Manifolds 2017

http://www.staff.science.uu.nl/∼crain101/

Lecturer: Marius Crainic

Contents

CHAPTER 1

Reminder on Topology; topological manifolds

Here is a very brief reminder on the basic notions from Topology. For those which are not so familiar with these basics, one may skip the later parts of this section (most notably the part on partitions of unity) and return to it later on, when necessary.

1. The objects of Topology

First of all, the main objects of Topology: **a topological space** is a set X endowed with a **topology**, i.e. a collection $\mathcal{T}$ of subsets of X (called **the opens of the topological space**, or simply **open in** X) such that $\emptyset$ and X are open in X, arbitrary unions of opens are open and finite intersections of opens are open. We usually omit $\mathcal{T}$ from the notations, and we simply say that X is a topological space; hence that means that X is a set and we can talk about the subsets of X that are open (in X).

A topology on X allows us to make sense of the central phenomena of Topology: "two points being close to each other". Formally, we can make sense of neighborhoods in a topological space X: given $x \in X$, a **neighborhood** (in X) of x is any subset $V \subset X$ that contains at least an open neighborhood of x, i.e. an open U with $x \in U$.

In turn, this allows us to talk about convergence: a sequence $(x_n)_{n\geq 1}$ of elements of X **converges** (in the topological space X) to $x \in X$ if for any neighborhood V of x there exists an integer n_V such that $x_n \in V$ for all $n \geq n_V$.

The notion of neighborhoods also allows to talk also about an important property one requires on topological spaces in order to exclude pathological examples- Hausdorffness: a topological space X called **Hausdorff** if for any $x, y \in X$ distinct, there are neighborhoods U of x and V of y such that $U \cap V = \emptyset$.

2. The morphisms/isomomorphisms of Topology

The relevant maps in Topology (the "morphism" in the category of topological spaces) are the continuous ones: a map $f : X \to Y$ between topological spaces is called **continuous** if for any U-open in Y, its pre-image $f^{-1}(U)$ is open in X. "Isomorphism" between topological spaces are known under the name of **homeomorphisms**: they are the bijections $f : X \to Y$ with the property that both f as well as f^{-1} are continuous.

Remark 1.1. Note that, while proving that two topological spaces are homeomorphic (i.e there exists a homeomorphism between them) is relatively easy in principle (one just has to produce ONE single homeomorphism between them- and for that it is often enough to follows ones intuition), proving that two spaces are not homeomorphic is much harder. One way to proceed is by understanding the specific "topological properties" of the spaces under discussion (such as Hausdorffness, compactness, etc); if one of them has such a topological property and the other one does not, then they cannot be homeomorphic. A more advanced approach consists of constructing topological invariants of algebraic nature (such as numbers, groups, etc)- and that is what Algebraic Topology is about.

3. Metric topologies; bases

One of the largest class of topological spaces are metric spaces (X, d): any metric $d : X \times X \to \mathbb{R}$ induces a topology $\mathcal{T}_d$ on X: a subset $U \subset X$ is open iff for any $x \in U$ there exists $r > 0$ such that U contains the d-ball of center x and radius r:

$$B_d(x, r) := \{y \in X : d(x, y) < r\}. \tag{3.1}$$

In general, a topological space X is called **metrizable** if there is a metric d on X such that the original topology on X coincides with $\mathcal{T}_d$ (note also that, if such a d exists, in general it is far from being unique). One of the most interesting questions about topological spaces is to decide whether they are metrizable or not; **metrizability theorems** aim at finding simple topological conditions that imply metrizability.

Considering the Euclidean metric d on $\mathbb{R}^m$, or on any subset $A \subset \mathbb{R}^m$, we see that A is endowed with a canonical topology- called **the Euclidean topology** on the subset $A \subset \mathbb{R}^m$ (exercise: show that also the square metric induces the same topology). Note that, in this case, the resulting notion of convergence (and continuity) coincides with the one from Analysis.

Remark 1.2. Continuing the previous remark, let us point out that showing that two Euclidean spaces $\mathbb{R}^m$ and $\mathbb{R}^n$ of different dimensions $m \neq n$ are not homeomorphic is non-trivial. When $m = 1$, this can be done using the notion of connectedness but, for $m, n \geq 2$, one has to appeal to tools from Algebraic Topology.

We now return to general metric spaces. A metric d on X allows us to talk about the open balls $B_d(x, r)$ for $x \in X$, $r \in \mathbb{R}_+$ (see (3.1)), giving rise to the collection of open balls induced by d:

$$\mathcal{B}_d = \{B_d(x, r) : x \in X, r \in \mathbb{R}_+\}.$$

This is not a topology on X, but $\mathcal{T}_d$ is the smallest topology containing $\mathcal{B}_d$.

Recall also (simple exercise) that any family $\mathcal{B}$ of subsets of X gives rise to a topology $\mathcal{T}(\mathcal{B})$ on X, defined as the smallest one containing $\mathcal{B}$. It is called **the topology generated by** $\mathcal{B}$. In general, the members of $\mathcal{T}(\mathcal{B})$ are arbitrary unions of finite intersections of members of $\mathcal{T}$.

Depending on the properties of $\mathcal{B}$, the members of $\mathcal{T}(\mathcal{B})$ may have simpler descriptions. The most common case is when $\mathcal{B}$ is a **topology basis**, i.e. satisfies the following axioms: any $x \in X$ is contained in at least one member B of $\mathcal{B}$ and, for any $B_1, B_2 \in \mathcal{B}$ and any $x \in B_1 \cap B_2$, there exists $B \in \mathcal{B}$ containing x with $B \subset B_1 \cap B_2$. In this case, for $U \subset X$, the following are equivalent:

(0) U belongs to $\mathcal{T}(\mathcal{B})$.
(1) for any $x \in U$ there exists $B \in \mathcal{B}$ s.t. $x \in B \subset U$.
(2) U is a union of members of $\mathcal{B}$.

For instance, for any metric d on X, the collection $\mathcal{B}_d$ is a topology basis, and (1) is precisely the original definition of $\mathcal{T}_d$.

One can change a bit the point of view and, starting with a topology $\mathcal{T}$ on X, look for collections $\mathcal{B}$ generating $\mathcal{T}$, i.e. such that $\mathcal{T} = \mathcal{T}(\mathcal{B})$. Of course, one possibility is to take $\mathcal{B} = \mathcal{T}$, but this is the least interesting one. The more interesting choices are the ones for which $\mathcal{B}$ is smaller- e.g. countable. And here is the precise terminology: given a topological space X, a **basis for the topological space** X is any collection $\mathcal{B}$ of subsets of X with the property it is a topology basis and $\mathcal{T} = \mathcal{T}(\mathcal{B})$. As above, for a collection $\mathcal{B}$ of subsets of X, the following are equivalent:

(0) $\mathcal{B}$ is a basis for the space X.
(1) for any open U in X and any $x \in U$ there exists $B \in \mathcal{B}$ s.t. $x \in B \subset U$.
(2) any open in X is a union of members of $\mathcal{B}$.

(in particular, each of the conditions (1) and (2) imply that $\mathcal{B}$ is a topology basis).

Repeating what we said before, but with a slightly different wording, we have that for any metric d on X, the metric topology admits $\mathcal{B}_d$ as basis. Another possible basis for the space X (endowed with the topology $\mathcal{T}_d$), slightly smaller, is

$$\mathcal{B}_d = \{B_d(x, \frac{1}{n}) : x \in X, n \in X\}.$$

For the Euclidean metric d_{Eucl} on $\mathbb{R}^m$ we can do even better:

$$\mathcal{B}_{\mathbb{Q}} := \{B_{d_{\text{Eucl}}}(q, \frac{1}{n}) : q \in \mathbb{Q}^m, n \in \mathbb{N}\}$$

is still a basis for the Euclidean topology on $\mathbb{R}^m$, but it it "much smaller": it is countable.

In general, one says that a topological space X is **second countable** if it admits a basis $\mathcal{B}$ which is countable.

4. Topological manifolds

The second countability condition is a very subtle one and turns out to be of capital importance in establishing some central results in Topology and Geometry- such as metrizability and embedding theorems. In particular, it is part of the basic axioms for the notion of manifolds. For now:

DEFINITION 1.3. *A d-dimensional **topological manifold** is any topological space X satisfying the following properties:*

- (TM0) *it is locally Euclidean (of dimension d), i.e.: any point $x \in X$ admits a neighborhood X which is homeomorphic to an open subset of $\mathbb{R}^d$.*
- (TM1) *it is Hausdorff.*
- (TM2) *it is second countable.*

***A (d-dimensional) chart of** M is any homeomorphism*

$$\chi : U \to \Omega \subset \mathbb{R}^d$$

*from an open subset U of M (called **the domain of the chart**) and an open subset Ω in $\mathbb{R}^d$. Hence axiom (TM0) says that M can be covered by (domains of) d-dimensional charts.*

You should convince yourself (or remember) why some of the toy examples of topological spaces such as spheres, tori, Moebius band, etc are topological manifolds. Also, in all these examples, one should concentrate first on the condition (TM0). Note however that, while the notion of dimension is intuitively clear (at least in all examples), handling it theoretically is not such a piece of cake; see Remark 1.2. This is due to the fact that there is no obvious topological characterization of the (intuitive notion) of dimension. This will be much less of a problem as soon as we move to (differentiable) manifolds.

REMARK 1.4. Since the notion of "topological space" is build on the notion of "open", so are most of the basic definitions in Topology- such as continuity, Hausdorffness, compactness, etc etc. However, under rather mild assumptions, such definitions can be rephrased more intuitively, using sequences. The main "mild assumption" that we have in mind here is that of "first countability"; please see the basic course on Topology. This condition weaker even than the second countability condition. For instance, metric topologies are always first countable but may fail to be second countable. For our purpose, it is enough to know that either of the conditions (TM0) or (TM2) implies 1st countability (and, if you look at the definitions, you will see that this statement is completely trivial).

What is interesting to know here is that, when restricting to spaces X which are first countable, many of the basic notions can be reformulated in terms of sequences. E.g.:

- X is Hausdorff iff any convergent sequence in X has at most one limit.
- $f : X \to Y$ is continuous iff it is sequential continuous i.e.: if $(x_n)_{n\geq 1}$ is a sequence converging in X to $x \in X$, then $(f(x_n))_{n\geq 1}$ converges in Y to $f(x)$.

5. Inside a topological space

Recall that, given a space X, a subset $A \subset X$ is said to be **closed in** X if its complement $X \setminus A$ is open. Of course, knowing the closed subsets of X is equivalent to knowing the open ones- hence one could have introduced the notion of topology completely in terms of closed subsets (which would then be the axioms?). Opens are preferred because some of the the most important properties can be described more directly in terms of opens (and perhaps also because they are closer in spirit to the notion of "ball" in a metric space). However, closed subsets often have some very nice properties- e.g. when talking about compactness.

Given the axioms of a topology (namely the fact that arbitrary unions of opens is open or, equivalently, that arbitrary intersections of closeds is closed), it follows that for any subset A of a topological space X one can talk about:

- the largest open contained in A- and this is called **the interior of** A (in the space X), and denoted $\text{Int}(A)$.
- the smallest closed containing A- and this is called **the closure of** A (in the space A), and denoted $\text{Cl}(A)$,

Recall also that, under the first countability axiom (in particular, for topological manifolds), the closure has a particularly nice description in terms of sequences:

$$\text{Cl}(A) = \{x \in X : \exists \text{ a sequence in } A \text{ converging to } x\}.$$

6. Construction of topological spaces

We have already seen two (related) ways of constructing topologies on a set X: the metric topology $\mathcal{T}_d$ induced by any metric d on X, and the topology $\mathcal{T}(\mathcal{B})$ generated by any family $\mathcal{B}$ of subsets of X (with the particularly nice situation when $\mathcal{B}$ is a topology basis).

There are various other important constructions of topologies out of the old ones. For instance, given any two topological spaces X and Y, the Cartesian product

$$X \times Y = \{(x, y) : x \in X, y \in Y\}$$

carries a canonical topology, called **the product topology**. There is a slight complication: while we would like that the products of opens is open,

$$\mathcal{B}_{X\times Y} := \{U \times V : U - \text{open in } X, V - \text{open in } Y\}$$

is not a topology on $X \times Y$; instead, it is a topology basis, and the product topology is defined as the topology generated by $\mathcal{B}_{X\times Y}$. Equivalently, and more conceptually, it is the smallest topology on $X \times Y$ with the property that the projections

$$\text{pr}_X : X \times Y \to X, \text{pr}_Y : X \times Y \to Y$$

are continuous.

The last description is more conceptual because it follows a general philosophy that one should apply when looking for topologies: require that the most interesting maps that you have around to be continuous, and look for "the best (least boring)" topology that does that (usually "the best" means "the largest" or "the smallest").

Another example of this philosophy is **the induced topology**: given a topological space X, any subset $A \subset X$ carries a canonical, induced, topology: it is the smallest topology with the property that the canonical inclusion

$$i : A \to X, \quad i(a) = a$$

is continuous (why looking for the largest topology with this property would be "boring"?). Explicitly, the opens in A (endowed with the topology induced from X) are the intersections $A \cap U$ of A with opens U of X.

Yet another example is that of **quotient topology**. In some sense, it is at the other extreme than the previous example. While before we started with an inclusion $I : A \to X$, we now start with a surjection

$$\pi : X \to Y,$$

where X is a topological space and Y is just a set (on which we would like to induce a topology). This time, looking for "the most interesting" topology on Y, we are lead to looking at the largest topology on Y with the property that π is continuous (why?). We obtain the quotient topology on Y: a subset $U \subset Y$ is an open of this topology if an donly if $\pi^{-1}(U)$ is open in X (check that this is, indeed, a topology on Y).

The terminology "quotient" comes from the fact that, typically, the situation of having a surjection $\pi : X \to Y$ arises when starting with X and an equivalence relation R on X. Then, with the intuition that we want to glue the points of X that are equivalent (w.r.t. the equivalence relation R), we obtain the quotient space

$$Y = X/R$$

(abstractly made of R-equivalence classes $[x]_R = \{y \in X : (x, y) \in R\}$ of points $x \in X$) together with the canonical projection

$$\pi_R : X \to X/R, \quad \pi_R(x) = [x]_R.$$

Therefore, starting with an equivalence relation R on a topological space X, we see that the resulting quotient X/R carries a canonical (quotient) topology.

One of the most interesting examples of quotient topologies is the canonical topology on the projective space

$$\mathbb{P}^m := \{l : l - \text{line through the origin in } \mathbb{R}^{m+1}\}$$

(i.e. the set of all 1-dimensional vector subspaces of $\mathbb{R}^{m+1}$). We can put ourselves in the previous situation by considering

$$\pi : \mathbb{R}^{m+1} \setminus \{0\} \to \mathbb{P}^m, x \mapsto l_x,$$

where l_x is the line through the origin and x (i.e. the vector subspace $\mathbb{R} \cdot x$ spanned by x). In terms of equivalence relations, we deal with the equivalence relation on $\mathbb{R}^{m+1} \setminus \{0\}$ given by:

$$x \sim y \Longleftrightarrow l_x = l_y \Longleftrightarrow y = \lambda \cdot x \text{ for some } \lambda \in \mathbb{R}.$$

Using the Euclidean topology on $\mathbb{R}^{m+1} \setminus \{0\}$ we obtain a natural topology on $\mathbb{P}^m$; endowed with this topology, $\mathbb{P}^m$ is called **the projective space** (of dimension m). You should convince yourself that convergence in this topology corresponds to the intuitive idea of "lines getting close to each other".

Finally, given the notion of induced topology, one can make use of that of embeddings: an **embedding** of a topological space X into a topological space Y is any map $i : X \to Y$ that is continuous, injective and, when interpreted as a continuous map $i : X \to i(X)$ and we endow $i(X) \subset Y$ with the topology induced from Y, it is a homemorphism (note that the last map is automatically continuous and bijective, but that does not imply that its inverse is continuous as well!). Next to metrizability theorems (see above), one of the most interesting problems in Topology/Geometry is that of deciding whether a space X can be embedded in an Euclidean space; results in this direction are usually labelled as **embedding theorems**. Looking at the notion of topological manifold, it is worth pointing out that, due also to the axioms (TM1) and

(TM2), it follows that any topological manifold is metrizable and can be embedded in some Euclidean space!

7. Topological properties

As we have pointed out in Remark 1.1, to distinguish topological spaces from each other (or to understand better each specific one), it is useful to isolate the various topological properties that spaces may have. By a topological property we mean any property that can be described by only using the notion of opens or, equivalently, any property that is preserved via homeomorphisms. We have already mentioned several such properties: Hausdorffness and second countability. Here we recall a few more.

The first one is that of connectedness: a space X is called **connected** if it cannot be written as $X = U \cup V$ with U, V-disjoint non-empty opens in X. Or, equivalently, if the only subsets of X that are both open and closed are $\emptyset$ and X. In general, if X is not connected, it can be "broken" into connected pieces; more precisely, recall that a **connected component** of a space X is any connected subset $C \subset X$ which (when endowed with the induced topology) is connected, and which is maximal (w.r.t. the inclusion) with this property. Then the set of connected components defines a partition of X by closed subspace. In examples, the partition into connected components is usually easy to guess intuitively; here is a simple exercise that can be used as a recipe to confirm such guesses: assume that we manage to write X as

$$X = X_1 \cup \ldots \cup X_k, \quad \text{with } X_i \cap X_j = \emptyset \text{ for } i \neq j.$$

Assume also that all the X_is are open or, equivalently (why?), that all the X_is are closed. Then $\{X_1, \ldots, X_k\}$ must coincide with the partition into connected components.

Remark 1.5. *Of course, the number of connected components may sometimes be infinite (even non-countable). Note however that, for topological manifolds M, due to the second countability axiom, the number of connected components if always at most countable (and finite if M is compact). Actually, one often restricts the attention to connected manifolds.*

Another important topological property is that of compactness. While this is a property that one usually encounters in the first courses in Analysis (compacts in $\mathbb{R}^m$ being the subsets $A \subset \mathbb{R}^m$ that are closed and bounded), the fact that this is a topological property (i.e. can be described by appealing only to the notion of opens in A, without any reference to the Euclidean metric or to the way that A sits inside $\mathbb{R}^m$) is not at all obvious. That makes the resulting general definition less intuitive and a bit hard to digest at first: a topological space X is said to be **compact** if for any open cover

$$\mathcal{U} = \{U_i : i \in I\}$$

of X (i.e. each U_i is open in X, their union is X, and I is an indexing set), one can extract a finite subcover, i.e. there exists $i_1, \ldots, i_k \in I$ such that $\{U_{i_1}, \ldots, U_{i_k}\}$ is still a cover of X- i.e.

$$X = U_{i_1} \cup \ldots \cup U_{i_k}.$$

Here is the list of the most important properties of compactness:

(1) Compact inside Hausdorff is closed: if X is a topological space, $A \subset X$ is endowed with the induced topology (see above) then:

$$A - \text{compact}, X - \text{Hausdorff} \Longrightarrow A - \text{is closed in} X.$$

(2) Closed inside compact is compact: if X is a topological space, $A \subset X$ is endowed with the induced topology (see above) then:

$$A - \text{is closed in } X, X - \text{compact} \Longrightarrow A - \text{is compact}$$

(3) any compact Hausdorff space is automatically normal. Recall here that a topological space X is said to be **normal** if for any $A, B \subset X$ closed disjoint subsets, one can find opens in X, U containing A and V containing B, such that $U \cap V = \emptyset$.
(4) Product of compacts is compact: if X and Y are compact spaces then $X \times Y$, endowed with the product topology (see above), is compact.
(5) Continuous applied to compact is compact: if $f : X \to Y$ is continuous and $A \subset X$ (with the induced topology) is compact, then so is $f(A) \subset Y$.
(6) In particular: quotients of compacts are compacts.
(7) A continuous bijection from a compact space to a Hausdorff one is automatically a homeomorphism! Or, slightly more generally: an continuous injection of a compact space into a Hausdorff one is an embedding (see above).

EXERCISE 1.1. Assuming that you already know that the unit interval $[0, 1]$ (endowed with the Euclidean topology) show (using the properties mentioned above) that a subset $A \subset \mathbb{R}^m$ (endowed with the Euclidean topology) is compact if and only if it is closed and bounded in $\mathbb{R}^m$.

A related topological property is the local version of compactness: one says that a space X is **locally compact** if any point $x \in X$ admits a compact neighborhood. If X is also Hausdorff, it follows that any point in X admits "arbitrarily small compact neighborhoods", i.e. for any neighborhood U of x in X there exists a compact neighbhborhood of x, contained in U. In general, Hausdorff locally compact spaces can be compactified by adding one extra-point. More on the 1-point compactification can be found in the lecture note on Topology. While, but axiom (MT0), topological manifolds are automatically locally compact, what is interesting for us to recall here is the interaction between (MT1), (MT2) and local compactness: the existence of exhaustions. This is Theorem 4.37 in the notes on Topology:

THEOREM 1.6. *Any locally compact, Hausdorff, 2nd countable space X admits an exhaustion, i.e. a family $\{K_n : n \in \mathbb{Z}_+\}$ of compact subsets of X such that $X = \cup_n K_n$ and $K_n \subset \overset{\circ}{K}_{n+1}$ for all n.*

PROOF. Let $\mathcal{B}$ be a countable basis and consider $\mathcal{V} = \{B \in \mathcal{B} : \overline{B} - \text{compact}\}$. Then $\mathcal{V}$ is a basis: for any open U and $x \in X$ we choose a compact neighborhood N inside U; since $\mathcal{B}$ is a basis, we find $B \in \mathcal{B}$ s.t. $x \in B \subset N$; this implies $\overline{B} \subset N$ and then $\overline{B}$ must be compact; hence we found $B \in \mathcal{V}$ s.t. $x \in B \subset U$. In conclusion, we may assume that we have a basis $\mathcal{V} = \{V_n : n \in \mathbb{Z}_+\}$ where $\overline{V}_n$ is compact for each n. We define the exhaustion $\{K_n\}$ inductively, as follows. We put $K_1 = \overline{V}_1$. Since $\mathcal{V}$ covers the compact K_1, we find i_1 such that

$$K_1 \subset V_1 \cup V_2 \cup \ldots \cup V_{i_1}.$$

Denoting by D_1 the right hand side of the inclusion above, we put

$$K_2 = \overline{D}_1 = \overline{V}_1 \cup \overline{V}_2 \cup \ldots \cup \overline{V}_{i_1}.$$

This is compact because it is a finite union of compacts. Since $D_1 \subset K_2$ and D_1 is open, we must have $D_1 \subset \overset{\circ}{K}_2$; since $K_1 \subset D_1$, we have $K_1 \subset \overset{\circ}{K}_2$. Next, we choose $i_2 > i_1$ such that

$$K_2 \subset V_1 \cup V_2 \cup \ldots \cup V_{i_2},$$

we denote by D_2 the right hand side of this inclusion, and we put

$$K_3 = \overline{D}_2 = \overline{V}_1 \cup \overline{V}_2 \cup \ldots \cup \overline{V}_{i_2}.$$

As before, K_3 is compact, its interior contains D_2, hence also K_2. Continuing this process, we construct the family K_n, which clearly covers X. □

8. The algebra of continuous functions

Given a topological space X, an "observable on X" has a precise meaning: it is a continuous function

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

The set of all such continuous functions is denoted by

$$\mathcal{C}(X).$$

One of the simplest but most fundamental ideas in various parts of Geometry is that of understanding a space X via the associated "object" $\mathcal{C}(X)$. This will allow one to consider "more relevant observables": e.g. for subspaces $X \subset \mathbb{R}^m$, one can consider only fs that are smooth, or polynomials. Or even to handle "spaces" which, although are quite intuitive, are not topological spaces in the strict sense of the word. All together, this point of view gives rise to several directions in Geometry: Differential Geometry (where the key-word is "smooth" instead of "continuous"), Algebraic Geometry (where the key-word is "polynomial", or "complex analytic"), Noncommutative Geometry (where X does not even make sense, but $\mathcal{C}(X)$ does).

Of course, what makes these work is the rich structure that $\mathcal{C}(X)$ posses- making the "object" $\mathcal{C}(X)$ (a priory just a set) into a more interesting mathematical object. We recall here the most important part of the algebraic structure present on $\mathcal{C}(X)$: it is an algebra. Recall here:

Definition 1.7. *A (real)* ***algebra*** *is a vector space A over $\mathbb{R}$ together with an operation*

$$A \times A \to A, \quad (a, b) \mapsto a \cdot b$$

which is unital in the sense that there exists an element $1 \in A$ such that

$$1 \cdot a = a \cdot 1 = a \quad \forall\, a \in A,$$

and which is $\mathbb{R}$-bilinear and associative, i.e., for all $a, a', b, b', c \in A$, $\lambda \in \mathbb{R}$,

$$(a + a') \cdot b = a \cdot b + a' \cdot b,\ \ a \cdot (b + b') = a \cdot b + a \cdot b',$$

$$(\lambda a) \cdot b = \lambda(a \cdot b) = a \cdot (\lambda b),,$$

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c\ .$$

We say that A is commutative if $a \cdot b = b \cdot a$ for all $a, b \in \mathcal{A}$.

Similarly one talks about complex algebras: then A is a vector space over $\mathbb{C}$ and $\lambda \in \mathbb{C}$.

For a topological space X, the algebra structure on $\mathcal{C}(X)$ is defined simply by pointwise addition and multiplication: for $f, g \in \mathcal{C}(X)$ and $\lambda \in \mathbb{R}$, $f + g, f \cdot g, \lambda \cdot f \in \mathcal{C}(X)$ are given by:

$$(f + g)(x) = f(x) + g(x), \quad (f \cdot g)(x) = f(x)g(x), \quad (\lambda \cdot f)(x) = \lambda f(x).$$

And, considering the space $\mathcal{C}(X, \mathbb{C})$ of $\mathbb{C}$-valued continuous functions on X, one obtains a complex algebra.

The fact that, under certain assumptions, a topological space X can be recovered from the algebra $\mathcal{C}(X)$, is the content of the Gelfand-Naimark theorem. While we refer to the basic course on Topology for the full statement and details, here is the very brief summary:

Theorem 1.8 (informative version of Gelfand Naimark theorem). *There is a way to associate to any algebra A a topological space $X(A)$ (called the spectrum of A) so that, when applied to $A = \mathcal{C}(X)$- the algebra of continuous functions on a compact Hausdorff space X, one recovers X (i.e. $X(\mathcal{C}(X))$ is homeomorphic to X).*

And, with the mind at the fact that we may want to consider more restrictive conditions than continuity (e.g. smoothness), here is the resulting relevant abstract notion:

Definition 1.9. *Given an algebra A (over the base field $\mathbb{R}$ or $\mathbb{C}$), a* ***subalgebra*** *of A is any vector subspace $B \subset A$, containing the unit 1 of A and such that*

$$b \cdot b' \in B \quad \forall\, b, b' \in B.$$

When we want to be more specific about the base field, we talk about real or complex subalgebras.

For instance, for $X \subset \mathbb{R}^m$, when looking at smooth or polynomial functions, we obtain a sequence of sub-algebras:

$$\mathcal{C}^{\mathrm{polyn}}(X) \subset \mathcal{C}^{\infty}(X) \subset \mathcal{C}(X).$$

Finally, when looking at a subset

$$\mathcal{A} \subset \mathcal{C}(X),$$

(subalgebra or not), there are several interesting properties that turn out to be interesting- and we say that:

(1) $\mathcal{A}$ is **point separating** if for any $x, y \in X$ distinct there exists $f \in \mathcal{A}$ such that $f(x) \neq f(y)$ or, equivalently, if there exists $f \in \mathcal{A}$ such that $f(x) = 0$ and $f(y) = 1$.
(2) $\mathcal{A}$ is **normal** if for any two disjoint closed subset $A, B \subset X$, there exists $f \in \mathcal{C}(X)$ such that $f|_A = 0$, $f|_B = 1$.
(3) $\mathcal{A}$ is **closed under sums** if $f + g \in \mathcal{A}$ whenever $f, g \in \mathcal{A}$.
(3) $\mathcal{A}$ is **closed under quotients** if $f/g \in \mathcal{A}$ whenever $f, g \in \mathcal{A}$ and g is nowhere vanishing.

For instance, the Stone-Weierstrass theorem (which will not be used in the rest of the course) says that, if X is a compact Hausdorff space, then any point-separating sub-algebra $\mathcal{A} \subset \mathcal{C}(X)$ is dense in $\mathcal{C}(X)$; with particular cases of the type: real valued continuous functions on $[01, 1]$ (or other similar spaces) can be approximated by polynomial functions.

Note that, for a general topological space X, even the entire $\mathcal{A} = \mathcal{C}(X)$ need not be point separating or normal. Actually, it is a rather simple exercise to check that the point separation of $\mathcal{C}(X)$ implies that X must be Hausdorff, while the normality of $\mathcal{C}(X)$ implies that the topological space X must be normal (i.e., as recalled above: any two disjoint closed subsets $A, B \subset X$ can be separated topologically: there exist opens $U, V \subset X$ containing A and B, respectively, with $U \cap V = \emptyset$). What is far less obvious (actually one of the most non-trivial basic results in Topology) is the converse, known as the Urysohn lemma: if a topological space X is Hausdorff and normal then $\mathcal{C}(X)$ is normal; more precisely, for any two disjoint closed subsets $A, B \subset X$ there exists

$$f : X \to [0, 1] \text{ continuous and such that } f|_A = 0, f|_B = 1.$$

This will not be used later in the course; we mention it here just for completeness.

9. Partitions of unity

Finally, one more basic topic from Topology- but this time one that is difficult to appreciate (and perhaps even to digest) without entering the realm of Differential Geometry and/or Analysis: partitions of unity. To be able to talk about partitions of unity that are not just continuous (as we will be interested only on smooth functions), we can place ourselves in the following setting: X is a topological space and

$$\mathcal{A} \subset \mathcal{C}(X)$$

is a given vector subspace; we will be looking at partitions of unity that belong to $\mathcal{A}$. For the main definition, we first need to recall the notion of support: given $\eta : X \to \mathbb{R}$ continuous, **the support of η in X**, denoted $\mathrm{supp}_X(\eta)$ or simply $\mathrm{supp}(\eta)$ is the closure in X of the set $\eta \neq 0$ of points of X on which η does not vanish.

Given an open $U \subset X$, we say that η **is supported in** U if $\mathrm{supp}(\eta) \subset U$. This condition allows one to promote functions that are defined only on U, $f : U \to \mathbb{R}$, to functions on X, at least after multiplying by η; namely, $\eta \cdot f$, a priory defined only on U, if extended to X by declaring it to be zero outside U, the resulting function

$$\eta \cdot f : X \to \mathbb{R}$$

will be continuous (check this and, by looking at examples, convince yourselves that this does not work if the condition $\text{supp}(\eta) \subset U$ is replaced by the weaker one that $\{\eta \neq 0\} \subset U$).

We first recall the notion of finite partitions of unity:

Definition 1.10. *Let X be a topological space, $\mathcal{U} = \{U_1, \ldots, U_n\}$ a finite open cover of X. A continuous partition of unity subordinated to $\mathcal{U}$ is a family of continuous functions $\eta_i : X \to [0,1]$ satisfying:*

$$\eta_1 + \ldots + \eta_k = 1, \quad \mathit{supp}(\eta_i) \subset U_i.$$

Given $\mathcal{A} \subset \mathcal{C}(X)$, we say that $\{\eta_i\}$ is an $\mathcal{A}$-partition of unity if $\eta_i \in \mathcal{A}$ for all i.

Theorem 1.11. *Let X be a topological space and assume that $\mathcal{A} \subset \mathcal{C}(X)$ is normal and is closed under sums and quotients. Then, for any finite open cover $\mathcal{U}$, there exists an $\mathcal{A}$-partition of unity subordinated to $\mathcal{U}$.*

In particular if X is Hausdorff and normal, by the Uryshon Lemma (to ensure that $\mathcal{A} := \mathcal{C}(X)$ is normal), any finite open cover $\mathcal{U}$ admits a continuous partition of unity subordinated to $\mathcal{U}$.

Proof (sketch; for more details, see the lecture notes on Topology). The main ingredients are:

(St1) the remark made above that the normality of $\mathcal{A}$ implies that X is a normal space (simple exercise).

(St2) the fact that, in a normal space X, whenever we have $A \subset U$ with A-closed in X and U-open in X, one can find a smaller open V such that

$$A \subset V \subset \overline{V} \subset U$$

(short proof, but a bit tricky).

(St3) the shrinking lemma: for any finite open cover $\mathcal{U} = \{U_1, \ldots, U_k\}$ of a normal space X one can find another cover $\mathcal{V} = \{V_1, \ldots, V_n\}$ such that

$$\overline{V}_i \subset U_i \quad \forall\, i \in \{1, \ldots, k\}.$$

(this follows by applying the previous step inductively, starting with $U = U_1$ $A = X \setminus (U_2 \cup \ldots \cup U_k)$.

Now the proof of the theorem. Apply the shrinking lemma twice and choose open covers $\mathcal{V} = \{V_i\}$, $\mathcal{W} = \{W_i\}$, with $\overline{V}_i \subset U_i$, $\overline{W}_i \subset V_i$. For each i, we use the separation property of $\mathcal{A}$ for the disjoint closed sets $(\overline{W}_i, X - V_i)$. We find $f_i : X \to [0,1]$ that belongs to $\mathcal{A}$, with $f_i = 1$ on $\overline{W}_i$ and $f_i = 0$ outside V_i. Note that

$$f := f_1 + \ldots + f_k$$

is nowhere zero. Indeed, if $f(x) = 0$, we must have $f_i(x) = 0$ for all i, hence, for all i, $x \notin W_i$. But this contradicts the fact that $\mathcal{W}$ is a cover of X. From the properties of $\mathcal{A}$, each

$$\eta_i := \frac{f_i}{f_1 + \ldots + f_k} : X \to [0,1]$$

is continuous. Clearly, their sum is 1. Finally, $\text{supp}(\eta_i) \subset U_i$ because $\overline{V}_i \subset U_i$ and$\{x : \eta_i(x) \neq 0\} = \{x : f_i(x) \neq 0\} \subset V_i$. □

And here is a nice application of the existence of (finite) partitions of unity:

Theorem 1.12. *Any compact topological manifold M can be embedded in some Euclidean space $\mathbb{R}^m$.*

PROOF. Cover M by opens that are homeomorphic to $\mathbb{R}^d$, where d is the dimension of M. Using that M is compact, we find an open cover $\mathcal{U} = \{U_1, \ldots, U_n\}$ together with homeomorphisms $\chi_i : U_i \to \mathbb{R}^d$. Since M is compact it is also normal hence we find a partition of unity $\{\eta_1, \ldots, \eta_n\}$ subordinated to $\mathcal{U}$. Each of the functions $\eta_i \cdot \chi_i : U_i \to \mathbb{R}^d$ is extended to M by declaring it to be zero outside U_i; by the previous comments, the resulting functions $\tilde{\chi}_i : M \to \mathbb{R}^d$ are continuous. Consider now

$$i = (\eta_1, \ldots, \eta_k, \tilde{\chi}_1, \ldots, \tilde{\chi}_k) : M \to \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{k \text{ times}} \times \underbrace{\mathbb{R}^d \times \ldots \times \mathbb{R}^d}_{k \text{ times}} = \mathbb{R}^{k(d+1)}.$$

One check directly that i is injective; since M is compact and $\mathbb{R}^{k(d+1)}$ is Hausdorff, by the properties recalled on compactness, i will be an embedding. □

Finite partitions of unity are useful mainly when working over compacts (so that one can ensure finite open covers). For the more general case one first has to make precise sense of "infinite sums $\sum_i \eta_i$". For that first recall that, given a topological space X, a family $\mathcal{S}$ of subsets of X is said to be **locally finite** (in X) if for any $x \in X$ there exists a neighborhood V of x which intersect only a finite number of members of $\mathcal{S}$. Given a family $\{\eta_i\}_{i \in I}$ (I some indexing set) of continuous functions $\eta_i : X \to \mathbb{R}$, we say that $\{\eta_i\}_{i \in I}$ **is locally finite** in X if their supports (in X) $\text{supp}(\eta_i)$ form a locally finite family of subsets of X. Note that in this case the sum

$$\sum_{i \in I} \eta_i : X \to \mathbb{R}$$

can be defined pointwise (at any $x \in X$ only a finite number of terms do not vanish), and the resulting function is continuous. For a subset $\mathcal{A} \subset \mathcal{C}(X)$, we say that is **closed under locally finite sums** if for any locally finite family $\{\eta_i\}$ with $\eta_i \in \mathcal{A}$, $\sum_i \eta_i$ is again in $\mathcal{A}$.

With these, we can not talk about infinite partitions of unity:

DEFINITION 1.13. *Let X be a topological space, $\mathcal{U} = \{U_i : i \in I\}$ an open cover of X. A (continuous) partition of unity subordinated to $\mathcal{U}$ is a locally finite family of continuous functions $\eta_i : X \to [0, 1]$ satisfying:*

$$\sum_{i \in I} \eta_i = 1, \quad supp(\eta_i) \subset U_i.$$

Given $\mathcal{A} \subset \mathcal{C}(X)$, we say that $\{\eta_i\}$ is an $\mathcal{A}$-partition of unity if $\eta_i \in \mathcal{A}$ for all i.

The existence of partitions of unity subordinated to arbitrary open covers forces a topological property of X called paracompactness: we say that a topological space X is **paracompact** if for any open cover $\mathcal{U}$ of X, there exists a locally finite open cover $\mathcal{V}$ that is a refinement of $\mathcal{U}$ in the sense that any $V \in \mathcal{V}$ is included inside some $U \in \mathcal{U}$. The existence of arbitrary partitions of unity is ensured by the following:

THEOREM 1.14. *Let X be a paracompact Hausdorff space and assume that $\mathcal{A} \subset \mathcal{C}(X)$ is normal, closed under locally finite sums and closed under quotients.*

Then, for any open cover $\mathcal{U}$ of X, there exists an $\mathcal{A}$-partition of unity subordinated to $\mathcal{U}$.

Since paracompact spaces are automatically normal, hence we can use Uryshon's lemma, it follow that in a paracompact Hausdorff space a for any open cover there exists a continuous partition of unity subordinated to the cover.

The proof of the previous theorem is almost identical with the one from the finite case- just that one now has to establish an infinite version of the shrinking lemma (and that is where paracompactness enters), To apply the previous theorem, there are two points that may be difficult to check: the paracompactness of X and, when working with arbitrary $\mathcal{A}$, that $\mathcal{A}$ is normal. For the first one, the following comes in handy:

THEOREM 1.15. *Any Hausdorff, locally compact and 2nd countable space is paracompact.*

In particular, topological manifolds are automatically compact. One can actually show that, under the axioms (TM0) and (TM1), the axiom (TM2) on second countability is equivalent to the fact that M is paracompact and has a countable number of connected components.

PROOF. We use an exhaustion $\{K_n\}$ of X (Theorem 1.6). Let $\mathcal{U}$ be an open cover of X. For each $n \in \mathbb{Z}_+$ there is a finite family $\mathcal{V}_n$ which covers $K_n - \mathrm{Int}(K_{n-1})$, consisting of opens V with the properties: $V \subset \mathrm{Int}(K_{n+1}) - K_{n-1}$, $V \subset U$ for some $U \in \mathcal{U}$. Indeed, for any $x \in K_n - \mathrm{Int}(K_{n-1})$ let V_x be the intersection of $\mathrm{Int}(K_{n+1}) - K_{n-1}$ with any member of $\mathcal{U}$ containing x; since $K_n - \mathrm{Int}(K_{n-1})$ is compact, just take a finite subcollection $\mathcal{V}_n$ of $\{V_x\}$, covering $K_n - \mathrm{Int}(K_{n-1})$. Set $\mathcal{V} = \cup_n \mathcal{V}_n$; it covers X since each $K_n - K_{n-1} \subset K_n - \mathrm{Int}(K_{n-1})$ is covered by $\mathcal{V}_n$. Finally, it is locally finite: if $x \in X$, choosing n and V such that $V \in \mathcal{V}_n$, $x \in V$, we have $V \subset \mathrm{Int}(K_{n+1}) - K_{n-1}$, hence V can only intersect members of $\mathcal{V}_m$ with $m \leq n+1$ (a finite number of them!). □

Finally, to check the normality of $\mathcal{A}$ needed in Theorem 1.14, the following comes in handy:

THEOREM 1.16. *Let X be a Hausdorff paracompact space and $\mathcal{A} \subset \mathcal{C}(X)$ closed under locally finite sums and under quotients. If X is also locally compact, then the following are equivalent:*

1. *$\mathcal{A}$ is normal.*
2. *for any $x \in M$ and any open neighborhood U of x, there exists $f \in \mathcal{A}$ supported in U with $f(x) > 0$.*

In particular, for a topological manifold M, checking that a subset $\mathcal{A} \subset \mathcal{C}(M)$ is normal is a local matter- and that is very useful since, locally, topological manifolds look just like Euclidean spaces.

PROOF. That 1 implies 2 is clear: apply the separation property to $\{x\}$ and $X - V$. Assume 2. We claim that for any $C \subset X$ compact and any open U such that $C \subset U$, there exists $f \in \mathcal{A}$ supported in U, such that $f|_C > 0$. Indeed, by hypothesis, for any $c \in C$ we can find an open neighborhood V_c of c and $f_c \in \mathcal{A}$ positive such that $f_c(c) > 0$; then $\{f_c \neq 0\}_{c \in C}$ is an open cover of C in X, hence we can find a finite subcollection (corresponding to some points $c_1, \ldots, c_k \in C$) which still covers C; finally, set $f = f_{c_1} + \ldots + f_{c_k}$.

To prove 1, let $A, B \subset X$ be two closed disjoint subsets. As terminology, $D \subset X$ is called relatively compact if $\overline{D}$ is compact. Since X is locally compact, any point has arbitrarily small relatively compact open neighborhoods. For each $y \in X - A$, we choose such a neighborhood $D_y \subset X - A$. For each $a \in A$, since $a \in X - B$, applying step (St2) from the proof of Theorem 1.11, we find an open D_a such that $a \in D_a \subset X - B$. Again, we may assume that $\overline{D}_a$ is relatively compact. Then $\{D_x : x \in X\}$ is an open cover of X; let $\mathcal{U} = \{U_i : i \in I\}$ be a locally finite refinement. We split the set of indices as $I = I_1 \cup I_2$, where I_1 contains those i for which $U_i \cap A \neq \emptyset$, while I_2 those for which $U_i \subset X - A$. Using the shrinking lemma (the infinite version of the one described in (St3) of he proof of Theorem 1.11) we can also choose an open cover of X, $\mathcal{V} = \{V_i : i \in I\}$, with $\overline{V}_i \subset U_i$. Note that, by construction, each U_i (hence also each V_i) is relatively compact. Hence, by the claim above, we can find $\eta_i \in \mathcal{A}$ such that

$$\eta_i|_{\overline{V}_i} > 0, \quad \mathrm{supp}(\eta_i) \subset U_i.$$

Finally, we define

$$f(x) = \frac{\sum_{i\in I_1} \eta_i(x)}{\sum_{i\in I} \eta_i(x)}$$

From the properties of $\mathcal{A}$, $f \in \mathcal{A}$. Also, $f|_A = 1$. Indeed, for $a \in A$, a cannot belong to the U_i's with $i \in I_2$ (i.e. those $\subset X - A$); hence $\eta_i(a) = 0$ for all $i \in I_2$, hence $f(a) = 1$. Finally, $f|_B = 0$. To see this, we show that $\eta_i(b) = 0$ for all $i \in I_1$, $b \in B$. Assume the contrary. We find $i \in I_1$ and $b \in B \cap U_i$. Now, from the construction of $\mathcal{U}$, $U_i \subset D_x$ for some $x \in X$. There are two cases. If $x = a \in A$, then the defining property for D_a, namely $D_a \cap B = \emptyset$, is in contradiction with our assumption ($b \in B \cap U_i$). If $x = y \in X - A$, then the defining property for D_y, i.e. $D_y \subset X - A$, is in contradiction with the fact that $i \in I_1$ (i.e. $U_i \cap A \neq \emptyset$). $\square$

CHAPTER 2

Reminder on Analysis

The relationship between Analysis and Differential Geometry is subtle. On one hand, Differential Geometry relies on the very basics of Analysis. On the other hand, various notions/results from Analysis become much more transparent/intuitive once the geometric perspective/intuition is brought into picture. In some sense, in many cases, the geometric point of view indicates the (expected) results while analysis provides the tools to prove them.

1. $\mathbb{R}^m$

The basic playground for multivariate analysis is the standard Euclidean space

$$\mathbb{R}^m = \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{m \ times}.$$

Despite its simplicity, this "space" has many different (but related) structures- and often the problem with handling $\mathbb{R}^m$ comes from the fact that it may not be completely clear which of the structures present on $\mathbb{R}^m$ is relevant for the specific discussions. Here are some of the many interesting structures present on $\mathbb{R}^m$:

- it is a vector space. When we want to emphasize this structure, we will denoted by
$$v = (v_1, \ldots, v_m) \in \mathbb{R}^m$$
its elements and we will think of them as "vectors"/"directions". Intrinsic in this notation is the presence of yet another piece of structure: it is not just a vector space- it comes with a preferred (canonical) basis:
$$e_1, \ldots, e_m \in \mathbb{R}^m;$$
in coordinates, e_i has 1 on the i-th position and 0 everywhere else.
- it is a vector space endowed with an inner product:
$$\langle v, w \rangle = \sum_{i=1}^{m} v_i \cdot w_i,$$
hence it is also a normed vector space, with the norm:
$$||v|| = \sqrt{\langle v, v \rangle}.$$
- it is a topological space- endowed with the standard Euclidean topology. When we want to emphasize this structure, we will denote by
$$x = (x_1, \ldots, x_m) \in \mathbb{R}^m$$
its elements and we will think of them as "points". For instance, when looking at a circle in $\mathbb{R}^2$, the vector space space structure on $\mathbb{R}^2$ is not so relevant, and we think of the circle as made by points rather then vectors. Also, when talking about the continuity of a function $f : \mathbb{R}^m \to \mathbb{R}$, the vector space structure of $\mathbb{R}^m$ is not relevant (though it may be useful).

Recall also that the topology on $\mathbb{R}^m$ is a shadow of yet another structure: $\mathbb{R}^m$ is also a metric space, with the standard Euclidean metric:

$$d(x, y) = \sqrt{\sum_{i=1}^{m} (x_i - y_i)^2}.$$

We say a "shadow" because uses part of what the metric allows us top talks about: points being "close to each other" (or, more precisely: convergence and continuity). In particular, there are several other natural metrics on $\mathbb{R}^m$ that induce the same Euclidean topology- e.g. the so called square metric

$$d'(x, y) = \max\{|x_1 - y_1|, \ldots, |x_m - y_m|\}.$$

- even when thinking about $\mathbb{R}^m$ as a topological space, so of its elements as points, each point $x \in \mathbb{R}^m$ can be represented using "canonical coordinates"- used already above. Again, the coordinates are not so relevant/important: they are useful and can be used, but they are not intrinsic to the structure. For instance, a circle in $\mathbb{R}^2$ can be described using coordinates by the equation $x^2+y^2 = 1$, but the circle itself can be drawn without any coordinate axes at our disposal.

  Note that the standard coordinates we mentioned are the simplest illustration of the notion of "chart"- to be discussed in a bit more detail below, and essential in defining the notion of manifold.
- it is a topological space "on which analysis can be performed" (... i.e. a manifold).
- etc.

Of course, all these are inter-related but, in each situation, it is important to realize which of these structures really matter. In particular, whenever one encounters a definition or result, it is instructive to figure out whether the elements in $\mathbb{R}^m$ that show up play the role of points and which ones of vectors, and how much the definition/result depends on the coordinates. This is the first step towards a geometric understanding of Analysis.

2. The differential and the inverse function theorem

One can talk about various notions of derivatives of a function f at a point

$$x \in \mathbb{R}^m$$

whenever we have a function f defined on a neighborhood of x- so that the expressions $f(y)$ used below makes sense for all y near x or, equivalently, $f(x + v)$ makes is defined for small vectors v.

Typically one assumes that f is defined on an open subset $\Omega \subset \mathbb{R}^m$ and takes values in some other Euclidean space $\mathbb{R}^n$,

$$f : \Omega \to \mathbb{R}^n,$$

so that it makes sense to talk about derivatives of f at any point in its domain, $x \in \Omega$.

The most intrinsic notion of derivative is that of "**total derivative**", also called **the differential of** f (at the given point $x \in \Omega$). This notion arises when trying to approximate f, near x, by simpler (linear-like) functions. Understanding f near x is about understanding

$$v \mapsto f(x + v)$$

for $v \in \mathbb{R}^m$ near 0. It sends $v = 0$ to $f(x)$, hence the best one can hope for is to approximate

$$v \mapsto f(x + v) - f(x)$$

by functions that are linear in v. Think that we try to write the last expression as a function linear in v, plus one that is qudratic in v, etc (plus eventually an "error term"), but we are interested only in the linear term A. We see we are looking for a linear map

$$A \in \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)$$

with the property that

$$\lim_{v \to 0} \frac{f(x+v) - f(x) - A(v)}{||v||} = 0. \tag{2.1}$$

It is easy to see that, if such a map A exists, then it is unique. And that is what the differential of f at x is.

Definition 2.1. *We say that $f : \Omega \to \mathbb{R}^n$ is* **differentiable at** x *if there exists a linear map $A \in Lin(\mathbb{R}^m, \mathbb{R}^n)$ satisfying (2.1). The linear map A (necessarily unique) is called* **the differential of f at the point** x *(or the total derivative of f at x) and is denoted*

$$(df)_x : \mathbb{R}^m \to \mathbb{R}^n.$$

We say that f is **differentiable** *if it differentiable at all points x in its domain Ω. We say that f is* **of class** C^1 *if it is differentiable and the resulting map*

$$df : \Omega \to Lin(\mathbb{R}^m, \mathbb{R}^n), \quad x \mapsto (df)_x$$

is continuous (recall here that, via the matrix representation of linear maps, $Lin(\mathbb{R}^m, \mathbb{R}^n)$ can be interpreted as the Euclidean space $\mathbb{R}^{m \cdot n}$).

