky enough to have a well-defined vector space and we are content with that. To define lengths we need an additional structure called *metric tensor*, that we will introduce later on in the subsequent chapters.

3.4.3. Differential of a map. We now introduce some kind of derivative of a smooth map, called *differential*. The differential is neither a number, nor a matrix of numbers in any sense: it is "only" a linear function between tangent spaces that approximates the smooth map at every point, in some sense.

Let $f\colon M \to N$ be a smooth map between smooth manifolds. The *differential* of f at a point $p \in M$ is the map

$$df_p\colon T_pM \longrightarrow T_{f(p)}N$$

that sends a curve γ with $\gamma(0) = p$ to the curve $f \circ \gamma$.

The map df_p is well-defined, because smooth maps send tangent curves to tangent curves, as one sees by taking charts. Alternatively, we may use derivations: the map df_p sends a derivation $v \in T_pM$ to the derivation $df_p(v) = v'$ that acts as $v'(g) = v(g \circ f)$.

Exercise 3.4.4. The function v' is indeed a derivation. The two definitions of df_p are equivalent; using the second one we see that df_p is linear.

The definition of df_p is clearly *functorial*, that is we have

$$d(g \circ f)_p = dg_{f(p)} \circ df_p, \qquad d(\mathrm{id}_M)_p = \mathrm{id}_{T_pM}.$$

This implies in particular that the differential df_p of a diffeomorphism $f: M \to N$ is invertible at every point $p \in M$.

When $M \subset \mathbb{R}^m$ and $N \subset \mathbb{R}^n$ are open subsets, the differential df_p of a smooth map $f: M \to N$ is a linear map

$$df_p : \mathbb{R}^m \longrightarrow \mathbb{R}^n$$

because we have the natural identifications $T_pM = \mathbb{R}^m$ and $T_{f(p)}N = \mathbb{R}^n$. It is an exercise to check that df_p is just the ordinary differential of Section 1.3.1.

3.4.4. On charts. A constant refrain in differential topology and geometry is that an abstract highly non-numerical definition becomes a more concrete numerical object when read on charts. If $\varphi: U \to V$ and $\psi: W \to Z$ are charts of M and N with $f(U) \subset W$, then we may consider the commutative diagram

$$\begin{tikzcd} U \arrow[r, "f"] \arrow[d, "\varphi"'] & W \arrow[d, "\psi"] \\ V \arrow[r, "F"'] & Z \end{tikzcd}$$

where $F = \psi \circ f \circ \varphi^{-1}$ is the map f read on charts. By taking differentials we find for every $p \in U$ another commutative diagram of linear maps

$$\begin{tikzcd} T_pM \arrow[r, "df_p"] \arrow[d, "d\varphi_p"'] & T_{f(p)}W \arrow[d, "d\psi_{f(p)}"] \\ \mathbb{R}^m \arrow[r, "dF_{\varphi(p)}"'] & \mathbb{R}^n \end{tikzcd}$$

and $dF_{\varphi(p)}$ should be thought as "the differential df_p read on charts". Note that the vertical arrows are isomorphisms, so one can fully recover df_p by looking at $dF_{\varphi(p)}$. In particular $dF_{\varphi(p)}$ has the same rank of df_p, and is injective/surjective $\Longleftrightarrow$ df_p is.

It is convenient to look at $dF_{\varphi(p)}$ because it is a rather familiar object: being the differential of a smooth map $F: V \to Z$ between open sets $V \subset \mathbb{R}^m$ and $Z \subset \mathbb{R}^n$, the differential $dF_{\varphi(p)}$ is a quite reassuring Jacobian $n \times m$ matrix whose entries vary smoothly with respect to the point $\varphi(p) \in V$.

Example 3.4.5. The Veronese embedding $f\colon \mathbb{RP}^1 \hookrightarrow \mathbb{RP}^2$ is

$$f\big([x_0, x_1]\big) = [x_0^2, x_0x_1, x_1^2],$$

see Exercise 2.5.3. The map sends the open subset $U_0 = \{x_0 \neq 0\} \subset \mathbb{RP}^1$ into $W_0 = \{x_0 \neq 0\} \subset \mathbb{RP}^2$. We use the coordinate charts $\varphi\colon U_0 \to \mathbb{R}, [1, t] \mapsto t$ and $\psi\colon W_0 \to \mathbb{R}^2, [1, t, u] \mapsto (t, u)$. Read on these charts the map f transforms into a map $F = \psi \circ f \circ \varphi^{-1}\colon \mathbb{R} \to \mathbb{R}^2$, that is

$$F(t) = (t, t^2).$$

Its differential is $(1, 2t)$, so in particular it is injective. Analogously the chart $U_1 = \{x_1 \neq 0\} \subset \mathbb{RP}^1$ injects into $W_2 = \{x_2 \neq 0\} \subset \mathbb{RP}^2$ like $t \mapsto (t^2, t)$. We have discovered that df_p is injective for every $p \in \mathbb{RP}^1$.

Exercise 3.4.6. For every k, n and every $p \in \mathbb{RP}^n$, show that the differential df_p of the Veronese embedding $f\colon \mathbb{P}^n \hookrightarrow \mathbb{P}^N$ of Exercise 2.5.3 is injective.

3.4.5. Products. Let $M \times N$ be a product of smooth manifolds of dimensions m and n. For every $(p, q) \in M \times N$ there is a natural identification

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

This identification is immediate using the definition of tangent spaces via curves, since a curve in $M \times N$ is the union of two curves in M and N.

Exercise 3.4.7. The Segre embedding $f\colon \mathbb{RP}^1 \times \mathbb{RP}^1 \hookrightarrow \mathbb{RP}^3$ is

$$[x_0, x_1] \times [y_0, y_1] \longmapsto [x_0y_0, x_0y_1, x_1y_0, x_1y_1].$$

See Section 2.1.5. Prove that for every $(p, q) \in \mathbb{RP}^1 \times \mathbb{RP}^1$ the differential $df_{(p,q)}$ is injective.

3.4.6. Velocity of a curve. If $\gamma\colon I \to M$ is a curve, for every $t \in I$ we get a differential $d\gamma_t\colon T_t\mathbb{R} \to T_{\gamma(t)}M$. Since $T_t\mathbb{R} = \mathbb{R}$ we may simply write $d\gamma_t\colon \mathbb{R} \to T_{\gamma(t)}M$ and it makes sense to define the *velocity* of γ at the time t as the tangent vector

$$\gamma'(t) = d\gamma_t(1).$$

In fact, if we use the description of T_pM via curves, this definition is rather tautological. Recall as we said above that there is no norm in $T_{\gamma(t)}M$, hence there is no way to quantify the "speed" of $\gamma'(t)$ as a number – except when it is zero.

3.4.7. Inverse Function Theorem. The Inverse Function Theorem 1.3.3 applies to this context. We say that $f\colon M \to N$ is a *local diffeomorphism* at $p \in M$ if there is an open neighbourhood $U \subset M$ of p such that $f(U) \subset N$ is open and $f|_U\colon U \to f(U)$ is a diffeomorphism.

Theorem 3.4.8. *A smooth map $f\colon M \to N$ is a local diffeomorphism at $p \in M \iff$ its differential df_p is invertible.*

Proof. Apply Theorem 1.3.3 to $\psi \circ f \circ \varphi^{-1}$ for some charts φ, ψ. $\square$

Exercise 3.4.9. Consider the map $S^n \to \mathbb{RP}^n$ that sends x to $[x]$. Prove that it is a local diffeomorphism.

3.5. Smooth coverings

In the smooth manifolds setting it is natural to consider topological coverings that are also compatible with the smooth structures, and these are called *smooth coverings*.

3.5.1. Definition. Let M and N be two smooth manifolds of the same dimension.

Definition 3.5.1. A *smooth covering* is a local diffeomorphism $f\colon M \to N$ between smooth manifolds that is also a topological covering.

For instance, the map $\mathbb{R} \to S^1, t \mapsto e^{it}$ is a smooth covering of infinite degree, and the map $S^n \to \mathbb{RP}^n$ of Exercise 3.4.9 is a smooth covering of degree two. To construct a local diffeomorphism that is not covering, pick any covering $M \to N$ (for instance, a diffeomorphism) and remove some random closed subset from the domain.

3.5.2. Surfaces. As an example, one may use a bit of complex analysis to construct many non-trivial smooth coverings between smooth surfaces.

Exercise 3.5.2. Let $p(z) \in \mathbb{C}[z]$ be a complex polynomial of some degree $d \geq 1$. Consider the set $S = \{z \in \mathbb{C} \mid p'(z) = 0\}$, that has cardinality at most $d-1$. The restriction

$$p\colon \mathbb{C} \setminus p^{-1}\big(p(S)\big) \longrightarrow \mathbb{C} \setminus p(S)$$

is a smooth covering of degree d.

For instance, the map $f(z) = z^n$ is a degree-n smooth covering $f\colon \mathbb{C}^* \to \mathbb{C}^*$ where we indicate $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$.

3.5.3. From topological to smooth coverings. Let $\tilde{M} \to M$ be a covering of topological spaces. If M has a smooth manifold structure, we already know from Exercise 1.2.3 that $\tilde{M}$ is a topological manifold; more than that:

Proposition 3.5.3. *There is a unique smooth structure on $\tilde{M}$ such that $p\colon \tilde{M} \to M$ is a smooth covering.*

Proof. For every chart $\varphi\colon U \to V$ of M and every open subset $\tilde{U} \subset \tilde{M}$ such that $p|_{\tilde{U}}\colon \tilde{U} \to U$ is a homeomorphism, we assign the chart $\varphi \circ p|_{\tilde{U}}$ to $\tilde{M}$. These charts form a smooth atlas on $\tilde{M}$ and p is a smooth covering. Conversely, since p is a local diffeomorphism the smooth structure of $\tilde{M}$ is uniquely determined (exercise). $\square$

As a consequence, much of the machinery on topological coverings summarised in Section 1.2.2 apply also to smooth coverings. For instance, if M is a connected smooth manifold, there is a bijective correspondence between the conjugacy classes of subgroups of $\pi_1(M)$ and the smooth coverings $\tilde{M} \to M$ considered up to isomorphism, where two smooth coverings $p\colon \tilde{M} \to M, p'\colon \tilde{M}' \to M$ are isomorphic if there is a diffeomorphism $f\colon \tilde{M} \to \tilde{M}'$ such that $p = p' \circ f$.

3.5.4. Smooth actions. We keep adapting the topological definitions of Section 1.2.6 to this smooth setting. A *smooth action* of a group G on a smooth manifold M is a group homomorphism

$$G \longrightarrow \mathrm{Diffeo}(M)$$

where $\mathrm{Diffeo}(M)$ is the group of all the self-diffeomorphisms $M \to M$. All the results stated there apply to this smooth setting. In particular we have the following.

Proposition 3.5.4. *Let G act smoothly, freely, and properly discontinuously on a smooth manifold M. The quotient $M/_G$ has a unique smooth structure such that $p\colon M \to M/_G$ is a smooth regular covering.*

Every smooth regular covering between smooth manifolds arise in this way.

Proof. We already know that p is a covering and $M/_G$ is a topological manifold. The smooth structure is constructed as follows: for every chart $U \to V$ on M such that $p|_U$ is injective, we add the chart $\varphi \circ p^{-1}\colon p(U) \to V$ to M. We get a smooth atlas on M because G acts smoothly. □

For instance, if M is a smooth manifold and $\iota\colon M \to M$ a fixed-point free involution (a diffeomorphism ι such that $\iota^2 = \mathrm{id}$), then $M/_\iota = M/_G$ where $G = \langle \iota \rangle$ has order two is a smooth manifold and $M \to M/_\iota$ a degree-two covering. This applies for instance to

$$\mathbb{RP}^n = S^n/_\iota$$

where ι is the antipodal map. Every degree-two covering in fact arises in this way, because every degree-two covering is regular (every index-two subgroup is normal).

3.5.5. The *n*-dimensional torus. Here is one example. Let $G = \mathbb{Z}^n$ act on $\mathbb{R}^n$ by translations, that is $g(v) = v + g$. The action is free and properly discontinuous, hence the quotient $T^n = \mathbb{R}^n/_{\mathbb{Z}^n}$ is a smooth manifold called the *n-dimensional torus*. The manifold is in fact diffeomorphic to the product

$$\underbrace{S^1 \times \cdots \times S^1}_{n}$$

via the map

$$f(x_1, \ldots, x_n) = \left(e^{2\pi x_1 i}, \ldots, e^{2\pi x_n i}\right).$$

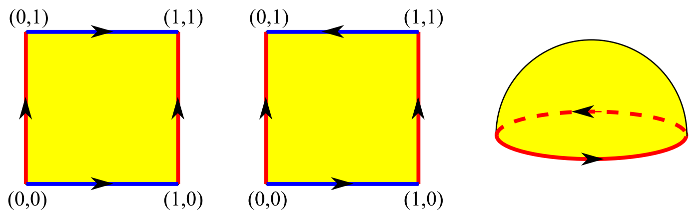

Figure 3.6. Some fundamental domains for the torus, the Klein bottle, and the projective plane. The surface is obtained from the domain by identifying the boundary curves with the same colours, respecting arrows.

The map f is defined on $\mathbb{R}^n$ but it descends to the quotient T^n, and is invertible there. The n-torus T^n is compact and its fundamental group is $\mathbb{Z}^n$.

3.5.6. Lens spaces. Let $p \geqslant 1$ and $q \geqslant 1$ be two coprime integers and define the complex number $\omega = e^{\frac{2\pi i}{p}}$. We identify $\mathbb{R}^4$ with $\mathbb{C}^2$ and see the three-dimensional sphere S^3 as

$$S^3 = \big\{(z,w) \in \mathbb{C}^2 \bigm| |z|^2 + |w|^2 = 1\big\}.$$

The map

$$f(z,w) = (\omega z, \omega^q w)$$

is a linear isomorphism of $\mathbb{C}^2$ that consists geometrically of two simultaneous rotations on the coordinate real planes $w = 0$ e $z = 0$. The map f preserves S^3, it has order p and none of its iterates $f, f^2, \ldots, f^{p-1}$ has a fixed point in S^3. Therefore the group $\Gamma = \langle f \rangle$ generated by f acts freely on S^3, and also properly discontinuously because it is finite. The quotient

$$L(p,q) = S^3/_\Gamma$$

is therefore a smooth manifold covered by S^3 called *lens space*. Its fundamental group is isomorphic to the cyclic group $\Gamma \cong \mathbb{Z}/_{p\mathbb{Z}}$. Note that the manifold depends on both p and q.

3.5.7. Fundamental domains. Let G be a group acting smoothly, freely, and properly discontinuously on a manifold M. Sometimes we can visualise the quotient manifold $M/_G$ by drawing a *fundamental domain* for the action.

A *fundamental domain* is a closed subset $D \subset M$ such that:

- every orbit intersects D in at least one point;
- every orbit intersects $\mathrm{int}(D)$ in at most one point.

For instance, Figure 3.6 shows some fundamental domains for:

- the action of $\mathbb{Z}^2$ to $\mathbb{R}^2$ via translations, yielding the torus $T = \mathbb{R}^2/_{\mathbb{Z}^2}$;
- the action of G on $\mathbb{R}^2$, yielding the *Klein bottle* $K = \mathbb{R}^2/_G$. Here G is the group of affine isometries generated by the maps

$$f(x,y) = (x+1, y), \qquad g(x,y) = \big(\tfrac{1}{2} - x, y+1\big);$$

- the action of the antipodal map ι on S^2 yielding $\mathbb{RP}^2 = S^2/_\iota$.

We will encounter the Klein bottle again in Section 3.6.5.

3.6. Orientation

Some (but not all) manifolds can be equipped with an additional structure called an *orientation*. An orientation is a way of distinguishing your left hand from your right hand, through a fixed convention that holds coherently in the whole universe you are living in.

3.6.1. Oriented manifolds. Let M be a smooth manifold. We say that a compatible atlas on M is *oriented* if all the transition functions φ_{ij} have orientation-preserving differentials. That is, for every p in the domain of φ_{ij} the differential $d(\varphi_{ij})_p$ has positive determinant, for all i, j. Note that this determinant varies smoothly on p and never vanishes because φ_{ij} is a diffeomorphism: hence if the domain is connected and the determinant is positive at one point p, it is so at every point of the domain by continuity.

Definition 3.6.1. An *orientation* on M is an equivalence class of oriented atlases (compatible with the smooth structure of M), where two oriented atlases are considered as equivalent if their union is also oriented.

There are two important issues about orientations: the first is that a manifold M may have no orientation at all (see Exercise 3.6.7 below), and the second is that an orientation for M is never unique, as the following shows.

Exercise 3.6.2. If $\mathcal{A} = \{\varphi_i\}$ is an oriented atlas for M, then $\mathcal{A}' = \{r \circ \varphi_i\}$ is also an oriented atlas, where $r\colon \mathbb{R}^n \to \mathbb{R}^n$ is a fixed reflection along some hyperplane $H \subset \mathbb{R}^n$. The two oriented atlases are not orientably compatible.

We say that the orientations on M induced by $\mathcal{A}$ and $\mathcal{A}'$ are *opposite*. If M admits some orientation, we say that M is *orientable*.

Exercise 3.6.3. The sphere S^n is orientable.

Exercise 3.6.4. If M and N are orientable, then $M \times N$ also is.

3.6.2. Tangent spaces. We now exhibit an equivalent definition of orientation that involves tangent spaces. Recall the notion of orientation for vector spaces from Section 2.6.1.

Let M be a smooth manifold. Suppose that we assign an orientation to the vector space T_pM for every $p \in M$. We say that this assignment is *locally coherent* if the following holds: for every $p \in M$ there is a chart $\varphi\colon U \to V$ with $p \in U$ whose differential $d\varphi_q\colon T_qM \to T_{\varphi(q)}\mathbb{R}^n = \mathbb{R}^n$ is orientation-preserving (that is, it sends a positive basis of T_qM to a positive one of $\mathbb{R}^n$), for all $q \in U$.

Here is a new definition of orientation on M.

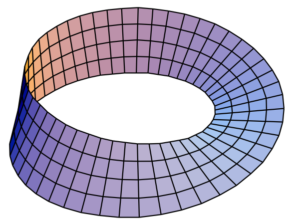

Figure 3.7. The Möbius strip is a non-orientable surface.

Definition 3.6.5. An *orientation* for M is a coherent assignment of orientations on all the tangent spaces T_pM.

We have two distinct notions of orientation on M, and we now show that they are equivalent. We see immediately how to pass from the first to the second: for every $p \in M$ there is some chart $\varphi\colon U \to V$ in the oriented atlas with $p \in U$ and we assign an orientation to T_pM by saying that a basis in T_pM is positive $\Longleftrightarrow$ its image in $\mathbb{R}^n$ along $d\varphi_p$ is. The orientation of T_pM is well-defined because it is chart-independent: every other chart of the oriented atlas differs by composition with a φ_{ij} with positive differentials. We leave to the reader as an exercise to discover how to go back from the second definition to the first.

Proposition 3.6.6. *A connected smooth manifold M has either two orientations or none.*

Proof. Let $\mathcal{A}$ be an oriented atlas, and $\mathcal{A}'$ its opposite. Suppose that we have a third oriented atlas $\mathcal{A}''$. We get a partition $M = S \sqcup S'$ where S (S') is the set of points $p \in M$ where the orientation induced by $\mathcal{A}''$ on T_pM coincides with that of $\mathcal{A}$ ($\mathcal{A}'$). Both sets S, S' are open, so either $M = S$ or $M = S'$, and hence $\mathcal{A}''$ is compatible with either $\mathcal{A}$ or $\mathcal{A}'$. $\square$

Exercise 3.6.7. The *Möbius strip* shown in Figure 3.7 is non-orientable. (A rigorous definition and proof will be exhibited soon, but it is instructive to guess why that surface is not orientable only by looking at the picture.)

3.6.3. Orientation-preserving maps. Let $f\colon M \to N$ be a local diffeomorphism between two oriented manifolds M and N. We say that f is *orientation-preserving* if the differential $df_p\colon T_pM \to T_{f(p)}N$ is an orientation-preserving isomorphism for every $p \in M$. That is, we mean that it sends positive bases to positive bases. Analogously, the map f is *orientation-reversing* if df_p is so for every $p \in M$, that is it sends positive bases to negative bases.

Exercise 3.6.8. If M is connected, every local diffeomorphism $f\colon M \to N$ between oriented manifolds is either orientation-preserving or reversing.

As a consequence, if M is connected, to understand whether $f: M \to N$ is orientation-preserving or reversing it suffices to examine df_p at any single point $p \in M$.

Exercise 3.6.9. The orthogonal reflection π along a linear hyperplane $H \subset \mathbb{R}^{n+1}$ restricts to an orientation-reversing diffeomorphism of S^n

Hint. Suppose $H = \{x_1 = 0\}$, pick $p = (0, \ldots, 0, 1)$, examine $d\pi_p$. □

Corollary 3.6.10. *The antipodal map $\iota: S^n \to S^n$ is orientation-preserving $\iff$ n is odd.*

Proof. The map ι is a composition of $n+1$ reflections along the coordinate hyperplanes. □

Remark 3.6.11. Let M be connected and oriented and $f: M \to M$ be a diffeomorphism. The condition of f being orientation-preserving or reversing is independent of the chosen orientation for M (exercise). A manifold M that admits an orientation-*reversing* diffeomorphism $M \to M$ is called *mirrorable*. For instance, the sphere S^n is mirrorable. Not all the orientable manifolds are mirrorable! This phenomenon is sometimes called *chirality*.

3.6.4. Orientability of projective spaces. We now determine whether $\mathbb{RP}^n$ is orientable or not, as a corollary of the following general fact.

Proposition 3.6.12. *Let $\pi: \tilde{M} \to M$ be a regular smooth covering of manifolds. The manifold M is orientable $\iff$ $\tilde{M}$ is orientable and all the deck transformations are orientation-preserving.*

Proof. If M is orientable, there is a locally coherent way to orient all the tangent spaces T_pM, which lifts to a locally coherent orientation of the tangent spaces $T_{\tilde{p}}\tilde{M}$, by requiring $d\pi_{\tilde{p}}$ to be orientation-preserving $\forall\, \tilde{p} \in \tilde{M}$. Every deck transformation τ is orientation preserving because $\pi \circ \tau = \pi$.

Conversely, suppose that $\tilde{M}$ is orientable and all the deck transformations are orientation-preserving. We can assign an orientation on T_pM by requiring that $d\pi_{\tilde{p}}$ be orientation-preserving for some lift $\tilde{p}$ of p: the definition is lift-independent since the deck transformations are orientation-preserving and act transitively on $\pi^{-1}(p)$ because π is regular. □

Corollary 3.6.13. *The real projective space $\mathbb{RP}^n$ is orientable $\iff$ n is odd.*

Proof. We have $\mathbb{RP}^n = S^n/_{\iota}$ and the deck transformation ι is orientation-preserving $\iff$ n is odd. □

Exercise 3.6.14. The projective plane $\mathbb{RP}^2$ contains an open subset diffeomorphic to the Möbius strip.

On the other hand, the n-torus and the lens spaces are orientable, because they are obtained by quotienting an orientable manifold ($\mathbb{R}^n$ or S^3) via

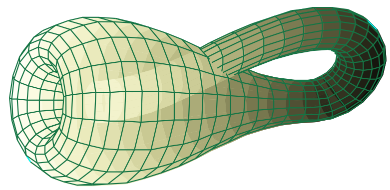

Figure 3.8. The Klein bottle immersed non-injectively in $\mathbb{R}^3$.

an group of orientation-preserving diffeomorphisms acting freely and properly discontinuously.

Example 3.6.15. We may redefine the Möbius strip as

$$S = S^1 \times (-1, 1)/_\iota$$

where ι is the involution $\iota(e^{i\theta}, t) = (e^{i(\theta+\pi)}, -t)$. The non-orientability of S is now a consequence of Proposition 3.6.12.

3.6.5. The Klein bottle. Inspired by Example 3.6.15, we now define another non-orientable surface K, called the *Klein bottle*. This is the quotient

$$K = T/_\iota$$

of the torus $T = S^1 \times S^1$ via the fixed-point free involution

$$\iota\big(e^{i\theta}, e^{i\varphi}\big) = \big(e^{i(\theta+\pi)}, e^{-i\varphi}\big).$$

Since ι is orientation-reversing, the Klein bottle is not orientable. It has infinite fundamental group $\pi_1(K)$ with an index-two normal subgroup isomorphic to $\pi_1(T) = \mathbb{Z}\times\mathbb{Z}$. This shows in particular that K is not homeomorphic to $\mathbb{RP}^2$.

We will soon see that, as opposite to the Möbius strip, the Klein bottle cannot be embedded in $\mathbb{R}^3$, and the best that we can do is to "immerse" it in $\mathbb{R}^3$ non-injectively as shown in Figure 3.8. The notions of immersion and embedding will be introduced in Section 3.8.

Exercise 3.6.16. Verify that this Klein bottle is indeed diffeomorphic to the Klein bottle already introduced in Section 3.5.7. Convince yourself that by glueing the opposite sides of the central square in Figure 3.6 you get a surface homeomorphic to that shown in Figure 3.8.

3.6.6. Orientable double cover. Non-orientable manifolds are fascinating objects, but we will see in the next chapters that it is often useful to assume that a manifold is orientable, just to make life easier. So, if you ordered an orientable manifold and you received a non-orientable one by mistake, what can you do? The best that you can do is to transform it into an orientable one by substituting it with an appropriate double cover. We now describe this operation.

We say that a manifold N is *doubly covered* by another manifold $\tilde N$ if there is a covering $\tilde N \to N$ of degree two.

Proposition 3.6.17. *Every non-orientable connected manifold M is canonically doubly covered by an orientable connected manifold $\tilde{M}$.*

Proof. We define $\tilde{M}$ as the set of all pairs (p, o) where $p \in M$ and o is an orientation for T_pM. By sending (p, o) to p we get a 2-1 map $\pi\colon \tilde{M} \to M$. We now assign to the set $\tilde{M}$ a structure of smooth connected orientable manifold and prove that π is a smooth covering.

For every chart $\varphi_i\colon U_i \to V_i$ on M we consider the set $\tilde{U}_i \subset \tilde{M}$ of all pairs (p, o) where $p \in U_i$ and o is the orientation induced by transferring back that of $\mathbb{R}^n$ via $d\varphi_p$. We also consider the map $\tilde{\varphi}_i\colon \tilde{U}_i \to V_i$, $\tilde{\varphi}_i = \varphi_i \circ \pi$. We now show that the maps

$$\tilde{\varphi}_i\colon \tilde{U}_i \longrightarrow V_i$$

constructed in this way form an oriented smooth atlas for the set $\tilde{M}$, recall the definition in Section 3.1.5.

To prove that this is an oriented smooth atlas, we first note that the sets $\tilde{U}_i$ cover $\tilde{M}$ and every $\tilde{\varphi}_i$ is a bijection. Then, we must show that for every i, j the images of $\tilde{U}_i \cap \tilde{U}_j$ along $\tilde{\varphi}_i$ and $\tilde{\varphi}_j$ are open subsets (if not empty) and the transition map $\tilde{\varphi}_{ij}$ is orientation-preservingly smooth.

We consider a point $(p, o) \in \tilde{U}_i \cap \tilde{U}_j$. The charts φ_i and φ_j both send o to the canonical orientation of $\mathbb{R}^n$, therefore the transition map φ_{ij} has positive determinant in $\varphi_i(p)$ and hence in the whole connected component W of $\varphi_i(U_i \cap U_j)$ containing $\varphi_i(p)$. This implies that $\tilde{\varphi}_i(\tilde{U}_i \cap \tilde{U}_j)$ contains the open set W. Moreover $\tilde{\varphi}_{ij}$ is orientation-preserving on W.

Now that $\tilde{M}$ is a smooth manifold, we check that π is a smooth covering: for every $p \in M$ we pick any chart $\varphi_i\colon U_i \to V_i$ with $p \in U_i$ and note that $\varphi_i' = r \circ \varphi_i$ is also a chart for any reflection r of $\mathbb{R}^n$; the two charts define two open subsets $\tilde{U}_i, \tilde{U}_i'$ of $\tilde{M}$, each projected diffeomorphically to U_i via π.

Actually, it still remains to prove that $\tilde{M}$ is connected: if it were not, it would split into two components, each diffeomorphic to M via π, but this is excluded because $\tilde{M}$ is orientable and M is not. □

For instance: the Klein bottle is covered by the torus, the projective spaces are covered by spheres, and the Möbius strip is covered by the annulus $S^1 \times (-1, 1)$, with degree two in all the cases.

Corollary 3.6.18. *Every simply connected manifold is orientable.*

Proof. A simply connected manifold has no non-trivial covering! □

Corollary 3.6.19. *The complex projective spaces $\mathbb{CP}^n$ are all orientable.*

Remark 3.6.20. The orientability of $\mathbb{CP}^n$ can be checked also by noting that $\mathbb{C}^n$ has a natural orientation and that the transition maps between the coordinate charts are holomorphic and hence orientation-preserving.

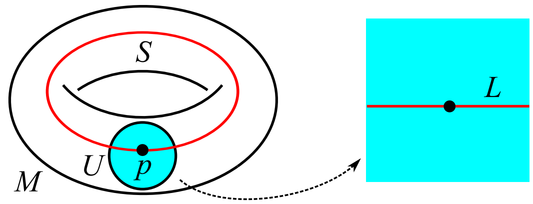

Figure 3.9. A smooth submanifold $S \subset M$ looks locally like a linear subspace $L \subset \mathbb{R}^m$.

3.7. Submanifolds

One of the fundamental aspects of smooth manifolds is that they contain plenty of manifolds of smaller dimension, called *submanifolds*.

3.7.1. Definition. Let M be a smooth m-manifold.

Definition 3.7.1. A subset $S \subset M$ is a n-dimensional *smooth submanifold* (shortly, a *n-submanifold*) if for every $p \in S$ there is a chart $\varphi\colon U \to \mathbb{R}^m$ with $p \in U$ that sends $U \cap S$ onto some linear n-subspace $L \subset \mathbb{R}^m$.

That is, the subset S looks locally like a vector n-subspace in $\mathbb{R}^m$, on some chart. Of course we must have $n \leq m$. See Figure 3.9.

A smooth submanifold $S \subset M$ is itself a smooth n-manifold: an atlas for S is obtained by restricting all the diffeomorphisms $U \to \mathbb{R}^m$ as above to $U \cap S$, composed with any linear isomorphism $L \to \mathbb{R}^n$. The transition maps are restrictions of smooth functions to linear subspaces and are hence smooth.

If we use the definition of tangent spaces via curves, we see immediately that for every $p \in S$ there is a canonical inclusion $i\colon T_pS \hookrightarrow T_pM$. Via derivations, the inclusion is $i(v)(f) = v(f|_S)$. We will see T_pS as a linear n-subspace of T_pM.

When $m = n$, a submanifold $N \subset M$ is just an open subset of M.

Example 3.7.2. Every linear subspace $L \subset \mathbb{R}^n$ is a submanifold.

Example 3.7.3. The graph S of a smooth function $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is a n-submanifold of $\mathbb{R}^n \times \mathbb{R}^m$ diffeomorphic to $\mathbb{R}^n$. The map $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$ that sends (x, y) to $\big(x, y + f(x)\big)$ is a diffeomorphism that sends the linear space $L = \{y = 0\}$ to S.

As a consequence, a subset $S \subset \mathbb{R}^n$ that is locally the graph of some smooth function is a submanifold. For instance, the sphere $S^n \subset \mathbb{R}^{n+1}$ can be seen locally at every point (up to permuting the coordinates) as the graph of the smooth function $x \mapsto \sqrt{1 - \|x\|^2}$ and is hence a n-submanifold in $\mathbb{R}^{n+1}$.

If $S \subset \mathbb{R}^n$ is a k-submanifold, the tangent space T_pS at a point $p \in S$ may be represented very concretely as a k-dimensional vector subspace of $T_p\mathbb{R}^n = \mathbb{R}^n$.

Exercise 3.7.4. For every $p \in S^n$ we have

$$T_pS^n = p^\perp$$

where $p^\perp$ indicates the vector space orthogonal to p. (We will soon deduce this exercise from a general theorem.)

Example 3.7.5. A projective k-dimensional subspace S of $\mathbb{RP}^n$ or $\mathbb{CP}^n$ is the zero set of some homogeneous linear equations. It is a smooth submanifold, because read on each coordinate chart it becomes a linear k-subspace in $\mathbb{R}^n$ or $\mathbb{C}^n$. It is diffeomorphic to $\mathbb{RP}^k$ or $\mathbb{CP}^k$.

Exercise 3.7.6. Let M, N be smooth manifolds. For every $p \in M$ the subset $\{p\} \times N$ is a submanifold of $M \times N$ diffeomorphic to N.

3.8. Immersions, embeddings, and submersions

We now study some particular kinds of nice maps called *immersions*, *embeddings*, and *submersions*.

3.8.1. Immersions. A smooth map $f: M \to N$ between smooth manifolds of dimension m and n is an *immersion* at a point $p \in M$ if the differential

$$df_p: T_pM \longrightarrow T_{f(p)}N$$

is injective. This implies in particular that $m \leq n$.

It is a remarkable fact that every immersion may be described locally in a very simple form, on appropriate charts. This is the content of the following proposition.

Proposition 3.8.1. *Let $f: M \to N$ be an immersion at $p \in M$. There are charts $\varphi: U \to \mathbb{R}^m$ and $\psi: W \to \mathbb{R}^n$ with $p \in U \subset M$ and $f(U) \subset W \subset N$ such that $\psi \circ f \circ \varphi^{-1}(x_1, \ldots, x_m) = (x_1, \ldots, x_m, 0, \ldots, 0)$.*

The proposition can be memorised via the following commutative diagram:

(5)
$$\begin{tikzcd} U \arrow[r, "f"] \arrow[d, "\varphi"'] & W \arrow[d, "\psi"] \\ \mathbb{R}^m \arrow[r, "F"'] & \mathbb{R}^n \end{tikzcd}$$

where $F(x_1, \ldots, x_m) = (x_1, \ldots, x_m, 0, \ldots, 0)$. Read on some charts, every immersion looks like F.

Proof. We can replace M and N with any open neighbourhoods of p and $f(p)$, in particular by taking charts we may suppose that $M \subset \mathbb{R}^m$ and $N \subset \mathbb{R}^n$ are some open subsets.

We know that $df_p \colon \mathbb{R}^m \to \mathbb{R}^n$ is injective. Therefore its image L has dimension m. Choose an injective linear map $g \colon \mathbb{R}^{n-m} \to \mathbb{R}^n$ whose image is in direct sum with L and define

$$G \colon M \times \mathbb{R}^{n-m} \longrightarrow \mathbb{R}^n$$

by setting $G(x, y) = f(x) + g(y)$. Its differential at $(p, 0)$ is $dG_{(p,0)} = (df_p, g)$ and it is an isomorphism. By the Implicit Function Theorem the map G is a local diffeomorphism at $(p, 0)$. Therefore there are open neighbourhoods U_1, U_2, W of $p, 0, f(p)$ such that

$$G|_{U_1 \times U_2} \colon U_1 \times U_2 \to W$$

is a diffeomorphism, and we call ψ its inverse. Now for every $x \in U_1$ we get

$$\psi\big(f(x)\big) = \psi\big(G(x, 0)\big) = (x, 0).$$

Therefore we get the commutative diagram

$$\begin{tikzcd} U_1 \arrow[r, "f"] \arrow[d, equal] & W \arrow[d, "\psi"] \\ U_1 \arrow[r, "F"'] & U_1 \times U_2 \end{tikzcd}$$

with $F(x) = (x, 0)$ as required. To conclude, we may take neighbourhoods U_1, U_2 diffeomorphic to $\mathbb{R}^m$, $\mathbb{R}^{n-m}$ and the diagram transforms into (5). $\square$

A map $f \colon M \to N$ is an *immersion* if it is so at every $p \in M$. An immersion is locally injective because of Proposition 3.8.1, but it may not be so globally: see for instance Figure 3.10-(left).

3.8.2. Embeddings. We have discovered that an immersion has a particularly nice local behaviour. We now introduce some special type of immersions that also behave nicely globally.

Definition 3.8.2. A smooth map $f \colon M \to N$ is an *embedding* if it is an immersion and a homeomorphism onto its image.

The latter condition means that $f \colon M \to f(M)$ is a homeomorphism, so in particular f is injective. We note that f may be an injective immersion while not being a homeomorphism onto its image! A counterexample is shown in Figure 3.10-(right). We really need the "homeomorphism onto its image" condition here, injectivity is not enough for our purposes.

The importance of embeddings relies in the following.

Proposition 3.8.3. *If $f \colon M \to N$ is an embedding, then $f(M) \subset N$ is a smooth submanifold and $f \colon M \to f(M)$ a diffeomorphism.*

Figure 3.10. A non-injective immersion $S^1 \to \mathbb{R}^2$ (left) and an injective immersion $\mathbb{R} \to \mathbb{R}^2$ that is not an embedding (right).

Proof. For every $p \in M$ there are open neighbourhoods $U \subset M$, $V \subset N$ of p, $f(p)$ such that $f|_U : U \to V \cap f(M)$ is a homeomorphism.

By Proposition 3.8.1, after taking a smaller V there is a chart that sends $\big(V, V \cap f(M)\big)$ to $(\mathbb{R}^n, L)$ for some linear subspace L. Therefore $f(M)$ is a smooth submanifold, and f is a diffeomorphism onto $f(M)$. □

Figure 3.10-(right) shows that the image of an injective immersion needs not to be a submanifold. Conversely:

Exercise 3.8.4. If $S \subset N$ is a smooth submanifold, then the inclusion map $i: S \hookrightarrow N$ is an embedding.

We now look for a simple embedding criterion. Recall that a map $f: X \to Y$ is *proper* if $C \subset Y$ compact implies $f^{-1}(C) \subset X$ compact.

Exercise 3.8.5. A proper injective immersion $f: M \to N$ is an embedding.

In particular, if M is compact then f is certainly proper, and we can conclude that every injective immersion of M is an embedding. This is certainly a fairly simple embedding criterion.

Example 3.8.6. Fix two positive numbers $0 < a < b$ and consider the map $f: S^1 \times S^1 \to \mathbb{R}^3$ given by

$$f(e^{i\theta}, e^{i\varphi}) = \big((a\cos\theta + b)\cos\varphi, (a\cos\theta + b)\sin\varphi, a\sin\theta\big).$$

Using the coordinates θ and φ, the differential is

$$\begin{pmatrix} -a\sin\theta\cos\varphi & -(a\cos\theta + b)\sin\varphi \\ -a\sin\theta\sin\varphi & (a\cos\theta + b)\cos\varphi \\ a\cos\theta & 0 \end{pmatrix}$$

and it has rank two for all θ, φ. Therefore f is an injective immersion and hence an embedding since $S^1 \times S^1$ is compact. The image of f is the standard torus in space already shown in Figure 3.3.

Example 3.8.7. Let p, q be two coprime integers. The map $g: S^1 \to S^1 \times S^1$ given by

$$g(e^{i\theta}) = \big(e^{ip\theta}, e^{iq\theta}\big)$$

is injective (exercise) and its differential in the angle coordinates is $(p, q) \neq (0, 0)$. Therefore g is an embedding.

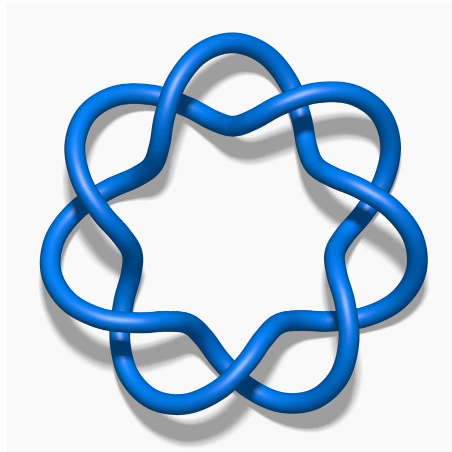

Figure 3.11. A knot is an embedding $S^1 \hookrightarrow \mathbb{R}^3$. This is a torus knot: what are the parameters p and q here?

The composition $f \circ g \colon S^1 \to \mathbb{R}^3$ with the map f of the previous example is also an embedding, and its image is called a *torus knot*: see an example in Figure 3.11. More generally, a *knot* is an embedding $S^1 \hookrightarrow \mathbb{R}^3$.

Exercise 3.8.8. Let p, q be two real numbers with irrational ratio p/q. The map $h \colon \mathbb{R} \to S^1 \times S^1$ defined by

$$h(t) = \left(e^{ipt}, e^{iqt}\right)$$

is an injective immersion but is not an embedding. Its image is in fact a dense subset of the torus.

Exercise 3.8.9. If M is compact and N is connected, and $\dim M = \dim N$, every embedding $M \to N$ is a diffeomorphism.

3.8.3. Submersions. We now describe some maps that are somehow dual to immersions. A smooth map $f \colon M \to N$ is a *submersion* at a point $p \in M$ if the differential df_p is surjective. This implies that $m \geq n$. Again, every such map has a simple local form.

Proposition 3.8.10. *Let $f \colon M \to N$ be a submersion at $p \in M$. There are charts $\varphi \colon U \to \mathbb{R}^m$ and $\psi \colon W \to \mathbb{R}^n$ with $p \in U \subset M$ and $f(U) \subset W \subset N$ such that $\psi \circ f \circ \varphi^{-1}(x_1, \ldots, x_m) = (x_1, \ldots, x_n)$.*

The proposition can be memorised via the following commutative diagram:

$$\begin{tikzcd} U \arrow[r, "f"] \arrow[d, "\varphi"'] & W \arrow[d, "\psi"] \\ \mathbb{R}^m \arrow[r, "F"'] & \mathbb{R}^n \end{tikzcd}$$

where $F(x_1, \ldots, x_m) = (x_1, \ldots, x_n)$. Read on some charts, every submersion looks like F.

Proof. The proof is very similar to that of Proposition 3.8.1. We can replace M and N with any open neighbourhoods of p and $f(p)$, in particular by taking charts we suppose that $M \subset \mathbb{R}^m$ and $N \subset \mathbb{R}^n$ are open subsets.

We know that $df_p \colon T_pM \to T_{f(p)}N$ is surjective, hence its kernel K has dimension $m-n$. Choose a linear map $g \colon \mathbb{R}^m \to \mathbb{R}^{m-n}$ that is injective on K and define

$$G \colon M \longrightarrow N \times \mathbb{R}^{m-n}$$

by setting $G(x) = \big(f(x), g(x)\big)$. Its differential at p is $dG_p = (df_p, g)$ and is an isomorphism. By the Implicit Function Theorem the map G is a local diffeomorphism at p.

Therefore there are open neighbourhoods U, W_1, W_2 of $p, f(p), 0$ such that $G(U) = W_1 \times W_2$ and $G|_U$ is a diffeomorphism. Now $f\big(G^{-1}(x, y)\big) = x$ and we conclude similarly as in the proof of Proposition 3.8.1. □

A smooth map $f \colon M \to N$ is a *submersion* if it is so at every $p \in M$.

3.8.4. Regular values. We have proved that the image of an embedding is a submanifold, and now we show that (somehow dually) the preimage of a submersion is also a submanifold. In fact, one does not really need the map to be a submersion: some weaker hypothesis suffices, that we now introduce.

Let $f \colon M \to N$ be a smooth map between manifolds of dimension $m \geq n$ respectively. A point $p \in M$ is *regular* if the differential df_p is surjective (that is if f is a submersion at p), and *critical* otherwise.

Proposition 3.8.11. *The regular points form an open subset of M.*

Proof. Read on charts, the differential df_p becomes a $n \times m$ matrix that depends smoothly on the point p. The matrices with maximum rank m form an open subset in the set of all $n \times m$ matrices. □

A point $q \in N$ is a *regular value* if the counterimage $f^{-1}(q)$ consists entirely of regular points, and it is *singular* otherwise. The map f is a submersion $\Longleftrightarrow$ all the points in the codomain are regular values.

Proposition 3.8.12. *If $q \in N$ is a regular value, then $S = f^{-1}(q)$ is either empty or a smooth $(m-n)$-submanifold. Moreover for every $p \in S$ we have*

$$T_pS = \ker df_p.$$

Proof. Thanks to Proposition 3.8.10 there are charts at p and $f(p)$ that transform f locally into a projection $\pi \colon \mathbb{R}^m \to \mathbb{R}^n$. On these charts $f^{-1}(q)$ is the linear subspace $\ker \pi$, hence a $(m-n)$-submanifold. The tangent space at p is $\ker \pi = \ker df_p$. □

Using this proposition we can re-prove that the sphere S^n is a submanifold of $\mathbb{R}^{n+1}$: pick the smooth map $f(x) = \|x\|^2$ and note that $S^n = f^{-1}(1)$. The gradient df_x is $(2x_1, \ldots, 2x_n)$, hence every non-zero point $x \in \mathbb{R}^{n+1}$ is regular for f, and therefore every non-zero point $y \in \mathbb{R}$ is a regular value: in particular 1 is regular and the proposition applies.

We can also deduce Exercise 3.7.4 quite easily: for every $x \in S^n$ we get

$$T_xS^n = \ker df_x = \ker(2x_1, \ldots, 2x_n) = x^\perp.$$

3.9. Examples

Some familiar spaces are actually smooth manifolds in a natural way. We list some of them and state a few results that will be useful in the sequel.

3.9.1. Matrix spaces. The vector space $M(m, n)$ of all $m \times n$ matrices is isomorphic to $\mathbb{R}^{mn}$ and inherits from it a structure of smooth manifold. The subset consisting of all the matrices with maximal rank is open, and is hence also a smooth manifold.

In particular, the set $M(n)$ of all the square $n \times n$ matrices is a smooth manifold, and the set $\mathrm{GL}(n, \mathbb{R})$ of all the invertible $n \times n$ matrices is a smooth manifold, both of dimension n^2. We do not forget that $M(n)$ is a vector space: hence for every $A \in M(n)$ we have a natural identification $T_AM(n) = M(n)$, and also $T_A\mathrm{GL}(n, \mathbb{R}) = M(n)$ for every $A \in \mathrm{GL}(n, \mathbb{R})$.

The subspaces $S(n)$ and $A(n)$ of all the symmetric and antisymmetric matrices are submanifolds of dimension $\frac{(n+1)n}{2}$ and $\frac{(n-1)n}{2}$ respectively.

A less trivial example is the set of $n \times n$ matrices with unit determinant:

$$\mathrm{SL}(n, \mathbb{R}) = \big\{A \in M(n) \bigm| \det A = 1\big\}.$$

Proposition 3.9.1. *The set* $\mathrm{SL}(n, \mathbb{R})$ *is a submanifold of* $M(n)$ *of codimension 1. We have*

$$T_I\mathrm{SL}(n, \mathbb{R}) = \big\{A \in M(n) \bigm| \mathrm{tr}A = 0\big\}.$$

Proof. The determinant is a smooth map $\det\colon M(n) \to \mathbb{R}$. We show that $1 \in \mathbb{R}$ is a regular value. For every $A \in \mathrm{SL}(n, \mathbb{R})$ and $B \in M(n)$ we easily get

$$\det(A + tB) = \det(I + tBA^{-1}) = 1 + t\mathrm{tr}(BA^{-1}) + o(t^2).$$

Therefore $d\det_A(B) = \mathrm{tr}(BA^{-1})$ and by taking $B = A$ we deduce that $d\det A$ is surjective. Hence 1 is a regular value, so by Proposition 3.8.12 the preimage $\mathrm{SL}(n, \mathbb{R})$ is a smooth submanifold and $T_I\mathrm{SL}(n, \mathbb{R}) = \ker d\det_I$ is as stated. $\square$

3.9.2. Orthogonal matrices. Another important example is the set of all the orthogonal matrices

$$\mathrm{O}(n) = \big\{A \in M(n) \bigm| {}^tAA = I\big\}.$$

Proposition 3.9.2. *The set* $\mathrm{O}(n)$ *is a submanifold of* $M(n)$ *of dimension* $\frac{(n-1)n}{2}$. *We have*

$$T_I\mathrm{O}(n) = A(n).$$

Proof. Consider the smooth map

$$\begin{aligned} f\colon \quad M(n) &\longrightarrow S(n), \\ A &\longmapsto {}^{\mathrm{t}}AA. \end{aligned}$$

Note that $\mathrm{O}(n) = f^{-1}(I)$. We now show that $I \in S(n)$ is a regular value. For every $A \in \mathrm{O}(n)$ we have

$$\begin{aligned} f(A+tB) &= {}^{\mathrm{t}}(A+tB)(A+tB) = {}^{\mathrm{t}}AA + t({}^{\mathrm{t}}BA + {}^{\mathrm{t}}AB) + t^2\,{}^{\mathrm{t}}BB \\ &= I + t({}^{\mathrm{t}}BA + {}^{\mathrm{t}}AB) + o(t). \end{aligned}$$

and hence

$$df_A(B) = {}^{\mathrm{t}}BA + {}^{\mathrm{t}}AB.$$

For every symmetric matrix $S \in S(n)$ there is a B such that ${}^{\mathrm{t}}BA + {}^{\mathrm{t}}AB = S$ (exercise). Therefore df_A is surjective for all $A \in \mathrm{O}(n)$ and hence I is a regular value.

We deduce from Proposition 3.8.12 that $\mathrm{O}(n) = f^{-1}(I)$ is a smooth manifold of dimension $\dim M(n) - \dim S(n) = \frac{(n-1)n}{2}$. Moreover, we have

$$T_I\mathrm{O}(n) = \ker df_I = \{B \mid {}^{\mathrm{t}}B + B = 0\} = A(n).$$

The proof is complete. □

3.9.3. Fixed rank. We now exhibit some natural submanifolds in the space $M(m,n)$ of all $m \times n$ matrices. For every $0 \leq k \leq \min\{m,n\}$, we define $M_k(m,n) \subset M(m,n)$ to be the subset consisting of all the matrices having rank k.

Proposition 3.9.3. *The subspace* $M_k(m,n)$ *is a submanifold in* $M(m,n)$ *of codimension* $(m-k)(n-k)$.

Proof. Consider a matrix $P_0 \in M_k(m,n)$. Up to permuting rows and columns, we may suppose that $P_0 = \left(\begin{smallmatrix} A_0 & B_0 \\ C_0 & D_0 \end{smallmatrix}\right)$ where $A_0 \in \mathrm{GL}(k,\mathbb{R})$.

On an open neighbourhood of P_0 every matrix P is also of this type $P = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ with $A \in \mathrm{GL}(k,\mathbb{R})$ and if we set $Q = \left(\begin{smallmatrix} A^{-1} & -A^{-1}B \\ 0 & I_{n-k} \end{smallmatrix}\right) \in \mathrm{GL}(n,\mathbb{R})$ we find

$$PQ = \begin{pmatrix} I_k & 0 \\ CA^{-1} & D - CA^{-1}B \end{pmatrix}.$$

Since $\mathrm{rk}P = \mathrm{rk}PQ$, we deduce that

$$\mathrm{rk}P = k \iff D = CA^{-1}B.$$

Therefore $M_k(m,n)$ is a manifold parametrised locally by (A,B,C), of codimension $(m-k)(n-k)$. □

3.9.4. Square roots. Let $S^+(n) \subset S(n)$ be the open subset of all positive-definite symmetric matrices. We will neeed the following.

Proposition 3.9.4. *Every $S \in S^+(n)$ has a unique square root $\sqrt{S} \in S^+(n)$, that depends smoothly on S.*

Proof. The existence and uniqueness of $\sqrt{S}$ are consequences of the spectral theorem. Smoothness may be proved by showing that the map $f: S^+(n) \to S^+(n)$, $A \mapsto A^2$ is a submersion: being a 1-1 correspondence, it is then a diffeomorphism.

To show that f is a submersion, up to conjugacy we may suppose that D is diagonal, and write

$$f(D + tM) = (D + tM)^2 = D^2 + t(DM + MD) + o(t).$$

We have

$$(DM + MD)_{ij} = D_{ii}M_{ij} + M_{ij}D_{jj} = (D_{ii} + D_{jj})M_{ij}.$$

Since $D_{ii} > 0$ for all i, if $M \neq 0$ then $DM + MD \neq 0$, so df_D is injective and hence invertible. $\square$

3.9.5. Some matrix decompositions. It is often useful to decompose a matrix into a product of matrices of some special types. Let $T(n)$ be the set of all upper triangular matrices with positive entries on the diagonal.

Proposition 3.9.5. *For every $A \in \mathrm{GL}(n, \mathbb{R})$ there are unique $O \in \mathrm{O}(n)$ and $T \in T(n)$ such that $A = OT$. Both O and T depend smoothly on A.*

Proof. Write $A = (v^1 \dots v^n)$ and orthonormalise its columns via the Gram–Schmidt algorithm to get $O = (w^1 \dots w^n)$. The algorithm may in fact be interpreted as a multiplication by some T. Conversely, if $A = OT$ then O is uniquely determined: the vector w^{i+1} must be the unit vector orthogonal to $\mathrm{Span}(v^1, \dots, v^i)$ on the same side as v^{i+1}. $\square$

Corollary 3.9.6. *We have the diffeomorphisms*

$$\mathrm{GL}(n, \mathbb{R}) \cong \mathrm{O}(n) \times T(n) \cong \mathrm{O}(n) \times \mathbb{R}^{\frac{n(n+1)}{2}}.$$

In particular there is a smooth strong deformation retraction of $\mathrm{GL}(n, \mathbb{R})$ onto the compact subset $\mathrm{O}(n)$. We also deduce a similar result for $\mathrm{SL}(n, \mathbb{R})$. Let $ST(n) \subset T(n)$ be the submanifold of all upper triangular matrices with positive entries on the diagonal and unit determinant.

Corollary 3.9.7. *We have the diffeomorphisms*

$$\mathrm{SL}(n, \mathbb{R}) \cong \mathrm{SO}(n) \times ST(n) \cong \mathrm{SO}(n) \times \mathbb{R}^{\frac{n(n+1)}{2}-1}.$$

The decomposition $M = OT$ is nice, but we will later need one that is "more invariant".

Proposition 3.9.8. *For every $A \in \mathrm{GL}(n, \mathbb{R})$ there are unique $O \in \mathrm{O}(n)$ and $S \in S^+(n)$ such that $A = OS$. Both O and S depend smoothly on A.*

Proof. Pick $S = \sqrt{{}^{\mathrm{t}}AA}$. Write $O = AS^{-1}$ and note that O is orthogonal:

$$ {}^{\mathrm{t}}OO = {}^{\mathrm{t}}S^{-1}\,{}^{\mathrm{t}}AAS^{-1} = S^{-1}S^2S^{-1} = I. $$

Conversely, if $A = OS$ then ${}^{\mathrm{t}}AA = {}^{\mathrm{t}}S\,{}^{\mathrm{t}}OOS = S^2$. □

The decomposition $A = OS$ is also known as the *polar decomposition* and is "more invariant" than $A = OT$ because it satisfies the following property:

Proposition 3.9.9. *If $A' = PAQ$ for some orthogonal matrices $P, Q \in \mathrm{O}(n)$, then the corresponding O' and S' are $O' = POQ$ and $S' = Q^{-1}SQ$.*

Proof. From $A = OS$ we deduce

$$ PAQ = (POQ)(Q^{-1}SQ). $$

Here $POQ \in \mathrm{O}(n)$ and $Q^{-1}SQ \in S^+(n)$. □

3.9.6. Connected components.

Recall that every $A \in \mathrm{O}(n)$ has $\det A = \pm 1$. We define

$$ \mathrm{SO}(n) = \big\{ A \in \mathrm{O}(n) \mid \det A = 1 \big\} $$

Proposition 3.9.10. *The manifold $\mathrm{O}(n)$ has two connected components, one of which is $\mathrm{SO}(n)$.*

Proof. We first prove that $\mathrm{SO}(n)$ is path-connected. Let R_θ be the θ-rotation 2×2 matrix. Linear algebra shows that every matrix $A \in \mathrm{SO}(n)$ is similar $A = M^{-1}BM$ via a matrix $M \in \mathrm{SO}(n)$ to a $B \in \mathrm{SO}(n)$ of type

$$ B = \begin{pmatrix} R_{\theta_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & R_{\theta_m} \end{pmatrix} \quad \text{or} \quad B = \begin{pmatrix} R_{\theta_1} & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & R_{\theta_m} & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} $$

depending on whether $n = 2m$ or $n = 2m + 1$, for some angles $\theta_1, \ldots, \theta_m$. By sending continuously the angles to zero we get a path connecting B to I_n and by conjugating everything with M we get one connecting A to I_n.

Finally, two matrices in $\mathrm{O}(n)$ with determinant 1 and -1 cannot be path-connected because the determinant is a continuous function. □

Corollary 3.9.11. *The manifold $\mathrm{GL}(n, \mathbb{R})$ has two connected components, consisting of matrices with positive and negative determinant, respectively.*

Corollary 3.9.12. *The manifold $\mathrm{SL}(n, \mathbb{R})$ is connected.*

3.9.7. Grassmannians. Let V be a real vector space of dimension n, and fix $1 \leq k \leq n$. We introduced and studied the Grassmannian $\mathrm{Gr}_k(V)$ in Section 2.5. We now show that $\mathrm{Gr}_k(V)$ has a natural smooth manifold structure.

We consider $\mathrm{Gr}_k(V)$ as a subset of $\mathbb{P}\big(\Lambda_k(V)\big)$ via the Plücker embedding.

Proposition 3.9.13. *The Grassmannian* $\mathrm{Gr}_k(V)$ *is a compact smooth submanifold of* $\mathbb{P}\big(\Lambda_k(V)\big)$ *of dimension* $(n-k)k$.

Proof. Consider any k-plane $W \in \mathrm{Gr}_k(V)$, and pick a basis $v_1, \ldots, v_k$ for W, so that in fact $W = [v_1 \wedge \ldots \wedge v_k]$ via the Plücker embedding. Complete to a basis $v_1, \ldots, v_n$ for V. Set $Z = \mathrm{Span}(v_{k+1}, \ldots, v_n)$. Then $W \oplus Z = V$.

Define the open subset $U \subset \Lambda_k(V)$ as

$$U = \big\{[T] \bigm| T \wedge v_{k+1} \wedge \ldots \wedge v_n \neq 0\big\}.$$

The open set U contains W. Clearly $U \cap \mathrm{Gr}_k(V)$ consists of all the k-subspaces W' such that $W' \oplus Z = V$.

Consider now the map

$$\begin{aligned} F\colon \underbrace{Z \times \cdots \times Z}_{k} &\longrightarrow U \\ (z_1, \ldots, z_k) &\longmapsto \big[(v_1 + z_1) \wedge \ldots \wedge (v_k + z_k)\big]. \end{aligned}$$

Linear algebra shows that F is injective and its image is $U \cap \mathrm{Gr}_k(V)$. The map F is an immersion at W (exercise: use on both sides the basis induced by $v_1, \ldots, v_n$) and F is proper (exercise). Therefore $\mathrm{Gr}_k(V)$ is a submanifold near W of dimension $k(n-k)$. Since W is generic, the subset $\mathrm{Gr}_k(V)$ is a submanifold. It is compact because it is the image of the map

$$\begin{aligned} G\colon \mathrm{O}(n) &\longrightarrow \mathbb{P}\big(\Lambda_k(V)\big) \\ A &\longmapsto \big[A^1 \wedge \ldots \wedge A^k\big] \end{aligned}$$

where A^i is the i-th column of A. The proof is complete. □

Exercise 3.9.14. Show that the Grassmannian $\mathrm{Gr}_k(V)$ is connected.

3.10. Homotopy and isotopy

There are plenty of smooth maps $M \to N$ between two given smooth manifolds, and in some cases it is natural to consider them up to some equivalence relation. We introduce here a quite mild relation called *smooth homotopy* and a stronger one, that works only for embeddings, called *isotopy*.

3.10.1. Smooth homotopy. We introduce the following notion.

Definition 3.10.1. A *smooth homotopy* between two given smooth maps $f, g\colon M \to N$ is a smooth map $F\colon M \times \mathbb{R} \to N$ such that $F(x, 0) = f(x)$ and $F(x, 1) = g(x)$ for all $x \in M$.

In general topology, a homotopy is just a continuous map $F\colon X\times[0,1]\to Y$ where X, Y are topological spaces. In this smooth setting we must (a bit reluctantly) substitute $[0,1]$ with $\mathbb{R}$ because we need the domain to be a smooth manifold. Anyway, the behaviour of $F(x,t)$ when $t\notin[0,1]$ is of no interest for us, and we may require $F(x,\cdot)$ to be constant outside that interval:

Proposition 3.10.2. *If F is a smooth homotopy between f and g, then there is another smooth homotopy F' such that $F'(x,t)$ equals $f(x)$ for all $t\leq 0$ and $g(x)$ for all $t\geq 1$.*

Proof. Take a smooth transition function $\Psi\colon\mathbb{R}\to\mathbb{R}$ as in Section 1.3.6, such that $\Psi(t)=0$ for all $t\leq 0$ and $\Psi(t)=1$ for all $t\geq 1$. Define $F'(x,t)=F\big(x,\Psi(t)\big)$. □

Two smooth maps $f,g\colon M\to N$ are *smoothly homotopic* if there is a smooth homotopy between them.

Proposition 3.10.3. *Being smoothly homotopic is an equivalence relation.*

Proof. The only non-trivial part is the transitive property. Let F be a smooth homotopy between f and g, and G be a smooth homotopy between g and h. We must glue them to an isotopy H between f and g.

To do this smoothly, we first modify F and G as in the proof of Proposition 3.10.2, taking a transition function Ψ such that $\Psi(x)=0$ for all $x\leq\frac{1}{3}$ and $\Psi(x)=1$ for all $x\geq\frac{2}{3}$. Now $F(x,\cdot)$ and $G(x,\cdot)$ are constant outside $\left[\frac{1}{3},\frac{2}{3}\right]$ and can be glued by writing

$$H(x,t)=\begin{cases} F(x,2t) & \text{for } t\leq\frac{1}{2},\\ G(x,2t-1) & \text{for } t\geq\frac{1}{2}.\end{cases}$$

The map H is smooth and the proof is complete. □

Example 3.10.4. Let M be a smooth manifold. Any two maps $f,g\colon M\to\mathbb{R}^n$ are smoothly homotopic: indeed, every $f\colon M\to\mathbb{R}^n$ is smoothly homotopic to the constant map $c(x)=0$, simply by taking

$$F(x,t)=tf(x).$$

3.10.2. Isotopy. We now introduce an enhanced version of smooth homotopy, called *isotopy*, that is nicely suited to embeddings.

Definition 3.10.5. An *isotopy* between two embeddings $f,g\colon M\to N$ is a smooth homotopy $F\colon M\times\mathbb{R}\to N$ between them, such that $F_t(x)=F(x,t)$ is an embedding $F_t\colon M\to N$ for all $t\in[0,1]$.

We can prove as above that the isotopy between embeddings is an equivalence relation. Being isotopic is much stronger than being homotopic: for instance two embeddings $f,g\colon M\to\mathbb{R}^n$ are always smoothly homotopic, but they may not be isotopic in many interesting cases.

As an example, two knots $f, g\colon S^1 \hookrightarrow \mathbb{R}^3$ may not be isotopic. The *knot theory* is an area of topology that studies precisely this phenomenon: its main (and still unachieved) goal would be to classify all knots up to isotopy in a satisfactory way.

Another interesting challenge is to study the set of all self-diffeomorphisms $M \to M$ of one fixed manifold M up to isotopy. Note that if M is compact and connected, every level F_t in one such isotopy is a diffeomorphism by Exercise 3.8.9. This is already a fundamental and non-trivial problem when $M = S^n$ is a sphere; the one-dimensional case is the only one that can be solved easily:

Proposition 3.10.6. *Every self-diffeomorphism $\varphi\colon S^1 \to S^1$ is isotopic either to the identity or to a reflection $z \mapsto \bar{z}$, depending on whether φ is orientation-preserving or not.*

Proof. Suppose that $\varphi\colon S^1 \to S^1$ is orientation-preserving. We lift φ to a map $\tilde\varphi\colon \mathbb{R} \to \mathbb{R}$ between universal covers, and note that $\tilde\varphi'(x) > 0$ for all $x \in \mathbb{R}$. Consider the map

$$\tilde F_t(x) = t\tilde\varphi(x) + (1-t)x.$$

Since $\tilde F_t(x + 2k\pi) = \tilde F_t(x) + 2k\pi$ the map descends to a map $F_t\colon S^1 \to S^1$. When $t \in [0,1]$ we get $\tilde F_t'(x) = t\tilde\varphi'(x) + (1-t) > 0$, hence each F_t is an embedding. Therefore F_t is an isotopy between id and φ. □

Here is another interesting question, that we will be able to solve in the positive in the next chapters.

Question 3.10.7. Let M be a connected n-manifold. Are two orientation-preserving embeddings $f, g\colon \mathbb{R}^n \hookrightarrow M$ always isotopic?

3.11. The Whitney embedding

We now show that every manifold may be embedded in some Euclidean space. This result was proved by Whitney in the 1930s.

3.11.1. Borel and zero-measure subsets. We start with some preliminaries that are of independent interest.

Let M be a smooth n-manifold. As in every topological space, a *Borel subset* of M is any subspace $S \subset M$ that can be constructed from the open sets through the operations of relative complement, countable unions and intersections.

Exercise 3.11.1. A subset $S \subset M$ is Borel $\Longleftrightarrow$ its image along any chart is a Borel subset of $\mathbb{R}^n$.

Let $S \subset M$ be a Borel set. Although there is no notion of measure for S, we may still say that S has *measure zero* if the image $\varphi(U \cap S)$ along any chart $\varphi\colon U \to V$ has measure zero, with respect to the Lebesgue measure in

$\mathbb{R}^n$. Note that any diffeomorphism sends zero-measure sets to zero-measure sets (Remark 1.3.6), so it suffices to check this for a set of charts covering S.

Proposition 3.11.2. *Let $f: M \to N$ be a smooth map between manifolds of dimensions m, n. If $m < n$, the image of f is a zero-measure set.*

Proof. This holds on charts by Corollary 1.3.8. □

In particular, the image of f has empty interior.

3.11.2. The compact case. We now prove that every compact manifold embeds in some Euclidean space. Not only the statement seems very strong, but its proof is actually relatively easy.

Theorem 3.11.3. *Every compact smooth manifold M embeds in some $\mathbb{R}^n$.*

Proof. Since M is compact, it has a finite adequate atlas $\{\varphi_i : U_i \to \mathbb{R}^m\}$ that consists of some k charts (see Section 3.3.1). The open subsets $V_i = \varphi_i^{-1}(B^m)$ also cover M. Let $\lambda: \mathbb{R}^m \to \mathbb{R}$ be a bump function with $\lambda(x) = 1$ if $\|x\| \leq 1$, see Section 1.3.5.

For every $i = 1, \ldots, k$ we define the smooth map $\lambda_i: M \to \mathbb{R}$ by setting $\lambda_i(p) = \lambda(\varphi_i(p))$ if $p \in U_i$ and zero otherwise. Note that $\lambda_i \equiv 1$ on V_i and $\lambda_i \equiv 0$ outside U_i. Analogously we define the smooth map $\psi_i: M \to \mathbb{R}^m$ by setting $\psi_i(p) = \lambda_i(p)\varphi_i(p)$ when $p \in U_i$ and zero otherwise.

Let $n = k(m+1)$. We define $F: M \to \mathbb{R}^n$ by setting

$$F(p) = \big(\psi_1(p), \ldots, \psi_k(p), \lambda_1(p), \ldots, \lambda_k(p)\big).$$

The codomain is indeed $\mathbb{R}^m \times \ldots \times \mathbb{R}^m \times \mathbb{R} \times \ldots \times \mathbb{R} = \mathbb{R}^n$. We now show that F is an injective immersion, and hence an embedding since M is compact.

Since the covering is adequate, for every $p \in M$ there is at least one i such that $\lambda_i = 1$ on a neighbourhood of p. In particular $\psi_i = \varphi_i$ is a local diffeomorphism at p, its differential has rank m, and hence also the differential of F has rank m. Therefore F is an immersion.

If $\lambda_i(p) = \lambda_i(q) = 1$, then $\psi_i = \varphi_i$ and therefore $\psi_i(p) = \psi_i(q)$ implies $p = q$. This shows injectivity. □

We now want to improve the theorem in two directions: we remove the compactness hypothesis, and we prove that the dimension $n = 2m+1$ suffices.

3.11.3. Immersions. Let M be a manifold of dimension m, not necessarily compact. We know from Proposition 3.3.8 that every continuous map $f: M \to \mathbb{R}^n$ into a Euclidean space can be perturbed to a smooth map. We now show that if $n \geq 2m$ the map can be perturbed to an immersion.

Theorem 3.11.4. *Let $f: M \to \mathbb{R}^n$ be a continuous map, and $n \geq 2m$. For every $\varepsilon > 0$ there is an immersion $F: M \to \mathbb{R}^n$ with $\|F(p) - f(p)\| < \varepsilon \; \forall p \in M$.*

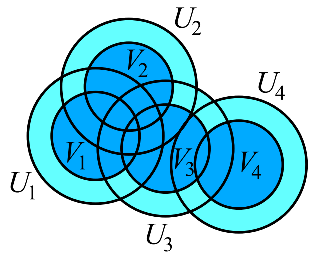

Figure 3.12. We pass from F^{i-1} to F^i by modifying the function only in U_i, with the purpose to get an immersion on $\bar{V}_i$.

Proof. By Proposition 3.3.8, we may suppose that f is smooth.

Let $\{\varphi_i : U_i \to \mathbb{R}^m\}$ be an adequate atlas, with countably many indices $i = 1, 2, \ldots$ The open subsets $V_i = \varphi_i^{-1}(B^m)$ also form a covering of M. Let $\psi_i : M \to \mathbb{R}^m$ be defined as in the proof of Theorem 3.11.3, so that $\psi_i = \varphi_i$ on V_i and $\psi_i \equiv 0$ outside U_i. We set

$$M_i = \bigcup_{j=1}^{i} V_j$$

and note that $\{\bar{M}_i\}$ is a covering of M with compact subsets.

We define a sequence $F^0, F^1, \ldots$ of maps $F^i : M \to \mathbb{R}^n$ such that:

(1) $\|F^i(p) - f(p)\| < \varepsilon$ for all $p \in M$,
(2) $F^i \equiv F^{i-1}$ outside of U_i,
(3) dF^i_p is injective for all $p \in \bar{M}_i$.

