Chapter 1

Reminder on differential calculus

What you should know or be able to do after this chapter

- Know the definition of the differential, and be able to use it.
- Be able to compute the differential or partial derivatives of a function, when given an explicit expression.
- Be able to convert between the different expressions of the differential (linear map $\leftrightarrow$ Jacobian matrix $\leftrightarrow$ partial derivatives).
- Know that a differentiable map has partial derivatives, but be able to give an example of a map which has partial derivatives, and no differential.
- Prove the classical result on the differentiability of a composition of differentiable functions.
- Be able to apply this result to an explicit example (with no error on the point at which each differential must be computed!).
- Know the definition of the gradient and Hessian.
- Know the definitions of homeomorphism and diffeomorphism.
- When you want to prove that a function is locally invertible, think to the local inversion theorem, and be able to apply it correctly.
- When you want to parametrize a set defined by an equation, think to the implicit function theorem, and be able to apply it correctly.
- Propose examples which show that the assumption "$\partial_y f(x_0, y_0)$ is bijective" is necessary.
- Know the definition of an immersion and a submersion.
- Be able to apply the normal form theorems on explicit examples.
- When you want to upper bound the values of a differentiable function, or the difference between its values, think to the mean value inequality, and be able to apply it.

1.1 Definition of differentiability

Let $(E, ||.||_E)$, $(F, ||.||_F)$, and $(G, ||.||_G)$ be normed vector spaces. We denote the set of continuous linear mappings from E to F by $\mathcal{L}(E, F)$ [1].

[1] Recall that when E is of finite dimension, all linear mappings from E to F are continuous. This is no longer true if E is of infinite dimension.

Definition 1.1 : differentiability at a point

Let $U \subset E$ be an open set, and $f : U \to F$ be a function.
If x is a point in U, we say that f is *differentiable at* x if there exists $L \in \mathcal{L}(E, F)$ such that

$$\frac{||f(x+h) - f(x) - L(h)||_F}{||h||_E} \to 0 \quad \text{as } ||h||_E \to 0,$$

(or, equivalently, $f(x+h) = f(x) + L(h) + o(||h||_E)$).
We then call L the *differential of* f *at* x and denote it $df(x)$.

Remark

If $(E, ||.||_E) = (\mathbb{R}, |.|)$, then the differential, when it exists, takes the form

$$h \in \mathbb{R} \quad \to \quad hz_x \in F,$$

for a certain element z_x in F. In this case, we write

$$f'(x) = z_x.$$

We then recover the well-known formula:

$$f(x+h) = f(x) + f'(x)h + o(h) \quad \text{as } h \to 0.$$

Definition 1.2 : functions of class C^n

Let $U \subset E$ be an open set, and $f : U \to F$ a function.
The function f is said to be *differentiable on* U if it is differentiable at every point of U.
It is *of class* C^1 if it is differentiable and $df : U \to \mathcal{L}(E, F)$ is a continuous mapping.
More generally, for any $n \geq 1$, it is *of class* C^n if it is differentiable and df is of class C^{n-1}.
It is *of class* C^∞ if it is of class C^n for every $n \geq 1$.

We won't revisit the basic properties related to differentiability (e.g., the sum of differentiable functions is differentiable, etc.), except for the one on functions defined by composition.

Theorem 1.3 : composition of differentiable functions

Let $U \subset E, V \subset F$ be open sets. Let $f : U \to V$ and $g : V \to G$ be two functions. Let $x \in U$.
If f is differentiable at x and g is differentiable at $f(x)$, then

- $g \circ f$ is differentiable at x;
- $d(g \circ f)(x) = dg(f(x)) \circ df(x)$.

1.2 Partial derivatives

In differential geometry, it is common to perform explicit calculations involving differentials of functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. For this purpose, it is useful to represent differentials as matrices of size $m \times n$ (or vectors if $m = 1$) whose coordinates can be computed. The concept of *partial derivatives* allows us to achieve this.

Definition 1.4 : partial derivative

Let $n \in \mathbb{N}^*$. Let U be an open subset of $\mathbb{R}^n$, and $f : U \to \mathbb{R}$ a function.

Let $x = (x_1, \dots, x_n) \in U$. For any $i = 1, \dots, n$, we say that f is *differentiable with respect to its i-th variable* at x if the function

$$y \to f(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots)$$

is differentiable at x_i. We then denote the derivative as $\partial_i f(x)$, $\partial_{x_i} f(x)$, or $\frac{\partial f}{\partial x_i}(x)$.

Remark

If f is differentiable at x, then it is also differentiable at x with respect to each of its variables. The converse is not necessarily true.

Remark

More generally, if $E_1, \dots, E_n, F$ are normed vector spaces, U is an open subset of $E_1 \times \cdots \times E_n$, and $f : U \to F$ is a function, we can define, for all $x = (x_1, \dots, x_n) \in U$ and $i = 1, \dots, n$, the partial derivative of f with respect to x_i,

$$\partial_{x_i} f(x) \in \mathcal{L}(E_i, F).$$

Now let $n, m \in \mathbb{N}^*$ be integers, U an open subset of $\mathbb{R}^n$, and $f : U \to \mathbb{R}^m$ a differentiable function. For any x, $df(x)$ is a linear mapping from $\mathbb{R}^n \to \mathbb{R}^m$; we denote $Jf(x)$ its matrix representation in the canonical bases. If we identify $\mathbb{R}^n$ (respectively $\mathbb{R}^m$) with the set of column vectors of size n (respectively m), then

$$\forall u \in \mathbb{R}^n, \quad df(x)(u) = Jf(x) \times u.$$

The matrix $Jf(x)$ is called the *Jacobian matrix* of f at the point x.

Proposition 1.5

Let $f_1, \dots, f_m : U \to \mathbb{R}$ be the components of f. Then, for any x,

$$Jf(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \dots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}.$$

Proof. Fix $x = (x_1, \dots, x_n) \in U$. Let $\nu \in 1, \dots, n$. Denote e_ν as the ν-th vector of the canonical basis of $\mathbb{R}^n$ (i.e., the vector whose coordinates are all 0 except the ν-th one, which is 1).

According to the definition of the differential,

$$\begin{aligned} f(x_1, \dots, x_{\nu-1}, y, x_{\nu+1}, \dots) &= f(x + (y - x_\nu)e_\nu) \\ &= f(x) + (y - x_\nu) df(x)(e_\nu) + o(y - x_\nu) \\ &\qquad \text{as } y \to x_\nu. \end{aligned}$$

For any $\mu \in 1, \dots, m$, we have

$$f_\mu(x_1, \dots, x_{\nu-1}, y, x_{\nu+1}, \dots) = f_\mu(x) + (y - x_\nu)(df(x)(e_\nu))_\mu + o(y - x_\nu) \quad \text{as } y \to x_\nu.$$

Thus, according to the definition of the partial derivative,

$$\begin{aligned} \partial_\nu f_\mu(x) &= \lim_{y \to x_\nu} \frac{f_\mu(x_1, \dots, x_{\nu-1}, y, x_{\nu+1}, \dots) - f_\mu(x)}{y - x_\nu} \\ &= (df(x)(e_\nu))_\mu. \end{aligned}$$

By the definition of the Jacobian matrix, $(Jf(x))_{\mu,\nu} = (df(x)(e_\nu))_\mu$, so

$$(Jf(x))_{\mu,\nu} = \partial_\nu f_\mu(x).$$

□

Example 1.6

Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be such that, for every $(x_1, x_2) \in \mathbb{R}^2$,

$$f(x_1, x_2) = (x_1 x_2, x_1 + x_2).$$

It is differentiable. Its Jacobian matrix is

$$\forall (x_1, x_2) \in \mathbb{R}^2, \quad Jf(x_1, x_2) = \begin{pmatrix} x_2 & x_1 \\ 1 & 1 \end{pmatrix}$$

and its differential is

$$\forall (x_1, x_2), (h_1, h_2) \in \mathbb{R}^2, \quad df(x_1, x_2)(h_1, h_2) = (h_1 x_2 + h_2 x_1, h_1 + h_2).$$

In the particular case where $m = 1$, the Jacobian matrix has a single row:

$$\forall x \in U, \quad Jf(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) & \frac{\partial f}{\partial x_2}(x) & \dots & \frac{\partial f}{\partial x_n}(x) \end{pmatrix}.$$

Its transpose is then called the *gradient*:

$$\forall x \in U, \quad \nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}.$$

For all $x \in U, h = (h_1, \dots, h_n) \in \mathbb{R}^n$,

$$df(x)(h) = Jf(x) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) h_i = \langle \nabla f(x), h \rangle,$$

where the notation "$\langle .,. \rangle$" denotes the usual scalar product in $\mathbb{R}^n$.

Still assuming $m = 1$, let us consider the case where f is twice differentiable. Its second differential can also be represented by a matrix. Indeed, for any x, $d^2 f(x) = d(df)(x)$ belongs to $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$. The map

$$(h, l) \in \mathbb{R}^n \times \mathbb{R}^n \quad \to d^2 f(x)(h)(l) \tag{1.1}$$

is therefore bilinear. As stated in the following property, it is even a quadratic form (i.e., it is symmetric), and the matrix associated with it in the canonical basis has a simple expression in terms of the partial derivatives of f.

Proposition 1.7 : Hessian matrix

Let $x \in U$. The map defined in (1.1) is a symmetric bilinear form. The matrix representing it in the canonical basis is

$$H(f)(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x) & \dots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}. \tag{1.2}$$

It is called the *Hessian matrix* of f at point x.

Exercise 1 : Proof of Proposition 1.7

1. Prove Equation (1.2).

In the rest of the exercise, we show that $H(f)(x)$ is symmetric. For this, we fix $i,j \in \{1,\dots,n\}$ such that $i \neq j$ and show

$$\frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(x) = \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}(x).$$

We denote e_i, e_j the i-th and j-th vectors of the canonical basis. For any $t,u \in \mathbb{R}$ such that $x+te_i+ue_j \in U$, we define

$$\phi(t,u) = f(x+te_i+ue_j) - f(x+te_i) - f(x+ue_j) + f(x).$$

2. a) Show that, for all t,u close enough to 0,

$$\phi(t,u) = \int_0^u \left[\frac{\partial f}{\partial x_j}(x+te_i+se_j) - \frac{\partial f}{\partial x_j}(x+se_j)\right] ds.$$

   b) Let $\epsilon > 0$ be any positive number. Show that, for all t,s close enough to 0,

$$\left|\frac{\partial f}{\partial x_j}(x+te_i+se_j) - \frac{\partial f}{\partial x_j}(x+se_j) - t\frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(x)\right| \leq \epsilon\left(|t|+|s|\right).$$

   c) Deduce from the previous question that, for all t,u close enough to 0,

$$\left|\phi(t,u) - tu\frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(x)\right| \leq \epsilon(|t|\,|u|+|u|^2).$$

   d) Show that, for all t,u close enough to 0,

$$\left|\phi(t,u) - tu\frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}(x)\right| \leq \epsilon(|t|\,|u|+|t|^2).$$

   e) Conclude.

1.3 Local inversion

Definition 1.8 : homeomorphism

Let U,V be two topological spaces[a]. A map $\phi : U \to V$ is a *homeomorphism* from U to V if it satisfies the following three properties:

1. ϕ is a bijection from U to V;
2. ϕ is continuous on U;
3. ϕ^{-1} is continuous on V.

[a] Readers not familiar with the concept of "topological space" can limit themselves to the case where U and V are two metric spaces, or even to the case where U and V are subsets, respectively, of $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$ for $n_1, n_2 \in \mathbb{N}$.

Definition 1.9 : diffeomorphism

Let $n \in \mathbb{N}^*$ be an integer, $U, V \subset \mathbb{R}^n$ be two open sets. A map $\phi : U \to V$ is a *diffeomorphism* if it satisfies the following three properties:

1. ϕ is a bijection from U to V;
2. ϕ is C^1 on U;
3. ϕ^{-1} is C^1 on V.

If, moreover, ϕ and ϕ^{-1} are C^k for an integer $k \in \mathbb{N}^*$, we say that ϕ is a C^k*-diffeomorphism*.

Theorem 1.10 : local inversion

Let $n, k \in \mathbb{N}^*$ be integers, $U, V \subset \mathbb{R}^n$ be two open sets, and $x_0 \in U$. Let $\phi : U \to V$ be a C^k map.
If $d\phi(x_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ is bijective, then there exist $U_{x_0} \subset U$ an open neighborhood of x_0 and $V_{\phi(x_0)} \subset V$ an open neighborhood of $\phi(x_0)$ such that ϕ is a C^k-diffeomorphism from U_{x_0} to $V_{\phi(x_0)}$.

For the proof of this result, one can refer to [Paulin, 2009, p. 250].

An important consequence of the local inversion theorem is the implicit functions theorem, which allows to parameterize the set of solutions of an equation.

Theorem 1.11 : implicit functions

Let $n, m \in \mathbb{N}^*$. Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set, $f : U \to \mathbb{R}^m$ be a C^k map for an integer $k \in \mathbb{N}^*$, and (x_0, y_0) be a point in U such that

$$f(x_0, y_0) = 0.$$

If $\partial_y f(x_0, y_0) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ is bijective, then there exist

- an open neighborhood $U_{(x_0,y_0)} \subset U$ of (x_0, y_0),
- an open neighborhood $V_{x_0} \subset \mathbb{R}^n$ of x_0,
- a map $g : V_{x_0} \to \mathbb{R}^m$ of class C^k

such that, for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$,

$$\big((x, y) \in U_{(x_0,y_0)} \text{ and } f(x, y) = 0\big) \quad \iff \quad (x \in V_{x_0} \text{ and } y = g(x))\,.$$

To get an intuitive feeling on this theorem, the condition "$f(x, y) = 0$" should be interpreted as an equation depending on a parameter x, whose unknown is y. The theorem states that, in the neighborhood of (x_0, y_0), the equation has, for each value of the parameter x, a unique solution (which is $g(x)$) and that this solution is C^k relatively to x.

Example 1.12

There exists an open neighborhood $U_{(1,1/2)} \subset \mathbb{R}^2$ of $(1, 1/2)$ and an open neighborhood $U_1 \subset \mathbb{R}$ of 1 such that the solutions of the equation

$$\cos(\pi x) - \cos(\pi y) + 3x^2y^2 + \frac{x^4}{4} = 0$$

for $(x, y) \in U_{(1,1/2)}$ are exactly the points of the set $\{(x, g(x))\}$ for a certain function $g : U_1 \to \mathbb{R}$ of class C^∞.

