

Manifolds 2014

JAC / FM

Contents

1	Introduction; topological manifolds	2
2	Differentiable manifolds	6
3	Sard's theorem and the Inverse Function Theorem	12
3.1	The Inverse Function Theorem	13
4	Rank theorems	15
5	Embeddings	18
6	Bump functions and partitions of unity	21
6.1	Useful functions	21
6.2	Partition of unity	24
7	The tangent bundle	27
7.1	Operations on vector fields	32
8	Tensors	33
8.1	Miscellany on the exterior algebra	33
8.2	The cotangent bundle	33
8.3	General tensor fields	34
9	Vector fields and differential equations	35
9.1	Lie derivatives	38
10	Differential forms	40
10.1	Basic theorems	40
10.2	The de Rham complex	40
10.3	The Poincaré Lemma	41

11 Integration	44
11.1 The boundary of a chain	46
11.2 Connection to the fundamental group	48
A Definitions	50
B Formulas	51
B.1 Local description of (co-)vector fields	51
B.2 Pullbacks and pushforwards	52
B.3 Lie derivatives	53
C A concrete exercise	56

1 Introduction; topological manifolds

Manifolds are geometric objects that locally look like some \mathbb{R}^n for some natural number n . In that respect, they are very easy to understand. The interesting things happen when we glue several pieces of \mathbb{R}^n together.

In Calculus, manifolds (in those rare occasions they were encountered) usually appeared embedded in some larger \mathbb{R}^n or \mathbb{C}^n . This is not the path we're going to take - for us, manifolds will be objects in their own right, existing independently of some "ambient" space. The way to do this is, exactly, glueing.

We start with the definition of a *topological manifold*.

Definition 1.1. A *topological manifold* M is a topological space such that each $x \in M$ has a neighbourhood U such that U is homeomorphic to \mathbb{R}^n . ■

Here are some examples of manifolds.

Example 1.2. Setting $M = \mathbb{R}^n$, set $U = \mathbb{R}^n$ for all $x \in M$. ★

Example 1.3. Any open ball B in \mathbb{R}^n : If $x \in B$, let $U = B$. Then I claim that $U \approx \mathbb{R}^n$: just use the map $\vec{x} \mapsto \frac{1}{1-|\vec{x}|}\vec{x}$. ★

Example 1.4. By the previous example, it follows that any open $U \subseteq \mathbb{R}^n$ is a manifold, by definition of an open set as a union of balls. ★

Example 1.5. Similarly, any open subset of a manifold is a manifold. ★

Example 1.6. Up to homeomorphism, there are only two 1-dimensional manifolds. Namely, the circle S^1 and the real line \mathbb{R} . The first is compact, the other is not. ★

Example 1.7. By stereographic projection, every n -sphere can be covered by two sheets homeomorphic to \mathbb{R}^n . See Figure 1. ★

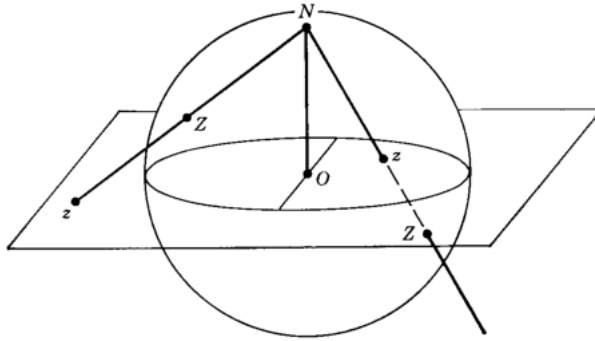


Figure 1: Stereographic projection.

Example 1.8. The \aleph_0 number of compact surfaces. These are 2-spheres attached with n handles, for some natural number n . ★

Example 1.9. The *non-orientable* manifolds. An example of a non-orientable manifold is the Möbius band in Figure 2 with the boundary removed. If we keep the boundary, the figure is an example of a *manifold with boundary*.

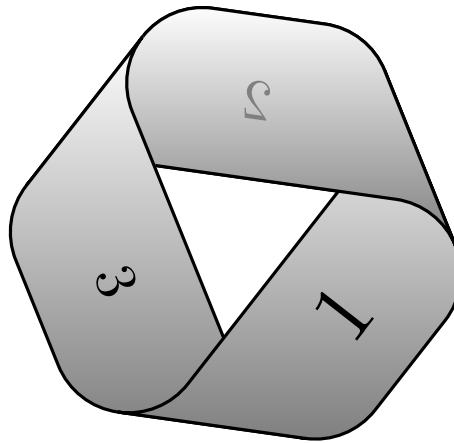


Figure 2: The Möbius band.