We say that f is of class C^2 if it differentiable and df is of class C^1; proceeding inductively, we can talk about f being of class C^k for any $k \in \mathbb{N}$. We say that f is **smooth** *if it is of class C^k for all k.*

Recall here also the chain rule that allows one to compute the differential of a composition of two functions:

Proposition 2.2 (the chain rule). *Given opens $\Omega \subset \mathbb{R}^m$, $\Omega' \subset \mathbb{R}^n$ and functions*

$$\Omega \xrightarrow{f} \Omega' \xrightarrow{g} \mathbb{R}^p,$$

if f is differentiable at $x \in \Omega$ and g is differentiable at $f(x) \in \Omega'$, then $g \circ f$ is differentiable at x and

$$(d(g \circ f))_x = (dg)_{f(x)} \circ (df)_x.$$

Despite the fact that the differential $(df)_x$ arises as "the linear approximation" of f near x, it contains a great deal of information of f near x- and that makes it extremely useful. Probably the best and most fundamental illustration is the inverse function theorem. Recall here that

Definition 2.3. *A map $f : \Omega \to \Omega'$ between two opens $\Omega \subset \mathbb{R}^n$ and $\Omega' \subset \mathbb{R}^m$ is said to be a* **diffeomorphism** *if it is bijective and both f and f^{-1} are smooth.*

We say that f is a **local diffeomorphism** *around $x \in \Omega$ if there exist opens $\Omega_x \subset \Omega$ and $\Omega'_{f(x)} \subset \Omega'$ with $x \in \Omega_x$, such that $f|_{\Omega_x} : \Omega_x \to \Omega'_{f(x)}$ is a diffeomorphism.*

It is interesting to draw an analogy with Topology (see also the next section), where the main objects are topological spaces, the relevant maps are the continuous ones and two spaces are "isomorphic in Topology" (the precise term being: homeomorphic) if there exists a bijection f between them such that both f as well as f^{-1} are continuous. However, in topology is it usually very hard to prove that two given spaces are not homeomorphic (and one often has to appeal to methods from Algebraic Topology); for instance, just the simple the fact that $\mathbb{R}^n$ and $\mathbb{R}^m$ are homeomorphic only when $n = m$ is very hard to prove. In contrast, the similar statements for diffeomorphisms are much easier to prove thanks to the notion of differential. Indeed, using the chain rule, the following should be a rather easy exercise:

EXERCISE 2.1. Show that if a map $f : \Omega \to \Omega'$ between two opens $\Omega \subset \mathbb{R}^m$ and $\Omega' \subset \mathbb{R}^n$ is a local diffeomorphism around $x \in \Omega$, then

$$(df)_x : \mathbb{R}^m \to \mathbb{R}^n$$

is a linear isomorphism. Deduce that if two opens $\Omega \subset \mathbb{R}^m$ and $\Omega' \subset \mathbb{R}^n$ are diffeomorphic, then $m = n$.

Although this clearly shows the usefulness of the differential, its great power is due to the inverse function theorem (and its immediate consequences, such as the implicit function theorem -see below). Indeed, we see that a condition on the differential of f at a single (given) point x tells us information about f around x:

THEOREM 2.4 (The inverse function theorem). *Given a smooth map $f : \Omega \to \Omega'$ between two opens $\Omega \subset \mathbb{R}^m$ and $\Omega' \subset \mathbb{R}^n$, if f is differentiable at a point $x \in \Omega$ and $(df)_x$ is an isomorphism, then f is a local diffeomorphism around x.*

3. Directional/partial derivatives; the implicit function theorem

Note that, when talking about the differential $(df)_x(v)$ (hence $f : \Omega \to \mathbb{R}^n$, with $\Omega \subset \mathbb{R}^m$ open), $x \in \Omega$ should be thought of as a point, while $v \in \mathbb{R}^n$ as a direction (vector). This becomes more apparent if we reformulate the total derivative as a directional derivative.

DEFINITION 2.5. *With $f : \Omega \to \mathbb{R}^n$ and $x \in \Omega$ as above, and an arbitrary vector $v \in \mathbb{R}^m$,* ***the derivative of f at x in the direction*** *v is the defined as the vector*

$$\partial_v(f)(x) = \frac{\partial f}{\partial v}(x) := \frac{d}{dt}\Big|_{t=0} f(x + tv) = \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t} \in \mathbb{R}^n.$$

When this derivative exists, we say that f is differentiable at x in the v-direction.

The relationship with the total differential is immediate: just replace in (2.1) v (small enough) by tv with $v \in \mathbb{R}^m$ fixed (but arbitrary) and $t \in \mathbb{R}$ approaching 0; using that A is linear, we find that:

$$(df)_x(v) = \frac{\partial f}{\partial v}(x).$$

In particular, if f is differentiable at x then it is differentiable in all directions. The converse is not true; however, one can show that if f is of class C^1 if and only if all the directional derivatives $\frac{\partial f}{\partial v}$ exist and are continuous (see also the discussion below on partial derivatives).

Applying the previous definition to $v \in \{e_1, \ldots, e_m\}$, a vector in the standard basis of $\mathbb{R}^m$, we obtain the partial derivatives

$$\frac{\partial f}{\partial x_i}(x) := \frac{\partial f}{\partial e_i}(x) = \frac{d}{dy}\Big|_{y=x_i} f(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_m) \in \mathbb{R}^n.$$

Its components are the partial derivatives of the components f_i of f:

$$\frac{\partial f}{\partial x_i}(x) = \left(\frac{\partial f_1}{\partial x_j}(x), \ldots, \frac{\partial f_n}{\partial x_j}(x)\right) \quad \text{(where } f = (f_1, \ldots, f_n)\text{)}.$$

These partial derivatives contain the same information as $(df)_x$, just in a less intrinsic way; however, they allow one to handle $(df)_x$ more concretely, via matrices. For that recall that, due to the fact that the standard Euclidean spaces come with a preferred basis, linear maps

$$A : \mathbb{R}^m \to \mathbb{R}^n$$

can be represented as matrices

$$A = (A^i_j)_{1\leq i\leq n, 1\leq j\leq m} = \begin{pmatrix} A^1_1 & \ldots & A^1_m \\ \ldots & \ldots & \ldots \\ A^n_1 & \ldots & A^n_m \end{pmatrix}$$

To make a distinction between the matrix A and the linear map A, one may want to denote by $\hat{A}$ the linear map, at least for a while. Then the relationship between the two is (by definition):

$$\hat{A}(e_j) = \sum_{i=1}^{n} A^i_j e_i$$

or, on a general vector $v = v^1 e_1 + \ldots + v^m e_m \in \mathbb{R}^m$, one has

$$\hat{A}(v) = \sum_{i=1}^{n}(\sum_{j=1}^{m} A^i_j v^j) e_i.$$

To write also this formula in terms of matrix multiplication, we interpret any $v \in \mathbb{R}^m$ as a row matrix and we denote by v^T its transpose (column matrix):

$$v^T = \begin{pmatrix} v^1 \\ \ldots \\ v^m \end{pmatrix}$$

With this, the previous formula becomes

$$\hat{A}(v)^T = A \cdot v^T$$

It follows immediately that the standard multiplication of matrices,

$$(A \cdot B)^i_j = \sum_k A^i_k B^k_j,$$

corresponds to the composition of linear maps:

$$\widehat{AB} = \hat{A} \circ \hat{B}.$$

All together, there should be no confusion in identifying A with $\hat{A}$ even notationally.

In the case of the differential $(df)_x$, to see the matrix representing it we write

$$(df)_x(e_j) = \frac{\partial f}{\partial x_j}(x) = \left(\frac{\partial f_1}{\partial x_j}(x), \ldots, \frac{\partial f_n}{\partial x_j}(x)\right) = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_j}(x) e_j$$

i.e., in the matrix notation,

$$(df)_x = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \ldots & \frac{\partial f_1}{\partial x_m}(x) \\ \ldots & \ldots & \ldots \\ \frac{\partial f_n}{\partial x_1}(x) & \ldots & \frac{\partial f_n}{\partial x_m}(x) \end{pmatrix}.$$

Note that, with this, the fact that $(df)_x$ is an isomorphism is equivalent to the fact that the matrix above is invertible. More generally, the rank of $(df)_x$ as a linear map coincides with the rank as a matrix.

With this:

PROPOSITION 2.6. *A function $f : \Omega \to \mathbb{R}^n$ is of class C^1 if and only if all the partial derivatives $\frac{\partial f}{\partial x_i}$ exist and are continuous functions on Ω.*

In this case we can further look at partials derivatives of order two etc. Hence the higher, order k, partial derivatives are defined inductively:

$$\frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}} = \frac{\partial}{\partial x_{i_1}} \left(\frac{\partial}{\partial x_{i_2}} \left(\ldots \left(\frac{\partial f}{\partial x_{i_k}} \right) \right) \right).$$

The previous proposition that extends to a characterization of f being of class C^l; in particular, f is smooth if and only if all its higher partial derivatives exist. We will denote by

$$\mathcal{C}^\infty(\mathbb{R}^m)$$

the space (algebra!) of smooth functions on $\mathbb{R}^m$. By **smooth partitions of unity** on $\mathbb{R}^m$ (subordinated to an open cover) is any partition of unity whose members η_i are smooth- i.e. Definition 1.10 (finite case) and Definition 1.13 (general case) applied at $\mathcal{A} := \mathcal{C}^\infty(\mathbb{R}^m) \subset \mathcal{C}(\mathbb{R}^m)$.

THEOREM 2.7. *Any open cover of $\mathbb{R}^m$ admits a smooth partition subordinated to it.*

PROOF. We want to use Theorem 1.14 for $\mathcal{A} := \mathcal{C}^\infty(\mathbb{R}^m)$. This is clearly closed under quotients and, for the same reason that locally finite sums of continuous functions is continuous, it is closed under locally finite sums. We still have to check that $\mathcal{A}$ is normal; for that we invoke theorem 1.16 to reduce everything to showing that for any $x \in \mathbb{R}^m$ and any ball centered at x, $B(x, \epsilon)$, there exists a smooth functions $f : \mathbb{R}^m \to [0, 1]$ such that $f(x) > 0$ and f is supported in the ball. It is clear that we may assume that $x = 0$. Also, by rescaling the argument of f (i.e. multiply it by a constant) we may assume that $\epsilon = 1$. Then set $f(x) = g(x_1^2 + \ldots + x_m^2)$ where $g : \mathbb{R} \to [0, 1]$ is any smooth function with $g(0) > 0$ and $g = 0$ outside $[-\frac{1}{2}, \frac{1}{2}]$. That such a function exists should be clear by thinking of its graph. The following exercise provides and explicit formula. □

EXERCISE 2.2. Show that

$$g_0 : \mathbb{R} \to \mathbb{R}, \quad g_0(x) = \begin{cases} e^{-\frac{1}{x}} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

is a smooth function. Then show that

$$g : \mathbb{R} \to \mathbb{R}, \quad g(x) = g_0(x + \frac{1}{2}) g_0(\frac{1}{2} - x).$$

is a smooth function with the properties required at the end of the previous proof.

We now recall the implicit function theorem- which is one the important and rather immediate consequences of the inverse function theorems. The importance is rather geometric, as it arises when looking at curves, surfaces (or higher dimensional ... submanifolds) in $\mathbb{R}^m$. While such subspaces are usually given by equations of type $f(x_1, \ldots, x_m) = 0$ (think e.g. of $x^2 + y^2 = 1$, defining the unit circle in the plane), one would like to express some of the coordinates x_i in terms of the others (or, equivalently, describe our subspace as a graph).

THEOREM 2.8. *Assume that $m = n + d \geq n$, let $f : \Omega \to R^n$ be a smooth map defined on an open $\Omega \subset \mathbb{R}^m$, $p \in \Omega$, and assume that the matrix*

$$\left(\frac{\partial f_i}{\partial x_j}(p) \right)_{1 \leq i,j \leq n}.$$

is non-singular. Then there exists a functions $par_1, \ldots, par_n$ depending on d variables, defined in a neighborhood of $(p_{n+1}, \ldots, p_{n+d})$ such that, for $x = (x_1, \ldots, x_{n+d})$ near p, one has

$$f(x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+d}) = 0 \Longleftrightarrow \begin{cases} x_1 = par_1(x_{n+1}, \ldots, x_{n+d}) \\ \ldots \\ x_n = par_n(x_{n+1}, \ldots, x_{n+d}) \end{cases}$$

REMARK 2.9. Note that this theorem is not completely canonical: it gives preference to the first n components of $f = (f_1, \ldots, f_{n+d})$. E.g. there is an obvious modification in which the starting assumption is that

$$\left(\frac{\partial f_{d+i}}{\partial x_j}(p)\right)_{1\leq i,j\leq n}.$$

is non-singular (what is the conclusion in this case?). And such modifications are necessary even when looking at the simplest examples. E.g., for the unit circle, i.e. for $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + y^2 - 1$, the condition

$$\frac{\partial f}{\partial x}(x, y) \neq 0$$

(necessary for the theorem) is valid at almost all the points (x, y) in the circle (and, indeed, we can always solve $x = \pm\sqrt{1 - y^2}$), except for the points $(0, 1)$ and $(0, -1)$ (and, indeed, there is a problem there: the smoothness of $\pm\sqrt{1 - y^2}$!). However, at those points one can switch the roles of x and y- and, indeed, around those points which are problematic for $\pm\sqrt{1 - y^2}$, one can write $y = \pm\sqrt{1 - x^2}$-smooth in x.

A more intrinsic version of the theorem (and with exactly the same proof) can be obtained by requiring that $(df)_x$ has maximal rank n, without specifying which minor is non-singular. the conclusion will be that there exists a permutation $\sigma \in S_{n+d}$ such that $f(x) = 0$ near a is equivalent to

$$x_{\sigma(1)} = par_1(x_{\sigma(n+1)}, \ldots, x_{\sigma(n+d)}), \ldots, x_{\sigma(n)} = par_n(x_{\sigma(n+1)}, \ldots, x_{\sigma(n+d)}),$$

for $par_1, \ldots, par_n$ as above. But perhaps the most geometric formulation is what is know as the submersion theorem- see below.

PROOF. Consider the map

$$\tilde{f} : \Omega \to \mathbb{R}^{n+d}, \quad \tilde{f}(x_1, \ldots, x_{n+d}) = (f_1(x_1, \ldots, x_{n+d}), \ldots, f_n(x_1, \ldots, x_{n+d}), x_{n+1}, \ldots, x_{n+d}).$$

Then the non-singularity condition in the statement precisely means that $(d\tilde{f})_p$ is non-singular. Hence, by the inverse function theorem, we find a smooth inverse $\tilde{g}$ of $\tilde{f}$, defined near $\tilde{f}(p)$. Given the form of $\tilde{f}$, it follows that $\tilde{g}$ is of a similar form:

$$\tilde{g}(y_1, \ldots, y_{n+d}) = (g_1(y_1, \ldots, y_{n+d}), \ldots, g_n(y_1, \ldots, y_{n+d}), y_{n+1}, \ldots, y_{n+d}).$$

That $\tilde{g} \circ \tilde{f}$ and $\tilde{f} \circ \tilde{g}$ are the identity maps (near p, and $\tilde{f}(p)$, respectively) translates into:

$$g_i(f(x_1, \ldots, x_{n+d}), x_{n+1}, \ldots, x_{n+d}) = x_i, \quad f_i(g(y_1, \ldots, y_{n+d}), y_{n+1}, \ldots, y_{n+d}) = y_i \tag{3.1}$$

for $1 \leq i \leq n$. The first equation shows that:

$$f(x) = 0 \Longrightarrow x_i = g_i(\underbrace{0, \ldots, 0}_{n \text{ times}}, x_{n+1}, \ldots, x_{n+d}), \quad (1 \leq i \leq n)$$

hence there is an obvious candidate for $par_i(x_{n+1}, \ldots, x_{n+d})$: the right hand side of the last equation. The fact that, indeed, $f(par_1(y), \ldots, par_n(y), y) = 0$ for $y = (y_{n+1}, \ldots, y_{n+d}) \in \mathbb{R}^d$ close to $(p_{n+1}, \ldots, p_{n+d})$ is just the second set of equations in (3.1) applied when $y_1 = \ldots = y_n = 0$. □

4. Local coordinates/charts; the immersion/submersion theorems

The standard coordinates in $\mathbb{R}^m$, despite being the obvious ones, are often not the best ones to use in specific problems- and often when dealing with (algebraic or differential) equations or computing integrals, one proceeds to a change of variables (i.e. passing to more convenient coordinates). As a baby example: looking at the curve in $\mathbb{R}^2$ defined by

$$5x^2 + 2xy + 2y^2 = 1,$$

a change of coordinates of type

$$x = \frac{u+v}{3}, \quad \frac{u-2v}{3} \tag{4.1}$$

brings us to the simpler looking equation $u^2 + v^2 = 1$. A very common change of coordinates in $\mathbb{R}^2$ is the passing to polar coordinates:

$$x = r\cos(\theta), \quad y = r\sin(\theta). \tag{4.2}$$

Definition 2.10. *Given $p \in \mathbb{R}^m$, a chart of $\mathbb{R}^m$ around p is a diffeomorphism*

$$\chi = (\chi_1, \ldots, \chi_m) : U \to \Omega \subset \mathbb{R}^m$$

between an open neighborhood $U \subset \mathbb{R}^m$ of p and an open subset $\Omega \subset \mathbb{R}^m$.

The open U is called the domain of the chart and, for $p \in U$, $(\chi_1(p), \ldots, \chi_m(p))$ are called the coordinates of p w.r.t. the chart (U, χ).

For instance the change of coordinates (4.1) is about the chart

$$\chi : \mathbb{R}^2 \to \mathbb{R}^2, \quad \chi(x, y) = (2x + y, x - y) \tag{4.3}$$

while the chart underlying the polar coordinates is

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

In general, given a smooth function

$$f : \mathbb{R}^m \to \mathbb{R}^n$$

and a point $p \in \mathbb{R}^m$, whenever we have two new charts χ and χ' around p and $f(p)$, respectively, one can represent the function f using the new resulting coordinates: **the representation of f w.r.t. the charts χ and χ'** is

$$f_{\chi}^{\chi'} = \chi' \circ f \circ \chi^{-1}.$$

Of course, for the standard charts one obtains back f. If just χ', or just χ, is the standard chart then we use the notations f_{χ} and $f^{\chi'}$, respectively.

For instance, for the function

$$f(x, y) = 5x^2 + 2xy + 2y^2,$$

with respect to the new chart (4.3) one obtains

$$f_{\chi}(u, v) = u^2 + v^2.$$

In general, it is interesting to try to write smooth functions in the simplest possible way, modulo change of coordinates. The simplest types of functions for which this is possible are the most "non-singular" ones. More precisely, given

$$f : U \to \mathbb{R}^n$$

a smooth map defined on an open $U \subset \mathbb{R}^m$ and given $x \in U$, the "non-singular behaviour" that we require is that

$$(df)_p : \mathbb{R}^m \to \mathbb{R}^n$$

has maximal rank. It is interesting to consider the cases $m \geq n$ and $m \leq n$ separately. The first case brings us to the more canonical version of the implicit function theorem:

THEOREM 2.11 (the submersion theorem). *Assume that f is a* ***submersion*** *at a given point $p \in U$ in the sense that $(df)_p : \mathbb{R}^m \to \mathbb{R}^n$ is surjective.*

Then there exists a chart χ of $\mathbb{R}^m$ around p such that, around $\chi(p)$, $f_\chi = f \circ \chi^{-1}$ is given by

$$f_\chi(x_1, \ldots, x_n, x_{n+1}, \ldots, x_m) = (x_1, \ldots, x_n).$$

PROOF. Since the matrix representing $(df)_p$ is of maximal rank, one of its maximal minors (an $n \times n$ matrix) is invertible; we may assume that the invertible minor is precisely the one made of the first n rows (why?)- which is also the hypothesis of the implicit function theorem (Theorem 2.8). Looking at the proof of the theorem, one remarks that the desired chart is $\chi = \tilde{g}$. □

A similar argument gives rise to the following:

THEOREM 2.12 (the immersion theorem). *Assume that f is an immersion at a given point $p \in U$ in the sense that $(df)_p : \mathbb{R}^m \to \mathbb{R}^n$ is injective. Then there exists a chart χ' of $\mathbb{R}^n$ around $f(p)$ such that, in a neighborhood p, $f^{\chi'} = \chi' \circ f$ is given by*

$$f^{\chi'}(x_1, \ldots, x_m) = (x_1, \ldots, x_m, \underbrace{0, \ldots, 0}_{n-m \text{ zeros}}). \tag{4.4}$$

More precisely, denoting $q = f(p)$, there exist:

- *a chart $\chi' : U'_q \to \Omega'_q$ of $\mathbb{R}^n$ around q,*
- *a neighborhood Ω_p of p in $\mathbb{R}^m$, inside the domain of f*

such that (4.4) holds on Ω_p. Furthermore, one may choose χ' and Ω_p so that:

$$f(\Omega_p) = \{u \in U'_q : \chi'_{m+1}(u) = \ldots = \chi'_n(u) = 0\}.$$

PROOF. Let us give a proof that makes reference to $(df)_p$ as a linear map and not as a matrix. Since $(df)_p : \mathbb{R}^m \to \mathbb{R}^n$ is injective, we find a second linear map $B : \mathbb{R}^{n-m} \to \mathbb{R}^n$ such that

$$((df)_p, B) : \mathbb{R}^m \times \mathbb{R}^{n-m} \to \mathbb{R}^n$$

is an isomorphism. Consider then

$$h : U \times \mathbb{R}^{n-m} \to \mathbb{R}^n, \quad h(x_1, x_2) = f(x_1) + B(x_2).$$

We see that h satisfies the hypothesis of the inverse function theorem at the point $(x, 0)$. Hence it is a diffeomorphism around a neighborhood of $(x, 0)$. We denote by

$$\chi' : U' \to \Omega'$$

its inverse. Note that:

(1) U' is an open neighborhood of $f(p)$ in $\mathbb{R}^n$.
(2) Ω' is an open neighborhood of $(p, 0)$ in $\mathbb{R}^n$, contained in $U \times \mathbb{R}^{n-m}$.
(3) the intersection of $\Omega' \subset \mathbb{R}^n$ with $\mathbb{R}^m \times \{0\}$,

$$\Omega := \{u \in \mathbb{R}^m : (u, 0) \in \Omega'\},$$

is an open neighborhood of p included in the domain of f.

Note that, since $\chi'(h(x_1, x_2) = 0$ for all $(x_1, x_2) \in \Omega'$ and $h(x, 0) = f(x)$, we have $\chi^{'}(f(x)) = (x, 0)$ for all $x \in \Omega$. This proves the main part of the theorem; for the last part, note that we have, by the first part, that $f(\Omega)$ is inside the zero set of $\chi_2^{'} : U^{'} \to \mathbb{R}^{n-m}$ (the second component of the chart $\chi^{'}$ w.r.t. the decomposition $\mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^{n-m}$). For the reverse inclusion, let $x \in U'$ with $\chi_2^{'}(x) = 0$; since $h \circ \chi' = \mathrm{id}$ on U', we obtain

$$x = h(\chi'(x)) = h(\chi_1^{'}(x), 0) = f(\chi_1^{'}(x))$$

where, for the last equality, we used the explicit formula for h. Moreover, since $\chi'(x) \in \Omega'$ and since $\chi'(x) = (\chi_1^{'}(x), 0)$, by the definition of Ω, we have $\chi_1^{'}(x) \in \Omega$. With the previous equality in mind, we obtain $x \in f(\Omega)$. □

For later use let us introduce the notion of smoothness defined on arbitrary subsets $M \subset \mathbb{R}^m$.

Definition 2.13. *Given $M \subset \mathbb{R}^m$ and a function $f : M \to \mathbb{R}^n$, we say that f* ***is differentiable at*** *$p \in M$ if there exists an open neighborhood $\tilde{U}$ of p in $\mathbb{R}^m$ and a function*

$$\tilde{f} : \tilde{U} \to \mathbb{R}^n$$

which is differentiable (in the usual sense) at p and such that, one the resulting open neighborhood $U = \tilde{U} \cap M$ of p in M, $\tilde{f}$ coincides with f.
(in short: near $p \in M$, f admits a smooth extension to an open neighborhood of p in $\mathbb{R}^m$).

As before, we say that f is differentiable if it is differentiable at each point $p \in M$; similarly we talk about f being of class C^k or smooth, etc.

Definition 2.14. *A diffeomorphism between $M \subset \mathbb{R}^m$ and $N \subset \mathbb{R}^n$ is any bijection $f : M \to N$ with the property that both f and f^{-1} are smooth.*

And here is a nice application of the existence of smooth partitions of unity.

Exercise 2.3. Show that if $M \subset \mathbb{R}^m$ is a closed subset that any smooth function $f : M \to \mathbb{R}^n$ admits a smooth extension $\tilde{f} : \mathbb{R}^m \to \mathbb{R}^n$.
(Hint: $\mathbb{R}^m \setminus M$ is open; get an open cover of $\mathbb{R}^m$ out of one of M.)

5. Embedded submanifolds of $\mathbb{R}^m$

We now move to the notion of (smooth) embedded submanifolds of $\mathbb{R}^m$. In low dimensions, these are curves (1-dimensional) and surfaces (2-dimensional); for an arbitrary dimension d we will be talking about d-dimensional submanifolds of $\mathbb{R}^m$. For instance, the standard sphere

$$S^d = \{(x_0, \ldots, x_d) \in \mathbb{R}^{d+1} : \sum_i (x_i)^2 = 1\}$$

will be such a smooth d-dimensional submanifold (... of $\mathbb{R}^{d+1}$). Already when looking at the simplest examples one sees that such subspaces may (naturally) be described in several different (but equivalent) ways. E.g., already for the unit circle in the plane, one has the standard descriptions:

- implicit (by equations): $x^2 + y^2 = 1$.
- parametric: $x = \cos(t), y = \sin(t)$ with $t \in \mathbb{R}$.

Due to this, the notion of embedded submanifold of $\mathbb{R}^m$ can be introduced in several equivalent ways. Let us start with a subset

$$M \subset \mathbb{R}^m$$

and an integer d (which will be the dimension of M), and we concentrate around a given point $p \in M$. We try to describe in different ways (e.g. parametric or implicit) our M around p, i.e.

a neighborhood U of p in M. Since M is endowed with the topology induced from $\mathbb{R}^m$, we will be looking at

$$U = M \cap \tilde{U}, \quad \text{with } \tilde{U} - \text{open neighborhood of } p \text{ in } \mathbb{R}^m.$$

We can look for:

(1) **a smooth d-dimensional chart of** M around p- be which we mean a d-dimensional chart (see Definition 1.3),
$$\chi : U \to \Omega \subset \mathbb{R}^d$$
defined on an open neighborhood U of p in M (hence $p \in U \subset M \subset \mathbb{R}^m$) and which is a diffeomorphism (in the sense of Definition 2.14).

(2) **a d-dimensional parametrization of** M **around** x- by which we mean an injective map
$$par : \Omega \to M$$
defined on an open $\Omega \subset \mathbb{R}^d$, which is a homeomorphism into an open neighborhood U of p in M, and which satisfies the regularity condition that, as a map from Ω to $\mathbb{R}^m$, par is an immersion.

(3) **a d-dimensional implicit equation of** M **around** p- by which we mean a submersion
$$eq : U^{'} \to \mathbb{R}^{m-d}$$
defined on an open neighborhood $U^{'}$ if p in $\mathbb{R}^m$ and which describes M near p by the equation $eq = 0$. The last part means that the induced neighborhood $U = M \cap U^{'}$ of p in M coincides with the zero-set of eq:
$$M \cap U^{'} = \{q \in U^{'} : eq(q) = 0\}.$$

The following shows that these points of view are equivalent.

THEOREM 2.15. *For a subset $M \subset \mathbb{R}^m$, $d \in \mathbb{Z}_+$ and $p \in M$, the following are equivalent:*

(1) *M admits a smooth d-dimensional chart around p.*
(2) *M admits a d-dimensional parametrization around p.*
(3) *M can be described around p by a d-dimensional implicit equation.*

REMARK 2.16. Here are two comments that may be useful for following the idea of the proof. The first one is that the relationship between the chart χ and the parametrization par is the obvious one: they are inverse to each other. The second comment is that all these conditions are equivalent to yet another one: $\mathbb{R}^m$ admits a chart around p which takes M (near p) to $\mathbb{R}^d \times \{0\} \subset \mathbb{R}^m$. More precisely:

(4) there exists, around p, **a chart of $\mathbb{R}^m$ adapted to** M, i.e. a chart
$$\chi^{'} : U^{'} \to \Omega^{'} \subset \mathbb{R}^m,$$
which takes $M \cap U^{'}$ into the intersection of $\Omega^{'}$ with the hyperplane $\mathbb{R}^d \times \{0\}$:
$$\chi^{'}(M \cap U^{'}) = \Omega^{'} \cap (\mathbb{R}^d \times \{0\}).$$
(note: in particular, the restriction of the chart $\chi^{'}$ to $M \cap U^{'}$ becomes a chart for M). Equivalently, the induced neighborhood $M \cap U^{'}$ of p in M is characterized by the equations $\chi_i^{'} = 0$ for $i > d$:
$$M \cap U^{'} = \{p' \in U^{'} : \chi_{d+1}^{'}(p') = \ldots = \chi_m^{'}(p') = 0.\}$$

This gives us the following insight: the chart $\tilde{\chi}$ is precisely (χ, eq) (with carefully chosen domains); i.e. the (inverse of the) parametrization and the implicit equation are put together to form the larger chart $\tilde{\chi}$ for $\mathbb{R}^m$.

PROOF. For keeping track of notations note that, throughout the proof, we look around the given point $p \in M \subset \mathbb{R}^m$ and around the corresponding point

$$x = \chi(p) = par^{-1}(p) \in \Omega \subset \mathbb{R}^d.$$

Therefore, we will deal with

- neighborhoods U_p of p in M, and $U_p^{'}$ of p in $\mathbb{R}^m$.
- neighborhoods Ω_x of x in $\mathbb{R}^d$.

The points in the neighborhoods of p will be denoted by p', while the ones in the neighborhoods of x by x'; for them, we may be looking at similar neighborhoods $U_{p'}$, $U_{p'}^{'}$ and $\Omega_{x'}$.

We first prove that (1) implies (2). We start with the the chart $\chi : U \to \Omega \subset \mathbb{R}^d$ defined in a neighborhood U of p in M. Setting $par = \chi^{-1}$ we have to check that par is a homeomorphism -which is clear by construction (it has the continuous χ as inverse)- and that, as a map $\Omega \to \mathbb{R}^m$, it is an immersion. For the last part use that the composition

$$U \xrightarrow{\chi} \Omega \xrightarrow{par} \mathbb{R}^m$$

is the inclusion $U \subset \mathbb{R}^m$ and then apply the chain rule to deduce that, for each point $p' \in U$, $(d\chi)_{par(p')} \circ (d\,par)_{p'}$ is the identity- hence, in particular, $(d\,par)_{p'}$ will be injective.

We now prove that (2) implies both (1) as well as (3). Hence we start with a parametrization $par : \Omega \to U \subset M$; as above, we set $\chi = par^{-1} : U \to \Omega$. To get (2), we still have to check that χ is smooth in the sense of Definition 2.14: i.e., around any point $p' \in U$, it is obtained by restricting a smooth map defined on an open $U_{p'}^{'} \subset \mathbb{R}^m$. For that we use the immersion theorem (Theorem 2.12) applied to $par : \Omega \to \mathbb{R}^m$ around

$$x' = \chi(p') \in \Omega.$$

We find:

- an open neighborhood $\Omega_{x'}$ of x' in $\Omega \subset \mathbb{R}^d$
- a diffeomorphism $\chi^{'} : U_{p'}^{'} \to \Omega_{p'}^{'}$ from an open neighborhood $U_{p'}^{'} \subset \mathbb{R}^m$ of p' to an open $\Omega_{p'}^{'} \subset \mathbb{R}^m$,

so that, on $\Omega_{x'}$, $\chi^{'} \circ par$ becomes the inclusion on the first factors. We now write $\chi^{'} = (\chi_1^{'}, \chi_2^{'})$ where we use again the decomposition $\mathbb{R}^m = \mathbb{R}^d \times \mathbb{R}^{m-d}$. We deduce that $\chi_1^{'}(par(x'')) = x''$ for all $x'' \in \Omega_{x'}$. Since $x'' = \chi(par(x''))$ for all $x'' \in \Omega$, we deduce that $\chi_1^{'}(p'') = \chi(p'')$ for all $p'' \in par(\Omega_{p'})$. In this way, on the neighborhood $par(\Omega_{p'})$ of p' in U, χ is now the restriction of a smooth function defined on an open neighborhood of p' in $\mathbb{R}^m$- namely $\chi_1^{'} : U_{p'}^{'} \to \mathbb{R}^d$.

To prove (3) (still assuming (2)), we use $p' = p$ in the previous reasoning and the resulting diffeomorphism $\chi^{'} : U_p^{'} \to \Omega_p^{'}$; in principle, the desired function f will be $\chi_2^{'}$, but we have to choose the domain of definition carefully. For that we use the last part of Theorem 2.12 which says that we may assume that

$$par(\Omega_x) = \{x' \in U_p^{'} : \chi_{d+1}^{'}(x') = 0, \ldots, \chi_m^{'}(x') = 0\}.$$

Since this is open in M, we can write it as $M \cap W_p$ for some open $W_p \subset \mathbb{R}^m$. Considering now

$$U^{'} := U_p^{'} \cap W_p, \quad eq = \chi_2^{'}|_{U'} : U^{'} \to \mathbb{R}^{m-d}$$

one checks right away that $M \cap U^{'}$ is the zero set of eq (why is eq a submersion?).

EXERCISE 2.4. Conclude now that also condition (4) above holds true.

We are now left with proving that (3) implies (1). Let $eq : U^{'} \to \mathbb{R}^{m-d}$ satisfying the conditions from the hypothesis. Note that if we replace $U^{'}$ by a smaller open neighborhood of p in $\mathbb{R}^m$ (and eq by its restriction), those conditions will still be satisfied. Therefore, using the submersion theorem applied to eq, we may assume that we also find a diffeomorphism $\chi' : U' \to \Omega'$ into an open subset $\mathbb{R}^m$, such that $eq = \chi'_2$. This chart will then take the zero set of eq into the zero set of the second projection $\mathrm{pr}_2 : \Omega' \to \mathbb{R}^{m-d}$, i.e. into

$$\Omega := \{u \in \mathbb{R}^d : (u, 0) \in \Omega'\}.$$

We deduce that the restriction of χ' to $U = M \cap U'$,

$$\chi := \chi'|_U : U \to \Omega \subset \mathbb{R}^d,$$

is a smooth chart of M (around p). □

DEFINITION 2.17. *A d-dimensional **embedded submanifold of** $\mathbb{R}^m$ is any subset $M \subset \mathbb{R}^m$ which, around each $p \in M$, satisfies one of the equivalent properties from the previous theorem.*

EXAMPLE 2.18. Returning to the circle S^1,

- $h(x, y) = x^2 + y^2 - 1$ serves as a (1-dimensional) implicit equation (around any point!)
- $p(t) = (\cos(t), \sin(t))$, when considered on sufficiently small intervals (on which it is injective) serves as paramatrization of S^1 around any point in S^1.
- as smooth (1-dimensional) charts one could use two projections $pr_1, pr_2 : S^1 \to \mathbb{R}$, restricted to the appropriate domains (so that they become homeomorphisms). Another possible choice of charts is given by the stereographic projections (see the lecture notes on Topology).

EXERCISE 2.5. Generalize this discussion to the spheres S^d of arbitrary dimension.

Let us point out the following important corollary of the previous discussion.

COROLLARY 2.19. *Assume that $M \subset \mathbb{R}^m$ can be written as the zero-set of a submersion $f : \Omega \to \mathbb{R}^n$ defined on an open subset $\Omega \subset \mathbb{R}^m$. Then M is an $(m-n)$-dimensional submanifold of $\mathbb{R}^m$.*

EXERCISE 2.6. Denote by $\mathcal{M}_m(\mathbb{R})$ the set of mxm matrices with real coefficients, and let $O(m)$ be the subset consisting of orthonormal matrices

$$O(m) = \{A \in \mathcal{M}_m(\mathbb{R}) : A \cdot A^T = I_m.\}$$

Identifying $\mathcal{M}_m(\mathbb{R})$ with $\mathbb{R}^{m^2}$, show that $O(m)$ is a submanifold of dimension $\frac{m(m-1)}{2}$.

6. From directional derivatives to tangent spaces

The point if view provided by the directional derivatives brings us closer to the intrinsic nature of (p, v) when talking about $(df)_p(v)$: that of tangent vector. The key point is that $(df)_p(v)$ depends only on the behaviour of f near p, in "the direction of v"- and how we realize that "direction" is less important. This is best seen by looking at arbitrary paths through p with the original speed v, i.e. any smooth map

$$\gamma : (-\epsilon, \epsilon) \to \mathbb{R}^m$$

(with $\epsilon > 0$) satisfying

$$\gamma(0) = p, \quad \frac{d\gamma}{dt}(0) = v. \tag{6.1}$$

For instance, one could take $\gamma(t) = p + tv$, but the point is that the variation of $f(p + tv)$ at $t = 0$ does not depend on this specific choice of γ.

LEMMA 2.20. *If f is differentiable at p then, for any path γ satisfying (6.1), one has*

$$(df)_p(v) = \frac{\partial f}{\partial v}(p) = \frac{d}{dt}\Big|_{t=0} f(\gamma(t)).$$

In particular, if f is constant along such a path γ, then $(df)_p(v) = 0$.

More formally, one defines the tangent space of $\mathbb{R}^n$ at x as

$$T_p\mathbb{R}^m = \{(p, v) : v \in \mathbb{R}^m\},$$

and one thinks that $(df)_p(v)$ is defined for any tangent vector $(p, v) \in T_p\mathbb{R}^m$ with p in the domain of f. For curves $\gamma : I \to \mathbb{R}^m$ (smooth, defined on an open interval $I \subset \mathbb{R}$), one interprets the derivative $\frac{d\gamma}{dt}(t)$ as a tangent vector at $\gamma(t)$ to $\mathbb{R}^m$; when we want to emphasize this interpretation, we will use the notation $\dot{\gamma}\,(t)$ for the derivative. Hence, strictly speaking,

$$\dot{\gamma}\,(t) = \left(\gamma(t), \frac{d\gamma}{dt}(t)\right) \in T_{\gamma(t)}\mathbb{R}^m.$$

On such vectors, df measures the variation of f along γ:

$$(df)_{\gamma(t)}\left(\dot{\gamma}\,(t)\right) = \frac{d}{dt}\Big|_{t=0} f(\gamma(t)). \tag{6.2}$$

This point of view becomes extremely useful when looking at more general subspaces

$$M \subset \mathbb{R}^m.$$

DEFINITION 2.21. *Let $M \subset \mathbb{R}^m$ and consider a point $p \in M$.*

A ***vector tangent to*** M ***at*** *p is any vector $v \in \mathbb{R}^m$ which can be realized as the speed at $t = 0$ of a smooth curve $\gamma : (-\epsilon, \epsilon) \to \mathbb{R}^m$ which passes through p at $t = 0$ and which lies entirely in M (i.e. there exists γ satisfying (6.1) and taking values in M).*

The set of such vectors is denoted by T_pM.

When looking at tangent spaces and we want to stress the point $x \in M$ at which we consider the tangent space we will also specify p as a first coordinate, i.e.:

$$T_pM = \{(p, v) : v - \text{is a tangent vector at } p \text{ to } M\}.$$

Although we use the name "tangent *space*", in general, T_pM is just a subset of $\mathbb{R}^m$ (... but it is a vector subspace if M is "nice").

EXERCISE 2.7. Compute T_pM when:

(1) $M \subset \mathbb{R}^2$ is the unit circle and $p = (1, 0)$.
(2) $M \subset R^2$ is the union of the coordinate axes and $p = (0, 0)$.

EXERCISE 2.8. Assume that $M \subset \mathbb{R}^m$ is defined by an equation $f(x) = 0$, where $f : \mathbb{R}^m \to \mathbb{R}$ is a smooth function. For $x \in \mathbb{R}^m$ we denote by $\text{Ker}_x(df)$ the kernel (= the zero set) of the differential $(df)_x : \mathbb{R}^m \to \mathbb{R}$. Show that, in general,

$$T_x(M_f) \subset \text{Ker}_x(df),$$

but the inclusion may be strict. Then prove this inclusion becomes and equality when

$$f = (x_1)^2 + \ldots + (x_n)^2 - 1.$$

EXERCISE 2.9. With the notations from the previous exercise show that for all $x \in M_f$ at which f is a submersion

$$T_x(M_f) = \text{Ker}_x(df).$$

We now return to our discussion on differentials/directional derivatives, recast in terms of tangent spaces. Namely, (6.2) gives us right away:

COROLLARY 2.22. *Given $M \subset \mathbb{R}^m$, $p \in M$ and a function $\tilde{f} : \mathbb{R}^m \subset \mathbb{R}^n$ differentiable at p then, for any vector $v \in \mathbb{R}^m$ tangent to M at p, $(d\tilde{f})_p(v)$ depends only on $\tilde{f}|_M$.*

Of course, this corollary applies to any function $\tilde{f} : U \to \mathbb{R}^n$ defined on an open neighborhood $U \subset \mathbb{R}^m$ of p, with the conclusion that $(d\tilde{f})_p(v)$ only depends on the values of $\tilde{f}|_M$ near p. This shows how to define the differential of a function $f : M \to \mathbb{R}^m$ which is differentiable at $p \in M$ in the sense of Definition 2.13: for $f : M \to \mathbb{R}^n$ that is differentiable at p, one has a well-defined differential

$$(df)_p : T_pM \to \mathbb{R}^n,$$

defined using an extension $\tilde{f}$ of f near p, but independent of the extension.

EXERCISE 2.10. Show that if $M \subset \mathbb{R}^m$ and $N \subset \mathbb{R}^n$ and $f : M \to N$ is smooth at $p \in M$, then $(df)_p$ takes values in $T_{f(p)}N$. Then prove the chain rule in this context and deduce that, if f is a diffeomorphism, then $(df)_p$ is a bijection between T_pM and $T_{f(p)}N$.

Finally, let us look at tangent spaces of submanifolds of $\mathbb{R}^m$.

PROPOSITION 2.23. *If $M \subset \mathbb{R}^m$ is a d-dimensional embedded submanifold then, for any $p \in M$, the tangent space T_pM is a d-dimensional vector subspace of $\mathbb{R}^m$. Furthermore:*

(1) *if $eq : \tilde{U} \to \mathbb{R}^{m-d}$ is an implicit equation defining M around p, then T_pM coincides with the kernel of differential,*

$$(d\,eq)_p : \mathbb{R}^m \to \mathbb{R}^{m-d}.$$

(2) *if $par : \Omega \to M$ is a parametrization of M around p, then T_pM coincides with the image of the differential of par interpreted as a map $par : \Omega \to \mathbb{R}^m$,*

$$(d\,par)_p : T_p\Omega \to \mathbb{R}^m.$$

PROOF. Exercise. □

EXERCISE 2.11. Compute again the tangent spaces of the spheres, but applying now the previous proposition.

7. Exercises given on September 12:

First solve exercise 2.1. Then the following:

EXERCISE 2.12. Consider two smooth functions

$$U \xrightarrow{f} U' \xrightarrow{g} \mathbb{R}^p,$$

defined on opens $U \subset \mathbb{R}^m$, $U' \subset \mathbb{R}^n$. Using the interpretation of linear maps as matrices (as made precise on page 23) show that the chain rule becomes:

$$\frac{\partial g \circ f}{\partial x_i}(x) = \sum_{k=1}^{n} \frac{\partial g}{\partial y_k}(f(p)) \frac{\partial f_k}{\partial x_i}(x)$$

for all $x \in U$ and $1 \leq i \leq m$.

EXERCISE 2.13. Show that for any function $g : R \to \mathbb{R}$, the function

$$\tilde{g} : \mathbb{R}^m \to \mathbb{R}, \quad \tilde{g}(x_1, \ldots, x_m) = g((x_1)^2 + \ldots + (x_m)^2)$$

is not a submersion at $x = 0$.

EXERCISE 2.14. Assume that $f : U_0 \to \mathbb{R}^n$ is a smooth map, $U \subset \mathbb{R}^m$ open, $p \in U$. Let

$$\chi : U \to \Omega \subset \mathbb{R}^m, \quad \chi' : U' \to \Omega' \subset \mathbb{R}^n$$

be charts, of $\mathbb{R}^m$ around p and of $\mathbb{R}^n$ around $f(p)$, respectively. What is the (maximal) domain of definition of $f_{\chi}^{\chi'}$?

EXERCISE 2.15. Consider

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x, y) = \sqrt{x^2 + 2xy + 2y^2}$$

and look around $p = (1, 0)$. Find a chart χ of $\mathbb{R}^2$ around p and a chart χ of $\mathbb{R}$ around $g(p) = 1$ such that, w.r.t. these charts,

$$f_{\chi}^{\chi'}(u, v) = u^2 + v^2.$$

EXERCISE 2.16. Show that

$$g : \mathbb{R} \to \mathbb{R}^2, \quad f(t) = (\cos(t), \sin(t))$$

is an immersion at each point. Then, looking around $t = 0$, find a chart χ' of $\mathbb{R}^2$ around $f(0) = (1, 0)$ such that, w.r.t. this chart,

$$f_{\chi'}(t) = (t, 0).$$

EXERCISE 2.17. Show that if $f : U \to \mathbb{R}^n$, $p \in U$ satisfy the conclusion of the submersion theorem, then f must be a submersion at p. Similarly for the immersion theorem.
(Hint: try it! If it really doesn't work, then look at the next exercise).

EXERCISE 2.18. Assume that $f : U \to \mathbb{R}^n$ is a smooth map, $U \subset \mathbb{R}^m$ open, $p \in U$. Let χ be a chart of $\mathbb{R}^m$ around p and let χ' be a chart of $\mathbb{R}^n$ around $f(p)$. Show that f is a submersion/immersion at p if and only if $f_{\chi}^{\chi'}$ is a submersion/immersion at $\chi(p)$.

EXERCISE 2.19. Consider the stereographic projection w.r.t. the north pole p_N, denoted

$$\chi_N : S^2 \setminus \{p_N\} \to \mathbb{R}^2$$

and similarly the one w.r.t. the south pole, denoted χ_S. Show that

$$\chi_S \circ \chi_N^{-1} : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}^2 \setminus \{0\}$$

is a diffeomorphism.

EXERCISE 2.20. Don't forget to read the proof of theorem 2.12 (short) and do Exercise 2.3.

EXERCISE 2.21. For any $\epsilon > 0$ describe a smooth function

$$f : \mathbb{R}^n \to [0, 1]$$

with the property that $f(0) > 0$ and whose support (in $\mathbb{R}^m$) is contained in the ball $\{x \in \mathbb{R}^n : ||x|| < \epsilon\}$.
(Hint: page 24- but give the precise formulas/details).

EXERCISE 2.22 (HOMEWORK). *Assume that $f : U \to U'$ and $g : U' \to U''$ are two smooth functions, with $U \subset \mathbb{R}^m$, $U' \subset \mathbb{R}^n$ and $U'' \subset \mathbb{R}^p$ opens. Show that if $g \circ f$ is a local diffeomorphism around a given point $x \in U$, then:*

(1) *f is an immersion at x and g is a submersion at $f(x)$.*
(2) *however, it may happen that f is not a submersion at x and g is not an immersion at $g(x)$ (describe an example!).*
(3) *if, furthermore, f is a submersion at x or g is an immersion at $f(x)$, then both f and g are local diffeomorphisms (around x and $f(x)$, respectively).*

CHAPTER 3

Smooth manifolds

1. Charts and smooth atlases

Recall from Definition 1.3 that, given a topological space M, an m-dimensional chart is a homeomorphism χ between an open U in M and an open subset $\chi(U)$ of $\mathbb{R}^m$,

$$\chi : U \to \chi(U) \subset \mathbb{R}^m.$$

Given such a chart, each point $p \in U$ is determined/paramterized by its coordinates w.r.t. χ:

$$(\chi_1(p), \ldots, \chi_m(p)) \in \mathbb{R}^m$$

(a more intuitive notation would be: $(x^1_\chi(p), \ldots, x^m_\chi(p))$).

Given a second chart

$$\chi' : U' \to \chi'(U') \subset \mathbb{R}^m,$$

the map

$$c_{\chi,\chi'} := \chi' \circ \chi^{-1} : \chi(U \cap U') \to \chi'(U \cap U')$$

is a diffeomorphism between two opens in $\mathbb{R}^m$. It will be called **the change of coordinates map** from the chart χ to the chart χ'. The terminology is motivated by the fact that, denoting

$$c_{\chi,\chi'} = (c_1, \ldots, c_m),$$

the coordinates of a point $p \in U \cap U'$ w.r.t. χ' can be expressed in terms of those w.r.t. χ by:

$$\chi'_i(p) = c_i(\chi_1(p), \ldots, \chi_m(p)).$$

Definition 3.1. *We say that two charts (U, χ) and (U', χ') are* ***smoothly compatible*** *if the change of coordinates map $c_{\chi,\chi'}$ (a map between two opens in $\mathbb{R}^d$) is a diffeomorphism.*

Definition 3.2. *A (m-dimensional)* ***smooth atlas*** *on a topological space M is a collection $\mathcal{A}$ of (m-dimensional) charts of M with the following properties:*

- *any point in M belongs to the domain of at least one chart from $\mathcal{A}$.*
- *each two charts from $\mathcal{A}$ are smoothly compatible.*

The charts of an atlas are used to transfer the smoothness properties that make sense on (opens inside) $\mathbb{R}^m$ to M; the compatibility of the charts ensures that the resulting notions (now on M) do not depend on the choice of the charts from the atlas. Hence one may say that an atlas on M allows us to put a "smooth structure" on M. For instance, given a smooth m-dimensional atlas $\mathcal{A}$ on the space M, a function $f : M \to \mathbb{R}$ is called **smooth w.r.t. the atlas $\mathcal{A}$** if for any chart (U, χ) that belongs to $\mathcal{A}$,

$$f \circ \chi^{-1} : \chi(U) \to \mathbb{R}$$

is smooth in the usual sense ($\chi(U)$ is an open in $\mathbb{R}^m$!). We will temporarily denote by

$$\mathcal{C}^\infty(M, \mathcal{A})$$

the set of such smooth functions.

