University of Luxembourg

Differential Geometry

Master in Mathematics

Norbert Poncin

2015

UNIVERSITÉ DU LUXEMBOURG

Attribution

The first TikZ versions of all figures and illustrations in Chapters 1-6 were prepared by K. Freis and A. Goedert. Chapter 7 is based on joint work with T. Covolo and S. Schouten.

Course description

Course name

Differential Geometry

ECTS

5

Teaching units

45 hours, 4 TU on even weeks and 2 on odd weeks

Course type

Interactive lectures and exercise sessions

Prerequisites

None

Learning outcomes

On successful completion of the course, the student should be able to:

(i) Explain the main definitions and results of Differential Geometry

(ii) Comment on new concepts, like the category of smooth manifolds, embedded submanifolds, smooth scalar observables, their derivatives, vector bundles and vector fields, differential equations on manifolds, tensor fields, Lie derivatives, differential forms, integral calculus on manifolds...

(iii) Apply the new techniques and solve related problems

(iv) Structure the acquired abilities and summarize essential aspects adopting a higher standpoint

(v) Give a talk for peers or students on a related topic and write scientific texts or lecture notes, observing modern standards in scientific writing, in Didactics and in Pedagogy

(vi) Provide evidence for mastery of the Mathematical Method

Objective

The objective is to provide students with the opportunity to become familiar with a very active area of mathematics that has wide application throughout science. Beyond this goal, special emphasis is placed on the mathematical method, i.e., the optimal technique for learning and applying mathematics, especially with regard to solving real-life problems using mathematical tools. This method is actually the most important goal of any course in mathematics.

Description

Differential Geometry has applications in numerous areas of science, including Einstein's general relativity, string theory, black holes and galaxy clusters, probability, engineering, economics, modeling and design, wireless communications and image processing, biology, chemistry, geology... The main concept of Differential Geometry is differential and in particular smooth manifolds – roughly, higher-dimensional analogs of curves and surfaces. In general relativity, for example, the Universe is often modeled as a four-dimensional smooth manifold equipped with a specific metric. To be able to work scientifically in these new spaces, a generalization of fundamental chapters of mathematical analysis, such as differential calculus and integration theory, is required.

Audience

A special emphasis of this geometry course is on communication skills and the mathematical method. Therefore, the course is enriching for students of the Master's Degree in Secondary Education in Mathematics as well as for students of the Master's Degree in Mathematics.

Evaluation

Oral exam

Warning

The script is aimed at students who have attended the oral lectures. The notes are in standard mathematical text form and only partially take into account the pedagogical approach of the lectures.

Chapter 1
Nonlinear Analysis

The trilogy of theorems we discuss in this chapter will be used throughout the course and in particular in the sections on embedded submanifolds of Cartesian space.

1 Preliminaries

1.1 Taylor's theorem

Let us first fix the notations. If $f \in C^k(\Omega)$, $k \geq 1$, $\Omega \subset \mathbb{R}^p$ open, and if $\alpha = (\alpha_1, \ldots, \alpha_p) \in \mathbb{N}^p$ is a multi-index of order $|\alpha| := \alpha_1 + \ldots + \alpha_p \leq k$, we set

$$\partial_x^\alpha f := \partial_{x^1}^{\alpha_1} \ldots \partial_{x^p}^{\alpha_p} f$$

and

$$\alpha! := \alpha_1! \cdot \ldots \cdot \alpha_p! \; .$$

Similarly, if $h \in \mathbb{R}^p$, we write

$$h^\alpha := (h^1)^{\alpha_1} \cdot \ldots \cdot (h^p)^{\alpha_p} \; .$$

Taylor's theorem can now be formulated as follows:

Theorem 1 (Taylor's theorem). *If $f \in C^k(\Omega, \mathbb{R})$, $k \geq 1$, $\Omega \subset \mathbb{R}^p$ open, and if, for a given $x \in \Omega$ and a given $h \in \mathbb{R}^p$, the segment $\{x + th : t \in [0,1]\}$ is included in Ω, then there exists an intermediate point $x + \theta h$, $\theta \in]0,1[$, between x and $x + h$, such that*

$$f(x+h) = \sum_{|\alpha|<k} \frac{1}{\alpha!} (\partial_x^\alpha f)(x) h^\alpha + \sum_{|\alpha|=k} \frac{1}{\alpha!} (\partial_x^\alpha f)(x + \theta h) h^\alpha \; .$$

Observe that, in the case $p = 1$, we recover the well-known formula

$$f(x+h) = f(x) + f'(x)h + \frac{1}{2!} f''(x)h^2 + \ldots + \frac{1}{(k-1)!} f^{(k-1)}(x)h^{k-1} + \frac{1}{k!} f^{(k)}(x+\theta h)h^k \; .$$

1.2 The Cauchy-Schwarz inequality

We also recall the Cauchy-Schwarz inequality. If $x, x' \in \mathbb{R}^p$, then

$$|\langle x, x' \rangle| \leq |x| \cdot |x'| \,, \tag{1}$$

where $\langle x, y \rangle$ denotes the standard scalar product in $\mathbb{R}^p$ and $|x|, |x'|$ the corresponding norm (of course $|\langle x, x' \rangle|$ denotes the absolute value (standard norm in $\mathbb{R}$) of the scalar product).

1.3 Vector-valued functions of several variables

A function $f : \Omega \subset \mathbb{R}^p \to \mathbb{R}^q$ from an open subset Ω of R^p to $\mathbb{R}^q$ is of class C^k, $k \geq 0$, (resp., of class C^∞ – we also say smooth) if and only if its canonical coordinate functions $f^i : \Omega \subset \mathbb{R}^p \to \mathbb{R}$, $i \in \{1, \ldots, q\}$, are all of class C^k (resp., smooth) in Ω. The derivative of $f \in C^1(\Omega, \mathbb{R}^q)$ at $x_0 \in \Omega$ is the linear map

$$f'(x_0) = \left(\left(\partial_{x^j} f^i\right)(x_0)\right)_{ij} : \mathbb{R}^p \to \mathbb{R}^q. \tag{2}$$

A function $f : \Omega_1 \subset \mathbb{R}^p \to \Omega_2 \subset \mathbb{R}^p$ is a diffeomorphism of class C^k, $k \geq 1$ – we also say a coordinate transformation of class C^k – if $f : \Omega_1 \to \Omega_2$ is a bijection, $f \in C^k(\Omega_1, \mathbb{R}^q)$, and $f^{-1} \in C^k(\Omega_2, \mathbb{R}^p)$.

2 Banach fixed point theorem

Let $f : S \to S$ be a map from a set S to itself. A fixed point of f is an element $s \in S$ such that $f(s) = s$. A fixed point theorem provides conditions under which f has at least one or a unique fixed point.

In the Banach fixed point theorem, S is a complete metric space. A metric space is a set S endowed with a metric or distance d, i.e. a map $d : S \times S \to \mathbb{R}_+$, such that, for any $s, t, u \in S$,

1. $d(s, t) = 0$ if and only if $s = t$,
2. $d(s, t) = d(t, s)$, and
3. $d(s, u) \leq d(s, t) + d(t, u)$.

In a metric space S, we can define the concept of limit of a sequence of points. A sequence $s_n \in S$ converges to $s \in S$ if the sequence $d(s_n, s)$ of real numbers converges to 0, when $n \to +\infty$. If it exists, the limit of a sequence s_n is unique. Indeed, if s, t are two limits, it follows from the properties of d that

$$d(s, t) \leq d(s_n, s) + d(s_n, t) \to 0$$

so that $s = t$.

A sequence s_n of a metric space S is a Cauchy sequence if $d(s_p, s_q) \to 0$, when $\inf(p,q) \to +\infty$. Of course, any converging sequence $s_n \to s$ is a Cauchy sequence, since

$$d(s_p, s_q) \leq d(s_p, s) + d(s_q, s) \to 0\,,$$

when $\inf(p,q) \to +\infty$. A metric space is called complete, if the converse holds true, i.e. if any Cauchy sequence converges.

Theorem 2. *In a complete metric space S (with metric d), any contraction – i.e. any map $f : S \to S$ for which there exists $\theta \in [0,1[$ such that, for any $s,t \in S$,*

$$d(f(s), f(t)) \leq \theta\; d(s,t) \;–$$

has a unique fixed point s. Moreover, s can be found as follows: start with an arbitrary element $s_0 \in S$ and define a sequence $s_n = f(s_{n-1})$, $n \geq 1$; this sequence converges to the fixed point s.

Proof. We first prove uniqueness. Let s,t be two fixed points. Then

$$d(s,t) = d(f(s), f(t)) \leq \theta\; d(s,t)\;,$$

so that $(1-\theta)d(s,t) = 0$ and $s = t$.

As for existence, we will prove that the sequence constructed in the statement of the theorem is a Cauchy sequence and that its limit is a fixed point.

For that purpose let us first recall the following result regarding the geometric series:

$$\sum_{n=1}^{+\infty} \theta^n = \frac{\theta}{1-\theta}\;,$$

if $|\theta| < 1$; it follows that the sequence of partial sums is a Cauchy sequence.

Note now that

$$d(s_{n+1}, s_n) = d(f(s_n), f(s_{n-1})) \leq \theta\; d(s_n, s_{n-1}) \leq \ldots \leq \theta^n d(s_1, s_0)\;.$$

Hence, for $p \leq q$,

$$\begin{aligned} d(s_p, s_q) &\leq d(s_p, s_{p+1}) + d(s_{p+1}, s_{p+2}) + \ldots + d(s_{q-1}, s_q) \\ &\leq (\theta^p + \theta^{p+1} + \ldots + \theta^{q-1}) d(s_1, s_0) \to 0\;, \end{aligned}$$

when $\inf(p,q) \to +\infty$, so that the sequence s_n is Cauchy and converges to a limit s.

To see that this limit s is a fixed point, observe first that f is continuous. Indeed, if $x \to x_0$ in S, then

$$d(f(x), f(x_0)) \leq \theta\; d(x, x_0) \to 0\;.$$

Therefore,

$$s \leftarrow s_n = f(s_{n-1}) \to f(s)\;,$$

so that $f(s) = s$, in view of the uniqueness of a limit. □

3 Implicit function theorem

3.1 Statement and proof

Example 1. Let $f(x,y) = y - \sin y - x$. It is not possible to solve the equation $f(x,y) = 0$ explicitly with respect to y, i.e. to compute a solution of the type $y = g(x)$. However, it can easily be seen that, for any $x \in \mathbb{R}$, the function $f(x,-)$ is continuous and strictly increasing, and more precisely that it is a bijection $f(x,-) : \mathbb{R} \to \mathbb{R}$. This means that the equation $f(x,y) = 0$ has, for any $x \in \mathbb{R}$, a unique solution $y = g(x)$, which, since it is not known explicitly, is called the implicit function defined by the equation $f(x,y) = 0$.

More generally, let $\Omega_1 \subset \mathbb{R}^p$ and $\Omega_2 \subset \mathbb{R}^q$ be two open subsets, and let $f : \Omega_1 \times \Omega_2 \to \mathbb{R}^q$ be a function. The objective of the implicit function theorem is to solve the equation $f(x,y) = 0$ implicitly in the form $y = g(x)$, and to study the differentiability of g for a given differentiability of f. Note that $f(x,y) = 0$ contains q scalar equations and that y contains q scalar unknown variables, so that the stated problem actually makes sense.

Theorem 3 (Implicit function theorem). *If*

$$f : \Omega_1 \times \Omega_2 \subset \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^q$$

is of class C^k, $k \geq 1$, and if, for some $(x_0, y_0) \in \Omega_1 \times \Omega_2$, the derivative $(\partial_y f)(x_0, y_0)$ (which is a $q \times q$ matrix with entries in $\mathbb{R}$) is invertible, then the equation

$$f(x,y) = f(x_0, y_0)$$

can be solved implicitly with respect to y in a neighborhood of (x_0, y_0).

More precisely, there exist open neighborhoods $\omega_1 \ni x_0$ in Ω_1 and $\omega_2 \ni y_0$ in Ω_2, as well as a function $g : \omega_1 \to \omega_2$ of class C^k, such that

$$x \in \omega_1, y \in \omega_2, f(x,y) = f(x_0, y_0) \Leftrightarrow x \in \omega_1, y = g(x) \; .$$

Eventually, the derivative of the implicit function $g : \omega_1 \subset \mathbb{R}^p \to \omega_2 \subset \mathbb{R}^q$ can be obtained from the derivatives of the given function

$$f : \Omega_1 \times \Omega_2 \subset \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^q \; :$$

for any $x \in \omega_1$, we have

$$\partial_x g = - \; (\partial_y f)^{-1}(x, g(x)) \; (\partial_x f)(x, g(x)) \; . \tag{3}$$

Proof. Let us first observe that we can, without loss of generality, assume that $(x_0, y_0) = (0,0)$ and that $f(x_0, y_0) = f(0,0) = 0$. Indeed, otherwise we use the coordinate transformation $x = x_0 + X, y = y_0 + Y$, so that $f(x,y) = f(x_0, y_0)$ reads

$$F(X,Y) := f(x_0 + X, y_0 + Y) - f(x_0, y_0) = 0 \; .$$

(It is easily checked that the new function $F(X,Y)$ verifies the conditions of the theorem for $\Omega_1 \times \Omega_2$ replaced by $(\Omega_1 - x_0) \times (\Omega_2 - y_0)$ and for (x_0, y_0) replaced by $(0,0)$.)

The basic idea of the proof is to introduce an auxiliary function that will allow to reduce the problem to an application of Banach's fixed point theorem. Indeed, if we consider the auxiliary function

$$\phi(x,y) = y - (\partial_y f)^{-1}(0,0) f(x,y) \in C^k(\Omega_1 \times \Omega_2, \mathbb{R}^q) \ , \tag{4}$$

we have

$$f(x,y) = 0 \Leftrightarrow \phi(x,y) = y \ ,$$

so that instead of looking for the solutions of $f(x,y) = 0$, we now look for the fixed points of the parametric function $\phi(x,-)$.

The main condition in the Banach fixed point theorem is the requirement that $\phi(x,-)$ be a contraction, i.e. that

$$|\phi(x,y) - \phi(x,y')| \leq \theta \, |y - y'| \ ,$$

with $\theta \in [0,1[$ (although the appropriate complete metric space will be found not until later, it is already clear from the context that the good notion of distance is the standard metric of $\mathbb{R}^q$).

We now prove this inequality. Since $y_0 = 0 \in \Omega_2 \subset \mathbb{R}^q$, there exists an open ball $b(0,\eta) := \{y \in \mathbb{R}^q : |y| < \eta\}$ centered at 0, $\eta > 0$, such that $\bar{b}(0,\eta) \subset \Omega_2$. For any $x \in \Omega_1$ and any $y, y' \in b(0,\eta)$, we then have, in view of Taylor's theorem applied at order $k = 1$,

$$\phi^i(x,y') - \phi^i(x,y) = \sum_{k=1}^{q} (\partial_{y^k}\phi^i)(x,z)(y'^k - y^k) \ ,$$

where z is intermediate between y and y', so that $|z| < \eta$ (we omit the dependence of z on i and x). It now follows from the Cauchy-Schwarz inequality (1) that

$$|\phi^i(x,y') - \phi^i(x,y)| = |\sum_{k=1}^{q} (\partial_{y^k}\phi^i)(x,z)(y'^k - y^k)|$$

$$\leq \sqrt{\sum_k \left((\partial_{y^k}\phi^i)(x,z)\right)^2} \, |y' - y| \leq \sup_{|z|\leq\eta} \sqrt{\sum_k \left((\partial_{y^k}\phi^i)(x,z)\right)^2} \, |y' - y| \ ,$$

as any $\mathbb{R}$-valued function that is continuous on a compact subset of $\mathbb{R}^q$ is bounded on this subset. When passing to the squares, we get

$$|\phi^i(x,y') - \phi^i(x,y)|^2 \leq \left(\sup_{|z|\leq\eta} \sum_k \left((\partial_{y^k}\phi^i)(x,z)\right)^2 \right) \, |y' - y|^2$$

$$\leq \left(\sup_{|z|\leq\eta} \sum_{ik} \left((\partial_{y^k}\phi^i)(x,z)\right)^2 \right) \, |y' - y|^2 \ ,$$

since the square of the supremum is the supremum of the squares (the same holds true for the square root). Sum now over i,

$$\sum_{i=1}^{q} |\phi^i(x,y') - \phi^i(x,y)|^2 \leq q \left(\sup_{|z|\leq\eta} \sum_{ik} \left((\partial_{y^k}\phi^i)(x,z)\right)^2 \right) |y'-y|^2 ,$$

and pass to the square root,

$$|\phi(x,y') - \phi(x,y)| \leq \sqrt{q} \left(\sup_{|z|\leq\eta} \sqrt{\sum_{ik} \left((\partial_{y^k}\phi^i)(x,z)\right)^2} \right) |y'-y|$$

$$= \sqrt{q} \left(\sup_{|z|\leq\eta} ||\partial_z\phi(x,z)|| \right) |y'-y| ,$$

if $x \in \Omega_1$ and $|y|, |y'| \leq \eta$ (possibly modulo replacement of η by a smaller η).

Observe now that it follows from Definition (4) that

$$(\partial_z\phi)(0,0) = \mathrm{id} - (\partial_z f)^{-1}(0,0)\ (\partial_z f)(0,0) = 0 .$$

Hence, $||\partial_z\phi(x,z)||$ is a positive continuous function in $\Omega_1 \times \Omega_2$, which vanishes at $(0,0)$: in a sufficiently small neighborhood

$$\{x \in \Omega_1 : |x| \leq \eta\} \times \{z \in \Omega_2 : |z| \leq \eta\}$$

of $(0,0)$ this function is smaller than any strictly positive constant, in particular smaller than $\frac{1}{2\sqrt{q}} > 0$ (if the radius η that appears here and the radius η that appeared above are not equal, we replace both by their infimum). It follows that

$$|\phi(x,y') - \phi(x,y)| \leq \sqrt{q} \left(\sup_{|z|\leq\eta} ||\partial_z\phi(x,z)|| \right) |y'-y| \leq \frac{1}{2}|y'-y| , \tag{5}$$

if $|x|, |y|, |y'| \leq \eta$.

The latter result means that, for any x such that $|x| \leq \eta$, the map $\phi(x,-)$ is a contraction, but we still have to check that $\phi(x,-)$ maps $\bar{b} := \{y \in \mathbb{R}^q : |y| \leq \eta\}$ to itself, and that $\bar{b}$ is a complete metric space. The last requirement is satisfied, since any compact metric space is complete. As for the first, observe that, for any $(x,y) \in \Omega_1 \times \Omega_2$, we have

$$|\phi(x,y)| \leq |\phi(x,y) - \phi(x,0)| + |\phi(x,0)| . \tag{6}$$

In view of the contraction property,

$$|\phi(x,y) - \phi(x,0)| \leq \frac{1}{2}|y| \leq \frac{1}{2}\eta , \tag{7}$$

if $|x|, |y| \leq \eta$. Further, as $\phi(x,0)$ is continuous at 0 and $\phi(0,0) = 0$, we have

$$|\phi(x,0)| \leq \frac{1}{2}\eta , \tag{8}$$

provided $|x|$ is small enough, say $|x| \leq \eta' \leq \eta$. It follows from (6), (7), and (8) that

$$|\phi(x,y)| \leq \eta\,,\ \text{if } |x| \leq \eta', |y| \leq \eta\,. \tag{9}$$

Finally, if we set $\bar{\mathfrak{b}} := \{x \in \mathbb{R}^p : |x| \leq \eta'\}$, the map $\phi(x,-)$ is, in view of (5) and (9), for any $x \in \bar{\mathfrak{b}}$, a contraction

$$\phi(x,-) : \bar{b} \to \bar{b}$$

from the complete metric space $\bar{b} = \{y \in \mathbb{R}^q : |y| \leq \eta\}$ into itself.

It now follows from Banach's fixed point theorem that, for any $x \in \mathfrak{b}$, $\phi(x,-)$ has a unique fixed point in $\bar{b}$. When denoting this point by $g(x)$, we get a map $g : \mathfrak{b} \to \bar{b}$. A refinement of the fixed point theorem – that we omit here for simplicity – shows that this map g is actually continuous. Hence, we get

$$x \in \mathfrak{b}, y \in \bar{b},\ f(x,y) = 0 \Leftrightarrow x \in \mathfrak{b}, y \in \bar{b},\ \phi(x,y) = y \Leftrightarrow x \in \mathfrak{b},\ y = g(x)\,.$$

It suffices now to consider the open neighborhood $\omega_1 := g^{-1}b \subset \mathfrak{b}$ (resp., $\omega_2 := b$) of $0 \in \mathbb{R}^p$ (resp., $0 \in \mathbb{R}^q$). Indeed, it is easily seen that

$$x \in \omega_1, y \in \omega_2,\ f(x,y) = 0 \Leftrightarrow x \in \omega_1,\ y = g(x)\,.$$

This completes the proof of the implicit function theorem, except that we have still to show that $g : \omega_1 \to \omega_2$ is of the same class C^k as f, and that we must explain the derivation formula for implicit functions. We will not give the proof regarding the differentiability of g. As for the derivative of g, remark that $f(x,g(x)) = 0$, for all $x \in \omega_1$. It then follows from the derivation theorem for composite functions that, in ω_1,

$$(\partial_x f)(x,g(x)) + (\partial_y f)(x,g(x))\ \ \partial_x g = 0\,.$$

Since $\partial_y f$ is invertible at $(0,0)$, i.e. since the continuous function $\det(\partial_y f)$ does not vanish at $(0,0)$, this determinant remains nonzero in $\omega_1 \times g(\omega_1)$ (possibly modulo replacement of ω_1 by a smaller ω_1 – note that $g(0) = 0$ and that g is continuous). Therefore, for any $x \in \omega_1$,

$$\partial_x g = -\ \ (\partial_y f)^{-1}(x,g(x))\ \ (\partial_x f)(x,g(x))\,.$$

□

There exist refinements of the implicit function theorem that we do not mention in these lecture notes.

3.2 Application

It can be shown that the integration of the differential equation

$$x(\cos\frac{y}{x})y' = y(\cos\frac{y}{x}) - x\,, \tag{10}$$

can be reduced to the integration of the equation

$$x - Ce^{-\sin\frac{y}{x}} = 0\ , \tag{11}$$

where C is an arbitrary nonzero constant.

We first study Equation (11) using the implicit function theorem (ImFT). Set

$$f(x,y) = x - Ce^{-\sin\frac{y}{x}}\ .$$

For any $C \in \mathbb{R}^\times := \mathbb{R}\setminus\{0\}$, we have $f \in C^\infty(\mathbb{R}^\times \times \mathbb{R})$. Further, for any $(x,y) \in \mathbb{R}^\times \times \mathbb{R}$,

$$f'_x(x,y) = 1 - Ce^{-\sin\frac{y}{x}}(\cos\frac{y}{x})\frac{y}{x^2}$$

and

$$f'_y(x,y) = Ce^{-\sin\frac{y}{x}}(\cos\frac{y}{x})\frac{1}{x}\ .$$

The ImFT allows to solve, implicitly and locally, the equation $f(x,y) = f(x_0,y_0)$, hence, the equation (11), i.e. $f(x,y) = 0$, if we find a point $(x_0,y_0) \in \mathbb{R}^\times \times \mathbb{R}$ such that $f(x_0,y_0) = 0$, i.e. if we find a specific solution of Equation (11) (and if, in addition, the conditions of the ImFT are satisfied). The specific solution $(C,0) \in \mathbb{R}^\times \times \mathbb{R}$ is readily guessed. Since $f'_y(C,0) = 1 \neq 0$ the ImFT applies: there is a neighborhood $\omega_1 \ni C$, a neighborhood $\omega_2 \ni 0$, and an implicit function $g \in C^\infty(\omega_1)$ such that, for any $x \in \omega_1$,

$$f(x,g(x)) = 0,\ \text{i.e}, \ x - Ce^{-\sin\frac{g(x)}{x}} = 0\ . \tag{12}$$

We now show that, in conformance with our above remark, $g(x)$ is in ω_1 also a solution of Equation (10). Observe first that, for all $x \in \omega_1$,

$$g'(x) = -\frac{1 - Ce^{-\sin\frac{g(x)}{x}}(\cos\frac{g(x)}{x})\frac{g(x)}{x^2}}{Ce^{-\sin\frac{g(x)}{x}}(\cos\frac{g(x)}{x})\frac{1}{x}} = -\frac{1 - (\cos\frac{g(x)}{x})\frac{g(x)}{x}}{\cos\frac{g(x)}{x}}\ ,$$

in view of the ImFT and Equation (12). Hence, $g(x)$ is actually a solution of (10) in $\omega_1 \ni C$. It follows that Equation (10) admits a solution in a neighborhood of any nonzero real number C.

3.3 Exercises

1. The integration of the differential equation

$$xy'(2y - x) = y^2 \tag{13}$$

can be reduced to the integration of the equation

$$y^2 - xy - Cx = 0\ , \tag{14}$$

where C denotes an arbitrary real constant.

a. Prove that, for $C \neq 0$, Equation (14) defines an implicit function g_1 (resp., g_2) from a neighborhood of C into a neighborhood of $\frac{1+\sqrt{5}}{2}C$ (resp., $\frac{1-\sqrt{5}}{2}C$), and show that g_1 (resp., g_2) is a solution of Equation (13) in a neighborhood of C.

b. Solve Equation (14) explicitly with respect to y (what is possible in the present example) and verify that Equation (13) actually admits two solutions in the neighborhood of any nonzero C.

2. The integration of the differential equation

$$xyy' = x^2 + y^2 \tag{15}$$

can be reduced to the integration of the equation

$$y^2 - 2x^2 \ln|\frac{x}{C}| = 0 \ , \tag{16}$$

where C is an arbitrary strictly positive constant.

a. Prove that Equation (16) defines an implicit function g_1 (resp., g_2) from a neighborhood of eC into a neighborhood of $eC\sqrt{2}$ (resp., $-eC\sqrt{2}$), and show that g_1 (resp., g_2) is a solution of Equation (15) in a neighborhood of eC.

b. Solve Equation (16) explicitly with respect to y (what is possible in the present example) and verify that Equation (15) actually admits two solutions in the neighborhood of eC.

4 Inverse function theorem

4.1 Statement and proof

The next proposition is well-known:

Proposition 1. *If $f : \Omega_1 \subset \mathbb{R}^p \to \Omega_2 \subset \mathbb{R}^q$, where Ω_1 and Ω_2 are open in $\mathbb{R}^p$ and $\mathbb{R}^q$ respectively, is a diffeomorphism of class C^1, then, for any $x_0 \in \Omega_1$, the derivative $f'(x_0) : \mathbb{R}^p \to \mathbb{R}^q$ is a vector space isomorphism, $p = q$, and*

$$(f^{-1})'(f(x_0)) = (f'(x_0))^{-1}.$$

Proof. Indeed, as $f \circ f^{-1} = \mathrm{id}_{\Omega_2}$ and $f^{-1} \circ f = \mathrm{id}_{\Omega_1}$, we have

$$f'(x_0) \circ (f^{-1})'(f(x_0)) = \mathrm{id}_{\mathbb{R}^q} \quad \text{and} \quad (f^{-1})'(f(x_0)) \circ f'(x_0) = \mathrm{id}_{\mathbb{R}^p} \ .$$

□

Conversely,

Theorem 4 (Inverse function theorem). *If $f : \Omega \subset \mathbb{R}^p \to \mathbb{R}^q$ is of class C^k, $k \geq 1$, and if $f'(x_0)$, $x_0 \in \Omega$, is a vector space isomorphism, then $p = q$, and there is an open subset $\omega \ni x_0$ and an open subset $\omega' \ni f(x_0)$, such that $f : \omega \to \omega'$ is a diffeomorphism of class C^k. Moreover, for any $y \in \omega'$,*

$$(f^{-1})'(y) = (f'(x))^{-1} \ , \tag{17}$$

where x is of course the point $x = f^{-1}(y)$ that corresponds to y.

Remark that it thus follows from the invertibility (nonsingularity) of the derivative $f'(x_0) = (\partial_x f)(x_0)$ of f at a point x_0, that the original function f is *locally* a diffeomorphism. This is actually the best possible result, since it is clear that the derivative of a function at a point encodes information about the behavior of the original function only in the neighborhood of the considered point.

Proof. The main challenge in the proof of the inverse function theorem (InFT) is actually to show that f is locally invertible, i.e. that the equation $y = f(x)$ can locally be solved with respect to x. It thus suffices to apply the ImFT to

$$\phi(x, y) = y - f(x) : \Omega \times \mathbb{R}^p \to \mathbb{R}^p \ ,$$

but be aware of the fact that, as we solve with respect to x, the roles of x and y are exchanged in comparison with previous applications of the ImFT.

Set $y_0 := f(x_0)$, and note that $\phi \in C^k(\Omega \times \mathbb{R}^p, \mathbb{R}^p)$, that

$$(\partial_x \phi)(x_0, y_0) = -(\partial_x f)(x_0) = -f'(x_0)$$

is invertible, and that $\phi(x_0, y_0) = 0$, so that the ImFT solves the equation $\phi(x, y) = 0$, i.e. $y = f(x)$, with respect to x: there exists an open neighborhood $\omega_1 \ni x_0$ in Ω, an open neighborhood $\omega_2 \ni f(x_0)$ in $\mathbb{R}^p$, as well as a function $g \in C^k(\omega_2, \omega_1)$ such that

$$x \in \omega_1, y \in \omega_2, y = f(x) \Leftrightarrow y \in \omega_2, x = g(y) \ ,$$

or, equivalently,

$$x \in \omega_1, f(x) \in \omega_2, y = f(x) \Leftrightarrow y \in \omega_2, x = g(y) \ ,$$

or, as well,

$$x \in \omega_1 \cap f^{-1}(\omega_2), y = f(x) \Leftrightarrow y \in \omega_2, x = g(y) \ .$$

Since f is in particular continuous, the intersection $\omega_1 \cap f^{-1}(\omega_2)$ is open in Ω, and the preceding conclusion, together with the facts that f and g are of class C^k, thus means that f is a diffeomorphism of class C^k from the open neighborhood $\omega := \omega_1 \cap f^{-1}(\omega_2)$ of x_0 onto the open neighborhood $\omega' := \omega_2$ of $f(x_0)$. As for Formula (17), it suffices to differentiate the equality $f(g(y)) = y$, $y \in \omega'$. Indeed, we then get

$$f'(g(y)) \circ g'(y) = \mathrm{id}_{\mathbb{R}^p} \ ,$$

so that, since $g = f^{-1}$, we obtain

$$(f^{-1})'(y) = (f'(f^{-1}(y)))^{-1} \ ,$$

for any $y \in \omega'$. $\square$

4.2 Exercise

Prove that if $f : \Omega \subset \mathbb{R}^p \to \mathbb{R}^p$ is of class C^1 and $f'(x)$ is, for any $x \in \Omega$, a vector space isomorphism (i.e. is bijective / invertible / nonsingular – since $f'(x)$ is a linear map by definition), then f is an open map, i.e. a map that sends open subsets of Ω to open subsets of $\mathbb{R}^p$.

Hint: To show that $f(O)$, O open in Ω, is open in $\mathbb{R}^p$, it suffices to prove that, for any $y_0 \in f(O)$, there is an open neighborhood of y_0 that is contained in $f(O)$. To see this, it suffices to apply the InFT to a preimage x_0 of y_0 and to f restricted to O.

5 Constant rank theorem

5.1 Subimmersions, immersions, submersions

We need some prerequisites.

Definition 1. *Let $f : \Omega \subset \mathbb{R}^p \to \mathbb{R}^q$ be of class C^1. For any $x_0 \in \Omega$, the rank $\rho_{x_0} f$ of f at x_0 is the rank $\rho(f'(x_0))$ of the linear map or matrix $f'(x_0)$.*

A well-known result of Linear Algebra states that, for any linear map $\ell \in \mathrm{Hom}_{\mathbb{F}}(V, V')$ between two vector spaces V and V' over a field $\mathbb{F}$, we have

$$\rho\,\ell := \dim \operatorname{im} \ell = \dim V - \dim \ker \ell, \tag{18}$$

where notation is self-explaining.

Exercise. Prove that if $\ell \in \mathrm{Hom}_{\mathbb{F}}(V, V')$ and $\ell' \in \mathrm{Hom}_{\mathbb{F}}(V', V'')$, then

$$\rho(\ell' \circ \ell) \leq \inf(\rho\ell, \rho\,\ell') \ .$$

Apply this result and show that the rank is invariant under diffeomorphisms, i.e. if Ω_1, Ω_2 (resp., Ω_3, Ω_4) are open in $\mathbb{R}^p$ (resp., $\mathbb{R}^q$), if $\phi : \Omega_1 \to \Omega_2$ and $\phi' : \Omega_3 \to \Omega_4$ are diffeomorphisms of class C^1, and if $f : \Omega_2 \to \Omega_3$ is of class C^1, then for any $x \in \Omega_1$,

$$\rho_x(\phi' \circ f \circ \phi) = \rho_{\phi(x)} f. \tag{19}$$

Definition 2. *Let $f : \Omega \subset \mathbb{R}^p \to \mathbb{R}^q$ be of class C^1. Then f is an immersion (resp., a submersion) at a point $x_0 \in \Omega$, if its derivative $f'(x_0)$ at x_0 is injective (resp., surjective), i.e. if $\rho_{x_0} f$ coincides with the dimension of the source space (so that $p \leq q$) (resp., the dimension of the target space (so that $q \leq p$)). Further we say that f is an immersion (resp., a submersion), if it is an immersion (resp., a submersion) at any point $x_0 \in \Omega$.*

Exercise. Prove that

$$f : \mathbb{R}^\times \times \mathbb{R} \ni (\rho, \theta) \mapsto (\rho \cos \theta, \rho \sin \theta) \in \mathbb{R}^2$$

is an immersion and a submersion. Conclude that f is locally a diffeomorphism (recall also the open subsets of $\mathbb{R}^2$ between which the transition f from polar to cartesian coordinates is usually considered).

Remark that, the rank is lower semi-continuous, i.e. that, in a neighborhood of any point, it cannot decrease. It follows that the rank of an immersion or a submersion at x_0 is locally constant at x_0, i.e. constant in a neighborhood of x_0.

Definition 3. *A map $f : \Omega \subset \mathbb{R}^p \to \mathbb{R}^q$ of class C^1 is a subimmersion at a point $x_0 \in \Omega$ (resp., subimmersion), if its rank is locally constant at x_0 (resp., locally constant).*

Thus, immersions and submersions are special subimmersions.

5.2 Statement and proof

We are now prepared to state the constant rank theorem. It claims that any subimmersion f, i.e. any function of locally constant rank ρ, has locally, up to diffeomorphisms of the source and the target, the very simple canonical form

$$f(x^1, \ldots, x^p) = (x^1, \ldots, x^\rho, 0, \ldots, 0) \; .$$

Theorem 5. *Let $f : \Omega \subset \mathbb{R}^p \to \mathbb{R}^q$ be a subimmersion of class C^k, $k \geq 1$, at $x_0 \in \Omega$. Then, there are open subsets*

$$U \ni x_0, U' \supset f(U) \ni f(x_0), \omega \subset \mathbb{R}^p, \omega' \subset \mathbb{R}^q \,,$$

and diffeomorphisms $\phi : U \to \omega$ and $\phi' : U' \to \omega'$ of class C^k, such that, for all $x := (x^1, \ldots, x^p) \in \omega$,

$$\left(\phi' \circ f \circ \phi^{-1}\right)(x^1, \ldots, x^p) = (x^1, \ldots, x^\rho, 0, \ldots, 0), \tag{20}$$

where $\rho = \rho_{x_0} f$. Further,

$$\left(\phi' \circ f \circ \phi^{-1}\right)(\omega) = \{y \in \omega' : y^{\rho+1} = \ldots = y^q = 0\}. \tag{21}$$

Observe that the requirement that f must be a subimmersion at x_0 is necessary in view of the diffeomorphism invariance of the rank.

Let us also mention that if the subimmersion f is an immersion (resp., a submersion) at x_0, the rank $\rho_{x_0} f$ coincides with the dimension of the source space, what entails that $p \leq q$ (resp., target space, what entails that $q \leq p$), and the source (resp., target) diffeomorphism can be suppressed. Hence, the local canonical form of an immersion is

$$(\phi' \circ f)\,(x^1, \ldots, x^p) = (x^1, \ldots, x^p, 0, \ldots, 0) \in \mathbb{R}^q, \tag{22}$$

and the local canonical form of a submersion is

$$\left(f \circ \phi^{-1}\right)(x^1, \ldots, x^p) = (x^1, \ldots, x^q) \in \mathbb{R}^q. \tag{23}$$

Observe that in Equation (22) (resp., Equation (23)), the function "f" is linear and injective (resp., linear and surjective). As $f'(x_0)$ is also a linear injection (resp., a linear surjection), these results show that, up to a diffeomorphism, the behavior of f in the neighborhood of x_0 is the same as that of its derivative $f'(x_0)$ at x_0.

Proof. Since $\rho_{x_0} f = \rho((\partial_x f)(x_0)) = \rho$, the Jacobian matrix $(\partial_x f)(x_0)$ contains a nonvanishing subdeterminant of dimension $\rho \times \rho$. We can assume without loss of generality that this subdeterminant occupies the left top corner (otherwise it suffices to exchange the variables). It is therefore natural to introduce the following notation:

$$\mathbb{R}^p \ni x = (x', x'') \in \mathbb{R}^\rho \times \mathbb{R}^{p-\rho},\ \mathbb{R}^q \ni y = (y', y'') \in \mathbb{R}^\rho \times \mathbb{R}^{q-\rho}, \text{ and}$$

$$\mathbb{R}^q \ni f(x) = (f'(x', x''), f''(x', x'')) \in \mathbb{R}^\rho \times \mathbb{R}^{q-\rho}\ .$$

Now

$$\det((\partial_{x'} f')(x_0)) \neq 0$$

(let us emphasize that here f' is not the derivative of f).

Set now

$$\phi(x', x'') = (f'(x', x''), x'')\ . \tag{24}$$

The function ϕ is actually a diffeomorphism of class C^k from an open neighborhood U of x_0 to an open neighborhood ω of $\phi(x_0)$. In view of the inverse function theorem, it suffices to check that $(\partial_x \phi)(x_0)$ is nonsingular. This requirement is obviously satisfied, since

$$(\partial_x \phi)(x_0) = \begin{pmatrix} (\partial_{x'} f')(x_0) & (\partial_{x''} f')(x_0) \\ 0 & \mathrm{id} \end{pmatrix}\ .$$

We now have

$$(f \circ \phi^{-1})(f'(x', x''), x'') = (f'(x', x''), f''(x', x''))\ ,$$

so that, when denoting the new coordinates $(f'(x', x''), x'')$ by (u', u''), we get, for any $u = (u', u'') \in \omega$,

$$(f \circ \phi^{-1})(u', u'') = (u', g(u', u'')) \; , \tag{25}$$

where

$$g(u', u'') = f''(\phi^{-1}(u', u'')) \in \mathbb{R}^{q-\rho} \; .$$

We have still to use the assumption that $\rho_x f = \rho$, for any $x \in U$ (modulo a possible replacement of U by a smaller U). Since the rank is invariant under diffeomorphism, we have as well

$$\rho = \rho_{(u',u'')}(f \circ \phi^{-1}) = \rho(\partial_{(u',u'')}(f \circ \phi^{-1})) = \rho \begin{pmatrix} \mathrm{id}_{\rho\times\rho} & 0 \\ \partial_{u'} g & \partial_{u''} g \end{pmatrix} \; ,$$

for any $(u', u'') \in \omega$. It follows that $\partial_{u''} g = 0$, for all $u = (u', u'') \in \omega$. Indeed, if at some $u \in \omega$, one of the entries of $\partial_{u''} g$, say $\partial_{u''^j} g^i$, does not vanish, then the subdeterminant

$$\begin{pmatrix} \mathrm{id}_{\rho\times\rho} & 0 \\ * & \partial_{u''^j} g^i \end{pmatrix}$$

of dimension $(\rho+1)\times(\rho+1)$ is nonzero. Since this is a contradiction, $\partial_{u''} g = 0$ at all the points $u \in \omega$. It follows that g is independent of u''. Here we actually apply the well-known result stating that, if all the derivatives $\partial_{x^i} f$, $i \in \{1, \ldots, p\}$, of a differentiable function f vanish in a connected open subset $O \subset \mathbb{R}^p$, then this function f is constant in O. To exclude all problems related to connectedness, we observe that $\omega \ni \phi(x_0)$ contains a neighborhood of $\phi(x_0)$ of the type $\omega' \times \omega''$, $\omega' \subset \mathbb{R}^\rho$, $\omega'' \subset \mathbb{R}^{p-\rho}$ connected, and we replace ω by the smaller $\omega := \omega' \times \omega''$ and U by the smaller $U := \phi^{-1}(\omega' \times \omega'')$. Now, for any fixed $u' \in \omega'$, all the derivatives $\partial_{u''^j} g^i$ of g^i vanish in the connected ω'', so that g^i is constant in ω''; eventually g is independent of u'' in $\omega = \omega' \times \omega''$, i.e. $g = g(u')$, $u' \in \omega' \subset \mathbb{R}^\rho$.