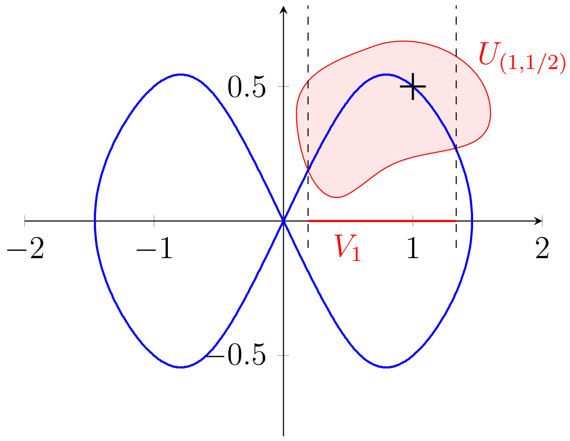

Figure 1.1: In blue, $\{(x,y)\in\mathbb{R}^2, \cos(\pi x)-\cos(\pi y)+3x^2y^2+\frac{x^4}{4}=0\}$. This set is not the graph of a function. However, the part of the set inside $U_{(1,1/2)}$ coincides with the graph of a function $g : V_1 \to \mathbb{R}$.

This is proven by applying the implicit functions theorem to

$$f : (x,y)\in\mathbb{R}\times\mathbb{R} \quad\to\quad \cos(\pi x)-\cos(\pi y)+3x^2y^2+\frac{x^4}{4}\in\mathbb{R}.$$

The bijectivity assumption of $\partial_y f(1,1/2)$ is indeed satisfied:

$$\partial_y f(1,1/2)=\pi+3\neq 0.$$

The set of solutions to the equation is represented in Figure 1.1.

Proof of the implicit function theorem. Let us define

$$\begin{array}{rccc}\phi & : & U & \to & \mathbb{R}^n\times\mathbb{R}^m\\ & & (x,y) & \to & (x,f(x,y)).\end{array}$$

This is a C^k function, and for all $(h,l)\in\mathbb{R}^n\times\mathbb{R}^m$,

$$\begin{aligned}d\phi(x_0,y_0)(h,l)&=(h,df(x_0,y_0)(h,l))\\&=(h,\partial_x f(x_0,y_0)(h)+\partial_y f(x_0,y_0)(l)).\end{aligned}$$

The map $d\phi(x_0,y_0)$ is injective. Indeed, for all $(h,l)\in\mathbb{R}^n\times\mathbb{R}^m$ such that $d\phi(x_0,y_0)(h,l)=0$,

$$h=0 \text{ and } \partial_y f(x_0,y_0)(l)=0.$$

Since $\partial_y f(x_0,y_0)$ is bijective, this implies $l=0$. Thus, $d\phi(x_0,y_0)$ is an injective map from $\mathbb{R}^n\times\mathbb{R}^m$ to $\mathbb{R}^n\times\mathbb{R}^m$. Therefore, it is bijective (its domain and codomain have the same dimension).

We apply the local inversion theorem at (x_0,y_0). There exists an open neighborhood $U_{(x_0,y_0)}$ of (x_0,y_0), an open neighborhood V of $\phi(x_0,y_0)=(x_0,0)$ such that ϕ is a C^k-diffomorphism from $U_{(x_0,y_0)}$ to V. Let

$$\psi : V\to U_{(x_0,y_0)}$$

be its inverse.

For all $(x,y)\in V$, we write $\psi(x,y)=(\psi_1(x,y),\psi_2(x,y))\in\mathbb{R}^n\times\mathbb{R}^m$. For all $(x,y)\in V$,

$$\begin{aligned}(x,y)&=\phi\circ\psi(x,y)\\&=\phi(\psi_1(x,y),\psi_2(x,y))\\&=(\psi_1(x,y),f(\psi_1(x,y),\psi_2(x,y))).\end{aligned}$$

Therefore,

$$\psi_1(x, y) = x.$$

We set

$$V_{x_0} = \{x \in \mathbb{R}^n, (x, 0) \in V\};$$
$$g : x \in V_{x_0} \to \psi_2(x, 0) \in \mathbb{R}^m.$$

As required, V_{x_0} is an open neighborhood of x_0 and g is C^k. For all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$,

$$\begin{aligned}
\big((x, y) \in U_{(x_0,y_0)} \;&\text{ and } f(x, y) = 0\big) \\
&\iff \big((x, y) \in U_{(x_0,y_0)} \text{ and } \phi(x, y) = (x, 0)\big) \\
&\iff \big((x, y) \in U_{(x_0,y_0)} \text{ and } (x, 0) \in V \text{ et } (x, y) = \psi(x, 0)\big) \\
&\iff ((x, 0) \in V \text{ and } (x, y) = \psi(x, 0) = (x, \psi_2(x, 0))) \\
&\iff (x \in V_{x_0} \text{ and } y = g(x)) .
\end{aligned}$$

□

1.4 Immersions and submersions

We now introduce two particular categories of differentiable functions: *immersions* and *submersions*. These functions will have an important role in the remainder of the course because they represent two of the main ways of showing that a given set is a submanifold.

Let $n, m \in \mathbb{N}^*$ be integers. Let $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ be a C^k map (for some $k \geq 1$), with U an open set.

Definition 1.13 : immersions and submersions

For any point $x \in U$, we say that f is an *immersion at* x if $df(x) : \mathbb{R}^n \to \mathbb{R}^m$ is injective. We say that f is an *immersion* if it is an immersion at every point $x \in U$.
For any point $x \in U$, we say that f is a *submersion at* x if $df(x) : \mathbb{R}^n \to \mathbb{R}^m$ is surjective. We say that f is a *submersion* if it is a submersion at every point $x \in U$.

Remark

The function f can only be an immersion if $n \leq m$ and a submersion if $n \geq m$.

If f is an immersion at a point x, it is injective in a neighborhood of x (a consequence of Theorem 1.14). However, being an immersion is a significantly stronger property than local injectivity. Similarly, a submersion is locally surjective, but not all locally surjective functions are submersions.

When $n \leq m$, the simplest immersion from $\mathbb{R}^n$ to $\mathbb{R}^m$ is the function

$$(x_1, \dots, x_n) \in \mathbb{R}^n \quad \to \quad (x_1, \dots, x_n, 0, \dots, 0) \in \mathbb{R}^m.$$

The following theorem asserts that, in the neighborhood of every point, up to a change of coordinates in the codomain (i.e., a transformation of the codomain by a diffeomorphism), all immersions are equal to this one.

Theorem 1.14 : normal form of immersions

Suppose that $0_{\mathbb{R}^n} \in U$ and $f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m}$.
If f is an immersion at $0_{\mathbb{R}^n}$, there exists a neighborhood U' of $0_{\mathbb{R}^n}$ and a C^k-diffeomorphism ψ from a neighborhood of $0_{\mathbb{R}^m}$ to a neighborhood of $0_{\mathbb{R}^m}$ such that

$$\forall (x_1, \dots, x_n) \in U', \quad \psi \circ f(x_1, \dots, x_n) = (x_1, \dots, x_n, 0, \dots, 0).$$

Proof. Suppose that f is an immersion at $0_{\mathbb{R}^n}$.

Let $e_1, \dots, e_n$ be the vectors of the canonical basis of $\mathbb{R}^n$, and $\epsilon_1, \dots, \epsilon_m$ be those of the canonical basis of $\mathbb{R}^m$. Let us first prove the result under the assumption that

$$\forall r \in \{1, \dots, n\}, \quad df(0_{\mathbb{R}^n})(e_r) = \epsilon_r.$$

Define

$$\begin{array}{rccc} \phi & : & \mathbb{R}^m & \to & \mathbb{R}^m \\ & & (x_1, \dots, x_m) & \to & f(x_1, \dots, x_n) + (0, \dots, 0, x_{n+1}, \dots, x_m). \end{array}$$

We have $\phi(0) = 0$. Moreover, ϕ is a C^k map, and for any $h = (h_1, \dots, h_m) \in \mathbb{R}^m$,

$$\phi(0_{\mathbb{R}^m})(h) = df(0_{\mathbb{R}^n})(h_1, \dots, h_n) + (0, \dots, 0, h_{n+1}, \dots, h_m).$$

From this formula, it can be verified that $d\phi(0)(\epsilon_r) = \epsilon_r$ for all $r = 1, \dots, m$, meaning that $d\phi(0) = \mathrm{Id}_{\mathbb{R}^m}$. In particular, $d\phi(0)$ is bijective.

According to the inverse function theorem, there exist open neighborhoods V_1, V_2 of $0_{\mathbb{R}^m}$ such that ϕ is a C^k-diffeomorphism between them. Let $\psi : V_2 \to V_1$ be its inverse. For any $x = (x_1, \dots, x_n) \in U' \overset{def}{=} f^{-1}(V_2)$,

$$f(x_1, \dots, x_n) = \phi(x_1, \dots, x_n, 0, \dots, 0),$$

so

$$\psi \circ f(x_1, \dots, x_n) = (x_1, \dots, x_n, 0, \dots, 0).$$

This completes the proof of the theorem under the assumption that $df(0)(e_r) = \epsilon_r$ for all $r = 1, \dots, n$.

Now, let's drop this assumption. For any $r \in \{1, \dots, n\}$, denote $v_r = df(0_{\mathbb{R}^n})(e_r)$. As $df(0_{\mathbb{R}^n})$ is injective, the family $(v_1, \dots, v_n)$ is linearly independent; it can be completed to a basis of $\mathbb{R}^m$, denoted by $(v_1, \dots, v_m)$. Let $L \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ be such that

$$\forall r \in \{1, \dots, m\}, \quad L(v_r) = \epsilon_r.$$

It is a bijection since it sends a basis to a basis.

Let $\tilde{f} = L \circ f$. We have $\tilde{f}(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m}$ and $d\tilde{f}(0_{\mathbb{R}^n}) = L \circ df(0_{\mathbb{R}^n})$. In particular, $\tilde{f}(0_{\mathbb{R}^n})$ is an immersion at 0. For any $r \in \{1, \dots, n\}$,

$$d\tilde{f}(0_{\mathbb{R}^n})(e_r) = L(df(0_{\mathbb{R}^n})(e_r)) = L(v_r) = \epsilon_r.$$

Thus, the function $\tilde{f}$ satisfies our previous assumption. Consequently, there exist U' an open neighborhood of $0_{\mathbb{R}^n}$ and $\tilde{\psi}$ a diffeomorphism between two neighborhoods of $0_{\mathbb{R}^m}$ such that, for all $(x_1, \dots, x_n) \in U'$,

$$\tilde{\psi} \circ \tilde{f}(x_1, \dots, x_n) = (x_1, \dots, x_n, 0, \dots, 0),$$
$$\text{meaning } (\tilde{\psi} \circ L) \circ f(x_1, \dots, x_n) = (x_1, \dots, x_n, 0, \dots, 0).$$

We set $\psi = \tilde{\psi} \circ L$ to conclude. □

A similar result holds for submersions and has a similar proof. When $n \geq m$, the simplest submersion from $\mathbb{R}^n$ to $\mathbb{R}^m$ is the projection onto the first m coordinates:

$$(x_1, \dots, x_n) \in \mathbb{R}^n \quad \to \quad (x_1, \dots, x_m) \in \mathbb{R}^m.$$

Subject to a change of coordinates in the domain, all submersions are locally equal to this one.

Theorem 1.15 : normal form of submersions

Suppose that $0_{\mathbb{R}^n} \in U$ and $f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m}$.
If f is a submersion at $0_{\mathbb{R}^n}$, there exist U_1, U_2 open neighborhoods of $0_{\mathbb{R}^n}$ and a C^k diffeomorphism $\phi : U_1 \to U_2$ such that

$$\forall (x_1, \dots, x_n) \in U_1, \quad f \circ \phi(x_1, \dots, x_n) = (x_1, \dots, x_m).$$

1.5 Mean value inequality

Let's conclude this chapter with a useful inequality, the mean value inequality.

Let $(E, ||.||_E)$ and $(F, ||.||_F)$ be normed vector spaces. We equip $\mathcal{L}(E, F)$ with the uniform norm: for any $u \in \mathcal{L}(E, F)$,

$$||u||_{\mathcal{L}(E,F)} = \sup_{x\in E\setminus\{0\}} \frac{||u(x)||_F}{||x||_E}.$$

Theorem 1.16 : mean value inequality

Let $U \subset E$ be a convex open set, and $f : U \to F$ a differentiable function.
Suppose there exists $M \in \mathbb{R}^+$ such that

$$\forall x \in U, \quad ||df(x)||_{\mathcal{L}(E,F)} \leq M.$$

Then,

$$\forall x, y \in U, \quad ||f(x) - f(y)||_F \leq M||x - y||_E.$$

For the proof of this result, one can refer to [Paulin, 2009, p. 237].

Remark

Be careful not to forget the convexity assumption. The theorem may be false if it is not satisfied.
For example, the function $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ defined by $f(x) = -1$ for all $x < 0$ and $f(x) = 1$ for all $x > 0$ satisfies

$$|f'(x)| \leq 0 \quad \text{for all } x \in \mathbb{R} \setminus \{0\}$$

(as its derivative is zero).
However, it is not true that $|f(x) - f(y)| = 0$ for all $x, y \in \mathbb{R} \setminus \{0\}$.

Exercise 2 : classical application of the mean value inequality

Let $n, m \in \mathbb{N}^*$ be integers. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable function such that, for any $x \in \mathbb{R}^n$,

$$||df(x)||_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)} \leq 1.$$

Show that, for any $x \in \mathbb{R}^n$,

$$||f(x)|| \leq ||f(0)|| + ||x||.$$

Chapter 2

Submanifolds of $\mathbb{R}^n$

What you should know or be able to do after this chapter

- Have an intuition of what is a submanifold of $\mathbb{R}^n$. In particular, from a drawing of a subset of $\mathbb{R}^2$ or $\mathbb{R}^3$, be able to guess with confidence whether it represents a submanifold or not.
- Know the four definitions of a submanifold of $\mathbb{R}^n$.
- When given the explicit expression of a set, be able to prove that it is a submanifold of $\mathbb{R}^n$, choosing the most appropriate of the four definitions.
- Know the definition of $\mathbb{S}^{n-1}$.
- Be able to prove that a set is a submanifold using the fact that it is a product of submanifolds.
- Understand the proof that $O_n(\mathbb{R})$ is a submanifold (i.e. be able to do it again alone, given only the definition of $\tilde{g}$).
- Be able to use the submersion definition of submanifolds to prove that sets are not submanifolds.
- Propose a definition of the tangent space to a submanifold, then remember the "true" one.
- Given a picture of a submanifold of $\mathbb{R}^2$ or $\mathbb{R}^3$, be able to draw (a plausible version of) the tangent space at any point.
- Given the explicit expression of a submanifold, be able to compute its tangent space, choosing the most appropriate of the four formulas.
- Know the tangent space to the sphere.
- Know that the tangent space of a product submanifold is the product of the tangent spaces.
- Be able to use the tangent space to prove that sets are not submanifolds (when possible).
- Be able to show that a map between submanifolds is C^r, using the facts that compositions of C^r maps are C^r and that, on a C^k-submanifold, projections onto a coordinate are C^k.