There is one aspect that requires a bit of attention: the fact that two different atlases may give rise to the same "smooth structure". Therefore the following definition.

Definition 3.3. *We say that two smooth atlases $\mathcal{A}_1$ and $\mathcal{A}_2$ are* ***smoothly equivalent*** *if any chart in $\mathcal{A}_1$ is smoothly compatible with any chart in $\mathcal{A}_2$ (or, shorter: if $\mathcal{A}_1 \cup \mathcal{A}_2$ is again a smooth atlas).*

Exercise 3.1. Show that, indeed, for any two atlases $\mathcal{A}_1$ and $\mathcal{A}_2$ on M, one has:

$$\mathcal{C}^\infty(M, \mathcal{A}_1) = \mathcal{C}^\infty(M, \mathcal{A}_2) \Longleftrightarrow \mathcal{A}_1 \text{ is smoothly equivalent to } \mathcal{A}_2.$$

If we want to avoid working with equivalence classes of atlases, one may restrict the attention only to atlases that are "largest".

Definition 3.4. *An atlas $\mathcal{A}$ is called* ***maximal*** *if there is no larger smooth atlas containing it. Given an arbitrary (smooth) atlas $\mathcal{A}$, one defines the associated maximal atlas $\mathcal{A}^{max}$ by adding to $\mathcal{A}$ all the charts that are smoothly compatible with all the charts of $\mathcal{A}$.*

Exercise 3.2. Show that, indeed, $\mathcal{A}^{\max}$ is an atlas, and that it is maximal. Also deduce that

$$\mathcal{C}^\infty(M, \mathcal{A}) = \mathcal{C}^\infty(M, \mathcal{A}^{\max}).$$

Example 3.5. On $\mathbb{R}^m$ one can consider a very small atlas $\mathcal{A}_0$- the one consisting only of the identity map

$$\chi = \mathrm{Id} : U = \mathbb{R}^m \to \mathbb{R}^m.$$

Then $\mathcal{A}_0^{\max}$ consists of all diffeomorphisms between opens in $\mathbb{R}^m$. Another natural atlas would be $\mathcal{A}_1$-consisting of all inclusions of opens $U \subset \mathbb{R}^m$. Or $\mathcal{A}_2$ consisting of all inclusions $B \subset \mathbb{R}^m$ of arbitrary open balls in $\mathbb{R}^m$. Then $\mathcal{A}_0$ and $\mathcal{A}_2$ are disjoint, they are both inside $\mathcal{A}_1$, all three of them are equivalent, and

$$\mathcal{A}_0^{\max} = \mathcal{A}_1^{\max} = \mathcal{A}_2^{\max}.$$

Exercise 3.3. Show that the unit circle S^1 (with the topology induced from $\mathbb{R}^2$) admits an atlas made of two charts, but does not admit an atlas made of a single chart.

The following simple exercise shows that any equivalence class of atlases can be represented by an unique maximal atlas.

Extra-exercise: Show that, for any atlas $\mathcal{A}$, $\mathcal{A}$ is (smoothly) equivalent to $\mathcal{A}^{\max}$; then show that for two atlases $\mathcal{A}_1$ and $\mathcal{A}_2$, one has

$$\mathcal{A}_1 \text{ is smoothly equivalent to } \mathcal{A}_2 \Longleftrightarrow \mathcal{A}_1^{\max} = \mathcal{A}_2^{\max}.$$

2. Smooth structures; manifolds

The main point of the equivalence of atlases is that two atlases that are equivalent give rise to the same notion of smoothness on M. Therefore:

Definition 3.6. *An m-dimensional smooth structure on a topological space M is an equivalence class of of m-dimensional smooth atlases on M.*

The main point of mentioning maximal atlases is the fact that in each equivalence class of an atlas there is a unique maximal one. Therefore one may say that:

a smooth structure on M is a maximal atlas on M.

However, one should keep in mind that very often a smooth structure is exhibited by describing a smaller atlas (ideally "the smallest possible one"). E.g., already for $\mathbb{R}^m$, the simplest choice is the single chart $\mathrm{Id} : \mathbb{R}^m \to \mathbb{R}^m$ (see the previous example). In any case, it is clear that:

any atlas $\mathcal{A}$ on M induces a smooth structure on M;

it may be interpreted as the equivalence class $[\mathcal{A}]$ (with respect to the notion of smooth equivalence of atlases), or as the maximal atlas $\mathcal{A}^{\max}$ associated to $\mathcal{A}$, respectively.

With these:

DEFINITION 3.7. *A* ***smooth*** m***-dimensional manifold*** *is a Hausdorff, second countable topological space M together with am m-dimensional smooth structure on M.*

The charts belonging to the maximal atlas corresponding to the smooth structure are called ***the charts of the smooth manifold*** *M.*

EXERCISE 3.4. In the definition of smooth manifolds show that the condition that M is second countable is equivalent to the fact that the smooth structure on M can be defined by an atlas that is at most countable.

3. Smooth maps

Having introduced the main objects of Differential Geometry, we now move to the maps between them.

DEFINITION 3.8. *Let $f : M \to N$ be a map between two manifold M and N of dimensions m and n, respectively. Given charts (U, χ) and (U', χ') of M and N, respectively, the representation of f with respect to these two charts is defined as*

$$f_{\chi,\chi'} := \chi' \circ f \circ \chi^{-1}.$$

This map makes sense when applied to a point of type $\chi(p) \in \mathbb{R}^d$ with $p \in U$ with the property that $f(p) \in U'$, i.e. $p \in U$ and $p \in f^{-1}(U')$. Therefore, it is a map

$$f_{\chi,\chi'} : \text{Domain}(f_{\chi,\chi'}) \to \mathbb{R}^n.$$

whose domain is the following open subset of $\mathbb{R}^m$:

$$\text{Domain}(f_{\chi,\chi'}) = \chi(U \cap f^{-1}(U')) \subset \mathbb{R}^m.$$

DEFINITION 3.9. *Let M and N be two manifolds and*

$$f : M \to N$$

a map between them. We say that f is ***smooth*** *if its representation $f_{\chi,\chi'}$ with respect to any chart χ of M and χ' of N, is smooth (in the usual sense of Analysis).*

We say that f is a ***diffeomorphism*** *if f is bijective and both f and f^{-1} are smooth.*

EXERCISE 3.5. With the notation from the previous definitions, let $\mathcal{A}_M$ and $\mathcal{A}_N$ be two arbitrary (i.e. not necessarily maximal) atlases inducing the smooth structure on M and N, respectively. Show that, to check that f is smooth, it suffices to check that $f_{\chi,\chi'}$ is smooth for $\chi \in \mathcal{A}_M$ and $\chi' \in \mathcal{A}_N$.

Like in the case of $\mathbb{R}^m$, there are certain types of smooth maps that deserve separate names.

DEFINITION 3.10. *Let $f : M \to N$ be a smooth map between two manifolds. We say that f is an* ***immersion at*** *p if its local representations $f_{\chi,\chi'}$ w.r.t. any chart χ of M around p and χ' of N around $f(p)$ is an immersion at $\chi(p)$ (as a map from an open inside $\mathbb{R}^m$ to one inside $\mathbb{R}^n$- see Theorem 2.12).*

We say that f is an ***immersion*** *if it is an immersion at all points $p \in M$.*

Similarly we define the notion of ***submersion*** *and that of* ***local diffeomorphism****.*

EXERCISE 3.6. Show that, in the previous definition, it is enough to check the required condition for (single!) one chart χ of M around p and one chart χ' of M' around $f(p)$.

Note that, since any open U inside a manifold M has a canonical induced smooth structure made up of the charts of M whose domain is inside U (simple exercise!), the notion of local diffeomorphism at p can also be rephrased without reference to charts as: there exists an open neighborhood U_p of p in M and $U'_{f(p)}$ of $f(p)$ in N such that $f|_{U_p} : U_p \to U'_{f(p)}$ is a diffeomorphism. The rephrasing of the immersion and submersion conditions independently of charts (therefore providing another proof of (actually insight into) the previous exercise) will be possible once we discuss that notion of tangent spaces.

From the submersion and immersion theorems on Euclidean spaces, i.e. Theorem 2.11 and Theorem 2.12 from the previous chapter, we immediately deduce:

THEOREM 3.11 (the submersion theorem). *If $f : M \to N$ is a smooth map between two manifolds which is a submersion at a point $p \in M$, then there exist a charts χ of M around p and χ' of N around $f(p)$ such that $f_{\chi}^{\chi'} = \chi' \circ f \circ \chi^{-1}$ is given by*

$$f_{\chi}^{\chi'}(x_1, \ldots, x_n, x_{n+1}, \ldots, x_m) = (x_1, \ldots, x_n).$$

THEOREM 3.12 (the immersion theorem). *If $f : M \to N$ is a smooth map between two manifolds which is a submersion at a point $p \in M$, then there exist a charts χ of M around p and χ' of N around $f(p)$ such that $f_{\chi}^{\chi'} = \chi' \circ f \circ \chi^{-1}$ is given by*

$$f_{\chi}^{\chi'}(x_1, \ldots, x_m) = (x_1, \ldots, x_m, 0, \ldots, 0).$$

Moreover, the charts $\chi : U \to \Omega \subset \mathbb{R}^m$ and $\chi' : U' \to \Omega' \subset \mathbb{R}^n$ can be chosen so that

$$f(U) = \{q \in U' : \chi'_{m+1}(q) = \ldots = \chi'_n(q) = 0\}.$$

Combining also with the inverse function theorem we deduce:

LEMMA 3.13. *Consider a smooth function $f : M \to N$, $p \in M$. Then the following are equivalent:*

1. *f is a local diffeomorphism around p.*
2. *f is both a submersion as well as an immersion at p.*
3. *f is a submersion at p and $dim(M) = dim(N)$.*
4. *f is an immersion at p and $dim(M) = dim(N)$.*

Moreover, f is a diffeommorphism if and only if it is a local diffeomorphism as well as a bijection.

PROOF. The equivalence between (2), (3) and (4) is immediate from linear algebra: for a linear map $A : V \to W$ between two finite dimensional vector spaces, looking at the conditions: A is injective, A is surjective, $\dim(V) = \dim(W)$, any two implies the third. Since these actually imply that A is an isomorphism, making use also of the inverse function theorem, we obtain the equivalence with (1) as well.

The last part should be clear (otherwise should be treated as an easy exercise!). □

4. Variations

There are several rather obvious variations on the notion of smooth manifold. For instance, keeping in mind that

$$\text{smooth} = \text{of class } \mathcal{C}^\infty,$$

one can consider a $\mathcal{C}^k$-version of the previous definitions for any $1 \leq k \leq \infty$. E.g., instead of talking about smooth compatibility of two charts, one talks about $\mathcal{C}^k$-compatibility, which means that the change of coordinates is a $\mathcal{C}^k$-diffeomorphism. One arises at the notion of **manifold of class $\mathcal{C}^k$**, or **$\mathcal{C}^k$-manifold**. For $k = \infty$ we recover smooth manifolds, while for $k = 0$ we recover topological manifolds.

Yet another possibility is to require "more than smoothness"- e.g analyticity. That gives rise to the notion of **analytic manifold**. Looking at manifolds of dimension $m = 2n$, hence modelled by $\mathbb{R}^m = \mathcal{C}^n$, one can also restrict even further- to maps that are holomorphic. That gives rise to the notion of **complex manifold**. Etc.

Another possible variation is to change the "model space" $\mathbb{R}^m$. The simplest and most standard replacement is by the upper-half planes

$$\mathbb{H}^m = \{(x_1, \dots, x_m) \in \mathbb{R}^m : x_m \geq 0\}.$$

This gives rise to the notion of **smooth manifold with boundary**: what changes in the previous definition is the fact that the charts are homeomorphisms into opens inside $\mathbb{H}^m$.

REMARK 3.14. *Another variation that one sometimes encounters in the literature, usually in order to avoid reference to Topology, is the definition of manifolds as a structure on a set M rather than a topological space M. One problem with such an approach is that its "directness" (avoiding topology) is only apparent, and the various topological notions that inevitably show up over and over again would look rather mysterious and even more complicated. Another problem with such an approach (even more important than the previous one) is the fact that the various manifolds that we encounter are, before anything, topological spaces in a very natural way. Not making any reference to such natural topologies would be artificial and less intuitive.*

However, here is an exercise that indicates how one could (but should not) proceed.

EXERCISE 3.7. Let M be a set and let $\mathcal{A}$ be a collection of bijections

$$\chi : U \to \chi(U) \subset \mathbb{R}^m$$

between subsets $U \subset M$ and opens $\chi(U) \subset \mathbb{R}^m$. We look for topologies on M with the property that each $\chi \in \mathcal{A}$ becomes a homeomorphism. We call them topologies compatible with $\mathcal{A}$.

(i) If the domains of all the $\chi \in \mathcal{A}$ cover M, show that M admits at most one topology compatible with $\mathcal{A}$.

(ii) Assume that, furthermore, for any $\chi, \chi' \in \mathcal{A}$, $\chi' \circ \chi^{-1}$ (defined on $\chi(U \cap U')$, where U is the domain of χ and U' is of χ') is a homeomorphism between opens in $\mathbb{R}^m$. Show that M admits a topology compatible with $\mathcal{A}$.

(Hint: try to define a topology basis).

5. Example: the Euclidean spaces $\mathbb{R}^m$

The **standard smooth structure on the Euclidean space** $\mathbb{R}^m$ is the one defined by the atlas consisting of the (single) chart $\mathrm{Id} : \mathbb{R}^m \to \mathbb{R}^m$.

In general, $\mathbb{R}^m$ has many smooth structures.

EXERCISE 3.8. For a homeomorphism $\chi : \mathbb{R}^m \to \mathbb{R}^m$, consider

$$\mathcal{A}_\chi = \{\chi\},$$

(the atlas on $\mathbb{R}^m$ consisting of one single chart, namely χ itself).

Show that, for a general homeomorphism χ, $\mathcal{A}_\chi$ defines a smooth structure different than the standard one, but $\mathbb{R}^m$ endowed with the resulting smooth structure is diffeomorphic to the standard one.

As the previous exercise shows, the interesting question for $\mathbb{R}^m$, and as a matter of fact for any topological space M, is:

how many non-diffeomorphic smooth structures does a topological space M admit?

Even for $M = \mathbb{R}^m$ this is a highly non-trivial question, despite the fact that the answer is deceivingly simple: the spaces $\mathbb{R}^m$ with $m \neq 4$ admit only one such smooth structure, while $\mathbb{R}^4$

admits an infinite number of them (called "exotic" smooth structures on $\mathbb{R}^4$)!

In the rest of the course, whenever making reference to an Euclidean space $\mathbb{R}^m$ *as a manifold, we always consider it endowed with the standard smooth structure. And similarly for opens* $U \subset \mathbb{R}^m$ (using the atlas consisting of the inclusion $U \subset \mathbb{R}^m$). In particular, this is relevant when talking about smooth functions whose domain or codomain are (opens inside) Euclidean spaces. For instance, when looking at smooth maps

$$\gamma : I \to M$$

defined on an open interval $I \subset M$ we are talking about **(smooth) curves** in M. Here M is a manifold and the smoothness of γ is checked using charts (U, χ) for M and the resulting representation of γ in the chart χ:

$$\gamma^{\chi} := \chi \circ \gamma : I_{\chi} \subset \mathbb{R}^m$$

(with domain $I_{\chi} = \gamma^{-1}(U) \subset I$).

Similarly, given a map

$$f : M \to \mathbb{R}^n$$

defined on a manifold M (and $n \geq 1$ integer), we may talk about its smoothness. Using the minimal atlas of $\mathbb{R}^n$ (and Exercise 3.5!) we see that the smoothness of f is checked by using charts (U, χ) for M and looking at the representation of f w.r.t. f

$$f_{\chi} = f \circ \chi^{-1} : \chi(U) \to \mathbb{R}^n$$

(without making reference to the chart that we use for $\mathbb{R}^n$ which is, of course, $\mathrm{Id} : \mathbb{R}^n \to \mathbb{R}^n$). In particular, for $n = 1$ we can talk about real-valued functions of M, and this gives rise to **the algebra of smooth functions on** M:

$$\mathcal{C}^{\infty}(M) \subset \mathcal{C}(M).$$

Exercise 3.9. Show that, indeed, $\mathcal{C}^{\infty}(M)$ is a sub-algebra of $\mathcal{C}(M)$. Then show that $\mathcal{C}^{\infty}(M)$ is closed under locally finite sums.

For the purpose of obtaining **smooth partitions of unity**, i.e. partitions of unity that are smooth, one is left with showing that $\mathcal{C}^{\infty}(M)$ is normal- and this is where Theorem 1.16, from Chapter I, comes in handy. Actually, the proof of the existence of smooth partitions of unity on $\mathbb{R}^m$ (Theorem 2.7, Chapter I) applies right away to any manifold M; we obtain:

Theorem 3.15. *On any manifold* M *and for any open cover* $\mathcal{U}$*, there exists a smooth partition of unity on* M *subordinated to* $\mathcal{U}$*.*

Corollary 3.16. *Any compact manifold* M *can be embedded in* R^n *for a sufficiently large integer* n*, i.e.* M *is diffeomorphic to an embedded submanifold* M' *of* $\mathbb{R}^n$*.*

6. Example: The spheres S^m

The next example we would like to mention is that of the m-dimensional sphere, realized as the subspace of $\mathbb{R}^{m+1}$ at distance 1 from the origin:

$$S^m = \{(x_0, \ldots, x_m) \in \mathbb{R}^{m+1} : (x_0)^2 + (x_1)^2 + \ldots + (x_m)^2 = 1\}.$$

As with any subspace of an Euclidean space, we endow it with the Euclidean topology. This example will also fit into the general discussion of embedded submanifolds of Euclidean spaces (see below) but it is instructive to consider it separately.

Intuitively it should be clear that, locally, S^m looks like (opens inside) $\mathbb{R}^m$- and that is something that we use everyday (we do live on some sort of sphere, remember?). For instance, drawing small disks on the sphere and projecting them on planes through the origin, one can easily build charts. For instance,

$$U = \{(x_0, \ldots, x_m) \in S^m : x_m > 0\}$$

(not even so small!) and the projection into the horizontal plane $\mathbb{R}^m \times \{0\} \subset \mathbb{R}^{m+1}$ gives rise to

$$\chi : U \to \mathbb{R}^m, \quad \chi(x_0, \ldots, x_m) = (x_0, \ldots, x_m).$$

If you follow your intuition to build such charts, than any charts that you obtain should be smoothly compatible- therefore you will obtain a natural smooth structure on S^m.

Here is a more precise, and more elegant way to do it: via the stereographic projections. This allows one to define the smooth structure made of (only!) two atlases. For that we use the stereographic projections w.r.t. the north and the south poles,

$$p_N = (0, \ldots, 0, 1), \quad p_S = (0, \ldots, 0, -1) \in S^n,$$

respectively. The one w.r.t. to p_N is the map

$$\chi_N : S^m \setminus \{p_N\} \to \mathbb{R}^n$$

which associates to a point $p \in S^n$ the intersection of the line $p_N p$ with the horizontal hyperplane (see Figure 1).

Figure 1.

Computing the intersections, we find the precise formula:

$$\chi_N : S^m \setminus \{p_N\} \to \mathbb{R}^n, \quad \chi_N(p_0, p_1, \ldots, p_m) = \left(\frac{p_0}{1 - p_m}, \ldots, \frac{p_0}{1 - p_m}\right)$$

(while for χ_S we find a similar formula, but with +s instead of the −s). Reversing the process (i.e. computing its inverse), we find

$$\chi_N^{-1} : \mathbb{R}^m \to S^m \setminus \{p_N\}, \quad \chi_N^{-1}(x_1, \ldots, x_n) = \left(\frac{2x_1}{|x|^2 + 1}, \ldots, \frac{2x_m}{|x|^2 + 1}, \frac{|x|^2 - 1}{|x|^2 + 1}\right)$$

and we deduce that χ_N is a homeomorphism. And similarly for χ_S. Computing the change of coordinates between the two charts we find

$$\chi_S \circ \chi_N^{-1} : \mathbb{R}^m \setminus \{0\} \to \mathbb{R}^m \setminus \{0\}, \quad \chi_S \circ \chi_N^{-1}(x) = \frac{x}{|x|^2}.$$

This is clearly smooth, composed with itself is the identity, therefore it is a diffeomorphism (... therefore also solving Exercise 2.19 from Chapter 1). We deduce that the two charts

$$(S^m \setminus \{p_N\}, \chi_N), \quad (S^m \setminus \{p_S\}, \chi_S) \tag{6.1}$$

define a smooth m-dimensional atlas on S^m. The resulting smooth structure is called **the standard smooth structure on S^m.**

EXERCISE 3.10. Show that, if we add to (6.1) the chart (U, χ) described above, we obtain the same smooth structure on S^m. Or, in other words, (U, χ) belongs to the maximal atlas induced by (6.1).

EXERCISE 3.11. Using your intuition, describe a smooth structure on the 2-torus. Then show that your smooth structure can be described by an atlas consisting of two only charts.
(you should think on the picture. If you want to write down explicit formulas, you can use the following concrete realizations of the 2-torus:

$$T^2 = \{(x, y, z) \in \mathbb{R}^3 : (\sqrt{x^2 + y^2} - R)^2 + z^2 = r^2\}, \tag{6.2}$$

with $R > r > 0$).

As for the Euclidean spaces, it becomes very interesting (and exciting!) when we ask about the existence of other smooth structures on S^m (endowed with the Euclidian topology). Of course, for the reasons we explained, we only look at non-diffeomorphic smooth structures. Again, the answers are deceivingly simple (and the proofs higher non-trivial):

- except for the exceptional case $m = 4$ (see below), for $m \leq 6$ the standard smooth structure on S^m is the only one we can find (up to diffeomorphisms).
- S^7 admits precisely 28 (!?!) non-diffeomorphic smooth structures.
- S^8 admits precisely 2.
- ...
- and S^{11} admits 992, while S^{12} only one!
- and S^{31} more than 16 million, while S^{61} only one (and actually, next to the case S^m with $m \leq 6$, S^{61} is the only odd-dimensional sphere that admits only one smooth structure).
- ...

Moreover, some of the exotic (i.e. non-diffeomorphic to the standard) spheres can be described rather simply. For instance, fixing $\epsilon > 0$ small enough,

$$W_k := \{(z_1, z_2, z_3, z_4, z_5) : z_1^{6k-1} + z_2^3 + z_3^2 + z_4^2 + z_5^2\} \subset \mathbb{C}^5 \cong \mathbb{R}^{10}$$

(one for each integer $k \geq 1$) are all homeomorphic to S^7, they are all endowed with smooth structures induced from the standard (!) one on $\mathbb{R}^{10}$, but the first 28 of them are each two non-diffeomorphic (and after that they start repeating). In this way one obtains a rather explicit description of all the 28 smooth structures on S^7.

While the number of smooth structures on S^m is well understood for $m \neq 4$, the case of S^4 remains a mystery: is there just one smooth structure? Is there a finite number of them (and how many?)? Or there is actually an infinite number of them? The smooth Poincare conjecture says that there is only one; however, nowadays it is believed that the conjecture is false.

7. The projective spaces $\mathbb{P}^m$

Probably the simplest example of a smooth manifold that does not sit *naturally* inside an Euclidean space (therefore for which the abstract notion of manifold is even more appropriate) is the m-dimensional projective space $\mathbb{P}^m$. Recall that it consists of all lines through the origin in $\mathbb{R}^{m+1}$:

$$\mathbb{P}^m = \{l \subset \mathbb{R}^{m+1} : l - \text{one dimensional vector subspace}\}.$$

As you have seen in the course on Topology (but we recall below), it comes with a natural topology. This topology can be described in several ways. The bottom line is that each point $x \in \mathbb{R}^{m+1} \setminus \{0\}$ defines a line: the one through x and the origin:

$$l_x = \mathbb{R}x = \{\lambda x : \lambda \in \mathbb{R}\} \subset \mathbb{R}^{m+1};$$

and each line arises in this way. More formally, we have a natural surjective map

$$\pi : \mathbb{R}^{m+1} \setminus \{0\} \to \mathbb{P}^m, \quad x \mapsto l_x,$$

and then we can endow $\mathbb{P}^m$ with the induced quotient topology: the smallest one which makes π into a continuous map. Explicitly, a subset U of $\mathbb{P}^m$ is open if and only if $\pi^{-1}(U)$ is open in $\mathbb{R}^{m+1}$ (think on a picture!).

Noticing that two different points x, y define the same line if and only if $x = \lambda y$ for some $\lambda \in \mathbb{R}^*$, we see that we are in the situation of group actions: the multiplicative group $\mathbb{R}^*$ acts on $\mathbb{R}^{m+1} \setminus \{0\}$ and $\mathbb{P}^m$ is the resulting quotient

$$\mathbb{P}^m = \left(\mathbb{R}^{m+1} \setminus \{0\}\right) / \mathbb{R}^*,$$

endowed with the quotient topology. And one can do a bit better by making use of the fact that each line intersects the unit sphere, i.e. can be written as l_x with $x \in S^m$. The restriction of π defines a surjection

$$\pi_0 : S^m \to \mathbb{P}^m$$

which is continuous (or, if we did not define the smooth structure on $\mathbb{P}^m$ yet, can be used to define it again as a quotient topology). This "improvement" shows in particular that $\mathbb{P}^m$ is compact (as the image of a compact by a continuous map).

Moreover, since $l_x = l_y$ with $x, y \in S^m$ happens only when $y = \pm x$, we see that we deal with an action of $\mathbb{Z}_2$ on S^m and

$$\mathbb{P}^m = S^m / \mathbb{Z}_2.$$

As a quotient of a Hausdorff space modulo an action of a finite group, $\mathbb{P}^m$ is Hausdorff (however, that should be clear to you right away!).

From now on we will use the more standard notation

$$[x_0 : x_1 : \ldots : x_m]$$

to represent that elements of $\mathbb{P}^m$ (hence $[x]$ is the line through x). The notation $[\cdot]$ may remind you that we may think of a line as equivalence classes, while the use of : in the notation should suggest "division". Keep in mind that

$$[x_0 : x_1 : \ldots : x_m] = [\lambda \cdot x_0 : \lambda \cdot x_1 : \ldots : \lambda \cdot x_m]$$

with $\lambda \in \mathbb{R}^*$, are the only type of identities that we have in $\mathbb{P}^m$.

The natural smooth structure is obtained by starting from a simple observation: the last equality allows us in principle to make the first coordinate equal to 1:

$$[x_0 : x_1 : \ldots : x_m] = \left[1 : \frac{x_1}{x_0} : \ldots : \frac{x_m}{x_0}\right],$$

i.e. to use just coordinates from $\mathbb{R}^m$; with one little problem- when $x_0 = 0$ (this "little problem" is what forces us to use more than one chart). We arrive at a very natural chart for $\mathbb{P}^m$: with domain

$$U_0 = \{[x_0 : x_1 : \ldots : x_m] \in \mathbb{P}^m : x_0 \neq 0\}$$

and defined as

$$\chi_0 : U_0 \to \mathbb{R}^m,\ \chi_0([x_0 : x_1 : \ldots : x_m]) = \left(\frac{x_1}{x_0}, \ldots, \frac{x_m}{x_0}\right).$$

And similarly when trying to make the other coordinates equal to 1:

$$\chi_i : U_i \to \mathbb{R}^m,\ \chi_0([x_0 : x_1 : \ldots : x_m]) = \left(\frac{x_0}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_m}{x_i}\right),$$

defined on U_i defined by $x_i \neq 0$.

Definition 3.17. *Show that*

$$(U_0, \chi_0), (U_1, \chi_1), \ldots, (U_m, \chi_m)$$

defines a smooth structure on $\mathbb{P}^m$.

Note that this is also a first example in which, at least intuitively, we need many more than just two charts to form an atlas. The question, valid for any manifold M, asking what is the minimal number of charts one needs to obtain an atlas of M, is a very interesting one. In general one can show that one can always find an atlas consisting of no more than $m + 1$ charts where m is the dimension of M (not very deep, but non-trivial either!). Therefore, denoting by $N_0(M)$ the minimal number of charts that we can find, one always has $N_0(M) \leq \dim(M) + 1$. The atlas described above confirms this inequality for $\mathbb{P}^m$. However, in concrete examples, we can always do better. E.g. we have seen that $N_0(S^m) = 2$ (well, why can't it be 1?). The same holds for all compact orientable surfaces. Thinking a bit (but not too long) about $\mathbb{P}^m$ one may expect that $N_0(\mathbb{P}^m) = m + 1$; however, that is not the case. The precise computation was carried out by M. Hopkins, except for the cases $n = 31$ and $n = 47$. For those cases we know that $N_0(\mathbb{P}^{31})$ is either 3 or 4, while $N_0(\mathbb{P}^{46})$ is either 5 or 6. For all the other cases, writing $m = 2^k a - 1$ with a odd, one has

$$N_0(\mathbb{P}^m) = \begin{cases} \max\{2, a\} & \text{if } k \in \{1, 2, 3\} \\ \text{the least integer} \geq \frac{m+1}{2(k+1)} & \text{otherwise} \end{cases}$$

Returning to the basics, note that there is a complex analogue of $\mathbb{P}^m$; for that reason $\mathbb{P}^m$ is sometimes denoted $\mathbb{RP}^m$ and called the real projective space. The m-dimensional complex projective space $\mathbb{CP}^m$ is defined completely analogously but using complex lines $l_z \subset \mathbb{C}^{m+1}$, i.e. 1-dimensional complex subspaces of $\mathbb{C}^{m+1}$ (and 1-dimensional is in the complex sense). Again, one can write

$$\mathbb{CP}^m = \{[z_0 : z_1 : \ldots : z_m] : (z_0, z_1, \ldots, z_m) \in \mathbb{C}^m \setminus \{0\}$$

with the (only) equalities:

$$[z_0 : x_1 : \ldots : z_m] = [\lambda \cdot z_0 : \lambda \cdot z_1 : \ldots : \lambda \cdot z_m] \quad \text{for } \lambda \in mathbbC^*.$$

Analogously to the real case,

$$\mathbb{CP}^m = \left(\mathbb{C}^{m+1} \setminus \{0\}\right)/\mathbb{C}^*.$$

We can also intersect the (complex) lines with the sphere $S^{2m+1} \subset \mathbb{C}^m$ i.e., in formulas, realize each line as as $[z]$ with $z \in S^{2m+1}$; two like this, $[z_1]$ and $[z_2]$, are equivalent if and only if $z_2 = \lambda \cdot z_1$, where this time $\lambda \in S^1$; i.e. the group $\mathbb{Z}_2$ from the real case is replaced by the group S^1 of complex numbers of norma 1 (endowed with the usual multiplication). We obtain

$$\mathbb{CP}^m = S^{2m}/S^1.$$

As for the smooth structure one proceeds completely analogously, keeping in mind the identification $\mathbb{C}^m = \mathbb{R}^{2m}$: we get a (smooth) atlas made of $m + 1$ charts. Note that the change of coordinates is actually given by holomorphic maps- therefore $\mathbb{CP}^m$ is also a complex manifold (see Section 4). Note that m is the complex dimension of $\mathbb{CP}^m$; as a manifold, it is $2m$-dimensional.

As a curiosity: for the minimal number of charts needed to cover $\mathbb{CP}^m$, the answer is much simpler in the complex case:

$$N_0(\mathbb{CP}^m) = \begin{cases} m+1 & \text{if } m \text{ is even} \\ \frac{m+1}{2} & \text{if } m \text{ is odd} \end{cases}.$$

8. Embedded submanifolds of $\mathbb{R}^m$

First of all, we would like to point out that all embedded submanifolds of $\mathbb{R}^m$ in the sense of Analysis (as discussed in Section 5 of Chapetr II) can naturally be turned into manifolds in the general sense of this chapter. We will change a bit the notation from Chapter II: while there we had d-dimensional submanifolds M or $\mathbb{R}^m$, we will now be talking about n-dimensional submanifolds N of $\mathbb{R}^m$ (hence d and M there now become n and N, respectively).

THEOREM 3.18. *If $N \subset \mathbb{R}^m$ is an n-dimensional embedded submanifold of $\mathbb{R}^m$ then the collection of n-dimensional charts of N (in the sense of (1) from Section 5, Chapter II) define an atlas, and therefore a smooth structure on N.*

The resulting smooth structure will be called the induced smooth structure on N. The proof of the theorem is immediate since the composition and inverses of diffeomorphisms (between subsets of Euclidean spaces) are automatically diffeomorphisms.

EXERCISE 3.12. Show that the induced smooth structure on $S^m \subset \mathbb{R}^{m+1}$ coincides with the one defined by the stereographic projections.

While in Chapter 2 (Section 5) we had four characterization of submanifolds, labelled there as (1)-(4), it was the first one that made the previous theorem immediate. As we pointe out in Chapter 2, condition (4) was the one that contained most information (as it used both the parametrizations as well as the implicit equations); hence there is no wonder that (4) is usually the most useful one. Moreover, it will also be the one that will allow us to proceed more generally and talk about embedded submanifolds N of an arbitrary manifold M.

Let us now check that the the general notion of smoothness from this chapter is compatible with the one from Chapter 2.

PROPOSITION 3.19. *Let $f : M \to N$ be a function between two embedded submanifolds*

$$M \subset \mathbb{R}^{\tilde{m}}. \quad N \subset \mathbb{R}^{\tilde{n}}.$$

Then f is smooth as a function between subsets of Euclidean spaces (Definition 2.13, Chapter 2) if and only if it is smooth as a map between manifolds (Definition 3.9).

PROOF. Let us call by "A-smoothness" the notion from Chapter 2 and smoothness the one from this chapter. It is clear that the two coincide when M and N are open in their ambient Euclidean spaces.

Assume first that f is A-smooth. We will consider charts $\tilde{\chi} : \tilde{U} \to \Omega$ of $\mathbb{R}^{\tilde{m}}$ adapted to M (i.e. satisfying condition (4) in the definition of smoothness from Chapter 2, and similarly $\tilde{\chi}' : \tilde{U}' \to \Omega'$ adapted to N; they induce charts $\chi := \tilde{\chi}|_{U\cap M}$ for M, and similarly χ' for N. To prove that f is smooth around a point $p \in M$ it suffices to show that for any $\tilde{\chi}$ around p and $\tilde{\chi}'$ around $f(p)$, as above, $f_{\chi}^{\chi'}$ is smooth. Since f is A-smooth, we may assume that $f|_U = \tilde{f}|_U$ with $\tilde{f} : \tilde{U} \to \tilde{U}'$ is smooth; in turn, $\tilde{f}$ induces $F := \tilde{f}_{\tilde{\chi}}^{\tilde{\chi}'} : \Omega \to \Omega'$, whose restriction to $\Omega \cap (\mathbb{R}^m \times \{0\} \subset \mathbb{R}^{\tilde{m}}$ is precisely $f_{\chi}^{\chi'}$ (and takes values in $\Omega' \cap (\mathbb{R}^n \times \{0\} \subset \mathbb{R}^{\tilde{n}})$. Hence we are in the situation of having a smooth map $F : \Omega \to \Omega'$ between opens in $\mathbb{R}^{\tilde{m}}$ and $\mathbb{R}^{\tilde{n}}$, taking $\Omega_0 = \Omega \cap (\mathbb{R}^m \times \{0\}$ to $\Omega_0' = \Omega' \cap (\mathbb{R}^n \times \{0\}$ and we want to show that $F|_{\Omega_0} : \Omega_0 \to \Omega_0'$ is smooth as a map between opens in $\mathbb{R}^m$ and $\mathbb{R}^n$; that should be clear.

Assume now that f is smooth. Tho check that it is A-smooth, we fix $p \in M$ and consider charts $\tilde{\chi}$ and $\tilde{\chi}'$ and proceed as above and with the same notation. This time but we do not have $\tilde{f}$, but having it is equivalent to having F extending $F_0 = f_{\chi}^{\chi'}$. Then we are in a similar general situation as above, when we have a smooth map $F_0 : \Omega_0 \to \Omega_0'$ between opens in $\mathbb{R}^m$ and $\mathbb{R}^n$ and we want to extend it to a smooth map F between opens in $\mathbb{R}^{\tilde{m}}$ and $\mathbb{R}^{\tilde{n}}$, defined in a neighborhood of $\chi(p)$. Using the decomposition $\mathbb{R}^{\tilde{m}} = \mathbb{R}^m \times \mathbb{R}^{\tilde{m}-m}$ and similarly for $\mathbb{R}^{\tilde{n}}$, we just set $F(x, y) = (F_0(x), 0)$. □

Note that in most text-books the notion of submersion and immersion is defined using tangent spaces; we will come back to this once tangent spaces will be discussed in full generality. Here we discuss the case of embedded submanifolds of Euclidean spaces, since their tangent spaces have already been discussed. Start with two embedded submanifolds

$$M \subset \mathbb{R}^{\tilde{m}}. \quad N \subset \mathbb{R}^{\tilde{n}}.$$

For $p \in M$ consider the tangent space $T_pM \subset \mathbb{R}^{\tilde{m}}$ as defined in Chapter 2 (Definition 2.21). For a smooth map $f : M \to N$ and $p \in M$ we consider the differential f at p (cf. e.g. the first part of Exercise 2.10):

$$(df)_p : T_pM \to T_{f(p)}N.$$

Exercise 3.13. For a smooth map $f : M \to N$ as above (between embedded submanifolds of Euclidean spaces) and $p \in M$,

- f is a submersion at p if and only if $(df)_p : T_pM \to T_{f(p)}N$ is surjective.
- f is an immersion at p if and only if $(df)_p : T_pM \to T_{f(p)}N$ is injective.

9. Embedded submanifolds and embeddings

As we have already mentioned, the characterization (4) of embedded submanifolds of Euclidean spaces (Section 5 of Chapter 2) allows us to proceed more generally and talk about embedded submanifolds N of an arbitrary manifold M. Indeed, condition (4) has an obvious generalization to this context:

Definition 3.20. *Given an m-dimensional manifold M and a subset $N \subset M$, we say that N is an embedded n-dimenisonal submanifold of M if for any $p \in N$, there exists a chart*

$$\chi : U \to \Omega \subset \mathbb{R}^m$$

for M around p with the property that

$$\chi(U \cap N) = \Omega \cap (\mathbb{R}^n \times \{0\}) \subset \mathbb{R}^m$$

or, more explicitly,

$$\chi(U \cap N) = \{p \in U : \chi_{n+1}(p) = \ldots = \chi_m(p) = 0\}. \tag{9.1}$$

*The charts (U, χ) of M satisfying this equality are called **charts of M adapted to N**.*

Any chart χ as in the definition induces an open $\Omega_N \subset \mathbb{R}^n$ defined by

$$\chi(U \cap N) = \Omega_N \times \{0\},$$

and then a chart

$$\chi_N : U \cap N \to \Omega_N \subset \mathbb{R}^n$$

for N. With these we obtain an induced smooth structure on N.

Exercise 3.14. Check that, indeed, the collection of all charts of type χ_N obtained from charts χ of M that are adapted to N, define a smooth structure on N.

EXERCISE 3.15. Assume that $N \subset M$ is an embedded submanifold and let P be another manifold. Show that:

(1) a function $f : P \to N$ is smooth if and only if it is smooth as an M-valued function.
(2) if a function $f : N \to P$ admits a smooth extension to M, then f is smooth.
(3) if N is closed in M, than the previous statement holds with "if and only if".

(Note: the last point is more difficult because the hint is: use partitions of unity for M (and don't forget that $M \setminus N$ is already open in M).

Here is an extremely very useful method for constructing embedded submanifolds (... whose usefulness is heard to overestimate). In principle, it should work in the explicit examples given by equations. The setting is as follows: we are looking at a smooth map

$$f : M \to N$$

and we are interested in its fiber above a point $q \in N$:

$$F = f^{-1}(q), \quad (\text{with } q \in N).$$

The condition that we need here is that q is a **regular value of** f in the sense that f is a submersion at all points $p \in f^{-1}(q)$.

THEOREM 3.21 (the regular value theorem). *If $q \in N$ is a regular value of a smooth map*

$$f : M \to N,$$

then the fiber above q, $f^{-1}(q)$, is an embedded subamnifold of M of dimension

$$dim(f^{-1}(q)) = dim(M) - dim(N).$$

PROOF. In principle, this is just another face of the submersion theorem. Let $d = m - n$, where m and n are the dimension of M and N, respectively. We check the submanifold condition around an arbitrary point $p \in F = f^{-1}(q)$. For that we apply the submersion theorem to f near p to find charts $\chi : U \to \Omega$ and $\chi' : U' \to \Omega'$ of M around p and of N around $f(p)$, respectively, such that

$$f_{\chi}^{\chi'}(x) = (x_1, \ldots, x_n).$$

After changing χ and χ' by a translation, we may assume that $\chi(p) = 0$ and $\chi'(q) = 0$. We claim that, up to a reindexing of the coordinates, χ is a chart of M adapted to F. Indeed, we have

$$\chi(U \cap F) = \{\chi(p') : p' \in U, f(p') = q\} = \{x \in U : f_{\chi}^{\chi'}(x) = 0\},$$

or, using the form of $f_{\chi}^{\chi'}$,

$$\chi(U \cap F) = \{x = (x_1, \ldots, x_m) \in U : x_1 = \ldots = x_n = 0\},$$

proving that F is an $m - n$-dimensional embedded submanifold of M. □

EXERCISE 3.16. Show that

$$M = \{(x, y, z) \in \mathbb{R}^3 : x^3 + y^3 + z^3 - 3xyz = 1\}$$

is a 2-dimensional embedded submanifold of $\mathbb{R}^3$.

The notion of embedded submanifold allows us to introduce the smooth version of the notion of topological embedding. Recall that a map $f : N \to M$ between two topological spaces is a topological embedding if it is injective and, as a map from N to $f(N)$ (where the second space is now endowed with the topology induced from M), is a homeomorphism. The difference between the topological and the smooth case is that, while subspaces of a topological space inherit a natural induced topology, for smooth structures we need to restrict to embedded submanifolds.

Definition 3.22. *A smooth map $f : N \to M$ between two manifolds N and M is called a* ***smooth embedding*** *if:*

(1) *$f(N)$ is an embedded submanifold of M.*
(2) *as a map from N to $f(N)$, f is a diffeomorphism.*

Here is an useful criterion for smooth embeddings.

Theorem 3.23. *A map $f : N \to M$ between two manifolds is a smooth embedding if and only if it is both a topological embedding as well as an immersion.*

Proof. The direct implication should be clear, so we concentrate on the converse. We first show that $f(N)$ is an embedded submanifold. We check the required condition at an arbitrary point $q = f(p) \in M$, with $p \in N$. For that use the immersion theorem (Theorem 3.12) around p; we consider the resulting charts $\chi : U \to \Omega \subset \mathbb{R}^n$ for N around p and $\chi' : U' \to \Omega' \subset \mathbb{R}^m$ for M around q; in particular,

$$f(U) = \{q \in U^{'} : \chi^{'}_{m+1}(q) = \ldots = \chi^{'}_{n}(q) = 0\}.$$

Since f is a topological embedding, $f(U)$ will be open in $f(N)$, i.e. of type

$$f(U) = f(U) \cap W$$

for some open neighborhood W of $q \in M$. It should be clear now that $U'' := U' \cap W$ and $\chi'' := \chi^{'}|_{U''}$ defines a chart for M adapted to N.

We still have to show that, as a map $f : N \to f(N)$, f is a diffeomorphism (where $f(N)$ is with the smooth structure induced from M). It should be clear that this map continues to be an immersion. Since N and $f(N)$ have the same dimension (an injective maps between vectors spaces of the same dimension are automatically bijective), it follows from Lemma 3.13 that $f : N \to f(N)$ is a diffeomorphism. □

And here is a very useful consequence:

Corollary 3.24. *Let*

$$f : N \to M$$

be a smooth map between two manifolds N and M, with the domain N being compact. Then f is an embedding if and only if it is an injective submersion.

Proof. Use one of the main properties of compact spaces: injective maps from compacts to Hausdorff spaces are automatically topological embeddings! □

In particular, using the same arguments as in the topological case (see Theorem 1.12 from Chapter 1) we find:

Theorem 3.25. *Any (smooth) compact manifold can be smoothly embedded into some Euclidean space.*

Proof. We use the same map as in the proof of Theorem 1.12 from Chapter 1, just that we now make use of smooth partitions of unity (which exist by Theorem 3.15). It suffices to show that the resulting topological embedding

$$i = (\eta_1, \ldots, \eta_k, \tilde{\chi}_1, \ldots, \tilde{\chi}_k) : M \to \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{k \text{ times}} \times \underbrace{\mathbb{R}^d \times \ldots \times \mathbb{R}^d}_{k \text{ times}} = \mathbb{R}^{k(d+1)}.$$

is also an immersion. We check this at an arbitrary point $p \in M$. Since $\sum_i \eta_i = 1$, we find an i such that $\eta_i(p) \neq 0$. We may assume that $i = 1$. In particular, p must be in U_1, and then we can choose the chart χ_1 around p and look at the representation i_{χ_1} with respect to this chart; we will check that it is an immersion at $x := \chi_1(p)$. Hence assume that the differential of i_{χ_1} at

x kills some vector $v \in \mathbb{R}^d$ (and we want to prove that $v = 0$). Then the differential of all the components of i_{χ_1} (taken at x) must kill v. But looking at those components, we remark that:

- the first $\mathbb{R}$-component is $\eta := \eta_1 \circ \chi_1^{-1} : U_1 \to \mathbb{R}$.
- the first $\mathbb{R}^d$ component is the linear map $f : U_1 \to \mathbb{R}^d, f(u) = \eta(u) \cdot u$.

Hence we must have in particular:

$$(d\eta)_x(v) = 0, \quad (df)_x(v) = 0.$$

From the formula of f we see that $(df)_x(v) = (d\eta)_x(v) \cdot x + \eta(x) \cdot v$; hence, using that previous equations we find that $\eta(x) \cdot v = 0$. But $\eta(x) = \eta_1(p)$ was assumed to be non-zero, hence $v = 0$ as desired. □

10. General (immersed) submanifolds

In Topology the leading principle for talking about "a subspace A of a space X" was to make sure that the inclusion $i : A \to X$ was continuous, and we chose "the best possible topology" on A doing that- and that gave rise to the induced (subspace) topology on any subset $A \subset X$. Looking for the analogous notion of "subspace" in the smooth context (i.e. a notion of "submanifold"), the first remark is that "inclusions" in the smooth context are expected to be immersions. Furthermore, starting with a subset N of a manifold M, there are several possible ways to proceed:

(1) try to make the set N into a manifold (in particular into a topological space) such that the inclusion $i : N \hookrightarrow M$ is smooth (immersion).
(2) in the previous point choose "the best possible" manifold structure on N.
(3) try to make the space N, endowed with the topology induced from M, into a manifold such that the inclusion $i : N \hookrightarrow M$ is smooth (immersion)?

The previous section took care of the last possibility- which gave rise to the notion of embedded submanifold. The first possibility (i.e.e item (1) above) gives rise to the notion of immersed submanifold.

DEFINITION 3.26. *Given a manifold M, and* ***immersed submanifold*** *of M is a subset $N \subset M$ together with a structure of smooth manifold on N, such that the inclusion $i : N \to M$ is an immersion.*

We emphasize: an immersed submanifold is not just the subset $N \subset M$, but also the auxiliary data of a smooth structure on N (and, as the next exercise shows, given N there may be several different smooth structures on N that make it into an immersed submanifold). Moreover, the required smooth structure on N induces, in particular, a topology on N; but we insist: this topology does not have to coincide with the one induced from M!

EXERCISE 3.17. Consider the figure eight in the plane $\mathbb{R}^2$. Show that it is not an embedded submanifold of $\mathbb{R}^2$, but it has at least two different smooth structures that make it into an immersed submanifold of $\mathbb{R}^2$.

Immersed sub manifolds may look a bit strange/pathological, but they do arise naturally and one does have to deal with them. In general, any injective immersion

$$f : N \to M$$

gives rise to such an immersed submanifold: $f(N)$ together with the smooth structure obtained by transporting the smooth structure from N via the bijection $f : N \to f(N)$ (i.e. the charts of $f(N)$ are those of type $\chi \circ f^{-1}$ with χ a chart of N). And often immersed sub manifolds do arise naturally in this way. For instance, the two immersed submanifold structure on the figure eight arise from two different injective immersions $f_1, f_2 : \mathbb{R} \to \mathbb{R}^2$ (with the same image: the figure eight), as indicated in the picture.

If N (... together with a smooth structure on it) is an immersed submanifold of M, the immersion theorem tells us that there exist charts (U, χ) of M around arbitrary points $p \in N$ such that $U_N := N \cap \{q \in U : \chi_{n+1}(q) = \ldots = \chi_m(q) = 0\}$ is open in N and $\chi_N = \chi|_{U_N}$ is a chart for N. However, when saying that "U_N is open in N", we are making use of the topology on N induced by the smooth structure we considered on N, and not with respect to the topology induced from M (look again at the previous exercise!). This is the difference with the notion of "embedded submanifold".

Proposition 3.27. *If N is an embedded submanifold of the manifold M, then the set N admits precisely one smooth structure such that the inclusion into N becomes an immersion.*

Thinking about item (2) above, a rather natural class of subsets N of the manifold M to consider is obtained by requiring that the set N admits a unique smooth structure that makes it into an immersed submanifold of M. Such Ns will be called **immersed submanifolds with unique smooth structure**. This is precisely the class of immersed submanifolds which are encoded just in the subset $N \subset M$ and no extra-data. By the previous proposition, embedded submanifolds have this property; however, the converse is not true- for instance the dutch figure eight is not embedded, but does have the uniqueness property (exercise).

However, it turns out that there is a smaller and more interesting class of immersed submanifolds that have the uniqueness property (and still includes the embedded submanifolds, as well as most of the other interesting examples).