Set now

$$\psi(y', y'') = (y', y'' - g(y')) \; . \tag{26}$$

Observe that $\psi \in C^k(\omega' \times \mathbb{R}^{q-\rho}, \mathbb{R}^q)$ and that $\partial_y \psi$ is nonsingular at any $y = (y', y'') \in \omega' \times \mathbb{R}^{q-\rho}$. Since ψ is thus an open map, it is obvious that ψ is a diffeomorphism of class C^k from $\omega' \times \mathbb{R}^{q-\rho}$ onto $\psi(\omega' \times \mathbb{R}^{q-\rho})$ (its inverse is given by $\psi^{-1}(v', v'') = (v', v'' + g(v'))$). It finally follows from (25) that, for any $u \in \omega = \omega' \times \omega''$,

$$(\psi \circ f \circ \phi^{-1})(u', u'') = (u', 0) \; . \tag{27}$$

This completes the proof of the first part of the constant rank theorem. As for the second part, the inclusion $\subset$ is obvious. To find the inclusion $\supset$ it suffices to chase through the main equations of the proof (and to note that in the proof notation is different from that in the statement of the theorem). $\square$

The proof of the fact that, in case f is an immersion (resp., a submersion), the source (resp., target) diffeomorphism can be avoided (see remark below the constant rank theorem), is not instructive and will not be given here.

Chapter 2
Topological spaces

We assume that the reader is already familiar with topological spaces. Since manifolds are specific topological spaces, the most important definitions are recalled below, also to ensure independent readability of the present text.

1 Sets

In the following, we need some set theoretical concepts.

The basic operations on sets, the union and the intersection, can be extended in an obvious way to families $(S_i)_i$ of subsets of a given set S. These generalizations have similar properties than the underlying usual operations. For instance, the union (resp. the intersection) is distributive with respect to the intersection (resp. the union), and the complement of a union (resp. an intersection) coincides with the intersection (resp. the union) of the complements. Nevertheless the empty family of subsets of S deserves some attention, as the union (resp. the intersection) of the empty family $(S_i)_{i\in\emptyset}$ is the empty subset (resp. the total set S).

Further, the properties of images and preimages also apply to these extensions. Let us recall that if f denotes a map $f : S \to S'$ from a set S to a set S', and if $X, X' \subset S$ and $Y, Y' \subset S'$, we have

$$f^{-1}(Y \cup Y') = f^{-1}(Y) \cup f^{-1}(Y'), f^{-1}(Y \cap Y') = f^{-1}(Y) \cap f^{-1}(Y'), f^{-1}(S' \setminus Y) = S \setminus f^{-1}(Y) \quad (1)$$

and

$$f(X \cup X') = f(X) \cup f(X'), \quad f(X \cap X') \subset f(X) \cap f(X'), \quad f(S \setminus X) \supset f(S) \setminus f(X). \quad (2)$$

Moreover,

$$f^{-1}(f(X)) \supset X \quad \text{and} \quad f(f^{-1}(Y)) \subset Y. \quad (3)$$

Of course, if f is a bijection, we get

$$f(X \cap X') = f(X) \cap f(X') \quad \text{and} \quad f(S \setminus X) = S' \setminus f(X),$$

since $(f^{-1})^{-1} = f$, as well as

$$f^{-1}(f(X)) = X \quad \text{and} \quad f(f^{-1}(Y)) = Y\,.$$

2 Exercise

Prove the preceding results.

3 Topological Spaces

Often problems and their solutions do not depend on the exact shape of an involved object, but on the way the object is composed, that is, on its topological structure. For example, the square and the circle have the same topological structure, while a circle and a line have a different topological structure. Similarly, a strip whose ends have been glued together and the Möbius strip where the ends have been glued together only after half a turn are not topologically equivalent. Intuitively, two "spaces" have the same topological structure or are topologically equivalent if one can be deformed into the other without cutting or gluing. Therefore, even a coffee mug and a donut are equivalent from the point of view of topology, since a sufficiently flexible donut can be reshaped into a coffee mug by forming a handle from one of its halves and a cylinder from the other, and then making an indentation in the cylinder.

The mathematical concept that best abstracts the idea of a reversible deformation that does not involve cutting or gluing is – as is easy to understand – a continuous bijective mapping whose inverse is also continuous. Such a mapping is called a homeomorphism. Since the natural generalization of the ordinary idea of a continuous function is the requirement that the preimage of an open subset is also open, we must first define open subsets. In the Cartesian space $\mathbb{R}^n$, the standard definition is that a subset $\Omega \subset \mathbb{R}^n$ is open if every point $x \in \Omega$ is the center of a sphere that is contained in Ω. These subsets of $\mathbb{R}^n$ have basic properties: every union (resp. every finite intersection) of open subsets is again an open subset. These minimal requirements lead to the definition of open subsets or – better – a topology in more abstract spaces than $\mathbb{R}^n$.

Definition 1. *A topological space is a set X together with a collection T of subsets of X that satisfies the following axioms:*

- (O_1) *The total set X is an element of T.*
- (O_2) *The union of every collection of sets in T is also in T.*
- (O_3) *The intersection of two sets in T is in T.*

The collection T is called a topology on X and the elements of X are called points. The sets in T are referred to as open sets and their complements in X are the closed sets. Let us mention that the union of the empty collection is the empty set and the intersection of the empty collection is X.

Remarks.

- Axioms (O_1) and (O_3) can be replaced by a unique axiom, say (O_1'), which asks that every finite intersection of elements of T be again an element of T.
- There is a dual definition of a topological space based upon the fundamental properties of closed subsets.

- In the following a topological space will be denoted by (X, T) or simply by X, if no confusion regarding the considered topology is possible.

Examples.

- The discrete topology of a set X is the topology in which every subset of X is open.
- The trivial topology of X is the topology in which only the empty set and the whole space X are open.
- The standard topology of $\mathbb{R}$ is made of the unions of the bounded open intervals. If the unions of the subsets of a collection satisfy the defining axioms $(\mathrm{O}_1) - (\mathrm{O}_3)$ of a topology, we call these subsets, a basis of the topology.
- Every metric space can be given the metric topology, in which the basic open sets are the open balls defined by the metric.
- The Zariski topology of $\mathbb{R}^2$ is defined by means of its closed subsets, which are just the plane $\mathbb{R}^2$ itself, every algebraic curve $p(x, y) = 0$, where p denotes a polynomial, every point, as well as all finite unions of such subsets. This topology can be extended to $\mathbb{R}^n$ and to more general spaces, and is of special importance in Algebraic Geometry.

Definition 2. *A map $f : X \to Y$ from a topological space X to another topological space Y is continuous if, for every open subset V of Y, the preimage*

$$f^{-1}(V) := \{x \in X : f(x) \in V\}$$

is an open subset of X. The set of continuous maps between the topological spaces X and Y is denoted by $C^0(X, Y)$. A continuous map $\phi : X \to Y$ is a homeomorphism if it is bijective and its inverse is continuous as well. Two topological spaces related by a homeomorphism are topologically equivalent.

Often axioms $(\mathrm{O}_1) - (\mathrm{O}_3)$ are too weak to allow efficient investigation of a given problem. We then add additional requirements, such as e.g. the condition that the topology must admit a countable basis of open subsets, see above (let us recall that a set E is countable if there is a bijection $\phi : E \to P$, where $P \subset \mathbb{N}$). A topological space with a countable basis of open subsets is said to be second countable. Another type of frequently used restrictions are separation axioms. They allow distinguishing e.g. distinct points by topological means, in particular separating points by neighborhoods.

Definition 3. *In a topological space X, a neighborhood of a subset P is a subset N, such that $N \supset U \supset P$, where U is an open subset of X.*

In particular, any open subset that contains P is a neighborhood of P.

In the following, we essentially use Hausdorff's separation axiom (Felix Hausdorff, 1868 – 1942, German mathematician) and work in Hausdorff spaces, in which points can actually be separated by neighborhoods. This implies for instance uniqueness of limits of sequences.

Definition 4. *A topological space is a Hausdorff space if any two distinct points admit disjoint neighborhoods.*

Examples.

- Almost all spaces encountered in mathematical analysis are Hausdorff. Of course, all metric spaces, in particular $\mathbb{R}^n$, are Hausdorff spaces.

- But then we also understand that pseudometric spaces are typically not Hausdorff. Also the Zariski topology of $\mathbb{R}^2$, see above, does not satisfy Hausdorff's axiom. Indeed, two non empty open subsets are never disjoint, as the complement of their intersection, i.e. the union of their complements is, by definition of the topology, a finite union of algebraic curves and points.

4 Exercises

1. Let S be a subset of a topological space X. The collection of traces $S \cap U$ on S of the open subsets U of X is a topology on S. We refer to this topology as the induced or relative topology and to a subset endowed with the relative topology as a topological subspace. Prove that the topology axioms actually hold true.

2. Prove that any subspace of a Hausdorff (resp. second countable) space is itself Hausdorff (resp. second countable). We say that the Hausdorff property (resp. the existence of a countable basis) is hereditary.

3. Let B be a collection of subsets of a set X. Prove that X has a unique topology with basis B if and only if any finite intersection of elements of B is a union of elements of B, i.e. if and only if X is a union of elements of B and any intersection of two elements of B is a union of elements of B. *Hint*: Use the fact that the axioms $(O_1) - (O_3)$ are equivalent to $(O_1'), (O_2)$.

4. Let
$$(X_i, T_i)_i\,,\ i \in I := \{1, \ldots, n\}\,,\ n \in \mathbb{N}^*\,,$$
be a finite number of topological spaces. We denote by $X := \Pi_{i\in I} X_i$ the Cartesian product of the X_i. Prove that the subsets $O := \Pi_{i\in I} O_i\,,\ O_i \in T_i\,,$ form a basis of a topology T of X. In the following, we refer to (X, T) as the product space and to T as the product topology. *Hint*: It suffices to observe that
$$(\Pi_i O_i) \cap (\Pi_j \Omega_j) = \Pi_k (O_k \cap \Omega_k)\,.$$

5. Prove that if $f \in C^0(X, Y)$ and $g \in C^0(Y, Z)$, then $g \circ f \in C^0(X, Z)$, where X, Y, Z denote of course topological spaces.

6. Let X and Y be topological spaces, let $f : X \to Y$, and let S (resp. S') be a topological subspace of X (resp. of X', such that $f(X) \subset S'$). Prove that, if $f \in C^0(X, Y)$, then the restriction of f to S is $f|_S \in C^0(S, Y)$
$$(\text{resp. } f \in C^0(X, Y) \Leftrightarrow f \in C^0(X, S'))\,.$$

Chapter 3
Manifolds

1 Smooth Manifolds

Manifolds are higher dimensional (dimension $n \geq 0$) analogs of curves ($n = 1$) and surfaces ($n = 2$). They are of importance in most branches of Mathematics and numerous areas in Theoretical Physics, e.g. in Mechanics, General Relativity, String Theory, ...

We often tend to think of manifolds as "surfaces" that are embedded in a Cartesian space and more precisely as embedded submanifolds of $\mathbb{R}^n$. However, manifolds are in fact more abstract objects. For instance, space-time is a 4-dimensional (pseudo-Riemannian) manifold that exists without living in a bigger space.

In order to understand the definition of a manifold, consider the 2-dimensional sphere $S^2 \subset \mathbb{R}^3$, say the surface of the Earth. A subset $U_\alpha \subset S^2$ can be represented by a chart, mathematically a bijection $\varphi_\alpha : U_\alpha \to \varphi_\alpha(U_\alpha) \subset \mathbb{R}^2$. Hence, a sphere or—more generally—a manifold looks locally like (a subset of) a Cartesian space, but its global structure is more complicated.

In order to represent it completely, we need a family of charts, i.e. an atlas $(U_\alpha, \varphi_\alpha)_\alpha$. Obviously, the region $U_\alpha \cap U_\beta \subset S^2$ can be mapped to $\mathbb{R}^2$ by both charts,

$$\varphi_\alpha : U_\alpha \to \varphi_\alpha(U_\alpha) \subset \mathbb{R}^2 \quad \text{and} \quad \varphi_\beta : U_\beta \to \varphi_\beta(U_\beta) \subset \mathbb{R}^2.$$

The map

$$\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\beta \cap U_\alpha)$$

is the transition map from chart $(U_\alpha, \varphi_\alpha)$ to chart (U_β, φ_β). Transition maps encode the information how the manifold can be reconstructed from its parts by gluing them together.

Further, it is natural to think that in the case of the sphere or even of a more general "smooth manifold" the chart maps $\varphi_\alpha : U_\alpha \to \varphi(U_\alpha) \subset \mathbb{R}^2$ should be "smooth" bijections with "smooth" inverse, i.e. diffeomorphisms, so that the transition maps

$$\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\beta \cap U_\alpha)$$

are diffeomorphisms as well. However, as $U_\alpha \subset S^2$ is not an open subset of a *Cartesian space*, smoothness of $\varphi_\alpha : U_\alpha \to \varphi_\alpha(U_\alpha)$ has no meaning so far, even if we assume that $\varphi_\alpha(U_\alpha) \subset \mathbb{R}^2$ is open. On the other hand, smoothness of

$$\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\beta \cap U_\alpha)\,,$$

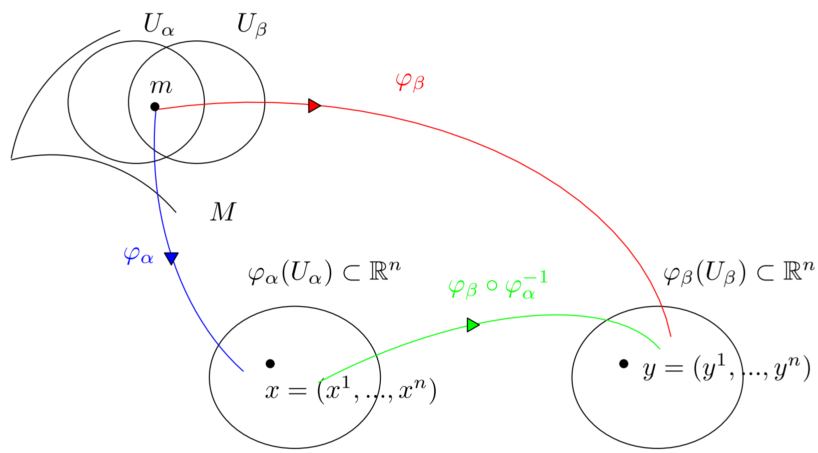

Figure 1: Manifold and charts

which is a map between two subsets of $\mathbb{R}^2$, is a well-known concept, at least if we assume that $\varphi_\alpha(U_\alpha \cap U_\beta) \subset \mathbb{R}^2$ and $\varphi_\beta(U_\beta \cap U_\alpha) \subset \mathbb{R}^2$ are open subsets.

Hence, the following definitions.

Definition 1. *A chart of a set M is a pair (U, φ), where $\varphi : U \to \varphi(U) \subset \mathbb{R}^n$ is a bijection from a subset $U \subset M$ onto an open subset $\varphi(U) \subset \mathbb{R}^n$. The components of the image*

$$\varphi(m) = (\varphi^1(m), \ldots, \varphi^n(m)) =: x = (x^1, \ldots, x^n)$$

of a point $m \in U$ are the coordinates of m in the considered chart or coordinate system.

Definition 2. *A smooth n-dimensional atlas $(n \in \mathbb{N})$ of a set M is a collection of charts $(U_\alpha, \varphi_\alpha)_\alpha$ of M, such that*

- *the U_α cover M,*
- *the images $\varphi_\alpha(U_\alpha \cap U_\beta)$ are open in $\mathbb{R}^n$, and*
- *the transition or coordinate transformation maps*

  $$\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \mapsto \varphi_\beta(U_\beta \cap U_\alpha)$$

  are smooth.

Examples.

1. Every n-dimensional real vector space V admits a smooth n-dimensional atlas. It suffices to choose a basis $(e_i)_i$ and to consider the isomorphism

$$\varphi : V \ni v = \sum_i x^i e_i \mapsto x := (x^1, \ldots, x^n) \in \mathbb{R}^n \,.$$

Observe that this atlas is made of a unique chart. In particular, for $V = \mathbb{R}^n$, we may take the global chart $(\mathbb{R}^n, \mathrm{id})$. Also any open subset $\Omega \subset \mathbb{R}^n$ admits such an atlas (Ω, id).

2. The sphere
$$S^{n-1} = \{x \in \mathbb{R}^n : \sum_i (x^i)^2 = 1\} \subset \mathbb{R}^n$$
has a smooth $(n-1)$-dimensional atlas. In order to simplify notations, consider the case $n = 3$. We set $x := x^1\,, y := x^2\; z = x^3$. The atlas is made *for instance* of the six charts induced by the two projections $\varphi_{1\pm} : (x, y, z) \to (y, z)$ onto the hyperplane $x = 0$ of the two hemispheres without boundary $H_{1\pm}$ defined by this hyperplane, and the similar projections $\varphi_{2\pm}$ and $\varphi_{3\pm}$. Indeed, the $(H_{i\pm}, \varphi_{i\pm})_i$ are charts the domains of which cover S^2. Moreover, the image $\varphi_{2+}(H_{2+} \cap H_{1-})$ for instance, is the set
$$\Omega_{2^+1^-} = \{(x, z) : x^2 + z^2 < 1, x < 0\}\,,$$
which is open in $\mathbb{R}^2$, and
$$\varphi_{1^-}(\varphi_{2^+}^{-1}(x, z)) = (\sqrt{1 - x^2 - z^2}, z)$$
is a smooth bijection between the open subsets $\Omega_{2^+1^-}$ and $\Omega_{1^-2^+}$, the inverse of which is smooth as well.

3. The extended complex plane $M = \mathbb{C} \cup \{\infty\}$ admits an atlas with two charts, $U_1 = \mathbb{C}$, $\varphi_1(z) = (x, y) \in \mathbb{R}^2$ and $U_2 = (\mathbb{C} \setminus \{0\}) \cup \{\infty\}$, $\varphi_2(z) = 1/z$, where it is understood that $1/z$ is viewed as an element of $\mathbb{R}^2$ and that $1/\infty = 0$. Further, the image $\varphi_1(U_1 \cap U_2)$ for instance, is $\mathbb{R}^2 \setminus \{0\}$ and the transition map
$$\varphi_2(\varphi_1^{-1}(x, y)) = (\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2})$$
is a smooth permutation of $\mathbb{R}^2 \setminus \{0\}$, the inverse of which is also smooth.

4. The n-dimensional real projective space $\mathbb{R}P^n$ has an n-dimensional smooth atlas that contains $n + 1$ charts. This case will be detailed in the lectures.

If we add a geographic chart to a geographic atlas, we get of course another, maybe even better, atlas. The union of a mathematical chart (U, φ) of a set M and a mathematical atlas $(U_\alpha, \varphi_\alpha)_\alpha$ of M is again an atlas if and only if the $\varphi(U \cap U_\alpha)$ and the $\varphi_\alpha(U_\alpha \cap U)$ are open subsets of $\mathbb{R}^n$ and the transition maps $\varphi \circ \varphi_\alpha^{-1}$ and $\varphi_\alpha \circ \varphi^{-1}$ are smooth on their domains.

Definition 3. *A chart and an atlas of a set M (resp. two atlases of M) are compatible if their union is an atlas of M.*

On the one hand, compatibility of atlases is obviously an equivalence relation. On the other, we implicitly think of an “n-dimensional smooth manifold” as a set endowed with an n-dimensional smooth atlas. However, the “manifold” is completely and equivalently represented by each atlas of a same equivalence class. Hence, the following definition.

Definition 4. *An n-dimensional smooth manifold is a set together with an equivalence class of n-dimensional smooth atlases.*

To define a smooth manifold structure on a set M, we consequently only need to specify *one* atlas of M, which then in turn defines an equivalence class. The preceding definition just means that the same smooth structure on M can be defined by many atlases. For instance, the global charts (V, φ) and (V, φ') of a finite-dimensional real vector space V induced by two bases $(e_i)_i$ and $(e'_i)_i$ form two equivalent atlases, which therefore define the same smooth manifold structure on V. Indeed, the transition map $\varphi' \circ \varphi^{-1}$ maps the coordinates x of a vector v in the basis $(e_i)_i$ onto the coordinates $x' = A^{-1}x$ of v in $(e'_i)_i$, where A is the transition matrix from the first to the second basis.

Let us explicitly mention that all manifolds considered below are smooth and finite-dimensional. Further, instead of manifold structure, we also use the terminology smooth structure and differentiable structure.

We already mentioned that the coordinate maps

$$\varphi_\alpha : U_\alpha \ni m \mapsto (\varphi^1_\alpha(m), \ldots, \varphi^n_\alpha(m)) = (x^1, \ldots, x^n) \in \varphi_\alpha(U_\alpha)$$

should be diffeomorphisms between the manifolds U_α and $\varphi_\alpha(U_\alpha)$, but that the concept of smooth map between manifolds is not yet defined. Actually, the map φ_α can even not yet be a homeomorphism, as so far we have no topology on U_α or M. We now show that an atlas of M defines a topology on M that only depends on the smooth structure of M, i.e. on the considered equivalence class of atlases.

Definition 5. *Let M be a manifold and let $(U_\alpha, \varphi_\alpha)_\alpha$ be an atlas of M. A subset $W \subset M$ is an open subset of M, if all its representations $\varphi_\alpha(W \cap U_\alpha) \subset \mathbb{R}^n$ are open in the metric topology of the Cartesian space $\mathbb{R}^n$.*

Theorem 1. *The collection of open subsets of a manifold M, defined in Definition 5 by means of an atlas of M, is a topology on M that is independent of the considered atlas.*

Proof. Let $(W_i)_i$ be a family of open subsets of M. Due to the properties of images, we have

$$\varphi_\alpha(\cup_i W_i \ \cap U_\alpha) = \cup_i \, \varphi_\alpha(W_i \cap U_\alpha)$$

and

$$\varphi_\alpha(\cap_i W_i \ \cap U_\alpha) = \cap_i \, \varphi_\alpha(W_i \cap U_\alpha).$$

The conclusion regarding the topological structure follows if we interpret the family of W_i as finite in the second case.

As for the independence of the atlas chosen in the equivalence class of the manifold, it suffices to show that any open subsets of the topology T_2, defined by another compatible atlas $(V_a, \psi_a)_a$ of M, is an open subset in the topology T_1 defined by $(U_\alpha, \varphi_\alpha)_\alpha$ and vice versa. Let us for instance prove that $W \in T_1$ is open in T_2, i.e. that $\psi_a(W \cap V_a)$ is open in $\mathbb{R}^n$.

Any $y \in \psi_a(W \cap V_a)$ is the image $\psi_a(x)$ of some $x \in W \cap V_a$. As we have for sure to use the compatibility of the two atlases, observe that there is U_β that contains x, so that

$$y \in \psi_a(W \cap U_\beta \cap V_a) \subset \psi_a(W \cap V_a)\,.$$

If we prove now that $(U, \varphi) := (W \cap U_\beta, \varphi_\beta|_{W \cap U_\beta})$ is a chart that is compatible with the "corresponding" atlas $(U_\alpha, \varphi_\alpha)_\alpha$, then, since compatibility is an equivalence, it is also compatible with the atlas $(V_a, \psi_a)_a$, and therefore $\psi_a(U \cap V_a) = \psi_a(W \cap U_\beta \cap V_a)$ is open in $\mathbb{R}^n$, which completes the proof.

Note first that $(U, \varphi) := (W \cap U_\beta, \varphi_\beta|_{W \cap U_\beta})$ is a chart since $W \in T_1$. As for the compatibility, we have to show that $\varphi_\alpha(U_\alpha \cap U)$ and $\varphi(U \cap U_\alpha)$ are open in $\mathbb{R}^n$ and that $\varphi \circ \varphi_\alpha^{-1}$ and $\varphi_\alpha \circ \varphi^{-1}$ are smooth on their domains. We prove the last pair of assertions. As

$$\varphi(U \cap U_\alpha) = \varphi_\beta(W \cap U_\beta \cap U_\alpha) = \varphi_\beta(W \cap U_\beta) \cap \varphi_\beta(U_\alpha \cap U_\beta),$$

where the first (resp. second) image is open due to the fact that $W \in T_1$ (resp. the definition of an atlas), the first statement follows. The second is also clear, because $\varphi_\alpha \circ \varphi^{-1}$ is the restriction of $\varphi_\alpha \circ \varphi_\beta^{-1}$ to the preceding open subset. □

We are now prepared to prove the expected

Theorem 2. *Let M be a manifold and let $(U_\alpha, \varphi_\alpha)_\alpha$ be an atlas of M. Every coordinate map $\varphi_\alpha : U_\alpha \to \varphi_\alpha(U_\alpha)$ is a homeomorphism (it is of course understood that U_α and $\varphi_\alpha(U_\alpha)$ carry the topologies that are induced by those of M and $\mathbb{R}^n$ respectively).*

Proof. First consider an open subset of a topological space and provide it with the induced topology. Then the open subsets of this topological subspace are exactly the open subsets of the whole topological space that are included in the subspace.

Let W be an open subset of U_α. As

$$(\varphi_\alpha^{-1})^{-1}(W) = \varphi_\alpha(W) = \varphi_\alpha(W \cap U_\alpha) \subset \varphi_\alpha(U_\alpha)$$

is open, the map φ_α^{-1} is continuous.

Take now an open subset Ω of $\varphi_\alpha(U_\alpha)$. In order to see that $\varphi_\alpha^{-1}(\Omega) \subset U_\alpha$ is open, we have to show that $\varphi_\beta(\varphi_\alpha^{-1}(\Omega) \cap U_\beta)$ is open in $\mathbb{R}^n$. Since we must of course use the transition diffeomorphism, it is natural to write

$$\varphi_\beta(\varphi_\alpha^{-1}(\Omega) \cap U_\beta) = \varphi_\beta(\varphi_\alpha^{-1}(\Omega) \cap U_\alpha \cap U_\beta)$$

$$= \varphi_\beta(\varphi_\alpha^{-1}(\Omega) \cap \varphi_\alpha^{-1}\varphi_\alpha(U_\alpha \cap U_\beta)) = \varphi_\beta\varphi_\alpha^{-1}(\Omega \cap \varphi_\alpha(U_\alpha \cap U_\beta)).$$

The conclusion that φ_α is continuous now follows from the fact that the transition diffeomorphism is also a homeomorphism. □

2 Exercises

1. Prove that every open subset U of every manifold M is a manifold, and more precisely that, if $(U_\alpha, \varphi_\alpha)_\alpha$ is an atlas of M, the restrictions $(U \cap U_\alpha, \varphi_\alpha|_{U \cap U_\alpha})_\alpha$ form an atlas of U.

2. Show that if M (resp. M') is a manifold of dimension n (resp. n') with atlas $(U_\alpha, \varphi_\alpha)_\alpha$ (resp. $(U'_\beta, \varphi'_\beta)_\beta$), then $M \times M'$ is an $(n + n')$-dimensional manifold for the atlas

$$(U_\alpha \times U'_\beta, \varphi_\alpha \times \varphi'_\beta)_{(\alpha,\beta)}\,.$$

   We refer to $M \times M'$ equipped with this differential structure as the product manifold $M \times M'$.

3. The torus $T^2 := S^1 \times S^1$, where $S^1 \subset \mathbb{R}^2$ is the circle with center $(0,0)$ and radius 1, is a manifold of dimension 2, and, more generally, the "n-dimensional" torus $T^n := (S^1)^{\times n}$ is an n-dimensional smooth manifold.

4. The manifold topology can be defined just as well as the topology for which the chart domains are a basis. Prove that the unions of the chart domains actually form a topology and that this topology coincides with the above-defined manifold topology.

3 Submanifolds of $\mathbb{R}^n$

Submanifolds of $\mathbb{R}^n$ are an important class of manifolds. Actually every abstract manifold can be viewed as submanifold of a Cartesian space with sufficiently high dimension.

The next subsection is a summary of the results now needed from Chapter 1, which we recall here for simplicity.

3.1 Subimmersions, immersions, submersions, diffeomorphisms

Definition 6. *A map $f : \Omega \subset \mathbb{R}^p \to \mathbb{R}^q$ from an open subset Ω of the Cartesian space $\mathbb{R}^p$ to the Cartesian space $\mathbb{R}^q$ is smooth if and only if its canonical coordinate functions $f^i : \Omega \subset \mathbb{R}^p \to \mathbb{R}$, $i \in \{1, \ldots, q\}$, are all smooth in Ω. The derivative of f at $x_0 \in \Omega$ is the linear map*

$$f'(x_0) = \left(\left(\partial_{x^j} f^i\right)(x_0)\right)_{ij} : \mathbb{R}^p \to \mathbb{R}^q. \tag{1}$$

Definition 7. *Let $f : \Omega \subset \mathbb{R}^p \to \mathbb{R}^q$ be a smooth map. For every $x_0 \in \Omega$, the rank $\rho_{x_0} f$ of f at x_0 is the rank $\rho(f'(x_0))$ of the linear map or matrix $f'(x_0)$.*

A well-known result from Linear Algebra states that, for every linear map $\ell \in \mathrm{Hom}_{\mathbb{F}}(V, V')$ between two vector spaces V and V' over a field $\mathbb{F}$, we have

$$\rho\,\ell := \dim \mathrm{im}\,\ell = \dim V - \dim \ker \ell, \tag{2}$$

where notations are self-explaining.

Exercise. Prove that if $\ell \in \mathrm{Hom}_{\mathbb{F}}(V, V')$ and $\ell' \in \mathrm{Hom}_{\mathbb{F}}(V', V'')$, then $\rho(\ell' \circ \ell) \le \inf(\rho\ell, \rho\,\ell')$. Apply this result and show that the rank is invariant under diffeomorphisms, i.e. if Ω_1, Ω_2 (resp. Ω_3, Ω_4) are open in $\mathbb{R}^p$ (resp. $\mathbb{R}^q$), if $\varphi : \Omega_1 \to \Omega_2$ and $\varphi' : \Omega_3 \to \Omega_4$ are diffeomorphisms, and if $f : \Omega_2 \to \Omega_3$ is smooth, then for every $x \in \Omega_1$, we have

$$\rho_x(\varphi' \circ f \circ \varphi) = \rho_{\varphi(x)} f. \tag{3}$$

Definition 8. *Let $f : \Omega \subset \mathbb{R}^p \to \mathbb{R}^q$ be smooth. The map f is an immersion (resp. a submersion) at a point $x_0 \in \Omega$ if its derivative $f'(x_0)$ at x_0 is injective (resp. surjective), i.e. if $\rho_{x_0} f$ coincides with the dimension of the source space (so that $p \le q$) (resp. the dimension of the target space (so that $q \le p$)). The map f is an immersion (resp. a submersion), if it is an immersion (resp. a submersion) at every point $x_0 \in \Omega$.*

Notice that the rank is lower semi-continuous, i.e. that in a neighborhood of any point it cannot decrease. It follows that the rank of an immersion or a submersion is locally constant, i.e. constant in a neighborhood of any point.

Definition 9. *A smooth map $f : \Omega \subset \mathbb{R}^p \to \mathbb{R}^q$ is a subimmersion at a point $x_0 \in \Omega$ (resp. subimmersion) if its rank is locally constant at x_0 (resp. locally constant).*

Therefore immersions and submersions are special subimmersions. We are now prepared to recall the Constant Rank Theorem, see lectures in Analysis. It states that every subimmersion has locally, up to diffeomorphisms of the source and the target, a very simple canonical form.

Theorem 3. *Let $f : \Omega \subset \mathbb{R}^p \to \mathbb{R}^q$ be a subimmersion at $x_0 \in \Omega$. Then there are open subsets*

$$U \ni x_0, U' \supset f(U) \ni f(x_0), \omega \subset \mathbb{R}^p, \omega' \subset \mathbb{R}^q$$

and diffeomorphisms $\varphi : U \to \omega$ and $\varphi' : U' \to \omega'$, such that, for all $x := (x^1, \ldots, x^p) \in \omega$, we have

$$\left(\varphi' \circ f \circ \varphi^{-1}\right)(x^1, \ldots, x^p) = (x^1, \ldots, x^\rho, 0, \ldots, 0)\,, \tag{4}$$

where $\rho = \rho_{x_0} f$. Further

$$\left(\varphi' \circ f \circ \varphi^{-1}\right)(\omega) = \{y \in \omega' : y^{\rho+1} = \ldots = y^q = 0\}\,. \tag{5}$$

Note that the requirement that f be a subimmersion at x_0 is necessary because of the diffeomorphism invariance of the rank.

Remember that if the subimmersion f is an immersion (resp. a submersion) at x_0, the rank $\rho_{x_0} f$ coincides with the dimension of the source space, which entails that $p \le q$ (resp. of the target space, which entails that $q \le p$), and the source (resp. target) diffeomorphism can be suppressed. Hence, the local canonical form of an immersion is

$$(\varphi' \circ f)(x^1, \ldots, x^p) = (x^1, \ldots, x^p, 0, \ldots, 0) \in \mathbb{R}^q \tag{6}$$

and the local canonical form of a submersion is

$$\left(f \circ \varphi^{-1}\right)(x^1, \ldots, x^p) = (x^1, \ldots, x^q) \in \mathbb{R}^q. \tag{7}$$

Observe that in Equation (6) (resp. Equation (7)), the map "f" is linear and injective (resp. linear and surjective). As $f'(x_0)$ is also a linear injection (resp. a linear surjection), these results show that, up to a diffeomorphism, the behavior of f in the neighborhood of x_0 is the same as that of its derivative $f'(x_0)$ at x_0.

The previous results concern functions f the derivative $f'(x_0)$ of which is a linear injection or a linear surjection. The next theorem, called Inverse Function Theorem, asserts that functions f such that $f'(x_0)$ is a linear bijection, i.e. a vector space isomorphism, are locally at x_0 diffeomorphisms. More precisely, we have the

Theorem 4. *If $f : \Omega \subset \mathbb{R}^p \to \mathbb{R}^q$ is smooth and if $f'(x_0)$, $x_0 \in \Omega$, is a vector space isomorphism, then $p = q$ and there is an open subset $U \ni x_0$, such that $f(U)$ is open and $f : U \to f(U)$ is a diffeomorphism.*

The Inverse Function Theorem is actually a consequence of the Constant Rank Theorem.

Proof. Since f is both an immersion and a submersion at x_0, it follows from the above remarks that $p = q$. Moreover, Equation (5) implies that $f(U) = U'$, which is open, and Equation (4) entails that $\varphi' \circ f \circ \varphi^{-1} = \mathrm{id}_\omega$, so that $f = \varphi'^{-1} \circ \varphi$, which is a diffeomorphism between U and $f(U)$. □

The reverse result is also valid:

Proposition 1. *If $f : \Omega_1 \subset \mathbb{R}^p \to \Omega_2 \subset \mathbb{R}^q$, where Ω_1 and Ω_2 are open in $\mathbb{R}^p$ and $\mathbb{R}^q$ respectively, is a diffeomorphism, then, for every $x_0 \in \Omega_1$, the derivative $f'(x_0) : \mathbb{R}^p \to \mathbb{R}^q$ is a vector space isomorphism, $p = q$, and*

$$(f^{-1})'(f(x_0)) = (f'(x_0))^{-1}.$$

Proof. Indeed, as $f \circ f^{-1} = \mathrm{id}_{\Omega_2}$ and $f^{-1} \circ f = \mathrm{id}_{\Omega_1}$, we have

$$f'(x_0) \circ (f^{-1})'(f(x_0)) = \mathrm{id}_{\mathbb{R}^q} \quad \text{and} \quad (f^{-1})'(f(x_0)) \circ f'(x_0) = \mathrm{id}_{\mathbb{R}^p}\,.$$

□

3.2 Embedded submanifolds of $\mathbb{R}^n$

There exist several more or less restrictive concepts of submanifolds. Embedded submanifolds of $\mathbb{R}^n$ provide examples of (sub)manifolds that live in a Cartesian space $\mathbb{R}^n$ and carry the relative topology. Although we will not study submanifolds of an arbitrary manifold M, we would like to mention that such general submanifolds may be endowed with a richer topology than the one inherited from M and even need not always be subsets of M.

Take two positive integers $p \leq n$. The definition of an embedded p-dimensional submanifold N of $\mathbb{R}^n$ requires that locally $N \subset \mathbb{R}^n$ looks like $\mathbb{R}^p \times \{0\} \subset \mathbb{R}^n$.

Definition 10. *A subset $N \subset \mathbb{R}^n$ is an embedded p-dimensional submanifold of $\mathbb{R}^n$, if, for every $x \in N$, there is an open subset $U \ni x$ of $\mathbb{R}^n$ and a diffeomorphism $f : U \subset \mathbb{R}^n \to f(U) \subset \mathbb{R}^n$, $f(U)$ open, such that $f(U \cap N) = f(U) \cap (\mathbb{R}^p \times \{0\})$.*

To clarify the naturalness of the next theorem, we remind the reader that there are two natural approaches to "surfaces" in $\mathbb{R}^n$. For instance, in $\mathbb{R}^2$, the circle with center $(0, 0)$ and radius 1 can as well be described by the Cartesian equation $x^2 + y^2 = 1$ as by the parametric representation $(x, y) = (\cos t, \sin t)$, $t \in [0, 2\pi[$.

Let us briefly examine further examples. If we write $m = (x, y, z)$, the system of equations

$$f^1(m) = ax + by + cz + d = 0 \quad \text{and} \quad f^2(m) = a'x + b'y + c'z + d' = 0\,,$$

where $a, b, c, d, a', b', c', d' \in \mathbb{R}$, defines a line L in $\mathbb{R}^3$, if the derivatives $(f^1)'(m) \in \mathbb{R}^{3*}$ and $(f^2)'(m) \in \mathbb{R}^{3*}$ are linearly independent. Note that $L = \cap_{i=1}^2 (f^i)^{-1}\{0\}$ and that the number of equations coincides with the codimension of L. On the other hand,

$$(x, y, z) = (a\cos t, a\sin t, bt),\; a, b \in \mathbb{R}_0^+,\; t \in \mathbb{R}\,,$$

is the parametric representation of a helix H with radius a that rises by $2\pi b$ units per turn. Observe that the map

$$\psi : \mathbb{R} \ni t \to (a\cos t, a\sin t, bt) \in H$$

is a bijection and an immersion, and that the number of parameters gives the dimension of H.

