Marius Crainic

Manifolds 2022-2023

http ://www.staff.science.uu.nl/∼crain101/manifolds-2022/

September, 2022

Springer

Contents

Chapter 1
Reminders on Topology and Analysis

1.1 Reminder 1: 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.

1.1.1 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 $\mathscr{T}$ of subsets of X (called **the opens of the topological space**, or simply **opens 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 $\mathscr{T}$ 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". First of all 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 $U \cap V = \emptyset$. Hence, intuitively, this means that "if two are distinct, then they cannot be too close to each other" (yes, not having this sounds pathological but, since this condition is not automatic, it is often imposed precisely to avoid "strange/pathological spaces).

1.1.2 The morphisms/isomomorphisms of Topology

The relevant maps (the only ones that really matter) in Topology 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.

In the language of "Category Theory", Topology is the category whose objects are topological spaces, and whose morphisms (between objects) are the continuous maps.

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 follow 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.

1.1.3 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 $\mathscr{T}_d$ 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:

$$B_d(x,r) := \{y \in X : d(x,y) < r\}. \tag{1.1.1}$$

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 $\mathscr{T}_d$ (note also that, if such a d exists, in general it is far from being unique; e.g. already $2d$, $3d$, $\frac{d}{d+1}$ would do the same job). 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 (1.1.1)), giving rise to the collection of open balls induced by d:

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

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

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

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

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

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

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

(in particular, each of the conditions (1) and (2) imply that $\mathscr{B}$ 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 $\mathscr{B}_d$ as basis. Another possible basis for the space X (endowed with the topology $\mathscr{T}_d$), slightly smaller, is

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

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

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

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 $\mathscr{B}$ which is countable.

1.1.4 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 **topological *m*-dimensional manifold** is a topological space X satisfying the following:

(TM0): any point $x \in X$ admits a neighborhood X which is homeomorphic to an open subset of $\mathbb{R}^m$.
(TM1): it is Hausdorff.
(TM2): it is second countable.

A homeomorphism

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

from an open subset U of X to an open subset Ω in $\mathbb{R}^m$ is called a ***m*-dimensional topological chart for** X, and U is called the domain of the chart- so that axiom (TM0) can also be read as:

(TM0): X can be covered by (domains) of m-dimensional topological charts.

You should convince yourself (or remember) why some of the usual 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) (... as the labelling indicates). 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 built 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 is 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 $f(x)$.

1.1.5 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 $\mathrm{Int}(A)$.
- the smallest closed containing A- and this is called **the closure of** A (in the space X), and denoted $\mathrm{Cl}(A)$,

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

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

1.1.6 Construction of topological spaces

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

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

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 $\mathscr{B}_{X\times Y}$. Equivalently, and more conceptually, it is the smallest topology on $X \times Y$ with the property that the projections

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

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 would looking for the largest topology with this property 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 a 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!

1.1.7 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

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

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:
$$X - \text{compact} \Longrightarrow X - \text{normal}.$$
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:

$$X, Y - \text{compact} \Longrightarrow X \times Y - \text{compact}.$$

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 $f(A) \subset Y$:

$$f : X \to Y \text{ continuous}, A - \text{compact inside } X \Longrightarrow f(A) - \text{compact}.$$

6. In particular: quotients of compacts are compacts.
7. A continuous bijection from a compact space to a Hausdorff one is automatically a homeomorphism:

$$(\text{continuous } f) : (\text{compact space } X) \to \text{ (Hausdorff space } Y) \Longrightarrow f \text{ is a homeomorphism}.$$

More generally: a continuous injection from compact to Hausdorff is automatically an embedding (see above).

Exercise 1.6. Assuming that you already know that the unit interval $[0,1]$ (endowed with the Euclidean topology) is compact, use the properties listed above to deduce that: for subsets of $\mathbb{R}^m$ endowed with the Euclidean topology:

$$A \subset \mathbb{R}^m \text{ is compact} \Longleftrightarrow A \text{ is closed and bounded in } \mathbb{R}^m.$$

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": 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.

For topological manifolds, axiom (MT0) ensures that they are automatically locally compact. But also axioms (MT1), (MT2) interact nicely with local compactness: they ensure the existence of "exhaustions". This is Theorem 4.37 in the notes on Topology:

Theorem 1.7. *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 $\mathscr{B}$ be a countable basis and consider $\mathscr{V} = \{B \in \mathscr{B} : \overline{B} - \text{compact}\}$. Then $\mathscr{V}$ is a basis: for any open U and $x \in X$ we choose a compact neighborhood N inside U; since $\mathscr{B}$ is a basis, we find $B \in \mathscr{B}$ s.t. $x \in B \subset N$; this implies $\overline{B} \subset N$ and then $\overline{B}$ must be compact; hence we found $B \in \mathscr{V}$ s.t. $x \in B \subset U$. In conclusion, we may assume that we have a basis $\mathscr{V} = \{V_n : n \in \mathbb{Z}_+\}$ 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 $\mathscr{V}$ 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. 😀

1.1.8 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

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

One of the simplest but most fundamental ideas in various parts of Geometry is that of understanding a space X via the associated "object" $\mathscr{C}(X)$. 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 $\mathscr{C}(X)$ does).

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

Definition 1.8. A (real) **algebra** is a vector space A over $\mathbb{R}$ together with an operation

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

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 $a, b \in A$.

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 $\mathscr{C}(X)$ is defined simply by pointwise addition and multiplication: for $f, g \in \mathscr{C}(X)$ and $\lambda \in \mathbb{R}$, $f+g, f \cdot g, \lambda \cdot f \in \mathscr{C}(X)$ 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 $\mathscr{C}(X, \mathbb{C})$ 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 $\mathscr{C}(X)$, 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 1.9 (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 $A = \mathscr{C}(X)$- the algebra of continuous functions on a compact Hausdorff space X, one recovers X (i.e. $X(\mathscr{C}(X))$ is homeomorphic to X).*

Remark 1.10 (Some details). The spectrum $X(A)$ of an algebra A is defined as the set of characters on A, i.e. maps

$$\chi : A \to \mathbb{R}$$

which preserve the algebra structure, i.e. which are linear, multiplicative ($\chi(ab) = \chi(a)\chi(b)$ for all $a, b \in A$) and send the unit of A to $1 \in \mathbb{R}$. The topology on $X(A)$ is "the best" for which all the evaluation maps

$$\mathrm{ev}_a : X(A) \to \mathbb{R}, \quad \chi \mapsto \chi(a) \quad \text{(one for each } a \in A\text{)}$$

are continuous.