See Figure 3.12. Since $\{U_i\}$ is locally finite, the maps F^i stabilise on every compact set and converge to an immersion $F : M \to \mathbb{R}^n$ as required.

We define F^i inductively on i as follows. We set $F^0 = f$ and

$$F^i = F^{i-1} + A_i \psi_i$$

for some appropriate matrix $A = A_i \in M(n, m)$ that we now choose accurately so that the conditions (1-3) will be satisfied.

We note that F^i satisfies (2). Condition (1) is also fine as long as $\|A\|$ is sufficiently small. To get (3) we need a bit of work. By the inductive hypothesis dF^{i-1}_p is injective for all $p \in \bar{M}_{i-1}$, and it will keep being so if $\|A\|$ is sufficiently small. It remains to consider the points $p \in \bar{M}_i \setminus \bar{M}_{i-1}$.

At every $p \in \bar{V}_i$ we have $\psi_i = \varphi_i$ and

$$dF^i_p = dF^{i-1}_p + A d(\varphi_i)_p.$$

Therefore dF^i_p is not injective if and only if

$$A = B - d(F^{i-1} \circ \varphi_i^{-1})_{\varphi_i(p)}$$

for some matrix $B \in M(n, m)$ of rank $k < m$.

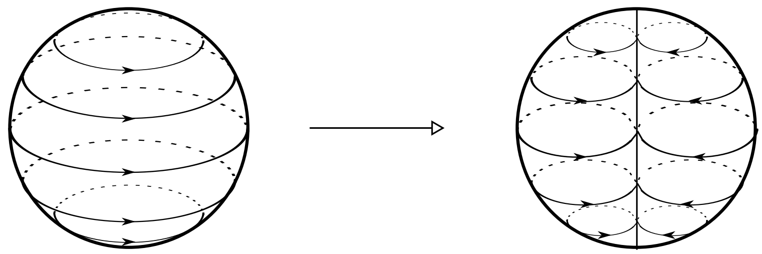

Figure 3.13. Can you perturb this continuous map $f : S^2 \to \mathbb{R}^3$ to an immersion? Probably not... At every horizontal level except the poles, the map is as in Figure 3.14 below. The map f is an immersion everywhere except at the poles, but it seems hard to eliminate the singular points at the poles just by perturbing f. If we are allowed to raise the dimension of the target, then f can certainly be perturbed to an immersion $S^2 \to \mathbb{R}^4$ and to an embedding $S^2 \to \mathbb{R}^5$ by Whitney's Theorems 3.11.4 and 3.11.7, although both perturbations may be hard to see...

By Proposition 3.9.3, the space $M_k(m, n)$ of all rank-k matrices is a manifold of dimension $mn - (m - k)(n - k)$. For every $k < m$ consider the map

$$\begin{aligned} \Psi : B^m \times M_k(n, m) &\longrightarrow M(n, m) \\ (x, B) &\longmapsto B - d(F^{i-1} \circ \varphi_i^{-1})_x. \end{aligned}$$

The dimensions of the domain and codomain are

$$m + mn - (m - k)(n - k), \qquad mn.$$

Since $n \geq 2m$ and $k \leq m - 1$ we have

$$m - (m - k)(n - k) \leq m - 1 \cdot (n - m + 1) = 2m - n - 1 < 0.$$

By Proposition 3.11.2 the image of Ψ has zero measure for all k. Therefore it suffices to pick A with small $\|A\|$ and away from these zero-measure sets. $\square$

In particular, every continuous map $\mathbb{R} \to \mathbb{R}^2$ or $S^1 \to \mathbb{R}^2$ can be perturbed to an immersion. If S is a surface, every continuous map $S \to \mathbb{R}^4$ can be perturbed to an immersion.

We cannot remove the condition $n \geq 2m$ in general. For instance, no map $S^1 \to \mathbb{R}$ can be perturbed to an immersion, because there are no immersions $S^1 \to \mathbb{R}$ at all. The dimensions $m = 2$ and $n = 3$ seem also problematic: as a challenging example, consider the continuous map $f : S^2 \to \mathbb{R}^3$ drawn in Figure 3.13. Can you perturb f to an immersion?

Remark 3.11.5. The proof of Theorem 3.11.4, especially in the choice of the matrix A, suggests that any "generic" smooth perturbation of f should be an immersion. This suggestion can be made precise by endowing the space of all maps $M \to \mathbb{R}^n$ with the appropriate topology: we do not pursue this here.

Corollary 3.11.6. *Every m-manifold M immerses in $\mathbb{R}^{2m}$.*

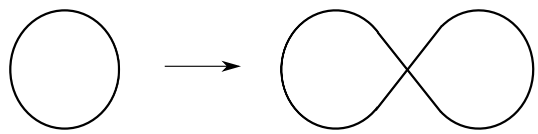

Figure 3.14. This immersion $S^1 \to \mathbb{R}^2$ cannot be perturbed to an embedding.

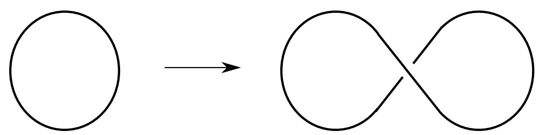

Figure 3.15. It suffices to raise the dimension of the target by one, and the immersion can now be perturbed to an injective immersion.

Proof. Pick a constant map $f\colon M \to \mathbb{R}^{2m}$ and apply Theorem 3.11.4. $\square$

3.11.4. Injective immersions. Can we perturb an immersion $M^m \to \mathbb{R}^n$ to an *injective* immersion? This may not be possible in some cases, see Figure 3.14. In fact, Figure 3.15 suggests that we could achieve injectivity just by adding one dimension to the codomain: the immersion can be perturbed to be injective in $\mathbb{R}^3$, not in $\mathbb{R}^2$. We now show that this is a general principle.

Theorem 3.11.7. *Let $f\colon M \to \mathbb{R}^n$ be an immersion, and $n \geq 2m+1$. For every $\varepsilon > 0$ there is an injective immersion $F\colon M \to \mathbb{R}^n$ with $\|F(p) - f(p)\| < \varepsilon \ \forall p \in M$.*

Proof. We adapt the proof of Theorem 3.11.4 to this context. By Proposition 3.8.1 the map f is locally injective, so by Proposition 3.3.1 we can find an adequate atlas $\{\varphi_i\colon U_i \to \mathbb{R}^m\}$ such that $f|_{U_i}$ is injective for all i.

We define again $V_i = \varphi_i^{-1}(B^m)$ and $M_i = \cup_{j\leq i} V_j$. Let $\lambda_i\colon M \to \mathbb{R}$ be a bump function with $\lambda_i \equiv 1$ on V_i and $\lambda_i \equiv 0$ outside U_i.

We now construct a sequence $F^0, F^1, \ldots$ of immersions $F^i\colon M \to \mathbb{R}^n$, that satisfy the following conditions:

(1) $\|F^i(p) - f(p)\| < \varepsilon$ for all $p \in M$,
(2) $F^i \equiv F^{i-1}$ outside of U_i,
(3) $F^i|_{U_j}$ is injective for all j,
(4) F^i is injective on $\bar{M}_i$.

Again, we conclude that F^i converge to some F, that is an injective immersion.

We set $F^0 = f$. Given F^{i-1}, we define

$$F^i = F^{i-1} + \lambda_i v_i$$

where $v = v_i \in \mathbb{R}^n$ is some vector that we now determine. If $\|v\|$ is sufficiently small, then F^i is an immersion and (1) is satisfied. Moreover (2) is automatic.

Now let $U \subset M \times M$ be the open subset

$$U = \big\{(p, q) \in M \times M \;\big|\; \lambda_i(p) \neq \lambda_i(q)\big\}.$$

We define $\Psi \colon U \to \mathbb{R}^n$ by setting

$$\Psi(p, q) = -\frac{F^{i-1}(p) - F^{i-1}(q)}{\lambda_i(p) - \lambda_i(q)}.$$

We deduce that $F^i(p) = F^i(q)$ if and only if one of the following holds:

(a) $(p, q) \in U$ and $v = \Psi(p, q)$, or
(b) $(p, q) \notin U$ and $F^{i-1}(p) = F^{i-1}(q)$.

Since $\dim U = 2m$, the image $\Psi(U)$ form a zero-measure subset and we may require that v be disjoint from it. This excludes (a) and therefore F^i is injective where F^{i-1} is injective: we get (3).

To show (4), suppose that $F^i(p) = F^i(q)$ for some $p, q \in \bar{M}_i$. We must have $\lambda_i(p) = \lambda_i(q)$ and $F^{i-1}(p) = F^{i-1}(q)$. If $\lambda_i(p) = 0$, then $p, q \in \bar{M}_{i-1}$ and we get $p = q$ by the induction hypothesis. If $\lambda_i(p) > 0$, then $p, q \in U_i$ and we get $p = q$ by the induction hypothesis again. □

3.11.5. Embeddings. We now want to make one step further, and promote injective immersions to embeddings. The following result is the main achievement of this section.

Theorem 3.11.8 (Whitney embedding Theorem). *For every smooth m-manifold M there is a proper embedding $M \hookrightarrow \mathbb{R}^{2m+1}$.*

Proof. Pick a smooth exhaustion $g \colon M \to \mathbb{R}_{>0}$ from Proposition 3.3.9 and consider the proper map $f \colon M \to \mathbb{R}^{2m+1}$, $f(p) = \big(g(p), 0, \ldots, 0\big)$. By applying Theorems 3.11.4 and 3.11.7 with any fixed $\varepsilon > 0$ we can perturb f to an injective immersion, that is easily seen to be still proper. Being proper, it is an embedding by Exercise 3.8.5. □

Concerning properness, we note the following.

Exercise 3.11.9. An embedding $i \colon M \hookrightarrow \mathbb{R}^n$ is proper $\Longleftrightarrow$ $i(M)$ is a closed subset of $\mathbb{R}^n$.

Corollary 3.11.10. *Every m-manifold M is diffeomorphic to a closed submanifold of $\mathbb{R}^{2m+1}$.*

For instance, every surface embeds properly in $\mathbb{R}^5$.

CHAPTER 4

Bundles

We introduce here a notion that is ubiquitous in modern geometry, that of a *bundle*. We start with the more general concept of *fibre bundle*, and then we turn to *vector bundles*.

4.1. Fibre bundles

In the previous chapter we have introduced the *immersions* $M \to N$, and we have proved that they behave nicely near each point $p \in M$. After that, we have discussed the enhanced notion of *embedding* that is also nice at every point $q \in N$.

Here we do more or less the same thing with *submersions*. These are maps that behave nicely at every point $p \in M$, and we would like them to be nice also at every point $q \in N$. Following this path we are led quite naturally to the notion of *fibre bundle*.

4.1.1. Definition. We work as usual in the smooth manifolds context.

Definition 4.1.1. Let F be a smooth manifold. A *smooth fibre bundle* with fibre F is a smooth map

$$\pi \colon E \longrightarrow B$$

between two smooth manifolds E, B called the *total space* and the *base space*, that satisfies the following *local triviality* condition. Every $p \in B$ has an open *trivialising* neighbourhood $U \subset B$ whose counterimage $\pi^{-1}(U)$ is diffeomorphic to a product $U \times F$, via a map $\varphi \colon \pi^{-1}(U) \to U \times F$ such that the following diagram commute:

$$\begin{tikzcd} \pi^{-1}(U) \arrow[r, "\varphi"] \arrow[d, "\pi"'] & U \times F \arrow[dl, "\pi_1"] \\ U & \end{tikzcd}$$

where $\pi_1 \colon U \times F \to U$ is the projection onto the first factor.

The definition might look slightly technical, but on the contrary is indeed very natural: in a fibre bundle $E \to B$, every fibre is diffeomorphic to F, and locally the fibration looks like a product $U \times F$ projecting onto the first factor.

Example 4.1.2. The *trivial bundle* is the product $E = B \times F$, with the projection $\pi \colon E \to B$ onto the first factor.

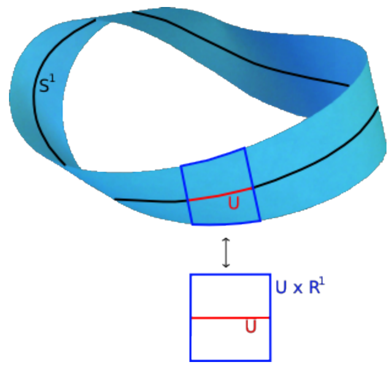

Figure 4.1. The Möbius strip is the total space of a fibre bundle with base a circle and fibre $\mathbb{R}$. Although it is locally trivial (as every fibre bundle), it is globally non-trivial: the fibre $\mathbb{R}$ makes a "twist" when transported all through the base circle.

immersion	submersion	local diffeomorphism	smooth homotopy
embedding	fibre bundle	smooth covering	isotopy

Table 4.1. We summarise here some of the most important definitions in differential topology. Every notion in the second row is an improvement of the one above.

The prototype of a non-trivial fibre bundle is the *Möbius strip* shown in Figure 4.1, which is the total space of a fibre bundle with $F = \mathbb{R}$ and $B = S^1$.

If the fibre F is diffeomorphic to the line $\mathbb{R}$, the circle S^1, the sphere S^n, the torus T, etc. we say correspondingly that E is a *line, circle, sphere*, or *torus bundle* over B. For instance, the Möbius strip is a line bundle over S^1.

Two fibre bundles $\pi\colon E \to B$ and $\pi'\colon E' \to B$ are *isomorphic* if there is a diffeomorphism $\psi\colon E \to E'$ such that $\pi = \pi' \circ \psi$. We say that a fibre bundle is *trivial* if it is isomorphic to the trivial bundle.

Remark 4.1.3. Every fibre bundle is a submersion, but not every submersion is a fibre bundle. Table 4.1 summarises some important definitions that we have introduced up to now. Recall that immersions and submersions are somehow dual notions, and every concept in the second row is an improvement of the one lying above.

Example 4.1.4. Both the torus T and the Klein bottle K are total spaces of fibre bundles over S^1 with fibre S^1. A fibration on the torus is $(e^{i\theta}, e^{i\varphi}) \mapsto e^{i\theta}$ and is clearly trivial. Recall from Section 3.6.5 that $K = T/_\iota$ with $\iota\big(e^{i\theta}, e^{i\varphi}\big) = \big(e^{i(\theta+\pi)}, e^{-i\varphi}\big)$. A fibration on the Klein bottle is $(e^{i\theta}, e^{i\varphi}) \mapsto e^{2i\theta}$. It is not trivial, because K is not diffeomorphic to $S^1 \times S^1$. See Figure 4.2.

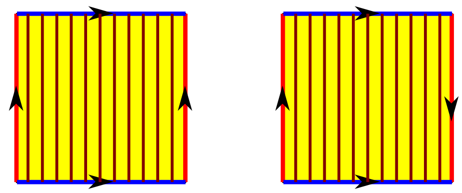

Figure 4.2. The torus and the Klein bottles are both total spaces of circle fibrations over the circle. The first is trivial, the second is not.

4.1.2. Sections. A *section* of a fibre bundle $E \to B$ is a smooth map $s \colon B \to E$ such that $\pi \circ s = \mathrm{id}_B$.

Example 4.1.5. On a trivial fibre bundle $B \times F \to B$ every map $f \colon B \to F$ determines a section $s(p) = \big(p, f(p)\big)$, and every section is obtained in this way, so sections and maps $B \to F$ are roughly the same thing.

On non-trivial bundles sections are more subtle: there are fibre bundles that have no sections at all. We will often confuse a section s with its image $s(B)$; we can do this without creating any ambiguity since $s(B)$ determines s.

Exercise 4.1.6. Show that any two sections on the Möbius strip bundle intersect. This also implies that the bundle is non-trivial.

4.2. Vector bundles

A *vector bundle* is a particular fibre bundle where every fibre has a structure of finite-dimensional real vector space. This is an extremely useful concept in differential topology and geometry.

4.2.1. Definition. A *smooth vector bundle* is a smooth fibre bundle $E \to M$ where the fibre $E_p = \pi^{-1}(p)$ of every point $p \in M$ has an additional structure of a real vector space of some dimension k, compatible with the smooth structure in the following way: every $p \in M$ must have a trivialising open neighbourhood U such that the following diagram commutes

$$\begin{tikzcd} \pi^{-1}(U) \arrow[r, "\varphi"] \arrow[d, "\pi"'] & U \times \mathbb{R}^k \arrow[dl, "\pi_1"] \\ U & \end{tikzcd}$$

via a diffeomorphism φ that sends every fibre E_p to $\mathbb{R}^k \times \{p\}$ isomorphically as vector spaces. Note that the dimensions k and n of the fibre and of M may be arbitrary.

The simplest example of a vector bundle over M is the trivial one $M \times \mathbb{R}^k$. In general, the natural number $k > 0$ is the *rank* of the vector bundle. A vector bundle with rank $k = 1$ is called a *line bundle*. Vector bundles arise quite naturally in various contexts, as we will soon see.

Exercise 4.2.1. Recall that $\mathbb{RP}^n$ may be interpreted as the space of all the vector lines $l \subset \mathbb{R}^{n+1}$. Consider the space

$$E = \left\{ (l, v) \in \mathbb{RP}^n \times \mathbb{R}^{n+1} \;\middle|\; v \in l \right\}.$$

This is a smooth $(n+1)$-submanifold of $\mathbb{RP}^n \times \mathbb{R}^{n+1}$ and the map $\pi\colon E \to \mathbb{RP}^n$ that sends (l, v) to l is a smooth line bundle with fibre $F = \mathbb{R}$, called the *tautological line bundle*.

4.2.2. Morphisms. A *morphism* between two vector bundles $E \to M$ and $E' \to M'$ is a commutative diagram

$$\begin{tikzcd} E \arrow[r, "F"] \arrow[d, "\pi"'] & E' \arrow[d, "\pi"] \\ M \arrow[r, "f"'] & M' \end{tikzcd}$$

where F and f are smooth maps, and F is a linear map on each fibre (that is $F|_{E_p}\colon E_p \to E'_{f(p)}$ is linear for each $p \in M$).

Note that the dimensions of the manifolds M, M' and of their fibres are arbitrary, so this is a quite general notion. As usual, we say that a morphism is an isomorphism if it is invertible on both sides: this is in fact equivalent to requiring that both maps f and F be diffeomorphisms.

In some cases we might prefer to consider vector bundles on a fixed base manifold M, and in that setting it is natural to consider only morphisms where f is the identity map on M.

4.2.3. The zero-section. As opposite to more general fibre bundles, every vector bundle $E \to M$ has a canonical section $s\colon M \to E$, called the *zero-section*, defined as $s(p) = 0$ where 0 is the zero in the vector space E_p, for all $p \in M$. It is convenient to identify the image $s(M)$ of the zero-section with M itself.

We will always consider the base space M embedded canonically in E through its zero-section.

4.2.4. Manipulations of vector bundles. Roughly speaking, every operation on vector spaces translates into one on vector bundles over a fixed base manifold M. For instance, given two vector bundles $E \to M$ and $E' \to M$ we may define:

- their sum $E \oplus E' \to M$,
- the dual $E^* \to M$,
- their tensor product $E \otimes E' \to M$.

To do so we simply need to perform these operations fibrewise. If E_p, E'_p are the fibres over p in E, E', then the fibre of $E \oplus E'$ is by definition $E_p \oplus E'_p$.

Of course, to complete the construction we need to build a natural smooth structure on $E \oplus E'$, and this is done as follows: if $U \times \mathbb{R}^k$ and $U \times \mathbb{R}^h$ are local trivialisations of E and E', then $U \times (\mathbb{R}^k \oplus \mathbb{R}^h)$ is a local trivialisation for $E \oplus E'$ and we equip it with the obvious product smooth structure.

The dual and tensor product bundles are defined analogously. More vector bundles may be constructed by combining these operations.

Example 4.2.2. The vector bundle $\mathrm{Hom}(E, E') \to M$ is by definition the vector bundle $E^* \otimes E' \to M$. The fiber over $p \in M$ is $\mathrm{Hom}(E_p, E'_p) = E^*_p \otimes E'_p$, see Corollary 2.1.13.

4.2.5. Subbundle and quotient bundle. The notion of vector subspace translates into that of *subbundle*. A h-dimensional *subbundle* of a given vector bundle $\pi\colon E \to M$ is a submanifold $E' \subset E$ that is also a h-dimensional vector bundle over M. That is, we require that $E'_p = E_p \cap E'$ be a vector subspace of E_p for every $p \in M$, and the projection $\pi|_{E'}\colon E' \to M$ be a vector bundle.

Example 4.2.3. The line bundle of Exercise 4.2.1 is a subbundle of the trivial bundle $\mathbb{RP}^n \times \mathbb{R}^{n+1}$.

If E' is a subbundle of E, we can define the *quotient bundle* $E/_{E'} \to M$, whose fibre over $p \in M$ is the quotient vector space $E_p/_{E'_p}$. The smooth structure may not look obvious at this point: we will return on this later in Section 4.4. The resulting maps

$$\begin{tikzcd} E' \arrow[r] \arrow[d, "\pi"'] & E \arrow[r] \arrow[d, "\pi"] & E/_{E'} \arrow[d, "\pi"] \\ M \arrow[r, "\mathrm{id}"'] & M \arrow[r, "\mathrm{id}"'] & M \end{tikzcd}$$

are bundle morphisms.

4.2.6. Restriction and pull-back. So far we have only described some manipulations of vector bundles on a fixed base manifold M. Some interesting operations arise also by varying the base manifold.

For instance we can change the base while keeping the fibres fixed: if $N \subset M$ is a submanifold, then every vector bundle $E \to M$ restricts to a vector bundle $E|_N \to N$ with the same fibres E_p in the obvious way. We call this operation the *restriction* to a submanifold. We get a bundle morphism

$$\begin{tikzcd} E|_N \arrow[r] \arrow[d, "\pi"'] & E \arrow[d, "\pi"] \\ N \arrow[r, hook] & M \end{tikzcd}$$

More generally, let $f\colon N \to M$ be any smooth map and $E \to M$ be a vector bundle. The *pull-back* of f is a new vector bundle $f^*E \to N$ constructed as follows: the total space is

$$f^*E = \big\{(p, v) \in N \times E \bigm| f(p) = \pi(v)\big\} \subset N \times E.$$

The map $\pi\colon f^*E \to N$ is $\pi(p, v) = p$. The fibre $(f^*E)_p$ over p is naturally identified with $E_{f(p)}$ and is hence a vector space.

Proposition 4.2.4. *The total space f^*E is a smooth submanifold of $N \times E$ and $f^*E \to N$ is a vector bundle.*

Proof. By restricting to a trivialising neighbourhood for E it suffices to consider the case where $N = \mathbb{R}^n$, $M = \mathbb{R}^m$, and $E = \mathbb{R}^m \times \mathbb{R}^k$. We get

$$f^*E = \big\{(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^k \;\big|\; f(x) = y\big\}.$$

Everything now follows from Example 3.7.3. $\square$

We draw the commutative diagram

$$\begin{tikzcd} f^*E \arrow[r, dotted] \arrow[d, dotted] & E \arrow[d, "\pi"] \\ N \arrow[r, "f"'] & M \end{tikzcd}$$

The dotted arrows indicate the maps that are induced by pulling-back π along f. The restriction is a particular kind of pull-back where $N \subset M$ is a submanifold and f is the inclusion map.

Exercise 4.2.5. If f is constant, then f^*E is trivial.

4.3. Tangent bundle

We now introduce the most important vector bundle on a smooth n-manifold M, the *tangent bundle*. We will also define some of its relatives, like the *cotangent*, the *normal*, and the more general *tensor bundle*.

4.3.1. Definition. Let M be a smooth manifold. As a set, the *tangent bundle* of M is the union

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

of all its tangent spaces. There is an obvious projection $\pi\colon TM \to M$ that sends T_pM to p.

The set TM has a natural structure of smooth manifold induced from that of M as follows: every chart $\varphi\colon U \to V$ of M induces an isomorphism $d\varphi_p\colon T_pM \to \mathbb{R}^n$ for every $p \in U$. Therefore it induces an overall identification $\varphi_*\colon \pi^{-1}(U) \to V \times \mathbb{R}^n$ via

$$\varphi_*(v) = \big(\varphi(p), d\varphi_p(v)\big)$$

where $p = \pi(v)$, for every $v \in \pi^{-1}(U)$. We define an atlas on TM by taking all the charts φ_* of this type. We have just defined the *tangent bundle*

$$TM \longrightarrow M$$

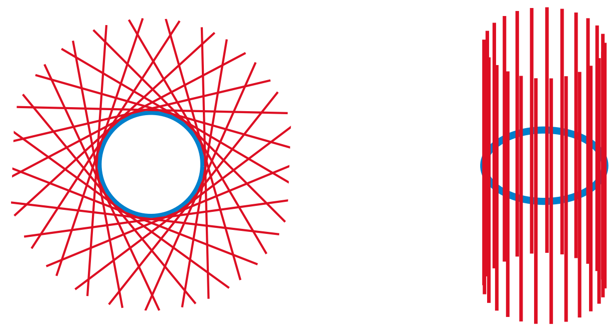

Figure 4.3. The tangent bundle of S^1 is isomorphic to the trivial one.

of M. If $\dim M = n$, then $\dim TM = 2n$. We think of M embedded in TM as the zero-section, as usual with vector bundles.

Example 4.3.1. The tangent bundle of an open subset $U \subset \mathbb{R}^n$ is canonically identified with the trivial bundle

$$TU = U \times \mathbb{R}^n$$

because every tangent space in U is canonically identified with $\mathbb{R}^n$.

More generally, we can write the tangent bundle TM of a submanifold $M \subset \mathbb{R}^n$ of any dimension $m < n$ quite explicitly:

Example 4.3.2. The tangent bundle of a submanifold $M \subset \mathbb{R}^n$ is naturally a submanifold $TM \subset \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$, defined by

$$TM = \big\{(p, v) \bigm| p \in M, v \in T_pM\big\}.$$

For instance, we have

$$TS^n = \big\{(x, v) \bigm| \|x\| = 1, v \in x^\perp\big\}.$$

Example 4.3.3. As suggested by Figure 4.3, the tangent bundle of S^1 is trivial. A bundle isomorphism $f\colon S^1 \times \mathbb{R} \to TS^1$ is the following:

$$f(e^{i\theta}, t) = \big(e^{i\theta}, te^{i(\theta + \frac{\pi}{2})}\big)$$

Is the tangent bundle of S^2 also trivial? And that of S^3?

Exercise 4.3.4. The tangent bundle TM is always an orientable manifold (even when M is not!).

Every smooth map $f\colon M \to N$ induces a morphism of tangent bundles

$$\begin{tikzcd} TM \arrow[r, "f_*"] \arrow[d, "\pi"'] & TN \arrow[d, "\pi"] \\ M \arrow[r, "f"'] & N \end{tikzcd}$$

by setting $f_*(v) = df_p(v)$ where $p = \pi(v)$ for all $v \in TM$. The restriction of f_* to each fibre T_pM is the differential $df_p\colon T_pM \to T_{f(p)}N$.

If f is a diffeomorphism, then f_* is an isomorphism.

4.3.2. Cotangent bundle. The *cotangent bundle* T^*M of a smooth manifold M is by definition the dual of the tangent bundle TM. The fibre T_p^*M at $p \in M$ is the dual of the tangent space T_pM and is called the *cotangent space* at p.

The cotangent bundle has some curious features that are lacking in the tangent bundle. One is the following: every smooth function $f\colon M \to \mathbb{R}$ induces a differential $df_p\colon T_pM \to \mathbb{R}$ at every $p \in M$, which is an element

$$df_p \in T_p^*M$$

of the *cotangent* space. We can therefore interpret the family of differentials $\{df_p\}_{p\in M}$ as a section of the cotangent bundle, and call it simply df.

We have discovered that every smooth function $f\colon M \to \mathbb{R}$ induces a section df of the cotangent bundle called its *differential*.

Remark 4.3.5. When $M = \mathbb{R}^n$, both the tangent and the cotangent space at every $p \in M$ are identified to $\mathbb{R}^n$ and the differential df is simply the gradient ∇f, that assigns a vector $(\nabla f)_p \in \mathbb{R}^n$ to every point $p \in \mathbb{R}^n$. Note however that the tangent and cotangent spaces at a point $p \in M$ are *not* canonically identified on a general smooth manifold M. A map $f\colon M \to \mathbb{R}$ induces a section of the cotangent bundle, not of the tangent bundle!

4.3.3. Normal bundle. Let M be a smooth manifold and $N \subset M$ a submanifold. We can find two natural vector bundles based on N: the tangent bundle TN and the restriction $TM|_N$ of the tangent bundle of M to N. The first is naturally a subbundle of the second, since at every $p \in N$ we have a natural inclusion $T_pN \subset T_pM$.

The *normal bundle* at N is the quotient

$$\nu N = TM|_N/_{TN}.$$

An interesting feature of the normal bundle is that the total space νN has the same dimension of the ambient space M. Indeed if $\dim M = m$ and $\dim N = n$,

$$\dim \nu N = (m-n) + n = m.$$

This preludes to an important topological application of νN that will be revealed in the next chapters.

Example 4.3.6. On a submanifold $M \subset \mathbb{R}^n$ we may use the Euclidean scalar product to identify ν_pM with $T_pM^\perp$ for every $p \in M$. We get an orthogonal decomposition

$$T_pM \oplus \nu_pM = \mathbb{R}^n$$

for every p. Therefore

$$\nu M = \big\{(p, v) \bigm| p \in M, v \in \nu_p M\big\} \subset \mathbb{R}^n \times \mathbb{R}^n.$$

For instance we have

$$\nu S^n = \big\{(x, v) \bigm| \|x\| = 1, v \in \mathrm{Span}(x)\big\}.$$

It is easy to deduce that the normal bundle of S^n inside $\mathbb{R}^{n+1}$ is trivial. Therefore we get a connected sum of bundles

$$TS^n \oplus \nu S^n = S^n \times \mathbb{R}^{n+1}$$

where two of them νS^n and $S^n \times \mathbb{R}^{n+1}$ are trivial, but the third one TS^n may not be trivial, as we will see.

4.3.4. Tensor bundle. For every $h, k \geq 0$ we may construct the *tensor bundle* $\mathcal{T}_h^k(M)$ via tensor products of the tangent and cotangent bundles:

$$\mathcal{T}_h^k(M) = \underbrace{T(M) \otimes \cdots \otimes T(M)}_{h} \otimes \underbrace{T^*(M) \otimes \cdots \otimes T^*(M)}_{k}.$$

The fiber over p is the tensor space $\mathcal{T}_h^k(T_pM)$. We define analogously the symmetric and antisymmetric tensor bundles

$$S^k(M), \qquad \Lambda^k(M)$$

whose fibres over p are $S^k(T_pM)$ and $\Lambda^k(T_pM)$. In particular $\mathcal{T}_1(M)$ is the tangent bundle and $\mathcal{T}^1(M) = S^1(M) = \Lambda^1(M)$ is the cotangent bundle. We also define the trivial tensor bundle $\mathcal{T}_0^0(M) = M \times \mathbb{R}$, coherently with the fact that a tensor of type $(0, 0)$ is just a scalar in $\mathbb{R}$.

4.4. Sections

The most important feature of vector bundles is that they contain plenty of sections. Sections are not as exoteric as they might look like: in fact, many mathematical entities that will be introduced in this book – like *vector fields*, *differential forms*, and *metric tensors* – are sections in some appropriate vector bundles, so it makes perfectly sense to study them in more detail. The effort we are making now in treating these abstract objects in full generality will be soon rewarded.

4.4.1. Vector space. Let $\pi\colon E \to M$ be a vector bundle. The space of all sections $s\colon M \to E$ is usually denoted by

$$\Gamma(E).$$

This space is naturally a vector space: the sum $s + s'$ of two sections s and s' is defined by setting $(s + s')(p) = s(p) + s'(p)$ for every $p \in M$, using the vector space structure of E_p, and the product with scalars is analogous. The zero of $\Gamma(E)$ is of course the zero-section.

Moreover, for every smooth function $f\colon M \to \mathbb{R}$ and every section s we can define a new section fs by setting $(fs)(p) = f(p)s(p)$. Therefore $\Gamma(E)$ is also a module over the ring $C^\infty(M)$.

If E and E' are two bundles over M, with sections s and s', then one can define the sections $s \oplus s'$ and $s \otimes s'$ of $E \oplus E'$ and $E \otimes E'$ in the obvious way, by setting $(s \oplus s')(p) = \big(s(p), s'(p)\big)$ and $(s \otimes s')(p) = s(p) \otimes s'(p)$.

4.4.2. Extensions of sections. We now show that vector bundles have plenty of sections, and we do this by proving that every "locally defined" section may be extended to a global one.

Let $\pi\colon E \to M$ be a vector bundle and s be a section. On a trivialising neighbourhood U, we get a diffeomorphism $\varphi\colon \pi^{-1}(U) \to U \times \mathbb{R}^k$ and hence

$$\varphi\big(s(p)\big) = \big(p, s'(p)\big)$$

for some smooth map $s'\colon U \to \mathbb{R}^k$. In other words, every smooth section s can be read as a function $s'\colon U \to \mathbb{R}^k$ on every trivalising neighbourhood U.

The fact that sections look locally like functions has some interesting consequences: for instance, we now show that sections defined only partially may be extended globally.

Let $S \subset M$ be any subset. We say that a smooth map $s\colon S \to E$ is a *partial section* if $\pi \circ s = \mathrm{id}_S$. Recall from Definition 3.3.4 the correct meaning of "smooth" here.

Proposition 4.4.1. *If $S \subset M$ is a closed subset, every partial section $s\colon S \to E$ may be extended to a global one $M \to E$.*

Proof. We adapt the proof Proposition 3.3.5 to this context. Locally, sections are like maps $U \to \mathbb{R}^k$ and can hence be extended. Therefore for every $p \in S$ there are an open trivialising neighbourhood U and a local extension $g_p\colon U_p \to E$ of s. We then proceed with a partition of unity following the same proof of Proposition 3.3.5. $\square$

Remark 4.4.2. By construction, we may suppose (if needed) that s vanishes outside of any given neighbourhood of S.

Exercise 4.4.3. Let $E \to M$ be a vector bundle of rank $k \geq 1$. If M is not a finite collection of points, the vector space $\Gamma(E)$ has infinite dimension.

4.4.3. Zeroes. Let $\pi\colon E \to M$ be a vector bundle over some smooth manifold M. We say that a section $s\colon M \to E$ *vanishes* at a point $p \in M$ if $s(p) = 0$. In that case p is called a *zero* of s. The section is *nowhere vanishing* if $s(p) \neq 0$ for all $p \in M$.

Here is one important thing to keep in mind about sections of vector bundles: although there are plenty of them, it may be hard – and sometimes impossible – to construct one that is nowhere vanishing. As an example:

Exercise 4.4.4. The Möbius strip line bundle $E \to S^1$ has no nowhere-vanishing section.

4.4.4. Frames. Let $\pi\colon E \to M$ be a rank-k vector bundle. A *frame* for π consists of k sections $s_1, \ldots, s_k$ such that the vectors $s_1(p), \ldots, s_k(p)$ are independent, and hence form a basis for E_p, for every $p \in M$.

On a frame, every s_i is in particular a nowhere-vanishing section: therefore finding a frame is even harder than constructing a nowhere-vanishing section. In fact, the following shows that frames exist only on very specific bundles.

Proposition 4.4.5. *A bundle has a frame $\Longleftrightarrow$ the bundle is trivial.*

Proof. On a trivial bundle $E = M \times \mathbb{R}^k$, the sections $s_i(p) = (p, e_i)$ with $i = 1, \ldots, k$ form a frame. Conversely, a frame $s_1, \ldots, s_k$ on $\pi\colon E \to M$ provides a bundle isomorphism $F\colon M \times \mathbb{R}^k \to E$ by writing

$$F\big(p, (\lambda_1, \ldots, \lambda_k)\big) = \lambda_1 s_1(p) + \ldots + \lambda_k s_k(p).$$

The proof is complete. $\square$

In light of this result, a frame is also called a *trivialisation* of the bundle. A nontrivial bundle $E \to M$ has no global frame, but it has many local frames: we define a *local frame* to be a frame on a trivialising open set $U \subset M$. Every trivialising open set has a local frame, induced by the trivialising chart.

4.4.5. Subbundles demystified. Frames are useful tools, for instance we use them now to clarify the notion of subbundle.

Lemma 4.4.6. *Let $E \to M$ be a bundle and $E' \subset E$ a subset. Define $E'_p = E_p \cap E'$. The following are equivalent:*

(1) *E' is a rank-h subbundle;*
(2) *every $p \in M$ has a trivialising neighbourhood U and a frame $s_1, \ldots, s_k$ for $E|_U$ such that $E'_q = \operatorname{Span}\big(s_1(q), \ldots, s_h(q)\big)$ for all $q \in U$;*

Proof. (1)$\Rightarrow$(2). Pick a neighbourhood U that trivialises both E and E'. The bundle $E|_U$ is like $U \times \mathbb{R}^k$. Since $E'|_U$ is also trivial, it has a frame $s_1, \ldots, s_h$ in U. Choose some fixed vectors $s_{h+1}, \ldots, s_k \in \mathbb{R}^n$ so that the k vectors $s_1(p), \ldots, s_h(p), s_{h+1}, \ldots, s_k$ are independent. After shrinking U, the vectors $s_1(q), \ldots, s_h(q), s_{h+1}, \ldots, s_k$ remain independent for all $q \in U$ and thus $s_1, \ldots, s_k$ is a frame for $E|_U$.

(2)$\Rightarrow$(1). The neighbourhood U trivialises also E'. $\square$

This shows in particular that a subbundle $E' \subset E$ looks locally like $U \times \mathbb{R}^h \times \{0\} \subset U \times \mathbb{R}^h \times \mathbb{R}^{k-h}$ above $U \subset M$. In particular the quotient bundle $E/_{E'}$ looks locally as $U \times \mathbb{R}^{k-h}$, and these identifications may be used to assign a smooth atlas to $E/_{E'}$, as we mentioned in Section 4.2.5.

4.4.6. Tensor fields. We now introduce the most important types of sections in differential topology and geometry: these appear everywhere, and will be ubiquitous also in this book.

Let M be a smooth manifold. A *tensor field* of type (h,k) is a section s of the tensor bundle $\mathcal{T}_h^k(M)$ of M, that is

$$s \in \Gamma\big(\mathcal{T}_h^k(M)\big).$$

In other words, we have a tensor $s(p) \in \mathcal{T}_h^k(T_pM)$ that varies smoothly with the point $p \in M$.

Since $\mathcal{T}_0^0(M) = M \times \mathbb{R}$ is the trivial line bundle, a tensor field of type $(0,0)$ is just a smooth function $s\colon M \to \mathbb{R}$.

A tensor field of type $(1,0)$ assigns a tangent vector at every point and is called a *vector field*: vector fields are extremely important in differential topology and we will study them in the next chapter with some detail.

A tensor field of type $(0,1)$ may be called a *covector field*, but the term *1-form* is more often employed. More generally, a *k-form* is a section of the antisymmetric tensor bundle $\Lambda^k(M)$. These are also important objects and we will dedicate the Chapter 7 to them.

A symmetric tensor field of type $(0,2)$ assigns a bilinear symmetric form to every tangent space: this notion will open the doors to *differential geometry*.

Most of the operations that we defined on tensors apply naturally to tensor fields. For instance, the tensor product $s \otimes s'$ of two tensor fields s and s' of type (h,k) and (h',k') is a tensor field of type $(h+h', k+k')$, and the contraction of a tensor field of type (h,k) is a tensor field of type $(h-1,k-1)$.

4.4.7. Coordinates. Let s be a tensor field of type (h,k) on M and let $\varphi\colon U \to V$ be a chart. We now want to express s in coordinates with respect to the chart φ.

As we already noticed, for every $p \in U$ the differential $d\varphi_p$ identifies the tangent space T_pM with $\mathbb{R}^n$, and we deduce from that an identification of the tensor space $\mathcal{T}_h^k(T_pM)$ with $\mathcal{T}_h^k(\mathbb{R}^n)$. The tensor field s, restricted to U, may therefore be represented as a smooth map

$$s'\colon V \longrightarrow \mathcal{T}_h^k(\mathbb{R}^n).$$

How can we write such a map? The vector space $\mathcal{T}_h^k(\mathbb{R}^n)$ has a canonical basis that consists of the elements

$$\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_h} \otimes \mathbf{e}^{j_1} \otimes \cdots \otimes \mathbf{e}^{j_k}$$

where $1 \leq i_1, \ldots, i_h, j_1, \ldots, j_k \leq n$ and $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is the canonical basis of $\mathbb{R}^n$, see Section 2.2.2. Therefore s' may be written uniquely as

$$s'(x) = s^{i_1,\ldots,i_h}_{j_1,\ldots,j_k}(x)\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_h} \otimes \mathbf{e}^{j_1} \otimes \cdots \otimes \mathbf{e}^{j_k}$$

where the coefficients vary smoothly with respect to $x \in V$. Shortly, the coordinates of s with respect to φ are the coefficients

$$s^{i_1,\dots,i_h}_{j_1,\dots,j_k}$$

that depend smoothly on a point x.

4.4.8. Changes of coordinates. If we pick another chart around a point $p \in M$, the same tensor field s is represented via some different coordinates

$$\hat{s}^{i_1,\dots,i_h}_{j_1,\dots,j_k}$$

and the transformation law relating the two different coordinates is prescribed by Proposition 2.2.11. It is convenient here to denote the coordinates of the two charts by $x_1, \dots, x_n$ and $\hat{x}_1, \dots, \hat{x}_n$ respectively, so that the differential of the transition map may be written simply as

$$\frac{\partial \hat{x}_i}{\partial x_j}.$$

The transformation law says that

$$\hat{s}^{i_1\dots i_h}_{j_1\dots j_k} = \frac{\partial \hat{x}_{i_1}}{\partial x_{l_1}} \cdots \frac{\partial \hat{x}_{i_h}}{\partial x_{l_h}} \frac{\partial x_{m_1}}{\partial \hat{x}_{j_1}} \cdots \frac{\partial x_{m_k}}{\partial \hat{x}_{j_k}} s^{l_1\dots l_h}_{m_1\dots m_k}.$$

For instance, for a vector field we have

$$\hat{s}^i = \frac{\partial \hat{x}_i}{\partial x_j} s^j$$

while for a covector field we get

$$\hat{s}_j = \frac{\partial x_i}{\partial \hat{x}_j} s_i.$$

Note that everything is designed so that every two repeated indices stay one on the top and the other on the bottom, in every formula. This is a convention that helps us to prevent mistakes; another trick consists of replacing the notations $\mathbf{e}_i$ and $\mathbf{e}^j$ with the symbols $\frac{\partial}{\partial x_i}$ and dx^j. We will explain this in the subsequent chapters.

4.5. Riemannian metric

It is sometimes useful to equip a vector bundle with some additional structure, called *Riemannian metric*. Not only this structure is interesting in its own right, but it is also useful as an auxiliary tool.

4.5.1. Definition. Let $\pi: E \to M$ be a vector bundle. Consider the bundle $E^* \otimes E^* \to M$. Remember that the fibre above $p \in M$ is the space $E_p^* \otimes E_p^*$ of all tensors on E_p of type $(0, 2)$. Remember also that scalar products are particular kinds of symmetric tensors of type $(0, 2)$.

Definition 4.5.1. A *Riemannian metric* in π is a section g of $E^* \otimes E^*$ such that $g(p)$ is a positive-definite scalar product on E_p for every $p \in M$.

In other words, a Riemannian metric is a positive-definite scalar product $g(p)$ on each fibre E_p, that varies smoothly with p. On a trivialising chart U the bundle E looks like $U \times \mathbb{R}^k$ and g can be represented concretely as a positive-definite symmetric matrix g_{ij} smoothly varying with $p \in U$.

Proposition 4.5.2. *Every vector bundle has a Riemannian metric.*

Proof. We fix an open covering U_i of trivialising sets. Above every U_i the bundle is like $U_i \times \mathbb{R}^k$, so we can identify $E_p = \mathbb{R}^k$ for every $p \in U_i$ and assign it the Euclidean scalar product, that we name $g(p)_i$.

To patch the $g(p)_i$ altogether, we pick a partition of unity $\{\rho_i\}$ subordinate to the covering. For every $p \in M$ we define

$$g(p) = \sum_i \rho_i(p) g(p)_i.$$

This is a positive-definite scalar product, because a linear combination of positive definite scalar products with positive coefficients is always a positive-definite scalar product. $\square$

Example 4.5.3. The *Euclidean metric* on the trivial bundle $M \times \mathbb{R}^k$ is the assignment of the Euclidean scalar product on every fibre $\mathbb{R}^k$.

If $E \to M$ has a Riemannian metric, then every subbundle and every restriction to a submanifold also inherits a Riemannian metric.

4.5.2. Orthonormal frames. Let $E \to M$ be a vector bundle equipped with a Riemannian metric. An *orthonormal frame* is a frame $s_1, \ldots, s_k$ where $s_1(p), \ldots, s_k(p)$ form an orthonormal basis for every $p \in M$.

Proposition 4.5.4. *Every frame transforms canonically into an orthonormal frame via the Gram – Schmidt algorithm.*

Proof. This sentence already says everything. The Gram – Schmidt algorithm transforms $s_1(p), \ldots, s_k(p)$ into k orthonormal vectors in a way that depends smoothly on p, as one can see on a chart. $\square$

Corollary 4.5.5. *A bundle has an orthonormal frame $\Longleftrightarrow$ it is trivial.*

Proof. We already know that a bundle has a frame $\Longleftrightarrow$ it is trivial. $\square$

4.5.3. Isotopies. We will soon need an appropriate notion of isotopy between bundle isomorphisms.

Let $E \to M$ and $E' \to M$ be two vector bundles, and $f, g\colon E \to E'$ be two isomorphisms. An *isotopy* between f and g is a smooth map

$$F\colon E \times \mathbb{R} \longrightarrow E'$$

such that each $F_t = F(\cdot, t)$ is an isomorphism, and $F_0 = f$, $F_1 = g$.

4.5.4. Isometries. An *isometry* between vector bundles E, E' with Riemannian metrics g, g' is an isomorphism $F\colon E \to E'$ that preserves the metric, that is with $g'\big(F(v), F(w)\big) = g(v, w)$ for all $v, w \in E_p$ and all $p \in M$.

The following proposition says that, maybe a bit surprisingly, isometry between vector bundles is not a stronger relation than isomorphism. This fact extends the well-known linear algebra theorem that says that two real vector spaces equipped with positive definite scalar products are isometric if and only if they are isomorphic.

Proposition 4.5.6. *Two isomorphic vector bundles equipped with arbitrary Riemannian metrics are always isometric, via an isometry that is isotopic to the initial isomorphism.*

Proof. We may reduce to the case where $\pi\colon E \to M$ is a vector bundle and g, g' are two arbitrary Riemannian metrics on it; we must construct an isomorphism $E \to E$ relating g and g', isotopic to the identity.

Let U be a trivialising neighbourhood. Pick two orthonormal frames s_i and s_i' for g and g' on U. We may represent every isomorphism of $E|_U$ with respect to these frames as a matrix $A(p) \in \mathrm{GL}(n, \mathbb{R})$ that depends smoothly on $p \in U$. The isomorphism is an isometry $\Longleftrightarrow A(p) \in \mathrm{O}(n)$ for every $p \in U$.

Let $A = A(p)$ represent the identity isomorphism in these basis. Use Proposition 3.9.8 to decompose A as $A = OS$ with $O \in \mathrm{O}(n)$ and $S \in S^+(n)$. The matrix $O(p)$ defines an isometry relating g and g'.

The remarkable aspect of this definition is that, by Proposition 3.9.9, the isometry defined by $O(p)$ does not depend on the orthogonal frames s_i and s_i' chosen above! Therefore by covering M with charts we get a well-defined global isometry $E \to E$ relating g and g'.

An isotopy between O and the identity is $B(p) = O(p)\big(tI + (1-t)S(p)\big)$, using that $S^+(n)$ is convex. This is well defined again by Proposition 3.9.9. $\square$

This shows in particular that every bundle $E \to M$ with any Riemannian metric g is locally Euclidean: for every trivialising subset $U \subset M$ the bundle $E|_U$ is isometric to $U \times \mathbb{R}^k$ equipped with the Euclidean metric.

4.5.5. Unitary sphere bundle. Let $\pi\colon E \to M$ be a vector bundle. Let us equip it with a Riemannian metric g. Every fibre E_p has a positive-definite scalar product $g(p)$ and hence every vector $v \in E_p$ has a *norm*

$$\|v\| = \sqrt{g(v,v)}.$$

The associated *unitary sphere bundle* is the submanifold

$$S(E) = \big\{ v \in E \;\big|\; \|v\| = 1 \big\}.$$

The projection π restricts to a projection $\pi\colon S(E) \to M$ whose fibre $S(E)_p$ is the unitary sphere in E_p.

Proposition 4.5.7. *The projection $\pi\colon S(E) \to M$ is indeed a sphere bundle. It does not depend, up to isotopy, on the chosen metric g.*

By "isotopy" we mean that the sphere bundles constructed from two metrics g and g' are related by a self-isomorphism of $E \to M$ isotopic to the identity.

Proof. We have to prove the local triviality. On a trivialising open set U the bundle E is isometric to the Euclidean $U \times \mathbb{R}^k$, so $S(E)|_U$ is like $U \times S^{k-1}$.

If we pick another metric g', we get an E' isometric to E by Proposition 4.5.6. Therefore $S(E')$ is isotopic to $S(E)$. □

4.5.6. Orthogonal bundle. Let $E \to M$ be a vector bundle equipped with a Riemannian metric. For every subbundle $E' \to M$ we have an *orthogonal bundle* $(E')^\perp \to M$, whose fiber $(E')^\perp_p$ is the orthogonal subspace to $E'_p \subset E_p$ with respect to the metric.

The orthogonal bundle is canonically isomorphic to the normal bundle $E/_{E'}$ and may be seen as a realisation of it as a subbundle of E.

Example 4.5.8. If the tangent bundle TM of a manifold M is equipped with a Riemannian metric, the normal bundle νN of any submanifold $N \subset M$ may be seen (using the metric) as a subbundle of $TM|_N$, so that we have an orthogonal sum

$$TM|_N = TN \oplus \nu N.$$

4.5.7. Dual vector bundle. Here is another instance where a Riemannian metric may be used as an auxiliary tool, to prove theorems.

Proposition 4.5.9. *Every vector bundle $E \to M$ is isomorphic to its dual $E^* \to M$.*

Proof. Pick a Riemannian metric on M. The scalar product on E_p may be used to identify E_p with its dual E^*_p as described in Section 2.3.3. This furnishes the bundle isomorphism. □

Example 4.5.10. A Riemannian metric on the tangent bundle TM determines an identification of the tangent and the cotangent bundle over M. More generally, it furnishes some bundle isomorphisms

$$\mathcal{T}^k_h(\mathbb{R}^n) \cong \mathcal{T}_{h+k}(\mathbb{R}^n) \cong \mathcal{T}^{h+k}(\mathbb{R}^n).$$

4.5.8. Shrinking vector bundles. A Riemannian metric may be used to *shrink* a vector bundle as follows. We will need this technical operation in the next chapters.

Lemma 4.5.11. *Let $E \to M$ be a vector bundle. For every neighbourhood $W \subset E$ of the zero-section M, there is an embedding $g\colon E \to W$ with*

- $g|_M = \mathrm{id}_M$,
- $g(E_p) \subset E_p$ *for every* $p \in M$.

Proof. Fix a Riemannian metric on E. Using a partition of unity, we can prove (exercise) that there is a smooth positive function $\varepsilon\colon M \to \mathbb{R}$ such that W contains all the vectors $v \in E_p$ with $\|v\| < \varepsilon(p)$, for all $p \in M$. Define

$$g(v) = \varepsilon\big(\pi(v)\big)\frac{v}{\sqrt{1+\|v\|^2}}.$$

This map fulfills the requirements. $\square$

4.5.9. Trivialising sums. The tangent bundle TS^n of a sphere is often non-trivial, but it suffices to add the normal bundle of S^n in $\mathbb{R}^{n+1}$ to get a trivial bundle, that is:

$$TS^n \oplus \nu S^n = S^n \times \mathbb{R}^{n+1}.$$

This is in fact an instance of a more general phenomenon:

Exercise 4.5.12. For any vector bundle $E \to M$ there is another vector bundle $E' \oplus M$ such that $E \oplus E' \to M$ is trivial.

TBD

4.6. Homotopy invariance

We have encountered in the previous pages a formidable tool for creating new vector bundles from old ones, the *pull-back*, that transports a bundle $E \to M$ back to $f^*E \to N$ along any smooth map $f\colon N \to M$. We now show that (if N is compact) the resulting bundle depends only on the homotopy class of f. This homotopy invariance of pull-backs has important consequences.

4.6.1. Bundle isomorphism. Let $E \to M$ and $F \to M$ be two vector bundles of the same rank r on the same manifold M. How can we tell if the two bundles are isomorphic? This is a fairly non-obvious problem in general, so for the moment we just rephrase it in a different form.

Recall that $\mathrm{Hom}(E, F) \to M$ is the vector bundle whose fiber over $p \in M$ is the space $\mathrm{Hom}(E_p, F_p)$ of all homomorphisms $E_p \to F_p$. Let $\mathrm{Isom}(E_p, F_p) \subset \mathrm{Hom}(E_p, F_p)$ be the open dense subset consisting of all *invertible* homomorphisms. Set

$$\mathrm{Isom}(E, F) = \bigcup_{p\in M} \mathrm{Isom}(E_p, F_p).$$

Proposition 4.6.1. *The subset* $\mathrm{Isom}(E,F) \subset \mathrm{Hom}(E,F)$ *is open and dense. The restriction* $\mathrm{Isom}(E,F) \to M$ *is a fibre bundle with fibre* $\mathrm{GL}(r,\mathbb{R})$.

Proof. On a trivialising set $U \subset M$ for both E and F the two bundles are both like $U \times \mathbb{R}^r$ and $\mathrm{Isom}(E,F)$ is like $U \times \mathrm{GL}(r,\mathbb{R})$. □

Note that $\mathrm{Isom}(E,F)$ is just a fibre bundle, not a vector bundle. Here is a rephrasing of the isomorphism problem. The proof is obvious.

Proposition 4.6.2. *The bundles* $E \to M$ *and* $F \to M$ *are isomorphic* $\Longleftrightarrow$ *the fibre bundle* $\mathrm{Isom}(E,F)$ *has a section.*

4.6.2. Homotopy invariance. We can now prove the invariance of pull-backs under smooth homotopies. Let $f, g \colon N \to M$ be two smooth maps between manifolds, and $E \to M$ a vector bundle.

Theorem 4.6.3. *If* N *is compact and the maps* f, g *are smoothly homotopic, the pull-back vector bundles* f^*E *and* g^*E *are isomorphic.*

Proof. Let $\Phi \colon N \times \mathbb{R} \to M$ be a smooth homotopy between f and g. Set $f_t = \Phi(\cdot, t)$ and consider the pull-back vector bundle $f_t^*E \to N$. We now show that for every $t_0 \in \mathbb{R}$ there is an $\varepsilon > 0$ such that the bundles f_t^*E for $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$ are all isomorphic to $f_{t_0}^*E$. By the compactness of $[0,1]$ we then conclude that f_0^*E and f_1^*E are isomorphic.

Consider the bundles Φ^*E and $\pi^*(f_{t_0}^*E)$ over $N \times \mathbb{R}$, where $\pi \colon N \times \mathbb{R} \to N$ is the projection. Their restrictions to $N = N \times \{t\}$ are f_t^*E and $f_{t_0}^*E$, hence they are isomorphic at $t = t_0$. Finally, consider the fibre bundle

$$\mathrm{Isom}\big(\Phi^*E, \pi^*(f_{t_0}^*E)\big) \to N \times \mathbb{R}.$$

This fibre bundle clearly has a section on $N \times \{t_0\}$. Since N is compact, the section extends to some neighbourhood $N \times (t_0 - \varepsilon, t_0 + \varepsilon)$. To show this, we can pick any extension to a global section in the vector bundle

$$\mathrm{End}\big(\Phi^*E, \pi^*(f_{t_0}^*E)\big) \to N \times \mathbb{R}$$

and note that since Isom is open in End this extended section lies entirely in Isom for small ε. By Proposition 4.6.1, the vector bundles f_t^*E and $f_{t_0}^*E$ are isomorphic $\forall t \in (t_0 - \varepsilon, t_0 + \varepsilon)$. □

4.6.3. Vector bundles over contactible manifolds. Here is an important consequence of Theorem 4.6.3.

Corollary 4.6.4. *Every vector bundle over a contractible manifold is trivial.*

Proof. Let M be contractible, that is a constant map $f \colon M \to M$ is homotopic to the identity $\mathrm{id} \colon M \to M$. Let $E \to M$ be a vector bundle. Then $E \cong f^*E$ is trivial by Exercise 4.2.5. □

This shows in particular the following.

Corollary 4.6.5. *Let $E \to M$ be a vector bundle. Any contractible open set $U \subset M$ is trivialising, that is, $E|_U \cong U \times \mathbb{R}^r$.*

For instance, every open subset $U \subset M$ that is diffeomorphic to $\mathbb{R}^n$ trivialises the bundle.

CHAPTER 5

The basic toolkit

We now introduce some fundamental notions that apply to every context in differential topology: we start with *vector fields*, their flows and Lie brackets; then we turn to *distributions*, *foliations*, and the Fobenius Theorem; finally, we introduce the two most important tools to understand embedded submanifolds, namely *tubular neighbourhoods* and *transversality*.

5.1. Vector fields

5.1.1. Definition. Let M be a smooth manifold. A section $X\colon M \to TM$ of the tangent bundle is called a *vector field*: it assigns a tangent vector $X(p) \in T_p(M)$ to every point $p \in M$ that varies smoothly with p.

Some vector fields on the torus are drawn in Figure 5.1. Recall that a zero of X is a point p such that $X(p) = 0$. Note that the vector fields in the figure have no zeroes.

Example 5.1.1. When $n = 2m - 1$ is odd, the following is a nowhere-vanishing vector field on $S^n \subset \mathbb{R}^{2m}$:

$$(x_1, \ldots, x_{2m}) \longmapsto (-x_2, x_1, \ldots, -x_{2m}, x_{2m-1}).$$

Exercise 5.1.2. Write a smooth vector field on S^n that vanishes only at the poles $(\pm 1, 0, \ldots, 0)$.

We denote by $\mathfrak{X}(M)$ the set of all the vector fields on M. Recall from Section 4.4 that $\mathfrak{X}(M) = \Gamma(TM)$ is a vector space and also a $C^\infty(M)$-module.

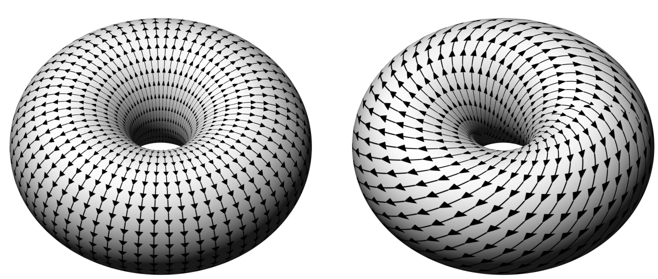

Figure 5.1. Nowhere-vanishing vector fields on the torus.

5.1.2. Diffeomorphisms. Many of the mathematical objects that we define are naturally transported along smooth maps $f: M \to N$, either from M to N or vice-versa from N to M, but this is *not* the case with vector fields: there is no meaningful way to transport a vector field along a generic map f, neither forward from M to N nor backwards from N to M.

On the other hand, every intrinsic (that is, coordinates-independent) notion can be transported in both directions if $f: M \to N$ is a diffeomorphism. In that case, every vector field X in M induces a vector field Y on N via differentials, that is by imposing:

$$Y\big(f(p)\big) = df_p\big(X(p)\big) \quad \text{for every } p \in M.$$

This gives an isomorphism between $\mathfrak{X}(M)$ and $\mathfrak{X}(N)$.

5.1.3. On charts. If X is a vector field on M and $\varphi: U \to V \subset \mathbb{R}^n$ is a chart, we can restrict X to a vector field on U and then transport it into a vector field in V. As we noticed in Section 4.4.7, the transported vector field assumes the familiar form of a smooth map $V \to \mathbb{R}^n$ because $T(V) = V \times \mathbb{R}^n$, and we may write it as a vector

$$\big(X^1(x), \ldots, X^n(x)\big)$$

in $\mathbb{R}^n$ that varies smoothly on $x \in V$. Here X^i is the i-coordinate of X in the chosen chart, a real number that depend smoothly on $x \in V$. We can use the Einstein notation and write the transported vector field in V more concisely as

$$X^i \mathbf{e}_i.$$

It turns out that it is more comfortable to use the symbol $\frac{\partial}{\partial x_i}$ instead of $\mathbf{e}_i$, and we write instead

$$X^i \frac{\partial}{\partial x_i}.$$

Why do we prefer the awkward notation $\frac{\partial}{\partial x_i}$ to $\mathbf{e}_i$? The partial derivative symbol is appropriate here for three reasons: (i) it is coherent with the interpretation of tangent vectors as derivations, (ii) there is no risk of confusing it with anything else, and more importantly (iii) it helps us to write the coordinate changes correctly via the chain rule. Indeed, if we pick another chart we get different coordinates

$$\bar{X}^i \frac{\partial}{\partial \bar{x}_i}$$

and we know from Section 4.4.8 that the coordinates of a vector change contravariantly, hence

$$\bar{X}^j = X^i \frac{\partial \bar{x}_j}{\partial x_i}. \tag{6}$$

Thanks to the partial derivative notation, there is no need to remember the formula by heart: it suffices to apply formally the chain rule and we get

$$X^i \frac{\partial}{\partial x_i} = X^i \frac{\partial \bar{x}_j}{\partial x_i} \frac{\partial}{\partial \bar{x}_j}.$$

This gives (6). Beware that one possible source of confusion is that the coordinates of a vector change contravariantly, while the vectors themselves of the basis change *covariantly*: indeed we have

$$\frac{\partial}{\partial \bar{x}_j} = \frac{\partial x_i}{\partial \bar{x}_j} \frac{\partial}{\partial x_i}$$

and the change of basis matrix here is the *inverse* of the one that we find in (6). Luckily, we can relax: the partial derivative notation helps us to write the correct form in any context.

5.1.4. Vector fields on subsets. Let M be a smooth manifold. It is sometimes useful to have vector fields defined not on the whole of M, but only on some subset $S \subset M$. By definition, a vector field in S is a smooth partial section $S \to TM$ of the tangent bundle, see Section 4.4.2. The following example may be quite common.

Example 5.1.3. If $f \colon N \hookrightarrow M$ is an embedding, every vector field X in N induces a vector field Y on the image $S = f(N)$ by setting

$$Y\big(f(p)\big) = df_p(X(p)).$$

We now rephrase Proposition 4.4.1 in this context:

Proposition 5.1.4. *If $S \subset M$ is a closed subset, every vector field on S may be extended to a global one on M.*

We may also require that the extended vector field vanishes outside of an arbitrary neighbourhood of S.

Corollary 5.1.5. *Let $N \subset M$ be a compact submanifold. Every vector field in N extends to a vector field in M that vanishes outside of any given neighbourhood of N.*

5.1.5. Straightening. Let X be a vector field on a smooth manifold M, and $p \in M$ a point. Among the infinitely many possible charts near p, is there one that transports X into a reasonably nice vector field in $\mathbb{R}^n$? The answer is positive if X does not vanish at p.

Proposition 5.1.6 (Straightening vector fields). *If $X(p) \neq 0$, there is a chart $U \to V$ with $p \in U$ that transports X into $\frac{\partial}{\partial x_1}$.*

Proof. By taking a chart we may suppose that $M = \mathbb{R}^n$, $p = 0$, and $X(p) = \frac{\partial}{\partial x_1}$. We now use the flow $F(x, t)$ to construct a chart that straightens the field X. We set

$$\psi(x_1, \ldots, x_n) = F\big((0, x_2, \ldots, x_n), x_1\big).$$

The differential $d\psi_0$ is the identity, because $\psi(0, x_2, \ldots, x_n) = (0, x_2, \ldots, x_n)$ and $\gamma(t) = \psi(t, 0, \ldots, 0)$ is an integral curve of X, hence $\gamma'(0) = \frac{\partial}{\partial x_1}$.

Therefore ψ is a local diffeomorphism that sends the lines $x + te_1$ to integral curves of X, so it sends the vector field $\frac{\partial}{\partial x_1}$ to X. □

5.2. Flows

It is hard to overestimate the importance of vector fields in differential topology: they appear naturally everywhere, not only as intrinsically interesting objects, but also as very powerful tools to prove deep theorems.

In this section, we show that a vector field X on a smooth manifold M defines an infinitesimal way to deform M through a *flow* which moves every point of p along an *integral curve*, a curve that is tangent to X at every point.

Flows are powerful tools, and we will use them here to promote isotopies to *ambient isotopies* on every compact manifold.

5.2.1. Integral curves. Let M be a smooth manifold and X a given vector field on M. An *integral curve* of X is a curve $\gamma\colon I \to M$ such that

$$\gamma'(t) = X\big(\gamma(t)\big)$$

for all $t \in I$.

Example 5.2.1. The curve $\gamma(t) = (\cos t, \sin t, \ldots, \cos t, \sin t)$ is an integral curve of the vector field in S^n described in Example 5.1.1.

An integral curve $\gamma\colon I \to M$ is *maximal* if there is no other integral curve $\eta\colon J \to M$ with $I \subsetneq J$ and $\gamma(t) = \eta(t)$ for all $t \in I$. Every integral curve can be extended to a maximal one by enlarging the domain as much as possible. A straightforward application of the Cauchy – Lipschitz Theorem 1.3.5 proves the existence and uniqueness of maximal integral curves:

Proposition 5.2.2. *Let X be a vector field in M. For every $p \in M$ there is a unique maximal integral curve $\gamma\colon I \to M$ with $\gamma(0) = p$.*

Proof. Pick a chart $\varphi\colon U \to \mathbb{R}^n$ and translate locally everything into $\mathbb{R}^n$. The vector field X transforms into a smooth map $\mathbb{R}^n \longrightarrow \mathbb{R}^n$, that we still denote by X for simplicity. An integral curve γ satisfies $\gamma'(t) = X\big(\gamma(t)\big)$. The local existence and uniqueness of γ follows from the Cauchy – Lipschitz Theorem 1.3.5. The maximal integral curve is also clearly unique. □

5.2.2. Flows. One very nice feature of the Cauchy – Lipschitz Theorem is that the unique solution depends smoothly on the initial data. In this topological context, this implies that all the integral curves on a fixed vector field may be gathered into a single smooth family, as follows.

Let X be a vector field on a smooth manifold M.

Theorem 5.2.3. *There is a unique open neighbourhood U of $M \times \{0\}$ inside $M \times \mathbb{R}$ and a unique smooth map $\Phi\colon U \to M$ such that the following holds: for every $p \in M$ the set $I_p = \{t \in \mathbb{R} \mid (p,t) \in U\}$ is an open interval and $\gamma_p\colon I_p \to M$, $\gamma_p(t) = \Phi(p,t)$ is the maximal integral curve with $\gamma_p(0) = p$.*

Proof. For every $p \in M$ there is a maximal integral curve $\gamma_p\colon I_p \to M$ with $\gamma_p(0) = p$. We define

$$U = \{(p,t) \mid t \in I_p\}, \qquad \Phi(p,t) = \gamma_p(t).$$

The Cauchy –Lipschitz Theorem 1.3.5, applied locally at every point (p,t), implies that U is open and Φ is smooth. $\square$

The map Φ is the *flow* associated to the vector field X. If the open maximal set U is the whole of $M \times \mathbb{R}$ we say that the vector field X is *complete*.

Example 5.2.4. Pick $M = \mathbb{R}^n$ and $X = \frac{\partial}{\partial x_1}$ constantly. In this case we have $U = M \times \mathbb{R}$ and $\Phi(x,t) = x + te_1$, so X is complete. If we remove from M a random closed subset, the resulting vector field X is probably not complete anymore.

Here is a simple completeness criterion.

Lemma 5.2.5. *If $M \times (-\varepsilon, \varepsilon) \subset U$ for some $\varepsilon > 0$, then X is complete.*

Proof. We fix an arbitrary point $p \in M$ and we must prove that $I_p = \mathbb{R}$. Pick any $t \in I_p$. The integral curves emanating from p and $\Phi(p,t)$ differ only by a translation of the domain: hence $I_p = I_{\Phi(p,t)} + t$ and

$$\Phi\big(\Phi(p,t), u\big) = \Phi(p, t+u) \tag{7}$$

for every $u \in I_{\Phi(p,t)}$. By hypothesis $(-\varepsilon, \varepsilon) \subset I_{\Phi(p,t)}$ and hence $(t-\varepsilon, t+\varepsilon) \subset I_p$. Since this holds for every $t \in I_p$ we get $I_p = \mathbb{R}$. $\square$

Corollary 5.2.6. *Every vector field on a compact M is complete.*

Proof. By compactness any neighbourhood U of $M \times \{0\}$ in $M \times \mathbb{R}$ must contain $M \times (-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$. $\square$

Let now X be a complete vector field on a smooth manifold M and Φ be its flow. We denote by $\Phi_t\colon M \to M$ the level map $\Phi_t(p) = \Phi(p,t)$.

Proposition 5.2.7. *The map Φ_t is a diffeomorphism for all $t \in \mathbb{R}$. Moreover*

$$\Phi_{-t} = \Phi_t^{-1}, \qquad \Phi_{t+s} = \Phi_t \circ \Phi_s$$

for all $t, s \in \mathbb{R}$.

Proof. The equality (7) implies that $\Phi_{t+s} = \Phi_t \circ \Phi_s$ for all $t, s \in \mathbb{R}$. This in turn gives $\Phi_{-t} = \Phi_t^{-1}$ and hence Φ_t is a diffeomorphism. $\square$

A smooth map $\Phi\colon M \times \mathbb{R} \to M$ with these properties is also called a *one-parameter group of diffeomorphisms*. Indeed we may consider this family as a group homomorphism $\mathbb{R} \to \mathrm{Diffeo}(M)$, $t \mapsto \Phi_t$ where $\mathrm{Diffeo}(M)$ is the group of all diffeomorphisms $M \to M$.