In the whole chapter, let $k, n \in \mathbb{N}^*$ be fixed integers.

2.1 Definition

The simplest example of a submanifold of $\mathbb{R}^n$ is

$$\mathbb{R}^d \times \{0\}^{n-d} = \{(x_1, \dots, x_d, 0, \dots, 0) | x_1, \dots, x_d \in \mathbb{R}\},$$

where d is any integer between 0 and n. The concept of a *submanifold of* $\mathbb{R}^n$ generalizes this example: a set is a submanifold if it is locally the image of $\mathbb{R}^d \times \{0\}^{n-d}$ under a diffeomorphism from $\mathbb{R}^n$ to $\mathbb{R}^n$. Let's formalize this definition and provide other equivalent definitions.

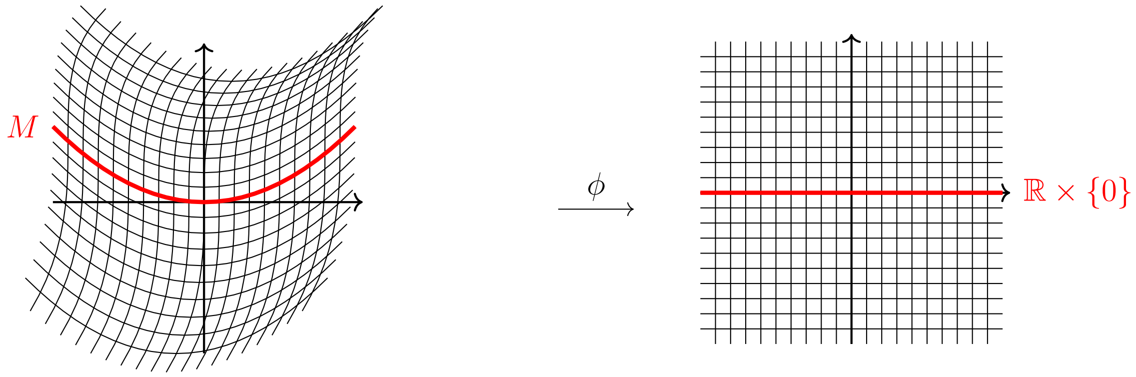

Figure 2.1: Illustration of property 1 in definition 2.1: there exists a local diffeomorphism from $\mathbb{R}^2$ to $\mathbb{R}^2$ that maps the set M onto $\mathbb{R} \times \{0\}$.

Definition 2.1 : submanifolds

Let $d \in \{0, 1 \dots, n\}$.
Let $M \subset \mathbb{R}^n$. We say that the set M is a *submanifold of* $\mathbb{R}^n$ *of dimension* d *and class* C^k if it satisfies one of the following properties.

1. (Definition by diffeomorphism)

   For every $x \in M$, there exists a neighborhood $U \subset \mathbb{R}^n$ of x, a neighborhood $V \subset \mathbb{R}^n$ of 0, and a C^k-diffeomorphism $\phi : U \to V$ such that
   $$\phi(M \cap U) = (\mathbb{R}^d \times \{0\}^{n-d}) \cap V.$$

2. (Definition by immersion)

   For every $x \in M$, there exists a neighborhood $U \subset \mathbb{R}^n$ of x, an open set V in $\mathbb{R}^d$, a C^k function $f : V \to \mathbb{R}^n$ such that f is a homeomorphism between V and $f(V)$,
   $$M \cap U = f(V)$$
   and, denoting a as the unique pre-image of x under f, f is an immersion at a.

3. (Definition by submersion)

   For every $x \in M$, there exists a neighborhood $U \subset \mathbb{R}^n$ of x, a C^k function $g : U \to \mathbb{R}^{n-d}$ that is a submersion at x such that
   $$M \cap U = g^{-1}(\{0\})$$

4. (Definition by graph)

   For every $x \in M$, there exists a neighborhood $U \subset \mathbb{R}^n$ of x, an open set V in $\mathbb{R}^d$, a C^k function $h : V \to \mathbb{R}^{n-d}$, and a coordinate system[a] in which
   $$\begin{aligned} M \cap U &= \mathrm{graph}(h) \\ &\overset{def}{=} \{(x_1, \dots, x_d, h(x_1, \dots, x_d)), (x_1, \dots, x_d) \in V\}. \end{aligned}$$

[a] A *coordinate system* is the specification of a basis $(e_1, \dots, e_n)$ for $\mathbb{R}^n$. In this system, the notation $(x_1, \dots, x_n)$ denotes the point $x_1 e_1 + \dots + x_n e_n$.

Theorem 2.2

The four properties in Definition 2.1 are equivalent.

Among the four equivalent definitions in the theorem, the definition by diffeomorphism (property 1, illustrated in figure 2.1) is the one that most clearly reveals the connection between a general submanifold and the "model" submanifold $\mathbb{R}^d \times \{0\}^{n-d}$. However, it is not the most convenient to manipulate: when proving that a given set is a submanifold, the definitions by immersion, submersion, or graph are generally more convenient, as we will see in Section 2.2.

Remark

Pay attention to the fact that, in the definition by submersion (property 3), the function g maps into $\mathbb{R}^{n-d}$ and not into $\mathbb{R}^d$.
In a very informal way, in this definition, a submanifold is defined as the set of points in $\mathbb{R}^n$ that satisfy a set of scalar equations

$$g(x)_1 = 0, g(x)_2 = 0, ...$$

Intuitively, we expect the set of solutions to have $n - e$ "degrees of freedom", where e is the number of equations. For the submanifold defined in this way to be of dimension d, we need to have $e = n - d$, meaning that g maps into $\mathbb{R}^{n-d}$.

We advise the reader to study the examples in Section 2.2 before reading the proof of Theorem 2.2.

Proof of Theorem 2.2.
$\boxed{1 \Rightarrow 3}$: Assume that M satisfies Property 1. We show that it satisfies Property 3.
Let $x \in M$. Consider U a neighborhood of x in $\mathbb{R}^n$, V a neighborhood of 0 in $\mathbb{R}^n$, and $\phi : U \to V$ a C^k-diffeomorphism such that

$$\phi(M \cap U) = (\mathbb{R}^d \times \{0\}^{n-d}) \cap V.$$

Denote $\mathrm{pr}_2 : \mathbb{R}^n \to \mathbb{R}^{n-d}$ the projection onto the last $n - d$ coordinates and define

$$g = \mathrm{pr}_2 \circ \phi : U \to \mathbb{R}^{n-d}.$$

It is a submersion at x because $dg(x)(\mathbb{R}^n) = \mathrm{pr}_2(d\phi(x)(\mathbb{R}^n)) = \mathrm{pr}_2(\mathbb{R}^n) = \mathbb{R}^{n-d}$ (recall that ϕ is a diffeomorphism, and thus, $d\phi(x)$ is bijective, meaning $d\phi(x)(\mathbb{R}^n) = \mathbb{R}^n$).
We verify that $M \cap U = g^{-1}(\{0\})$.
For every $x' \in M \cap U$, $\phi(x') \in \phi(M \cap U) = (\mathbb{R}^d \times \{0\}^{n-d}) \cap V \subset \mathbb{R}^d \times \{0\}^{n-d}$, so $\mathrm{pr}_2 \circ \phi(x') = 0$, i.e., $g(x') = 0$.
On the other hand, if $x' \in g^{-1}(\{0\})$, then $\mathrm{pr}_2(\phi(x')) = 0$, so $\phi(x') \in \mathbb{R}^d \times \{0\}^{n-d}$. Since $x' \in U$, $\phi(x') \in V$, and thus, $\phi(x') \in (\mathbb{R}^d \times \{0\}^{n-d}) \cap V = \phi(M \cap U)$, implying $x' \in M \cap U$.
$\boxed{3 \Rightarrow 4}$: Assume that M satisfies Property 3. We show that it satisfies Property 4.
Let $x \in M$. Consider U a neighborhood of x in $\mathbb{R}^n$, and $g : U \to \mathbb{R}^{n-d}$ a C^k map, submersive at x, such that

$$M \cap U = g^{-1}(\{0\}).$$

Let $(e_1, \dots, e_n)$ be an orthonormal basis of $\mathbb{R}^n$ such that

$$\mathrm{Vect}\{dg(x)(e_{d+1}), \dots, dg(x)(e_n)\} = \mathbb{R}^{n-d}. \tag{2.1}$$

(Such a basis exists because $dg(x) : \mathbb{R}^n \to \mathbb{R}^{n-d}$ is surjective.) We now use the coordinate system defined by this basis. In this system, we denote

$$x = (x_1, \dots, x_n).$$

According to Equation (2.1), the derivative of g with respect to $(x_{d+1}, \dots, x_n)$ is surjective from $\mathbb{R}^{n-d}$ to $\mathbb{R}^{n-d}$, hence bijective. Thus, by the implicit function theorem (Theorem 1.11), there exist $U' \subset U$ a neighborhood of x, V a neighborhood of $(x_1, \dots, x_d)$, and $h : V \to \mathbb{R}^{n-d}$ of class C^k such that

$$U' \cap g^{-1}(\{0\}) = \{(t, h(t)), t \in V\}.$$

Hence we have $M \cap U' = U' \cap g^{-1}(\{0\}) = \mathrm{graph}(h)$.
$\boxed{4 \Rightarrow 2}$: Let's assume that M satisfies Property 4, and show that it satisfies Property 2.

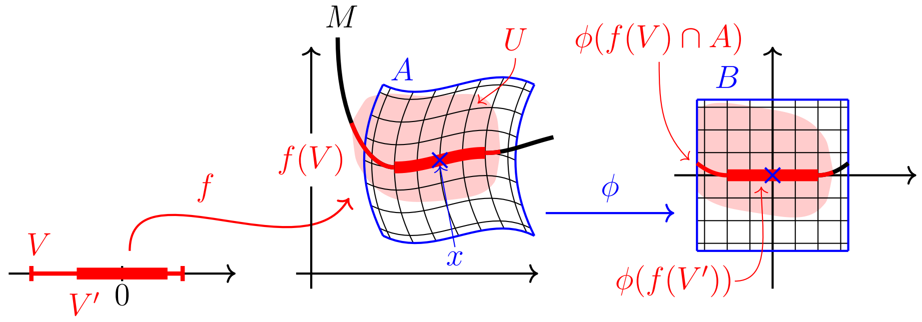

Figure 2.2: Illustration of the objects used in the proof of the implication $\boxed{2 \Rightarrow 1}$ of Theorem 2.2

Let $x \in M$. Without loss of generality, we can assume $x = 0$ to simplify notation. Let U be a neighborhood of $x = 0$ in $\mathbb{R}^n$, V an open set in $\mathbb{R}^d$, and $h : V \to \mathbb{R}^{n-d}$ be a C^k function such that, in a suitably chosen coordinate system,

$$M \cap U = \text{graph}(h) = \{(t, h(t)) \mid t \in V\}.$$

Note that $0 \in V$ and $h(0) = 0$, since $x = 0$ belongs to $M \cap U$.

Define

$$\begin{array}{rcccc} f & : & V & \to & \mathbb{R}^n \\ & & t & \to & (t, h(t)). \end{array}$$

This is a C^k map. It is an immersion at 0 because, for any $t \in \mathbb{R}^d$, $df(0)(t)$ is given by

$$(t_1, \dots, t_d, dh(0)(t)),$$

which can only be zero if $t = 0$.

We have $f(0) = 0 = x$ and f is a homeomorphism between V and $f(V)$ (its inverse is the projection onto the first d coordinates, which is continuous). Furthermore,

$$M \cap U = \text{graph}(h) = f(V).$$

$\boxed{2 \Rightarrow 1}$: Let's assume that M satisfies Property 2, and show that it satisfies Property 1.

Let $x \in M$. Let U, V be neighborhoods of x and 0 in $\mathbb{R}^n$ and $\mathbb{R}^d$ respectively, and let $f : V \to \mathbb{R}^n$ be a C^k map, which is a homeomorphism from V to $f(V)$, such that

$$M \cap U = f(V)$$

and f is immersive at a, where a is the unique preimage of x under f. Without loss of generality, we can assume, for simplicity, that $a = 0$, i.e., $f(0) = x$.

According to the normal form theorem for immersions (Theorem 1.14), there exist a neighborhood $V' \subset V$ of $0_{\mathbb{R}^d}$ and a C^k diffeomorphism $\phi : A \to B$ between a neighborhood A of x and a neighborhood B of $0_{\mathbb{R}^n}$ such that

$$\forall (t_1, \dots, t_d) \in V', \quad \phi \circ f(t_1, \dots, t_d) = (t_1, \dots, t_d, 0, \dots, 0). \tag{2.2}$$

An illustration of the various definitions in this proof is given in Figure 2.2.

Let $E \subset A \cap U$ be a neighborhood of x such that

- $f^{-1}(f(V) \cap E) \subset V'$ (such a neighborhood exists because f is a homeomorphism onto its image, so f^{-1} is well-defined and continuous on $f(V)$);
- $\phi(E) \subset V' \times \mathbb{R}^{n-d}$ (it also exists because ϕ is continuous, $V' \times \mathbb{R}^{n-d}$ is open and $\phi(x) = \phi \circ f(0) = 0 \in V' \times \mathbb{R}^{n-d}$).

Let $F = \phi(E)$.