For instance, when $A = \mathscr{C}(X)$ for a compact Hausdorff space X, then any point $x \in M$ gives rise to a character χ_x on $\mathscr{C}(X)$, namely the evaluation at x, and the resulting map

$$X \to X(A), \quad x \mapsto \chi_x$$

is the one the realises the desired homeomorphism (of course, there are things to prove along the way). Let us give here a direct argument showing that, if X is a compact space, then any character on $\mathscr{C}(X)$,

$$\chi : \mathscr{C}(X) \to \mathbb{R},$$

is necessarily of type χ_x for some $x \in M$ (proving that the previous map is surjective). □

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 1.11. 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 subalgebras:

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

Finally, when looking at a subset

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

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

(1) $\mathscr{A}$ is **point separating** if for any $x, y \in X$ distinct there exists $f \in \mathscr{A}$ such that $f(x) \neq f(y)$ or, equivalently, if there exists $f \in \mathscr{A}$ such that $f(x) = 0$ and $f(y) = 1$.
(2) $\mathscr{A}$ is **normal** if for any two disjoint closed subset $A, B \subset X$, there exists $f \in \mathscr{A}$ such that $f|_A = 0$, $f|_B = 1$.
(3) $\mathscr{A}$ is **closed under sums** if $f + g \in \mathscr{A}$ whenever $f, g \in \mathscr{A}$.
(3) $\mathscr{A}$ is **closed under quotients** if $f/g \in \mathscr{A}$ whenever $f, g \in \mathscr{A}$ 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 $\mathscr{A} \subset \mathscr{C}(X)$ is dense in $\mathscr{C}(X)$; with particular cases of the type: real valued continuous functions on $[0,1]$ (or other similar spaces) can be approximated by polynomial functions.

Note that, for a general topological space X, even the entire $\mathscr{A} = \mathscr{C}(X)$ need not be point separating or normal. Actually, it is a rather simple exercise to check that the point separation of $\mathscr{C}(X)$ implies that X must be Hausdorff, while the normality of $\mathscr{C}(X)$ 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 $\mathscr{C}(X)$ 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.

1.1.9 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

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

is a given vector subspace; we will be looking at partitions of unity that belong to $\mathscr{A}$. 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:

$$\mathrm{supp}_X(\eta) := \overline{\{\eta \neq 0\}} = \overline{\{x \in X : \eta(x) \neq 0\}}^X.$$

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 $\mathrm{supp}(\eta) \subset U$ is replaced by the weaker one that $\{\eta \neq 0\} \subset U$).

We now move to partitions of unity; we start with the finite ones.

Definition 1.12. Let X be a topological space, $\mathscr{U} = \{U_1, \ldots, U_n\}$ a finite open cover of X. A continuous partition of unity subordinated to $\mathscr{U}$ is a family of continuous functions $\eta_i : X \to [0,1]$ satisfying:

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

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

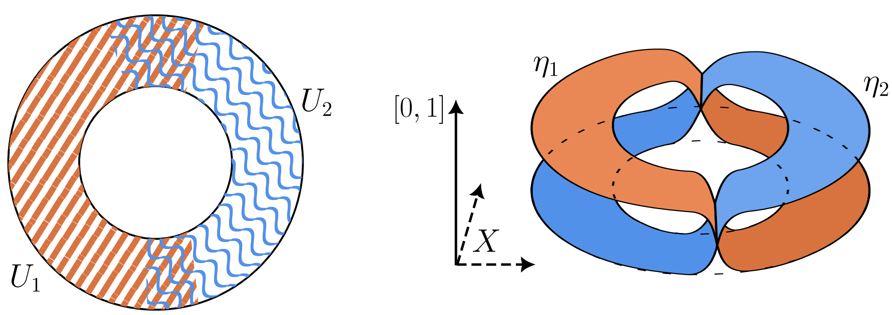

Fig. 1.1 On the left, an annulus X is covered by two open sets U_1 and U_2. The graph on the right shows two functions $\eta_i : X \to [0,1]$ that form a partition of unity subordinate to this cover.

Theorem 1.13. *Let X be a topological space and assume that $\mathscr{A} \subset \mathscr{C}(X)$ is normal and is closed under sums and quotients. Then, for any finite open cover $\mathscr{U}$, there exists an $\mathscr{A}$-partition of unity subordinated to $\mathscr{U}$.*

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

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

(St1) the remark made above that the normality of $\mathscr{A}$ 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 $\mathscr{U} = \{U_1, \ldots, U_k\}$ of a normal space X one can find another cover $\mathscr{V} = \{V_1, \ldots, V_n\}$ 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 $\mathscr{V} = \{V_i\}$, $\mathscr{W} = \{W_i\}$, with $\overline{V}_i \subset U_i$, $\overline{W}_i \subset V_i$. For each i, we use the separation property of $\mathscr{A}$ for the disjoint closed sets $(\overline{W}_i, X - V_i)$. We find $f_i : X \to [0,1]$ that belongs to $\mathscr{A}$, 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 $\mathscr{W}$ is a cover of X. From the properties of $\mathscr{A}$, each

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

is continuous. Clearly, their sum is 1. Finally, $\mathrm{supp}(\eta_i) \subset U_i$ 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 1.14. *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 $\mathscr{U} = \{U_1, \ldots, U_n\}$ 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 $\mathscr{U}$. 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 $\mathscr{S}$ 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 $\mathscr{S}$. 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) $\mathrm{supp}(\eta_i)$ 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 $\mathscr{A} \subset \mathscr{C}(X)$, we say that is **closed under locally finite sums** if for any locally finite family $\{\eta_i\}$ with $\eta_i \in \mathscr{A}$, $\sum_i \eta_i$ is again in $\mathscr{A}$.

With these, we can now talk about infinite partitions of unity:

Definition 1.15. Let X be a topological space, $\mathscr{U} = \{U_i : i \in I\}$ an open cover of X. A (continuous) partition of unity subordinated to $\mathscr{U}$ is a locally finite family of continuous functions $\eta_i : X \to [0,1]$ satisfying:

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

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

If we are pragmatic and we only care about what is directly applicable later on in this course, the result to have in mind is:

Theorem 1.16. *Let X be a Hausdorff, locally compact and 2nd countable space, $\mathscr{A} \subset \mathscr{C}(X)$ and assume that:*

- *$\mathscr{A}$ is closed under locally finite sums and under quotients and*
- *$\mathscr{A}$ satisfies: for any $x \in M$ and any open neighborhood U of x, there exists $f \in \mathscr{A}$ supported in U with $f(x) > 0$.*

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

For the curious student, here is the more detailed discussion to which the previous theorem belongs (with the explanation of how the proof goes). 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 $\mathscr{U}$ of X, there exists a locally finite open cover $\mathscr{V}$ that is a refinement of $\mathscr{U}$ in the sense that any $V \in \mathscr{V}$ is included inside some $U \in \mathscr{U}$. The existence of arbitrary partitions of unity is ensured by the following:

Theorem 1.17. *Let X be a paracompact Hausdorff space and assume that $\mathscr{A} \subset \mathscr{C}(X)$ is normal, closed under locally finite sums and closed under quotients.*