★

Example 1.10. The last example (for now) is the projective plane over the reals, called $\mathbb{P}_{\mathbb{R}}^2$. By definition, as a set, it is the set of lines through the origin in \mathbb{R}^3 . However, it has a nice manifold structure. First, note that a line through the origin in \mathbb{R}^3 is determined by giving a point on the 2-sphere S^2 . This point is however not unique: the antipodal point gives the same line. This means that we can identify $\mathbb{P}_{\mathbb{R}}^2$ with the quotient space S^2/\sim where $x \sim -x$. We now give the quotient the quotient topology.

This is not completely enlightening though. There is no obvious local homeomorphisms with \mathbb{R}^2 , nor do we really have a picture of how $\mathbb{P}_{\mathbb{R}}^2$ looks like. The second problem can be solved like this: Note that any line not in the plane $z = 0$, has a *unique* representative in the open north hemisphere. Similarly, every line in the plane $z = 0$, but with $y \neq 0$, has a unique representative on the equator minus a point. Thus $\mathbb{P}_{\mathbb{R}}^2$ can be written as the union of a three cells, namely a B_2 (a 2-ball), a B_1 and a B_0 . So $\mathbb{P}_{\mathbb{R}}^2 = B_2 \cup B_1 \cup B_0$, but glued in a special way. We can think of this glueing as “adding points/lines at infinity”.

There is a natural choice of coordinates on $\mathbb{P}_{\mathbb{R}}^2$. We define *homogeneous coordinates*: a point P can be represented by a 3-tuple $[x_0 : x_1 : x_2]$, and this tuple is unique up to multiplication by $\mathbb{R} \setminus \{0\}$. In other words, every point has a representative of the form $[x_0 : x_1 : x_2]$, where $[x_0 : x_1 : x_2] = [\lambda x_0 : \lambda x_1 : \lambda x_2]$. Define the “basic open sets U_i ” as $U_i := \{P \in \mathbb{P}_{\mathbb{R}}^2 \mid x_i \neq 0\}$. There is a natural bijection between U_0 , say, and \mathbb{R}^2 . In U_0 , every point has a *unique* representative of the form $[1 : x_1 : x_2]$, and this can be mapped to $(x_1, x_2) \in \mathbb{R}^2$. Now it is an **exercise** to show that the set U_i is open in the quotient topology.

It is easy to see that the U_i cover $\mathbb{P}_{\mathbb{R}}^2$, and so this defined the structure of a topological manifold on $\mathbb{P}_{\mathbb{R}}^2$. ★

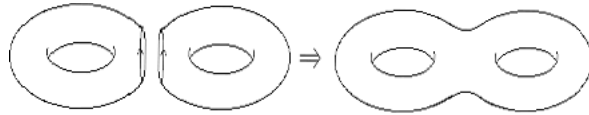


Figure 3: The connected sum of two tori.

Example 1.11. One can make new manifolds from old ones by means of the *connected sum*.

One starts with two manifolds M and N , removes an open disk from both of them. The boundary of a disk is a S^1 , and we can identify the two

boundaries by an arbitrary homeomorphism $\varphi : S^1 \rightarrow S^1$. We glue to get $M \# N$. See Figure 3. ★

For the Möbius band example to be a manifold, we have to define a manifold with boundary:

Definition 1.12. Let M be a topological space. We say that M is a *manifold with boundary* if every point of x has a neighbourhood homeomorphic to either \mathbb{R}^n or $H_n := \{(x^1, \dots, x^n) : x^n \geq 0\}$, the upper half-plane. ■

We call the n in the definition the *dimension* of the manifold. However, to prove that the dimension is well-defined, reduces to proving that if $\mathbb{R}^n \approx \mathbb{R}^m$, then $m = n$. This is a very non-trivial problem, best proved with tools from algebraic topology, using such tools as homology groups and long exact sequences.

2 Differentiable manifolds

To get a course worth something, one has to have something more than just a topological manifold. We do this by adding structure to M , thereby defining the notion of a *differentiable* manifold.