It is indeed a remarkable fact that by constructing vector fields on a compact manifold M we get plenty of one-parameter families of diffeomorphisms for M.

Example 5.2.8. The vector field on S^n constructed in Example 5.1.1 generates the flow

$$\Phi(x_1, \ldots, x_{2m}, t) = \big(x_1 \cos t - x_2 \sin t, x_2 \cos t + x_1 \sin t, \ldots\big).$$

5.3. Ambient isotopy

The previous discussion on flows and diffeomorphisms leads us naturally to define a stronger form of isotopy, called *ambient isotopy*, that involves a smooth distortion of the ambient space.

5.3.1. Definition. Let M be a smooth manifold.

Definition 5.3.1. An *ambient isotopy* in M is an isotopy F between the identity $\mathrm{id}\colon M \to M$ and some diffeomorphism $\varphi\colon M \to M$, such that every level $F_t\colon M \to M$ is a diffeomorphism.

For instance, every flow Φ generated by some complete vector field X on M is an ambient isotopy between the identity Φ_0 and the diffeomorphism Φ_1.

Let now M, N be two manifolds. We say that two embeddings $f, g\colon M \to N$ are *ambiently isotopic* if there is an ambient isotopy F on N wth $F_0 = \mathrm{id}$ and $F_1 = \varphi$ such that $g = \varphi \circ f$. We check that this notion is indeed stronger than that of an isotopy.

Proposition 5.3.2. *If f, g are ambiently isotopic, they are isotopic.*

Proof. An isotopy G_t between f and g is $G_t(x) = F_t(f(x))$. □

Informally, two embeddings f and g are ambiently isotopic if they related by an isotopy that "moves the whole of N". We now use the flows to show that, if M is compact, the two notions actually coincide.

Theorem 5.3.3. *If M is compact, any two embeddings $f, g\colon M \to N$ are isotopic $\iff$ they are ambiently isotopic.*

Proof. Let $F\colon M \times \mathbb{R} \to N$ be an isotopy relating f and g. We define

$$G\colon M \times \mathbb{R} \longrightarrow N \times \mathbb{R}$$

by setting $G(p, t) = \big(F(p, t), t\big)$. We note that G is time-preserving and proper (because M is compact). Moreover

$$dG_{(p,t)} = \begin{pmatrix} d(F_t)_p & * \\ 0 & 1 \end{pmatrix}$$

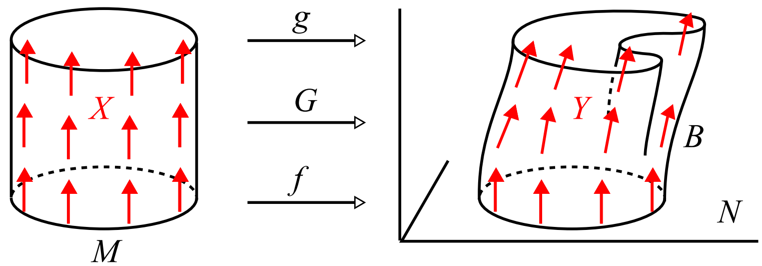

Figure 5.2. The vertical vector field X on $M \times [0, 1]$ is transported via G into a vector field Y defined only on the compact set B.

and hence G is an injective immersion. Being proper, the map G is an embedding (see Exercise 3.8.5) and therefore its image $G(M \times \mathbb{R})$ is a submanifold of $N \times \mathbb{R}$.

The vertical vector field $X = \frac{\partial}{\partial t}$ on $M \times [0, 1]$ is transported via G into a vector field Y defined only on the compact set $B = G\big(M \times [0, 1]\big)$, by setting $Y\big(G(p, t)\big) = dG_{(p,t)}\big(\frac{\partial}{\partial t}\big)$ as in Example 5.1.3. See Figure 5.2.

The vector field Y is defined only on the compact subset $B \subset N \times \mathbb{R}$, but we may extend it to a vector field Y on the whole of $N \times \mathbb{R}$ with the property that $Y = \frac{\partial}{\partial t}$ outside of some compact neighbourhood V of B. To show this, we first extend Y to a vector field that vanishes outside V, and then modify everywhere its t-coordinate to be constantly 1.

We now consider the flow Φ of Y in $N \times \mathbb{R}$. The vector field Y is complete: to show this, we note that V is compact and $\Phi_t(p, u) = (p, u + t)$ outside V, and these two facts easily imply that there is an $\varepsilon > 0$ such that Φ is defined at every time $|t| < \varepsilon$, so Lemma 5.2.5 applies.

Since the t-component of Y is constantly 1 we get

$$\Phi_t(p, 0) = \big(H(p, t), t\big)$$

for some smooth map $H\colon N \times \mathbb{R} \to N$. We write $H_t(p) = H(p, t)$ and note that $H_t\colon N \to N$ is diffeomorphism for every t, since Φ_t is. Moreover $H_0 = \mathrm{id}$ and hence H furnishes an ambient isotopy. Finally, we have $H\big(f(p), t\big) = F(p, t)$ for every $(p, t) \in M \times [0, 1]$ because $Y = dG\big(\frac{\partial}{\partial t}\big)$ on B. Therefore H is an ambient isotopy relating f and g. $\square$

Corollary 5.3.4. *Every connected smooth manifold M is* homogeneous, *that is for every two points $p, q \in M$ there is a diffeomorphism $f\colon M \to M$ isotopic to the identity such that $f(p) = q$.*

Proof. There is a smooth arc $\gamma\colon \mathbb{R} \to M$ with $\gamma(0) = p$ and $\gamma(1) = q$ (exercise). This arc may be interpreted as an isotopy between two embeddings $\{\mathrm{pt}\} \to M$ that send a single point to p and to q, respectively. This isotopy may be promoted to an ambient isotopy, that sends p to q. $\square$

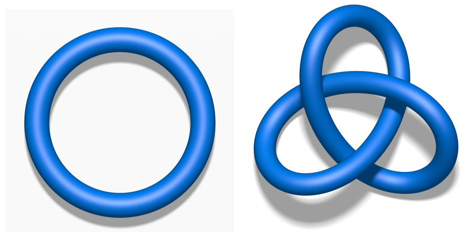

Figure 5.3. The trivial and the trefoil knot are not isotopic. This is certainly true... but how can we prove it?

How can we prove that two given homotopic embeddings are actually not isotopic? For instance, how can we prove the intuitive fact that the two knots in Figure 5.3 are not isotopic? If they were isotopic, they would also be ambient isotopic, and hence in particular they would have homeomorphic complements. One can then try to calculate the fundamental groups of the complement and prove that they are not isomorphic: this strategy actually works for the two knots depicted in the figure.

5.4. Lie brackets

We now introduce an operation on vector fields called *Lie bracket*. The Lie bracket $[X, Y]$ of two vector fields X and Y in M is a third vector field that measures the "lack of commutativity" of X and Y.

5.4.1. Vector fields as derivations. Let X be a vector field on a smooth manifold M. For every open subset $U \subset M$ and every smooth function $f \in C^\infty(U)$ we may define a new function $Xf \in C^\infty(U)$ by setting

$$(Xf)(p) = X(p)(f)$$

for every $p \in U$. Recall that $X(p) \in T_pM$ is a derivation and hence transforms any locally defined function f into a real number $X(p)(f)$, so the definition of Xf makes sense.

In coordinates, the vector field X is written as

$$X^i \frac{\partial}{\partial x_i}$$

and the new function Xf is simply

$$X^i \frac{\partial f}{\partial x_i}.$$

This shows in particular that Xf is smooth.

We have just discovered that we can employ vector fields to "derive" functions. We use the term "derivation" here, because the Leibniz rule

$$X(fg) = (Xf)g + f(Xg)$$

is satisfied by construction for every functions f and g defined on some common open set $U \subset M$. Of course the derived function Xf depends heavily on the vector field X.

Another way of seeing Xf is as the result of a contraction of the differential df, a tensor field of type $(0,1)$, with X, a tensor field of type $(1,0)$. The result is a tensor field Xf of type $(0,0)$, that is a smooth function.

5.4.2. Lie brackets. Let X and Y be two vector fields on a smooth manifold M. The *Lie bracket* $[X,Y]$ of X and Y is a new vector field, uniquely determined by requiring that

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

for every function f defined on any open subset $U \subset M$.

Proposition 5.4.1. *The vector field* $[X,Y]$ *is well-defined.*

Proof. For the moment, the bracket $[X,Y] = XY - YX$ is just an operator on smooth functions defined on any open subset $U \subset M$. For every $f,g \in C^\infty(U)$ we get

$$\begin{aligned} XY(fg) &= X\big((Yf)g\big) + X\big(f(Yg)\big) \\ &= (XYf)g + (Yf)(Xg) + (Xf)(Yg) + f(XYg), \\ YX(fg) &= (YXf)g + (Xf)(Yg) + (Yf)(Xg) + f(YXg) \end{aligned}$$

from which we deduce that

$$[X,Y](fg) = \big([X,Y]f\big)g + f\big([X,Y]g\big).$$

We have proved that $[X,Y]$ is also a derivation. This allows us to define $[X,Y]$ as a vector field, by setting

$$[X,Y](p)(f) = [X,Y](f)(p)$$

for every $p \in M$ and every f defined near p. The proof is complete. □

5.4.3. Lie algebra. We introduce an important concept.

Definition 5.4.2. A *Lie algebra* is a real vector space A equipped with an antisymmetric bilinear operation $[,]$ called *Lie bracket* that satisfies the *Jacobi identity*

$$\big[[x,y],z\big] + \big[[y,z],x\big] + \big[[z,x],y\big] = 0$$

for every $x,y,z \in A$.

Let M be a smooth manifold. Recall that $\mathfrak{X}(M)$ is the vector space consisting of all the vector fields in M.

Exercise 5.4.3. The space $\mathfrak{X}(M)$ with the Lie bracket $[,]$ is a Lie algebra.

5.4.4. In coordinates. The definition of the Lie bracket is quite abstract and is now due time to write an explicit formula that is valid in coordinates with respect to any chart.

Exercise 5.4.4. In coordinates we get

$$[X, Y]^i = X^j \frac{\partial Y^i}{\partial x_j} - Y^j \frac{\partial X^i}{\partial x_j}.$$

The reader may also wish to define $[X, Y]$ directly via this formula, but in that case she needs to verify that this definition is chart-independent, a fact that is not immediately obvious: for instance if we eliminate one of the two members then the definition is not chart-independent anymore.

In the definition of the Lie bracket of two vector fields we have seen the appearance of a recurrent theme in differential topology and geometry: the eternal quest for intrinsic (that is, chart-independent) definitions. One may fulfil this task either working entirely in coordinates, or using some more abstract arguments as we just did. As usual, both viewpoints are important.

The following exercises may be solved working in coordinates.

Exercise 5.4.5. For every $X, Y \in \mathfrak{X}(M)$ and $f, g \in C^\infty(M)$ we have

$$[fX, gY] = fg[X, Y] + f(Xg)Y - g(Yf)X.$$

Exercise 5.4.6. On an open set of $\mathbb{R}^n$, for every i, j we have

$$\left[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right] = 0.$$

More generally, we have

$$\left[\frac{\partial}{\partial x_i}, Y^j \frac{\partial}{\partial x_j}\right] = \frac{\partial Y^j}{\partial x_i} \frac{\partial}{\partial x_j} = \frac{\partial Y}{\partial x_i}.$$

Exercise 5.4.7. Let A, B be two $n \times n$ matrices. Consider the vector fields in some open subset of $\mathbb{R}^n$ defined as

$$X(x) = Ax, \qquad Y(x) = Bx.$$

Their Lie bracket is

$$[X, Y](x) = (BA - AB)x.$$

Exercise 5.4.8. Let $N \subset M$ be a submanifold. If X, Y are vector fields on N, and $\bar{X}, \bar{Y}$ are any extensions of X, Y to some open subset $U \subset M$ containing N, then at every point $p \in N$ we get

$$[X, Y](p) = [\bar{X}, \bar{Y}](p).$$

The previous exercise is in fact a special case of the following. If $f : M \to N$ is any smooth map between manifolds, two vector fields $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$ are *f-related* if $df_p(X(p)) = Y(f(p))$ for all $p \in M$.

Exercise 5.4.9. If X_1, X_2 are f-related to Y_1, Y_2 respectively, then $[X_1, X_2]$ is f-related to $[Y_1, Y_2]$.

We now introduce a more geometric interpretation of the Lie bracket.

5.4.5. Non-commuting flows. Let X and Y be two vector fields on a smooth manifold M, and let F, G be their corresponding flows. Consider a point $p \in M$. In general, the two flows do not commute, that is $F_s \circ G_t(p)$ may be different from $G_t \circ F_s(p)$ whenever they are defined. We now show that the Lie bracket $[X, Y]$ at p measures this possible lack of commutation.

Proposition 5.4.10. *On any chart, we have*

$$G_t \circ F_s(p) - F_s \circ G_t(p) = st[X, Y](p) + o(s^2 + t^2).$$

Note that the whole expression makes sense only on a chart, that is on some open subset $V \subset \mathbb{R}^n$ with $p \in V$. On a general smooth manifold M the points $G_t(F_s(p))$ and $F_s(G_t(p))$ are probably distinct points in M and there is no way of estimating their "distance". The expression is however very useful because it holds on every possible chart.

Proof. We fix p and consider the smooth function

$$\Psi(s, t) = G_t \circ F_s(p) - F_s \circ G_t(p).$$

Consider its Taylor expansion

$$\begin{aligned}\Psi(s, t) =& \Psi(0, 0) + s\frac{\partial \Psi}{\partial s}(0, 0) + t\frac{\partial \Psi}{\partial t}(0, 0) \\ &+ \frac{s^2}{2}\frac{\partial^2 \Psi}{\partial s^2}(0, 0) + st\frac{\partial^2 \Psi}{\partial s \partial t}(0, 0) + \frac{t^2}{2}\frac{\partial^2 \Psi}{\partial t^2}(0, 0) + o(s^2 + t^2).\end{aligned}$$

The crucial fact here is that $\Psi(s, 0) = \Psi(0, t) = 0$ for all s, t. Since $\Psi \equiv 0$ on the axis $s = 0$ and $t = 0$, all the terms in the Taylor expansion above vanish except the mixed one $\frac{\partial^2 \Psi}{\partial s \partial t}(0, 0)$, that we now calculate. We have

$$\frac{\partial}{\partial t}\big(G_t \circ F_s(p)\big) = Y\big(G_t \circ F_s(p)\big)$$

and hence

$$\left(\frac{\partial}{\partial t} G_t \circ F_s(p)\right)(s, 0) = Y\big(F_s(p)\big)$$

which gives

$$\left(\frac{\partial^2}{\partial s \partial t} G_t \circ F_s(p)\right)(0, 0) = \frac{\partial}{\partial s} Y\big(F_s(p)\big)(0) = X^j \frac{\partial Y}{\partial x_j}.$$

Therefore

$$\frac{\partial^2 \Psi}{\partial s \partial t}(0, 0) = X^j \frac{\partial Y}{\partial x_j} - Y^j \frac{\partial X}{\partial x_j} = [X, Y](p)$$

by Exercise 5.4.4. The proof is complete. □

We say that two vector fields X and Y *commute* if $[X,Y]=0$ everywhere. The corresponding flows F and G *commute* if

$$F_s \circ G_t(p) = G_t \circ F_s(p)$$

for every p,s,t such that both members are defined. These two notions of commutativity coincide:

Proposition 5.4.11. *Two vector fields commute $\Longleftrightarrow$ their flows do.*

Proof. If the flows commute, then $[X,Y]=0$ because of Proposition 5.4.10. Conversely, suppose that $[X,Y]=0$.

Consider a point $p\in M$. If $X(p)=Y(p)=0$, we get $F_s(p)=G_t(p)=p$ and we are done. Otherwise, suppose that $X(p)\neq 0$. On a chart we can straighten X and get $X=\frac{\partial}{\partial x_1}$ and $F_s(p)=p+se_1$.

Now $[X,Y]=0$ and Exercise 5.4.6 imply that

$$\frac{\partial Y}{\partial x_1}=0.$$

The field Y is hence invariant by translations along e_1. Therefore $G_t(p+se_1)=G_t(p)+se_1$, that is G_t commutes with F_s.

We have proved that the flows commute for every $p\in M$ when the times s and t are sufficiently small. This implies easily that they commute at all times s,t such that the flows are defined (exercise). □

5.4.6. Multiple straightenings. Can we straighten two or more vector fields simultaneously? It should not be a surprise now that the answer depends on their Lie brackets. Let $X_1,\ldots,X_k$ be vector fields on a smooth manifold M, and $p\in M$ be a point.

Proposition 5.4.12. *Suppose that $X_1(p),\ldots,X_k(p)$ are independent vectors. There is a chart $U\to V$ that transports $X_1,\ldots,X_k$ into $\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_k}$ $\Longleftrightarrow$ $[X_i,X_j]=0$ for all i,j on some neighbourhood of p.*

Proof. If there is a chart of this type, then clearly $[X_i,X_j]=0$. We now prove the converse and suppose $[X_i,X_j]=0$ for all i,j.

By taking a chart we may suppose that M is an open set in $\mathbb{R}^n$, $p=0$, and $X_i(0)=\frac{\partial}{\partial x_i}$ for all $i=1,\ldots,k$. Let F^i_t be the flow of X_i. Define

$$\psi(x_1,\ldots,x_n)=F^k_{x_k}\circ\cdots\circ F^1_{x_1}(0,\ldots,0,x_{k+1},\ldots,x_n).$$

The differential $d\psi_0$ is the identity, because

$$\psi(0,\ldots,0,x_{k+1},\ldots,x_n)=(0,\ldots,0,x_{k+1},\ldots,x_n)$$

and $\gamma_i(t)=\psi(te_i)$ with $i=1,\ldots,k$ is an integral curve for X_i, so $\gamma_i'(0)=\frac{\partial}{\partial x_i}$.

We deduce that ψ is a local diffeomorphism. It is clear that ψ sends the lines $x+te_k$ to integral curves for X_k, so it sends $\frac{\partial}{\partial x_k}$ to X_k. Since $[X_i,X_j]=0$, the flows F^i_t commute and we can permute them in the definition of ψ at our pleasure: so the same argument shows that ψ sends $\frac{\partial}{\partial x_i}$ to X_i for all i. □

5.4.7. Lie derivative. We have just noted that a vector field X may be used to derive functions. Can we also use X to derive other objects, for instance another vector field Y or more generally any tensor field s? The answer is positive, and this operation is called the *Lie derivative*.

We first recall that every diffeomorphism $f\colon M \to N$ induces an isomorphism between the corresponding tensor bundles

$$f_*\colon \mathcal{T}_h^k M \longrightarrow \mathcal{T}_h^k N$$

induced from that of the tangent bundles $f_*\colon TM \to TN$, and we may use f_* to transfer tensor fields from M to N and viceversa.

Let now X be a vector field on a smooth manifold M, and let s be any tensor field on M, of some type (h, k). The *Lie derivative* $\mathcal{L}_X s$ is a new tensor field of the same type (h, k), morally obtained by deriving s along X, and defined as follows.

Let F_t be the flow generated by X. For every point $p \in M$, there is a sufficiently small $\varepsilon > 0$ such that $F_t(p)$ is defined on a neighbourhood of p and F_t is a local diffeomorphism at p for all $|t| < \varepsilon$. Therefore $(F_t)_*(s)$ is another tensor field defined on a neighbourhood of $F_t(p)$, that varies smoothly in t, and we now want to compare s and $(F_t)_*(s)$.

We note that $(F_{-t})_*$ transports the tensor $s\big(F_t(p)\big)$ that lies in the tensor space at $F_t(p)$ into the tensor space at p and can hence be compared with $s(p)$. More specifically the tensor

$$(F_{-t})_*\big(s(F_t(p))\big)$$

lies in $\mathcal{T}_h^k(T_pM)$ for every t and varies smoothly in t, so it makes sense to define its derivative

$$\big(\mathcal{L}_X s\big)(p) = \frac{d}{dt}\Big|_{t=0}(F_{-t})_*\big(s(F_t(p))\big).$$

We have defined a linear map

$$\mathcal{L}_X\colon \Gamma\big(\mathcal{T}_k^h(M)\big) \longrightarrow \Gamma\big(\mathcal{T}_k^h(M)\big)$$

that "derives" any tensor field along X.

Exercise 5.4.13. The following holds:

- if $f \in C^\infty(M)$, then $\mathcal{L}_X f = Xf$;
- if Y is a vector field, then $\mathcal{L}_X Y = [X, Y]$;
- for every tensor fields S and T of any types we have

$$\mathcal{L}_X(S \otimes T) = (\mathcal{L}_X S) \otimes T + S \otimes (\mathcal{L}_X T).$$

The Lie derivative $\mathcal{L}_X s$ measures how s changes along X, in fact it follows readily from the definition that $\mathcal{L}_X s \equiv 0$ on $M \iff$ the tensor field s is invariant under the flow F_t wherever it is defined.

It is important to note here that, as opposite to the directional derivative in $\mathbb{R}^n$, the value of $\mathcal{L}_X s$ at a point p depends on the local behaviour of X near

p, but not on the directional vector $X(p)$ alone! To get a derivation that, like the directional derivative in $\mathbb{R}^n$, depends in p only on the directional vector based at p, we need to introduce a new structure called *connection*. We will do this later on in this book.

5.5. Foliations

We now introduce some higher-dimensional analogues of vector fields and integral curves, where we replace vectors with k-dimensional subspaces, and integral curves with k-dimensional submanifolds.

5.5.1. Foliations. Let M be a smooth n-manifold. An *immersed submanifold* in M is the image of an immersion $S \to M$.

Definition 5.5.1. A k-dimensional *foliation* is a partition $\mathscr{F} = \{\lambda_i\}$ of M into injectively immersed k-dimensional connected submanifolds $\lambda_i \subset M$ called *leaves*, such that the following holds: for every $p \in M$ there is a chart $\varphi\colon U \to \mathbb{R}^n$ with $p \in U$ such that $\varphi(\lambda_i \cap U)$ is the union of some parallel horizontal affine k-planes (that is, of type $\{x_{k+1} = c_{k+1}, \ldots, x_n = c_n\}$), for every i.

In other words, at every point p there is a chart φ that transforms the partition $\mathscr{F}$ near p into the standard one of parallel horizontal k-planes in $\mathbb{R}^n$. We say that such a chart φ is *compatible* with the foliation.

Mettere figura.

Remark 5.5.2. For a fixed leaf λ_i, the image $\varphi(\lambda_i \cap U)$ along a compatible chart φ may consist of infinitely many k-planes. These are countable, because λ_i is the image of an immersed submanifold $S \to M$ and S is second countable.

We also note that a foliation contains uncountably many leaves: this is a consequence of the previous remark, or of the more general fact that the union of countably many immersed manifolds of smaller dimension than M has measure zero.

Example 5.5.3. The following are foliations:

(1) the partition of $\mathbb{R}^n$ into all the affine spaces parallel to a fixed vector subspace $L \subset \mathbb{R}^n$;
(2) if $E \to B$ is a fibre bundle, the partition of E into the fibres E_p;
(3) for a fixed *slope* $\nu \in \mathbb{R}$, the family of all curves $\alpha\colon \mathbb{R} \to S^1 \times S^1$ of type $\alpha(t) = \left(e^{2\pi i t}, e^{2\pi i(\nu t+\mu)}\right)$ as μ varies.

Exercise 5.5.4. In the last example, the leaves are compact $\Longleftrightarrow \lambda \in \mathbb{Q}$. If $\lambda \in \mathbb{R} \setminus \mathbb{Q}$ every leaf is dense.

We now furnish an equivalent definition of foliation.

Definition 5.5.5. A k-dimensional *foliation* in M is an atlas $\{\varphi_i \colon U_i \to \mathbb{R}^n\}$ compatible with the smooth structure of M whose transition maps φ_{ij} are all locally of the following form:

$$\varphi_{ij}(x, y) = \big(\varphi^1_{ij}(x, y), \varphi^2_{ij}(y)\big).$$

Here we represent $\mathbb{R}^n$ as $\mathbb{R}^k \times \mathbb{R}^{n-k}$, both as a domain and as a codomain.

In other words, we require that the last $n-k$ coordinates of φ_{ij} should depend locally only on the last $n-k$ coordinates of the point. By "locally" we mean as usual that every point p in the domain of φ_{ij} has a neighbourhood such that φ_{ij} is of that form.

The two definitions look very different but are indeed equivalent! If $\mathscr{F}$ is a foliation in the partition sense, by considering only charts that are compatible with $\mathscr{F}$ we get an atlas as in Definition 5.5.5 (exercise). Conversely, given an atlas $\{\varphi_i\}$ of this kind, the transition maps preserve locally the k-dimensional affine horizontal subspaces $\{y = c\}$ which hence glue to form immersed submanifolds in M.

To construct the immersed manifolds rigorously, we proceed as follows. We assign to $\mathbb{R}^k$ and $\mathbb{R}^{n-k}$ respectively the Euclidean and the discrete topology, and we give the product topology to $\mathbb{R}^k \times \mathbb{R}^{n-k}$. Note that this topology is finer than the Euclidean one. We now use this model to define a finer topology on M, by declaring a set in M to be open if it intersects every chart U_i into a subset whose image in $\varphi_i(U_i) \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$ is open in the new finer topology.

The manifold M with the finer topology decomposes into (uncountably many) connected components $\{M_j\}$. The atlas $\{\varphi_i \colon U_i \to \mathbb{R}^n\}$ furnishes to every M_j a structure of smooth manifold: the only tricky part here is to prove second countability, and is left as an exercise. *Hint:* Select a countable sub-atlas $\{\varphi_i\}$ and prove that every leaf "propagates" only to countably many nearby ones at each step.

5.5.2. Distributions. Let M be a smooth n-manifold. Here is another natural geometric definition.

Definition 5.5.6. A k-*distribution* in M is a rank-k subbundle D of the tangent bundle TM.

In other words, a distribution is a collection of k-subspaces $D_p \subset T_pM$ that varies smoothly with p. See Lemma 4.4.6.

Example 5.5.7. If $\mathscr{F}$ is a k-dimensional foliation on M, the k-spaces tangent to the leaves of $\mathscr{F}$ form a k-distribution.

A distribution that is tangent to some foliation $\mathscr{F}$ is called *integrable*. Note that a diffeomorphism $\varphi \colon M \to M'$ transforms a distribution D on M into one D' on M' in the obvious way, by setting $D'_{\varphi(p)} = d\varphi_p(D_p)\ \forall p \in M$. The integrability condition may also be expressed without using foliations:

Proposition 5.5.8. *A distribution D is integrable $\iff \forall p \in M$ there is a chart $\varphi\colon U \to \mathbb{R}^n$ with $p \in U$ that transforms D into a constant distribution.*

A constant distribution in $\mathbb{R}^n$ is $D_p \equiv L$ for some fixed subspace $L \subset \mathbb{R}^n$.

Proof. ($\Rightarrow$). If D is tangent to a foliation $\mathscr{F}$, any chart compatible with $\mathscr{F}$ transforms D into a constant one.

($\Leftarrow$). All these charts define a foliation in the sense of Definition 5.5.5. $\square$

5.5.3. The Frobenius Theorem. We now state and prove a theorem that characterises the integrable distributions via the Lie bracket of vector fields.

A vector field X on a manifold M is *tangent* to a distribution D if $X(p) \in D_p$ for all $p \in M$. A distribution D is *involutive* if whenever X, Y are two vector fields tangent to D, their Lie bracket $[X, Y]$ is also tangent.

Theorem 5.5.9 (Frobenius Theorem). *A distribution D on a manifold M is integrable $\iff$ it is involutive.*

Proof. If D is integrable, it is tangent to a foliation $\mathscr{F}$. For every $p \in M$, a chart $U \to \mathbb{R}^n$ compatible with $\mathscr{F}$ transforms the leaves of $\mathscr{F}$ into horizontal leaves $\{x_{k+1} = c_{k+1}, \ldots, x_n = c_n\}$ and hence it transforms D into the constantly horizontal distribution $D_p = \{x_{k+1} = \ldots = x_n = 0\}$. If X, Y are vector fields tangent to D, then read on U they are of the form

$$X = \sum_{i=1}^k X^i \frac{\partial}{\partial x_i}, \qquad Y = \sum_{i=1}^k Y^i \frac{\partial}{\partial x_i}$$

and by Exercise 5.4.4 we get $[X, Y]^i = 0$ for all $i > k$. Therefore $[X, Y]$ is also tangent to D and D is involutive.

Conversely, suppose that D is involutive. For every $p \in M$ we pick a chart $\varphi\colon U \to \mathbb{R}^n$ that transforms p in 0 and D_p into the horizontal space $D_0 = \{x_{k+1} = \ldots = x_n = 0\}$. We can suppose that U is small enough so that for every $p \in U$ the chart φ transports D_p into a k-space $D_{\varphi(p)}$ that is transverse to the vertical space $V = \{x_1 = \ldots = x_k = 0\}$. Therefore we can find a local frame on D that read on U is of the type

$$X_1 = \frac{\partial}{\partial x_1} + \sum_{i=k+1}^n X_1^i \frac{\partial}{\partial x_i}, \quad \ldots, \quad X_k = \frac{\partial}{\partial x_k} + \sum_{i=k+1}^n X_k^i \frac{\partial}{\partial x_i}.$$

Exercise 5.4.4 gives $[X_i, X_j]^l = 0$ for all $i, j, l = 1, \ldots, k$, hence $[X_i, X_j]$ is tangent to the vertical space V at every point. Since D is involutive, the vector field $[X_i, X_j]$ must be tangent to D and this implies that $[X_i, X_j] = 0$.

We have discovered that $X_1, \ldots, X_k$ are commuting vector fields and by Proposition 5.4.12 we can transform them via a chart into the coordinate ones $X_i = \frac{\partial}{\partial x_i}$. In this chart the distribution is constant so Proposition 5.5.8 applies. The proof is complete. $\square$

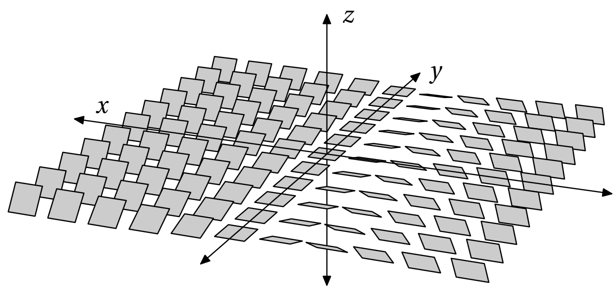

Figure 5.4. A non-integrable plane distribution in $\mathbb{R}^3$.

As an example, the vector fields in $\mathbb{R}^3$

$$X_1 = \frac{\partial}{\partial x}, \qquad X_2 = \frac{\partial}{\partial y} + x\frac{\partial}{\partial z}$$

do not commute since $[X_1, X_2] = \frac{\partial}{\partial z}$. Therefore they generate a non-integrable plane distribution in $\mathbb{R}^3$, drawn in Figure 5.4.

The following criterium may be useful in some cases.

Exercise 5.5.10. A distribution D in M is involutive $\Longleftrightarrow$ for every $p \in M$ there is a local frame $X_1, \ldots, X_k$ for D such that $[X_i, X_j]$ is tangent to D $\forall i, j$.

Hint. Use Exercise 5.4.5 □

5.6. Tubular neighbourhoods

Let M be a compact smooth m-manifold. Among all the open neighbourhoods of a given point $p \in M$, the simplest ones are undoubtedly those that are diffeomorphic to $\mathbb{R}^m$. These are certainly not unique, and there is no canonical way to choose a preferred one; however, we will prove in this section that these are unique up to isotopy, thus answering to Question 3.10.7.

More generally, we will show that not only points, but any submanifold $N \subset M$ has a similar kind of nice open neighbourhood, called *tubular neighbourhood*. The idea that we have in mind is that, for a curve on the plane, a tubular neighbourhood should look like in Figure 5.5, and for a knot $K \subset \mathbb{R}^3$ it should be a little open tube around K. As in Figure 5.5, a tubular neighbourhood should be a bundle over N.

We prove here the existence and uniqueness (up to isotopy) of tubular neighbourhoods for any submanifold $N \subset M$.

5.6.1. Definition. Let M be a m-manifold and $N \subset M$ a n-submanifold. A *tubular neighbourhood* for N is a vector bundle $E \to N$ together with an embedding $i: E \hookrightarrow M$ such that:

- $i|_N = \mathrm{id}_N$, where we identify N with the zero-section in E;
- $i(E)$ is an open neighbourhood of N.

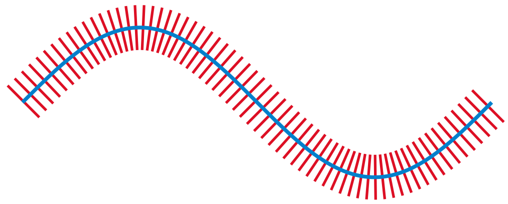

Figure 5.5. A tubular neighbourhood of a curve on the plane.

We usually call a tubular neighbourhood simply the image $i(E)$ of E in N, but keeping in mind that it has a bundle structure with base N.

The second hypothesis implies that $\dim E = \dim M$, so E must have rank $m - n$. Recall that the normal bundle νN of N inside M has precisely that rank, so it seems a promising candidate.

5.6.2. Existence. We now prove the existence of tubular neighbourhoods in two steps: in the first we only consider the case $M = \mathbb{R}^m$.

Proposition 5.6.1. *Every submanifold $N \subset \mathbb{R}^m$ has a tubular neighbourhood with $E = \nu N$.*

Proof. As shown in Example 4.3.6, we have

$$\nu N = \big\{(p, v) \bigm| p \in N, v \in \nu_p N\big\} \subset N \times \mathbb{R}^m \subset \mathbb{R}^m \times \mathbb{R}^m.$$

We have identified $\nu_p N$ with $T_p N^\perp$. We now define the smooth map

$$\begin{aligned} f\colon \quad \nu N &\longrightarrow \mathbb{R}^m, \\ (p, v) &\longmapsto p + v. \end{aligned}$$

See Figure 5.6. We now study the differential $df_{(p,0)}$ at each $p \in N$. We have

$$T_{(p,0)}\nu N = T_p N \oplus \nu_p N$$

and with this identification the differential $df_{(p,0)}$ is just the identity. In particular, it is invertible, so f is an immersion at every point in N.

There is (exercise) a continuous positive function $r\colon N \to \mathbb{R}$ such that f is an embedding on $B\big(p, r(p)\big) \cap \nu N$, for every $p \in N$. Define

$$U = \big\{(p, v) \in \nu N \mid \|v\| < \tfrac{1}{2} r(p)\big\}.$$

One checks easily that $f|_U$ is an embedding. By shrinking νN as in Lemma 4.5.11 we can embed $i\colon \nu N \hookrightarrow U$ keeping N fixed, and the composition $f \circ i$ is a tubular neighbourhood for N. □

We now turn to a more general case.

Theorem 5.6.2. *Let M be a manifold. Every submanifold $N \subset M$ has a tubular neighbourhood with $E = \nu N$.*

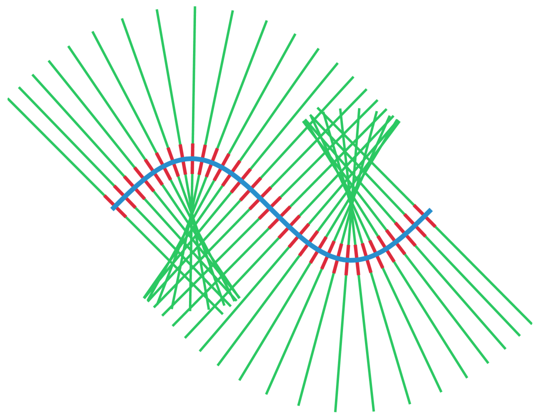

Figure 5.6. To construct a tubular neighbourhood, we map the normal bundle in $\mathbb{R}^n$ and pick a sufficiently small neighbourhood so that this map is an embedding.

Proof. We may embed M in some $\mathbb{R}^k$ thanks to Whitney's Theorem 3.11.8. Now for every $p \in N$ we have the vector space inclusions

$$T_pN \subset T_pM \subset \mathbb{R}^k.$$

We identify ν_pN with the orthogonal complement of T_pN inside T_pM, so that

$$T_pN \oplus \nu_pN = T_pM \subset \mathbb{R}^k.$$

We consider the smooth map

$$\begin{aligned} f\colon \quad \nu N &\longrightarrow \mathbb{R}^k, \\ (p, v) &\longmapsto p + v. \end{aligned}$$

Let W be a tubular neighbourhood of M in $\mathbb{R}^k$, with bundle projection $\pi\colon W \to M$. We set $U = f^{-1}(W)$ and define the map

$$\begin{aligned} f\colon \quad U &\longrightarrow M, \\ (p, v) &\longmapsto \pi(p + v). \end{aligned}$$

As above, the differential is just the identity and we conclude that $f \circ i$ is a tubular neighbourhood for N for some appropriate bundle shrinking i. $\square$

5.6.3. Uniqueness. It is a remarkable and maybe surprising fact that, despite their quite general definition, tubular neighbourhoods are actually unique if one considers them up to isotopy.

We first clarify what we mean by "isotopy" here. Let M be a manifold and $N \subset M$ a submanifold. Two tubular neighbourhoods $i_0\colon E^0 \to M$ and $i_1\colon E^1 \to M$ are *isotopic* if there are a bundle isomorphism $\psi\colon E^0 \to E^1$ and an isotopy F relating the embeddings i_0 and $i_1 \circ \psi$ that keeps N pointwise fixed, that is such that $F(p, t) = p$ for all $p \in N$ and all t.

Note that each embedding $F_t = F(\cdot, t)$ is a tubular neighbourhood of N, so F indeed describes a smooth path of varying tubular neighbourhoods.

Theorem 5.6.3. *Let M be a manifold and $N \subset M$ a submanifold. Every two tubular neighbourhoods of N are isotopic.*

To warm up, we start by proving the following.

Proposition 5.6.4. *Every embedding $f : \mathbb{R}^n \hookrightarrow \mathbb{R}^n$ with $f(0) = 0$ is isotopic to its differential df_0 via an isotopy that fixes 0 at each time.*

Proof. The isotopy for $t \in (0, 1]$ is simply defined as follows:

$$F(x, t) = \frac{f(tx)}{t}.$$

We extend it to the time $t = 0$ by writing the first-order Taylor expansion

$$f(x) = h_1(x)x_1 + \ldots + h_n(x)x_n$$

where $h_i(0) = \frac{\partial f}{\partial x_i}(0)$ for all i. For every $t \in (0, 1]$ we get

$$F(x, t) = h_1(tx)x_1 + \ldots + h_n(tx)x_n$$

and this expression makes sense also for $t = 0$, yielding the equality $F(x, 0) = df_0(x)$. The proof is complete.[1] □

We can now prove Theorem 5.6.3.

Proof. Let E^0 and E^1 be two tubular neighbourhoods of N. We see E^1 as embedded directly in M, and we want to modify the embedding $f : E^0 \to M$ via an isotopy so that matches it with E^1.

We first prove that after an isotopy we may suppose that $f(E^0) \subset E^1$. Indeed, Lemma 4.5.11 provides a shrinkage $g : E^0 \to E^0$ with $f \circ g(E^0) \subset E^1$, and we may construct an isotopy F between f and $f \circ g$ simply by writing $F(v, t) = f\big((1 - t)v + tg(v)\big)$.

Now that $f(E^0) \subset E^1$, we can construct the isotopy $F : E^0 \times [0, 1] \to M$ by mimicking the proof of Proposition 5.6.4: we simply write

$$F(v, t) = \frac{f(tv)}{t}.$$

Here $f(tv)$ is a particular vector in E^1 and hence its division by t makes sense. This is certainly an isotopy for $t \in (0, 1]$, and we now extend it to $t = 0$ similarly to what we did above.

Consider a $v \in E^0$, with $p = \pi(v) \in N$. The point p has an open neighbourhood U above which E^1 is trivialised as $U \times \mathbb{R}^{m-n}$. There are also

[1] To be precise, we should substitute t with $\rho(t)$ via a transition function ρ to get an isotopy defined for all $t \in \mathbb{R}$. We will tacitly assume this in other points in this book.

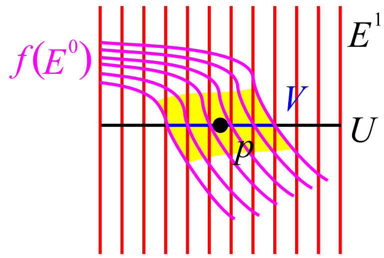

Figure 5.7. By continuity, we can find two neighbourhoods $V \subset U$ of p above which both E^0 and E^1 trivialise, and a $r > 0$ such that $f\big(V \times B(0,r)\big) \subset U \times \mathbb{R}^{m-n}$ (the yellow zone).

a smaller neighbourhood $V \subset U$ and a $r > 0$ such that $E^0|_V$ is also trivialised as $V \times \mathbb{R}^{m-n}$ and moreover

$$f\big(V \times B(0,r)\big) \subset U \times \mathbb{R}^{m-n}.$$

This holds by continuity. See Figure 5.7. We may represent f on $V \times B(0,r)$ as a map

$$f(x,y) = \big(f_1(x,y), f_2(x,y)\big).$$

We have $f(x,0) = (x,0)$. Since $f_2(x,0) = 0$ we can write

$$f_2(x,y) = h_1(x,y)y_1 + \ldots + h_{m-n}(x,y)y_{m-n}$$

with

$$h_i(x,0) = \frac{\partial f_2}{\partial y_i}(x,0).$$

We can then represent F as

$$\begin{aligned} F(x,y,t) &= \left(f_1(x,ty), \frac{1}{t}f_2(x,ty)\right) \\ &= \big(f_1(x,ty), h_1(x,ty)y_1 + \ldots + h_{m-n}(x,ty)y_{m-n}\big). \end{aligned}$$

This map is well-defined and smooth also at $t = 0$. The map at $t = 0$ is

$$F_0(x,y) = F(x,y,0) = \left(x, \frac{\partial f_2}{\partial y}(x,0)y\right).$$

It sends every fibre of E^0 to a fibre of E^1 via a linear map, which is in fact an isomorphism because f is an embedding and hence

$$df_{(x,0)} = \begin{pmatrix} I_n & * \\ 0 & \frac{\partial f_2}{\partial y}(x,0) \end{pmatrix}$$

is an isomorphism. Therefore $F_0 \colon E^0 \to E^1$ is a bundle isomorphism. $\square$

We have proved that the tubular neighbourhood of a submanifold $N \subset M$ is unique up to isotopy and bundle isomorphisms: in particular, this shows that every tubular neighbourhood of N is isomorphic to the normal bundle νN.

5.6.4. Embedding open balls. The uniqueness theorem for tubular neighbourhoods is quite powerful, and it has some remarkable consequences already when N is a point.

Proposition 5.6.5. *Let M be a connected smooth n-manifold. Two embeddings $f, g\colon \mathbb{R}^n \hookrightarrow M$ are always isotopic, possibly after pre-composing g with a reflection in $\mathbb{R}^n$.*

Proof. We may see both f and g as tubular neighbourhoods of $f(0)$ and $g(0)$. Since connected manifolds are homogeneous (Corollary 5.3.4), after an ambient isotopy we may suppose that $f(0) = g(0)$. By the uniqueness of the tubular neighbourhood, the map f is isotopic to $g \circ \psi$ for some linear isomorphism $\psi\colon \mathbb{R}^n \to \mathbb{R}^n$. By Corollary 3.9.11 we may isotope ψ to be either the identity or a reflection. $\square$

The oriented case is more elegant:

Proposition 5.6.6. *Let M be an oriented connected smooth n-manifold. Two orientation-preserving embeddings $f, g\colon \mathbb{R}^n \hookrightarrow M$ are always isotopic.*

5.6.5. Hypersurfaces. Let M be a smooth manifold. A *hypersurface* in M is a submanifold $N \subset M$ of codimension 1.

Proposition 5.6.7. *Let M be orientable. The normal bundle of a hypersurface $N \subset M$ is trivial $\Longleftrightarrow$ N is also orientable.*

Proof. Fix an orientation for M. The normal bundle is a line bundle, and it is trivial $\Longleftrightarrow$ it has a nowhere-vanishing section.

If N is orientable, we fix an orientation. The two orientations of M and N induce a locally coherent orientation on the normal line νN_p for every $p \in N$, which distinguishes between "positive" and "negative" normal vectors, see Exercise 2.6.2. Fix a Riemannian metric on νN, and pick all the positive vectors of norm one: they form a nowhere-vanishing section.

On the other hand, if the normal bundle is trivial, the normal orientation and the orientation of M induce an orientation on N. $\square$

5.6.6. Continuous maps are homotopic to smooth maps. By combining the tubular neighbourhoods and Whitney's Embedding Theorem, we may now prove that every continuous map between smooth manifolds is homotopic to a smooth map. Let M and N be two smooth manifolds.

Proposition 5.6.8. *Let $f\colon M \to N$ be a continuous map, whose restriction to some (possibly empty) closed subset $S \subset M$ is smooth. The map f is continuously homotopic to a smooth map $g\colon M \to N$ with $f(x) = g(x)$ for all $x \in S$, via a homotopy that fixes S pointwise.*

Proof. By Whitney's Embedding Theorem 3.11.8 we may suppose that $N \subset \mathbb{R}^n$ for some n. Let νN be a tubular neighbourhood of N. For every $p \in N$ we let $r(p)$ be the distance from p to the boundary of the open set νN.

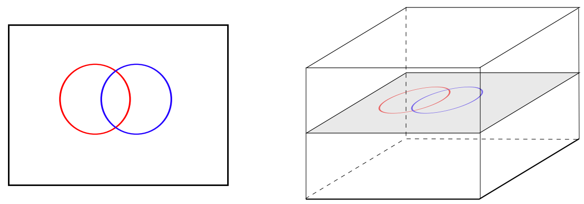

Figure 5.8. Transversality depends on the ambient space: the two curves are transverse in $\mathbb{R}^2$, not in $\mathbb{R}^3$.

By Proposition 3.3.8 there is a smooth map $h\colon M \to \mathbb{R}^n$ with $|h(p) - f(p)| < r(f(p))$. The homotopy $H(p,t) = (1-t)f(p) + th(p)$ lies entirely in νN and hence can be composed with the projection $\pi\colon \nu N \to N$ to give a homotopy $G(p,t) = \pi(H(p,t))$ between f and the smooth $g = \pi \circ h$. $\square$

The proof shows also that g may be chosed to be arbitrarily close to f, but to express "closeness" rigorously we need to see N embedded in some $\mathbb{R}^n$.

Corollary 5.6.9. *Two smooth maps $f, g\colon M \to N$ are continuously homotopic $\iff$ they are smoothly homotopic.*

Proof. Every continuous homotopy $F\colon M \times [0,1] \to N$ can be extended to a continuous map $F\colon M \times \mathbb{R} \to N$ and then be homotoped to a smooth map $G\colon M \times \mathbb{R} \to N$ by keeping $F|_{M\times\{0\}}$ and $F|_{M\times\{1\}}$ fixed. $\square$

5.7. Transversality

We now show that any two smooth maps (and in particular, submanifolds) can be perturbed to cross nicely. The notion of "nice crossing" is surprisingly simple to define and is called *transversality*.

5.7.1. Definition. Let $f\colon M \to N$ and $g\colon W \to N$ be two smooth maps between manifolds, sharing the same target N.

Definition 5.7.1. We say that f and g are *transverse* if for every $p \in M$ and $q \in W$ with $f(p) = g(q)$ we have

$$\operatorname{Im} df_p + \operatorname{Im} dg_q = T_{f(p)}N.$$

In this case we write $f \pitchfork g$.

If $M \subset N$ is a submanifold and f is the inclusion map, we say that g is transverse to M and we write $g \pitchfork M$. Similarly, if both f and g are inclusions, we say that M is transverse to W and we write $M \pitchfork W$.

Set $m = \dim M$, $w = \dim W$, and $n = \dim N$. Note that if $m + w < n$ then $f \pitchfork g \iff$ the maps f and g have disjoint images. See Figure 5.8.

If $W = \{q\}$ is a point, then $f \pitchfork g \iff g(q)$ is a regular value for f.

5.7.2. Fibre bundles. Here is a basic example.

Proposition 5.7.2. *Let $\pi\colon E \to M$ be a fibre bundle. A map $f\colon N \to E$ is transverse to a fibre E_q $\Longleftrightarrow$ q is a regular value for $\pi \circ f$.*

Proof. Pick $p \in N$ with $f(p) \in E_q$. We have $T_{f(p)}E_q = \ker d\pi_{f(p)}$, so

$$\operatorname{Im} df_p + T_{f(p)}E_q = T_{f(p)}E \Longleftrightarrow \operatorname{Im} d(\pi \circ f)_p = T_qN.$$

The proof is complete. $\square$

Exercise 5.7.3. A submanifold $W \subset E$ is the image of a section of a bundle $E \to M$ $\Longleftrightarrow$ it intersects transversely every fibre E_q in a single point.

5.7.3. Intersections. We now extend a theorem from the context of regular values to the wider one of transverse maps.

Proposition 5.7.4. *Let $M \subset N$ be a submanifold and $g\colon W \to N$ a smooth map. If $g \pitchfork M$ then $X = g^{-1}(M)$ is a submanifold of codimension $n - m$.*

Proof. Pick $p \in X$. We look only at a neighbourhood of $q = g(p) \in M$ and after taking a chart we may suppose that $N = \mathbb{R}^n$, $q = 0$, and $M = \mathbb{R}^m \subset \mathbb{R}^n$ embedded as the first m coordinates.

Consider the projection $\pi\colon \mathbb{R}^n \to \mathbb{R}^{n-m}$ onto the last coordinates. Near p we have $X = (\pi \circ g)^{-1}(0)$ and by transversality $\pi \circ g$ is a submersion at p. Therefore X is a submanifold in W, of codimension $n - m$. $\square$

In particular, the intersection $X = M \cap W$ of two transverse submanifolds $M, W \subset N$ is a submanifold with $\operatorname{codim} X = \operatorname{codim} M + \operatorname{codim} W$. We may write $X = M \pitchfork W$. The intersection looks locally as expected:

Proposition 5.7.5. *Every point $p \in X$ has a neighbourhood U and a chart $\varphi\colon U \to \mathbb{R}^n$ that transforms $U \cap M$ and $U \cap W$ into the linear subspaces of the first m and last w coordinates.*

Proof. We work locally, so we can suppose $N = \mathbb{R}^n$ and $p = 0$. If $\dim X = 0$, the map $f\colon M \times W \to \mathbb{R}^n$, $(x, y) \mapsto x + y$ has $df_{(0,0)} = \mathrm{id}$ and hence is a local diffeomorphism, whose local inverse furnishes the desired chart.

In general, we follow a different proof. Locally, we may suppose that $M = \mathbb{R}^m \subset \mathbb{R}^n$ is the space of the first m coordinates. Then we can straighten N keeping M and all its affine translates fixed: details are left as an exercise. $\square$

5.7.4. Thom's Transversality Theorem. We now state a general theorem, that will easily imply that every map can be perturbed to be transverse.

Theorem 5.7.6. *Let $F\colon M \times S \to N$ be a smooth map between manifolds. If F is transverse to some submanifold $Z \subset N$, then $F_s = F(\cdot, s)\colon M \to N$ is also transverse to Z for almost every $s \in S$.*

We mean as usual that the thesis holds for all the values $s \in S$ that lie outside of some zero measure subset.

Proof. Since $F \pitchfork Z$, the preimage $W = F^{-1}(Z) \subset M \times S$ is a smooth submanifold. Consider the projection $\pi\colon M \times S \to S$ and particularly its restriction $\pi|_W\colon W \to S$. We now claim that if s is a regular value for $\pi|_W$ then $F_s \pitchfork Z$. From this we conclude: by Sard's Lemma almost every $s \in S$ is a regular value for $\pi|_W$.

Consider a point $(p, s) \in W$. Since s is regular for $\pi|_W$ we have

$$T_{(p,s)}W + T_{(p,s)}(M \times \{s\}) = T_{(p,s)}(M \times S).$$

Since $F \pitchfork Z$ we have

$$dF_{(p,s)}\big(T_{(p,s)}(M \times S)\big) + T_{F(p,s)}Z = T_{F(p,s)}N.$$

By combining the two equations we get

$$\begin{aligned} T_{F(p,s)}N &= dF_{(p,s)}(T_{(p,s)}W) + dF_{(p,s)}\big(T_{(p,s)}(M \times \{s\})\big) + T_{F(p,s)}Z \\ &= dF_{(p,s)}\big(T_{(p,s)}(M \times \{s\})\big) + T_{F(p,s)}Z \\ &= d(F_s)_p(T_pM) + T_{F(p,s)}Z. \end{aligned}$$

In the second equality we have eliminated the first addendum since it is equal to the third. We have proved that $F_s \pitchfork Z$. $\square$

5.7.5. Consequences. We now draw some consequences from Thom's Transversality Theorem. Here is an amazingly simple application.

Corollary 5.7.7. *Let M be a manifold and $f\colon M \to \mathbb{R}^n$ be a smooth map. Let $Z \subset \mathbb{R}^n$ be a submanifold. For almost all $s \in \mathbb{R}^n$, the translated map*

$$f_s(p) = f(p) + s$$

is transverse to Z.

Proof. The map $F\colon M \times \mathbb{R}^n \to \mathbb{R}^n$, $F(p, s) = f(p) + s$ is a submersion and is hence clearly transverse to any submanifold $Z \subset \mathbb{R}^n$. So Thom's Transversality Theorem applies. $\square$

Corollary 5.7.8. *Let $M, N \subset \mathbb{R}^n$ be any two submanifolds. For almost every $s \in \mathbb{R}^n$ the translate $M + s$ and N are transverse.*

This is interesting already in the case $M = N$. Here is a perturbation theorem for a map between two arbitrary manifolds.

Corollary 5.7.9. *Let $f\colon M \to N$ be a smooth map between manifolds and $W \subset N$ be a submanifold. There is a $g\colon M \to N$ homotopic to f that is transverse to W.*

Proof. Consider N embedded in some $\mathbb{R}^n$ and pick a tubular neighbourhood $\nu N \subset \mathbb{R}^n$ of N with projection $\pi\colon \nu N \to N$. Using a partition of unity, pick a smooth positive function $r\colon N \to \mathbb{R}$ such that $B(q, r(q)) \subset \nu N$ for every $q \in N$. We define the map

$$F\colon M \times B^n \longrightarrow N, \quad F(p, s) = \pi\big(f(p) + r(f(p))s\big).$$

Here $B^n \subset \mathbb{R}^n$ is the unit ball as usual. The map F is a submersion and is hence transverse to any $W \subset N$. Therefore for some $s \in B^n$ the map $g = F_s$ is transverse to W and is homotopic to f through F_{ts}. □

5.7.6. Perturbations. We now show that two maps can always be perturbed to be transverse. We will use tubular neighbourhoods as an essential tool: we start with the following case.

Lemma 5.7.10. *Let $\pi\colon E \to M$ be a vector bundle and $f\colon N \to E$ a smooth map. There is a section $s\colon M \to E$ transverse to f.*

Proof. The product case $E = M \times \mathbb{R}^k$ is particularly simple. Consider a constant section $s(p) = v$ with $v \in \mathbb{R}^k$. We know that $s \pitchfork f \iff v$ is a regular value for the map $\pi_2 \circ f$ where $\pi_2\colon M \times \mathbb{R}^k \to \mathbb{R}^k$ is the projection onto the second factor. By the Sard Lemma, there is a regular value v.

We have covered the product case and we now prove the lemma in general. Exercise 4.5.12 furnishes a bundle $\pi'\colon E' \to M$ such that $E \oplus E' \to M$ is trivial. We consider $E \oplus E'$ as a bundle over E, and construct the pullback bundle $f^*(E \oplus E') \to M$ and its induced map $F\colon f^*(E \oplus E') \to E \oplus E'$.

Since $E \oplus E' \to M$ is trivial, we know by the previous discussion that there is a section $s\colon M \to E \oplus E'$ transverse to F. We get the commutative diagram:

$$\begin{tikzcd}
f^*(E \oplus E') \arrow[r, "F"] \arrow[d, "\pi"'] & E \oplus E' \arrow[d, "\pi"] & M \arrow[l, "s"'] \arrow[dl, "s'"] \\
N \arrow[r, "f"'] & E &
\end{tikzcd}$$

It only remains to prove that $s' = \pi \circ s$ is transverse to f. Suppose that $f(p) = s'(q)$ for some $p \in N$ and $q \in M$. Now $s(q) = (f(p), v)$ for some v in the fibre of $f(p)$, and we also have $F(p, v) = s(q)$. By hypothesis $F \pitchfork s$ so

$$\operatorname{Im} dF_{(p,v)} + \operatorname{Im} ds_q = T_{(f(p),v)}(E \oplus E').$$

By projecting with the differential of π we get

$$\operatorname{Im} df_p + \operatorname{Im} ds'_q = T_{f(p)}E.$$

Therefore $f \pitchfork s'$. The proof is complete. □

We immediately get the following. Let M, N, and W be some manifolds.

Corollary 5.7.11. *Let $i\colon M \hookrightarrow N$ be an embedding and $f\colon W \to N$ a smooth map. There is an embedding $j\colon M \hookrightarrow N$ isotopic to i and transverse to f.*

Proof. Let νM be a tubular neighbourhood of $i(M)$. By the previous lemma there is a section $j\colon M \to \nu M$ transverse to f, isotopic to i. □

If M is compact we can promote the isotopy between i and j to an ambient isotopy of N, as usual. (Actually, it is possible to construct an ambient isotopy between two sections of a tubular neighbourhood even without this compactness hypothesis.) Here is a case of a particular interest:

Corollary 5.7.12. *Any two submanifolds $N, W \subset M$ can be made transverse after modifying the embedding of anyone of them by an isotopy.*

We can also prove a similar theorem when both maps f and g are arbitrary. Of course we must replace "isotopy" by "homotopy" since these maps are arbitrary and need not be embeddings.

Corollary 5.7.13. *Let $f: M \to N$ and $g: W \to N$ be any two smooth maps between manifolds. The map g is homotopic to a map h transverse to f.*

Proof. Consider the commutative diagram:

$$\begin{tikzcd}
M \times W \arrow[r, "f_1"] \arrow[d, "\pi"'] & N \times W \arrow[d, "\pi"] & W \arrow[l, "g_1"'] \arrow[dl, "g"] \\
M \arrow[r, "f"'] & N &
\end{tikzcd}$$

where $f_1(p, q) = \big(f(p), q\big)$, $g_1(q) = \big(g(q), q\big)$, and each π is a projection onto the first factor. The map g_1 is an embedding and can hence be isotoped to a map h_1 that is transverse to f_1. By composing with π we get a homotopy between g and a map $h = \pi \circ h_1$ that is transverse to f. $\square$

CHAPTER 6

Cut and paste

> Cutting and gluing are simple geometrical constructions which, given some smooth manifolds (possibly with boundaries or corners) and additional data where necessary, give rise to new manifolds. On account of their perspicuity, these methods were much used in the days of topology of surfaces, and they remain a very powerful tool
>
> C. T. C. Wall, 1960

In this chapter we address the following question: how can we construct new smooth manifolds? The most effective techniques known consist in building more complicated smooth manifolds out of simpler pieces, glued altogether along smooth maps. A piece is usually a *manifold with boundary*, and the pieces are glued along (portions of) their boundaries. We introduce here the most important decompositions of this kind, the *triangulations* and the *handle decompositions*. We then use these to classify all compact surfaces.

6.1. Manifolds with boundary

We introduce a variation of the definition of smooth manifold that allows the presence of some particular *boundary points*. This is a very natural notion and is present everywhere in differential topology and geometry.

Most of the definitions and theorems about smooth manifolds also apply to manifolds with boundary, with appropriate modifications.

6.1.1. Definition. Consider the upper half-space

$$\mathbb{R}^n_+ = \big\{ x \in \mathbb{R}^n \bigm| x_n \geq 0 \big\}$$

in $\mathbb{R}^n$. Its *boundary* is the horizontal hyperplane $\partial\mathbb{R}^n_+ = \{x_n = 0\}$, while its *interior* is the open subset $\mathbb{R}^n_+ \setminus \partial\mathbb{R}^n_+ = \{x_n > 0\}$.

We now redefine the notions of charts and atlases in a more general context that allows the presence of boundary points: everything will be like in Section 3.1.1, only with $\mathbb{R}^n_+$ instead of $\mathbb{R}^n$.

Let M be a topological space. A $\mathbb{R}^n_+$-*chart* is a homeomorphism $\varphi\colon U \to V$ from an open set $U \subset M$ onto an open set $V \subset \mathbb{R}^n_+$. A *smooth* $\mathbb{R}^n_+$-*atlas* in M is a set $\{\varphi_i\}$ of $\mathbb{R}^n_+$-charts with $\cup U_i = M$ such that the transition maps φ_{ij} are smooth where they are defined. Note that the domain of φ_{ij} is an open subset of $\mathbb{R}^n_+$ and may not be open in $\mathbb{R}^n$, so the correct notion of smoothness is that stated in Definition 3.3.4.

Definition 6.1.1. A *smooth manifold with boundary* is a topological space M equipped with a smooth $\mathbb{R}^n_+$-atlas.

We will drop the symbol $\mathbb{R}^n_+$ from the notation. As in Section 3.1.1, two compatible atlases are meant to give the same smooth structure.

6.1.2. The boundary. Let M be a smooth manifold with boundary. The points $p \in M$ that are sent to $\partial\mathbb{R}^n_+$ via some chart form the *boundary* ∂M. There is no possible ambiguity here, since if one chart sends p inside $\partial\mathbb{R}^n_+$, then all charts do (exercise).

The boundary ∂M is naturally a $(n-1)$-dimensional smooth manifold *without* boundary. Indeed by restricting the charts to ∂M we get an atlas for ∂M with values onto some open sets of the hyperplane $\partial\mathbb{R}^n_+$, that we identify with $\mathbb{R}^{n-1}$.

Example 6.1.2. Every open subset $U \subset \mathbb{R}^n_+$ is a smooth manifold with boundary $\partial U = U \cap \partial\mathbb{R}^n_+$. The atlas consists of just the identity chart.

The *interior* of M is $\text{int}(M) = M \setminus \partial M$. It is a manifold without boundary. The notions of smooth maps and diffeomorphisms extend to this new boundary context without any modification.

6.1.3. Regular domains. We now describe one important source of examples. Let M be a smooth n-manifold without boundary.

Definition 6.1.3. A *regular domain* is a subset $D \subset M$ such that for every $p \in D$ there is a chart $\varphi\colon U \to V$ with $p \in U$ and $V \subset \mathbb{R}^n$ that sends $U \cap D$ onto an open subset of $\mathbb{R}^n_+$.

Every regular domain D has a natural structure of manifold with boundary, obtained by taking as an atlas all the charts φ of this type.

Exercise 6.1.4. For every $a < b$, the closed segment $[a, b]$ is a domain in $\mathbb{R}$ and hence a manifold with boundary consisting of the points a and b.

Here is a concrete way to construct regular domains:

Proposition 6.1.5. *Let M be a manifold without boundary and $f : M \to \mathbb{R}$ a smooth function. If y_0 is a regular value, then $D = f^{-1}(-\infty, y_0]$ is a regular domain with $\partial D = f^{-1}(y_0)$.*

Proof. Consider a point $p \in D$. If $f(p) < y_0$, the point p has an open neighbourhood fully contained in D that can be sent inside the interior of $\mathbb{R}^n_+$ via some chart.

If $f(p) = y_0$, by Proposition 3.8.10 there are charts $\varphi\colon U \to \mathbb{R}^n$ and $\psi\colon W \to \mathbb{R}$ with $p \in U$ and $f(U) \subset W$ such that $\psi \circ f \circ \varphi^{-1}(x_1, \ldots, x_n) =$

x_n and we may also require that $\varphi(p) = 0$ and ψ is orientation-reversing. Therefore $\varphi(U \cap D) = \mathbb{R}^n_+$. $\square$

Corollary 6.1.6. *The unit disc*

$$D^n = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$$

is a domain in $\mathbb{R}^n$ *with boundary* $\partial D^n = S^{n-1}$.

Proof. We pick $f(x) = \|x\|^2$ and get $D^n = f^{-1}(-\infty, 1]$. Every non-zero value is regular. $\square$

Remark 6.1.7. The square $[-1,1] \times [-1,1]$ is not a regular domain in $\mathbb{R}^2$, because it has corners. More generally, the product $M \times N$ of two manifolds with boundary is *not* necessarily a manifold with boundary, because if $\partial M \neq \varnothing$ and $\partial N \neq \varnothing$ then some corners arise. However, if $\partial M = \varnothing$ then $M \times N$ is naturally a manifold with boundary and

$$\partial(M \times N) = M \times \partial N.$$

For instance, the cylinder $S^1 \times [-1,1]$ is a surface with boundary, and the boundary consists of the two circles $S^1 \times \{\pm 1\}$. More generally $S^m \times D^n$ is a manifold with boundary and

$$\partial(S^m \times D^n) = S^m \times S^{n-1}.$$

6.1.4. Tangent space. The definition of tangent space via derivations also extends *verbatim* to manifolds with boundary. For every point $p \in \mathbb{R}^n_+$, included those on the boundary, we get $T_p\mathbb{R}^n_+ = \mathbb{R}^n$. For a general n-manifold M with boundary, the space T_pM is a n-dimensional vector space at every $p \in M$, included the boundary points.

At every boundary point $p \in \partial M$ the tangent space $T_p\partial M$ is naturally a hyperplane inside T_pM, that divides T_pM into two components, the "interior" and "exterior" tangent vectors, according to whether they point towards the interior of M or the exterior. This subdivision between interior and exterior is obvious in $\mathbb{R}^n_+$ and transferred to M unambiguously via charts.

As in the boundaryless case, every smooth map $f\colon M \to N$ induces a differential $df_p\colon T_pM \to T_{f(p)}N$ at every point $p \in M$. Note that a smooth map f may send a boundary point to an interior point, or an interior point to a boundary point.

6.1.5. Orientation. One nice feature of manifolds with boundary is that an orientation on M induces one on its boundary ∂M.

Let M be an oriented manifold with boundary of dimension $n \geq 2$. Recall that an orientation on M is a locally coherent way of assigning an orientation to all the tangent spaces T_pM. For every $p \in \partial M$, we choose an exterior vector $v \in T_pM$ and note that

$$T_pM = \mathrm{Span}(v) \oplus T_p\partial M.$$

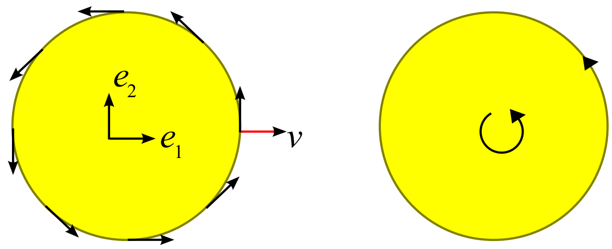

Figure 6.1. The canonical orientation on the disc (given by the canonical basis e_1, e_2) induces the counterclockwise orientation on the boundary circle (left). We may write conveniently the orientations on a surface and on a curve using (curved) arrows (right)

With this subdivision, the orientation on T_pM induces one on $T_p\partial M$: we say that a basis $v_2 . \ldots . v_n$ for $T_p\partial M$ is positive $\Longleftrightarrow$ the basis $v, v_2, \ldots, v_n$ is positive for T_pM. By looking on a chart we see that this is a locally coherent assignment that does not depend on the choice of the exterior vector v.

We now consider the one-dimensional case, that is slightly different. First, we define an *orientation* on a point to be the assignment of a sign ± 1. When not mentioned, a point is equipped with the $+1$ orientation: points are in fact the only manifolds that have a canonical orientation!

If M^1 is an oriented 1-manifold, we orient every boundary point $p \in \partial M^1$ as 1 or -1 depending on whether the vectors pointing outside in the line T_pM are positive or negative.

Every domain in $\mathbb{R}^n$ is naturally oriented by the canonical basis $e_1, \ldots, e_n$, so for instance the disc D^n has a canonical orientation. This canonical orientation induces an orientation on the boundary sphere S^{n-1}. The case $n = 2$ is shown in Figure 6.1.

6.1.6. Immersions, embeddings, submanifolds. Let M, N be manifolds with boundary. We define an *immersion* as usual as a map $f : M \to N$ with injective differentials, and then an *embedding* as an injective immersion $f : M \to N$ that is a homeomorphism onto its image.

Definition 6.1.8. Let N be a manifold. A *submanifold* is the image of an embedding $f : M \hookrightarrow N$.

The reader should note that, as opposite to Definition 3.7.1, we are not saying that a submanifold should look locally like some simple model. This is by far not the case here: Figure 6.2 shows that many different kinds of local configurations arise already when one embeds a segment in the half-plane $\mathbb{R}^2_+$. In higher dimensions things may also get more complicated.

In some cases, we may require the submanifold to satisfy some requirements. For instance, a submanifold $M \subset N$ is *neat* if

$$\partial M = M \cap \partial N$$

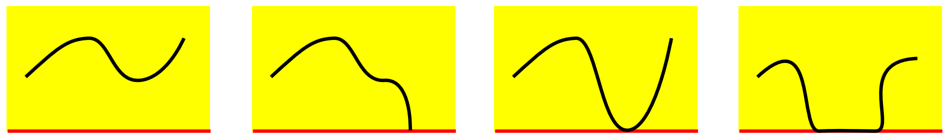

Figure 6.2. Different kinds of compact 1-dimensional submanifolds inside the half-plane $\mathbb{R}^2_+$.

and moreover M meets ∂N transversely, that is at every $p \in \partial M$ we have $T_pM + T_p\partial N = T_pN$.

6.1.7. Homotopy, isotopy, ambient isotopy. The notions of homotopy, isotopy, and ambient isotopy also extend *verbatim* to manifolds with boundary.

Some important theorems also hold, with the same proofs, for manifolds with boundary: if M is a manifold with boundary, it may be embedded in $\mathbb{R}^n$ via some proper map (Theorem 3.11.8), and if M is compact every two isotopic embeddings $f, g\colon M \to N$ are also ambiently isotopic, for every N without boundary (Theorem 5.3.3).

6.1.8. Fibre bundles. The theory of bundles extends to manifolds with boundary with minor obvious modifications. On a fibre bundle $E \to M$, we can allow M to have boundary, and in that case the trivialising neighborhoods will be diffeomorphic to open subsets of $\mathbb{R}^n_+$, or we can allow the fibre F to have boundary; however, some care is needed if *both* M and F have boundary, because some corners would arise and E would not be a smooth manifold.