The map ϕ is a C^k-diffeomorphism from E to F. Let's show that

$$\phi(M \cap E) = (\mathbb{R}^d \times \{0\}^{n-d}) \cap F. \tag{2.3}$$

For any $x' \in M \cap E$, we have $x' \in M \cap U = f(V)$, so $x' = f(t)$ for some $t \in V$. As $x' \in f(V) \cap E$, t is an element of V' according to the definition of E. Thus, by Equation (2.2), $\phi(x') = \phi(f(t)) \in \mathbb{R}^d \times \{0\}^{n-d}$. Moreover, $\phi(x') \in \phi(E) = F$. Therefore, $\phi(x') \in (\mathbb{R}^d \times \{0\}^{n-d}) \cap F$, which shows

$$\phi(M \cap E) \subset (\mathbb{R}^d \times \{0\}^{n-d}) \cap F.$$

Conversely, if $(t_1, \dots, t_d, 0, \dots, 0) \in (\mathbb{R}^d \times \{0\}^{n-d}) \cap F$, then $t \stackrel{\text{def}}{=} (t_1, \dots, t_d)$ is an element of V' (because $F = \phi(E) \subset V' \times \mathbb{R}^{n-d}$). Therefore, according to Equation (2.2),

$$(t_1, \dots, t_d, 0, \dots, 0) = \phi(f(t)).$$

As $f(t) \in f(V) \subset M$ and $f(t) \in \phi^{-1}(F) = E$, this shows that

$$(t_1, \dots, t_d, 0, \dots, 0) \in \phi(M \cap E).$$

Hence the inclusion $\phi(M \cap E) \supset (\mathbb{R}^d \times \{0\}^{n-d}) \cap F$, which completes the proof of Equation (2.3). □

2.2 Examples and counterexamples

As seen in the previous section, for any $d \in 0, \dots, n$,

$$\mathbb{R}^d \times \{0\}^{n-d}$$

is a submanifold of $\mathbb{R}^n$ (of class C^∞ and of dimension d).

Open sets provide another simple example of submanifolds: any non-empty open set in $\mathbb{R}^n$ is a submanifold of dimension n of $\mathbb{R}^n$.

2.2.1 Sphere

Definition 2.3

The *unit sphere in* $\mathbb{R}^n$ is the set

$$\mathbb{S}^{n-1} = \{(x_1, \dots, x_n) \in \mathbb{R}^n | x_1^2 + \dots + x_n^2 = 1\}.$$

Proposition 2.4

The set $\mathbb{S}^{n-1}$ is a submanifold of $\mathbb{R}^n$, of class C^∞, and of dimension $n-1$[a].

[a] It is precisely denoted $\mathbb{S}^{n-1}$ instead of $\mathbb{S}^n$ because its dimension is $n-1$.

Proof. We will use the definition by submersion (Property 3 of Definition 2.1).

Let $x \in \mathbb{S}^{n-1}$. Consider $g : (t_1, \dots, t_n) \in \mathbb{R}^n \to t_1^2 + \dots + t_n^2 - 1 \in \mathbb{R}$. This is a C^∞ function. It is a submersion at x. Indeed, $dg(x)$ is a linear map from $\mathbb{R}^n$ to $\mathbb{R}$, so it is either the zero map or a surjective map. Now,

$$\forall t = (t_1, \dots, t_n) \in \mathbb{R}^n, \quad dg(x)(t_1, \dots, t_n) = 2(x_1 t_1 + \dots + x_n t_n).$$

Since $x_1^2 + \dots + x_n^2 = 1$, x is not the zero vector, so $dg(x)$ is not the zero map; it is surjective.

Moreover, the definition of g implies that

$$\mathbb{S}^{n-1} = g^{-1}(\{0\}).$$

Property 3 of Definition 2.1 is therefore satisfied (with $U = \mathbb{R}^n$). □

2.2.2 Product of submanifolds

Proposition 2.5

Let $n_1, n_2 \in \mathbb{N}^*, d_1 \in \{0, \dots, n_1\}, d_2 \in \{0, \dots, n_2\}$. If M_1 is a submanifold of $\mathbb{R}^{n_1}$ of class C^k and dimension d_1, and M_2 is a submanifold of $\mathbb{R}^{n_2}$ of class C^k and dimension d_2, then

$$M_1 \times M_2 \overset{def}{=} \{(x_1, x_2), x_1 \in M_1, x_2 \in M_2\}$$

is a submanifold of $\mathbb{R}^{n_1+n_2}$ of dimension $d_1 + d_2$.

Proof. We use the definition by immersion (Property 2 of Definition 2.1). Let $x = (x_1, x_2) \in M$.

As M_1 is a submanifold, there exists a neighborhood U_1 of x_1, an open set V_1 in $\mathbb{R}^{d_1}$, and $f_1 : V_1 \to \mathbb{R}^{n_1}$ of class C^k, which is a homeomorphism onto its image, such that

$$M_1 \cap U_1 = f_1(V_1)$$

and f_1 is immersive at $f_1^{-1}(x_1)$.

Define similarly U_2, V_2, and $f_2 : V_2 \to \mathbb{R}^{n_2}$.

The function $f : (t_1, t_2) \in V_1 \times V_2 \to (f_1(t_1), f_2(t_2)) \in \mathbb{R}^{n_1+n_2}$ is of class C^k. It is a homeomorphism onto its image. Indeed, it is continuous (as each of its components is continuous, since f_1 and f_2 are continuous). It is surjective onto its image (from the definition of the image), and also injective (this can be checked from the injectivity of f_1 and f_2). Therefore, it is a bijection. Denoting f_1^{-1} and f_2^{-1} the respective inverses of f_1 and f_2), the inverse of f is

$$\begin{array}{cccc} f^{-1} & : & f(V_1 \times V_2) & \to & V_1 \times V_2 \\ & & (z_1, z_2) & \to & (f_1^{-1}(z_1), f_2^{-1}(z_2)), \end{array}$$

which is continuous because f_1^{-1} and f_2^{-1} are continuous.

Furthermore,

$$\begin{aligned} (M_1 \times M_2) \cap (U_1 \times U_2) &= (M_1 \cap U_1) \times (M_2 \cap U_2) \\ &= f_1(V_1) \times f_2(V_2) \\ &= f(V_1 \times V_2). \end{aligned}$$

Finally, f is immersive at $f^{-1}(x) = (f_1^{-1}(x_1), f_2^{-1}(x_2))$. Indeed, for any $t = (t_1, t_2) \in \mathbb{R}^{n_1+n_2}$,

$$df(f^{-1}(x_1), f^{-1}(x_2))(t_1, t_2) = (df_1(f_1^{-1}(x_1))(t_1), df_2(f_2^{-1}(x_2))(t_2)),$$

which equals 0 only if $t_1 = 0$ and $t_2 = 0$, since $df_1(f_1^{-1}(x_1))$ and $df_2(f_2^{-1}(x_2))$ are injective.

Thus, the set $M_1 \times M_2$ satisfies Property 2 of Definition 2.1. □

Example 2.6 : torus

The set $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$ is a submanifold of $\mathbb{R}^4$, of dimension 2. It is called a *torus of dimension* 2.

2.2.3 $O_n(\mathbb{R})$

Let $\mathbb{R}^{n\times n}$ denote the set of $n \times n$ matrices with real coefficients. If we reindex the coordinates, this set can also be viewed as $\mathbb{R}^{n^2}$. Several important subsets of $\mathbb{R}^{n\times n}$ have a submanifold structure. Here, we focus on the orthogonal group.

Definition 2.7 : orthogonal group

The *orthogonal group* is defined as

$$O_n(\mathbb{R}) = \{A \in \mathbb{R}^{n\times n}, I_n = {}^tAA\}.$$

Proposition 2.8

The set $O_n(\mathbb{R})$ is a submanifold of $\mathbb{R}^{n\times n}$, of class C^∞ and of dimension $\frac{n(n-1)}{2}$.

Proof. We will use the definition by submersion. Let $G \in O_n(\mathbb{R})$. We must express $O_n(\mathbb{R})$ as $g^{-1}(\{0\})$, where g is a C^∞ function, submersive at G.

A first idea is to define

$$g : A \in \mathbb{R}^{n\times n} \to {}^tAA - I_n \in \mathbb{R}^{n\times n}.$$

The definition of the orthogonal group implies that $O_n(\mathbb{R}) = g^{-1}(\{0\})$. However, this function is not a submersion at G. Indeed,

$$\forall A \in \mathbb{R}^{n\times n}, \quad dg(G)(A) = {}^tGA + {}^tAG,$$

so $dg(G)(\mathbb{R}^{n\times n})$ is contained in Sym_n, the set of symmetric matrices of size $n \times n$. We even have $dg(G)(\mathbb{R}^n) = \mathrm{Sym}_n$ because, for any $S \in \mathrm{Sym}_n$,

$$dg(G)\left(\frac{GS}{2}\right) = \frac{{}^tGGS + {}^tS{}^tGG}{2} = \frac{S + {}^tS}{2} = S.$$

In particular, $dg(G)(\mathbb{R}^{n\times n}) \neq \mathbb{R}^{n\times n}$.

Therefore, we define instead

$$\tilde{g} = \mathrm{Tri} \circ g : \mathbb{R}^{n\times n} \to \mathbb{R}^{\frac{n(n+1)}{2}},$$

where Tri is the function that extracts the upper triangular part of an $n \times n$ matrix:

$$\forall A \in \mathbb{R}^{n\times n}, \quad \mathrm{Tri}(A) = (A_{ij})_{i\leq j} \in \mathbb{R}^{\frac{n(n+1)}{2}}.$$

The function $\tilde{g}$ is C^∞. It is a submersion at G:

$$\begin{aligned} d\tilde{g}(G)(\mathbb{R}^{n\times n}) &= (\mathrm{Tri} \circ dg(G))(\mathbb{R}^{n\times n}) \\ &= \mathrm{Tri}(dg(G)(\mathbb{R}^{n\times n})) \\ &= \mathrm{Tri}(\mathrm{Sym}_n) \\ &= \mathbb{R}^{\frac{n(n+1)}{2}}. \end{aligned}$$

Furthermore, for any matrix $A \in \mathbb{R}^{n\times n}$, ${}^tAA = I_n$ if and only if ${}^tAA - I_n = 0$, which is equivalent to $\mathrm{Tri}({}^tAA - I_n) = 0$, since ${}^tAA - I_n$ is a symmetric matrix. Thus,

$$O_n(\mathbb{R}) = \tilde{g}^{-1}(\{0\}),$$

so $O_n(\mathbb{R})$ indeed satisfies Property 3, with $U = \mathbb{R}^{n\times n}$ and $d = n - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$. □

2.2.4 Equation solutions and images of maps

Proposition 2.9

Let $d \in \{0, \dots, n\}$. Let U be an open subset of $\mathbb{R}^n$, and

$$g : U \to \mathbb{R}^{n-d}$$

a C^k function. Assume that g is a submersion over $g^{-1}(\{0\})$ (meaning that g is a submersion at x for all $x \in g^{-1}(\{0\})$).
Then $g^{-1}(\{0\})$ is a submanifold of $\mathbb{R}^n$, of class C^k and dimension d.

Proof. This is a direct application of Definition 2.1, "submersion" version. □

We have already seen two examples of submanifolds defined as in Proposition 2.9:

- the sphere $\mathbb{S}^{n-1}$ is equal to $g^{-1}(\{0\})$ for the function $g : x \in \mathbb{R}^n \to ||x||^2 - 1 \in \mathbb{R}$;
- the orthogonal group $O_n(\mathbb{R})$ is equal to $g^{-1}(\{0\})$ for the function $g : A \in \mathbb{R}^{n\times n} \to \mathrm{Tri}({}^tAA - I_n) \in \mathbb{R}^{\frac{n(n+1)}{2}}$.

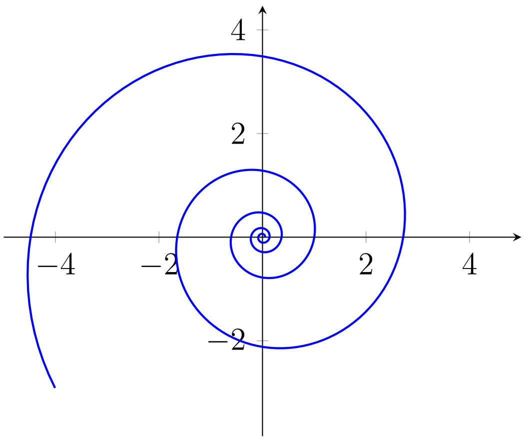

Figure 2.3: Image of the map f defined in Example 2.11

Proposition 2.10

Let $d \in \{0, \dots, n\}$. Let U be an open subset of $\mathbb{R}^d$, and $f : U \to \mathbb{R}^n$ be C^k. Assume that f is an immersion, and is a homeomorphism from U to $f(U)$.
Then $f(U)$ is a submanifold of $\mathbb{R}^n$, of class C^k and dimension d.

Proof. This is a direct application of Definition 2.1, "immersion" version. □

Example 2.11 : spiral

Let's define

$$\begin{array}{rcccl} f & : & \mathbb{R} & \to & \mathbb{R}^2 \\ & & \theta & \to & \left(e^\theta \cos(2\pi\theta), e^\theta \sin(2\pi\theta)\right). \end{array}$$

Its image $f(\mathbb{R})$ is a submanifold. It is represented in Figure 2.3.
Indeed, for any $\theta \in \mathbb{R}$,

$$f'(\theta) = e^\theta \left((\cos(2\pi\theta), \sin(2\pi\theta)) + 2\pi\left(-\sin(2\pi\theta), \cos(2\pi\theta)\right)\right),$$

which never vanishes (we observe, for example, that $\langle f'(\theta), (\cos(2\pi\theta), \sin(2\pi\theta))\rangle = e^\theta \neq 0$ for any $\theta \in \mathbb{R}$). Thus, the map f is an immersion. Moreover, it is a homeomorphism from $\mathbb{R}$ to $f(\mathbb{R})$. Indeed, it is continuous, injective[a] and therefore bijective onto $f(\mathbb{R})$. For any $\theta \in \mathbb{R}$,

$$e^{2\theta} = ||f(\theta)||^2,$$

so $\theta = \frac{1}{2}\log\left(||f(\theta)||^2\right)$. As a consequence, the inverse of f is given by the following explicit expression:

$$\begin{array}{rcccl} f^{-1} & : & f(\mathbb{R}) & \to & \mathbb{R} \\ & & (x, y) & \to & \frac{1}{2}\log(x^2 + y^2). \end{array}$$

From this expression, we see that f^{-1} is the restriction to $f(\mathbb{R})$ of a continuous function on $\mathbb{R}^2 \setminus (0, 0)$, so f^{-1} is continuous.

[a] For any θ_1, θ_2, if $f(\theta_1) = f(\theta_2)$, then $e^{2\theta_1} = ||f(\theta_1)||^2 = ||f(\theta_2)||^2 = e^{2\theta_2}$, so $\theta_1 = \theta_2$.

2.2.5 Submanifolds of dimension 0 and n

Proposition 2.12

Let M be any subset of $\mathbb{R}^n$. The following properties are equivalent:

1. M is a C^k-submanifold of $\mathbb{R}^n$ with dimension n ;
2. M is an open subset of $\mathbb{R}^n$.

Proof. $\boxed{1 \Rightarrow 2}$: We assume that M is a C^k-submanifold with dimension n, and show that it is an open set.