Given two open subsets U, V of M and homeomorphisms $x : U \rightarrow x(U) \subseteq \mathbb{R}^n$ and $y : V \rightarrow y(V) \subseteq \mathbb{R}^n$, one has transition maps on the intersection $U \cap V$:

$$\begin{array}{ccc} U \cap V & \xrightarrow{y} & y(U \cap V) \subseteq \mathbb{R}^n \\ \downarrow x & \nearrow x \circ y^{-1} & \\ \mathbb{R}^n \supseteq x(U \cap V) & & \end{array}$$

We say that the two charts (x, U) and (y, V) are *C^∞ -related* if the map $x \circ y^{-1}$ and the map $y \circ x^{-1}$ are C^∞ -maps as maps between subsets of \mathbb{R}^n .

Furthermore, we say that a set of charts $\{(x_i, U_i)\}_{i \in I}$ is an *C^∞ -atlas* for M if all the x_i, x_j are C^∞ -related and the union $\bigcup_i U_i$ cover M .

Here a uniqueness result:

Lemma 2.1. *Given an atlas \mathcal{A} on M , there exists a unique maximal atlas \mathcal{A}' containing \mathcal{A} .*

Proof. Define \mathcal{A}' to be the set of all charts y which are C^∞ -related to all charts $x \in \mathcal{A}$. Then \mathcal{A}' contains two types of charts: those who belonged to \mathcal{A} , and possibly new ones. Every two charts in \mathcal{A}' are C^∞ -related, by definition of atlas, and every pair with one in each are C^∞ -related by definition of \mathcal{A}' . Further, every pair of charts with both in \mathcal{A}' are C^∞ -related: one can intersect their domains with a chart in \mathcal{A} , compose, and conclude.

Clearly \mathcal{A}' is the unique maximal atlas containing \mathcal{A} . \square

Thus:

Definition 2.2. A *differentiable manifold* is a pair (M, \mathcal{A}) , where M is a topological manifold, and \mathcal{A} is a maximal atlas (of C^∞ -related transition functions). \blacksquare

From now on, when we say “manifold”, we will always mean “differentiable manifold”. Note that it is defined as a *pair*, and we cannot forget the atlas, because there are, famously, homeomorphic topological manifolds with different differentiable structures (for example, Milnor’s “exotic spheres”, the first one being the 28 different differentiable structures on S^7).

Here’s an example (and probably the only explicit calculation we will do in the course):

Example 2.3. Recall that S^2 (the zero set of the equation $x_1^2 + x_2^2 + x_3^2 = 1$ in \mathbb{R}^3) can be covered by two charts, by stereographic projection (see Figure 1), by first projecting from the north pole, and then from the south pole. Call these maps φ and ψ , respectively. We are going to compute the transition maps, to verify that this atlas is actually a C^∞ -atlas.

We need to find explicit formulae for φ and ψ . The first map has domain $S^2 \setminus \{N\}$ and image \mathbb{R}^2 . But we want to compute $\psi \circ \varphi^{-1}$, so we start with a point (a, b) in the plane \mathbb{R}^2 (embedded in \mathbb{R}^3 by $z = 0$). We want to connect it with the north pole $(0, 0, 1)$ by a line, and compute its intersection with the 2-sphere. The line is given parametrically as

$$t(a, b, 0) + (1 - t)(0, 0, 1) = (at, bt, 1 - t).$$

To find the intersection with the sphere, we must compute when the right hand side has norm 1:

$$a^2 t^2 + b^2 t^2 + 1 - 2t + t^2 = 1$$

Getting rid of the 1 and cancelling the t , we end up with the condition

$$t = \frac{2}{1 + \|a\|},$$

where $\|a\| = \sqrt{a^2 + b^2}$. In other words, explicitly, we have

$$\varphi^{-1}(a, b) = \left(\frac{2a}{1 + \|a\|}, \frac{2b}{1 + \|a\|}, \frac{\|a\| - 1}{1 + \|a\|} \right).$$