In view of the above, the following theorem seems perfectly logical:

Theorem 5. *Let $N \subset \mathbb{R}^n$ be a subset of $\mathbb{R}^n$. The next four statements are equivalent:*

- *N is an embedded p-dimensional submanifold of $\mathbb{R}^n$.*
- *For every $x \in N$, there is an open subset $U \ni x$ of $\mathbb{R}^n$ and $n-p$ smooth functions*

$$f^i : U \subset \mathbb{R}^n \to \mathbb{R},\ i \in \{1, \dots, n-p\},$$

such that the derivatives

$$(f^i)'(x) \in \mathbb{R}^{n*},\ i \in \{1, \dots, n-p\},$$

are linearly independent and

$$U \cap N = \cap_{i=1}^{n-p} (f^i)^{-1}\{0\}\,.$$

- *For every $x \in N$, there is an open subset $U \ni x$ of $\mathbb{R}^n$ and a submersion $f : U \subset \mathbb{R}^n \to \mathbb{R}^{n-p}$, such that*

$$U \cap N = f^{-1}\{0\}\,.$$

- *For every $x \in N$, there are open subsets $U \ni x$ of $\mathbb{R}^n$ and $\Omega \ni 0$ of $\mathbb{R}^p$, as well as a homeomorphism $\psi : \Omega \subset \mathbb{R}^p \to U \cap N \subset \mathbb{R}^n$, which is smooth, is an immersion at 0, and maps 0 to x.*

Observe that all assertions of this theorem are local. The second and third statements, which are equivalent (it suffices to set $f = (f^1, \dots, f^{n-p})$), correspond to the aforementioned description by Cartesian equations. The fourth item is consistent with the possibility of parametric representations. Further, it is understood that in the requirement that $\psi : \Omega \subset \mathbb{R}^p \to U \cap N \subset \mathbb{R}^n$ be a homeomorphism, the subset $U \cap N \subset \mathbb{R}^n$ is endowed with the relative topology. The immersion-condition in the last item is essential. This will follow from the proof of the theorem, but can also be seen from the example $\psi : \mathbb{R} \ni t \to (t^2, t^3) \in \mathbb{R}^2$ (the subset $N := \psi(\mathbb{R}) \subset \mathbb{R}^2$ does not satisfy the conditions of Definition 10 at $(0,0) \in N$ and ψ is not an immersion at 0).

As for the proof, we already noticed that Item (2) is equivalent to Item (3). Due to the Constant Rank Theorem, Item (3) leads to a diffeomorphism $f : U \subset \mathbb{R}^n \to \omega \subset \mathbb{R}^n$, so it should imply Item (1). As the diffeomorphism f of Item (1) maps $U \cap N$ to the open subset $\Omega := f(U) \cap \mathbb{R}^p$ of $\mathbb{R}^p$, the restriction of its inverse is the homeomorphism ψ needed in Item (4). Remark that $\psi = f^{-1} \circ i$, where

$$i : (x^1, \dots, x^p) \in \Omega \subset \mathbb{R}^p \to (x^1, \dots, x^p, 0, \dots, 0) \in f(U) \subset \mathbb{R}^n$$

is the canonical inclusion, so an immersion. Finally, we will prove, using as above-mentioned (in particular) the immersion-property, that Item (4) implies Item (2). Hence, we get $(2) \Leftrightarrow (3) \Rightarrow (1) \Rightarrow (4) \Rightarrow (2)$, where we still have to explain further the implications $(3) \Rightarrow (1)$ and mainly $(4) \Rightarrow (2)$.

Proof. **Implication** $(\mathbf{3}) \Rightarrow (\mathbf{1})$. As already said, for every $x \in N$, there are open subsets $U \ni x$ and ω in $\mathbb{R}^n$ and a diffeomorphism $\varphi : U \to \omega$, such that in $\omega = \varphi(U)$, we have

$$(f \circ \varphi^{-1})(x^1, \dots, x^n) = (x^1, \dots, x^{n-p})\,.$$

As

$$U \cap N = \{(y^1, \dots, y^n) \in U : f(y^1, \dots, y^n) = 0\}\,,$$

it follows that

$$\varphi(U \cap N) = \varphi(U) \cap (\{0\} \times \mathbb{R}^p)\,.$$

Implication $(4) \Rightarrow (2)$. Let $m \in N$ and let $\psi : \Omega \subset \mathbb{R}^p \to U \cap N \subset \mathbb{R}^n$ be the corresponding homeomorphism. As

$$\psi'(0) : \mathbb{R}^p \ni x' \to \begin{pmatrix} A \\ B \end{pmatrix} x' = \begin{pmatrix} Ax' \\ Bx' \end{pmatrix} \in \mathbb{R}^n,$$

where A (resp. B) is a real $p \times p$ (resp. $(n-p) \times p$) matrix, has rank p, we can assume—modulo a renumbering of the coordinates—that A is nonsingular, so that

$$\pi_1 \circ \psi'(0) = (\pi_1 \circ \psi)'(0) : \mathbb{R}^p \ni x' \to Ax' \in \mathbb{R}^p,$$

where

$$\pi_1 : \mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{n-p} \ni (x', x'') \to x' \in \mathbb{R}^p$$

is the canonical projection (we will denote the projection onto $\mathbb{R}^{n-p}$ by π_2), is a vector space isomorphism. It therefore follows from the Inverse Function Theorem that $\pi_1 \circ \psi : \Omega \subset \mathbb{R}^p \to \mathbb{R}^p$ is locally a diffeomorphism or coordinate transformation, i.e. that there are open subsets $\omega \ni 0$ and $\omega' \ni m' = \pi_1(m)$ of $\mathbb{R}^p$, such that

$$\psi_1 := \pi_1 \circ \psi : \omega \ni x' \to \psi_1(x') =: y' \in \omega'$$

is a transformation of coordinates or, better, of parameters. When combining the homeomorphism $\psi : \Omega \ni x' \to \psi(x') \in U \cap N$ with this transformation of parameters, we get a simpler parametric representation

$$\psi \circ \psi_1^{-1} : \omega' \ni y' = \psi_1(x') \to \psi(x') = (\psi_1(x'), \psi_2(x')) = (y', \psi_2(\psi_1^{-1}(y')) \in U' \cap N,$$

where $U' \ni m$ is open in $\mathbb{R}^n$ and where $U' \cap N = \psi \circ \psi_1^{-1}(\omega')$ is the image of ω' by the bijection $\psi \circ \psi_1^{-1}$. It is interesting to observe that $U' \cap N$ appears as the graph of $\psi_2 \circ \psi_1^{-1}$. In order to get the equations f^i, $i \in \{1, \ldots, n-p\}$, it suffices to set

$$f^i(y) = \psi^{p+i}(\psi_1^{-1}(y')) - y^{d+i}\,,$$

for every

$$y = (y', y'') \in U'' := U' \cap (\omega' \times \mathbb{R}^{n-p})\,.$$

Indeed, the derivatives at m of these smooth functions f^i are obviously linearly independent and

$$U'' \cap N = \cap_{i=1}^{n-p} (f^i)^{-1}(0)\,.$$

□

3.3 Embedded submanifolds versus abstract manifolds

It is possible to construct manifolds whose topology is not even Hausdorff or second-countable. In order to make further progress in our theory, we exclude such exotic cases.

Remark. In the following all manifolds are implicitly assumed to be Hausdorff and second countable.

We now prove the result mentioned at the beginning of the previous subsection.

Proposition 2. *Any embedded p-dimensional submanifold of $\mathbb{R}^n$ is a smooth p-dimensional manifold whose topology is induced by the topology of $\mathbb{R}^n$.*

In view of the above remark, to prove this theorem we must not only construct a smooth p-dimensional atlas, but we must also show that the topology of the manifold is Hausdorff and second-countable. However, as these properties are hereditary, the conclusion immediately follows from the fact that $\mathbb{R}^n$ is Hausdorff and second countable. For example, the open spheres $b(x, r)$ of $\mathbb{R}^n$ with center $x \in \mathbb{Q}^n$ and radius $r \in \mathbb{Q}$ form a countable basis of open subsets of $\mathbb{R}^n$.

Actually, there are no more abstract manifolds than embedded submanifolds of $\mathbb{R}^n$. Indeed, Whitney's embedding theorem states (Hassler Whitney, 1907 - 1989, American Mathematician) that any smooth (Hausdorff and second-countable) manifold of dimension p can be embedded in $\mathbb{R}^{2p+1}$, and even in $\mathbb{R}^{2p}$ if $p > 2$. Consequently, any manifold can be treated, if we wish, as an object that lives in a larger Cartesian space (although this view is not always of advantage). The proof of Whitney's embedding theorem is complicated and will not be given in these notes.

Hereafter we detail the proof of Proposition 2.

Proof. Let N be an embedded p-dimensional submanifold of $\mathbb{R}^n$. The construction of an atlas follows from the observation that the parametrizations provided by Item (4) of Theorem 5 clearly correspond to local coordinate maps.

Let $x \in N$. There are open subsets $U'_x \ni x$ of $\mathbb{R}^n$ and Ω_x of $\mathbb{R}^p$, as well as a smooth map ψ_x that is a homeomorphism $\psi_x : \Omega_x \subset \mathbb{R}^p \to U'_x \cap N \subset \mathbb{R}^n$ and an immersion (it suffices to restrict the subsets Ω_x and U'_x). Set now

$$U_x := \psi_x(\Omega_x) = U'_x \cap N \quad \text{and} \quad \varphi_x = \psi_x^{-1} : U_x \subset N \to \Omega_x \subset \mathbb{R}^p\,.$$

The charts $(U_x, \varphi_x)_{x\in N}$ form a smooth p-dimensional atlas of N. Indeed, the family $(U_x)_{x\in N}$ is a cover of N and the images $\varphi_x(U_x \cap U_y)$ are open in $\mathbb{R}^p$. The only problem is to prove that the transition bijections

$$\varphi_y \circ \varphi_x^{-1} : \varphi_x(U_x \cap U_y) \to \varphi_y(U_y \cap U_x)$$

are smooth. Indeed, the source of the bijection

$$\varphi_y = \psi_y^{-1} : U'_y \cap N \subset N \to \Omega_y \subset \mathbb{R}^p$$

is not an open subset of a Cartesian space and smoothness of φ_y has no meaning so far. The way out is to use the bijections

$$f : U' \cap N \to f(U' \cap N) = f(U') \cap \mathbb{R}^p \subset \mathbb{R}^p,$$

which are the restrictions of the diffeomorphisms $f : U' \subset \mathbb{R}^n \to f(U') \subset \mathbb{R}^n$ given by Definition 10. Indeed, when writing

$$\varphi_y \circ \varphi_x^{-1} = \varphi_y \circ f^{-1} \circ f \circ \varphi_x^{-1} = (f \circ \varphi_y^{-1})^{-1} \circ (f \circ \varphi_x^{-1}),$$

we deal with bijections $f \circ \varphi_z^{-1}$ between open subsets of $\mathbb{R}^p$, so that smoothness of the building blocks of $\varphi_y \circ \varphi_x^{-1}$ makes sense. The difficulties concerning the matching of the involved subsets can be solved via anew restrictions of these. The $f \circ \varphi_z^{-1}$ are smooth in $\Omega_z \subset \mathbb{R}^p$ and valued in $\mathbb{R}^p$. Moreover, for every $X \in \Omega_y$, we have

$$(f \circ \varphi_y^{-1})'(X) = f'(\psi_y(X)) \circ \psi_y'(X) \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^p).$$

As the derivative of ψ_y is injective and that of f bijective, the map $(f \circ \varphi_y^{-1})'(X)$ is injective and therefore bijective (since the injectivity of a linear map implies its surjectivity if the dimensions of the source and target spaces coincide). It now follows from the Inverse Function Theorem, that $f \circ \varphi_y^{-1}$ is a local diffeomorphism. This finally implies that $\varphi_y \circ \varphi_x^{-1}$ is smooth in the neighborhood of each point.

As for the manifold topology, denote $T(N)$ (resp. $I(N)$) the manifold topology of N (resp. the topology induced on N by $\mathbb{R}^n$). From Exercise 4 it follows that $T(N) \subset I(N)$. Conversely, if $U \cap N \in I(N)$, i.e. if U is open in $\mathbb{R}^n$, the image

$$\varphi_x(U \cap N \cap U_x) = \psi_x^{-1}(U \cap U_x' \cap N)$$

is open in $\mathbb{R}^p$, so that, in view of Definition 5, we also get $I(N) \subset T(N)$. $\square$

3.4 Exercises

1. The preceding result yields that embedded submanifolds of $\mathbb{R}^n$ are manifolds that live in the ambient space $\mathbb{R}^n$ and carry the induced topology. The next exercise gives an example of a manifold that is included in the 2-dimensional Cartesian space and the topology of which is richer than the relative topology.

   Let l_1 and l_2 be two lines of $\mathbb{R}^2$ that intersect at a point x_0. Denote by l_2^- and l_2^+ the two open half-lines defined by x_0 and consider the set $N = l_1 \cup l_2^- \cup l_2^+$. Construct an (obvious) atlas for $N \subset \mathbb{R}^2$ and show that the resulting manifold topology is richer than the induced one.

2. Prove that every manifold is locally compact and locally connected. A bit of topology is needed to solve this quite simple problem.

Chapter 4
Derivatives of smooth maps between manifolds

1 Smooth maps between manifolds

In order to do Analysis on manifolds, we must define smoothness of a map between manifolds.

Take for instance the rotation

$$f : S^1 \ni m \to f(m) \in S^1$$

with center $(0,0)$ and angle $\frac{\pi}{2}$. In standard polar coordinates, or, more precisely, in coordinate charts

$$\varphi : S^1 \setminus \{(1,0)\} \ni m \to \theta \in]0, 2\pi[\quad \text{and} \quad \varphi' : S^1 \setminus \{(0,1)\} \ni m \to \theta' \in]\frac{\pi}{2}, \frac{5\pi}{2}[\,,$$

the rotation f (locally) reads

$$\varphi' f \varphi^{-1} : \mathbb{R} \supset]0, 2\pi[\ni \theta \to \theta' = \theta + \frac{\pi}{2} \in \mathbb{R} \,.$$

The map $\varphi' f \varphi^{-1}$ (note that we omit the symbol for the composition of functions) is the local form of f in the considered charts.

We say that a map $f : M \to M'$ between two manifolds M and M' is smooth, if, in the neighborhood of every point of M, it is smooth in coordinates, i.e. it has a smooth local form $\varphi' f \varphi^{-1}$:

Definition 1. *Let M (resp. M') be an n-dimensional (resp. n'-dimensional) manifold. A map $f : M \to M'$ is a smooth map from M to M', if for any $m \in M$, there exist charts (U, φ) of M around m and (U', φ') of M' around $f(m)$, such that $f(U) \subset U'$ and*

$$\varphi' f \varphi^{-1} : \varphi(U) \subset \mathbb{R}^n \to \mathbb{R}^{n'}$$

is smooth in the sense of Analysis. The set of smooth maps from the manifold M to the manifold M' is denoted by $C^\infty(M, M')$.

The next proposition claims that the existence of charts, in which the local form of f is smooth, implies that this property holds for all charts.

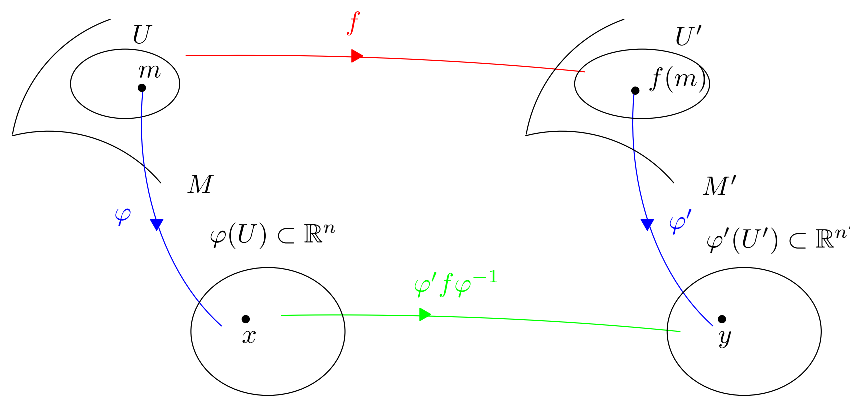

Figure 2: Smooth map between manifolds

Proposition 1. *If $f : M \to M'$ is a smooth map between two manifolds M and M', then, for any charts (U, φ) of M and (U', φ') of M', such that $f(U) \subset U'$, the local form $\varphi' f \varphi^{-1}$ of f is smooth.*

Proof. In order to show that $\varphi' f \varphi^{-1}$ is smooth in $\varphi(U)$, it suffices to prove that it is smooth in a neighborhood of every point $\varphi(m_0)$, $m_0 \in U$, of $\varphi(U)$. Smoothness of f implies that there are charts (V, ψ) of M around m_0 and (V', ψ') of M' around $f(m_0)$, such that $f(V) \subset V'$ and $\psi' f \psi^{-1}$ is smooth in $\psi(V)$. In the neighborhood $\varphi(U \cap V)$ of $\varphi(m_0)$, we then have

$$\varphi' f \varphi^{-1} = (\varphi' \psi'^{-1})(\psi' f \psi^{-1})(\psi \varphi^{-1}) \, ,$$

so that the LHS is smooth in this neighborhood. □

As continuity of a map between manifolds is defined as for any map between topological spaces, we need to check that smoothness actually implies continuity.

Proposition 2. *For all manifolds M and M', we have $C^\infty(M, M') \subset C^0(M, M')$, where it is understood that M and M' are endowed with their canonical manifold topology.*

Proof. Let $f \in C^\infty(M, M')$ and $m \in M$. There are charts (U, φ) of M around m and (U', φ') of M' around $f(m)$, such that $f(U) \subset U'$ and $\varphi' f \varphi^{-1} : \varphi(U) \subset \mathbb{R}^n \to \varphi'(U') \subset \mathbb{R}^{n'}$ is smooth and therefore continuous. As the coordinate maps $\varphi : U \to \varphi(U)$ are homeomorphisms for the induced topologies, it follows that the restriction of f to U, i.e.

$$f|_U = \varphi'^{-1}(\varphi' f \varphi^{-1})\varphi : U \to U',$$

is continuous. Hence, for every $m \in M$, there is $U_m \in \mathfrak{T}_m(M)$, such that $f|_{U_m} \in C^0(U_m, M')$, where $\mathfrak{T}_m(M)$ denotes the set of those open subsets of the manifold topology of M that contain m. But then, if $V \in \mathfrak{T}(M')$, we have

$$f^{-1}(V) = \cup_{m \in M} f|_{U_m}^{-1}(V) \in \mathfrak{T}(M) \, ,$$

so that $f \in C^0(M, M')$. $\square$

Just as for other types of spaces (e.g. vector spaces, topological spaces, ...), there is a concept of equivalence (isomorphism, homeomorphism, ...) for manifolds. Equivalent or diffeomorphic manifolds are manifolds M and M' related by a diffeomorphism $\phi : M \to M'$, i.e. a smooth bijection with smooth inverse. We denote by $\mathrm{Diff}(M, M')$ the set of diffeomorphisms from M to M'. Intuitively, two manifolds are diffeomorphic if each one can be smoothly deformed to the other.

As recalled above, chart maps $\varphi : M \supset U \to \varphi(U) \subset \mathbb{R}^n$ are homeomorphisms, and, as mentioned in Chapter 3, they should be diffeomorphisms $\varphi \in \mathrm{Diff}(U, \varphi(U))$. Since U and $\varphi(U)$ are manifolds (see Chapter 3, Section 2, Exercise 1), the set $\mathrm{Diff}(U, \varphi(U))$ is now defined. We actually have the

Proposition 3. *Let M be a manifold and let $(U_\alpha, \varphi_\alpha)_\alpha$ be an atlas of M . Every coordinate map $(U_\alpha, \varphi_\alpha)$ is a diffeomorphism: $\varphi_\alpha \in \mathrm{Diff}(U_\alpha, \varphi_\alpha(U_\alpha))$.*

Proof. Obvious. $\square$

We now examine the smoothness of maps made of smooth building blocks. The proofs of these results are simple but instructive. They are left to the reader.

Proposition 4. *Let M, M', and M'' be manifolds. If $f \in C^\infty(M, M')$ and $g \in C^\infty(M', M'')$, then $g \circ f \in C^\infty(M, M'')$.*

It is well-known that the restriction of a linear (resp. continuous) map to a vector (resp. topological) subspace is still linear (resp. continuous), and that a linear (resp. continuous) map viewed as valued in a vector (resp. topological) subspace that contains all images, is also linear (resp. continuous) (see e.g. Chapter 2, Section 4, Exercise 6). Similar results are true in the smooth category.

Proposition 5. *Let M (resp. M') be a manifold, let $f \in C^\infty(\mathbb{R}^r, M')$ (resp. $f \in C^\infty(M, \mathbb{R}^s)$), and let N (resp. N') be an embedded submanifold of $\mathbb{R}^r$ (resp. of $\mathbb{R}^s$, such that $f(M) \subset N'$). Then $f|_N \in C^\infty(N, M')$ (resp. $f \in C^\infty(M, N')$). Analogous results are valid for $f \in C^\infty(M, M')$ and restrictions of the source (resp. target) manifold to an open subset of M (resp. an open subset of M' that contains $f(M)$).*

The set of smooth functions of a manifold M, i.e. of smooth maps from M to the target manifold $M' = \mathbb{R}$, is denoted by $C^\infty(M)$. The addition and multiplication of $\mathbb{R}$ induce an addition $f+g$ and a multiplication $f.g$ of functions, as well as a multiplication λf of functions by reals.

Proposition 6. *The set $C^\infty(M)$ of smooth functions of a manifold M is an associative commutative unital algebra for the canonical operations $f + g$, λf, and $f.g$.*

The function algebra $C^\infty(M)$ is actually a fundamental object associated to the manifold M. Indeed, the algebraic structure of $C^\infty(M)$ characterizes the manifold structure of M. More precisely:

Theorem 1. *Two manifolds M and M' are diffeomorphic if and only if their associative algebras of functions $C^\infty(M)$ and $C^\infty(M')$ are isomorphic.*

Proof. If $\phi : M \to M'$ is a diffeomorphism, then, obviously, the map

$$\phi_* : C^\infty(M) \ni f \to f \circ \phi^{-1} \in C^\infty(M')$$

is an associative algebra isomorphism. The proof of the converse result is not obvious and will not be given. □

2 Exercises

1. Prove Propositions 4, 5, and 6. Hint: To prove Proposition 4, choose first a chart (U', φ') around $f(m)$, then a chart (U, φ) around m such that $U \subset f^{-1}(U')$. In the proof of Proposition 6, choose the same chart for f and g.

2. Show that the set $\mathrm{Diff}(M)$ of diffeomorphisms of M (i.e. from the manifold M to itself) is a group for the composition of maps.

3. Prove that the projections $\pi_1 : M \times M' \to M$ and $\pi_2 : M \times M' \to M'$ (M, M': manifolds) are smooth maps (see Chapter 3, Section 2, Exercise 2).

4. Explain that, if $f_1 \in C^\infty(M_1, M_1')$ and $f_2 \in C^\infty(M_2, M_2')$, then

$$f_1 \times f_2 : M_1 \times M_2 \ni (m_1, m_2) \to (f_1(m_1), f_2(m_2)) \in M_1' \times M_2'$$

   is also smooth.

3 Tangent space

In Lagrangian Mechanics, the configurations of a double pendulum, i.e. of a pendulum with another pendulum attached to its end, are described by two angles $q = (q^1, q^2) \in [0, 2\pi] \times [0, 2\pi]$. Hence, the configuration space M is $[0, 2\pi] \times [0, 2\pi]$ modulo identification of the extreme values 0 and 2π, in other words $M = T^2 = S^1 \times S^1 \subset \mathbb{R}^3$. The representative point q of the considered dynamical system runs through a curve $\alpha : I \ni t \to q(t) \in M \subset \mathbb{R}^3$ of M, where $I \subset \mathbb{R}$ is the (open) observation interval, and the generalized velocity $\dot{q}(0) = d_t\alpha|_{t=0}$ at time 0 is a tangent vector of the curve α at the point $q(0) = \alpha(0) =: m$. Of course, a vector that is tangent to a curve of M at m is also tangent to the manifold M at m, and the "tangent space" T_mM of M at m can be viewed as the set of the tangent vectors $d_t\alpha|_{t=0}$ of all the curves $\alpha \in C^\infty(I, M)$ of M, such that $\alpha(0) = m$.

If M is an abstract manifold that does not sit in an ambient space, our mental picture of a tangent vector, as well as the definition of the derivative $d_t\alpha|_{t=0}$ of a curve $\alpha \in C^\infty(I, M)$, where I denotes an open interval of the real line, meet a problem. A natural idea is to look at a curve $\alpha \in C^\infty(I, M)$ locally, in coordinates φ, and to end up, roughly, with a usual curve $\varphi\alpha \in C^\infty(I, \mathbb{R}^n)$, which *has* a tangent vector $d_t(\varphi\alpha)|_{t=0}$. However, obviously, different curves $\beta \neq \alpha$, i.e. $\varphi\beta \neq \varphi\alpha$, can have the same tangent vector

$$d_t(\varphi\beta)|_{t=0} = d_t(\varphi\alpha)|_{t=0}\,.$$

Therefore, a tangent vector is characterized, not by a single curve, but by a class of curves, and we can define a tangent vector of an abstract manifold M at a point $m \in M$ as the class

$[\alpha]$ of all the curves $\alpha \in C^\infty(I, M)$ of M, such that $\alpha(0) = m$ and whose local forms have the same tangent vector $d_t(\varphi\alpha)|_{t=0}$.

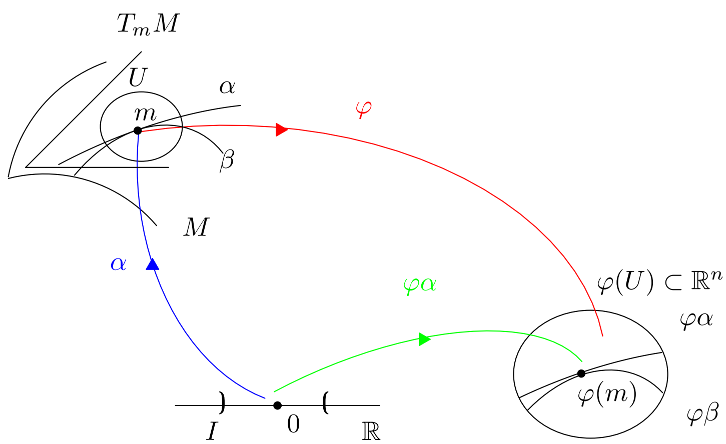

Figure 3: Tangent space of a manifold

We now make these ideas more precise.

Let m be a point of an n-dimensional manifold M. We denote by $\mathcal{C}$ the set of all curves $\alpha \in C^\infty(I, M)$ of M (where I is an open interval of $\mathbb{R}$ that contains 0) that pass through m at $t = 0$, i.e. satisfy $\alpha(0) = m$. Due to the continuity of α, there is, for any charts (U, φ) and (V, ψ) of M around m, an open interval $J \ni 0$ of $\mathbb{R}$, such that

$$\varphi\alpha \in C^\infty(J, \varphi(U \cap V)) \quad \text{and} \quad \psi\alpha \in C^\infty(J, \psi(V \cap U))\,.$$

Definition 2. *Let M, m, and $\mathcal{C}$ be the just defined objects. Two curves $\alpha, \beta \in \mathcal{C}$ are tangent at m, if there is a chart (U, φ) of M around m, such that*

$$d_t(\varphi\alpha)|_{t=0} = d_t(\varphi\beta)|_{t=0}. \tag{1}$$

Note that the same condition is then satisfied for every chart (V, ψ) of M around m. Indeed, we have $\psi\alpha = (\psi\varphi^{-1})(\varphi\alpha)$ in J, so that

$$d_t(\psi\alpha)|_{t=0} = (\psi\varphi^{-1})'(\varphi(m))\, d_t(\varphi\alpha)|_{t=0}\,, \tag{2}$$

and as the same result holds for β, the conclusion follows from Equation (1). It is now clear that the relation "tangent at m" is an equivalence in $\mathcal{C}$.

Definition 3. *The notations are again those introduced above. A tangent vector of M at m is an equivalence class $[\alpha]$, $\alpha \in \mathcal{C}$, of the equivalence relation "tangent at m" in $\mathcal{C}$. The set of all tangent vectors of M at m is denoted by T_mM.*

The mental picture of the preceding construction suggests that every chart (U, φ) of M at m defines a 1-to-1 correspondence

$$T_m\varphi : T_mM \to \mathbb{R}^n$$

($T_m\varphi$ is merely a notation) between the tangent vectors of M at m and the vectors of $\mathbb{R}^n$. Of course, the set T_mM of tangent *vectors* at m should be a *vector space*. We will use the bijection just described to transport the vector space structure of $\mathbb{R}^n$ to T_mM, i.e. for every vectors $[\alpha_i] \in T_mM$, $i \in \{1, \ldots, N\}$, and every real numbers $\lambda_i \in \mathbb{R}$, we set

$$\sum_i \lambda_i[\alpha_i] := (T_m\varphi)^{-1}\left(\sum_i \lambda_i\, T_m\varphi[\alpha_i]\right). \tag{3}$$

Theorem 2. *Let (U, φ) be a chart of M around m. The map*

$$T_m\varphi : T_mM \ni [\alpha] \to d_t(\varphi\alpha)|_{t=0} \in \mathbb{R}^n \tag{4}$$

is a well-defined bijection. Moreover, the set T_mM admits a unique vector space structure, such that, for every chart (U, φ) of M around m, the bijection $T_m\varphi$ is a vector space isomorphism.

Proof. It is obvious that $T_m\varphi$ is well-defined and injective. Let $v \in \mathbb{R}^n$ and set

$$\alpha(t) := \varphi^{-1}(\varphi(m) + tv)\,,$$

for t close to 0. We have $\alpha \in \mathcal{C}$ and

$$T_m\varphi[\alpha] = d_t(\varphi\alpha)|_{t=0} = v\,,$$

so that $T_m\varphi$ is also surjective.

We now choose a chart (U, φ) of M around m and define a vector space structure on T_mM via Definition 3, which, of course, turns $T_m\varphi$ into a vector space isomorphism. If (V, ψ) is another chart of M around m, the combination of Equations (2) and (4) shows that

$$T_m\psi = (\psi\varphi^{-1})'(\varphi(m))T_m\varphi\,,$$

so that $T_m\psi$ is built from two vector space isomorphisms and is therefore itself a vector space isomorphism. The uniqueness of the vector space structure having the required property is obvious. □

Remark. In view of the preceding theorem, the tangent *space* T_mM of a manifold M at a point $m \in M$ is a real vector *space* with the same dimension as the underlying manifold. For $M = \mathbb{R}^n$, the isomorphism

$$T_m\mathbb{R}^n \ni [\alpha] \simeq d_t\alpha|_{t=0} \in \mathbb{R}^n \tag{5}$$

is canonical (the atlas contains only one chart) and we often identify both spaces. If it is necessary to remember the point at which the vectors are tangent to $\mathbb{R}^n$, we view $T_m\mathbb{R}^n$ as the space

$$T_mM \simeq \{m\} \times \mathbb{R}^n$$

of vectors of $\mathbb{R}^n$ with origin m. A similar identification is used for real finite-dimensional vector spaces $M = V$: $T_mV \simeq V$. Further, it is clear from the above definitions that for every open subset U of a manifold M, we have $T_mU \simeq T_mM$, for all $m \in U$ – a fact that is also corroborated by intuition.

4 Derivative of a smooth map between manifolds

We are now prepared to define the derivative at a point $m \in M$ of a smooth map $f \in C^\infty(M, M')$ between two manifolds M ($\dim M = n$) and M' ($\dim M' = n'$). Take a chart (U, φ) of M at m and a chart (V, ψ) of M' at $f(m)$, such that $f(U) \subset V$. Just as the diffeomorphisms $\varphi \in \mathrm{Diff}(U, \varphi(U))$ and $\psi \in \mathrm{Diff}(V, \psi(V))$ allow us to construct the map

$$f = \psi^{-1}(\psi f \varphi^{-1})\varphi$$

locally from its coordinate form $\psi f \varphi^{-1} \in C^\infty(\varphi(U), \psi(V))$, the isomorphisms $T_m\varphi \in \mathrm{Isom}(T_mM, \mathbb{R}^n)$ and $T_{f(m)}\psi \in \mathrm{Isom}(T_{f(m)}M', \mathbb{R}^{n'})$ allow us to define the derivative

$$T_m f := (T_{f(m)}\psi)^{-1}\,(\psi f \varphi^{-1})'(\varphi(m))\,T_m\varphi$$

of f at m by means of the derivative $(\psi f \varphi^{-1})'(\varphi(m)) \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^{n'})$ of the coordinate form.

Definition 4. *Let $f \in C^\infty(M, M')$ and $m \in M$. If (U, φ) is a chart of M at m and (V, ψ) a chart of M' at $f(m)$, such that $f(U) \subset V$, the derivative or tangent map of f at m is defined by*

$$T_m f := (T_{f(m)}\psi)^{-1}\,(\psi f \varphi^{-1})'(\varphi(m))\,T_m\varphi. \tag{6}$$

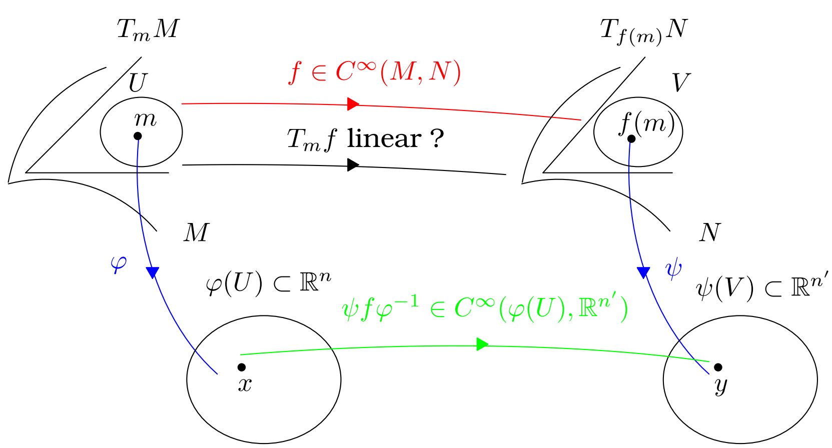

Figure 4: Derivative of a smooth map

This definition implies that for $f = \varphi \in C^\infty(U, \varphi(U))$, the derivative $T_m\varphi$, $m \in U$, given by Equation (6), coincides with the isomorphism $T_m\varphi$, denoted by the same symbol, defined by Equation (4) and used in Definition (6). Further, it is easily checked that if $f \in C^\infty(\Omega, \mathbb{R}^{n'})$, Ω open in $\mathbb{R}^n$, the "Geometry"-derivative T_mf, $m \in \Omega$, coincides with the "Analysis"-derivative $f'(m) \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^{n'})$. As T_mf,

$$f \in C^\infty(M, M'),$$

$m \in M$, thus extends the usual derivative, it should be a linear map between vector spaces attached to M and M'. Indeed, Equation (6) implies that

$$T_m f \in \mathrm{Hom}_{\mathbb{R}}(T_mM, T_{f(m)}M'). \tag{7}$$

This last result exhibits the functorial character of T (we encourage the reader to inform himself about the concepts of category and functor).

Of course, the derivative $T_m f$ of $f \in C^\infty(M, M')$ at a point $m \in M$ should only depend on f and m, and not on the charts chosen to compute it. To convince oneself of this independence, it is sufficient to apply the definition to the computation of $T_m f\,[\alpha]$, $[\alpha] \in T_m M$. We get

$$T_{f(m)}\psi\; T_m f\,[\alpha] = (\psi f \varphi^{-1})'(\varphi(m))\; T_m\varphi\,[\alpha]$$

$$= (\psi f \varphi^{-1})'(\varphi(m))\; d_t(\varphi\alpha)|_{t=0} = d_t(\psi f \alpha)|_{t=0} = T_{f(m)}\psi\,[f\alpha],$$

as $f\alpha$ is a curve of M' that passes at $t = 0$ through $f(m)$. Since $T_{f(m)}\psi$ is an isomorphism, we get:

Proposition 7. *Let $f \in C^\infty(M, M')$ and $m \in M$. The tangent map $T_m f$ is given in a coordinate-free way by*

$$T_m f : T_m M \ni [\alpha] \to [f\alpha] \in T_{f(m)} M'.$$

The next proposition extends the known result concerning the derivation of composite maps.

Proposition 8. *Let $f \in C^\infty(M, M')$, $g \in C^\infty(M', M'')$, and $m \in M$. The derivative at m of $g \circ f \in C^\infty(M, M'')$, see Proposition 4, is given by*

$$T_m(g \circ f) = T_{f(m)} g \circ T_m f. \tag{8}$$

Proof. Let $[\alpha] \in T_m M$. On the one hand $T_m(gf)\,[\alpha] = [(gf)\alpha]$, on the other

$$T_{f(m)} g\, T_m f\,[\alpha] = T_{f(m)} g\,[f\alpha] = [g(f\alpha)]\,.$$

□

We also generalize the relations between diffeomorphisms f and isomorphisms $f'(x)$, see Chapter 3, Section 3.1, Theorem 4, and Proposition 1.

Proposition 9. *If $f \in \mathrm{Diff}(M, M')$, then, for every $m \in M$, we have*

$$T_m f \in \mathrm{Isom}(T_m M, T_{f(m)} M') \quad \textit{and} \quad T_{f(m)} f^{-1} = (T_m f)^{-1}\,.$$

Proof. We only need to check that the two composite maps built from

$$T_m f \in \mathrm{Hom}_{\mathbb{R}}(T_m M, T_{f(m)} M') \quad \text{and} \quad T_{f(m)} f^{-1} \in \mathrm{Hom}_{\mathbb{R}}(T_{f(m)} M', T_m M)$$

are equal to identity. For instance,

$$T_m f \circ T_{f(m)} f^{-1} = T_{f(m)}(f \circ f^{-1}) = T_{f(m)}\, \mathrm{id}_{M'} = \mathrm{id}_{T_{f(m)} M'}\,.$$

The second verification is similar. □

Proposition 10. *If $f \in C^\infty(M, M')$, $m \in M$ and $T_m f \in \mathrm{Isom}(T_m M, T_{f(m)} M')$, there is an open subset W of M around m, such that $f(W)$ is open in M' and $f \in \mathrm{Diff}(W, f(W))$.*

Proof. As Equation (6) yields that

$$(\psi f \varphi^{-1})'(\varphi(m)) \in \mathrm{Isom}(\mathbb{R}^n, \mathbb{R}^{n'})\,,$$

the aforementioned Theorem 4 asserts that there is an open subset $\Omega \ni \varphi(m)$ of $\varphi(U)$, such that $\Omega' := (\psi f \varphi^{-1})(\Omega)$ is an open subset of $\psi(V)$ and $\psi f \varphi^{-1} \in \mathrm{Diff}(\Omega, \Omega')$. Since the coordinate maps are also diffeomorphisms, $\varphi \in \mathrm{Diff}(U, \varphi(U))$ and $\psi \in \mathrm{Diff}(V, \psi(V))$, it follows that

$$f = \psi^{-1}(\psi f \varphi^{-1})\varphi \in \mathrm{Diff}(\varphi^{-1}(\Omega), \psi^{-1}(\Omega'))\,.$$

□

The concept of velocity, see first paragraph, leads to a practical technique for computing tangent maps. Note that the velocity $d_t\alpha$ of a curve α at time $t \in I$ is so far only defined for curves $\alpha \in C^\infty(I, \mathbb{R}^n)$ of $\mathbb{R}^n$. Actually we did not make a difference between this tangent vector $d_t\alpha \in \mathbb{R}^n$ and the derivative $\alpha'(t) \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}, \mathbb{R}^n)$, because of the canonical isomorphism $\mathrm{Hom}_{\mathbb{R}}(\mathbb{R}, \mathbb{R}^n) \simeq \mathbb{R}^n$. In order to define the velocity $d_t\alpha$ of a curve $\alpha \in C^\infty(I, M)$ at $t \in I$, for any manifold M, note that we should have $d_t\alpha \in T_{\alpha(t)}M$ and that $T_t\alpha \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}, T_{\alpha(t)}M)$.