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

Since paracompact spaces are automatically normal, hence we can use Uryshon's lemma, it follows 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 $\mathscr{A}$, that $\mathscr{A}$ is normal. For the first one, the following comes in handy:

Theorem 1.18. *Any Hausdorff, locally compact and 2nd countable space is paracompact.*

In particular, topological manifolds are automatically paracompact. 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 1.7). Let $\mathscr{U}$ be an open cover of X. For each $n \in \mathbb{Z}_+$ there is a finite family $\mathscr{V}_n$ which covers $K_n - \text{Int}(K_{n-1})$, consisting of opens V with the properties: $V \subset \text{Int}(K_{n+1}) - K_{n-1}$, $V \subset U$ for some $U \in \mathscr{U}$. Indeed, for any $x \in K_n - \text{Int}(K_{n-1})$ let V_x be the intersection of $\text{Int}(K_{n+1}) - K_{n-1}$ with any member of $\mathscr{U}$ containing x; since $K_n - \text{Int}(K_{n-1})$ is compact, just take a finite subcollection $\mathscr{V}_n$ of $\{V_x\}$, covering $K_n - \text{Int}(K_{n-1})$. Set $\mathscr{V} = \cup_n \mathscr{V}_n$; it covers X since each $K_n - K_{n-1} \subset K_n - \text{Int}(K_{n-1})$ is covered by $\mathscr{V}_n$. Finally, it is locally finite: if $x \in X$, choosing n and V such that $V \in \mathscr{V}_n$, $x \in V$, we have $V \subset \text{Int}(K_{n+1}) - K_{n-1}$, hence V can only intersect members of $\mathscr{V}_m$ with $m \leq n+1$ (a finite number of them!). 😀

Finally, to check the normality of $\mathscr{A}$ needed in Theorem 1.17, the following comes in handy:

Theorem 1.19. *Let X be a Hausdorff paracompact space and $\mathscr{A} \subset \mathscr{C}(X)$ closed under locally finite sums and under quotients. If X is also locally compact, then the following are equivalent:*

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

In particular, for a topological manifold M, checking that a subset $\mathscr{A} \subset \mathscr{C}(M)$ 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 $f \in \mathscr{A}$ 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 $f_c \in \mathscr{A}$ 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 1.13, 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 $\mathscr{U} = \{U_i : i \in I\}$ 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 1.13) we can also choose an open cover of X, $\mathscr{V} = \{V_i : i \in I\}$, 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 $\eta_i \in \mathscr{A}$ such that

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

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 $\mathscr{A}$, $f \in \mathscr{A}$. 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 $\mathscr{U}$, $U_i \subset D_x$ 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. $U_i \cap A \neq \emptyset$). 😀

1.2 Reminder 2: 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.

1.2.1 $\mathbb{R}^n$

The basic playground for multivariate analysis is the standard Euclidean space

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

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

- it is a vector space. When we want to emphasize this structure, we will denoted by

$$v = (v_1, \ldots, v_n) \in \mathbb{R}^n$$

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:

$$e_1, \ldots, e_m \in \mathbb{R}^n;$$

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:

$$\langle v, w \rangle = \sum_{i=1}^{n} v_i \cdot w_i,$$

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

$$x = (x_1, \ldots, x_n) \in \mathbb{R}^n$$

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 $f : \mathbb{R}^n \to \mathbb{R}$, the vector space structure of $\mathbb{R}^n$ is not relevant (though it may be useful).
Recall also that the topology on $\mathbb{R}^n$ is a shadow of yet another structure: $\mathbb{R}^m$ is also a metric space, with the standard Euclidean metric:

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

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 $\mathbb{R}^n$ that induce the same Euclidean topology- e.g. the so called square metric

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

- even when thinking about $\mathbb{R}^n$ as a topological space, so of its elements as points, each point $x \in \mathbb{R}^n$ 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.

1.2.2 The differential and the inverse function theorem

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

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

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, $f(x+v)$ is defined for small vectors v.

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

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

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 $v \in \mathbb{R}^n$ 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

$$A \in \text{Lin}(\mathbb{R}^n, \mathbb{R}^k)$$

with the property that

$$\lim_{v \to 0} \frac{f(x+v) - f(x) - A(v)}{||v||} = 0. \tag{1.2.1}$$

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.

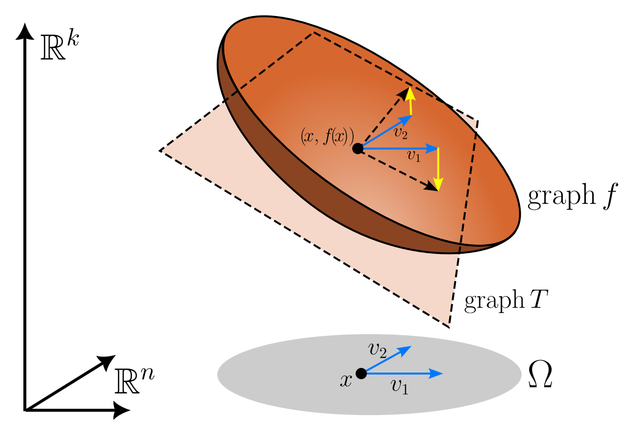

Fig. 1.2 Visualization of the total derivative for a function $f : \Omega \to \mathbb{R}^k$ defined over an open $\Omega \subseteq \mathbb{R}^n$, in this case with $n = 2$ and $k = 1$. Infinitesimal movements through a point $x \in \Omega$ are represented by the blue horizontal vectors v_i and the resulting infinitesimal movement in the codomain $\mathbb{R}^k$ is represented by the yellow vertical vectors $D_x f(v_i)$. They sum up to dashed vectors of the form $(v_i, D_x f(v_i)) \in \mathbb{R}^n \times \mathbb{R}^k$ that are tangent to the graph of f in the point $(x, f(x))$. The graph of the best affine approximation T of f in x (which sends any $\tilde{x} \in \mathbb{R}^n$ to $T(\tilde{x}) = f(x) + D_x f(\tilde{x} - x)$) is an affine plane spanned by the dashed vectors.

Definition 1.20. We say that $f : \Omega \to \mathbb{R}^k$ is **differentiable at** x if there exists a linear map $A \in \text{Lin}(\mathbb{R}^n, \mathbb{R}^k)$ satisfying (1.2.1). 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

$$D_x f = (Df)_x : \mathbb{R}^n \to \mathbb{R}^k.$$

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

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

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

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^l for any $l \in \mathbb{N}$. We say that f is **smooth** if it is of class C^l for all l.