To compute ψ , one start with a point on the sphere, and find a formula for the the intersection with the x, y -plane. So let (x, y, z) be on the sphere. Then $\psi(x, y, z)$ is on the line from S (the south pole) to the plane $z = 0$. This line is parametrically given by

$$t(x, y, z) + (1 - t)(0, 0, -1) = (tx, ty, tz - 1 + t).$$

The z -coordinate is zero when $t = \frac{1}{z+1}$. Thus

$$\psi(x, y, z) = \left(\frac{x}{z + 1}, \frac{y}{z + 1} \right)$$

Finally, we can compute the composition $\psi \circ \varphi^{-1}$:

$$\begin{aligned}
\psi \circ \varphi^{-1}(a, b) &= \psi \left(\frac{2a}{1 + \|a\|}, \frac{2b}{1 + \|a\|}, \frac{\|a\| - 1}{1 + \|a\|} \right) \\
&= \left(\frac{2a}{1 + \|a\|} / \frac{2\|a\|}{1 + \|a\|}, \frac{2b}{1 + \|a\|} / \frac{2\|a\|}{1 + \|a\|} \right) \\
&= \left(\frac{a}{\|a\|}, \frac{b}{\|a\|} \right).
\end{aligned}$$

So the transition maps are just inversion about the origin in the (a, b) -plane. These are clearly C^∞ functions. ★

The moral of the story is that we could just as well define S^2 as two copies of $\mathbb{R}^2 \setminus \{0\}$ glued together with the inversion map, because that's exactly how the differentiable structure is defined, and that is all we care about.

Here's another example:

Example 2.4. Recall that we could write the projective plane $\mathbb{P}_{\mathbb{R}}^2$ as the union of three open sets U_i , defined by the non-vanishing of one of the homogeneous coordinates. Here one computes (easily) that the transition maps $\varphi_1 \circ \varphi_0^{-1} : \mathbb{R}_{(a,b)}^2 \setminus \{a = 0\} \rightarrow \mathbb{R}_{(a,b)}^2 \setminus \{a = 0\}$ are given by

$$(a, b) \mapsto \left(\frac{1}{a}, \frac{b}{a} \right).$$

★

It is time to define *maps* between them.

Definition 2.5. Let M^n, N^m be manifolds (of dimension n, m , respectively), and let $f : M \rightarrow N$ be a continuous map between them. We say that f is *differentiable at p* if for all charts (x, U) with $p \in U$ and for all charts (y, V) with $f(p) \in V$, the function $y \circ f|_U \circ x^{-1}$ is differentiable (as a function between open subsets of \mathbb{R}^n and \mathbb{R}^m).

$$\begin{array}{ccc}
& U & \xrightarrow{f|_U} V \\
x^{-1} \nearrow & & \searrow y \\
x(U) & \xrightarrow{y \circ f \circ x^{-1}} & y(V)
\end{array}$$

If f is differentiable at all points $p \in M$, then we say that f is *differentiable*. ■

Thus we have a category **Diff** of differentiable manifolds. Its objects are the differentiable manifolds and the maps are differentiable maps. This gives us at once the notion of an *isomorphism* of manifolds: it is just a pair of differentiable maps that compose to the identity in each direction.

We want to define some properties of maps in **Diff**. To start off, we introduce some notation. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be differentiable map, then we define

$$D_i(f)(a) = \lim_{h \rightarrow 0} \frac{f(a^1, \dots, a^i + h, \dots, a^n)}{h}.$$

This is nothing but the i 'th partial derivative. The reason we're not using the usual Leibniz notation is because we want to reserve it for something else. Having established the D_i notation, one recalls the chain rule: Given two maps $g : \mathbb{R}^m \rightarrow \mathbb{R}^n$ and $f : \mathbb{R}^n \rightarrow \mathbb{R}$, the D_j of the composite is given as

$$D_j(f \circ g)(a) = \sum_{i=1}^n D_i(f \circ g)(g(a)) \cdot D_j(g^i)(a)$$

The classical Leibniz notation will be reserved for tangent fields on manifolds. So let $f : M \rightarrow \mathbb{R}$ be a map from a manifold to the real line, and let (x, U) be a chart, and let $p \in U$. Then we define

$$\frac{\partial f}{\partial x^i}(p) = \left. \frac{\partial f}{\partial x^i} \right|_p := D_i(f \circ x^{-1})(x(p))$$

Thus $\left. \frac{\partial f}{\partial x^i} \right|_p$ measures the rate of change of the function $f : M \rightarrow \mathbb{R}$ with respect to the coordinate system (x, U) .

Thus we may ask: what happens if we change charts?