Definition 3.28. *An immersed submanifold N of a manifold M is called* ***an initial submanifold*** *if the following condition holds: for any*

- *manifold P,*
- *map $f : P \to N$,*

one has that f is smooth if and only if it is smooth as a map into M.

Proposition 3.29. *For a subset N of a manifold M one has the following:*

embedded submanifold $\Longrightarrow$ initial submanifold $\Longrightarrow$ submanifold with unique smooth structure.

Proof. Consider the inclusion $i : N \hookrightarrow M$. For the first implication the main point is to show that if N is embedded and $f : P \to N$ has the property that $i \circ f : P \to M$ is smooth, then f is smooth. To check that f is smooth, we we use arbitrary charts χ of P and charts χ' of M adapted to N. Note that

$$(i \circ f)_{\chi}^{\chi'} = f_{\chi}^{\chi'_N} : \Omega \to \mathbb{R}^n \cong \mathbb{R}^n \times \{0\} \subset \mathbb{R}^m,$$

for some open $\Omega \subset \mathbb{R}^p$. Hence it suffices to remark that if a function $g : \Omega \to \mathbb{R}^n$ is smooth as a map with values in the larger space $\mathbb{R}^m$, then it is smooth.

For the second part assume that N is initial. In particular, it comes with a smooth structure defined by some maximal atlas $\mathcal{A}$, making the inclusion into M an immersion. To prove the uniqueness property, we assume that we have a second smooth structure with the same property, with associated maximal atlas denoted $\mathcal{A}'$. Applying the initial condition to $P = (N, \mathcal{A}')$ with f being the inclusion into M, we find that the identity map $\mathrm{id} : (N, \mathcal{A}') \to (N, \mathcal{A})$ is smooth. Applying the definition of smoothness, we see that any chart in $\mathcal{A}'$ must be compatible with any chart in $\mathcal{A}$. Therefore $\mathcal{A}' = \mathcal{A}$. □

Using similar arguments, you should try the following:

Exercise 3.18. Let N be a subset of the manifold M. Then:

(i) For any topology on N, the resulting space N can be given at most one smooth structure that makes it into an immersed submanifold of M.
(ii) If we endow N with the quotient topology, then the space N can be given a smooth structure that makes it into an immersed submanifold of M if and only if N is an embedded submanifold of M; moreover, in this case the smooth structure on N is unique.

11. Homework for September 27

In general, for any two manifolds M and N of dimensions M and n, their product has a canonical structure of $(m+n)$-dimensional manifold: if $\chi : U \to \Omega \subset \mathbb{R}^m$ is a smooth chart of M and $\chi' : U' \to \Omega' \subset \mathbb{R}^n$ one of N, then

$$\chi \times \chi' : U \times U' \to \Omega \times \Omega' \subset \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}, \quad (p,q) \mapsto (\chi(p), \chi'(q))$$

will define a smooth chart of $M \times N$ (we do not ask you to prove anything here).

Take now $M = N = S^1$ the unit circle with the standard smooth structure, and we consider the resulting manifold $S^1 \times S^1$. While $S^1 \subset \mathbb{R}^2$, $S^1 \times S^1$ sits naturally inside $\mathbb{R}^2 \times \mathbb{R}^2 = \mathbb{R}^4$. We now want to embed it in $\mathbb{R}^3$, i.e. find a submanifold T of $\mathbb{R}^3$ that is diffeomorphic to $S^1 \times S^1$.

Exercise 3.19. *We define $T \subset \mathbb{R}^3$ as the image of the map:*

$$f : \mathbb{R}^2 \to \mathbb{R}^3, \quad f(\alpha, \beta) = ((2+\cos\alpha)\cos\beta, (2+\cos\alpha)\sin\beta, \sin\alpha).$$

Do the following:

(1) *write T as $g^{-1}(0)$ (the zero-set of g) for some smooth map $g : \mathbb{R}^3 \to \mathbb{R}$ which is a submersion at each point in T. This ensures that T is indeed a submanifold of $\mathbb{R}^3$.*

(2) *show that f is an immersion.*

(3) *using f find a homeomorphism $f_0 : S^1 \times S^1 \to T$ and prove that it is an immersion.*

(4) *then show that $f_0 : S^1 \times S^1 \to T$ is a diffeomorphism.*

(5) *if you haven't done that already: draw a picture of T inside $\mathbb{R}^3$.*

(Hint: try to avoid ugly computations and use what we have already discussed in the class).

12. More examples: classical groups and ... Lie groups

One very interesting (and special) class of manifolds are the so-called Lie groups: they are both groups, as well as manifolds, and the two structures are compatible (see below for the precise definition). The classical examples are groups of matrices (endowed with the usual product of matrices).They sit inside the space of all $n \times n$ matrices (for some n natural number)- which is itself and Euclidean space (just that the variables are arranged in a table rather than in a row):

$$\mathcal{M}_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$$

for matrices with real entries, and

$$\mathcal{M}_n(\mathbb{C}) \cong \mathbb{C}^{n^2} \cong \mathbb{R}^{2n^2}$$

for matrices with complex entries. Since we are interested in groups w.r.t. multiplication of matrices (and in a group one can talk about inverses), we have to look inside the so called **general linear group**:

$$GL_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) : \det(A) \neq 0\},$$

and similarly $GL_n(\mathbb{C})$. Since the determinant is a polynomial function, it is also continuous, hence the groups GL_n are opens inside the Euclidean spaces $\mathcal{M}_n$- therefore they are manifolds in a canonical way (and the usual entries of matrices serves as global coordinates). Inside these there are several interesting groups such as:

- **the orthogonal group**:
$$O(n) = \{A \in GL_n(\mathbb{R}) : A \cdot A^T = I\}.$$
- **the special orthogonal group**:
$$SO(n) = \{A \in O(n) : \det(A) = 1\}.$$
- **the special linear group**:
$$SL_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) : \det(A) = 1\}.$$
- **the unitary group**:
$$U(n) = \{A \in GL_n(\mathbb{C}) : A \cdot A^* = I\}.$$
- **the special unitary group**:
$$SU(n) = \{A \in GL_n(\mathbb{C}) : A \cdot A^* = I, \det(A) = 1\}.$$
- **the symplectic group**:
$$Sp_n(\mathbb{R}) := \{A \in GL_{2n}(\mathbb{R}) : A^T J A = J\},$$
where

$$J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} \in GL_{2n}(\mathbb{R}). \tag{12.1}$$

Recall that A^T denotes the transpose of A, A^* the conjugate transpose.

We would like to point out that they are carry a natural smooth structure (making them into Lie groups); actually, they are all embedded submanifolds of the corresponding (Euclidean) spaces of matrices. Since they are all given by (algebraic) equations, there is a natural to proceed: use the regular value theorem (Theorem 3.21). As an illustration, let us look at the case of $O(n)$.

EXAMPLE 3.30 (the details in the case of $O(n)$). The equation defining $O(n)$ suggest that we should be using the function

$$f : GL_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R}), \quad f(A) = A \cdot A^T.$$

However, it is important to note that the values of this function takes values in a smaller Euclidean space. Indeed, since in general, $(A \cdot B)^T = B^T \cdot A^T$, the matrices of type $A \cdot A^T$ are always symmetric. Therefore, denoting by $\mathcal{S}_n$ the space of symmetric $n \times n$ matrices (again an Euclidean space, of dimension $\frac{n(n+1)}{2}$), we will deal with f as a map

$$f : GL_n(\mathbb{R}) \to \mathcal{S}_n.$$

(If we didn't remark that f was taking values in $\mathcal{S}_n$, we would not have been able to apply the regular value theorem; however, the problem that we would have encountered would clearly indicate that we have to return and make use of $\mathcal{S}_n$ from the beginning; so, after all, there is no mystery here).

Now, f is a map from from an open in an Euclidean space (and we could even use f defined on the entire $\mathcal{M}_n(\mathbb{R})$), with values in another Euclidean space; and it is clearly smooth since it is given by polynomial expressions. So, at each point $A \in GL_n(\mathbb{R})$, the differential of f at A,

$$(df)_A : \mathcal{M}_n(\mathbb{R}) \to \mathcal{S}_n$$

can be computed by the usual formula:

$$(df)_A(X) = \frac{d}{dt}\Big|_{t=0} f(A+tX) = \frac{d}{dt}\Big|_{t=0} \left(A \cdot A^T + t(A \cdot X^T + A^T \cdot X) + t^2 X \cdot X^T\right) = A \cdot X^T + X \cdot A^T.$$

Since $O(n) = f^{-1}(\{I\})$, we have to show that $(df)_A$ is surjective for each $A \in O(n)$. I.e., for $Y \in \mathcal{S}_n$, show that the equation $A \cdot X^T + X \cdot A^T = Y$ has a solution $X \in \mathcal{M}_n(\mathbb{R})$. Well, it does, namely: $X = \frac{1}{2}YA$ (how did we find it?). Therefore $O(n)$ is a smooth sub manifold of GL_n of dimension $n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$.

Note that it also follows that the tangent space of $O(n)$ at the identity,

$$T_I O(n) \subset T_I GL_n(\mathbb{R}) = \mathcal{M}_n(\mathbb{R}),$$

coincides with the kernel of $(df)_I$, i.e. it is the space of antisymmetric matrices, usually denoted:

$$o(n) := \{X \in \mathcal{M}_n(\mathbb{R}) : X + X^T = 0\}.$$

The tangent space of such groups, taken at the identity matrix, play a very special role and are known under the name of *Lie algebra of the group* (well, the name also refers to some extra-structure they inherit- but we will come back to that later, after we discuss tangent vectors and vector fields). For the other groups one proceeds similarly and we find that they are all embedded submanifolds of GL_ns, with:

- $SO(n)$: of dimension $\frac{n(n-1)}{2}$, with
$$T_I SO(n) = o(n) = \{X \in \mathcal{M}_n(\mathbb{R}) : X + X^T = 0\}.$$
- $SL_n(\mathbb{R})$: of dimension $n^2 - 1$, with
$$T_I SL_n(\mathbb{R}) = sl_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) : Tr(X) = 0\}.$$
- $U(n)$: of dimension n^2, with
$$T_I U(n) = u(n) = \{X \in \mathcal{M}_n(\mathbb{C}) : X + X^* = 0\}.$$
- $SU(n)$: of dimension $n^2 - 1$, with
$$T_I SU(n) = su(n) = \{X \in \mathcal{M}_n(\mathbb{C}) : X + X^* = 0, Tr(X) = 0\}.$$

EXERCISE 3.20. Do the same for the other groups in the list. At least for one more.

EXERCISE 3.21. The aim of this exercise is to show how $GL_n(\mathbb{C})$ sits inside $GL_{2n}(\mathbb{R})$, and similarly for the other complex groups. Well, there is an obvious map

$$j : GL_n(\mathbb{C}) \to GL_{2n}(\mathbb{R}), \quad A + iB \mapsto \begin{pmatrix} A & B \\ -B & A \end{pmatrix}.$$

Show that:

(1) j is an embedding, with image $\{X \in GL_{2n}(\mathbb{R}) : J \cdot X = X \cdot J\}$, where J is the matrix (12.1).
(2) j takes $U(n)$ into $O(n)$.
(3) actually j takes $U(n)$ into $SO(n)$.

(Hints:

- for (2): first check that $j(Z^*) = j(Z)^T$ for all $Z \in GL_n(\mathbb{C})$.
- for (3): First prove that $\det(j(Z)) = |\det(Z)|^2$ for all $Z \in GL_n(\mathbb{C})$. For this: show that there exists a matrix $X \in GL_n(\mathbb{R})$ such that

$$\begin{pmatrix} A & B \\ -B & A \end{pmatrix} = \begin{pmatrix} X & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & X^{-1} \\ -I & X \end{pmatrix} \begin{pmatrix} B & 0 \\ 0 & B \end{pmatrix})$$

Instead of handling each of the groups separately, there is a much stronger result that ensures that such groups are smooth manifolds:

THEOREM 3.31. *Any closed subgroup of $GL_n(\mathbb{R})$ or $GL_n(\mathbb{C})$ is automatically an embedded submanifold.*

Of course, this theorem is very useful and one should at least be aware of it. The proof is not tremendously difficult, but it would require one entire lecture- and that we cannot afford. But, in principle, the idea is simple: starting with $G \subset GL_n(\mathbb{R})$ closed subgroups, one looks at $\mathfrak{g} := T_I G \subset \mathcal{M}_n$, one shows that it is a vector subspace, and one uses the exponential of matrices ($\exp : \mathcal{M}_n \to GL_n$) restricted to $\mathfrak{g}$ to produce a chart for G around $\exp(0) = I$ (the identity matrix); for charts around other points, one uses the group structure to translate the chart around I to a chart around any point in G.

Finally, since we have mentioned "Lie groups", here is the precise definition:

DEFINITION 3.32. *A Lie group G is a group which is also a manifold, such that the two structures are compatible in the sense that the multiplication and the inversion operations,*

$$m : G \times G \to G, m(g, h) = gh, \quad \iota : G \to G, \iota(g) = g^{-1},$$

are smooth.

An isomorphism between two Lie groups G and H is any group homomorphism $f : G \to H$ which is a diffeomorphism.

Of course, all the groups that we mentioned are Lie groups (just think e.g. that the multiplication is given by a polynomial formula!); and, the conclusion of the last theorem is that all closed subgroups are actually Lie groups.

EXAMPLE 3.33. The unit circle S^1, identified with the space of complex numbers of norm one,

$$S^1 = \{z \in \mathbb{C} : |z| = 1\},$$

is a Lie group with respect to the usual multiplication of complex numbers. Actually, we see that

$$S^1 = SU(1).$$

Also, looking at $SO(2)$, we see that it consists of matrices of type

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.$$

and we obtain an obvious diffeomorphism (an isomorphism of Lie groups) $SO(2) \cong S^1$.

EXAMPLE 3.34. Similarly, the 3-sphere S^3 can be made into a (non-commutative, this time) Lie group. For that we replace $\mathbb{C}$ by the space of quaternions:

$$\mathbb{H} = \{x + iy + jz + kt : x, y, z, t \in \mathbb{R}\}$$

where we recall that the product in $\mathbb{H}$ is uniquely determined by the fact that it is $\mathbb{R}$-bilinear and $i^2 = j^2 = k^2 = -1, ij = k, jk = i, ki = j$. Recall also that for

$$u = x + iy + jz + kt \in \mathbb{H}$$

one defines

$$u^* = x - iy - jz - kt \in \mathbb{H}, \quad |u| = \sqrt{uu^*} = \sqrt{x^2 + y^2 + z^2 + t^2} \in \mathbb{R}.$$

Then, the basic property $|u \cdot v| = |u| \cdot |v|$ still holds and we see that, identifying S^3 with the space of quaternionic numbers of norm 1, S^3 becomes a Lie group.

EXERCISE 3.22. Show that

$$F : S^3 \to SU(2), \quad F(\alpha, \beta) = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}$$

where we interpret S^3 as $\{(\alpha, \beta) \in \mathbb{C}^2 : |\alpha|^2 + |\beta|^2 = 1\}$, is an isomorphism of Lie groups.

By analogy with the previous example, one may expect that S^3 is isomorphic also to $SO(3)$ (at least the dimensions match!). However, the relation between these two is more subtle:

EXERCISE 3.23. Recall that we view S^3 inside $\mathbb{H}$. We also identify $\mathbb{R}^3$ with the space of pure quaternions

$$\mathbb{R}^3 \xrightarrow{\sim} \{v \in \mathbb{H} : v + v^* = 0\}, \ (a, b, c) \mapsto ai + bj + ck.$$

For each $u \in S^3$, show that

$$A_u(v) := u^* v u$$

defines a linear map $A_u : \mathbb{R}^3 \to \mathbb{R}^3$ which, as a matrix, gives an element

$$A_u \in SO(3).$$

Then show that the resulting map

$$\phi : S^3 \to SO(3), \ u \mapsto A_u$$

is smooth, is a group homomorphism, is a surjective local diffeomorphism, but each fiber has two elements (it is a 2-1 cover).

Then deduce that $SO(3)$ is diffeomorphic to the real projective space $\mathbb{P}^3$.

REMARK 3.35. One may wonder: which spheres can be made into Lie groups? Well, it turns out that S^0, S^1 and S^3 are the only ones!

On the other hand, it is interesting to understand what happens with the arguments we used for S^1 and S^3 (and which were very similar to each other) in higher dimensions. The main point there was the multiplication on $\mathbb{C}$ and $\mathbb{H}$ and the presence of a norm such that

$$|x \cdot y| = |x| \cdot |y| \tag{12.2}$$

(so that, for two elements in the sphere, i.e. of norm one, their product is again in the sphere). So, to handle S^n similarly, we would need a "normed division algebra" structure on $\mathbb{R}^{n+1}$, by which we mean a multiplication "$\cdot$" on $\mathbb{R}^{n+1}$ that is bilinear and a norm satisfying the previous condition. Again, it is only on $\mathbb{R}$, $\mathbb{R}^2$ and $\mathbb{R}^4$ that such a multiplication exists. If we do not insist on the associativity of the multiplication, there is one more possibility: $\mathbb{R}^8$ (the so called octonions). But nothing else! And this was known since the 19th century! But why did people care about such operations in the 19th century? Well ... it was number theory and the question of which numbers (integers) can be written as a sum of two, three, etc squares. For sum of two squares the central formula which shows that a product of two numbers that can be written as a sum of two squares can itself be written as a sum of two squares is:

$$(x^2 + y^2)(a^2 + b^2) = (xa - yb)^2 + (xb + ya)^2.$$

Or, in terms of the complex numbers $z_1 = x + iy, z_2 = a_i b$, the norm equation (12.2). The search for similar "magic formulas" for sum of three squares never worked, but it did for four:

$$(x^2 + y^2 + z^2 + t^2)(a^2 + b^2 + c^2 + d^2) =$$
$$(xa + yb + zc + td)^2 + (xb - ya - zd + tc)^2 +$$
$$+(xc + yd - za - tb)^2 + (xd - yc + zb - ta)^2.$$

This is governed by the quaternions and its norm equation (12.2).

CHAPTER 4

Tangent spaces

So far, the fact that we dealt with general manifolds M and not just embedded submanifolds of Euclidean spaces did not pose any serious problem: we could make sense of everything right away, by moving (via charts) to opens inside Euclidean spaces.

However, when trying to make sense of "tangent vectors", the situation is quite different. Indeed, already intuitively, when trying to draw a "tangent space" to a manifold M, we are tempted to look at M as part of a bigger (Euclidean) space and draw planes that are "tangent" (in the intuitive sense) to M. The fact that *the tangent spaces of M are intrinsic to M and not to the way it sits inside a bigger ambient space* is remarkable and may seem counterintuitive at first. And it shows that the notion of tangent vector is indeed of a geometric, intrinsic nature. Given our preconception (due to our intuition), the actual general definition may look a bit abstract. For that reason, we describe several different (but equivalent) approaches to tangent spaces. Each one of them starts with one intuitive perception of tangent vectors on $\mathbb{R}^m$, and then proceeds by concentrating on the main properties (to be taken as axioms) of the intuitive perception.

However, before proceeding, it is useful clearly state what we expect:

- for M-manifold, $p \in M$, have a vector space T_pM (the tangent space of M at p).
- for smooths map $F : M \to N$ (between manifolds), $p \in M$, have an induced linear map
$$(dF)_p : T_pM \to T_{F(p)}N,$$
called the differential of F at p. Of course, when $F = \mathrm{Id} : M \to M$ is the identity map, $(dF)_p$ should be the identity map of T_pM.
- The chain rule: for $M \stackrel{F}{\to} N \stackrel{G}{\to} P$ smooth maps between manifold and $p \in M$,
$$(d\,G \circ F)_p = (dG)_{F(p)} \circ (dF)_p.$$

And, of course, for embedded submanifolds $M \subset \mathbb{R}^{\tilde{m}}$, these tangent spaces should be (canonically) isomorphic to the previously defined tangent spaces T_pM. Also, as for opens inside Euclidean spaces, we expect that for any open $U \subset M$ and $p \in U$, one has $T_pU = T_pM$.

Exercise 4.1. Show that, whatever construction of the tangent spaces T_pM we give so that it satisfies the previous properties, for any diffeomorphism $F : M \to N$ and any $p \in M$, the differential of F at p,
$$(dF)_p : T_pM \to T_{F(p)}N,$$
is a linear isomorphism. Deduce that for any m-dimensional manifold M, T_pM is an m-dimensional vector space.

However, even before these properties, one should keep in mind that the way we should think intuitively about tangent spaces is:

T_pM is made of speeds (at $t = 0$) of curves γ in M passing through p at $t = 0$.

This is clear and precise in the case when $M \subset \mathbb{R}^{\tilde{m}}$, where the embedding was used to make sense of "speeds". For conciseness, we introduce the notation

$$\textbf{Curves}_p(M) := \{\gamma : (-\epsilon, \epsilon) \to M \textbf{ smooth, with } \epsilon > 0, \gamma(0) = p\}$$

(valid for all manifold M and $p \in M$). Hence, in general, we still want that

$$T_pM = \{\text{"}\frac{d\gamma}{dt}(0)\text{"} : \gamma \in \text{Curves}_p(M)\}, \tag{0.3}$$

the only problem being that "$\frac{d\gamma}{dt}(0)$" does not make sense yet.

One way to proceed is by "brute force", starting from the following remark: although "the speed of γ" (at $t = 0$) is not defined yet, we can make sense right away of the property that two such curves $\gamma_1, \gamma_2 \in \text{Curves}_p(M)$ have the same speed at $t = 0$: if their representations γ_1^χ and γ_2^χ with respect to a/any chart χ around p have this property:

$$\frac{d\gamma_1^\chi}{dt}(0) = \frac{d\gamma_2^\chi}{dt}(0).$$

(recall that, for $\chi : U \to \Omega \subset \mathbb{R}^m$, $\gamma_i^\chi = \chi \circ \gamma_i$ is a curve in $\Omega \subset \mathbb{R}^m$). When this happens, we write

$$\gamma_1 \sim_p \gamma_2.$$

Exercise 4.2. Explain the "a/any" in the discussion above and show that $\sim_p$ is an equivalence relation.

The "brute force" approach to tangent spaces would then be to define:

- T_pM as the quotient of $\text{Curves}_p(M)$ modulo the equivalence relation $\sim_p$.
- the speed $\frac{d\gamma}{dt}(0)$ as the $\sim_p$ equivalence class of γ.

Then equality (0.3) that corresponds to our intuition is forced in tautologically. Although this does work, and it is sometimes taken as definition in some text-books/lecture notes, we find it too abstract for such an intuitive concept. Instead, we will discuss the tangent space via charts, as well as via derivations. Nevertheless, one should always keep in mind the intuition via speeds of paths (and make it precise in whatever model we use).

1. Tangent vectors via charts

The brief philosophy of this approach is:

while a chart χ for M allows one to represent points p by coordinates, it will allow us to represent tangent vectors $v \in T_pM$ by vectors in $\mathbb{R}^m$

The actual definition can be discovered as the outcome of "wishful thinking": we want to extend the definition of T_pM from the case of embedded submanifolds of Euclidean spaces to arbitrary manifolds M, so that the properties that we have seen inside Euclidean spaces continue to hold (such as the chain rule- see the general comments itemized above). We see that, for a general manifold M and $p \in M$, choosing a chart (U, χ) for M around p and viewing χ as a smooth map from (an open inside) M to (an open inside) $\mathbb{R}^m$, the wishful thinking tells us that we expect an induced map (even an isomorphism)

$$(d\chi)_p : T_pM \to \mathbb{R}^m.$$

Therefore, an element $v \in T_pM$ can be represented w.r.t. to the chart χ by a vector in the standard Euclidean space:

$$v^\chi := (d\chi)_p(v) \in \mathbb{R}^m.$$

What happens if we change the chart χ by another chart χ' around p? Then $v^{\chi'}$ should be changed in a way that is dictated by the change of coordinates

$$c = c_{\chi,\chi'} = \chi' \circ \chi^{-1}.$$

Indeed, since $\chi' = c \circ \chi$ and we want the chain rule still to hold, we would have:

$$v^{\chi'} = (d\chi')_p(v) = (d\,c)_{\chi(p)}\left((d\chi)_p(v)\right),$$

i.e.

$$v^{\chi'} = (d\,c)_{\chi(p)}(v^{\chi}). \tag{1.1}$$

Therefore, whatever the meaning of T_pM is, we expect that the elements $v \in T_pM$ can be represented w.r.t local charts χ (around p) by vectors $v^{\chi} \in \mathbb{R}^m$ which, when we change the chart, transform according to (1.1). Well, if that is what we want from T_pM, let us just define it that way!

Definition 4.1. *Given an m-dimensional manifold M and $p \in M$, a* ***tangent vector of*** *M* ***at*** *p is any function*

$$v : \{\textit{charts of } M \textit{ around } p\} \to \mathbb{R}^m, \quad \chi \mapsto v^{\chi},$$

with the property that, for any two charts χ and χ' one has

$$v^{\chi'} = (d\,c)_{\chi(p)}(v^{\chi}), \tag{1.2}$$

where $c = c_{\chi,\chi'}$ is the change of coordinates from χ to χ'.

We denote by T_pM the vector space of all such tangent vectors of M at p (a vector space using the vector space structure on $\mathbb{R}^m$, i.e. $(v+w)^{\chi} := v^{\chi} + w^{\chi}$, etc).

Let us explain, right away, that this model for the tangent space serves the original purpose: we can make sense of

$$\frac{d\gamma}{dt}(0) \in T_pM$$

for $\gamma \in \mathrm{Curves}_p(M)$. The actual definition should be clear: the representation of this vector with respect to a chart χ should be the speed (at $t = 0$) of the representation γ^{χ} of γ with respect to χ. In more detail: if $\chi : U \to \Omega$ is a chart of M around p, consider the resulting curve $\gamma^{\chi} = \chi \circ \gamma$ in $\Omega \subset \mathbb{R}^m$, and take its derivatives at 0- which is a vector in $\mathbb{R}^m$. This gives an operation

$$\{\text{charts of } M \text{ around } p\} \to \mathbb{R}^m, \quad \chi \mapsto \frac{d\gamma^{\chi}}{dt}(0),$$

which is clearly linear and satisfies the Leibniz identity; therefore we obtain a tangent vector denoted

$$\frac{d\gamma}{dt}(0) \in T_pM.$$

In a short formula,

$$\frac{d\gamma}{dt}(0)^{\chi} := \frac{d\gamma^{\chi}}{dt}(0).$$

Of course, there is nothing special for $t = 0$, and in the same way one talks about

$$\frac{d\gamma}{dt}(t) \in T_{\gamma(t)}M.$$

for all t in the domain of γ.

Staring for a minute at the definition of tangent vectors $v \in T_pM$, including (1.2), we see that once we know v^{χ_0} for one single chart around p, we completely know v. More precisely:

LEMMA 4.2. *If χ_0 is a chart of M around p, then*

$$T_pM \to \mathbb{R}^m, \quad v \mapsto v^{\chi_0}$$

is an isomorphism of vector spaces.

PROOF. As we have already remarked, given $v_0 \in \mathbb{R}^m$ then the $v \in T_pM$ such that $v^{\chi_0} = v_0$ is unique: it must be given by the formula

$$v^{\chi} = (d\, c_{\chi_0,\chi})_{\chi_0(p)}(v_0),$$

for any other chart χ around p. Strictly speaking we still prove (1.2) but that follows from the chain rule and the remark that $c_{\chi_0,\chi'} = c_{\chi,\chi'} \circ c_{\chi_0,\chi}$. □

DEFINITION 4.3. *For a chart χ of M around p, we denote by*

$$\left(\frac{\partial}{\partial \chi_1}\right)_p, \ldots, \left(\frac{\partial}{\partial \chi_m}\right)_p \in T_pM$$

called ***the canonical basis of*** T_pM ***with respect to*** χ*, the basis that corresponds to the canonical one $e_1, \ldots, e_m$ of $\mathbb{R}^m$ via the isomorphism from the previous lemma.*

In other words,

$$\left(\frac{\partial}{\partial \chi_i}\right)_p \in T_pM$$

is the unique tangent vector at p with

$$\left(\frac{\partial}{\partial \chi_i}\right)_p^{\chi} = e_i.$$

EXERCISE 4.3. If χ and χ' are two charts of M around p, $c = c_{\chi,\chi'}$, show that

$$\left(\frac{\partial}{\partial \chi_i}\right)_p = \sum_j \frac{\partial c_j}{\partial x_i}(\chi(p)) \left(\frac{\partial}{\partial \chi'_j}\right)_p. \tag{1.3}$$

Finally, let us show show that, as we wanted, smooth maps induce differentials at the level of tangent spaces.

LEMMA 4.4. *Given a smooth map $f : M \to N$ between two manifolds, and given $p \in M$, there exists a unique linear map*

$$(df)_p : T_pM \to T_{f(p)}N, \quad v \mapsto (df)_p(v)$$

with the property that, for any chart χ of M around p and χ' of N around $f(p)$, one has:

$$((df)_p(v))^{\chi'} = (d\, f_{\chi,\chi'})_{\chi(p)}(v^{\chi})$$

for all $v \in T_pM$.

PROOF. Fixing v and χ, the required condition forces the definition of $(d\,f)_p(v)$; to see that one ends up with a tangent vector to N at $f(p)$ (i.e. the condition (1.2) is satisfied in this context) one uses again the chain rule. One still has to check that the resulting map does not depend on the choice of χ- but that follows again from the chain rule. □

DEFINITION 4.5. *For a smooth map $f : M \to N$ and $p \in M$, the resulting linear map $(df)_p : T_pM \to T_{f(p)}N$ is called* ***the differential of*** f ***at the point*** p.

It should be clear (or an easy exercise) that the chain rule continues to hold.

EXERCISE 4.4. Show that $(df)_p$ is uniquely characterized also by the condition that, for any $\gamma \in \text{Curves}_p(M)$, one has:

$$(df)_p(\frac{d\gamma}{dt}(0)) = \frac{df\circ\gamma}{dt}(0).$$

EXERCISE 4.5. For any curve $\gamma \in \text{Curves}_p(M)$ defined on the interval $I = (-\epsilon, \epsilon)$, viewed as a smooth map $\gamma : I \to M$ between manifolds, show that

$$(d\gamma)_0\left(\left(\frac{\partial}{\partial t}\right)_0\right) = \frac{d\gamma}{dt}(0) \quad \text{(an equality of vectors in } \in T_pM).$$

2. Tangent vectors as directional derivatives

Here is another approach to the tangent spaces; the brief philosophy of this approach is:

tangent vectors at p allow us to take directional derivatives (at p) of smooth functions

For this we return to the differential

$$(df)_p(v)$$

of a smooth function $f : \Omega \to \mathbb{R}$ ($\Omega \subset \mathbb{R}^m$ open) at a point $p \in \Omega$, applied to a (tangent) vector $v \in \mathbb{R}^m$, that we reinterpreted it as a v-derivative at p

$$(df)_p(v) = \frac{\partial f}{\partial v}(p) = \partial_v(f)(p).$$

Hence, if we want to understand what the (tangent) vector v really does (at p), one may say that it defines a function

$$\partial_v = \frac{\partial}{\partial v} : \mathcal{C}^\infty(\Omega) \to \mathbb{R};$$

What are the main properties of this function? Well, it is clearly linear and, moreover, it acts on the product of functions according to the Leibniz rule at p:

$$\partial_v(fg) = f(p)\partial_v(g) + g(p)\partial_v(g)$$

for all $f, g \in \mathcal{C}^\infty(\Omega)$. Note that here we are making use of the algebra structure on $C^\infty(\Omega)$: we make reference to the vector space structure (linearity) as well as to the multiplication (to make sense of the Leibniz identity at p).

The main point of this discussion is that, conversely, any linear map

$$\partial : \mathcal{C}^\infty(\Omega) \to \mathbb{R}$$

which is linear and satisfies the Leibniz identity at p is of type ∂_v for a unique vector $v \in \mathbb{R}^m$ (this will follow from the discussion below). This gives another perspective/possible approach to tangent spaces.

DEFINITION 4.6. *Given an m-dimensional manifold M and $p \in M$, a **derivation of** M **at** p is any linear map*

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

that is linear and satisfies the Leibniz identity at p.

We denote by $T_p^{\boldsymbol{deriv}}M$ the vector space of all such derivations M at p (a vector space using the usual addition and multiplication by scalars of linear maps).

EXAMPLE 4.7. When $M = \Omega$ is an open in $\mathbb{R}^m$ (endowed with the canonical smooth structure), $p \in \Omega$, then the operation of taking the usual partial derivatives at p,

$$\left(\frac{\partial}{\partial x_i}\right)_p : \mathcal{C}^\infty(U) \to \mathbb{R}, \quad f \mapsto \frac{\partial f}{\partial x_i}(p),$$

are derivations at p- hence they can be interpreted as vectors

$$\left(\frac{\partial}{\partial x_1}\right)_p, \ldots, \left(\frac{\partial}{\partial x_m}\right)_p \in T_p^{\text{deriv}}\Omega.$$

As it will follow from the discussion below, they form a basis of $T_p^{\text{deriv}}\Omega$.

Similarly, for any manyfold M, a chart $\chi : U \to \Omega \subset \mathbb{R}^m$ of M around p induces **partial derivatives at p w.r.t. the chart** χ,

$$\frac{\partial f}{\partial \chi_i}(p) := \frac{\partial f_\chi}{\partial x_i}(\chi(p)),$$

where $f_\chi = f \circ \chi^{-1} : \Omega \to \mathbb{R}$. And then vectors

$$\left(\frac{\partial}{\partial \chi_1}\right)_p, \ldots, \left(\frac{\partial}{\partial \chi_m}\right)_p \in T_p^{\text{deriv}}M$$

which form a basis of $T_p^{\text{deriv}}M$ (but this is still to be proven).

Let us point out, right away, that also this model for tangent spaces serves the original purpose: we can make sense of speeds

$$\dot{\gamma}(0) \in T_p^{\text{deriv}}M.$$

Here we use the notation $\dot{\gamma}(0)$ instead of $\frac{d\gamma}{dt}(0)$ to (temporarily) avoid overlaps with the previous section. The actual definition should be clear: just consider the variation of functions along γ (at $t = 0$):

$$\partial_{\dot{\gamma}(0)} : \mathcal{C}^\infty(M) \to \mathbb{R}, \quad \partial_{\dot{\gamma}(0)}(f) := \left.\frac{d}{dt}\right|_{t=0} f(\gamma(t)).$$

Of course, one can use the derivative at any t on the domain of γ, giving rise to

$$\partial_{\dot{\gamma}(t)} \in T_{\gamma(t)}^{\text{deriv}}M.$$

There is also a very natural interaction (which turns out to be an isomorphism) with the tangent space from the previous section. To see this, fix $v \in T_pM$; we try to define $\partial_v(f)$ for $f \in C^\infty(M)$. Once a chart $\chi : U \to \Omega \subset \mathbb{R}^m$ of M around p is chosen, we can use the representation of f with respect to χ, $f_\chi : \Omega \to \mathbb{R}$, as well as the one of v, $v^\chi \in \mathbb{R}^m$. Using again the chain rule, we see that

$$\partial_v(f) := \frac{\partial f_\chi}{\partial v^\chi}(\chi(p)) = (df_\chi)_{\chi(p)}(v^\chi)$$

actually does not depend on the choice of χ. Therefore we obtain

$$\partial_v : \mathcal{C}^\infty(M) \to \mathbb{R};$$

it should be clear that it is a derivation of M at p, hence

$$\partial_v \in T_p^{\text{deriv}}M.$$

For speeds of curves, when $v = \frac{d\gamma}{dt}(0)$ as defined in the previous section, one obtains $\partial_{\dot{\gamma}(0)}$ that we have just discussed. Similarly the vectors $\left(\frac{\partial}{\partial \chi}\right)_p$ from the previous section (Definition 4.3) give rise to the similar vectors defined in this section (Example 4.7). In general, we obtain a linear map

$$I_p : T_pM \to T_p^{\text{deriv}}M, \quad v \mapsto \partial_v.$$

THEOREM 4.8. *The map $I_p : T_pM \to T_p^{deriv}M$ is a linear isomorphism.*

At this point we could just prove this theorem and then transfer all the properties/constructions that we know for T_pM to $T_p^{\text{deriv}}M$, via I. It is true that the proof of the theorem is not completely trivial (well, it is not very difficult either, but requires some preparation). However, more importantly, it is instructive to look at $T_p^{\text{deriv}}M$ independently and check that it does satisfy the properties that we want. That will reveal the advantages of $T_p^{\text{deriv}}M$ as well as the further nature of tangent spaces.

Regarding the nature of $T_p^{\text{deriv}}M$, note that it is completely algebraic. See e.g. the following exercise.

EXERCISE 4.6. Let A be an algebra (with unit 1_A). Recall that a character on A is any map

$$\chi : A \to \mathbb{R}$$

which is a map of algebras, i.e. it is linear, multiplicative ($\chi(ab) = \chi(a)\chi(b)$ for all $a, b \in A$) and $\chi(1_A) = 1$. We denote by $X(A)$ the set of all characters on A.

For instance, when $A = \mathcal{C}^\infty(M)$ is the algebra of smooth functions on a manifold M, then any $p \in M$ gives rise to the character

$$\chi_p : \mathcal{C}^\infty(M) \to \mathbb{R}, \quad \chi_p(f) = f(p).$$

(and, even more is true: any character arises in this way!).

Motivated by this example, for any algebra A and any character χ on A we introduce the notion of χ-derivation on A, by which we mean a linear map $\chi : A \to \mathbb{R}$ satisfying the derivation identity:

$$\partial(ab) = \chi(a)\partial(b) + \partial(a)\chi(b)$$

for all $a, b \in A$. We denote by $\text{Der}_\chi(A)$ the (vector) space of all such derivations. Intuitively, this may be interpreted as "the tangent space of $X(A)$ at χ".

Assume now that M is a manifold, and we want to apply the previous discussion to two algebras: $A = \mathcal{C}(M)$ and $A^\infty = \mathcal{C}^\infty(M)$. Let us mention that the construction $M \ni p \mapsto \chi_p$ defines a bijection between M and both $X(A)$ as well as $X(A^\infty)$ (we are not asking you to prove this here- for details please see "Inleiding Topologie"). Any way, the point is that characters do not see the difference between A and A^∞. However, derivations (i.e. tangent spaces) do. More precisely, while

$$\text{Der}_{\chi_p}(\mathcal{C}^\infty(M)) = T_p^{\text{deriv}}M,$$

we are asking you to prove that $\text{Der}_{\chi_p}(\mathcal{C}(M)) = 0$.

(Hint: the proof is at least ten times shorter than the bla-bla (but otherwise interesting) of the exercise)

EXERCISE 4.7. When $A = \mathbb{R}[X_1, \ldots, X_m]$ is the algebra of polynomials in m variables show that any character is the evaluation χ_x at some point $x \in \mathbb{R}^m$, and then show that $\text{Der}_{\chi_x}(A)$ is a finite dimensional vector space and exhibit a basis.

Then consider another A: the algebra of polynomial functions on the circle S^1. And, again, consider the characters χ_p for $p \in S^1$ and compute $\text{Der}_{\chi_p}(A)$.

We now return to our tangent spaces $T_p^{\text{deriv}}M$ and the properties that we expect from them:

EXERCISE 4.8. Show that any smooth map $F : M \to N$ between two manifolds induces, at any $p \in M$, a linear map

$$(dF)_p : T_p^{\text{deriv}}M \to T_{F(p)}^{\text{deriv}}N, \quad (dF)_p(\partial)(f) = \partial(f \circ F).$$

Then state and prove the chain rule and deduce that any local diffeomorphism induces an isomorphism at the level of the (deriv-)tangent spaces.

We now discuss another aspect that $T_p^{\text{deriv}}M$ reveals: the local nature of derivations.

LEMMA 4.9. *Any derivation of M at p, $\partial \in T_p^{deriv}M$, has the following local property: for $f_1, f_2 \in \mathcal{C}^\infty(M)$ one has:*

$$f_1 = f_2 \textit{ in a neighborhood of } p \Longrightarrow \partial(f_1) = \partial(f_2).$$

PROOF. Let $f = f_1 - f_2$. We know that $f = 0$ in a neighborhood U of p and we want to prove that $\partial(f) = 0$. We may assume that U is chosen so that it is diffeomorphic to a ball $B(0, \epsilon) \subset \mathbb{R}^m$, by a diffeomorphism that takes p to 0. As we have already noticed, one can find a smooth function on $\mathbb{R}^m$ that is supported inside $B(0, \epsilon)$ and is non-zero at 0 (e.g. take $x \mapsto g(\frac{1}{\epsilon^2}||x||^2)$, where $g : R \to \mathbb{R}$ is the function from Exercise 2.2). Moving from $B(0, \epsilon)$ to U, and extending by zero outside U, we find a function $\eta \in C^\infty(M)$ which is supported inside U and $\eta(p) \neq 0$. Then $\eta f = 0$, and applying the Leibniz identity for ηf we find $\eta(p)\partial(f) = 0$ hence, since $\eta(p) \neq 0$, one must have $\partial(f) = 0$. ☐

REMARK 4.10. The previous lemma can be packed into a more conceptual conclusion: any derivations ∂ at p descends to the quotient space

$$\mathcal{C}_p^\infty(M) := \mathcal{C}^\infty(M)/I_p = C^\infty(M)/\sim_p .$$

where I_p is the space of functions that vanish around p and $\sim_p$ is the associated equivalence relation:

$$f_1 \sim_p f_2 \Longleftrightarrow f_1 = f_2 \text{ in a neighborhood of } p.$$

For $f \in \mathcal{C}^\infty(M)$, its equivalence class is denoted

$$\text{germ}_p(f) \in \mathcal{C}_p^\infty(M)$$

and is called **the germ of f at p**. The space of germs at p, i.e. our quotient $\mathcal{C}_p^\infty(M)$, is still an algebra (since I_p is an ideal of $\mathcal{C}^\infty(M)$); explicitly, the operations are induced from $\mathcal{C}^\infty(M)$:

$$\text{germ}_p(f + g) = \text{germ}_p(f) + \text{germ}_p(g), \quad \textit{etc.}$$

The locality property from the previous lemma can now be interpreted as an isomorphism

$$T_p^{\text{deriv}}M \cong \text{Der}_p(\mathcal{C}_p^\infty(M))$$

between the tangent space and the space of all derivations of $\mathcal{C}_p^\infty(M)$ at p (with respect to the evaluation at p).

Assume that $U \subset M$ is an open containing p. Then any derivation $\partial \in T_p^{\text{deriv}}U$ gives a derivation $\tilde{\partial} \in T_p^{\text{deriv}}M$ simply by

$$\tilde{\partial}(f) := \partial(f|_U).$$

Of course, the resulting map

$$T_p^{\text{deriv}}U \to T_p^{\text{deriv}}M, \quad \partial \mapsto \tilde{\partial}$$

is just the map induced (cf. Exercise 4.8 above) by the inclusion $U \hookrightarrow M$.

LEMMA 4.11. *Given $p \in U \subset M$ with U open, the canonical map $T_p^{deriv}U \to T_p^{deriv}M$ is a linear isomorphism.*

PROOF. For injectivity, let $\partial : \mathcal{C}^\infty(U) \to \mathbb{R}$ be a derivation at p such that $\partial(\tilde{f}|U) = 0$ for all $\tilde{f} \in \mathcal{C}^\infty(M)$; we must show that $\partial = 0$. This follows immediately from the following:

- $\partial(f)$ depends only on f in an arbitrarily small neighborhood of p. This is the content of the previous lemma.

- for any $f \in \mathcal{C}^\infty(U)$, there exists $\tilde{f} \in \mathcal{C}^\infty(M)$ s.t. $f = \tilde{f}$ in some neighborhood of p.

To prove the last item, we take $\tilde{f} := \eta \cdot f$ where $\eta \in \mathcal{C}^\infty(M)$ is supported inside U (so that $\tilde{f}$ is defined and smooth on the entire M) and $\eta = 1$ in some (smaller) neighborhood of p. The existence of η is a local problem- therefore it suffice to build a smooth function η on $\mathbb{R}^m$ that is supported in the ball $B(0,1)$ and is 1 in a neighborhood of 0 (say on $B(0, \frac{1}{3})$); for that one takes again η of type $\eta(x) = g(||x||^2)$ where $g : \mathbb{R} \to \mathbb{R}$ is any smooth function that is 1 when $|t| < \frac{1}{3}$ and is 0 when $t \geq \frac{1}{2}$.

For the surjectivity, we start with $\tilde{\partial} : \mathcal{C}^\infty(M) \to \mathbb{R}$ and we want to build ∂. By now the definition should be clear: or $f \in \mathcal{C}^\infty(U)$, choose any $\tilde{f} \in \mathcal{C}^\infty(M)$ that coincides with f near p (possible by the first part) and set $\partial(f) := \tilde{\partial}(\tilde{f})$. The previous lemma shows that the definition of $\partial(f)$ does not depend on the choice of $\tilde{f}$. And this can used also to check that ∂ is still a derivation at p. E.g., for the Leibniz identity, choosing extensions $\tilde{f}$ or f and $\tilde{g}$ or g, then use the extension $\tilde{f}\tilde{g}$ or fg and use the Leibniz identity for $\tilde{\partial}(\tilde{f}\tilde{g})$. □

We are now ready to prove the main theorem of this section,

Proof of Theorem 4.8. Fix a chart $\chi : U \to \mathbb{R}^m$ around p that takes p to $0 \in \mathbb{R}^m$. Since the induced basis of T_pM (of Definition 4.3) are sent by I_p into the similar vectors $\left(\frac{\partial}{\partial \chi}\right)_p$ of this section (see Example 4.3), it suffices to show that the last ones form a basis of $T_p^{\text{deriv}}M$. Or, by the previous lemma, of $T_p^{\text{deriv}}U$. But the chart χ takes $T_p^{\text{deriv}}U$ isomorphically into $T_0^{\text{deriv}}\Omega$ (this follows using e.g. the chain rule- see Exercise 4.8) and, o course, takes our vectors into $\left(\frac{\partial}{\partial x_i}\right)_0 \in T_0^{\text{deriv}}\Omega$. Using again the previous lemma, it suffices to show that

$$\left(\frac{\partial}{\partial x_i}\right)_0 : \mathcal{C}^\infty(\mathbb{R}^m) \to \mathbb{R}$$

form a basis of $T_p^{\text{deriv}}\mathbb{R}^m$. The linear independence is immediate: if $\sum_i \lambda_i \left(\frac{\partial}{\partial x_i}\right)_0 = 0$, applying this to x_i (i.e. to the coordinate function $x \mapsto x_i$) we find $\lambda_i = 0$.

We still have to check that any derivation at 0,

$$\partial : \mathcal{C}^\infty(\mathbb{R}^m) \to \mathbb{R},$$

is a linear combination of the standard partial derivatives at 0:

$$\partial = \sum_i \lambda_i \left(\frac{\partial}{\partial x_i}\right)_0 \quad \text{with } \lambda_i \in \mathbb{R}.$$

We will show that $\lambda_i := \partial(x_i) \in \mathbb{R}$ do the job. To show this, we use the fact that any $f \in \mathcal{C}^\infty(\mathbb{R}^m)$ can be written as

$$f(x) = f(0) + \sum_i x_i g_i(x), \quad \text{with } g_i \in \mathcal{C}^\infty(\mathbb{R}^m) \tag{2.1}$$

(see below). Using the Leibniz identity (which also implies that ∂ is 0 on constant functions) we obtain

$$\partial(f) = \sum_i \lambda_i g_i(0),$$

while $\frac{\partial f}{\partial x_i}(0) = g_i(0)$. Hence we are left with proving that any f can be written in the form (2.1). For this we use the identity

$$h(1) - h(0) = \int_0^1 \frac{dh}{dt}(t)dt$$

for all smooth functions $h : [0,1] \to \mathbb{R}$; of course, we want to choose h such that $h(1) = f(x)$ and $h(0) = f(0)$. Choose then $h(t) = f(tx)$ and we obtain the desired identity (2.1) with

$$g_i = \int_0^1 \frac{\partial f}{\partial x_i}(tx)dt.$$

□

3. Tangent spaces: conclusions

Let us put together the main conclusions on tangent spaces.

3.1. Via charts, or as derivations: In the previous two sections we have seen (for any m-dimensional manifold M, $p \in M$) two ways of looking at the tangent space T_pM:

T1: *its elements $v \in T_pM$ can be represented w.r.t. any chart χ around p by vectors $v^\chi \in M$; the passing from one chart to the other was given by*

$$v^{\chi'} = (d\,c)_{\chi(p)}(v^\chi), \quad \text{where } c = c_{\chi,\chi'} = \chi' \circ \chi^{-1}.$$

T2: *its elements $v \in T_pM$ can be interpreted as derivations on $\mathcal{C}^\infty(M)$ at p.*

The equivalence between the two is given by the isomorphism $I_p : T_pM \to T_p^{\text{deriv}}M$ of Theorem 4.8. From now on we will identify the two (via I_p) and use only the notation T_pM. The upshot is that, when describing a vector $v \in T_pM$ we have the choice to describe it w.r.t. a chart, or as a derivation at p.

3.2. Speeds: Most importantly, any curve $\gamma \in \text{Curves}_p(M)$ has a speed

$$\frac{d\gamma}{dt}(0) \in T_pM$$

Explicitly:

- w.r.t. a chart χ, it is $\frac{d\gamma^\chi}{dt}(0) \in \mathbb{R}^m$.
- as a derivation at p it sends a smooth function f to $\frac{df\circ\gamma}{dt}(0)$.

(and, with the same descriptions, one has $\frac{d\gamma}{dt}(t) \in T_{\gamma(t)}M$ for any t in the domain of γ).