We now introduce an important case where the fibre F is a disc.

6.1.9. The unit disc bundle. Let $E \to M$ be a vector bundle over a manifold M without boundary. Fix a Riemannian metric g for E. The *unit disc bundle* is the submanifold with boundary

$$D(E) = \big\{ v \in E \ \big|\ \|v\| \leq 1 \big\}.$$

The projection π restricts to a projection $\pi\colon D(E) \to M$ and one sees as in Proposition 4.5.7 that this is a disc bundle (a fibre bundle with $F = D^k$) and that it does not depend on g up to isotopy (that is, up to an isomorphism of $E \to M$ that is isotopic to the identity).

The boundary of $D(E)$ is the unit sphere bundle $S(E)$. The interior of $D(E)$ may be given a bundle structure isomorphic to $E \to M$.

6.1.10. Closed tubular neighbourhoods. Let M be a m-manifold and $N \subset \mathrm{int}(M)$ be a compact submanifold without boundary. Since N avoids ∂M, it has a tubular neighbourhood $\nu N \subset M$.

Definition 6.1.9. A *closed tubular neighbourhood* of N in M is the unit disc bundle of any tubular neighbourhood of N.

To better distinguish a tubular neighbourhood from a *closed* tubular neighbourhood, we can call the first an *open* tubular neighbourhood. We will use the notation νN for both; note that the interior of a closed tubular neighbourhood may in turn be given the structure of an open tubular neighbourhood, so one can switch easily from open to closed and vice-versa.

The closed tubular neighbourhood of a compact submanifold is also compact: for this reason it is sometimes better to work with closed tubular neighbourhoods; for instance, we may promote isotopy to ambient isotopy:

Proposition 6.1.10. *A compact submanifold $M \subset \text{int}(N)$ without boundary has a unique closed tubular neighbourhood up to ambient isotopy in N.*

Proof. We already know that tubular neighbourhoods are isotopic, and hence also the closed tubular neighbourhoods are. Since these are compact, the isotopy may be promoted to an ambient isotopy. $\square$

6.1.11. Collar. Let M be a manifold with boundary, and N be the union of some connected components of ∂M. A *collar* of N in M is an embedding

$$i\colon N \times [0,1) \hookrightarrow M$$

such that $i(p,0) = p$ for every $p \in N$. The collars should be interpreted as the tubular neighbourhoods of the boundary.

Proposition 6.1.11. *The manifold N has a unique collar up to isotopy.*

The proof is the same as that for tubular neighbourhoods, and we omit it. We can define analogously a *closed collar* to be an embedding of $N \times [0,1]$ as above; if N is compact, the closed collar is unique up to ambient isotopy.

Exercise 6.1.12. For every manifold M the inclusion $\text{int}(M) \hookrightarrow M$ is a homotopy equivalence.

Hint. Use a collar for ∂M to define the homotopy inverse. $\square$

6.1.12. One-dimensional manifolds. We leave to the reader to solve the following exercise, that fully classifies all connected one-dimensional manifolds.

Exercise 6.1.13. Every connected one-dimensional manifold is diffeomorphic to one of the following:

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

In particular S^1 is the unique connected compact one-dimensional manifold without boundary.

6.1.13. Discs. Let M be a n-manifold. We define a *disc* in M to be an embedding $f\colon D^n \hookrightarrow \mathrm{int}(M)$. As an example, a closed tubular neighbourhood of a point is a disc. We can now prove this remarkable theorem.

Theorem 6.1.14 (The Disc Theorem). *Let M be a connected smooth n-manifold. Two discs $f, g\colon D^n \hookrightarrow M$ are always ambiently isotopic, possibly after pre-composing g with a reflection.*

Proof. Since $B^n = \mathrm{int}(D^n)$ is diffeomorphic to $\mathbb{R}^n$, the restrictions $f|_{B^n}$ and $g|_{B^n}$ are isotopic by Proposition 5.6.5. Now we can shrink isotopically f and g to the maps $\bar{f}(v) = f(\frac{v}{2})$ and $\bar{g}(v) = g(\frac{v}{2})$ and deduce that f and g are also isotopic. Since D^n is compact, isotopy is promoted to ambient isotopy. $\square$

With a little abuse we sometimes call a *disc* the image of an embedding $f\colon D^n \hookrightarrow M$. With this interpretation, which disregards the parametrisation, two discs are always ambiently isotopic. The reader should appreciate how powerful this theorem is, already in the only apparently simpler case $M = \mathbb{R}^n$, for instance in dimension $n = 2$.

The Disc Theorem was proved by Palais in 1960.

6.1.14. Spheres. We end this section by describing how every sphere decomposes beautifully into two simple submanifolds with boundary.

For every $0 < k < n$ we identify $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$ and write a point of $\mathbb{R}^n$ as (x, y) with $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^{n-k}$. By radial expansion we may easily construct a homeomorphism between D^n and $D^k \times D^{n-k}$, which restricts to a homeomorphism between S^{n-1} and the topological boundary of $D^k \times D^{n-k}$. The latter in turn decomposes into two closed subsets

$$S^{k-1} \times D^{n-k}, \qquad D^k \times S^{n-k-1}$$

whose intersection is $S^{k-1} \times S^{n-k-1}$. Having understood this simple topological phenomenon, we write an analogous decomposition of S^{n-1} in the smooth setting. We write

$$S^{n-1} = \big\{(x, y) \bigm| \|x\|^2 + \|y\|^2 = 1\big\}.$$

We now consider the subsets

$$A = \big\{(x, y) \in S^{n-1} \bigm| \|x\|^2 \leq \tfrac{1}{2}\big\}, \quad B = \big\{(x, y) \in S^{n-1} \bigm| \|y\|^2 \leq \tfrac{1}{2}\big\}.$$

These are both domains, since $\frac{1}{2}$ is a regular value for the maps $(x, y) \mapsto \|x\|^2$ or $\|y\|^2$ on S^{n-1} (exercise). The common boundary

$$A \cap B = \big\{(x, y) \in S^{n-1} \bigm| \|x\|^2 = \|y\|^2 = \tfrac{1}{2}\big\}$$

is diffeomorphic to $S^{k-1} \times S^{n-k-1}$ via the map $(x, y) \mapsto (\sqrt{2}x, \sqrt{2}y)$. We now identify the domains: the map

$$A \longrightarrow D^k \times S^{n-k-1}, \quad (x, y) \longmapsto \left(\sqrt{2}x, \frac{y}{\|y\|}\right)$$

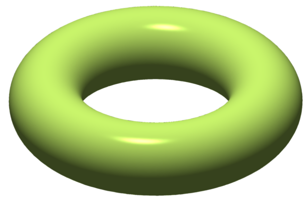

Figure 6.3. A solid torus $D^2 \times S^1$. Its complement inside S^3 is another solid torus: can you see it?

is a diffeomorphism, with inverse $(x, y) \mapsto \frac{\sqrt{2}}{2}\left(x, (2 - \|x\|^2)y\right)$.

We have discovered that S^{n-1} decomposes into two domains $A \cong D^k \times S^{n-k-1}$ and $B \cong S^{k-1} \times D^{n-k}$ with common boundary $S^{k-1} \times S^{n-k-1}$. We also note that A and B are closed tubular neighborhoods of the spheres

$$S^n \cap \{x = 0\} \cong S^{n-k-1}, \qquad S^n \cap \{y = 0\} \cong S^{k-1}.$$

The 3-manifold $S^1 \times D^2$ is a *solid torus*. The 3-sphere S^3 decomposes into two solid tori $S^1 \times D^2$ and $D^2 \times S^1$ along their common boundary $S^1 \times S^1$. See Figure 6.3.

6.2. Cut and paste

We now introduce some basic cut and paste manipulations that allow to modify the topology of a smooth manifold.

6.2.1. Punctures. Let M be a connected smooth n-manifold, possibly with boundary. The simplest topological modification we can make on M is to remove a point $p \in \text{int}(M)$. By Corollary 5.3.4, the new manifold $M \setminus \{p\}$ does not depend (up to diffeomorphism) on p, and we say that it is obtained by *puncturing* M.

A variation of this modification consists of picking a disc $D \subset M$ and removing its interior: the new manifold

$$M' = M \setminus \text{int}(D)$$

has the same boundary components as M, plus one new sphere ∂D. The manifold M' does not depend (up to diffeomorphisms) on the chosen disc D by the Disc Theorem 6.1.14.

Exercise 6.2.1. The manifolds $M \setminus \{p\}$ and $M' \setminus \partial D$ are diffeomorphic.

Exercise 6.2.2. If $M = S^n$, we get $M \setminus \{p\} \cong \mathbb{R}^n$ and $M' \cong D^n$.

Exercise 6.2.3. If $M = D^n$ then $M' \cong S^{n-1} \times [-1, 1]$.

Exercise 6.2.4. If $\dim M \geq 3$, then $\pi_1(M') \cong \pi_1(M \setminus \{p\}) \cong \pi_1(M)$.

Hint. Use Van Kampen. □

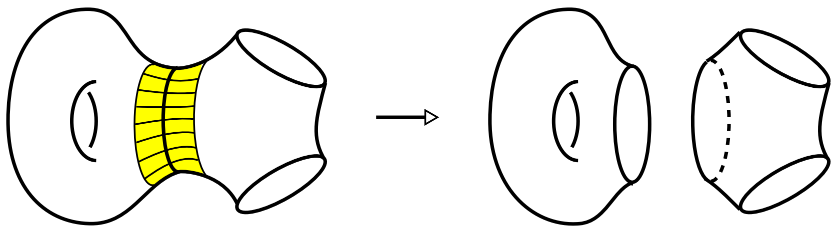

Figure 6.4. How to cut a manifold along a two-sided hypersurface.

6.2.2. Removing submanifolds. We now extend the above manipulation from points to arbitrary compact submanifolds.

Let M be a smooth manifold and $N \subset \text{int}(M)$ a compact submanifold of some codimension $k \geq 1$. The complement $M \setminus N$ is a new manifold. Again, a variation consists in taking a closed tubular neighbourhood νN and considering

$$M' = M \setminus \text{int}(\nu N).$$

The manifold M' has a new compact boundary component $\partial \nu N$, which is a S^{k-1}-bundle over N. The manifold M' only depends on N and not on the tubular neighbourhood νN since it is unique up to ambient isotopy.

This operation is particularly interesting if N has codimension 1 and is *two-sided*, that is has trivial normal bundle $\nu N \cong N \times \mathbb{R}$. For instance, this holds if both M and N are orientable: see Proposition 5.6.7. In this case the new manifold M' has two new boundary components, both diffeomorphic to N. See Figure 6.4. We say that M' is obtained by *cutting* M along N.

Example 6.2.5. By cutting S^n along its equator S^{n-1} we get two discs.

If M, N are connected and N has codimension one, the new manifold M' may be connected or not; in the first case, we say that N is *non-separating*, and *separating* in the second.

6.2.3. Pasting along the boundary. Pasting is of course the inverse of cutting. Let M be a (possibly disconnected) manifold, let N_1, N_2 be two boundary components of M, and $\varphi\colon N_1 \to N_2$ be a diffeomorphism. We now define a new manifold M' obtained by *pasting M along φ*.

A naïve construction would be to define M' as $M/_\sim$ where $\sim$ is the equivalence relation that identifies $p \sim \varphi(p)$ for all $p \in N_1$. The result is indeed a topological manifold, but it is not obvious to assign a smooth atlas to $M/_\sim$. So we abandon this route, and we define M' instead by overlapping open collars as suggested by Figure 6.5.

Here are the details. We identify two disjoint closed collars of N_1 and N_2 in M with $N_1 \times [0,1]$ and $N_2 \times [0,1]$, where $N_i = N_i \times \{0\}$. The manifold M' is obtained from M by first removing N_1 and N_2, and then identifying the open

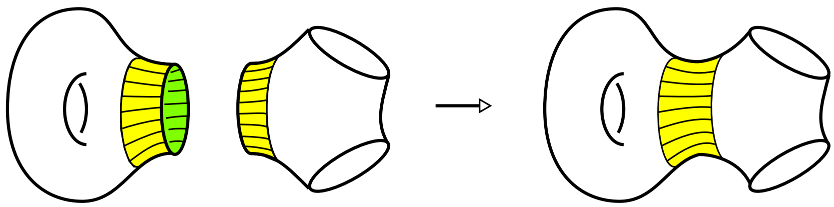

Figure 6.5. How to paste two boundary components N_1 and N_2 via a diffeomorphism φ. To get a new smooth manifold, we pick two collars and we make them overlap.

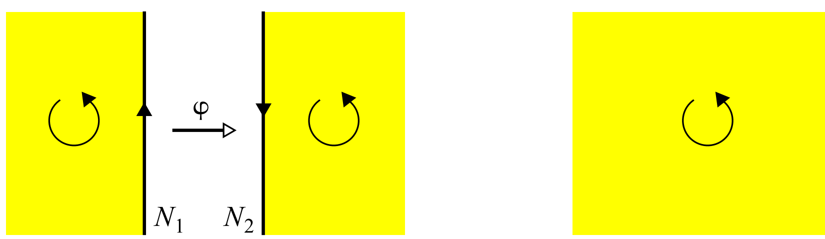

Figure 6.6. If the gluing map φ is orientation-*reversing*, the orientations extend to the new manifold M'.

subsets $N_1 \times (0,1)$ and $N_2 \times (0,1)$ via the map $\Phi\colon (p,t) \mapsto \big(\varphi(p), 1-t\big)$. The smooth structure on M' is now easily induced by that of M.

Proposition 6.2.6. *The manifold M' depends up to diffeomorphism only on M and on the isotopy class of φ.*

Proof. Different closed collars are ambiently isotopic and hence produce diffeomorphic manifolds M'. If F is an isotopy between $\varphi_0 = F_0$ and $\varphi_1 = F_1$, a diffeomorphism between the resulting manifolds M'_0 and M'_1 is constructed as follows: it is the identity outside the collar, and $(p,t) \mapsto \big(F_t(\varphi_0^{-1}(p)), t\big)$ on the collars. $\square$

Remark 6.2.7. Suppose that M is oriented. Both N_1 and N_2 inherit an orientation. If φ is orientation-*reversing*, then Φ is orientation-*preserving* and hence the orientation of M induces naturally an orientation on M'. So, if you want orientations to extend, you need to glue along orientation-reversing maps φ. See Figure 6.6.

Exercise 6.2.8. The smooth manifold M' is homeomorphic to the topological manifold $M/_\sim$ obtained from M by identifying $p \sim \varphi(p)$ for every $p \in N_1$.

In light of this fact, we will often think of M' simply as the topological space $M/_\sim$, equipped with a smooth atlas induced by φ.

6.2.4. Self-diffeomorphisms. Proposition 6.2.6 suggests that it is important to understand the self-diffeomorphisms of a manifold up to isotopy. We now state a couple of basic results on this quite difficult problem.

Let N be a connected smooth orientable manifold. We denote by $\mathrm{Diffeo}(N)$ the group of all self-diffeomorphisms of N. If N is orientable, then the group decomposes into

$$\mathrm{Diffeo}(N) = \mathrm{Diffeo}^+(N) \sqcup \mathrm{Diffeo}^-(N)$$

where $\mathrm{Diffeo}^\pm(N)$ is the subset of all self-diffeomorphisms that preserve/invert the orientation of N. We say that N is *mirrorable* if $\mathrm{Diffeo}^-(N)$ is non-empty. We say that two self-diffeomorphisms $\varphi, \psi \in \mathrm{Diffeo}(N)$ are *cooriented* if they either both preserve or both invert the orientation.

Exercise 6.2.9. If φ, ψ are isotopic, they are cooriented.

The converse is also sometimes true.

Proposition 6.2.10. *Two cooriented diffeomorphisms of S^1 are isotopic.*

Proof. Let $\varphi_0, \varphi_1 \colon S^1 \to S^1$ be two cooriented diffeomorphisms. They lift to smooth maps $\tilde\varphi_0, \tilde\varphi_1 \colon \mathbb{R} \to \mathbb{R}$ between their universal covers, that are monotone (that is, $\tilde\varphi_0'(t), \tilde\varphi_1'(t) > 0$ (or < 0) $\forall t$) and periodic (that is, $\varphi_i(t + 2\pi) = \varphi_i(t) + 2\pi\ \forall t$). The convex combination

$$\tilde\varphi_t(x) = (1 - t)\tilde\varphi_0(x) + t\tilde\varphi_1(x)$$

is also periodic and monotone, hence it descends to a monotone map $\varphi_t \colon S^1 \to S^1$. Each φ_t is hence a covering, but since it is homotopic to φ_0 it is a diffeomorphism: we get an isotopy between φ_0 and φ_1. □

This fact has important consequences when we want to glue two surfaces along their boundaries. Let Σ_1 and Σ_2 be two surfaces with boundary as in Figure 6.5, and we want to glue them along a diffeomorphism $\varphi \colon C_1 \to C_2$, between two connected boundary components C_1 and C_2 of Σ_1 and Σ_2, both diffeomorphic to a circle S^1. The proposition tells us that there are only two possible gluing maps φ up to isotopy.

6.2.5. Doubles. Here is a simple kind of pasting that applies to every manifold with boundary.

The *double DM* of a manifold M with boundary is obtained by taking two identical copies M_1, M_2 of M and defining $\varphi \colon \partial M_1 \to \partial M_2$ as the *identity map*, that is the one that sends every point in ∂M_1 to its corresponding point in ∂M_2. Then DM is obtained by pasting $M_1 \sqcup M_2$ along φ.

The doubled manifold DM has no boundary. If M is compact, then DM also is.

Exercise 6.2.11. The double of D^n is diffeomorphic to S^n. The double of a cylinder $S^1 \times [0, 1]$ is diffeomorphic to a torus $S^1 \times S^1$. What is the double of a Möbius strip?

6.2.6. Exotic spheres. We now investigate the following apparently innocuous construction: we pick a self-diffeomorphism $\varphi\colon S^{n-1} \to S^{n-1}$ and we glue two copies of D^n along φ, thus getting a new manifold M without boundary. What kind of smooth manifold M do we get?

Exercise 6.2.11 says that if $\varphi = \mathrm{id}$ then M is diffeomorphic to S^n. More generally, in the topological category, the answer does not depend on φ.

Proposition 6.2.12. *The manifold M is homeomorphic to S^n. If φ extends to a self-diffeomorphism of D^n, then M is also diffeomorphic to S^n.*

Proof. By Exercise 6.2.8 the manifold M is homeomorphic to the topological manifold $D_1 \cup_\varphi D_2$ obtained by identifying p with $\varphi(p)$. We define a continuous map

$$F\colon D_1 \cup_{\mathrm{id}} D_2 \longrightarrow D_1 \cup_\varphi D_2$$

by coning φ, that is: if $v \in D_1$ then $F(v) = v$, while if $v \in D_2$ we set

$$F(v) = \begin{cases} |v|\varphi\big(\frac{v}{|v|}\big) & \text{if } v \neq 0 \\ 0 & \text{if } v = 0 \end{cases}$$

The map F is a homeomorphism. By Exercise 6.2.11 we have $D_1 \cup_{\mathrm{id}} D_2 \cong S^n$, and this completes the proof that M is homeomorphic to S^n.

If φ extends to a diffeomorphism $\Phi\colon D^n \to D^n$, we can replace $F|_{D_2}$ with Φ and get a diffeomorphism. More precisely, to get a smooth map we need to smoothen it at the equator ∂D^n like we do when we compose two smooth isotopies (details are left as an exercise). $\square$

Corollary 6.2.13. *If $n = 2$ then M is diffeomorphic to S^2.*

Proof. Up to isotopy, the gluing map $\varphi\colon S^1 \to S^1$ is either the identity or a reflection $z \mapsto \bar{z}$, and they both extend to self-diffeomorphisms of D^2. $\square$

The striking fact here is that when $n \geq 7$ the smooth manifold M may *not* be diffeomorphic to S^n, despite being homeomorphic to it. This implies in particular that there are some crazy self-diffeomorphisms of S^n that are not isotopic neither to the identity nor to a reflection, and moreover they do not extend to self-diffeomorphisms of D^n.

Remark 6.2.14. A smooth manifold diffeomorphic but not homeomorphic to S^n is called an *exotic sphere*. In dimension $n \geq 7$ there are many exotic spheres, and they are all constructed in this way. On the other hand, there are no exotic spheres in dimensions $n = 1, 2, 3, 5, 6$. The dimension 4 remains a total mystery: we do not know if there are exotic spheres, and if there are, they are certainly *not* constructed in this way (that is, by gluing two discs). Even more puzzling, we know that the number of exotic spheres (considered up to diffeomorphism) is finite in every dimension – for instance these are 27 in dimension 7 – except in dimension four, where the number of exotic spheres could be any value from 0 to ∞, both extremes included, as far as we know.

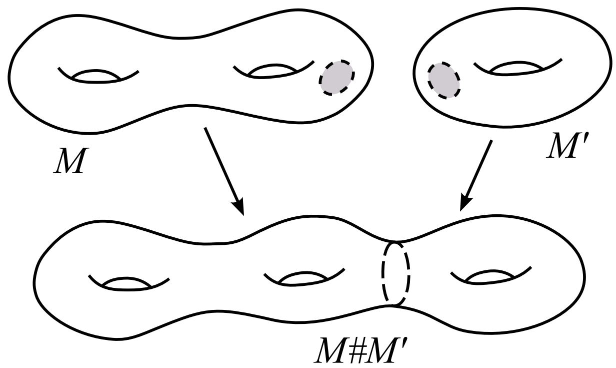

Figure 6.7. The connected sum of two compact surfaces.

6.3. Connected sums and surgery

We now introduce some more elaborate manipulations. The most important ones are the *connected sum* that "connects" two manifolds along a tube, and the more general *surgery* that roughly replaces a k-sphere (with trivial normal bundle) with a $(n-k-1)$-sphere. The boundary versions of these manipulations are also important.

6.3.1. Definition. Let M_1 and M_2 be two connected oriented n-manifolds, possibly with boundary. We now define a new oriented manifold $M_1\#M_2$ called the *connected sum* of M_1 and M_2.

We define the orientation-reversing diffeomorphism of the punctured disc

$$\alpha\colon \operatorname{int}(D^n)\setminus\{0\}\longrightarrow \operatorname{int}(D^n)\setminus\{0\},\quad \alpha(v)=\big(1-|v|\big)\frac{v}{|v|}.$$

We pick two arbitrary embeddings

$$f_1\colon D^n\hookrightarrow \operatorname{int}(M_1),\qquad f_2\colon D^n\hookrightarrow \operatorname{int}(M_2)$$

such that f_1 is orientation-preserving and f_2 is orientation-reversing. Then we glue the punctured manifolds $M_1\setminus f_1(0)$ and $M_2\setminus f_2(0)$ via the diffeomorphism

$$f_2\circ\alpha\circ f_1^{-1}\colon f_1\big(\operatorname{int}(D^n)\setminus\{0\}\big)\longrightarrow f_2\big(\operatorname{int}(D^n)\setminus\{0\}\big).$$

The resulting smooth manifold is the *connected sum* of M_1 and M_2 and is denoted as

$$M_1\#M_2.$$

Since $f_2\circ\alpha\circ f_1^{-1}$ is orientation-preserving, the manifold $M_1\#M_2$ is naturally oriented. You may visualise an example in Figure 6.7. By the Disc Theorem 6.1.14 the manifold $M_1\#M_2$ does not depend, up to orientation-preserving diffeomorphisms, on the maps φ_1 and φ_2.

Remark 6.3.1. The connected sum $M_1\#M_2$ may also be described as a two-steps cut-and-paste operation, where:

(1) first, we remove $f_i\big(\mathrm{int}(D_i)\big)$ from M_i, thus creating a new boundary component $f_i(\partial D_i)$ for M_i, $\forall i = 1, 2$;
(2) then, we paste the two new boundary components via the diffeomorphism $f_2 \circ f_1^{-1} \colon \partial D_1 \to \partial D_2$.

We leave as an exercise to prove that this definition of $M_1 \# M_2$ is equivalent to the one given above. In light of the exotic spheres construction, it is important to require the gluing map to be $f_2 \circ f_1^{-1}$ and not any map.

We may see $\#$ as a binary operation on the set[1] of all oriented connected n-manifolds considered up to diffeomorphism.

Proposition 6.3.2. *The connected sum is commutative and associative, and S^n is the neutral element. That is, there are diffeomorphisms*

$$M \# N \cong N \# M, \qquad M \#(N \# P) \cong (M \# N) \# P, \qquad M \# S^n \cong M.$$

Proof. Commutativity is obvious. Associativity holds because we can separate the discs using isotopies, so that both connected sums can be performed simultaneously.

To construct $M \# S^n$ we follow Remark 6.3.1. We choose $\varphi_2 \colon D^n \hookrightarrow S^n$ to be the standard parametrisation of the upper hemisphere. The two-steps operation consists of substituting the upper hemisphere with the lower one along the same map, and this does not change the manifold M. □

The connected sum may be defined also for non-oriented manifolds, but in this case the resulting manifold $M \# N$ is not unique: there are two possibilities, and these may produce non-diffeomorphic manifolds in some cases. We have used orientations here only to simplify the theory.

6.3.2. Compact surfaces. Enough for the theory, we need examples. One-dimensional manifolds are not very exciting, so we turn to surfaces. We already know some compact connected surfaces:

$$S^2, \quad \mathbb{RP}^2, \quad D^2, \quad S^1 \times [0,1], \quad S^1 \times S^1, \quad M$$

where M is the compact Möbius strip, considered with its (connected!) boundary. Can we add more surfaces to this list?

Definition 6.3.3. The *genus-g surface* S_g is the connected sum

$$S_g = \underbrace{T \# \ldots \# T}_{g}$$

of g copies of the torus $T = S^1 \times S^1$.

[1] The suspicious reader may object that smooth manifolds do not form a set. However, if we consider them up to diffeomorphism, we may use Whitney's embedding theorem and see them as subsets of some $\mathbb{R}^n$, and the subsets of $\mathbb{R}^n$ of course form a set.

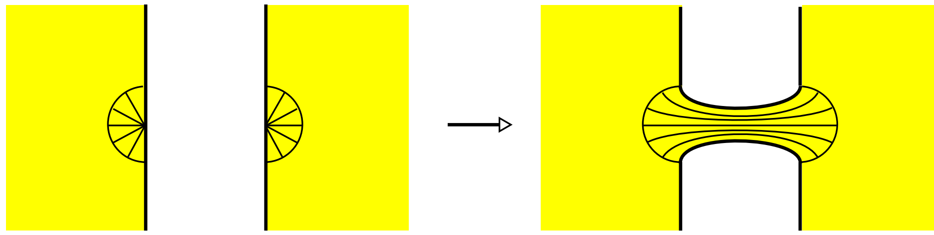

Figure 6.8. The ∂-connected sum of two manifolds.

By convention, the surface of genus zero S_0 is the sphere S^2, and that of genus one S_1 is the torus. We have

$$S_g \# S_h \cong S_{g+h}.$$

Figure 6.7 shows that $S_2 \# S_1 \cong S_3$. Note that the torus T is mirrorable, so each time we make a connected sum with T it is not really important which orientation we put on T.

6.3.3. ∂-connected sum. A ∂-connected sum is an operation similar to the connected sum, where a bridge is added to connect two portions of the boundaries as in Figure 6.8.

The construction goes as follows. We consider the *half-disc* $D^n_+ = D^n \cap \mathbb{R}^n_+$. We define $D^{n-1} = D^n_+ \cap \{x_n = 0\}$ and $\operatorname{int}(D^n_+) = D^n_+ \cap \{|x| < 1\}$. We consider the same orientation-reversing diffeomorphism as above

$$\alpha \colon \operatorname{int}(D^n_+) \setminus \{0\} \longrightarrow \operatorname{int}(D^n_+) \setminus \{0\}, \quad \alpha(v) = \big(1 - \|v\|\big)\frac{v}{\|v\|}.$$

Let M_1 and M_2 be two oriented n-manifolds with boundary. Pick two embedded half-discs

$$f_1 \colon D^n_+ \hookrightarrow M_1, \qquad f_2 \colon D^n_+ \hookrightarrow M_2$$

such that $f_i^{-1}(\partial M_i) = D^{n-1}$ as in Figure 6.8-(left). We require f_1 to be orientation-preserving and f_2 orientation-reversing. Then we glue the manifolds $M_1 \setminus f_1(0)$ and $M_2 \setminus f_2(0)$ via the diffeomorphism

$$f_2 \circ \alpha \circ f_1^{-1} \colon f_1\big(\operatorname{int}(D^n_+) \setminus \{0\}\big) \longrightarrow f_2\big(\operatorname{int}(D^n_+) \setminus \{0\}\big).$$

The resulting oriented smooth manifold with boundary is the *∂-connected sum* of M_1 and M_2 and is denoted as

$$M_1 \#_\partial M_2.$$

See Figure 6.8. As above one proves that the resulting manifold depends only on the connected components of ∂M_1 and ∂M_2 intersecting the half-discs. In particular, if both M_1 and M_2 have connected boundary, then $M_1 \#_\partial M_2$ is uniquely determined.

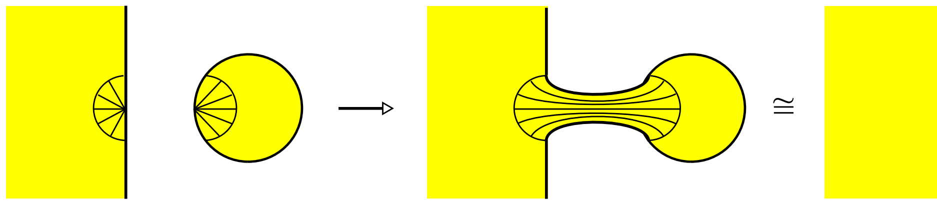

Figure 6.9. The ∂-connected sum with a disc does not change the manifold up to diffeomorphism.

Proposition 6.3.4. *If ∂M_1 and ∂M_2 are connected, we have*

$$\partial(M_1 \#_\partial M_2) \cong \partial M_1 \# \partial M_2.$$

In general we have $M \#_\partial D^n \cong M$.

Proof. The manipulation restricted to the boundaries is a connected sum, so the first isomorphism holds. The second is sketched in Figure 6.9, and we leave the tedious exercise of writing the correct diffeomorphism to the courageous reader. $\square$

6.3.4. Pasting manifolds along submanifolds. We now introduce a generalisation of the connected sum, in which we glue manifolds along disc bundles instead of just discs.

Pick $0 \leq k < n$. Let M_1 and M_2 be two n-manifolds, and let $N_1 \subset \mathrm{int}(M_1)$ and $N_2 \subset \mathrm{int}(M_2)$ be two diffeomorphic compact k-submanifolds without boundary, with closed tubular neighbourhoods $\nu N_1 \subset \mathrm{int}(M_1)$ and $\nu N_2 \subset \mathrm{int}(M_2)$. We suppose that the two tubular neighbourhoods are also isomorphic, and we fix a disc bundles isomorphism

$$\varphi \colon \nu N_1 \longrightarrow \nu N_2.$$

As above, we define the self-diffeomorphism

$$\alpha \colon \mathrm{int}(\nu N_1) \setminus N_1 \longrightarrow \mathrm{int}(\nu N_1) \setminus N_1, \qquad \alpha(v) = \big(1 - \|v\|\big)\frac{v}{\|v\|}.$$

We now glue the manifolds $M_1 \setminus N_1$ and $M_2 \setminus N_2$ via the diffeomorphism

$$\varphi \circ \alpha \colon \mathrm{int}(\nu N_1) \setminus N_1 \longrightarrow \mathrm{int}(\nu N_2) \setminus N_2.$$

The resulting manifold M is obtained by *pasting* M_1 and M_2 along the submanifolds N_1 and N_2. It is an operation that can be done as soon as the submanifolds N_1 and N_2 have isomorphic tubular neighbourhoods; note however that, as opposite to connected sum, the choice of the isomorphism φ is important here, because two different isomorphisms may not be isotopic in many interesting cases, not even if they are co-oriented.

Remark 6.3.5. As in Remark 6.3.1, the construction of M may be described alternatively as a two-steps cut-and-paste operation, where:

(1) first, we remove from M_i the open submanifold $\mathrm{int}(\nu N_i)$, thus creating a new boundary component $\partial \nu N_i$;
(2) then, we paste the two new boundary components via φ.

6.3.5. Surgery. There is a particular type of pasting that is so important to deserve a separate name.

Let M be a n-manifold, and $S \subset M$ be a k-sphere (that is, a submanifold diffeomorphic to S^k) with trivial normal bundle, for some $0 \leq k \leq n-1$. As in Section 6.1.14, we see S^n inside $\mathbb{R}^{k+1} \times \mathbb{R}^{n-k}$ and consider $S^k = S^n \cap \{y = 0\}$. We have seen that the normal bundle $\nu S^k \subset S^n$ is also trivial.

We can therefore paste M and S^n along the k-spheres S and S^k. To do so, we must choose a disc bundle isomorphism $\varphi\colon \nu S \to \nu S^k$. This operation is called a *surgery* along the sphere S. The resulting manifold M' depends on the chosen φ.

Remark 6.3.6. We have seen in Section 6.1.14 that S^n decomposes into $S^k \times D^{n-k}$ and $D^{k+1} \times S^{n-k-1}$. Therefore, by Remark 6.3.5, a surgery may also be described as follows: whenever we find a domain in M diffeomorphic to $S^k \times D^{n-k}$, we first remove its interior, thus creating a new boundary $S^k \times S^{n-k-1}$, and then glue $D^{k+1} \times S^{n-k-1}$ to it via the identity map. Shortly: we substitute a $S^k \times D^{n-k}$ inside M with $D^{k+1} \times S^{n-k-1}$.

Remark 6.3.7. A surgery along a 0-sphere is like a connected sum: we replace $S^0 \times D^n$, that is two disjoint discs, with $D^1 \times S^{n-1}$, that is a tube. When both points in S^0 are contained in the same connected component, this may be interpreted as a *self*-connected sum of that component.

The inverse operation of a surgery along a k-sphere is naturally a surgery along a $(n-k-1)$-sphere.

6.3.6. Pasting along submanifolds in the boundary. There is of course a boundary version of pasting along submanifolds, where the submanifolds lie in the boundary. This operation generalises the ∂-connected sum and will be fundamental in the next section.

Let M_1 and M_2 be two n-manifolds with boundary, and let $N_1 \subset \partial M_1$ and $N_2 \subset \partial M_2$ be two compact k-submanifolds of the boundary. We require that N_1 and N_2 have no boundary, and that they have isomorphic closed tubular neighbourhoods $\varphi\colon \nu N_1 \to \nu N_2$ in ∂M_1 and ∂M_2.

We now define a new manifold M' obtained by *pasting* M_1 and M_2 along the submanifolds N_1 and N_2. The operation is the same as above, only with half-discs instead of disc bundles.

Each $\nu N_i \subset M_i$ is a D^{n-k-1}-bundle over N_i, and using collars we may extend it to a half-disc D^{n-k}_+-bundle $\bar\nu N_i$ that is a "half"-tubular neighbourhood of N_i in M_i. The diffeomorphism φ also extends to $\varphi\colon \bar\nu N_1 \to \bar\nu N_2$. We glue the manifolds $M_1 \setminus N_1$ and $M_2 \setminus N_2$ via the diffeomorphism

$$\varphi \circ \alpha\colon \operatorname{int}(\bar{\nu} N_1) \setminus N_1 \longrightarrow \operatorname{int}(\bar{\nu} N_2) \setminus N_2$$

where α and $\operatorname{int}(\nu N_i)$ are defined on every fibre D^{n-k}_+ as we did for ∂-connected sums.

The ∂-connected sum corresponds to the case where N_1 and N_2 are points.

6.4. Handle decompositions

We now show that every compact manifold M decomposes into finitely many simple blocks, called *handles*. This important procedure is called a *handle decomposition*.

6.4.1. Handles. We have described in the previous section the operation of pasting two manifolds along submanifolds in their boundaries. We now introduce a particular, but very important, case.

Let M be a n-manifold with boundary. Let $S \subset \partial M$ be a $(k-1)$-sphere with trivial normal boundary $\nu S \subset \partial M$, with $0 < k < n$. A *k-handle addition* on M is the operation that consists of pasting M with D^n, along the $(k-1)$-spheres S and $S^{k-1} \subset S^{n-1}$. As in Section 6.1.14, we see D^n inside $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$ with coordinates (x, y), and $S^{k-1} = S^{n-1} \cap \{y = 0\}$.

The result is a new smooth manifold M'. As in Section 6.1.14, we may identify νS^{k-1} with $S^{k-1} \times D^{n-k}$, so that M' depends on the diffeomorphism

$$\varphi\colon S^{k-1} \times D^{n-k} \longrightarrow \nu S.$$

We also define the extremal cases $k = 0$ and $k = n$. The addition of a 0-handle to M is simply the addition of a disjoint connected component D^n, with no attachment. On the contrary, on a n-handle the n-sphere S is a connected component of ∂M, and we attach D^n along a diffeomorphism $\varphi\colon S^{n-1} \to S$.

6.4.2. Alternative description. To better visualise what is going on, we furnish an alternative description of a k-handle addition, drawn in Figure 6.10.

Let $S \subset \partial M$ be a $(k-1)$-sphere with trivial normal boundary. It has a half-tubular neighbourhood in M is diffeomorphic to $S \times \mathbb{R}^{n-k} \times \mathbb{R}_+$ and we identify it with the manifold with boundary

$$U = \big\{(x, y) \in \mathbb{R}^{n-k} \times \mathbb{R}^k \;\big|\; \|y\| \geq 1\big\}$$

via the map $(u, v, t) \mapsto (v, (t+1)u)$. Now we have

$$S = \{x = 0, \|y\| = 1\}, \qquad \partial U = U \cap \partial M = \big\{\|y\| = 1\big\}.$$

Let $\rho\colon [-1, 1] \to \mathbb{R}_+$ be a continuous positive function that is smooth on $(-1, 1)$ and such that all derivatives of ρ tend to $\pm\infty$ as $t \to \pm 1$ (corresponding signs). We define a bigger manifold M' by substituting U with the bigger set

$$U' = U \cup \big\{\|y\| < 1, \|x\| < \rho(\|y\|)\big\}.$$

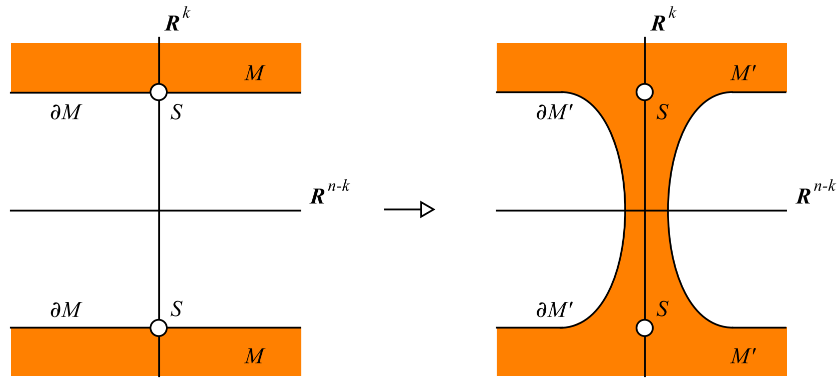

Figure 6.10. An alternative description of the attachment of a k-handle to M.

Exercise 6.4.1. The manifold M' is diffeomorphic to M with a k-handle attached to S.

See Figure 6.10. Note that with this description the original manifold M is naturally a submanifold of M'.

6.4.3. Topological handles. We can make one further step towards visualization and intuition by using *topological handles*. These capture the topological structure of M' while being a little bit imprecise on its smooth structure. See Figure 6.11.

A topological handle is what we get if we take $\rho(t) = 1$ constantly in the previous construction. The result is not smooth, but it still works up to homeomorphisms.

In other words, we use $D^k \times D^{n-k}$ instead of D^n. This is not a smooth manifold because of its corners; its topological boundary decomposes into the horizontal $D^k \times S^{n-k-1}$ and the vertical $S^{k-1} \times D^{n-k}$. For every embedding

$$\varphi \colon S^{k-1} \times D^{n-k} \hookrightarrow \partial M$$

we define a new topological space

$$M' = M \cup_\varphi \left(D^k \times D^{n-k}\right)$$

obtained by attaching $D^k \times D^{n-k}$ to M along φ. This operation is the attachment of a *topological k-handle* to M. The attaching of a handle or a topological handle along the same map φ produce homeomorphic manifolds M': the only difference between the two constructions is that in the topological setting the smooth structure on M' is not obvious to see – some new corners arise that should be smoothened, see Figure 6.11. From now on, we will always think as a handle as a topological handle whose corners have been smoothened.

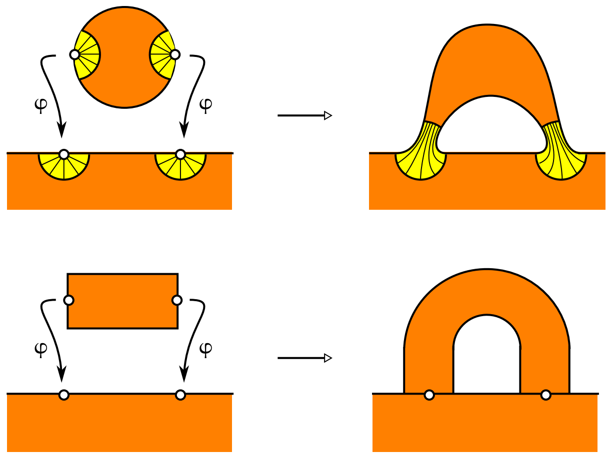

Figure 6.11. The attachment of a 1-handle and of a topological 1-handle along the same map φ. The resulting topological manifold is the same in both constructions, but the smooth structure is well-defined only with the first. For practical purposes, we usually think of a handle as a topological handle whose corners have been somehow "smothened."

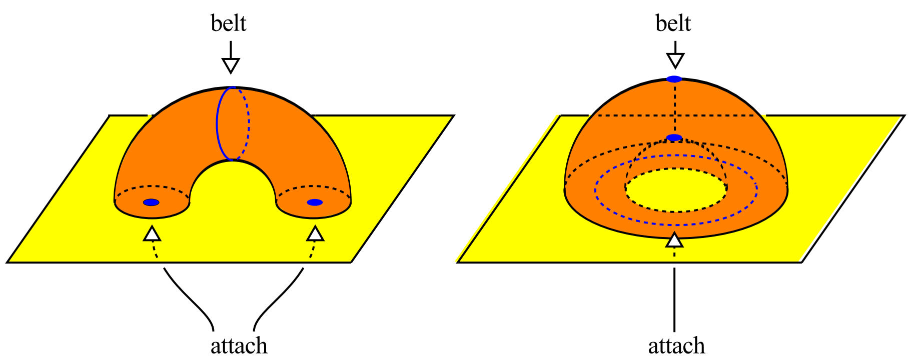

Figure 6.12. A three-dimensional topological 1-handle (left) and 2-handle (right), with the attaching and belt spheres in blue.

One should think of a topological k-handle $D^k \times D^{n-k}$ as a thickened k-dimensional disc. Here is some useful terminology: the number k is the *index* of the handle; the sphere $S^{k-1} \times \{0\}$ is the *attaching sphere*, while the sphere $\{0\} \times S^{n-k-1}$ is the *belt sphere*. The discs $D^k \times \{0\}$ and $\{0\} \times D^{n-k}$ are the *attaching* and *belt discs*. See some examples in Figure 6.12.

Remark 6.4.2. If M' is obtained from M by the attachment of a k-handle to the $(k-1)$-sphere $S \subset M$, the new boundary $\partial M'$ is obtained from the old ∂M by surgery along the sphere S. This follows readily from the definition.

6.4.4. Handle decomposition. Let M be a compact smooth n-manifold, possibly with boundary. A *handle decomposition* for M is the realisation of M

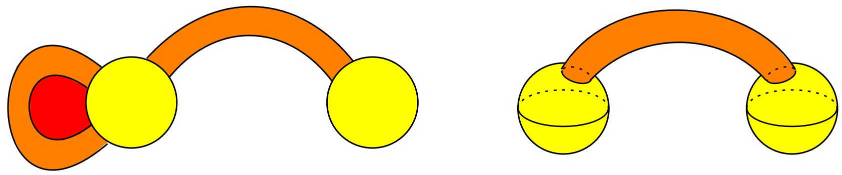

Figure 6.13. Some handle decompositions in dimension two and three. On the left, we have two 0-handles (yellow), one 1-handle (orange), and one 2-handle (red) in dimension two. On the right, we have two 0-handles (yellow) and one 1-handle (orange) in dimension three.

as the result of a finite number of operations

$$\varnothing = M_0 \rightsquigarrow M_1 \rightsquigarrow \cdots \rightsquigarrow M_k = M$$

where each M_{i+1} is obtained by attaching some handle to M_i. Since the only handle that can be attached to the empty set is a 0-handle, the manifold M_1 is the result of a 0-handle attachment to $\varnothing$ and is hence a n-disc.

Example 6.4.3. The sphere S^n, and more generally each of the exotic spheres described in Section 6.2.6, decomposes into two n-discs. We may interpret this decomposition as a n-handle attached to a 0-handle. Therefore S^n has a handle decomposition with one 0-handle and one n-handle.

Conversely, if a compact manifold M without boundary decomposes into two handles only, then these must be a 0- and a n-handle, and so M is either S^n or an exotic sphere (in all cases, it is homeomorphic to S^n).

6.4.5. Reordering handles. More examples are shown in Figure 6.13. In both examples in the figure the handle decomposition goes as follows: we first attach some 0-handles (that is, we create discs out of nothing), then we attach some 1-handles, then we attach some 2-handles. We think at the 1-handles in the (left) figure as attached simultaneously. We now show that every handle decomposition can be modified to be of this type.

Proposition 6.4.4. *Every handle decomposition can be modified so that we first attach all 0-handles, then all 1-handles, then all 2-handles ... and so on.*

Proof. Suppose that M_{i+1} is obtained from M_i by attaching a k-handle H^k, and M_{i+2} is obtained from M_{i+1} by attaching a h-handle H^h. We write

$$M_{i+1} = M_i \cup_\varphi H^k, \qquad M_{i+2} = M_{i+1} \cup_\psi H^h.$$

We show below that if $h \leq k$ then H^h can be slid away from H^k as in Figure 6.14. After this move, the handles H^h and H^k are disjoint and hence we can obtain the same manifold M_{i+2} by first attaching H^h and then H^k.

By applying finitely many exchanges of this type we transform every handle decomposition into one where handles are attached with non-decreasing index.

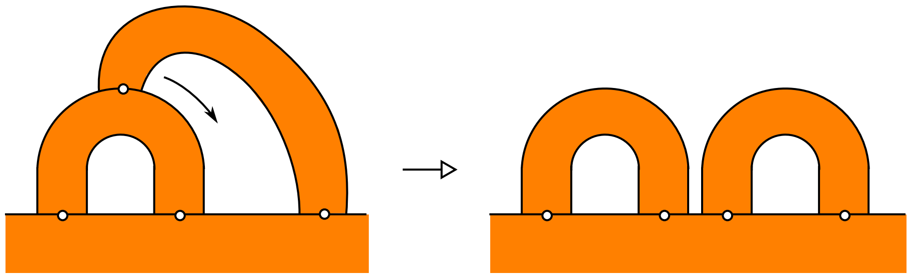

Figure 6.14. If $h \leq k$, we can always slide a k-handle away from a previously attached h-handle. Here $h = k = 1$.

Moreover, the handles with the same index can be slid to be disjoint, and hence can be thought to be attached simultaneously. This proves the proposition.

We now show how to slide H^h aways from H^k. The attaching sphere of H^h is a $(h-1)$-sphere $S \subset \partial M_{i+1}$, while the belt sphere of H^k is a $(n-k-1)$-sphere $S' \subset \partial M_{i+1}$. If $h \leq k$, we have $(h-1)+(n-k-1) < n-1$. By transversality, we may isotope S away from S'.

The handles H^k and H^h intersect ∂M into two closed tubular neighbourhoods of S' and S. Since $S' \cap S = \varnothing$, we can isotope the tubular neighbourhood of S to be disjoint from that of S'. That is, we can slide the handle H^h away from H^k, as stated. $\square$

As stressed in the proof, the handles of the same index are disjoint and can be attached simultaneously, as in Figure 6.13.

Our next goal is to show that every compact smooth manifold decomposes into handles. To this purpose we study the critical points of functions $M \to \mathbb{R}$ and we introduce the *Morse functions*, that are of independent interest.

6.4.6. Hessian at the critical points. Let M be a manifold without boundary and $f : M \to \mathbb{R}$ be a smooth function. We know that its differential df is a section of the cotangent bundle, that is a tensor field of type $(0,1)$ on M. On a chart, the differential is just the gradient.

Can we define a kind of second derivative of f, that behaves like the Hessian when read on a chart? For instance, this might be some tensor field of type $(0,2)$? The answer is unfortunately negative in general: there is no way to define a Hessian unambiguously; to get a Hessian we need to equip M with some additional structure, like the *connections* introduced in Chapter 9.

Despite these premises, a Hessian is however defined at the critical points of f. If p is a critical point, then we can define a symmetric bilinear form

$$\mathrm{Hess}\,(f)_p \colon T_pM \times T_pM \longrightarrow \mathbb{R}$$

as follows. Given $v, w \in T_pM$, extend them to two arbitrary vector fields X, Y in some neighbourhood of p. Then we set

$$\mathrm{Hess}\,(f)_p(v, w) = X\big(Y(f)\big)(p).$$

Exercise 6.4.5. The map $\mathrm{Hess}\,(f)_p$ is well-defined, bilinear, and symmetric.

It is crucial here that $df_p = 0$. Alternatively, we can also define the Hessian in coordinates: we pick $p = 0$ for simplicity and get

$$f(x) = f(0) + \tfrac{1}{2}\,{}^{\mathrm{t}}xHx + o\big(\|x\|^2\big).$$

On some other chart with variables $\bar{x}$, we get $x = J\bar{x} + o\big(\|\bar{x}\|\big)$ where J is the differential of the coordinates change at $x = 0$ and therefore

$$\begin{aligned} f(x) &= f(0) + \tfrac{1}{2}\,{}^{\mathrm{t}}\big(J\bar{x} + o\big(\|\bar{x}\|\big)\big)H\big(J\bar{x} + o\big(\|\bar{x}\|\big)\big) + o\big(\|x\|^2\big) \\ &= f(0) + \tfrac{1}{2}\,{}^{\mathrm{t}}\bar{x}\,{}^{\mathrm{t}}JHJ\bar{x} + o\big(\|\bar{x}\|^2\big). \end{aligned}$$

Therefore H changes to ${}^{\mathrm{t}}JHJ$ and hence describes a chart-independent bilinear form on T_pM. Of course the two definitions given coincide (exercise).

6.4.7. Non-degenerate critical points. Let M be a manifold without boundary and $f\colon M \to \mathbb{R}$ a smooth function. We say that a critical point $p \in M$ for f is *non-degenerate* if the bilinear form $\mathrm{Hess}\,(f)_p$ on T_pM is non-degenerate. We now study the non-degenerate critical points. We start by exhibiting an alternative definition.

Proposition 6.4.6. *A critical point p is non-degenerate $\Longleftrightarrow$ the section df of T^*M is transverse to the zero-section at p.*

Proof. On a chart, we have $f\colon U \to \mathbb{R}$ for some open set $U \subset \mathbb{R}^n$. We see df as the gradient $\nabla f\colon U \to \mathbb{R}^n$. Now ∇f is transverse to the zero-section at $p \in U \Longleftrightarrow$ the differential of ∇f is invertible in p. The differential of ∇f is $\mathrm{Hess}\,(f)_p$, so we are done. □

Corollary 6.4.7. *Non-degenerate critical points are isolated.*

If p is a non-degenerate critical point, then $\mathrm{Hess}\,(f)_p$ is a scalar product on T_pM and has some signature $(k, n-k)$ for some $0 \leq k \leq n$. The integer $n-k$ is the *index* of the critical point p. The Morse Lemma determines the behaviour of f near p, according to its index.

Lemma 6.4.8 (Morse Lemma). *Let p be a non-degenerate critical point of index $n-k$. On some appropriate chart near p the function f is read as*

$$f(x) = f(p) + x_1^2 + \ldots + x_k^2 - x_{k+1}^2 - \ldots - x_n^2.$$

Proof. On a chart we get $f\colon \mathbb{R}^n \to \mathbb{R}$ with $p = 0$. Since 0 is a critical point, Taylor's Theorem 1.3.1 gives

$$f(x) = f(0) + \tfrac{1}{2}\sum_{i,j=1}^{n} h_{ij}(x)x_ix_j$$

for some smooth maps h_{ij}. After substituting both h_{ij} and h_{ji} with $\frac{1}{2}(h_{ij}+h_{ji})$ we get $h_{ij}=h_{ji}$. Now $h_{ij}(0)$ is non-degenerate with signature $(k, n-k)$.

To transform f into the desired form, we follow the usual procedure to diagonalise scalar products, and extend it smoothly on a neighbourhood of 0. We proceed by induction: suppose that on some coordinates we write

$$f(x) = \pm x_1^2 \pm \cdots \pm x_{r-1}^2 + \sum_{i,j\geq r} h_{ij}(x)x_ix_j.$$

Since $h_{ij}(0)$ has maximal rank, after a linear change of coordinates we may suppose that $h_{rr}(x) \neq 0$ at $x=0$ and hence on some small neighbourhood around 0. We pick new coordinates

$$\begin{cases} y_i &= x_i \quad \text{for } i \neq r, \\ y_r &= \sqrt{|h_{rr}(x)|}\left(x_r + \sum_{i>r} \frac{h_{ir}(x)x_i}{h_{rr}(x)}\right). \end{cases}$$

With these new coordinates we easily get

$$f(y) = \pm y_1^2 \pm \cdots \pm y_r^2 + \sum_{i,j>r} h'_{ij}(y)y_iy_j$$

for some functions h'_{ij} defined near p, and we conclude by induction. $\square$

6.4.8. Morse functions. Let M be a manifold without boundary. A *Morse function* on M is a function $f\colon M \to \mathbb{R}$ whose critical points are all non-degenerate. That is, the differential df is transverse to the zero-section.

We now prove that there are plenty of Morse functions. Via the Whitney embedding theorem, we may suppose that $M \subset \mathbb{R}^m$ for some m.

Proposition 6.4.9. *Let $M \subset \mathbb{R}^m$ be a submanifold and $f\colon M \to \mathbb{R}$ any smooth function. For almost every $a \in \mathbb{R}^m$, the modified function*

$$f_a\colon M \longrightarrow \mathbb{R}, \qquad f_a(x) = f(x) - \langle a, x\rangle$$

is a Morse function.

Proof. For $a \in \mathbb{R}^n$, we define the maps

$$L_a\colon M \longrightarrow \mathbb{R}, \qquad L_a(x) = \langle a, x\rangle.$$

$$\pi\colon M \times \mathbb{R}^m \longrightarrow T^*M, \qquad (p,a) \longmapsto \big(p, (dL_a)_p\big).$$

Of course $(dL_a)_p(v) = \langle a, v\rangle$. The map π is a fibre bundle (exercise). For every $a \in \mathbb{R}^n$ we get a commutative diagram

$$\begin{tikzcd} E \arrow[r, dashed, "g"] \arrow[d, dashed] & M \times \mathbb{R}^n \arrow[d, "\pi"] & M \arrow[l, "h_a"'] \arrow[dl, "dL_a"] \\ M \arrow[r, "df"'] & T^*M & \end{tikzcd}$$

where $E \to M$ is the pull-back of π along df, and $h_a(p) = (p,a)$.

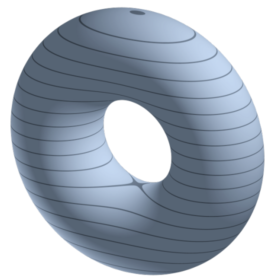

Figure 6.15. On this torus, the height function $f(x, y, z) = z$ is a Morse function with four non-degenerate critical points of index 0, 1, 1, and 2. The level sets $f^{-1}(t)$ are manifolds, except when t is a critical value.

By Proposition 5.7.2, the maps g and h_a are transverse $\iff$ a is a regular value for $\pi_2 \circ g$ where $\pi_2(p, a) = a$. By Sard's Lemma, this holds for almost every a and hence $g \pitchfork h_a$. As in the end of the proof of Lemma 5.7.10, this implies that $df \pitchfork dL_a$. Therefore $d(f - L_a)$ is transverse to the zero-section, that is $f_a = f - L_a$ is a Morse function. $\square$

Corollary 6.4.10. *Let $f: M \to \mathbb{R}$ a smooth function. For every $\varepsilon > 0$ there is a Morse function $g: M \to \mathbb{R}$ with $|f(p) - g(p)| < \varepsilon$ for all $p \in M$.*

Proof. Embed M in a small ball of $\mathbb{R}^n$ and apply Proposition 6.4.9. $\square$

We have proved in particular that every M has some Morse function $f: M \to \mathbb{R}$. It is sometimes useful to add the following requirement.

Proposition 6.4.11. *Every manifold M without boundary has a Morse function $f: M \to \mathbb{R}$ where distinct critical points have distinct images.*

Proof. Pick a Morse function $f: M \to \mathbb{R}$. At a critical point $p \in M$, choose a bump function $\rho: M \to \mathbb{R}$ that is constantly $c > 0$ on a small neighbourhood of p and is 0 outside a slightly bigger neighbourhood, disjoint from all the other critical points. Modify f to $f + \rho$. If c is small and $d\rho$ is uniformly small, the function $f + \rho$ is still Morse with the same critical points. However, the value of p has changed by c. By choosing appropriate c we can separate the images of all the critical points. $\square$

6.4.9. Existence of handle decompositions. We have introduced Morse functions as a fundamental tool to prove the following remarkable theorem.

Theorem 6.4.12. *Every compact manifold M without boundary has a handle decomposition.*

Proof. Let $f: M \to \mathbb{R}$ be a Morse function, where critical points have distinct images. Since M is compact, it has finitely many critical points. For

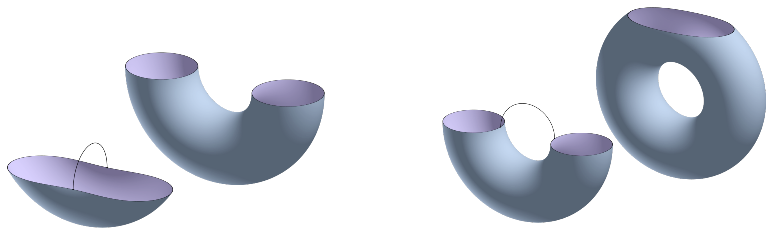

Figure 6.16. Each time a non-degenerate critical point of index k is crossed, a k-handle is added. We show here the two critical points of index 1, and the core segment of the 1-handle in each case.

instance, Figure 6.15 shows a Morse function on the torus with four critical points. For every $a \in \mathbb{R}$ we define

$$M_a = f^{-1}(-\infty, a].$$

When a is regular, M_a is a domain in M, that is a submanifold with boundary. Consider two regular values $a < b$. We now prove two facts:

(1) If $[a, b]$ contains no critical values, then M_a and M_b are diffeomorphic.
(2) If $[a, b]$ contains a single critical value, image of a critical point of index k, then M_b is diffeomorphic to M_a with a k-handle attached.

An example is shown in Figure 6.16. When a crosses a critical point of index k, a k-handle is attached to M_a. So the torus decomposes into one 0-handle, two 1-handles, and one 2-handle. The claims (1) and (2) clearly imply that M decomposes into handles, one for each critical point of M.

We first prove (1). Fix an arbitrary Riemannian metric on M, that is on the tangent bundle TM. Every T_pM is equipped with a scalar product $\langle , \rangle$, and we use it to transform the covector field df into a vector field ∇f in the usual way, by requiring that

$$df_p(v) = \big\langle \nabla f(p), v \big\rangle.$$

The field ∇f vanishes at the critical points. On a curve $\gamma \colon I \to M$ we get

$$(f \circ \gamma)'(t) = df_{\gamma(t)}\big(\gamma'(t)\big) = \big\langle \nabla f, \gamma'(t) \big\rangle.$$

Let $\rho \colon M \to \mathbb{R}$ be a smooth function that equals $1/\langle \nabla f, \nabla f \rangle$ on the compact set $f^{-1}[a, b]$ and which vanishes outside some bigger compact subset. We define a new vector field

$$X(p) = \rho(p) \nabla f(p).$$

Since M is compact, the vector field X is complete and generates a flow Φ. Consider an integral curve $\gamma(t) = \Phi(p, t)$. If $\gamma(t) \in f^{-1}[a, b]$ then

$$(f \circ \gamma)'(t) = \big\langle \nabla f, \gamma'(t) \big\rangle = \big\langle \nabla f, X \big\rangle = 1.$$

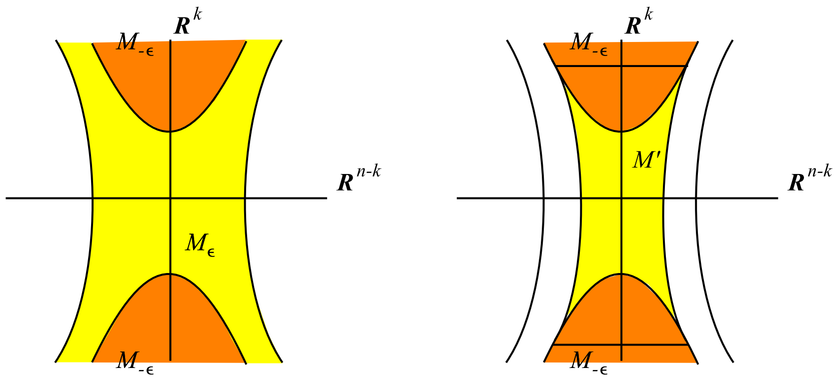

Figure 6.17. The manifolds M_ε and $M_{-\varepsilon}$ intersect the chart $\mathbb{R}^{n-k} \times \mathbb{R}^k$ as shown here (left). We replace M_ε with a diffeomorphic submanifold M', still containing $M_{-\varepsilon}$, so that the yellow zone $M' \setminus M_{-\varepsilon}$ lies entirely in this chart (right).

Therefore the flow defines a diffeomorphism

$$M_a \longrightarrow M_b, \qquad p \longmapsto \Phi(p, b-a).$$

We turn to (2). Let $p \in M$ be the unique critical point in $f^{-1}[a,b]$. We suppose for simplicity that $f(p) = 0$. By (1) we may choose $a = -\varepsilon$ and $b = \varepsilon$ for some small $\varepsilon > 0$. By the Morse Lemma, on a chart $U \cong \mathbb{R}^n = \mathbb{R}^{n-k} \times \mathbb{R}^k$ the function f is

$$f(x) = \|x\|^2 - \|y\|^2$$

where $(x, y) \in \mathbb{R}^{n-k} \times \mathbb{R}^k$ and $p = (0,0)$. The manifolds M_ε and $M_{-\varepsilon}$ intersect the chart $\mathbb{R}^{n-k} \times \mathbb{R}^k$ as in Figure 6.17-(left).

We now substitute M_ε with a diffeomorphic submanifold M' that still contains $M_{-\varepsilon}$, and which has the additional property that $M' \setminus M_{-\varepsilon}$ lies entirely in the chart $\mathbb{R}^{n-k} \times \mathbb{R}^k$ as shown in Figure 6.17-(right). To this purpose, we pick a function $\phi\colon \mathbb{R} \to \mathbb{R}$ such that

$$\phi(0) > \varepsilon, \quad \phi(t) = 0 \ \forall t \geq 2\varepsilon, \quad -1 < \phi'(t) \leq 0 \ \forall t.$$

We now define another function $F\colon M \to \mathbb{R}$, by requiring that $F(p) = f(p)$ outside the chart, and

$$F(x,y) = f(x,y) - \phi\big(2\|x\|^2 + \|y\|^2\big)$$

inside the chart. We then define

$$M' = F^{-1}(-\infty, -\varepsilon].$$

Clearly $M' \supset M_{-\varepsilon}$ and $M' \setminus M_{-\varepsilon}$ is contained in the chart. We show that

$$M_\varepsilon = F^{-1}(-\infty, \varepsilon].$$

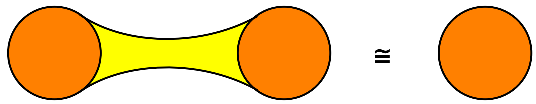

Figure 6.18. A 1-handle attached to two distinct 0-handles: the result is diffeomorphic to a disc.

Indeed, we obviously have $M_\varepsilon \subset F^{-1}(-\infty, \varepsilon]$, and conversely if $F(x, y) \leq \varepsilon$ and $\phi(2\|x\|^2 + \|y\|^2) > 0$ we get $2\|x\|^2 + \|y\|^2 < 2\varepsilon$; therefore

$$f(x, y) = \|x\|^2 - \|y\|^2 \leq \|x\|^2 + \tfrac{1}{2}\|y\|^2 < \varepsilon.$$

We verify easily that $dF = 0 \iff df = 0$, hence F has the same critical points as f. Since $F(p) < -\varepsilon$, the function F has no critical values in $[-\varepsilon, \varepsilon]$ and (1) implies that M' and M_ε are diffeomorphic.

Finally, we need to show that M' is diffeomorphic to $M_{-\varepsilon}$ with a (yellow) k-handle attached, as suggested by Figure 6.17-(right). To this purpose we fix $y_0 \in \mathbb{R}^k$ and study the horizontal slice

$$M' \cap \big\{y = y_0\big\} = \Big\{(x, y) \ \Big|\ y = y_0,\ \|x\|^2 \leq \|y_0\|^2 + \phi\big(2\|x\|^2 + \|y_0\|^2\big) - \varepsilon\Big\}.$$

This is easily seen to be a disc with radius $r(y_0) > 0$ depending smoothly on y_0. When $\|y_0\|^2 > 2\varepsilon$ we get $r(y_0) = \sqrt{\|y_0\|^2 - \varepsilon}$.

One concludes by showing that Figure 6.17-(right) is in fact diffeomorphic to Figure 6.10-(right). Therefore M' is $M_{-\varepsilon}$ with a k-kandle attached. The explicit diffeomorphism is left as an exercise. $\square$

6.5. Classification of surfaces

In the previous section we have shown a powerful construction that allows to decompose every compact smooth manifold without boundary into simple pieces called handles. We now use this construction to classify all compact surfaces.

6.5.1. The main theorem. We start by solving the most interesting case. Recall from Section 6.3.2 that we defined the genus-g surface S_g as the connected sum of g tori.

Theorem 6.5.1. *Every compact connected and orientable surface S without boundary is diffeomorphic to S_g, for some $g \geq 0$.*

Proof. We pick a handle decomposition of S. This consists of some 0-handles, then 1-handles attached to these 0-handles, and finally 2-handles attached to the result.

We first make an observation that is valid in all dimensions: if we attach a 1-handle to two distinct 0-handles as in Figure 6.18, this is equivalent to making a boundary connected sum of two discs, so the result is again a disc.

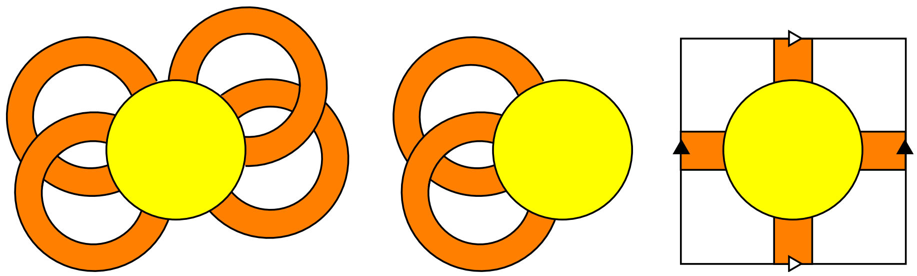

Figure 6.19. The 0-handle and some 1-handles (left). Two interlaced 1-handles (centre). Two interlaced handles form a handle decomposition of a holed torus, seen here as a square with opposite edges identified, with the white hole removed (right).

Therefore we can replace the two 0-handles and the 1-handle altogether with a singe 0-handle, thus simplifying the handle decomposition.

After finitely many such moves, we may suppose that in the handle decomposition of S every 1-handle is attached twice to the same 0-handle. Since S is connected, this easily implies that there is only one 0-handle.

A dual argument works for the 2-handles. Note that every 1-handle is incident to two 2-handles, attached to the two long sides of the 1-handle. If the 2-handles are distinct, then the 1-handle together with the two incident 2-handles form again a picture like in Figure 6.18, and can thus be replaced by a single disc, that is a single 2-handle. After finitely many moves of this type, we easily end with a single 2-handle.

We have simplified the handle decomposition of S so that it has only one 0- and one 2-handle. If there are no 1-handles, then S decomposes into a 0- and a 2-handle and is hence diffeomorphic to S^2 by Corollary 6.2.13.

Suppose that there are 1-handles. Every 1-handle is a topological rectangle attached to the 0-handle along its short sides, as in Figure 6.19-(left). Up to diffeomorphism, there are two ways of attaching a 1-handle: with or without a twist. However, twists produce Möbius strips, which are excluded since S is orientable. So every 1-handle is attached without a twist, as in the figure.

Since there is only one 2-handle, the union of the 0- and 1-handles is a surface with connected boundary. This implies that every 1-handle must be interlaced with some other 1-handle as in Figure 6.19-(centre). Let $S' \subset S$ be the subsurface consisting of the 0-handle and these two 1-handles. Figure 6.19-(right) shows that S' is diffeomorphic to a torus with a hole. Therefore if we substitute S' with a single 0-handle, that is a disc, we find a simpler handle decomposition of a new surface S'' such that

$$S = S'' \# T.$$

We conclude by induction on the number of 1-handles that S is a connected sum of some g tori. $\square$

In the next chapters we will prove that S_g is not diffeomorphic to $S_{g'}$ if $g \neq g'$, so the genus of a surface fully characterises the surface up to diffeomorphism.

CHAPTER 7

Differential forms

In a smooth manifold there is no notion of distance between points, angle between intersecting curves, volume of domains, etc. To get all these natural geometric concepts, we need to equip the manifold with an additional structure: as we will see in the next chapters, it suffices to choose a *metric tensor* to recover them all. Here we study a somehow weaker, and quite different, structure called *differential form*.