Proposition 2.6. *Let $f : M \rightarrow \mathbb{R}$ be a map and let $(x, U), (y, V)$ be two overlapping coordinate charts. Then*

$$\frac{\partial f}{\partial y^i} = \sum_{j=1}^n \frac{\partial f}{\partial x^j} \frac{\partial x^j}{\partial y^i}.$$

Or in matrix notation:

$$\left[\frac{\partial f}{\partial y^i} \right]_{i=1 \dots n} = \left[\frac{\partial x^j}{\partial y^i} \right]_{i=1 \dots n, j=1 \dots n} \left[\frac{\partial f}{\partial x^j} \right]_{j=1 \dots n}.$$

Proof. This is a definition-chase and the chain rule. Explicitly, by definition:

$$\begin{aligned}
\left. \frac{\partial f}{\partial y^i} \right|_p &= D_i(f \circ y^{-1})(y(p)) = D_i((f \circ x^{-1}) \circ (x \circ y^{-1}))(y(p)) \\
&= \sum_{j=1}^n D_j(f \circ x^{-1})(x(p)) D_i((x \circ y^{-1})^j)(y(p)) \\
&= \sum_{j=1}^n \left. \frac{\partial f}{\partial x^j} \right|_p \cdot \left. \frac{\partial x^j}{\partial y^i} \right|_p
\end{aligned}$$

The chain rule appear in the middle. \square

Remark. From now on, all maps will be differentiable, unless otherwise stated. Thus, when referring to a “map” above, it should be read as “differentiable map” (or better, “a map in the category Diff ”).

Let $f : M \rightarrow N$ be a map of manifolds, and let $(x, U), (y, V)$ be charts on M, N , respectively. One can define the *Jacobian matrix of f at p* (if $p \in U$):

$$J_f(p) := \left[\left. \frac{\partial(y^i \circ f)}{\partial x^j} \right|_p \right].$$

Then we define the *rank* of $f : M \rightarrow N$ to be the rank of the Jacobian matrix.

Proposition 2.7. *The rank is well-defined.*

Proof. Suppose (x', U') is another chart containing p , and (y', V') is another chart containing $f(p)$. Then by restricting both x and x' to $U \cap U'$, we can assume that $U = U'$ and $V = V'$. Then the rank of f at p is both the rank of the Jacobian

$$\left[\left. \frac{\partial(y^i \circ f)}{\partial x^j} \right|_p \right]$$

and the Jacobian

$$\left[\left. \frac{\partial((y')^i \circ f)}{\partial (x')^j} \right|_p \right].$$

Consider the diagram:

$$\begin{array}{ccccc}
 & & M & \xrightarrow{f} & N \\
 & & \uparrow & & \uparrow \\
 & & U & \xrightarrow{f|_U} & V \\
 & \swarrow x' & \downarrow x & & \downarrow y & \searrow y' \\
 \mathbb{R}^n & \xrightarrow{x \circ x'^{-1}} & \mathbb{R}^n & \xrightarrow{y \circ f \circ x^{-1}} & \mathbb{R}^n & \xleftarrow{y \circ y'^{-1}} & \mathbb{R}^n
 \end{array}$$

Then xx'^{-1} , yfx^{-1} , and yy'^{-1} are maps from (open subsets of) \mathbb{R}^n to \mathbb{R}^n . Consider the composition $(y'y^{-1}) \circ (yfx^{-1}) \circ (xx'^{-1}) = y'fx'^{-1}$. Applying the chain rule, we get that the second Jacobian matrix is just a conjugate of the first Jacobian by invertible matrices. These have the same rank. \square

We say that a point $p \in M$ is *critical for f* if $\text{rank } f < m$ (the dimension of N). Otherwise it is *regular*. We say that a point $q \in N$ is regular if all the points in the preimage $f^{-1}(q)$ are regular points.

Example 2.8. Let $M = N = \mathbb{R}$ and let $f : M \rightarrow N$. Then a point $p \in \mathbb{R}$ is regular if and only if $f'(p) \neq 0$. In other words, critical points correspond to either maxima, minima or plateaus of the graph of f inside \mathbb{R}^2 .