And this gives the best way to think (and even to work with) tangent spaces:

T0: $T_pM = \{\frac{d\gamma}{dt}(0) : \mathit{Curves}_p(M)\}$, *where we know that two such curves have the same speed at $t = 0$ if and only if that happens for their representations w.r.t. a/any chart.*

3.3. Basis with respect to a chart: The tangent space is an m-dimensional vector space-hence isomorphic to $\mathbb{R}^m$. Each chart χ around p gives rise to a basis of T_pM (hence to a specific isomorphism with $\mathbb{R}^m$):

$$\left(\frac{\partial}{\partial \chi_1}\right)_p, \ldots, \left(\frac{\partial}{\partial \chi_m}\right)_p \in T_pM,$$

called the canonical basis w.r.t. χ:

T1: w.r.t. χ, $\left(\frac{\partial}{\partial \chi_i}\right)_p$ corresponds to $e_i \in \mathbb{R}^m$.

T2: as derivations, they are given by the partial derivatives w.r.t. χ:

$$\frac{\partial f}{\partial \chi_i}(p) := \frac{\partial f_\chi}{\partial x_i}(\chi(p)).$$

T0: as speeds, $\left(\frac{\partial}{\partial \chi_i}\right)_p$ are induced by the paths

$$t \mapsto \chi^{-1}(\chi(p) + te_i).$$

See also Definition 4.3 and Example 4.7. The way that these vectors change when we change the chart is easiest read off from the point of view of [T1], which gives the formula (1.3). Or, using the partial derivatives given by the point of view of [T2], we have the more compact formula:

$$\left(\frac{\partial}{\partial \chi_i}\right)_p = \sum_{j=1}^{m} \frac{\partial c_j}{\partial \chi_i}(\chi(p)) \left(\frac{\partial}{\partial \chi'_j}\right)_p = \sum_{j=1}^{m} \frac{\partial \chi'_j}{\partial \chi_i}(p) \left(\frac{\partial}{\partial \chi'_j}\right)_p \quad \text{(where } c = \chi' \circ \chi^{-1}\text{)}.$$

Or, if one denotes by x and y the charts χ and χ', respectively, and c by y but interpreted as a function of x, one come across the more compact (but a bit sloppy) formula:

$$\frac{\partial}{\partial x_i} = \sum_j \frac{\partial y_j}{\partial x_i} \frac{\partial}{\partial y_j}.$$

3.4. Differentials: While, intuitively, the tangent T_pM is "the linearization of M near p", we can also linearize smooth maps. And this was one of the main properties that we wanted from tangent spaces (and the property that allows one to use them in order to compare different manifolds). More precisely, one can talk about the differential

$$(dF)_p : T_pM \to T_{F(p)}N \quad \text{for any smooth map } f : M \to N.$$

Again, this can described from the three points of view:

T0: It sends the speed (at $t = 0$) of $\gamma \in \mathrm{Curves}_p(M)$ to the one of $F \circ \gamma \in \mathrm{Curves}_{F(p)}(N)$.

T1: For $v \in T_pM$ represented w.r.t. χ by v^χ, $(dF)_p(v)$ is represented w.r.t. a/any chart χ' around $f(p)$ by the image of v^χ by $(dF_{\chi,\chi'})_{\chi(p)}$.

T2: $(dF)_p(v)$, as a derivation, acts on a function $g \in \mathcal{C}^\infty(M)$ by acting with v on $g \circ F$.

Or, if we want to write down $(dF)_p$ w.r.t. bases induced by charts χ around p and χ' around $F(p)$:

$$\left(\frac{\partial}{\partial \chi_i}\right)_p \xmapsto{(dF)_p} \sum_{j=1}^{m} \frac{dF^j_{\chi,\chi'}}{dx_i}(\chi(p)) \left(\frac{\partial}{\partial \chi'_j}\right)_{F(p)},$$

where we use the representation $F_{\chi,\chi'} = \chi' \circ F \circ \chi^{-1}$ of F w.r.t. the two bases. Or, in other words, the matrix corresponding to $(dF)_p$ w.r.t. the two bases is precisely the matrix of the partial derivatives of $F_{\chi,\chi'}$ at $\chi(p)$.

Finally, there is the following characterization of immersions/submersions which is often taken as a definition (we did not do that because we did not want to wait until the tangent spaces were defined. However, this should certainly be the way to think about immersions/submersions from now on!). For embedded submanifolds of Euclidean spaces, this was already pointed out in the previous chapter (Exercise 3.13 there). The following exercise (the general case) should be rather immediate, but it is instructive to do it using the various properties we mentioned above (e.g. without mentioning any formula!).

Exercise 4.9. Show that a smooth map $f : M \to N$ between two manifolds is an immersion/submersion at a point $p \in M$ if and only if its differential at p, $(df)_p : T_pM \to T_{f(p)}N$, is injective/surjective.

3.5. Submanifolds: Let us restrict for simplicity to embedded submanifolds. Given such a submanifold $N \subset M$ of a manifold M, the inclusion $i : N \hookrightarrow M$ is an immersion, therefore, for each $p \in N$,

$$(di)_p : T_pN \to T_{i(p)}M$$

is an injection. In the same way that we do not write $i(p)$ but p, this injection will be viewed as an inclusion

$$T_pN \subset T_{i(p)}M.$$

From the point of view of curves this is completely clear: since any curve in N is also one in M, the speed $\frac{d\gamma}{dt}(0) \in T_pN$ is now identified with $\frac{d\gamma}{dt}(0) \in T_pM$ (we do not even have a notation to distinguish the two!). Also, as derivations, if $v \in T_pN$, it acts on $f \in \mathcal{C}^\infty(M)$ by acting with v on $f|_N$.

In this way tangent spaces give us (linear) information on the way that N sits inside M (near p).

Finally, here is the full version of the regular value theorem proven in Chapter 3 (Theorem 3.21 there), this time with exra-information on the tangent spaces.

THEOREM 4.12 (the regular value theorem with tangent spaces). *If $q \in N$ is a regular value of a smooth map*

$$F : M \to N,$$

then the fiber above q, $F^{-1}(q)$, is an embedded submanifold of M of dimension

$$dim(F^{-1}(q)) = dim(M) - dim(N)$$

and the tangent spaces $T_p\left(F^{-1}(q)\right)$, as subspaces of T_pM, coincide with the kernel of the differential $(dF)_p : T_pM \to T_{F(p)}N$:

$$T_p\left(F^{-1}(q)\right) = Ker(dF)_p \quad ((\text{for all } p \in F^{-1}(q)).$$

PROOF. Nice exercise (hint: just look back at our reminder on Analysis). □

3.6. Back to Euclidean spaces and their embedded submanifolds: Back to Analysis, i.e. for a manifold of type $\mathbb{R}^k$ (or an open inside it), the general/abstract theory of this chapter tells us that the tangent spaces $T_p\mathbb{R}^k$ come with a canonical basis

$$\left(\frac{\partial}{\partial x_1}\right)_p, \ldots, \left(\frac{\partial}{\partial x_m}\right)_p. \tag{3.1}$$

(the smooth structure on $\mathbb{R}^k$ can be given by one, global, chart- the identity).

As derivations at p, these correspond to the standard partial derivatives. With respect to the canonical chart, they are represented by the canonical basis $e_1, \ldots, e_k$ of $\mathbb{R}^k$. As speeds, they correspond to the paths $t \mapsto p + te_i$. In any case, we have an isomorphism that will be treated from now on as an identification:

$$T_p\mathbb{R}^k = \mathbb{R}^k,$$

but we will still keep the notation (3.1) for the canonical basis. This is in order to indicate that we are thinking about/using tangent vectors.

With the identification above, whenever we have a smooth map $f : M \to \mathbb{R}^k$, we talk about its differentials as maps $(df)_p : T_pM \to \mathbb{R}^k$.

EXERCISE 4.10. For $f \in \mathcal{C}^\infty(M)$ and $v \in T_pM$, check that the action of v (as a derivation) on f is precisely

$$\partial_v(f) = (df)_p(v).$$

EXERCISE 4.11. Check that, via this identification, the speeds $\frac{d\gamma}{dt}(t)$ discussed in this chapter are identified with the standard ones from Analysis.

As in the previous exercise for curves, we have a similar problem for embedded submanifolds:

EXERCISE 4.12. For an embedded submanifold $M \subset \mathbb{R}^k$, $p \in M$, we now have two tangent spaces,

$$T_pM \subset \mathbb{R}^k,$$

namely the one from this chapter $T_pM \subset T_p\mathbb{R}^k$ (see subsection 3.5) combined with the identification $T_p\mathbb{R}^k = \mathbb{R}^k$, and the one from Analysis using derivatives of curves. Show that the two coincide. Try to find different arguments; e.g. one using the previous exercise, one using the regular values theorem, etc.

By now, it should also be clear that for smooth maps between embedded submanifolds of Euclidean spaces, the differential discussed in this chapter becomes (identified with) the one from Analysis (e.g. as recalled in Chapter 2, Exercise 2.10).

4. Vector fields

We now discuss vector fields. The brief philosophy is:

While for embedded submanifolds $M \subset \mathbb{R}^k$ each single tangent space $T_pM \subset \mathbb{R}^k$ was interesting (e.g. because it reflects the position of M inside $\mathbb{R}^k$ near p), for a general manifold M this is less so. Instead, it is the entire family $\{T_pM\}_{p\in M}$ that is interesting. I.e. not one single tangent vector, but a family $\{v_p\}_{p\in M}$ of vectors (of course, "varying smothly" with respect to p). I.e. vector fields.

Looking back at the previous sections of this chapter, we see that we insisted in making sense of T_pM as vector spaces independently of the way that M may sit in some larger Euclidean space. Now, just as a vector space, T_pM is rather boring: it is isomorphic to $\mathbb{R}^m$. And the same is true for any finite dimensional vector space V; however, one should keep in mind that, for a bare m-dimensional vector space V, realizing an explicit isomorphism with $\mathbb{R}^m$ amounts to extra choices (namely a basis of V). In particular, for the tangent spaces T_pM, we obtained identifications with $\mathbb{R}^m$ once we fixed a chart χ (and these identifications work for p in the domain of χ only!). Actually this is a "problem" (or better: "interesting phenomena") not only for general Ms, but even for embedded submanifolds $M \subset \mathbb{R}^k$: while isomorphisms

$$\mathbb{R}^m \cong T_pM \subset \mathbb{R}^k$$

can be chosen for each p, often cannot be chosen so that they depend smoothly on $p \in M$. The "hairy ball theorem" implies that this is the case already for $M = S^2 \subset \mathbb{R}^3$:

EXERCISE 4.13. Assume that for each $p \in S^2$ one can find a linear isomorphism

$$\Phi_p : \mathbb{R}^2 \to T_pS^2 \subset \mathbb{R}^3$$

so that the Φ_ps vary smoothly with respect to p, in the sense that the map

$$\Phi : S^2 \times \mathbb{R}^2 \to \mathbb{R}^3, \quad (p, v) \mapsto \Phi_p(v)$$

is smooth. Show that there exists a no-where vanishing vector field on S^2, i.e. a smooth map

$$X : S^2 \to \mathbb{R}^3,$$

nowhere vanishing, such that $X(p) \in T_pS^2$ for all $p \in S^2$. Try now to construct such an X! Then state the "hairy ball theorem".

More generally, for an embedded submanifold $M \subset \mathbb{R}^k$, a vector field on M is a function

$$M \ni p \mapsto X(p) \in T_pM \subset \mathbb{R}^k$$

which, when interpreted as a function with values in $\mathbb{R}^k$, is smooth. What if an embedding into some Euclidean space is not fixed (or if we use a different one)? First of all, we can still look at maps

$$X : M \ni p \mapsto X_p \in T_pM$$

(interpreted also as families $X = \{X_p\}_{p \in M}$). These will be called **set-theoretical vector fields** on M. To make sense of their smoothness, we use charts $\chi : U \to \Omega \subset \mathbb{R}^m$ of M. Recall that any such chart induces a basis $\left(\frac{\partial}{\partial \chi_i}\right)_p$ of T_pM hence any set-theoretical vector field X can be written as

$$X_p = \sum_{i=1}^{m} X^i_\chi(\chi(p)) \left(\frac{\partial}{\partial \chi_i}\right)_p \tag{4.1}$$

for all $p \in M$, where each X^i_χ is a function

$$X^i_\chi : \Omega \to \mathbb{R}.$$

These will be called **the coordinates (functions) of X w.r.t. the chart χ.**

DEFINITION 4.13. *A vector field on a manifold is any set-theoretical vector field X with the property that its coordinate functions X^i_χ w.r.t any chart χ are smooth.*
We denote by $\mathfrak{X}(M)$ the set of all vector fields on M.

Note the algebraic structure on $\mathfrak{X}(M)$ that is present right away:

- it is a vector space, with the addition an multiplication of scalars defined pointwise:

$$(X + Y)_p := X_p + Y_p, \quad (\lambda \cdot X)_p := \lambda \cdot X_p.$$

- it is a $\mathcal{C}^\infty(M)$-module, i.e. there is an operation of multiplication of vector fields by smooth functions,

$$\mathcal{C}^\infty(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M), \quad (f, X) \mapsto f \cdot X.$$

  Again, this is defined pointwise by the obvious formula:

$$(f \cdot X)_p := f(p) \cdot X_p.$$

Next, we look for alternative ways to characterize the smoothness of vector fields; as a bonus, these will reveal more of the structure that vector fields carry.

First of all, while tangent vectors could be interpreted as derivations (at given points $p \in M$), for any $X \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$ one has $\partial_{X_p}(f) \in \mathbb{R}$ for each $p \in M$, therefore a function

$$L_X(f) : M \to \mathbb{R}, \quad p \mapsto \partial_{X_p}(f) = (df)_p(X_p).$$

This is called **the Lie derivative of f along the vector field** X. Of course, we expect that $L_X(f)$ is again smooth.

PROPOSITION 4.14. *A set-theoretical vector field X on M is smooth if and only if*

$$L_X(f) \in C^\infty(M) \text{ for all } f \in C^\infty(M).$$

PROOF. Assume first that X is smooth in sense of the definition. For $f \in C^{\infty}(M)$, we can check the smoothness of $L_X(f)$ locally: for an arbirary chart χ we have to make sure that $L_X(f) \circ \chi^{-1}$ is smooth. But applying the definitions, we see that this expression is precisely

$$\sum_i X^i_{\chi} \frac{\partial f_{\chi}}{\partial x_i}$$

and then smoothness follows.

For the converse, by the type of arguments we have already seen, it suffices to show that for any point $p \in M$ there exists a chart around p such that the coefficients X_i^{χ} are smooth. For an arbitrary chart $\chi : U \to \Omega \subset \mathbb{R}^m$ around p, we choose $f_i \in \mathcal{C}^{\infty}(M)$ such that, in a smaller neighborhood $U_0 \subset U$ of p, f_i coincides with χ_i. By hypothesis, $L_X(f_i)$ is smooth; but over U_0, this function is precisely X_i^{χ}. Hence taking $\chi_0 := \chi|_{U_0}$, the coefficients $X_i^{\chi_0}$ □

The interpretation of tangent vectors at p as derivations at p can be pushed further to a similar characterization of vector fields. The algebraic objects are **derivations of** $\mathcal{C}^{\infty}(M)$, by which we mean linear maps

$$L : \mathcal{C}^{\infty}(M) \to \mathcal{C}^{\infty}(M)$$

satisfying the derivation low (Leibniz identity):

$$L(fg) = L(f)g + fL(g) \quad \text{for all } f, g \in \mathcal{C}^{\infty}(M).$$

We denote by $\mathrm{Der}(\mathcal{C}^{\infty}(M))$ the (vector) space of such derivations (note: again, this is a purely algebraic construction, that applies to any algebra A).

THEOREM 4.15. *For any $X \in \mathfrak{X}(M)$, the Lie derivative along X,*

$$L_X : \mathcal{C}^{\infty}(M) \to \mathcal{C}^{\infty}(M), \quad f \mapsto L_X(f)$$

is a derivation. Moreover,

$$I : \mathfrak{X}(M) \to Der(\mathcal{C}^{\infty}(M)), \quad X \mapsto L_X$$

is an isomorphism of vector spaces.

PROOF. The fact that L_X is a derivation follows right away from the similar property of the ∂_{X_p}s. The injectivity of I follows from the fact that, if a tangent vector $v_p \in T_pM$ is non-zero, there exists a smooth function f on M such that $\partial_{v_p}(f) \neq 0$ (why?).

For the surjectivity of I, let L be a derivation. Then, for each $p \in M$,

$$X_p : \mathcal{C}^{\infty}(M) \to \mathbb{R}, \quad f \mapsto L(f)(p)$$

is a derivation at p, hence defines a tangent vector $X_p \in T_pM$. Since$L_X(f) = L(f)$ is smooth for any smooth f, the previous proposition implies that X is smooth. □

And here is one clear advantage of the point of view of derivations: the presence of yet another structure on $\mathfrak{X}(M)$, namely a bracket operation

$$\mathfrak{X}(M) \times \mathfrak{X}(M) \to \in \mathfrak{X}(M).$$

More precisely, given two derivations on $\mathcal{C}^{\infty}(M)$, necessarily of type L_X and L_Y, their composition

$$L_X \circ L_Y : \mathcal{C}^{\infty}(M) \to \mathcal{C}^{\infty}(M)$$

is not a derivation anymore. Indeed, what we have is:

$$(L_X \circ L_Y)(fg) = (L_X \circ L_Y)(f)g + f(L_X \circ L_Y)(g) + L_X(f)L_Y(g) + L_Y(f)L_X(g),$$

i.e. a "defect term" $L_X(f)L_Y(g) + L_Y(f)L_X(g)$ which spoils the Leibniz identity. However, the other composition $L_Y \circ L_X$ comes with the same "defect term"; therefore, their difference

$$[L_X, L_Y] := L_X \circ L_Y - L_Y \circ L_X : \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$$

is again a derivation.

DEFINITION 4.16. *Given $X, Y \in \mathfrak{X}(M)$, we denote by $[X, Y] \in \mathfrak{X}(M)$, called* ***the Lie bracket of** X **and** Y, the (unique) vector field on M with the property that*

$$L_{[X,Y]} = L_X \circ L_Y - L_Y \circ L_X.$$

The following exercise should convince you of the advantage of the point of view of derivations.

EXERCISE 4.14. Compute the coordinate functions of $[X, Y]$ with respect to a chart χ, in terms of the coordinates functions of X and Y defined by (4.1).

EXERCISE 4.15. In this exercise we point out the main properties of the Lie bracket operation $[\cdot, \cdot]$ on $\mathfrak{X}(M)$.

First of all, the way it interacts with the other structure present on on $\mathfrak{X}(M)$- show that:

- w.r.t. the vector space structure: it is a bilinear and skew-symmetric operation
$$[\cdot, \cdot] : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \in \mathfrak{X}(M).$$
- w.r.t. the $\mathcal{C}^\infty(M)$-module structure: it satisfies the derivation rule:
$$[X, f \cdot Y] = f \cdot [X, Y] + L_X(f) \cdot Y,$$
for all $X, Y \in XX(M)$, $f \in \mathcal{C}^\infty(M)$.

And then a property that $[\cdot, \cdot]$ has on its own- **the Jacobi identity**:

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0 \quad \text{for all } X, Y, Z \in \mathfrak{X}(M).$$

EXERCISE 4.16. Here we point out the main properties of the operation $[\cdot, \cdot]$ on $\mathfrak{X}(M)$.

First of all, to see the way it interacts with the other structures present on $\mathfrak{X}(M)$, show that:

- w.r.t. the vector space structure: it is a bilinear and skew-symmetric operation
$$[\cdot, \cdot] : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M).$$
- w.r.t. the $\mathcal{C}^\infty(M)$-module structure: it satisfies the derivation rule:
$$[X, f \cdot Y] = f \cdot [X, Y] + L_X(f) \cdot Y,$$
for all $X, Y \in XX(M)$, $f \in \mathcal{C}^\infty(M)$.

Then show the following property that $[\cdot, \cdot]$ has on its own:

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

for all $X, Y, Z \in \mathfrak{X}(M)$ (this is known as **the Jacobi identity**).

Back to the the very definition of vector fields and of $\mathfrak{X}(M)$, here is another way to characterize smoothness. The idea is to put all the spaces T_pM (disjointly) in one big space

$$TM := \sqcup_{p \in M} T_pM = \{(p, v_p) : p \in M, v_p \in T_pM\}$$

and make this into a manifold, so that the smoothness of a set-theoretical vector field X is the same as the smoothness of X viewed as a map between manifolds:

$$X : M \to TM, \quad p \mapsto X_p.$$

To put a smooth structure on TM, i.e. to exhibit charts, we need to understand how to parametrize the points of TM by coordinates. Of course, we should start from a chart

$$\chi : U \to \Omega \subset \mathbb{R}^m$$

for M. The point is that such a chart also allows us to parametrize tangent vectors $v_p \in T_pM$, for $p \in U$, by coordinates $\hat{\chi}_i(v_p)$ w.r.t. to the induced basis $\left(\frac{\partial}{\partial \chi_i}\right)_p$ of T_pM: just decompose v_p as

$$v_p = \sum_i \lambda_i \left(\frac{\partial}{\partial \chi_i}\right)_p \quad \text{and set } \hat{\chi}_i(v_p) := \lambda_i. \tag{4.2}$$

Therefore, on the subset $TU \subset TM$, we have a natural "chart":

$$\tilde{\chi} : TU \to \Omega \times \mathbb{R}^m \subset \mathbb{R}^{2m},$$
$$\tilde{\chi}(p, v_p) = (\chi_1(p), \ldots, \chi_m(p), \hat{\chi}_1(v_p), \ldots, \hat{\chi}_m(v_p)).$$

Proposition 4.17. *TM can be made into a manifold in a unique way so that, for any chart (U, χ) of M, $(TU, \tilde{\chi})$ is a chart of TM.*

Moreover, this is also the unique smooth structure with the property that a set-theoretical tangent vector X on M is smooth if and only if it is smooth as a map $X : M \to TM$.

Proof. We first compute the change of coordinates between charts of type $(TU, \tilde{\chi})$. We start with two charts (U, χ) and (U', χ') for M, with change of coordinates $c = \chi' \circ \chi^{-1}$. We want to compute $\tilde{c} := \tilde{\chi}' \circ \tilde{\chi}^{-1}$. Hence we try to write $\tilde{\chi}'(p)$ as a function depending on $\tilde{\chi}(p)$. While the first m coordinates are $c(\chi(p))$, we need to understand the last m. For that we use:

$$\left(\frac{\partial}{\partial \chi_i}\right)_p = \sum_j \frac{\partial c_j}{\partial x_i}(\chi(p)) \left(\frac{\partial}{\partial \chi'_i}\right)_p .$$

(see (1.3)). TRhen, for a tangent vector $p \in M$, replacing this in (4.2), we find that

$$\hat{\chi'}_j(v_p) = \sum_i \hat{\chi}_i(v_p) \frac{\partial c_j}{\partial x_i}(\chi(p)).$$

Denoting by $(x, \hat{x})$ the coordinates in $\mathbb{R}^{2m} = \mathbb{R}^m \times \mathbb{R}^m$, we find that

$$\tilde{c}(x, \hat{x}) = (c(x), \sum_i \hat{x}_i \frac{\partial c}{\partial x_i}(x).$$

Hence the changes of coordinates is smooth.

However, strictly speaking, to show that TM is a manifold, there are still a few things to be done. First of all, we should have made it into a topological space. But, as we discussed at the end of Section 4 from Chapter 3, the topology can actually be recovered from an atlas. Let us give the details, independently of our earlier discussion. Given

$$\text{a chart } \chi : U \to \Omega \text{ for } M, \quad \tilde{\Omega} - \text{open in } \Omega \times \mathbb{R}^m \tag{4.3}$$

we would like

$$\tilde{\chi}^{-1}(\tilde{\Omega}) \subset TM$$

to be open in M. We denote by $\mathcal{B}$ the collection of subsets of TM of this type. We will declare a subset $D \subset TM$ to be open if for each $(p, v_p) \in D$ there exists $B \in \mathcal{B}$ such that

$$(p, v_p) \in B \subset D.$$

Of course, what is happening here is that $\mathcal{B}$ is a topology basis, and we consider the associated topology (hence one can also say that a subset of TM is open iff it is an union of members of $\mathcal{B}$).

It should be clear that the resulting topology is Hausdorff. For 2nd countability, note that, in (4.3), we can restrict ourselves to χ belonging to an atlas inducing the smooth structure on M and, for each χ, to $\tilde{\Omega}$ belonging to a basis of the Euclidean topology on $\Omega \times \mathbb{R}^m$. As long as the domains of the charts from $\mathcal{A}$ for a basis for the topology of M, the resulting $\mathcal{B}$ would still be a topology basis for the topology we defined on TM. Since M and the $\Omega \times \mathbb{R}^m$s are 2nd countable, we are able to produce a countable basis for the topology on TM.

The fact that a set-theoretical vector field is smooth if and only if it is smooth as a map $X : M \to TM$ should be clear (one just has to check it locally, and then it is really the definition of the smoothness of X). The uniqueness is left as an exercise for the interested student. □

Note the object that results from this discussion: TM. It is now a manifold that relates to M by an obvious "projection map"

$$\pi : TM \to M, \quad (p, v_p) \mapsto p.$$

And the fibers $\pi^{-1}(p)$ are vector spaces (precisely the tangent spaces of M). This is precisely what vector bundles are- but we will return to them later on.

4.1. Some interesting questions/phenomena with vector fields. (... for the curious students) Here are some very simple questions about vector fields whose answer is non-trivial question but very exciting. Let us start from the "hairy ball theorem" which can be seen as an interesting property of the two-sphere S^2: it does not admit a nowhere vanishing vector field. The first question in our list is now obvious:

Question: *Which manifolds admit a non-where vanishing vector field?*

It is worth looking at this question even for the spheres $S^m \subset \mathbb{R}^{m+1}$. Then the question becomes pretty elementary: does there exist a smooth function

$$X : S^m \to \mathbb{R}^{m+1} \setminus \{0\}$$

with the property that

$$x_0 \cdot X_0(x) + x_1 \cdot X_1(X) + \ldots + X_m \cdot X_m(x) = 0 \quad \forall\, x = (x_0, x_1, \ldots, x_m) \in S^{m+1}?$$

For S^1 the answer is clearly yes (think on the picture!) and, as we have already mentioned, for S^2 the answer is no. What about S^3? Well, using quaternions as in Example 3.34 we see that the answer is yes. More precisely, viewing $S^3 \subset \mathbb{H} \cong \mathbb{R}^4$, the multiplication by i, j and k provide three nowhere vanishing (and even linearly independent!) vector fields on S^3:

$$X^i(x) = i \cdot x, \quad X^j(x) = j \cdot x, \quad X^k(x) = k \cdot x.$$

(why are these vector fields on spheres?).

It turns out that, among the spheres, it is precisely the odd dimensional ones that do admit such vector fields.

For general manifolds the answer to the previous question is closely related to the topology of M- namely to the Euler number $\chi(M)$ of M. We will briefly discuss the Euler number later on when we will introduce DeRham cohomology and we will see that $\chi(S^m) = 1 + (-1)^m$. For now, let us mention that this number is, in principle, easy to compute using a triangulation of M, by the Euler formula

$$\chi(M) = \#\,\text{vertices} - \#\,\text{edges} + \#\,\text{faces} - \ldots.$$

(using this you should convince yourself that, indeed, $\chi(S^m) = 1 + (-1)^m$). The answer to the previous question is: a compact manifold M admits a no-where vanishing vector field if and only $\chi(M) = 0$!

What about more than one no-where vanishing vector fields? Of course, the question is interesting only if we look for linearly independent ones- and then the number of such vector fields will be bounded from above by the dimension of the manifold. E.g., on S^3, there are at most three such; and the vector field X^i, X^j and X^k described above show that the upper bound can be achieved. More generally:

Question: *An m-dimensional manifold M is said to be parallelizable if it admits m vector fields which, at each point, are linearly independent.*

For instance, we already know that S^1 and S^3 are parallelizable. One possible explanation for this is the fact that both S^1 as well as S^3 are Lie groups (and they are the only spheres that can be made into Lie groups- see Example 3.34 from Chapter 3). More precisely:

EXERCISE 4.17. *Show that all Lie groups are parallelizable. Here is how to do it. Let G be a k-dimensional Lie group. Consider its tangent space at the identity, $\mathfrak{g} := T_eG$ (a k-dimensional vector space). For $v \in \mathfrak{g}$, we define the vector field $\vec{v}$ on G by*

$$\vec{v}_g := (dL_g)_e(v)$$

where $L_g : G \to G$ is given by $L_g(h) = gh$ and $(dL_g)_e : \mathfrak{g} \to T_gG$ is its differential.

Show that if $\{v_1, \ldots, v_k\}$ is a basis of the vector space $\mathfrak{g}$, then $\vec{v}_1, \ldots, \vec{v}_k$ are k-linearly independent vector fields on G.

While the only spheres that can be made into Lie groups are S^0, S^1 and S^3, are there any other spheres that are parallelizable? Can one reproduce the argument that we used for S^3 so that it applies to other spheres S^m? We see that we need some sort of "product operation" on $\mathbb{R}^{m+1}$, so that we can define vector fields X^l by

$$X^l(x) = e_l \cdot x, \quad l \in \{1, \ldots, m\}$$

where $\{e_0, e_1, \ldots, e_m\}$ is the canonical basis. We end up again with the question mentioned in Example 3.34 from Chapter 3, of whether $\mathbb{R}^{m+1}$ can be made into a normed division algebra. And, as we mentioned there, the only possibilities were $\mathbb{R}, \mathbb{R}^2$ and R^4 for associative products, and also $\mathbb{R}^8$(octonions) in full generality. However, while the failure of associativity was a problem in making the corresponding sphere into a Lie group, it is not difficult to see that it does not pose a problem in producing the desired vector fields. Therefore: yes, the arguments that used $\mathbb{H}$ works also for octonions, hence also S^7 is parallelizable.

And, as you may expect by now, it turns out that S^0, S^1, S^3 and S^7 are the only spheres that are parallelizable!

So far we have looked at the extreme cases: one single nowhere vanishing vector field, or the maximal number of linearly independent vector fields. Of course, the most general question is:

Question: *Given a manifold M, what is the maximal number of linearly independent vector fields that one can find on M?*

Already for the case of the spheres S^m, this deceivingly simple question turns out to be highly non-trivial. The solution (found half way in the 20th century) requires again the machinery of

Algebraic Topology. But here is the answer: write $m+1 = 2^{4a+r}m'$ with m' odd, $a \geq 0$ integer, $r \in \{0,1,2,3\}$. Then the maximal number of linearly independent vector fields on S^m is

$$r(m) = 8a + 2^r - 1.$$

Note that $r(m) = 0$ if m is even- and that corresponds to the fact that there are no nowhere vanishing vector fields on even dimensional spheres. The parallelizability of S^m is equivalent to $r(m) = m$, i.e. $8a + 2^r = 2^{4a+r}m'$, which is easily seen to have the only solutions

$$m = 1, \quad a = 0, \quad r \in \{0,1,2,3\},$$

giving $m = 0, 2, 4, 8$. Hence again the spheres S^0, S^1, S^3 and S^7.

4.2. Push-forwards of vector fields. We now discuss how vector fields can be pushed forward from one manifold to another. The best scenario is when we start with a diffeomorphism $F : M \to N$ between two manifolds; then induces a push-forward operation:

$$F_* : \mathfrak{X}(M) \to \mathfrak{X}(N), \quad X \mapsto F_*(X)$$

where, for $X \in \mathfrak{X}(M)$, $F_*(X) \in \mathfrak{X}(N)$ is defined by describing how it acts on functions $f \in \mathcal{C}^\infty(N)$: one first moves f to N via F, i.e. consider $f \circ F \in \mathcal{C}^\infty(M)$, then one acts by X, i.e. consider $L_X(f \circ F) \in \mathcal{C}^\infty(M)$, then one moves back to N to obtain $L_X(f \circ F) \circ F^{-1} \in \mathcal{C}^\infty(N)$. One checks right away that this operation on X is a derivation, hence there exists a unique vector field $F_*(X) \in \mathfrak{X}(N)$ such that

$$L_{F_*(X)}(f) = L_X(f \circ F) \circ F^{-1} \quad \text{for all } f \in \mathcal{C}^\infty(N).$$

A more explicit description of $F_*(X)$ (that gives $F_*(X)_q \in T_qN$ for each $q \in N$) is obtained using the formula $L_Y(g)(p) = (dg)_p(Y_p)$ for vector fields Y and smooth functions g.

Exercise 4.18. Deduce that, for all $q \in N$,

$$F_*(X)_q = (dF)_p(X_p) \quad \text{where } p = F^{-1}(q). \tag{4.4}$$

Also show that, if $G : N \to P$ is another diffeomorphism, then $(G \circ F)_* = G_* \circ F_*$.

Proposition 4.18. *The push forward operation preserves the Lie bracket of vector fields: for any $X_1, X_2 \in \mathfrak{X}(M)$, one has*

$$F_*([X_1, X_2]) = [F_*(X_1), F_*)(X_2)].$$

Proof. We start with the definition of $[F_*(X_1), F_*)(X_2)]$:

$$L_{[F_*(X_1),F_*)(X_2)]}(f) = L_{F_*(X_1)}\left(L_{F_*)(X_2)}(f)\right) - L_{F_*(X_2)}\left(L_{F_*)(X_1)}(f)\right)$$

(for all $f \in \mathcal{C}^\infty(N)$, in which we plug in the definition of $F_*(X_i)$ to obtain

$$L_{F_*(X_1)}\left(L_{X_2}(f \circ F) \circ F^{-1}\right) - \text{ the similar one} = L_{X_1}\left(L_{X_2}(f \circ F)\right) \circ F - \text{ the similar one ;}$$

using again the formula defining $[X_1, X_2]$, and then the one for F_*, the last expression is

$$L_{[X_1,X_2]}(f \circ F) \circ F^{-1} = L_{F_*([X_1,X_2])}(f).$$

□

Exercise 4.19. By similar arguments defined also a pull-back operation $F^* : \mathfrak{X}(N) \to \mathfrak{X}(M)$. And show that this makes sense not only for diffeomorphisms $F : M \to N$, but for all local diffeomorphisms. Also: is it true that $(G \circ F)^* = G^* \circ F^*$?

The construction of push-forwards cannot be extended for arbitrary smooth maps $F : M \to N$. Instead, one can talk about a vector field $X \in XX(M)$ being "related via F" to a vector field $Y \in \mathfrak{X}(N)$:

DEFINITION 4.19. *Given a smooth map $F : M \to N$, we say that $X \in \mathfrak{X}(M)$ is **F-projectable** to $Y \in \mathfrak{X}(N)$ if*

$$(dF)_p(X_p) = Y_{F(p)} \quad \text{for all } p \in M.$$

Note that, when F is a diffeomorphism, this corresponds precisely to $Y = F_*(X)$. However, in general, there may be different vector fields on M that are F-projectable to the same vector field on N (think e.g. of the case when N is a point) or there may be vector fields on M that are not projectable to any vector field on N. However, Proposition 4.18 still has a version that applies to this general setting:

EXERCISE 4.20. *For any smooth map $F : M \to N$, if $X_i \in \mathfrak{X}(M)$ are F-projectable to $Y_i \in \mathfrak{X}(N)$ for $i \in \{1, 2\}$, then $[X_1, X_2]$ is F-projectable to $[Y_1, Y_2]$. (Hint: proceed locally and remember the formula that you obtained when you solved Exercise 4.14).*

5. The Lie algebra of a Lie group

We now discuss the infinitesimal counterpart of Lie groups.

DEFINITION 4.20. *A Lie algebra is a vector space $\mathfrak{a}$ endowed with an operation*

$$[\cdot, \cdot] : \mathfrak{a} \times \mathfrak{a} \to \mathfrak{a}$$

which is bi-linear, antisymmetric and which satisfies the Jacobi identity:

$$[[u, v], w] + [[v, w], u] + [[w, u], v = 0 \quad \text{for all } u, v, w \in \mathfrak{a}.$$

EXAMPLE 4.21 (commutators of matrices). The space $\mathcal{M}_n(\mathbb{R})$ of $n \times n$ matrices, together with the commutator of matrices,

$$[A, B] := AB - BA$$

is a Lie algebra. More generally, for any vector space V, the space $\mathrm{Lin}(V, V)$ of linear maps from V can be made into a Lie algebra by the same formula (just that AB becomes the composition of A and B).

EXAMPLE 4.22 (vector fields). As we have already pointed out, for any manifold M, the space of vector fields $\mathfrak{X}(M)$ endowed with the Lie bracket of vector fields becomes a (infinite dimensional) Lie algebra.

EXERCISE 4.21. What is the relationship between the previous two examples?

We now explain how a Lie group G gives rise to a finite dimensional Lie algebra $\mathfrak{g}$. As a linear space, it is defined as the tangent space of G at the identity element $e \in G$:

$$\mathfrak{g} := T_eG.$$

For the Lie bracket, we will related $\mathfrak{g}$ to vector fields on G, and then use the Lie bracket of vector fields. Hence, as a first step, we note that any $v \in \mathfrak{g}$ gives rise to a vector field

$$\vec{v} \in \mathfrak{X}(G).$$

defined as follows. For $g \in G$, to define $\vec{v}_g \in T_gG$, we use the left translation by g:

$$L_g : G \to G, \quad L_g(h) = gh.$$

This is smooth (and even a diffeomorphism!) and make use of its differential at e, $(dL_g)_e : \mathfrak{g} \to T_gG$. Define then

$$\vec{v}_g := (dL_g)_e(v) \in T_gG.$$

EXERCISE 4.22. Show that, indeed, $\vec{v}$ is smooth.

PROPOSITION 4.23. *With the same notations as above (a Lie group G and $\mathfrak{g} := T_eG$), for any $u, v \in \mathfrak{g}$, the Lie bracket $[\overrightarrow{u}, \overrightarrow{v}] \in \mathfrak{X}(G)$ is again of type $\overrightarrow{w}$ for some $w \in \mathfrak{g}$. Denoting the resulting w by $[v, w] \in \mathfrak{g}$, we obtain an operation*

$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

that makes $\mathfrak{g}$ into a Lie algebra.

PROOF. Note that the map

$$\mathrm{j} : \mathfrak{g} \to \mathfrak{X}(G), \quad v \mapsto \overrightarrow{v} \tag{5.1}$$

is injective. That is simply because $\overrightarrow{v}_e = v$. This implies the uniqueness of w; and it also indicates how to define it:

$$w := [\overrightarrow{u}, \overrightarrow{v}]_e.$$

We still have to show that $[\overrightarrow{u}, \overrightarrow{v}] = \overrightarrow{w}$. For that we note that

$$(L_a)_*(\overrightarrow{v}) = \overrightarrow{v}$$

for all $a \in G$ and all $v \in \mathfrak{g}$, where we use the push-forward as defined in subsection 4.2. Indeed, using the formula (4.4) for the push-forward and then the definition of $\overrightarrow{v}$:

$$(L_a)_*(\overrightarrow{v})_g = (dL_a)_{a^{-1}g}\left((dL_{a^{-1}g})_e(v)\right)$$

and then, using the chain rule and $L_a \circ L_{a^{-1}g} = L_g$, we end up with $\overrightarrow{v}_g$.

Since $(L_a)_*(\overrightarrow{v}) = \overrightarrow{v}$ and similarly for u, Proposition 4.18 implies that $[\overrightarrow{u}, \overrightarrow{v}] = (L_a)_*([\overrightarrow{u}, \overrightarrow{v}])$. Evaluating this at $a \in G$ and using again formula formula (4.4) (where $F = L_a, q = a$ so that $p = L_a^{-1}(a) = e$):

$$[\overrightarrow{u}, \overrightarrow{u}]_a = (dL_a)_e([\overrightarrow{u}, \overrightarrow{u}]_e) = (dL_a)_e(w_e) = \overrightarrow{w}_a\ .$$

To fact that the resulting operation $[\cdot, \cdot]$ on $\mathfrak{g}$ is a bilinear and skew-symmetric is immediate. For the Jacobi identity, since the map j from (5.1) is injective and satisfies $j([v, w]) = [j(v), j(w)]$, it suffices to remember that the Lie bracket of vector fields does satisfy the Jacobi identity. □

DEFINITION 4.24. ***The Lie algebra of the Lie group*** *G is defined as $\mathfrak{g} = T_eG$ endowed with the Lie algebra structure from the previous proposition.*

We now compute some examples.

EXAMPLE 4.25. *We start with the general linear group $GL_n(\mathbb{R})$, and we claim that its Lie algebra is just the algebra $\mathcal{M}_n(\mathbb{R})$ of $n \times n$ matrices endowed with the commutator bracket.*

PROOF. As we already mentioned, $\mathcal{M}_n(\mathbb{R})$ is seen as an Euclidean space; we denote coordinates functions by

$$x^i_j : \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$$

(sending a matrix to the element on the position (i, j)). We also remarked already that, while $GL_n = GL_n(\mathbb{R})$ sits openly inside $\mathcal{M}_n(\mathbb{R})$, its tangent space at the identity matrix I (and similarly at any point) is canonically identified with $\mathcal{M}_n(\mathbb{R})$: any $X \in \mathcal{M}_n(\mathbb{R})$ is identified with the speed at $t = 0$ of $t \mapsto (I + tX)$. After left translating, we find that the corresponding vector field $\overrightarrow{X} \in \mathfrak{X}(GL_n)$ is given, at an arbitrary point $A \in GL_n$, by

$$\overrightarrow{X}_A = \frac{d}{dt}\Big|_{|t=0} A \cdot (I + tX) \in T_A GL_n. \tag{5.2}$$

Let $X, Y \in \mathcal{M}_n(\mathbb{R})$ and let $[X, Y] \in \mathcal{M}_n(\mathbb{R})$ be the Lie algebra bracket- hence defined by $\overrightarrow{[X,Y]} = [\vec{X}, \vec{Y}]$, where the last bracket is the Lie bracket of vector fields on GL_n. To compute it, we use the defining equation for the Lie bracket of two vector fields:

$$L_{[\vec{X},\vec{Y}]}(F) = L_{\vec{X}} L_{\vec{Y}}(F) - L_{\vec{Y}} L_{\vec{X}}(F) \tag{5.3}$$

for all smooth functions

$$F : GL_n \to \mathbb{R}.$$

From the previous description of $\vec{X}$ we deduce that

$$L_{\vec{X}}(F)(A) = \frac{d}{dt}\Big|_{t=0} F(A \cdot (I + tX)) = \sum_{i,j,k} \frac{\partial F}{\partial x^i_j}(A)\, A^i_k X^k_j.$$

In particular, applied to a coordinate function $F = x^i_j$, we find

$$L_{\vec{X}}(x^i_j)(A) = \sum_k A^i_k X^k_j$$

or, equivalently, $L_{\vec{X}}(x^i_j) = \sum_k x^i_k X^k_j$. Therefore

$$L_{\vec{Y}}(L_{\vec{X}}(x^i_j)) = \sum_k L_{\vec{Y}}(x^i_k) X^k_j = \sum_{k,l} x^i_l Y^l_k X^k_j = \sum_l x^i_l (Y \cdot X)^l_j,$$

which is precisely $L_{\overrightarrow{Y \cdot X}}(x^i_j)$. Since this holds for all coordinate functions, we have

$$L_{\vec{Y}} \circ L_{\vec{X}} = L_{\overrightarrow{Y \cdot X}};$$

we deduce that $[\vec{X}, \vec{Y}] = \overrightarrow{Y \cdot X} - \overrightarrow{X \cdot Y}$ and then that $[X, Y]$ is $X \cdot Y - Y \cdot X$, i.e. the usual commutator of matrices. ☐

Example 4.26. In section 12 of Chapter 3 we have looked at several classical Lie groups $G \subset GL_n$ and we have computed their tangent spaces at the identity matrix I, as subspaces of $T_I GL_n = \mathcal{M}_n(\mathbb{R})$:

- for $O(n)$ and $SO(n)$ we obtained
$$o(n) = \{X \in \mathcal{M}_n(\mathbb{R}) : X + X^T = 0\}.$$
- for $SL_n(\mathbb{R})$ we obtained
$$sl_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) : Tr(X) = 0\}.$$
- $U(n)$ we obtained
$$u(n) = \{X \in \mathcal{M}_n(\mathbb{C}) : X + X^* = 0\}.$$
- for $SU(n)$ we obtained
$$su(n) = \{X \in \mathcal{M}_n(\mathbb{C}) : X + X^* = 0, Tr(X) = 0\}.$$

We claim that the resulting Lie brackets on these subspaces of $\mathcal{M}_n(\mathbb{R})$ are still given by the commutators of matrices (just for fun, you should check that, indeed, all these subspaces are closed under taking commutators).

Of course, all these examples can be treated at once: given $G \subset GL_n$ which is a Lie subgroup (i.e. a subgroup and an embedded sub manifold), the resulting Lie algebra structure on $\mathfrak{g} = T_I G \subset T_I GL_n = \mathcal{M}_n(\mathbb{R})$ is just the commutator bracket.

PROOF. Note that for $X \in \mathfrak{g}$ we have two induced vector fields: one on G, denoted as above by $\vec{X} \in \mathfrak{X}(G)$ and then, since $\mathfrak{g} \subset \mathcal{M}_n(\mathbb{R})$ we will have a similar one on GL_n; to distinguish the two we will denote the second one by $\overset{\rightarrowtail}{X} \in \mathfrak{X}(GL_n)$. Note that, denoting by $i : G \hookrightarrow GL_n$ the inclusion, we are in the situation of Definition 4.19: $\vec{X}$ is i-projectable to $\overset{\rightarrowtail}{X}$. Hence, by Exercise 4.20, $[\vec{X}, \vec{Y}]$ is i-projectable to $[\overset{\rightarrowtail}{X}, \overset{\rightarrowtail}{Y}]$ (for all $X, Y \in \mathfrak{g}$). By the previous example, the last expression is $\overset{\rightarrowtail}{[X, Y]_{\text{comm}}}$ where, for clarity, we denote by $[X, Y]_{\text{comm}}$ the commutator bracket of the matrices X and Y. Hence, for the Lie bracket of X and Y as elements of the Lie algebra $\mathfrak{g}$ of G we find

$$[X, Y] = [\overset{\rightarrowtail}{X}, \overset{\rightarrowtail}{Y}]_I = \left(\overset{\rightarrowtail}{[X, Y]_{\text{comm}}}\right)_I = [X, Y]_{\text{comm}}.$$

□

REMARK 4.27 (for the extra-curious students). The Lie algebra $\mathfrak{g}$ of a Lie group G contains (almost) all the information about G! This may sound surprising at first, since $\mathfrak{g}$ is only "the linear approximation" of G and only around the identity $e \in G$. The fact that it is enough to consider only the identity e can be explained by the fact that, using the group structure (and the resulting left translations) one can move around from e to any other element in g. The fact that "the linear approximation" contains (almost) all the information is more subtle and the real content of this is to be found on the Lie bracket structure of $\mathfrak{g}$.

And here is what happens precisely:

- As we have seen, any Lie group has an associated (finite dimensional) Lie algebra.
- Any finite dimensional Lie algebra $\mathfrak{g}$ comes from a Lie group.
- For any finite dimensional Lie algebra $\mathfrak{g}$ there exists and is unique a Lie group $G(\mathfrak{g})$ whose Lie algebra is $\mathfrak{g}$ and which is 1-connected (i.e. both connected as well as simply-connected). The Lie group $G(\mathfrak{g})$ can be constructed explicitly out of $\mathfrak{g}$. And any other connected Lie group which has $\mathfrak{g}$ as Lie algebra is a quotient of $G(\mathfrak{g})$ by a discrete subgroup.
- Explanation for the 1-connectedness condition: starting with a Lie group G, the connected component of the identity element, is itself a Lie group with the same Lie algebra as G. And so is the universal cover of G. I.e. one can always replace G by a Lie group that is 1-connected without changing its Lie algebra.
- the construction $G \mapsto \mathfrak{g}$ gives an equivalence between (the category of) 1-connected Lie groups and (the category of) finite dimensional Lie algebras. At the level of morphisms even more is true: while a morphism of Lie groups $F : G \to H$ induces a morphism of the corresponding Lie algebras, $f : \mathfrak{g} \to \mathfrak{h}$, if G is 1-connected (but H is arbitrary), then F can be recovered from f (i.e.: any morphism f of Lie algebras comes from an unique morphism F of Lie groups!). That is remarkable since f is just the linearization of F at the identity $f = (dF)_e$.

6. Integral curves (and flows)

The brief philosophy is:

If your world is a manifold M and your are sitting at a point p, then a tangent vector $X_p \in T_pM$ gives you a direction in which you can start walking. A vector field however gives you an entire path to walk on.

In the ideal situation (no friction etc, no opposition from your side), and if you think of a vector field as a wind, its integral curve at p is the trajectory that you will follow, blown by the wind.

DEFINITION 4.28. *Given a vector field $X \in \mathfrak{X}(M)$,* ***an integral curve of*** *X is any curve $\gamma : I \to M$ defined on some open interval $I \subset \mathbb{R}$ ($I = \mathbb{R}$ not excluded) such that*

$$\frac{d\gamma}{dt}(t) = X_{\gamma(t)} \quad \text{for all } t \in I.$$

We say that γ ***starts at*** *p if $0 \in I$ and $\gamma(0) = p$.*

EXAMPLE 4.29. Consider $M = \mathbb{R}^2$ and the vector field X given by

$$X = \frac{\partial}{\partial x} + \frac{\partial}{\partial y}.$$

Then a curve in $\mathbb{R}^2$, written as $\gamma(t) = (x(t), y(t))$, is an integral curve if and only if

$$\dot{x}(t) = 1, \quad \dot{y}(t) = 1.$$

Hence we find that the integral curves of X are those of type

$$\gamma(t) = (t + a, t + b)$$

with $a, b \in \mathbb{R}$ constants. Prescribing the starting point of γ is the same thing as fixing the constants a and b (and then we have an unique integral curve).

A more interesting vector field is:

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

Then the equations to solve are $\dot{x} = y$ and $\dot{y} = -x$. Hence $\ddot{x} = -x$ from which one can deduce (or guess, and then use the uniqueness recalled below) that

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

for some (arbitrary) constants $a, b \in \mathbb{R}$ (and, again, prescribing the starting point of γ is the same thing as fixing the constants a and b).