A differential form may be used to talk about volumes, but not yet about distances or angles. This apparently weaker structure has however some important applications that go beyond volumes and integration: it may be manipulated quite easily – for instance, it can be pulled back via any smooth maps, whereas metric tensors cannot – and can also be "differentiated" in a very natural way. This differentiation will lead in the next chapter to a rich and beautiful algebraic theory called *De Rham cohomology*.

7.1. Differential forms

We introduce the differential k-forms.

7.1.1. Definition. Let M be a smooth n-manifold. A *differential k-form* (shortly, a *k-form*) is a section ω of the alternating bundle

$$\Lambda^k(M)$$

over M, see Section 4.3.4. In other words, for every $p \in M$ we have an antisymmetric multilinear form

$$\omega(p)\colon \underbrace{T_pM \times \cdots \times T_pM}_{k} \longrightarrow \mathbb{R}$$

that varies smoothly with $p \in M$.

Example 7.1.1. A 1-form is a section of $\Lambda^1(M) = T^*M$, that is a field of covectors. As an important example, the differential df of a smooth function $f\colon M \to \mathbb{R}$ is a 1-form, see Section 4.3.2. This example is not exhaustive: we will see that some 1-forms are not the differential of any function.

By Corollary 2.4.10, every k-form with $k > n$ is necessarily trivial. The vector space of all the k-forms on M is denoted by

$$\Omega^k(M) = \Gamma(\Lambda^k M).$$

7.1.2. Exterior product. Recall from Section 2.4.3 that the exterior algebra $\Lambda^*(V)$ of a real vector space V is equipped with the exterior product $\wedge$. Let now ω and η be a k-form and a h-form on a manifold M. Their *exterior product* is the $(k+h)$-form $\omega \wedge \eta$ defined pointwise by setting

$$(\omega \wedge \eta)(p) = \omega(p) \wedge \eta(p).$$

As in Section 2.4.3, the space

$$\Omega^*(M) = \bigoplus_{k \geq 0} \Omega^k(M)$$

is an anticommutative associative algebra, that is

$$\omega \wedge \eta = (-1)^{hk} \eta \wedge \omega$$

and if k is odd we get

$$\omega \wedge \omega = 0.$$

This holds in particular for every 1-form ω.

7.1.3. In coordinates. As usual, differential forms may be written quite conveniently in coordinates.

Let $U \subset \mathbb{R}^n$ be an open set. Recall that for some notational reasons it is preferable to denote the canonical basis of $\mathbb{R}^n$ by

$$\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}.$$

For similar reasons, we will now write the dual basis of $(\mathbb{R}^n)^* = \mathbb{R}^n$ as

$$dx^1, \ldots, dx^n.$$

We have seen in Section 2.4.4 that the vector space $\Lambda^k(\mathbb{R}^n)$ has dimension $\binom{n}{k}$ and a basis consists of all the elements

$$dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$

where $1 \leq i_1 < \ldots < i_k \leq n$ vary. Therefore we can write any k-form ω in U in the following way:

$$\omega = \sum_{i_1 < \cdots < i_k} f_{i_1, \ldots, i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$

where $f_{i_1,\ldots,i_k}$ is some smooth function on U. The notation is appropriate because we can also interpret dx^i as the differential of the linear map $x \mapsto x_i$.

Example 7.1.2. The differential of a function $f \colon U \to \mathbb{R}$ is

$$df = \frac{\partial f}{\partial x_1} dx^1 + \ldots + \frac{\partial f}{\partial x_n} dx^n.$$

Example 7.1.3. The following are 1-forms in $\mathbb{R}^3$:

$$x^2 dy - xe^y dz, \qquad xdx + ydy + zdz$$

and the following are 2-forms:

$$xdx \wedge dy + x^3 dy \wedge dz, \qquad xdy \wedge dz - ydx \wedge dz + zdx \wedge dz.$$

Remark 7.1.4. Every n-form in $U \subset \mathbb{R}^n$ is of type

$$f dx^1 \wedge \cdots \wedge dx^n$$

for some smooth function $f : U \to \mathbb{R}$. Therefore n-forms on open sets $U \subset \mathbb{R}^n$ are somehow like smooth functions on U, but one should not go too far with this analogy, because forms and functions are intrinsically different objects!

It is sometimes useful to write a form as a linear combination of elements $dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ without the hypothesis $i_1 < \ldots < i_k$. One has to take care that the notation is not unique in this case, for instance

$$\omega = dx \wedge dy = -dy \wedge dx = \frac{1}{2} dx \wedge dy - \frac{1}{2} dy \wedge dx.$$

It suffices to keep in mind the following relations:

$$dx^i \wedge dx^j = -dx^j \wedge dx^i, \qquad dx^i \wedge dx^i = 0.$$

Example 7.1.5. With these rules in mind, it is also easy to write the wedge product of two differential forms. For instance:

$$(xz^2 dy + xdz) \wedge (e^y dx \wedge dz) = -xe^y z^2 dx \wedge dy \wedge dz.$$

7.1.4. Change of coordinates. On a chart, every form ω may be expressed uniquely as a linear combination

$$\omega = \sum_{i_1 < \cdots < i_k} f_{i_1,\ldots,i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$

If we use another chart, with variables $\bar{x}$, we get

$$\omega = \sum_{i_1 < \cdots < i_k} \bar{f}_{i_1,\ldots,i_k} d\bar{x}^{i_1} \wedge \cdots \wedge d\bar{x}^{i_k}$$

for some new functions $\bar{f}$. How can we pass from one expression to the other? The differentials dx^i are elements of $(\mathbb{R}^n)^*$ and hence change contravariantly, that is we have

$$d\bar{x}^i = \frac{\partial \bar{x}^i}{\partial x^j} dx^j.$$

The notation dx^i is designed to help us to write this equation correctly. We can then plug this expression in the linear combination to pass from one notation to the other.

Example 7.1.6. Consider the 2-form $\omega = z dx \wedge dy$ on the open set $U = \{x, y, z > 0\}$. We change the coordinates via $x = \bar{x}^2$, $y = \bar{y} + \bar{z}$, $z = \bar{y}$. Then

$$dx = 2\bar{x} d\bar{x}, \qquad dy = d\bar{y} + d\bar{z}, \qquad dz = d\bar{y}$$

and by substituting we see that ω in the new coordinates is read as

$$\omega = (\bar{y})(2\bar{x}d\bar{x}) \wedge (d\bar{y} + d\bar{z}) = 2\bar{x}\bar{y} d\bar{x} \wedge d\bar{y} + 2\bar{x}\bar{y} d\bar{x} \wedge d\bar{z}.$$

An interesting case occurs when we consider n-forms in a n-dimensional manifold. Here on a chart we have

$$\omega = f dx^1 \wedge \cdots \wedge dx^n$$

and Proposition 2.4.15 yields the following simple formula:

$$\omega = f \det \left(\frac{\partial x^i}{\partial \bar{x}^j} \right) d\bar{x}^1 \wedge \cdots \wedge d\bar{x}^n. \tag{8}$$

This equality is very much similar to the change of coordinates formula for integration given in Section 1.3.8, and this is in fact the most important feature of differential forms: they can be meaningfully integrated on manifolds, as we will soon see.

7.1.5. Support. Let M be a n-manifold and ω be a k-form on M. We define the *support* of ω to be the closure in M of the set of all the points p such that $\omega(p) \neq 0$. Using bump functions, one can easy construct plenty of non-trivial k-forms in $\mathbb{R}^n$ having compact support.

Moreover, for every k-form ω on M and every open covering U_i of M, we can pick a partition of unity ρ_i subordinate to the covering and write

$$\omega = \sum_i \rho_i \omega.$$

The support of $\rho_i \omega$ is contained in U_i for every i, and this possibly infinite sum makes sense because it is finite at every point $p \in M$. One can in this way write every k-form ω as a (possibly infinite, but locally finite) sum of compactly supported k-forms $\rho_i \omega$. If ω is already compactly supported, the sum is finite.

7.1.6. Pull-back. When we introduced tensors in Chapter 2, the roles of covariance and contravariance were somehow interchangeable, because one can switch the spaces V and V^* thanks to the canonical isomorphism $V = V^{**}$. This symmetry is now broken when we talk about manifolds and tensor fields, and it turns out that *contravariant* tensor fields are sometimes preferable.

We explain this phenomenon. Let $f: M \to N$ be any smooth map between two manifolds. We have already alluded to the fact that a covariant tensor field like a vector field cannot be transported along f in general, neither forward from M to N nor backwards from N to M. On the other hand, every contravariant tensor field α of some type $(0, k)$ on N may be transported back to a tensor field $f^*\alpha$ of the same type $(0, k)$ on M, by setting

$$f^*\alpha(p)(v_1, \ldots, v_k) = \alpha\big(f(p)\big)\big(df_p(v_1), \ldots, df_p(v_k)\big) \tag{9}$$

for every $p \in M$ and every $v_1, \ldots, v_k \in T_pM$. The tensor field $f^*\alpha$ is the *pull-back* of α along f. If α is (anti-)symmetric, then $f^*\alpha$ also is.

In particular, the pull-back of a k-form ω in N is a k-form $f^*\omega$ in M. We get a morphism of algebras

$$f^*\colon \Omega^*(N) \longrightarrow \Omega^*(M).$$

In particular, we have

$$f^*(\omega \wedge \eta) = f^*(\omega) \wedge f^*(\eta). \tag{10}$$

As usual, we can describe this operation in coordinates: let $f\colon U \to V$ be a smooth map between two open subsets $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$, and

$$\omega = \sum_{i_1 < \ldots < i_k} g_{i_1,\ldots,i_k} dx^{i_1} \wedge \ldots \wedge dx^{i_k}$$

be a k-form in V. We get

$$f^*\omega = \sum_{i_1 < \ldots < i_k} \big(g_{i_1,\ldots,i_k} \circ f\big) df_{i_1} \wedge \cdots \wedge df_{i_k}$$

where $f_i\colon U \to \mathbb{R}$ is the i-th coordinate of f and df_i its differential. This equality is proved (exercise) by showing that it satisfies (9), using (10).

Example 7.1.7. Consider $f\colon \mathbb{R}^3 \to \mathbb{R}^2$, $f(x, y, z) = (xy, yz)$ and the 2-form $\omega = xdx \wedge dy$ on $\mathbb{R}^2$. We get

$$\begin{aligned} f^*\omega &= xydf_1 \wedge df_2 = xy(ydx + xdy) \wedge (zdy + ydz) \\ &= xy^2zdx \wedge dy + xy^3dx \wedge dz + x^2y^2dy \wedge dz. \end{aligned}$$

7.2. Integration

We now show that k-forms are designed to be integrated along k-submanifolds.

7.2.1. Integration. Consider a n-form

$$\omega = f dx^1 \wedge \cdots \wedge dx^n$$

on some open subset $V \subset \mathbb{R}^n$, having compact support. We define the *integral* of ω over V simply and naïvely as

$$\int_V \omega = \int_V f.$$

Let now $\psi\colon V \to V'$ be an orientation-preserving diffeomorphism between open sets in $\mathbb{R}^n$, and denote by $\psi_*\omega = (\psi^{-1})^*\omega$ the n-form transported along ψ. Here is the crucial property that characterises differential forms:

Proposition 7.2.1. *We have*

$$\int_V \omega = \int_{V'} \psi_* \omega.$$

Proof. Combine (8), where $\det > 0$ since ψ is orientation-preserving, with the change of coordinates law for multiple integrals, see Section 1.3.8. □

It is really important that ψ be orientation-preserving: if ψ is orientation-reversing, then a minus sign appears in the equality. Encouraged by this result, we now want to extend integration of forms from open subsets of $\mathbb{R}^n$ to arbitrary oriented manifolds.

Let M be an oriented n-manifold and ω be a n-form over M with compact support. We now define the *integral of* ω *over* M, that is a number

$$\int_M \omega$$

as follows. If the support of ω is fully contained in the domain U of a chart $\varphi\colon U \to V$, then we set

$$\int_M \omega = \int_V \varphi_* \omega.$$

The definition is well-posed because it is chart-independent thanks to Proposition 7.2.1. More generally, if the support of ω is not contained in the domain of any chart, we pick an oriented atlas $\{\varphi_i\colon U_i \to V_i\}$ on M and a partition of unity ρ_i subordinated to the covering U_i. We decompose ω as a finite sum $\omega = \sum_i \rho_i \omega$ and define

$$\int_M \omega = \sum_i \int_M \rho_i \omega.$$

Proposition 7.2.2. *This definition is well-posed.*

Proof. Let $\{\varphi'_j\colon U'_j \to V'_j\}$ be another compatible oriented atlas and ρ'_j a partition of unity subordinated to U'_j. For every i we find

$$\int_M \rho_i \omega = \int_M \Big(\sum_j \rho'_j\Big)\rho_i \omega = \sum_j \int_M \rho'_j \rho_i \omega$$

and therefore

$$\sum_i \int_M \rho_i \omega = \sum_{i,j} \int_M \rho'_j \rho_i \omega.$$

Analogously we get

$$\sum_j \int_M \rho'_j \omega = \sum_{i,j} \int_M \rho'_j \rho_i \omega$$

and therefore the definition is well-posed. □

The following properties follow readily from the definitions. Let ω be a compactly supported n-form on an oriented n-manifold M. We denote by $-M$ the manifold M with the opposed orientation.

Proposition 7.2.3. *We have*

$$\int_{-M} \omega = -\int_M \omega.$$

If $f: M \to N$ is an orientation-preserving diffeomorphism, then

$$\int_M \omega = \int_N f_* \omega.$$

Remark 7.2.4. In Remark 7.1.4 we observed that on a chart a n-form looks like a function, but we warned the reader that the two notions are quite different on a general manifold M. As opposite to n-forms, functions in M cannot be integrated in any meaningful way; conversely, the value $\omega(p)$ of a n-form ω at $p \in M$ is not a number, in any reasonable sense. Shortly: functions can be evaluated at points, and n-forms can be integrated on sets, but not the converse.

7.2.2. Examples. In practice, nobody uses partitions of unity to integrate a n-form on a manifold, because the formulas get too complicated. Instead, we prefer to subdivide the manifold into small pieces where the n-form may be integrated easily. We explain briefly the details.

Let M be a smooth n-manifold. Recall the notion of Borel subset from Section 3.11.1. If ω is a compactly supported n-form on M, we can define the integral $\int_S \omega$ over a Borel set $S \subset M$ using a partition of unity as we did above.

Proposition 7.2.5. *If the support of ω is contained in a Borel set S that is a countable disjoint union of Borel sets S_i, then*

$$\int_S \omega = \sum_i \int_{S_i} \omega.$$

Proof. The equality holds for Borel sets in $\mathbb{R}^n$ because it is a property of Lebesgue integration; via a partition of unity we can extend it to M. □

Recall that having measure zero is a well-defined property for Borel subsets of any smooth manifold. If the complement of $S \subset M$ has measure zero, then

$$\int_M \omega = \int_S \omega$$

because the integral over $M \setminus S$ is zero. So we can remove from M any zero-measure set to get a more comfortable domain S and integrate ω there.

Example 7.2.6. Consider the n-dimensional torus $T = S^1 \times \cdots \times S^1$ where every point has some coordinates $(\theta^1, \ldots, \theta^n)$, and the n-form

$$\omega = d\theta^1 \wedge \cdots \wedge d\theta^n.$$

We have

$$\int_T \omega = \int_U \omega = \int_{(0,2\pi)\times\cdots\times(0,2\pi)} 1 = (2\pi)^n$$

by using the open chart $U = (0, 2\pi) \times \cdots \times (0, 2\pi)$ whose complement has measure zero.

In all our discussion we have implicitly considered only manifolds of dimension $n \geq 1$. However, it will be soon important to consider also points: we define the integral of a 0-form, that is of a function f, over an oriented point p to be $\pm f(p)$ according to the orientation of p.

7.2.3. Integration on submanifolds. By combining pull-backs and integration, we get a nice new tool: we can integrate k-forms along k-submanifolds.

Let M be a smooth manifold and ω be a fixed compactly supported k-form on M. For every oriented submanifold $S \subset M$ of dimension k, we may define the integral of ω along S as follows:

$$\int_S \omega = \int_S i^*\omega$$

where $i\colon S \hookrightarrow M$ is the inclusion map. Quite remarkably, we can use ω to assign a real number to every k-submanifold $S \subset M$.

Remark 7.2.7. More generally, the k-form ω needs not to have compact support: it suffices that the intersection of the support of ω with S is compact, and in that case the integral makes sense. For instance, this holds for every $\omega \in \Omega^k(M)$ if S is itself compact.

Shortly: functions can be evaluated at points, and k-forms can be integrated along oriented k-submanifolds.

Exercise 7.2.8. Consider the torus $T = S^1 \times S^1$ with coordinates (θ^1, θ^2) and the 1-form $\omega = d\theta^1$. Consider the 1-submanifold $\gamma_i = \{\theta^i = 0\}$ for $i = 1, 2$, oriented like S^1. We have

$$\int_{\gamma_1} \omega = 0, \qquad \int_{\gamma_2} \omega = 2\pi.$$

7.2.4. Volume form. As we anticipated in the introduction of this chapter, a smooth manifold is not equipped with any canonical notion of "volume" for its Borel subsets. We can add this geometric structure to the manifold, by selecting a preferred differential form called a *volume form*.

Let M be an oriented n-manifold.

Definition 7.2.9. A *volume form* in M is a n-form ω such that

$$\omega(p)(v_1, \ldots, v_n) > 0$$

for every $p \in M$ and every positive basis $v_1, \ldots, v_n$ of T_pM.

Let ω be a volume form on M and $S \subset M$ be a Borel set with compact closure. It makes sense to define the *volume* of S as

$$\mathrm{Vol}(S) = \int_S \omega.$$

Here is the crucial property of volume forms:

Proposition 7.2.10. *We have* $\mathrm{Vol}(S) \geq 0$. *If S has non-empty interior, then* $\mathrm{Vol}(S) > 0$.

Proof. If we use only orientation-preserving charts, the form ω transforms into n-forms $f\,dx^1 \wedge \cdots dx^n$ with $f(x) > 0$ for every x. $\square$

As in ordinary Lebesgue measure theory, we can now define $\mathrm{Vol}(S)$ for every Borel set S, as the supremum of the volumes of the Borel sets with compact closure contained in S. The volume may (or may not) be infinite if S has not compact closure. We have obtained a measure on all the Borel sets in M, that is we have the countable additivity

$$\mathrm{Vol}(S) = \sum \mathrm{Vol}(S_i)$$

whenever S is the disjoint union of countably many Borel sets S_i.

Of course different selections of the volume form ω give rise to different measures, and there is no way to choose a "preferred" volume form ω on an arbitrary oriented manifold M.

Proposition 7.2.11. *If ω is a volume form and $f : M \to \mathbb{R}$ is a strictly positive function, then $\omega' = f\omega$ is another volume form. Every volume form ω' may be constructed from ω in this way.*

Proof. The first assertion is obvious, and the converse follows from the fact that $\Lambda^n(T_pM)$ has dimension 1 and hence for every ω, ω' we may define $f(p)$ as the unique positive number such that $\omega'(p) = f(p)\omega(p)$. $\square$

We also note that volume forms always exist:

Proposition 7.2.12. *If M is oriented, there is always a volume form on M.*

Proof. Pick an oriented atlas $\{\varphi_i : U_i \to V_i\}$ and a partition of unity ρ_i subordinate to the covering $\{U_i\}$. We define

$$\omega(p) = \sum_i \rho_i(p)\varphi_i^*(dx^1 \wedge \cdots \wedge dx^n)$$

and get a volume form ω. Indeed for every $p \in M$ and positive basis $v_1, \ldots, v_n$ at T_pM the number $\omega(p)(v_1, \ldots, v_n)$ is a finite sum of strictly positive numbers with strictly positive coefficients $\rho_i(p)$, so it is strictly positive. $\square$

7.2.5. Euclidean volume form. The *Euclidean volume form* on $\mathbb{R}^n$ is

$$\omega_E = dx^1 \wedge \ldots \wedge dx^n$$

which acts as

$$\omega_E(p)(v_1, \ldots, v_n) = \det \begin{pmatrix} v_1 \cdots v_n \end{pmatrix}$$

at every $p \in \mathbb{R}^n$. It has the characterising property that $\omega_E(p)(v_1, \ldots, v_n) = 1$ for every positive orthonormal basis $v_1, \ldots, v_n$. The measure it defines in $\mathbb{R}^n$ is of course the ordinary Lebesgue measure.

More generally, we may define a *Euclidean volume form* ω on every oriented k-submanifold $M \subset \mathbb{R}^n$ as follows: for every $p \in M$ we set

$$\omega(p)(v_1, \ldots, v_k) = \omega_E(v_1, \ldots, v_n) = \det \begin{pmatrix} v_1 \cdots v_n \end{pmatrix}$$

where $v_{k+1}, \ldots, v_n$ is any orthonormal basis of the normal space N_pM. Again $\omega(p)$ is characterised by the property that $\omega(p)(v_1, \ldots, v_k) = 1$ on every positive orthonormal basis $v_1, \ldots, v_k$ for T_pM.

Note that we are using the Euclidean scalar product here to define ω. A volume form on a smooth manifold N does *not* induce in general a volume form on lower-dimensional submanifolds M. Some scalar product is needed here, as we will see in the next chapters.

Example 7.2.13. Consider the n-form ω in $\mathbb{R}^{n+1} \setminus \{0\}$ given by

$$\omega = \frac{1}{r} \sum_{i=1}^{n+1} (-1)^{i-1} x_i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^{n+1}$$

where

$$r = \sqrt{x_1^2 + \ldots + x_{n+1}^2}.$$

Consider the sphere $S(0, r)$ centred in 0 and of radius $r > 0$. We consider r as a function on $\mathbb{R}^{n+1} \setminus \{0\}$, so dr is a 1-form, and we discover easily that

$$dr \wedge \omega = dx^1 \wedge \cdots \wedge dx^{n+1}.$$

This fact implies that the restriction of ω to $S(0, r)$ is the Euclidean volume form on the sphere, for every $r > 0$. So, the Euclidean volume form on S^2 is

$$\omega = dy \wedge dz + dz \wedge dx + dx \wedge dy.$$

7.3. Stokes' Theorem

At various places in this book we introduce some objects, typically some tensor fields, and then we try to "derive" them in a meaningful way. We now show that differential forms can be derived quite easily, through an operation called *exterior derivative*, that transforms k-forms into $(k + 1)$-forms and extends the differential of functions (that transform functions, that is 0-forms, into 1-forms).

We end up by proving Stokes' Theorem, that relates elegantly exterior derivatives and integration along manifolds with boundary.

7.3.1. Exterior derivative. Let ω be a k-form in a smooth manifold M. We now define the *exterior derivative* $d\omega$, a new $(k+1)$-form on M.

We start by considering the case where M is an open set in $\mathbb{R}^n$. We have

$$\omega = \sum_{i_1<\cdots<i_k} f_{i_1,\ldots,i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$

and we define

$$d\omega = \sum_{i_1<\cdots<i_k} df_{i_1,\ldots,i_k} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

Recall that $df_{i_1,\ldots,i_k}$ is a 1-form, hence $d\omega$ is a $(k+1)$-form. When ω is a 0-form, that is a function $\omega = f$, then $d\omega$ is the ordinary differential.

Example 7.3.1. Consider the form $\omega = xydx + xydz$ in $\mathbb{R}^3$. We get

$$d\omega = xdy \wedge dx + ydx \wedge dz + xdy \wedge dz.$$

We now extend this definition to an arbitrary smooth manifold M, as usual by considering charts: we just define $d\omega$ on any open chart as above.

Proposition 7.3.2. *The definition of $d\omega$ using charts is well-posed. The derivation induces a linear map*

$$d\colon \Omega^k(M) \longrightarrow \Omega^{k+1}(M)$$

such that, for every $\omega \in \Omega^k(M)$ and $\eta \in \Omega^h(M)$ the following holds:

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta, \tag{11}$$

$$d(d\omega) = 0. \tag{12}$$

Proof. We first prove the properties on a fixed chart, and later we use these properties to show that the definition of $d\omega$ is chart-independent and hence well-posed.

Linearity of d is obvious, and using it we may suppose that $\omega = fdx^I$ and $\eta = gdx^J$ where I, J are some multi-indices. We get

$$\begin{aligned} d(\omega \wedge \eta) &= d(fg) \wedge dx^I \wedge dx^J = df \wedge dx^I \wedge gdx^J + dg \wedge fdx^I \wedge dx^J \\ &= d\omega \wedge \eta + (-1)^k \omega \wedge d\eta. \end{aligned}$$

If $\omega = fdx^I$ then

$$d(d\omega) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} dx^i \wedge dx^j \wedge dx^I = 0$$

because $dx^i \wedge dx^j = -dx^j \wedge dx^i$ so the terms cancel in pairs.

Finally, we can prove that the definition is chart-independent, via the following trick: on open subsets $U \subset \mathbb{R}^n$, the derivation d may be characterised (exercise) as the unique linear map $d\colon \Omega^k(U) \to \Omega^{k+1}(U)$ that is the ordinary differential for $k = 0$ and that satisfies (11) and (12). Therefore two definitions of d on overlapping charts must coincide in their intersection. $\square$

The following exercise says that the exterior derivative commutes with the pull-back.

Exercise 7.3.3. If $\varphi\colon M \to N$ is smooth and $\omega \in \Omega^k(N)$, we get

$$d(\varphi^*\omega) = \varphi^*(d\omega).$$

Hint. Prove it when $\omega = f$ is a function, and when $\omega = df$ is the differential of a function. Use Proposition 7.3.2 to extend it to any $\omega = f_I dx^I$. $\square$

7.3.2. Action on vector fields. We may characterise the exterior derivative of k-forms by describing their actions on vector fields. For instance, the differential df of a function f acts on vector fields $X \in \mathcal{X}(M)$ as

$$df(X) = X(f).$$

Concerning 1-forms, we get the following:

Exercise 7.3.4. If $\omega \in \Omega^1(M)$ is a 1-form and and $X, Y \in \mathcal{X}(M)$ are vector fields, we get

$$d\omega(X,Y) = X\big(\omega(Y)\big) - Y\big(\omega(X)\big) - \omega\big([X,Y]\big).$$

Hint. Again, everything is local, so work in coordinates. $\square$

A similar formula holds also for the differential $d\omega$ of a k-form.

7.3.3. Gradient, curl, and divergence. We now show that the inspiring formula $d(d\omega) = 0$ generalises a couple of familiar equalities about functions and vector fields in $\mathbb{R}^3$.

Let $U \subset \mathbb{R}^3$ be an open set. Recall that the *gradient* of a function $f : U \to \mathbb{R}$ is the vector field

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}\right).$$

If X is a vector field in U, its *divergence* is the function

$$\mathrm{div} X = \frac{\partial X^1}{\partial x_1} + \frac{\partial X^2}{\partial x_2} + \frac{\partial X^3}{\partial x_3}$$

while its *curl* is the vector field

$$\mathrm{rot} X = \left(\frac{\partial X^3}{\partial x_2} - \frac{\partial X^2}{\partial x_3}, \frac{\partial X^1}{\partial x_3} - \frac{\partial X^3}{\partial x_1}, \frac{\partial X^2}{\partial x_1} - \frac{\partial X^1}{\partial x_2}\right).$$

In U we may interpret a vector field X as a 1-form

$$\omega = X^1 dx^1 + X^2 dx^2 + X^3 dx^3$$

and vice-versa. We can also interpret a vector field X as a 2-form

$$\omega = X^1 dx^2 \wedge dx^3 + X^2 dx^3 \wedge dx^1 + X^3 dx^1 \wedge dx^2$$

and viceversa. Finally, we can interpret a 3-form as a function. Beware that this interpretation is not allowed in an arbitrary smooth manifold.

Exercise 7.3.5. With this interpretation, the equality $d(df) = 0$ for every function f in U is equivalent to

$$\mathrm{rot}(\nabla f) = 0$$

while the equality $d(d\omega) = 0$ for every 1-form ω is equivalent to

$$\mathrm{div}(\mathrm{rot} X) = 0$$

for every vector field X on U.

7.3.4. Stokes' Theorem. We first note that the whole theory of differentiable forms and integration applies also to manifolds with boundary with no modification. Then we remark a fascinating analogy: when we talk about forms ω we have

$$d(d\omega) = 0$$

while when we deal with manifolds M with boundary we also get

$$\partial(\partial M) = 0.$$

Note also that d transforms a k-form into a $(k+1)$-form, while ∂ transforms a $(k+1)$-manifold into a k-manifold. The operations d and ∂ are beautifully connected by the Stokes' Theorem.

Let M be an oriented $(n+1)$-manifold with (possibly empty) boundary, and equip ∂M with the orientation induced by M.

Theorem 7.3.6 (Stokes' Theorem). *For every compactly supported n-form ω in an oriented $(n+1)$-manifold M possibly with boundary, we have*

$$\int_M d\omega = \int_{\partial M} \omega.$$

Proof. We first prove the theorem for $M = \mathbb{R}^{n+1}_+$. We have

$$\omega = \sum_{i=1}^{n+1} \omega_i$$

with

$$\omega_i = f_i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^{n+1}$$

where the hat indicates that the i-th term is missing. By linearity it suffices to prove the theorem for each ω_i individually. We have

$$d\omega_i = df_i \wedge dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^{n+1} = (-1)^{i-1} \frac{\partial f_i}{\partial x_i} dx^1 \wedge \cdots \wedge dx^{n+1}.$$

If $i \leq n$, we have

$$\begin{aligned}
\int_{\mathbb{R}^{n+1}_+} d\omega_i &= (-1)^{i-1} \int_{\mathbb{R}^{n+1}_+} \frac{\partial f_i}{\partial x_i} dx^1 \wedge \cdots \wedge dx^{n+1} \\
&= (-1)^{i-1} \int_{\mathbb{R}^{n+1}_+} \frac{\partial f_i}{\partial x_i} dx^1 \cdots dx^{n+1} \\
&= (-1)^{i-1} \int_{\mathbb{R}^{n}_+} \left(\int_{\mathbb{R}} \frac{\partial f_i}{\partial x_i} dx^i \right) dx^1 \cdots \widehat{dx^i} \cdots dx^{n+1} = 0.
\end{aligned}$$

When the $\wedge$ is not present in the expression, it means that we are just doing the usual Lebesgue integration of functions on some Euclidean space. In the last equality we have used that

$$\begin{aligned}
\int_{\mathbb{R}} \frac{\partial f_i}{\partial x_i} dx^i = \lim_{t \to \infty} \big[& f_i(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_{n+1}) \\
& - f_i(x_1, \ldots, x_{i-1}, -t, x_{i+1}, \ldots, x_{n+1}) \big] = 0 - 0 = 0
\end{aligned}$$

because f_i has compact support. On the other hand, we also have

$$\int_{\partial \mathbb{R}^{n+1}_+} \omega_i = 0$$

because ω_i contains dx^{n+1} whose pull-back to $\partial \mathbb{R}^{n+1}_+$ vanishes.

If $i = n+1$, we get

$$\begin{aligned}
\int_{\mathbb{R}^{n+1}_+} d\omega_{n+1} &= (-1)^n \int_{\mathbb{R}^n} \left(\int_0^{+\infty} \frac{\partial f_{n+1}}{\partial x_{n+1}} dx^{n+1} \right) dx^1 \cdots dx^n \\
&= (-1)^n \int_{\mathbb{R}^n} \big(0 - f_{n+1}(x_1, \ldots, x_n, 0) \big) dx^1 \cdots dx^n \\
&= (-1)^{n+1} \int_{\mathbb{R}^n} f_{n+1}(x_1, \ldots, x_n, 0) dx^1 \cdots dx^n \\
&= \int_{\partial \mathbb{R}^{n+1}_+} f_{n+1} dx^1 \wedge \cdots \wedge dx^n = \int_{\partial \mathbb{R}^{n+1}_+} \omega_{n+1}.
\end{aligned}$$

We must justify the suspicious disappearance of the $(-1)^{n+1}$ sign in the last equality. The space $\mathbb{R}^n$ is identified naturally to $\partial \mathbb{R}^{n+1}_+$ via the map $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_n, 0)$. However, the orientation on $\partial \mathbb{R}^{n+1}_+$ induced by that of $\mathbb{R}^{n+1}_+$ coincides with that of $\mathbb{R}^n$ only when n is odd, as one can easily check. This explains the sign cancelation.

We have proved the theorem for $M = \mathbb{R}^{n+1}_+$. In general, we pick an atlas $\{\varphi_i \colon U_i \to V_i\}$ with $V_i \subset \mathbb{R}^{n+1}_+$ and a partition of unity ρ_i subordinate to U_i, so that $\omega = \sum_i \rho_i \omega$ is a finite sum (because ω has compact support). By linearity, it suffices to prove the theorem for each addendum $\rho_i \omega$, but in this case via φ_i we can transport it to a form in $\mathbb{R}^{n+1}_+$ and we are done. $\square$

Corollary 7.3.7. *If M is an oriented n-manifold without boundary, for every compactly supported $(n-1)$-form ω we have*
$$\int_M d\omega = 0.$$

7.3.5. Some consequences. Some familiar theorems in multivariate analysis in $\mathbb{R}$, $\mathbb{R}^2$, or $\mathbb{R}^3$ may be seen as particular instances of Stokes' Theorem.

In the line $\mathbb{R}$, Stokes' Theorem is just the fundamental theorem of calculus. A bit more generally, we may consider an embedded oriented arc $\gamma \subset \mathbb{R}^3$ with endpoints p and q and a smooth function f defined on it. Stokes says that
$$\int_\gamma df = f(q) - f(p).$$
So in particular the result depends only on the endpoints of γ, not of γ itself.

In the plane $\mathbb{R}^2$, we may consider a 1-form
$$\omega = f\,dx + g\,dy$$
and calculate
$$d\omega = \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dx \wedge dy.$$
For every compact domain $D \subset \mathbb{R}^2$ bounded by a simple closed curve $C = \partial D$, Stokes' Theorem transforms into *Green's Theorem*:
$$\int_C f\,dx + g\,dy = \int_D \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dxdy.$$

In the space $\mathbb{R}^3$, the boundary ∂D of a compact domain $D \subset \mathbb{R}^3$ is some surface, and we pick a vector field X on D. After interpreting X as a 2-form as in Section 7.3.3, we apply Stokes' Theorem and get the *Divergence Theorem*:
$$\int_D \mathrm{div}X = \int_{\partial D} X \cdot \mathbf{n}$$
where $\mathbf{n}$ is the normal vector to ∂D.

Finally, we can also consider an oriented surface $S \subset \mathbb{R}^3$ with some (possibly empty) boundary ∂S, and a vector field X in $\mathbb{R}^3$ supported on S. By interpreting X as a 1-form and applying Stokes' Theorem we get the *Kelvin – Stokes Theorem*:
$$\int_S \mathrm{rot}X \cdot \mathbf{n} = \int_{\partial S} X \cdot \mathbf{t}$$
where $\mathbf{n}$ is the unit normal field to S and $\mathbf{t}$ is the unit tangent field to ∂S, both oriented coherently with the orientations of S and $\mathbb{R}^3$.

We have proudly proved all these theorems (and many more!) at one time.

CHAPTER 8

De Rham cohomology

We now exploit the relation $d(d\omega) = 0$ on differential forms to build an algebraic construction called *De Rham cohomology*. This algebraic construction has some similarities with the fundamental group: it assigns groups to manifolds, and it is functorial, that is smooth maps induce groups homomorphisms. It can be used in particular to distinguish manifolds.

Cohomology is however different from fundamental groups, and may be used to fulfill some tasks that the fundamental group is unable to accomplish. For instance, we will use it to prove that the smooth manifolds

$$S^4, \qquad S^2 \times S^2, \qquad \mathbb{CP}^2$$

are pairwise non-homeomorphic, and not even homotopy equivalent, although they are all simply-connected compact four-manifolds.

8.1. Definition

In all this chapter, manifolds are allowed to have boundary even when not mentioned. When we want to consider manifolds without boundary, we will say it explicitly.

8.1.1. Closed and exact forms.

Let M be a smooth n-manifold.

Definition 8.1.1. A k-form ω on M is *closed* if $d\omega = 0$, and is *exact* if there is a $(k-1)$-form η such that $\omega = d\eta$.

Since $d(d\eta) = 0$, every exact form is also closed, but the converse does not always hold, and this is the key point that motivates everything that we are going to say in this chapter. We now list some motivating examples.

Example 8.1.2. Every n-form ω in M is closed, since $d\omega$ is a $(n+1)$-form, and every $(n+1)$-form is trivial on M. On the other hand, if M is compact, oriented, and without boundary, and ω is a volume form, then ω is not exact: if $\omega = d\eta$ by Stokes' Theorem we would get

$$\int_M \omega = \int_M d\eta = 0$$

but the integral of a volume form is always strictly positive, a contradiction.

Example 8.1.3. On the torus $T = S^1 \times S^1$ with coordinates θ^1, θ^2, the 1-form $\omega = d\theta^1$ of Exercise 7.2.8 is closed but is not exact: indeed note that θ^1 is only a locally defined function (whose value has a 2π indeterminacy); this suffices for getting closeness $d(d\theta^1) = 0$ but not for exactness. If we had $\omega = df$ for a true function f, then the integral of ω over the curve γ_2 would vanish by Stokes' Theorem, a contradiction.

Example 8.1.4. Pick $U = \mathbb{R}^2 \setminus \{0\}$. Using polar coordinates ρ, θ we may define the closed non-exact form $\omega = d\theta$ on U, like in the previous example. In Euclidean coordinates the form is

$$\omega = \frac{x dy - y dx}{x^2 + y^2}$$

and the skeptic reader may check that $d\omega = 0$ via direct calculation. As above, the 1-form is not exact because its integral above the curve $S^1 \subset U$ is $2\pi \neq 0$.

In the last example, it is tempting to think that ω is not exact because there is a "hole" in U where the origin has been removed (note that ω does not extend to the origin). We will confirm this intuition in the next pages: closed non-exact forms detect some kinds of topological holes in the manifold M, and this precious information is efficiently organised into the more algebraic *De Rham cohomology*.

8.1.2. De Rham cohomology. Let M be a smooth manifold. We define

$$Z^k(M), \qquad B^k(M)$$

respectively as the vector subspaces of $\Omega^k(M)$ consisting of all the closed and all the exact k-forms.

As we said, we have the inclusion $B^k(M) \subset Z^k(M)$ and hence we may define the *De Rham cohomology group* as the quotient

$$H^k(M) = Z^k(M)/_{B^k(M)}.$$

This is actually a vector space, but the term "group" is usually employed in analogy with some more general constructions where all these spaces are modules over some ring.

An element in $H^k(M)$ is usually denoted as a k-form ω, and sometimes as a class $[\omega]$ of k-forms when we feel the need to be more rigorous.

8.1.3. The Betti numbers. The *k-th Betti number* of M is the dimension

$$b^k(M) = \dim H^k(M).$$

Of course this number may be infinite, but we will see that it is finite in the most interesting cases. This is a remarkable and maybe unexpected fact, since both $Z^k(M)$ and $B^k(M)$ are typically infinite-dimensional.

The Betti number $b^k(M)$ depends only on M and is hence a numerical invariant of the smooth manifold M. That is, two diffeomorphic manifolds have the same Betti numbers.

Proposition 8.1.5. *For every $k > \dim M$ we have $b^k(M) = 0$.*

Proof. There are no k-forms on M for $k > n$. □

8.1.4. The Euler characteristic. Let M be a smooth n-manifold whose Betti numbers b^k are all finite. The *Euler characteristic* of M is the integer

$$\chi(M) = \sum_{i=0}^{n} (-1)^i b^i(M).$$

This is an ubiquitous invariant, defined also for more general topological spaces.

8.1.5. The zeroest group. As a start, we may easily identify $H^0(M)$ for any smooth manifold M.

We first make a general remark: if M has finitely many connected components $M_1, \ldots, M_h$, we naturally get

$$H^k(M) = H^k(M_1) \oplus \cdots \oplus H^k(M_h).$$

For this reason, we usually suppose that M be connected.

Proposition 8.1.6. *If M is connected, there is a natural isomorphism*

$$H^0(M) \cong \mathbb{R}.$$

Proof. The space $Z^0(M)$ consists of all the functions $f: M \to \mathbb{R}$ such that $df = 0$, and $B^0(M)$ is trivial. By taking charts, we see that $df = 0 \iff f$ is locally constant (that is, every $p \in M$ has a neighbourhood where f is constant) $\iff f$ is constant, since M is connected. Therefore $H^0(M) = Z^0(M)$ consists of the constant functions and is hence naturally isomorphic to $\mathbb{R}$. □

For a possibly disconnected M, we get the following.

Corollary 8.1.7. *The Betti number $b^0(M)$ equals the number of connected components of M.*

8.1.6. The cohomology algebra. Let M be a smooth manifold. We may define the vector space

$$H^*(M) = \bigoplus_{k \geq 0} H^k(M).$$

Proposition 8.1.8. *The exterior product $\wedge$ descends to $H^*(M)$ and gives it the structure of an associative algebra.*

Proof. If $\omega \in Z^k(M)$ and $\eta \in Z^h(M)$ then

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta = 0$$

and hence $\omega \wedge \eta \in Z^{k+h}(M)$. If moreover $\omega \in B^k(M)$, that is $\omega = d\zeta$, we get

$$\omega \wedge \eta = d\zeta \wedge \eta = d(\zeta \wedge \eta) - (-1)^{k-1} \zeta \wedge d\eta = d(\zeta \wedge \eta)$$

and hence $\omega \wedge \eta \in B^{k+h}(M)$. Therefore the product passes to the quotients $H^k(M)$ and $H^h(M)$. □

If $\omega \in H^p(M)$ and $\eta \in H^q(M)$, then $\omega \wedge \eta \in H^{p+q}(M)$. As for $\Omega^*(M)$, the algebra $H^*(M)$ is anticommutative, that is

$$\omega \wedge \eta = (-1)^{pq} \eta \wedge \omega.$$

In particular, if p is odd we get

$$\omega \wedge \omega = 0.$$

8.1.7. Functoriality. Every smooth map $f: M \to N$ induces a linear map

$$f^* : \Omega^k(N) \longrightarrow \Omega^k(M)$$

by pull-back. The map commutes with d and hence it sends close forms to close forms, and exact forms to exact forms. Therefore it induces a map

$$f^* : H^k(N) \longrightarrow H^k(M)$$

and more generally a morphism of algebras

$$f^* : H^*(N) \longrightarrow H^*(M).$$

We may say that cohomology is a *contravariant functor*, where *contravariant* means that arrows are reversed (we go backwards from $H^k(N)$ to $H^k(M)$), and *functor* means that $(f \circ g)^* = g^* \circ f^*$ and $\mathrm{id}_M^* = \mathrm{id}_{H^*(M)}$.

The reader should compare this functor with the *covariant functor* furnished by the fundamental group, that sends pointed topological spaces (X, x_0) to groups $\pi_1(X, x_0)$.

8.1.8. The line. The De Rham cohomology of $\mathbb{R}$ can be calculated easily.

Proposition 8.1.9. *We have $H^0(\mathbb{R}) = \mathbb{R}$ and $H^k(\mathbb{R}) = 0$ for all $k > 0$.*

Proof. There are no k-forms with $k \geq 2$, so the only thing to prove is that $H^1(\mathbb{R}) = 0$. Given a 1-form $\omega = f(x)dx$, we can define

$$F(x) = \int_0^x f(t)dt$$

and we get $\omega = dF$. Therefore every 1-form is exact and $H^1(\mathbb{R}) = 0$. $\square$

We say that the cohomology of a manifold M is *trivial* if $H^0(M) = \mathbb{R}$ and $H^k(M) = 0$ for all $k > 0$. We will soon discover that the cohomology of $\mathbb{R}^n$ is also trivial for every n.

8.1.9. Integration along submanifolds. Let M be a n-manifold and $S \subset M$ an oriented compact $k-$submanifold without boundary. Remember that every k-form $\omega \in \Omega^k(M)$ may be integrated over S, so furnishing a linear map

$$\int_S : \Omega^k(M) \longrightarrow \mathbb{R}.$$

By Stokes' Theorem, the integral of an exact form vanishes, and hence this linear map descends to a map in cohomology

$$\int_S : H^k(M) \longrightarrow \mathbb{R}.$$

This shows in particular that if the integral of a k-form ω is non-zero on some oriented compact k-submanifold S, then ω is non-trivial in $H^k(M)$.

8.2. The Poincaré Lemma

One important feature of the fundamental group is that the it is unaffected by homotopies. We prove here the same thing for the De Rham cohomology. As a consequence, we will show that the cohomology of $\mathbb{R}^n$ is trivial, as that of any contractible manifold. This fact is known as the *Poincaré Lemma*.

8.2.1. Cochain complexes. Some of the properties of De Rham cohomology may be deduced by purely algebraic means, and work in more general contexts. For these reasons we now reintroduce cohomologies with a purely algebraic language.

A *cochain complex* C is a sequence of vector spaces $C^0, C^1, C^2, \ldots$ with linear maps $d^k\colon C^k \to C^{k+1}$ such that $d^{k+1} \circ d^k = 0$ for all k. We usually indicate d^k by d and write the cochain complex as

$$C^0 \xrightarrow{d} C^1 \xrightarrow{d} C^2 \xrightarrow{d} \ldots$$

The elements in $Z^k = \ker d^k$ are called *cocycles*, and those in $B^k = \operatorname{Im} d^{k-1}$ are the *coboundaries*. The *cohomology* of C is constructed as above as $H^k = Z^k/_{B^k}$ for every $k \geq 0$. We may indicate it as $H^k(C)$ to stress its dependence on the cochain complex C.

Of course when $C^k = \Omega^k(M)$ we obtain the De Rham cohomology of M, but this general construction applies to many other contexts, so it makes sense to consider it abstractly.

Remark 8.2.1. A *chain complex* is a sequence of vector spaces $C_0, C_1, \ldots$ equipped with maps $d_k\colon C_k \to C_{k-1}$ such that $d \circ d = 0$. The theory of chain complexes is similar and somehow dual to that of *co*chain complexes: one defines the *cycles* as $Z_k = \ker d_k$, the *boundaries* as $B_k = \operatorname{Im} d_{k+1}$, and the *homology group* $H_k = Z_k/_{B_k}$.

A *morphism* between two cochain complexes C and D is a map $f^k\colon C^k \to D^k$ for all $k \geq 0$ such that the following diagram commutes

$$\begin{array}{ccccccccc}
\cdots & \xrightarrow{d} & C^{k-1} & \xrightarrow{d} & C^k & \xrightarrow{d} & C^{k+1} & \xrightarrow{d} & \cdots \\
 & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle f} & & \\
\cdots & \xrightarrow{d} & D^{k-1} & \xrightarrow{d} & D^k & \xrightarrow{d} & D^{k+1} & \xrightarrow{d} & \cdots
\end{array}$$

We have denoted f^k simply by f. Since f commutes with d, it sends cocycles to cocycles and coboundaries to coboundaries, and hence induces a homomorphism $f_*\colon H^k(C) \to H^k(D)$ for every k.

8.2.2. Cochain homotopy. We introduce an algebraic notion of homotopy that will reflect the notion of homotopy between maps. Let $f, g\colon C \to D$ be two morphisms between cochain complexes. A *cochain homotopy* between them is a linear map $h^k\colon C^k \to D^{k-1}$ for all $k \geq 1$ such that

$$f^k - g^k = d^{k-1} \circ h^k + h^{k+1} \circ d^k$$

for all $k \geq 0$. Shortly, we may write

$$f - g = d \circ h + h \circ d. \tag{13}$$

We may visualise everything by drawing the following diagram:

$$\begin{tikzcd}
\cdots \arrow[r, "d"] & C^{k-1} \arrow[r, "d"] \arrow[dl, "h"'] \arrow[d, shift right, "g"'] \arrow[d, shift left, "f"] & C^k \arrow[r, "d"] \arrow[dl, "h"'] \arrow[d, shift right, "g"'] \arrow[d, shift left, "f"] & C^{k+1} \arrow[r, "d"] \arrow[dl, "h"'] \arrow[d, shift right, "g"'] \arrow[d, shift left, "f"] & \cdots \arrow[dl, "h"'] \\
\cdots \arrow[r, "d"] & D^{k-1} \arrow[r, "d"] & D^k \arrow[r, "d"] & D^{k+1} \arrow[r, "d"] & \cdots
\end{tikzcd}$$

Note that this diagram is *not* commutative. Two cochain maps f, g are *cochain homotopic* if there is a cochain homotopy between them. The relevance of cochain homotopies relies in the following fact.

Proposition 8.2.2. *If two cochain maps f, g are cochain homotopic, they induce the same maps in cohomology.*

Proof. For every $a \in C^k$ we have

$$f(a) - g(a) = d(h(a)) + h(d(a)).$$

If $a \in Z^k(C)$ we get $d(a) = 0$ and hence

$$f(a) - g(a) = d(h(a)) \in B^k(D).$$

Therefore f and g induce the same maps on cohomology. $\square$

Having settled the basic algebraic machinery, we now turn back to De Rham cohomology.

8.2.3. Products with a line. We now prove that M and $M \times \mathbb{R}$ have the same cohomology. Since we already know the cohomology of $\mathbb{R}$, this will imply that $\mathbb{R}^n$ and $\mathbb{R}$ have the same cohomology.

Let M be a smooth manifold and $t_0 \in \mathbb{R}$ a point. We have two maps

$$\pi\colon M \times \mathbb{R} \longrightarrow M, \qquad s\colon M \longrightarrow M \times \mathbb{R}.$$

The first is the projection, the second is $s(p) = (p, t_0)$. These induce

$$\pi^*\colon H^*(M) \longrightarrow H^*(M \times \mathbb{R}), \qquad s^*\colon H^*(M \times \mathbb{R}) \longrightarrow H^*(M).$$

Lemma 8.2.3. *The maps s^* and π^* are isomorphisms and $s^* = (\pi^*)^{-1}$.*

Proof. We have $\pi \circ s = \mathrm{id}_M$ and functioriality gives $s^* \circ \pi^* = \mathrm{id}_{H^*(M)}$. However $s \circ \pi \neq \mathrm{id}_{M\times\mathbb{R}}$, and the map

$$\pi^* \circ s^* \colon \Omega^*(M \times \mathbb{R}) \to \Omega^*(M \times \mathbb{R})$$

is not the identity in general. We now construct a cochain homotopy

$$h \colon \Omega^k(M \times \mathbb{R}) \longrightarrow \Omega^{k-1}(M \times \mathbb{R})$$

between $\pi^* \circ s^*$ and the identity: this implies by Proposition 8.2.2 that $\pi^* \circ s^*$ induces the identity map on cohomology, and concludes the proof.

We define h as follows:

$$(h\omega)(p,t)(v_1,\ldots,v_{k-1}) = \int_{t_0}^{t} \omega(p,u)\left(\frac{\partial}{\partial t}, v_1, \ldots, v_{k-1}\right) du.$$

Here we have identified the tangent spaces of (p,t) and (p,u) in the obvious way. We need to prove that h is a cochain homotopy, that is

$$(dh + hd)(\omega) = (\mathrm{id} - \pi^* \circ s^*)(\omega)$$

for every k-form ω. Since this is a local property, we may pick a chart and suppose that $M = \mathbb{R}^n$. We use coordinates $(x_1, \ldots, x_n, t)$ for $M \times \mathbb{R}$. Every k-form in $M \times \mathbb{R}$ may be written uniquely as a linear combination of k-forms of two types:

(1) $f dx^I$,
(2) $g dt \wedge dx^J$

where the multi-indices I and J have order k and $k-1$ respectively. By linearity we may suppose that ω is of type (1) or (2). We get:

$$\begin{aligned}
(\pi^* \circ s^*)(f dx^I) &= f(x, t_0) dx^I, \\
(\pi^* \circ s^*)(g dt \wedge dx^J) &= 0, \\
h(f dx^I) &= 0, \\
h(g dt \wedge dx^J) &= \left(\int_{t_0}^{t} g(x,u) du\right) dx^J.
\end{aligned}$$

There are two cases:

(1) We have $\omega = f dx^I$ and hence

$$\begin{aligned}
(dh + hd)(\omega) &= hd\omega = h\left(df \wedge dx^I\right) = h\left(\frac{\partial f}{\partial t} dt \wedge dx^I\right) \\
&= \big(f(x,t) - f(x,t_0)\big) dx^I, \\
(\mathrm{id} - \pi^* \circ s^*)(\omega) &= \big(f(x,t) - f(x,t_0)\big) dx^I.
\end{aligned}$$

(2) We have $\omega = g dt \wedge dx^J$ and hence

$$
\begin{aligned}
dh(\omega) &= d\left(\left(\int_{t_0}^{t} g(x,u)du\right) dx^J\right) \\
&= g dt \wedge dx^J + \sum_{i=1}^{n} \int_{t_0}^{t} \frac{\partial g}{\partial x_i} dx^i \wedge dx^J, \\
hd(\omega) &= h\left(-\sum_{i=1}^{n} \frac{\partial g}{\partial x_i} dt \wedge dx^i \wedge dx^J\right) \\
&= -\sum_{i=1}^{n} \int_{t_0}^{t} \frac{\partial g}{\partial x_i} dx^i \wedge dx^J, \\
(dh + hd)(\omega) &= \omega, \\
(\mathrm{id} - \pi^* \circ s^*)(\omega) &= \omega.
\end{aligned}
$$

The proof is complete. □

We have proved with some effort that products with lines do not affect the cohomology. This fact has many nice consequences.

8.2.4. Poincaré Lemma. The first immediate corollary of Lemma 8.2.3 is the following. Let $k \geq 1$.

Corollary 8.2.4 (Poincaré's Lemma). *Every closed k-form in $\mathbb{R}^n$ is exact.*

Proof. We know from Proposition 8.1.9 that the cohomology of $\mathbb{R}$ is trivial, and Lemma 8.2.3 applied inductively on n gives $H^k(\mathbb{R}^n) = H^k(\mathbb{R})$ for all k. □

In other words, we have $H^0(\mathbb{R}^n) = \mathbb{R}$ and $H^k(\mathbb{R}^n) = 0$ for all $k > 0$.

8.2.5. Homotopy invariance. Lemma 8.2.3 has applications that go far beyond the Poincaré Lemma. Let M and N be two smooth manifolds of dimensions m and n.

Corollary 8.2.5. *Two homotopic smooth maps $f, g\colon M \to N$ induce the same homomorphisms $f^* = g^*\colon H^*(N) \to H^*(M)$ in De Rham cohomology.*

Proof. Let F be the homotopy between f and g. By Corollary 5.6.9 we may suppose that F is smooth. We have

$$f = F \circ s_0, \qquad g = F \circ s_1$$

where $s_t(p) = (p, t)$. In cohomology we have

$$f^* = s_0^* \circ F^*, \qquad g^* = s_1^* \circ F^*.$$

From Lemma 8.2.3 we get $s_0^* = (\pi^*)^{-1} = s_1^*$ and hence $f^* = g^*$. □

We discover in particular that cohomology is a homotopy invariant.

Corollary 8.2.6. *Two homotopically equivalent manifolds have isomorphic De Rham cohomologies.*

Proof. If $f : M \to N$ and $g : N \to M$ are homotopy equivalences, then $f \circ g \sim \mathrm{id}_N$ and $g \circ f = \mathrm{id}_M$ and hence $f^* \circ g^* = \mathrm{id}$ and $g^* \circ f^* = \mathrm{id}$. □

In particular, two homeomorphic manifolds have the same De Rham cohomology. This is a quite remarkable fact: the cohomology groups $H^*(M)$ are defined in an analytic way through k-forms, but the result is in fact independent of the smooth structure. The following corollary strengthens the Poincaré Lemma.

Corollary 8.2.7. *Every contractible manifold has trivial cohomology.*

Proof. The point (or $\mathbb{R}$, if you prefer) has trivial cohomology. □

8.2.6. Closed orientable manifolds. We now use the De Rham cohomology to prove a non-trivial topological fact.

Proposition 8.2.8. *A compact oriented manifold M without boundary with* $\dim M \geq 1$ *is never contractible.*

Proof. The manifold M has a volume form ω by Proposition 7.2.12, and Example 8.1.2 shows that ω is closed but not exact. Therefore $H^n(M) \neq 0$ for $n = \dim M$. In particular the cohomology of M is not trivial. □

Note that the hypothesis "compact" and "without boundary" are both necessary, as the counterexamples $\mathbb{R}^n$ and D^n show. The orientability hypothesis may be removed, but more work is needed for that (for instance, one may use a different kind of cohomology).

With the same techniques, we can in fact prove more.

Proposition 8.2.9. *A compact oriented manifold M without boundary is never homotopy equivalent to any manifold N with* $\dim N < \dim M$.

Proof. If $m = \dim M$, we have $H^m(M) \neq 0$ and $H^m(N) = 0$. □

8.3. The Mayer – Vietoris sequence

We have calculated the De Rham cohomology of contractible spaces, and we are ready for more complicated manifolds. The main tool for calculating $H^*(M)$ for general manifolds M is the *Mayer – Vietoris sequence*, and we introduce it here.

8.3.1. Exact sequences. We now introduce some algebra. A (finite or infinite) sequence of real vector spaces and linear maps

$$\ldots \longrightarrow V_{i-1} \xrightarrow{f_{i-1}} V_i \xrightarrow{f_i} V_{i+1} \longrightarrow \ldots$$

is *exact* if $\operatorname{Im} f_i = \ker f_{i+1}$ for all i such that f_i and f_{i+1} are both defined. The vector spaces V_i may have infinite dimension, although in most cases they will be finite: see Section 2.1.6 for the appropriate definitions in the infinite-dimensional case.

For instance, the following sequence

$$0 \longrightarrow V \xrightarrow{f} W$$

is exact $\iff$ f is injective, and

$$V \xrightarrow{f} W \longrightarrow 0$$

is exact $\iff$ g is surjective. The sequence

$$0 \longrightarrow U \xrightarrow{f} V \xrightarrow{g} W \longrightarrow 0$$

is exact $\iff$ f is injective, g is surjective, and $\operatorname{Im} f = \ker g$. An exact sequence of this type is called a *short exact sequence*.

Exercise 8.3.1. If a sequence

$$\ldots \longrightarrow V_{i-1} \xrightarrow{f_{i-1}} V_i \xrightarrow{f_i} V_{i+1} \longrightarrow \ldots$$

is exact, then the following sequences are also exact:

$$\ldots \longleftarrow V_{i-1}^* \xleftarrow{f_{i-1}^*} V_i^* \xleftarrow{f_i^*} V_{i+1}^* \longleftarrow \ldots$$

$$\ldots \longrightarrow V_{i-1} \otimes W \xrightarrow{f_{i-1} \otimes \mathrm{id}} V_i \otimes W \xrightarrow{f_i \otimes \mathrm{id}} V_{i+1} \otimes W \longrightarrow \ldots$$

for every vector space W.

Exercise 8.3.2. For every finite exact sequence of finite-dimensional spaces

$$0 \longrightarrow V_1 \xrightarrow{f_1} V_2 \xrightarrow{f_2} \ldots \xrightarrow{f_{k-1}} V_k \longrightarrow 0$$

we have

$$\sum_{i=1}^k (-1)^i \dim V_i = 0.$$

8.3.2. The long exact sequence. The notion of exact sequence applies also to other algebraic notions like groups, modules, etc. and also to cochain complexes: a short exact sequence of cochain complexes is an exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

where A, B, C are cochain complexes and f, g are morphisms. Exactness means that f is injective, g is surjective, and $\operatorname{Im} f = \ker g$. That is, we have a big planar commutative diagram of morphisms

(14)
$$\begin{array}{ccccccccc}
 & & \vdots & & \vdots & & \vdots & & \\
 & & \downarrow d & & \downarrow d & & \downarrow d & & \\
0 & \longrightarrow & A^{k-1} & \xrightarrow{f} & B^{k-1} & \xrightarrow{g} & C^{k-1} & \longrightarrow & 0 \\
 & & \downarrow d & & \downarrow d & & \downarrow d & & \\
0 & \longrightarrow & A^{k} & \xrightarrow{f} & B^{k} & \xrightarrow{g} & C^{k} & \longrightarrow & 0 \\
 & & \downarrow d & & \downarrow d & & \downarrow d & & \\
0 & \longrightarrow & A^{k+1} & \xrightarrow{f} & B^{k+1} & \xrightarrow{g} & C^{k+1} & \longrightarrow & 0 \\
 & & \downarrow d & & \downarrow d & & \downarrow d & & \\
 & & \vdots & & \vdots & & \vdots & &
\end{array}$$

where every horizontal line is a short exact sequence of vector spaces.

Theorem 8.3.3. *Every short exact sequence of cochain complexes*

(15) $$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

induces naturally an exact sequence in cohomology

(16) $$\cdots \longrightarrow H^k(A) \xrightarrow{f_*} H^k(B) \xrightarrow{g_*} H^k(C) \xrightarrow{\delta} H^{k+1}(A) \longrightarrow \cdots$$

for some appropriate morphism δ.

Proof. The morphism

$$\delta\colon H^k(C) \longrightarrow H^{k+1}(A)$$

is defined as follows. Given a cocycle $\gamma \in C^k$, by surjectivity of g there is a $\beta \in B^k$ with $g(\beta) = \gamma$. We have

$$g(d\beta) = dg(\beta) = d\gamma = 0$$

because γ is a cocycle. Since $\operatorname{Im} f = \ker g$ there is an $\alpha \in A^{k+1}$ such that $f(\alpha) = d\beta$, and we set

$$\delta(\gamma) = \alpha.$$

There are now a number of things to check, and we leave to the reader the pleasure of proving all of them through "diagram chasing." Here are they:

- α is a cocycle, that is $d\alpha = 0$;
- the class $[\alpha] \in H^{k+1}(A)$ does not depend on the choices of β and α;
- if γ is a coboundary then α also is.

This shows that δ is well-defined. Finally, we have to show that the sequence (16) is exact. Have fun! $\square$

The induced sequence (16) is called the *long exact sequence* associated to the short exact sequence (15).

8.3.3. The Mayer – Vietoris sequence. It is now time to go back to smooth manifolds and their De Rham cohomology.

Let M be a smooth manifold, and $U, V \subset M$ be two open subsets covering M, that is with $U \cup V = M$. The inclusions

$$\begin{tikzcd} & U \arrow[dr, "l"] & \\ U \cap V \arrow[ur, "i"] \arrow[dr, "j"'] & & M \\ & V \arrow[ur, "m"'] & \end{tikzcd}$$

induce the morphisms in cohomology

$$\begin{tikzcd} & H^k(U) \arrow[dl, "i^*"'] & \\ H^k(U \cap V) & & H^k(M) \arrow[ul, "l^*"'] \arrow[dl, "m^*"] \\ & H^k(V) \arrow[ul, "j^*"] & \end{tikzcd}$$

Theorem 8.3.4 (Mayer – Vietoris Theorem). *There is an exact sequence*

$$\cdots \longrightarrow H^k(M) \xrightarrow{(l^*,m^*)} H^k(U) \oplus H^k(V) \xrightarrow{i^*-j^*} H^k(U \cap V) \xrightarrow{\delta} H^{k+1}(M) \longrightarrow \cdots$$

for some canonically defined map δ.

Proof. This is the long exact sequence obtained via Theorem 8.3.3 from the short exact sequence of cochain complexes

$$0 \longrightarrow \Omega^*(M) \xrightarrow{(l^*,m^*)} \Omega^*(U) \oplus \Omega^*(V) \xrightarrow{i^*-j^*} \Omega^*(U \cap V) \longrightarrow 0.$$

We only need to check that this short sequence is exact. Note that the morphisms l^*, m^*, i^*, and j^* are just restrictions of k-forms to open subsets. There are three things to check:

- The map (l^*, m^*) is clearly injective.
- If (α, β) is such that $i^*(\alpha) = j^*(\beta)$, then α and β agree on $U \cap V$ and hence are restrictions of a global form in M.
- To prove that $i^* - j^*$ is surjective, pick a partition of unity ρ_U, ρ_V subordinate to $\{U, V\}$. Given $\omega \in \Omega^k(U \cap V)$, note that $\rho_V \omega$ extends smoothly to U simply by setting it constantly zero on $U \backslash V$. Therefore $\rho_V \omega \in \Omega^k(U)$ and $\rho_U \omega \in \Omega^k(V)$ and we can write
$$(i^* - j^*)(\rho_V \omega, -\rho_U \omega) = (\rho_U + \rho_V)\omega = \omega.$$

The proof is complete. $\square$

The exact sequence resulting from Theorem 8.3.4 is called the *Mayer – Vietoris long exact sequence* induced by the covering $\{U, V\}$ of M. Recall that $H^k(M) = 0$ whenever $k > n = \dim M$, so the Mayer – Vietoris sequence is finite. It starts and ends as follows:

$$0 \longrightarrow H^0(M) \longrightarrow H^0(U) \oplus H^0(V) \longrightarrow \cdots \longrightarrow H^n(U \cap V) \longrightarrow 0.$$

The morphisms i^*, j^*, l^*, m^* are simply restrictions of k-forms. The morphism δ is a bit more complicated, and for many applications we do not really need to understand it, so the reader may decide to jump to the next section. Just in case, here is a description of δ. Let ρ_U, ρ_V be a partition of unity subordinated to the covering $\{U, V\}$. Given a k-form ω in $U \cap V$, we may consider the $(k+1)$-form

$$\eta = -d\rho_V \wedge \omega = d\rho_U \wedge \omega.$$

The forms $d\rho_V$ and $d\rho_U$ have their support in $U \cap V$, hence the support of η is also in $U \cap V$. The two expressions coincide since $d\rho_U + d\rho_V = 0$.

Proposition 8.3.5. *We have $\delta(\omega) = \eta$.*

Proof. The proofs of Theorems 8.3.3 and 8.3.4 show that $\delta(\omega)$ is constructed by picking the counterimage $(-\rho_V\omega, \rho_U\omega)$ of ω, then differentiating

$$\big(-d(\rho_V\omega), d(\rho_U\omega)\big) = (-d\rho_V \wedge \omega, d\rho_U \wedge \omega)$$

using $d\omega = 0$, and finally noting that the pair is the image of η. $\square$

8.3.4. Cohomology of spheres. As a reward for all the effort that we made with short and long sequences, we can now easily calculate the De Rham cohomology of spheres.

Proposition 8.3.6. *For every $n \geq 1$ we have*

$$H^0(S^n) \cong H^n(S^n) \cong \mathbb{R}, \qquad H^k(S^n) = 0 \;\; \forall k \neq 0, n.$$

Proof. Using stereographic projections along opposite poles we may cover S^n as $S^n = U \cup V$ with $U \cong V \cong \mathbb{R}^n$ and also $U \cap V \cong S^{n-1} \times \mathbb{R}$. By homotopy equivalence, we have $H^*(U \cap V) \cong H^*(S^{n-1})$.

We first examine the case $n = 1$. Remember that $H^k(M) = 0$ whenever $k > \dim M$. The Mayer – Vietoris sequence is

$$0 \longrightarrow H^0(S^1) \longrightarrow H^0(\mathbb{R}^1) \oplus H^0(\mathbb{R}^1) \longrightarrow H^0(S^0) \stackrel{\delta}{\longrightarrow} H^1(S^1) \longrightarrow 0$$

which translates as

$$0 \longrightarrow \mathbb{R} \longrightarrow \mathbb{R} \oplus \mathbb{R} \longrightarrow \mathbb{R} \oplus \mathbb{R} \longrightarrow H^1(S^1) \longrightarrow 0.$$

since S^0 has two connected components. Exercise 8.3.2 gives $H^1(S^1) \cong \mathbb{R}$.

We now consider the case $n \geq 2$. The Mayer – Vietoris sequence breaks into pieces since $H^k(\mathbb{R}^n) \oplus H^k(\mathbb{R}^n) = 0$ for all $k > 0$. It starts with

$$0 \longrightarrow H^0(S^n) \longrightarrow H^0(\mathbb{R}^n) \oplus H^0(\mathbb{R}^n) \longrightarrow H^0(S^{n-1}) \xrightarrow{\delta} H^1(S^n) \longrightarrow 0$$

which translates as

$$0 \longrightarrow \mathbb{R} \longrightarrow \mathbb{R} \oplus \mathbb{R} \longrightarrow \mathbb{R} \longrightarrow H^1(S^n) \longrightarrow 0.$$

Therefore $H^1(S^n) = 0$. Then for every $2 \leq k \leq n$ we get

$$0 \longrightarrow H^{k-1}(S^{n-1}) \xrightarrow{\delta} H^k(S^n) \longrightarrow 0$$

and therefore $H^k(S^n) \cong H^{k-1}(S^{n-1})$. We conclude by induction on n. $\square$

8.3.5. Complex projective spaces. The De Rham cohomology of the complex projective spaces is quite different from that of the spheres, and is in fact very interesting:

Proposition 8.3.7. *We have*

$$H^k(\mathbb{CP}^n) = \begin{cases} \mathbb{R} & \text{if } k \text{ is even and } k \leq 2n, \\ 0 & \text{if } n \text{ otherwise.} \end{cases}$$

Proof. Consider a complex hyperplane $H \subset \mathbb{CP}^n$ and a point $p \in \mathbb{CP}^n$ not contained in H. Pick the open sets

$$U = \mathbb{CP}^n \setminus H, \qquad V = \mathbb{CP}^n \setminus \{p\}.$$

We have the diffeomorphisms

$$U \cong \mathbb{R}^{2n}, \qquad U \cap V \cong \mathbb{R}^{2n} \setminus \{p\} \cong S^{2n-1} \times \mathbb{R}.$$

The pencil of complex lines passing through p gives V the structure of a $\mathbb{C}$-bundle over $H \cong \mathbb{CP}^{n-1}$. In particular, we have the homotopy equivalences

$$U \sim \{\mathrm{pt}\}, \qquad U \cap V \sim S^{2n-1}, \qquad V \sim \mathbb{CP}^{n-1}.$$

The Mayer – Vietoris sequence gives

$$H^{k-1}(S^{2n-1}) \longrightarrow H^k(\mathbb{CP}^n) \longrightarrow H^k(\mathbb{CP}^{n-1}) \longrightarrow H^k(S^{2n-1})$$

for every $k \geq 1$. When $k < 2n-1$, we deduce that

$$H^k(\mathbb{CP}^n) \cong H^k(\mathbb{CP}^{n-1}).$$

When $k = 2n-1$ we get

$$0 = H^{2n-2}(S^{2n-1}) \longrightarrow H^{2n-1}(\mathbb{CP}^n) \longrightarrow H^{2n-1}(\mathbb{CP}^{n-1}) = 0$$

and therefore $H^{2n-1}(\mathbb{CP}^n) = 0$. Finally, the sequence ends with

$$0 \longrightarrow H^{2n-1}(S^{2n-1}) \longrightarrow H^{2n}(\mathbb{CP}^n) \longrightarrow 0$$

that gives $H^{2n}(\mathbb{CP}^n) = \mathbb{R}$. We conclude by induction on n, starting with $\mathbb{CP}^1 \cong S^2$. $\square$

Corollary 8.3.8. *The manifolds S^{2n} and $\mathbb{CP}^n$ are not diffeomorphic, and in fact not even homotopy equivalent, when $n > 1$.*

8.4. Compactly supported forms

We now introduce a variation of De Rham cohomology that considers only forms with compact supports. We will see that this variation has a somehow dual behaviour with respect to De Rham cohomology.