One can also think of critical points as points where the inverse function theorem fails (which is the topic of the next lecture). See Figure 4. \star

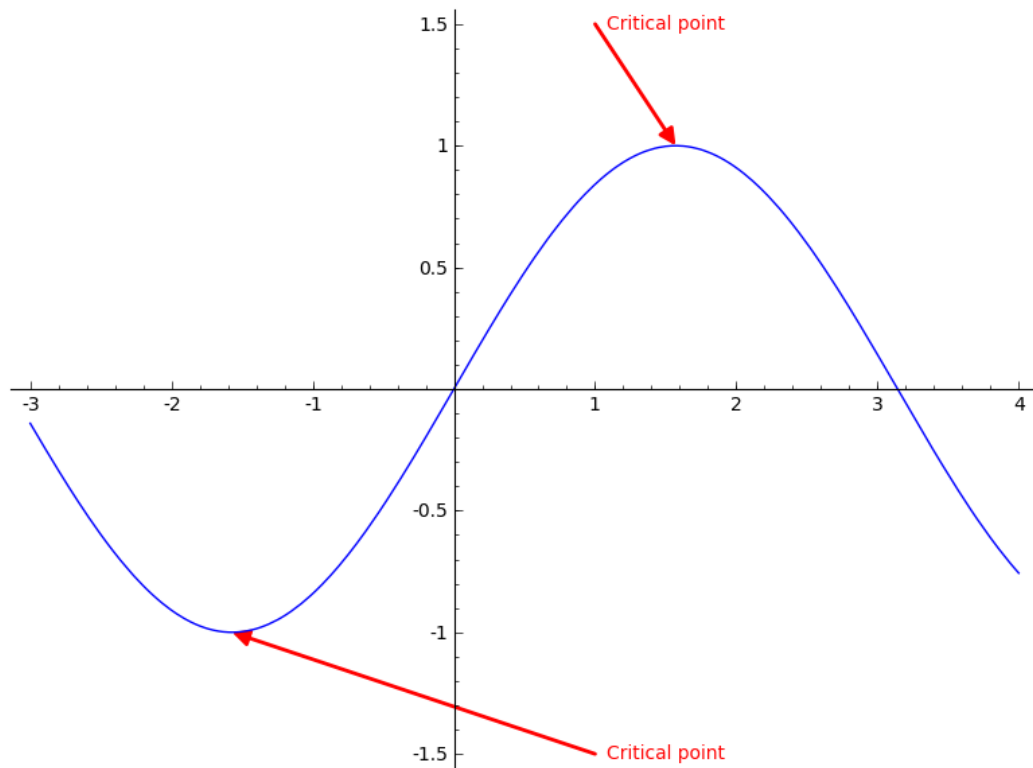


Figure 4: Critical points of the sine function.

3 Sard's theorem and the Inverse Function Theorem

Last time we spoke about critical points, i.e. those points where the Jacobian of a map had non-maximal rank (the maximal rank of a map $\mathbb{R}^n \rightarrow \mathbb{R}^m$ is n). Sard's theorem says that most points $p \in M$ are non-critical (=regular).

Think of the case $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. The graph of such a function is a surface in \mathbb{R}^3 . In this case, the theorem says that the maxima of this function constitute no "volume".

To be able to state the theorem, we have to present some definitions, the first one being, that of "measure 0". We start with a subset $A \subset \mathbb{R}^n$. We say that it has *measure 0* if it can be covered by a sequence of (open) rectangles

R_{ab} in such a way that

$$\sum_{i=1}^n \text{vol}(R_{ab}) < \epsilon$$

for every $\epsilon > 0$. Here's an example:

Example 3.1. Consider the real line $\{y = 0\} \subseteq \mathbb{R}^2$. Intuitively, this has zero volume, *as a subset of \mathbb{R}^2* . Let $B_{an, \frac{1}{n^3}}$ be the rectangle centered at origo with width n and height $\frac{1}{n^3}$. Since the height is finite, clearly the union of all these rectangles (for $n \in \mathbb{N}$) cover the real line. The sum of the volumes is

$$\sum_{i=1}^{\infty} \frac{a}{n^2} = \frac{a\pi^2}{6},$$

so choosing $a < \frac{6\epsilon}{\pi^2}$, gives total volume less than ϵ . ★

Now, a subset A of a manifold M has *measure 0* if there exists a countable sequence of charts (x_i, U_i) with $A \subseteq \bigcup U_i$ such that each $x_i(A \cap U_i)$ has measure 0. To see that this notion is well-defined, one turns to Lemma 6 in Spivak, which says that smooth functions take measure 0 sets to measure 0 sets.

Now, Sard's theorem says the following:

Theorem 3.2. *If $f : M \rightarrow N$ is a C^∞ -map of n -manifolds, and M has at most countably many components (e.g. one), then the critical values of f form a set of measure zero in N .*

This, in some sense, is similar to a theorem in algebraic geometry, which says that the smooth points of a variety are dense in the Zariski topology.