Note that in both examples the integral curves that we were obtaining are defined on the entire $I = \mathbb{R}$. However, this need not always be the case. For instance, for

$$X = -x^2\frac{\partial}{\partial x} - y\frac{\partial}{\partial y},$$

starting with any constants $a, b \in \mathbb{R}$ one has the integral curve given by

$$\gamma(t) = (\frac{a}{at+1},\ be^{-t});$$

but, for $a > 0$, the largest interval containing 0 on which this is defined is $(-\frac{1}{a}, \infty)$. However, still as above, for any point $p = (a, b) \in \mathbb{R}^2$, one finds an integral curve that starts at p.

Of course, in the previous examples, the existence of integral curves starting at a given (arbitrary) point is no accident:

PROPOSITION 4.30 (local existence and uniqueness). *Given a vector field $X \in \mathfrak{X}(M)$:*

1. *for any $p \in M$ there exists an integral curve $\gamma : I \to M$ of X that starts at p.*
2. *any two integral curves of X that coincide at a certain time t_0 must coincide in a neighborhood of t_0.*

Since this is a local statement, we just have to look in a chart. Then, as we shall see, what we end up with is a system of ODEs and the previous proposition becomes the standard local existence/uniqueness result for ODEs. Here are the details. Fix a chart $\chi : U \to \Omega \subset M$ of M around p. Then, as in (4.1), X can be written on U as

$$X_p = \sum_{i=1}^{m} X_\chi^i(\chi(p)) \left(\frac{\partial}{\partial \chi_i}\right)_p$$

with the coordinate functions X_χ^i of X w.r.t. χ being smooth functions on Ω. Also, a curve γ in U becomes, via χ, a curve γ_χ in Ω. Since

$$\gamma(t) = \chi^{-1}(\gamma_\chi(t)),$$

its derivatives becomes

$$\frac{d\gamma}{dt}(t) = \frac{d\gamma_\chi^i}{dt}(t) \left(\frac{\partial}{\partial \chi_i}\right)_{\gamma(t)}.$$

Hence the integral curve condition becomes the system of ODEs:

$$\frac{d\gamma_\chi^i}{dt}(t) = X_\chi^i(\gamma_\chi^1(t), \dots, \gamma_\chi^m(t)) \quad i \in \{1, \dots, m\}$$

or, more compactly,

$$\frac{d\gamma_\chi}{dt}(t) = X_\chi(\gamma_\chi(t)),$$

where $X^\chi = (X_\chi^1, \dots, X_\chi^m) : \Omega \to \mathbb{R}^m$. And here is the standard result on such ODEs:

THEOREM 4.31. *Let $F : \Omega \to \mathbb{R}^m$ be a smooth function. Then, for any $x \in \Omega$, the following ordinary differential equation with initial condition:*

$$\frac{d\gamma}{dt}(t) = F(\gamma(t)), \quad \gamma(0) = x \tag{6.1}$$

has a solution γ defined on an open interval containing $0 \in \mathbb{R}$, and any two such solutions must coincide in a neighborhood of 0.

Furthermore, for any $x_0 \in \Omega$ there exists an open neighborhood Ω_0 of x_0 in Ω, $\epsilon > 0$ and a smooth map

$$\phi : (-\epsilon, \epsilon) \times \Omega_{x_0} \to \Omega,$$

such that, for any $x \in \Omega_0$, $\phi(\cdot, x) : (-\epsilon, \epsilon) \to \Omega$ is a solution of (6.1).

We see that this the first part of the theorem immediately implies (well, it is actually the same as) the previous proposition. For a more global version of the proposition, we have to look at maximal integral curves.

DEFINITION 4.32. *Given a vector field $X \in \mathfrak{X}(M)$, a* ***maximal integral curve*** *of X is any integral curve $\gamma : I \to M$ which admits no extension to a strictly larger interval $\tilde{I}$ and which is still an integral curve of X.*

With this, the local result implies quite easily the following one, with a more global flavour.

COROLLARY 4.33. *Given a vector field $X \in \mathfrak{X}(M)$, for any $p \in M$ there exists an unique maximal integral curve of X,*

$$\gamma_p : I_p \to M,$$

that starts at p.

PROOF. The main remark is that if γ_1 and γ_2 are two integral curves defined on intervals I_1 and I_2, respectively, then the set I of points in $I_1 \cap I_2$ on which the two coincide is closed in $I_1 \cap I_2$ (because it is defined by an equation) and open by the second part of the previous proposition. Hence, since $I_1 \cap I_2$ is connected, if $I \neq \emptyset$ (i.e. the two coincide at some t), then $I = I_1 \cap I_2$ (i.e. the two must coincide on their common domain of definition). Therefore, to obtain I_p and γ_p we can just put together all the integral curves that start at p. $\square$

The best possible scenario is:

DEFINITION 4.34. *We say that $X \in \mathfrak{X}(M)$ is **complete** if all the maximal integral curves are defined on the entire $\mathbb{R}$.*

As we shall see below, all the vector fields that are compactly supported (i.e. which vanish outside a compact subset of M) are complete. In particular, when M is compact, all vector fields on M are complete.

We now put together all the maximal integral curves in one object: the flow of X. First of all, we define the domain of the flow of X as

$$\mathcal{D}(X) := \{(p,t) \in M \times \mathbb{R} : t \in I_p\} \subset M \times \mathbb{R}$$

and then **the flow of** X is defined as the resulting map

$$\phi_X : \mathcal{D}(X) \to M, \quad \phi_X(p,t) := \gamma_p(t).$$

For complete vector fields one has $\mathcal{D}(X) = M \times \mathbb{R}$. What we can say in general is that

$$M \times \{0\} \subset \mathcal{D}(X) \subset M \times \mathbb{R}.$$

and $\mathcal{D}(X)$ is open in $M \times \mathbb{R}$. Actually, applying the second part of the Theorem 4.31, we deduce:

COROLLARY 4.35. *For any $X \in \mathfrak{X}(M)$, the domain $\mathcal{D}(X)$ of its flow is open in $M \times \mathbb{R}$ and the flow ϕ_X is a smooth map.*

One of the main uses of the flow of a vector field X comes from the maps one obtains whenever one fixes a time t; we will use the notation:

$$\phi_X^t(\cdot) := \phi_X(\cdot, t)$$

(whenever the right hand side is defined). The best scenario is when X is complete, when each ϕ^t will be defined on the entire M:

THEOREM 4.36. *If $X \in \mathfrak{X}(M)$ is complete then $\phi_X^t : M \times M$ satisfy:*

$$\phi_X^t \circ \phi_X^s = \phi_X^{t+s}, \quad \phi_X^0 = Id$$

(for all $t, s \in \mathbb{R}$). In particular, each ϕ_X^t is a diffeomorphism and

$$\left(\phi_X^t\right)^{-1} = \phi_X^{-t}.$$

PROOF. Fixing $x \in M$ and $s \in \mathbb{R}$, the two curves

$$t \mapsto \ \phi_X^{t+s}(p), \quad t \mapsto \phi_X^t(\phi_X^s(p))$$

are both integral curves of X and they coincide at $t = 0$- hence they coincide everywhere. That ϕ_X^0 is the identity is clear; we see that the last part follows by taking $s = -t$ in the first part. $\square$

For general (possibly non-complete) vector fields one has to be a bit more careful.

DEFINITION 4.37. *For $X \in \mathfrak{X}(M)$ and $t \in \mathbb{R}$, define **the flow of** X **at time** t as the map*

$$\phi_X^t : \mathcal{D}_t(X) \to M, \quad p \mapsto \phi_X^t(p) := \phi_X(p,t) = \gamma_p(t),$$

defined on:

$$\mathcal{D}_t(X) := \{p \in M : t \in I_p\}.$$

Theorem 4.38. *For any* $X \in \mathfrak{X}(M)$*:*

(1) *for any* $t \in \mathbb{R}$*, the domain* $\mathcal{D}_t(X)$ *of* ϕ_X^t *is open in* M*,* ϕ_X^t *takes values in* $\mathcal{D}_{-t}(X)$*, and*
$$\phi_X^t : \mathcal{D}_t(X) \to \mathcal{D}_{-t}(X)$$
is a diffeomorphism.

(2) *For* $t, s \in \mathbb{R}$ *one has*
$$\phi^t \circ \phi^s = \phi^{t+s}$$
in the sense that, for each $p \in M$ *on which the left hand side is defined, also* $\phi^{t+s}(p)$ *is defined and the two expressions coincide.*

Proof. The fact that $\mathcal{D}_t(X)$ is open and ϕ_X^t is smooth follows right away from the previous the standard local result (Theorem 4.31). For the last part we consider again the two curves from the proof of Theorem 4.36; this time we have to be more careful with their domain of definition. Hence we fix $p \in M$, $s \in I(p)$ and we look at the curves:
$$a : -s + I(p) \to M, \quad a(t) = \phi_X^{t+s}(p),$$
$$b : I(\phi_X^s(p)) \to M, \quad b(t) = \phi_X^t(\phi_X^s(p)).$$
Note that both a and b are integral curves of X, and their are both maximal (the second one is maximal from its very definition; for the first one, if $\tilde{a}$ was an extension to $J \supset -s + I(p)$, then $s + J \ni t \mapsto \tilde{a}(t - s)$ would be an integral curve defined on $s + J \supset I(p)$, and starting at p). We then obtain the last part of the theorem together with the equality
$$I(\phi_X^s(p)) = I(p) - s$$
for all s. Since $0 \in I(p)$ we find that $-s \in I(\phi_X^s(p))$, hence
$$\phi_X^s(\mathcal{D}_s) \subset \mathcal{D}_{-s}$$
and, for $p \in \mathcal{D}_s$, $\phi_X^{-s}(\phi_X^s(p)) = p$. We still have to show that, in the previously centered inclusion, equality holds. Using the inclusion for $-s$ instead of s and applying ϕ_X^s we obtain
$$\phi_X^s\left(\phi_X^{-s}(\mathcal{D}_{-s})\right) \subset (\phi_X^s(\mathcal{D}_s);$$
but the first term equals to $\phi_X^0(\mathcal{D}_s) = \mathcal{D}_s$, hence we obtain the reverse inclusion and this finishes the proof. □

Theorem 4.39. *If* M *is compact then any vector field* $X \in \mathfrak{X}(M)$ *is complete.*

More generally, for any M*, any* $X \in \mathfrak{X}(M)$ *that is compactly supported (i.e. is zero outside some compact subset of* M*) is complete.*

Proof. When M is compact, since $\mathcal{D}(X)$ is an open in $M \times \mathbb{R}$ containing $M \times \{0\}$, the tube lemma implies that there exists $\epsilon > 0$ such that:

(6.2)
$$M \times (-\epsilon, \epsilon) \subset \mathcal{D}(X).$$

(proof: for any $p \in M$, using that ϕ_X is continuous at $(p, 0)$ and that $\mathcal{D}(X)$ is open, find $\epsilon_p > 0$ and U_p open containing p with $U_p \times (-\epsilon_p, \epsilon_p) \subset \mathcal{D}(X)$. Then $\{U_p : p \in M\}$ is an open cover of M hence, by compactness, M is covered by a finite number of them, say the ones corresponding to $p_1, \ldots, p_k \in M$; set ϵ as the smallest of the epsilons corresponding to the points p_i.)

We claim that the existence of ϵ such that the inclusion (6.2) holds implies (without further using the compactness of M!) that X is complete. Indeed, we would have that $\phi_X^t(p)$ is defined for all $t \in (-\epsilon, \epsilon)$ and all $p \in M$. But then so would be $\phi_X^t(\phi_X^t(p))$, hence $\phi_X^{2t}(p)$, hence (6.2) holds also with 2ϵ instead of ϵ. Repeating the process, we find that $\mathbb{R} \times M \subset \mathcal{D}(X)$, hence X must be complete.

For the last part (M general and X is compactly supported) it is enough to remark that an inclusion of type (6.2) can still be achieved: if X is zero outside the compact K then the tube lemma implies that $(-\epsilon, \epsilon) \times K \subset \mathcal{D}(X)$ for some $\epsilon > 0$, while all the points (t, p) with $p \notin K$ are clearly in $\mathcal{D}(X)$ since X vanishes at p (hence the integral curve of X starting at p is simply $\gamma(t) = p$, defined for all t). □

For latter reference, let us also cast the last part of the argument into:

COROLLARY 4.40. *If $X \in \mathfrak{X}(M)$ has the property that $M \times (-\epsilon, \epsilon) \subset \mathcal{D}(X)$ for some $\epsilon > 0$, then X is complete.*

EXERCISE 4.23. Let $F : M \to N$ be a smooth function. Recall that $X \in \mathfrak{X}(M)$ is said to be F-projectable to $Y \in \mathfrak{X}(N)$ if

$$(dF)_p(X_p) = Y_{F(p)} \quad \text{for all } p \in M.$$

If X is F-projectable to Y show that, for any $t \in \mathbb{R}$ one has

$$F(\mathcal{D}_t(X)) \subset \mathcal{D}_t(Y)$$

and, one $\mathcal{D}_t(X)$,

$$F \circ \phi_X^t = \phi_Y^t \circ F.$$

7. Lie derivatives along vector fields

The general philosophy of taking Lie derivatives L_X along vector fields is rather simple:

for any "type of objects" ξ on manifolds M (functions, vector fields, etc etc), which are natural (in the sense that a diffeomorphism $\phi : M \to N$ allows one pull-back the objects from N to ones on M), the Lie derivative $\mathcal{L}_X(\xi)$ of ξ along X measures the variation of ξ along the flow of X:

$$\mathcal{L}_X(\xi) := \left.\frac{d}{dt}\right|_{t=0} (\phi_X^t)^*(\xi). \tag{7.1}$$

This is a very general principle that is applied over and over in Differential Geometry, depending on the "type of objects" one is interested in. Here we illustrate this principle for functions and vector fields, and compute the outcome.

For functions, there is an obvious pull-back operation associated to any smooth function $F : M \to N$ (diffeomorphism or not):

$$F^* : \mathcal{C}^\infty(N) \to \mathcal{C}^\infty(M), \quad F^*(f) = f \circ F.$$

Therefore, when $F = \phi_X^t : M \to M$ is the flow of a complete vector field, the guiding equation (7.1) for $\xi = f \in \mathcal{C}^\infty(M)$ makes sense pointwise as

$$\mathcal{L}_X(f)(p) := \left.\frac{d}{dt}\right|_{t=0} f(\phi_X^t(p)). \tag{7.2}$$

Even more, written in this way, the completeness of X is not even necessary for this formula to make sense (indeed, we only need $\phi_X^t(p)$ to be defined for t near 0, which is always the case).

PROPOSITION 4.41. *For $X \in \mathfrak{X}(M)$ and $f \in \mathcal{C}^\infty(M)$, $\mathcal{L}_X(f)$ defined by (7.2) is precisely $L_X(f) \in \mathcal{C}^\infty(M)$.*

PROOF. The expression in the right hand side of (7.2) is precisely

$$(df)_p \left(\left.\frac{d}{dt}\right|_{t=0} \phi_X^t(p) \right) = (df)_p(X_p) = L_X(f)(p).$$

□

The discussion is similar for vector fields. In this case, a diffeomorphism $F : M \to N$ allows one to pull-back vector fields on N to vector fields on M,

$$F^* : \mathfrak{X}(N) \to \mathfrak{X}(M)$$

where, for $Y \in \mathfrak{X}(N)$, its pull-back $F^*(Y) \in \mathfrak{X}(M)$ is defined by

$$F^*(Y)_p := (dF^{-1})_{F(p)}(Y_{F(p)}).$$

With this, when $F = \phi_X^t : M \to M$ is the flow of a complete vector field, the guiding equation (7.1) for $\xi = Y \in \mathfrak{X}(M)$ makes sense pointwise as

$$\mathcal{L}_X(Y)(p) := \left.\frac{d}{dt}\right|_{t=0} (d\phi_X^{-t})_{\phi_X^t(p)}(Y_{\phi_X^t(p)}). \tag{7.3}$$

And, as for functions, this formula makes sense without any completeness assumption on X.

Proposition 4.42. *For $X \in \mathfrak{X}(M)$ and any other $Y \in \mathfrak{X}(M)$, $\mathcal{L}_X(f)$ defined by (7.3) is precisely the Lie bracket $[X, Y] \in \mathfrak{X}(M)$.*

Proof. Let $Z = \mathcal{L}_X(Y)$. For $f \in \mathcal{C}^\infty(M)$ we compute $L_Z(f)(p) = (df)_p(Z_p)$ and we find

$$\left.\frac{d}{dt}\right|_{t=0} (df)_p \left((d\phi_X^{-t})_{\phi_X^t(p)}(Y_{\phi_X^t(p)})\right) = \left.\frac{d}{dt}\right|_{t=0} d\left(f \circ \phi_X^{-t}\right)_{\phi_X^t(p)} \left(Y_{\phi_X^t(p)}\right).$$

Note that $f \circ \Phi_X^{-t}$ is f at $t = 0$ and has the derivative w.r.t. t at $t = 0$ equal to $-L_X(f)$, hence

$$f \circ \Phi_X^{-t} = f - tL_X(f) + t^2 \cdot ?,$$

where ? is smooth. Continuing the computation started above, we find

$$\left.\frac{d}{dt}\right|_{t=0} (df - td\,(L_X(f)))_{\phi_X^t(p)} \left(Y_{\phi_X^t(p)}\right) = \left.\frac{d}{dt}\right|_{t=0} (df)_{\phi_X^t(p)} \left(Y_{\phi_X^t(p)}\right) - (dL_X(f))\,(Y_p) =$$

$$\left.\frac{d}{dt}\right|_{t=0} L_Y(f)(\phi_X^t(p)) - L_Y(L_X(f))(p) = L_X(L_Y(f))(p) - L_Y(L_X(f))(p),$$

i.e. $L_Z(f)(p) = L_{[X,Y]}(f)(p)$ for all p and f. This implies that $Z = [X, Y]$. □

We obtain the following characterization of the commutation relation $[X, Y] = 0$; for simplicity, we restrict here to the case of complete vector fields.

Corollary 4.43. *For $X, Y \in \mathfrak{X}(M)$ complete, one has:*

$$[X, Y] = 0 \quad \Longleftrightarrow \quad \Phi_X^t \circ \Phi_Y^s = \Phi_Y^s \circ \Phi_X^t \quad \forall\ t, s \in \mathbb{R}.$$

Exercise 4.24. Using the previous Proposition and Exercise 4.23 prove again that if $F : M \to N$ is smooth and $X_1, X_2 \in \mathfrak{X}(M)$ are F-projectable to Y_1 and $Y_2 \in \mathfrak{X}(N)$, respectively, then $[X_1, X_2]$ is F-projectable to $[Y_1, Y_2]$.

8. Lie groups again: the exponential map

Proposition 4.44. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. For any $u \in \mathfrak{g}$, the vector field $\vec{u}$ is complete and its flow $\phi_{\vec{u}}^t$ satisfies*

$$\phi_{\vec{u}}^t(ag) = a\phi_{\vec{u}}^t(g).$$

In particular, the flow can be reconstructed from what it does at time $t = 1$, at the identity element of G, i.e. from

$$exp : \mathfrak{g} \to G, \quad exp(u) := \phi_{\vec{u}}^1(e),$$

by the formula

$$\phi^t_{\vec{u}}(a) = a\, exp(t\, u).$$

Moreover, the exponential is a local diffeomorphism around the origin: it sends some open neighborhood of the origin in $\mathfrak{g}$ diffeomorphically into an open neighborhood of the identity matrix in G.

PROOF. For the first part the key remark is that if $\gamma : I \to G$ is an integral curve of $\vec{u}$ defined on some interval I, i.e. if

$$\frac{d\gamma}{dt}(t) = \vec{u}\,(\gamma(t)) \quad \forall\, t \in I$$

then, for any $a \in G$, the left translate $a \cdot \gamma : t \mapsto a \cdot \gamma(t)$ is again an integral curve (of the same $\vec{u}$); this follows by writing $a \cdot \gamma = L_a \circ \gamma$ and using the chain rule. Hence, if γ_g is the maximal integral curve starting at $g \in G$ then $a \cdot \gamma_g$ will be the one starting at $a \cdot g$- and that proves the first identity (for as long as both terms are defined). And we also obtain that $I_g = I_{ag}$ for all $a, g \in G$, hence $I_g = I_e$ for all $g \in G$. Using Corollary 4.40, we obtain that $\vec{u}$ is complete.

We are left with the last part. For that it suffices to show that $(d\exp)_0$ is an isomorphism. But

$$(d\exp)_0 : \mathfrak{g} \to T_eG = \mathfrak{g}$$

sends $u \in \mathfrak{g}$ to

$$\left.\frac{d}{dt}\right|_{t=0} \exp(tu) = \left.\frac{d}{dt}\right|_{t=0} \phi^t_{\vec{u}}(e) = \vec{u}_e = u,$$

i.e. the differential is actually the identity map. □

EXAMPLE 4.45. In Example 4.25 we have seen that the Lie algebra of the general linear group $GL_n(\mathbb{R})$ is just the algebra $\mathcal{M}_n(\mathbb{R})$ of $n \times n$ matrices endowed with the commutator bracket. We claim that the resulting exponential map is just the usual exponential map for matrices:

$$\mathrm{Exp} : \mathcal{M}_n(\mathbb{R}) \to GL_n(\mathbb{R}),\ X \mapsto e^X := \sum_{k=0}^{\infty} \frac{X^k}{k!}$$

(where we work with the conventions that $X^0 = I$ the identity matrix and $0! = 1$). Let us compute $\exp(X)$ for an arbitrary $X \in \mathcal{M}_n(\mathbb{R}) = T_I GL_n$; according to the definition, it is $\gamma(1)$, where γ is the unique path in GL_n satisfying

$$\gamma(0) = I,\ \frac{d\gamma}{dt}(t) = \vec{X}\,(\gamma(t)).$$

The last equality takes place in $T_{\gamma(t)}GL_n$, which is canonically identified with $\mathcal{M}_n(\mathbb{R})$ (as in Example 4.25 , $Y \in \mathcal{M}_n(\mathbb{R})$ is identified with the speed at $\epsilon = 0$ of $\epsilon \mapsto \gamma(t) + \epsilon Y$). By this identification, $\frac{d\gamma}{dt}(t)$ goes to the usual derivative of γ (as a path in the Euclidean space $\mathcal{M}_n(\mathbb{R})$) and $\vec{X}\,(\gamma(t))$ goes to $\gamma(t) \cdot X$. Hence γ is the solution of

$$\gamma(0) = I,\ \gamma'(t) = X \cdot \gamma(t),$$

which is precisely what $t \mapsto e^{tX}$ does. We deduce that $\exp(X) = e^X$.

9. For the curious students: application to closed subgroups of GL_n

As an application of the exponential map we present here, for the curious students, a proof of Theorem 3.31 from Chapter 3; actually, of an improved version of it.

THEOREM 4.46. *Any closed subgroup G of GL_n is an embedded Lie subgroup of GL_n (i.e. an embedded submanifold which, together with the resulting smooth structure, becomes a Lie group). Moreover, it Lie algebra is*

$$\mathfrak{g} := \{X \in \mathcal{M}_n(\mathbb{R}) : e^{tX} \in G \;\; \forall \;\; t \in \mathbb{R}\} \subset \mathcal{M}_n(\mathbb{R})$$

(endowed with the commutator bracket).

PROOF. For the first part we have to prove that G is an embedded submanifold of GL_n; for simplicity, assume we work over $\mathbb{R}$. We will show that $\mathfrak{g}$ is a linear subspace of $\mathcal{M}_n(\mathbb{R})$; then we will choose a complement $\mathfrak{g}'$ of $\mathfrak{g}$ in $\mathcal{M}_n(\mathbb{R})$ and we show that one can find an open neighborhood V of the origin in $\mathfrak{g}$, and a similar one V' in $\mathfrak{g}'$, so that

$$\phi : V \times V' \to GL_n, \phi(X, X') = e^X \cdot e^{X'}$$

is a local diffeomorphism onto an open neighborhood of the identity matrix and so that

$$G \cap \phi(V \times V') = \{\phi(v, 0) : v \in V\}. \tag{9.1}$$

This means that the submanifold condition is verified around the identity matrix and then, using left translations, it will hold at all points of G. Note also that the differential of ϕ at 0 (with ϕ viewed as a map defined on the entire $\mathfrak{g} \times \mathfrak{g}'$) is the identity map:

$$(d\phi)_0 : \mathfrak{g} \times \mathfrak{g}' \to T_I GL_n = \mathcal{M}_n(\mathbb{R}), (X, X') \mapsto \frac{d}{dt}_{t=0} e^{tX} e^{tX'} = X + X'.$$

In particular, ϕ is a indeed a local diffeomorphism around the origin. Hence the main condition that we have to take care of is (9.1), i.e. that if $X \in V$, $X' \in V'$ satisfy $e^X \cdot e^{X'} \in G$, then $X' = 0$. Hence it suffices to show that one can find V' so that

$$X' \in V', e^{X'} \in G \Longrightarrow X' = 0$$

(then just choose any V small enough such that ϕ is a diffeomorphism from $V \times V'$ into an open neighborhood of the identity in GL_n). For that we proceed by contradiction. If such V' would not exist, we would find a sequence $Y_n \to 0$ in $\mathfrak{g}'$ such that $e^{Y_n} \in G$. Since Y_n converges to 0, we find a sequence of integers $k_n \to \infty$ such that $X_n := k_n Y_n$ stay in a closed bounded region of $\mathfrak{g}'$ not containing the origin (e.g. on $\{X \in \mathfrak{g}' : 1 \leq ||X|| \leq 2\}$, for some norm on $\mathfrak{g}'$); then, after eventually passing to a subsequence, we may then assume that $X_n \to X \in \mathfrak{g}'$ non-zero. In particular, X cannot belong to $\mathfrak{g}$. Setting $t_n = \frac{1}{k_n}$, we see that $e^{t_n X_n} = e^{Y_n} \in G$, and then the desired contradiction will follow from:

LEMMA 4.47. *Let $X_n \to X$ be a sequence of elements in $\mathfrak{gl}_n$ and $t_n \to 0$ a sequence of non-zero real numbers. Suppose $e^{t_n X_n} \in G$ for all n: then $e^{tX} \in G$ for all t, i.e., $X \in \mathfrak{g}$.*

For a proof of this Lemma, please see the book by Sternberg on Differential Geometry (Lemma 4.2, Chapter V). This Lemma can also be used to prove the claim we made (and used) at the start of the argument: that $\mathfrak{g}$ is a linear subspace of $\mathcal{M}_n(\mathbb{R})$. The $\mathbb{R}$-linearity follows immediately from the definition. Assume now that $X, Y \in \mathfrak{g}$ and we show that $X + Y \in \mathfrak{g}$. Write now

$$e^{tX} e^{tY} = e^{f(t)}, \quad \text{with} \quad \lim_{t \to 0} \frac{f(t)}{t} = X + Y$$

(exercise: why is this possible?). Taking t_n a sequence of real numbers converging to 0 and $X_n = \frac{f(t_n)}{t_n}$ in the lemma, to conclude that $X + Y \in \mathfrak{g}$.

This closes the proof of the fact that G is a submanifold of GL_n (hence also a Lie group). It should now also be clear that $\mathfrak{g}$ is the Lie algebra of G since $\phi|_{V\times\{0\}}$ is precisely the restoration of the exponential map e to the open neighborhood V of the origin in $\mathfrak{g}$. □

EXERCISE 4.25. In particular, it follows that $\mathfrak{g}$ is closed under the commutator bracket of matrices:

$$X, Y \in \mathfrak{g} \Longrightarrow [X, Y] \in \mathfrak{g}.$$

However, prove this directly, using the previous Lemma, by an argument similar to the one we used to prove that $\mathfrak{g}$ is a linear subspace of gl_n.

CHAPTER 5

Differential forms

1. Differential forms of degree one (1-forms)

For a (real) vector space V, its dual V^* is the vector space consisting of all linear maps $\xi : V \to \mathbb{R}$. While the elements of V are vectors of V, the elements of V^* are thought of as "covectors of V". And vectors and covectors pair via the evaluation map

$$\langle \cdot, \cdot \rangle : V^* \times V \to \mathbb{R}, \quad (\xi, v) \mapsto \langle \xi, v \rangle := \xi(v).$$

Moreover, the equations

$$\langle e^i, e_j \rangle = \delta^i_j$$

show how to produce a basis $e^1, \ldots, e^m$ of V^* from a basis $e_1, \ldots, e_m$ of M (and the other way around).

Differential 1-forms on a manifold M arise similarly, in duality with vector fields. First of all, for each $p \in M$, we define **the cotangent space** of M at p as the dual of the tangent space

$$T^*_p M := (T_p M)^*.$$

The elements of $T^*_p M$ are called **cotangent vectors at** p.

EXAMPLE 5.1. For any smooth function $f : M \to \mathbb{R}$ its differential at any point $p \in M$,

$$(df)_p : T_p M \to \mathbb{R}$$

can now be interpreted as a cotangent vector at p,

$$(df)_p \in T^*_p M.$$

Of course, for this to make sense we only need f to be defined on an open containing p.

While a chart χ around p gives rise to a basis of $T_p M$ (for p in the domain of χ), it will also give rise to a basis of $T^*_p M$: just apply the previous example to the components $\chi_i : U \to \mathbb{R}$ of χ to obtain cotangent vectors at p

$$(d\chi_1)_p, \ldots, (d\chi_m)_p.$$

To see that this is a basis one can just remark that this is dual to the similar basis of $T_p M$; more precisely, it follows right away from the definitions of all the term involved that

$$\langle (d\chi_i)_p, \left(\frac{d}{d\chi_j}\right)_p \rangle = \delta^i_j$$

for all i and j.

Similar to set-theoretical vector fields on M we can now talk about set-theoretical cotangent vector fields on M, by which we mean maps

$$\omega : M \ni p \mapsto \omega_p \in T^*_p M.$$

To make sense of their smoothness one proceeds like for vector fields: with respect to any chart $\chi : U \to \Omega \subset \mathbb{R}^m$, we can write

$$\omega_p = \sum_i \omega_\chi^i(\chi(p))(d\chi_i)_p,$$

with $\omega_\chi^i : \Omega \to \mathbb{R}$ called the coefficients of ω with respect to the chart χ.

DEFINITION 5.2. *A **differential** 1**-form** on M is any map*

$$\omega : M \ni p \mapsto \omega_p \in T_p^*M$$

whose coefficients w.r.t. any chart are smooth. We denote by $\Omega^1(M)$ the vector space of all such 1*-forms (with the vector space structure defined by $(\omega + \omega^{'})_p = \omega_p + \omega^{'}_p$, etc).*

EXAMPLE 5.3 (differentials of functions). Continuing the previous example, we see that for any function $f \in \mathcal{C}^\infty(M)$ its differential now becomes a 1-form

$$df \in \Omega^1(M)$$

(why is it smooth?). In this way, the differentiation of $\mathbb{R}$-valued functions becomes a linear operator

$$d : \mathcal{C}^\infty(M) \to \Omega^1(M).$$

EXAMPLE 5.4 (expressions of type $\sum f_i dg_i$). As for vector fields, 1-forms ω can be multiplied by smooth functions f to produce a new 1-form $f \cdot \omega$ (whose value at $p \in M$ is $f(p) \cdot \omega_p$). More abstractly, one has an operation

$$\mathcal{C}^\infty(M) \times \Omega^1(M) \to \Omega^1(M)$$

which makes the space of 1-forms into a $C^\infty(M)$-module.

Combining with differentials of functions, we see see that expressions of type

$$\sum_{i=1}^k f_i \cdot dg_i$$

now make sense as 1-forms. Actually, any 1-form is locally representable by such an expression (of course, in a non-unique way). Looking in a chart, this becomes a property of $\mathbb{R}^m$; and it should be clear that any 1-form ω on $\mathbb{R}^m$ can be uniquely written as

$$\omega = \sum f_i \cdot dx_i \tag{1.1}$$

with $f_i \in \mathcal{C}^\infty(\mathbb{R}^m)$ (where x_i are the coordinate functions $\in \mathcal{C}^\infty(\mathbb{R}^m)$),

EXERCISE 5.1. Show that:

(1) For any function $f \in \mathcal{C}^\infty(\mathbb{R}^m)$, one has

$$df = \sum_i \frac{\partial f}{\partial x_i} \cdot dx_i.$$

(2) Show that the 1-form on $\mathbb{R}^2$ given by

$$\omega = ydx - xdy$$

is not exact, i.e. cannot be written as df for some smooth function f.

EXERCISE 5.2. Show that, for a 1-form $\omega \in \Omega^1(M)$ to be exact, i.e. to be of type df for some smooth function f, a necessary (but not sufficient) condition is that

$$\omega([X, Y]) = L_X(\omega(Y)) - L_Y(\omega(X)) \quad \text{for all } X, Y \in \mathfrak{X}(M).$$

EXAMPLE 5.5 (restrictions to submanifolds). There is a "restriction operation" of 1-forms $\omega \in \Omega^1(M)$ to arbitrary submanifolds $N \subset M$. Indeed, since any tangent vector $X_p \in T_pN$ can be viewed also as a tangent vector to M, one can just evaluate ω_p on any $X_p \in T_pN$. The resulting 1-form on N is called the restriction of ω to N and is denoted $\omega|_N$. Hence

$$(\omega|_N)_p (X_p) := \omega_p(X_p)$$

defines the restriction operation

$$\mathrm{restr} : \Omega^1(M) \to \Omega^1(N), \quad \omega \mapsto \omega|_N.$$

This operation is particularly useful when N is an embedded submanifold of an Euclidean space $\mathbb{R}^k$: one can produce 1-forms on N by starting with 1-forms on $\mathbb{R}^k$ (necessarily of type (1.1)); and, even more, any 1-form on N arises in this way (but this last statement is not completely obvious). In other words, one can represent 1-form on N by 1-forms on $\mathbb{R}^k$. E.g., for $N = S^2$, we can now talk about

$$(x^2 + y) \cdot dx + dy + xyz \cdot dz \text{ on } S^2,$$

by which we mean the restriction of $(x^2 + y) \cdot dx + dy + xyz \cdot dz$ from $\mathbb{R}^3$ to $S^2 \subset \mathbb{R}^3$.

However, one should be aware that different forms on $\mathbb{R}^k$ may give the same restriction to N. E.g., for $S^2 \subset \mathbb{R}^3$, there are expressions that look like being non-zero (and they are non-zero as 1-forms on $\mathbb{R}^3$) but which represent the zero form on S^2. For instance,

$$x \cdot dx + y \cdot dy + z \cdot dz \text{ on } S^2$$

is zero. Indeed, viewing it as a 1-form on S^2, it means we evaluate it only on vectors tangent to S^2; and, for $p = (x, y, z) \in S^2$ and an arbitrary tangent vector

$$V_p = a\left(\frac{\partial}{\partial x}\right)_p + b\left(\frac{\partial}{\partial y}\right)_p + c\left(\frac{\partial}{\partial z}\right)_p \in T_pS^2$$

one has

$$(x \cdot dx + y \cdot dy + z \cdot dz)(V_p) = ax + by + cz = 0$$

since V was tangent to S^2. In other words,

$$(x \cdot dx + y \cdot dy + z \cdot dz)|_{S^2} = 0.$$

EXERCISE 5.3. Show that the 1-form on S^1 given by

$$\omega = (ydx - xdy)\,|_{S^1}$$

is not exact, i.e. cannot be written as df with $f \in \mathcal{C}^\infty(S^1)$.

EXERCISE 5.4. Show that for any $f \in \mathcal{C}^\infty(M)$ and any chart (U, χ) one has, on U,

$$df = \sum_i \frac{\partial f}{\partial \chi_i} \cdot \chi_i.$$

Deduce that for any two coordinate charts (U, χ) and (U', χ') one has, on $U \cap U'$,

$$d\chi_i' = \sum_j \frac{\partial \chi_i'}{\partial \chi_j} \cdot d\chi_j. \tag{1.2}$$

Then, using the duality between $\frac{\partial}{\partial \chi_i}$ and $d\chi_i$, show again that, over $U \cap U'$, one has:

$$\frac{\partial}{\partial \chi_i} = \sum \frac{\partial \chi_j'}{\partial \chi_i} \cdot \frac{\partial}{\partial \chi_j'}.$$

And here is one very nice feature of 1-forms: unlike vector fields, 1-forms can be pulled-back along *any* smooth map. Actually, an instance of this operation is the restriction to submanifolds that we have just described- which corresponds to pull-backs via the inclusion map.

To explain this, we start with a smooth map $F : M \to N$ and $\omega \in \Omega^1(N)$. Then at each point $p \in M$ we can apply the differential $(dF)_p : T_pM \to T_{F(p)}N$ followed by $\omega_{F(p)} : T_{F(p)}N \to \mathbb{R}$ to obtain a covector on M at p:

$$T_pM \ni X_p \mapsto \omega_{F(p)}\left((dF)_p(X_p)\right) \in \mathbb{R}$$

(compare with vector fields!). The resulting 1-form on M is denoted by

$$F^*(\omega) \in \Omega^1(M)$$

and is called **the pull-back of** ω **by** F.

Exercise 5.5. Show that, indeed, $F^*(\omega)$ is smooth. Also check that, when $M \subset N$ and F is the inclusion, then $F^*(\omega) = \omega|_M$.

Therefore, any smooth map $F : M \to N$ gives rise to an operation

$$F^* : \Omega^1(N) \to \Omega^1(M).$$

Exercise 5.6. Show that for any $F : M \to N$ smooth, the following diagram is commutative:

$$\begin{array}{ccc} \mathcal{C}^\infty(N) & \xrightarrow{d} & \Omega^1(N) \\ {\scriptstyle F^*}\downarrow & & \downarrow{\scriptstyle F^*} \\ \mathcal{C}^\infty(M) & \xrightarrow{d} & \Omega^1(M) \end{array} .$$

While covectors of a vector space V take vectors to real numbers, 1-forms take vector fields to smooth functions: given ω as above and $X \in \mathfrak{X}(M)$, evaluating ω_p on X_p for each $p \in M$ we obtain a smooth function

$$\omega(X) \in \mathcal{C}^\infty(M);$$

(why smooth?). When we vary X it is clear that the resulting map

$$\mathfrak{X}(M) \to \mathcal{C}^\infty(M), \quad X \to \omega(X)$$

is $C^\infty(M)$-linear, i.e. it is linear and

$$\omega(f \cdot X) = f \cdot \omega(X) \quad \text{for all } f \in \mathcal{C}^\infty(M), X \in \mathfrak{X}(M).$$

We denote by

$$\mathrm{Hom}_{\mathcal{C}^\infty}(\mathfrak{X}(M), \mathcal{C}^\infty(M))$$

the space of all $\mathcal{C}^\infty(M)$-linear maps from $\mathfrak{X}(M)$ to $\mathcal{C}^\infty(M)$.

Theorem 5.6. *Interpreting any* $\omega \in \Omega^1(M)$ *as a map* $\mathfrak{X}(M) \to \mathcal{C}^\infty(M)$*, we obtain an isomorphism*

$$\Omega^1(M) \cong Hom_{\mathcal{C}^\infty}(\mathfrak{X}(M), \mathcal{C}^\infty(M)).$$

Proof. Note first that, by the same arguments as in Proposition 4.14 from Chapter 4, for a set-theoretical 1-form ω its smoothness is equivalent to the fact that $\omega(X)$ is a smooth function on M for all $X \in \mathfrak{X}(M)$.

Another ingredient that we need in the proof is the fact that, for each $p \in M$, the evaluation of vector fields at p,

$$\mathrm{ev}_p : \mathfrak{X}(M) \to T_pM, \quad X \mapsto X_p,$$

is surjective. To show this, we start with an arbitrary tangent vector $v \in M$ (p is fixed now). Using a coordinate chart around p, we find a vector field Y defined on an open neighborhood U of p such that $Y_p = v$. As in the proof of Lemma 4.9 from Chapter 4, choose a smooth function η on M, supported inside U and with $\eta(p) = 1$. Then $X := \eta \cdot Y$ is well-defined and smooth on M and $X_p = \eta(p) \cdot v = v$.

We now move to the actual proof. For the injectivity of the map we have to show that, if $\omega_p(X_p) = 0$ for all $X \in \mathfrak{X}(M)$ and all $p \in M$, then $\omega_p = 0$ for all $p \in M$. But this clearly follows from the surjectivity of ev_p.

We now prove the surjectivity of the map in the statement. Let $A : \mathfrak{X}(M) \to \mathcal{C}^\infty(M)$ be a $\mathcal{C}^\infty(M)$-linear map. We fix $p \in M$ arbitrary and we define

$$\omega_p : T_pM \to \mathbb{R}, \quad \omega_p(v) := A(Y)(p)$$

where A is any vector field with $Y_p = v$. If we check that this definition does not depend on the choice of Y then we are done: linearity of each ω_p follows right away, while the resulting set-theoretical 1-form is also smooth since, by construction, $\omega(X) = A$.

To see that $A(Y)(p)$ does not depend on the choice of Y, assume that Y' is another vector field with $Y'_p = v$. Then $X := Y - Y'$ vanishes at p and we have to show that $A(X)(p) = 0$. Using the $\mathcal{C}^\infty(M)$-linearity of A, this will follow from the following:

LEMMA 5.7. *If a vector field $X \in \mathfrak{X}(M)$ vanishes at a point $p \in M$, then we can write*

$$X = \sum_{j=0}^{k} f_j \cdot X^j$$

with $k \geq 1$ integer, $X_j \in \mathfrak{X}(M)$ and f_j smooth functions that vanish at p.

PROOF. We first look in a coordinate chart (U, χ) around p which sends p to $\chi(p) = 0 \in \mathbb{R}^m$. Then writing

$$X|_U = \sum_{i=1}^{m} F_i \cdot \frac{\partial}{\partial \chi_i},$$

the coefficients F_i are smooth functions on U that vanish at p. By the the proof of Theorem 4.8 from Chapter 4 applied to $F_i \circ \chi^{-1}$, we see that each F_i can be written as a finite sum $\sum_j \chi_j \cdot g_i^j$ with $g_i^j : U \to \mathbb{R}$ smooth. From this we immediately deduce that we can find some smooth functions $f_1, \ldots, f_k$ on U and vector fields $Y_1, \ldots, Y_k$ (on U) such that

$$X|_U = \sum_{j=1}^{k} f_j \cdot Y^j.$$

To go from U to the entire M, the main idea is to write $X = \eta \cdot X + (1 - \eta) \cdot X$, for some carefully chosen smooth function η.

Note that, after eventually making U smaller, we may assume that $f_i = \tilde{f}_i|_U$ for some smooth function $\tilde{f}_i \in \mathcal{C}^\infty(M)$; to see this, just use the second dotted item in the proof of Lemma 4.11 from Chapter 4. Furthermore, as in that proof, we choose a smooth function $\eta \in C^\infty(M)$ which is supported inside U and is 1 in some neighborhood V of p. Since η is supported inside U, each

$\eta|_U \cdot Y^j \in \mathfrak{X}(U)$ can be promoted to a vector field $X^j \in \mathfrak{X}(M)$ by declaring it to be zero outside U. With the inspiration from the "main idea" mentioned above, we now remark that

$$X = (1-\eta)\cdot X + \sum_{j=1}^{k} \tilde{f}_j \cdot \tilde{X}^j$$

on the entire M. Indeed, at $p \in U$ the right hand side is

$$(1-\eta(p))\cdot X_p + \eta(p)\cdot \sum_{j=1}^{k} f_j(p)\cdot X_p^j = (1-\eta(p))\cdot X_p + \eta(p)\cdot X_p = X_p,$$

while at p outside U we obtain $(1-0)\cdot X_p + \sum_j \tilde{f}_j(p)\cdot 0 = X_p$.

Therefore the desired decomposition is achieved if we take $f_0 = 1-\eta$ and $X^0 = X$. □

□

And, still as in the case of tangent vectors, one can form the cotangent bundle

$$T^*M := \sqcup_{p\in M} T_p^*M = \{(p,\omega_p) : p \in M, \omega_p \in T_p^*M\}$$

and put a smooth structure on T^*M so that the smoothness of a set-theoretical 1-form ω is equivalent to the smoothness of ω as a map from M to T^*M.

Exercise 5.7. Fill in the details, i.e. state and prove a version of Proposition 4.17 but for T^*M instead of TM.

2. Differential forms: the linear algebra behind them

The passage from 1-forms to arbitrary k-forms is basically a passage from linear maps $T_pM \to \mathbb{R}$ to k-multilinear skew-symmetric maps. Here we discuss the linear algebra involved.

Definition 5.8. *Given a vector space V and $k \in \mathbb{N}$ a natural number, a k-covector on V is any map*

$$\eta : \underbrace{V \times \ldots \times V}_{k \text{ times}} \to \mathbb{R}$$

that is multiplinear (i.e. linear in each argument). We say that ω if a linear k-form on V if

$$\omega(v_\sigma(1), \ldots, v_{\sigma(k)}) = sign(\sigma)\omega(v_1, \ldots, v_k)$$

for any $v_1, \ldots, v_k \in V$ and any permutation $\sigma \in S_k$. We denote by $\Lambda^k V^*$ the (vector) space of all such linear k-forms on V.

Exercise 5.8. Show that ω is skew-symmetric if and only if the expression $\omega(v_1, \ldots, v_k)$ changes sign whenever two entries v_i and v_j are interchanged.

Although our interest is $\Lambda^k V^*$, it will be useful to consider more general k-multiplinear maps (without being skew-symmetric); we denote the space of such by $T^k V^*$. Note that, while

$$\Lambda^k V^* \subset T^k V^*,$$

there is also a skew-symmetrization map going the other way,

$$\text{Alt} : T^k V^* \to \Lambda^k V^*$$

where for $\omega \in T^k V^*$, $\text{Alt}(\omega)$ is given by

$$\text{Alt}(\omega)(v_1, \ldots, v_k) := \frac{1}{k!}\sum_{\sigma\in S_k} \text{sign}(\sigma)\omega(v_\sigma(1), \ldots, v_{\sigma(k)}).$$

EXERCISE 5.9. Show that, for $\omega \in T^k V^*$, one has:

$$\omega \in \Lambda^k V^* \Longleftrightarrow \mathrm{Alt}(\omega) = \omega.$$

The main operation on forms is that of wedge-product. Before giving the precise definition/formula, note first that there a pretty obvious operation involving the spaces $T^k V^*$, namely:

$$T^k V^* \times T^l V^* \to T^{k+l} V^*, \quad (\omega, \eta) \mapsto \omega \bullet \eta,$$

where $\omega \bullet \eta$ is given by

$$(\omega \bullet \eta)(v_1, \ldots, v_{k+l}) := \omega(v_1, \ldots, v_k) \cdot \eta(v_{k+1}, \ldots, v_{k+l}).$$

EXERCISE 5.10. Do the following:

(1) If $\omega \in T^2 V^*$ is symmetric and $\eta \in T^l V^*$ is arbitrary, show that $\mathrm{Alt}(\omega \bullet \eta) = 0$.
(2) Prove the same but now assuming that $\omega \in T^k V^*$ satisfies $\mathrm{Alt}(\omega) = 0$.
(3) Deduce that for any $\omega \in T^k V^*$ and $\eta \in T^l V^*$ one has $\mathrm{Alt}(\omega \bullet \eta) = \mathrm{Alt}(\mathrm{Alt}(\omega) \bullet \eta)$.

However, if ω and η are skew-symmetric, $\omega \cdot \eta$ may fail to be skew-symmetric. Therefore, to obtain a skew-symmetric element, we will apply the skew-symmetrization map Alt. To avoid un-necessary coefficients in the final formula we will do a small rescaling (see also below):

DEFINITION 5.9. *Given $\omega \in \Lambda^k V^*$ and $\eta \in \Lambda^l V^*$, define the element*

$$\omega \wedge \eta \in \Lambda^{k+l} V^*,$$

called the ***wedge product of*** ω ***and*** η*, by*

$$\omega \wedge \eta = \frac{(k+l)!}{k! \cdot l!} Alt(\omega \bullet \eta).$$

REMARK 5.10. When computing $\mathrm{Alt}(\omega \cdot \eta)$, the fact that ω and η are skew-symmetric will be reflected in the fact that many of the terms will repeat. When we avoid repetitions, we get sums not over all permutations $\sigma \in S_{k+l}$ but only over those with the property that

$$\sigma(1) < \ldots < \sigma(k), \quad \sigma(k+1) < \ldots < \sigma(k+l).$$

Such permutations will be called (k, l)**-shuffles** and we denote by $S_{k,l} \subset S_{k+l}$ the resulting set. All together, we find the following explicit formula for the wedge product:

$$(\omega \wedge \eta)(v_1, \ldots, v_{k+l}) := \sum_{\sigma \in S_{k,l}} \omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \eta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)}).$$

For instance, for $\omega, \eta \in V^*$ one obtains $\omega \wedge \eta \in \Lambda^2 V^*$ given by

$$(\omega \wedge \eta)(v, w) = \omega(v)\eta(w) - \omega(w)\eta(v).$$

Or, for $\omega, \in \Lambda^2 V^*, \eta \in V^*$,

$$(\omega \wedge \eta)(u, v, w) = \omega(u, v)\eta(w) + \omega(v, w)\eta(u) - \omega(u, w)\eta(v).$$

The reason for the coefficients $k!$ used in the definition of the wedge was precisely obtain such formulas without coefficients (incidentally, let us point out that the resulting formula now makes send over any field, not necessarily of characteristic zero). Actually, replacing $k!$ by any coefficients c_k would still produce (modified) wedge-products with similar properties (see the next proposition) but, in some (precise) sense, they are all "isomorphic". However, simple choices of c_ks, such as $c_k = 1$, would produce uglier final formulas.