8.4.1. Definition. Let M be a smooth manifold. For every $k \geq 0$ we define the vector subspace

$$\Omega^k_c(M) \subset \Omega^k(M)$$

that consists of all the k-forms having compact support. Of course if M is compact we have $\Omega^k_c(M) = \Omega^k(M)$. The differential restrict to a map

$$d\colon \Omega^k_c(M) \longrightarrow \Omega^{k+1}_c(M)$$

with $d^2 = 0$. As above, we get a cochain complex $\Omega^*_c(M)$, and its cohomology is called the *De Rham cohomology with compact support*

$$H^k_c(M).$$

Of course when M is compact we get nothing new, but $H^k_c(M)$ may differ from $H^k(M)$ when M is not compact, as we now show.

8.4.2. The zeroest group. We now study $H^0_c(M)$ and notice immediately a difference between the compact and the non compact case.

As with De Rham cohomology, if M has finitely many connected components $M_1, \ldots, M_k$ we get $H^0_c(M) = H^0_c(M_1) \oplus \cdots \oplus H^0_c(M_k)$, so one usually considers only connected manifolds.

Proposition 8.4.1. *Let M be connected. If M is compact then $H^0_c(M) = \mathbb{R}$, while if M is not compact then $H^0_c(M) = 0$.*

Proof. The space H^0_c consists of all the compactly supported constant functions. Non-trivial such functions exist only if M is compact. □

As in the De Rham cohomology, we have $H^k_c(M) = 0$ for every $k > \dim M$.

8.4.3. The line. As usual we start by considering the line $\mathbb{R}$.

Proposition 8.4.2. *We have $H^1_c(\mathbb{R}) \cong \mathbb{R}$ and $H^k_c(\mathbb{R}) = 0$ for all $k \neq 1$.*

Proof. We already know that $H^k_c(\mathbb{R}) = 0$ for $k = 0$ and $k \geq 2$, so we turn to the case $k = 1$. The integration map

$$\int_{\mathbb{R}}\colon H^1_c(\mathbb{R}) \longrightarrow \mathbb{R}$$

is surjective. If $\omega = g(x)dx$ is such that $\int \omega = 0$, we may define $f(x) = \int_{-\infty}^x g(t)dt$ and get a compactly supported f with $\omega = df$. Therefore the integration map is also injective. □

We note that $H^i_c(\mathbb{R}) \cong H^{1-i}(\mathbb{R})$. This is not an accident, as we will see.

8.4.4. Functoriality? If $f\colon M \to N$ is a *proper* map, then the pull-back $f^*\omega$ of $\omega \in \Omega^k_c(N)$ is compactly supported also in M and we get a morphism

$$f^*\colon \Omega^k_c(N) \longrightarrow \Omega^k_c(M).$$

However, if f is not proper the pull-back is not defined in this context. So we can say that contravariant functoriality holds only for proper maps.

On the other hand, the compactly supported cohomology demonstrates some covariant behaviour: every inclusion map $i\colon U \hookrightarrow M$ of some open subset U induces the *extension morphism*

$$i_*\colon \Omega^k_c(U) \longrightarrow \Omega^k_c(M)$$

defined simply by extending k-forms to be zero outside of U. This does not work for general k-forms (extensions would not be smooth, nor continuous).

8.4.5. Integration along fibres. Let $\pi\colon M \to N$ be a submersion between oriented manifolds without boundary of dimension $m \geq n$.

For every $p \in N$ the fibre $F = \pi^{-1}(p)$ is a manifold of dimension $h = m-n$, with an orientation induced by that of M and N as follows: for every $p \in M$ we say that $v_1, \ldots, v_h \in T_pF$ is a positive basis if it may be completed to a positive basis $v_1, \ldots, v_m$ of T_pM such that $v_{h+1}, \ldots, v_m$ project to a positive basis of $T_{\pi(p)}N$.

We now define a map

$$\pi_*\colon \Omega^k_c(M) \longrightarrow \Omega^{k-h}_c(N)$$

called *integration along fibres*, as follows. For every $p \in N$ and $v_1, \ldots, v_{k-h} \in T_p(N)$ we set

$$\pi_*(\omega)(p)(v_1, \ldots, v_{k-h}) = \int_{\pi^{-1}(p)} \beta$$

where β is the k-form on the oriented k-submanifold $F = \pi^{-1}(p)$ defined as

$$\beta(q)(w_1, \ldots, w_h) = \omega(w_1, \ldots, w_h, \tilde{v}_1, \ldots, \tilde{v}_{k-h})$$

where $\tilde{v}_i$ is any vector in $T_q(F)$ such that $d\pi_q(\tilde{v}_i) = v_i$.

Proposition 8.4.3. *The form β is well-defined.*

Proof. For any other lift $\tilde{v}_i'$ we get $\tilde{v}_i' = \tilde{v}_i + \lambda_1 w_1 + \ldots + \lambda_h w_h$ and hence

$$\omega(w_1, \ldots, w_h, \ldots, \tilde{v}_i', \ldots) = \omega(w_1, \ldots, w_h, \ldots, \tilde{v}_i, \ldots)$$

since $\omega(w_1, \ldots, w_h, \ldots, \lambda_j w_j, \ldots) = 0$. □

Proposition 8.4.4. *The linear map π_* commutes with differentials and hence descends to a map in cohomology*

$$\pi_*\colon H^k_c(M) \longrightarrow H^{k-h}_c(N).$$

Proof. We must prove that $\pi_*(d\omega) = d\pi_*(\omega)$ for every $\omega \in H^k_c(M)$. Via some charts, the submersion π is locally like a projection

$$\pi\colon U \times V \longrightarrow U$$

where $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^h$ are open subsets. As a start, we suppose that the support of ω lies entirely in $U \times V$. We use variables $x_1, \ldots, x_n$ for U and $y_1, \ldots, y_h$ for V. We have

$$\omega = \sum_{I,J} f_{I,J} dx^I \wedge dy^J.$$

By linearity we may suppose

$$\omega = f dx^I \wedge dy^J.$$

If $J = \{1, \ldots, h\}$ we get

$$\pi_*(\omega) = \left(\int_V f(x,y) dy^J \right) dx^I$$

and hence

$$\begin{aligned} d\pi_*(\omega) &= \sum_{i=1}^h \frac{\partial}{\partial x_i} \left(\int_V f(x,y) dy^J \right) dx^i \wedge dx^I \\ &= \left(\int_V \sum_{i=1}^h \frac{\partial}{\partial x_i} f(x,y) dy^J \right) dx^i \wedge dx^I = \pi_* d(\omega). \end{aligned}$$

If $J \neq \{1, \ldots, h\}$ we get $\pi_*(\omega) = 0$ and also $\pi_*(d\omega) = 0$ (exercise).

For a general form $\omega \in \Omega^k_c(M)$, the compact support of ω may be covered by some r charts and one concludes with a partition of unity ρ_i since

$$d\pi_*(\omega) = \sum_{i=1}^r d\pi_*(\rho_i \omega) = \sum_{i=1}^r \pi_* d(\rho_i \omega) = \pi_* d\omega.$$

We have only used that d and π_* are linear. The proof is complete. □

We have discovered that every submersion $f\colon M \to N$ between oriented manifolds induces a linear map

$$\pi_*\colon H^k_c(M) \longrightarrow H^{k-h}_c(N).$$

The map π_* is called *integration along fibres*.

8.4.6. Smooth coverings. Let $M \to N$ be a smooth covering between smooth n-manifolds. A covering is a submersion, and the integration along fibres is a map

$$\pi_*\colon H^k_c(M) \longrightarrow H^k_c(N).$$

In this case the integration along the fibres is just a summation, that is

$$\pi_*(\omega)(p)(v_1, \ldots, v_n) = \sum_{\pi(q)=p} \omega(q)(\tilde{v}_1, \ldots, \tilde{v}_n)$$

where $v_i \in T_pN$ and $\tilde{v}_i = d\pi_q^{-1}(v_i)$. Here is a remarkable application.

Proposition 8.4.5. *If $\pi\colon M \to N$ is a covering of finite degree d, then $\pi^*\colon H_c^k(N) \to H_c^k(M)$ is injective.*

Proof. We have $\frac{1}{d}\pi_* \circ \pi^* = \mathrm{id}$ on $H_c^k(N)$. □

If the covering has infinite degree the maps in cohomology need not to be injective, as the universal covering $\mathbb{R} \to S^1$ easily shows.

8.4.7. Poincaré Lemma. We now prove the appropriate version of the Poincaré Lemma for $H_c^k(\mathbb{R}^n)$.

Let M be a smooth manifold. Let $\eta \in \Omega_c^1(\mathbb{R})$ have $\int \eta = 1$, so that in particular it generates $H_c^1(\mathbb{R}) = \mathbb{R}$. Consider the morphism

$$\begin{aligned} \iota\colon H_c^k(M) &\longrightarrow H_c^{k+1}(M \times \mathbb{R}) \\ \omega &\longmapsto \omega \wedge \eta. \end{aligned}$$

Lemma 8.4.6. *This morphism is an isomorphism.*

Proof. We consider the projection

$$\pi\colon M \times \mathbb{R} \longrightarrow M.$$

By integrating along fibres we get a map

$$\pi_*\colon H_c^{k+1}(M \times \mathbb{R}) \longrightarrow H_c^k(M).$$

We want to show that π_* inverts ι. We have $\pi_* \circ \iota = \mathrm{id}$ already in $\Omega_c^k(M)$. On forms, we have $\iota \circ \pi_* \neq \mathrm{id}$ and we construct a chain homotopy to prove that $\iota \circ \pi_* = \mathrm{id}$ in cohomology. We need a map

$$h\colon \Omega_c^k(M \times \mathbb{R}) \longrightarrow \Omega_c^{k-1}(M \times \mathbb{R}).$$

The map is defined as follows:

$$\begin{aligned} (h\omega)(p,t)(v_1,\ldots,v_{k-1}) = &\int_{-\infty}^t \omega(p,u)\left(\frac{\partial}{\partial t}, v_1, \ldots, v_{k-1}\right) du \\ &- E(t)\int_{\mathbb{R}} \omega(p,u)\left(\frac{\partial}{\partial t}, v_1, \ldots, v_{k-1}\right) du \end{aligned}$$

where

$$\eta = e(t)dt, \qquad E(t) = \int_{-\infty}^t e(u)du.$$

We now prove that

$$dh + hd = \mathrm{id} - \iota \circ \pi_*. \tag{17}$$

This will conclude the proof. Since this is a local property, we pick a chart and use coordinates $x_1, \ldots, x_n, t$. By linearity, there are two cases to consider:

(1) $\omega = f dx^I$,
(2) $\omega = g dt \wedge dx^J$.

We get

$$(\iota \circ \pi_*)(f\,dx^I) = 0,$$
$$(\iota \circ \pi_*)(g\,dt \wedge dx^J) = \left(\int_{\mathbb{R}} g(p,u)du\right) dx^J \wedge \eta.$$

The map h sends the forms of type (1) to zero, and those of type (2) to

$$h(g\,dt \wedge dx^J) = \left(\int_{-\infty}^{t} g(p,u)du - E(t)\int_{\mathbb{R}} g(p,u)du\right) dx^J.$$

Here are the two cases:

(1) If $\omega = f\,dx^I$ we get

$$\begin{aligned}(dh+hd)(\omega) &= hd\omega = h(df \wedge dx^I) = h\left(\frac{\partial f}{\partial t}dt \wedge dx^I\right)\\ &= \left(\int_{-\infty}^{t} \frac{\partial f}{\partial t}(p,u)du - E(t)\int_{\mathbb{R}} \frac{\partial f}{\partial t}(p,u)du\right) dx^I\\ &= f(p,t)dx^I = \omega,\\ (\mathrm{id} - \iota \circ \pi_*)(\omega) &= \omega.\end{aligned}$$

(2) If $\omega = g\,dt \wedge dx^J$ we get

$$\begin{aligned}dh(\omega) &= d\left(\int_{-\infty}^{t} g(p,u)du - E(t)\int_{\mathbb{R}} g(p,u)du\right) dx^J\\ &= \omega + \sum_{j=1}^{n}\left(\int_{-\infty}^{t} \frac{\partial g}{\partial x_j}(p,u)du\right) dx^j \wedge dx^J\\ &\quad - \left(\int_{\mathbb{R}} g(p,u)du\right) \eta \wedge dx^J\\ &\quad - E(t)\sum_{j=1}^{n}\left(\int_{\mathbb{R}} \frac{\partial g}{\partial x_j}(p,u)du\right) dx^j \wedge dx^J,\\ hd(\omega) &= \sum_{i=1}^{n} h\left(\frac{\partial g}{\partial x_i}dx^i \wedge dt \wedge dx^J\right)\\ &= -\sum_{j=1}^{n}\left(\int_{-\infty}^{t} \frac{\partial g}{\partial x_j}(p,u)du\right) dx^j \wedge dx^J\\ &\quad + E(t)\sum_{j=1}^{n}\left(\int_{\mathbb{R}} \frac{\partial g}{\partial x_j}(p,u)du\right) dx^j \wedge dx^J,\\ (dh+hd)(\omega) &= \omega - \left(\int_{\mathbb{R}} g(p,u)du\right) \eta \wedge dx^J,\\ (\mathrm{id} - \iota \circ \pi_*)(\omega) &= \omega - \left(\int_{\mathbb{R}} g(p,u)du\right) \eta \wedge dx^J.\end{aligned}$$

The proof is complete. $\square$

As a corollary, we can compute the compactly supported cohomology of Euclidean spaces. This result is also known as the Poincaré Lemma.

Corollary 8.4.7. *We have $H_c^n(\mathbb{R}^n) = \mathbb{R}$ and $H_c^k(\mathbb{R}^n) = 0$ for all $k \neq n$.*

We keep observing that $H_c^k(\mathbb{R}^n) = H^{n-k}(\mathbb{R}^n)$ for all n and k. We also note that the compactly supported cohomology is evidently *not* invariant under homotopy equivalence.

8.4.8. The Mayer – Vietoris sequence. Proving the Poincaré Lemma in this compactly supported context was not easy; on the other hand the Mayer – Vietoris sequence is almost straightforward.

Let M be a smooth manifold, and $U, V \subset M$ be two open subsets covering M. The inclusions

$$\begin{tikzcd} & U \arrow[dr, "l"] & \\ U \cap V \arrow[ur, "i"] \arrow[dr, "j"'] & & M \\ & V \arrow[ur, "m"'] & \end{tikzcd}$$

induce the extension morphisms in cohomology

$$\begin{tikzcd} & H_c^k(U) \arrow[dr, "l_*"] & \\ H_c^k(U \cap V) \arrow[ur, "i_*"] \arrow[dr, "j_*"'] & & H_c^k(M) \\ & H_c^k(V) \arrow[ur, "m_*"'] & \end{tikzcd}$$

Theorem 8.4.8 (Mayer – Vietoris Theorem). *There is an exact sequence*

$$\cdots \longrightarrow H_c^k(U \cap V) \xrightarrow{(-i_*, j_*)} H_c^k(U) \oplus H_c^k(V) \xrightarrow{l_* + m_*} H_c^k(M) \xrightarrow{\delta} H_c^{k+1}(U \cap V) \longrightarrow \cdots$$

for some canonically defined map δ.

Proof. The sequence of complexes

$$0 \longrightarrow \Omega_c^*(U \cap V) \xrightarrow{(-i_*, j_*)} \Omega_c^*(U) \oplus \Omega_c^*(V) \xrightarrow{l_* + m_*} \Omega_c^*(M) \longrightarrow 0$$

is easily seen to be exact: use a partition of unity to show that $l_* + m_*$ is surjective. $\square$

Note that this Mayer – Vietoris sequence is different in nature from the one that we obtained from Theorem 8.3.4.

Exercise 8.4.9. Use the Mayer – Vietoris sequence to confirm that

$$H^0_c(S^n) = H^0(S^n) = \mathbb{R}, \qquad H^n_c(S^n) = H^n(S^n) = \mathbb{R},$$
$$H^k_c(S^n) = H^k(S^n) = 0 \text{ if } k \neq 0, n.$$

We cannot refrain from noting again that $H^k_c(S^n) = H^{n-k}(S^n)$. As in ordinary De Rham cohomology, we can write δ explicitly. Let ρ_U, ρ_V be a partition of unity subordinate to U, V. Given $\omega \in H^k_c(M)$ we can define

$$\eta = d\rho_V \wedge \omega = -d\rho_U \wedge \omega \in H^{k+1}_c(U \cap V).$$

Exercise 8.4.10. We have $\delta(\omega) = \eta$.

8.4.9. Countably many connected components. We end this section by pointing out another difference between $H^k(M)$ and $H^k_c(M)$.

Exercise 8.4.11. Let M have countably many connected components M_1, $M_2, \ldots$ We have

$$H^k(M) = \prod_i H^k(M_i), \qquad H^k_c(M) = \bigoplus_i H^k_c(M_i).$$

Remember that $\prod_i V_i$ is the space of all sequences $(v_1, v_2, \ldots)$ while $\oplus_i V_i$ is the subspace of all sequences having only finitely many non-zero elements.

8.5. Poincaré duality

We have already noted that $H^k(M) \cong H^{n-k}_c(M)$ on many n-manifolds M, and we now prove this equality in a much wider generality.

We stress the fact that all the manifolds considered in this section have no boundary!

8.5.1. The Poincaré bilinear map. Let M be an oriented smooth manifold without boundary. We define the *Poincaré bilinear map*

$$H^k(M) \times H^{n-k}_c(M) \longrightarrow \mathbb{R}$$

by sending the pair (ω, η) to the real number

$$\langle \omega, \eta \rangle = \int_M \omega \wedge \eta.$$

The map is well-defined since $\omega \wedge \eta$ has compact support. As every bilinear form, it induces a map

$$\mathrm{PD}\colon H^k(M) \longrightarrow H^{n-k}_c(M)^*$$

that sends ω to the functional $\eta \mapsto \langle \omega, \eta \rangle$. We dedicate this section to proving the following.

Theorem 8.5.1 (Poincaré duality). *The map* PD *is an isomorphism.*

As usual, we will need a bit of homological algebra.

8.5.2. The Five Lemma. The following lemma is solved by diagram chasing, and we leave it to the reader as an exercise – there is certainly much more fun in trying to solve it alone than in reading a boring sequence of implications.

Exercise 8.5.2 (The Five Lemma). Given the following commutative diagram of abelian groups and morphisms

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "\alpha"] & B \arrow[r, "g"] \arrow[d, "\beta"] & C \arrow[r, "h"] \arrow[d, "\gamma"] & D \arrow[r, "i"] \arrow[d, "\delta"] & E \arrow[d, "\epsilon"] \\
A' \arrow[r, "j"'] & B' \arrow[r, "k"'] & C' \arrow[r, "l"'] & D' \arrow[r, "m"'] & E'
\end{tikzcd}$$

in which the rows are exact, if $\alpha, \beta, \delta, \epsilon$ are isomorphisms then γ also is.

8.5.3. Induction on open subsets. Let M be a smooth manifold. We want to prove the Poincaré duality Theorem by induction on open subsets of M, starting with those diffeomorphic to $\mathbb{R}^n$ and then passing to more complicated ones in a controlled way. We will need the following.

Let $\mathcal{A}$ be the collection of open subsets in M determined by the rules:

- $\mathcal{A}$ contains all the open subsets diffeomorphic to $\mathbb{R}^n$,
- if $U, V, U \cap V \in \mathcal{A}$, then $U \cup V \in \mathcal{A}$,
- if $U_i \in \mathcal{A}$ are pairwise disjoint, then $\cup U_i \in \mathcal{A}$.

Note that in the last point there can be infinitely many disjoint sets U_i (they are always countable, since M is second countable).

Lemma 8.5.3. *We have $M \in \mathcal{A}$.*

Proof. The proof is subdivided into steps.

(1) If $U_1, \ldots, U_k \in \mathcal{A}$ and all their intersections lie in $\mathcal{A}$, then also $U_1 \cup \cdots \cup U_k \in \mathcal{A}$.
(2) If $\{U_i\} \subset \mathcal{A}$ is a locally finite countable family, with $\overline{U_i}$ compact for all i, and such that all the finite intersections also lie in $\mathcal{A}$, then $\cup U_i \in \mathcal{A}$.
(3) If $U \subset M$ is diffeomorphic to an open subset $V \subset \mathbb{R}^n$, then $U \in \mathcal{A}$.
(4) $M \in \mathcal{A}$.

Point (1) is a simple exercise (prove it by induction on k). Concerning (2), we may suppose that $U = \cup U_i$ is connected, and note that every U_i intersects only finitely many U_j.

We define some new open subsets by setting $W_0 = U_0$ and defining W_{i+1} as the union of all the U_j that intersect W_i and are not contained in $\cup_{a \leq i} W_a$. Every W_i contains finitely many U_j and hence $W_i \in \mathcal{A}$ by (1). Note that $W_i \cap W_{i+2} = \varnothing$ for all i. We set

$$Z_0 = \sqcup_i W_{2i}, \qquad Z_1 = \sqcup_i W_{2i+1}.$$

We have $Z_0, Z_1 \in \mathcal{A}$ and also $Z_0 \cap Z_1 \in \mathcal{A}$, so $U = Z_0 \cup Z_1 \in \mathcal{A}$.

About (3), we note that V is covered by products $(a_1, b_1) \times \cdots \times (a_n, b_n)$ whose closure is contained in V. Every finite intersection is again a product, so all these sets and their intersections are diffeomorphic to $\mathbb{R}^n$ and hence lie in $\mathcal{A}$. This covering can be made locally finite using an exhaustion of V by compact sets. Now (2) applies and we get $U \in \mathcal{A}$.

Finally, by taking an adequate atlas for M (see Proposition 3.3.1) we find a locally finite covering U_i such that every U_i is diffeomorphic to $\mathbb{R}^n$ and has compact closure. The intersections are diffeomorphic to open subsets of $\mathbb{R}^n$ and hence are in $\mathcal{A}$ by (3). We conclude again by (2). □

We have also proved that every open subset of M is contained in $\mathcal{A}$.

8.5.4. Proof of the Poincaré duality. We can now prove Theorem 8.5.1.

Proof. Let $\mathcal{B}$ be the collection of the open subsets U of M where Poincaré duality holds. Our aim is of course to prove that $M \in \mathcal{B}$.

If $U \cong \mathbb{R}^n$, then $U \in \mathcal{B}$. Indeed, we only have to prove that $\mathrm{PD}\colon H^0(\mathbb{R}^n) \to H^n_c(\mathbb{R}^n)^*$ is an isomorphism. Both spaces have dimension one, so it suffices to check that the map is not trivial: if η is a compactly supported n-form over $\mathbb{R}^n$ with $\int \eta = 1$ and 1 is the constant function we get $\langle 1, \eta\rangle = 1$ and hence $1 \in H^0(\mathbb{R}^n)$ is mapped to a nontrivial element $\mathrm{PD}(1) \in H^n_c(\mathbb{R}^n)^*$.

If $U, V, U\cap V \in \mathcal{B}$, then $U \cup V \in \mathcal{B}$. To show this, we consider the following diagram that contains both Mayer – Vietoris sequences:

$$\begin{tikzcd}
\arrow[r] & H^{k-1}(U\cap V) \arrow[r] \arrow[d, "\mathrm{PD}"'] & H^k(U\cup V) \arrow[r] \arrow[d, "\mathrm{PD}"'] & H^k(U)\oplus H^k(V) \arrow[r] \arrow[d, "\mathrm{PD}"'] & {} \\
\arrow[r] & H^{n-k+1}_c(U\cap V)^* \arrow[r] & H^{n-k}_c(U\cup V)^* \arrow[r] & H^{n-k}_c(U)^*\oplus H^{n-k}_c(V)^* \arrow[r] & {}
\end{tikzcd}$$

The bottom row is obtained by dualising the Mayer – Vietoris exact sequence in the compactly supported cohomology. We leave as an exercise to show that this diagram commutes up to sign (use Proposition 8.3.5 and Exercise 8.4.10). By the Five Lemma, if PD is an isomorphism for U, V, and $U \cap V$, then it is so also for $U \cup V$.

If $U = \sqcup_i U_i$ and $U_i \in \mathcal{B}$, then $U \in \mathcal{B}$. This is a consequence of Exercise 8.4.11 and of the natural equality $(\oplus_i V_i)^* = \prod_i V_i^*$.

By Proposition 8.5.3 we have $M \in \mathcal{B}$ and we are done. □

8.5.5. Betti numbers. As a first consequence of Poicaré Duality, for every orientable manifold M we have

$$\dim H^k(M) = \dim H^{n-k}_c.$$

When M is compact, this becomes

$$b^k = \dim H^k(M) = \dim H^{n-k}(M) = b^{n-k}.$$

In particular we have $b^0 = b^n = 1$. In fact we can prove that all these numbers are finite.

Proposition 8.5.4. *If M is compact then b^k is finite.*

Proof. If M is orientable, we have the canonical Poincaré isomorphisms

$$H^k(M) \cong H^{n-k}(M)^*, \qquad H^{n-k}(M) \cong H^k(M)^*.$$

By combining them we deduce that the canonical embedding $H^k(M) \hookrightarrow H^k(M)^{**}$ is an isomorphism, and we know that this holds if and only if the vector space is finite-dimensional.

If M is non-orientable, it has an orientable double cover and we conclude using Proposition 8.4.5. $\square$

Proposition 8.5.5. *If M is compact orientable and n is odd, then $\chi(M) = 0$.*

Proof. We have $b^i = b^{n-i}$, so everything cancels. $\square$

8.5.6. Orientability. We now show that cohomology distinguishes between orientable and non-orientable manifolds. Let M be a connected smooth n-manifold.

Proposition 8.5.6. *If M is oriented, the map*

$$\int_M : H^n_c(M) \longrightarrow \mathbb{R}$$

is an isomorphism.

Proof. We have $\mathbb{R} = H^0(M) = H^n_c(M)^* = H^0(M)^*$ so $H^n_c(M) \cong \mathbb{R}$. Moreover $\int_M$ is surjective. $\square$

Proposition 8.5.7. *We have*

$$H^n_c(M) = \begin{cases} \mathbb{R} & \text{if } M \text{ is orientable,} \\ 0 & \text{otherwise.} \end{cases}$$

Proof. If M is not orientable, it has an orientable double cover $\pi\colon \tilde{M} \to M$, with orientation-reversing deck involution $\iota\colon \tilde{M} \to \tilde{M}$. The induced map

$$\pi^*\colon H^n_c(M) \to H^n_c(\tilde{M})$$

is injective by Proposition 8.4.5. Moreover, for every n-form $\omega \in \Omega^n(M)$, the pull-back $\pi^*\omega$ is ι-invariant, but since ι reverses the orientation of $\tilde{M}$ we get

$$\int_{\tilde{M}} \pi^*\omega = \int_{-\tilde{M}} \iota^*\pi^*\omega = -\int_{\tilde{M}} \pi^*\omega.$$

Hence this integral vanishes, and by the previous proposition we get $\pi^*\omega = 0$ in cohomology. Since π^* is injective, we get $H^n_c(M) = 0$. $\square$

8.5.7. Real projective spaces. We can now easily calculate the De Rham cohomology of $\mathbb{RP}^n$.

Proposition 8.5.8. *We have* $H^0(\mathbb{RP}^n) = \mathbb{R}$, $H^k(\mathbb{RP}^n) = 0\ \forall k \neq 0, n$, *and*

$$H^n(\mathbb{RP}^n) = \begin{cases} \mathbb{R} & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

Proof. This works for every manifold M that is covered by S^n. Since the pull-back $\pi^*\colon H^k(M) \to H^k(S^n)$ is injective, the only indeterminacy is for $k = n$ and is determined by whether M is orientable or not. □

The proof also shows the following. Remember the lens spaces $L(p, q)$.

Corollary 8.5.9. *We have*

$$H^0\big(L(p,q)\big) = H^3\big(L(p,q)\big) = \mathbb{R}, \qquad H^1\big(L(p,q)\big) = H^2\big(L(p,q)\big) = 0.$$

8.5.8. Signature. If M is an oriented compact manifold of even dimension $2n$, Poincaré duality furnishes a non-degenerate bilinear form

$$H^n(M) \times H^n(M) \longrightarrow \mathbb{R}$$

that is symmetric or antisymmetric, according to whether n is even or odd. This is because of the formula $\omega \wedge \eta = (-1)^{n^2} \eta \wedge \omega$.

When M has dimension $4m$, the non-degenerate bilinear form on H^{2m} is symmetric and hence has a *signature* (p, m), see Section 2.3.1. The *signature* of M is the integer

$$\sigma(M) = p - m.$$

A nice feature of this invariant is that it reacts to orientation reversals.

Proposition 8.5.10. *We have* $\sigma(-M) = -\sigma(M)$

Proof. We have $\int_M \omega = -\int_{-M} \omega$, hence the orientation reversal modifies the bilinear form by a sign and its signature changes from (p, m) to (m, p). □

Recall that an orientable manifold M is mirrorable if it has an orientation-reversing diffeomorphism.

Corollary 8.5.11. *A mirrorable orientable* $4m$*-manifold* M *has* $\sigma(M) = 0$.

We deduce that for every $m \geq 1$ the manifold $\mathbb{CP}^{2m}$ is not mirrorable: its middle Betti number is $b^{2m} = 1$ and hence its signature is $\sigma = \pm 1$. In particular the complex projective plane $\mathbb{CP}^2$ is not mirrorable (while the complex projective line $\mathbb{CP}^1 \cong S^2$ is mirrorable).

8.5.9. The Künneth formula. We now prove an elegant formula that relates the cohomology of a product $M \times N$ with the cohomologies of the factors. This formula is known as the *Künneth formula*.

Let M and N be two smooth manifolds. The two projections

$$\pi_M \colon M \times N \longrightarrow M, \qquad \pi_N \colon M \times N \longrightarrow N$$

give rise to a bilinear map

$$\begin{aligned} \Omega^k(M) \times \Omega^h(N) &\longrightarrow \Omega^{k+h}(M \times N) \\ (\omega, \eta) &\longmapsto \pi_M^* \omega \wedge \pi_N^* \eta \end{aligned}$$

that passes to a bilinear map

$$H^k(M) \times H^h(N) \longrightarrow H^{k+h}(M \times N).$$

By the universal property of tensor products, this induces a linear map

$$H^k(M) \otimes H^h(N) \longrightarrow H^{k+h}(M \times N).$$

These linear maps when k and h vary can be grouped altogether as

$$\Psi \colon H^*(M) \otimes H^*(N) \longrightarrow H^*(M \times N).$$

We will henceforth suppose that the Betti numbers of N are all finite: this holds for instance if N is compact, but also for many other manifolds.

Theorem 8.5.12 (Künneth's formula). *The map Ψ is an isomorphism.*

Before entering into the proof, we note that this implies that

$$H^k(M \times N) \cong \bigoplus_{p+q=k} H^p(M) \otimes H^q(N).$$

Proof. As in the proof of Poincaré Duality, we define $\mathcal{B}$ to be the set of all the open subsets $U \subset M$ such that the theorem holds for the product $U \times N$. Our aim is to show that $M \in \mathcal{B}$.

If $U \cong \mathbb{R}^n$, this is the Poincaré Lemma, more specifically Lemma 8.2.3.

If $U, V, U \cap V \in \mathcal{B}$, then $U \cup V \in \mathcal{B}$. To show this, we fix $k \geq 0$, pick $p \leq k$ and consider the Mayer – Vietoris sequence

$$\cdots \longrightarrow H^{p-1}(U \cap V) \longrightarrow H^p(U \cup V) \longrightarrow H^p(U) \oplus H^p(V) \longrightarrow \cdots$$

If we tensor it with $H^{k-p}(N)$ and sum over $p = 0, \ldots, k$ we still get an exact sequence by Exercise 8.3.1. Here it is:

$$\begin{gathered} \cdots \longrightarrow \oplus_{p=0}^k \big(H^{p-1}(U \cap V) \otimes H^{k-p}(N)\big) \longrightarrow \oplus_{p=0}^k \big(H^p(U \cup V) \otimes H^{k-p}(N)\big) \\ \longrightarrow \oplus_{p=0}^k \big(H^p(U) \otimes H^{k-p}(N)\big) \oplus \oplus_{p=0}^k \big(H^p(V) \otimes H^{k-p}(N)\big) \longrightarrow \cdots \end{gathered}$$

We now send via Ψ this sequence to the Mayer – Vietoris sequence for $M \times N$:

$$\cdots \to H^{k-1}\big((U \cap V) \times N\big) \to H^k\big((U \cup V) \times N\big) \to H^k(U \times N) \otimes H^k(V \times N) \to \cdots$$

The resulting diagram commutes (exercise) and has two exact rows. Using the Five Lemma we conclude that $U \cup V \in \mathcal{B}$.

If $U = \sqcup_i U_i$ and $U_i \in \mathcal{B}$, then $U \in \mathcal{B}$. This is a consequence of Exercise 2.1.16 and of the fact that $\dim H^p(N) < \infty$ for all p.

By Proposition 8.5.3 we have $M \in \mathcal{B}$ and we are done. □

Remark 8.5.13. When $M = N = \mathbb{Z}$, the map Ψ is not an isomorphism (exercise). We really need one of the factor to have finite-dimensional cohomology here.

Corollary 8.5.14. *Let M and N be manifolds with finite cohomology (for instance, they are compact). For every k we have:*

$$b^k(M \times N) = \sum_{i=0}^{k} b^i(M) b^{k-i}(N).$$

Corollary 8.5.15. *The torus $T = S^1 \times S^1$ has Betti numbers*

$$b^0 = 1, \qquad b^1 = 2, \qquad b^2 = 1.$$

Exercise 8.5.16. The Betti numbers of $T^n = \underbrace{S^1 \times \cdots \times S^1}_{n}$ are

$$b^k(T^n) = \binom{n}{k}.$$

Exercise 8.5.17. The Betti numbers of $S^2 \times S^2$ are

$$b^0 = 1, \quad b^1 = 0, \quad b^2 = 2, \quad b^3 = 0, \quad b^4 = 1.$$

We deduce from the exercise that the compact four-manifolds

$$S^4, \qquad \mathbb{CP}^2, \qquad S^2 \times S^2$$

are pairwise not homotopy equivalent (although they are all simply connected) because their second Betti number is respectively 0, 1, and 2.

Exercise 8.5.18. If M and N are manifolds with finite Betti numbers, then

$$\chi(M \times N) = \chi(M) \cdot \chi(N).$$

8.5.10. Connected sums. The following exercises can be solved using the Mayer – Vietoris sequence carefully.

Exercise 8.5.19. Let M be a smooth connected n-manifold without boundary and N be obtained from M by removing a point. We have:

$$b^i(N) = b^i(M) \quad \forall i \leq n-2$$

$$b^{n-1}(N) = \begin{cases} b^{n-1}(M) & \text{if } M \text{ is compact and oriented,} \\ b^{n-1}(M) + 1 & \text{otherwise,} \end{cases}$$

$$b^n(N) = \begin{cases} b^n(M) - 1 & \text{if } M \text{ is compact and oriented,} \\ b^n(M) & \text{otherwise,} \end{cases}$$

Hint. Use the Mayer – Vietoris sequence with $M = U \cup V$, $U = N$, and V an open ball containing the removed point. □

Note that in all cases we get $\chi(N) = \chi(M) - 1$ when they are defined.

Exercise 8.5.20. Let $M \# N$ be the connected sum of two oriented connected compact manifolds M and N without boundary. We have

$$\begin{aligned} b^i(M\#N) &= 1 \quad \text{for } i = 0, n, \\ b^i(M\#N) &= b^i(M) + b^i(N) \quad \text{for } 0 < i < n. \end{aligned}$$

We can finally calculate the cohomology of a genus-g surface S_g.

Corollary 8.5.21. *The Betti numbers of S_g are*

$$b^0 = 1, \qquad b^1 = 2g, \qquad b^2 = 1.$$

Therefore $\chi(S_g) = 2 - 2g$.

8.6. Intersection theory

We now combine transversality and De Rham cohomology to build a geometric theory on submanifolds called *intersection theory*.

As in the previous section, all the manifolds considered here are without boundary. We will be mostly interested in compact ones.

8.6.1. Poincaré dual of an oriented subsurface.

Let M be an oriented compact connected smooth n-manifold without boundary. Let $S \subset M$ be an oriented compact k-dimensional submanifold. We have already observed that integration along S yields a linear map

$$\int_S : H^k(M) \longrightarrow \mathbb{R}.$$

By Poincaré Duality, this linear map corresponds to some cohomology element $\omega_S \in H^{n-k}(M)$ called the *Poincaré dual* of S, characterised by the equality

$$\int_M \omega_S \wedge \eta = \int_S \eta$$

for every $\eta \in H^k(M)$. We have just discovered that we can naturally transform oriented compact submanifolds S into cohomology classes ω_S. For example:

- the Poincaré dual of M itself is $\omega_M = 1 \in H^0(M) = \mathbb{R}$,
- the Poincaré dual of a point $p \in M$ is $\omega_p = 1 \in H^n(M) = \mathbb{R}$.

We now want to construct the $(n-k)$-form ω_S explicitly. To this purpose we consider vector bundles.

8.6.2. Thom forms. Let $\pi\colon E \to N$ be an oriented rank-r vector bundle over a connected compact n-manifold N. Consider a closed form $\omega \in \Omega^r_c(E)$.

Proposition 8.6.1. *The integral*
$$\int_{E_p} \omega$$
is independent of $p \in N$.

Proof. Two points $p, q \in N$ are connected by an embedded arc α, and $\pi^{-1}(\alpha)$ is a manifold with boundary $E_p \cup E_q$. Use Stokes. $\square$

The closed form $\omega \in \Omega^r_c(E)$ is a *Thom form* if
$$\int_{E_p} \omega = 1.$$

Proposition 8.6.2. *Thom forms exist.*

Proof. We pick
$$\eta(x) = \rho\big(\|x\|^2\big)\, dx^1 \wedge \cdots \wedge dx^r \in \Omega^r(\mathbb{R}^r)$$
where ρ is non-negative and compactly supported, rescaled so that $\int_{\mathbb{R}^r} \eta = 1$. We fix a Riemannian metric on E. On a trivialising neighbourhood U the bundle is isometric to $U \times \mathbb{R}^r$ and we equip it with the closed form $\pi_2^*\eta$ where π_2 is the projection onto $\mathbb{R}^r$. Since η is $\mathrm{O}(r)$-invariant, all these r-forms match to a Thom form ω in E. $\square$

We consider as usual N embedded in E via the zero-section $i\colon N \hookrightarrow E$. Here is the reason why we are interested in Thom forms:

Proposition 8.6.3. *If* $\omega \in \Omega^r_c(E)$ *is a Thom form, then*
$$\int_E \omega \wedge \eta = \int_N \eta$$
for every closed form $\eta \in \Omega^n(E)$.

Proof. The map $i \circ \pi\colon E \to E$ is homotopic to the identity, hence in cohomology we get $[\eta] = (i \circ \pi)^*[\eta]$ and therefore $\eta = \pi^* i^* \eta + d\phi$. Then
$$\int_E \omega \wedge \eta = \int_E \omega \wedge \pi^* i^* \eta + \int_E \omega \wedge d\phi.$$
The second addendum vanishes because $\omega \wedge d\phi = \pm d(\omega \wedge \phi)$ and Stokes applies. We study the first addendum locally. On a trivialising chart $U \to V$ the bundle is like $V \times \mathbb{R}^r$ with $V \subset \mathbb{R}^m$. We use the variables x^i and y^j for $\mathbb{R}^m$ and $\mathbb{R}^r$. We have
$$\pi^* i^* \eta = \sum_I f^I(x) dx^I.$$

This gives
$$\int_{V\times\mathbb{R}^r}\omega\wedge\eta=\int_V\left(\int_{\mathbb{R}^r}\omega\right)\sum_I f^I(x)=\int_V\eta$$
because ω is a Thom form, and therefore
$$\int_E\omega\wedge\eta=\int_N\eta.$$
The proof is complete. □

We now turn back to our oriented compact connected n-manifold M and compact oriented k-submanifold $S\subset M$. Let $\nu S\subset M$ be any tubular neighbourhood. Every Thom form in νS is compactly supported and hence extends to a form in M, thus representing an element in $H^{n-k}(M)$.

Corollary 8.6.4. *Any Thom form in νS represents the Poincaré dual ω_S.*

Proof. Let ω be a Thom form in νS. For every closed $\eta\in\Omega^k(M)$ we get
$$\int_M\omega\wedge\eta=\int_E\omega\wedge\eta=\int_S\eta.$$
The proof is complete. □

Summing up, the Poincaré dual of a submanifold $S\subset M$ may be represented as a $(n-k)$-form supported in an arbitrarily small tubular neighbourhood of S, that gives 1 when integrated along any fibre: we should think at this as a kind of "bump form" concentrated near S.

8.6.3. Transverse intersection. Let N be an oriented connected compact manifold, and let $M,W\subset N$ be two oriented compact transverse submanifolds. Recall that $X=M\cap W$ is also a submanifold with $\operatorname{codim}X=\operatorname{codim}M+\operatorname{codim}W$. We also have
$$\nu X=\nu M\oplus\nu W.$$
The manifold X is naturally oriented: the bundles νM and νW are oriented, and hence so is the bundle νX and finally the manifold X.

The following proposition is the core of intersection theory: it shows that, via Poincaré duality, transverse intersection of oriented submanifolds corresponds to wedge products of forms:

Proposition 8.6.5. *We have $\omega_X=\omega_M\wedge\omega_W$.*

Proof. If ω_M,ω_W are Thom forms in $\nu M,\nu W$, the wedge product $\omega_M\wedge\omega_W$ in a Thom form in $\nu X=\nu M\oplus\nu W$. □

Example 8.6.6. Let $S,T\subset\mathbb{CP}^n$ be two transverse projective subspaces, of complex codimension s and t. Their intersection is a projective subspace $X=S\cap T$ of complex codimension $s+t$. All these are naturally oriented and their Poincaré dual forms are
$$\omega_S\in H^{2s}(\mathbb{CP}^n)=\mathbb{R},\qquad\omega_T\in H^{2t}(\mathbb{CP}^n)=\mathbb{R},\qquad\omega_X\in H^{2s+2t}(\mathbb{CP}^n)=\mathbb{R}.$$

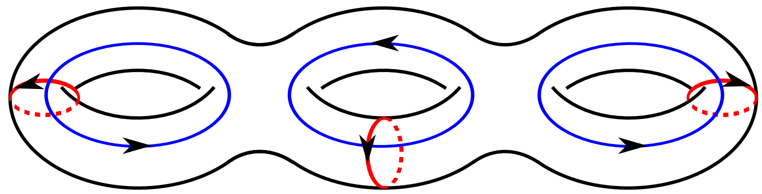

Figure 8.1. A symplectic basis for $H^1(S_3) \cong \mathbb{R}^6$ consists of the Poincaré duals of the oriented curves $\alpha_1, \alpha_2, \alpha_3$ (red) and $\beta_1, \beta_2, \beta_3$ (blue).

The proposition says that

$$\omega_X = \omega_S \wedge \omega_T.$$

If $s + t = n$ then X is a point and therefore $\omega_X = 1$. This shows in particular that the class ω_S is non-trivial, and is hence a generator of $H^{2s}(\mathbb{CP}^n)$.

8.6.4. Algebraic intersection. Let N and $M, W \subset N$ be as above. The case where M and W have complementary dimension is of particular interest. Here $X = M \cap W$ is a collection of oriented points p, each equipped with a sign ± 1 depending on whether the orientation of $T_pM \oplus T_pW$ matches with that of T_pN. We define the *algebraic intersection* $i(M, W)$ of M and W to be the sum of these values ± 1.

The n-form $\omega_M \wedge \omega_W \in H^n_c(N) = \mathbb{R}$ may be considered canonically as a real number. Proposition 8.6.5 says that

$$i(M, W) = \omega_M \wedge \omega_W.$$

This relation is of the highest importance when N has even dimension $2k$ and $\dim M = \dim W = k$, because it furnishes a concrete way to represent and calculate the intersection form in $H^k(N)$.

Example 8.6.7. We examine the genus-g surface S_g. The intersection form on $H^1(S_g) \cong \mathbb{R}^{2g}$ is non-degenerate and antisymmetric. Consider the $2g$ oriented curves $\alpha_1, \ldots, \alpha_g, \beta_1, \ldots, \beta_g$, shown in Figure 8.1. Their algebraic intersections are

$$i(\alpha_i, \alpha_j) = i(\beta_i, \beta_j) = 0 \ \forall i \neq j, \quad i(\alpha_i, \beta_j) = \delta_{ij}.$$

The intersection form on their dual $2g$ classes is antisymmetric, and hence it forms the antisymmetric matrix $J = \left(\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right)$. Since J is an invertible matrix, we can deduce by elementary linear algebra that these $2g$ classes form a basis of $H^1(S_g)$. A basis with such an intersection matrix is called a *symplectic basis*.

8.6.5. Homotopy invariance. Let M be an oriented connected compact n-manifold. The Poincaré dual may in fact be defined not only for submanifolds, but also for every smooth map $f\colon S \to M$ where S is a k-dimensional oriented manifold. Every such map f induces a linear functional

$$H^k(M) \longrightarrow \mathbb{R}$$
$$\eta \longmapsto \int_S f^*\eta$$

which is by Poincaré Duality an element $\omega_f \in H^{n-k}(M)$. Two homotopic maps $f, g\colon S \to M$ induce the same functional $\omega_f = \omega_g$. In particular, we get:

Corollary 8.6.8. *Isotopic oriented submanifolds have equal Poincaré duals.*

This has some important concrete consequences. Let $S, T \subset M$ be two compact submanifolds of complementary dimension. We may isotope them to some transverse submanifolds S', T', and define

$$i(S, T) = i(S', T').$$

TBD Mettere a posto gli esempi.

This map is independent of the S', T' chosen since it equals $\omega_S \wedge \omega_T$.

Example 8.6.9. The algebra $H^*(\mathbb{CP}^n)$ is isomorphic to

$$H^*(\mathbb{CP}^n) \cong \mathbb{R}[x]/_{(x^{n+1})}$$

where $x = \omega_H \in H^2(\mathbb{CP}^n)$ is the dual form to any hyperplane $H \subset \mathbb{CP}^n$.

Example 8.6.10. We know that $M = S^2 \times S^2$ has $H^2(M) = \mathbb{R}^2$. If we pick $S = S^2 \times \{p\}$ and $S' = \{q\} \times S^2$ oriented as S^2 we find two transverse surfaces in M with algebraic intersection $+1$. The two spheres form a basis of $H^2(M)$ and the intersection form in this basis is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

spiegare meglio e aggiungere curve in $\mathbb{CP}^2$.

CHAPTER 9

Riemannian manifolds

We have warned the reader multiple times that a smooth manifold M lacks many natural geometric notions, such as distance between points, length of curves, volumes, angles, geodesics. It is now due time to introduce all these concepts, by enriching M with an additional structure, called *metric tensor*. The manifold M equipped with a metric tensor is called a *Riemannian manifold*.

9.1. The metric tensor

It is a quite remarkable fact that all the various natural geometric notions that we are longing for can be introduced by equipping a smooth manifold with a single additional structure, that of a *metric tensor*.

9.1.1. Definition. Let M be a smooth manifold. A *metric tensor* is a Riemannian metric g on the tangent bundle TM, see Section 4.5. That is, it is a section g of the symmetric bundle

$$S_2(M)$$

such that $g(p)$ is positive-definite scalar product for every $p \in M$. Said again in other words, for every $p \in M$ we have a positive-definite scalar product

$$g(p)\colon T_pM \times T_pM \longrightarrow \mathbb{R}$$

that varies smoothly with p.

Example 9.1.1. The *Euclidean metric tensor* g_E on $\mathbb{R}^n$ is

$$g_E(x,y) = \sum_{i=1}^n x_i y_i$$

where we have identified $T_p\mathbb{R}^n$ with $\mathbb{R}^n$, as usual.

Definition 9.1.2. A *Riemannian manifold* is a pair (M,g) where M is a smooth manifold and g is a metric tensor on M.

For instance, the pair $(\mathbb{R}^n, g_E)$ is a Riemannian manifold called the *Euclidean space*.

Remark 9.1.3. We have shown in Section 4.5 that every bundle carries a Riemannian metric. Therefore every smooth manifold M has a metric tensor. The metric tensor is however not unique in any reasonable sense.

9.1.2. In coordinates. Let (M, g) be a Riemannian manifold and $\varphi\colon U \to V$ a chart. The tensor g on U may be transported along φ into a metric tensor $\varphi_* g$ on V, whose coordinates are denoted by

$$g_{ij}(p).$$

Here $g_{ij}(p)$ is a positive-definite symmetric matrix that depends smoothly on p. For instance, the Euclidean metric tensor is $g_{ij} = \delta_{ij}$.

9.1.3. Isometries. Every category has its own morphisms; in the presence of Riemannian metrics, one typically introduces only isomorphisms.

Let (M, g) be a Riemannian manifold. At every point $p \in M$ the tangent space T_pM is equipped with the scalar product $g(p)$, that we also denote for simplicity with the familiar symbol $\langle , \rangle$.

Definition 9.1.4. A diffeomorphism $f\colon M \to N$ between two Riemannian manifolds (M, g) and (N, h) is an *isometry* if

$$\langle v, w\rangle = \langle df_p(v), df_p(w)\rangle$$

for every $p \in M$ and $v, w \in T_pM$.

Two Riemannian manifolds M and N are *isometric* if there is an isometry relating them. A smooth map $f\colon M \to N$ is a *local isometry* at $p \in M$ if there are open neighbourhoods U and V of p and $f(p)$ such that $f(U) = V$ and $f|_U\colon U \to V$ is an isometry.

9.1.4. Submanifolds. Let (M, g) be a Riemannian manifold. Here is a simple albeit crucial observation: every submanifold $N \subset M$, of any dimension, inherits a metric tensor $g|_N$ simply by restricting g to the subspace $T_pN \subset T_pM$ at every $p \in N$. Therefore every smooth submanifold of a Riemannian manifold is itself naturally a Riemannian manifold.

In particular, every submanifold $S \subset \mathbb{R}^n$ inherits a Riemannian manifold structure by restricting g_E to S. Using Whitney's Embedding Theorem, we find here another proof that every manifold M carries a Riemannian structure.

A fundamental example is of course the sphere $S^n \subset \mathbb{R}^{n+1}$.

9.1.5. Products. The product $M\times N$ of two Riemannian manifolds (M, g) and (N, h) carries a natural Riemannian structure $g\times h$. Recall that $T_{(p,q)}M \times N = T_pM \times T_qN$ and define

$$\langle (v_1, w_1), (v_2, w_2)\rangle = \langle v_1, v_2\rangle + \langle w_1, w_2\rangle$$

for every $v_1, v_2 \in T_pM$ and $w_1, w_2 \in T_qN$.

Example 9.1.5. The torus $T = S^1 \times S^1$ with the product metric is the *flat torus*. It is important to note that the flat torus is *not* isometric to the torus of Figure 3.3. The first is flat, but the second is not: we will introduce the notion of *curvature* to explain that.

9.1.6. Length of curves. As we promised, we now start to show how the metric tensor alone generates a wealth of fundamental geometric concepts. We start by defining the lengths of smooth curves.

Let $\gamma\colon I \to M$ be a smooth curve in a Riemannian manifold M. We define its *length* as

$$L(\gamma) = \int_I \|\gamma'(t)\| dt.$$

Here of course the norm of a vector $v \in T_pM$ is

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

A *reparametrisation* of the curve γ is obtained by picking an interval diffeomorphism $\varphi\colon J \to I$ and setting $\eta = \gamma \circ \varphi$.

Proposition 9.1.6. *The length of γ is independent of the parametrisation.*

Proof. We have

$$L(\gamma) = \int_I \|\gamma'(t)\| dt = \int_J \big\|\gamma'\big(\varphi(u)\big)\big\| |\varphi'(u)| du = \int_J \|\eta'(u)\| du = L(\eta).$$

The proof is complete. □

More generally, the length $L(\gamma)$ is also invariant if we pre-compose γ with a smooth surjective monotone map $\varphi\colon J \to I$, that is with $\varphi'(t) \geq 0$ everywhere (or $\varphi'(t) \leq 0$ everywhere). With some abuse of language we also call this change of variables a *reparametrisation*.

9.1.7. Metric space. A connected Riemannian manifold (M, g) is also a metric space, with the following distance: for every $p, q \in M$ we define $d(p, q)$ as the infimum of the lengths of all the paths connecting p to q, that is

$$d(p, q) = \inf \big\{ L(\gamma) \bigm| \gamma\colon [a, b] \to M,\ \gamma(a) = p,\ \gamma(b) = q \big\}.$$

Proposition 9.1.7. *This is a distance, compatible with the topology of M.*

Proof. We clearly have $d(p, p) = 0$. We now prove that $p \neq q \Rightarrow d(p, q) > 0$. Pick a small open chart $\varphi\colon U \to V$ with $p \in U$, $\varphi(p) = 0$, and $q \notin U$. Choose a disc $D \subset V$ of some small radius r centred at the origin. The transported metric tensor on D is some g_{ij} depending smoothly on $x \in D$.

For every $x \in D$ and $v \in T_x\mathbb{R}^n$, we indicate with $\|v\|_E$ and $\|v\|_g$ the Euclidean and g-norm of v. Since D is compact, there are $M > m > 0$ with

$$m\|v\|_E < \|v\|_g < M\|v\|_E$$

for every $x \in D$ and every $v \in T_x\mathbb{R}^n$. Let α be a curve in V that goes from 0 to some point in ∂D. We know that the Euclidean length of α is $\geq r$, and we deduce that the g-length of α is $> rm$. Since every curve γ connecting p and q must cross $\varphi^{-1}(\partial D)$, we deduce that $L(\gamma) \geq rm$ and hence $d(p, q) \geq rm$.

We clearly have $d(p, q) = d(q, p)$. To show transitivity, we note that if γ is a curve from p to q and η is a curve from q to r, we can concatenate γ

and η to a *smooth* curve from p to r: to get smoothness it suffices to priorly reparametrise γ and η using transition functions.

In our discussion, we have also shown that for every neighbourhood U of p there is an $\varepsilon > 0$ such that the d-ball of radius ε is entirely contained in U. Conversely, it is also clear that an open d-ball is open in the topology of M. Therefore d is compatible with the topology of M. □

Remark 9.1.8. The infimum defining $d(p, q)$ may not be a minimum! On $M = \mathbb{R}^2 \setminus \{0\}$ with the Euclidean metric tensor, we have $d\big((1,0),(-1,0)\big) = 2$ but there is no curve in M joining $(1,0)$ and $(-1,0)$ having length precisely 2.

9.1.8. Volume form. An oriented Riemannian manifold (M, g) has a natural volume form ω, defined as follows. At every point $p \in M$, the tangent space T_pM is equipped with an orientation and a positive-definite scalar product $g(p)$, and as in Section 2.6.3 we define ω unambiguously by requiring

$$\omega(p)(v_1, \ldots, v_n) = 1$$

on every positive orthornormal basis $v_1, \ldots, v_n$ of T_pM. To show that ω varies smoothly with p, we calculate ω on coordinates.

Proposition 9.1.9. *If g_{ij} is a metric tensor on $U \subset \mathbb{R}^n$, then*

$$\omega = \sqrt{\det g_{ij}}\, dx^1 \wedge \ldots \wedge dx^n.$$

Proof. Let $v^1, \ldots, v^n$ be a positive g-orthonormal basis for $(\mathbb{R}^n)^*$. We get

$$\omega = v^1 \wedge \ldots \wedge v^n = \det A\, dx^1 \wedge \ldots \wedge dx^n$$

where $v^i = A^i_j e^j$. Now $A^l_i g^{ij} A^k_j = \delta^{lk}$ gives $(\det A)^2 \det g^{-1} = 1$ and hence we get $\det A = \sqrt{\det g}$. □

In particular the *volume* of a Borel subset $S \subset U$ is

$$\mathrm{Vol}(S) = \int_S \sqrt{\det g_{ij}}\, dx^1 \cdots dx^n.$$

This expression is of course chart-independent.

9.2. Connections

We now want to define geodesics. It would be natural to try to define them as curves that minimise locally the distance; however, differential geometers usually prefer to take a different perspective: they introduce geodesics as curves whose tangent vectors do not "deviate" from the trajectory, that is that go as "straight" as possible.

To formalise this notion of "deviation" we need somehow to connect nearby tangent vectors via a structure called *connection*. This structure has many interesting features that go beyond the definition of geodesics: it is also a way to derive vector fields along tangent vectors, and for that reason it is also called with another appropriate name: *covariant derivative*. The two notions

– connection and covariant derivative – are in fact the same thing, a powerful structure that can be employed for different purposes, whose application goes even beyond the realm of riemannian manifolds.

9.2.1. Definition. As we said in the previous chapters, one of the main themes in differential topology is the quest for a correct notion of derivation of vector (more generally, tensor) fields on a smooth manifold M. Without equipping M with an additional structure, the best thing that we can do is to derive a vector field Y with respect to another vector field X via the *Lie derivative* $L_X(Y) = [X, Y]$.

As we have already noted, the definition of $L_X(Y)$ is *local*, in the sense that its value at $p \in M$ depends only on the values of X and Y in any neighbourhood of p, but is *not* a *pointwise* definition, in the sense that it does not depend on the vector $v = X(p)$ alone, as it happens in the usual directional derivative of smooth functions in $\mathbb{R}^n$. We are then urged to introduce a somehow stronger notion of derivation that depends only on the tangent vector $v = X(p)$.

Let M be a smooth manifold.

Definition 9.2.1. A *connection* ∇ is an operation that associates to every $v \in T_pM$ at every $p \in M$, and to every vector field X defined on a neighbourhood of p, another tangent vector

$$\nabla_v X \in T_pM$$

called the *covariant derivative* of X along v, such that the following holds:

(1) if X and Y agree on a neighbourhood of p, then $\nabla_v X = \nabla_v Y$;
(2) we have linearity in both terms:
$$\nabla_v(\lambda X + \mu Y) = \lambda \nabla_v(X) + \mu \nabla_v(Y),$$
$$\nabla_{\lambda v + \mu w} X = \lambda \nabla_v(X) + \mu \nabla_w(X),$$
where $\lambda, \mu \in \mathbb{R}$ are arbitrary scalars;
(3) the Leibnitz rule holds:
$$\nabla_v(fX) = v(f)X(p) + f(p)\nabla_v X$$
for every function f defined in a neighbourhood of p;
(4) ∇ depends smoothly on p.

We explain the last condition. For every two vector fields X, Y defined in a common open subset $U \subset M$, we require

$$\nabla_{Y(p)} X$$

to be another vector field in U. That is we require $\nabla_{Y(p)} X$ to vary smoothly with respect to the point $p \in U$.

We note that in fact (3) implies (1), as one sees easily by taking f to be a bump function that is constantly 1 in a neighbourhood of p.

9.2.2. Christoffel symbols. On a chart, we may consider the coordinate vector fields $e_i = \frac{\partial}{\partial x_i}$. We get

$$\nabla_{e_i} e_j = \Gamma_{ij}^k e_k$$

where we have used the Einstein summation convention, for some real numbers Γ_{ij}^k that depend smoothly on p because of the smoothness assumption (4).

The smooth functions Γ_{ij}^k are called the *Christoffel symbols* of the connection. On a chart, these determine the connection completely: indeed, for every vector field $X = X^j e_j$ and tangent vector $v = v^i e_i$ at some point we get

$$\begin{aligned}\nabla_v X = v^i \nabla_{e_i}(X^j e_j) &= v^i \frac{\partial X^j}{\partial x_i} e_j + v^i X^j \nabla_{e_i} e_j \\ &= v^i \frac{\partial X^j}{\partial x_i} e_j + v^i X^j \Gamma_{ij}^k e_k.\end{aligned}$$

We may rewrite this equality as

$$\nabla_v X = \left(v^i \frac{\partial X^k}{\partial x_i} + v^i X^j \Gamma_{ij}^k \right) e_k. \tag{18}$$

Therefore the covariant derivative ∇_v is the usual directional derivative along v plus a correction term that is encoded by the Christoffel symbols Γ_{ij}^k. In particular we have

$$\nabla_{e_i} X = \frac{\partial X}{\partial x_i} + X^j \Gamma_{ij}^k e_k.$$

Note that the directional derivative is not a chart-independent operation! You may think at Γ_{ij}^k as a correction term that transforms it into a chart-independent one.

Conversely, on any open subset $U \subset \mathbb{R}^n$, for every choice of smooth maps $\Gamma_{ij}^k \colon U \to \mathbb{R}$ there is a connection ∇ whose Christoffel symbols are Γ_{ij}^k. The connection ∇ is defined via (18), and one readily verifies that the axioms (1-4) are satisfied.

Of course when the connection is read on another chart the Christoffel symbols modify in some appropriate way:

Exercise 9.2.2. If the coordinates change as

$$\frac{\partial}{\partial \hat{x}_i} = \frac{\partial x_k}{\partial \hat{x}_i} \frac{\partial}{\partial x_k}$$

the Christoffel symbols modify accordingly as follows:

$$\hat{\Gamma}_{ij}^k = \frac{\partial x_p}{\partial \hat{x}_i} \frac{\partial x_q}{\partial \hat{x}_j} \Gamma_{pq}^r \frac{\partial \hat{x}_k}{\partial x_r} + \frac{\partial \hat{x}_k}{\partial x_m} \frac{\partial^2 x_m}{\partial \hat{x}_i \partial \hat{x}_j}.$$

The second derivatives are there to warn us that the Christoffel symbols Γ_{ij}^k are not the coordinates of any tensor. A connection is not a tensor field in any sense.

9.2.3. Curves suffice. We know that $\nabla_v X \in T_pM$ depends only on the behaviour of X on any neighbourhood of p. In fact, its restriction to a smaller subset suffices to determine $\nabla_v X$.

Proposition 9.2.3. *The covariant derivative $\nabla_v X \in T_pM$ depends only on v and the restriction of X to any curve tangent to v.*

Proof. On a chart (18) shows that $\nabla_v X$ depends only on $v, X(p)$, and the directional derivative of X along v. This proves the assertion. □

In particular, two vector fields that coincide on some curve tangent to v have the same covariant derivative along v.

9.2.4. Vector fields along curves. Proposition 9.2.3 leads us naturally to the following definition.

Definition 9.2.4. Let M be a manifold and $\gamma\colon I \to M$ a curve. A *vector field* along γ is a smooth map $X\colon I \to TM$ with $X(t) \in T_{\gamma(t)}M$ for all $t \in I$.

The vector field X is *tangent* to γ if $X(t)$ is a multiple of $\gamma'(t)$ for all t. For instance, the *velocity field* of γ is the vector field $\gamma'(t)$ and is of course tangent to γ.

If γ is an embedding, we may interpret X as a vector field on its support, but this interpretation fails if γ is only an immersion.

Let ∇ be a fixed connection on M. Let $\gamma\colon I \to M$ be an immersed curve, that is we have $\gamma'(t) \neq 0$ for all $t \in I$. For every vector field X along γ, we define another vector field $\frac{DX}{dt}$ on γ called its *derivative*, as follows.

If I is a compact interval and γ is an embedding, we consider X as a vector field defined on $\gamma(I)$, we extend X arbitrarily to an open neighbourhood of $\gamma(I)$, and for every $t \in I$ we define

$$\frac{DX}{dt} = \nabla_{\gamma'(t)} X.$$

The vector field $\frac{DX}{dt}$ does not depend on the extension of X outside γ thanks to Proposition 9.2.3.

In general, the curve γ is an immersion and hence it is an embedding on every sufficiently small neighbourhood of every point $t_0 \in I$. Therefore we may define $\frac{DX}{dt}(t_0)$ as above for every $t_0 \in I$.

Everything can be written more explicitly on a chart. On an open subset $V \subset \mathbb{R}^n$ we have $\gamma(t) = (\gamma^1(t), \ldots, \gamma^n(t))$ and $X = X^i(t)e_i$. We get

$$\frac{DX}{dt} = \frac{dX}{dt} + \gamma'(t)^i X^j(t) \Gamma^k_{ij}(\gamma(t)) e_k. \tag{19}$$

Remark 9.2.5. One may use (19) to define $\frac{DX}{dt}$ for any smooth curve γ, not only immersions. We will not need this.

9.2.5. Parallel transport. We have just defined a way to derive vector fields along immersed curves, and we now investigate the vector fields whose derivative vanishes at every point of the curve.

Let M be a smooth manifold equipped with a connection ∇. Let $\gamma\colon I \to M$ be an immersed curve. A vector field X along γ is *parallel* if

$$\frac{DX}{dt} = 0$$

for all $t \in I$. Here is a very important existence and uniqueness property:

Proposition 9.2.6. *For every $t_0 \in I$ and every $v \in T_{\gamma(t_0)}M$ there is a unique parallel vector field X on γ with $X(t_0) = v$.*

Proof. We easily reduce to the case where $\gamma(I)$ is entirely contained in the domain U of a chart $\varphi\colon U \to V$. Using (19), the problem reduces to solving a system of n linear differential equations in $X^k(t)$ with $k = 1, \ldots, n$, that is:

$$\frac{dX^k}{dt} + \gamma'(t)^i X^j(t)\Gamma_{ij}^k(\gamma(t)) = 0. \tag{20}$$

The system has a unique solution satisfying the initial condition $X^k(t_0) = v^k$ for all k. The solution exists for all $t \in I$ because the system is linear. □

For every $t \in I$, we think at the vector $X(t)$ as the one obtained from $v = X(t_0)$ by *parallel transport* along γ. We have just discovered a very nice (and maybe unexpected) feature of connections: they may be used to transport tangent vectors along curves.

It is sometimes useful to denote the parallel-transported vector $X(t)$ as

$$X(t) = \Gamma(\gamma)_{t_0}^t(v)$$

to stress the dependence on all the objects involved. We get a map

$$\Gamma(\gamma)_{t_0}^t\colon T_{\gamma(t_0)}M \longrightarrow T_{\gamma(t)}M$$

called the *parallel transport map*.

Proposition 9.2.7. *The parallel transport map is a linear isomorphism.*

Proof. The map is linear because (20) is a linear system of differential equations. It is an isomorphism because its inverse is $\Gamma(\gamma)_t^{t_0}$. □

Note that

$$\Gamma(\gamma)_{t_0}^{t_2} = \Gamma(\gamma)_{t_1}^{t_2} \circ \Gamma(\gamma)_{t_0}^{t_1}$$

for every triple $t_0, t_1, t_2 \in I$. The smooth dependence on initial values tells us that $\Gamma(\gamma)_t^{t'}$ depends smoothly on t and t', when read on charts.

We now understand where the name "connection" comes from: the operator ∇ can be used to connect via isomorphisms all the tangent spaces T_pM at the points $p = \gamma(t)$ visited by any immersed curve γ. It is important to stress here that the isomorphisms depend heavily on the chosen curve γ: two

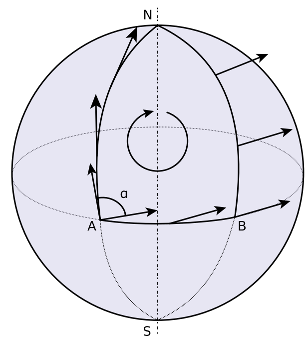

Figure 9.1. By parallel-transporting a vector along the edges of a spherical triangle in S^2, from A to N to B and back to A, we transform it into a new one rotated by some angle α. Here α is proportional to the area of the triangle, and in general it is connected to the curvature of the manifold. The connection ∇ that we are using here is the one naturally associated to the metric, to be defined in Section 9.3.

distinct immersed curves γ_1 and γ_2, both connecting the same points p and q, produce in general two different isomorphisms between the tangent spaces T_pM and T_qM. This may hold also if γ_1 and γ_2 are homotopic. As we will see, the *curvature* of ∇ measures precisely this discrepancy. See Figure 9.1.

Remark 9.2.8. A continuous map $\gamma\colon I \to M$ is *a piecewise immersion* if it is a concatenation of finitely many immersions. Parallel transport extends to piecewise smooth immersed curves in the obvious way, see Figure 9.1.

9.2.6. Connections form an affine space. Does every smooth manifold admit some connection ∇? And if it does, how many connections are there? The answer to the first question is positive but we postpone it to the next section. We can easily answer the second one here.

Recall that a tensor field T of type $(1, 2)$ on M is a bilinear map

$$T(p)\colon T_pM \times T_pM \longrightarrow T_pM$$

that depends smoothly on p.

Proposition 9.2.9. *If ∇ is a connection on M and $T \in \Gamma\big(\mathcal{T}_1^2(M)\big)$ is a tensor field of type $(1, 2)$, then the operator $\nabla' = \nabla + T$, defined as*

$$\nabla'_v X = \nabla_v X + T(p)(v, X(p))$$

is also a connection. Every connection ∇' on M arises in this way.

In the expression we have $p \in M$, $v \in T_pM$, and X is a vector field defined in a neighbourhood of p, as usual.