We will not prove Sard's theorem here, but just note that it is at least intuitively plausible.

3.1 The Inverse Function Theorem

We are used to thinking about the derivative of a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ as a linear approximation to f near a point p , i.e. a linear map $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Linear maps live in *vector spaces* and not on manifolds, so we should really think of the derivative as some kind of *functor* taking maps $f : U \rightarrow V$ (where U, V are open subsets of \mathbb{R}^n and \mathbb{R}^m , respectively) to *vector space maps* $Df(p) : T(\mathbb{R}^n) \rightarrow T(\mathbb{R}^m)$, where the T means that we are thinking of just the *vector space* \mathbb{R}^n .

For a map $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and a point $p \in \mathbb{R}^n$, we define the vector space map $Df(p)$ as follows:

Definition 3.3. If it exists, the map $Df(p) : T_p(\mathbb{R}^n) \rightarrow T_{f(p)}(\mathbb{R}^m)$ is the unique linear map satisfying

$$\lim_{h \rightarrow 0} \frac{|f(a+h) - f(a) - D_p f(h)|}{|h|} = 0$$

If we choose the standard coordinate systems on \mathbb{R}^n and \mathbb{R}^m , then $D_p f(p)$ has a matrix $[D_j f^i(a)]$ of partial derivatives (the Jacobian). ■

In particular, the chain rule takes a very nice form with this notation. It just says that given two maps, the derivative of the composition is just the composition of the derivatives:

$$\begin{array}{ccccc} \mathbb{R}^n & \xrightarrow{f} & \mathbb{R}^m & \xrightarrow{g} & \mathbb{R}^p \\ & & \searrow_{g \circ f} & & \end{array}$$

This gives that

$$\begin{array}{ccccc} T_p(\mathbb{R}^n) & \xrightarrow{D_p f} & \mathbb{R}^m & \xrightarrow{D_{f(p)} g} & \mathbb{R}^p \\ & & \searrow_{D_p(g \circ f)} & & \end{array}$$

is commutative.

Now we state the inverse function theorem:

Theorem 3.4. Suppose $p \in U \subseteq \mathbb{R}^n$ and that $f : U \rightarrow \mathbb{R}^m$ is a smooth function. Suppose further that the Jacobian is non-singular at p , i.e. that p is a regular value of f (or equivalently that $Df(p)$ is an isomorphism of vector spaces).

Then there exists a neighbourhood V of p such that $f|_V$ is invertible and such that $f|_V^{-1}$ is smooth. Moreover, the following equality holds:

$$J_{f|_V}^{-1}(f(p)) = J_f(p)^{-1}.$$

In words (that will have meaning later), this says that if the tangent map is an isomorphism at a point p , then f is locally a diffeomorphism near p . This is a theorem special to differential geometry. Similar situations does not occur in algebraic or complex geometry, because inverse functions are usually not algebraic/complex.

4 Rank theorems

In this lecture we will elaborate on some “structure theorems” on maps $f: M \rightarrow N$, given restrictions on the rank of f , the first one being the following:

Theorem 4.1. *Let $f: M^n \rightarrow N^m$ be a map of manifolds.*

1. *If $f: M^n \rightarrow N^m$ has rank k at p , then there charts (x, U) and (y, V) containing p and $f(p)$, respectively, such that*

$$y \circ f \circ x^{-1}(a) = (a^1, \dots, a^k, \psi^{k+1}(a), \dots, \psi^m(a))$$

for all $a \in U$.

2. *If furthermore, $f: M^n \rightarrow N^m$ has rank k in a neighbourhood of p , then there are charts (x, U) and (y, V) containing p and $f(p)$, respectively, such that*

$$y \circ f \circ x^{-1}(a) = (a^1, \dots, a^k, 0, \dots, 0)$$

for all $a \in U$.

Proof. 1. Start by choosing any coordinate system (u, U') around p . Since the rank of the Jacobian $\left[\frac{\partial(y^\alpha \circ f)}{\partial u^\beta} \Big|_p \right]$ is k at p , there is some $k \times k$ -minor that is non-zero. So, after permuting coordinates, we can assume that this minor is the upper-left block of the Jacobian.