Definition 5. *For every curve $\alpha \in C^\infty(I, M)$ of M and every point $t \in I$, the velocity or tangent vector $d_t\alpha$ of α at t is*

$$d_t\alpha := T_t\alpha\,(1) \in T_{\alpha(t)}M.$$

Remarks.

1. Let, as usual, $f \in C^\infty(M, M')$ and $m \in M$, so that $T_m f \in \mathrm{Hom}_{\mathbb{R}}(T_mM, T_{f(m)}M')$. In order to compute $T_m f\,(X_m)$, $X_m \in T_mM$, by means of a velocity, we need a curve $\alpha \in C^\infty(I, M)$ of M that passes through m at $t = 0$ and whose velocity or tangent vector at $t = 0$ is equal to X_m. Indeed, in this case,

$$T_m f\,(X_m) = T_{\alpha(0)} f\,(d_t\alpha|_{t=0}) = T_{\alpha(0)} f\, T_0\alpha\,(1) = T_0(f\alpha)\,(1) = d_t(f\alpha)|_{t=0}.$$

The point is that, if M' is a vector space or even $\mathbb{R}^{n'}$, the map $f\alpha \in C^\infty(I, \mathbb{R}^{n'})$ is a map between (subsets of) Cartesian spaces, and the RHS of the last equation is a derivative in the sense of Analysis.

2. In view of the identification (5), we have

$$T_{\alpha(0)}\mathbb{R}^n \ni [\alpha] = d_t\alpha|_{t=0} \in \mathbb{R}^n\,.$$

This result is also valid for a curve β of an arbitrary manifold M, which is defined on an interval $I \ni 0$. In fact, as just mentioned, if id denotes the curve $\mathrm{id} : I \ni t \to t \in \mathbb{R}$, we have

$$T_0\mathbb{R} \ni [\mathrm{id}] = d_t\,\mathrm{id}\,|_{t=0} = 1 \in \mathbb{R}\,.$$

Hence

$$T_{\beta(0)}M \ni [\beta] = [\beta\,\mathrm{id}] = T_0\beta\,[\mathrm{id}] = T_0\beta\,(1) = d_t\beta|_{t=0} \in T_{\beta(0)}M.$$

5 Exercises

1. Prove that for every open subset U of a manifold M, the tangent space T_mU, $m \in U$, is isomorphic to the tangent space T_mM.

   Suggestion: If $i : U \hookrightarrow M$ denotes the canonical injection, it follows from Equation (6) that $T_m i$ is a canonical isomorphism.

2. We know that if N denotes an embedded submanifold of $\mathbb{R}^n$ of dimension p, there is, for any $x \in N$, an open subset $U \ni x$ of $\mathbb{R}^n$ and a submersion $f : U \subset \mathbb{R}^n \to \mathbb{R}^{n-p}$, such that $U \cap N = f^{-1}(0)$ (resp. there are open subsets $U \ni x$ of $\mathbb{R}^n$ and $\Omega \ni 0$ of $\mathbb{R}^p$, and a homeomorphism $\psi : \Omega \subset \mathbb{R}^p \to U \cap N \subset \mathbb{R}^n$, which is smooth, is an immersion at 0, and maps 0 to x). Show that, if f (resp. ψ) is such a submersion (resp. immersion), the tangent space of N at x is given by

$$T_xN = \ker f'(x) \quad (\text{resp.} \quad T_xN = \operatorname{im} \psi'(0)). \tag{9}$$

   Suggestion: Observe that the LHS and RHS of Equation (9) are vector spaces of dimension p. Hence, it suffices to prove that $T_xN \subset \ker f'(x)$ (resp. $\operatorname{im} \psi'(0) \subset T_xN$). In order to compute $f'(x)(X_x)$, $X_x \in T_xN$, note that $X_x = [\alpha] = d_t\alpha|_{t=0}$, where α is a curve $\alpha \in C^\infty(I, U \cap N)$, such that $\alpha(0) = x$ (resp. in order to compute $\psi'(0)(\mathbb{R}^p)$, remark that

$$\mathbb{R}^p = T_0\mathbb{R}^p = T_0\Omega \ni X = [\alpha] = d_t\alpha|_{t=0} \, ,$$

   where α is a curve $\alpha \in C^\infty(I, \Omega)$, such that $\alpha(0) = 0$).

3. Prove that the orthogonal group

$$\mathrm{O}(n) = \{A \in \mathrm{gl}(n, \mathbb{R}) : \tilde{A}A - \mathrm{id} = 0\}$$

   is an $n(n-1)/2$-dimensional, smooth, Hausdorff, and second countable manifold.

   Suggestion: As $\mathrm{O}(n) = f^{-1}(0)$, where

$$f : \mathbb{R}^{n^2} \ni A \to \tilde{A}A - \mathrm{id} \in \mathbb{R}^{n(n+1)/2}$$

   (as a matter of fact, the matrix $\tilde{A}A - \mathrm{id}$ is symmetric), it is enough to ensure that f is a submersion at any point $A \in \mathrm{O}(n)$. Indeed, then $\mathrm{O}(n)$ is an embedded submanifold of $\mathbb{R}^{n^2}$ that has the announced dimension, hence, it is a Hausdorff and second countable manifold. Compute $f'(A)(H)$, $H \in \mathbb{R}^{n^2}$, with the help of the velocity of a curve. Note that based on the result obtained, we get $f'(A)(AK) = 2K$, for any symmetric matrix K.

6 Tangent and cotangent bundles

6.1 Model of the cotangent space

We already mentioned that the algebra of smooth functions of a manifold M deserves special attention. In the following, we refer to the tangent map at $m \in M$ or derivative at m of a *function* $f \in C^\infty(M)$ as the differential of f at m and we write $T_mf =: (d\,f)_m$. Remark that

$$(d\,f)_m \in \mathrm{Hom}_{\mathbb{R}}(T_mM, \mathbb{R}) = T_m^*M \, .$$

The dual space T_m^*M of T_mM is called the cotangent space of M at m.

Local coordinates of M around m, say $(U, \varphi = (x^1, \ldots, x^n))$, induce a basis of T_mM and of T_m^*M, by simple transport of the bases of the isomorphic vector spaces $\mathbb{R}^n$ and $\mathbb{R}^{n*}$. Indeed, if

$$e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^{\tilde{}}$$

denotes the ith vector of the canonical basis of $\mathbb{R}^n$ and if

$$\varepsilon^i = (0, \ldots, 0, 1, 0, \ldots, 0)$$

is the ith vector of the dual basis in $\mathbb{R}^{n*}$, since

$$T_m\varphi \in \mathrm{Isom}(T_mM, \mathbb{R}^n) \quad \text{and} \quad T_m^{\sim}\varphi \in \mathrm{Isom}(\mathbb{R}^{n*}, T_m^*M),$$

the vectors

$$t_i := (T_m\varphi)^{-1}e_i \quad \text{and} \quad c^i := T_m^{\sim}\varphi\, \varepsilon^i$$

are bases of the tangent and cotangent spaces respectively. These bases are dual, as

$$c^i(t_j) = T_m^{\sim}\varphi\, \varepsilon^i((T_m\varphi)^{-1}e_j) = \varepsilon^i(T_m\varphi\,(T_m\varphi)^{-1}e_j) = \delta^i_j,$$

where δ^i_j is Kronecker's symbol.

Since $x^i = \varphi^i = \varepsilon^i \circ \varphi : U \to \mathbb{R}$ is a smooth function $x^i \in C^\infty(U)$, we have

$$(d\,x^i)_m = T_{\varphi(m)}\varepsilon^i \circ T_m\varphi = \varepsilon^i \circ T_m\varphi = T_m^{\sim}\varphi\, \varepsilon^i = c^i.$$

Hence the

Proposition 11. *If $(U, \varphi = (x^1, \ldots, x^n))$ is a coordinate chart of M, and if we write $d\,x^i$ instead of*

$$(d\,x^i)_m = T_m^{\sim}\varphi\, \varepsilon^i\,,$$

the differentials $(d\,x^1, \ldots, d\,x^n)$ form a basis of the cotangent space T_m^*M of M at every point $m \in U$.

Recall that in the previous proposition the dependence of $d\,x^i$ on m is implicit. This simplifying convention is prevalent in textbooks.

In Differential Geometry we prefer global viewpoints to computations in local coordinates. Nevertheless, local computations *are* important.

In order to find the local coordinate form of the differential of a function $f \in C^\infty(M)$, remember first that if V is a real finite-dimensional vector space with basis e_i and if $\ell \in V^*$ is a linear form in the vector space V^* with dual basis ε^i, we have $\ell = \sum_i \ell(e_i)\varepsilon^i$. When working in a chart $(U, \varphi = (x^1, \ldots, x^n))$ of M and applying the preceding remark to $(d\,f)_m \in T_m^*M$, with $m \in U$, so that $(t_1, \ldots, t_n)$ is a basis of T_mM and $(d\,x^1, \ldots, d\,x^n)$ a basis of T_m^*M, we get

$$(d\,f)_m = \textstyle\sum_i (d\,f)_m(t_i)d\,x^i = \sum_i (f\varphi^{-1})'(\varphi(m))\; T_m\varphi\,(t_i)d\,x^i$$

$$= \textstyle\sum_i (f\varphi^{-1})'(\varphi(m))(e_i)d\,x^i = \sum_i \partial_{x^i}(f\varphi^{-1})|_{\varphi(m)}d\,x^i.$$

In a chart, we can identify points m with their coordinates

$$(x^1, \ldots, x^n) = (x^1(m), \ldots, x^n(m)) = \varphi(m)\,.$$

We then recover the concept of differential known from Mechanics:

Proposition 12. *Let $f \in C^\infty(M)$ and let $(U, (x^1, \ldots, x^n))$ be a coordinate chart of M. For every $m \in U$, the differential of f at m is given by*

$$(d\,f)_m = \sum_i \partial_{x^i} f|_m d\,x^i. \tag{10}$$

Corollary 1. *The differential at $m \in M$,*

$$d_m : C^\infty(M) \ni f \to (d\,f)_m \in T^*_m M,$$

is a linear operator that satisfies the Leibniz rule, i.e., for every $f, g \in C^\infty(M)$, we have

$$d(f.g)_m = (d\,f)_m.g(m) + f(m).(d\,g)_m\ .$$

Moreover d_m vanishes on functions that are constant in an open neighborhood of m.

Note that the last property implies that d_m is a local operator, i.e. that the value of d_m only depends on the values of f around m, or, more precisely, that, if $f = g$ (resp. $f = 0$) in an open neighborhood of m, then $(d\,f)_m = (d\,g)_m$ (resp. $(d\,f)_m = 0$).

The differential d_m at m leads to a model of the cotangent space $T^*_m M$, $m \in M$. Indeed, as d_m is a linear operator, it induces a vector space isomorphism

$$\tilde{d}_m : C^\infty(M)/\ker d_m \to \operatorname{im} d_m \subset T^*_m M\ .$$

If d_m is surjective, the LHS quotient is a model of $T^*_m M$.

Theorem 3. *Let M be a smooth manifold of dimension n and let $m \in M$. The differential d_m at m induces an isomorphism between the real n-dimensional cotangent vector space $T^*_m M$ and the quotient space of the vector space of smooth functions of M by the vector subspace $\ker d_m$ of those functions whose differential at m vanishes:*

$$T^*_m M \simeq C^\infty(M)/\ker d_m\ . \tag{11}$$

Remark. Bump functions and partitions of unity (we encourage the reader to familiarize themselves with these concepts) are crucial technical tools that make it possible to move from local considerations to global ones. For every *smooth* manifold M (let us stress that the Hausdorff property is of importance here), every point $m \in M$, and every open neighborhood U of m, there is a smooth function $\gamma \in C^\infty(M)$, valued in $[0, 1]$, with support

$$\operatorname{supp} \gamma = \overline{\{m \in M : \gamma(m) \neq 0\}} \subset U\,,$$

and which is equal to 1 in a neighborhood of m. Such functions are referred to as bump functions or plateau functions. The existence of plateau functions will not be proven in these notes.

In the following we give the proof of Theorem 3.

Proof. Let $(U, \varphi = (x^1, \ldots, x^n))$ be a chart of M around m. For every

$$\alpha_m = \sum_i \alpha_{i,m} d\,x^i \in T^*_m M,\ \alpha_{i,m} \in \mathbb{R}\,,$$

the function $l \in C^\infty(U)$, $l(x) = \sum_i \alpha_{i,m} x^i$, has differential $(d\,l)_m = \alpha_m$. If γ is a plateau function with support in U and value 1 around m, we have $f := \gamma l \in C^\infty(M)$ and $(d\,f)_m = (d\,l)_m = \alpha_m$. □

6.2 Model of the tangent space

In Mechanics, we associate to every vector a directional derivative:

$$\vec{v} \mapsto \left(f \mapsto \vec{v} \cdot \vec{\nabla} f = \sum_i v^i \partial_{x^i} f\right).$$

The extension of this construction to tangent vectors in $T_m M$, $m \in M$, leads to an algebraic characterization of the tangent space.

Let $X_m \in T_m M$ and $f \in C^\infty(M)$. As $(d\,f)_m \in T_m^* M$, we have $X_m(d\,f)_m \in \mathbb{R}$. In order to see that this corresponds to the just mentioned directional derivative, it suffices to write the last value in local coordinates $(U, \varphi = (x^1, \ldots, x^n))$ of M around m. As $X_m = \sum_i X_m^i t_i$ and $(d\,f)_m = \sum_j \partial_{x^j} f|_m d\,x^j$, we get

$$X_m = \sum_i X_m^i t_i \mapsto \left(f \mapsto X_m(d\,f)_m = (d\,f)_m(X_m) = \sum_j \partial_{x^j} f|_m d\,x^j \left(\sum_i X_m^i t_i\right) = \sum_i X_m^i \partial_{x^i} f|_m\right). \tag{12}$$

Remember that the result

$$X_m(d\,f)_m = (d\,f)_m(X_m)$$

is the direct consequence of the identification

$$(T_m^* M)^* \simeq T_m M\,,$$

which is valid for finite-dimensional vector spaces.

Note that the map

$$L_{X_m} : C^\infty(M) \ni f \mapsto X_m(d\,f)_m \in \mathbb{R}$$

is linear and satisfies Leibniz' rule. We call such a map a derivation of $C^\infty(M)$ at m and we denote the set of all these derivations by $\mathrm{Der}_m(C^\infty(M))$. The set $\mathrm{Der}_m(C^\infty(M))$ is clearly a vector space. Hence, we get a map

$$L_m : T_m M \ni X_m \mapsto L_{X_m} = X_m \circ d_m \in \mathrm{Der}_m(C^\infty(M))\,. \tag{13}$$

It will be shown that this map is a vector space isomorphism.

Theorem 4. *Let M be a smooth manifold of dimension n and let $m \in M$. The directional derivative L_m at m is an isomorphism between the real n-dimensional tangent vector space $T_m M$ and the vector space $\mathrm{Der}_m(C^\infty(M))$ of derivations of $C^\infty(M)$ at m:*

$$T_m M \simeq \mathrm{Der}_m(C^\infty(M))\,. \tag{14}$$

Because of Equation (12), the identification of $X_m \in T_m M$ and $L_{X_m} \in \mathrm{Der}_m(C^\infty(M))$ implies in particular the identification of the basis tangent vectors t_i with the derivations $\partial_{x^i}|_m$ at m.

Proposition 13. *If $(U, \varphi = (x^1, \ldots, x^n))$ is a coordinate chart of M and if we write ∂_{x^i} instead of*

$$\partial_{x^i}|_m = (T_m \varphi)^{-1} e_i\,,$$

the derivations $(\partial_{x^1}, \ldots, \partial_{x^n})$ form a basis of the tangent space $T_m M$ of M at every point $m \in U$ and every tangent vector $X_m \in T_m M$ reads

$$X_m = \sum_i X_m^i \partial_{x^i}\,, \tag{15}$$

with $X_m^i \in \mathbb{R}$.

Observe that Theorem 4 provides an algebraic characterization of the tangent space at a point of a manifold. In order to prove this result, we need the following lemma.

Lemma 1. *Every derivation of $C^\infty(M)$ at m is a local operator and vanishes on constant functions.*

Proof. Let $\delta_m \in \mathrm{Der}_m(C^\infty(M))$ and let $f \in C^\infty(M)$ vanish in an open neighborhood U of m. The proof uses a bump function γ with support in U and value 1 in a neighborhood of m. Since $f = (1-\gamma)f$, we get

$$\delta_m(f) = \delta_m((1-\gamma)f) = \delta_m(1-\gamma).f(m) + (1-\gamma(m)).\delta_m(f) = 0\,.$$

As for the last statement, note that $\delta_m(c) = c\,\delta_m(1)$, for every constant function $c \in C^\infty(M)$, and that

$$\delta_m(1) = \delta_m(1.1) = \delta_m(1).1 + 1.\delta_m(1) = 2\delta_m(1) = 0\,.$$

□

We are now able to understand the proof of Theorem 4.

Proof. It is obvious that the map L_m is linear. As concerns injectivity, note that $L_m X_m = 0$ implies that $X_m(d\,f)_m = 0$, for all $f \in C^\infty(M)$. Since d_m is surjective, this means that X_m vanishes on $T^*_m M$. Hence $X_m = 0$. Surjectivity of L_m is less obvious. Let $\delta_m \in \mathrm{Der}_m(C^\infty(M))$. Because of Theorem 3, we have

$$T_m M \simeq (T^*_m M)^* \simeq \mathrm{Hom}_{\mathbb{R}}(C^\infty(M)/\ker d_m, \mathbb{R})\,.$$

We will prove that δ_m descends to the quotient $C^\infty(M)/\ker d_m$. After that δ_m is well-defined on $T^*_m M$ and $\delta_m(d\,f)_m = \delta_m(f)$, for all $f \in C^\infty(M)$. Finally $\delta_m \in T_m M$ and

$$L_m \delta_m = L_{\delta_m} = \delta_m \circ d_m = \delta_m\,.$$

This then completes the proof.

It remains to show that, for every $f \in C^\infty(M)$, we have $\delta_m f = 0$, if $(d\,f)_m = 0$. Take a coordinate chart $(U, \varphi = (x^1, \ldots, x^n))$ at m and set $\varphi(m) = x_0$. When using a Taylor expansion in the neighborhood of x_0, say in $\varphi(U)$ (we agree to restrict U if necessary), we get

$$f(x) - f(x_0) = \sum_i (x^i - x_0^i)\,((\partial_{x^i} f)(x_0) + \varepsilon_i(x - x_0)) =$$

$$\sum_i (x^i - x_0^i)\varepsilon_i(x - x_0) =: \sum_i p^i(x) g_i(x).$$

It follows that $f|_U = f(m) + \sum_i p^i g_i$. If γ denotes again a bump function, we obtain

$$\delta_m(f) = \delta_m(\gamma^2 f|_U) = \delta_m(f(m)) + \textstyle\sum_i \delta_m(\gamma p^i \gamma g_i)$$

$$= \textstyle\sum_i \left(\delta_m(\gamma p^i).\gamma(m) g_i(m) + \gamma(m) p^i(m).\delta_m(\gamma g_i)\right) = 0,$$

as $p^i(m) = g_i(m) = 0$. □

6.3 Tangent and cotangent bundles

In Mechanics, a vector field is a vector $\vec{v} = \vec{v}(m)$ that depends on the point m where it is "measured", i.e. is a smooth map

$$\vec{v} : \mathbb{R}^3 \ni m \mapsto \vec{v}(m) \in \{m\} \times \mathbb{R}^3 \simeq T_m\mathbb{R}^3$$

(of course, the field, e.g. a constant [with respect to time] fluid velocity, could be defined only in an open subset of $\mathbb{R}^3$). In Differential Geometry, we investigate vector fields on manifolds. A contravariant vector field (resp. covariant vector field) on a manifold M is a map

$$X : M \ni m \mapsto X_m \in T_mM \quad \text{(resp.} \quad \alpha : M \ni m \mapsto \alpha_m \in T^*_mM)\,.$$

As the target space of a map must be independent of its variable (here m), we define

$$TM := \coprod_{m\in M} T_mM \quad \text{(resp. } T^*M := \coprod_{m\in M} T^*_mM),$$

where $\coprod$ denotes the disjoint union. The set TM (resp. T^*M) is the tangent bundle (resp. cotangent bundle) of the manifold M.

The tangent and cotangent bundles are prototypes of vector bundles. Roughly, if to every point m of a manifold M (imagine a "horizontal" line segment [resp. a "horizontal" surface]), we attach a vector space, e.g. T_mM or T^*_mM (pictured as a "vertical" line segment [resp. a "reversed vertical" triangle] over m), the amalgamation of all these vector spaces is a vector bundle (hence, an often used mental picture of a vector bundle is a rectangle over a line segment).

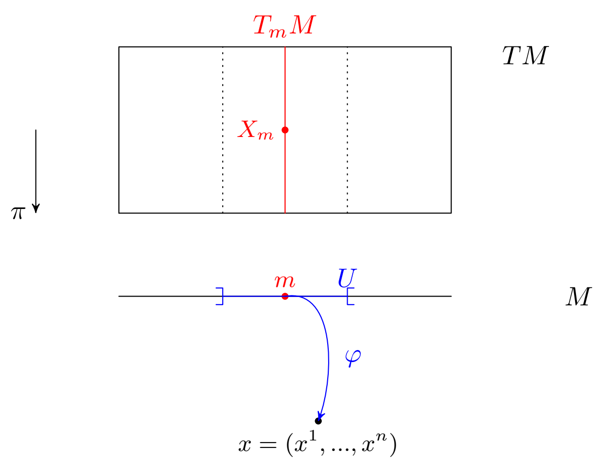

Figure 5: Tangent bundle of a manifold

More precisely, a vector bundle is made of three ingredients, the amalgamation or total space (on Figure 5: TM), the underlying manifold or base space (on Figure 5: M), and the projection that associates to every vector in the total space the corresponding base point (on Figure 5: π). The preimage of a base point by the projection is referred to as the fiber of the bundle at this point (on Figure 5: $\pi^{-1}\{m\} = T_mM$ is the fiber of TM at m).

In order to have the possibility to consider smooth contravariant (resp. covariant) vector fields

$$X : M \ni m \mapsto X_m \in TM \quad \text{(resp.} \quad \alpha : M \ni m \mapsto \alpha_m \in T^*M)\,,$$

we must endow their target TM (resp. T^*M) with a manifold structure, i.e. we must assign coordinates to every vector in TM (resp. in T^*M).

Each vector of TM belongs to a single tangent space T_mM ($m \in M$). Denote the vector under consideration by X_m and note that it has a unique decomposition

$$X_m = \sum_i X_m^i \partial_{x^i}$$

in the basis $(\partial_{x^1}, \ldots, \partial_{x^n})$, which is induced by a chosen coordinate chart

$$(U, \varphi) = (U, (x^1, \ldots, x^n))$$

of M around m. However, the map

$$\phi : \pi^{-1}(U) = \amalg_{\mu \in U} T_\mu M \supset T_mM \ni X_m \mapsto (X_m^1, \ldots, X_m^n) \in \mathbb{R}^n$$

is not a coordinate map of TM, since each tuple $(X_m^1, \ldots, X_m^n) \in \mathbb{R}^n$ is the image by ϕ of one vector in every tangent space $T_\mu M$ ($\mu \in U$), i.e. of infinitely many vectors in $\pi^{-1}(U)$. To fix the problem that the coordinate tuple does not inform about which point of M the considered vector is located over, we use the map

$$\Phi : \pi^{-1}(U) = \amalg_{\mu \in U} T_\mu M \supset T_mM \ni X_m \mapsto (x^1, \ldots, x^n, X_m^1, \ldots, X_m^n) \in \varphi(U) \times \mathbb{R}^n\,, \tag{16}$$

where $(x^1, \ldots, x^n)$ are the coordinates of m in the chosen chart (U, φ) of M. This map is clearly a bijection valued in the open subset $\varphi(U) \times \mathbb{R}^n \subset \mathbb{R}^{2n}$, so that $(\pi^{-1}(U), \Phi)$ is a chart of TM – induced by the chart (U, φ) of M. The coordinates $(x^1, \ldots, x^n)$ of $m = \pi(X_m)$ are the base coordinates of X_m and give the position of its base point m; the coordinates $(X_m^1, \ldots, X_m^n)$ of X_m are the fiber coordinates of X_m and give the position of X_m in the fiber T_mM of TM at m.

If the chart (U, φ) runs through an atlas of M, the chart $(\pi^{-1}(U), \Phi)$ runs through an atlas of TM. To prove this assertion, we need to verify the three atlas axioms.

(i) The cover condition is fulfilled:

$$\cup\, \pi^{-1}(U) = \pi^{-1}(\cup\, U) = \pi^{-1}(M) = TM\,.$$

(ii) The mental completion of Figure 5 illustrates that

$$\Phi(\pi^{-1}(U) \cap \pi^{-1}(U')) = \Phi(\pi^{-1}(U \cap U')) = \varphi(U \cap U') \times \mathbb{R}^n\,,$$

where the RHS is open since we start from an atlas of M.

(iii) It remains to be checked whether the coordinate transformations are smooth. Let

$$(U', \varphi') = (U', (y^1, \ldots, y^n))$$

be a second coordinate chart of M around m and let $(\pi^{-1}(U'),\Phi')$ be the induced chart of TM :

$$\Phi' : \pi^{-1}(U') = \amalg_{\mu\in U'} T_\mu M \supset T_m M \ni X_m \mapsto (y^1,\ldots,y^n,Y^1_m,\ldots,Y^n_m) \in \varphi'(U') \times \mathbb{R}^n ,$$

where $(y^1,\ldots,y^n)$ are the coordinates of m in (U',φ') and where $(Y^1_m,\ldots,Y^n_m)$ are the coordinates of X_m in the basis $(\partial_{y^1},\ldots,\partial_{y^n})$ of T_mM, which is induced by (U',φ'), i.e. where

$$X_m = \sum_j Y^j_m \partial_{y^j} .$$

The corresponding coordinate transformation in TM is

$$\Phi'\Phi^{-1} : \varphi(U\cap U')\times\mathbb{R}^n \ni (x^1,\ldots,x^n,X^1_m,\ldots,X^n_m) \mapsto (y^1,\ldots,y^n,Y^1_m,\ldots,Y^n_m) \in \varphi'(U\cap U')\times\mathbb{R}^n .$$

All its component maps must be smooth. The n first component maps are the components of the coordinate transformation

$$\varphi'\varphi^{-1} : \varphi(U\cap U') \ni (x^1,\ldots,x^n) \mapsto (y^1,\ldots,y^n) \in \varphi'(U\cap U')$$

in M and are smooth since we start from an atlas of M. As for the n last component maps, it follows from the chain rule that

$$X_m = \sum_i X^i_m \partial_{x^i} = \sum_{ij} X^i_m \partial_{x^i} y^j \partial_{y^j} = \sum_j Y^j_m \partial_{y^j} ,$$

so that

$$Y^j_m = \sum_i X^i_m \partial_{x^i} y^j ,$$

which is a component map that is smooth with respect to the x^i and the X^i_m, since the component maps y^j are smooth as stated.

From (i), (ii) and (iii) it follows that the pairs $(\pi^{-1}(U),\Phi)$, induced by the charts (U,φ) of an n-dimensional smooth atlas of the base manifold M form a $2n$-dimensional smooth atlas of the tangent bundle TM of M, so that the tangent bundle of a manifold is a manifold of double dimension

However, since by convention we only use manifolds whose topology is Hausdorff and second countable, we still have to prove these properties.

First, let us note that the projection $\pi : TM \supset T_mM \ni X_m \mapsto m \in M$ has the coordinate form

$$\pi(x^1,\ldots,x^n,X^1_m,\ldots,X^n_m) = (x^1,\ldots,x^n),$$

which implies that π is smooth (and is a submersion). Take now two vectors X, X' in TM and try to separate them. If their projections $m := \pi(X)$ and $m' := \pi(X')$ are different, we can separate m and m' by open subsets U and U' of M. But then $\pi^{-1}(U)$ and $\pi^{-1}(U')$ are open in TM and separate X and X'. If $m = m'$, the vectors X and X' are in a chart domain $\pi^{-1}(U)$, which is homeomorphic to an open subset of $\mathbb{R}^{2n}$, so the vectors can be separated there.

As for second countability, note that a manifold, say M', which is (a priori not second countable and is) a countable union of chart domains, say U'_α $(\alpha \in \mathbb{N})$, *is* second countable.

Indeed, as any chart domain U'_α is homeomorphic to an open subset of $\mathbb{R}^{n'}$, it is also second countable. Let now W be an open subset of M'. Since $W = \cup_{\alpha\in\mathbb{N}} W \cap U'_\alpha$ and $W \cap U'_\alpha$ is a union of elements of the countable basis of U'_α, the union over $\alpha \in \mathbb{N}$ of these countable bases is a countable basis of M'. But then, as the base manifold M of the tangent bundle TM is second countable, it is a countable union of chart domains, so that TM is also a countable union of chart domains.

Hence the following result:

Theorem 5. *If M is a smooth n-dimensional (Hausdorff and second countable) manifold, its tangent bundle TM is a smooth (Hausdorff and second countable) manifold of dimension $2n$.*

Because of this theorem, we can consider smooth vector fields, i.e. smooth maps

$$X : M \ni m \mapsto X_m \in T_m M \subset TM\,.$$

These basic objects will be investigated in the next chapter. Moreover, the elementary idea that the derivative of a smooth function is again a smooth function, can now be extended to maps between manifolds. In fact:

Proposition 14. *If $f \in C^\infty(M, M')$, its derivative or tangent map*

$$Tf : TM \supset T_m M \ni X_m \mapsto T_m f(X_m) \in T_{f(m)}M' \subset TM'$$

is smooth.

The proof of this proposition is easy and will not be given.

Of course, the cotangent bundle T^*M can also be endowed with a manifold structure.

Theorem 6. *If M is a smooth n-dimensional (Hausdorff and second countable) manifold, its cotangent bundle T^*M is a smooth (Hausdorff and second countable) manifold of dimension $2n$.*

Proof. The proof of Theorem 6 is analogous to that of Theorem 5. If we denote the projection $T^*M \to M$ by π^*, every chart $(U, \varphi = (x^1, \dots, x^n))$ of M induces a chart $((\pi^*)^{-1}(U), \Phi^*)$ of T^*M, whose coordinate map is defined by

$$\Phi^* : (\pi^*)^{-1}(U) \supset T^*_m M \ni \alpha^m \mapsto (x^1, \dots, x^n, \alpha^m_1, \dots, \alpha^m_n) \in \varphi(U) \times \mathbb{R}^n, \tag{17}$$

where $x = (x^1, \dots, x^n)$ are the coordinates of the projection $m = \pi^*(\alpha^m)$ in the considered chart of M and where $(\alpha^m_1, \dots, \alpha^m_n)$ are the components of α in the induced basis dx^i of $T^*_m M$. The transition maps are here

$$\Phi'^*(\Phi^*)^{-1} : (x, \alpha^m_1, \dots, \alpha^m_n) \mapsto (y(x), \partial_{y^1} x^j|_{y=y(x)}\alpha^m_j, \dots, \partial_{y^n} x^j|_{y=y(x)}\alpha^m_j).$$

They are obviously smooth with respect to $x = (x^1, \dots, x^n)$ and $(\alpha^m_1, \dots, \alpha^m_n)$. □

Although in this Lecture script no general vector bundles or even fiber bundles are treated—these geometric objects are similar to vector bundles, except that the "space" over every base point, the fiber over this point, is not a vector space, but an arbitrary manifold—some additional comments on bundles are appropriate. The equations (16) and (17) show that a fiber bundle is locally diffeomorphic to a product manifold, although globally it can be a more complex topological object. For example, the Möbius strip (August Ferdinand Möbius, 1790 - 1868, German mathematician and theoretical astronomer) is a fiber bundle over the base manifold S^1. Its fiber $]-1,1[$, whose 0 point is glued to every point $m \in S^1$, undergoes a half rotation when m passes through S^1. It is obvious that the Möbius strip has a smooth 2-dimensional atlas consisting of two charts, for instance

$$\varphi_1 : U_1 \to]0, 2\pi[\times]-1,1[\quad \text{and} \quad \varphi_2 : U_2 \to]-\pi,\pi[\times]-1,1[\ .$$

The transition diffeomorphism is

$$\varphi_2\varphi_1^{-1} : \left.\begin{array}{c}]0,\pi[\times]-1,1[\\ ]\pi,2\pi[\times]-1,1[\end{array}\right\} \ni (x,X) \mapsto \left\{\begin{array}{c} (x,X) \in]0,\pi[\times]-1,1[\\ (x,-X) \in]\pi,2\pi[\times]-1,1[\end{array}\right. .$$

Consequently, the Möbius bundle is locally diffeomorphic to product manifolds, although globally it is topologically more complicated. Although the aim of this remark is to emphasize that fiber bundles are in general only locally trivial, i.e. are in general only locally diffeomorphic to a product manifold, it should be mentioned that there are of course exceptions. For example, since $\mathbb{R}^n$ is a manifold with a single global chart $(\mathbb{R}^n, \mathrm{id})$, Equation (16) shows that the tangent bundle $T\mathbb{R}^n$ is globally diffeomorphic to the product manifold $\mathbb{R}^n \times \mathbb{R}^n \simeq \mathbb{R}^{2n}$.

Chapter 5
Differential Equations on Manifolds

1 Definition

In Chapter 4, we defined a smooth vector field on a manifold M as a smooth map $X : M \ni m \mapsto X_m \in TM$, such that $X_m \in T_mM$, for all $m \in M$. The last requirement admits the equivalent formulation used in

Definition 1. *Let TM be the tangent bundle of a manifold M and let $\pi : TM \to M$ be the corresponding projection onto the base. A smooth vector field of M is a smooth map $X : M \to TM$, such that $\pi \circ X = \mathrm{id}_M$.*

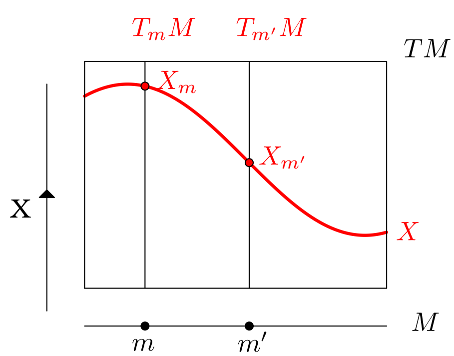

Figure 6: Vector field of a manifold

Figure 6 suggests to interpret a smooth vector field of a manifold M as a smooth section of the tangent bundle TM of M. The standard notation for the set of smooth vector fields of M is $\mathrm{Vect}(M)$ and the one for the set of smooth sections of TM is $\Gamma(TM)$: $\mathrm{Vect}(M) = \Gamma(TM)$.

In particular, a smooth vector field of $\mathbb{R}^n$ is a smooth map $X : \mathbb{R}^n \ni m \mapsto X_m \in T_m\mathbb{R}^n \subset T\mathbb{R}^n$ and, since $T_m\mathbb{R}^n \simeq \mathbb{R}^n$ and $T\mathbb{R}^n \simeq \mathbb{R}^{2n}$, we get

$$\mathrm{Vect}(\mathbb{R}^n) \simeq C^\infty(\mathbb{R}^n, \mathbb{R}^n)\,,$$

so we recover the original definition of a vector field of $\mathbb{R}^n$ as a family of vectors of $\mathbb{R}^n$ indexed by the point of $\mathbb{R}^n$.

In the following, all vector fields of M or, equivalently, sections of TM will implicitly be assumed to be smooth.

2 Playing with local forms

The following observations provide insight into the practical handling of coordinate forms.

Remember that if $f \in C^\infty(M, M')$ and if $(U', \varphi' = (y^1, \ldots, y^{n'}))$ is a chart of M' and $(U, \varphi = (x^1, \ldots, x^n))$ a chart of M, such that $f(U) \subset U'$, the local form of f in these coordinate systems is $\varphi' f \varphi^{-1}$. If we agree to identify a point of a chart domain with its coordinates in the corresponding chart map, this local form reads

$$(\varphi' f \varphi^{-1})(x) = (y^1(f(x)), \ldots, y^{n'}(f(x))) = (f^1(x), \ldots, f^{n'}(x)) =: \mathbf{f}(x),$$

where x denotes the coordinates in the source chart of the variable $m \in U$ of f and where $\mathbf{f}(x)$ is the point of $\mathbb{R}^{n'}$ made of the coordinates in the target chart of the image $f(x) \in U'$. If, for instance, $f \in C^\infty(M)$, its local form is

$$(f\varphi^{-1})(x) = f(x)\,.$$

Smoothness of f just means that these coordinate forms $\mathbf{f}(x)$ or $f(x)$ are smooth.

If we use the notation of Chapter 4, Subsection 6.3, the coordinate form of a vector field $X \in \mathrm{Vect}(M) \subset C^\infty(M, TM)$ is

$$(\Phi X \varphi^{-1})(x) = (x, X^1(x), \ldots, X^n(x)) \simeq (X^1(x), \ldots, X^n(x)) =: \mathbf{X}(x),$$

where the $X^i(x)$ are the components of X_m in the basis of T_mM induced by the coordinates x^i. Smoothness of X means that these components X^i are smooth with respect to $x \in \varphi(U)$ or, equivalently, smooth with respect to $m \in U$. In other words, smoothness of a vector field of M means that, in every coordinate chart $(U, \varphi = (x^1, \ldots, x^n))$ of an atlas of M, this field reads

$$X|_U = \sum_i X^i \partial_{x^i}, \quad \text{with} \quad X^i \in C^\infty(U), \forall i. \tag{1}$$

Often we also refer to *this* formula as the local form of a vector field.

These examples show that a diffeomorphism $\varphi \in \mathrm{Diff}(U, \varphi(U))$ transforms a function $f \in C^\infty(U)$ into a function $\varphi_* f \in C^\infty(\varphi(U))$, defined by

$$(\varphi_* f)(x) := f(x) = (f\varphi^{-1})(x), \tag{2}$$

and changes a vector field $X \in \mathrm{Vect}(U)$ into a vector field $\varphi_* X \in \mathrm{Vect}(\varphi(U)) \simeq C^\infty(\varphi(U), \mathbb{R}^n)$, defined by

$$(\varphi_* X)_x := \mathbf{X}(x) = T_{\varphi^{-1}(x)}\varphi\, X_{\varphi^{-1}(x)}. \tag{3}$$

Indeed, for any $x \in \varphi(U) \rightleftarrows m \in U$, we have

$$T_{\varphi^{-1}(x)}\varphi\, X_{\varphi^{-1}(x)} = T_m\varphi X_m = T_m\varphi \sum_i X^i_m \partial_{x^i}|_m = T_m\varphi \sum_i X^i_m (T_m\varphi)^{-1}(e_i)$$

$$= T_m\varphi(T_m\varphi)^{-1}\sum_i X^i_m e_i = \mathbf{X}(x).$$

These examples are only the shadow of the general fact that diffeomorphisms act on most objects and transform them into objects of the same type. For instance, if $\phi \in \mathrm{Diff}(M, M')$ and $X \in \mathrm{Vect}(M)$, it is natural to define the action $\phi_* X \in \mathrm{Vect}(M')$ of ϕ on X or pushforward of X by ϕ using exactly the same formula as above, i.e. to set

$$(\phi_* X)_{m'} := T_{\phi^{-1}(m')}\phi\, X_{\phi^{-1}(m')} \in T_{m'}M', \tag{4}$$

for all $m' \in M'$. Since $\phi_* X = (T\phi)X\phi^{-1}$, the vector field $\phi_* X$ is smooth (recall that we have chosen to consider only smooth vector fields in these notes). It is instructive to check this fact by looking at the local form of $\phi_* X$.