Let x be any point of M. We use the "diffeomorphism" definition of submanifolds: let $U \subset \mathbb{R}^n$ be a neighborhood of x, $V \subset \mathbb{R}^n$ a neighborhood of 0, and $\phi : U \to V$ a C^k-diffeomorphism such that

$$\phi(M \cap U) = (\mathbb{R}^n \times \{0\}^{n-n}) \cap V = V.$$

Since ϕ is a bijection from U to V, this equality implies that $M \cap U = U$. Therefore, M contains U, a neighborhood of x. Since this property is true at any point x, M is an open set.

$\boxed{2 \Rightarrow 1}$: We assume that M is an open set, and show that it is a submanifold with dimension n.

Let x be a point in M. We show that M satisfies the "diffeomorphism" definition of submanifolds. We set $U = B(x, r)$, for $r > 0$ small enough so that $U \subset M$. We also set $V = B(0, r)$ and $\phi : y \in U \to y - x \in V$. This map is a diffeomorphism (with reciprocal $(y \in V \to y + x \in U)$). It holds

$$\phi(M \cap U) = \phi(U) = V = (\mathbb{R}^n \times \{0\}^{n-n}) \cap V.$$

□

Proposition 2.13

Let M be any subset of $\mathbb{R}^n$. The following properties are equivalent:

1. M is a C^k-submanifold of $\mathbb{R}^n$ with dimension 0 ;
2. M is a discrete set.[a]

[a] The set M is *discrete* if, for any $x \in M$, there exists $U \subset \mathbb{R}^n$ a neighborhood of x such that $M \cap U = \{x\}$.

Proof. $\boxed{1 \Rightarrow 2}$: We assume that M is a C^k-submanifold with dimension 0, and show that it is a discrete set.

Let x be any point of M. Let us show that there exists U a neighborhood of x such that $M \cap U = \{x\}$.

We use the "diffeomorphism" definition of submanifolds: let $U \subset \mathbb{R}^n$ be a neighborhood of x, $V \subset \mathbb{R}^n$ a neighborhood of $(0, \dots, 0)$ and $\phi : U \to V$ a C^k-diffeomorphism such that

$$\phi(M \cap U) = (\mathbb{R}^0 \times \{0\}^n) \cap V = \{(0, \dots, 0)\}.$$

As ϕ is injective and $\phi(M \cap U)$ contains only one point, $M \cap U$ itself must be a singleton. Since it contains x, $M \cap U = \{x\}$.

$\boxed{2 \Rightarrow 1}$: We assume that M is a discrete set, and show that it is a submanifold of $\mathbb{R}^n$, of dimension 0.

Let x be any point in M. We show that M satisfies the "diffeomorphism" definition of submanifolds in the neighborhood of x.

Let $U \subset \mathbb{R}^n$ be a neighborhood of x such that $M \cap U = \{x\}$. Let us set $V = \{u - x, u \in U\}$ (the translation of U by $-x$) and $\phi : y \in U \to y - x \in V$. This is a C^∞-diffeomorphism (with reciprocal $(y \in V \to y + x \in U)$). It holds

$$\phi(M \cap U) = \phi(\{x\}) = \{\phi(x)\} = \{(0, \dots, 0)\} = (\mathbb{R}^0 \times \{0\}^n) \cap V.$$

□

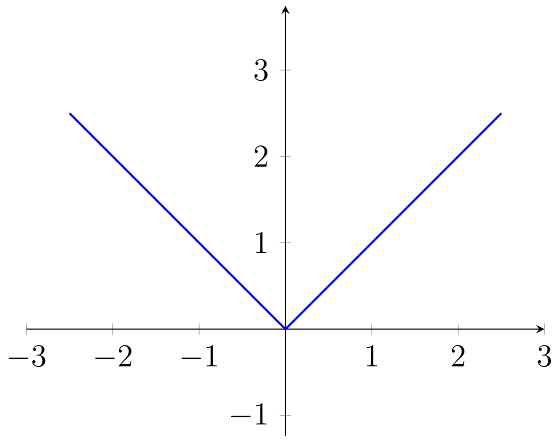

Figure 2.4: The graph of the absolute value is not a submanifold of $\mathbb{R}^2$.

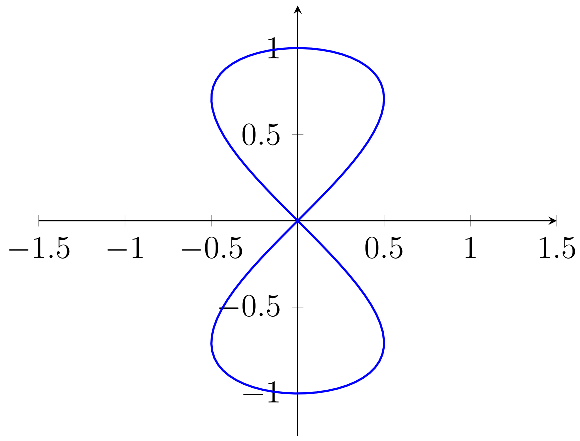

Figure 2.5: The "eight" is not a submanifold of $\mathbb{R}^2$.

2.2.6 Two counterexamples

The graph of the absolute value (Figure 2.4) is not a submanifold of $\mathbb{R}^2$. Intuitively, the reason is that this graph has a “non-regular” point at $(0,0)$.

To prove this rigorously, the simplest way is to proceed by contradiction. Assume that it is a submanifold and denote its dimension by d. Then, according to the "submersion" definition of submanifolds (Property 3 of Definition 2.1), there exists $U \subset \mathbb{R}^2$ a neighborhood of $(0,0)$ and $g : U \to \mathbb{R}^{2-d}$ a function, at least C^1, submersive at $(0,0)$, such that

$$\{(t,|t|), t \in \mathbb{R}\} \cap U = g^{-1}(\{0\}). \tag{2.4}$$

Such a map g must satisfy, for all t close enough to 0,

$$\begin{aligned} \text{if } t \leq 0, \quad & 0 = g(t,|t|) = g(t,-t), \\ \text{if } t \geq 0, \quad & 0 = g(t,|t|) = g(t,t). \end{aligned}$$

Differentiating these two equalities, we get:

$$\begin{aligned} \partial_1 g(0,0) - \partial_2 g(0,0) = 0; \\ \partial_1 g(0,0) + \partial_2 g(0,0) = 0. \end{aligned}$$

This implies that $\partial_1 g(0,0) = \partial_2 g(0,0) = 0$, i.e., $dg(0,0) = 0$. As $dg(0,0)$ is surjective, this is impossible, unless $\mathbb{R}^{2-d} = \{0\}$, i.e., $d = 2$. But if $d = 2$, then $g^{-1}(\{0\}) = U$, so Equality (2.4) implies that the graph of the absolute value contains a neighborhood of $(0,0)$ in $\mathbb{R}^2$, which is not true. Thus, we reach a contradiction.

The "eight" (Figure 2.5) is also not a submanifold of $\mathbb{R}^2$. Here, the reason is that the eight is a regular curve but with a point of "self-intersection" at zero. This can be rigorously demonstrated using the same method as before.

Remark

This example highlights the importance of the property "f is a homeomorphism onto its image" in the "immersion" definition of submanifolds (Property 2 of Definition 2.1), as well as in Proposition 2.10. Indeed, the eight is equal to $f(]-\pi;\pi[)$, where f is the map

$$\begin{array}{cccc} f & : &]-\pi;\pi[& \to & \mathbb{R}^2 \\ & & \theta & \to & (\sin(\theta)\cos(\theta), \sin(\theta)), \end{array}$$

which is an immersion, and a bijection between $]-\pi;\pi[$ and $f(]-\pi;\pi[)$, but not a homeomorphism (its inverse is not continuous).

2.3 Tangent spaces

2.3.1 Definition

Intuitively, the tangent space to a submanifold M at a point x is the set of directions an ant could take while moving on the surface of M starting from the point x. More formally, the definition is as follows.

Definition 2.14 : tangent space

Let M be a submanifold of $\mathbb{R}^n$, and x a point on M.
The tangent space to M at x, denoted T_xM, is the set of vectors $v \in \mathbb{R}^n$ such that there exists an open interval I containing 0 and $c : I \to \mathbb{R}^n$ a C^1 function satisfying

- $c(t) \in M$ for all $t \in I$;
- $c(0) = x$;
- $c'(0) = v$.

Proposition 2.15

Keeping the notation from the previous definition, the set T_xM is a vector subspace of $\mathbb{R}^n$, with the same dimension as M.

Proof. This is a consequence of the following theorem. □

The four equivalent definitions of submanifolds (Definition 2.1) each provide a way to explicitly compute the tangent space.

Theorem 2.16 : computing the tangent space

Let M be a submanifold of $\mathbb{R}^n$, and x a point on M. Let d be the dimension of M.

1. (Computation by diffeomorphism)

   If U and V are neighborhoods of x and 0 in $\mathbb{R}^n$, respectively, and $\phi : U \to V$ is a C^k-diffeomorphism such that $\phi(x) = 0$ and $\phi(M \cap U) = (\mathbb{R}^d \times \{0\}^{n-d}) \cap V$, then

   $$T_xM = d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d}).$$

2. (Computation by immersion)

   If U is a neighborhood of x in $\mathbb{R}^n$, V an open set in $\mathbb{R}^d$, and $f : V \to \mathbb{R}^n$ a C^k map, which is a homeomorphism between V and $f(V)$, such that $M \cap U = f(V)$ and f is an immersion at $z_0 \overset{def}{=} f^{-1}(x)$, then

   $$T_xM = df(z_0)(\mathbb{R}^d)(= \mathrm{Im}(df(z_0)))$$

3. (Computation by submersion)

   If U is a neighborhood of x and $g : U \to \mathbb{R}^{n-d}$ a C^k map surjective at x such that $M \cap U = g^{-1}(\{0\})$, then
$$T_x M = \mathrm{Ker}(dg(x)).$$

4. (Computation by graph)

   If U is a neighborhood of x, V an open set in $\mathbb{R}^d$, and $h : V \to \mathbb{R}^{n-d}$ is a C^k map such that, in a well-chosen coordinate system, $M \cap U = \mathrm{graph}(h)$, then
$$T_x M = \{(t_1, \dots, t_d, dh(x_1, \dots, x_d)(t_1, \dots, t_d)), t_1, \dots, t_d \in \mathbb{R}\}.$$

Proof. Let's begin with Property 1. Let U, V, and ϕ be as stated in the property.

First, let's prove the inclusion $T_x M \subset d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d})$. Let v be an arbitrary element in $T_x M$; we will show that it belongs to $d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d})$.

Let c be as in the definition of the tangent space, i.e. a C^1 map from an open interval I containing 0 to $\mathbb{R}^n$, with images in M, such that $c(0) = x$ and $c'(0) = v$.

For any t close enough to 0, $c(t)$ belongs to U, so $\phi(c(t))$ is well-defined. Moreover, since $\phi(M \cap U) \subset \mathbb{R}^d \times \{0\}^{n-d}$, we must have
$$0 = \phi(c(t))_{d+1} = \dots = \phi(c(t))_n.$$
Differentiating these equalities at $t = 0$ gives:
$$\begin{aligned} 0 &= d\phi(c(0))(c'(0))_{d+1} = d\phi(x)(v)_{d+1}, \\ &\dots \\ 0 &= d\phi(x)(v)_n. \end{aligned}$$
Therefore, $d\phi(x)(v) \in \mathbb{R}^d \times \{0\}^{n-d}$, i.e., $v \in d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d})$.

Now, let's prove the other inclusion: $d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d}) \subset T_x M$. Let $v \in d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d})$; we will show that $v \in T_x M$.

Denote
$$w = d\phi(x)(v) \in \mathbb{R}^d \times \{0\}^{n-d}.$$
We must find a function c as in the definition of the tangent space. We will define it as the preimage by ϕ of a function γ with images in $\mathbb{R}^n$ such that $\gamma(0) = 0$ and $\gamma'(0) = w$.

Choose an open interval I containing 0 small enough, and define
$$\begin{array}{cccc} \gamma & : & I & \to & \mathbb{R}^n \\ & & t & \to & tw. \end{array}$$
This is a C^∞ function satisfying
$$\gamma(0) = 0 \quad \text{and} \quad \gamma'(0) = w.$$
If I is small enough, $\gamma(I) \subset V$. Thus, we can define
$$c = \phi^{-1} \circ \gamma : I \to \mathbb{R}^n.$$
This is a C^k function. It takes values in M because $\gamma(t) \in \mathbb{R}^d \times \{0\}^{n-d}$ for all $t \in I$ (since $w \in \mathbb{R}^d \times \{0\}^{n-d}$). Therefore,
$$c(t) \in \phi^{-1}\left(\left(\mathbb{R}^d \times \{0\}^{n-d}\right) \cap V\right) = M \cap U.$$
Moreover,
$$c(0) = \phi^{-1}(\gamma(0)) = \phi^{-1}(0) = x$$
and
$$w = \gamma'(0)$$

$$
\begin{aligned}
&= (\phi \circ c)'(0) \\
&= d\phi(c(0))(c'(0)) \\
&= d\phi(x)(c'(0)).
\end{aligned}
$$

Therefore,

$$c'(0) = d\phi(x)^{-1}(w) = v.$$

the map c satisfies the properties required in the definition of the tangent space. Therefore,

$$v \in T_x M.$$

This completes the proof of the equality

$$T_x M = d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d}).$$

Before proving the remaining three properties of the theorem, let's observe that the equality we have just obtained already shows that $T_x M$ is a vector subspace of $\mathbb{R}^n$ of dimension d. Indeed, it is the image of a vector subspace of dimension d of $\mathbb{R}^n$ ($\mathbb{R}^d \times \{0\}^{n-d}$) under a linear isomorphism ($d\phi(x)^{-1}$).

This observation simplifies the proof of properties 2, 3, and 4. Indeed, the sets

$$
\begin{gathered}
df(z_0)(\mathbb{R}^d), \operatorname{Ker}(dg(x)) \\
\text{and } \{(t_1, \ldots, t_d, dh(x_1, \ldots, x_d)(t_1, \ldots, t_d)), t_1, \ldots, t_d \in \mathbb{R}\},
\end{gathered}
$$

which appear in these properties, are vector subspaces of $\mathbb{R}^n$ of dimension d (the first is the image of $\mathbb{R}^d$ by an injective linear map, the second is the kernel of a surjective linear map from $\mathbb{R}^n$ to $\mathbb{R}^{n-d}$, and the third is generated by the following free family of d elements:

$$
\begin{gathered}
(1, 0, \ldots, 0, dh(x_1, \ldots, x_d)(1, 0, \ldots, 0)), \\
\ldots, \\
(0, \ldots, 0, 1, dh(x_1, \ldots, x_d)(0, \ldots, 0, 1))).
\end{gathered}
$$

To show that they are equal to $T_x M$, it is therefore sufficient to prove either

- that they contain $T_x M$,
- or that they are included in $T_x M$.