In this way we obtain operations

$$\cdot \wedge \cdot : \Lambda^k V^* \times \eta \in \Lambda^l V^* \to \Lambda^{k+l} V^*.$$

Here are the main properties of these operations:

PROPOSITION 5.11. *The wedge operation is bilinear, and:*

(1) *graded-commutativity: for any $\omega \in \Lambda^k V^*$ and $\eta \in \Lambda^l V^*$,*

$$\eta \wedge \omega = (-1)^{kl} \omega \wedge \eta.$$

(2) *associativity: for any linear forms ω, η and ζ one has*

$$(\omega \wedge \eta) \wedge \zeta = \omega \wedge (\eta \wedge \zeta).$$

PROOF. For graded-commutativity, we look at $\eta \wedge \omega(v_1, \ldots, v_{k+l})$. The sum defining it has expressions of type

$$\text{sign}(\sigma)\eta(v_{\sigma(1)}, \ldots, v_{\sigma(l)}) \cdot \omega(v_{\sigma(l+1)}, \ldots, v_{\sigma(l+k)})$$

where σ is an (l, k)-shuffle. Now, any such σ defines another permutation $\hat{\sigma}$ by

$$\hat{\sigma}(1) = \sigma(l+1), \ldots, \hat{\sigma}(k) = \sigma(l+k), \quad \hat{\sigma}(k+1) = \sigma(1), \ldots, \hat{\sigma}(k+l) = \sigma(l)$$

which is a (k, l)-shuffle. A moment of reflection shows that

$$\text{sign}(\hat{\sigma}) = (-1)^{kl}\text{sign}(\sigma).$$

All these clearly imply the graded-commutativity.

For associativity, up to a constant that is the same on both sides, we have to show that

$$\text{Alt}(\text{Alt}\,(\omega \bullet \theta) \bullet \zeta) = \text{Alt}(\omega \bullet \text{Alt}\,(\theta \bullet \zeta))\,.$$

Using the last part of Exercise 5.10 to rewrite the left hand side, and the similar argument for the right hand side, what we have to prove becomes:

$$\text{Alt}((\omega \bullet \theta) \bullet \zeta) = \text{Alt}(\omega \bullet (\theta \bullet \zeta))\,.$$

This follows right away since the operation $\bullet$ is clearly associative. □

EXAMPLE 5.12. For any linear 1-forms $\xi_1, \ldots, \xi_k \in V^*$, one obtains a k-form

$$\xi_1 \wedge \ldots \wedge \xi_k \in \Lambda^k V^*.$$

As we shall soon see, any linear k-form can be written as a sum of expressions of this type. Even more, this operation allows us to produce an explicit basis for the vector space $\Lambda^k V^*$ by starting with a basis

$$e_1, \ldots, e_m \in V$$

of V. First of all, we consider the dual basis

$$e^1, \ldots, e^m \in V^*.$$

Then we start taking wedges of such elements; to obtain k-forms, we need to use k such (co)vectors. Given the graded commutativity of the wedge-operations, we can always re-order the indices in an increasing fashion. Therefore, we will consider sets of increasing indices

$$I = (i_1, \ldots, i_k), \quad \text{with} 1 \leq i_1 < \ldots < i_k \leq m;$$

the set of such Is will be denoted by $\text{Ord}_k(m)$. And, for $I \in \text{Ord}_k(m)$, consider

$$e^I := e^{i_1} \wedge \ldots \wedge e^{i_k} \in \Lambda^k V^*.$$

LEMMA 5.13. *The vectors e^I, with $I \in Ord_k(m)$, form a basis of $\Lambda^k V^*$. In particular, for $k \leq m$ (where m is the dimension of V) one has*

$$dim(\Lambda^k V^*) = Card\,(Ord_k(m)) = \frac{m!}{k!(m-k)!},$$

while for $k > m$ one has $\Lambda^k V^* = 0$.

Proof. Note that, for any I as above and any other $J = (j_1, \ldots, j_k) \in \mathrm{ord}_k(M)$ one has

$$e^I(e_{j_1}, \ldots, e_{j_k}) = \delta^I_J$$

(i.e. 1 when $I =$ and zero otherwise). From this we deduce right away that the e^Is are linearly independent: if $\sum_I \lambda_I e^I = 0$ with $\lambda_I \in \mathbb{R}$, evaluating on $(e_{j_1}, \ldots, e_{j_k})$ we find that $\lambda_J = 0$ for all J.

To see that our family spans the entire $\Lambda^k V^*$, it suffices to show that, for any $\omega \in \Lambda^k V^*$),

$$\omega = \sum_{I \in \mathrm{Ord}_k(m)} \omega(e_{i_1}, \ldots, e_{i_k}) \cdot e^I.$$

To check this equality, we first note that the difference of the two sides, call it $\eta \in \Lambda^k V^*$, is zero when evaluated on all elements of type $(e_{i_1}, \ldots, e_{i_k})$. But this is easily seen to imply that $\eta = 0$: for an arbitrary expression $\eta(v_1, \ldots, v_k)$, decomposing each v_i with respect to the basis $\{e_1, \ldots, e_m\}$ and using the multilinearity of η and then the antisymmetry, we obtain a sum of expressions of type $\eta(e_{i_1}, \ldots, e_{i_k})$. □

3. Differential forms

We now pass to differential forms on an arbitrary m-dimensional manifold M. As for 1-forms, one first talks about **set-theoretical k-forms**, by which we mean maps

$$\omega : M \ni p \mapsto \omega_p \in \Lambda^k T^*_p M.$$

And, still as for 1-forms, one makes sense of their smoothness using charts (U, χ): any such chart gives rise to the basis

$$\left(\frac{\partial}{\partial \chi_1}\right)_p, \ldots, \left(\frac{\partial}{\partial \chi_m}\right)_p$$

of $T_p M$, the dual basis

$$(d\chi_1)_p, \ldots, (d\chi_m)_p$$

of $T^*_p M$ and then, using Lemma 5.13, to the basis of $\Lambda^k T^*_p M$ given by

$$(d\chi)^I_p = (d\chi_{i_1})_p \wedge \ldots \wedge (d\chi_{i_k})_p$$

with $I \in \mathrm{Ord}_k(1, \ldots, m)$. Therefore, we can write our ω as

$$\omega_p = \sum_{I \in \mathrm{Ord}_k(m)} \omega^I_\chi(\chi(p))(d\chi)^I_p,$$

where each coefficient ω^I_χ is a function on $\Omega = \chi(U) \subset \mathbb{R}^m$. We say that ω is **smooth** if all its coefficients, with respect to any chart, are smooth functions.

A differential k-form on M, or a differential form of degree k, is any set-theoretical k-form that is smooth; we denote by

$$\Omega^k(M)$$

the space of all such k-forms.

It should be clear now that $\Omega^k(M)$ is a vector space, and it comes with a structure of $\mathcal{C}^\infty(M)$-module,

$$\mathcal{C}^\infty(M) \times \Omega^k(M) \to \Omega^k(M), \quad (f, \omega) \mapsto f \cdot \omega.$$

While, point-wise (at some $p \in M$), ω takes k tangent vectors (at p) to produce numbers, globally (when varying p), it takes k vector fields and produces a function. More explicitly, for $X^1, \ldots, X^k \in \mathfrak{X}(M)$, one obrains the function

$$\omega(X^1, \ldots, X^k) : M \to \mathbb{R}, \quad p \mapsto \omega_p(X^1_p, \ldots, X^k_p).$$

EXERCISE 5.11. Check that the smoothness of ω is equivalent to the fact that $\omega(X^1, \ldots, X^k)$ is smooth for any $X^1, \ldots, X^k \in \mathfrak{X}(M)$.

PROPOSITION 5.14. *The previous interpretation of k-forms gives a 1-1 correspondence between k-forms on M and multilinear maps*

$$\omega : \underbrace{\mathfrak{X}(M) \times \mathfrak{X}(M)}_{k-times} \to \mathcal{C}^\infty(M)$$

with the following two properties:

- *ω is skew-symmetric.*
- *ω is $\mathcal{C}^\infty(M)$-linear in each argument X_i:*

$$\omega(X_1, \ldots, X_{i_1}, f \cdot X_i, X_{i+1}, \ldots, X_k) = f \cdot \omega(X_1, \ldots, X_k).$$

PROOF. For the moment, let denoted by $\hat{\omega}$ the version of ω that acts on vector fields. It is clear that it is skew-symmetric and $\mathcal{C}^\infty(M)$-multilinear. For the converse, let

$$A : \underbrace{\mathfrak{X}(M) \times \mathfrak{X}(M)}_{k-\text{times}} \to \mathcal{C}^\infty(M)$$

a map with the same properties and we want to show that $A = \hat{\omega}$ for some differential k-form ω. As in the previous sections, for $p \in M$ we define ω_p on tangent vectors $v_p^1, \ldots, v_p^k \in T_pM$ by:

$$\omega_p(v_p^1, \ldots, v_p^k) := A(X^1, \ldots, X^k)(p),$$

where $X^i \in \mathfrak{X}(M)$ are chosen so that, at p, they give back the vectors v_p^i. The main issue is, as for the similar discussion for 1-forms, to see that this definition does not depend on the choice of the X^is. But that follows exactly as for 1-forms (or one could even use that result inductively). The rest should be clear. □

EXAMPLE 5.15. In Exercise 5.2 we have seen that, for a 1-form $\omega \in \Omega^1(M)$ to be exact one would need in particular that the expression

$$-\omega([X, Y]) + L_X(\omega(Y)) - L_Y(\omega(X))$$

to be zero for all $X, Y \in \mathfrak{X}(M)$. Looking at the expression above one can check right away that it is skew-symmetric and $\mathcal{C}^\infty(M)$-bilinear in X and Y. Therefore it defines a 2-form on M- that is usually denoted by

$$d\omega \in \Omega^2(M);$$

as we have seen, it arises as the obstruction to ω being exact.

Note that, in this way, we extended the differential of functions to 1-forms:

$$\mathcal{C}^\infty(M) = \Omega^0(M) \xrightarrow{d} \Omega^1(M) \xrightarrow{d} \Omega^2(M).$$

This, and its further continuation will be discussed in more detail in the next section ("De Rham differential").

Next, the wedge-product discussed in the previous section, applied at each point $p \in M$, gives rise to similar operations

$$\cdot \wedge \cdot : \Omega^k(M) \times \Omega^l(M) \to \Omega^{k+l}(M), \quad (\omega, \eta) \mapsto \omega \wedge \eta,$$

i.e. defined by

$$(\omega \wedge \eta)_p := \omega_p \wedge \eta_p.$$

Or, using the interpretation of k-forms as in the previous proposition,

$$(\omega \wedge \eta)(X_1, \ldots, X_{k+l}) = \sum_{\sigma \in S_{k,l}} \omega(X_{\sigma(1)}, \ldots, X_{\sigma(k)})\eta(V_{\sigma(k+1)}, \ldots, V_{\sigma(k+l)}).$$

Of course, one still has:

PROPOSITION 5.16. *The wedge operation is bilinear, and:*

(1) *graded-commutativity: for any* $\omega \in \Omega^k(M)$ *and* $\eta \in \Omega^l(M)$,

$$\eta \wedge \omega = (-1)^{kl}\omega \wedge \eta.$$

(2) *associativity: for any differential forms* ω, η *and* ζ *on* M *one has*

$$(\omega \wedge \eta) \wedge \zeta = \omega \wedge (\eta \wedge \zeta).$$

PROOF. Just apply point-wise the similar linear algebra result. □

And, still as for 1-forms, forms of arbitrary degree can be pulled-back via any smooth map $F : M \to N$, giving rise to operations:

$$F^* : \Omega^k(N) \to \Omega^k(M).$$

Explicitly, for $\omega \in \Omega^k(N)$, **the pull-back of** ω **by** F, $F^*\omega \in \Omega^k(M)$, is given by

$$F^*(\omega)_p(X_p^1, \ldots, X_p^k) := \omega_{F(p)}\left((dF)_p(X_p^1), \ldots, (dF)_p(X_p^k)\right).$$

Again, when M is a submanifold of N and F is the inclusion map, the pull-back of ω via the inclusion will be denoted by

$$\omega|_M \in \Omega^k(M)$$

and will be called **the restriction of** ω **to** M. And of particular interest is the case when M is a submanifold of some Euclidean space $\mathbb{R}^d$; then:

- k-forms on $\mathbb{R}^d$ are always of type
$$\sum_{1 \le i_1 < \ldots < i_k \le d} f_{i_1,\ldots,i_k} \cdot dx_{i_1} \wedge \ldots \wedge dx_{i_k}$$
with $f_{i_1,\ldots,i_k}$ smooth functions on $\mathbb{R}^d$.
- one obtains k-forms on M by restricting such k-forms on $\mathbb{R}^d$ to M.

EXERCISE 5.12. Show that the pull-back operation preserves the wedge-products: if $F : M \to N$ is smooth then, for any ω and η differential forms on N, one has

$$F^*(\omega \wedge \eta) = F^*(\omega) \wedge F^*(\eta).$$

REMARK 5.17 (The graded world). It is customary to formally put together all the spaces $\Omega^k(M)$ and consider

$$\Omega^\bullet(M) := \bigoplus_k \Omega^k(M).$$

The notation "$\bullet$" is to indicate that we deal with a graded object; its elements are (finite) sums of homogeneous ones (the ones that live in one single degree). With this, the wedge product becomes now (after extending bi-linearly) a product

$$\cdot \wedge \cdot : \Omega^\bullet(M) \times \Omega^\bullet(M) \to \Omega^\bullet(M)$$

making $\Omega(M)$ into a (graded) algebra. Note the similarity with the algebra of polynomials in several variables, where each polynomial is a sum of homogeneous ones. And, as there, on sums of homogeneous elements, the wedge product is:

$$(\omega_0+\omega_1+\ldots)\wedge(\eta_0+\eta_1+\ldots) = (\omega_0 \wedge \eta_0)+(\omega_0 \wedge \eta_1 + \omega_1 \wedge \eta_0)+(\omega_0 \wedge \eta_2 + \omega_1 \wedge \eta_1 + \omega_2 \wedge \eta_0)+\ldots$$

And here is one very useful slogan/principle when working with graded objects:

The sign rule: *looking at an expression involving elements that have a degree, when interchanging two such elements, say of degrees k and l, the expression changes sign by* $(-1)^{kl}$.

E.g., graded commutativity becomes precisely the one from Proposition 5.16.

4. The exterior (DeRham) differential

In Example 5.3 we mentioned the differential of functions (0-forms) as 1-forms, while in Example 5.15 we indicated a similar operation on from 1-forms to 2-forms. We now extend those to arbitrary degrees, i.e. a sequence of operators d:

$$\Omega^0(M) \xrightarrow{d} \Omega^1(M) \xrightarrow{d} \ldots \xrightarrow{d} \Omega^k(M) \xrightarrow{d} \Omega^{k+1}(M) \xrightarrow{d} \ldots$$

or, equivalently, an operator on the full algebra of differential forms (cf. Remark 5.17),

$$d : \Omega^{\bullet}(M) \to \Omega^{\bullet+1}(M)$$

which rises the degrees by 1 (as indicated by the notation $\bullet + 1$"). From the graded point of view (Remark 5.17) one should think of the symbol d as having degree 1 (and this will dictate some signs later on).

There are quite a few different ways to proceed to introduce d. But, before anything, let us point out the properties that we may want to be satisfied by d; we will see that some of them already force what d should be, or at least its uniqueness. Already at the start, we have to decide whether we want to discuss d (properties, definitions, existence, etc) on a fixed manifold M, or for all manifolds at once.

Of course, first of all we want:

DeRham-0: *d is $\mathbb{R}$-linear and, on 0-forms (functions), it is the usual differential of functions.*

Next, by thinking about the behaviour of d on the product of functions, we would like a similar Leibniz identity w.r.t. the wedge product $\omega \wedge \eta$:

DeRham-1: *d satisfies the (graded) Leibniz identity:*

$$d(\omega \wedge \eta) = d(\omega) \wedge \eta + (-1)^k \omega \wedge d(\eta).$$

for all $\omega \in \Omega^k(M)$, $\eta \in \Omega^l(M)$.

For an explanation of the sign, please see the sign rule from Remark 5.17 and remember that the symbol d is assigned degree 1.

One way to come across d is by applying inductively the reasoning described in Example 5.15: if d was defined on forms up to degree $k-1$, and we wonder when a k-form $\omega \in \Omega^k(M)$ is exact (i.e. of the type $d\eta$ for some $\eta \in \Omega^{k-1}(M)$), looking for "the natural conditions" that ω must satisfy, one discovers an expression $d\omega$ whose non-vanishing is an obstruction to ω being exact; this will be d on the k-forms. Of course, underlying the entire idea is one very simple property that d will have:

DeRham-2: *$d \circ d = 0$.*

EXAMPLE 5.18. On $M = \mathbb{R}^m$, if we assume these properties for d, we already know how to compute d on arbitrary forms. E.g., for

$$\omega = (x^2 + yz)dx \wedge dy,$$

applying the Leibniz rule we find

$$d\omega = d(x^2 + yz) \wedge dx \wedge dy + (x^2 + yz)d(dx \wedge dz).$$

For the term $d(x^2 + yz)$ we get, using "DeRham-0", $2xdx + ydz + zdy$. The term $d(dx \wedge dz)$ is, by Leibniz and $d^2 = 0$, zero. We obtain

$$d\omega = 2xdx \wedge dx \wedge dy + ydz \wedge dx \wedge dy + zdy \wedge dx \wedge dy.$$

Remembering that $dx \wedge dx = 0$ (by graded commutativity of the wedge), and similarly $zdy \wedge dx \wedge dy = 0$, we are left with

$$d\omega = ydz \wedge dx \wedge dy,$$

or, to keep the natural order of the variables,

$$d\omega = ydx \wedge dy \wedge dz.$$

Actually, a similar reasoning forces the definition of d on all k-forms on $\mathbb{R}^m$: any k-form is a sum of forms of type

$$fdx_{i_1} \wedge \ldots \wedge dx_{i_k}$$

and, by the same arguments as above,

$$d(fdx_{i_1} \wedge \wedge dx_{i_k}) = \sum_i \frac{\partial f}{\partial x_i} dx_i \wedge dx_{i_1} \wedge \wedge dx_{i_k}.$$

The fact that the d defined in this way does indeed satisfy (DeRham-1) and (DeRham-2) is now a simple check forms of this type.

Of course, exactly the same applies to any open subset $U \subset \mathbb{R}^m$ or to any domain U of a coordinate chart (U, χ) in any manifold M (we would use then the 1-forms $d\chi_i$ instead of dx_i); given the uniqueness of d satisfying the axioms, it follows that d doesn't even depend on the actual χ that we use to write down the formulas.

COROLLARY 5.19. *If M is a manifold, then for any $U \subset M$ that is the domain of a coordinate chart, there is and is unique an operator*

$$d_U : \Omega^\bullet(U) \to \Omega^{\bullet+1}(U)$$

satisfying (DeRham-0), (DeRham-1) and (DeRham-2). It is given by

$$d_U(fd\chi_{i_1} \wedge \wedge d\chi_{i_k}) = \sum_i \frac{\partial f}{\partial \chi_i} d\chi_i \wedge d\chi_{i_1} \wedge \wedge d\chi_{i_k}$$

where $\chi : U \to \Omega \subset \mathbb{R}^m$ is any coordinate chart defined on U (but independent of the actual χ).

Therefore, an obvious way to proceed to define d on a general M is to "proceed locally". That would mean that we would be looking for operators d that are local in the following sense:

DeRham-3: *Locality: if $U \subset M$ is open and $\omega \in \Omega^\bullet(M)$ is supported inside U, then so is $d\omega$.*

As for functions, the support of ω is the closure of the set of points $p \in M$ at which $\omega_p \neq 0$.

EXERCISE 5.13. Show that a linear operator $d : \Omega^\bullet(M) \to \Omega^{\bullet+1}(M)$ is local if and only if for any open $U \subset M$ there is a linear operator $d|_U : \Omega^\bullet(U) \to \Omega^{\bullet+1}(U)$ with the property that

$$(d\,\omega)|_U = d|_U(\omega|_U) \quad \text{for all } \omega \in \Omega^\bullet(U).$$

$$\begin{array}{ccc} \Omega^\bullet(M) & \xrightarrow{d} & \Omega^{\bullet+1}(M) \\ {\scriptstyle \text{restr}}\downarrow & & \downarrow{\scriptstyle \text{restr}} \\ \Omega^\bullet(U) & \xrightarrow{d|_U} & \Omega^{\bullet+1}(U) \end{array}.$$

Then show that, given two local operators d and d', $d = d'$ if and only if M can be covered by opens U on which $d|_U = d^{'}|_U$.

COROLLARY 5.20. *On any manifold M there exists and is unique an operator $d : \Omega^\bullet(M) \to \Omega^{\bullet+1}(M)$ which is local and so that for any domain U of a coordinate chart, $d|_U$ coincides with d_U from the previous corollary. Moreover, d will automatically satisfy (DeRham-0), (DeRham-1) and (DeRham-2).*

PROOF. The definition is forced: for $p \in M$, choosing a coordinate chart (U, χ) around p we must have

$$(d\omega)_p = (d_U \omega|_U)_p.$$

Given the uniqueness property of the d_Us it follows that the right hand side does not depend on the coordinate chart that we use, hence it can be used to define $d\omega$ unambiguously. Since the conditions (DeRham-1) and (DeRham-2) are local and they hold for the d_Us, it follows that they hold for d as well. □

It is interesting to note that locality can be obtained from the rest of the axioms:

LEMMA 5.21. *On a manifold M, if $D : \Omega^\bullet(M) \to \Omega^{\bullet+1}(M)$ satisfies the condition that*

$$D(f \cdot \omega) = fD(\omega) + df \wedge \omega \quad \textit{for all } f \in \mathcal{C}^\infty(M), \omega \in \Omega^\bullet(M)$$

(in particular if D satisfies (DeRham-0) and (DeRham-1)) then D is local.

PROOF. We fix ω, we denote by A its support and we claim that the support of $d\omega$ is inside A. Note that A is the largest closed subset of M outside which ω vanishes. Therefore suffices to show that $d(\omega)$ vanishes outside A. Let $p \in M \setminus A$; since A is closed, $M \setminus A$ is open, hence we find a smooth function $\eta \in \mathcal{C}^\infty(M)$ that is zero on A and is 1 in a neighborhood of p. But then $\omega = (1 - \eta) \cdot \omega$ and, using the condition from the statement we obtain

$$d\omega = (1 - \eta)d\omega - d\eta \wedge d\omega.$$

Evaluating at the point p we obtain that $d\omega$ vanishes at p. □

In particular,

COROLLARY 5.22. *On any manifold M there exists and is unique an operator $d : \Omega^\bullet(M) \to \Omega^{\bullet+1}(M)$ which satisfies (DeRham-0) , (DeRham-1) and (DeRham-2).*

PROOF. One just needs to remark that, if d satisfies (DeRham-1) and (DeRham-2) then so do all its restrictions $d|_U$; for domains U of coordinate charts, due to the uniqueness of d_U, we obtain that $d|_U = d_U$. □

DEFINITION 5.23. *The resulting operator $d : \Omega^\bullet(M) \to \Omega^{\bullet+1}(M)$ is called the **DeRham operator on** M (or the exterior differential).*

While so far we concentrated on building d on a given manifold M, one can proceed slightly differently and handle all the manifolds at once; then the natural axiom to impose is the behaviour of d when we change the manifold, i.e. its compatibility with pull-backs. In low degree this was seen in Exercise 5.6. It means that, for any smooth map $F : M \to N$ one has a commutative diagram

$$\begin{array}{ccc} \Omega^{\bullet}(N) & \xrightarrow{d} & \Omega^{\bullet+1}(N) \\ \downarrow F^* & & \downarrow F^* \\ \Omega^{\bullet}(M) & \xrightarrow{d} & \Omega^{\bullet+1}(M) \end{array} ,$$

i.e. $F^* \circ d = d \circ F^*$. Note that this condition forces automatically that each d is local. Therefore:

COROLLARY 5.24. *If we are looking for operators $d_M : \Omega^{\bullet}(M) \to \Omega^{\bullet+1}(M)$, one for each manifold M, such that they are compatible with pull-backs, we have uniqueness and existence under any of the following requirements:*

(1) *on $\mathbb{R}^m$, $d_{\mathbb{R}^m}$ is the operator previously discussed.*
(2) *all d_M satisfy (DeRham-0) and (DeRham-1).*

Of course, for any manifold M, d_M will be precisely the DeRham operator.

And here is another way to introduce the DeRham operator: just right away, by explicit formulas. One way to find those formulas would be by returning to the previous idea of proceeding like in Example 5.15 inductively, by looking for the natural conditions for a form to be exact.

PROPOSITION 5.25. *One has the following explicit formula for the DeRham operator on any manifold M: for $\omega \in \Omega^k(M)$, $d\omega \in \Omega^{k+1}(M)$ is given by*

$$d\omega(X_0, \ldots, X_k) = \sum_{i=0}^{k} (-1)^i L_{X_i}(\omega(X_0, \ldots, \hat{X}_i, \ldots, X_k)) + \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_0, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots X_k)) \quad (4.1)$$

where the elements under the hat are deleted.

PROOF. Note that the previous formula is skew-symmetric and $\mathcal{C}^{\infty}(M)$-multilinear (check that!). Hence it defines an operator d_M for each M. Looking back at the previous corollary, we see that all that is used is the compatibility with pull-backs by diffeomorphisms and by inclusions of opens inside manifolds. That is clearly true for the d_Ms. Hence we are left with checking that, for $\mathbb{R}^m$, this formula gives back the standard DeRham operator. Moreover, a simple check shows that $d(f \cdot \omega) = f \cdot d\omega + df \wedge \omega$ for any function f, hence it suffices to restrict our attention to forms on $\mathbb{R}^m$ of type $\omega = dx_{i_1} \wedge \ldots \wedge dx_{i_k}$ and we are left with showing that

$$d_{\mathbb{R}^m}(dx_{i_1} \wedge \ldots \wedge dx_{i_k}) = 0$$

(where recall that $d_{\mathbb{R}^m}$ now denotes the operator from the statement). This can be done in several ways:

- direct computation (straightforward but not so nice).
- check it first for $k = 1$ (obvious) and then check the Leibniz identity on expressions of form $\omega \wedge \eta$ where ω is a 1-form (and this computation is not that ugly and can easily be carried out on a general manifold M).

□

EXERCISE 5.14. Check by direct computation that, defining d by the formula (4.1), then it satisfies the Leibniz identity on expressions of form $\omega \wedge \eta$ where ω is a 1-form.

5. Cartan's magic formula $di_X + i_X d = \mathcal{L}_X$

We first discuss the Lie derivative of forms $\omega \in \Omega^\bullet(M)$ along vector fields $V \in \mathfrak{X}(M)$, giving rise to operations

$$\mathcal{L}_V : \Omega^\bullet(M) \to \Omega^\bullet(M).$$

These Lie derivatives fit in the general philosophy explained and illustrated in Section **??** of the previous chapter:

Definition 5.26. *For $\omega \in \Omega^k(M)$ and $V \in \mathfrak{X}(M)$ define **the Lie derivative of** ω **along** V as*

$$\mathcal{L}_V(\omega) := \left.\frac{d}{dt}\right|_{t=0} (\phi_V^t)^*\omega \in \Omega^k(M).$$

As before, when V is not complete, one has to look at each $p \in M$ and not that the resulting formula

$$\mathcal{L}_V(\omega)(X_p^1, \ldots, X_p^k) = \left.\frac{d}{dt}\right|_{t=0} \omega_{\phi_V^t(p)}\left((d\phi_V^t)_p(X_p^1), \ldots, (d\phi_V^t)_p(X_p^k)\right) \tag{5.1}$$

makes sense independent of whether V is complete or not.

Not also that, from the graded point of view, the operation $\mathcal{L}_V$ has degree 0 (because it does not change the degree of the forms).

Proposition 5.27. *For any vector field $V \in \mathfrak{X}(M)$:*

(1) *$\mathcal{L}_V : \Omega^\bullet \to \Omega^\bullet(M)$ is a degree zero derivation, i.e. it is linear and satisfies*

$$\mathcal{L}_V(\omega \wedge \eta) = \mathcal{L}_V(\omega) \wedge \eta + \omega \wedge \mathcal{L}_V(\eta)$$

for any two differential forms ω and η.

(2) *$\mathcal{L}_V$ commutes with d: $\mathcal{L}_V \circ d = d \circ \mathcal{L}_V$.*

(3) *$\mathcal{L}_V$ is the only degree zero derivation on $\Omega^\bullet(M)$ which commutes with d and which, on functions, it is the usual derivative L_V.*

Moreover, it can be described by the explicit formula:

$$\mathcal{L}_V(\omega)(X_1, \ldots, X_k) = L_V(\omega(X_1, \ldots, X_k)) + \sum_{i=1}^{k} \omega(X_1, \ldots, X_{i-1}, [V, X_i], X_{i+1}, \ldots, X_k).$$

Proof. The derivation identity follows from the standard one and the fact that $(\phi_V^t)^*$ is compatible with the wedge operation:

$$\left.\frac{d}{dt}\right|_{t=0} (\phi_V^t)^*(\omega\wedge\eta) = \left.\frac{d}{dt}\right|_{t=0} \left((\phi_V^t)^*\omega \wedge (\phi_V^t)^*\eta\right) = \left(\left.\frac{d}{dt}\right|_{t=0} (\phi_V^t)^*\omega\right)\wedge\eta+\omega\wedge\left(\left.\frac{d}{dt}\right|_{t=0} (\phi_V^t)^*\eta\right).$$

That $\mathcal{L}_V$ commutes with d follows from the similar property of the pull-backs $(\phi - V^t)^*$.

For the uniqueness in (3), assume that D is another operators with the same properties. It follows that $\mathcal{L}_V(df) = dL_V(f)$, and we obtain that $D(\omega) = \mathcal{L}_V(\omega)$ on differential forms of type

$$\omega = \sum_i f_i \wedge dg_i^1 \wedge \ldots \wedge dg_i^k.$$

It would suffices to show that any k-form can be written in this way. This is true and it would follow for instance if we proved that M can be embedded in some Euclidean space (why would that be enough?). To avoid using such a result, one can first restrict to manifolds on which it is certainly true: domain of coordinate charts. In more detail: due to the Leibniz identity it follows (as in the discussions about d) that D is local, we can talk about $D|_U$ and, on domains of coordinate charts, $D|_U$ still has the same properties as D. Hence $D|_U = \mathcal{L}_V|_U$, and then $D = \mathcal{L}_V$.

For the final formula, we start from the explicit formula (5.1) and we use that $\frac{d}{dt}|_{t=0}(\phi_V^t)^*(X) = [V, X]$ (Proposition 4.42 in the previous chapter) or, equivalently,

$$\frac{d}{dt}\bigg|_{t=0}(d\phi_V^{-t})_{\phi_V^t(p)}(X_{\phi_V^t(p)}) = [V, X]_p$$

for all X. This allows us to write

$$(d\phi_V^{-t})_{\phi_V^t(p)}(X_{\phi_V^t(p)}) = X_p + t[V, X]_p + o(t^2)$$

for t near 0, and then

$$(d\phi_V^t)_p(X_p) = X_{\phi_V^t(p)} + t[V, X]_{\phi_V^t(p)} + o(t^2).$$

Inserting this in (5.1) and using also Proposition 4.41 from the previous chapter (applied to V and $f = \omega(X_1, \ldots, X_k)$ we find the formula from the statement. □

Another interesting operators induced by a vector field $V \in \mathfrak{X}(M)$, but simpler this time, is the interior product:

Definition 5.28. *For $\omega \in \Omega^k(M)$ and $V \in \mathfrak{X}(M)$ define* ***the interior product*** *ω* ***by*** *V, denoted $i_V(\omega)$ as the differential form of degree $k-1$ given by*

$$i_V(\omega)(X_1, \ldots, X_{k-1}) = \omega(V, X_1, \ldots, X_{k_1}).$$

This defines an operator

$$i_V : \Omega^\bullet(M) \to \Omega^{\bullet-1}(M);$$

since it lowers the degree by 1, we say it is of degree -1.

Proposition 5.29. *For any vector field $V \in \mathfrak{X}(M)$,i_V is a degree -1 derivation, i.e. it is linear and satisfies*

$$i_V(\omega \wedge \eta) = i_V(\omega) \wedge \eta + (-1)^k \omega \wedge i_V(\eta)$$

for any two differential forms $\omega \in \Omega^k(M)$ and η. Actually, it is the only degree -1 derivation on $\Omega^\bullet(M)$ which, on 1-forms, is the evaluation on V.

Proof. The derivation formula is a completely straightforward computation. For the uniqueness, one can just imitate the arguments used for the uniqueness of $\mathcal{L}_V$, just that it is a bit simpler now (being degree -1, it kills of the function, and the condition from the statement prescribes i_V on 1-forms hence, by the Leibniz identity, on all the sums of wedges of 1-forms. One can either show that any differential form can be written as such a sum, or just provide a local argument completely similar to that for $\mathcal{L}_V$. □

Next: the Cartan magic formula, which says that $\mathcal{L}_V$ is the graded commutator of d and i_V; given that d has degree 1 and i_V has degree -1, that means:

Theorem 5.30. *For any vector field $V \in \mathfrak{X}(M)$ one has*

$$\mathcal{L}_V = d \circ i_V + i_V \circ d.$$

Proof. A simple computation using the Laibinz identities for d and i_V (with special care at the signs involved) shows that $d \circ i_V + i_V \circ d$ is a derivation of degree zero. Since $d \circ d = 0$, we see that it also commutes with d. Moreover, on functions f it gives $i_V d(f) = L_V(f)$. Hence, by the uniqueness from Proposition 5.27, it must coincide with $\mathcal{L}_V$. □

EXERCISE 5.15. Assume that we have a right action of the circle S^1 on a manifold M, i.e. a smooth map

$$M \times S^1 \to M, \quad (\lambda, p) \mapsto \lambda \cdot p$$

such that $p \cdot p = p$ and $(p \cdot \lambda) \cdot \lambda' = p \cdot (\lambda\lambda')$ for all $\lambda, \lambda' \in S^1$, $p \in M$. Define the following vector field on M:

$$V_p = \frac{d}{dt}\Big|_{t=0} p \cdot e^{2\pi\, it}.$$

Assume now that $\omega \in \Omega^1(M)$ is a 1-form that is invariant in the sense that

$$R_\lambda^*(\omega) = \omega \quad \text{for all } \lambda \in S^1,$$

where $R_\lambda : M \to M$ is the right multiplication by λ. Show that:

(1) Also $d\omega$ is S^1-invariant.
(2) If $i_V(\omega) = 1$ then $i_V(d\omega) = 0$.

6. Integration on manifolds

In this section we explain how to extend standard integration from $\mathbb{R}^m$ to more general manifolds. By standard integration we mean the integral

$$\int : \mathcal{C}_c^\infty(\mathbb{R}^m) \to \mathbb{R} \quad f \mapsto \int f = \int_{\mathbb{R}^m} f(x)dx = \int_{\mathbb{R}^m} f(x_1, \ldots, x_m)dx_1 \ldots dx_m$$

defined (only) on the space of compactly supported smooth functions. For $U \subset \mathbb{R}^m$ open and $f \in \mathcal{C}_c^\infty(\Omega)$, we will use the notation

$$\int_\Omega f = \int_\Omega f(x)dx$$

for the integral $\int \tilde{f}$ where $\tilde{f}$ is f extended by zero outside Ω.

Actually, what we need to know about the standard integration is that:

- it is linear in f.
- change of variables formula: if $c : \Omega \to \Omega'$ is a diffeomorphism between two opens in $\mathbb{R}^m$ then, for any $f \in \mathcal{C}_c^\infty(\Omega')$,

$$\int_{\Omega'} f(x')dx' = \int_\Omega |\mathrm{Jac}_x(c)| \cdot f(c(x))dx$$

or, more compactly,

$$\int_{\Omega'} f = \int_\Omega |\mathrm{Jac}(c)| \cdot f \circ c.$$

Recall that the Jacobian of c is the function

$$\mathrm{Jac}(c) : \Omega \to \mathbb{R}, \quad x \mapsto \mathrm{Jac}_x(c) := \det\left(\frac{\partial c_i}{\partial x_j}(x)\right).$$

Note that its presence in the change of variable formula shows that the integral of a function f is not independent of the coordinates. For that reason, the objects that can (naturally) be integrated on manifolds will not be functions (we will be able to integrate functions only when a certain extra-choice will be made). Instead, one has to look for objects which locally (in coordinate charts) can be represented by smooth functions, and so that when we change the coordinates the representing functions change by a Jacobian factor. This brings us to differential forms of top degree m (also making more sense of the symbol of the expression "$f(x_1, \ldots, x_m)dx_1 \ldots dx_m$" in the integration).

To give the details, we fix an m-dimensional manifold M and we look at differential forms of top degree

$$\omega \in \Omega^m(M).$$

Note that, for any coordinate chart $\chi : U \to \Omega \subset \mathbb{R}^m$, one can write

$$\omega|_U = f_\chi^\omega \circ \chi \cdot d\chi_1 \wedge \ldots \wedge d\chi_m$$

for a unique function $f_\chi^\omega \in \mathcal{C}^\infty(\Omega')$; this will be called **the representing function of ω w.r.t. the chart** (U, χ). How does the representing function change when we change the chart? Well, if $\chi' : U' \to \Omega'$ is another chart, denoting by $c = \chi' \circ \chi^{-1} : \Omega \to \Omega'$ the change of coordinates, we know that

$$d\chi_i^{'} = \sum_j \frac{\partial \chi_i^{'}}{\partial \chi_j} \cdot d\chi_j = \sum_j \frac{\partial c_i}{\partial x_j} \circ \chi \cdot d\chi_j.$$

From this, and the graded commutativity of the wedge-product, we deduce that

$$d\chi_1^{'} \wedge \ldots \wedge d\chi_m^{'} = \det\left(\frac{\partial \chi_i^{'}}{\partial \chi_j}\right) \cdot d\chi_1 \wedge \ldots \wedge d\chi_m,$$

where the coefficient is

$$\det\left(\frac{\partial \chi_i^{'}}{\partial \chi_j}\right) = \det\left(\frac{\partial c_i}{\partial x_j} \circ \chi\right) = \mathrm{Jac}(c \circ \chi).$$

We deduce right away that

$$f_\chi = \mathrm{Jac}(c) \cdot f_{\chi'} \circ c.$$

This is "almost" the expression in the change of variables formula- only the absolute value is missing. For that reason we stop for a bit and look into this issue.

6.1. Digression: oriented manifolds. We start with the linear discussion. For a finite dimensional vector space V, we denote by $\mathrm{Fr}(V)$ the set of all frames (i.e. ordered bases) of V. For two frames $u, u' \in \mathrm{Fr}(V)$ we denote by $[u : u']$ the matrix of coordinate changes from u to u' (an $m \times m$ invertible matrix with entires real numbers, where $m = \dim(V)$) and by $\mathrm{sign}(u : u')$ the sign of its determinant:

$$\mathrm{sign}(u : u') := \mathrm{sign}\left(\det([u : u'])\right) \in \{-1, +1\}.$$

An orientation on V is basically a partition of the set of frames into two: positively and negatively oriented frames, so that two frames u and u' are of the same type if and only if $\mathrm{sign}(u : u') > 0$.

One way to formalize this is as follows: call an **orientation** on V any map

$$\mathcal{O} : \mathrm{Fr}(V) \to \{-1, +1\}$$

with the property that

$$\mathcal{O}(u') = \mathrm{sign}(u : u')\mathcal{O}(u) \cdot \quad \text{for all } u, u' \in \mathrm{Fr}(V).$$

We also say that $(V, \mathcal{O})$ is an oriented vector space, and then

$$\mathrm{Fr}(V) = \mathrm{Fr}_+(V, \mathcal{O}) \cup \mathrm{Fr}_-(V, \mathcal{O}), \quad \text{where } \mathrm{Fr}_\pm(V, \mathcal{O}) = \{u \in \mathrm{Fr}(V) : \mathcal{O}(u) = \pm 1\}$$

(positively and negatively oriented frames of $(V, \mathcal{O})$).

Note that any $u \in \mathrm{Fr}(V)$ gives rise to a unique orientation $\mathcal{O}_u$ by the condition that u is positively oriented for $\mathcal{O}_u$; explicitly,

$$\mathcal{O}_u : \mathrm{Fr}(V) \to \{-1, +1\}, \quad u' \mapsto \mathrm{sign}(u : u').$$

And, moreover, for $u_0, u_1 \in \mathrm{Fr}(V)$,

$$\mathcal{O}_{u_0} = \mathcal{O}_{u_1} \Longleftrightarrow \mathrm{sign}(u_0 : u_1) > 0.$$

Hence an orientation on V can also be seen as an equivalence class of frames with respect to the equivalence relation based on the condition $\text{sign}(u_0 : u_1) > 0$.

Moving to a manifold M, we look at families

$$\mathcal{O} = \{\mathcal{O}_p : p \in M\} \quad \text{with } \mathcal{O}_p - \text{orientation on } T_pM. \tag{6.1}$$

We say that $\mathcal{O}$ is **smooth** if for any $p_0 \in M$ there is a chart (U, χ) around p_0 such that, for any $p \in U$,

$$\left(\frac{\partial}{\partial \chi_1}\right)_p, \ldots, \left(\frac{\partial}{\partial \chi_m}\right)_p$$

is a (positively) oriented frame of $(T_pM, \mathcal{O}_p)$. Such a chart is said to be **adapted to** $\mathcal{O}$.

DEFINITION 5.31. *An* ***orientation on the manifold*** *M is any family $\mathcal{O}$ of orientations $\mathcal{O}_p$ on the tangent spaces T_pM which is smooth in the sense described above.*

We also say that $(M, \mathcal{O})$ is an oriented manifold, and the charts adapted to $\mathcal{O}$ will be called ***positively oriented charts*** *of $(M, \mathcal{O})$.*

Of course, on an oriented manifold, we will only use positively oriented charts. From the previous discussion it should be clear now that

$$\text{Jac}(\chi' \circ \chi^{-1}) > 0 \quad \text{for all } \chi, \chi' - \text{positively oriented.}$$

EXERCISE 5.16. Let M be a manifold.

(1) for an atlas $\mathcal{A}$ on M show that the existence of an orientation $\mathcal{O}_{\mathcal{A}}$ on M with the property that each $\chi \in \mathcal{A}$ is adapted to $\mathcal{O}_{\mathcal{A}}$ is equivalent to

$$\text{Jac}(\chi' \circ \chi^{-1}) > 0 \quad \text{for all } \chi, \chi' \in \mathcal{A}. \tag{6.2}$$

(2) show that two atlases $\mathcal{A}$ and $\mathcal{A}'$ with this property induce the same orientation iff

$$\text{Jac}(\chi' \circ \chi^{-1}) > 0 \quad \text{for all } \chi \in \mathcal{A}, \chi' \in \mathcal{A}'. \tag{6.3}$$

In other words, to chose an orientation on M is the same thing as choosing an atlas $\mathcal{A}$ satisfying (6.2), and two such atlases $\mathcal{A}$ and $\mathcal{A}'$ induce the same orientation on M if an only if (6.3) holds.

EXERCISE 5.17. If M is connected and orientable (i.e. it admits an orientation), show that it admits exactly two orientations.

End of the digression on oriented manifolds

Returning to our discussion on integration, it should be clear now that the choice of an orientation on M is precisely what we need to solve our small problem with the absolute value of the Jacobians; here is the final outcome:

THEOREM 5.32. *On any oriented manifold $(M, \mathcal{O})$, there is a well-defined map*

$$\int_M : \Omega_c^m(M) \to \mathbb{R}$$

which is uniquely determined by the following two properties:

(1) *$\int_M$ is linear.*
(2) *for any positively oriented chart $\chi : U \to \Omega$, if $\omega \in \Omega_c^m(M)$ is supported in U then*

$$\int_M \omega = \int_\Omega f_\chi^\omega(x)dx \quad (\textit{standard integration over } \Omega \subset \mathbb{R}^m), \tag{6.4}$$

where f_χ^ω is the representing function of ω w.r.t. the chart (U, χ).

PROOF. We divide the proof into three steps:

Step 1: integrating over domains of charts: It should now be clear from the previous discussions that for any $U \subset M$ which is inside the domain of some chart of M one obtains an integration map

$$\int_U : \Omega_c^m(U) \to \mathbb{R}$$

(defined on the space of m-forms whose support is compact and included in U); it is defined simply by the right hand side of (6.4), but the point is that it does not depend on the choice of a positively oriented chart χ defined on U.

Note also that for any open $V \subset U$, one has a commutative diagram

$$\begin{array}{ccc} \Omega_c^m(V) & \hookrightarrow & \Omega_c^m(U) \\ & {\scriptstyle \int_V} \searrow \quad \swarrow {\scriptstyle \int_U} & \\ & \mathbb{R} & \end{array}, \text{ i.e. } \quad \int_V \omega = \int_U \omega \quad \text{for all } \omega \in \Omega_c^m(V).$$

This implies that we can safely write $\int_M \omega$ for any $\omega \in \Omega_c^m(M)$ which is supported inside the domain of a chart: it actually does not depend on which domain we use.

Step 2: from domains of charts to the entire M: Now we claim that for any $\omega \in \Omega_c^m(M)$, can be written as a finite sum

$$\omega = \sum_{i=1}^{k} m\omega^i$$

where each $\omega_i \in \Omega_c^m(M)$ is supported in the domain of a chart (and k is some integer). Once we prove this, we will define

$$\int_M \omega := \sum_i \int_M \omega_i. \tag{6.5}$$

To prove the claim one uses a partition of unity argument: if K is the support of ω, since K is compact it can be covered by a finite number of opens $U_1, \ldots, U_k$, each of them a domain of a coordinate chart, and each one of them relatively compact (in the sense that they have compact closures) has compact closure. Then $\{M \setminus K, U_1, \ldots, U_k\}$ is an open cover of M and we choose a partition of unity $\eta_0, \eta_1, \ldots, \eta_k$ with η_0 supported inside $M \setminus K$ and η_i supported inside U_i. Note that, since U_i is relatively compact, each η_i is compactly supported. Also, since η_0 is supported in $M \setminus K$, one has $\eta_0 \cdot \omega = 0$. Therefore, multiplying the identity $1 = \sum_{i=0}^{k} \eta_i$ by ω we obtain

$$\omega = \sum_{i=1}^{k} \omega_i \quad \text{where } \omega_i = \eta_i \cdot \omega \in \Omega_c^m(M).$$

Step 3: $\int_M \omega$ does not depend on the decomposition $\omega = \sum_i \omega_i$: Taking now (6.5) as the definition of $\int_M$, we have to check that the outcome does not depend on the way we write ω as a sum of forms supported inside domains of charts.

To prove this, assume that we have two such writings

$$\omega = \sum_{i=1}^{k} \omega_i = \sum_{j=1}^{k'} \omega_j^{'}$$

Let K be a compact that contains the supports of all ω_is and ω_js. As before, we can find a finite family $\{\eta_\alpha\}_{\alpha \in I}$ of compactly supported smooth functions, each supported inside a chart domain V_α, such that $\sum_\alpha \eta_\alpha = 1$. Then $\omega_i = \sum_\alpha \eta_\alpha \cdot \omega_i$. Note also that we already know that $\int_M$

behaves linearly when looking at forms that are supported inside a coordinate chat. Applying this for the coordinate charts U_i, and then V_α, we find that $\sum_{i=1}^{k}\int_M \omega_i$ equals to

$$\sum_{i=1}^{k}\sum_{\alpha}\int_M \eta_\alpha\cdot\omega_i = \sum_{\alpha}\sum_{i=1}^{k}\int_M \eta_\alpha\cdot\omega_i = \sum_{\alpha}\int_M \eta_\alpha\cdot\omega.$$

Of course, the same applies to $\sum_{j=1}^{k'}\omega_j'$, therefore proving the claim.

Finally, note also that $\int_M$ defined in this way is obviously linear. □

THEOREM 5.33 (Stokes- version without bounday). *If M is an m-dimensional oriented manifold then, for any $\eta\in\Omega_c^{m-1}(M)$ one has*

$$\int_M d\eta = 0.$$

In particular, on a compact oriented manifold, $\int_M \omega = 0$ whenever ω is exact.

PROOF. Writing (as above) η as a sum of forms supported in domains of coordinate charts, we may assume that $M=\mathbb{R}^m$. Then η is of type

$$\eta = \sum_i f_i\cdot dx_1\wedge\ldots dx_{i-1}\wedge dx_{i+1}\wedge\ldots\wedge dx_m, \tag{6.6}$$

so that we can write

$$d\eta = \left(\sum_i (-1)^{i-1}\frac{\partial f_i}{\partial x_i}\right) dx_1\wedge\ldots\wedge dx_m.$$

Hence it suffices to show that, for any $f\in\mathcal{C}^\infty(\mathbb{R}^m)$ compactly supported and any i one has

$$\int_{\mathbb{R}^m}\frac{\partial f}{\partial x_i}dx_1\wedge\ldots\wedge dx_m = 0.$$

Choosing an interval $[a,b]$ such that the support of f is strictly inside $[a,b]^m$, everything follows from

$$\int_a^b \frac{\partial f_i}{\partial x_i}dx_i = f(b)-f(a) = 0. \tag{6.7}$$

□

EXERCISE 5.18. Do again Exercise 5.3.

6.2. Volume forms; integration of functions. We now return to manifolds M that are compact without boundary. We have seen that an orientation $\mathcal{O}$ on M allows us to talk about the integral of m-forms (m being the dimension of M),

$$\int_M : \Omega^m(M)\to\mathbb{R}.$$

Therefore, an integration of functions could be defined as soon as we fix a top-form

$$\mu\in\Omega^m(M),$$

by considering the integral of $f\cdot\mu$:

$$\mathcal{C}^\infty(M)\to\mathbb{R},\quad f\mapsto\int_M f\cdot\mu.$$

In some sense, this is precisely how the standard integration on $\mathbb{R}^m$ (of functions!) is obtained: it is associated to the choice of the the top-form

$$\mu = dx_1\wedge\ldots\wedge dx_m$$

(note the compatibility with the standard notation $\int_{R^m} f(x)dx_1 \dots dx_m$ for the integration of functions!).

Something similar can be done on any manifold, once a "volume form" is fixed.

DEFINITION 5.34. *A volume form on an m-dimensional manifold M is any top form*

$$\mu \in \Omega^m(M)$$

with the property that $\mu_p \neq 0$ for all $p \in M$.