Of course, the local form of a derivative

$$T_m f = (T_{f(m)}\varphi')^{-1}\,(\varphi' f \varphi^{-1})'(\varphi(m))\; T_m\varphi,$$

with self-explanatory notations, is the linear map or matrix

$$(\varphi' f \varphi^{-1})'(\varphi(m)) = \partial_x \mathbf{f}(x) = (\partial_{x^j} f^i(x))_{ij},$$

where the RHS is evaluated at $x = \varphi(m)$. More precisely, we have the

Proposition 1. *If $f \in C^\infty(M, M')$ and $m \in M$, and if $(U, \varphi = (x^1, \ldots, x^n))$ is a chart of M at m and $(U', \varphi' = (y^1, \ldots, y^{n'}))$ is a chart of M', such that $f(U) \subset U'$, the derivative*

$$T_m f \in \mathrm{Hom}_{\mathbb{R}}(T_m M, T_{f(m)}M')$$

at m of f is characterized, in the induced bases $\partial_{x^i}|_m$ of $T_m M$ and $\partial_{y^i}|_{f(m)}$ of $T_{f(m)}M'$, by the derivative

$$\partial_x \mathbf{f}(x) = (\partial_{x^j} f^i(x))_{ij}$$

at $x = \varphi(m)$ of the coordinate form of f in the considered charts.

Proof. Remember that the matrix of a linear map $\ell \in \mathrm{Hom}_{\mathbb{R}}(V, V')$, where V and V' are finite-dimensional real vector spaces, in two bases $(t_j)_j$ of V and $(t'_i)_i$ of V', is given by $c'^i(\ell(t_j))$, where $(c'^i)_i$ is the dual basis of $(t'_i)_i$. It then suffices to note that in the present case

$$t_j = \partial_{x^j}|_m = (T_m\varphi)^{-1}(e_j) \quad \text{and} \quad c'^i = (dy^i)_{f(m)} = \widetilde{T_{f(m)}}\varphi'(\varepsilon^i)\,,$$

and to use the above definition of $T_m f$. □

The local form of

$$(\phi_* X)_{m'} := T_m\phi\, X_m \in T_{m'}M'\,,$$

where $m = \phi^{-1}(m')$ (see above), can now easily be found. We get

$$\partial_x \boldsymbol{\phi}(x)\mathbf{X}(x)|_{x=\boldsymbol{\phi}^{-1}(y)}\,.$$

Note that the coordinate form of the "composite object"

$$T_{\phi^{-1}(m')}\phi\, X_{\phi^{-1}(m')}$$

is the "object" which is composed in the same way of the local forms of the components. As a byproduct, we again find that $\phi_* X$ is smooth.

3 Algebraic structures on $\mathrm{Vect}(M)$

Just as addition and multiplication in $\mathbb{R}$ induce similar operations on functions $f : M \ni m \mapsto f(m) \in \mathbb{R}$, the vector space structures of the tangent spaces T_mM ($m \in M$) induce a vector space structure on vector fields $X : M \ni m \mapsto X_m \in T_mM \subset TM$. If $r \in \mathbb{R}$ and $X, Y \in \mathrm{Vect}(M)$, the vector fields $X + Y \in \mathrm{Vect}(M)$ and $rX \in \mathrm{Vect}(M)$ are of course defined pointwise, i.e. defined by $(X + Y)_m = X_m + Y_m$ and $(rX)_m = r\,X_m$. The smoothness of $X+Y$ and rX follows from the above remarks on the smoothness of a vector field. Of course, the new operations inherit the properties of the inducing vector space operations, so that $\mathrm{Vect}(M)$ is indeed a real vector space for the induced operations.

It is possible to substitute functions $f \in C^\infty(M)$ for reals $r \in \mathbb{R}$, i.e. to set $(fX)_m = f(m)\,X_m$, but the "scalars" f then live in a ring $C^\infty(M)$ (even in an associative commutative unital algebra) and not in a field $\mathbb{R}$. Apart from this difference, this leads to a similar structure on $\mathrm{Vect}(M)$ as before, which we call the module structure of $\mathrm{Vect}(M)$ over $C^\infty(M)$.

Proposition 2. *The set $\mathrm{Vect}(M)$ of vector fields of a manifold M is a vector space over the field $\mathbb{R}$ of real numbers and a module over the ring $C^\infty(M)$ of smooth functions of M.*

The algebraic interpretation of the tangent space, $T_mM \simeq \mathrm{Der}_m(C^\infty(M))$, $m \in M$, leads to an algebraic characterization of vector fields.

Theorem 1. *The vector space $\mathrm{Vect}(M)$ is isomorphic with the vector space $\mathrm{Der}(C^\infty(M))$ of derivations of the algebra $C^\infty(M)$:*

$$\mathrm{Vect}(M) \simeq \mathrm{Der}(C^\infty(M))\,.$$

Let us first clarify that:

Definition 2. *A derivation of the algebra $C^\infty(M)$ is an endomorphism $\delta \in \mathrm{End}(C^\infty(M))$ that satisfies the Leibniz rule, i.e. the condition $\delta(f.g) = \delta(f).g + f.\delta(g)$, for all $f, g \in C^\infty(M)$. The vector space of all derivations of $C^\infty(M)$ is denoted by $\mathrm{Der}(C^\infty(M))$.*

Proof of Theorem 1. We need to define an isomorphism

$$L : \mathrm{Vect(M)} \ni X \mapsto L_X \in \mathrm{Der}(C^\infty(M))\,,$$

that is, for each $X \in \mathrm{Vect}(M)$ and each $f \in C^\infty(M)$ we must define a function $L_X f \in C^\infty(M)$ and show that $L_X f$ is linear in f and satisfies Leibniz's rule. To define the function $L_X f$, we set

$$(L_X f)(m) := L_{X_m} f = (d\,f)_m(X_m) \in \mathbb{R}\,,$$

for each $m \in M$. The properties of L follow from the similar properties of

$$L_m : T_mM \ni X_m \mapsto L_{X_m} \in \mathrm{Der}_m(C^\infty(M))\,.$$

The detailed proof of the last statement is not particularly exciting and is not given here. $\square$

In addition to its real vector space and function-module structures, the set $\mathrm{Vect}(M)$ has a third important algebraic structure which we now describe.

If $\star$ is an associative algebra multiplication on a vector space $(A, +, \cdot)$, it induces the commutator-bracket-multiplication

$$[-,-]_c : A \times A \ni (a,b) \mapsto a \star b - b \star a \in A \tag{5}$$

on A. However, the commutator bracket is not a new associative multiplication on A since

$$[a,[b,c]] \neq [[a,b],c]$$

but

$$[a,[b,c]] = [[a,b],c] + [b,[a,c]]\,. \tag{6}$$

The property (6) which means that $[a,-]$ acts as a derivation on the product $[b,c]$, is referred to as the Jacobi identity. Further, it is clear that the multiplication $[-,-]$ is antisymmetric, i.e. that

$$[b,a] = -[a,b]\,. \tag{7}$$

This observation motivates the following

Definition 3. *A Lie algebra (LA for short) structure on a vector space $(V, +, \cdot)$ is a bilinear map*

$$[-,-] : V \times V \to V$$

on V that satisfies

- $[v,u] = -[u,v]$ *(antisymmetry (AS for short)) and*
- $[u,[v,w]] = [[u,v],w] + [v,[u,w]]$ *(Jacobi identity (JI for short)),*

for all $u, v, w \in V$.

Proposition 3. *Every associative algebra structure $\star$ on a vector space $(A, +, \cdot)$ induces a Lie algebra structure on $(A, +, \cdot)$ which is given by the commutator bracket $[-,-]_c$ defined by Equation (5).*

Now note that the usual composition $\circ$ of maps provides the vector space $\mathrm{End}(V)$ of endomorphisms of an underlying vector space V with an associative algebra structure. Therefore, the commutator bracket

$$[\ell, \ell']_c = \ell \circ \ell' - \ell' \circ \ell\,,$$

$\ell, \ell' \in \mathrm{End}(V)$, defines a LA structure on the endomorphism space. Now, if one chooses as the underlying vector space V the space $C^\infty(M)$ of functions of a manifold, one obtains a LA structure on the space $\mathrm{End}(C^\infty(M))$ and also on its subspace $\mathrm{Der}(C^\infty(M))$ of derivations of functions. To prove that this subspace inherits the LA structure from the endomorphism space, we need only show that the subspace is closed under the commutator bracket, i.e. that if

$$\delta, \delta' \in \mathrm{Der}(C^\infty(M)) \subset \mathrm{End}(C^\infty(M))\,,$$

then

$$[\delta, \delta']_c \in \mathrm{Der}(C^\infty(M))\,,$$

or, more explicitly, the commutator bracket of two endomorphisms satisfying the Leibniz rule itself satisfies this rule. The interesting aspect here is that it turns out that the commutator bracket satisfies the Leibniz rule, but not the two terms whose difference it is. Verifying all this is easy, we leave it to the reader.

Let us now recall the vector space isomorphism

$$L : \mathrm{Vect}(M) \rightleftarrows \mathrm{Der}(C^\infty(M)) \tag{8}$$

from Theorem 1. Explicitly, the Lie derivative L is a bijection preserving internal addition and external multiplication by scalars. It allows us to naturally transfer the LA structure from the derivations to the vector fields by defining the bracket $[X, Y]$ of two vector fields $X, Y \in \mathrm{Vect}(M)$ by

$$[X, Y] := L^{-1}_{[L_X, L_Y]_c} \in \mathrm{Vect}(M)\,. \tag{9}$$

Since the inducing bracket $[-,-]_c$ on the derivations is a Lie bracket, i.e. satisfies the Jacobi identity and is antisymmetric, it is fairly obvious and easy to verify that the same holds for the induced bracket $[-,-]$ on the vector fields. Thus, the set $\mathrm{Vect}(M)$ of vector fields of a manifold M has a rich and interesting algebraic structure: it is a real vector space, a $C^\infty(M)$-module, and a real Lie algebra. Since the defining equation (9) can be equivalently written as

$$L_{[X,Y]} := [L_X, L_Y]_c\,,$$

the bijection L not only preserves addition and multiplication by scalars, but it also respects the Lie bracket: the bijection L is not only an isomorphism of vector spaces, but even an isomorphism of Lie algebras.

Theorem 2. *The real vector space* $\mathrm{Vect}(M)$ *of the vector fields of a manifold* M *is a Lie algebra with bracket* $[-,-]$ *which is isomorphic to the Lie algebra* $\mathrm{Der}(C^\infty(M))$ *of the derivations of the functions of* M *with the commutator bracket* $[-,-]_c$*, and the Lie algebra isomorphism between the two is given by the Lie derivative:*

$$L_{[X,Y]} = [L_X, L_Y]_c\,, \quad \textit{for all} \quad X, Y \in \mathrm{Vect}(M)\,. \tag{10}$$

We often identify a vector field with the corresponding derivation. For example, if $(U, (x^1, \ldots, x^n))$ is a coordinate chart of M the local form of a vector field $X \in \mathrm{Vect}(M)$ in these coordinates is

$$X|_U = \sum_i X^i \partial_{x^i}\,,$$

where the LHS is a vector field on U and the RHS is the corresponding derivation of $C^\infty(U)$ (see Equation (1)). More generally, if we identify X with L_X, the equation (10) becomes

$$[X, Y] = [X, Y]_c = X \circ Y - Y \circ X\,,$$

where X, Y are considered as vector fields in the LHS and as derivations in the subsequent parts of the last equation. From here it follows that if

$$X|_U = \sum_i X^i \partial_{x^i} \quad \text{and} \quad Y|_U = \sum_j Y^j \partial_{x^j}\,,$$

then

$$[X,Y]|_U = \sum_i X^i \partial_{x^i} \left(\sum_j Y^j \partial_{x^j}(-) \right) - \sum_j Y^j \partial_{x^j} \left(\sum_i X^i \partial_{x^i}(-) \right) = \ldots = \sum_j \left(X(Y^j) - Y(X^j) \right) \partial_{x^j} \, , \tag{11}$$

where X, Y are viewed as vector fields in the LHS and as derivations or Lie derivatives in the RHS. The same type of proof allows us to determine the behavior of the $\mathbb{R}$-bilinear bracket $[-,-]$ of vector fields with respect to the module structure of the vector fields over the functions. If $f, g \in C^\infty(M)$, then

$$[fX, gY] = fX\,(gY(-)) - gY\,(fX(-)) = \ldots = fg\,[X,Y] + f\,X(g)\,Y - g\,Y(f)\,X\,, \tag{12}$$

where X, Y are considered as vector fields except when acting on a function where they are considered as derivations. We encourage the reader to either memorize the formulas (11) and (12) or practice finding them quickly when needed.

4 Differential equations on manifolds

Consider a particle p moving in a manifold M. Let $m \in M$ be an arbitrary initial position and let $\phi_t(m) \in M$ be the position of p at time $t \in \mathbb{R}$. It is clear that

$$\phi_0(m) = m \quad \text{and} \quad \phi_t(\phi_s(m)) = \phi_{t+s}(m). \tag{13}$$

If

$$\phi \in C^\infty(\mathbb{R} \times M, M) \text{ and, for every fixed } t \in \mathbb{R},\ \phi_t \in \mathrm{Diff}(M), \tag{14}$$

it follows from the previous properties that the maps ϕ_t form a subgroup of the group $\mathrm{Diff}(M)$, or, better, a 1-parameter group of diffeomorphisms.

Such a 1-parameter group $\{\phi_t : t \in \mathbb{R}\}$ of diffeomorphisms of M defines a vector field of M. Indeed, for every fixed $m \in M$, the map $\phi(m) = \phi_t(m) \in C^\infty(\mathbb{R}, M)$ is a curve of M that passes through m at $t = 0$, so

$$X_m := d_t|_{t=0}\phi_t(m) \in T_m M$$

defines a vector field X of M.

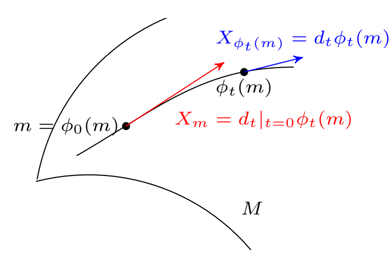

Figure 7: 1-parameter group and vector field

In the following, we examine the opposite problem, i.e. we start from a vector field and try to build a 1-parameter group of diffeomorphisms from it. More precisely, if $X \in \mathrm{Vect}(M)$, is it then possible to integrate X, i.e. to find curves $\alpha \in C^\infty(I, M)$ in M (I open interval in $\mathbb{R}$) such that

$$d_t\alpha = X_{\alpha(t)}, \forall t \in I\,?$$

Furthermore, do the solutions of this equation form a 1-parameter group of diffeomorphisms?

Definition 4. *An integral curve of a vector field $X \in \mathrm{Vect}(M)$ of a manifold M, is a curve $\alpha \in C^\infty(I, M)$ of M (I denotes an open interval of $\mathbb{R}$) such that*

$$d_t\alpha = X_{\alpha(t)}, \forall t \in I. \tag{15}$$

In view of the above remarks on local forms, it is clear that in local coordinates (U, φ) of M the equation (15) reads

$$d_t\boldsymbol{\alpha} = \mathbf{X}_{\boldsymbol{\alpha}(t)}, \forall t \in I, \tag{16}$$

provided that $\alpha(I) \subset U$. The last equation has the form $\mathcal{X}(t, \boldsymbol{\alpha}, d_t\boldsymbol{\alpha}) = 0$ of a known relation $\mathcal{X}$ between an unknown function $\boldsymbol{\alpha}$, its derivative $d_t\boldsymbol{\alpha}$ and its variable t, and is therefore an ordinary differential equation (ODE). As in fact $\mathcal{X}$ does not explicitly depend on the variable t in the present case, Equation (16) is a so-called autonomous ODE. In Analysis, there exists a local existence and uniqueness theorem for this type of equation. Of course, the statements of this theorem can be transferred to Equation (15). They then read as follows:

Theorem 3. *Let $X \in \mathrm{Vect}(M)$. For every $t_0 \in \mathbb{R}$ and $m_0 \in M$, there are open neighborhoods $]t_0 - \varepsilon, t_0 + \varepsilon[$, $\varepsilon > 0$, of t_0 in $\mathbb{R}$ and U of m_0 in M, such that for every value $m \in U$, there is an integral curve $\alpha(m) = \alpha(t, m)$ of X, which is defined in the interval $]t_0 - \varepsilon, t_0 + \varepsilon[$ – which is independent of m – and passes through the point m at time $t = t_0$. Moreover, $\alpha \in C^\infty(]t_0 - \varepsilon, t_0 + \varepsilon[\times U, M)$, and if $\beta \in C^\infty(I, M)$ and $\gamma \in C^\infty(J, M)$ are two integral curves of X that coincide at one point $t_1 \in I \cap J$, the they coincide everywhere they are both defined, i.e. on $I \cap J$.*

We could of course view the points m as the initial values of α at the initial time $t = t_0$. However, we restrict the use of the term "initial value" to the case $t = t_0 = 0$. Let now m be any point in M and consider all the integral curves $\alpha_\iota \in C^\infty(I_\iota, M)$ of X with initial value m. As any two such curves α_i and α_j coincide at the point $0 \in I_i \cap I_j$, they coincide everywhere in $I_i \cap I_j$. Hence, we can glue all the α_ι together – remark that, due to the previous local existence and uniqueness theorem, there is at least one such curve – and construct a unique maximal integral curve of X with initial value m. This maximal curve $\phi(m) = \phi_t(m)$ is defined in $I_m = \cup_\iota I_\iota$ by $\phi_t(m) = \alpha_\iota(t)$, if $t \in I_\iota$. It is obvious that we obtain this way a well-defined smooth integral curve of X with initial value m and that this curve is maximal and unique.

Proposition 4. *Let $X \in \mathrm{Vect}(M)$. For every point $m \in M$, there exists a unique maximal integral curve $\phi_t(m)$ of X with initial value m.*

Take now any integral curve $\phi_t(m)$ and any point $\phi_s(m)$ on it. *Observe that we assume here that $s \in I_m$.* As a consequence of the last proposition, there is a unique maximal integral curve $\phi_t(\phi_s(m))$ that admits the point $\phi_s(m)$ as initial value. Of course, we ask if $\phi_t(\phi_s(m)) = \phi_{t+s}(m)$ and if $I_{\phi_s(m)} = I_m - s$. These results are reworded in

Proposition 5. *Let $m \in M$. If $s \in I_m$, then*

$$t \in I_{\phi_s(m)} \Leftrightarrow t + s \in I_m,$$

and in this case, we have

$$\phi_t(\phi_s(m)) = \phi_{t+s}(m). \tag{17}$$

Thus Property (17) is valid if the inner map of the LHS is defined and either the LHS composite map is defined or the RHS map is defined.

Lemma 1. *Let* $X \in \mathrm{Vect}(M)$. *If* $\alpha = \alpha(t)$ *is an integral curve of* X *with domain* I, *then* $\alpha(\cdot + s) = \alpha(t+s)$ *is an integral curve of* X *with domain* $I - s$, *for every* $s \in \mathbb{R}$.

We first prove this lemma.

Proof. It is clear that $\alpha(\cdot + s) \in C^\infty(I - s, M)$. Further, for every $t \in I - s$, we have

$$d_t(\alpha(t+s)) = (d_\tau \alpha)(t+s).1 = X_{\alpha(t+s)}.$$

□

Now we establish the above proposition.

Proof. In view of the lemma, the RHS $\phi_{t+s}(m)$ of Equation (17) is an integral curve of X with domain $I_m - s$ and initial value $\phi_s(m)$ (remark that we used here the fact that $s \in I_m$). It follows that

$$I_m - s \subset I_{\phi_s(m)} \tag{18}$$

and that for every $t \in I_m - s$, we have

$$\phi_{t+s}(m) = \phi_t(\phi_s(m)). \tag{19}$$

When applying the lemma, for a translation by $-s$, to the LHS $\phi_t(\phi_s(m))$ (we use again the information $s \in I_m$), which is an integral curve of X with domain $I_{\phi_s(m)}$, we find that $\phi_{t-s}(\phi_s(m))$ is an integral curve of X that is defined in $I_{\phi_s(m)} + s$ and has initial value (the fact that 0 lies in the domain comes from Equation (18)) m (this results from Equation (19)). Hence, $I_{\phi_s(m)} + s \subset I_m$. □

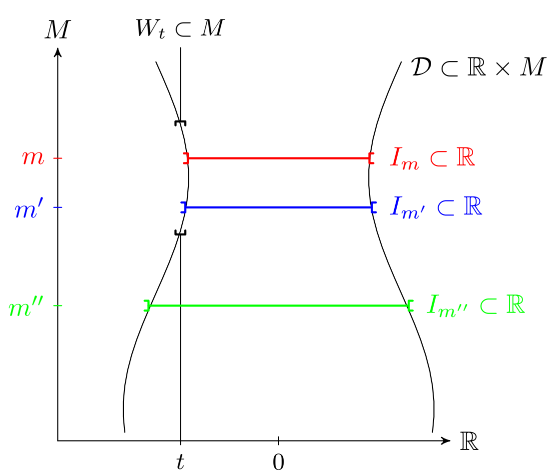

Figure 8: Domains $\mathcal{D}$, I_m and W_t

We understand that the image of several maximal integral curves of a vector field $X \in \mathrm{Vect}(M)$ resembles a liquid flow. Hence, the family of all maximal curves, $\phi = \phi(t,m) := \phi_t(m)$, $t \in I_m, m \in M$, is called the flow of the vector field X. The domain of the flow ϕ is $\mathcal{D} = \{(t,m) \in \mathbb{R} \times M : t \in I_m\}$ (see Figure 8). If we fix in $\phi = \phi(t,m)$ the variable m, we recover of course the integral curve $\phi_t(m)$ ($\phi_m = \phi_m(t)$ would be a better notation, but we'll stick to the traditional ones), which is defined in I_m, i.e. in the section of $\mathcal{D}$ at level m. If we fix t, we obtain a map $\phi_t = \phi_t(m)$ that is defined in the section $W_t \subset M$ of $\mathcal{D}$ at level t. Finally, we have $\phi : \mathcal{D} \subset \mathbb{R} \times M \to M$ and $\phi_t : W_t \subset M \to M$. As we thus get close to the concept of 1-parameter group of diffeomorphisms, see above, three natural questions arise: can we prove that $\mathcal{D}$ is open and ϕ smooth, is ϕ_t a diffeomorphism, and under which condition do we have $\mathcal{D} = \mathbb{R} \times M$ and $W_t = M$?

Theorem 4. *Let $X \in \mathrm{Vect}(M)$ and denote by I_m the domain of the maximal integral curve $\phi_t(m)$ of X with initial value $m \in M$. The source $\mathcal{D} = \{(t,m) \in \mathbb{R} \times M : t \in I_m\}$ of the flow $\phi(t,m)$ of X is an open subset of $\mathbb{R} \times M$ and $\phi \in C^\infty(\mathcal{D}, M)$.*

It suffices to prove that for every $(t_0, m_0) \in \mathcal{D}$, there are open neighborhoods $]t_0 - \varepsilon, t_0 + \varepsilon[$ of t_0 in $\mathbb{R}$ and U of m_0 in M, such that $]t_0 - \varepsilon, t_0 + \varepsilon[\times U \subset \mathcal{D}$ and $\phi \in C^\infty(]t_0 - \varepsilon, t_0 + \varepsilon[\times U, M)$.

At first sight this requirement seems to be a direct consequence of Theorem 3, which guarantees the existence of a flow

$$\alpha = \alpha(t,m) \in C^\infty(]t_0 - \varepsilon, t_0 + \varepsilon[\times U, M)\,.$$

However, for every $m \in U$, the integral curve $\alpha(t,m)$ *passes through* m *at time* $t = t_0$ *and not at time* $t = 0$. The curve $\alpha(t + t_0, m)$ is an integral curve, which is defined in $]-\varepsilon, \varepsilon[$ and *has initial value* m, so that $]-\varepsilon, \varepsilon[\subset I_m$, for all $m \in U$, and

$$\phi = \phi(t,m) = \alpha(t + t_0, m) \in C^\infty(]-\varepsilon, \varepsilon[\times U, M)\,.$$

Finally it follows from Theorem 3 that, for every $m_0 \in M$, there is an "open box" $]-\varepsilon, \varepsilon[\times U$ around $(0, m_0)$, which sits inside $\mathcal{D}$ and on which ϕ is smooth. In other words, the above requirement is a priori only satisfied for the points of the type $(0, m_0)$.

The following idea underlies the extension of this conclusion to any point $(t_0, m_0) \in \mathcal{D}$. As

$$\phi(t - t_0, m) \in C^\infty(]t_0 - \varepsilon, t_0 + \varepsilon[\times U, M),$$

if in addition $\phi(t_0, m) \in C^\infty(U, U)$, then, on the one hand

$$\phi(t - t_0, \phi(t_0, m)) \in C^\infty(]t_0 - \varepsilon, t_0 + \varepsilon[\times U, M),$$

and on the other, this last mapping coincides with $\phi = \phi(t,m)$. However, the implementation of this idea is somewhat tricky.

Proof. Let $(t_0, m_0) \in \mathcal{D}$. The proof consists of three steps.

Step 1: Construction of an "open box" $]-\varepsilon, \varepsilon[\times V$, such that $\phi(t,m) \in C^\infty(]-\varepsilon, \varepsilon[\times V, M)$ and $V \supset \phi([0, t_0], m_0)$

For any $t_\iota \in [0, t_0] \subset I_{m_0}$, we have $\phi(t_\iota, m_0) \in M$, and there exists, see above, a box $]-\varepsilon_\iota, \varepsilon_\iota[\times U_\iota$ around $(0, \phi(t_\iota, m_0))$, such that $\phi \in C^\infty(]-\varepsilon_\iota, \varepsilon_\iota[\times U_\iota, M)$. As $U_\iota \ni \phi(t_\iota, m_0)$, it suffices to set $V := \cup_\iota U_\iota$ and $]-\varepsilon, \varepsilon[:= \cap_\iota]-\varepsilon_\iota, \varepsilon_\iota[$. Indeed, the problem concerning the last intersection can be solved by using the fact that the open covering $J_\iota := \phi(t, m_0)^{-1}U_\iota$, $J_\iota \ni t_\iota$, of the compact subset $[0, t_0]$ contains a finite subcover.

Step 2: Construction of an open neighborhood $U \subset V$ of m_0, such that $\phi(t_0, m) \in C^\infty(U, V)$

Let $p \in \mathbb{N}^*$ be an integer, such that $\frac{t_0}{p} \in]-\varepsilon, \varepsilon[$. Then,

$$\phi(\tfrac{t_0}{p}, m) \in C^\infty(V, M),$$

and the i successive inverse images

$$V_i := \phi(\tfrac{t_0}{p}, m)^{-1} \ldots \phi(\tfrac{t_0}{p}, m)^{-1} V$$

provide an open subset $V_i \subset V$, for any $i \in \{1, \ldots, p\}$. It now follows from the conclusion of Step 1 that the open subset $U := \cap_i V_i \subset V$ contains m_0. Further,

$$\phi(\tfrac{t_0}{p}, \phi(\tfrac{t_0}{p}, \ldots \phi(\tfrac{t_0}{p}, m))) \in C^\infty(U, V),$$

for (at most) p iterations. In view of Proposition 5, this entails that $\phi(t_0, m) \in C^\infty(U, V)$.

Step 3: Application of the above-detailed basic idea

The conclusions of the steps 1 and 2 imply that

$$\phi(t - t_0, \phi(t_0, m)) \in C^\infty(]t_0 - \varepsilon, t_0 + \varepsilon[\times U, M)\,,$$

so that finally

$$\phi(t, m) \in C^\infty(]t_0 - \varepsilon, t_0 + \varepsilon[\times U, M)\,,$$

where U is a neighborhood of m_0. □

For any fixed $t \in \mathbb{R}$, the map $\phi_t = \phi(t, m)$ is defined in the section $W_t = \{m \in M : I_m \ni t\}$ of $\mathcal{D}$ at level t and $\phi_t \in C^\infty(W_t, M)$. Let us turn to the question whether ϕ_t is a diffeomorphism. Observe first that $t \in I_m \Leftrightarrow m \in W_t$. Equation (17) then yields $\phi_{-t}(\phi_t(m)) = m$, for any $m \in W_t$, and $\phi_t(\phi_{-t}(m)) = m$, for any $m \in W_{-t}$. Therefore, $\phi_t \in \mathrm{Diff}(W_t, W_{-t})$ and $\phi_t^{-1} = \phi_{-t}$.

Corollary 1. *For every $t \in \mathbb{R}$, the section W_t of $\mathcal{D}$ at level t is open in M, $\phi_t \in \mathrm{Diff}(W_t, W_{-t})$, and $\phi_t^{-1} = \phi_{-t}$.*

Remark. It is interesting to remember that, for any $m_0 \in M$, Closure Equation (17) is valid for t, s near 0 and m near m_0, i.e. in an "open box" around $(0, m_0)$. As a matter of fact, the flow is defined (and even smooth) in an "open box" $]-\varepsilon, \varepsilon[\times U$ around $(0, m_0)$, so that it suffices to take $t, s \in]-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}[$ (since then $s, t+s \in]-\varepsilon, \varepsilon[$) and $m \in U$.

The obvious result $\phi_0(m) = m$, for all $m \in M$, Closure Equation (17), Theorem 4, and Corollary 1, mean that the ϕ_t, $t \in \mathbb{R}$, form a local 1-parameter group of local diffeomorphisms of M. Indeed, the maximal integral curves $\phi_t(m) \in C^\infty(I_m, M)$ are not necessarily defined on the whole real line $\mathbb{R}$. However, if M is *compact*, we *have* $I_m = \mathbb{R}$, for all $m \in M$, i.e. $W_t = M$, for all $t \in \mathbb{R}$, or, equivalently, $\mathcal{D} = \mathbb{R} \times M$.

Proposition 6. *For every vector field $X \in \mathrm{Vect}(M)$ of a compact manifold M, the maximal integral curves $\phi_t(m)$, $m \in M$, of X are all defined on the whole real line and the ϕ_t, $t \in \mathbb{R}$, form a 1-parameter group of diffeomorphisms of M.*

Proof. It suffices to prove that $\mathbb{R} \subset I_m$, for all $m \in M$. First, we explain that compactness of M allows building a "stripe" $]-\varepsilon, \varepsilon[\times M$ that sits inside $\mathcal{D}$, so that $]-\varepsilon, \varepsilon[\subset I_m$, for all $m \in M$. Then, we show that the last claim entails that the same statement holds true for $]-2\varepsilon, 2\varepsilon[$, so that, by iteration, we get $\mathbb{R} \subset I_m$, for all $m \in M$.

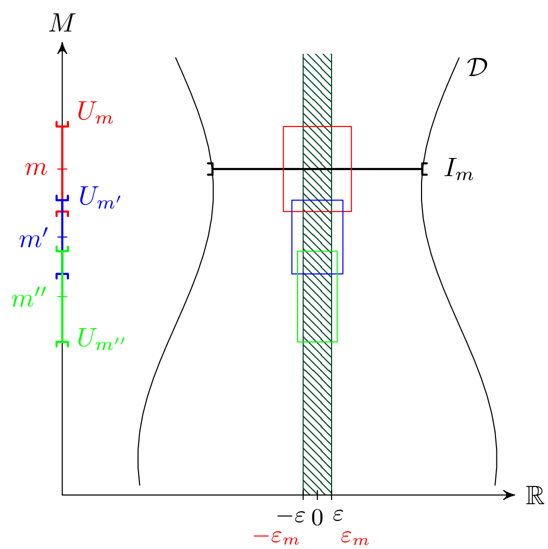

Figure 9: Problem of the possible existence of smaller and smaller boxes

We already observed that, for every $m \in M$, there is an open box $]-\varepsilon_m, \varepsilon_m[\times U_m \subset \mathcal{D}$. As the open cover U_m, $m \in M$, of the compact manifold M contains a finite covering U_i, $i \in \{1, \ldots, p\}$, the flow is defined in $]-\varepsilon, \varepsilon[\times M$, where $]-\varepsilon, \varepsilon[= \cap_i]-\varepsilon_i, \varepsilon_i[$, so that $]-\varepsilon, \varepsilon[\subset I_m$, for all $m \in M$.

In order to extend this interval to $]-2\varepsilon, 2\varepsilon[$, note that, for every $m \in M$, see Proposition 5, if $s \in I_m$, then

$$]-\varepsilon, \varepsilon[\subset I_m = s + I_{\phi_s(m)} \supset]s-\varepsilon, s+\varepsilon[.$$

Finally,

$$\tau \in]-2\varepsilon, 2\varepsilon[\Rightarrow \tfrac{\tau}{2} \in]-\varepsilon, \varepsilon[\Rightarrow \tau \in]\tfrac{\tau}{2}-\varepsilon, \tfrac{\tau}{2}+\varepsilon[\subset I_m,$$

due to the previous result (take $s = \frac{\tau}{2}$). Hence the conclusion. $\square$

Remarks.

1. If M is not compact, the flow of $X \in \mathrm{Vect}(M)$ is not necessarily defined in $\mathbb{R} \times M$.

2. For obvious reasons, the local diffeomorphisms ϕ_t, $t \in \mathbb{R}$, which are induced by a vector field $X \in \mathrm{Vect}(M)$ are often denoted by $\exp(tX)$. If we use the notation ϕ_t and must specify the underlying vector field, we write ϕ_t^X.

Exercise. Consider the vector field $X \in \mathrm{Vect}(\mathbb{R}^2)$ that is defined by $X_{(x,y)} = -y\partial_x + x\partial_y$. Show that the maximal integral curves $\phi_t(x, y)$ of X are

$$\phi_t(x, y) = (x \cos t - y \sin t, x \sin t + y \cos t)$$

and that the representation of the flow is a family of circles with center at the origin. Prove that the ϕ_t, $t \in \mathbb{R}$, form a 1-parameter group of diffeomorphisms and that this group is isomorphic to $\mathrm{SO}(2, \mathbb{R})$.

5 Lie derivative of a vector field

We defined the Lie derivative of a function $f \in C^\infty(M)$ in the direction of a vector field $X \in \mathrm{Vect}(M)$ (Marius Sophus Lie, 1842 - 1899, Norwegian mathematician) by

$$(L_X f)_m = X_m(d f)_m, \tag{20}$$

for all $m \in M$. As vector fields $X \in \mathrm{Vect}(M)$ are isomorphic to derivations $L_X \in \mathrm{Der}(C^\infty(M))$, we identify X with L_X most of the time and simply write Xf instead of $L_X f$.

When trying to define the Lie derivative of a vector field $Y \in \mathrm{Vect}(M)$ in the direction of a field $X \in \mathrm{Vect}(M)$ at a point $m \in M$, we have to compare the value $Y_m \in T_m M$ of Y at m with the value of Y at a point of M which is close to m in the direction given by X, i.e. with the value

$$Y_{\phi_t^X(m)} \in T_{\phi_t^X(m)} M,$$

$t \simeq 0$.

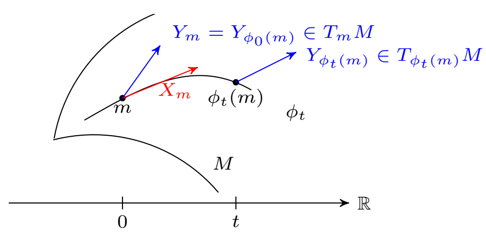

Figure 10: Problem of difference of values of Y

In order to subtract these values, we transport the second into the space $T_m M$ of the first, by means of

$$T\phi_{-t}^X : T_{\phi_t^X(m)} M \to T_m M.$$

Hence, we get

$$(L_X Y)_m = \lim_{t \to 0} \frac{T_{\phi_t^X(m)} \phi_{-t}^X Y_{\phi_t^X(m)} - Y_m}{t} = \lim_{t \to 0} \frac{(\phi_{-t,*}^X Y)_m - (\phi_{-0,*}^X Y)_m}{t} = d_t|_{t=0} (\phi_{-t,*}^X Y)_m,$$

where $\phi_{-t,*}^X$ denotes the pushforward by ϕ_{-t}^X, see Equation (4).

Of course, we have to check if the above derivative of

$$(t, m) \mapsto (\phi_{-t,*}^X Y)_m = T_{\phi_t^X(m)} \phi_{-t}^X Y_{\phi_t^X(m)}$$

actually exists. Let $m_0 \in M$ and let (t, m) vary in an open box $]-\varepsilon, \varepsilon[\times U$ around $(0, m_0)$, in which the flow $\phi^X(t, m)$ is smooth and the closure property

$$\phi^X_{-t}(\phi^X_t(m)) = \phi^X_{-t+t}(m) = m$$

is valid. It follows in particular that $\phi^X_t(m) \in W_{-t}$, where W_{-t} is the domain of $\phi^X_{-t} \in \mathrm{Diff}(W_{-t}, W_t)$. Further,

$$T_{\phi^X_t(m)}\phi^X_{-t} \in \mathrm{Isom}(T_{\phi^X_t(m)}M, T_mM).$$

As $Y_{\phi^X_t(m)} \in T_{\phi^X_t(m)}M$, we get a map

$$]-\varepsilon, \varepsilon[\times U \ni (t, m) \mapsto (\phi^X_{-t,*}Y)_m = T_{\phi^X_t(m)}\phi^X_{-t}Y_{\phi^X_t(m)} \in T_mM \subset TM.$$

In local coordinates $(x^1, \ldots, x^n)$ in U, this map reads

$$\partial_y \boldsymbol{\phi}(-t, y)|_{y=\boldsymbol{\phi}(t,x)} \mathbf{Y}_{\boldsymbol{\phi}(t,x)}, \tag{21}$$

with self-explaining notations. The preceding coordinate form is clearly smooth with respect to (t, x) and its derivative with respect to t at $t = 0$ is smooth with respect to x. Hence,

$$U \ni m \mapsto (L_XY)_m = d_t|_{t=0}\, (\phi^X_{-t,*}Y)_m = d_t|_{t=0}\, T_{\phi^X_t(m)}\phi^X_{-t}Y_{\phi^X_t(m)} \in T_mM \subset TM \tag{22}$$

is smooth, and therefore L_XY is defined at every point $m_0 \in M$ and smooth in the neighborhood of every point $m_0 \in M$. Finally $L_XY \in \mathrm{Vect}(M)$.

Definition 5. *The Lie derivative of a vector field $Y \in \mathrm{Vect}(M)$ with respect to a vector field $X \in \mathrm{Vect}(M)$ is the vector field $L_XY \in \mathrm{Vect}(M)$, which is defined for every $m \in M$ by*

$$(L_XY)_m = d_t|_{t=0}(\phi^X_{-t,*}Y)_m. \tag{23}$$

Note that the definition allows us to state that the Lie derivative is a local operator with respect to both arguments, see Equation (22)—a fact that is intuitively clear a priori.

The next theorem is quite amazing.

Theorem 5. *For every $X, Y \in \mathrm{Vect}(M)$, the Lie derivative $L_XY \in \mathrm{Vect}(M)$ of Y in the direction of X coincides with the Lie bracket $[X, Y] \in \mathrm{Vect}(M)$ of X and Y (see Equation (9)). Hence the Lie derivative*

$$L : \mathrm{Vect}(M) \times \mathrm{Vect}(M) \ni X, Y \mapsto L_XY \in \mathrm{Vect}(M)$$

endows the vector space $\mathrm{Vect}(M)$ with a Lie algebra structure.

For the definition of Lie brackets and Lie algebras, we refer the reader to Section 3.