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

Proposition 1.21 (the chain rule). *Given opens $\Omega \subset \mathbb{R}^n$, $\Omega' \subset \mathbb{R}^k$ and functions*

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

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 1.22. A map $f : \Omega \to \Omega'$ between two opens $\Omega \subset \mathbb{R}^n$ and $\Omega' \subset \mathbb{R}^k$ 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, where the main objects are topological spaces, the relevant maps are the continuous ones and two spaces are "isomorphic in Topology" (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 $\mathbb{R}^k$ are homeomorphic only when $n = k$ 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 1.23. Show that if a map $f : \Omega \to \Omega'$ between two opens $\Omega \subset \mathbb{R}^n$ and $\Omega' \subset \mathbb{R}^k$ is a local diffeomorphism around $x \in \Omega$, then

$$(Df)_x : \mathbb{R}^n \to \mathbb{R}^k$$

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

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 <u>at a single (given) point x</u> tells us information about f <u>around x</u>:

Theorem 1.24 (The inverse function theorem). *Given a smooth map $f : \Omega \to \Omega'$ between two opens $\Omega \subset \mathbb{R}^n$ and $\Omega' \subset \mathbb{R}^k$, if f is differentiable at a point $x \in \Omega$ and $(Df)_x$ is an isomorphism, then f is a local diffeomorphism around x.*

1.2.3 Directional/partial derivatives; the implicit function theorem

Note that, when talking about the differential $(Df)_x(v)$ (hence $f : \Omega \to \mathbb{R}^k$, with $\Omega \subset \mathbb{R}^n$ 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.

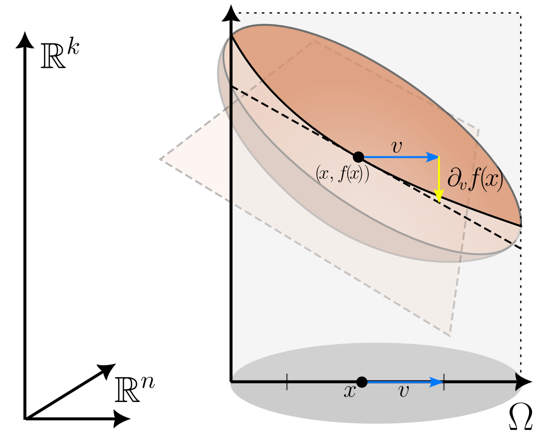

Fig. 1.3 The directional derivative $\partial_v f(x)$ of the function f from Fig. 1.2 can be seen to arise geometrically in the following way: The map $t \mapsto x + vt$ traces a line through Ω. The slice of the graph of f over this line is exactly the graph of the function $t \mapsto f(x+tv)$ if we label the horizontal axis by integer multiples of v. The directional derivative can now be defined as the normal derivative of this function since we have a one-dimensional domain.

Definition 1.25. With $f : \Omega \to \mathbb{R}^k$ and $x \in \Omega$ as above, and an arbitrary vector $v \in \mathbb{R}^n$, **the derivative of** f **at** x **in the direction** v is defined as the vector

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

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 (1.2.1) v (small enough) by tv with $v \in \mathbb{R}^n$ fixed (but arbitrary) and $t \in \mathbb{R}$ approaching 0; using that A is linear, we find that:

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

This relationship is visualized in Fig. 1.3. 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 $v \in \{e_1, \dots, e_n\}$, a vector in the standard basis of $\mathbb{R}^n$, we obtain the partial derivatives

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

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

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

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

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

can be represented as matrices

$$A = (A^i_j)_{1\le i\le k, 1\le j\le n} = \begin{pmatrix} A^1_1 & \dots & A^1_n \\ \dots & \dots & \dots \\ A^k_1 & \dots & A^k_n \end{pmatrix}$$

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):

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

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

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

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

$$v^T = \begin{pmatrix} v^1 \\ \dots \\ v^n \end{pmatrix}$$

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

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

i.e., in the matrix notation,

$$(Df)_x = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \dots & \dots & \dots \\ \frac{\partial f_k}{\partial x_1}(x) & \dots & \frac{\partial f_k}{\partial x_n}(x) \end{pmatrix}.$$

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 these:

Proposition 1.26. *A function $f : \Omega \to \mathbb{R}^k$ 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 l, partial derivatives are defined inductively:

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

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

$$\mathscr{C}^\infty(\mathbb{R}^n)$$

the space (algebra!) of smooth functions on $\mathbb{R}^n$. A **smooth partitions of unity** on $\mathbb{R}^n$ (subordinated to an open cover) is any partition of unity whose members η_i are smooth- i.e. Definition 1.12 (finite case) and Definition 1.15 (general case) applied at $\mathscr{A} := \mathscr{C}^\infty(\mathbb{R}^n) \subset \mathscr{C}(\mathbb{R}^n)$.

Theorem 1.27. *Any open cover of $\mathbb{R}^n$ admits a smooth partition subordinated to it.*

Proof. We want to use Theorem 1.17 for $\mathscr{A} := \mathscr{C}^\infty(\mathbb{R}^n)$. This is clearly closed under quotients and, for the same reason that locally finite sums of continuous functions are continuous, it is closed under locally finite sums. We still have to check the last condition on $\mathscr{A}$ or, equivalently: for any $x \in \mathbb{R}^n$ and any ball centered at x, $B(x,\varepsilon)$, there exists a smooth functions $f : \mathbb{R}^n \to [0,1]$ 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 $\varepsilon = 1$. Then set $f(x) = g(x_1^2 + \ldots + x_n^2)$ 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 1.28. 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 $\mathbb{R}^n$. While such subspaces are usually given by equations of type $f(x_1, \ldots, x_n) = 0$ (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).

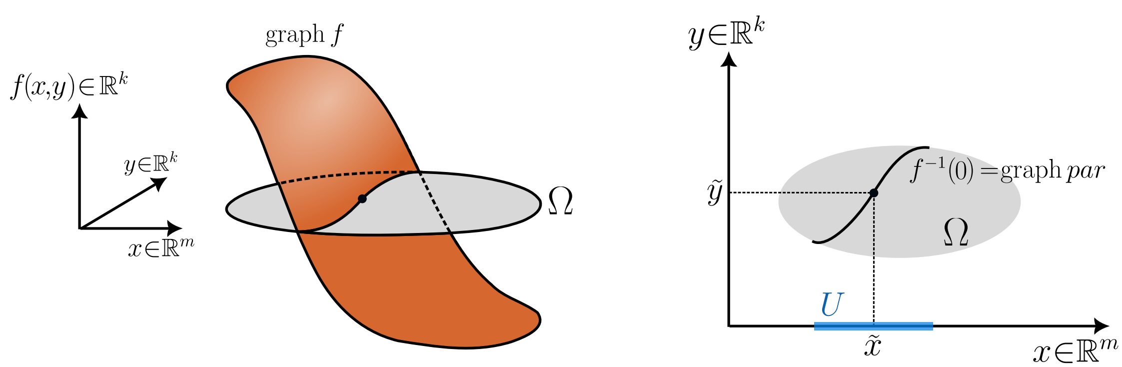

Fig. 1.4 The implicit function Theorem 1.29 illustrated for a concrete choice of $f : \Omega \to \mathbb{R}^k$. The theorem establishes that the preimage $f^{-1}(0)$ (under appropriate assumptions) can locally be written as a graph of a function $par : U \to \mathbb{R}^k$ over a subset of the variables. In other words, the condition that f vanishes *implicitly* defines the function par.