Proof. To prove that ∇' is a connection, we show that it satisfies the Leibniz rule (the other axioms are obvious). We have:

$$\begin{aligned}\nabla'_v(fX) &= \nabla_v(fX) + T(p)\big(v, f(p)X(p)\big)\\ &= v(f)X + f(p)\nabla_v X + f(p)T(p)\big(v, X(p)\big)\\ &= v(f)X + f(p)\nabla'_v X.\end{aligned}$$

Conversely, if ∇' is another connection, we consider the expressions in coordinates (18) for both $\nabla'_v X$ and $\nabla_v X$ and discover that

$$\nabla'_v X - \nabla_v X = v^i X^j \big((\Gamma')^k_{ij} - \Gamma^k_{ij}\big)e_k.$$

The right-hand expression describes a tangent vector at p that depends (linearly) only on the tangent vectors v and $X(p)$. If we indicate this vector as $T(p)(v, X(p))$, we get a tensor field T of type (1,2). In coordinates, we have

$$T^k_{ij} = (\Gamma')^k_{ij} - \Gamma^k_{ij}.$$

The proof is complete. □

We have just discovered that the space of all connections ∇ on M is naturally an affine space on the (infinite-dimensional) space $\Gamma\big(\mathcal{T}^2_1(M)\big)$.

Remark 9.2.10. We can use Exercise 9.2.2 to confirm that $T^k_{ij} = (\Gamma')^k_{ij} - \Gamma^k_{ij}$ are the coordinates of a tensor (the second partial derivatives cancel).

9.3. The Levi–Civita connection

We have already seen that on a Riemannian manifold M we can talk about distances between points, length of curves, and volumes. We now show that M also has a preferred connection, called the *Levi-Civita connection*. We will then use it to define geodesics in the next section.

9.3.1. Introduction. As we have seen, a smooth manifold M carries many different connections, and we are now looking at some reasonable way to discriminate between them. The main motivation is the following ambitious question: if M has a metric tensor g, is there a connection ∇ that is somehow more suited to g?

An elegant and useful way to understand a connection ∇ consists of examining some tensor fields that are associated canonically to ∇. We now introduce one of these.

9.3.2. Torsion. Let ∇ be a connection on a smooth manifold M. The *torsion* T of ∇ is a tensor field of type $(1, 2)$ defined as follows. For every $p \in M$ and $v, w \in T_pM$ we set

$$T(p)(v, w) = \nabla_v Y - \nabla_w X - [X, Y](p)$$

where X and Y are any vector fields defined in a neighbourhood of p extending the tangent vectors v and w. Of course we need to prove that this definition is well-posed, a fact that is not evident at all at first sight.

Proposition 9.3.1. *The tangent vector $T(p)(v,w)$ is independent of the extensions X and Y.*

Proof. In coordinates we have

$$\begin{aligned} T(p)(v,w) &= \left(v^i \frac{\partial Y^k}{\partial x_i} + v^i Y^j \Gamma_{ij}^k - w^i \frac{\partial X^k}{\partial x_i} - w^i X^j \Gamma_{ij}^k - v^i \frac{\partial Y^k}{\partial x_i} + w^i \frac{\partial X^k}{\partial x_i} \right) e_k \\ &= \left(v^i w^j \Gamma_{ij}^k - w^i v^j \Gamma_{ij}^k \right) e_k = v^i w^j (\Gamma_{ij}^k - \Gamma_{ji}^k) e_k. \end{aligned}$$

The proof is complete. □

During the proof, we have also shown that in coordinates we have

$$T_{ij}^k = \Gamma_{ij}^k - \Gamma_{ji}^k.$$

A connection ∇ is *symmetric* if its torsion vanishes, that is if $\Gamma_{ij} = \Gamma_{ji}$ on any coordinate chart. The torsion is clearly an antisymmetric tensor, that is $T(p)(v,w) = -T(p)(w,v)$ for all v, w. Finally, if we contract the torsion T with two vector fields X and Y we get the elegant equality of vector fields:

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

9.3.3. Bilinear operators on vector fields. We have already encountered in this book three bilinear operators

$$\mathfrak{X}(M) \times \mathfrak{X}(M) \longrightarrow \mathfrak{X}(M)$$

that are quite dissimilar in nature: these are $[,]$, ∇, and T. Given two vector fields X and Y, then we can define a third one Z by setting it to be equal to

$$[X,Y], \qquad \nabla_X Y, \qquad \text{or} \qquad T(X,Y).$$

The main difference between these three operators is the following:

- $[X,Y]$ at p depends on X and Y;
- $\nabla_X Y$ at p depends on $X(p)$ and Y;
- $T(X,Y)$ at p depends on $X(p)$ and $Y(p)$.

This also expresses the fact that the operator T is the only one that arises from a tensor field.

Remark 9.3.2. Some authors describe these differences by saying that the operator T is $C^\infty(M)$-bilinear, that is $T(fX, gY) = fgT(X,Y)$ for every $f, g \in C^\infty(M)$. Analogously, ∇ is $C^\infty(M)$-linear on its left, that is $\nabla_{fX} Y = f\nabla_X Y$.

9.3.4. Compatible connections. Let (M, g) be a Riemannian manifold. As we said, we would like to assign an appropriate conection ∇ to g. We start by defining a reasonable compatibility condition.

We say that a connection ∇ is *compatible* with g if every parallel transport isomorphism

$$\Gamma(\gamma)_{t_0}^{t_1} \colon T_{\gamma(t_0)}M \longrightarrow T_{\gamma(t_1)}M$$

is actually an isometry, for every immersed curve $\gamma \colon I \to M$ and every $t_0, t_1 \in I$.

We now express this condition in three more equivalent ways.

Proposition 9.3.3. *The connection ∇ is compatible if and only if*

$$\frac{d}{dt}\langle X, Y\rangle = \left\langle \frac{DX}{dt}, Y \right\rangle + \left\langle X, \frac{DY}{dt}\right\rangle \tag{21}$$

for every immersed curve $\gamma \colon I \to M$ and vector fields X, Y on it.

Proof. If (21) holds, for every parallel vector fields X, Y on γ we get that $\big\langle X(t), Y(t)\big\rangle$ is constant on t and hence the parallel transport along γ is an isometry. Therefore ∇ is compatible.

Conversely, suppose that ∇ is compatible. Pick an orthonormal basis $e_1, \ldots, e_n$ of T_pM and parallel-transport it along γ. Write

$$X(t) = X(\gamma(t)) = X^ie_i, \qquad Y(t) = Y(\gamma(t)) = Y^ie_i.$$

Using the Leibnitz rule we deduce that

$$\nabla_{\gamma'(t)}X = \frac{dX^i}{dt}e_i, \qquad \nabla_{\gamma'(t)}Y = \frac{dY^i}{dt}e_i$$

and hence

$$\frac{d}{dt}\big\langle X(t),Y(t)\big\rangle = \frac{d}{dt}(X^iY^i) = \frac{dX^i}{dt}Y^i + X^i\frac{dY^i}{dt} = \big\langle \nabla_{\gamma'(t)}X, Y\big\rangle + \big\langle X, \nabla_{\gamma'(v)}Y\big\rangle.$$

The proof is complete. □

We can easily translate this into a local condition. We interpret v as a derivation acting on the smooth function $\langle X, Y\rangle$.

Corollary 9.3.4. *The connection ∇ is compatible if and only if*

$$v\langle X, Y\rangle = \big\langle \nabla_v X, Y\big\rangle + \big\langle X, \nabla_v Y\big\rangle \tag{22}$$

for every tangent vector $v \in T_pM$ and every vector fields X, Y defined in a neighbourhood of p.

Expressed in coordinates, this is translated as follows.

Proposition 9.3.5. *The connection ∇ is compatible if and only if*

$$\frac{\partial g_{ij}}{\partial x_k} = \Gamma_{ki}^l g_{lj} + \Gamma_{kj}^l g_{li} \tag{23}$$

in coordinates at every chart.

Proof. We pick any chart and write (22). By linearity in v, we may suppose that $v = e_k$. We have $X = X^i e_i$ and $Y = Y^j e_j$. The equation transforms into

$$\frac{\partial}{\partial x_k}\left(g_{ij}X^iY^j\right) = \left(\frac{\partial X^i}{\partial x_k} + X^j\Gamma^i_{kj}\right) g_{il}Y^l + \left(\frac{\partial Y^i}{\partial x_k} + Y^j\Gamma^i_{kj}\right) g_{il}X^l.$$

After deriving the left member and simplifying this transforms into

$$\frac{\partial g_{ij}}{\partial x_k}X^iY^j = X^j\Gamma^i_{kj}g_{il}Y^l + Y^j\Gamma^i_{kj}g_{il}X^l.$$

After renaming indices, this holds for every X and Y precisely when

$$\frac{\partial g_{ij}}{\partial x_k} = \Gamma^l_{ki}g_{lj} + \Gamma^l_{kj}g_{li}.$$

The proof is complete. □

The proof also shows that if (23) holds on all the charts of an atlas, then it also does at any compatible chart.

9.3.5. The Levi-Civita connection. As promised, we now assign to any Riemannian manifold (M, g) a canonical connection ∇, called the *Levi-Civita connection*.

Theorem 9.3.6. *Every Riemannian manifold (M, g) has a unique symmetric compatible connection ∇. On any chart, its Christoffel symbols are*

$$\Gamma^l_{ij} = \frac{1}{2}g^{kl}\left(\frac{\partial g_{jk}}{\partial x_i} + \frac{\partial g_{ki}}{\partial x_j} - \frac{\partial g_{ij}}{\partial x_k}\right). \tag{24}$$

Proof. We start by proving uniqueness. Let ∇ be a symmetric compatible connection. On a chart, we use (23) three times with the indices i, j, k permuted cyclically, and using symmetry $\Gamma^k_{ij} = \Gamma^k_{ji}$ we get

$$\frac{\partial g_{jk}}{\partial x_i} + \frac{\partial g_{ki}}{\partial x_j} - \frac{\partial g_{ij}}{\partial x_k} = 2\Gamma^m_{ij}g_{mk}.$$

By multiplying both members with the inverse matrix g^{kl} we find

$$\Gamma^l_{ij} = \frac{1}{2}g^{kl}\left(\frac{\partial g_{jk}}{\partial x_i} + \frac{\partial g_{ki}}{\partial x_j} - \frac{\partial g_{ij}}{\partial x_k}\right).$$

This shows that Γ^l_{ij} and hence ∇ are uniquely determined.

Concerning existence, we now use (24) to define ∇ locally on a chart. The connection is clearly symmetric and one verifies easily that is also compatible using Proposition 9.3.5. Moreover, the resulting ∇ is actually chart-independent: if not, we would get two different symmetric and compatible connections on some open set, which is impossible. Therefore all the ∇ constructed along charts glue to a global ∇ on M. □

The unique symmetric compatible connection ∇ is called the *Levi-Civita connection*.

Example 9.3.7. If $U \subset \mathbb{R}^n$ is equipped with the Riemannian metric g, the Christoffel symbols $\Gamma_{ij}^k = 0$ vanish everywhere and the Levi-Civita connection coincides with the usual directional derivative.

We will since now equip every Riemannian manifold (M, g) with its Levi-Civita connection ∇.

Remark 9.3.8. While the compatibility assumption looks natural, the reasons for preferring a symmetric connection may look obscure at this point. We can single out three arguments in its favour: (i) this seems the only (or at least the simplest) way to get a canonical connection; (ii) we will see in the next section that, thanks to symmetry, the Levi-Civita connection extends in a very simple way to submanifolds; (iii) by picking a compatible connection with non-vanishing torsion things do not change too much, since (as we will see) we would get exactly the same geodesics (and defining geodesics is the main reason for introducing connections).

TBD torsione dopo

9.3.6. Submanifolds. Let M be a Riemannian manifold and $N \subset M$ a submanifold. The manifold N has an induced Riemannian structure, and we now investigate the relation between the corresponding Levi-Civita connections ∇^M and ∇^N. It turns out that ∇^N is very easily determined by ∇^M. This is particularly useful when the ambient space is $M = \mathbb{R}^m$ with the Euclidean metric tensor, since there ∇^M is the usual directional derivative and ∇^N assumes a simple and intuitive form.

Let $p \in N$ be a point and $v \in T_pN$ a tangent vector. Let X be a vector field (tangent to N) defined on a neighbourhood of p in N. Extend X arbitrarily to a vector field on a neighbourhood of p in M. Let $\pi\colon T_pM \to T_pN$ be the orthogonal projection.

Proposition 9.3.9. *The following holds:*

$$\nabla^N_v X = \pi\big(\nabla^M_v X\big).$$

Proof. We define a connection ∇ on N by setting $\nabla_v(X) = \pi(\nabla^M_v X)$ for every vector field X in some open subset of N, using some local extension of X in M. The vector $\nabla_v(X)$ does not depend on the extension (exercise) and ∇ is indeed a connection on N. It is compatible: by Corollary 9.3.4 we get

$$v\langle X, Y\rangle = \langle \nabla^M_v X, Y\rangle + \langle X, \nabla^M_v Y\rangle = \langle \nabla_v X, Y\rangle + \langle X, \nabla_v Y\rangle$$

for every vector fields X, Y on a neighbourhood of p in N, extended arbitrarily to a neighbourhood in M. The connection is symmetric: analogously we have

$$T(v, w) = \nabla_v Y - \nabla_w X - [X, Y](p) = \pi\big(\nabla^M_v Y - \nabla^M_w X - [X, Y](p)\big) = \pi(0) = 0$$

where we have used that $[X, Y](p)$ is tangent to N since both X and Y are. By the uniqueness of the Levi-Civita connection we have $\nabla = \nabla^N$. $\square$

Let $\gamma\colon I \to N$ be an immersed curve and X be a vector field on γ. We denote analogously by $\frac{D^MX}{dt}$ and $\frac{D^NX}{dt}$ the derivatives of γ with respect to the two connections ∇^M and ∇^N.

Corollary 9.3.10. *The following holds:*

$$\frac{D^NX}{dt} = \pi\left(\frac{D^MX}{dt}\right).$$

The case where $M = \mathbb{R}^m$ is equipped with the Euclidean metric and $N \subset \mathbb{R}^n$ is a submanifold is particularly interesting:

Corollary 9.3.11. *A vector field X on $\gamma\colon I \to N$ is parallel (on N) if and only if its derivative $X'(t)$ in $\mathbb{R}^m$ is orthogonal to $T_{\gamma(t)}N$ for every $t \in I$.*

9.4. Geodesics

We know that every Riemannian manifold (M, g) has a preferred connection ∇, and now we use ∇ to define geodesics. We end this section by showing that geodesics are precisely the curves that minimise the path length, at least locally (not necessarily globally).

9.4.1. Definition. Let M be a manifold equipped with a connection ∇.

Definition 9.4.1. A smooth immersed curve $\gamma\colon I \to M$ is a *geodesic* if the velocity field $\gamma'(t)$ is parallel along γ.

Recall that this means that $\frac{D\gamma'}{dt} = 0$ for every $t \in I$. A geodesic is *maximal* if it is not the restriction of a longer geodesic $\eta\colon J \to M$ with $I \subsetneq J$. Geodesics have many nice properties; the first important one is that they exist, and they are also unique once a starting point and a direction are fixed:

Proposition 9.4.2. *For every $p \in M$ and $v \in T_pM$ there is a unique maximal geodesic $\gamma\colon I \to M$ with $0 \in I$, $\gamma(0) = p$, and $\gamma'(0) = v$.*

In the proposition we also include the *trivial* constant geodesic $\gamma\colon \mathbb{R} \to M$, $\gamma(t) = p$, that corresponds to $v = 0$ (although this is not strictly speaking a geodesic according to our definition). The unique maximal geodesic γ tangent to v at $t = 0$ is sometimes denoted by γ_v.

Proof. In coordinates, an immersed curve $\gamma(t) = x(t)$ is a geodesic if and only if the following holds for all k, see (19):

$$\frac{d^2x_k}{dt^2} + \frac{dx_i}{dt}\frac{dx_j}{dt}\Gamma_{ij}^k = 0. \tag{25}$$

This is a second-order system of ordinary differential equations. The Cauchy–Lipschitz Theorem 1.3.5 ensures that the system has locally a unique solution with prescribed initial data $x(0) = p$ and $\frac{dx}{dt}(0) = v$. $\square$

The second-order system of differential equations (25) describe the geodesics in any coordinate system. Using the dot notation for time derivative, the equations may be written as

$$\ddot{x}_k + \dot{x}_i \dot{x}_j \Gamma_{ij}^k = 0. \tag{26}$$

To define geodesics we only need a connection ∇, not a Riemannian metric. We are of course mainly interested in the case where ∇ is the Levi-Civita connection of a Riemannian metric g. In that case the speed $\|\gamma'(t)\|$ of a geodesic γ makes sense, and it is clearly constant along t by (21). One may wonder if the same geodesic run at a different constant speed is still a geodesic: this is true thanks to the following fact, that holds for all connections ∇.

Proposition 9.4.3. *If γ is a geodesic, then $\eta(t) = \gamma(ct)$ is also a geodesic, for every non-zero $c \in \mathbb{R}$.*

Proof. If $\nabla_v X = 0$, then also

$$\nabla_{cv} cX = c^2 \nabla_v X = 0.$$

This concludes easily the proof. □

In particular, we have $\gamma_{cv}(t) = \gamma_v(ct)$.

Example 9.4.4. On $U \subset \mathbb{R}^n$ with the Euclidean metric, we have $\Gamma_{ij} = 0$ and hence the geodesics are precisely the straight lines run at constant speed.

Example 9.4.5. Let $N \subset \mathbb{R}^m$ be a submanifold, equipped with the induced Riemannian metric. By Corollary 9.3.11, an immersion $\gamma\colon I \to N$ is a geodesic if and only if $\gamma''(t)$ is orthogonal to $T_{\gamma(t)}N$ for all $t \in I$.

Example 9.4.6. By the previous example, every maximal circle on S^n run at constant speed is a geodesic. In other words, for every $p \in S^n$, every unitary vector $v \in T_pS^n = p^\perp$, and every $c > 0$, the curve $\gamma\colon \mathbb{R} \to S^n$ defined as

$$\gamma(t) = \cos(ct) \cdot p + \sin(ct) \cdot v$$

is a geodesic that starts from p in the direction v at speed c. To prove this it suffices to check that $\gamma(t) \in S^n$ and $\gamma''(t)$ is parallel to $\gamma(t)$, hence orthogonal to $T_{\gamma(t)}S^n$. By Proposition 9.4.2 these are precisely all the maximal geodesics in the sphere S^n.

9.4.2. Geodesic flow.

Let M be a smooth manifold equipped with a connection ∇. It would be nice if we could represent all the geodesics in M as the integral curves of some fixed vector field on M. However, this is clearly impossible! On a vector field, there is only one integral curve crossing each point p, but there are infinitely many geodesics through p, one for each direction $v \in T_pM$.

However, this strategy works if we just replace M with its tangent bundle TM. We can define a vector field X in TM as follows: for every $v \in TM$, let

$\gamma_v\colon I_v \to M$ be the unique maximal geodesic with $\gamma_v'(0) = v$. The derivative $\gamma_v'\colon I_v \to TM$ is a curve in TM, that we see as a canonical lift of γ_v from M to TM. We define $X(v) = d(\gamma_v')_0$.

The resulting vector field X on TM is smooth because the geodesic γ_v depends smoothly on the initial data. It is called the *geodesic vector field* on TM. Its maximal integral curves are precisely all the lifts of all the maximal geodesics in M. The vector field X generates a flow Φ on TM called the *geodesic flow*. The flow Φ moves the points in TM along the lifted geodesics.

The geodesic flow Φ is defined on some maximal open subset U of $TM\times\mathbb{R}$ containing $TM \times \{0\}$. We have $U \cap \big(\{v\} \times \mathbb{R}\big) = \{v\} \times I_v$. With moderate effort, mostly relying on theorems proved in the previous chapters, we have defined a quite general and fascinating geometric flow on (the tangent bundle of) every Riemannian manifold.

9.4.3. Exponential map. We now define a useful map that is tightly connected with the geodesic flow, called the *exponential map*. We start by defining the following subset of the tangent bundle:

$$V = \big\{ v \in TM \bigm| 1 \in I_v \big\} \subset TM.$$

Recall that $I_v \subset \mathbb{R}$ is the domain of γ_v. The *exponential map* is

$$\begin{aligned} \exp\colon V &\longrightarrow M \\ v &\longmapsto \gamma_v(1). \end{aligned}$$

For every $p \in M$ we define

$$V_p = V \cap T_pM, \qquad \exp_p = \exp|_{V_p}.$$

We see as usual M embedded in TM as the zero-section.

Proposition 9.4.7. *The domain V is an open neighbourhood of M and* exp *is smooth. Each V_p is open and star-shaped with respect to* 0*. We have*

$$\gamma_v(t) = \exp(tv)$$

for every $v \in TM$ and $t \in \mathbb{R}$ such that both members are defined.

Proof. Let U be the open domain of the geodesic flow Φ. We have $V = \big\{v \in TM \mid v\times\{1\} \in U\big\}$ and hence V is open. The map $\exp(v) = \pi(\Phi(v,1))$ is smooth. Star-shapeness and $\gamma_v(t) = \exp(tv)$ follow by Proposition 9.4.3. $\square$

Here is one important fact about the exponential map:

Proposition 9.4.8. *The map $\exp_p$ is a local diffeomorphism at $0 \in V_p$.*

Proof. We determine the endomorphism $d(\exp_p)_0\colon T_pM \to T_pM$. For every $v \in T_pM$ we have $\exp_p(tv) = \gamma_v(t)$ for all sufficiently small t. Therefore $d(\exp_p)_0(v) = \gamma_v'(0) = v$. We have proved that $d(\exp_p)_0 = \mathrm{id}$. In particular, it is invertible and hence $\exp_p$ is a local diffeomorphism at 0. $\square$

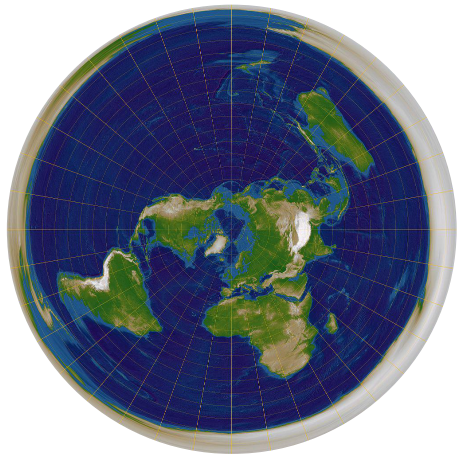

Figure 9.2. If we model the Earth as S^2 and look at the exponential map from the north pole N, the disc D of radius π in $T_N S^2$ is mapped to S^2 as shown here. The points in ∂D are all sent to the south pole.

The proposition says that the exponential map $\exp_p$ may be used as a parametrisation of a sufficiently small open neighbourhood of p. After many pages, we recover here a very intuitive idea: the tangent space T_pM should approximate the manifold near the point p. This idea may be realised concretely, via the exponential map, only after fixing a Riemannian metric on M.

Example 9.4.9. Consider the sphere S^n. Example 9.4.6 shows that for this Riemannian manifold we have $V = TM$ and

$$\exp(v) = \cos|v| \cdot p + \sin|v| \cdot \frac{v}{|v|}$$

for every $p \in S^n$ and $v \in T_pS^n$. Note that when $|v| = \pi$ we get $\exp(v) = -p$.

The map $\exp_p$ sends the open disc $D(0,\pi) \subset T_pM$ of radius π diffeomorphically onto $S^n \setminus \{-p\}$, while its boundary sphere $\partial D(0,\pi)$ goes entirely to the antipodal point $-p$. See Figure 9.2. Note in particular that $\exp_p$ is not a local diffeomorphism at the points in $\partial D(0,\pi)$. In general, it is guaranteed to be a local diffeomorphism only at the origin.

9.4.4. Normal coordinates. The exponential map furnishes some nice local parametrisations called *normal coordinates*, that we now investigate. These are very useful in many computations.

Let M be a Riemannian manifold and $p \in M$ a point. We fix an isometric isomorphism $\mathbb{R}^n \cong T_pM$. Let $r > 0$ be a sufficiently small radius such that the exponential map $\exp_p \colon B(0,r) \to M$ is defined and is an embedding. The image of $B(0,r)$ in M is called the *geodesic ball of radius r centred at p* and the coordinates $(x_1,\ldots,x_n)$ furnished by the parametrisation $\exp_p$ are the *normal coordinates* of the geodesic ball.

In normal coordinates, we represent a geodesic ball of radius r as $B(0,r) \subset \mathbb{R}^n$ with 0 corresponding to p. The metric g_{ij} varies smoothly in $B(0,r)$. The following is an immediate consequence of Proposition 9.4.7.

Proposition 9.4.10. *The geodesics emanated from the origin with speed c are Euclidean lines run with speed c. In particular at every $x \in B(0,r)$ we have the equality $x^i g_{ij}(x)x^j = x^i x^i$.*

As a consequence, we get the following.

Proposition 9.4.11. *At the origin we have:*

$$g_{ij}(0) = \delta_{ij}, \qquad \frac{\partial g_{ij}}{\partial x_k}(0) = 0,$$

$$\Gamma_{ij}^k(0) = 0, \qquad \frac{\partial \Gamma_{ij}^k}{\partial x_l}(0) + \frac{\partial \Gamma_{jl}^k}{\partial x_i}(0) + \frac{\partial \Gamma_{li}^k}{\partial x_j}(0) = 0.$$

Proof. The first equality follows from $d(\exp_p)_0 = \mathrm{id}$. The third and fourth follow from the geodesic equation (26), that is satisfied by all the lines $x(t) = tv$, $\forall v \in \mathbb{R}^n$. Plugging $x(t)$ in the equation we get

$$v_i v_j \Gamma_{ij}^k(0) = 0$$

for every $v \in \mathbb{R}^n$, and hence $\Gamma_{ij}^k(0) = 0$. By deriving the geodesic equation we get the more complicated third order equations

$$\dddot{x}_k + \ddot{x}_i \dot{x}_j \Gamma_{ij}^k + \dot{x}_i \ddot{x}_j \Gamma_{ij}^k + \dot{x}_i \dot{x}_j \frac{\partial \Gamma_{ij}^k}{\partial x_l} \dot{x}_l = 0.$$

If we substitute $x(t) = tv$ again we get

$$v_i v_j v_l \frac{\partial \Gamma_{ij}^k}{\partial x_l}(0) = 0$$

for every $v \in \mathbb{R}^n$, and we easily deduce the fourth equality. We then recover the second one from (23). $\square$

Of course the Christoffel symbols Γ_{ij}^k are guaranteed to vanish only at the origin, and not at the other points of $B(0,r)$. Proposition 9.4.10 can be upgraded to a stronger statement universally known as the *Gauss Lemma*.

Lemma 9.4.12 (Gauss Lemma). *At every $x \in B(0,r)$ we have the equality $x^i g_{ij}(x)y^j = x^i y^i$ for every $y \in \mathbb{R}^n$. In particular the spheres $\partial B(0,r')$ with $0 < r' < r$ are orthogonal to all the geodesics emanated from the origin.*

Proof. By the previous proposition, it suffices to consider the case $x^i y^i = 0$, that is y is tangent to $\partial B(0,x)$. We can also rescale y so that $x^i x^i = y^i y^i$. We must prove that $\langle x, y\rangle = x^i g_{ij}(x)y^j = 0$.

We want to extend x and y to two vector fields as in Figure 9.3. To do so, we define the curve $\gamma\colon (-\varepsilon, \varepsilon) \to B(0,r)$,

$$\gamma(t) = \cos t \cdot x + \sin t \cdot y.$$

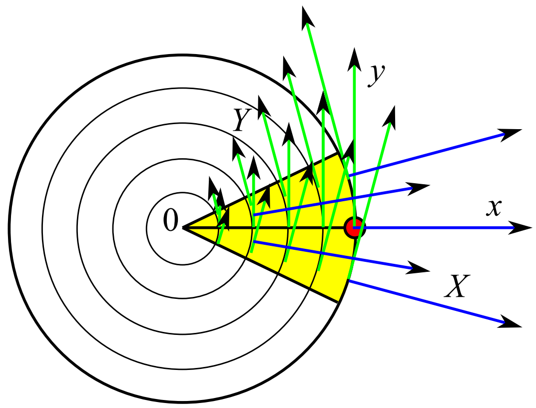

Figure 9.3. The Gauss Lemma says that, in normal coordinates, the vectors x and y are orthogonal. To prove this, we extend x and y to two commuting vector fields X (blue) and Y (green) defined on a (yellow) pencil of radial geodesics. Then we show that $\langle X, Y\rangle$ is constant along the rays, and hence vanishes everywhere.

We have $\gamma(0) = x$ and $\gamma'(0) = y$. Consider the embedding $F\colon (0,1] \times (-\varepsilon, \varepsilon) \to B(0,r)$,

$$F(s,t) = s\gamma(t).$$

We extend x and y to the vector fields $X = \frac{\partial F}{\partial s}$ and $Y = \frac{\partial F}{\partial t}$ on the image of F, see Figure 9.3. Note that $[X, Y] = 0$. We think of both vector fields depending on (s,t), so that $x = X(1,0)$ and $y = Y(1,0)$. At every point (s,t) we get

$$\frac{\partial}{\partial s}\langle X, Y\rangle = \langle \nabla_X X, Y\rangle + \langle X, \nabla_X Y\rangle.$$

We have $\nabla_X X = 0$ because X is the tangent field of the geodesic $s \mapsto s\gamma(t)$. Since $[X, Y] = 0$ and the torsion vanishes, we get $\nabla_X Y = \nabla_Y X$. Therefore

$$\frac{\partial}{\partial s}\langle X, Y\rangle = \langle X, \nabla_Y X\rangle = \frac{1}{2}\frac{\partial}{\partial t}\langle X, X\rangle = 0.$$

We have proved that $\frac{\partial}{\partial s}\langle X, Y\rangle = 0$ and hence $\langle X, Y\rangle$ is constant on the geodesic $s \mapsto sx$. Since we clearly have $\lim_{s\mapsto 0}\langle X, Y\rangle = 0$ we deduce that $\langle X, Y\rangle = 0$ everywhere and in particular $\langle x, y\rangle = 0$. The proof is complete. □

Every sphere $\partial B(0, r')$ with $0 < r' < r$ is called a *geodesic sphere of radius r'*. The Gauss Lemma says that g_{ij} at every point $x \neq 0$ decomposes orthogonally into a radial part that coincides with the Euclidean metric, and a tangential part, tangent to the geodesic sphere, that may however be arbitrary.

9.4.5. Minimising curves. We now start to study the tight connection between geodesics and distance between points.

Let M be a Riemannian manifold and $p, q \in M$ two points. We are interested in the smooth curves that connect p to q, that is the $\gamma\colon [a, b] \to M$ with $\gamma(a) = p$ and $\gamma(b) = q$. Recall that the length $L(\gamma)$ of γ is independent of its parametrisation. Recall also that $d(p, q)$ is the infimum of all the lengths of all the smooth curves connecting p and q. This infimum may not be realised in some cases; if it does, that is if there is a curve γ with $L(\gamma) = d(p, q)$, then the curve γ is called *minimising*.

Let $p \in M$ a point. Let $B \subset M$ be a geodesic ball centred at p with some radius r, and $q \in B$ be any other point. We know that B contains a radial geodesic $\gamma_{p,q}\colon [0, 1] \to B$ connecting p to q.

Proposition 9.4.13. *The geodesic $\gamma_{p,q}$ is a minimising curve. Every other minimising curve in M connecting p to q is obtained by reparametrising $\gamma_{p,q}$.*

Proof. Use the normal coordinates for B. Now $B = B(0, r)$ and the points p, q become $0, x \in B(0, r)$. Every curve γ in M connecting p to q contains an initial subcurve γ_* with support in $\overline{B(0, \|x\|)}$ and connecting 0 to some point in the sphere $\partial B(0, \|x\|)$.

By the Gauss Lemma the velocity $\gamma_*(t)'$ decomposes orthogonally into a radial and a tangential component. The integral of the norm of the radial component is at least r, since the radial component coincides with the Euclidean one. Therefore $L(\gamma_*) \geq r = L(\gamma_{p,q})$, and the equality holds if and only if there is no tangential component and the radial component is never decreasing, that is if $\gamma_*(t)$ is obtained by reparametrising $\gamma_{p,q}$. $\square$

Corollary 9.4.14. *A geodesic sphere of radius r around p consists precisely of all the points in M at distance r from p.*

For the same reason a geodesic ball centred at p of radius r consists precisely of the set $B(p, r)$ of all points in M at distance $< r$ from p. Conversely, if r is sufficiently small, every such set $B(p, r)$ is a geodesic ball.

It is a remarkable fact that the metric balls $B(p, r)$ with sufficiently small radius $r > 0$ are precisely the images of the balls $B(0, r) \subset T_pM$ along the exponential map.

9.4.6. Totally normal neighbourhoods. Let M be a Riemannian manifold. We have discovered that every point $p \in M$ has a neighbourhood U that is nice with respect to p, and now we want to be more democratic and show that we may pick a U that is also nice with respect to every point $q \in U$.

We say that an open subset $U \subset M$ is *totally normal* if for every $q \in U$ there is a geodesic ball centred at q containing U.

Proposition 9.4.15. *Every $p \in M$ has a totally normal neighbourhood U.*

Proof. Recall that $\exp\colon V \to M$ is defined on some open neighbourhood $V \subset TM$ of M. We consider the map

$$\begin{aligned} F\colon \quad V &\longrightarrow M \times M \\ (p, v) &\longmapsto (p, \exp_p(v)). \end{aligned}$$

We already know that $d(\exp_p)_0 = \mathrm{id}$. This implies easily that $dF_{(p,0)}$ is invertible and hence F is a local diffeomorphism at $(p, 0)$. Therefore there are a neighbourhood W of p and a $\delta > 0$ such that the restriction of F to

$$W' = \big\{(p, v) \bigm| p \in W, |v| < \delta\big\}$$

is a diffeomorphism onto its image $F(W')$. Pick a neighbourhood U of p such that $U \times U \subset F(W')$. □

If $U \subset M$ is a totally normal neighbourhood, then by Proposition 9.4.13 every two distinct points $p, q \in U$ are connected by a unique minimising geodesic $\gamma_{p,q}$ in M run at unit speed. The geodesic $\gamma_{p,q}$ varies smoothly in $p, q \in U$.

9.4.7. Locally minimising curves. We have defined geodesics as the solution of certain differential equations, and we can finally characterise them using only the distance between points.

Let M be a Riemannian manifold. We say that a curve $\gamma\colon I \to M$ is *locally minimising* if every $t \in I$ has a compact neighbourhood $[t_0, t_1] \subset I$ such that the restriction $\gamma|_{[t_0,t_1]}$ is minimising.

Exercise 9.4.16. If γ is minimising, it is also locally minimising.

Theorem 9.4.17. *A curve $\gamma\colon I \to M$ is locally minimising $\Longleftrightarrow$ it is obtained by reparametrising a geodesic.*

Proof. Let $\gamma\colon I \to M$ be a curve. For every t, pick a totally normal neighbourhood U containing $\gamma(t)$ and let $J \subset I$ be a neighbourhood of t such that $\gamma(J) \subset U$. Apply Proposition 9.4.13. □

The theorem is also true for piecewise immersions (see Remark 9.2.8), since using transition functions these can be reparametrised as smooth curves that have velocity zero at the angles. Geodesics are precisely the locally minimising curves, in a very robust manner.

9.4.8. Convex neighbourhoods. We now further improve the totally normal neighbourhoods by adding a quite natural requirement.

Definition 9.4.18. A subset $S \subset M$ of a Riemannian manifold M is *strictly convex* if any two points p, q in the closure $\bar{S}$ of S are joined by a unique minimising geodesic γ in M, and moreover its interior is contained in S.

We will prove that geodesic balls of sufficiently small radius are strictly convex. To this purpose, we will need the following.

Lemma 9.4.19. *For every point $p \in M$ there is a $r_0 > 0$ such that $B(p, r_0)$ is a geodesic ball, and every geodesic tangent to the geodesic sphere $\partial B(p, r)$ stays locally outside $B(p, r)$, for every $0 < r \leq r_0$.*

Proof. Use normal coordinates, that is represent $B(p, r)$ as $B(0, r) \subset \mathbb{R}^n$ for a small $r > 0$. For every $(x, v) \in B(0, r) \times S^{n-1}$ we have a geodesic $\gamma_{x,v} \colon J_{x,v} \to B(0, r)$ with $0 \in J_{x,v}$ and $\gamma'_{x,v}(0) = v$. Consider the smooth map

$$F(x, v) = \frac{\partial^2}{\partial t^2}\big(|\gamma_{x,v}(t)|^2\big)\Big|_{t=0}.$$

When $x = 0$, the geodesic is radial $\gamma_{0,v}(t) = tv$ and hence $F(0, v) = 2$. Therefore there is a $0 < r_0 < r$ such that $F(x, v) > 0$, and hence $|\gamma_{x,v}(t)|^2$ has a local minimum at $t = 0$, whenever $|x| \leq r_0$. This proves the lemma. □

Proposition 9.4.20. *For every point $p \in M$ there is a $r_0 > 0$ such that $B(p, r)$ is a strictly convex geodesic ball, for every $0 < r \leq r_0$.*

Proof. We know that there is a $r_1 > 0$ such that $B(p, r_1)$ is a geodesic ball and every geodesic tangent to the geodesic sphere $\partial B(p, r)$ stays locally outside the ball, for every $0 < r \leq r_1$.

Pick a $0 < r_0 < r_1/2$ such that every minimising geodesic $\gamma_{q,q'}$ with endpoints $q, q' \in \overline{B(p, r_0)}$ has length at most $r_1/2$. (We can do this because on a totally normal neighbourhood the minimising geodesic, and hence its length, varies smoothly on the points.) In particular $\gamma_{q,q'}$ is contained in $B(p, r_1)$.

If we represent $B(p, r_1)$ in normal coordinates, we see that the maximum of $|\gamma_{q,q'}(t)|^2$ must be at one of its endpoints, otherwise $\gamma_{q,q'}(t)$ would be tangent to a geodesic sphere locally from inside. Therefore $B(p, r)$ is strictly convex for every $r \leq r_0$. □

Convex subsets have two nice properties: they are closed under intersection, and they are contractible (exercise). These imply the following.

Proposition 9.4.21. *Every smooth manifold M has a locally finite covering $\{U_i\}$ such that every non-empty finite intersection of U_i's is contractible.*

Proof. Put an arbitrary metric on M and use convex neighbourhoods. □

9.5. Completeness

A riemannian manifold M is also a metric space, so it makes perfectly sense to consider whether it is *complete* or not – a notion that is meaningless for unstructured smooth manifolds. We prove here the *Hopf – Rinow Theorem*, that shows that completeness may actually be stated in equivalent ways, one of which involves only geodesics.

9.5.1. Geodesically complete manifolds. Let M be a riemannian manifold. We say that M is *complete* if its underlying metric space is. We say that M is *geodesically complete* if the exponential map $\exp_p$ is defined on the full tangent space for all $p \in M$. Equivalently, we are asking that every maximal geodesic $\gamma(t)$ in M be defined for all times $t \in \mathbb{R}$.

Recall that the distance $d(p, q)$ of two points $p, q \in M$ is the infimum of the lengths of all the curves γ joining p and q; if such an infimum is realised by γ, then γ is called *minimising* and we have discovered in the last section that γ must be a geodesic (after a reparametrisation). Here is one nice consequence of geodesical completeness:

Proposition 9.5.1. *If M is connected and geodesically complete, every two points $p, q \in M$ are joined by a minimising geodesic.*

Proof. Pick a geodesic ball $B(p, r)$ at p, with geodesic sphere $\partial B(p, r)$. If $q \in B(p, r)$ we are done. Otherwise, let $p_0 \in \partial B(p, r)$ be a point at minimum distance from q. Let $v \in T_pM$ be the unique vector with $\|v\| = 1$ and $\gamma_v(r) = p_0$.

By hypothesis, the geodesic $\gamma_v(t) = \exp_p(tv)$ exists for all $t \in \mathbb{R}$. Set $d = d(p, q)$. We now show that $\gamma_v(d) = q$. To do so, let $I \subset [0, d]$ be the subset of all times t such that $d(\gamma_v(t), q) = d - t$. This set is non-empty and closed, and using Theorem 9.4.17 we easily see that it is also open (exercise). Therefore $I = [0, d]$ and we are done. $\square$

Corollary 9.5.2. *If M is connected and geodesically complete, the exponential map $\exp_p \colon T_pM \to M$ is surjective at every $p \in M$.*

The exponential map $\exp_p$ of a geodesically complete riemannian manifold M sends the tangent space T_pM onto the whole manifold M. Recall that $\exp_p$ is a local diffeomorphism at the origin, but it may not be as nice at the other points.

9.5.2. The Hopf – Rinow Theorem. The following important theorem says that different notions of completeness are actually equivalent.

Theorem 9.5.3 (Hopf – Rinow). *Let M be a connected riemannian manifold. The following are equivalent:*

1. *M is geodesically complete,*
2. *a subset $K \subset M$ is compact $\iff$ it is closed and bounded;*
3. *M is complete.*

Proof. (1)$\Rightarrow$(2). Let $K \subset M$ be a subset. Compact always implies closed and bounded, so we prove the converse. Take a point $p \in M$. By hypothesis $\exp_p \colon T_pM \to M$ is surjective. If K is closed and bounded, there is a $r > 0$ such that $K \subset B(p, r)$ and hence K is contained in the compact set $\exp_p(\overline{B(p, r)})$. Since K is closed there, it is also compact.

(2)⇒(3). Every Cauchy sequence is bounded, so it has compact closure. Therefore it contains a converging subsequence, and hence it converges.

(3)⇒(1). Let $\gamma\colon I \to M$ be a maximal geodesic. We know that I is open, and since M is complete it is also closed: if $t_i \in I$ converges to some $t \in \mathbb{R}$, then $\gamma(t_i)$ is a Cauchy sequence and converges to some $p \in M$. Pick a totally normal neighbourhood V containing p. Every geodesic in V intersects ∂V; this implies that γ can be pursued on and hence $t \in I$. □

Corollary 9.5.4. *Compact riemannian manifolds are geodesically complete.*

Of course many interesting complete manifolds are not compact, for instance $\mathbb{R}^n$ and all the closed unbounded submanifolds in $\mathbb{R}^n$.

Corollary 9.5.5. *Every closed submanifold of a geodesically complete riemannian manifold is also geodesically complete.*

Corollary 9.5.6. *Every smooth manifold has a geodesically complete riemannian metric.*

Proof. By Whitney's Embedding Theorem, it is diffeomorphic to a closed submanifold of $\mathbb{R}^n$. □

9.6. Curvature

How can we distinguish two riemannian manifolds? Globally, they may have different topologies – and this is often detected by invariants like the fundamental group or De Rham cohomology – so we are now interested in constructing some *local* invariants. Can we measure locally how a riemannian manifold differs from being the familiar Euclidean space?

The answer to all these questions is *curvature*, and the most complete answer is a formidable tensor field called the *Riemann curvature tensor*. This tensor field is pretty complicated and one sometimes wish to examine some more reasonable tensor fields obtained from it via appropriate contractions: these are the *Ricci tensor* and finally the *scalar curvature*. A more geometric invariant which is in fact equivalent to the Riemann curvature tensor is the *sectional curvature*.

9.6.1. The Riemann curvature tensor. Let M be a riemannian manifold, equipped with its Levi-Civita connection ∇. We have already experienced with the torsion tensor T that one of the most efficient and natural ways to encode some information from ∇ is to build an appropriate tensor field. Tensor fields are lovely because they furnish some precise data at every single point $p \in M$. Of course the torsion tensor is useless here, since $T \equiv 0$ by assumption, so we must look for something else.

Recall that a tensor field of type $(1, n)$ on M is a multilinear map

$$\underbrace{T_pM \times \cdots \times T_pM}_{n} \longrightarrow T_pM$$

that depends smoothly on p.

Definition 9.6.1. The *Riemann curvature tensor* R is a tensor field on M of type $(1,3)$ defined as follows. For every point $p \in M$ and vectors $u, v, w \in T_pM$ we set

$$R(p)(u,v,w) = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y](p)} Z$$

where X, Y, Z are vector fields extending u, v, w on some neighbourhood of p.

Of course it is crucial here to prove that the definition is well-posed:

Proposition 9.6.2. *The tangent vector $R(p)(u,v,w)$ is independent of the extensions X, Y, and Z.*

Proof. Armed with patience and optimism, we write everything in coordinates and get

$$\begin{aligned}
\nabla_X \nabla_Y Z &= \nabla_X \left(Y^i \frac{\partial Z^k}{\partial x_i} e_k + Y^i Z^j \Gamma^k_{ij} e_k \right) \\
&= X^j \frac{\partial Y^i}{\partial x_j} \frac{\partial Z^k}{\partial x_i} e_k + X^j Y^i \frac{\partial^2 Z^k}{\partial x_j \partial x_i} e_k + X^j Y^i \frac{\partial Z^l}{\partial x_i} \Gamma^k_{jl} e_k \\
&\quad + X^m \frac{\partial Y^i}{\partial x_m} Z^j \Gamma^k_{ij} e_k + X^m Y^i \frac{\partial Z^j}{\partial x_m} \Gamma^k_{ij} e_k + X^m Y^i Z^j \frac{\partial \Gamma^k_{ij}}{\partial x_m} e_k \\
&\quad + X^m Y^i Z^j \Gamma^l_{ij} \Gamma^k_{lm} e_k.
\end{aligned}$$

If we calculate the difference $\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z$ the terms number 2, 3, and 5 cancel, and the terms 1 and 4 form precisely the expression

$$[X,Y]^i \frac{\partial Z^k}{\partial x_i} e_k + [X,Y]^i Z^j \Gamma^k_{ij} e_k = \nabla_{[X,Y]} Z.$$

From this we deduce that $R(p)(u,v,w)$ consists only of the terms number 6 and 7 that depend (linearly) on u, v, and w and not on their extensions. The proof is complete. $\square$

The tensor field R is therefore well-defined. To check that it is indeed smooth, we work on a chart and note that during the proof we have also found implicitly the coordinates of R in terms of the Christoffel symbols and their derivatives. After renaming indices we get

$$R^i{}_{jkl} = R(e_k, e_l, e_j)^i = \frac{\partial \Gamma^i_{lj}}{\partial x_k} - \frac{\partial \Gamma^i_{kj}}{\partial x_l} + \Gamma^i_{km} \Gamma^m_{lj} - \Gamma^i_{lm} \Gamma^m_{kj}. \tag{27}$$

In particular $R^i{}_{jkl}$ depends smoothly on the point. Note the particular ordering chosen for the indices: the letters j, k, l correspond to the vectors w, u, v respectively. Therefore we have

$$R(u,v,w)^i = R^i{}_{jkl} u^k v^l w^j.$$

The only example we make for the moment is rather trivial.

Example 9.6.3. If an open set $U \subset \mathbb{R}^n$ is equipped with the Euclidean metric, then $\Gamma_{ij}^k = 0$ and therefore $R^i{}_{jkl} = 0$ vanishes everywhere.

It is important to keep in mind that the definition of R is intrinsic, that is it only depends on the metric g and on nothing else: this implies for instance that the tensor field R is preserved by any isometry.

As every tensor field, the Riemann tensor gives a $C^\infty(M)$-multilinear map

$$R\colon \mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) \longrightarrow \mathfrak{X}(M)$$

that can be written elegantly as

$$R(X, Y, Z) = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z.$$

It is sometimes useful to consider another version of the Riemann tensor, where all the indices are in lower position:

$$R_{ijkl} = R^m{}_{jkl} g_{im}.$$

In this version the Riemann tensor is a tensor of type $(0, 4)$. Of course we can transform it back to the original $(1, 3)$ tensor using g^{lm}, so there is no loss of information in using one version instead of the other.

9.6.2. Holonomy along small quadrilaterals. The Riemann tensor has a simple and intuitive geometric interpretation that we now describe. Let $u, v \in T_pM$ be two tangent vectors. It is always possible (exercise: pick a chart) to extend them locally to two commuting vector fields X and Y. Extend w to any vector field Z. Now the formula simplifies

$$R(p)(u, v, w) = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z.$$

For sufficiently small $t > 0$, let $\gamma_{s,t}$ be the closed loop based in p constructed ad in Figure 9.4-(left) as the concatenation of four integral curves of X, Y, $-X$, and $-Y$, lasting precisely the time s, t, s, t respectively. Of course the loop $\gamma_{s,t}$ closes up because the two vector fields (and hence their flows) commute. We can now parallel-transport the vector w along the curve $\gamma_{s,t}$ as shown in the figure, to find at the end a new vector $h_{s,t}(w) \in T_pM$, called the *holonomy* of w along $\gamma_{s,t}$. TBD?

Exercise 9.6.4. We have

$$h_{s,t}(w) = w + R(u, v, w)st + O(s^2 + t^2).$$

Note the analogies with Proposition 5.4.10. The endomorphism

$$R(u, v, \cdot)\colon T_pM \longrightarrow T_pM$$

whose coordinates are $R^i{}_{jkl} u^k v^l$ measures the second-order deviation (from the identity) of the holonomy along a small quadrilateral with sides determined by u and v.

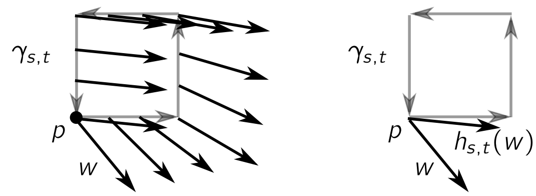

Figure 9.4. Given two commuting vector fields X and Y extending v and w, for every small $s, t > 0$ a quadrilateral loop $\gamma_{s,t}$ based in p is defined as the concatenation of four integral curves of $X, Y, -X$, and $-Y$ that last precisely the time s, t, s, t respectively. On a chart we may write X and Y as two coordinate vector fields, so that γ_t is a rectangle of sides $s \times t$ as in the picture. The holonomy along $\gamma_{s,t}$ is the parallel transport along $\gamma_{s,t}$ (left). The Riemann tensor measures the second-order deviation of the holonomy along $\gamma_{s,t}$ from being the identity (right).

9.6.3. Normal coordinates. Recall from Section 9.4.4 that the exponential map $\exp_p$ furnishes some nice normal coordinates around each point $p \in M$, such that $g_{ij} = \delta_{ij}$ and $\Gamma^k_{ij} = 0$ at the point. In these coordinates the expression (27) simplifies and we get

$$R^i{}_{jkl} = \frac{\partial \Gamma^i_{lj}}{\partial x_k} - \frac{\partial \Gamma^i_{kj}}{\partial x_l}. \tag{28}$$

Of course this equation is valid only at the point p. We can also deduce a reasonable expression for R_{ijkl} directly in terms of the metric tensor:

Proposition 9.6.5. *At the point p, in normal coordinates we have*

$$R_{ijkl} = \frac{1}{2}\left(\frac{\partial^2 g_{il}}{\partial x_j \partial x_k} + \frac{\partial^2 g_{jk}}{\partial x_i \partial x_l} - \frac{\partial^2 g_{jl}}{\partial x_i \partial x_k} - \frac{\partial^2 g_{ik}}{\partial x_j \partial x_l}\right). \tag{29}$$

Proof. Recall that in normal coordinates the first derivative of g in p vanishes. We get

$$\begin{aligned}
R_{ijkl} &= g_{im} R^m{}_{jkl} = g_{im}\left(\frac{\partial \Gamma^m_{lj}}{\partial x_k} - \frac{\partial \Gamma^m_{kj}}{\partial x_l}\right) \\
&= \frac{1}{2} g_{im} g^{hm}\left(\frac{\partial}{\partial x_k}\left(\frac{\partial g_{hj}}{\partial x_l} + \frac{\partial g_{hl}}{\partial x_j} - \frac{\partial g_{jl}}{\partial x_h}\right) - \frac{\partial}{\partial x_l}\left(\frac{\partial g_{hj}}{\partial x_k} + \frac{\partial g_{hk}}{\partial x_j} - \frac{\partial g_{jk}}{\partial x_h}\right)\right) \\
&= \frac{1}{2}\left(\frac{\partial^2 g_{il}}{\partial x_j \partial x_k} + \frac{\partial^2 g_{jk}}{\partial x_i \partial x_l} - \frac{\partial^2 g_{jl}}{\partial x_i \partial x_k} - \frac{\partial^2 g_{ik}}{\partial x_j \partial x_l}\right).
\end{aligned}$$

The proof is complete. □

Note the absence of repeated indices: the element R_{ijkl} is just the sum of four second partial derivatives of the metric g. Of course the use of normal coordinates is crucial here.

Recall that in normal coordinates we have $g_{ij}(0) = \delta_{ij}$ and $\frac{\partial g_{ij}}{\partial x_k}(0) = 0$. The first interesting terms in the Taylor expansion for g_{ij} are the second order derivatives, and these are determined precisely by the Riemann tensor:

Proposition 9.6.6. *In normal coordinates we have*

$$g_{ij}(x) = \delta_{ij} - \frac{1}{3}R_{ikjl}(0)x^k x^l + O(|x|^3).$$

Proof. By Proposition 9.4.11 we have

$$\frac{\partial \Gamma^i_{jk}}{\partial x_l}(0) + \frac{\partial \Gamma^i_{kl}}{\partial x_j}(0) + \frac{\partial \Gamma^i_{lj}}{\partial x_k}(0) = 0.$$

Combining this with (28) we get

$$R_{ijkl}(0) + R_{ikjl}(0) = R^i{}_{jkl}(0) + R^i{}_{kjl}(0) = -3\frac{\partial \Gamma^i_{jk}}{\partial x_l}(0).$$

Now we write the Taylor expansion

$$\begin{aligned}
g_{ij}(x) &= \delta_{ij} + \frac{1}{2}\frac{\partial^2 g_{ij}}{\partial x_l \partial x_k}(0)x^k x^l + O(|x|^3)\\
&= \delta_{ij} + \frac{1}{2}\frac{\partial}{\partial x_l}\left(\Gamma^m_{ki}g_{mj} + \Gamma^m_{kj}g_{mi}\right)\Big|_{x=0} x^k x^l + O(|x|^3)\\
&= \delta_{ij} + \frac{1}{2}\left(\frac{\partial \Gamma^j_{ki}}{\partial x_l}(0) + \frac{\partial \Gamma^i_{kj}}{\partial x_l}(0)\right) x^k x^l + O(|x|^3)\\
&= \delta_{ij} - \frac{1}{6}\left(R_{jkil}(0) + R_{jikl}(0) + R_{ikjl}(0) + R_{ijkl}(0)\right) x^k x^l + O(|x|^3)\\
&= \delta_{ij} - \frac{1}{6}\left(R_{iljk}(0) + R_{ikjl}(0)\right) x^k x^l + O(|x|^3)\\
&= \delta_{ij} - \frac{1}{3}R_{ikjl}(0)x^k x^l + O(|x|^3).
\end{aligned}$$

We have used that $g_{ij}(0) = \delta_{ij}$, $\frac{\partial g_{ij}}{\partial x_k}(0) = 0$, and the equalities $R_{jikl} + R_{ijkl} = 0$ and $R_{jkil} = R_{iljk}$ are easy consequences of Proposition 9.6.5. □

In normal coordinates, the Riemann tensor measures the second-order deviation of g_{ij} from the Euclidean metric δ_{ij}.

9.6.4. Symmetries. Being a $(1, 3)$-tensor field, we expect the Riemann tensor R to contain a tremendous amount of information on g, and this is what really happens. To help mastering this huge amount of data, we start by unraveling some symmetries.

Proposition 9.6.7. *The following symmetries hold in any coordinate chart:*

(1) $R_{ijkl} = -R_{jikl} = -R_{ijlk}$,
(2) $R_{ijkl} = R_{klij}$,
(3) $R^i{}_{jkl} + R^i{}_{klj} + R^i{}_{ljk} = 0$.

Before entering in the proof, note that these symmetries may be stated more intrinsically as follows: for every $p \in M$ and $u, v, w, z \in T_pM$ we get

(1) $R(p)(u, v, w, z) = -R(p)(v, u, w, z) = -R(p)(u, v, z, w)$,
(2) $R(p)(u, v, w, z) = R(p)(w, z, u, v)$,
(3) $R(p)(u, v, w) + R(p)(v, w, u) + R(p)(w, u, v) = 0$.

In the first two we interpret R as a $(0, 4)$ tensor field, while in the last we take the original $(1, 3)$ tensor field. We will use R slightly ambiguously in this way.

Proof. To prove the intrinsic version of the symmetries, we may take some normal coordinates at p. There R_{ijkl} has the convenient expression (29), which displays (1) and (2) immediately. Analogously for $R^i{}_{jkl}$ we use (28) to deduce (3) easily. The proof is complete. □

9.6.5. Sectional curvature. What kind of geometric information can we get from the Riemann tensor R? One answer to this question passes through the definition of *sectional curvature*.

Let M as usual be a Riemannian manifold and R be its Riemann curvature tensor field in the $(0, 4)$ version. Let $p \in M$ be a point and $\sigma \subset T_pM$ be a two dimensional linear subspace, that is a plane passing through the origin. We now assign to σ a number $K(\sigma)$ called the *sectional curvature* along σ, as follows.

Let $u, v \in \sigma$ be arbitrary generators. We define

$$K(\sigma) = \frac{R(p)(u, v, u, v)}{A^2(u, v)}$$

where

$$A^2(u, v) = \|u\|^2\|v\|^2 - \langle u, v\rangle^2$$

is the square of the area of the parallelogram spanned by u and v.

Proposition 9.6.8. *The sectional curvature $K(\sigma)$ is well-defined.*

Proof. The quantity $K(\sigma)$ does not change if we substitute (u, v) with one of the following:

$$(v, u), \qquad (\lambda u, v), \qquad (u + \lambda v, v).$$

By composing such moves we can transform (u, v) into any other basis. □

The Riemann tensor of course determines the sectional curvatures by definition; we now see that also the converse holds:

Proposition 9.6.9. *The sectional curvatures $K(\sigma)$ along planes $\sigma \subset T_pM$ determine the Riemann tensor $R(p)$.*

Proof. The sectional curvatures determine $R(p)(u, v, u, v)$ for all pairs of vectors $u, v \in T_pM$. The vector $R(p)(u + w, v, u + w, v)$ is therefore determined, and it equals

$$R(p)(u, v, u, v) + 2R(p)(u, v, w, v) + R(p)(w, v, w, v).$$

Therefore the sectional curvatures also determine $R(p)(u, v, w, v)\ \ \forall u, v, w$. Analogously, the vector $R(p)(u, v + z, w, v + z)$ is determined and it equals

$$R(p)(u, v, w, v) + R(p)(u, v, w, z) + R(p)(u, z, w, v) + R(p)(u, z, w, z)$$

so the sectional curvatures determine the value of

$$R(p)(u, v, w, z) + R(p)(u, z, w, v) = R(p)(u, v, w, z) - R(p)(u, z, v, w)$$

for all u, v, w, z. If we look at the three numbers

$$R(p)(u, v, w, z), \quad R(p)(u, w, z, v), \quad R(p)(u, z, v, w)$$

we see that their sum is zero and their differences are determined: hence the three numbers are also determined. □

Therefore we are not losing any information if we consider sectional curvatures instead of the Riemann tensor. Sectional curvatures have a nice geometric interpretation that we will describe soon. For the time being, we keep on manipulating the Riemann tensor.

9.6.6. Ricci tensor. The Riemann curvature tensor R is a tensor of type $(1, 3)$ and it is of course natural to study its contractions, that are tensor fields of type $(0, 2)$. There are three possible contractions of ${R^i}_{jkl}$, namely:

$$R^k{}_{kij}, \qquad R^k{}_{ikj}, \quad \text{and} \quad R^k{}_{ijk}.$$

Using the symmetries of R we see easily that the first vanishes and the remaining two differ only by a sign. Therefore there is essentially only one way to get a non-trivial tensor by contraction, and this yields the *Ricci tensor*:

$$R_{ij} = R^k{}_{ikj}.$$

This is a tensor field of type $(2, 0)$. Since Ricci has the same initial as Riemann, we still indicate it by R. To distinguish which is which it suffices to look at the number of indices, or arguments. The Ricci tensor of course also defines a $C^\infty(M)$-bilinear map

$$R\colon \mathfrak{X}(M) \times \mathfrak{X}(M) \longrightarrow C^\infty(M).$$

Proposition 9.6.10. *The Ricci tensor is symmetric.*

Proof. We have

$$R_{ij} = R^k{}_{ikj} = R_{hikj}g^{hk} = R_{kjhi}g^{hk} = R^h{}_{jhi} = R_{ji}.$$

The proof is complete. □

Like the metric tensor, the Ricci tensor is a symmetric tensor field of type $(0, 2)$. Note however that the Ricci tensor needs not to be positive-definite and may also be degenerate: indeed, on an open set $U \subset \mathbb{R}^n$ with the Euclidean metric, all the tensors that we introduce vanish, including Ricci.

What geometric information is carried by the Ricci tensor? In normal coordinates, it measures the first-order variation of the determinant of g_{ij}.

Proposition 9.6.11. *In normal coordinates we have*

$$\det g_{ij}(x) = 1 - \frac{1}{3}R_{ij}(0)x^i x^j + O(|x|^3).$$

Proof. Recall that for any $n \times n$ matrix A we have

$$\det(I + A) = 1 + \mathrm{tr}A + O(|A|^2).$$

Combining this with Proposition 9.6.6 we get

$$\det g_{ij}(x) = 1 - \frac{1}{3}R^i{}_{kil}(0)x^k x^l + O(|x|^3) = 1 - \frac{1}{3}R_{kl}(0)x^k x^l + O(|x|^3).$$

The proof is complete. □

Let ω be the volume form determined by g. As a consequence, the Ricci tensor measures the first-order variation of ω.

Corollary 9.6.12. *In normal coordinates we have*

$$\omega = \left(1 - \frac{1}{6}R_{ij}(0)x^i x^j + O(|x|^3)\right) dx^1 \wedge \cdots \wedge dx^n.$$

Proof. This follows by applying the formula

$$\omega = \sqrt{\det g_{ij}}dx^1 \wedge \cdots \wedge dx^n$$

together with $\sqrt{1+t} = 1 + \frac{1}{2}t + O(|t|^2)$. □

Remark 9.6.13. By the spectral theorem, at every point $p \in M$ we can find a basis for T_pM that is orthonormal for g_{ij} and orthogonal for R_{ij}. Therefore we can choose normal coordinates at p where $g_{ij}(0) = \delta_{ij}$ and $R_{ij}(0)$ is a diagonal matrix.

9.6.7. Scalar curvature. If you think that a tensor of type $(0, 2)$ is yet too complicated an invariant, you can still contract it and get an interesting number, called the *scalar curvature*.

The *scalar curvature* of a Riemannian manifold M at a point $p \in M$ is

$$R = g^{ij}R_{ij}.$$

This is the trace of the Ricci tensor; note that we need the metric g to raise an index in order to define the trace of a tensor of type $(0, 2)$ unambiguously. The scalar curvature is still indicated with the same letter R as the Riemann and Ricci curvature: the number of indices is enough to understand which is which.

What geometric information carries the scalar curvature? It brings some information on the volumes of small geodesic balls. Let $p \in M$ be a point and $B(p, r)$ a geodesic ball of radius r centered at p (remember that this notion is well defined only for sufficiently small $r > 0$). We recall that the volume of a Euclidean ball $B(0, r) \subset \mathbb{R}^n$ is

$$\mathrm{Vol}\big(B(0, r)\big) = V_n(r) = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2} + 1\right)} r^n$$

where Γ is Euler's gamma function.

Let $p \in M$ be any point in a Riemannian manifold M, and $B(p,r)$ a geodesic ball.

Proposition 9.6.14. *We have*

$$\operatorname{Vol}\big(B(p,r)\big) = V_n(r)\cdot\left(1-\frac{1}{6(n+2)}R(p)r^2+O(r^4)\right). \tag{30}$$

Proof. Following Remark 9.6.13, we work in normal coordinates around $p=0$ where the Ricci tensor $R_{ij}(0)$ is diagonal with entries $\lambda_1,\dots,\lambda_n$. The scalar curvature is its trace $R(0)=\lambda_1+\cdots+\lambda_n$. We have

$$\operatorname{Vol}\big(B(0,r)\big) = \int_{B(0,r)}\omega = \int_{B(0,r)}\left(1-\frac{1}{6}R_{ij}(0)x^ix^j+O(|x|^3)\right)dx^1\wedge\cdots\wedge dx^n.$$

We now compute

$$\begin{aligned}\int_{B(0,r)}R_{ij}(0)x^ix^j dx^1\wedge\cdots\wedge dx^n &= \int_{B(0,r)}(\lambda_1x_1^2+\cdots+\lambda_nx_n^2)dx^1\wedge\cdots\wedge dx^n\\ &=\sum_{i=1}^n\lambda_i\int_{B(0,r)}x_i^2dx^1\wedge\cdots\wedge dx^n\\ &=\left(\sum_{i=1}^n\lambda_i\right)\frac{1}{n}\int_{B(0,r)}\rho^2dx^1\wedge\cdots\wedge dx^n\end{aligned}$$

where $\rho^2=x_1^2+\cdots x_n^2$. Let $d\Omega$ be the volume form in the Euclidean S^{n-1}. The last expression equals

$$\begin{aligned}\frac{R(0)}{n}\int_{B(0,r)}\rho^2\cdot\rho^{n-1}d\rho\wedge d\Omega &= \frac{R(0)}{n}\left(\int_0^r\rho^{n+1}d\rho\right)\left(\int_{S^{n-1}}d\Omega\right)\\ &=\frac{R(0)}{n}\cdot\frac{r^{n+2}}{n+2}\operatorname{Vol}(S^{n-1}).\end{aligned}$$

With similar calculations, the volume of the Euclidean ball of radius r is

$$V_n(r)=\frac{r^n}{n}\operatorname{Vol}(S^{n-1})$$

and therefore the last expression equals

$$\frac{R(0)}{n+2}V_n(r)r^2.$$

Finally, we get

$$\operatorname{Vol}\big(B(0,r)\big)=V_n(r)\left(1-\frac{R(0)}{6(n+2)}r^2+O(r^4)\right).$$

The proof is complete. □

The scalar curvature measures (at the second order) the ratio between volumes of small geodesic balls and Euclidean balls with the same small radius. Note that this is an intrinsic property of a point $p \in M$, that is not dependent on a particular coordinate systems. In particular, if $R(p)$ is negative (respectively, positive), geodesic balls of small radius r centered at p have larger (respectively, smaller) volume than the Euclidean ones with the same radius r.

Example 9.6.15. On a surface, the equation (30) becomes

$$\mathrm{Vol}\big(B(p,r)\big) = \pi r^2 \left(1 - \frac{R(p)}{24} r^2 + O(r^4)\right) = \pi r^2 - \frac{R(p)}{24}\pi r^4 + O(r^6).$$

On a 3-manifold, we get

$$\mathrm{Vol}\big(B(p,r)\big) = \frac{4}{3}\pi r^3 \left(1 - \frac{R(p)}{30} r^2 + O(r^4)\right) = \frac{4}{3}\pi r^3 - \frac{2R(p)}{45}\pi r^5 + O(r^7).$$

9.6.8. Flatness. We have already noticed that the Riemann tensor measures the local deviation of g from the Euclidean metric. We now show that it does so in a complete way: we prove that a metric tensor g is locally Euclidean if and only if the Riemann tensor vanishes. Let us first fix some definitions.

We say that a Riemannian manifold M is *Euclidean* if it is locally isometric to $\mathbb{R}^n$, that is every $p \in M$ has an open neighbourhood $U(p) \subset M$ which is isometric to some open subset of the Euclidean $\mathbb{R}^n$.

We say that M is *flat* if its Riemann tensor $R^i{}_{jkl}$ vanishes everywhere.

Theorem 9.6.16. *A Riemannian manifold M is Euclidean $\Longleftrightarrow$ it is flat.*

Proof. We already know that Euclidean implies flat, so we prove the converse. Pick a point in M and represent a small neighbourhood of it via normal coordinates $B(0,r) \subset \mathbb{R}^n$. Pick a small cube $(-\varepsilon, \varepsilon)^n$ contained in $B(0,r)$.

We now extend the orthonormal basis $e_1, \ldots, e_n$ at 0 to a frame on the cube, as follows: we first parallel-transport the basis along x_1, then along x_2, and so on until x_n. At the i-th step the frame is defined only on the slice $S_i = \{x_{i+1} = \ldots = x_n = 0\}$ of the cube, and at the end it is defined everywhere. It is smooth because parallel transport depends smoothly on the initial data. We have thus constructed a frame $X_1, \ldots, X_n$ that is an orthonormal basis at every point of the cube, such that $X_i(0) = e_i$. By construction we have

$$\nabla_{e_i} X_k = 0 \quad \text{on } S_i \quad \forall k.$$

We now prove that in fact

$$\nabla_{e_j} X_k = 0 \quad \text{on } S_i \quad \forall k, \forall j \le i.$$

We show this by induction on i. The case $i = 1$ is done, so we suppose that it holds for i and prove it for $i+1$. We already know that $\nabla_{e_{i+1}} X_k = 0$ on S_{i+1}. If $j \le i$, by our induction hypothesis we have $\nabla_{e_j} X_k = 0$ on the hyperplane

S_i. To conclude it suffices to check that $\nabla_{e_{i+1}}(\nabla_{e_j} X_k) = 0$ on S_{i+1}. The coordinate fields $e_1, \ldots, e_n$ commute, hence flatness gives

$$\nabla_{e_{i+1}}(\nabla_{e_j} X_k) = \nabla_{e_j}(\nabla_{e_{i+1}} X_k) = \nabla_{e_j}(0) = 0.$$

The inductive proof is completed and when $i = n$ it shows that

$$\nabla_{e_j} X_k = 0 \quad \forall k, j$$

everywhere on the cube. Since ∇ is symmetric we have

$$[X_i, X_j] = \nabla_{X_i} X_j - \nabla_{X_j} X_i = 0 - 0 = 0.$$

By Proposition 5.4.12 there is a chart $\varphi\colon U \to V$ with $U \subset (-\varepsilon, \varepsilon)^n$ that straightens these vector fields, that is that transports X_i into e_i. The map φ is an isometry between U and V with its Euclidean metric, because it sends pointwise an orthonormal basis $X_1, \ldots, X_n$ into the orthonormal basis $e_1, \ldots, e_n$. □

9.6.9. Low dimensions. In dimensions 2 and 3 the information carried by the curvature tensors reduce considerably and is more manageable.