Proof. In view of Equations (21) and (22), the local coordinate form of L_XY is

$$(L_XY)_x = d_t|_{t=0}\, \partial_y \boldsymbol{\phi}(-t, y)|_{y=\boldsymbol{\phi}(t,x)} \mathbf{Y}_{\boldsymbol{\phi}(t,x)} = d_t|_{t=0}\, \left(\partial_x \boldsymbol{\phi}(t, x)\right)^{-1} \mathbf{Y}_{\boldsymbol{\phi}(t,x)}\,.$$

Let us recall that the definition $A^{-1}A = AA^{-1} = \mathrm{id}$ of the inverse of a matrix implies that the derivative of the inverse of an invertible matrix, which depends smoothly on a variable t, is given by $d_tA^{-1} = -A^{-1}d_tA\, A^{-1}$. We therefore get

$$\begin{aligned}(L_XY)_x &= -d_t|_{t=0}\, \partial_x\boldsymbol{\phi}(t, x)\; \mathbf{Y}_x + d_t|_{t=0} \mathbf{Y}_{\boldsymbol{\phi}(t,x)} = -\partial_x\mathbf{X}\; \mathbf{Y}_x + \partial_x\mathbf{Y}\; d_t|_{t=0}\boldsymbol{\phi}(t, x)\\ &= \partial_x\mathbf{Y}\; \mathbf{X}_x - \partial_x\mathbf{X}\; \mathbf{Y}_x = \textstyle\sum_i X^i_x\, \partial_{x^i}\mathbf{Y} - \sum_i Y^i_x\, \partial_{x^i}\mathbf{X} = \sum_j (X(Y^j) - Y(X^j))_x \partial_{x^j} = [X, Y]_x\end{aligned} \tag{24}$$

(see Equation (11)). $\square$

We know from Theorem 2 that

$$L : \mathrm{Vect}(M) \to \mathrm{Der}(C^\infty(M))$$

is a Lie algebra isomorphism, so that $L_{[X,Y]} = [L_X, L_Y]_c$. It follows that for every function $f \in C^\infty(M)$, we have

$$L_{[X,Y]}f = L_X(L_Y f) - L_Y(L_X f)\,.$$

The same equality is actually valid if we replace the function $f \in C^\infty(M)$ by a vector field $Z \in \mathrm{Vect}(M)$:

$$L_{[X,Y]}Z = L_X(L_Y Z) - L_Y(L_X Z)\,. \tag{25}$$

Indeed, in view of Theorem 5, the last equality reads

$$[[X,Y],Z] = [X,[Y,Z]] - [Y,[X,Z]]$$

and is nothing but the Jacobi identity of the Lie bracket of vector fields. In fact, Equation (25) is true for all tensor fields, and not only for functions (tensor fields of type $(0,0)$) and vector fields (tensor fields of type $(1,0)$). Of course, this claim cannot yet be understood, since tensor fields and their Lie derivatives have not been defined so far.

Since the bracket $[X,Y] = L_X Y$ of vector fields is a Lie bracket on $\mathrm{Vect}(M)$, its behavior with respect to the vector space structure of $\mathrm{Vect}(M)$ is clear: the Lie bracket of vector fields is an $\mathbb{R}$-bilinear map on $\mathrm{Vect}(M)$. The next proposition clarifies the behavior of the Lie bracket with respect to the $C^\infty(M)$-module structure of $\mathrm{Vect}(M)$.

Proposition 7. *For every $X, Y \in \mathrm{Vect}(M)$ and every $f, g \in C^\infty(M)$, we have*

$$[fX, gY] = fg[X,Y] + f(Xg)Y - g(Yf)X. \tag{26}$$

Proof. It suffices to show that the LHS and the RHS derivations coincide on functions. □

6 Exercises

Exercise 1

Above, we defined the actions of diffeomorphisms (resp. vector fields) on functions and vector fields, see Equations (2) and (3) (resp. Equations (20) and (23)). In the following exercises, we further investigate these actions, as well as the link between them.

- Let $\phi \in \mathrm{Diff}(M, M')$ and $\psi \in \mathrm{Diff}(M', M'')$.

    1. Prove that

    $$\phi_* : C^\infty(M) \ni f \to f \circ \phi^{-1} \in C^\infty(M')$$

    is an associative algebra isomorphism, and that

    $$(\psi \circ \phi)_* = \psi_* \circ \phi_*\ . \tag{27}$$

2. Show that
$$\phi_* : \mathrm{Vect}(M) \ni X \to T\phi \circ X \circ \phi^{-1} \in \mathrm{Vect}(M')$$
is a Lie algebra isomorphism and that the preceding formula for pushforwards and compositions is also valid in the case of vector fields. In order to show that
$$\phi_*[X, Y] = [\phi_* X, \phi_* Y], \tag{28}$$
$X, Y \in \mathrm{Vect}(M)$, prove first that $\phi_*(Xf) = (\phi_* X)(\phi_* f)$, $X \in \mathrm{Vect}(M)$, $f \in C^\infty(M)$, then use this result to show that the LHS and RHS derivations of the equation (28) coincide on each function.

- In local coordinates, we have $d_t|_{t=0}\, \phi^i_t(x) = X^i_x$ and this equation implies that, for an infinitesimal $t = \varepsilon$,
$$\phi^i_\varepsilon(x) = x^i + \varepsilon X^i_x.$$
Because of this result, the vector field X is called the infinitesimal generator of the diffeomorphism ϕ_t .

  Let us stress that the action of a vector field X on a field Y,
$$L_X Y = d_t|_{t=0}\, \phi^X_{-t,*} Y, \tag{29}$$
is defined as the derivative at time $t = 0$ of the action on Y of the diffeomorphism $\phi^X_{-t} = \exp(-tX)$. This observation is only the shadow of a more general link between the action of certain types of objects and the action of the corresponding infinitesimal objects.

  Prove for instance that we also have
$$L_X f = d_t|_{t=0}\, \phi^X_{-t,*} f. \tag{30}$$

- Above, we remarked that the actions ϕ_* of diffeomorphisms $\phi \in \mathrm{Diff}(M)$ on functions (resp. vector fields) are associative algebra (resp. Lie algebra) automorphisms $\phi_* \in \mathrm{Aut}(C^\infty(M))$ (resp. $\phi_* \in \mathrm{Aut}(\mathrm{Vect}(M))$) (an automorphism is an isomorphism from an object onto itself). The induced actions L_X of vector fields $X \in \mathrm{Vect}(M)$ on functions (resp. vector fields) are associative algebra (resp. Lie algebra) derivations $L_X \in \mathrm{Der}(C^\infty(M))$ (resp. $L_X \in \mathrm{Der}(\mathrm{Vect}(M))$). Explain this fact.

Exercise 2

Let $X, Y \in \mathrm{Vect}(M)$. Prove that, for all $m \in M$, if t is sufficiently close to 0, the equation
$$(\phi^X_{-t,*} L_X Y)_m = d_t\, (\phi^X_{-t,*} Y)_m, \tag{31}$$
which generalizes the definition (29) of the Lie derivative of vector fields, is valid.

Suggestion: Observe that
$$d_t\, (\phi^X_{-t,*} Y)_m = d_s|_{s=0}\, (\phi^X_{-t-s,*} Y)_m\ .$$

7 Geometric interpretation of the Lie bracket of vector fields

This section shows that the Lie bracket $[X, Y]$ of vector fields measures the non-commutativity of the corresponding "flows" ϕ_t^X and ϕ_s^Y.

We first need a new concept. The pushforward

$$X' := f_* X = Tf \circ X \circ f^{-1}$$

of a vector field $X \in \mathrm{Vect}(M)$ by a function $f \in C^\infty(M, M')$ does of course *not* make sense, if the function is not a diffeomorphism. However, there can exist vector fields $X \in \mathrm{Vect}(M)$ and $X' \in \mathrm{Vect}(M')$ that satisfy the equation

$$X' \circ f = Tf \circ X,$$

and that are therefore *related* by the function f. Hence, the following weak substitute for the pushforward of vector fields by diffeomorphisms.

Definition 6. *Let $f \in C^\infty(M, M')$. Two vector fields $X \in \mathrm{Vect}(M)$ and $X' \in \mathrm{Vect}(M')$ are f-related, if*

$$T_m f X_m = X'_{f(m)}, \tag{32}$$

for all $m \in M$.

Lemma 2. *Let $f \in C^\infty(M, M')$. If $X \in \mathrm{Vect}(M)$ and $X' \in \mathrm{Vect}(M')$ are f-related, the function f intertwines the flows of X and X', i.e., for every $m \in M$ and every $t \in I_m$, we have*

$$f(\phi_t^X(m)) = \phi_t^{X'}(f(m)). \tag{33}$$

Proof. Let $m \in M$. It suffices to check that the LHS is an integral curve of X' with initial value $f(m)$. □

Theorem 6. *The Lie bracket of two vector fields $X, Y \in \mathrm{Vect}(M)$ vanishes, i.e.*

$$[X, Y] = 0,$$

if and only if, for every $m \in M$,

$$\phi_t^X(\phi_s^Y(m)) = \phi_s^Y(\phi_t^X(m)),$$

if t and s are close to 0.

Theorem 6 holds in the general case, but for simplicity we assume that the flows are defined in $\mathbb{R} \times M$.

Proof. If $[X, Y] = 0$, Equation (31) implies that

$$d_t\,(\phi_{-t,*}^X Y)_m = (\phi_{-t,*}^X [X, Y])_m = 0,$$

so that

$$T\phi_{-t}^X Y_{\phi_t^X(m)} = (\phi_{-t,*}^X Y)_m = (\phi_{-0,*}^X Y)_m = Y_m,$$

or, equivalently,

$$T\phi_t^X Y_m = Y_{\phi_t^X(m)}.$$

Since the last equality means that Y is ϕ_t^X-related to itself, it follows from Lemma 2 that

$$\phi_t^X(\phi_s^Y(m)) = \phi_s^Y(\phi_t^X(m)).$$

Conversely, if the previous commutation relation is valid, it suffices to compute, for any fixed t, the derivative d_s of this relation at the point $s = 0$. This yields

$$T\phi_t^X(d_s|_{s=0}\ \phi_s^Y(m)) = d_s|_{s=0}\ \phi_s^Y(\phi_t^X(m)),$$

or even better

$$T\phi_t^X Y_m = Y_{\phi_t^X(m)}.$$

It follows that $[X,Y]_m = 0$. □

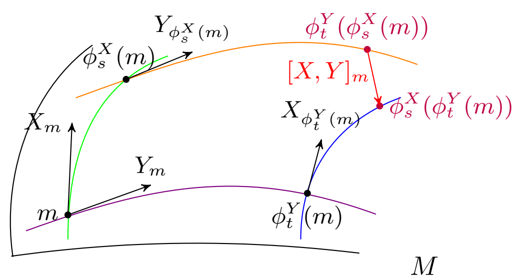

Figure 11: Commutation of flows and Lie bracket of vector fields

Exercise. Let $f \in C^\infty(M,M')$, $X,Y \in \mathrm{Vect}(M)$ and $X',Y' \in \mathrm{Vect}(M')$. Prove that, if X and X', as well as Y and Y' are f-related, then their brackets $[X,Y]$ and $[X',Y']$ are also f-related.

Suggestion: Suppose again, in order not to obscure the ideas, that the flows are defined in $\mathbb{R}\times M$ and compute $[X',Y']_{f(m)}$, $m \in M$, using the definitions, assumptions, and known results.

Chapter 6
Differential Calculus on Manifolds

1 Tensors calculus on vector spaces

This first section is a revision of a topic that is explained in most Bachelor programmes.

In the following, we only consider finite-dimensional real vector spaces. They will be denoted by $V, W, V_i, \ldots$

1.1 Vector law

It is well-known that the transition matrix from a basis (e_i) (resp. (e'_i)) of V to another basis (e'_i) (resp. (e_i)) of V is defined by

$$e'_j = A^i{}_j e_i \text{ (resp. } e_j = A'^i{}_j e'_i) ,$$

where the Einstein summation convention has been used. Note that one passes from one equation to the other by suppressing and adding dashes. Of course, we have $A' = A^{-1}$.

It is easily checked that the components in the two bases of a vector $v \in V$ satisfy the so-called "vector law"

$$v^i = A^i{}_j v'^j \text{ (resp. } v'^i = A'^i{}_j v^j) .$$

1.2 Bidual of a vector space

Let us recall that the dual of V is the vector space $V^* = \mathcal{L}_1(V, \mathbb{R})$ of linear forms on V, i.e. of $\mathbb{R}$-valued linear maps on V. The forms ε^j ($j \in \{1, \ldots, n\}$, $n = \dim V$), defined by

$$\varepsilon^j(e_i) = \delta^j{}_i ,$$

where $\delta^j{}_i$ is Kronecker's symbol, are a basis of V^*, the dual basis of (e_i). Hence $\dim V^* = \dim V = n$.

The bilinear map

$$b : V \times V^* \ni (v, \alpha) \to \alpha(v) \in \mathbb{R}$$

defines a vector space isomorphism, still denoted by b,

$$b : V \ni v \to (b(v) : V^* \ni \alpha \to \alpha(v) \in \mathbb{R}) \in (V^*)^* .$$

When identifying (for finite-dimensional vector spaces) V with its bidual $(V^*)^*$, we get

$$v(\alpha) = \alpha(v) .$$

1.3 Tensor algebra over a vector space

We just explained that

$$V = \mathcal{L}_1(V^*, \mathbb{R})$$

and that the component transformation law for $v \in V$ is

$$v^i = A^i_j v'^j \text{ and } v'^i = A'^i_j v^j \ .$$

As this law is 'contrary' to the law for basis vectors, we say that the elements of V are contravariant vectors.

Similarly

$$V^* = \mathcal{L}_1(V, \mathbb{R}) \ ,$$

and since the components of any $\alpha \in V^*$ in the dual basis are $\alpha_i = \alpha(e_i) = \alpha(A'^j{}_i e'_j)$, the component transformation law for elements $\alpha \in V^*$ is

$$\alpha_i = A'^j{}_i \alpha'_j \text{ and } \alpha'_i = A^j{}_i \alpha_j \ .$$

Since the components of $\alpha \in V^*$ are thus transformed in correspondence to the basis vectors, we say that the elements of V^* are covariant vectors or covectors.

Exercise. Prove that if, in a Euclidian space, we confine ourselves to orthonormal bases, the distinction of contravariant and covariant vectors is redundant.

In order to extend the above observations, we set

$$V \otimes V^* := \mathcal{L}_2(V^* \times V, \mathbb{R}) \ .$$

To examine the transformation law of the components of $T \in V \otimes V^*$, we first need a basis of this vector space. Observe that any basis (e_i) of V induces not only a basis (ε^j) of V^*, but also a basis

$$(e_i \otimes \varepsilon^j) \ (i, j \in \{1, \ldots, n\})$$

of $V \otimes V^*$. Just set

$$(e_i \otimes \varepsilon^j)(\alpha, v) = e_i(\alpha)\ \varepsilon^j(v) = \alpha_i v^j \ , \tag{1}$$

with self-explaining notations. It is easily checked that these bilinear forms are independent. As $\dim(V \otimes V^*) = n^2$, the conclusion follows.

Since, for $T \in V \otimes V^*$, we have

$$T(\alpha, v) = \alpha_i v^j T(\varepsilon^i, e_j) = T(\varepsilon^i, e_j)(e_i \otimes \varepsilon^j)(\alpha, v) \ ,$$

the transformation law of the components $t^i_j = T(\varepsilon^i, e_j)$ of $T \in V \otimes V^*$ will follow from the transformation law of the vectors ε^i of the dual basis (and the known law for the vectors e_j). The guess

$$\varepsilon^i = A^i_j \varepsilon'^j \text{ and } \varepsilon'^i = A'^i_j \varepsilon^j$$

is readily verified. Hence, we have

$$t^i_j = A^i_k A'^\ell{}_j t'^k_\ell \text{ and } t'^i_j = A'^i_k A^\ell{}_j t^k_\ell \ ,$$

so that the components t^i_j of the elements of $V \otimes V^*$ are contravariant in i and covariant in j. We therefore refer to the elements of $V \otimes V^*$ as tensors on V of contravariant and covariant degrees 1, as tensors on V of type $(1,1)$, or, still, as $(1,1)$-tensors on V.

More generally, the set

$$\otimes^p_q V := V \otimes \overset{(p)}{\ldots} \otimes V \otimes V^* \otimes \overset{(q)}{\ldots} \otimes V^* = \mathcal{L}_{p+q}(V^* \times \overset{(p)}{\ldots} \times V^* \times V \times \overset{(q)}{\ldots} \times V, \mathbb{R})$$

is an n^{p+q}-dimensional vector space. Its basis induced by a basis (e_i) of V is obtained as for $\otimes^1_1 V = V \otimes V^*$, the transformation law of the components $t^{i_1\ldots i_p}_{j_1\ldots j_q}$ of an element $T \in \otimes^p_q V$ is obvious and these elements are the p times contravariant and q times covariant tensors on V. Observe that $\otimes^1_0 V = V$, $\otimes^0_1 V = V^*$ and note that by convention $\otimes^0_0 V = \mathbb{R}$.

It follows that any physical quantity characterized in every basis (e_i) by an ordered set of n^{p+q} real numbers $t^{i_1\ldots i_p}_{j_1\ldots j_q}$ that satisfy the just mentioned tensor law of type (p,q), i.e. that satisfy the condition

$$t^{i_1\ldots i_p}_{j_1\ldots j_q} = A^{i_1}_{k_1} \ldots A^{i_p}_{k_p} A'^{\ell_1}_{j_1} \ldots A'^{\ell_q}_{j_q} t'^{k_1\ldots k_p}_{\ell_1\ldots \ell_q} ,$$

can be viewed as a tensor of type (p,q).

Exercise. Let $n = \dim V = 2$ and let $T \in \otimes^1_2 V$. Denote by (a,b) (resp. (α,β)) the basis of V (resp. the dual basis of V^*). Decompose T in the induced basis of $\otimes^1_2 V$.

We now endow the vector space

$$\otimes V = \oplus_{p,q\in\mathbb{N}} \otimes^p_q V$$

with an associative multiplication. Let $T \in \otimes^p_q V$ and $S \in \otimes^r_s V$. To define $T \otimes S \in \otimes^{p+r}_{q+s} V$, we set, see Equation (1),

$$(T\otimes S)(\alpha^1,\ldots,\alpha^{p+r},v_1,\ldots,v_{q+s}) =$$
$$T(\alpha^1,\ldots,\alpha^p,v_1,\ldots,v_q)\; S(\alpha^{p+1},\ldots,\alpha^{p+r},v_{q+1},\ldots,v_{q+s}) . \quad (2)$$

This multiplication $\otimes$ endows the space $\otimes V$ with an associative unital (noncommutative) graded algebra structure. Indeed, observe that, if $c \in \mathbb{R} = \otimes^0_0 V$, we have $c \otimes T = cT$, so that, in particular, $1 \otimes T = T \otimes 1 = T$.

Exercise. Let $T \in \otimes^2_1 V$ and $S \in \otimes^1_3 V$. Show that, if t^{ij}_a (resp. s^i_{abc}) are the components of T (resp. S), the components of $U := T \otimes S$ are given by $u^{ijk}_{abcd} = t^{ij}_a s^k_{bcd}$. What is the number of components of U if $\dim V = 3$?

1.4 Tensor product of vector spaces

We define the tensor product of a finite number of vector spaces $V_1,\ldots,V_p$ in a way similar to the tensor powers $\otimes^p_0 V$, i.e. we set

$$V_1 \otimes \ldots \otimes V_p := \mathcal{L}_p(V_1^* \times \ldots \times V_p^*, \mathbb{R}) .$$

Moreover, the tensor product $v_1 \otimes \ldots \otimes v_p \in V_1 \otimes \ldots \otimes V_p$ of vectors $v_i \in V_i$ is defined as the tensor product of tensors on V, see Equations (1) and (2):

$$(v_1 \otimes \ldots \otimes v_p)(\alpha^1,\ldots,\alpha^p) = \Pi_i v_i(\alpha^i) ,$$

where $\alpha^i \in V_i^*$.

The following vector space isomorphism is fundamental. For every vector space W, we have

$$\mathcal{L}_p(V_1 \times \ldots \times V_p, W) \simeq \mathcal{L}_1(V_1 \otimes \ldots \otimes V_p, W)\ , \tag{3}$$

i.e. the space of multilinear maps on a Cartesian product can be identified with the space of linear maps on the corresponding tensor product. Indeed, the map

$$\begin{aligned}&\mathcal{L}_1(V_1 \otimes \ldots \otimes V_p, W) \ni \ell \mapsto \\ &\quad (L : V_1 \times \ldots \times V_p \ni (v_1, \ldots, v_p) \mapsto \ell(v_1 \otimes \ldots \otimes v_p) \in W) \in \mathcal{L}_p(V_1 \times \ldots \times V_p, W)\end{aligned}$$

is obviously linear and injective and the source and target spaces have the same dimension. Hence, for any multilinear map L on the Cartesian product, there is a unique linear map ℓ on the tensor product, such that

$$\ell(v_1 \otimes \ldots \otimes v_p) = L(v_1, \ldots, v_p), \forall v_i \in V_i\ .$$

This fundamental property is often used to define maps the source space of which is a tensor product space. Indeed, the decomposition $T = \sum v_1 \otimes \ldots \otimes v_p$ of a tensor $T \in V_1 \otimes \ldots \otimes V_p$ as a finite sum of decomposable tensors is not unique. To understand this claim, it suffices to consider a basis (e_i^k) of each V_k and to write

$$T = \sum_{i_1 \ldots i_p} t^{i_1 \ldots i_p}\, e^1_{i_1} \otimes \ldots \otimes e^p_{i_p} = \sum_{i_1 \ldots i_p} (t^{i_1 \ldots i_p} e^1_{i_1}) \otimes \ldots \otimes e^p_{i_p} = \sum_{i_1 \ldots i_p} e^1_{i_1} \otimes \ldots \otimes (t^{i_1 \ldots i_p} e^p_{i_p})\ .$$

Exercise. Prove the following important isomorphisms:

$$V^* \otimes W \simeq \mathcal{L}_1(V, W) \text{ and } \left(\otimes^p_q V\right)^* \simeq \otimes^p_q V^*.$$

Hint: The first one is simply a consequence of the well-known isomorphism $\mathcal{L}_2(V \times W^*, \mathbb{R}) \simeq \mathcal{L}_1(V, \mathcal{L}_1(W^*, \mathbb{R}))$ and the second one is just a particular case of the isomorphism (3).

1.5 Skew-symmetric covariant tensor algebra over a vector space

Let us recall that the space

$$\otimes^p V^* := \otimes^p_0 V^* = V^* \overset{(p)}{\otimes \ldots \otimes} V^* = \mathcal{L}_p(V \times \ldots \times V, \mathbb{R})$$

of p times covariant tensors on V is just the space of p-linear forms on V. So the space of skew-symmetric p times covariant tensors on V,

$$\wedge^p V^* = V^* \overset{(p)}{\wedge \ldots \wedge} V^* := \mathcal{A}_p(V \times \ldots \times V, \mathbb{R})\ ,$$

is nothing but the space of skew-symmetric p-linear forms on V. Again, by convention, $\wedge^0 V^* = \mathbb{R}$. Furthermore, due to antisymmetry, such a tensor necessarily vanishes if $p > n = \dim V$. We denote by

$$\wedge V^* = \oplus^n_{p=0} \wedge^p V^*$$

the vector space of all skew-symmetric covariant tensors on V.

Exercise. Let $T \in \wedge^3 V^* \subset \otimes^3 V^*$. Explain why the components t_{abc} of T in the basis of $\otimes^3 V^*$ are skew-symmetric in a, b, c.

To define an associative algebra structure on $\wedge V^*$ and to get a basis of this space, we introduce the skew-symmetrization operator $\mathcal{A}$. First note that if $T \in \otimes^2 V^*$ and if we set

$$(\mathcal{A}T)(v, w) := \frac{1}{2}\left(T(v, w) - T(w, v)\right) ,$$

then $\mathcal{A}T \in \wedge^2 V^*$ and $\mathcal{A}T = T$, if T is antisymmetric from the beginning. More generally, the skew-symmetrization operator is defined by

$$\mathcal{A} : T \in \otimes^p V^* \mapsto$$

$$\left(\mathcal{A}T : V \times \ldots \times V \ni (v_1, \ldots, v_p) \mapsto \frac{1}{p!} \sum_{\sigma \in \mathcal{S}_p} \mathrm{sign}\sigma\, T(v_{\sigma_1}, \ldots, v_{\sigma_p}) \in \mathbb{R}\right) \in \wedge^p V^* ,$$

where $\mathcal{S}_p$ is the symmetric group of order p, i.e. the group of all the permutations of p different objects.

We are now ready to define the skew-symmetric tensor product $\wedge$ (also called "exterior product" or "wedge product") of skew-symmetric covariant tensors on V. Let $T \in \wedge^p V^* \subset \otimes^p V^*$ and $S \in \wedge^q V^* \subset \otimes^q V^*$ and set

$$T \wedge S = \frac{(p+q)!}{p!q!} \mathcal{A}(T \otimes S) \in \wedge^{p+q} V^* .$$

Then, for any $v_1, \ldots, v_{p+q} \in V$, we have

$$\begin{aligned} &(T \wedge S)(v_1, \ldots, v_{p+q}) \\ &= \frac{1}{p!q!} \sum_\sigma \mathrm{sign}\sigma\, (T \otimes S)(v_{\sigma_1}, \ldots, v_{\sigma_{p+q}}) \\ &= \frac{1}{p!q!} \sum_\sigma \mathrm{sign}\sigma\, T(v_{\sigma_1}, \ldots, v_{\sigma_p})\, S(v_{\sigma_{p+1}}, \ldots, v_{\sigma_{p+q}}) . \end{aligned} \tag{4}$$

When using the antisymmetry of T and of S, we can write this result also as follows:

$$(T \wedge S)(v_1, \ldots, v_{p+q}) = \sum_{\substack{\mu_1 < \ldots < \mu_p \\ \mu_{p+1} < \ldots < \mu_{p+q}}} \mathrm{sign}\mu\, T(v_{\mu_1}, \ldots, v_{\mu_p})\, S(v_{\mu_{p+1}}, \ldots, v_{\mu_{p+q}}) . \tag{5}$$

The permutations μ such that $\mu_1 < \ldots < \mu_p$ and $\mu_{p+1} < \ldots < \mu_{p+q}$ are called the (p, q)-shuffles. We often write $\mu \in \mathrm{Sh}(p, q)$. To understand the last claim, consider a (p, q)-shuffle $\mu \in \mathrm{Sh}(p, q)$ and a $(p + q)$-permutation $\sigma \in \mathcal{S}_{p+q}$ that can be obtained via permutation from this shuffle:

$$\begin{aligned} \mu'(\mu_1,\ \ldots,\ \mu_p) &= (\sigma_1,\ \ldots,\ \sigma_p) , \quad & \mu' \in \mathcal{S}_p , \\ \mu''(\mu_{p+1},\ \ldots,\ \mu_{p+q}) &= (\sigma_{p+1},\ \ldots,\ \sigma_{p+q}) , \quad & \mu'' \in \mathcal{S}_q . \end{aligned}$$

The antisymmetry of T and S implies that

$$\mathrm{sign}\mu\, T(v_{\mu_1}, \ldots, v_{\mu_p})\, S(v_{\mu_{p+1}}, \ldots, v_{\mu_{p+q}}) =$$

$$\text{sign}\mu\,\text{sign}\mu'\,\text{sign}\mu''\;T(v_{\sigma_1},\ldots,v_{\sigma_p})\;S(v_{\sigma_{p+1}},\ldots,v_{\sigma_{p+q}}) =$$
$$\text{sign}\sigma\;T(v_{\sigma_1},\ldots,v_{\sigma_p})\;S(v_{\sigma_{p+1}},\ldots,v_{\sigma_{p+q}})\;.$$

Therefore, in the last sum of Equation (4), the term corresponding to $\mu \in \mathrm{Sh}(p,q)$ appears $p!q!$ times. Hence the announced result (5).

Exercise.

- Consider the case $(p,q) = (2,1)$ and prove by direct computation that Equation (4) reduces to Equation (5).
- Let $T, S \in \wedge^2 V^*$ and compute explicitly $(T \wedge S)(v_1, v_2, v_3, v_4)$.

Remark that the wedge product is graded commutative, i.e. satisfies

$$T \wedge S = (-1)^{pq} S \wedge T\;,$$

where $T \in \wedge^p V^*$ and $S \in \wedge^q V^*$. In particular, for $\alpha, \beta \in \wedge^1 V^* = V^*$, we get

$$\alpha \wedge \beta = -\beta \wedge \alpha\;.$$

This graded commutativity is easily understood from the explicit form (5) of $T \wedge S$ and $S \wedge T$. Indeed, the term of $T \wedge S$ characterized by the shuffle $\mu_1 < \ldots < \mu_p, \mu_{p+1} < \ldots < \mu_{p+q}$ has up to sign the same value as the term of $S \wedge T$ characterized by the shuffle $\mu_{p+1} < \ldots < \mu_{p+q}, \mu_1 < \ldots < \mu_p$. The signature of these shuffles are $\text{sign}\mu$ and $(-1)^{pq}\,\text{sign}\mu$, respectively, which explains the result.

Exercise. Take $(p,q) = (3,1)$ and prove by direct computation that $T \wedge S = -S \wedge T$.

Finally, the wedge product $\wedge$ endows the vector space $\wedge V^*$ with a graded commutative associative unital algebra structure.

In the following we use the same notation as above.

Exercise. This exercise will guide the reader through the next proof. Let $n = \dim V = 2$ and let $T \in \wedge^2 V^* \subset \otimes^2 V^*$. Show that

$$\begin{aligned} T = \mathcal{A}T &= \mathcal{A}\left(t_{11}\varepsilon^1 \otimes \varepsilon^1 + t_{12}\varepsilon^1 \otimes \varepsilon^2 + t_{21}\varepsilon^2 \otimes \varepsilon^1 + t_{22}\varepsilon^2 \otimes \varepsilon^2\right) \\ &= \tfrac{1}{2}\left(t_{11}\varepsilon^1 \wedge \varepsilon^1 + t_{12}\varepsilon^1 \wedge \varepsilon^2 + t_{21}\varepsilon^2 \wedge \varepsilon^1 + t_{22}\varepsilon^2 \wedge \varepsilon^2\right) \\ &= \tfrac{1}{2}\left(t_{12}\varepsilon^1 \wedge \varepsilon^2 + t_{21}\varepsilon^2 \wedge \varepsilon^1\right) \\ &= t_{12}\varepsilon^1 \wedge \varepsilon^2\;. \end{aligned}$$

Let now $n = 3$ and prove that

$$(\varepsilon^1 \wedge \varepsilon^2)(e_1, e_2) = \varepsilon^1(e_1)\varepsilon^2(e_2) - \varepsilon^1(e_2)\varepsilon^2(e_1) = 1 \quad \text{and} \quad (\varepsilon^1 \wedge \varepsilon^2)(e_1, e_3) = 0\;.$$

The p-covariant skew-symmetric tensors

$$\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p} \qquad (1 \le i_1 < \ldots < i_p \le n)$$

form a basis of $\wedge^p V^*$. Indeed, for any $T \in \wedge^p V^* \subset \otimes^p V^*$, we have

$$\begin{aligned} T = \mathcal{A}T &= \mathcal{A}\left(\sum_{j_1,\ldots,j_p} t_{j_1\ldots j_p}\varepsilon^{j_1} \otimes \ldots \otimes \varepsilon^{j_p}\right) \\ &= \sum_{j_1,\ldots,j_p} t_{j_1\ldots j_p}\tfrac{1}{p!}\varepsilon^{j_1} \wedge \ldots \wedge \varepsilon^{j_p} \\ &= \sum_{i_1<\ldots<i_p} t_{i_1\ldots i_p}\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p}\ , \end{aligned}$$

so that the $\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p}$ actually generate $\wedge^p V^*$. To get the result in the last line, note first that – due to antisymmetry – any product $\varepsilon^{j_1} \wedge \ldots \wedge \varepsilon^{j_p}$ with at least two identical factors vanishes. Then observe that in any of the remaining products the factors can be written in the natural order. The signature generated in this way is annihilated by the also skew-symmetric component $t_{i_1\ldots i_p} = T(e_{i_1}, \ldots, e_{i_p})$. As any term of the last line is obtained that way $p!$ times, the result follows. Moreover, it is easily checked that the vectors $\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p}$ are independent. Indeed, if

$$\sum_{i_1<\ldots<i_p} t_{i_1\ldots i_p}\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p} = 0\ ,$$

we have

$$\sum_{i_1<\ldots<i_p} t_{i_1\ldots i_p}(\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p})(e_{k_1}, \ldots, e_{k_p}) = 0\ ,$$

for each $k_1 < \ldots < k_p$. Since $(\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p})(e_{k_1}, \ldots, e_{k_p})$ vanishes, except that it takes value 1, if $(i_1, \ldots, i_p) = (k_1, \ldots, k_p)$, the last equation reduces to $t_{k_1\ldots k_p} = 0$. It follows that the tensors $\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p}$, $1 \le i_1 < \ldots < i_p \le n$, form a basis of the space $\wedge^p V^*$ of skew-symmetric p-covariant tensors on V and that this space has dimension

$$\dim \wedge^p V^* = \binom{n}{p} = C_n^p\ .$$

Exercise. Let $n = \dim V = 3$. Decompose $T \in \wedge^p V^*$, for $p = 0, 1, 2, 3$, in the basis induced by the basis ε^i of V^*.

Of course, instead of considering skew-symmetric covariant tensors, we could just as well study skew-symmetric contravariant tensors. Moreover, symmetric contravariant or covariant tensors play an important role too, even if they are not mentioned in this text.

Finally, the above fundamental property concerning maps on Cartesian and tensor products remains valid: for any vector space W and for any skew-symmetric multilinear map $L \in \mathcal{A}_p(V \times \ldots \times V, W)$, there is a unique linear map $\ell \in \mathcal{L}_1(V \wedge \ldots \wedge V, W)$, such that

$$\ell(v_1 \wedge \ldots \wedge v_p) = L(v_1, \ldots, v_p), \forall v_i \in V\ .$$

2 Tensor calculus on manifolds

2.1 Tensor bundles

Let M be a smooth n-dimensional manifold (that is Hausdorff and second countable). For every $m \in M$, set $V := T_m M$, so that $V^* = T_m^* M$,

$$\otimes_q^p V = \otimes_q^p T_m M \quad \text{and} \quad \wedge^p V^* = \wedge^p T_m^* M\ .$$

Just as we have considered the tangent and cotangent bundles $TM = \coprod_{m\in M} T_mM$ and $T^*M = \coprod_{m\in M} T^*_mM$, we now study the disjoint unions

$$\otimes^p_q TM := \coprod_{m\in M} \otimes^p_q T_mM \quad \text{and} \quad \wedge^p T^*M := \coprod_{m\in M} \wedge^p T^*_mM\ .$$

Let us recall that every chart $(U, \varphi = (x^1, \ldots, x^n))$ of M induces a basis (∂_{x^i}) of T_mM and a basis $(\mathrm{d}\, x^i)$ of T^*_mM, for $m \in U$ (as usual, the dependence of these bases on m is not explicitly mentioned). Hence, every tangent or cotangent vector $X_m \in T_mM$, $\alpha_m \in T^*_mM$, $m \in U$, can be uniquely decomposed as

$$X_m = \sum_i X^i(m)\partial_{x^i} \quad \text{and} \quad \alpha_m = \sum_i \alpha_i(m)\,\mathrm{d}\, x^i\ .$$

Further, the chart φ of M over U allows us to define charts Φ of TM and Φ^* of T^*M over $\pi^{-1}(U)$, where π denotes the projection or "foot map" of the bundle under consideration. These charts are given by

$$\Phi : TM|_U := \coprod_{m\in U} T_mM \ni X_m \mapsto \Big(\varphi(m); \ldots, X^i(m), \ldots\Big) \in \varphi(U) \times \mathbb{R}^n\ ,$$

and

$$\Phi^* : T^*M|_U := \coprod_{m\in U} T^*_mM \ni \alpha_m \mapsto \Big(\varphi(m); \ldots, \alpha_i(m), \ldots\Big) \in \varphi(U) \times \mathbb{R}^n\ .$$

We have proved that the charts Φ of TM and Φ^* of T^*M, induced by all the charts φ of an atlas of M, form an atlas of TM and T^*M, respectively, and thus endow TM and T^*M with a smooth manifold structure of dimension $2n$ (that is Hausdorff and second countable).

Similar constructions go through for $\otimes^p_q TM$ and $\wedge^p T^*M$. Note first that, for $m \in U$, every (p, q)-tensor $T_m \in \otimes^p_q T_mM$ reads

$$T_m = \sum_{\substack{i_1,\ldots,i_p \\ j_1,\ldots,j_q}} t^{i_1\ldots i_p}_{j_1\ldots j_q}(m)\ \partial_{x^{i_1}} \otimes \ldots \otimes \partial_{x^{i_p}} \otimes \mathrm{d}\, x^{j_1} \otimes \ldots \otimes \mathrm{d}\, x^{j_q}\ , \tag{6}$$

and that every skew-symmetric covariant p-tensor $\omega_m \in \wedge^p T^*_mM$ can be written in the form

$$\omega_m = \sum_{i_1<\ldots<i_p} \omega_{i_1\ldots i_p}(m)\ \mathrm{d}\, x^{i_1} \wedge \ldots \wedge \mathrm{d}\, x^{i_p}.$$

Just as in the case of TM and T^*M, the maps

$$\Phi^\otimes : \otimes^p_q TM|_U := \coprod_{m\in U} \otimes^p_q T_mM \ni T_m \mapsto$$

$$\Big(\varphi(m); \ldots, t^{i_1\ldots i_p}_{j_1\ldots j_q}(m), \ldots\Big) \in \varphi(U) \times \mathbb{R}^{n^{p+q}} \simeq \varphi(U) \times \otimes^p_q \mathbb{R}^n$$

and

$$\Phi^\wedge : \wedge^p T^*M|_U := \coprod_{m\in U} \wedge^p T^*_mM \ni \omega_m \mapsto$$

$$\Big(\varphi(m); \ldots, \omega_{i_1\ldots i_p}(m), \ldots\Big) \in \varphi(U) \times \mathbb{R}^{\mathrm{C}^p_n} \simeq \varphi(U) \times \wedge^p(\mathbb{R}^n)^*\ ,$$

obtained if φ runs through an atlas of M, form an atlas of $\otimes_q^p TM$ and $\wedge^p T^*M$, respectively, and endow $\otimes_q^p TM$ and $\wedge^p T^*M$ with a smooth manifold structure of dimension $n + n^{p+q}$ and $n + \mathrm{C}_n^p$, respectively (which is Hausdorff and second countable).

It is straightforwardly checked that

Proposition 1. *If M denotes a smooth n-dimensional manifold, the manifolds $TM, T^*M, \otimes_q^p TM$ $(p, q \in \mathbb{N})$, and $\wedge^p T^*M$ $(p \in \{0, \ldots, n\})$ are vector bundles of rank n, n, n^{p+q}, and C_n^p, respectively.*

These vector bundles are called tensor bundles. Let us write down the exact definition of a vector bundle. See also Chapter 4 and think about the local (but not global) triviality of the Möbius strip.

Definition 1. *Let E and M be two manifolds and $\pi : E \to M$ a smooth surjective map from E onto M. The manifold E is a vector bundle of rank r over the base manifold M – with typical fiber $\mathbb{R}^r$ and projection π – if and only if the fibers $E_m := \pi^{-1}(m)$, $m \in M$, are r-dimensional real vector spaces, and, for every $m \in M$, there is an open neighborhood U in M and a diffeomorphism*

$$\Phi : \pi^{-1}(U) \to U \times \mathbb{R}^r \tag{7}$$

– called local trivialization –, such that, for every $p \in \pi^{-1}(U)$, we have

$$\Phi(p) =: (\pi(p), \phi(p))$$

and, for every $m \in U$, the restriction

$$\phi_m : E_m \to \mathbb{R}^r$$

of ϕ is a vector space isomorphism.