Theorem 1.29. *Let $f : \Omega \to R^k$ be a smooth map defined on an open $\Omega \subset \mathbb{R}^m \times \mathbb{R}^k$ whose elements we label as $(x,y) = (x_1,...,x_m,y_1,...,y_k)$. Furthermore let $(\tilde{x},\tilde{y}) \in \Omega$ be a point where $f(\tilde{x},\tilde{y}) = 0$ and the matrix*

$$\left(\frac{\partial f_i}{\partial y_j}(\tilde{x},\tilde{y})\right)_{1\leq i,j\leq k}$$

is non-singular. Then there exists a function $par : U \to \mathbb{R}^k$ defined in a neighborhood U of $\tilde{x}$ such that for all (x,y) near $(\tilde{x},\tilde{y})$, one has

$$f(x,y) = 0 \Longleftrightarrow y = par(x).$$

The matrix appearing in this theorem is exactly the Jacobian matrix of the map $y \mapsto f(\tilde{x},y)$ at $y = \tilde{y}$. You can convince yourself using Fig. 1.4 that its non-singularity is a neccessary condition to find a smooth function *par*: In the depicted situation, it corresponds exactly to a tangency of $f^{-1}(0)$ in the y-direction at $(\tilde{x},\tilde{y})$.

Remark 1.30. Note that this theorem is not completely canonical: It gives preference to the last k components of the arguments of f, i.e. depends on how we split the components of elements of Ω into x and y-components. For example, there is an obvious modification in which the starting assumption is that the Jacobian with respect to the *first* k components is regular, instead of the last ones. Such modifications are necessary even when looking at the simplest examples: E.g., for the unit circle where $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^2 + y^2 - 1$, the condition

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

(necessary for the theorem) is valid at *almost* all the points (x,y) in the circle (and, indeed, we can always solve $y = \pm\sqrt{1-x^2}$), except for the points $(1,0)$ and $(-1,0)$ (and, indeed, there is a problem there: we would need a function two take two values simultaneously close to $y = 0$ in order for it's graph to match $f^{-1}(0)$). However, at those points one can switch the roles of x and y- and, indeed, around those points which are problematic for $\pm\sqrt{1-x^2}$, one can write $x = \pm\sqrt{1-y^2}$.

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

$$\pi_y(\sigma \cdot p) = par(\pi_x(\sigma \cdot p)),$$

where π_x and π_y are the projections of $\mathbb{R}^{m+k}$ onto the first m and last k components, respectively, and we write $\sigma \cdot p$ for the result of permuting p by σ. But perhaps the most geometric formulation is what is know as the submersion theorem- see below.

Proof. Consider the map

$$F : \Omega \to \mathbb{R}^{m+k}, \quad F(x,y) := (x, f(x,y)).$$

Then the non-singularity condition in the statement precisely means that $(DF)_{(\tilde{x},\tilde{y})}$ is non-singular. Hence, by the inverse function theorem, we find a smooth inverse G of F, defined near $F(\tilde{x},\tilde{y})$. Given the form of F, it follows that G is of a similar form:

$$G(x,z) = (x, g(x,z)).$$

That $G \circ F$ and $F \circ G$ are the identity maps (near $(\tilde{x},\tilde{y})$, and $F(\tilde{x},\tilde{y})$, respectively) translates into

$$g(x, f(x,y)) = y \qquad \text{and} \qquad f(x, g(x,z)) = z. \tag{1.2.2}$$

The first equation shows that

$$f(x,y) = 0 \Longrightarrow g(x,0) = y,$$

hence we have an obvious candidate $par(x) := g(x,0)$. Note that the assumption $f(\tilde{x},\tilde{y}) = 0$ guarantees that g is defined for (x,z) near $(\tilde{x},0)$. The fact that, indeed, $f(x, par(x)) = 0$ for x close to $\tilde{x}$ is just the second equation in (1.2.2) applied when $z = 0$. 😀

1.2.4 Local coordinates/charts

The standard coordinates in $\mathbb{R}^n$, despite being "obvious", are often not the best ones to use in specific problems. E.g.: 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). Baby example: looking at the curve in $\mathbb{R}^2$ defined by

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

a change of coordinates of type

$$x = \frac{u+v}{3}, \quad y = \frac{u-2v}{3} \tag{1.2.3}$$

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:

$$x = r\cos(\theta), \quad y = r\sin(\theta). \tag{1.2.4}$$

To formalise such changes of coordinates, one talks about charts:

Definition 1.31. A **smooth chart of** $\mathbb{R}^n$ is a diffeomorphism

$$\chi = (\chi_1, \dots, \chi_n) : U \to \Omega \subset \mathbb{R}^n$$

between an open $U \subset \mathbb{R}^n$ and an open $\Omega \subset \mathbb{R}^n$. The open U is called the **domain of the chart** and, for $p \in U$,

$$(\chi_1(p), \dots, \chi_n(p))$$

are called the **coordinates of** p **w.r.t. the chart** (U, χ) and we also say that (U, χ) is a smooth chart around p.

For instance the change of coordinates (1.2.3) is about the chart

$$\chi : \mathbb{R}^2 \to \mathbb{R}^2, \quad \chi(x,y) = (2x+y, x-y) \tag{1.2.5}$$

so that, in the new coordinates, a point $p = (x,y)$ will have the coordinates (w.r.t. χ)

$$u(x,y) = 2x+y, \quad v(x,y) = x-y.$$

Similarly for the polar coordinates where, computing the inverse of $(r,\theta) \mapsto (r\cos(\theta), r\sin(\theta))$, one finds the chart

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

1.2.5 Changing coordinates to make functions simpler (the immersion/submersion theorem)

In general, given a smooth function

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

and a point $p \in \mathbb{R}^n$, 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 (the identity maps) 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 (1.2.5) 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

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

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

$$(Df)_p : \mathbb{R}^n \to \mathbb{R}^k$$

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

Theorem 1.32 (the submersion theorem). *Assume that f is a* ***submersion*** *at a given point $p \in U$ in the sense that $(Df)_p : \mathbb{R}^n \to \mathbb{R}^k$ is surjective. Then there exists a smooth chart χ of $\mathbb{R}^n$ around p such that, around $\chi(p)$, $f_\chi = f \circ \chi^{-1}$ is given by*

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

Proof. Since the matrix representing $(Df)_p$ is of maximal rank, one of its maximal minors (an $k \times k$ matrix) is invertible; we may assume that the invertible minor is precisely the one made of the last k rows (why?)- which is also the hypothesis of the implicit function theorem (Theorem 1.29). Looking at the proof of the theorem, one remarks that the desired chart is $\chi = \tilde{f}$.