Let S be a surface equipped with a Riemannian metric. At every point $p \in S$ the tangent plane T_pS has a sectional curvature $K(p)$, and the whole Riemann tensor is determined by this number by Proposition 9.6.9. Therefore all the information encoded by the Riemann tensor reduces to a more comfortable smooth function $K\colon S \to \mathbb{R}$, which is in fact equal to the scalar curvature R: on an orthonormal basis e_1, e_2 for T_pS we get

$$K(p) = R_{1212} = R(p).$$

Let M be a Riemannian 3-manifold. At a point $p \in M$ we fix an orthonormal basis e_1, e_2, e_3 for T_pM and note that the components R_{ijkl} of the Riemann tensor are determined by the Ricci tensor: at least two of the four indices i, j, k, l must coincide and therefore R_{ijkl} is either zero or equal to an entry of the Ricci tensor R_{ij}. Summing up, we have discovered the following.

Proposition 9.6.17. *The Riemann curvature tensor is determined by the scalar curvature in dimension $n = 2$ and by the Ricci tensor in dimension $n = 3$.*

CHAPTER 10

Lie groups

A *Lie group* is a group that is also a smooth manifold. Lie groups are everywhere: most symmetry groups that one encounters in geometry are naturally Lie groups. The fundamental examples are matrix groups like $\mathrm{GL}(n, \mathbb{R})$ and $\mathrm{O}(n)$.

10.1. Basics

We define the Lie groups and start to investigate their properties.

10.1.1. Definition. A *Lie group* is a smooth manifold G equipped with a group structure, such that the multiplication and inverse maps

$$G \times G \longrightarrow G, \qquad (g, h) \longmapsto gh,$$
$$G \longrightarrow G, \qquad g \longmapsto g^{-1}$$

are both smooth. This is equivalent to requiring the map $G \times G \to G$, $(g, h) \mapsto gh^{-1}$ to be smooth.

Here are some important examples.

Example 10.1.1 (Abelian). The first examples of Lie groups are $\mathbb{R}^n$ with the sum operation and S^1 with the product, where we see $S^1 \subset \mathbb{C}$ as the unit complex numbers. These Lie groups are abelian.

Example 10.1.2 (Linear and orthogonal groups). A more elaborated and equally important example is the *general linear group* $\mathrm{GL}(n, \mathbb{R})$ of all $n \times n$ invertible matrices with the product operation. This Lie group contains also many other interesting Lie groups, such as the *special linear group* $\mathrm{SL}(n, \mathbb{R})$, the *orthogonal group* $\mathrm{O}(n)$, and the *special orthogonal group* $\mathrm{SO}(n)$. We studied the topology of these manifolds in Section 3.9.

Example 10.1.3 (Products). The product $G \times H$ of two Lie groups is naturally a Lie group. For instance, the n-torus $S^1 \times \cdots \times S^1$ is an abelian compact Lie group of dimension n.

Example 10.1.4 (Affine transformations). Another example is the group $\mathrm{Aff}(\mathbb{R}^n) = \mathrm{GL}(n, \mathbb{R}) \rtimes \mathbb{R}^n$ of all affine transformations of $\mathbb{R}^n$. As a set, we have $\mathrm{Aff}(\mathbb{R}^n) = \mathrm{GL}(n, \mathbb{R}) \times \mathbb{R}^n$ and we use this bijection to assign a smooth manifold structure to $\mathrm{Aff}(\mathbb{R}^n)$. The group structure is not a direct product, but the group operations are smooth nevertheless.

A Lie group of dimension 0 is called *discrete*. Every countable group G like $\mathbb{Z}$ may be given the structure of a Lie group by assigning it the discrete topology. Of course a discrete Lie group is connected if and only if it is trivial.

10.1.2. Homomorphisms. A *Lie group homomorphism* is a smooth homomorphism $f\colon G \to H$ between Lie groups. As usual, this is an *isomorphism* if f is invertible, that is if f is a diffeomorphism, and an *automorphism* if in addition $G = H$. For instance, every conjugation $G \to G, x \mapsto g^{-1}xg$ by some fixed element $g \in G$ is an auotomorphism of the Lie group G.

Example 10.1.5. The Lie groups S^1 and $\mathrm{SO}(2)$ are isomorphic, via the map

$$S^1 \longrightarrow \mathrm{SO}(2), \qquad e^{i\theta} \longmapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

10.1.3. Left and right multiplication. If $g \in G$, the *left and right multiplications by g* are the maps

$$L_g\colon G \to G, \quad x \mapsto gx,$$
$$R_g\colon G \to G, \quad x \mapsto xg.$$

Both maps are diffeomorphisms, with inverses $L_{g^{-1}}$ and $R_{g^{-1}}$, but are not Lie group isomorphisms, unless $g = e$. The maps L_g and $R_{g'}$ commute for all $g, g' \in G$. Conjugation by g is just $L_{g^{-1}} \circ R_g$.

10.1.4. Lie subgroups. Let G be a Lie group. A *Lie subgroup* of G is the image of any injective Lie group homomorphism $H \hookrightarrow G$ that is also an immersion. We identify H with its image and write $H < G$. For instance, $\mathrm{O}(n)$ is a Lie subgroup of $\mathrm{GL}(n, \mathbb{R})$.

We require H to be "injectively immersed" in G instead of the stronger and nicer "embedded" because we do not want to rule out the following types of Lie subgroups:

Example 10.1.6. Pick $\lambda \in \mathbb{R} \setminus \mathbb{Q}$ and consider the injective immersion $\mathbb{R} \to S^1 \times S^1, t \mapsto (e^{2\pi i t}, e^{2\pi i \lambda t})$. The image is a dense Lie subgroup of $S^1 \times S^1$. See Exercise 5.5.4.

The reason for allowing non-embedded Lie subgroups will be apparent in the next section. We exhibit more examples.

Example 10.1.7. The Lie group $\mathrm{Aff}(\mathbb{R}^n)$ may be embedded as a Lie subgroup of $\mathrm{GL}(n+1, \mathbb{R})$, by representing the affine transformation $x \mapsto Ax + b$ via the matrix

$$\begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}.$$

Example 10.1.8. The *Heisenberg group* is the Lie subgroup of $\mathrm{SL}(3,\mathbb{R})$ formed by all the matrices

$$\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}$$

where $a, b, c \in \mathbb{R}$ vary. It is diffeomorphic to $\mathbb{R}^3$, but it is not abelian.

10.1.5. Identity connected component. Let G be a Lie group. We denote by $G^0 \subset G$ the connected component of G containing the identity $e \in G$. The following may be seen as the first interesting result in Lie groups theory. The proof mixes topological and group theory arguments.

Proposition 10.1.9. *The component G^0 is a normal Lie subgroup.*

Proof. For every $g \in G$, the left multiplication L_g is a diffeomorphism and hence permutes the connected components of G. If $g \in G^0$, then L_g sends e to g and hence sends G^0 to itself. Therefore $gh \in G^0$ for all $g, h \in G^0$, so G^0 is closed under multiplication.

Analogously, the inverse map $g \mapsto g^{-1}$ permutes the connected components of G and fixes e, hence leaves G^0 invariant. Therefore G^0 is a subgroup. Along the same line, for every $g \in G$ the conjugation $x \mapsto g^{-1}xg$ is a diffeomorphism that fixes e and hence leaves G^0 invariant. So G^0 is normal. $\square$

The quotient $G/_{G^0}$ is naturally a discrete Lie group.

Example 10.1.10. We have $\mathrm{O}(n)^0 = \mathrm{SO}(n)$, while $\mathrm{GL}(n,\mathbb{R})^0$ consists of all invertible matrices with positive determinant.

10.1.6. Identity neighbourhoods. Let G be a Lie group. If $U, V \subset G$ are subsets, we construct more subsets as follows:

$$UV = \{uv \mid u \in U, v \in V\}, \qquad U^{-1} = \{u^{-1} \mid u \in U\}.$$

If U, V are neighbourhoods of the identity, then both UV and U^{-1} also are. We can use this to prove the following.

Proposition 10.1.11. *If G is connected, any neighbourhood U of the identity generates G.*

Proof. We can suppose that U is open and $U = U^{-1}$, otherwise we substitute U with $U \cap U^{-1}$. The subgroup generated by U is $H = \cup_{n=1}^{\infty} U^n$. Each U^n is open, so H is an open subgroup of G. Its left cosets are also open. Since G is connected, we get $G = H$. $\square$

10.1.7. Universal cover. Let G be a connected Lie group, and $\tilde{G}$ be its universal cover. We show that the Lie group structure lifts from G to $\tilde{G}$.

Proposition 10.1.12. *There is a natural Lie group structure on $\tilde{G}$ such that the cover $\pi\colon \tilde{G} \to G$ is a Lie groups homomorphism.*

Proof. We fix an arbitrary identity $\tilde{e} \in \pi^{-1}(e)$. Since $\tilde{G}$ is simply connected, both the product $G \times G \to G$ and the inversion $G \to G$ lift to two smooth maps $\tilde{G} \times \tilde{G} \to \tilde{G}$ and $\tilde{G} \to \tilde{G}$ between the universal covers, such that $(\tilde{e}, \tilde{e})$ goes to $\tilde{e}$ and $\tilde{e}$ goes to $\tilde{e}$, respectively. These define a product and inverse structure on $\tilde{G}$. Using the unique lift property of paths we can prove that these indeed satisfy the group axioms (exercise). $\square$

We have discovered that every connected Lie group has a universal cover. The universal cover of S^1 is of course $\mathbb{R}$. For $n \geq 3$, the *spin group* is defined as the universal cover of $\mathrm{SO}(n)$:

$$\mathrm{Spin}(n) = \widetilde{\mathrm{SO}(n)}.$$

10.1.8. Coverings. Let a *covering of Lie groups* be a homomorphism of connected Lie groups $G \to H$ that is also a smooth covering. The universal cover $\tilde{G} \to G$ constructed above is one example. In general, it is quite easy to understand when a Lie group homomorphism is a covering.

Proposition 10.1.13. *A Lie group homomorphism $f\colon G \to H$ between connected Lie groups is a smooth covering $\Longleftrightarrow$ df_e is invertible.*

Proof. The implication $\Rightarrow$ is obvious, so we prove $\Leftarrow$. Since df_e is invertible, there are open neighbourhoods U and V of $e \in G$ and $e \in H$ such that f maps diffeomorphically U to V.

For every $h \in H$, and every $g \in f^{-1}(h)$, we define

$$V_h = L_h(V), \qquad U_g = L_g(U).$$

These are open neighbourhoods of h and g, and one sees easily that

$$f^{-1}(V_h) = \bigsqcup_{g \in f^{-1}(h)} U_g.$$

The restriction of f to U_g is a diffeomorphism onto V_h, therefore f is a smooth covering. $\square$

Here is a concrete way to build coverings of Lie groups:

Proposition 10.1.14. *Let G be a Lie group and $\Gamma < Z(G)$ be a discrete central subgroup. The quotient $G/_\Gamma$ is naturally a Lie group and $G \to G/_\Gamma$ is a regular covering of Lie groups, with deck transformation group Γ.*

Proof. The action of Γ on G by multiplication is smooth, free, and properly discontinuous (exercise). Proposition 3.5.4 applies. $\square$

We now want to prove a converse of this proposition.

Proposition 10.1.15. *Let G be a connected Lie group. Every discrete normal subgroup $\Gamma \subset G$ is central.*

Proof. Pick $\gamma \in \Gamma$. For every $g \in G$, choose a path $g_t \in G$ connecting $g_0 = e$ and $g_1 = g$. By normality $g_t^{-1}\gamma g_t$ is a path in Γ, that must be constant, so $g^{-1}\gamma g = \gamma$ for all $g \in G$. □

Here is a converse for Proposition 10.1.14:

Proposition 10.1.16. *Every covering of Lie groups $G \to H$ is as in Proposition 10.1.14. That is, $\Gamma = \ker G$ is discrete and central and $H = G/_{\Gamma}$.*

Proof. The kernel Γ is the fibre of e and is hence discrete. Being also normal, it is central by the previous proposition. □

By assembling all our discoveries, we obtain the following.

Corollary 10.1.17. *Every connected Lie group is a quotient $G/_{\Gamma}$ of a simply connected Lie group G along some discrete central subgroup Γ.*

The classification of connected Lie groups hence reduces to the classification of simply connected ones (and their discrete central subgroups). The classification of simply connected Lie groups is hence a fundamental topological problem, that is elegantly transformed into an algebraic one through the fundamental notion of *Lie algebra* that we introduce in the next section.

We close our investigation with a corollary.

Corollary 10.1.18. *The fundamental group of every Lie group is abelian.*

10.2. Lie algebra

One of the most important aspects of Lie groups G is the leading role played by the tangent space T_eG at the identity $e \in G$, that has a natural structure of *Lie algebra*, see Definition 5.4.2.

10.2.1. Left-invariant vector fields. Let G be a Lie group. We now consider the tangent space T_eG at the identity $e \in G$. We note that for every $g \in G$ the differential of L_g yields an isomorphism

$$(dL_g)_e \colon T_eG \longrightarrow T_gG$$

on tangent spaces. Therefore we can use left-multiplication to identify canonically all the tangent spaces to T_eG, and this is a crucial aspect of Lie groups.

In particular, every fixed vector $v \in T_eG$ extends canonically to a vector field X in G by left-multiplication, as follows:

$$X(g) = (dL_g)_e(v).$$

The vector field X is *left-invariant*, that is it is invariant under the diffeomorphisms L_h, for all $h \in G$. Indeed we have

$$X(hg) = (dL_{hg})_e(v) = (dL_h)_g \circ (dL_g)_e(v) = (dL_h)_g(X(g)).$$

Every left-invariant vector field is clearly constructed in this way. We have obtained a natural isomorphism between T_eG and the subspace of $\mathfrak{X}(G)$ consisting of all the left-invariant vector fields. (Recall that $\mathfrak{X}(G)$ is the space of all vector fields in G.) We will henceforth identify these two spaces along this isomorphism.

By replacing L_g with R_g in the construction we would get analogously a natural isomorphism between T_eG and the subspace of all *right*-invariant vector fields. Note that a left-invariant vector field is not necessarily right-invariant, so the two subspaces of $\mathfrak{X}(G)$ may differ.

10.2.2. Parallelizability. The first important consequence that we can draw form our discovery is the following.

Proposition 10.2.1. *Every Lie group G is parallelizable.*

Proof. Every basis $v_1, \ldots, v_n$ of T_eG extends by left-multiplication to n left-invariant vector fields $X_1, \ldots, X_n$ on G that trivialise the bundle. □

Corollary 10.2.2. *Every Lie group G is orientable.*

10.2.3. Lie algebra. Let G be a Lie group. We have identified T_eG with the subspace of left-invariant vector fields in $\mathfrak{X}(G)$. We now note the following.

Proposition 10.2.3. *If $X, Y \in \mathfrak{X}(G)$ are left-invariant, then $[X, Y]$ also is.*

Proof. If two vector fields X, Y are invariant under some diffeomorphism, then their bracket also is. □

This observation shows that the space T_eG of all left-invariant vector fields is closed under the Lie bracket $[,]$. In other words T_eG is a *Lie subalgebra* of $\mathfrak{X}(G)$, and it is such an important object that it deserves a new symbol:

$$\mathfrak{g} = T_eG.$$

This is the *Lie algebra* of the Lie group G. The Lie algebra of Lie groups like $\mathrm{GL}(n, \mathbb{R})$, $\mathrm{O}(n)$, etc. is usually denoted as $\mathfrak{gl}(n, \mathbb{R})$, $\mathfrak{o}(n)$, etc.

10.2.4. Examples. On $\mathbb{R}^n$, a vector field is left-invariant if and only if it is constant, and the bracket of two constant vector fields is zero. Therefore the Lie algebra of $\mathbb{R}^n$ is $\mathbb{R}^n$ with the trivial Lie bracket. A Lie algebra with trivial Lie bracket is called *abelian*.

Analogously, the Lie algebra of S^1 is $\mathbb{R}$ with trivial Lie bracket. The Lie algebra of a product of Lie groups is just the product of their Lie algebras: in particular the Lie algebra of $S^1 \times \cdots \times S^1$ is again $\mathbb{R}^n$ with the trivial Lie bracket.

A more interesting example is $\mathrm{GL}(n, \mathbb{R})$. Being an open subset of the vector space $M(n)$ of all $n \times n$ matrices, its Lie algebra $\mathfrak{gl}(n, \mathbb{R})$ is $M(n)$ as a vector space, and we only need to understand the Lie bracket.

Proposition 10.2.4. *The Lie bracket of* $A, B \in \mathfrak{gl}(n, \mathbb{R})$ *is*

$$[A, B] = AB - BA.$$

Proof. Since $\mathrm{GL}(n, \mathbb{R})$ is an open subset of $M(n)$, a vector field is simply a map $\mathrm{GL}(n, \mathbb{R}) \to M(n)$. Every vector $A \in M(n)$ tangent at the origin extends by left-multiplication to the vector field $X \mapsto XA$. Similarly to Exercise 5.4.7, one can check (exercise) that the bracket of two vector fields $X \mapsto XA$ and $X \mapsto XB$ is $X \mapsto X(AB - BA)$. □

In particular, the Lie algebra $\mathfrak{gl}(n, \mathbb{R})$ is non-abelian as soon as $n \geq 2$.

10.2.5. Homomorphisms. Every Lie group homomorphism $f: G \to H$ induces a linear map $f_*: \mathfrak{g} \to \mathfrak{h}$ which is just the differential $f_* = df_e$.

Proposition 10.2.5. *The map* $f_* : \mathfrak{g} \to \mathfrak{h}$ *is a Lie algebra homomorphism.*

Proof. The homomorphism f commutes with left-multiplication, that is

$$f \circ L_g = L_{f(g)} \circ f$$

for every $g \in G$. This implies that a left-invariant vector field $X \in \mathfrak{g}$ and its image $f_*(X) \in \mathfrak{h}$ are f-related. Exercise 5.4.9 says that for every $X, Y \in \mathfrak{g}$ the vector fields $[X, Y]$ and $[f_*(X), f_*(Y)]$ are also f-related, so $f_*([X, Y]) = [f_*(X), f_*(Y)]$ as required. □

During the proof we have also discovered that for every $X \in \mathfrak{g}$ the vector fields X and $f_*(X)$ are f-related.

10.2.6. Lie subgroups. A Lie subgroup $H < G$ is by definition the image of an injective immersion and homomorphism, so by the previous discussion the Lie algebra $\mathfrak{h}$ of H is naturally a Lie subalgebra of $\mathfrak{g}$.

This implies in particular that the Lie algebra of any Lie subgroup of $\mathrm{GL}(n, \mathbb{R})$ is completely determined as soon as we know its tangent space at the identity: there is no need of computing the Lie bracket again since it will always be $[A, B] = AB - BA$.

For instance, we know from Propositions 3.9.1 and 3.9.2 that

$$\mathfrak{sl}(n, \mathbb{R}) = \{A \in M(n, \mathbb{R}) \mid \mathrm{tr} A = 0\},$$
$$\mathfrak{o}(n, \mathbb{R}) = \mathfrak{so}(n, \mathbb{R}) = \{A \in M(n, \mathbb{R}) \mid {}^{t}A = -A\},$$

where both $\mathfrak{sl}(n, \mathbb{R})$ and $\mathfrak{o}(n, \mathbb{R})$ are subalgebras of $\mathfrak{gl}(n, \mathbb{R})$. One verifies easily they are indeed both closed under the Lie bracket multiplication.

10.2.7. From Lie subalgebras to Lie subgroups. Here is a striking application of the Frobenius Theorem.

Theorem 10.2.6. *Let* G *be a Lie group. For every subalgebra* $\mathfrak{h} \subset \mathfrak{g}$ *there is a unique connected Lie subgroup* $H < G$ *whose Lie algebra is* $\mathfrak{h}$.

Proof. A subalgebra $\mathfrak{h} \subset \mathfrak{g}$ is in particular a subspace of $\mathfrak{g} = T_eG$, and by left-multiplication it extends to a distribution D in G, defined as

$$D_g = (dL_g)_e(\mathfrak{h}) \subset T_gG \tag{31}$$

for every $g \in G$. Since $\mathfrak{h}$ is a subalgebra, the distribution D is involutive. To prove this, pick k left-invariant vector fields $X_1, \ldots, X_k$ generating $\mathfrak{h}$. By construction they are tangent to D. Since $\mathfrak{h}$ is a subalgebra, their brackets $[X_i, X_j]$ are still in $\mathfrak{h}$ and hence are also tangent to D. Now Exercise 5.5.10 shows that D is involutive.

By the Frobenius Theorem 5.5.9, there is a foliation $\mathscr{F}$ of G tangent to D. Let H be the leaf of $\mathscr{F}$ containing the identity e. It is an injectively immersed manifold in G, with tangent space $T_eH = \mathfrak{h}$. For every $g \in G$, the diffeomorphism L_g preserves D and hence permutes the leaves of $\mathscr{F}$. If $h \in H$, then $L_{h^{-1}}$ sends $h \in H$ to $e \in H$ and hence preserves the leaf H. This implies that H is a subgroup, and hence a Lie subgroup.

If $H < G$ is connected and its Lie algebra is $\mathfrak{h}$, then H in fact must be obtained from $\mathfrak{h}$ in the way just described. This shows uniqueness. □

We have discovered a beautiful natural 1-1 correspondence:

$$\big\{\text{connected Lie subgroups of } G\big\} \longleftrightarrow \big\{\text{Lie subalgebras of } \mathfrak{g}\big\}.$$

We note that the subgroup $H < G$ corresponding to $\mathfrak{h}$ is not guaranteed to be embedded, and there is no easy way to understand from $\mathfrak{h}$ alone whether $H < G$ is embedded or not. In fact, the pleasure of obtaining such a powerful and elegant theorem is the main reason for allowing non-embedded Lie subgroups in our definition.

10.2.8. Foliations. The proof of Theorem 10.2.6 also displays a nice geometric phenomenon that is worth emphasising. Let G be a Lie group. Given a subalgebra $\mathfrak{h} \subset \mathfrak{g}$, by left-multiplication we get an integrable distribution D as in (31), and hence a foliation $\mathscr{F}$ of G. We write $\mathscr{F}_{\mathfrak{h}}$ to stress its dependence on $\mathfrak{h}$. The construction implies easily the following fact.

Proposition 10.2.7. *Let $H < G$ be a Lie subgroup with Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$. The left cosets of H are unions of leaves of the foliation $\mathscr{F}_{\mathfrak{h}}$.*

Corollary 10.2.8. *Every embedded Lie subgroup $H < G$ is closed.*

Proof. Every embedded union of leaves in a foliation is closed. □

10.2.9. Local homomorphisms. We now pass from subgroups to homomorphisms; that is, we ask ourselves if every Lie algebra homomorphism should be induced by some Lie group homomorphism. This is true only locally.

A *local homomorphism* between two Lie groups G and H is a smooth map $f: U \to H$ defined on some neighbourhood U of $e \in G$, such that

$$f(ab) = f(a)f(b) \quad \forall a, b, ab \in U.$$

Here is a partial converse to Proposition 10.2.5.

Theorem 10.2.9. *Let G, H be Lie groups and $F: \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism. There is a local homomorphism $f: U \to H$ with $df_e = F$.*

Proof. The graph of the map F is

$$\mathfrak{f} = \big\{ \big(X, F(X)\big) \bigm| X \in \mathfrak{g} \big\} \subset \mathfrak{g} \times \mathfrak{h}$$

and it is a Lie subalgebra of $\mathfrak{g} \times \mathfrak{h}$, the Lie algebra of $G \times H$. By Theorem 10.2.6 there is a Lie subgroup $K \subset G \times H$ with Lie algebra $\mathfrak{f}$.

The projections $\pi_1: K \to G$ and $\pi_2: K \to H$ are Lie group homomorphisms. The differential of π_1 at $(e, e) \in K$ is invertible (it is $(X, F(X)) \mapsto X$) so π_1 is a local diffeomorphism at (e, e). Thus we can define on some open neighbourhood U of $e \in G$ the local homomorphism

$$f: U \to H, \quad f = \pi_2 \circ \pi_1^{-1}.$$

Its differential is clearly F. □

With similar techniques we obtain also a uniqueness result.

Proposition 10.2.10. *Let G, H be Lie groups. If G is connected, two homomorphisms $f, f': G \to H$ with the same differentials $f_* = f'_*$ must coincide.*

Proof. Following the previous proof, the graphs of f and f' are two connected Lie subgroups $K, K' \subset G \times H$ with the same Lie subalgebra $\mathfrak{f}$, and hence must coincide, that is $f = f'$. □

If G is simply connected, existence is also achieved.

Proposition 10.2.11. *Let G, H be Lie groups. If G is simply connected, every Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{h}$ is the differential of a unique Lie group homomorphism $G \to H$.*

Proof. In the proof of Theorem 10.2.9, the map $\pi_1: K \to G$ is a smooth covering by Proposition 10.1.13. Being G simply connected, the map π_1 is an isomorphism, so we can define $f = \pi_2 \circ \pi_1^{-1}: G \to H$ and conclude. □

10.2.10. Simply connected Lie groups. The results just stated have the following important consequence.

Corollary 10.2.12. *Two simply connected Lie groups are isomorphic $\Longleftrightarrow$ their Lie algebras are.*

Proof. Every isomorphism $\mathfrak{g} \to \mathfrak{h}$ gives rise to two homomorphisms $G \to H$ and $H \to G$, whose composition is the identity because its differential is. □

Remember that Corollary 10.1.17 reduces the problem of classifying connected Lie groups to the simply connected ones. Now Corollary 10.2.12 in turn translates this task into the purely algebraic problem of classifying all the Lie algebras (to be precise, only the Lie algebras that arise from some Lie groups are important for us).

Two Lie groups G, H are *locally isomorphic* if there are neighbourhoods U and V of $e \in G$ and $e \in H$ and a diffeomorphism $f: U \to V$ such that $f(ab) = f(a)f(b)$ whenever $a, b, ab \in U$.

Corollary 10.2.13. *Let G, H be two connected Lie groups. The following are equivalent:*

- *G and H are locally isomorphic;*
- *G and H have isomorphic universal covers;*
- *$\mathfrak{g}$ and $\mathfrak{h}$ are isomorphic Lie algebras.*

10.2.11. Abelian Lie groups. We now apply the techniques just introduced to classify all the abelian Lie groups. We will need the following.

Proposition 10.2.14. *The differentials of the multiplication $m: G \times G \to G$ and the inverse $i: G \to G$ are*

$$\mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}, \quad (X, Y) \longmapsto X + Y, \qquad \mathfrak{g} \longrightarrow \mathfrak{g}, \quad X \longmapsto -X.$$

Proof. For the first, by linearity it suffices to prove that $(X, 0) \mapsto X$, which is obvious since $ge = g$. The second follows from $m(g, i(g)) = g$. □

Here is a smart application.

Proposition 10.2.15. *If a Lie group G is abelian, then $\mathfrak{g}$ also is.*

Proof. Since G is abelian, the map $G \to G$, $g \mapsto g^{-1}$ is an endomorphism. Therefore its derivative $\mathfrak{g} \to \mathfrak{g}$, $X \mapsto -X$ is a Lie algebra endomorphism. Hence for every $X, Y \in \mathfrak{g}$ we get

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

which implies $[X, Y] = 0$. □

Recall that in every dimension n there is a unique abelian Lie algebra $\mathbb{R}^n$. We will also need the following.

Exercise 10.2.16. Let $\Gamma < \mathbb{R}^n$ be a discrete subgroup. There is a basis $v_1, \ldots, v_n$ of $\mathbb{R}^n$ where $v_1, \ldots, v_k$ generate Γ. In particular $\Gamma \cong \mathbb{Z}^k$.

Here is a complete classification of abelian Lie groups.

Theorem 10.2.17. *Every abelian Lie group is isomorphic to*

$$\underbrace{S^1 \times \cdots \times S^1}_{k} \times \mathbb{R}^{n-k}$$

for some $0 \leq k \leq n$.

Proof. By Proposition 10.2.15 the Lie algebra of an abelian group G is $\mathbb{R}^n$, which is also the Lie algebra of the Lie group $\mathbb{R}^n$. By Corollary 10.2.12 then $\tilde{G} = \mathbb{R}^n$, and by Corollary 10.1.17 we have $G = \mathbb{R}^n/_\Gamma$ for some discrete $\Gamma < \mathbb{R}^n$. Now Exercise 10.2.16 applies. □

10.3. Examples

Having proved a number of general theorems, it is due time to exhibit and study more examples of Lie groups.

10.3.1. Complex matrices. We introduce some Lie groups using complex matrices. To this purpose we identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ in the usual way, by sending $(z_1, \ldots, z_n)$ to $(\Re z_1, \Im z_1, \ldots, \Re z_n, \Im z_n)$. We consider every complex endomorphism of $\mathbb{C}^n$ as a particular real endomorphism of $\mathbb{R}^{2n}$ and thus see $M(n, \mathbb{C})$ as a linear subspace of $M(2n, \mathbb{R})$, and more than that as a subalgebra with respect to matrix multiplication.

Our first example is the *complex general linear group*

$$\mathrm{GL}(n, \mathbb{C}) = \big\{A \in M(n, \mathbb{C}) \bigm| \det A \neq 0\big\}.$$

This is an open subset of $M(n, \mathbb{C})$ and hence a Lie group of dimension $2n^2$. It is a Lie subgroup of $\mathrm{GL}(2n, \mathbb{R})$, with Lie algebra

$$\mathfrak{gl}(n, \mathbb{C}) = M(n, \mathbb{C})$$

where we see $M(n, \mathbb{C})$ as a Lie subalgebra of $M(2n, \mathbb{R})$, with the same Lie bracket $[A, B] = AB - BA$. Note the Lie subgroup inclusions:

$$\mathrm{GL}(n, \mathbb{R}) \subset \mathrm{GL}(n, \mathbb{C}) \subset \mathrm{GL}(2n, \mathbb{R}).$$

These Lie groups have dimensions n^2, $2n^2$, and $4n^2$ respectively. When $n = 1$ these reduce to

$$\mathbb{R}^* \subset \mathbb{C}^* \subset \mathrm{GL}(2, \mathbb{R}).$$

In the second inclusion, every element $\rho e^{i\theta} \in \mathbb{C}^*$ is interpreted as the product of a ρ-dilation with a θ-rotation:

$$\rho \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

The determinant is a Lie group homomorphism $\det\colon \mathrm{GL}(n, \mathbb{C}) \to \mathbb{C}^*$. As in the real case, the *complex special linear group* is its kernel

$$\mathrm{SL}(n, \mathbb{C}) = \big\{A \in \mathrm{GL}(n, \mathbb{C}) \bigm| \det A = 1\big\}.$$

This is a Lie subgroup, with Lie algebra

$$\mathfrak{sl}(n, \mathbb{C}) = \big\{A \in M(n, \mathbb{C}) \bigm| \mathrm{tr} A = 0\big\}.$$

The Lie group $\mathrm{GL}(n, \mathbb{C})$ contains the *unitary group* $\mathrm{U}(n)$, that consists of all unitary matrices:

$$\mathrm{U}(n) = \big\{A \in \mathrm{GL}(n, \mathbb{C}) \bigm| {}^{\mathrm{t}}\bar{A}A = I\big\}.$$

Exercise 10.3.1. The unitary group is a Lie subgroup of $\mathrm{GL}(n, \mathbb{C})$ of dimension n^2, whose Lie algebra consists of all the $n \times n$ skew-Hermitian matrices:

$$\mathfrak{u}(n) = \big\{A \in M(n, \mathbb{C}) \bigm| {}^{\mathrm{t}}\bar{A} + A = 0\big\}.$$

Hint. Adapt the proof of Proposition 3.9.2 to the complex case. □

Finally, the *special unitary group* is

$$\mathrm{SU}(n) = \left\{ A \in \mathrm{GL}(n, \mathbb{C}) \;\middle|\; {}^{\mathrm{t}}\bar{A}A = I, \det A = 1 \right\}.$$

Exercise 10.3.2. This is a Lie subgroup of dimension $n^2 - 1$ with Lie algebra

$$\mathfrak{su}(n) = \left\{ A \in M(n, \mathbb{C}) \;\middle|\; {}^{\mathrm{t}}\bar{A} + A = 0, \mathrm{tr} A = 0 \right\}.$$

We note that

$$\mathrm{SU}(n) = U(n) \cap \mathrm{SL}(n, \mathbb{C}).$$

Exercise 10.3.3. The Lie groups $\mathrm{GL}(n, \mathbb{C})$, $\mathrm{SL}(n, \mathbb{C})$, $\mathrm{U}(n)$, and $\mathrm{SU}(n)$ are all connected.

10.3.2. More matrix Lie groups. We further introduce some Lie subgroups of $\mathrm{GL}(n, \mathbb{R})$ that are widely used in geometry.

Example 10.3.4 (Indefinite orthogonal groups). Let $\mathbb{R}^{p,q}$ be the vector space $\mathbb{R}^{p+q}$ equipped with the standard scalar product with (p, q) signature:

$$\langle x, y \rangle = x_1 y_1 + \ldots + x_p y_p - x_{p+1} y_{p+1} - \ldots - x_n y_n.$$

Here $n = p + q$. Let $\mathrm{O}(p, q) \subset \mathrm{GL}(n, \mathbb{R})$ be the subgroup of all the isometries of $\mathbb{R}^{p,q}$, that is the isomorphisms that preserve the scalar product. That is,

$$\mathrm{O}(p, q) = \left\{ A \in \mathrm{GL}(n, \mathbb{R}) \;\middle|\; {}^{\mathrm{t}}A I_{p,q} A = I_{p,q} \right\}$$

where

$$I_{p,q} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}.$$

Similarly to the proof of Proposition 3.9.2, we check that $\mathrm{O}(p, q)$ is indeed a submanifold of $\mathrm{GL}(n, \mathbb{R})$ of dimension $\frac{n(n-1)}{2}$ with Lie algebra

$$\mathfrak{o}(p, q) = \left\{ A \in M(n) \;\middle|\; {}^{\mathrm{t}}A I_{p,q} + I_{p,q} A = 0 \right\}.$$

Every matrix in $\mathrm{O}(p, q)$ has determinant ± 1, and $\mathrm{SO}(p, q)$ is the index-two subgroup consisting of those with determinant 1. We have $\mathfrak{so}(p, q) = \mathfrak{o}(p, q)$.

The Lie groups $\mathrm{O}(p, q)$ and $\mathrm{O}(q, p)$ are isomorphic. If $p, q > 0$, the Lie group $\mathrm{O}(p, q)$ is not compact (exercise).

Example 10.3.5 (Indefinite unitary groups). Proceeding exactly as above with the standard hermitian product of signature (p, q) on $\mathbb{C}^{p+q}$, we construct the Lie groups

$$\mathrm{U}(p, q) = \left\{ A \in \mathrm{GL}(n, \mathbb{C}) \;\middle|\; {}^{\mathrm{t}}\bar{A} I_{p,q} A = I_{p,q} \right\}$$

with Lie algebra

$$\mathfrak{u}(p, q) = \left\{ A \in M(n, \mathbb{C}) \;\middle|\; {}^{\mathrm{t}}\bar{A} I_{p,q} + I_{p,q} A = 0 \right\}.$$

The matrices of $\mathrm{U}(p, q)$ with unit determinant form a Lie subgroup $\mathrm{SU}(p, q)$, with Lie algebra

$$\mathfrak{su}(p, q) = \left\{ A \in M(n, \mathbb{C}) \;\middle|\; {}^{\mathrm{t}}\bar{A} I_{p,q} + I_{p,q} A = 0, \mathrm{tr} A = 0 \right\}.$$

We have $\dim \mathrm{U}(p,q) = n^2$ and $\dim \mathrm{SU}(p,q) = n^2 - 1$, with $n = p+q$.

Example 10.3.6 (Symplectic groups). Let $\mathbb{R}^{2n}$ or $\mathbb{C}^{2n}$ be equipped with the standard symplectic (that is, antisymmetric and non-degenerate) form

$$\omega(x,y) = {}^{t}xJy$$

where

$$J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}.$$

Let $\mathrm{Sp}(2n,\mathbb{R})$ or $\mathrm{Sp}(2n,\mathbb{C})$ be group of all linear isomorphism preserving the symplectic form. That is,

$$\mathrm{Sp}(2n,\mathbb{R}) = \left\{ A \in \mathrm{GL}(n,\mathbb{R}) \;\middle|\; {}^{t}AJA = J \right\}.$$

The Lie algebra is

$$\mathfrak{sp}(2n,\mathbb{R}) = \left\{ A \in M(n) \;\middle|\; {}^{t}AJ + JA = 0 \right\}.$$

The complex case is analogous. The dimensions of $\mathrm{Sp}(2n,\mathbb{R})$ and $\mathrm{Sp}(2n,\mathbb{C})$ are $n(2n+1)$ and $2n(2n+1)$ respectively.

Example 10.3.7 (Affine extensions). For every Lie subgroup $G < \mathrm{GL}(n,\mathbb{R})$ we may consider its affine extension

$$G \rtimes \mathbb{R}^n = \left\{ x \mapsto Ax + b \;\middle|\; A \in G, b \in \mathbb{R}^n \right\} \subset \mathrm{Aff}(\mathbb{R}^n).$$

This is a Lie subgroup of $\mathrm{Aff}(\mathbb{R}^n)$, which is in turn a Lie subgroup of $\mathrm{GL}(n+1,\mathbb{R})$, recall Example 10.1.7. Its Lie algebra is the subalgebra of $\mathfrak{gl}(n+1,\mathbb{R})$ consisting of all matrices

$$\begin{pmatrix} A & b \\ 0 & 0 \end{pmatrix}$$

where $A \in \mathfrak{g}$ and $b \in \mathbb{R}^n$.

10.3.3. Low dimensions. We now try to embark on a more systematic classification of connected Lie groups with increasing dimension. We use the powerful *Lie groups – Lie algebra correspondence* proved in the previous pages, which can be reassumed as follows:

(i) Every connected Lie group is the quotient $G/_\Gamma$ of a simply connected Lie group G by a discrete central subgroup $\Gamma < G$.
(ii) Every simply connected Lie group G is totally determined by its Lie algebra $\mathfrak{g}$.

An optimistic strategy to produce all connected Lie groups would be the following:

(1) Classify all Lie algebras $\mathfrak{g}$.
(2) Try to build a simply connected Lie group G for each Lie algebra $\mathfrak{g}$.
(3) Quotient G by its central discrete subgroups.

Dimension one. The only one-dimensional Lie algebra is the abelian $\mathbb{R}$, so the 1-dimensional connected Lie groups are $\mathbb{R}$ and S^1.

Dimension two. In dimension two, we find two Lie algebras:

- The abelian $\mathbb{R}^2$.
- The Lie algebra $\mathfrak{aff}(\mathbb{R})$ of $\mathrm{Aff}(\mathbb{R})$.

Proposition 10.3.8. *These are the only two 2-dimensional Lie algebras up to isomorphism.*

Proof. Let $\mathfrak{a}$ be a 2-dimensional Lie algebra. Pick a basis $X, Y \in \mathfrak{a}$ and note that the whole structure is determined by the element $[X, Y]$. If $[X, Y] = 0$ then $\mathfrak{a}$ is abelian. Otherwise, after changing the basis we easily reduce to the case $[X, Y] = Y$ and we get $\mathfrak{aff}(\mathbb{R})$. Indeed, We see $\mathrm{Aff}(\mathbb{R}) \subset \mathrm{GL}(2, \mathbb{R})$ as the set of all matrices

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$$

with $a, b \in \mathbb{R}$. Its Lie algebra is generated by the matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

We have $[A, B] = B$, so $\mathfrak{aff}(\mathbb{R}) \cong \mathfrak{a}$. □

The simply connected Lie group with algebra $\mathfrak{aff}(\mathbb{R})$ is $\mathrm{Aff}(\mathbb{R})^0$. We can easily classify the two-dimensional connected Lie groups up to isomorphism:

Proposition 10.3.9. *The two-dimensional connected Lie groups are*

$$\mathbb{R}^2, \quad S^1 \times \mathbb{R}, \quad S^1 \times S^1, \quad \mathrm{Aff}(\mathbb{R})^0.$$

Proof. Since the centre of $\mathrm{Aff}(\mathbb{R})^0$ is trivial, there is no other connected Lie group with Lie algebra $\mathfrak{aff}(\mathbb{R})$ except $\mathrm{Aff}(\mathbb{R})^0$ itself. □

Dimension three. In dimension three we find many more Lie algebras. Here are some:

(1) The abelian $\mathbb{R}^3$.
(2) The *Heisenberg algebra*, which is the subalgebra of $\mathfrak{sl}(3, \mathbb{R})$ formed by the matrices
$$\begin{pmatrix} 0 & a & c \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix}$$
with $a, b, c \in \mathbb{R}$. This is the Lie algebra of the Heisenberg group.
(3) The direct product $\mathbb{R} \oplus \mathfrak{aff}(\mathbb{R})$.
(4) The Lie algebra of the affine isometries of $\mathbb{R}^2$.
(5) The Lie algebra of the affine isometries of $\mathbb{R}^{1,1}$.
(6) The Lie algebra $\mathfrak{sl}(2, \mathbb{R})$.
(7) The Lie algebra $\mathfrak{so}(3)$.

Each of these seven algebras is the Lie algebra of some Lie group. Unfortunately, this is not the end of the story: the are uncountably many Lie algebras in dimension three, as the following exercise shows. to be checked

Exercise 10.3.10. Consider $\mathbb{R}^3$ with basis X, Y, T and Lie bracket defined by

$$[T, X] = X, \qquad [T, Y] = tY, \qquad [X, Y] = 0.$$

This defines a Lie algebra $\mathfrak{g}_t$ for all $t \in \mathbb{R}$. If $tu \neq 1$ then $\mathfrak{g}_t$ and $\mathfrak{g}_u$ are not isomorphic. Every $\mathfrak{g}_t$ is a subalgebra of $\mathfrak{gl}(n, \mathbb{R})$ for some n and is hence the Lie algebra of some Lie subgroup of $\mathrm{GL}(n, \mathbb{R})$.

It is actually possible to classify all the three-dimensional Lie algebras: this was done by Bianchi in 1898 who subdivided them into 11 classes, two of which are continuous families. However, these examples already suggest that it is practically impossible to classify all connected Lie groups without adding further assumptions like, for instance, that the Lie group should be compact, or abelian, or some weaker assumption.

We now write some isomorphisms between some notable three-dimensional Lie algebras. Let $\times$ be the cross product of vectors in $\mathbb{R}^3$.

Proposition 10.3.11. *The Lie algebras* $\mathfrak{so}(3)$ *and* $\mathfrak{su}(2)$ *are both isomorphic to the algebra* $(\mathbb{R}^3, \times)$.

Proof. A basis for $\mathfrak{so}(3)$ is given by the matrices

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We have

$$[A, B] = C, \qquad [B, C] = A, \qquad [C, A] = B.$$

Therefore $\mathfrak{so}(3) \cong (\mathbb{R}^3, \times)$. Analogously $\mathfrak{su}(2)$ is generated by the matrices

$$A = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad B = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \qquad C = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

whose Lie brackets are again as above. □

This implies that $\mathrm{SO}(3)$ and $\mathrm{SU}(2)$ have the same universal cover. In fact, we will write an explicit double cover $\mathrm{SU}(2) \to \mathrm{SO}(3)$ soon.

Proposition 10.3.12. *The Lie algebras* $\mathfrak{sl}(2, \mathbb{R})$ *and* $\mathfrak{so}(2, 1)$ *are isomorphic.*

Proof. A basis for $\mathfrak{sl}(2, \mathbb{R})$ is

$$A = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad B = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad C = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

We have

$$[A, B] = C, \qquad [B, C] = -A, \qquad [C, A] = -B.$$

The Lie algebra $\mathfrak{so}(2,1)$ consists of matrices of the form

$$\begin{pmatrix} M & b \\ {}^{t}b & 0 \end{pmatrix}$$

with ${}^{t}M + M = 0$. A basis is

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Their Lie brackets are as above. □

The *derived algebra* $[\mathfrak{g}, \mathfrak{g}]$ of a Lie algebra $\mathfrak{g}$ is the subalgebra generated by all the brackets $[X, Y]$ as $X, Y \in \mathfrak{g}$ varies. The derived algebra is trivial $\Longleftrightarrow$ $\mathfrak{g}$ is abelian.

Exercise 10.3.13. In the seven Lie algebras listed above, the dimension of $[\mathfrak{g}, \mathfrak{g}]$ is zero for (1), one for (2, 3), two for (4,5), and three for (6,7).

10.4. The exponential map

Similar to Riemannian manifolds, Lie groups G are equipped with an *exponential map* $\mathfrak{g} \to G$. For matrix groups, this is the usual matrix exponential, and this finally explains the reason for adopting this name...

10.4.1. Definition. Let G be a Lie group. Pick an arbitrary left-invariant vector field $X \in \mathfrak{g}$.

Proposition 10.4.1. *The vector field X is complete.*

Proof. Let $\gamma_g \colon I_g \to G$ be the maximal integral curve of X at g. Since X is left-invariant, we have $\gamma_g = L_g \circ \gamma_e$ and $I_g = I_e$ for all $g \in G$. By Lemma 5.2.5 the vector field is complete. □

Being complete, the vector field $X \in \mathfrak{g}$ induces a flow $\Phi_X \colon G \times \mathbb{R} \to G$.

Definition 10.4.2. The *exponential map* $\exp \colon \mathfrak{g} \longrightarrow G$ is

$$\exp(X) = \Phi_X(e, 1).$$

The map exp is smooth because $\Phi_X(e, 1)$ depends smoothly on the initial values X of the system.

10.4.2. One-parameter subgroups. In the Riemannian case, the restrictions of the exponential map to the vector lines are geodesics; here, these are "one-parameter subgroups."

Let G be a Lie group. For every $X \in \mathfrak{g}$ we consider the curve $\gamma_X \colon \mathbb{R} \to G$,

$$\gamma_X(t) = \exp(tX).$$

As in the Riemannian case, by construction we have $\gamma_{\lambda X}(t) = \gamma_X(\lambda t)$.

Proposition 10.4.3. *The map $\gamma_X\colon \mathbb{R} \to G$ is the integral curve of the left-invariant field X with $\gamma_X(0) = e$. It is a Lie group homomorphism.*

Proof. We have

$$\gamma_X(t) = \exp(tX) = \Phi_{tX}(e, 1) = \Phi_X(e, t)$$

so γ_X is the integral curve for X with $\gamma_X(0) = e$. Since X is left-invariant,

$$\gamma_X(s)\gamma_X(t) = L_{\gamma_X(s)}(\gamma_X(t)) = \gamma_X(s + t).$$

Therefore γ_X is a Lie groups homomorphism. $\square$

A Lie group homomorphism $\mathbb{R} \to G$ is called a *one-parameter subgroup* of G. It turns out that every one-parameter subgroup arises in this way.

Proposition 10.4.4. *Every one-parameter subgroup of G is a γ_X for some element $X \in \mathfrak{g}$.*

Proof. Given $f\colon \mathbb{R} \to G$, we set $X = f_*(1)$. Since $f_* = (\gamma_X)_*$, we have $f = \gamma_X$ by Proposition 10.2.10. $\square$

The Lie algebra $\mathfrak{g}$ thus parametrises all the one-parameter subgroups in G.

10.4.3. Properties. We now list some properties of the exponential map.

Proposition 10.4.5. *Let G be a Lie group. The following holds.*

- *The differential $d\exp_0\colon \mathfrak{g} \to \mathfrak{g}$ is the identity. Hence the exponential map is a local diffeomorphism at 0.*
- *If $f\colon G \to H$ is a Lie group homomorphism, the following diagram commutes:*

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "f_*"] \arrow[d, "\exp"'] & \mathfrak{h} \arrow[d, "\exp"] \\ G \arrow[r, "f"'] & H \end{tikzcd}$$

Proof. Everything follows readily if we interpret $\mathfrak{g}$ and $\mathfrak{h}$ as sets of one-parameter subgroups. $\square$

In particular, if $H \subset G$ is a subgroup, the exponential map $\mathfrak{h} \to H$ is just the restriction of the exponential map $\mathfrak{g} \to G$.

10.4.4. Matrix exponential. We finally motivate the use of the term "exponential map". Recall that the exponential of a square matrix A is

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

If A and B commute, then $e^{A+B} = e^A e^B = e^B e^A$. In particular e^A is invertible with inverse e^{-A}.

Proposition 10.4.6. *The exponential map* $\exp\colon \mathfrak{gl}(n,\mathbb{R}) \to \mathrm{GL}(n,\mathbb{R})$ *is*

$$\exp(A) = e^A.$$

Proof. For every $A \in \mathfrak{gl}(n,\mathbb{R})$ consider the curve $\alpha\colon \mathbb{R} \to \mathrm{GL}(n,\mathbb{R})$, $\alpha(t) = e^{tA}$. We can differentiate it and find $\alpha'(t) = Ae^{tA}$. So α is a smooth curve and in fact a one-parameter subgroup of $\mathrm{GL}(n,\mathbb{R})$. By Proposition 10.4.4 we have $\alpha = \gamma_{\alpha'(0)} = \gamma_A$. In particular $e^A = \exp(A)$. □

By restriction, the same exponential map works for all the Lie subgroups of $\mathrm{GL}(n,\mathbb{R})$ like $\mathrm{SL}(n,\mathbb{R})$ or $\mathrm{O}(n)$. We discover in particular that the exponential of an antisymmetric matrix is orthogonal, and that of a traceless matrix has determinant one; these facts follow also from the following exercise.

Exercise 10.4.7. We have $e^{{}^tA} = {}^t(e^A)$ and $\det e^A = e^{\mathrm{tr} A}$.

From these examples we discover that, as in the Riemannian case, the exponential map needs not to be surjective, not even if G is connected.

Proposition 10.4.8. *The exponential map* $\mathfrak{sl}(2,\mathbb{R}) \to \mathrm{SL}(2,\mathbb{R})$ *is not surjective.*

Proof. If $g = \exp(A)$, it has a square root $\sqrt{g} = \exp\left(\frac{A}{2}\right)$. However

$$B = \begin{pmatrix} -4 & 0 \\ 0 & -\frac{1}{4} \end{pmatrix}$$

has no square root (exercise: use Jordan normal form). □

10.4.5. Applications. In the rest of this section we will use the exponential map to prove these remarkable non-trivial facts. Let G be a Lie group. Then:

(1) Every closed subgroup $H < G$ is a Lie subgroup.
(2) If $H \triangleleft G$ is closed and normal, the quotient $G/_H$ is a Lie group.
(3) The kernel and the image of any homomorphism $G \to H$ of Lie groups are Lie subgroups of G and H.

10.4.6. The closed subgroup theorem. As promised, we start by proving the following powerful theorem, which transforms a purely topological condition (closeness) into a much stronger differential one (being a smooth embedded submanifold).

Theorem 10.4.9. *Let G be a Lie group. Every closed subgroup $H \subset G$ is an embedded Lie subgroup.*

To prove this theorem we need a lemma. Recall that $\exp(X+Y) \neq \exp(X)\exp(Y)$ in general.

Lemma 10.4.10. *Let G be a Lie group. For every $X, Y \in \mathfrak{g}$ we have*

$$\exp(X+Y) = \lim_{n\to\infty}\left(\exp\frac{X}{n}\exp\frac{Y}{n}\right)^n.$$

Proof. When t is sufficiently small we have

$$\exp(tX)\exp(tY) = \exp\big(\psi(t)\big)$$

where ψ is the smooth map

$$\psi\colon \mathbb{R} \xrightarrow{\gamma_X\times\gamma_Y} G\times G \xrightarrow{m} G \xrightarrow{\exp^{-1}} \mathfrak{g}.$$

Here m is the multiplication and $\exp^{-1}$ is defined only in a neighbourhood of e. The map ψ is defined only near 0 and $\psi'(0) = X+Y$. Therefore we have

$$\psi(t) = t(X+Y) + t^2 Z(t)$$

for some smooth map Z defined only near 0. This implies

$$\exp(tX)\exp(tY) = \exp\big(\psi(t)\big) = \exp\big(t(X+Y)+t^2Z(t)\big).$$

If n is sufficiently big, we deduce that

$$\begin{aligned}\left(\exp\frac{X}{n}\exp\frac{Y}{n}\right)^n &= \left(\exp\left(\frac{1}{n}(X+Y)+\frac{1}{n^2}Z\left(\frac{1}{n}\right)\right)\right)^n \\ &= \exp\left(X+Y+\frac{1}{n}Z\left(\frac{1}{n}\right)\right).\end{aligned}$$

This completes the proof. $\square$

We can now turn back to the proof of Theorem 10.4.9

Proof. We must prove that $H\subset G$ is an embedded submanifold. Let $\mathfrak{h}\subset\mathfrak{g}$ be the subset defined as

$$\mathfrak{h} = \big\{X\in\mathfrak{g} \;\big|\; \exp(tX)\in H\ \forall t\in\mathbb{R}\big\}.$$

We first prove that $\mathfrak{h}$ is a subspace of $\mathfrak{g}$. To do so, we pick $X,Y\in\mathfrak{h}$, and prove that $X+Y\in\mathfrak{h}$. We know that $\exp\frac{tX}{n}, \exp\frac{tY}{n}\in H$, hence $\big(\exp\frac{tX}{n}\exp\frac{tY}{n}\big)^n\in H$. Since H is closed, by the previous lemma we get $\exp(t(X+Y))\in H$ for every $t\in\mathbb{R}$ and therefore $X+Y\in\mathfrak{h}$.

We now construct neighbourhoods U and W of $0\in\mathfrak{g}$ and $e\in G$ such that $\exp|_U\colon U\to W$ is a diffeomorphism and

$$\exp(\mathfrak{h}\cap U) = H\cap W. \tag{32}$$

This shows that H is an embedded submanifold near e, and hence everywhere by left multiplication.

Let $\mathfrak{k}\subset\mathfrak{g}$ be a complementary subspace for $\mathfrak{h}$. We leave as an exercise to prove that there is an open neighbourhood $U_\mathfrak{k}$ of $0\in\mathfrak{k}$ such that

$$H\cap\exp\big(U_\mathfrak{k}\setminus\{0\}\big) = \varnothing. \tag{33}$$

Instead of the exponential map, it is now convenient to consider the map

$$f\colon \mathfrak{h}\times\mathfrak{k}\longrightarrow\mathfrak{g}, \qquad f(X,Y) = \exp(X)\exp(Y).$$

We still have $df_0 = \mathrm{id}$, so there are neighbourhoods $U_\mathfrak{h}, U_\mathfrak{k}$ of $0\in\mathfrak{h},\mathfrak{k}$ such that

$$f\colon U_\mathfrak{h}\times U_\mathfrak{k}\longrightarrow G$$

is a diffeomorphism onto its image. We suppose that $U_{\mathfrak{f}}$ also satisfies (33). We now set $U = U_{\mathfrak{h}} \times U_{\mathfrak{f}}$ and prove that

$$f(\mathfrak{h} \cap U) = H \cap f(U). \tag{34}$$

We have $\mathfrak{h} \cap U = U_{\mathfrak{h}}$ and $\exp(U_{\mathfrak{h}}) \subset H$, therefore $f(\mathfrak{h} \cap U) \subset H \cap f(U)$. On the other hand, if $h \in H \cap f(U)$ then $h = \exp(X)\exp(Y)$ with $X \in U_{\mathfrak{h}}$ and $Y \in U_{\mathfrak{f}}$. Now $h, \exp(X) \in H$ implies that $\exp(Y) \in H$ and hence by (33) we get $Y = 0$. Therefore $h \in \exp(U_{\mathfrak{h}})$.

We have proved (34), which in turn implies (32) by taking $W = \exp(U)$. This concludes the proof. $\square$

By combining the theorem with Corollary 10.2.8 we get

Corollary 10.4.11. *Let G be a Lie group. A subgroup $H < G$ is an embedded Lie subgroup $\Longleftrightarrow$ it is closed.*

10.4.7. Kernel. Here is an immediate application of the closed subgroup theorem.

Proposition 10.4.12. *Let $f : G \to H$ be a homomorphism of Lie groups. The kernel* $\ker f$ *is an embedded Lie subgroup of G.*

Proof. It is closed since f is continuous. Theorem 10.4.9 applies. $\square$

We want to prove an analogous theorem for the image. It is more convenient to first study the quotients of Lie groups.

10.4.8. Quotient of Lie groups. We now recycle the proof of the closed subgroup theorem to obtain the following.

Theorem 10.4.13. *Let G be a Lie group and $H < G$ a closed subgroup. The quotient $G/_H$ has a natural structure of smooth manifold such that $\pi : G \to G/_H$ is a fibre bundle.*

Proof. We know that G is foliated into the the cosets of H. Since H is closed, it is embedded, and hence its cosets also are. We now need to show that the cosets fit like fibers in a bundle.

As in the proof of Theorem 10.4.9 we pick a complementary subspace $\mathfrak{f}$ for $\mathfrak{h} \subset \mathfrak{g}$ and consider the map

$$f : \mathfrak{f} \times \mathfrak{h} \longrightarrow \mathfrak{g}, \qquad f(X, Y) = \exp(X)\exp(Y).$$

Let $U_{\mathfrak{f}}, U_{\mathfrak{h}}$ be neighbourhoods of $0 \in \mathfrak{f}, \mathfrak{h}$ such that

$$f : U_{\mathfrak{f}} \times U_{\mathfrak{h}} \longrightarrow G$$

is a diffeomorphism onto its image and $\operatorname{Im} f \cap H = f(0 \times U_{\mathfrak{h}}) = \exp(U_{\mathfrak{h}})$. We now pick a smaller neighbourhood $U'_{\mathfrak{f}} \subset U_{\mathfrak{f}}$ such that $u_1, u_2 \in U'_{\mathfrak{f}} \Rightarrow u_1 - u_2 \in U_{\mathfrak{f}}$. This implies that

$$\exp(U'_{\mathfrak{f}})\big(\exp(U'_{\mathfrak{f}})\big)^{-1} \subset \exp(U_{\mathfrak{f}}).$$

We consider the multiplication map

$$m\colon \exp(U'_{\mathfrak{f}}) \times H \longrightarrow G, \qquad m(g,h) = gh.$$

The map m is injective: if $g_1h_1 = g_2h_2$, then $g_2g_1^{-1} = h_2^{-1}h_1 \in H$, but since $g_2g_1^{-1} \in \exp(U_{\mathfrak{f}})$ we deduce that $g_2g_1^{-1} = e$, so $g_1 = g_2$ and $h_1 = h_2$.

The map m is an open embedding, after replacing $U'_{\mathfrak{f}}$ with a smaller open neighbourhood: we have $dm_{(e,e)} = \mathrm{id}$, so $dm_{(e,g)}$ is invertible for every $g \in U'_{\mathfrak{f}}$ up to taking a smaller $U'_{\mathfrak{f}}$. Hence $dm_{(g,h)}$ is invertible by right-multiplication for every $h \in H$.

Finally, we assign to $G/_H$ its quotient topology. The map

$$U'_{\mathfrak{f}} \longrightarrow G/_H, \qquad X \longmapsto \exp(X)H$$

is a homeomorphism onto its image. More generally, for every $g \in G$ the map $U'_{\mathfrak{f}} \to G/_H, X \mapsto g\exp(X)H$ is a homeomorphism onto its image and we use these maps as charts to give $G/_H$ a smooth structure.

The space $G/_H$ is now a smooth manifold and the map $G \to G/_H$ is a fibre bundle, with fibre diffeomorphic to H. $\square$

When H is a normal subgroup, things of course improve.

Corollary 10.4.14. *Let G be a Lie group and $H \triangleleft G$ a closed normal subgroup. The quotient $G/_H$ has a natural structure of Lie group, and $G \to G/_H$ is a Lie group homomorphism.*

10.4.9. Image. After taking care of kernels and quotients, we can finally consider images of Lie group homomorphisms. It is remarkable how many non-trivial theorems are necessary to prove this reasonable-looking fact.

Proposition 10.4.15. *Let $f\colon G \to H$ be a homomorphism of Lie groups. The image* $\operatorname{Im} f$ *is a Lie subgroup of H.*

Proof. Since $\ker f$ is closed and normal, the quotient $G/_{\ker f}$ is a Lie group. The induced map $G/_{\ker f} \to H$ is an injective immersion: hence its image is an injectively immersed manifold and a subgroup of H, that is a Lie subgroup. $\square$

The image is of course not guaranteed to be embedded.

Remark 10.4.16. The use of the term *one-parameter subgroup* in Section 10.4.2 for any Lie group homomorphism $\mathbb{R} \to G$ is now fully legitimated, since its image is indeed a Lie subgroup of G.

10.5. Lie group actions

Lie groups arise often as symmetry groups, and are more generally designed to act on spaces of various kind.

10.5.1. Definition. Let M be a smooth manifold and G a Lie group. A *Lie group action* of G on M is a homomorphism

$$G \longrightarrow \mathrm{Diffeo}(M)$$

that is also smooth in the following sense: the induced map

$$G \times M \longrightarrow M, \qquad (g, x) \longmapsto g(x)$$

should be smooth. A manifold M equipped with a Lie group action of G is sometimes called a *G-manifold*.

Here are some important examples:

- The group $\mathrm{GL}(n, \mathbb{R})$ or $\mathrm{Aff}(\mathbb{R}^n)$ acts on $\mathbb{R}^n$.
- The group $\mathrm{O}(n)$ acts on $S^{n-1} \subset \mathbb{R}^n$.
- The group $\mathrm{U}(n)$ acts on $S^{2n-1} \subset \mathbb{C}^n$.
- Every Lie group G acts on itself by left-multiplication $g(x) = gx$, by right-multiplication $g(x) = xg^{-1}$, and by conjugation $g(x) = gxg^{-1}$.

An action of $\mathbb{R}$ on M was called a *one-parameter group of diffeomorphisms* in Section 5.2.2.

10.5.2. Lie algebras. As usual, Lie algebras are there to help us, by encoding elegantly the infinitesimal side of the story. Let $\rho \colon G \to \mathrm{Diffeo}(M)$ be a Lie group action on M. This induces a homomorphism

$$\rho_* \colon \mathfrak{g} \longrightarrow \mathfrak{X}(M)$$

as follows. For every $p \in M$ we have a map

$$G \longrightarrow M, \qquad g \longmapsto g(p)$$

whose image is the orbit of p. The differential of this map at $e \in G$ is a linear map $\mathfrak{g} \to T_p(M)$. By collecting all these linear maps as $p \in M$ varies we get our homomorphism $\rho_* \colon \mathfrak{g} \to \Gamma(TM) = \mathfrak{X}(M)$.

Exercise 10.5.1. For every $X \in \mathfrak{g}$, the vector field $\rho_*(X)$ on M is complete with flow $\Phi_t \colon M \to M$. We have $\Phi_t(p) = \exp(tX)(p)$ for every $p \in M$.

In some sense $\mathrm{Diffeo}(M)$ is an infinite-dimensional Lie group and $\mathfrak{X}(M)$ is its Lie algebra. A morphism ρ of Lie groups should then induce one ρ_* of Lie algebras: we leave a rigorous proof of this fact as an exercise.

Pare sia in realtà un anti-homomorphism. Controllare

Exercise 10.5.2. The homomorphism ρ_* is a Lie algebra homomorphism.

Exercise 10.5.3. Let ρ be the action of $\mathrm{GL}(n, \mathbb{R})$ on $\mathbb{R}^n$. For every $A \in \mathfrak{gl}(n, \mathbb{R}) = M(n)$ the vector field $\rho_*(A)$ is $\mathbb{R}^n \to \mathbb{R}^n$, $X \mapsto AX$.

10.5.3. Stabilisers and orbits. When dealing with group actions, the first thing to do is always to investigate stabilisers and orbits. Let a Lie group G act on a smooth manifold M.

Proposition 10.5.4. *For every $p \in M$ the stabiliser $G_p < G$ is an embedded Lie subgroup, whose Lie algebra is*

$$\mathfrak{g}_p = \{X \in \mathfrak{g} \mid \rho_*(X)(p) = 0\}.$$

Moreover the induced map

$$G/_{G_p} \longrightarrow M, \qquad g \longmapsto g(p)$$

is an injective immersion, whose image is the orbit of p.

Proof. The stabiliser G_p is closed (exercise), so it is an embedded Lie subgroup. By Exercise 10.5.1 we have $\rho_*(X)(p) = 0$ for every $X \in \mathfrak{g}_p$. Conversely, if $\rho_*(X)(p) = 0$ then $p = \Phi_t(p) = \exp(tX)(p)$ for all t and hence $\exp(tX) \in G_p$ for all t, so $X \in \mathfrak{g}_p$.

The map $G/_{G_p} \to M$ is smooth because $G \to M$ is. Its differential at e is injective because if $X \in \mathfrak{g} \setminus \mathfrak{g}_p$ then $\rho_*(X)(p) \neq 0$. It is hence injective everywhere by left-multiplication. □

We have discovered that stabilisers are Lie subgroups, and orbits are immersed submanifolds. The manifold M is hence partitioned into immersed submanifolds (the orbits) that may have varying dimension.

Example 10.5.5. Let S^1 act on $\mathbb{R}^2$ by rotations. The orbits are the circles centered at the origin, and the origin itself.

Example 10.5.6. Every similarity or congruence class of matrices in the space $M(n)$ of all $n \times n$ real matrices is an immersed submanifold. This holds because each such class is an orbit of the action of $\mathrm{GL}(n, \mathbb{R})$ by conjugation or congruence.

For the same reason, every conjugacy class in a Lie group G is an immersed submanifold.

As usual, one wonders whether injective immersions can be promoted to embeddings. The usual counterexample shows that non-embedded orbits may occur: the action

$$\mathbb{R} \longrightarrow \mathrm{Diffeo}(S^1 \times S^1), \qquad s \mapsto \big((e^{it}, e^{iu}) \mapsto e^{i(t+s)}, e^{i(u+\lambda s)}\big)$$

has dense orbits if $\lambda \notin \mathbb{Q}$. Things improve if an additional hypothesis is fulfilled.

10.5.4. Proper actions. Let G be a Lie group acting on a manifold M

Definition 10.5.7. The action is *proper* if the following map is:

$$G \times M \longrightarrow M \times M, \qquad (g, p) \longmapsto \big(g(p), p\big).$$

If the action is proper, the stabilisers $G_p < G$ are compact for every $p \in M$. The orbits are also nicer.

Proposition 10.5.8. *If the action is proper, orbits are embedded and closed.*

Proof. The induced map $G/_{G_p} \to M$, $g \mapsto g(p)$ is proper. By Exercise 3.8.5 A proper injective immersion is an embedding and has closed image. $\square$

If G is compact, then every action of G is proper.

10.5.5. Homogeneous spaces. Recall that a G-manifold is a manifold M equipped with the action of a Lie group G.

Definition 10.5.9. If the action is transitive, the G-manifold M is called a *homogeneous space*.

Example 10.5.10. Let G be a Lie group and $H < G$ a closed subgroup. The left action of G on $G/_H$ is transitive: hence $G/_H$ is a homogeneous space.

It turns out that every homogenous space is precisely of this form.

Proposition 10.5.11. *If G acts transtitively on M, for every $p \in M$ the map*

$$G/_{G_p} \longrightarrow M$$

is a G-equivariant diffeomorphism.

Proof. This is a corollary of Proposition 10.5.4. $\square$

In other words, a homogeneous space is just a quotient $G/_H$ of a Lie group G by a closed subgroup H. A homogeneous space is one where "all points look the same", since G act transitively on them.

10.5.6. Examples. There are many interesting examples of homogeneous spaces, and we list some here.

Example 10.5.12. The group $\mathrm{SO}(n)$ acts transitively on S^{n-1}, with stabiliser isomorphic to $\mathrm{SO}(n-1)$. Therefore we get the homogeneous space

$$\mathrm{SO}(n)/_{\mathrm{SO}(n-1)} \cong S^{n-1}.$$

By Theorem 10.4.13 we have a fibre bundle $\mathrm{SO}(n) \to S^{n-1}$ with fibre $\mathrm{SO}(n-1)$.

Example 10.5.13. The group $\mathrm{Isom}^+(\mathbb{R}^n)$ of the orientation-preserving Euclidean affine isometries acts transitively on $\mathbb{R}^n$ with stabiliser isomorphic to $\mathrm{SO}(n)$. We get the homogeneous space

$$\mathrm{Isom}^+(\mathbb{R}^n)/_{\mathrm{SO}(n)} \cong \mathbb{R}^n$$

and a fibre bundle $\mathrm{Isom}^+(\mathbb{R}^n) \to \mathbb{R}^n$ with fibre $\mathrm{SO}(n)$.

Example 10.5.14. The group $\mathrm{O}(n)$ acts on the grassmannian $\mathrm{Gr}_k(\mathbb{R}^n)$ with stabiliser isomorphic to $\mathrm{O}(k) \times \mathrm{O}(n-k)$. We get the homogeneous space

$$\mathrm{O}(n)/_{\mathrm{O}(k)\times\mathrm{O}(n-k)} \cong \mathrm{Gr}_k(\mathbb{R}^n)$$

and a fibre bundle $\mathrm{O}(n) \to \mathrm{Gr}_k(\mathbb{R}^n)$ with fibre $\mathrm{O}(k) \times \mathrm{O}(n-k)$.

In fact, we could have used this construction to *define* a natural smooth manifold structure on the grassmannian. We do this with another interesting set. A *flag* on a n-dimensional vector space V is a nested sequence

$$0 \subset V_1 \subset \ldots \subset V_n = V$$

of i-dimensional subspaces $V_i \subset V$. In the following example we build a natural smooth manifold structure on the set of all flags in V.

Example 10.5.15. The group $\mathrm{GL}(n,\mathbb{R})$ acts on the space F of all the flags in $\mathbb{R}^n$. The stabiliser of the coordinate flag $V_i = \mathrm{Span}(e_1, \ldots, e_i)$ is the closed subgroup $H < \mathrm{GL}(n,\mathbb{R})$ of all upper triangular invertible matrices. Therefore the space of all flags in $\mathbb{R}^n$ is naturally identified with the homogeneous manifold $\mathrm{GL}(n,\mathbb{R})/_H$.

Exercise 10.5.16. The group $\mathrm{SL}(2,\mathbb{C})$ acts transitively on $\mathbb{P}^1(\mathbb{C})$ as follows:

$$\rho\begin{pmatrix} a & b \\ c & d \end{pmatrix} : [w,z] \longmapsto [aw+bz, cw+dz].$$

The stabiliser is a Lie group diffeomorphic to $\mathbb{C}^* \times \mathbb{C}$. We get a fibre bundle $\mathrm{SL}(2,\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$ with fibre $\mathbb{C}^* \times \mathbb{C}$.