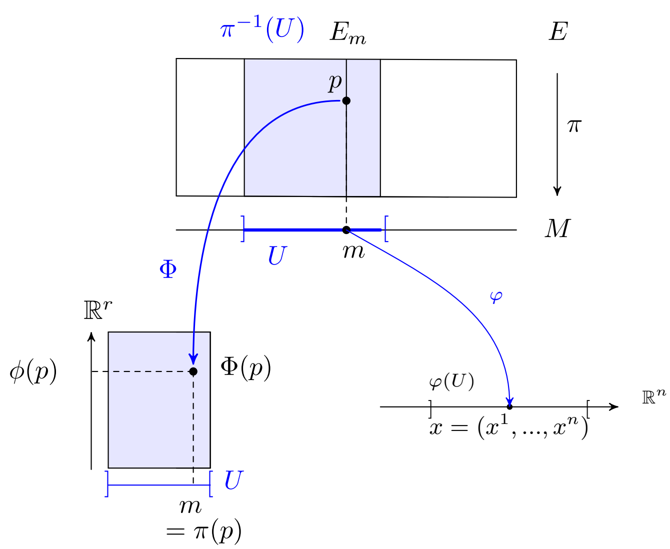

Figure 12: Vector bundle

Hence, roughly speaking, a vector bundle E is an amalgamation of vector spaces E_m ($m \in M$) that is locally trivial, i.e. that can locally be identified with a product manifold (but may globally have a more complicated structure). To get the local trivializations in the previous examples, it suffices to compose the chart diffeomorphisms $\Phi, \Phi^*, \Phi^{\otimes}, \Phi^{\wedge}$ with the diffeomorphism $\varphi^{-1} \times \mathrm{id}$.

Definition 2. *A smooth section of a vector bundle $\pi : E \to M$ is a smooth map $\sigma : M \ni m \mapsto \sigma_m \in E_m \subset E$. We denote by $\Gamma(E)$ the $\mathbb{R}$-vector space and $C^\infty(M)$-module of smooth sections of E.*

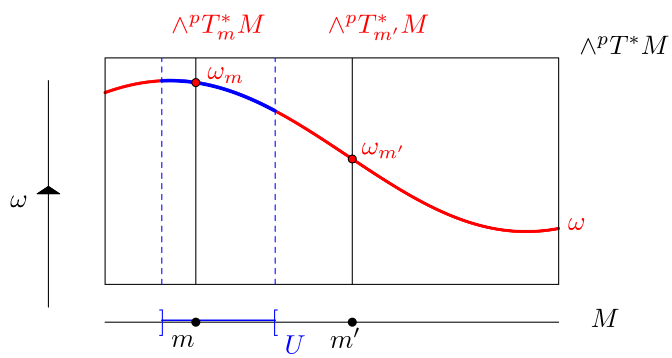

Figure 13: Differential p-form of a manifold

For instance, a section $\omega \in \Gamma(\wedge^p T^* M)$ is a smooth map

$$\omega : M \ni m \mapsto \omega_m \in \wedge^p T^*_m M \subset \wedge^p T^* M\,,$$

hence a map, whose local form over a chart $(U, \varphi = (x^1, \ldots, x^n))$ is given by

$$\omega|_U = \sum_{i_1 < \ldots < i_p} \omega_{i_1 \ldots i_p}\, \mathrm{d}\, x^{i_1} \wedge \ldots \wedge \mathrm{d}\, x^{i_p}\ , \tag{8}$$

where $\omega_{i_1 \ldots i_p} \in C^\infty(U)$. We study these tensor bundle sections in detail in the next section of this text.

The mentioned vector space and module structures on $\Gamma(E)$ are, as in the case of vector fields $\mathrm{Vect}(M) = \Gamma(TM)$, induced by the vector space operations in the fibers E_m, $m \in M$. More precisely, the sum of two sections $\sigma, \sigma' \in \Gamma(E)$ is defined by

$$\sigma + \sigma' :\ M \ni m \mapsto \sigma_m + \sigma'_m \in E_m \subset E\ .$$

Analogously, the multiplication of a section $\sigma \in \Gamma(E)$ by a scalar $\lambda \in \mathbb{R}$ is given by

$$\lambda\, \sigma :\ M \ni m \mapsto \lambda\, \sigma_m \in E_m \subset E$$

and the multiplication by a function $f \in C^\infty(M)$ by

$$f\sigma :\ M \ni m \mapsto f(m)\, \sigma_m \in E_m \subset E\ .$$

2.2 Differential forms

Let us recall that the differential $\mathrm{d}_m f$ at $m \in M$ of $f \in C^\infty(M)$ has been defined by

$$\mathrm{d}_m f := T_m f \in \mathrm{Hom}(T_m M, T_{f(m)}\mathbb{R}) = T_m^* M \ .$$

Hence,

$$\mathrm{d}\, f : M \ni m \mapsto \mathrm{d}_m f \in T_m^* M \subset T^* M$$

is a section of T^*M. Since $\mathrm{d}\, f$ admits – see Chapter 4 – in every chart $(U, \varphi = (x^1, \ldots, x^n))$ the local form

$$\mathrm{d}\, f|_U = \sum_i \partial_{x^i} f \, \mathrm{d}\, x^i \ ,$$

with $\partial_{x^i} f \in C^\infty(U)$, the section $\mathrm{d}\, f$ is smooth, i.e. $\mathrm{d}\, f \in \Gamma(T^*M) = \Gamma(\wedge^1 T^*M)$. As mentioned above, see Equation (8), an arbitrary smooth section $\alpha \in \Gamma(\wedge^1 T^* M)$ locally reads

$$\alpha|_U = \sum_i \alpha_i \, \mathrm{d}\, x^i \ ,$$

$\alpha_i \in C^\infty(U)$. Since

$$\alpha : M \ni m \mapsto \alpha_m \in T_m^* M = \mathcal{L}_1(T_m M, \mathbb{R}) \subset T^* M$$

assigns to every point $m \in M$ a *linear* form of the tangent space of M at m, we refer to it as a *differential* form, or, better, a differential 1-form of M . More generally,

Definition 3. *The sections $\omega \in \Gamma(\wedge^p T^* M)$, $p \in \{0, \ldots, n\}$, $n = \dim M$, are referred to as differential p-forms of M. The $\mathbb{R}$-vector space and $C^\infty(M)$-module $\Gamma(\wedge^p T^*M)$ of differential p-forms of a manifold M is also denoted by $\Omega^p(M)$. The direct sum*

$$\Omega(M) := \oplus_{p=0}^n \Omega^p(M)$$

is the vector space and module of all differential forms on M.

Differential forms are fundamental objects in Differential Geometry – which are for instance tightly connected with integration over manifolds.

Remember that every differential p-form $\omega \in \Omega^p(M)$ locally reads as specified by Equation (8). Note also that differential 0-forms are smooth maps

$$f : M \ni m \mapsto f(m) \in \wedge^0 T_m^* M = \mathbb{R} \ ,$$

so that $\Omega^0(M) = C^\infty(M)$.

Just as the vector space structures of the fibers $E_m = \wedge^p T_m^* M$ induce a vector space structure on $\Gamma(E) = \Gamma(\wedge^p T^* M) = \Omega^p(M)$ (see above), the graded commutative associative unital algebra structures 'wedge' on the $\wedge T_m^* M$ (see above) induce a similar structure on $\Omega(M)$. Indeed, for every $\omega \in \Omega^p(M)$ and $\omega' \in \Omega^q(M)$, we set

$$\omega \wedge \omega' : M \ni m \mapsto \omega_m \wedge \omega'_m \in \wedge^{p+q} T_m^* M \subset \wedge^{p+q} T^* M \ ,$$

so that $\omega \wedge \omega' \in \Omega^{p+q}(M)$ (for the smoothness of $\omega \wedge \omega'$, see Example 1). This wedge product is then extended by linearity to $\Omega(M)$.

Proposition 2. *The wedge product turns the vector space $\Omega(M)$ of differential forms of a manifold M into a graded commutative associative unital algebra.*

Example 1. Let M be of dimension $n = 3$ and let $(U, \varphi = (x^1, x^2, x^3))$ be a chart. Consider two differential forms $\omega \in \Omega^1(M), \omega' \in \Omega^2(M)$, which locally read

$$\omega|_U = \sum_i \omega_i \,\mathrm{d}\, x^i \quad \text{and} \quad \omega'|_U = \textstyle\sum_{k<l} \omega'_{k\,l} \,\mathrm{d}\, x^k \wedge \mathrm{d}\, x^l, \quad \omega_i, \omega_{k\,l} \in C^\infty(U) \ .$$

Their wedge product $\omega \wedge \omega' \in \Omega^3(M)$ is then given over U by

$$\omega \wedge \omega' : U \ni m \mapsto \sum_{i,k<l} \omega_i(m)\, \omega'_{k\,l}(m) \,\mathrm{d}\, x^i \wedge \mathrm{d}\, x^k \wedge \mathrm{d}\, x^l \ .$$

We note that, due to the antisymmetry of the wedge product of differential 1-forms, the only terms of the above sum that do not vanish are those in $\mathrm{d}\, x^1 \wedge \mathrm{d}\, x^2 \wedge \mathrm{d}\, x^3$, $\mathrm{d}\, x^2 \wedge \mathrm{d}\, x^1 \wedge \mathrm{d}\, x^3$ and $\mathrm{d}\, x^3 \wedge \mathrm{d}\, x^1 \wedge \mathrm{d}\, x^2$. Hence,

$$\omega \wedge \omega'|_U = (\omega_1 \omega'_{23} - \omega_2 \omega'_{13} + \omega_3 \omega'_{12}) \,\mathrm{d}\, x^1 \wedge \mathrm{d}\, x^2 \wedge \mathrm{d}\, x^3 \ ,$$

with $\omega_1 \omega'_{23} - \omega_2 \omega'_{13} + \omega_3 \omega'_{12} \in C^\infty(U)$, so that $\omega \wedge \omega'$ is actually smooth.

Exercise. Compute the explicit local form of $\omega \wedge \omega'$, $\omega \in \Omega^1(M)$, $\omega' \in \Omega^2(M)$, for a manifold M of dimension $n = 4$.

2.3 Interior product

Let $\omega \in \Omega^p(M)$ and $X \in \mathrm{Vect}(M)$. For every $m \in M$, we have $\omega_m \in \wedge^p T^*_m M = \mathcal{A}_p(T_m M^{\times p}, \mathbb{R})$ and $X_m \in T_m M$, so that

$$(i_X \omega)_m := \omega_m(X_m, \ldots) \in \mathcal{A}_{p-1}(T_m M^{\times (p-1)}, \mathbb{R}) = \wedge^{p-1} T^*_m M \tag{9}$$

defines a differential $(p-1)$-form $i_X \omega$ on M. For smoothness, see Equation (13).

Definition 4. *The differential form $i_X \omega \in \Omega^{p-1}(M)$, defined pointwise by Equation (9), is called the interior product of the differential form $\omega \in \Omega^p(M)$ by the vector field $X \in \mathrm{Vect}(M)$.*

The interior multiplication

$$i : \mathrm{Vect}(M) \times \Omega^p(M) \to \Omega^{p-1}(M)$$

is clearly $\mathbb{R}$- and $C^\infty(M)$-bilinear. Moreover:

Proposition 3. *The interior multiplication by $X \in \mathrm{Vect}(M)$ is a graded derivation of degree -1 of the algebra $(\Omega(M), \wedge)$, i.e.*

$$i_X(\omega \wedge \omega') = (i_X \omega) \wedge \omega' + (-1)^p \omega \wedge (i_X \omega') \ , \tag{10}$$

for every $\omega \in \Omega^p(M)$ and every $\omega' \in \Omega(M)$.

Exercise. Check the property for $\omega \in \Omega^1(M)$ and $\omega' \in \Omega^2(M)$.

Proof. It suffices to consider a form $\omega' \in \Omega^q(M)$. For every $m \in M$, we have

$$i_X(\omega \wedge \omega')|_m \in \mathcal{A}_{p+q-1}(T_m M^{\times(p+q-1)}, \mathbb{R})$$

and, for every $v_2, \ldots, v_{p+q} \in T_m M$,

$$\begin{aligned} i_X(\omega \wedge \omega')\,|_m\,(v_2, \ldots, v_{p+q}) &= (\omega_m \wedge \omega'_m)(X_m, v_2, \ldots, v_{p+q}) \\ &= \sum_{\substack{\mu_1 < \ldots < \mu_p \\ \mu_{p+1} < \ldots < \mu_{p+q}}} \operatorname{sign}\mu \;\; \omega_m(v_{\mu_1}, \ldots, v_{\mu_p}) \;\; \omega'_m(v_{\mu_{p+1}}, \ldots, v_{\mu_{p+q}}) \,, \end{aligned}$$

where we set $X_m =: v_1$. Since $1 = \mu_1$ or $1 = \mu_{p+1}$, this sum reads

$$\sum_{\substack{1 < \mu_2 < \ldots < \mu_p \\ \mu_{p+1} < \ldots < \mu_{p+q}}} \operatorname{sign}\mu \;\; \omega_m(X_m, v_{\mu_2}, \ldots, v_{\mu_p}) \;\; \omega'_m(v_{\mu_{p+1}}, \ldots, v_{\mu_{p+q}}) \tag{11}$$

$$+ \sum_{\substack{\mu_1 < \ldots < \mu_p \\ 1 < \mu_{p+2} < \ldots < \mu_{p+q}}} \operatorname{sign}\mu \;\; \omega_m(v_{\mu_1}, \ldots, v_{\mu_p}) \;\; \omega'_m(X_m, v_{\mu_{p+2}}, \ldots, v_{\mu_{p+q}}) \,. \tag{12}$$

In the sum (11), the factor $\omega_m(X_m, v_{\mu_2}, \ldots, v_{\mu_p})$ equals $(i_X\omega)_m(v_{\mu_2}, \ldots, v_{\mu_p})$ and the sum over the (p,q)-shuffles $\mu = (1, \mu_2, \ldots, \mu_{p+q})$ can be replaced by the sum over the $(p-1,q)$-shuffles $\mu' := (\mu_2, \ldots, \mu_{p+q})$. Since $\operatorname{sign}\mu = \operatorname{sign}\mu'$, we thus get

$$(i_X\omega) \wedge \omega'\,|_m\,(v_2, \ldots, v_{p+q}) \,.$$

Analogously, in the sum (12), the factor $\omega'_m(X_m, v_{\mu_{p+2}}, \ldots, v_{\mu_{p+q}})$ is equal to $(i_X\omega')_m(v_{\mu_{p+2}}, \ldots, v_{\mu_{p+q}})$. When replacing the (p,q)-shuffles $\mu = (\mu_1, \ldots, \mu_p, 1, \mu_{p+2}, \ldots, \mu_{p+q})$ by the $(p, q-1)$-shuffles $\mu' := (\mu_1, \ldots, \mu_p, \mu_{p+2}, \ldots, \mu_{p+q})$, we have to remark that $\operatorname{sign}\mu = (-1)^p \operatorname{sign}\mu'$. Hence, we get

$$(-1)^p \;\omega \wedge (i_X\omega')\,|_m\,(v_2, \ldots, v_{p+q}) \,.$$

□

Exercise. Let $X \in \operatorname{Vect}(M)$ and $\omega \in \Omega^p(M)$, with local forms $X|_U = \sum_j X^j \partial_{x^j}$ and

$$\omega|_U = \sum_{i_1 < \ldots < i_p} \omega_{i_1 \ldots i_p} \,\mathrm{d}\,x^{i_1} \wedge \ldots \wedge \mathrm{d}\,x^{i_p} \;,$$

in a coordinate patch $(U, (x^1, \ldots, x^n))$. Prove that the local form of $i_X\omega \in \Omega^{p-1}(M)$ is given by

$$(i_X\omega)|_U = \sum_{i_1 < \ldots < i_p} \sum_{k=1}^{p} (-1)^{k-1} \omega_{i_1 \ldots i_p} X^{i_k} \,\mathrm{d}\,x^{i_1} \wedge \ldots \wedge \widehat{\mathrm{d}\,x^{i_k}} \wedge \ldots \wedge \mathrm{d}\,x^{i_p} \;, \tag{13}$$

where ‘hat’ means that the corresponding differential is omitted.

2.4 Pullback and pushforward

2.4.1 Extensions of linear maps

Let us first recall an important elementary fact that will be needed in the following.

Proposition 4. *Every isomorphism $\ell \in \mathrm{Isom}(V,W)$ (resp. every linear map $\ell \in \mathcal{L}_1(V,W)$) between two vector spaces V and W, can be extended as a homomorphism to an isomorphism*

$$\ell^{\otimes} \in \mathrm{Isom}(\otimes_q^p V, \otimes_q^p W)$$

(resp. to a linear map

$$\ell^{\otimes} \in \mathcal{L}_1(\otimes_0^p V, \otimes_0^p W)$$

and a linear map

$$\ell^{\otimes} \in \mathcal{L}_1(\wedge^p V, \wedge^p W)) ,$$

for all $p, q \in \mathbb{N}$.

Proof. In the case ℓ is an isomorphism, it suffices to set, for $v_i \in V$, $\alpha^j \in V^*$,

$$L(v_1, \ldots, v_p, \alpha^1, \ldots, \alpha^q) := \ell(v_1) \otimes \ldots \otimes \ell(v_p) \otimes {}^t\ell^{-1}(\alpha^1) \otimes \ldots \otimes {}^t\ell^{-1}(\alpha^q) \in \otimes_q^p W ,$$

where ${}^t\ell^{-1} \in \mathrm{Isom}(V^*, W^*)$. If ℓ is just a linear map, then $q = 0$, because of the result we have to prove in this case. Due to the abovementioned fundamental property of the tensor product, there is a unique map $\ell^{\otimes} \in \mathcal{L}_1(\otimes_q^p V, \otimes_q^p W)$, such that

$$\ell^{\otimes}(v_1 \otimes \ldots \otimes v_p \otimes \alpha^1 \otimes \ldots \otimes \alpha^q) = \ell(v_1) \otimes \ldots \otimes \ell(v_p) \otimes {}^t\ell^{-1}(\alpha^1) \otimes \ldots \otimes {}^t\ell^{-1}(\alpha^q) .$$

It follows that the map $\ell^{\otimes}$ is in fact the extension of ℓ as a homomorphism. It is obvious that, in the isomorphism-case, we have

$$(\ell^{-1})^{\otimes} = (\ell^{\otimes})^{-1} .$$

The extension of ℓ to a linear map on the exterior power is similar; it suffices to replace the ordinary tensor product $\otimes$ by the skew-symmetric tensor product $\wedge$. Let us also stress that if $q = p = 0$, we take $\ell^{\otimes} = \mathrm{id}_{\mathbb{R}}$. □

We now briefly report on useful properties of the preceding extensions.

Lemma 1. *For every isomorphisms $\ell \in \mathrm{Isom}(V,W)$, $\ell' \in \mathrm{Isom}(W,Z)$ (resp. for every linear maps $\ell \in \mathcal{L}_1(V,W)$, $\ell' \in \mathcal{L}_1(W,Z)$), we have*

$$(\ell' \circ \ell)^{\otimes} = \ell'^{\otimes} \circ \ell^{\otimes} .$$

Proof. We first confine ourselves to the case $p = q = 1$. For $v \in V$ and $\alpha \in V^*$, we get, by definition,

$$(\ell' \circ \ell)^{\otimes}(v, \alpha) = (\ell' \circ \ell)(v) \otimes {}^t(\ell' \circ \ell)^{-1}(\alpha) .$$

Since

$${}^t(\ell' \circ \ell)^{-1} = {}^t(\ell^{-1} \circ \ell'^{-1}) = {}^t\ell'^{-1} \circ {}^t\ell^{-1} ,$$

we find

$$(\ell' \circ \ell)^{\otimes}(v, \alpha) = (\ell' \circ \ell)(v) \otimes ({}^t\ell'^{-1} \circ {}^t\ell^{-1})(\alpha) = (\ell'^{\otimes} \circ \ell^{\otimes})(v, \alpha) .$$

The proof goes through for arbitrary p and q. □

Lemma 2. *For every isomorphism $\ell \in \mathrm{Isom}(V,W)$ (resp. every linear map $\ell \in \mathcal{L}_1(V,W)$), we have*

$${}^t(\ell^{\otimes}) = ({}^t\ell)^{\otimes} =: {}^t\ell^{\otimes} .$$

Proof. We know that ℓ induces an isomorphism $\ell^{\otimes} : \otimes_q^p V \to \otimes_q^p W$ and consequently an isomorphism

$$^t(\ell^{\otimes}) : (\otimes_q^p W)^* = \otimes_q^p W^* \to (\otimes_q^p V)^* = \otimes_q^p V^* \,.$$

On the other hand, the isomorphism $^t\ell : W^* \to V^*$ can be extended to an isomorphism

$$(\,^t\ell)^{\otimes} : \otimes_q^p W^* \to \otimes_q^p V^* \,.$$

To prove the claim, we have to show that $^t(\ell^{\otimes})$ and $(\,^t\ell)^{\otimes}$ coincide on every element of their source space. Let us examine e.g. the case $p = 2$ and $q = 1$; the general case is completely analogous. When applying both maps to $\beta_1 \otimes \beta_2 \otimes w \in \otimes_1^2 W^*$, we obtain two elements in

$$\otimes_1^2 V^* = \mathcal{L}_3(V \times V \times V^*, \mathbb{R}) \simeq \mathcal{L}_1(V \otimes V \otimes V^*, \mathbb{R}) \,,$$

which we must apply to an arbitrary element of the type $v_1 \otimes v_2 \otimes \alpha$. Then,

$$\begin{aligned} ^t(\ell^{\otimes})(\beta_1 \otimes \beta_2 \otimes w)\,(v_1 \otimes v_2 \otimes \alpha) &= (\beta_1 \otimes \beta_2 \otimes w)\big(\ell^{\otimes}(v_1 \otimes v_2 \otimes \alpha)\big) \\ &= (\beta_1 \otimes \beta_2 \otimes w)(\ell(v_1) \otimes \ell(v_2) \otimes \,^t\ell^{-1}\alpha) \\ &= \beta_1(\ell(v_1))\,\beta_2(\ell(v_2))\,\alpha(\ell^{-1}(w)) \,, \end{aligned}$$

whereas

$$\begin{aligned} (\,^t\ell)^{\otimes}(\beta_1 \otimes \beta_2 \otimes w)\,(v_1 \otimes v_2 \otimes \alpha) &= \big(\,^t\ell(\beta_1) \otimes \,^t\ell(\beta_2) \otimes \,\ell^{-1}(w)\big)(v_1 \otimes v_2 \otimes \alpha) \\ &= \beta_1(\ell(v_1))\,\beta_2(\ell(v_2))\,\alpha(\ell^{-1}(w)) \,. \end{aligned}$$

□

2.4.2 Pullback of covariant tensor fields by a function

Let M and N be two manifolds. Every

$$f \in C^{\infty}(M, N)$$

allows us to pull every covariant tensor field on N back to M, i.e. to define a linear map

$$f^* : \Gamma(\otimes^p T^* M) \ni f^* T \leftarrow T \in \Gamma(\otimes^p T^* N) \,.$$

The pullback of differential forms,

$$f^* : \Omega^p(M) = \Gamma(\wedge^p T^* M) \ni f^*\omega \leftarrow \omega \in \Gamma(\wedge^p T^* N) = \Omega^p(N) \,,$$

will turn out to be of particular importance in the present context. We therefore detail this case.

Remark. The derivative of $f \in C^{\infty}(M, N)$ at $m \in M$ was denoted so far by $T_m f$. For $N = \mathbb{R}$, we mostly replaced $T_m f$ by $(d\,f)_m$ or $\mathrm{d}_m\, f$. To simplify the notations a little bit, we write f_{*m} instead of $T_m f$ in the following.

As we start with a field $\omega_n \in \wedge^p T_n^* N$ ($n \in N$) and aim to construct a field $(f^*\omega)_m \in \wedge^p T_m^* M$ ($m \in M$), we must look for a suitable 'means of transportation'. Since $f_{*m}^{\otimes} \in \mathcal{L}_1(\wedge^p T_m M, \wedge^p T_{f(m)} N)$ ($m \in M$), it is clear that $^t f_{*m}^{\otimes} \in \mathcal{L}_1(\wedge^p T_{f(m)}^* N, \wedge^p T_m^* M)$. Hence, the definition

$$(f^*\omega)_m = \,^t f_{*m}^{\otimes}\, \omega_{f(m)} \quad (m \in M) \,. \tag{14}$$

Observe that for a 0-form $h \in C^{\infty}(N)$, we have $^t f_{*m}^{\otimes} = \mathrm{id}$, so that

$$f^* h = h \circ f \,. \tag{15}$$

Proposition 5. *Let $M \xrightarrow{f} N \xrightarrow{g} P$ be smooth maps between manifolds. The pullback by the composite map is given by*

$$(g \circ f)^* = f^* \circ g^* \ .$$

Moreover, $f^* : \Omega(N) \to \Omega(M)$ is a unital $\mathbb{R}$-algebra homomorphism.

Proof. Let $\omega \in \Omega^p(P)$ and let $m \in M$. It follows from the definition of a pullback and from Lemma 1 that

$$\begin{aligned}
((g \circ f)^* \, \omega\,)_m &= {}^t(g \circ f)^{\otimes}_{*m} \, \omega_{g(f(m))} \\
&= {}^t\!\left(g^{\otimes}_{*f(m)} \circ f^{\otimes}_{*m}\right) \, \omega_{g(f(m))} \\
&= {}^tf^{\otimes}_{*m} \left({}^tg^{\otimes}_{*f(m)} \, \omega_{g(f(m))}\right) \\
&= {}^tf^{\otimes}_{*m} \left((g^*\omega)_{f(m)}\right) \\
&= ((f^* \circ g^*) \, \omega)_m \ .
\end{aligned}$$

As for the homomorphism property, it is clear from Definition (14) that f^* is $\mathbb{R}$-linear. Moreover, the constant function $1_N : N \ni n \mapsto 1 \in \mathbb{R}$ is the unit of $\Omega(N)$ and we get from Equation (15) that $f^* 1_N = 1_M$. Finally, since ${}^tf^{\otimes}_{*m} = {}^t(f^{\otimes}_{*m}) = ({}^tf_{*m})^{\otimes}$ and since the latter is defined as algebra homomorphism, we get

$$(f^* \, (\omega \wedge \omega'))_m = {}^tf^{\otimes}_{*m} \, (\omega \wedge \omega')_{f(m)} = {}^tf^{\otimes}_{*m} \, (\omega_{f(m)} \wedge \omega'_{f(m)}) =$$

$${}^tf^{\otimes}_{*m} \, \omega_{f(m)} \wedge {}^tf^{\otimes}_{*m} \, \omega'_{f(m)} = (f^* \, \omega \wedge f^* \, \omega')_m \ .$$

□

Exercise. Let $\omega \in \Omega^p(N)$, $f \in C^\infty(M, N)$ and consider a chart $(V, (y^1, \ldots, y^n))$ of N. Prove that if

$$\omega|_V = \sum_{j_1 < \ldots < j_p} \omega_{j_1 \ldots j_p} \, \mathrm{d}\, y^{j_1} \wedge \ldots \wedge \mathrm{d}\, y^{j_p} \ ,$$

then

$$(f^*\omega)|_{f^{-1}(V)} = \sum_{j_1 < \ldots < j_p} \omega_{j_1 \ldots j_p} \circ f \quad \mathrm{d}(y^{j_1} \circ f) \wedge \ldots \wedge \mathrm{d}(y^{j_p} \circ f) \ .$$

Consider now the case $p = n$ and $f = \phi \in \mathrm{Diff}(U, V)$, where $U \subset \mathbb{R}^n$. Show that, if the coordinates of U are denoted $x = (x^1, \ldots, x^n)$ and if we set $y^i(\phi(x)) = y^i(x)$, we obtain

$$(\phi^*\omega)_x = \det(\partial_x y) \; h(y(x)) \; \mathrm{d}\, x^1 \wedge \ldots \wedge \mathrm{d}\, x^n \ ,$$

where we wrote h instead of $\omega_{1 \ldots n}$.

2.4.3 Pushforward of contravariant tensor fields by a diffeomorphism

Let

$$f \in \mathrm{Diff}(M, N)$$

and try to define a pushforward map

$$f_* : \Gamma(\otimes^p TM) \ni T \mapsto f_* T \in \Gamma(\otimes^p TN) \ .$$

Here the known object is a field $T_m \in \otimes^p T_m M$ ($m \in M$) and the object we look for is $(f_* T)_n \in \otimes^p T_n N$ ($n \in N$). It is clear that $f^{\otimes}_{*m} \in \mathcal{L}_1(\otimes^p T_m M, \otimes^p T_{f(m)} N)$ ($m \in M$), so that it suffices to choose $m = f^{-1}(n)$ (here we use the assumption that f is a diffeomorphism). Hence the definition

$$(f_* T)_n = \left(f^{\otimes}_{*m} T_m\right)_{m=f^{-1}(n)} \quad (n \in N) .$$

Proposition 6. *Consider $M \xrightarrow{f} N \xrightarrow{g} P$ two diffeomorphisms between manifolds. The push-forward satisfies*

$$(g \circ f)_* = g_* \circ f_* .$$

Exercise. Prove the previous proposition.

2.5 Lie derivative of tensor fields

2.5.1 Definition and properties

The Lie derivative of a (p, q)-tensor field $T \in \Gamma(\otimes^p_q TM)$ is defined along the same lines as the Lie derivative of a vector field $Y \in \mathrm{Vect}(M) = \Gamma(TM) = \Gamma(\otimes^1_0 TM)$, i.e. of a $(1, 0)$-tensor field.

Take a tensor field $T \in \Gamma(\otimes^p_q TM)$, a vector field $X \in \mathrm{Vect}(M)$ and a point $m \in M$. To compute the value $(L_X T)_m$ at m of the Lie derivative of T with respect to X, i.e. the variation at m of T in the direction of X, we have to compare T_m and $T_{\varphi_t(m)}$, where φ_t is the flow of X and where $t \simeq 0$. The point is that $T_m \in \otimes^p_q T_m M$ and $T_{\varphi_t(m)} \in \otimes^p_q T_{\varphi_t(m)} M$. Hence, if we wish to subtract T_m from $T_{\varphi_t(m)}$, we first have to transfer $T_{\varphi_t(m)}$ into the space that contains T_m. Assume, for convenience, that X is complete. As then $\varphi_t \in \mathrm{Diff}(M)$, we have

$$(\varphi_{t*m})^{-1} = \varphi_{-t*\varphi_t(m)} \in \mathrm{Isom}(T_{\varphi_t(m)} M, T_m M)$$

and

$$\varphi^{\otimes}_{-t*} \in \mathrm{Isom}(\otimes^p_q T_{\varphi_t(m)} M, \otimes^p_q T_m M) ,$$

where, to simplify notations, we omitted the point $\varphi_t(m)$ at which the derivative φ_{-t*} is computed. Hence the definition

$$(L_X T)_m = \lim_{t \to 0} \frac{\varphi^{\otimes}_{-t*} T_{\varphi_t(m)} - T_m}{t} = d_t|_{t=0}\, \varphi^{\otimes}_{-t*} T_{\varphi_t(m)} \in \otimes^p_q T_m M . \tag{16}$$

As in the case of the Lie derivative of a vector field, the local form of $\varphi^{\otimes}_{-t*} T_{\varphi_t(m)}$ shows that this curve is smooth with respect to t and m, so that the derivative makes sense and defines a smooth tensor field $L_X T \in \Gamma(\otimes^p_q TM)$.

The Lie derivative of a differential form $\omega \in \Omega^p(M) = \Gamma(\wedge^p T^* M)$, i.e. of a skew-symmetric covariant tensor field, is defined analogously.

The Lie derivative has good properties.

Proposition 7. 1. *The Lie derivative preserves the type of the tensor field, i.e., for every $X \in \mathrm{Vect}(M)$, if $T \in \Gamma(\otimes^p_q TM)$, then $L_X T \in \Gamma(\otimes^p_q TM)$, and if $\omega \in \Omega^p(M)$, then $L_X \omega \in \Omega^p(M)$, ...*

2. *The Lie derivative is local in each argument and bilinear.*

3. *For every $X \in \mathrm{Vect}(M)$, every $T \in \Gamma(\otimes^p_q TM)$ and every $S \in \Gamma(\otimes^r_s TM)$, we have*

$$L_X(T \otimes S) = (L_X T) \otimes S + T \otimes (L_X S) \ ,$$

if $\omega \in \Omega^p(M)$ and $\omega' \in \Omega^q(M)$, we get

$$L_X(\omega \wedge \omega') = (L_X \omega) \wedge \omega' + \omega \wedge (L_X \omega') \ ,$$

and similar results are valid for other types of tensor fields and the corresponding tensor product. In other words, L_X, $X \in \mathrm{Vect}(M)$, is a (graded) derivation (of degree 0) of the algebras $(\Gamma(\otimes TM), \otimes)$, $(\Omega(M), \wedge)$, ...

4. *If $X \in \mathrm{Vect}(M)$, $T \in \Gamma(\otimes^p_q TM)$, $\omega^1, \ldots, \omega^p \in \Omega^1(M)$, and $X_1, \ldots, X_q \in \mathrm{Vect}(M)$, we get*

$$\begin{aligned} L_X\left(T(\omega^1, \ldots, \omega^p, X_1, \ldots, X_q)\right) = \ & (L_X T)(\omega^1, \ldots, \omega^p, X_1, \ldots, X_q) \\ & + \textstyle\sum_{i=1}^p T(\omega^1, \ldots, L_X \omega^i, \ldots, \omega^p, X_1, \ldots, X_q) \\ & + \textstyle\sum_{j=1}^q T(\omega^1, \ldots, \omega^p, X_1, \ldots, L_X X_j, \ldots, X_q) \ . \end{aligned} \tag{17}$$

5. *For every $T \in \Gamma(\otimes TM)$ and every $X, Y \in \mathrm{Vect}(M)$,*

$$L_{[X,Y]} T = L_X(L_Y T) - L_Y(L_X T) \ ,$$

i.e.

$$L_{[X,Y]} = [L_X, L_Y] \tag{18}$$

in $\Gamma(\otimes TM)$.

Proof. 1. By construction.

2. The expression $L_X T$ is linear with respect to T by construction. Linearity with respect to X comes – as in the case of the Lie derivative of vector fields – from the local form of $L_X T$, whereas locality is almost obvious.

3. The proofs of these results are similar. For instance,

$$\begin{aligned} L_X(T \otimes S)_m &= d_t|_{t=0}\, \varphi^{\otimes}_{-t*}(T \otimes S)_{\varphi_t(m)} \\ &= d_t\left(\varphi^{\otimes}_{-t*}\, T_{\varphi_t(m)} \otimes \varphi^{\otimes}_{-t*}\, S_{\varphi_t(m)}\right)\big|_{t=0} \\ &= \left(d_t\left(\varphi^{\otimes}_{-t*}\, T_{\varphi_t(m)}\right) \otimes \varphi^{\otimes}_{-t*}\, S_{\varphi_t(m)}\right)\big|_{t=0} \\ &\quad + \left(\varphi^{\otimes}_{-t*}\, T_{\varphi_t(m)} \otimes d_t\left(\varphi^{\otimes}_{-t*}\, S_{\varphi_t(m)}\right)\right)\big|_{t=0} \\ &= (L_X T)_m \otimes S_m + T_m \otimes (L_X S)_m \ . \end{aligned}$$

4. To avoid cumbersome notations, we prove the announced result for

$$\omega := T \in \Gamma(\otimes^0_1 TM) = \Gamma(T^*M) = \Omega^1(M) \ .$$

In this case, it reads

$$L_X(\omega(Y)) = (L_X \omega)(Y) + \omega(L_X Y) \ ,$$

where we wrote Y instead of X_1. Observe first that

$$\begin{aligned} \varphi^{\otimes}_{-t*}\omega_{\varphi_t(m)}(\varphi^{\otimes}_{-t*} Y_{\varphi_t(m)}) &= {}^t\varphi_{t*}\omega_{\varphi_t(m)}(\varphi_{-t*} Y_{\varphi_t(m)}) \\ &= \omega_{\varphi_t(m)}(\varphi_{t*}\varphi_{-t*} Y_{\varphi_t(m)}) \\ &= \varphi^{\otimes}_{-t*}(\omega_{\varphi_t(m)}(Y_{\varphi_t(m)})) \ , \end{aligned}$$

since φ_t and φ_{-t} are inverses and since the extension $\ell^\otimes$ of every linear map ℓ is identity on real numbers. If follows that

$$\begin{aligned}(L_X(\omega(Y)))_m &= d_t|_{t=0}\ \varphi^\otimes_{-t*}(\omega_{\varphi_t(m)}(Y_{\varphi_t(m)}))\\ &= d_t|_{t=0}\ \varphi^\otimes_{-t*}\omega_{\varphi_t(m)}(\varphi^\otimes_{-t*}Y_{\varphi_t(m)})\\ &= \left(d_t|_{t=0}\ \varphi^\otimes_{-t*}\omega_{\varphi_t(m)}\right)(Y_m) + \omega_m\left(d_t|_{t=0}\ \varphi^\otimes_{-t*}Y_{\varphi_t(m)}\right)\\ &= ((L_X\omega)(Y) + \omega(L_XY))_m\ .\end{aligned}$$

The extension to an arbitrary $T \in \Gamma(\otimes^p_q TM)$ is straightforward.

5. We prove the announced result at the end of Section 2.6.

□

2.5.2 Local forms

Equation (17) in Proposition 7 gives L_XT in terms of the Lie derivatives of a function ($C^\infty(M) = \Gamma(\otimes^0_0 TM)$), differential 1-forms ($\Omega^1(M) = \Gamma(\otimes^0_1 TM)$) and vector fields ($\mathrm{Vect}(M) = \Gamma(\otimes^1_0 TM)$). In this subsection, we provide the local forms of the Lie derivatives of a function, a vector field and a differential 1-form.

When applying Definition (16) of the Lie derivative of an arbitrary tensor field $T \in \Gamma(\otimes^p_q TM)$ in the direction of $X \in \mathrm{Vect}(M)$ to a function $f \in C^\infty(M)$, we get, for every $m \in M$,

$$(Xf)(m) := (L_Xf)_m = d_t|_{t=0}\ f(\varphi_t(m)) =$$
$$f_{*m}\ d_t|_{t=0}\ \varphi_t(m) = f_{*m}X_m = (\mathrm{d}_m f)(X_m) = X(\mathrm{d} f)|_m\ . \tag{19}$$

Is is now easily checked that the local expression of L_Xf in a chart $(U, (x^1, \ldots, x^n))$ of M is

$$(L_Xf)\,|_U = \sum_i X^i\partial_{x^i}f\ , \tag{20}$$

if $X\,|_U = \sum_i X^i\partial_{x^i}$. It is worth to remember that

$$Xf = L_Xf = (\mathrm{d}\, f)(X) = X(\mathrm{d}\, f)\ .$$

The relation

$$[fX, gY] = fg[X, Y] + f(Xg)Y - g(Yf)X$$

($f, g \in C^\infty(M)$, $X, Y \in \mathrm{Vect}(M)$) between the Lie and the module structure of $\mathrm{Vect}(M)$ allows to easily recover the local expression of L_XY. If $X\,|_U = \sum_i X^i\partial_{x^i}$ and $Y\,|_U = \sum_j Y^j\partial_{x^j}$, then

$$(L_XY)\,|_U = \sum_i\left(\sum_j X^j\partial_{x^j}Y^i - \sum_j Y^j\partial_{x^j}X^i\right)\partial_{x^i}\ . \tag{21}$$

When applied to $X, Y \in \mathrm{Vect}(M)$ and $\omega \in \Omega^1(M)$, such that $X|_U = \sum_j X^j\partial_{x^j}$, $Y|_U = \partial_{x^i}$, and $\omega\,|_U = \sum_j \omega_j\,\mathrm{d}\,x^j$, the property $L_X\,(Y(\omega)) = (L_XY)(\omega) + Y(L_X\omega)$ leads to the local form

$$(L_X\omega)\,|_U = \sum_i\left(\sum_k X^k\partial_{x^k}\,\omega_i + \sum_j \omega_j\partial_{x^i}X^j\right)dx^i\ . \tag{22}$$

Note that in the preceding computations we implicitly used the locality of the Lie derivative, i.e. the fact that

$$(L_XT)\,|_U = L_{X|_U}\,T|_U\ .$$

2.6 De Rham differential and Cartan calculus

We know that the differential of a function $f \in C^\infty(M) = \Omega^0(M)$ is a differential 1-form $\mathrm{d}\, f \in \Omega^1(M)$. Moreover, the linear map

$$\mathrm{d} : \Omega^0(M) \ni f \mapsto \mathrm{d}\, f \in \Omega^1(M)$$

is a derivation, i.e.

$$\mathrm{d}(fg) = (\mathrm{d}\, f)g + f(\mathrm{d}\, g) ,$$

for every $f, g \in C^\infty(M)$. This derivation d can be extended from functions, or differential 0-forms, to differential p-forms, $p > 0$.