Let's prove Property 2. Let U, V, and f be as in the statement of the property. We will show that

$$df(z_0)(\mathbb{R}^d) \subset T_x M. \tag{2.5}$$

Let $v \in df(z_0)(\mathbb{R}^d)$ be arbitrary; let's show that $v \in T_x M$. Let $a \in \mathbb{R}^d$ be such that $df(z_0)(a) = v$. Choose an interval $I \subset \mathbb{R}$ containing 0, small enough, and define

$$
\begin{array}{rcrcl}
c & : & I & \rightarrow & \mathbb{R}^n \\
 & & t & \rightarrow & f(z_0 + ta).
\end{array}
$$

the map c is well-defined if I is small enough, as $z_0 + ta \in V$ for all $t \in I$. It is a C^k (thus C^1) function. For all $t \in I$, $c(t) \in f(V) \subset M$. Moreover,

$$c(0) = f(z_0) = x$$

and

$$c'(0) = df(z_0)(a) = v.$$

This shows that $v \in T_x M$. Thus, Equation (2.5) is true.

Now let's prove Property 3. Let U and g be as in the statement of the property. We will show that

$$T_x M \subset \operatorname{Ker}(dg(x)).$$

Let $v \in T_xM$ be arbitrary. Let us show that v is in $\text{Ker}(dg(x))$. Let I be an interval in $\mathbb{R}$ containing 0, and $c : I \to \mathbb{R}^n$ as in the definition of the tangent space.

For any t close enough to 0, $c(t)$ is an element of U; it is also an element of M. Since $M \cap U = g^{-1}(\{0\})$,

$$0 = g(c(t)).$$

Differentiating this equality at 0,

$$0 = dg(c(0))(c'(0)) = dg(x)(v).$$

Therefore, $v \in \text{Ker}(dg(x))$.

Finally, let's prove Property 4. Let U, V, and h be as in the statement of this property. Let

$$E = \{(t_1, \dots, t_d, dh(x_1, \dots, x_d)(t_1, \dots, t_d)), t_1, \dots, t_d \in \mathbb{R}\}$$

We show that

$$E \subset T_xM.$$

Let $(t, dh(x_1, \dots, x_d)(t)) \in E$, with $t \in \mathbb{R}^d$. Let us show that this is an element of T_xM.

Choose an interval I in $\mathbb{R}$ containing 0 small enough, and define

$$\begin{array}{cccl} c & : & I & \to & \mathbb{R}^n \\ & & s & \to & ((x_1, \dots, x_d) + st, h((x_1, \dots, x_d) + st)). \end{array}$$

This function is well-defined if I is small enough, as $(x_1, \dots, x_d) + st$ belongs to V for all $s \in I$ (since V contains $(x_1, \dots, x_d)$ and is open). It is of class C^k (thus C^1). It is in the graph of h, and therefore in M. Moreover,

$$c(0) = (x_1, \dots, x_d, h(x_1, \dots, x_d)) = x$$

and

$$c'(0) = (t, dh(x_1, \dots, x_d)(t)).$$

This shows that $(t, dh(x_1, \dots, x_d)(t)) \in T_xM$.

□

To finish with the definitions, let's introduce the affine tangent space, which is simply the tangent space, translated so that it goes through the point x. This is not a notion that we will particularly use in the rest of the course, except in the figures: it is much more natural to draw tangent spaces that really touch[1] the submanifold they are associated with than tangent spaces which all contain 0.

Definition 2.17

If M is a submanifold of $\mathbb{R}^n$ and $x \in M$, the *affine tangent space to M at x* is the set

$$x + T_xM.$$

2.3.2 Examples

In this paragraph, we go back to the examples of submanifolds from Section 2.2 and compute their tangent spaces.

Proposition 2.18 : tangent space of the sphere

For any $x \in \mathbb{S}^{n-1}$,

$$T_x\mathbb{S}^{n-1} = \{x\}^\perp = \{t \in \mathbb{R}^n, \langle t, x \rangle = 0\}.$$

[1] The word "tangent" comes from the Latin verb *tangere*, which means "to touch".

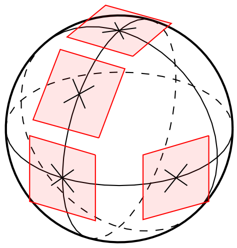

Figure 2.6: The sphere $\mathbb{S}^2$ and its affine tangent space at a few points.

Proof. Let's define, as in Subsection 2.2.1,

$$g \quad : \quad \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R} \\ (t_1, \dots, t_n) & \to & t_1^2 + \dots + t_n^2 - 1. \end{array}$$

It satisfies $\mathbb{S}^{n-1} = g^{-1}(\{0\})$ and is a submersion at x. According to Property 3 of Theorem 2.16,

$$T_x\mathbb{S}^{n-1} = \mathrm{Ker}(dg(x)).$$

Now, for any $t \in \mathbb{R}^n$, $dg(x)(t) = 2\langle x, t\rangle$. Therefore,

$$T_x\mathbb{S}^{n-1} = \{x\}^\perp.$$

□

Proposition 2.19 : tangent space of a product submanifold

Let $n_1, n_2 \in \mathbb{N}^*$. Assume M_1 is a submanifold of $\mathbb{R}^{n_1}$ and M_2 is a submanifold of $\mathbb{R}^{n_2}$. For any $x = (x_1, x_2) \in M_1 \times M_2$,

$$\begin{aligned} T_x(M_1 \times M_2) &= T_{x_1}M_1 \times T_{x_2}M_2 \\ &= \{(t_1, t_2), t_1 \in T_{x_1}M_1, t_2 \in T_{x_2}M_2\}. \end{aligned}$$

Proof. Let $x = (x_1, x_2) \in M_1 \times M_2$.

We will use the expression for the tangent space associated with the "immersion" definition of submanifolds (Property 2 of Theorem 2.16).

Let d_1 be the dimension of M_1. Assume U_1 is a neighborhood of x_1 in $\mathbb{R}^{n_1}$, V_1 a neighborhood of 0 in $\mathbb{R}^{d_1}$, and $f_1 : V_1 \to \mathbb{R}^{n_1}$ a map which is a homeomorphism onto its image, such that

$$M_1 \cap U_1 = f_1(V_1)$$

and f_1 is immersive at $z_1 = f^{-1}(x_1)$.

Define similarly $d_2, U_2, V_2, f_2 : V_2 \to \mathbb{R}^{n_2}$ and z_2.

According to Property 2 of Theorem 2.16, we have

$$T_{x_1}M_1 = df_1(z_1)(\mathbb{R}^{d_1}) \quad \text{and} \quad T_{x_2}M_2 = df_2(z_2)(\mathbb{R}^{d_2}).$$

Moreover, as shown in the proof of Proposition 2.5, the map $f : (t_1, t_2) \in V_1 \times V_2 \to (f_1(t_1), f_2(t_2)) \in \mathbb{R}^{n_1+n_2}$ is a homeomorphism onto its image, satisfies

$$f(V_1 \times V_2) = (M_1 \times M_2) \cap (U_1 \times U_2)$$

and is immersive at $(z_1, z_2) = f^{-1}(x)$. From Property 2 of Theorem 2.16, we have

$$\begin{aligned} T_x(M_1 \times M_2) &= df(z_1, z_2)(\mathbb{R}^{d_1+d_2}) \\ &= \{df(z_1, z_2)(t_1, t_2), t_1 \in \mathbb{R}^{d_1}, t_2 \in \mathbb{R}^{d_2}\} \\ &= \{(df_1(z_1)(t_1), df_2(z_2)(t_2)), t_1 \in \mathbb{R}^{d_1}, t_2 \in \mathbb{R}^{d_2}\} \\ &= df_1(z_1)(\mathbb{R}^{d_1}) \times df_2(z_2)(\mathbb{R}^{d_2}) \\ &= T_{x_1}M_1 \times T_{x_2}M_2. \end{aligned}$$

□

Example 2.20 : tangent space of the torus

For any $(x_1, x_2) \in \mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$,

$$T_{(x_1,x_2)}\mathbb{T}^2 = T_{x_1}\mathbb{S}^1 \times T_{x_2}\mathbb{S}^1 = \{x_1\}^\perp \times \{x_2\}^\perp.$$

If we fix θ_1, θ_2 such that $x_1 = (\cos(\theta_1), \sin(\theta_1)), x_2 = (\cos(\theta_2), \sin(\theta_2))$, we have

$$\begin{aligned}\{x_1\}^\perp &= (\sin(\theta_1), -\cos(\theta_1))\mathbb{R} \\ &= \{(t_1 \sin(\theta_1), -t_1 \cos(\theta_1)), t_1 \in \mathbb{R}\}\end{aligned}$$

and similarly for x_2. This allows us to write the previous expression for the tangent to the torus in a slightly more explicit way:

$$T_{(x_1,x_2)}\mathbb{T}^2 = \{(t_1 \sin(\theta_1), -t_1 \cos(\theta_1), t_2 \sin(\theta_2), -t_2 \cos(\theta_2)), t_1, t_2 \in \mathbb{R}\}.$$

Proposition 2.21 : tangent space of the orthogonal group

For any $G \in O_n(\mathbb{R})$,

$$T_G O_n(\mathbb{R}) = \{GR, R \in \mathbb{R}^{n\times n} \text{ is antisymmetric}\}.$$

Proof. Let $G \in O_n(\mathbb{R})$.

As shown in the proof of Proposition 2.8, $O_n(\mathbb{R})$ is equal to $\tilde{g}^{-1}(\{0\})$, where $\tilde{g}$ is defined as

$$\begin{array}{cccc}\tilde{g} & : & \mathbb{R}^{n\times n} & \to & \mathbb{R}^{\frac{n(n+1)}{2}} \\ & & A & \to & \mathrm{Tri}({}^tAA - I_n).\end{array}$$

The map $\tilde{g}$ is a submersion at G, with differential

$$d\tilde{g}(G) : A \in \mathbb{R}^{n\times n} \to \mathrm{Tri}({}^tGA + {}^tAG) \in \mathbb{R}^{\frac{n(n+1)}{2}}.$$

According to Property 3 of Theorem 2.16,

$$T_G O_n(\mathbb{R}) = \mathrm{Ker}(d\tilde{g}(G)) = \left\{A \in \mathbb{R}^{n\times n}, \mathrm{Tri}({}^tGA + {}^tAG) = 0\right\}.$$

Now, for any A,

$$\begin{aligned}\mathrm{Tri}({}^tGA + {}^tAG) = 0 &\iff {}^tGA + {}^tAG = 0 \\ &\qquad (\text{because } {}^tGA + {}^tAG \text{ is symmetric}) \\ &\iff ({}^tGA) + {}^t({}^tGA) = 0 \\ &\iff {}^tGA = R \text{ for some antisymmetric } R \\ &\iff A = GR \text{ for some antisymmetric } R \\ &\qquad (\text{because } G{}^tG = I_n).\end{aligned}$$

Therefore,

$$T_G O_n(\mathbb{R}) = \{GR, R \in \mathbb{R}^{n\times n} \text{ is antisymmetric}\}.$$

□

Proposition 2.22

Let $d \in \{0, \ldots, n\}$. Let U be an open set in $\mathbb{R}^n$, and $g : U \to \mathbb{R}^{n-d}$ be a C^k function. Assume that g is a submersion on $g^{-1}(\{0\})$.

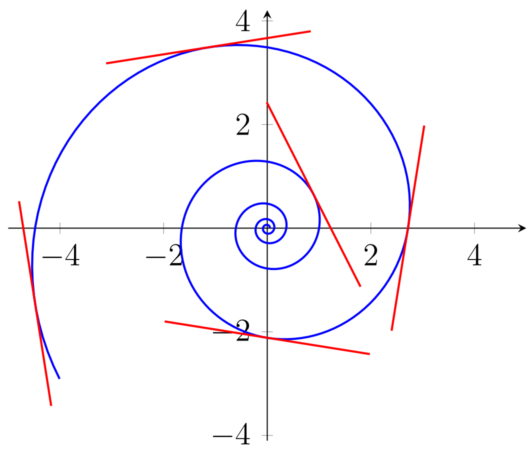

Figure 2.7: The spiral from Example 2.24 and its affine tangent space at a few points.

For any $x \in g^{-1}(\{0\})$,

$$T_x(g^{-1}(\{0\})) = \text{Ker}(dg(x)).$$

Proof. This is a direct application of Property 3 of Theorem 2.16. □

Proposition 2.23

Let $d \in \{0, \dots, n\}$. Let U be an open set in $\mathbb{R}^d$, and $f : U \to \mathbb{R}^n$ be an immersion, which is a homeomorphism from U to $f(U)$.
For any $x \in f(U)$,

$$T_x f(U) = df(z)(\mathbb{R}^d),$$

where z is the element of U such that $x = f(z)$.

Proof. This is a direct application of Property 2 of Theorem 2.16. □

Example 2.24 : tangent space of the spiral

Consider the map from Example 2.11:

$$\begin{array}{rcccc} f & : & \mathbb{R} & \to & \mathbb{R}^2 \\ & & \theta & \to & \left(e^\theta \cos(2\pi\theta), e^\theta \sin(2\pi\theta)\right). \end{array}$$

Let $(x, y) \in f(\mathbb{R})$. Denote $\theta \in \mathbb{R}$ the real number such that $(x, y) = f(\theta)$. According to Proposition 2.23:

$$\begin{aligned} T_{(x,y)} f(\mathbb{R}) &= f'(\theta)\mathbb{R} \\ &= e^\theta((\cos(2\pi\theta), \sin(2\pi\theta)) + 2\pi(-\sin(2\pi\theta), \cos(2\pi\theta)))\mathbb{R} \\ &= (x - 2\pi y, y + 2\pi x)\mathbb{R} \\ &= \{((x - 2\pi y)t, (y + 2\pi x)t), t \in \mathbb{R}\}. \end{aligned}$$

An illustration is shown on Figure 2.7.