A similar argument gives rise to the following:

Theorem 1.33 (the immersion theorem). *Assume that f is an immersion at a given point $p \in U$ in the sense that $(Df)_p : \mathbb{R}^n \to \mathbb{R}^k$ is injective. Then there exists a smooth chart χ' of $\mathbb{R}^k$ around $f(p)$ such that, in a neighborhood p, $f^{\chi'} = \chi' \circ f$ is given by*

$$f^{\chi'}(x_1, \ldots, x_n) = (x_1, \ldots, x_n, \underbrace{0, \ldots, 0}_{k-n \text{ zeros}}). \tag{1.2.6}$$

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

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

such that (1.2.6) holds on Ω_p. Furthermore, one may choose χ' and Ω_p so that:

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

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

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

is an isomorphism. Consider then

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

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 $\mathbb{R}^k$.
2. Ω' is an open neighborhood of $(p, 0)$ in $\mathbb{R}^k$, contained in $U \times \mathbb{R}^{k-n}$.

3. the intersection of $\Omega' \subset \mathbb{R}^k$ with $\mathbb{R}^n \times \{0\}$,
$$\Omega := \{u \in \mathbb{R}^n : (u,0) \in \Omega'\},$$
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 $\chi_2' : U' \to \mathbb{R}^{k-n}$ (the second component of the chart χ' w.r.t. the decomposition $\mathbb{R}^k = \mathbb{R}^n \times \mathbb{R}^{k-n}$). 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 $M \subset \mathbb{R}^n$.

Definition 1.34. Given $M \subset \mathbb{R}^n$ and a function $f : M \to \mathbb{R}^k$, we say that f **is smooth around** $p \in M$ if, in a neighborhood U of p in M, $f|_U$ admits a smooth extension to an open inside $\mathbb{R}^n$ containing U. When this happens around all $p \in M$, we say that f is **smooth**.

A **diffeomorphism** between $M \subset \mathbb{R}^n$ and $N \subset \mathbb{R}^k$ is any bijection $f : M \to N$ with both f and f^{-1} smooth.

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

Exercise 1.35. Show that if $M \subset \mathbb{R}^n$ is a closed subset that any smooth function $f : M \to \mathbb{R}^k$ admits a smooth extension $\tilde{f} : \mathbb{R}^n \to \mathbb{R}^k$. (Hint: $\mathbb{R}^n \setminus M$ is open; get an open cover of $\mathbb{R}^n$ out of one of M.)

1.2.6 Embedded submanifolds of $\mathbb{R}^L$

We now move to the notion of (smooth) embedded submanifolds of $\mathbb{R}^L$ [1]. In low dimensions, these are curves (1-dimensional) and surfaces (2-dimensional); for an arbitrary dimension m we will be talking about m-dimensional submanifolds of $\mathbb{R}^L$. For instance, the standard sphere
$$S^m = \{(x_0, \ldots, x_m) \in \mathbb{R}^{m+1} : \sum_i (x_i)^2 = 1\} \quad (L = m+1)$$
will be such a smooth m-dimensional submanifold. 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 $t \in \mathbb{R}$.

Accordingly, the notion of submanifold of $\mathbb{R}^L$ can be introduced in several ways that look differently (but which turn out to be equivalent).

We start with the definition that can be seen as just a small variation on the notion of topological manifold from Definition 1.3 just that, for $M \subset \mathbb{R}^L$ the axioms (TM1), (TM2) are automatically satisfied, and one can can further take advantage of the Euclidean space to talk about *smoothness* of charts- as in Definition 1.34.

Definition 1.36. An m-dimensional embedded submanifold of $\mathbb{R}^L$ is any subset $M \subset \mathbb{R}^L$ which, for each $p \in M$, satisfies **the** (m-dimensional) **manifold condition at** p in the following sense: there exists a topological chart of M (Definition 1.3)

[1] here L is an integer, possibly large, that will denote the dimension of the Euclidean space inside which our manifolds $M \subset \mathbb{R}^L$; we use here the letter L not only to suggest that L may be possibly large w.r.t. the dimension of M, but also to emphasise that the role of the dimension L is very different than that of the dimension m of M

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

($U \subset M$ open neighborhood of p, Ω open in $\mathbb{R}^m$) which is also a diffeomorphism (i.e. χ and χ^{-1} are smooth in the sense of Definition 1.34). These will also be called **smooth (m-dimensional) charts** for M.

Of course, when $M = \mathbb{R}^L$, the resulting notion of "smooth chart for $\mathbb{R}^L$" coincides with the one already introduced in Definition 1.31. For general M, a particularly nice class of smooth charts of M are the ones that can be obtained by restricting such charts of $\mathbb{R}^L$. More precisely, given a subset $M \subset \mathbb{R}^L$, a smooth chart of $\mathbb{R}^L$

$$\widetilde{\chi} : \widetilde{U} \to \widetilde{\Omega} \subset \mathbb{R}^L \tag{1.2.7}$$

said to be **adapted to** M if it takes $U := M \cap \widetilde{U}$ into $\Omega := \widetilde{\Omega} \cap (\mathbb{R}^m \times \{0\})$:

$$\widetilde{\chi}|_U : U \to \Omega. \tag{1.2.8}$$

Equivalently: inside $\widetilde{U} \subset \mathbb{R}^L$, the points that belong to M are characterised by the equations $\widetilde{\chi}_i = 0$ for $i > m$:

$$M \cap \widetilde{U} = \{q \in \widetilde{U} : \widetilde{\chi}_{m+1}(q) = \ldots = \widetilde{\chi}_L(q) = 0.\}$$

Note also that Ω may be, and will be, interpreted as an open in $\mathbb{R}^m$; in this way, any smooth chart (1.2.7) of $\mathbb{R}^L$ that is adapted to M induces a smooth chart (1.2.8) of M.

Not every smooth chart χ of M is induced by an adapted smooth chart $\widetilde{\chi}$ of $\mathbb{R}^L$. However:

Proposition 1.37. *For $M \subset \mathbb{R}^L$ and $p \in M$, the manifold condition for M at p is equivalent to the existence of a smooth chart of $\mathbb{R}^L$ around p, that is adapted to M.*

The proof will be done together with the proof of the following theorem. This theorem describes submanifolds parametrically (think of $x = \cos(t), y = \sin(t)$ for the circle) and by equations (think of $x^2 + y^2 = 1$ for the circle), taking care of the precise conditions.

Theorem 1.38. *Given a subset $M \subset \mathbb{R}^L$, $p \in M$, the following are equivalent:*

1. *M satisfies the m-dimensional manifold condition at p.*
2. *M admits an m-dimensional **parametrization around** p- by which we mean a homeomorphism*

$$par : \Omega \to U \subset M$$

between an open $\Omega \subset \mathbb{R}^m$ and an open neighborhood U of p in M satisfying the regularity condition that, as a map from Ω to $\mathbb{R}^L$, par is an immersion.