Theorem 1. *There exists a unique linear map*

$$\mathrm{d} : \Omega^p(M) \to \Omega^{p+1}(M) ,$$

$p \geq 0$, *called the de Rham differential or exterior differential, which satisfies the following requirements:*

1. *The de Rham differential extends the differential of functions.*
2. *For every $\omega \in \Omega^p(M)$ and every $\omega' \in \Omega(M)$, we have*

   $$\mathrm{d}(\omega \wedge \omega') = (\mathrm{d}\, \omega) \wedge \omega' + (-1)^p \omega \wedge (\mathrm{d}\, \omega') ,$$

   i.e. d is a graded derivation of degree 1 of the graded commutative algebra $(\Omega(M), \wedge)$ of differential forms of M.
3. *The map d is a differential on $\Omega(M)$ in the sense of Homological Algebra, i.e. an endomorphism of $\Omega(M)$, such that $\mathrm{d}^2 = \mathrm{d} \circ \mathrm{d} = 0$.*

Proof. 1. We first prove uniqueness of d.

If d exists, it is a local operator. This follows as usual from the derivation property. Indeed, if $\omega \in \Omega(M)$ and U is an open subset of M such that $\omega|_U = 0$, and if $m \in U$, we take a bump function γ around m, i.e. a function $\gamma \in C^\infty(M)$ such that $\operatorname{supp} \gamma \subset U$ and $\gamma = 1$ in a neighborhood $V \subset U$ of m. Then $\omega = (1 - \gamma)\omega$ and $\mathrm{d}\, \omega = \mathrm{d}(1 - \gamma) \wedge \omega + (1 - \gamma)\, \mathrm{d}\, \omega$, so that $(\mathrm{d}\, \omega)|_V = 0$. In view of this locality property, if two differential forms $\omega, \omega' \in \Omega(M)$ coincide in an open subset $U \subset M$, then their differentials coincide as well, i.e. $(\mathrm{d}\, \omega)|_U = (\mathrm{d}\, \omega')|_U$.

Let us also recall that local operators can be restricted to open subsets $U \subset M$. In the case of $\mathrm{d} : \Omega^p(M) \to \Omega^{p+1}(M)$, this means that there exists an operator $\mathrm{d}\,|_U : \Omega^p(U) \to \Omega^{p+1}(U)$ such that in particular $\mathrm{d}\,|_U \omega|_U = (\mathrm{d}\, \omega)|_U$ for every $\omega \in \Omega(M)$. Indeed, let $\omega_U \in \Omega^p(U)$ (observe that we denote by ω_U a form over U and by $\omega|_U$ the restriction to U of a form over M). To define $\mathrm{d}\,|_U \omega_U \in \Omega^{p+1}(U)$, we consider, for every $m \in U$, a bump function $\gamma_V \in C^\infty(M)$ around m that is equal to 1 in a neighborhood $V \subset U$ of m and equal to 0 in $M \setminus U$. Then $\gamma_V \omega_U \in \Omega^p(M)$ and we can set

$$(\mathrm{d}|_U\, \omega_U)|_V := \mathrm{d}(\gamma_V \omega_U)|_V \ \in \Omega^{p+1}(V) .$$

In every overlap $V \cap V'$ the two definitions coincide, i.e. $\mathrm{d}(\gamma_V \omega_U) = \mathrm{d}(\gamma_{V'} \omega_U)$, since in the overlap we have $\gamma_V \omega_U = \gamma_{V'} \omega_U$. Hence, we obtain a well-defined differential form $\mathrm{d}|_U\, \omega_U \in \Omega^{p+1}(U)$. It follows that, for the restriction $\omega|_U \in \Omega^p(U)$ of a form $\omega \in \Omega^p(M)$, we have

$$\mathrm{d}|_U\, \omega|_U = (\mathrm{d}\, \omega)|_U .$$

Take now a differential form $\omega \in \Omega^p(M)$, a chart $(U, (x^1, \ldots, x^n))$, a point $m \in U$ and a bump function γ_V as above. If

$$\omega|_U = \sum_{i_1 < \ldots < i_p} \omega_{i_1 \ldots i_p} \, \mathrm{d}\, x^{i_1} \wedge \ldots \wedge \mathrm{d}\, x^{i_p} \ , \tag{23}$$

we have necessarily

$$\begin{aligned}(\mathrm{d}\,\omega)|_V = (\mathrm{d}|_U\, \omega|_U)|_V \quad &= \quad \mathrm{d} \sum_{i_1 < \ldots < i_p} \gamma_V\, \omega_{i_1 \ldots i_p}\, \mathrm{d}(\gamma_V\, x^{i_1}) \wedge \ldots \wedge \mathrm{d}(\gamma_V\, x^{i_p})\,|_V \\ &= \sum_{i_1 < \ldots < i_p} \mathrm{d}(\gamma_V\, \omega_{i_1 \ldots i_p}) \wedge \mathrm{d}(\gamma_V\, x^{i_1}) \wedge \ldots \wedge \mathrm{d}(\gamma_V\, x^{i_p})\,|_V \\ &= \sum_{i_1 < \ldots < i_p} \mathrm{d}\,\omega_{i_1 \ldots i_p} \wedge \mathrm{d}\, x^{i_1} \wedge \ldots \wedge \mathrm{d}\, x^{i_p}\,|_V \ ,\end{aligned}$$

where we wrote, exactly as in Equation (23), d instead of $\mathrm{d}\,|_U$ (indeed, we write differentials of functions over U, e.g. $\mathrm{d}\, x^i$, in this a bit unprecise way as from the very beginning). Hence

$$(\mathrm{d}\,\omega)|_U = \sum_{i_1 < \ldots < i_p} \mathrm{d}\,\omega_{i_1 \ldots i_p} \wedge \mathrm{d}\, x^{i_1} \wedge \ldots \wedge \mathrm{d}\, x^{i_p} \ , \tag{24}$$

so that d is actually unique.

2. We now prove the existence of d. We first construct d in a chart $(U, (x^1, \ldots, x^n))$ and verify that this $\mathrm{d}_U : \Omega^p(U) \to \Omega^{p+1}(U)$ has all the required properties. Then, we show that the d_U-s can be glued, thus providing a $\mathrm{d} : \Omega^p(M) \to \Omega^{p+1}(M)$ that inherits the same properties.

2.a. For every chart $(U, (x^1, \ldots, x^n))$ and every $\omega_U \in \Omega^p(U)$, we have

$$\omega_U = \sum_{i_1 < \ldots < i_p} \omega_{i_1 \ldots i_p}\, \mathrm{d}\, x^{i_1} \wedge \ldots \wedge \mathrm{d}\, x^{i_p}$$

and we define $\mathrm{d}_U\, \omega_U$, see Equation (24), by

$$\mathrm{d}_U\, \omega_U := \sum_{i_1 < \ldots < i_p} \mathrm{d}\,\omega_{i_1 \ldots i_p} \wedge \mathrm{d}\, x^{i_1} \wedge \ldots \wedge \mathrm{d}\, x^{i_p} \ , \tag{25}$$

so that $\mathrm{d}_U \in \mathcal{L}(\Omega^p(U), \Omega^{p+1}(U))$ and coincides on $\Omega^0(U) = C^\infty(U)$ with the differential of functions. Further, for every $\omega = f\ \mathrm{d}\, x^{i_1} \wedge \ldots \wedge \mathrm{d}\, x^{i_p} \in \Omega^p(U)$ and every $\omega' = g\ \mathrm{d}\, x^{j_1} \wedge \ldots \wedge \mathrm{d}\, x^{j_q} \in \Omega^q(U)$ (to simplify notations we omit sums and indices), we find

$$\begin{aligned}\mathrm{d}_U(\omega \wedge \omega') &= \mathrm{d}_U \left((f\ \mathrm{d}\, x^{i_1} \wedge \ldots \wedge \mathrm{d}\, x^{i_p}) \ \wedge \ (g\ \mathrm{d}\, x^{j_1} \wedge \ldots \wedge \mathrm{d}\, x^{j_q}) \right) \\ &= \mathrm{d}_U \left((f\, g)\ \mathrm{d}\, x^{i_1} \wedge \ldots \wedge \mathrm{d}\, x^{i_p} \wedge \mathrm{d}\, x^{j_1} \wedge \ldots \wedge \mathrm{d}\, x^{j_q} \right) \\ &= ((\mathrm{d}\, f)\ g + \ f\ (\mathrm{d}\, g)) \wedge \mathrm{d}\, x^{i_1} \wedge \ldots \wedge \mathrm{d}\, x^{i_p} \wedge \mathrm{d}\, x^{j_1} \wedge \ldots \wedge \mathrm{d}\, x^{j_q} \\ &= (\mathrm{d}\, f \wedge \mathrm{d}\, x^{i_1} \wedge \ldots \wedge \mathrm{d}\, x^{i_p}) \wedge (g\ \mathrm{d}\, x^{j_1} \wedge \ldots \wedge \mathrm{d}\, x^{j_q}) \\ &\qquad + (-1)^p (f\ \mathrm{d}\, x^{i_1} \wedge \ldots \wedge \mathrm{d}\, x^{i_p}) \wedge (\mathrm{d}\, g \wedge \mathrm{d}\, x^{j_1} \wedge \ldots \wedge \mathrm{d}\, x^{j_q}) \\ &= (\mathrm{d}_U\, \omega) \wedge \omega' + (-1)^p \omega \wedge (\mathrm{d}_U\, \omega') \ .\end{aligned}$$

As for the property $\mathrm{d}_U^2 = 0$, note that

$$\begin{aligned}\mathrm{d}_U^2\,\omega &= \mathrm{d}_U\left(\mathrm{d}_U(f\ \mathrm{d}\,x^{i_1}\wedge\ldots\wedge\mathrm{d}\,x^{i_p})\right)\\ &= \mathrm{d}_U\left(\sum_k \partial_{x^k} f\ \mathrm{d}\,x^k\wedge\mathrm{d}\,x^{i_1}\wedge\ldots\wedge\mathrm{d}\,x^{i_p}\right)\\ &= \sum_{k,\ell}\partial_{x^\ell}\partial_{x^k} f\ \mathrm{d}\,x^\ell\wedge\mathrm{d}\,x^k\wedge\mathrm{d}\,x^{i_1}\wedge\ldots\wedge\mathrm{d}\,x^{i_p}\\ &= 0\ ,\end{aligned}$$

in view of the symmetry $\partial_{x^\ell}\partial_{x^k} f = \partial_{x^k}\partial_{x^\ell} f$ and the antisymmetry $\mathrm{d}\,x^\ell\wedge\mathrm{d}\,x^k = -\,\mathrm{d}\,x^k\wedge\mathrm{d}\,x^\ell$.

If $\omega\in\Omega^p(M)$, we have $\mathrm{d}_U\,\omega|_U\in\Omega^{p+1}(U)$ and $\mathrm{d}_V\,\omega|_V\in\Omega^{p+1}(V)$, for every chart domains $U,V\subset M$. We will prove that the latter forms coincide on $U\cap V$, so that they can be glued and define a $(p+1)$-form on $U\cup V$. When considering all chart domains of an atlas of M, we thus get a map $\mathrm{d}\in\mathcal{L}(\Omega^p(M),\Omega^{p+1}(M))$, which inherits the required properties from the underlying d_U-s.

Denote now by $(x^1,\ldots,x^n)$ (resp. $(y^1,\ldots,y^n)$) the coordinates in U (resp. V). Let us examine the case $\omega\in\Omega^1(M)$: $\omega|_U = \omega_i\ \mathrm{d}\,x^i$ and $\omega|_V = \omega'_j\ \mathrm{d}\,y^j$, where the sum symbols have been omitted. Hence, we have

$$\omega_i\ \mathrm{d}\,x^i = \omega'_j\ \mathrm{d}\,y^j$$

in $U\cap V$, and must show that, in this overlap,

$$\partial_{x^k}\omega_i\ \mathrm{d}\,x^k\wedge\mathrm{d}\,x^i = \mathrm{d}\,\omega_i\wedge\mathrm{d}\,x^i = \mathrm{d}\,\omega'_j\wedge\mathrm{d}\,y^j = \partial_{y^\ell}\omega'_j\ \mathrm{d}\,y^\ell\wedge\mathrm{d}\,y^j\ . \tag{26}$$

Since $\mathrm{d}\,x^i = A^i{}_k\,\mathrm{d}\,y^k = \partial_{y^k}x^i\,\mathrm{d}\,y^k$ and $\omega_i = A'^j{}_i\,\omega'_j = \partial_{x^i}y^j\,\omega'_j$, the first sum reads

$$\partial_{x^k}\partial_{x^i}y^j\partial_{y^\ell}x^k\partial_{y^m}x^i\omega'_j\,\mathrm{d}\,y^\ell\wedge\mathrm{d}\,y^m + \partial_{x^i}y^j\partial_{y^\ell}x^k\partial_{y^m}x^i\partial_{x^k}\omega'_j\,\mathrm{d}\,y^\ell\wedge\mathrm{d}\,y^m.$$

As the Jacobian matrices $\partial_x y$ and $\partial_y x$ are inverses, the last term is equal to $\mathrm{d}\,\omega'_j\wedge\mathrm{d}\,y^j$. The first term vanishes in view of the above-mentioned symmetry-antisymmetry argument. □

The next theorem describes the main properties of the de Rham differential.

Theorem 2. *Let M and N be two manifolds.*

1. *For every $f\in C^\infty(M,N)$,*
$$\mathrm{d}\circ f^* = f^*\circ\mathrm{d}\ .$$
2. *For every $X\in\mathrm{Vect}(M)$, we have in $\Omega(M)$,*
$$L_X = [i_X,\mathrm{d}] = i_X\circ\mathrm{d} + \mathrm{d}\circ i_X\ . \tag{27}$$
3. *For every $X\in\mathrm{Vect}(M)$,*
$$[L_X,\mathrm{d}] = L_X\circ\mathrm{d} - \mathrm{d}\circ L_X = 0\ . \tag{28}$$
4. *If $\omega\in\Omega^p(M)$, $X_i\in\mathrm{Vect}(M)$, $i\in\{0,\ldots,p\}$, we have*
$$\begin{aligned}(\mathrm{d}\,\omega)(X_0,\ldots,X_p) &= \sum_{i=0}^p(-1)^i X_i\left(\omega(X_0,\ldots,\hat{\imath},\ldots,X_p)\right)\\ &\quad + \sum_{i<j}(-1)^{i+j}\omega([X_i,X_j],X_0,\ldots,\hat{\imath},\ldots,\hat{\jmath},\ldots,X_p)\ ,\end{aligned} \tag{29}$$
where $\hat{\imath}$ means, as usual, that the corresponding argument, here X_i, is omitted.

Proof. 1. Let $g \in C^\infty(N)$. Since $f^* g = g \circ f$, we get

$$(f^* \mathrm{d}\, g)_m = {}^t f_{*m}(\mathrm{d}\, g)_{f(m)} = T_{f(m)} g \circ T_m f = (\mathrm{d}\ f^* g)_m\ ,$$

$m \in M$, so that the property is valid for 0-forms. Let now $\omega \in \Omega^p(N)$, $p > 0$, let $(V, (y^1, \ldots, y^s))$ be a chart of N, and let $U = f^{-1}(V)$. In U, we have

$$f^* \mathrm{d}\, \omega = f^* \mathrm{d}(g\, \mathrm{d}\, y^{j_1} \wedge \ldots \wedge \mathrm{d}\, y^{j_p}) = f^* \mathrm{d}\, g \wedge f^* \mathrm{d}\, y^{j_1} \wedge \ldots \wedge f^* \mathrm{d}\, y^{j_p}$$

and

$$\mathrm{d}\, f^* \omega = \mathrm{d}\, (f^* g \wedge f^* \mathrm{d}\, y^{j_1} \wedge \ldots \wedge f^* \mathrm{d}\, y^{j_p}) = (\mathrm{d}\, f^* g) \wedge f^* \mathrm{d}\, y^{j_1} \wedge \ldots \wedge f^* \mathrm{d}\, y^{j_p}\ .$$

Finally $f^* \mathrm{d}\, \omega = \mathrm{d}\, f^* \omega$ in U and thus everywhere in M.

2. Remember that i_X (resp. d) is a graded derivation of degree -1 (resp. 1) of $(\Omega(M), \wedge)$. It follows that the graded commutator

$$[i_X, \mathrm{d}] = i_X \circ \mathrm{d} - (-1)^{(-1)\cdot 1}\, \mathrm{d} \circ i_X = i_X \circ \mathrm{d} + \mathrm{d} \circ i_X$$

is a (graded) derivation (of degree $-1 + 1 = 0$). The same is true for L_X. These derivations coincide if they coincide on any coordinate patch. Considering their locality, it is therefore sufficient to prove that they coincide on functions and exact forms. It is clear that, for every function g, we have

$$L_X g = (\mathrm{d}\, g)(X) = i_X(\mathrm{d}\, g) = i_X(\mathrm{d}\, g) + \mathrm{d}(i_X g) = [i_X, \mathrm{d}]\, g\ .$$

Moreover, if we prove that $L_X(\mathrm{d}\, g) = \mathrm{d}(L_X g)$, we get

$$L_X(\mathrm{d}\, g) = \mathrm{d}(i_X(\mathrm{d}\, g)) = (i_X \circ \mathrm{d} + \mathrm{d} \circ i_X)(\mathrm{d}\, g) = [i_X, \mathrm{d}](\mathrm{d}\, g)\ .$$

Since

$$\begin{aligned}
(L_X(\mathrm{d}\, g))_m &= d_t|_{t=0}\, \varphi^{\otimes}_{-t*}(\mathrm{d}\, g)_{\varphi_t(m)} \\
&= d_t\, {}^t\varphi_{t*}(\mathrm{d}\, g)_{\varphi_t(m)}\, \big|_{t=0} \\
&= d_t\, (\varphi_t^* \,\mathrm{d}\, g)_m\, |_{t=0} \\
&= \partial_t\ \mathrm{d}\, \varphi_t^* g\, |_{t=0,m} \\
&= \mathrm{d}\ \partial_t\, (g \circ \varphi_t)\, |_{t=0,m} \\
&= (\mathrm{d}(L_X g))_m\ ,
\end{aligned}$$

for every $m \in M$, the proof is complete.

3. The result is a direct consequence of the previous one.

4. For $p = 0$ and $\omega = f \in \Omega^0(M)$, the announced result reads $(\mathrm{d}\, f)(X_0) = X_0(f)$, which is nothing but Equation (19). We now proceed by induction. It then follows from (27) and

(17) that

$$
\begin{aligned}
(\mathrm{d}\,\omega)(X_0, X_1, \ldots, X_p) &= (i_{X_0}\,\mathrm{d}\,\omega)(X_1, \ldots, X_p)\\
&= (L_{X_0}\omega)(X_1, \ldots, X_p) - (\mathrm{d}\; i_{X_0}\omega)(X_1, \ldots, X_p)\\
&= X_0(\omega(X_1, \ldots, X_p)) - \sum_{i=1}^{p} \omega(X_1, \ldots, [X_0, X_i], \ldots, X_p)\\
&\quad - \sum_{i=1}^{p} (-1)^{i-1} X_i\big((i_{X_0}\omega)(X_1, \ldots, \hat{\imath}, \ldots, X_p)\big)\\
&\quad - \sum_{1\le i<j\le p} (-1)^{i+j} (i_{X_0}\omega)([X_i, X_j], \ldots, \hat{\imath}, \ldots, \hat{\jmath}, \ldots)\\
&= \sum_{i=0}^{p} (-1)^i X_i\,(\omega(X_0, \ldots, \hat{\imath}, \ldots, X_p))\\
&\quad + \sum_{0\le i<j\le p} (-1)^{i+j} \omega([X_i, X_j], \ldots, \hat{\imath}, \ldots, \hat{\jmath}, \ldots)\;.
\end{aligned}
$$

□

Exercise. Write Equation (29) explicitly for $p = 1$ and $p = 2$.

Proposition 8. *For every vector fields $X, Y \in \mathrm{Vect}(M)$, the results*

$$[\mathrm{d}, i_X] = L_X, \quad [\mathrm{d}, L_X] = 0, \quad [i_X, i_Y] = 0, \quad [L_X, i_Y] = i_{[X,Y]} \quad \textit{and} \quad [L_X, L_Y] = L_{[X,Y]}\;, \tag{30}$$

concerning graded commutators of the degree -1 (resp. $0, 1$) derivations i_X (resp. L_X, d) of the algebra $(\Omega(M), \wedge)$, are valid and are referred to as the Cartan calculus on smooth manifolds.

Proof. The first and second results are known, see Equations (27) and (28); the third is a direct consequence of the skew-symmetry of differential forms. As for the fourth, note that both sides are graded derivations of the algebra $(\Omega(M), \wedge)$. It follows that they are local and that it thus suffices to prove that they coincide on functions and differential 1-forms. Since they have degree -1, they both vanish on functions, whereas, for $\omega \in \Omega^1(M)$, we get

$$i_{[X,Y]}\omega = \omega([X, Y]) = \omega(L_X Y) = L_X(\omega(Y)) - (L_X\omega)(Y) = [L_X, i_Y]\omega\;,$$

in view of Equation (17). We already observed that the last of the equations (30) is valid on functions and on vector fields, see Chapter 5. Since the graded commutator of graded derivations of $\Omega(M)$ is a graded Lie algebra bracket, it satisfies the graded Jacobi identity,

$$L_{[X,Y]} = [\mathrm{d}, i_{[X,Y]}] = [\mathrm{d}, [L_X, i_Y]] = [[\mathrm{d}, L_X], i_Y] + [L_X, [\mathrm{d}, i_Y]] = [L_X, L_Y]\;,$$

so that this result is valid on differential forms as well. This completes the proof of Proposition 8. However, we still have to prove Equation (18) of Proposition 7, i.e. we must show that the result is valid for every tensor field $T \in \Gamma(\otimes TM)$. Again, both sides $L_{[X,Y]}$ and $[L_X, L_Y]$ being derivations of $(\Gamma(\otimes TM), \otimes)$ and therefore local operators, we just need to check that they coincide on a coordinate patch $(U, x^1, \ldots, x^n)$. Since a tensor field reads there

$$T_U = \sum_{\substack{i_1,\ldots,i_p\\ j_1,\ldots,j_q}} t^{i_1\ldots i_p}_{j_1\ldots j_q}\, \partial_{x^{i_1}} \otimes \ldots \otimes \partial_{x^{i_p}} \otimes \mathrm{d}\,x^{j_1} \otimes \ldots \otimes \mathrm{d}\,x^{j_q}\;,$$

see Equation (6), these derivations coincide at T_U, as they do on functions, vector fields, and differential forms. □

Chapter 7
Integral Calculus on Manifolds

1 Orientable manifolds

The concept of orientability is well known for smooth surfaces: a smooth surface is orientable if it admits a continuous field of unit normal vectors. We now want to generalize the notion of orientability to smooth manifolds. What it means for a smooth manifold to be orientable becomes clear when we consider the non-orientable Möbius strip M.

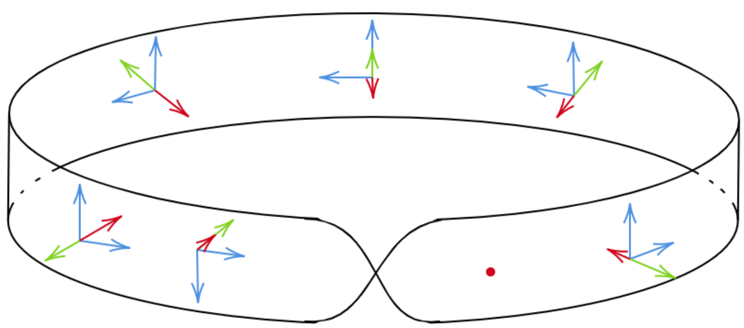

Figure 14: Non-orientable Möbius strip M

The blue arrows represent bases of the corresponding tangent spaces. Since the two leftmost bases $(\partial_{x^1}, \partial_{x^2})$ and $(\partial_{y^1}, \partial_{y^2})$, where the first (resp. second) vectors are horizontal (resp. vertical), are direct bases their transition matrix, which equals the Jacobian matrix of the coordinate transformation from x- to y-coordinates, satisfies $\det \partial_x y > 0$. However, as indicated in the above figure we cannot equip the whole manifold with coordinates that satisfy this condition. This means that there does not exist any atlas $\mathcal{A}_M = (U_\alpha, \varphi_\alpha)_\alpha$ satisfying

$$\det(\varphi_\beta \circ \varphi_\alpha^{-1})'(x) > 0$$

for all $x \in \varphi_\alpha(U_\alpha \cap U_\beta)$ and for all indices α and β. This is a defining criterion for non-orientability. Moreover, it can be observed that on non-orientable manifolds such as M there does not exist any nowhere vanishing smooth top-form, which constitutes an equivalent criterion for non-orientability. Indeed, the top-form represented by the green arrows is not smooth and the one indicated by the red arrows vanishes. We conclude that the orientable smooth manifolds are those that admit a nowhere vanishing (smooth) top-form, or, equivalently, those that can be equipped with an atlas whose Jacobian matrices have strictly positive determinants.

2 Integration over orientable manifolds

Let N be a p-dimensional smooth connected and *orientable* manifold and let Ω be a volume form of N, i.e. a nowhere vanishing (smooth) top-form. We fix an orientation of N, either Ω or $-\Omega$, say we pick Ω, and we choose a compatible atlas $\mathcal{A}_N$, i.e. an atlas where the determinant of the Jacobian matrix of each coordinate transformation is strictly positive and where

$$\Omega|_{\varphi(U)}(x) = X(x)\, dx^1 \wedge \cdots \wedge dx^p \tag{1}$$

with $X \in C^\infty(\varphi(U), \mathbb{R}_{>0})$ for every chart $(U, \varphi = (x^1, \ldots, x^p)) \in \mathcal{A}_N$. Note that (1) is obvious and that the assumption here is that the values of X are strictly positive. Observe also that if we choose a chart $(V, \psi = (y^1, \ldots, y^p)) \in \mathcal{A}_N$ such that $U \cap V \neq \emptyset$, if

$$\Omega|_{\psi(V)}(y) = Y(y)\ dy^1 \wedge \cdots \wedge dy^p$$

and if we denote the coordinate transformation

$$\varphi \circ \psi^{-1} : \psi(U \cap V) \to \varphi(U \cap V)$$

by $x = x(y)$, we get

$$dx^i = \sum_{\sigma_i=1}^{p} \partial_{y^{\sigma_i}} x^i dy^{\sigma_i}$$

and

$$\begin{aligned}
Y(y)\ dy^1 \wedge \cdots \wedge dy^p &= \Omega|_{\psi(U\cap V)}(y) \\
&= X(x(y)) \sum_{\sigma=(\sigma_1\cdots\sigma_p)} \partial_{y^{\sigma_1}} x^1 \cdots \partial_{y^{\sigma_p}} x^p\ dy^{\sigma_1} \wedge \cdots \wedge dy^{\sigma_p} \\
&= X(x(y)) \sum_{\sigma=(\sigma_1\cdots\sigma_p)\in\mathbb{S}_p} \partial_{y^{\sigma_1}} x^1 \cdots \partial_{y^{\sigma_p}} x^p\ \operatorname{sign}\sigma\ dy^1 \wedge \cdots \wedge dy^p \\
&= X(x(y))\ \det \partial_y x\ \ dy^1 \wedge \cdots \wedge dy^p\ ,
\end{aligned}$$

so that

$$Y(y) = X(x(y))\, \det \partial_y x\ . \tag{2}$$

The transformation law (2) of the volume form component is coherent with respect to our positivity assumptions.

We now come to the integral over the connected orientable manifold N associated with the volume form Ω.

Let us first consider a compactly supported continuous function $f \in C^0_c(N)$, whose support is contained in a chart domain U with coordinates $\varphi = (x^1, \ldots, x^p)$. If $f(x)$ is this function read in these coordinates and $X(x)\,dx^1\wedge\ldots\wedge dx^p$ is the coordinate form of the volume, we define the integral over N of f with respect to Ω by

$$\int_N f\,\Omega=\int_U f\,\Omega=\int_{\varphi(U)} f(x)\,X(x)\,dx^1\wedge\ldots\wedge dx^p:=\int_{\varphi(U)} f(x)\,X(x)\,dx^1\ldots dx^p\in\mathbb{R}\,, \tag{3}$$

where the RHS is the Lebesgue integral in $\mathbb{R}^p$, which makes sense as the integrated function is continuous and compactly supported in $\varphi(U)$.

We then pass to an arbitrary $f\in C^0_c(N)$ by means of a partition of unity $(U_\alpha,\varphi_\alpha,\pi_\alpha)_\alpha$ subordinate to the compatible atlas $\mathcal{A}_N$, i.e. we set

$$\int_N f\,\Omega:=\sum_\alpha\int_N(\pi_\alpha f)\,\Omega\in\mathbb{R}\,, \tag{4}$$

where the terms of the RHS are defined by (3) and the sum over α is finite because the cover is locally finite and the support of f is compact.

The integrals (3) and (4) only depend on f and neither on the chosen charts, nor on the partition of unity considered. The independence from the partition of unity follows from a standard proof from integration theory, which we will not repeat here, while the independence from the chosen chart can be easily checked. Indeed, if $(V,\psi=(y^1,\ldots,y^p))$ is another chart whose domain contains the support of f, we can write as well

$$\int_N f\,\Omega:=\int_{\psi(V)}\mathfrak{f}(y)\,Y(y)\,dy^1\ldots dy^p=\int_{\psi(U\cap V)}\mathfrak{f}(y)\,Y(y)\,dy^1\ldots dy^p\,, \tag{5}$$

where $\mathfrak{f}(y)$ is the function f read in the coordinates y. However, if we perform the coordinate transformation $x=x(y)\rightleftharpoons y=y(x)$ in the integral (3) which is also given by the Lebesgue integral

$$\int_{\varphi(U\cap V)} f(x)\,X(x)\,dx^1\ldots dx^p\,,$$

we get

$$\int_{\psi(U\cap V)} f(x(y))\,X(x(y))|\det\partial_y x|\;dy^1\ldots dy^p,$$

which, in view of the function transformation law $f(x(y))=\mathfrak{f}(y)$, the volume transformation law (2) and the positive sign of the Jacobian determinant, coincides with the integral (5). This also makes it understandable why it is so important that the manifold over which we integrate is orientable. The assumption of connectedness, on the other hand, is merely a simplifying assumption to limit the number of orientations of N to 2: connectedness can easily be avoided.

Note now that Equations (3) and (4) clearly define a positive linear form on $C^0_c(N)$.

Since the oriented smooth p-dimensional [Hausdorff and second countable] manifold N is a locally compact topological space and a countable union of compact subspaces, the general theory of Radon measures allows us to extend the positive linear form $\int_N -\,\Omega$ on $C^0_c(N)$ to a larger space $L^1_\Omega(N)\supset C^0_c(N)$. We denote μ this extension or measure and say that the functions $g\in L^1_\Omega(N)$ (we also write $g\in L^1_\mu(N)$) are integrable over N with respect to Ω (or μ) and their integral is defined by

$$\int_N g\,\Omega:=\mu(g)\in\mathbb{R}$$

(we also write $\int_N g\,\mu=\mu(g)\in\mathbb{R}$).

3 Integration over arbitrary manifolds

The Cartesian space $\mathbb{R}^n$ admits a canonical measure, the Lebesgue measure, which we denote by $\delta_0 = |dx^1 \wedge \ldots \wedge dx^n|$. The theorem that allows to change coordinates in a Lebesgue integral now reads as follows. Let $x = x(y) \rightleftharpoons y = y(x)$ be a diffeomorphism between two open subsets U and V of $\mathbb{R}^n$. We have $f(x) \in L^1_{\delta_0}(U)$ if and only if $f(x(y))|\det \partial_y x| \in L^1_{\delta_0}(V)$ and

$$\int_U f(x)\,|dx^1 \wedge \ldots \wedge dx^n| = \int_V f(x(y))|\det \partial_y x||dy^1 \wedge \ldots \wedge dy^n|\,. \tag{6}$$

The appropriate objects for integration over a (not necessarily orientable) smooth n-dimensional manifold M are 1-densities. Roughly, 1-densities are differential top forms up to sign. More precisely, a 1-density on the vector space T_mM ($m \in M$) is a map

$$\mathfrak{d} : \wedge^n T_mM \setminus \{0\} \to \mathbb{R}\,,$$

such that, for every $s \in \mathbb{R} \setminus \{0\}$ and every $\Pi \in \wedge^n T_mM \setminus \{0\}$, we have

$$\mathfrak{d}(s\Pi) = |s|^\lambda \mathfrak{d}(\Pi)\,,$$

with $\lambda = 1$. If λ is an arbitrary real number, $\mathfrak{d}$ is a λ-density of T_mM. It is clear that the set $D_\lambda(T_mM)$ of all λ-densities of T_mM is a real 1-dimensional vector space and that the disjoint union $D_\lambda(M) = \sqcup_m D_\lambda(T_mM)$ is a rank 1 vector bundle over M. Indeed, if $\omega \in \wedge^n T^*_mM$ is a nonzero top linear form of the tangent space, then $|\omega|^\lambda$ is a basis of $D_\lambda(T_mM)$. A λ-density field of M is then a smooth section $\delta \in \mathfrak{D}_\lambda(M) := \Gamma(D_\lambda(M))$ of the λ-density bundle. If no confusion is possible it is customary to speak about λ-densities instead of fields of such densities and about densities instead of 1-densities. From what has just been said it is obvious that over a coordinate chart $(U, \varphi = (x^1, \ldots, x^n))$ of M a λ-density δ reads

$$\delta|_{\varphi(U)}(x) = X(x)|dx^1 \wedge \ldots \wedge dx^n|^\lambda,$$

where $X(x)$ is smooth. Observe that if $(V, \psi = (y^1, \ldots, y^n))$ is another coordinate patch and if

$$\delta|_{\psi(V)}(y) = Y(y)|dy^1 \wedge \ldots \wedge dy^n|^\lambda\,,$$

then, in view of the computations in Section 2, the component transformation law for λ-densities is

$$Y(y) = X(x(y))|\det \partial_y x|^\lambda. \tag{7}$$

The point with densities is that for $\lambda = 1$ the basis vector $|\omega|$ is a volume element of the tangent space, viewed up to $\mathbb{Z}_2$-action. Whereas on non-orientable manifolds a global top differential form is either not smooth or has to vanish at some point, it is intuitively clear that a global smooth nowhere vanishing top differential form up to sign, i.e. a global smooth nevervanishing 1-density field, must exist even for non-orientable manifolds. It follows that the line bundle $D_1(M)$ is trivial.

Let us now come to the integral over a manifold M associated with a 1-density $\delta \in \mathfrak{D}_1(M)$. We know from what has been said above that it suffices to show that this density defines a positive linear form on $C^0_c(M)$.

It suffices to proceed as in the orientable case.

Let us first consider a continuous function $f \in C^0_c(M)$ that is compactly supported by a chart domain U with coordinates $\varphi = (x^1, \ldots, x^n)$. If $f(x)$ is this function read in these coordinates and $X(x)|dx^1 \wedge \ldots \wedge dx^n|$ is the coordinate form of the density δ, we define the measure or integral over M of f associated to δ by

$$\int_M f\,\delta := \int_{\varphi(U)} f(x)\,X(x)|dx^1 \wedge \ldots \wedge dx^n|\;, \tag{8}$$

where the RHS Lebesgue integral makes sense as the integrated function is continuous and compactly supported in $\varphi(U)$.

We then pass to an arbitrary $f \in C^0_c(M)$ by means of a partition of unity $(U_\alpha, \varphi_\alpha, \pi_\alpha)_\alpha$ subordinate to the charts of an atlas, i.e. we set

$$\int_M f\,\delta := \sum_\alpha \int_M (\pi_\alpha f)\,\delta\;. \tag{9}$$

This defines obviously a positive linear form and the integrals (8) and (9) only depend on f and neither on the chosen charts, nor on the partition of unity considered. Indeed, if $(V, \psi = (y^1, \ldots, y^n))$ is another chart that contains the support of f, we can write as well

$$\int_M f\,\delta := \int_{\psi(V)} \mathfrak{f}(y)\,Y(y)|dy^1 \wedge \ldots \wedge dy^n|\;. \tag{10}$$

When performing the coordinate transformation $x = x(y) \rightleftharpoons y = y(x)$ in the integral (8), see (6), we find

$$\int_{\psi(V)} f(x(y))\,X(x(y))|\det \partial_y x||dy^1 \wedge \ldots \wedge dy^n|\;,$$

which, in view of transformation law (7), coincides with the integral (10). This also allows us to understand why orientability is not needed if we integrate with respect to a 1-density.

4 Exercises

Additional exercises – related to the whole course – will be proposed in separate files provided in the UL Learning Management System MOODLE.

5 Disclaimer

The present text served as a reference for a Master Course given by the author at the University of Luxembourg. As the notes gradually grew over the years, some bibliographic data may have been lost or forgotten; in this case, the author is happy to add these references.

References

[BG92] M. Berger, B. Gostiaux, Géométrie différentielle: variétés, courbes et surfaces, PUF, ISBN:2 13 044708 2

[CJ04] D. Chruscinski, A. Jamiolkowski: Geometric Phases in Classical and Quantum Mechanics (chapter 1), Progress in Mathematical Physics, Birkhäuser, `ISBN:0 8176 4282 X`

[DWi99] De Wilde Marc, Géométrie différentielle, University of Liège, 1999

[Hit08] Hitchin Nigel, Differentiable Manifolds, Oxford, 2008

[Mic08] P.W. Michor: Topics in Differential Geometry (chapters 1-3), Graduate Studies in Mathematics, 93, American Mathematical Society, `ISBN-10:0821820036`, `ISBN-13:978-0821820032`

[Nak03] M. Nakahara: Geometry, Topology and Physics (chapter 5), Graduate Student Series in Physics, Institute of Physics Publishing, `ISBN-10:0750306068`, `ISBN-13:978-0750306065`

[nLab] Free Internet Encyclopedia 'nLab', `http://ncatlab.org/nlab/show/HomePage`

[Spi99] M. Spivak: A Comprehensive Introduction to Differential Geometry, 1, Publish or Perish, `ISBN-10:0914098705`, `ISBN-13:978-0914098706`

[War83] F.W. Warner: Foundations of Differentiable Manifolds and Lie groups (chapters 1-5), Graduate texts in Mathematics, 94, Springer, `ISBN-10:0387908943`, `ISBN-13:978-0387908946`

[Wiki] Free Internet Encyclopedia 'Wikipedia', `http://en.wikipedia.org/wiki/`