2.3.3 Application: proof that a set is not a submanifold

Let us go back to the second set considered in Subsection 2.2.6, the "eight", represented on Figure 2.5. This set is

$$M \overset{def}{=} \{f(\theta), \theta \in]-\pi; \pi[\}.$$

where f is defined as

$$\begin{array}{rccc} f & : &]-\pi;\pi[& \to & \mathbb{R}^2 \\ & & \theta & \to & (\sin(\theta)\cos(\theta),\sin(\theta)). \end{array}$$

Here, we prove that M is not a submanifold of $\mathbb{R}^2$ using a different technique from Subsection 2.2.6.

By contradiction, let us assume that it is a submanifold. We compute its tangent space at $(0,0)$.

First, we define

$$c_1 = f :]-\pi;\pi[\to \mathbb{R}^2.$$

It holds $c_1(t) \in M$ for all $t \in]-\pi;\pi[$, $c_1(0) = (0,0)$ and c_1 is C^1. Therefore,

$$(1,1) = c_1'(0) \in T_{(0,0)}M. \tag{2.6}$$

Second, we define

$$\begin{array}{rccc} c_2 & : &]-\pi;\pi[& \to & \mathbb{R}^2 \\ & & \theta & \to & (\sin(\theta)\cos(\theta),-\sin(\theta)). \end{array}$$

It holds $c_2(t) \in M$ for all $t \in]-\pi;\pi[$. Indeed, for any $t \in]-\pi;0[$, $c_2(t) = f(t+\pi) \in M$; $c_2(0) = f(0) \in M$ and, for any $t \in]0;\pi[$, $c_2(t) = f(t-\pi) \in M$. In addition, $c_2(0) = (0,0)$ and c_2 is C^1. Therefore,

$$(1,-1) = c_2'(0) \in T_{(0,0)}M. \tag{2.7}$$

As $T_{(0,0)}M$ is a vector subspace of $\mathbb{R}^2$, Equations (2.6) and (2.7) together imply that

$$T_{(0,0)}M = \mathbb{R}^2.$$

In particular, since the dimension of the tangent space is the same as the dimension of the submanifold, $\dim M = 2$. In virtue of Proposition 2.12, M must thus be an open set of $\mathbb{R}^2$. As this is not true (because, for instance, M contains no element of the form $(t,0)$, except $(0,0)$ itself, so it does not contain a neighborhood of $(0,0)$), we have reached a contradiction.

2.4 Maps between submanifolds

2.4.1 Definition of C^1 maps

In this section, we consider functions between two submanifolds $M \subset \mathbb{R}^{n_1}$ and $N \subset \mathbb{R}^{n_2}$:

$$f : M \to N.$$

If $M = \mathbb{R}^{d_1} \times \{0\}^{n_1-d_1}$ and $N = \mathbb{R}^{d_2} \times \{0\}^{n_2-d_2}$, f is essentially a function from $\mathbb{R}^{d_1}$ to $\mathbb{R}^{d_2}$. The notions of "differentiability" and "differential" are then well-defined for f, in accordance with Chapter 1.

However, if M is not a vector subspace of $\mathbb{R}^{n_1}$, this is no longer the case: Definition 1.1 involves linear maps between the domain and codomain, which do not exist if the sets are not vector spaces.

To give a meaning to the notion of "differentiability" for f, one can use the fact that M and N are identifiable with open sets in $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$ through diffeomorphisms. We say that f is differentiable if, when composed with these diffeomorphisms, it is a differentiable map from an open set in $\mathbb{R}^{d_1}$ to $\mathbb{R}^{d_2}$. This is, in a slightly different form, the content of the following definition.

Definition 2.25 : C^1 map from a submanifold to $\mathbb{R}^m$

Let $m \in \mathbb{N}$.
Consider M a C^k submanifold of $\mathbb{R}^n$, and a function

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

We say that f is *of class* C^1 if, for any integer $s \in \mathbb{N}^*$, any open set V in $\mathbb{R}^s$, and any C^1 function $\phi : V \to \mathbb{R}^n$ such that $\phi(V) \subset M$, the map

$$f \circ \phi : V \to \mathbb{R}^m$$

is of class C^1.

Remark

Similarly, one can define the notion of *function of class* C^r from M to $\mathbb{R}^m$, for any $r = 1, \dots, k$. Simply replace "C^1" with "C^r" in the above definition.
It can be shown that a function of class C^r is necessarily of class $C^{r'}$ for any $r' \leq r$.

Example 2.26 : projection onto a coordinate

Let $M \subset \mathbb{R}^n$ be a C^k-submanifold. For any $r = 1, \dots, n$, we define the projection onto the r-th coordinate

$$\begin{array}{rccl} \pi_r & : & M & \to & \mathbb{R} \\ & & (x_1, \dots, x_n) & \to & x_r. \end{array}$$

This is a C^k map.

Proof. Let $r \in \{1, \dots, n\}$. Let us fix $s \in \mathbb{N}^*$, V an open set in $\mathbb{R}^s$, and $\phi : V \to \mathbb{R}^n$ of class C^k such that $\phi(V) \subset M$. For any $x \in \mathbb{R}^s$, denote $\phi(x) = (\phi_1(x), \dots, \phi_n(x))$. The components $\phi_1, \dots, \phi_n$ are C^k. Hence, $\pi_r \circ \phi = \phi_r$ is C^k. □

Definition 2.27 : C^1 function between two submanifolds

Let M, N be two C^k submanifolds, respectively of $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$. Consider a function

$$f : M \to N.$$

Since $N \subset \mathbb{R}^{n_2}$, we can view f as a map from M to $\mathbb{R}^{n_2}$ rather than from M to N. We say that f is *of class* C^1 (more generally, C^r, for $r \in \{1, \dots, k\}$) between M and N if it is of class C^1 (more generally, C^r) when viewed as a map from M to $\mathbb{R}^{n_2}$.

Example 2.28 : projection on a product submanifold

Let A, B be two C^k-submanifolds, respectively of $\mathbb{R}^a$ and $\mathbb{R}^b$. Recall that $A \times B$ is a submanifold of $\mathbb{R}^{a+b}$ (Proposition 2.5).
We define the projection onto A as

$$\begin{array}{rccl} \pi_A & : & A \times B & \to A \\ & & (x_A, x_B) & \to x_A. \end{array}$$

This is a C^k function.
Similarly, the projection onto B is C^k.

Proof. Consider π_A as a function from $A \times B$ to $\mathbb{R}^a$ and show that this function is C^k. Take $s \in \mathbb{N}^*$, V an open set in $\mathbb{R}^s$, and $\phi : V \to \mathbb{R}^{a+b}$ a C^k map such that $\phi(V) \subset A \times B$.

For any $x \in \mathbb{R}^s$, denote $\phi(x) = (\phi_1(x), \dots, \phi_{a+b}(x))$. The functions $\phi_1, \dots, \phi_{a+b}$ are C^k. The function $\pi_A \circ \phi$ is given by

$$\forall x \in \mathbb{R}^s, \quad \pi_A \circ \phi(x) = \pi_A(\underbrace{\phi_1(x), \dots, \phi_a(x)}_{\text{element of } A}, \underbrace{\phi_{a+1}(x), \dots, \phi_{a+b}(x)}_{\text{element of } B})$$
$$= (\phi_1(x), \dots, \phi_a(x)).$$

Thus, $\pi_A \circ \phi$ is equal to $(\phi_1, \dots, \phi_a)$, which is C^k, and consequently, $\pi_A \circ \phi$ is C^k. □

Definitions 2.25 and 2.27 are more abstract than the definition of differentiability for a function from $\mathbb{R}^n$ to $\mathbb{R}^m$. However, one must not be intimidated. In practice, one rarely needs to resort to these definitions to show that a map is C^1 (or, more generally, C^r). Indeed, as is the case for maps from $\mathbb{R}^n \to \mathbb{R}^m$, basic operations preserve differentiability. For instance, if M is a submanifold and m an integer, the sum of two C^r functions from M to $\mathbb{R}^m$ is also C^r. Similarly, the product of two C^r functions from M to $\mathbb{R}$ is C^r. We will not state each of these properties here, only the one related to composition.

Proposition 2.29 : composition of C^1 functions

Let M, N, P be three C^k submanifolds of, respectively, $\mathbb{R}^{n_M}$, $\mathbb{R}^{n_N}$, and $\mathbb{R}^{n_P}$. Consider two functions

$$f_1 : M \to N \quad \text{and} \quad f_2 : N \to P.$$

If f_1 and f_2 are of class C^r, for some $r \in \{1, \ldots, k\}$, then

$$f_2 \circ f_1 : M \to P$$

is also of class C^r.

Proof. We view $f_2 \circ f_1$ as a function from M to $\mathbb{R}^{n_P}$ and show that this function is C^r. Let $s \in \mathbb{N}^*$ be an integer, V an open set in $\mathbb{R}^s$ and $\phi : V \to \mathbb{R}^{n_M}$ a C^r function such that $\phi(V) \subset M$. We must show that $f_2 \circ f_1 \circ \phi$ is of class C^r on V.

Since $f_1 : M \to N$ is of class C^r, it is also C^r when viewed as a function from M to $\mathbb{R}^{n_N}$. From Definition 2.25, $f_1 \circ \phi : V \to \mathbb{R}^{n_N}$ is C^r. Moreover, $(f_1 \circ \phi)(V) \subset f_1(M) \subset N$. As $f_2 : N \to P \subset \mathbb{R}^{n_P}$ is C^r, the function $f_2 \circ (f_1 \circ \phi)$ is C^r, also from Definition 2.25.

Since $f_2 \circ f_1 \circ \phi = f_2 \circ (f_1 \circ \phi)$, this proves that $f_2 \circ f_1 \circ \phi$ is C^r. □

Exercise 3

Show that the map

$$f : \begin{array}{ccc} \mathbb{S}^1 & \to & \mathbb{S}^1 \\ (x_1, x_2) & \to & (x_1^2, x_2\sqrt{1 + x_1^2}) \end{array}$$

is well-defined and C^∞.

Definition 2.30 : diffeomorphism between manifolds

Let M, N be two C^k submanifolds of $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$, respectively. Consider a map

$$\phi : M \to N.$$

For any $r \in \{1, \ldots, k\}$, we say that ϕ is a C^r-diffeomorphism between M and N if it satisfies the following three properties:

1. ϕ is a bijection from M to N;
2. ϕ is of class C^r on M;
3. ϕ^{-1} is of class C^r on N.

2.4.2 [More advanced] Differentials

Note that, contrarily to what we did for maps from $\mathbb{R}^n$ to $\mathbb{R}^m$, we have defined the notion of *differentiable function* between manifolds without introducing the notion of *differential*. Nevertheless, one can still define this notion; this is the aim of the following definition.

Definition 2.31 : differential on manifolds

Let M, N be two C^k submanifolds of, respectively, $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$. Let

$$f : M \to N$$

be a C^r function, where $r \in \{1, \dots, k\}$.
Let $x \in M$. For any $v \in T_xM$, fix I_v an open interval in $\mathbb{R}$ containing 0 and $c_v : I \to \mathbb{R}^{n_1}$ as in the definition of the tangent space (2.14), i.e., a C^1 function with values in M such that $c_v(0) = x$ and $c_v'(0) = v$.
The *differential of f at x*, denoted $df(x)$, is the following map:

$$\begin{array}{rcrcl} df(x) & : & T_xM & \to & T_{f(x)}N \\ & & v & \to & (f \circ c_v)'(0). \end{array}$$

The map $df(x)$ is well-defined: $f \circ c_v : I_v \to \mathbb{R}^{n_2}$ is a C^1 function, with values in N, such that $f \circ c_v(0) = f(x)$, so $(f \circ c_v)'(0)$ is indeed an element of $T_{f(x)}N$.

Remark

If M is an open subset of $\mathbb{R}^{n_1}$, then f, viewed as a function from this open subset of $\mathbb{R}^{n_1}$ to $\mathbb{R}^{n_2}$, is differentiable in the usual sense, and the differentials defined in Definitions 1.1 and 2.31 coincide, as in that case, denoting $df(x)$ the usual differential,

$$(f \circ c_v)'(0) = df(c_v(0))(c_v'(0)) = df(x)(v).$$

Theorem 2.32

We keep the notation from Definition 2.31.
The map $df(x)$ does not depend on the choice of intervals I_v and functions c_v.
Moreover, it is linear.

Proof. Let $v \in T_xM$. Show that $df(x)(v) = (f \circ c_v)'(0)$ does not depend on the choice of I_v and c_v. To do this, we will give an alternative expression for $df(x)(v)$ that does not involve I_v or c_v.

Let d_1 and d_2 be the dimensions of M and N. We use the "diffeomorphism" definition of submanifolds (Property 1 of Definition 2.1). Let $U_M, V_M \subset \mathbb{R}^{n_1}$ be neighborhoods of x and 0, respectively, and $\phi_M : U_M \to V_M$ be a C^k-diffeomorphism such that $\phi_M(x) = 0$ and

$$\phi_M(M \cap U_M) = (\mathbb{R}^{d_1} \times \{0\}^{n_1 - d_1}) \cap V_M.$$

Denote $\phi_{M,0}^{-1}$ the restriction of ϕ_M^{-1} to $(\mathbb{R}^{d_1} \times \{0\}^{n_1-d_1}) \cap V_M$. We have

$$\begin{aligned} df(x)(v) &= (f \circ c_v)'(0) \\ &= (f \circ \phi_{M,0}^{-1} \circ \phi_M \circ c_v)'(0) \\ &= ((f \circ \phi_{M,0}^{-1}) \circ \phi_M \circ c_v)'(0). \end{aligned}$$

The map $f \circ \phi_{M,0}^{-1}$ is defined on an open subset of $\mathbb{R}^{d_1}$ (actually, on $(\mathbb{R}^{d_1} \times \{0\}^{n_1-d_1}) \cap V_M$, but this is exactly an open set of $\mathbb{R}^{d_1}$ if one ignores the $(n_1 - d_1)$ zeros). It is of class C^r on this subset, since it is the composition of two C^r maps. Thus, the maps $f \circ \phi_{M,0}^{-1}, \phi_M$ and c_v are defined on open subsets of $\mathbb{R}^n$ (for different values of n) and differentiable in the usual sense. The usual theorem on the composition of differentials then gives

$$\begin{aligned} df(x)(v) &= (d(f \circ \phi_{M,0}^{-1})(\phi_M \circ c_v(0)) \circ d\phi_M(c_v(0)))(c_v'(0)) \\ &= d(f \circ \phi_{M,0}^{-1})(0) \circ d\phi_M(x)(v). \end{aligned}$$

As announced, this expression does not depend on c_v or I_v, which completes the first part of the proof.