3. *M can be described by an m-dimensional **implicit equation around** p- by which we mean a submersion*

$$eq : \widetilde{U} \to \mathbb{R}^{L-m}$$

defined on an open neighborhood $\widetilde{U}$ of p in $\mathbb{R}^L$ and which describes M near p by the equation $eq = 0$:

$$M \cap \widetilde{U} = \{q \in \widetilde{U} : eq(q) = 0\}.$$

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

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

Therefore, we will deal with

- neighborhoods U_p of p in M, and $\widetilde{U}_p$ of p in $\mathbb{R}^L$.
- neighborhoods Ω_x of x in $\mathbb{R}^m$.

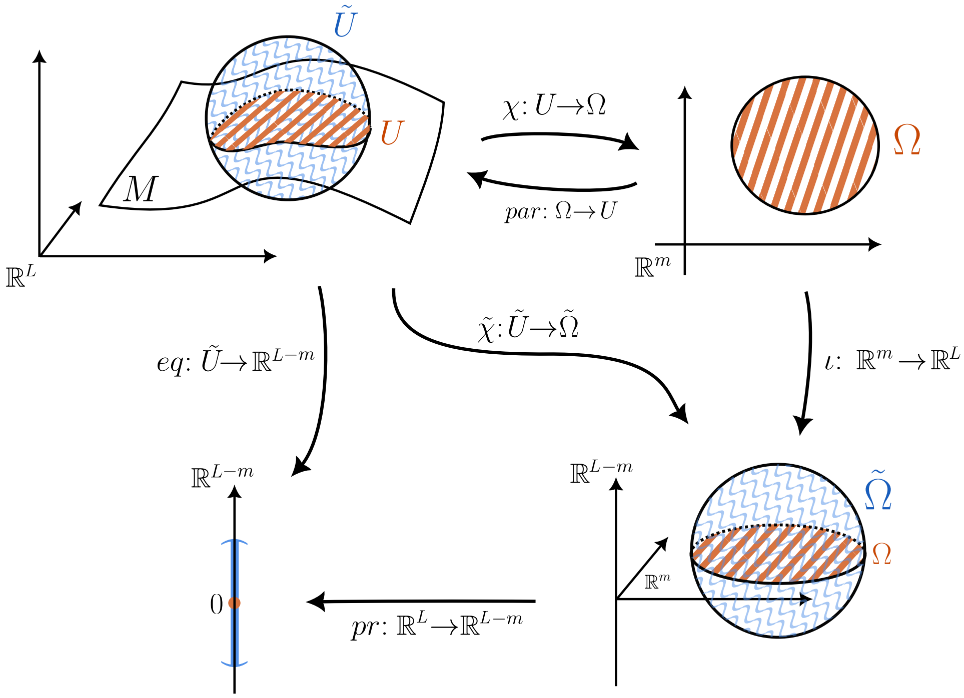

Fig. 1.5 The equivalent ways of phrasing the manifold condition in Theorem 1.38 involve (adapted) charts, parametrizations or implicit equations that are related as shown in this figure. Note that $U = \tilde{U} \cap M$ and $\Omega = \tilde{\Omega} \cap \mathbb{R}^m$. ι and pr are canonical inclusions and projections of the product space $\mathbb{R}^L = \mathbb{R}^m \times \mathbb{R}^{L-m}$ in the lower right. All these maps commute where they can be evaluated.

The points in the neighborhoods of p will be denoted by q, while the ones in the neighborhoods of x by y; for them, we may be looking at similar neighborhoods U_q, $\widetilde{U}_q$ and Ω_y.

We first prove that (1) implies (2). We start with the the chart $\chi : U \to \Omega \subset \mathbb{R}^m$ 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 $\Omega \to \mathbb{R}^L$, it is an immersion. For the last part use that the composition

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

is the inclusion $U \subset \mathbb{R}^L$ and then apply the chain rule to deduce that, for each point $q \in U$, $(D\chi)_{par(q)} \circ (D\,par)_q$ is the identity- hence, in particular, $(D\,par)_q$ 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 (1), we still have to check that χ is smooth in the sense of Definition 1.34: i.e., around any point $q \in U$, it is obtained by restricting a smooth map defined on an open $\widetilde{U}_q \subset \mathbb{R}^L$. For that we use the immersion theorem (Theorem 1.33) applied to $par : \Omega \to \mathbb{R}^L$ around

$$y = \chi(q) \in \Omega.$$

We find:

- an open neighborhood Ω_y of y in $\Omega \subset \mathbb{R}^m$
- a diffeomorphism $\widetilde{\chi} : \widetilde{U}_q \to \widetilde{\Omega}_q$ from an open neighborhood $\widetilde{U}_q \subset \mathbb{R}^L$ of p' to an open $\widetilde{\Omega}_q \subset \mathbb{R}^L$,

so that, on Ω_y, $\widetilde{\chi} \circ par$ becomes the inclusion on the first factors. We now write $\widetilde{\chi} = (\widetilde{\chi}_1, \widetilde{\chi}_2)$ where we use again the decomposition $\mathbb{R}^L = \mathbb{R}^m \times \mathbb{R}^{L-m}$. We deduce that $\widetilde{\chi}_1(par(z)) = z$ for all $z \in \Omega_y$. Since $z = \chi(par(z))$ for all $z \in \Omega$, we deduce that $\widetilde{\chi}_1(r) = \chi(r')$ for all $r \in par(\Omega_q)$. In this way, on the neighborhood $par(\Omega_q)$ of q in U, χ is now the restriction of a smooth function defined on an open neighborhood of q in $\mathbb{R}^L$- namely $\widetilde{\chi}_1 : \widetilde{U}_q \to \mathbb{R}^m$.

To prove (3) (still assuming (2)), we use $q = p$ in the previous reasoning and the resulting diffeomorphism $\widetilde{\chi} : \widetilde{U}_p \to \widetilde{\Omega}_p$; in principle, the desired function f will be $\widetilde{\chi}_2$, but we have to choose the domain of definition carefully. For that we use the last part of Theorem 1.33 which says that we may assume that

$$par(\Omega_x) = \{y \in \widetilde{U}_p : \widetilde{\chi}_{m+1}(y) = 0, \ldots, \widetilde{\chi}_L(y) = 0\}.$$

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

$$\widetilde{U} := \widetilde{U}_p \cap W_p, \quad eq = \widetilde{\chi}_2|_{\widetilde{U}} : \widetilde{U} \to \mathbb{R}^{L-m}$$

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

Exercise 1.39. Conclude now that $\widetilde{\chi}$ is actually an adapted chart.