EXAMPLE 5.35. Of course, on $\mathbb{R}^{m+1}$,

$$\mu_{\mathbb{R}^{m+1}} := dx_0 \wedge dx_1 \wedge \dots \wedge dx_m$$

is a volume form. Taking the interior product with the radial vector field $R = \sum_i x_i \cdot \frac{\partial}{\partial x_i}$ we obtain an m-form

$$i_R(\mu_{\mathbb{R}^{m+1}}) = \sum_{i=0}^{m} (-1)^i x_i \cdot dx_0 \wedge dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_m \in \Omega^m(\mathbb{R}^m).$$

Its restriction to the sphere $S^m \subset \mathbb{R}^{m+1}$,

$$\mu_{S^m} := \left(\sum_{i=0}^{m} (-1)^i x_i \cdot dx_0 \wedge dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_m \right) \Bigg|_{S^m} \in \Omega^m(S^m)$$

is a volume form on S^m.

EXERCISE 5.19. Show that, indeed, μ_{S^m} is a volume form on S^m.

The relationship with orientations is:

LEMMA 5.36. *A manifold M is orientable if and only if it admits a volume form $\mu \in \Omega^m(M)$. Moreover, any such volume form μ induces an orientation $\mathcal{O}_\mu$ on M, uniquely determined by the condition that, at any $p \in M$, a frame*

$$v_p^1, \dots, v_p^m$$

of T_pM is positively oriented if and only if

$$\mu_p(v_p^1, \dots, v_p^m) > 0.$$

Furthermore, using the integration induced by this orientation, the volume of (M, μ),

$$Vol(M, \mu) := \int_M \mu,$$

is strictly positive.

Therefore, once one fixes a volume form μ, one has an induced orientation $\mathcal{O}_\mu$, this orientation gives rise to an integration $\int_M : \Omega^m(M) \to \mathbb{R}$ and, using again μ, one obtains an integration of functions

$$f \mapsto \int_M f \cdot \mu.$$

And the integration of the constant function 1 is $\mathrm{Vol}(M, \mu)$.

PROOF. The orientation $\mathcal{O}_\mu$ can be defined, at each $p \in M$, as the map

$$\mathcal{O}_{\mu,p} : \mathrm{Fr}(T_pM) \to \{-1, +1\},$$

$$\mathcal{O}_{\mu,p}(v_p^1, \dots, v_p^m) := \mathrm{sign}(\mu(v_p^1, \dots, v_p^m)).$$

The fact that $\mu_p \neq 0$ ensures that the expression under the sign is non-zero.

The main thing still to prove is the converse: if M is orientable, then it admits a volume form. While volume forms clearly exist locally, the question is how to "glue them" to global volume forms. In principle, we would consider a family $\{\omega_i\}$ of volume forms on domains U_i of coordinate charts covering M indexed by a set I, a partition of unity $\{\eta_i : i \in I\}$ subordinated to the cover and set

$$\omega := \sum_i \eta_i \cdot \omega_i \in \Omega^m(M).$$

For this to be a volume form we would need to check that, for $p \in M$, ω_p is non-zero. Denoting by $J \subset I$ the set of indices j with the property that $p \in \text{supp}(\eta_j)$ (a finite number of them!),

$$c_j = \eta_j(p) \in \mathbb{R}, \quad V = T_pM, \quad , \mu_j = \omega_i, p \in \Lambda^m V^*,$$

we end with a finite sum

$$\sum_{j \in J} c_j \cdot \mu_j \in \Lambda^m V^*, \quad \text{with} \quad \sum_j c_j = 1, \quad \mu_j \neq 0$$

(an affine combination of non-zero elements in the top exterior power). The problem is that this sum may be zero- and it is here that we should the use the orientation $\mathcal{O}$ of M. Indeed, if we knew that V comes with an orientation and each μ_j was not only non-zero, but positively oriented, then the above sum is clearly still positively oriented, hence non-zero. Hence, all we have to do in the previous argument is to start right from the beginning with a family $\{\omega_i\}$ of volume forms on domains U_i which are positively oriented there with respect to the original orientation (i.e, at each $p \in U_i$, $\omega_{i,p}$ is so w.r.t. the orientation $\mathcal{O}_p$). □

6.3. Stokes on manifolds with boundary. The Stokes formula discussed above, when applied to functions on $M = \mathbb{R}$, or on intervals $(a, b) \subset \mathbb{R}$, boils down to the fact that if $f : \mathbb{R} \to \mathbb{R}$ is supported inside the interval (a, b) then

$$\int_{\mathbb{R}} \frac{\partial f}{\partial x}(x)dx = \int_a^b \frac{\partial f}{\partial x}(x)dx = f(b) - f(a) = 0.$$

However, such formulas are especially useful when f is supported inside (or just defined on) a closed interval $[a, b]$, without assuming that $f(a) = f(b) = 0$. This shows that a Stokes theorem would be (even) more interesting when formulated on manifolds with boundary.

The notion of manifold with boundary was briefly mentioned in Section 4 of Chapter 3: the main point was to allow charts that compare opens in M with opens inside the half space

$$\mathbb{H}^m = \{(x_1, \ldots, x_m) \in \mathbb{R}^m : x_m \geq 0\}.$$

Since any open inside $\mathbb{R}^m$ is diffeomorphic to an open inside

$$\text{Int}(\mathbb{H}^m) = \{(x_1, \ldots, x_m) \in \mathbb{R}^m : x_m > 00\},$$

that means that we are allowing more local models. Of course, the new addition is due to the presence of the boundary

$$\partial(\mathbb{H}^m) = \{(x_1, \ldots, x_m) \in \mathbb{R}^m : x_m = 0\}.$$

Definition 5.37. *Given a manifold with boundary, a point $p \in M$ is called a boundary point if there exists a chart $\chi : U \to \Omega \subset \mathbb{H}^m$ around p such that $\chi(p) \in \partial(\mathbb{H}^m)$.*
*The **bounday of** M, denoted ∂M, is the set of boundary points.*

One can check that, in the previous definition, the condition $\chi(p) \in \partial(\mathbb{H}^m)$ does not depend on the choice of χ. Also, it follows that

$$\chi|_{\partial(U)} : \partial(U) \to \partial(\Omega) = \Omega \cap \partial(\mathbb{H}^m)$$

(where the last space is open in $\partial(\mathbb{H}^m) \cong \mathbb{R}^{m-1}$) can be used as a chart for ∂M. These make ∂M into an $(m-1)$-dimensional manifold (fill in the details!).

The definitions we gave in the previous lectures for:

- smooth maps,
- tangent vectors and tangent spaces,
- differential forms and Deham differential,
- orientations

apply word by word to all manifolds with boundary. Care would be required when formulating the regular value theorem or when representing tangent vectors as speeds of curves (e.g., for points $p \in \partial M$, one has to allow also curves $\gamma : [0, \epsilon) \to M$). Note that, at points $p \in \partial M$, the tangent space $T_p\partial M$ (a tangent space of a manifold without boundary!) sits inside T_pM:

$$T_p\partial M \subset T_pM.$$

Even more, among the tangent vectors that are not tangent to the boundary, there is a special class- the ones pointing outwards.

DEFINITION 5.38. *For $p \in \partial M$, a tangent vector $v \in T_pM$ is said to be an outward tangent vector if $v \notin T_p\partial M$ and if there exists a curve $\gamma : [0, \epsilon) \to M$ with $\gamma(0) = p$ and such that $v = \frac{d\gamma}{dt}(0)$.*

In terms of coordinate charts, these are the tangent vectors $v \in T_pM$ with the property that, for any coordinate chart $\chi : U \to \Omega \subset \mathbb{H}^m$ around p, the corresponding vector

$$v^\chi \in \mathbb{R}^m$$

has strictly negative last coordinate (think on a picture!).

The notion of outward pointing vector is important for our discussion because it allows us to obtain an orientation of ∂M out of an orientation of M.

DEFINITION 5.39. *Given an oriented manifold with boundary $(M, \mathcal{O})$, define **the induced orientation** $\partial\mathcal{O}$ on the boundary as follows: for $p \in \partial M$, declare a frame $(v_1, \ldots, v_{m-1})$ of $T_p\partial M$ to be positively oriented if, for a (or, equivalently, for any) outwards pointing tangent vector $v \in T_pM$, $(v, v_1, \ldots, v_{m-1})$ is a positively oriented frame of $(T_pM, \mathcal{O}_p)$.*

With this, whenever we have an oriented manifold M with boundary, its boundary ∂M will be automatically considered to be oriented with the induced orientation. Therefore, the integrations $\int_M$ and $\int_{\partial M}$ are unambiguously defined.

THEOREM 5.40 (Stokes on manifolds with boundary). *For any m-dimensional oriented compact manifold with boundary M one has*

$$\int_M d\eta = \int_{\partial M} \eta|_{\partial M} \quad \textit{for all } \eta \in \Omega^{m-1}(M).$$

PROOF. We will prove the previous formula without assuming that M is compact, but under the condition that η is compactly supported. Covering the support of η by a finite number of coordinate charts $\chi_i : U_i \xrightarrow{\sim} \Omega_i$ where each Ω_i is either $\mathbb{R}^m$ or $\mathbb{H}^m$, and decomposing η as a sum of forms supported inside U_i, we may assume that $M = \mathbb{R}^m$ or $M = \mathbb{H}^m$. The first situation was already treated, so let us assume that $M = \mathbb{H}^m$. As in the case without boundary, we write η in the fom (6.6) and see that the terms corresponding to $i \neq m$ do not contribute to $\int d\eta$, and neither to $eta|_{\mathbb{H}^m}$. Hence we may assume that

$$\eta = f \cdot dx_1 \wedge \ldots \wedge dx_{m-1},$$

so that we can write

$$\int_{\mathbb{H}^m} d\eta = (-1)^{m-1} \int_{\mathbb{H}^m} \frac{\partial f}{\partial x_m}(x_1, \ldots, x_m) dx_1 \ldots dx_m.$$

Writing

$$\int_{\mathbb{H}^m} \frac{\partial f}{\partial x_m}(x_1, \ldots, x_m) dx_1 \ldots dx_m = \int_{\mathbb{R}^{m-1}} \left(\int_0^\infty \frac{\partial f}{\partial x_m}(x_1, \ldots, x_m) dx_m \right) dx_1 \ldots dx_{m-1}$$

we find that

$$\int_{\mathbb{H}^m} d\eta = (-1)^m \int_{\mathbb{R}^{m-1}} f(x_1, \ldots x_{m-1}, 0) dx_1 \ldots dx_{m-1}.$$

Note that, in the last equality, $\int_{R^{m-1}}$ is the standard integration- i.e. it uses the standard orientation on $\mathbb{R}^{m-1}$, for which the canonical basis $\{e_1, \ldots, e_{m-1}\}$ is positively oriented. However, the orientation induced by the one on $\mathbb{H}^m$ is defined by the condition that $\{-e_m, e_1, \ldots, e_{m-1}\}$ is positively oriented; hence the induced orientation differs from the canonical one by a factor of $(-1)^m$; hence

$$\int_{\mathbb{H}^m} d\eta = \int_{\partial\mathbb{H}^m} f(x_1, \ldots x_{m-1}, 0) dx_1 \ldots dx_{m-1} = \int_{\partial\mathbb{H}^m} \eta|_{\partial\mathbb{H}^m}.$$

□

7. De Rham cohomology

We now return to the DeRham differential and its property that $d \circ d = 0$. The principle underlying "Homological algebra" tells us that, whenever we have a sequence of operators d satisfying $d \circ d = 0$, the first thing we do is to look at the resulting cohomology.

Definition 5.41. *A cochain complex of vector spaces is a sequence*

$$\xrightarrow{d} C^0 \xrightarrow{d} C^1 \xrightarrow{d} C^2 \xrightarrow{d} \ldots$$

consisting of vector spaces and linear maps between them satisfying $d \circ d = 0$. We say, in short, that $(C^\bullet, d)$ ***is a complex****.*

Given such a complex, an element $\omega \in \mathcal{C}^k$ is called

- ***closed*** *if $d\omega = 0$; denote by Z^k the set of such.*
- ***exact*** *if $\omega = d\eta$ for some $\eta \in C^{k-1}$; denote by B^k the resulting set.*

While the condition $d \circ d = 0$ means simply that $B^k \subset Z^k$, we define the k-th cohomology space of $(C^\bullet, d)$ as the resulting quotient vector space

$$H^k(C^\bullet, d) := Z^k/B^k.$$

Of course, the central example for us is provided by differential forms and DeRham differential; the resulting complex

$$\Omega^0(M) \xrightarrow{d} \Omega^1(M) \xrightarrow{d} \Omega^2(M) \xrightarrow{d} \ldots$$

(with 0 in negative degrees) is called the **DeRham complex of the manifold** M.

Definition 5.42. *The* ***DeRham cohomology spaces of*** *M, denoted $H^k(M)$, are the cohomology spaces of the DeRham complex $(\Omega^\bullet(M), d)$. For a closed form $\omega \in \Omega^k(M)$ we will denote by*

$$[\omega] \in H^k(M)$$

the induced cohomology class.

When all the cohomology spaces are finite dimensional (e.g. when M is compact-see below) , define ***the Euler characteristic*** *of M as*

$$\chi(M) = \sum_i (-1)^i dim(H^i(M)).$$

EXERCISE 5.20. Show that for any connected compact manifold M, one has $H^k(M) = 0$ if $k > \dim(M)$ and $H^0(M) \cong \mathbb{R}$ if M is connected. What if M is not connected?

EXERCISE 5.21. Show that the wedge product of forms induces a well-defined operation

$$H^k(M) \times H^l(M) \to H^{k+l}(M), \quad ([\omega], [\eta]) \mapsto [\omega] \cdot [\eta] := [\omega \wedge \eta].$$

Deduce that

$$H^\bullet(M) := \bigoplus_k H^k(M)$$

has the structure of a ring (actually an algebra).

Next, the fact that differential forms can be pulled-back allows us to compare the cohomology of different manifolds. Underlying this discussion is the notion of chain map: given two complexes $(\mathcal{C}^\bullet, d_{\mathcal{C}})$ and $(\mathcal{D}^\bullet, d_{\mathcal{D}})$, a **chain map** between them is a sequence of linear maps $f : \mathcal{C}^k \to \mathcal{D}^k$ which commute with the operators d:

$$f \circ d_{\mathcal{C}} = d_{\mathcal{D}} \circ f.$$

In this situation it is an easy (but must do!) exercise to see that f induces maps

$$[f] : H^k(\mathcal{C}, d_{\mathcal{C}}) \to H^k(\mathcal{D}, d_{\mathcal{D}})$$

by sending the class of an element ω into the class of $f(\omega)$.

Again, the basic example is the pull-back map

$$F^* : \Omega^\bullet(N) \to \Omega^\bullet(M)$$

induced by a smooth map $F : M \to N$ between two manifolds. The maps induced in cohomology is denoted by the same letter:

$$F^* : H^\bullet(N) \to H^\bullet(M).$$

EXERCISE 5.22. Chek that $(G \circ F)^* = F^* \circ G^*$ in cohomology; deduce that if $F : M \to N$ is a diffeomorphism, than the map induced by F in cohomology is an isomorphism of vector spaces.

EXERCISE 5.23. With respect to the ring/algebra structure on cohomology from Exercise 5.21, show that the map $F^* : H^\bullet(N) \to H^\bullet(M)$ is a ring/algebra homomorphism.

So far we have discussed the cohomology groups H^k as "functors" that allow us to pass from the world of manifolds (with smooth maps between them) to that of vector spaces (with linear maps between them) (or, with the previous exercises in mind, to the category of algebras with algebra homomorphisms between them). We now look at some of the more specific properties, properties that are extremely useful for explicit computations.

7.1. Poincare lemma. In short, the Poincare lemma says that any closed form on $\mathbb{R}^m$ of strictly positive degree must be exact. In other words, it provides the first complete computation of cohomology:

THEOREM 5.43 (Poincare lemma). *One has*

$$H^k(\mathbb{R}^m) = \begin{cases} 0 & \text{if } k \neq 0 \\ \mathbb{R} & \text{if } k = 0 \end{cases}.$$

Moreover, we will see there is even more to learn from its proof.

PROOF. We proceed by induction on m. For $m = 0$ this is clearly true; we assume it to hold for m and prove it for $m+1$. Write

$$\mathbb{R}^{m+1} = \mathbb{R} \times \mathbb{R}^m$$

with coordinates denoted by $(t, x_1, \ldots, x_m)$. We will treat t as a special coordinate that we try to eliminate: we will show that any closed k-form on $\mathbb{R}^{m+1}$ is an exact one plus one that depends only on $(x_1, \ldots, x_m)$ (so that we can apply the induction hypothesis).

An arbitrary k-form on $\mathbb{R}^{m+1}$ can be decomposed as

$$\omega = \alpha_t + dt \wedge \beta_t \tag{7.1}$$

where

$$\mathbb{R} \ni t \mapsto \alpha_t \in \Omega^k(\mathbb{R}^m), \quad \mathbb{R} \ni t \mapsto \beta_t \in \Omega^{k-1}(\mathbb{R}^m)$$

are smooth families of forms on $\mathbb{R}^m$ and they are interpreted as forms on $\mathbb{R}^m \times \mathbb{R}$ without dt-terms. To such families one can apply the DeRham differential in two ways:

- apply the DeRham differential in $\mathbb{R}^m$ to each α_t. In this case we use the notation $d\alpha_t$- again a smooth family of forms, and again interpreted as a form on $\mathbb{R}^{m+1}$ without dt-component.
- interpret from the beginning α_t as a form on $\mathbb{R}^{m+1}$ and apply the DeRham differential for $\mathbb{R}^{m+1}$.

To distinguish the two, we will denote by d_{tot} the DeRham differential for $\mathbb{R}^{m+1}$. The relationship between the two is simply:

$$d_{\text{tot}}(\alpha_t) = d(\alpha_t) + dt \wedge \dot{\alpha}_t, \tag{7.2}$$

where $\dot{\alpha}_t$ is the derivative with respect to t (of $t \mapsto \alpha_t$). We now start with a $\omega \in \Omega^k(\mathbb{R}^{m+1})$ closed; writing it as (7.1), we find (using also that $dt \wedge dt = 0$)

$$0 = d_{\text{tot}}(\omega) = d(\alpha_t) + dt \wedge (\dot{\alpha}_t - d\beta_t)).$$

So, first of all,

$$d\alpha_t = 0 \quad \text{for all } t.$$

We will be using this only for $t = 0$ (i.e. α_0 will be the closed form on $\mathbb{R}^m$ we are looking for). Secondly, we obtain

$$\dot{\alpha}_t = d\beta_t \tag{7.3}$$

for all t and, therefore, we can write

$$\alpha_t = \alpha_0 + \int_0^t d\beta_t. \tag{7.4}$$

Denoting $\gamma_t = \int_0^t \beta_t$ (again a family of forms on $\mathbb{R}^m$, interpreted as a form on $\mathbb{R}^{m+1}$ with no dt-term), the general formula (7.2) gives us

$$d_{\text{tot}}(\gamma_t) = \int_0^t d\beta_t + dt \wedge \beta_t.$$

Plugging the resulting formula for $dt \wedge \beta_t$, as well as the formula (7.4) for α_t, in (7.1), we find:

$$\omega = \alpha_0 + d_{\text{tot}}(\gamma_t).$$

And this is precisely what we wanted to achieve (just to make sure: α_0 is a closed form on $\mathbb{R}^m$ hence, by induction hypothesis, it is $d\eta$ for some $\eta \in \Omega^k(\mathbb{R}^m)$; of course, since η has no dependence on t, one has $d\eta = d_{\text{tot}}\eta$, hence $\omega = d_{\text{tot}}(\eta + \gamma_t)$ will be exact.) $\square$

Remark 5.44. The previous proof is written in such a way that it is not so important that we work with $\mathbb{R}^m$: instead, we can replace it by an arbitrary manifold M and $\mathbb{R}^{m+1}$ by $M \times \mathbb{R}$; hence the argument above shows that if a manifold M has no cohomology in degree k, then the same is true for $M \times \mathbb{R}$; and with a bit more effort one can prove that, for any M, $M \times \mathbb{R}$ and M have isomorphic cohomologies (see below). However, let us state here what we get for general M without any effort (and which is at the heart of the homotopy invariance discussed below).

Corollary 5.45. *[of the proof] For any manifold M, the maps*

$$i_0^*, i_1^* : H^\bullet(M \times \mathbb{R}) \to H^\bullet(M)$$

induced by the "inclusions" $i_0, i_1 : M \to M \times \mathbb{R}$, $i_\epsilon(p) = (p, \epsilon)$ ($\epsilon \in \{0, 1\}$), coincide.

Proof. Decomposing a k-form on $M \times \mathbb{R}$ as before, i.e. as in (7.1), then $i_\epsilon^*\omega = \alpha_\epsilon$. Hence the statement is that, if ω is closed, then $\alpha_0 = \alpha_1$ is an exact form on M; but that is clear from (7.4) since we get

$$\alpha_1 - \alpha_0 = \int_0^1 d\beta_t = d\left(\int_0^1 \beta_t\right).$$

□

7.2. Homotopy invariance. We now push the previous discussion to a more conceptual level. While the heard work has already been done with the proof of the Poincare lemma, the conceptual gain is substantial: it says that $H^\bullet(M)$ depends only on the "homotopy type" of M (which is much weaker than the diffeomorphism type).

To explain this, we first recall the notion of homotopy and homotopy equivalence. These notions are usually introduced in the general context of topological spaces; of ocurse, here we restrict to the smooth context. Moreover, while for topological spaces (and also intuitively) it is natural to use the interval

$$I = [0, 1]$$

(e.g. to parametrize deformations), in the smooth context that would force us right away to allow for manifolds with boundary. Allowing manifold with boundary is not a real problem, but not absolutely necessary in order to achieve the main result; more precisely, we can just use $\mathbb{R}$ instead of $[0, 1]$.

Therefore, in what follows, one may assume that I is either $[0, 1]$ (if you are ok with manifolds with boundary), or $I = \mathbb{R}$ (if you want to stick to manifolds without boundary but being willing to a accept a less natural framework).

Definition 5.46. *Two smooth maps $F_0, F_1 : M \to N$ between two smooth manifolds M and N are said to be* ***homotopic*** *if there exists a smooth map*

$$\tilde{F} : M \times I \to N$$

such that $F_0(\cdot) = \tilde{F}(\cdot, 0)$, $F_1(\cdot) = \tilde{F}(\cdot, 1)$.

We say that a smooth map $F : M \to N$ is a ***homotopy equivalence*** *if there exists a smooth map $G : N \to M$ with the property that $G \circ F$ is homotopic to Id_M and $F \circ G$ is homotopic to Id_N. When such a map exists, we say that M and N are* ***homotopy equivalent****.*

While Topology studies topological spaces and continuous maps between them (the "category" of topological spaces) and Differential Geometry studies manifolds and smooth maps between them, Homotopy Theory studies topological spaces and homotopy classes of maps between them. "Isomorphisms" are: homeomorphisms in Topology, diffeomorphisms in Differential Geometry and ... homotopy equivalences in Homotopy Theory. Therefore, the next result is called "homotopy invariance":

Theorem 5.47 (homotopy invariance). *For any two manifolds M and N:*

(1) *If $F_0, F_1 : M \to N$ are homotopic, then they induce the same maps in cohomology.*
(2) *If $F : M \to N$ is a homotopy equivalence, then the maps induced in cohomology, $F^* : H^\bullet(N) \to H^\bullet(M)$, are isomorphisms.*

Proof. The first part follows right away from Corollary 5.45: if $\tilde{F} : M \times I \to N$ relates F_0 and F_1 as in the definition, we can write $F_0 = \tilde{F} \circ i_0$ and $F_1 = \tilde{F} \circ i_1$, where i_ϵ are the inclusions from the corollary. Therefore, in cohomology,

$$F_0^* = i_0^* \circ \tilde{F}^* = i_1^* \circ \tilde{F}^* = F_1^*.$$

And the second part follows easily from the first part: if $G : N \to M$ is as in the definition of homotopy equivalence we get

$$F^* \circ G^* = (G \circ F)^* = \mathrm{Id}$$

and similarly for $G^* \circ F^*$, hence F^* is an isomorohism. □

Example 5.48. $M = \mathbb{R}^m$ and $N = \{0\}$ are homotopic equivalent. In one direction, there is only one map $F : \mathbb{R}^m \to \{0\}$; in the other direction, we use the map G that maps N to the origin. The only thing we have to show is that

$$f : \mathbb{R}^m \to \mathbb{R}^m, \quad f(x) = 0,$$

is homotopic to the identity. For that take $\tilde{f}(x, t) = t \cdot x$.

Therefore, the previous theorem will imply the Poincare lemma.

Exercise 5.24. By a similar argument show that, for any manifold M, the projection pr : $M \times I \to M$ induces isomorphisms in cohomology

$$\mathrm{pr}^* : H^\bullet(M) \tilde{\to} H^\bullet(M \times I).$$

Example 5.49. Similarly, one gets

$$H^\bullet(\mathbb{R}^m \setminus \{0\}) \cong H^\bullet(S^{m-1})$$

(and the last group will be computed in the next subsection). In other words, we claim that the two spaces involved are homotopy equivalent. To see this, we consider the inclusion

$$i : S^{m-1} \hookrightarrow \mathbb{R}^m \setminus \{0\}$$

and, in the other direction, the rescaling

$$r : \mathbb{R}^m \setminus \{0\} \to S^{m-1}, \quad r(x) = \frac{1}{||x||} \cdot x.$$

While $r \circ i$ is the identity map of S^{m-1}, we claim that $i \circ r$ is homotopic to the identity of $\mathbb{R}^m \setminus \{0\}$. Indeed,

$$F : \mathbb{R}^m \setminus \{0\} \times I \to \mathbb{R}^m \setminus \{0\}, \quad F(x, t) = t \cdot x + (1 - t) \cdot i(r(x))$$

defined the desired homotopy.

7.3. Stokes theorem revisited. Stokes' theorem tells us that integration of top-forms interacts nicely with cohomology.

Corollary 5.50 (Stokes' theorem reformulated). *For a compact oriented manifold M, the integration $\int_M : \Omega^m(M) \to \mathbb{R}$ descends to a linear map*

$$\int_M : H^m(M) \to \mathbb{R}, \quad [\omega] \mapsto \int_M \omega.$$

Using orientations induced by a volume form μ (Lemma 5.36), the resulting integration gives $\int_M \mu > 0$ (why?), therefore:

COROLLARY 5.51. *On a compact manifold, any volume form $\mu \in \Omega^m(M)$ defines a non-trivial cohomology class*

$$[\mu] \neq 0 \in H^m(M).$$

EXAMPLE 5.52. In particular, we see that the m-th cohomology group of the sphere S^m is non-zero, with with the volume form μ_{S^m} from Example 5.35 inducing a non-zero cohomology class.

We no show that this fact alone has some very interesting applications.

COROLLARY 5.53 (Brower's fixed point theorem). *Any smooth map from the closed unit disk to itself, $f : D^m \to D^m$, has at least one fixed point.*

PROOF. Assume there is a map f with no fixed point. Then for each $x \in D^m$, the line through x and $f(x)$ intersects the boundary sphere S^{m-1} in two points. Let $r(x)$ the point with the property that x belongs to the segment $[f(x), r(x)]$. Then we obtain a smooth map

$$r : D^m \to S^{m-1} \quad \text{with} r|_{S^{m-1}} = \text{Id}$$

(why is the map r smooth)? Denoting by $i : S^{m-1} \hookrightarrow D^m$ the inclusion, we have $r \circ i = \text{Id}$ hence, if the degree $m-1$ cohomology, the composition

$$H^{m-1}(S^{m-1}) \xrightarrow{r^*} H^m(D^m) \xrightarrow{i^*} H^{m-1}(S^{m-1})$$

is the identity map. But, since $H^{m-1}(D^m) = 0$ (why?), that would imply that $H^{m-1}(S^{m-1}) = 0$. So no such f exists. □

COROLLARY 5.54 (Nowhere vanishing vector fields on spheres). *The sphere S^m admits a non-where vanishing vector field if and only if m is odd.*

PROOF. Assume that such a vector field exists V exists. Hence

$$V : S^m \to \mathbb{R}^m, \quad V(p) \perp p \text{ for all } p \in S^m.$$

By rescaling, we may assume $||V(p)|| = 1$ at all p. But then

$$H(p,t) = \cos(\pi\, t)\, p + \sin(\pi\, t)\, V(p) \in S^m \quad \text{for all } p \in S^m, t \in I$$

and we obtain a homotopy

$$H : S^m \times I \to S^m$$

between Id and $\tau := -\text{Id}$. Hence, by homotopy invariance, $-\tau$ will induce the identity map in the m-cohomology group of S^m; in particular,

$$[\tau^*(\mu_{S^m})] = [\mu_{S^m}].$$

But, given the explicit formula of μ_{S^m}, it is clear that

$$\tau^*(\mu_{S^m}) = (-1)^{m+1}\mu_{S^m}$$

already as forms (hence also at the level of cohomology classes). Since the cohomology class of μ_{S^m} is non-zero, we deduce that $1 = (-1)^{m+1}$, hence m must be odd. Of course, when m is odd, one can just build a simple nowhere vaninshing vector fields on S^m- e.g.

$$V := \left(x_0 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_0}\right) + \left(x_2 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_2}\right) + \dots$$

□

Remark 5.55 (for the interested students). For any oriented compact manifold M, using the product operation on cohomology induced by the wedge-product of forms (cf. e.g. Exercise 5.21) and the integration, we obtain a map

$$\beta : H^k(M) \times H^{m-k}(M) \to \mathbb{R}, \quad \beta([\omega], [\eta]) = \int_M \omega \wedge \beta$$

(for each k, where $m = \dim(M)$). This is know under the name of **the Poincare pairing**. The Poincare duality theorem says that this map is non-degenerate or, equivalently, that the induced map

$$H^k(M) \to H^{m-k}(M)^*, \quad u \mapsto \beta(u, \cdot)$$

is an isomorphism. Using the fact that the cohomology spaces of compact manifolds are finite dimensional (see below), this will imply that the groups $H^k(M)$ and $H^{m-k}(M)$ have the same dimension, for all k. In particular, we deduce that $\chi(M) = 0$ for all odd dimensional compact oriented manifolds M.

While the non-vanishing of $\int_M$ was enough to get such results, let us now describe the full information that it provides on the top cohomology spaces.

Theorem 5.56. *For any connected compact oriented manifold M,*

$$\int_M : H^m(M) \to \mathbb{R}, \quad [\omega] \mapsto \int_M \omega$$

is an isomorphism.

Proof. We have to show that a top-form $\omega \in \Omega^m(M)$ is zero in cohomology (exact) if and only if $\int_M \omega = 0$. Actually a version of this is true on any oriented M (compact or not), but for compactly supported form $\omega \in \Omega^m_c(M)$: $\omega = d\eta$ for some compactly supported η if and only if $\int_M \omega = 0$. We first prove that this is true for Euclidean spaces. On $\mathbb{R}$ the statement is immediate. In general, we proceed by induction on m; we assume it to hold for m and we prove it for $m+1$. Of course, this is similar to the proof of the Poincare lemma (Theorem 5.43), and we use the same notations as there. Hence let $\omega \in \Omega^{m+1}_c(\mathbb{R}^{m+1})$ whose integral is zero and write it in the form (7.1). Note that, from (7.3) and the fact that ω has compact support (hence $\alpha_t = 0$ for t very large or very small), it follows that

$$\int_{\mathbb{R}} d\beta_s = 0.$$

Hence the form $\int_{\mathbb{R}} \beta_s \in \Omega^m(\mathbb{R}^m)$ is closed, and clearly with compact supports. Moreover, its integral over $\mathbb{R}^m$ equals $\int_{\mathbb{R}^{m+1}} \omega = 0$. By the induction hypothesis, we can write it as

$$\int_{\mathbb{R}} \beta_s \, ds = d\gamma \quad \text{for some } \gamma \in \Omega^{m-1}_c(\mathbb{R}^{m-1}).$$

Consider now a function $\phi \in \mathcal{C}^\infty_c(\mathbb{R})$ with $\int_{\mathbb{R}} \phi(s)\, ds = 1$ and consider $\psi(t) = \int_{-\infty}^t \phi(s)\, ds$. And define the following m-form on $\mathbb{R}^{m+1}$:

$$\eta := \int_{-\infty}^t \beta_s \, ds - \psi(t) \cdot \int_{-\infty}^\infty \beta_s \, ds - \phi \cdot dt \wedge \gamma.$$

By the choice of ϕ and the fact that γ has compact support in $\mathbb{R}^m$, η is compactly supported. We now compute its differential; using $\psi'(t) = \phi(t)$ and $\int_{\mathbb{R}} d\beta_s = 0$ (see also 7.2), we find

$$d_{\text{tot}}\eta = \int_{-\infty}^t d\beta_s \, ds + dt \wedge \left(\beta_t - \phi(t) \int_{-\infty}^\infty \beta_s ds + \phi(t) d\gamma \right).$$

From the choice of γ, the last two terms that are wedged with dt cancel. And, using (7.3) again, $\int_{-\infty}^{t} d\beta_s\, ds = \alpha_t$. Hence we see that $d_{\text{tot}}\eta = \alpha_t + dt \wedge \beta_t = \omega$.

We now return to the original proof. Hence M is compact and oriented, $\int_M \omega = 0$ and we want to prove that ω is exact. The strategy is to write ω as a sum of forms, each supported in a coordinate chart $U_i \cong \mathbb{R}^m$, and each with zero integral; by the previous result for $\mathbb{R}^m$, we will be done. To implement this idea, we first cover M with a finite number of domains of coordinate charts $U_i \cong \mathbb{R}^m$, $i \in \{0, 1, \ldots, k\}$. After eventually adding some more opens, we may assume that

$$U_0 \cap U_1 \neq \emptyset, \quad U_1 \cap U_2 \neq \emptyset, \quad \ldots U_{k-1} \cap U_k \neq \emptyset.$$

(this is an exercise; hint: for U and V two consecutive domains, if they do not intersect, choose a path that starts in U and ends in V, then cover the path carefully so that each two consecutive charts intersect).

Choose now a partition of unity $\{\phi_0, \phi_1, \ldots, \phi_k\}$ subordinated to this cover and let $c_i = \int_M \phi_i \cdot \omega$. Choose also forms $\alpha^i \in \Omega^m(M)$, supported inside $U_i \cap U_{i+1}$, with $\int_M \alpha_i = 1$ and consider the following forms

$$\begin{aligned} \omega_0^{'} &= \phi_0 \cdot \omega - c_0\, \alpha_0 \\ \omega_1^{'} &= \phi_1 \cdot \omega - (c_0 + c_1)\, \alpha_1 + c_0\, \alpha_0 \\ \omega_2^{'} &= \phi_2 \cdot \omega - (c_0 + c_1 + c_2)\, \alpha_2 + (c_0 + c_1)\, \alpha_1, \quad \text{etc} \end{aligned}$$

Note that, since $\sum_i c_i = \int_M \omega = 0$, and $\sum \phi_i = 1$, the sum of these forms is precisely ω. Moreover, by construction, each $\omega_i^{'}$ is supported inside U_i. Hence this provides the desired decomposition of ω. □

7.4. Mayer-Vietoris. We now discuss another useful tool for computing cohomology- the Mayer-Vietoris sequence. Roughly speaking, it provides a way to compute that cohomology of a manifold M by breaking M into pieces. It is formulated in terms of long exact sequences in the following sense.

Definition 5.57. *A sequence of vector spaces with linear maps between them,*

$$\ldots \longrightarrow A \stackrel{f}{\longrightarrow} B \stackrel{g}{\longrightarrow} C \longrightarrow \ldots$$

is said to be ***exact at*** B *if* $Ker(g) = Im(g)$. *We say that it is an* ***exact sequence*** *is it is exact at all entries.*

Example 5.58. A sequence of vector spaces of type

$$A \stackrel{f}{\longrightarrow} B \longrightarrow 0$$

is exact if and only if f is surjective, while a sequence of type

$$0 \longrightarrow A \stackrel{f}{\longrightarrow} B$$

is exact if and only of f is injective.

Exercise 5.25. Exact sequences of type

$$0 \longrightarrow A \stackrel{f}{\longrightarrow} B \stackrel{g}{\longrightarrow} C \longrightarrow 0$$

are called **short exact sequences**. Show that, for such a short exact sequence, B is finite dimensional if and only if A and C are. Moreover, in this case prove that

$$\dim(B) = \dim(A) + \dim(C).$$

EXERCISE 5.26. Show that a sequence of vector spaces with linear maps between them,

$$\xrightarrow{d} C^{k-1} \xrightarrow{d} C^{k} \xrightarrow{d} C^{k+1} \xrightarrow{d} \ldots,$$

is exact if and only if it is a cochain complex with vanishing cohomology.

Moreover, when this happens, if the vector spaces C^i are finite dimensional and only a finite number of them are non-zero, show that

$$\sum_i (-1)^i \dim(C^i) = 0.$$

The Mayer-Vietoris theorem (and its application to the computation of cohomology of spheres) is probably the best illustration of the usefulness of long exact sequences.

To state the main result, assume we have a manifold M written as

$$M = U \cup V, \quad \text{with } U, V \subset M \text{ opens.}$$

There are some natural maps relating the cohomologies of M, U, V and $U \cap V$; they are built from the inclusions

$$i_U : U \to M, \quad i_V : U \to M,$$
$$j_U : U \cap V \to U, \quad j_U : U \cap V \to V.$$

Using the maps induced in cohomology, we consider

$$H^{\bullet}(M) \xrightarrow{(i_U^*, i_V^*)} H^{\bullet}(U) \oplus H^{\bullet}(V) \xrightarrow{j_U^* - j_V^*} H^{\bullet}(U \cap V).$$

Note that the composition of these two maps is zero.

THEOREM 5.59. *There exist linear maps*

$$\delta : H^{\bullet}(U \cap V) \to H^{\bullet+1}(M)$$

such that the following sequence is exact

$$\ldots \xrightarrow{\delta} H^k(M) \xrightarrow{(i_U^*, i_V^*)} H^k(U) \oplus H^{\bullet}(V) \xrightarrow{j_U^* - j_V^*} H^k(U \cap V) \xrightarrow{\delta} H^{k+1}(M) \longrightarrow \ldots.$$

PROOF. We first claim that, for each k, the sequence

$$0 \to \Omega^k(M) \xrightarrow{(i_U^*, i_V^*)} \Omega^k(U) \oplus \Omega^k(V) \xrightarrow{j_U^* - j_V^*} \Omega^k(U \cap V) \to 0 \tag{7.5}$$

is exact. Exactness at $\Omega^k(M)$ means the injectivity of the map (i_U^*, i_V^*), i.e. that, for $\omega \in \Omega^k(M)$, if $\omega|_U = 0$ and $\omega|_V = 0$ then $\omega = 0$. And this is clear.

Next, the exactness in the middle. Since the composition of the two maps is clearly zero, the image of the first is inside the kernel of the second. To prove the reverse inclusion, let

$$(\omega_U, \omega_V) \in \mathrm{Ker}(j_U^* - j_V^*).$$

That means that

$$\omega_U|_{U \cap V} = \omega_V|_{U \cap V}.$$

We have to show that there exists $\omega \in \Omega^k(M)$ whose restrictions to U and V is ω_U and ω_V, respectively. But that should be clear (the desired condition tells us how to define ω).

We are left with showing the exactness at $\Omega^k(U \cap V)$, i.e. the surjectivity of the map $j_U^* - j_V^*$. So, let $\omega \in \Omega^k(U \cap V)$ and we are looking for $\omega_U \in \Omega^k(U)$ and $\omega_V \in \Omega^k(V)$ so that

$$\omega = \omega_U|_{U \cap V} - \omega_V|_{U \cap V}.$$

For that choose partition of unity $\{\eta_U, \eta_V\}$ subordinated to $\{U, V\}$ and set $\omega_U = \eta_U \cdot \omega$, $\omega_V = -\eta_V \cdot \omega$.

Therefore (7.5) is exact. We now claim that the rest is a consequence of a general property of short exact sequence of complexes:

Lemma 5.60. *Consider a short exact sequence of chain complexes*

$$0 \to A^\bullet \xrightarrow{f} B^\bullet \xrightarrow{g} C^\bullet \to 0, \tag{7.6}$$

i.e. $(A^\bullet, d_A)$, $(B^\bullet, d_B)$ *and* $(C^\bullet, d_C)$ *are chain complexes,* f *and* g *are chain maps, and the previous sequence is exact in each degree. Then there exist linear maps*

$$\delta : H^\bullet(C, d_C) \to H^{\bullet+1}(A, d_A)$$

such that the following sequence is exact

$$\ldots \xrightarrow{\delta} H^k(A, d_A) \xrightarrow{[f]} H^k(B, d_B) \xrightarrow{[g]} H^k(C, d_C) \stackrel{\delta}{\to} H^{k+1}(A, d_A) \longrightarrow \ldots.$$

Proof. Let us describe the map δ on an arbitrary element $[c] \in H^k(C, d_C)$. Hence $c \in C^k$,

$$d_C(c) = 0.$$

The exactness of (7.6) at C ensures that g is surjective, hence we can write

$$c = g(b), \quad \text{with } b \in B^k.$$

Now, since $d_C(c) = 0$ and g is a chain map, we obtain $g(d_B(b)) = 0$, i.e. $d_B(b) \in \mathrm{Ker}(g)$. Using again the exactness of (7.6), this time at B (and in degree $k+1$) we can write

$$d_B(b) = f(a), \quad \text{with } a \in A^{k+1}.$$

Applying d_B we also find $f(d_A(a)) = 0$ and, since f is injective (the exactness of (7.6) at A), one has $d_A(a) = 0$. Hence a defines an element in the cohomology of A and we set:

$$\delta([c]) := [a].$$

We have to check that $[a]$ not depend on the choices involved: if $[c] = [c']$ and we also write

$$c' = g(b'), \quad d_B(b') = f(a'),$$

then a and a' represent the same cohomology class. To see this, since $[c] = [c']$, we can write $c' = c + d_C(z)$ with $z \in C^{k-1}$. Using again the surjectivity of g, write $z = g(y)$ for some $y \in B^{k-1}$. Then $c' = c + g(d_B(y))$ and, using the choice of b and b' we obtain

$$b' - b = d_B(y) \in \mathrm{Ker}(g)$$

hence we can write

$$b' = b + d_B(y) + f(x), \quad \text{with } x \in A^k.$$

Then $d_B(b') = d_B(b) + f(d_A(x))$ hence, using the way we chose a and a', $f(a') = f(a) + f(d_A(x))$. Hence, by the injectivity of f again, $a' = a + d_A(x)$, therefore $[a'] = [a]$ in cohomology.

With all the maps defined now, one still has to prove the exactness of the sequence. This is not so difficult and it is an extremely useful exercise (a really "must do" one!). Therefore, we check here the exactness at one spot only. Say at $H^k(C, d_C)$:

$$\mathrm{Im}[g] = \mathrm{Ker}(\delta).$$

(but we insist: it may be more instructive, and even easier, to do it yourself than reading the argument below!).

For the direct inclusion, we start with $[c] \in H^k(C, d_C)$ with the property that $c = g(b_0)$ for some $b_0 \in B^k$ closed. Recall that $\delta([c]) = [a]$ where a is the result of the choices of b with $c = g(b)$ and of a with $d_B(b) = f(a)$. But we can just choose $b = b_0$ and then, since $d_B(b_0) = 0$, one can choose $a = 0$. Hence $\delta([c]) = 0$. For the reverse inclusion, let $[c] \in C^k(C, d_C)$ such that $\delta([c]) = 0$. That means that we can write

$$c = g(b), \quad d_B(b) = f(a), \quad a = d_A(a_0)$$

for some $b \in B^k$, $a \in A^{k+1}$, $a_0 \in A^k$. The last two equations give $d_B(b) = fd_A(a_0) = d_Bf(a_0)$ hence, denoting $b' = b - f(a_0)$, we have

$$b = b' + f(a_0) \quad \mathrm{withd_B(b') = 0/}$$

Applying g and using that $g \circ f = 0$ and $c = g(b)$, we find

$$c = g(b') \quad \text{with } d_B(b') = 0.$$

Hence $[b'] \in H^k(B, d_B)$ is defined and is sent by $[g]$ to $[c]$. □

□

EXAMPLE 5.61. Let us compute the cohomology of the spheres S^m. We claim that

$$H^k(S^m) = \begin{cases} \mathbb{R} & \text{if } k = 0 \text{ or } k = m \\ 0 & \text{otherwise} \end{cases}.$$

The strategy is to decompose

$$S^m = U \cup V$$

where U and V are basically the half hemispheres, extended so that they become open. E.g. they could be taken the entire domains of the stereographic projections with respect to the north and south pole, or just

$$U = \{(x_0, \ldots, x_m) \in S^m : x_m > -\epsilon\}, \quad V = \{(x_0, \ldots, x_m) \in S^m : x_m < -\epsilon\},$$

with $\epsilon \in (0, 1)$ any number. Note that both U and V are diffeomorphic to a disk (hence also to a Euclidean space), hence

$$H^k(U) \cong H^k(V) \cong \begin{cases} \mathbb{R} & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases}.$$

On the other hand, $U \cap V$ is homotopy equivalent to S^{m-1}- more precisely, the inclusion

$$i : S^{m-1} \hookrightarrow U \cap V$$

is a homotopy equivalence (prove this, if necessary with some inspiration from Example 5.49). Hence

$$H^k(U \cap V) \cong H^k(S^{m-1}).$$

Let us look at the case $m = 1$ first. This is slightly different than the rest because, in this case, $S^{m-1} = S^0 = \{-1, 1\}$ has two connected components, hence $H^0(U \cap V) = \mathbb{R} \oplus \mathbb{R}$ (and zero in other degrees), hence the first part of the Mayer-Vietoris sequence,

$$0 \to H^0(S^1) \to H^0(U) \oplus H^0(V) \to H^0(U \cap V) \to H^1(S^1) \to 0$$

will become (isomorphic to)

$$0 \to \mathbb{R} \xrightarrow{i} \mathbb{R} \oplus \mathbb{R} \xrightarrow{j} \mathbb{R}^2 \to H^1(S^1) \to 0$$

Independent of what the maps i and j are, this can happen if and only of $H^1(S^1) \cong \mathbb{R}$ (exercise!). While $H^k(S^1) = 0$ for $k > 1$ (why?), we are done with the case $m = 1$. For $m \geq 2$, the Mayer-Vietoris will become

$$\xrightarrow{\delta} H^1(S^m) \longrightarrow 0 \longrightarrow H^1(S^{m-1}) \xrightarrow{\delta} H^2(S^m) \longrightarrow 0 \longrightarrow \ldots$$

$$\ldots$$

$$\ldots \xrightarrow{\delta} H^{k-1}(S^m) \to 0 \longrightarrow H^{k-1}(S^{m-1}) \xrightarrow{\delta} H^k(S^m) \longrightarrow 0 \longrightarrow \ldots$$

from which we deduce that

$$H^k(S^m) \cong H^{k-1}(S^{m-1}) \quad \forall k \geq 2, m \geq 1.$$

Hence

$$H^m(S^m) \cong H^{m-1}(S^{m-1}) \cong \ldots \cong H^1(S^1) \cong \mathbb{R}$$

and, similarly, $H^k(S^m) = 0$ for $k < m$, while for $k > 0$ we already know that the cohomology vanishes.

And here is another application of the Mayer-Vietoris sequence: the finite dimensionality of the cohomology groups of any manifold. However, for the proof we will need one extra-ingredient that will not be proven here:

PROPOSITION 5.62. *Any compact manifold admits a* ***finite good cover****, i.e. a finite open cover $\mathcal{U}$ with the property that, for each opens $U_1, \ldots, U_k \in \mathcal{U}$, the intersection $U_1 \cap \ldots \cap U_k$ is either empty or diffeomorphic to $\mathbb{R}^m$.*

Appealing to your intuition for imagining what a triangulation of a manifold is, a triangulation (which always exists) gives rise to such a finite good cover (one just has to "fatten" each simplex a bit). The standard way to prove the proposition above is by using a "Riemannian metric" on the manifold and taking geodesically convex neighborhoods of points. What we are able to prove here is:

COROLLARY 5.63. *Any manifold M which admits a finite good cover (in particular, any compact manifold) has finite dimensional cohomology group).*

This corollary allows us to define **the Euler characteristic** of any compact manifold M:

$$\chi(M) := \sum_i (-1)^i \dim\left(H^i(M)\right).$$

PROOF. We prove by induction on k the following statement: if M has a good cover consisting of k opens, then its cohomology is finite dimensional. The situation is clear when $k = 1$, since then $M \cong \mathbb{R}^m$. Assume that the statement is true for k and we prove it for $k+1$. Hence assume that M admits a good cover $\{U_1, \ldots, U_{k+1})$. Consider

$$U = U_1 \cup \ldots \cup U_k, \quad V = U_{k+1}.$$

Note that the induction hypothesis can be applied to U, V and $U \cap V$ (for the last one note that $\{U_1 \cap V, \ldots, U_k \cap V\}$ is a good cover); hence their cohomologies are finite dimensional. Given the Mayer-Vietoris, it suffices to show that, if one has an exact sequence

$$A \stackrel{f}{\to} B \stackrel{\delta}{\to} C \stackrel{g}{\to} D \stackrel{h}{\to} E$$

where A, B, D and E are finite dimensional, then so is C. To see this, note that g takes values in

$$D' = \mathrm{Ker}(h)$$

δ descends to the quotient

$$B' := B/\mathrm{Im}(f),$$

and the exactness of the sequence gives a short exact sequence

$$0 \to B' \stackrel{\delta}{\to} C \stackrel{g}{\to} D' \to 0,$$

where B' and D' are finite dimensional (as a subspaces of the finite dimensional B, and a quotient of the finite dimensional D). Form this is follows easily that C is finite dimensional, of dimension equal to the sums of the dimensions of B' and D'. ☐

Exercise 5.27. Using the last part of Exercise 5.26, shows that the Euler characteristic $\chi(M)$ of a compact manifold M can be computed with the help of a finite good cover $\mathcal{U} = \{U_i : i \in I\}$ as follows:

$$\chi(M) = \sum_k (-1)^k n_k(\mathcal{U})$$

where, for each k, $n_k(\mathcal{U})$ is the cardinality of the set

$$\{J \subset I : |J| = k, \bigcap_{j \in J} U_j \neq \emptyset\}.$$

Then use this to compute the Euler characteristic of S^2.