The linearity of $df(x)$ follows from the same argument. Indeed, our reasoning shows that

$$df(x) = d(f \circ \phi_{M,0}^{-1})(0) \circ d\phi_M(x),$$

i.e., $df(x)$ is the composition of two linear maps. Therefore, it is linear. □

As the notion of differentiability, the notion of differential for maps between manifolds is governed by almost the same rules as for maps between $\mathbb{R}^m$ and $\mathbb{R}^n$. Let's state, for example, the rule of composition of differentials.

Proposition 2.33

Let M, N, P be three C^k submanifolds of $\mathbb{R}^{n_M}$, $\mathbb{R}^{n_N}$, and $\mathbb{R}^{n_P}$, respectively. Consider two C^1 maps,

$$f_1 : M \to N \quad \text{and} \quad f_2 : N \to P.$$

For any $x \in M$,

$$d(f_2 \circ f_1)(x) = df_2(f_1(x)) \circ df_1(x).$$

Proof. Let $v \in T_xM$. Show that

$$d(f_2 \circ f_1)(x)(v) = df_2(f_1(x)) \circ df_1(x)(v).$$

Let I_v be an open interval in $\mathbb{R}$ containing 0, and let $c_v : I_v \to \mathbb{R}^{n_M}$ be a C^1 function such that $c_v(I_v) \subset M$, $c_v(0) = x$, and $c_v'(0) = v$. The definition of the differential gives

$$d(f_2 \circ f_1)(x)(v) = (f_2 \circ f_1 \circ c_v)'(0).$$

Let $w = (f_1 \circ c_v)'(0) = df_1(x)(v) \in \mathbb{R}^{n_N}$. The function $f_1 \circ c_v : I_v \to \mathbb{R}^{n_N}$ is C^1 and $f_1 \circ c_v(I_v) \subset N$. It satisfies $f_1 \circ c_v(0) = f_1(x)$ and, by definition of w, $(f_1 \circ c_v)'(0) = w$. The definition of the differential for f_2 then gives

$$df_2(f_1(x))(w) = (f_2 \circ f_1 \circ c_v)'(0).$$

Thus,

$$\begin{aligned} d(f_2 \circ f_1)(x)(v) &= df_2(f_1(x))(w) \\ &= df_2(f_1(x))(df_1(x)(v)) \\ &= [df_2(f_1(x)) \circ df_1(x)]\,(v). \end{aligned}$$

□

To give one more example of a standard result from differential calculus which straightforwardly generalizes to differential calculus on submanifolds, let us state the submanifold version of the local inversion theorem.

Theorem 2.34 : local inversion on submanifolds

Let M, N be two C^k submanifolds of $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$, respectively. Let $x_0 \in M$. For $r \in \{1, \dots, k\}$, consider a C^r map,

$$f : M \to N.$$

If $df(x_0) : T_{x_0}M \to T_{f(x_0)}N$ is bijective, then there exist U_{x_0} an open neighborhood of x_0 in M and $V_{f(x_0)}$ an open neighborhood of $f(x_0)$ in N such that f is a C^r-diffeomorphism from U_{x_0} to $V_{f(x_0)}$.

Proof. Let d be the dimension of M. Note that N has the same dimension as M: $df(x_0)$ is a bijective linear map between $T_{x_0}M$ and $T_{f(x_0)}N$, so

$$\dim T_{f(x_0)}N = \dim T_{x_0}M = d.$$

Let $U_M, V_M \subset \mathbb{R}^{n_1}$ be open neighborhoods of x_0 and 0, respectively, and $\phi_M : U_M \to V_M$ a C^k-diffeomorphism such that

$$\phi_M(M \cap U_M) = (\mathbb{R}^d \times \{0\}^{n_1 - d}) \cap V_M,$$

and $\phi_M(x_0) = 0$.

Similarly, let $U_N, V_N \subset \mathbb{R}^{n_2}$ be open neighborhoods of $f(x_0)$ and 0, and $\phi_N : U_N \to V_N$ a C^k-diffeomorphism such that

$$\phi_N(N \cap U_N) = (\mathbb{R}^d \times \{0\}^{n_2-d}) \cap V_N,$$

and $\phi_N(f(x_0)) = 0$.

The idea of the proof is to go back to the case where f is defined on an open subset of $\mathbb{R}^d$ and then apply the classical local inversion theorem. To do this, we "transfer" f to a map from $\mathbb{R}^d \times \{0\}^{n_1-d}$ to $\mathbb{R}^d \times \{0\}^{n_2-d}$ by composing it with the diffeomorphisms ϕ_M and ϕ_N.

More precisely, let $\phi_{M,0}^{-1}$ be the restriction of ϕ_M^{-1} to $(\mathbb{R}^d \times \{0\}^{n_1-d}) \cap V_M$. Define

$$g \overset{def}{=} \phi_N \circ f \circ \phi_{M,0}^{-1} : (\mathbb{R}^d \times \{0\}^{n_1-d}) \cap V_M \to (\mathbb{R}^d \times \{0\}^{n_2-d}) \cap V_N.$$

This definition is valid if we reduce U_M, V_M so that $f(U_M) \subset U_N$. The map g is C^r and its differential at 0 is injective: it is the composition of $d\phi_N(f(x_0))$, $df(x_0)$, and $d\phi_{M,0}^{-1}(0)$, all of which are injective. Since it goes from $\mathbb{R}^d$ to $\mathbb{R}^d$, it is bijective[2].

According to the classical local inversion theorem (Theorem 1.10), there exist E_M, E_N open neighborhoods of 0 in $\mathbb{R}^d$ such that g is a C^r-diffeomorphism from $E_M \times \{0\}^{n_1-d}$ to $E_N \times \{0\}^{n_2-d}$. Then f is a C^r-diffeomorphism from $U_{x_0} \overset{def}{=} \phi_M^{-1}(E_M \times \{0\}^{n_1-d})$ to $V_{f(x_0)} \overset{def}{=} \phi_N^{-1}(E_N \times \{0\}^{n_2-d})$: on these sets,

$$f = \phi_N^{-1} \circ g \circ \phi_M.$$

Since ϕ_M is a diffeomorphism (of class C^k hence also of class C^r) from U_{x_0} to $E_M \times \{0\}^{n_1-d}$, g is a C^r-diffeomorphism from $E_M \times \{0\}^{n_1-d}$ to $E_N \times \{0\}^{n_2-d}$, and ϕ_N^{-1} is a diffeomorphism (C^k hence also C^r) from $E_N \times \{0\}^{n_2-d}$ to $V_{f(x_0)}$, the map f is a composition of C^r-diffeomorphisms, hence a C^r-diffeomorphism. □

[2] We can see $\phi_N \circ f \circ \phi_{M,0}^{-1}$ as a map between two open subsets of $\mathbb{R}^d$.

Chapter 3

Solutions of some exercises

3.1 Exercise 1

1. Let $i, j \in \{1, \dots, n\}$ be fixed. From the definition of the differential,

$$d(df)(x)(e_i) = \lim_{t\to 0} \frac{df(x+te_i) - df(x)}{t} \quad (\in \mathcal{L}(\mathbb{R}^n, \mathbb{R})).$$

Since the map $(L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}) \to L(e_j) \in \mathbb{R})$ is continuous,

$$\begin{aligned} d(df)(x)(e_i)(e_j) &= \left(\lim_{t\to 0} \frac{df(x+te_i) - df(x)}{t}\right)(e_j) \\ &= \lim_{t\to 0} \left(\left(\frac{df(x+te_i) - df(x)}{t}\right)(e_j)\right) \\ &= \lim_{t\to 0} \frac{df(x+te_i)(e_j) - df(x)(e_j)}{t} \\ &= \lim_{t\to 0} \frac{\frac{\partial f}{\partial x_j}(x+te_i) - \frac{\partial f}{\partial x_j}(x)}{t} \\ &= \frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(x). \end{aligned}$$

2. a) Let $r > 0$ be such that $B(x, 2r) \subset U$. For any $t, u \in]-r; r[$, $f(x + te_i + ue_j)$ is well-defined. For any $t \in]-r; r[$, the map

$$\begin{array}{cccc} g_t & : &]-r;r[& \to & \mathbb{R} \\ & & s & \to & f(x+te_i+se_j) \end{array}$$

is differentiable. For each s, $g_t'(s) = \frac{\partial f}{\partial x_j}(x + te_i + se_j)$. Therefore,

$$\begin{aligned} f(x+te_i+ue_j) - f(x+te_i) &= g_t(u) - g_t(0) \\ &= \int_0^u g_t'(s)ds \\ &= \int_0^u \frac{\partial f}{\partial x_j}(x+te_i+se_j)ds. \end{aligned}$$

The same reasoning, but replacing t with 0, shows that

$$f(x+ue_j) - f(x) = \int_0^u \frac{\partial f}{\partial x_j}(x+se_j)ds.$$

If we substract this equality from the previous one, we obtain the result.

b) The map $\frac{\partial f}{\partial x_j}$ is differentiable at x (since df is differentiable). Therefore, for t, s going to 0,

$$\frac{\partial f}{\partial x_j}(x + te_i + se_j) = d\left(\frac{\partial f}{\partial x_j}\right)(x)(te_i + se_j) + o(|s| + |t|)$$
$$\text{and } \frac{\partial f}{\partial x_j}(x + se_j) = d\left(\frac{\partial f}{\partial x_j}\right)(x)(se_j) + o(s),$$

so that

$$\begin{aligned}
&\frac{\partial f}{\partial x_j}(x + te_i + se_j) - \frac{\partial f}{\partial x_j}(x + se_j)\\
&= d\left(\frac{\partial f}{\partial x_j}\right)(x)(te_i + se_j) - d\left(\frac{\partial f}{\partial x_j}\right)(x)(se_j) + o(|s| + |t|)\\
&= d\left(\frac{\partial f}{\partial x_j}\right)(x)(te_i) + o(|s| + |t|)\\
&\quad \text{(by linearity of the differential)}\\
&= t\frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(x) + o(|s| + |t|).
\end{aligned}$$

Consequently,

$$\begin{aligned}
\left|\frac{\partial f}{\partial x_j}(x + te_i + se_j) - \frac{\partial f}{\partial x_j}(x + se_j) - t\frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(x)\right| &= o(|s| + |t|)\\
&\leq \epsilon(|t| + |s|)\\
&\text{for all } t, s \text{ close enough to zero.}
\end{aligned}$$

c) Let $r > 0$ be such that the inequality from the previous question holds for all $t, s \in] - r; r[$. We combine Questions a) and b): for all $t, u \in] - r; r[$,

$$\begin{aligned}
&\left|\phi(t, u) - \int_0^u t\frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(x)ds\right|\\
&\leq \int_{[0;u]} \left|\frac{\partial f}{\partial x_j}(x + te_i + se_j) - \frac{\partial f}{\partial x_j}(x + se_j) - t\frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(x)\right| ds\\
&\quad \text{(by triangular inequality)}\\
&\leq \int_{[0;u]} \epsilon(|t| + |s|)ds\\
&= \epsilon\left(|t|\,|u| + \frac{|u|^2}{2}\right)\\
&\leq \epsilon\left(|t|\,|u| + |u|^2\right).
\end{aligned}$$

We obtain the result by noting that

$$\int_0^u t\frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(x)ds = tu\frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(x).$$

d) The definition of ϕ is invariant to exchanging t with u and i with j, so the same reasoning as before gives the same inequality as in the previous question, with t replaced by u and i by j.

e) Using the triangular inequality and the previous two questions, we get that, for all t, u close enough to 0,

$$\left|tu\frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(x) - tu\frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}(x)\right| \leq \epsilon(|u|^2 + 2|t|\,|u| + |t|^2).$$

In particular, for all t close enough to zero, setting $u = t$ and dividing by $|t|^2$,

$$\left|\frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(x) - \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}(x)\right| \leq 4\epsilon.$$

Since $\epsilon > 0$ is arbitrary, this shows that

$$\left| \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j}(x) - \frac{\partial}{\partial x_j} \frac{\partial f}{\partial x_i}(x) \right| = 0,$$

hence $\frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j}(x) = \frac{\partial}{\partial x_j} \frac{\partial f}{\partial x_i}(x)$.

3.2 Exercise 2

We apply the mean value inequality to $U = \mathbb{R}^n$ and $M = 1$:

$$\forall x, y \in \mathbb{R}^n, \quad ||f(x) - f(y)|| \leq ||x - y||.$$

In particular, for $y = 0$:

$$\forall x \in \mathbb{R}^n, \quad ||f(x) - f(0)|| \leq ||x||.$$

Using the triangular value inequality, it holds for all $x \in \mathbb{R}^n$ that

$$\begin{aligned} ||f(x)|| &\leq ||f(0)|| + ||f(x) - f(0)|| \\ &\leq ||f(0)|| + ||x||. \end{aligned}$$

3.3 Exercise 3

Showing that f is well-defined consists in showing that $f(x_1, x_2)$ indeed belongs to $\mathbb{S}^1$ for all $(x_1, x_2) \in \mathbb{S}^1$. Let us consider any $(x_1, x_2) \in \mathbb{S}^1$. It holds

$$\begin{aligned} (x_1^2)^2 + \left(x_2 \sqrt{1 + x_1^2} \right)^2 &= x_1^4 + x_2^2(1 + x_1^2) \\ &= x_1^2(x_1^2 + x_2^2) + x_2^2 \\ &= x_1^2 + x_2^2 \\ &= 1. \end{aligned}$$

Therefore, $f(x_1, x_2) \in \mathbb{S}^1$.

Let us now show that f is C^∞. From Definition 2.27, we must show that

$$\begin{array}{cccc} \tilde{f} : & \mathbb{S}^1 & \to & \mathbb{R}^2 \\ & (x_1, x_2) & \to & (x_1^2, x_2 \sqrt{1 + x_1^2}) \end{array}$$

is C^∞. From Example 2.26, we know that

$$\begin{array}{cccc} \pi_1 \times \pi_2 : & \mathbb{S}^1 & \to & \mathbb{R}^2 \\ & (x_1, x_2) & \to & (x_1, x_2) \end{array}$$

is C^∞. As $\tilde{f}$ is the composition of $\pi_1 \times \pi_2$ with the map

$$\begin{array}{cccc} g : & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (x_1, x_2) & \to & (x_1^2, x_2 \sqrt{1 + x_1^2}), \end{array}$$

which is C^∞ (it is a composition of $\sqrt{.} : \mathbb{R}_+^* \to \mathbb{R}$, which is C^∞ on this domain, and polynomial functions). From Proposition 2.29, $\tilde{f}$ is C^∞.