We are now left with proving that (3) implies (1). Let $eq : \widetilde{U} \to \mathbb{R}^{L-m}$ satisfying the conditions from the hypothesis. Note that if we replace $\widetilde{U}$ by a smaller open neighborhood of p in $\mathbb{R}^L$ (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 $\widetilde{\chi} : \widetilde{U} \to \widetilde{\Omega}$ into an open subset of $\mathbb{R}^L$, such that $eq = \widetilde{\chi}_2$. This chart will then take the zero set of eq into the zero set of the second projection $\mathrm{pr}_2 : \widetilde{\Omega} \to \mathbb{R}^{L-m}$, i.e. into

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

We deduce that the restriction of $\widetilde{\chi}$ to $U = M \cap \widetilde{U}$,

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

is a smooth chart of M (around p). 😀

Example 1.40. 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 1.41. Generalize this discussion to the spheres S^m of arbitrary dimension.

1.2.7 From directional derivatives to tangent spaces

The point of 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

$$\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$$

(with $\varepsilon > 0$) satisfying

$$\gamma(0) = p, \quad \frac{d\gamma}{dt}(0) = v. \tag{1.2.9}$$

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 1.42. *If f is differentiable at p then, for any path γ satisfying (1.2.9), one has*

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

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

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

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

Definition 1.43. Let $M \subset \mathbb{R}^n$ and consider a point $p \in M$. A **smooth curve in** M is any smooth map $\gamma : I \to \mathbb{R}^n$ defined on some interval $I \subset \mathbb{R}$, which takes values in M.

A **vector tangent to** M **at** p is any vector $v \in \mathbb{R}^n$ which can be realized as the speed at $t = 0$ of a smooth curve in M that passes through p at $t = 0$ (i.e. for which $0 \in I$ and $\gamma(0) = p$):

$$v = \frac{d\gamma}{dt}(0).$$

The set of such vectors is denoted by $T_p^{\text{geom}}M$; hence

$$T_p^{\text{geom}}M \subset \mathbb{R}^n.$$

Although we use the name ”tangent *space*”, in general (for completely random Ms inside $\mathbb{R}^n$), $T_p^{\text{geom}}M$ is just a subset of $\mathbb{R}^n$ (... but it is a vector subspace if M is ”nice”).

Exercise 1.44. Compute $T_p^{\text{geom}}M$ 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 1.45. Assume that $M \subset \mathbb{R}^n$ is defined by an equation $f(x) = 0$, where $f : \mathbb{R}^n \to \mathbb{R}^k$ is a smooth function. For $p \in \mathbb{R}^n$ we denote by $\text{Ker}_p(Df)$ the kernel (= the zero set) of the differential $(Df)_p : \mathbb{R}^n \to \mathbb{R}^k$. Show that, in general,

$$T_p^{\text{geom}}(M_f) \subset \text{Ker}_p(Df),$$

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

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

Exercise 1.46. With the notations from the previous exercise show that for all $p \in M_f$ at which f is a submersion

$$T_p^{\text{geom}}(M_f) = \text{Ker}_p(Df).$$

We now return to our discussion on differentials/directional derivatives, recast in terms of tangent spaces. Namely, Lemma 1.42 gives us right away:

Corollary 1.47. *Given $M \subset \mathbb{R}^n$, $p \in M$ and a function $\tilde{f} : \mathbb{R}^n \subset \mathbb{R}^k$ differentiable at p then, for any vector $v \in \mathbb{R}^n$ tangent to M at p, $(D\tilde{f})_p(v)$ depends only on $\tilde{f}|_M$.*

Of course, a similar conclusion holds slightly more generally, for any function $\tilde{f} : U \to \mathbb{R}^k$ defined on an open neighborhood $U \subset \mathbb{R}^n$ of p- the otcome being 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 $f : M \to \mathbb{R}^k$ which is differentiable at $p \in M$ in the sense of Definition 1.34: for $f : M \to \mathbb{R}^k$ that is differentiable at p, one has a well-defined differential

$$(Df)_p : T_p^{\text{geom}}M \to \mathbb{R}^k,$$

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

Exercise 1.48. Show that if $M \subset \mathbb{R}^n$ and $N \subset \mathbb{R}^k$ and $f : M \to N$ is smooth at $p \in M$, then $(Df)_p$ takes values in $T^{\text{geom}}_{f(p)}N$. Then prove the chain rule in this context and deduce that, if f is a diffeomorphism, then $(Df)_p$ is a bijection between $T^{\text{geom}}_p M$ and $T^{\text{geom}}_{f(p)}N$.

Finally, let us look at tangent spaces of submanifolds of $\mathbb{R}^n$.[2]

Proposition 1.49. *If $M \subset \mathbb{R}^n$ is a m-dimensional embedded submanifold then, for any $p \in M$, the tangent space of M at p is an m-dimensional vector subspace of $\mathbb{R}^n$, which can also be described as follows:*

1. *as the kernel of $(Deq)_p : \mathbb{R}^n \to \mathbb{R}^{L-m}$, where $eq : \widetilde{U} \to \mathbb{R}^{L-m}$ is any implicit equation defining M around p.*
2. *as the image of $(D\,par)_p : T_p\Omega \to \mathbb{R}^n$, where $par : \Omega \to M$ is any parametrization of M around p.*

Proof. Exercise. 😡

Exercise 1.50. Compute again the tangent spaces of the spheres, but applying now the previous proposition.

1.2.8 More exercises

Exercise 1.51. Consider two smooth functions

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

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

$$\frac{\partial g \circ f}{\partial x_i}(x) = \sum_{j=1}^{n} \frac{\partial g}{\partial y_j}(f(p)) \frac{\partial f_j}{\partial x_i}(x)$$

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

Exercise 1.52. Show that for any function $g : \mathbb{R} \to \mathbb{R}$, the function

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

is not a submersion at $x = 0$.

Exercise 1.53. Assume that $f : U_0 \to \mathbb{R}^k$ is a smooth map, $U \subset \mathbb{R}^n$ open, $p \in U$. Let

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

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

Exercise 1.54. Consider

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = 3 \cdot \sqrt[3]{x^2 + 2xy + 2y^2}$$

and look around $p = (1,0)$. Find a chart χ of $\mathbb{R}^2$ around p and a chart χ of $\mathbb{R}$ around $f(p) = 3$ such that, w.r.t. these charts,

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

[2] is this the right place?

Exercise 1.55. Show that

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

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 1.56. Show that if $f:U\to\mathbb{R}^k$, $p\in U$ 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 1.57. Assume that $f:U\to\mathbb{R}^k$ is a smooth map, $U\subset\mathbb{R}^n$ open, $p\in U$. Let χ be a chart of $\mathbb{R}^n$ around p and let χ' be a chart of $\mathbb{R}^k$ 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 1.58. 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 1.59. For any $\varepsilon>0$ describe a smooth function

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

with the property that $f(0)>0$ and whose support (in $\mathbb{R}^n$) is contained in the ball $\{x\in\mathbb{R}^n: ||x||<\varepsilon\}$.

Exercise 1.60. Assume that $f:U\to U'$ and $g:U'\to U''$ are two smooth functions, with $U\subset\mathbb{R}^n$, $U'\subset\mathbb{R}^k$ 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).