We define new local coordinates as follows:

$$\begin{aligned} x^\alpha &= y^\alpha \circ f & \text{for } \alpha = 1, \dots, k \\ x^r &= u^r & \text{for } r = k+1, \dots, n. \end{aligned}$$

Now consider the change of basis matrix:

$$\left[\frac{\partial x^i}{\partial u^j} \Big|_p \right] = \begin{bmatrix} \frac{\partial(y^\alpha \circ f)}{\partial u^\beta} & \mathbf{X} \\ 0 & 1 \end{bmatrix}$$

Since the upper corner has determinant non-zero, the whole has determinant non-zero. Hence, by the inverse function theorem, (x, U) is a diffeomorphism in a neighbourhood U of p . Thus $x = (x \circ u^{-1}) \circ u$

is a coordinate system near p , and in fact it is the coordinate system we want:

$$\begin{aligned} y \circ f \circ x^{-1}(a^1, \dots, a^n) &= y \circ f(f^{\alpha, -1}(y^{\alpha, -1}(a^1), \dots, u^{r-1}(a^n)) \\ &= (a^1, \dots, a^k, ?, \dots, ?). \end{aligned}$$

Where the questions marke denote $u^{-1}(a^r)$, which we don't care about. What is important, is that the coordinates have the desired form.

2. Start by choosing coordinate systems (x, U) and (v, V') as in 1). Since the rank of f is k in a neighbourhood, the all of the components of the lower right rectangle of the matrix

$$\left[\frac{\partial(v^i \circ f)}{\partial x^j} \right] = \begin{bmatrix} 1 & & & & \\ & 1 & & & 0 \\ & & 1 & & \\ & X & D_{k+1}\psi^{k+1} & \dots & D_n\psi^{k+1} \\ & & \vdots & & \vdots \\ & & D_{k+1}\psi^m & \dots & D_n\psi^m \end{bmatrix}$$

must vanish in a neighbourhood of p . In particular, this means that ψ is a function only of the first k coordinates of a , i.e. that it is constant along the last $n - k$ coordinates.

Now define new local coordinates on N by letting

$$\begin{aligned} y^\alpha &= v^\alpha & \text{for } r = 1, \dots, k \\ y^r &= v^r - \psi^r \circ (v^1, \dots, v^k) & \text{for } r = k + 1, \dots, m. \end{aligned}$$

Notice that the last line makes sense because ψ only depends on the first k coordinates. It is easy to see that the change of base matrix has non-zero Jacobian at $v(q)$, so (y, V) is actually a coordinate system, where V is a neighbourhood of $f(p)$.

Moreover:

$$\begin{aligned} y \circ f \circ x^{-1}(a^1, \dots, a^n) &= y \circ v^{-1} \circ v \circ f \circ x^{-1}(a^1, \dots, a^n) \\ &= y \circ v^{-1}(a^1, \dots, a^k, \psi^{k+1}(a), \dots, \psi^m(a)) \\ &= (a^1, \dots, a^k, 0, \dots, 0) \end{aligned}$$

The first equality is by part 1), and the second follows by definition of y .

□

Thus if the rank is constant in a neighbourhood, the theorem says that the map locally looks like an inclusion of the first k coordinates in \mathbb{R}^m .

If the rank is maximal, the theorem says even more:

Theorem 4.2. *1. If $m \leq n$ and $f : M^n \rightarrow N^m$ has rank m at p , then for any coordinate system (y, V) around $f(p)$, there is some coordinate system (x, U) around p with*

$$y \circ f \circ x^{-1}(a^1, \dots, a^n) = (a^1, \dots, a^m).$$

2. If $n \leq m$ and $f : M^n \rightarrow N^m$ has rank n at p , then for any coordinate system (y, V) around $f(p)$, there is a coordinate system (x, U) around p with

$$y \circ f \circ x^{-1}(a^1, \dots, a^n) = (a^1, \dots, a^n, 0, \dots, 0).$$

Proof. Part 1 is just the previous theorem. So we concentrate on part 2. [[comes later]]

□

5 Embeddings

What does it mean to put one manifold into another manifold? It turns out that this isn't completely trivial to define.

First off, we want the embedding to respect the smooth structure in some sense. This is achieved by requiring that the map has full rank everywhere:

Definition 5.1. A map $f : M^n \rightarrow N^m$ for $m \geq n$ is an *immersion* if $\text{rank } f = n$ everywhere. ■

Example 5.2. Consider the zero set of the equation $y^2 = x^3$, or equivalently, the image of the map $f(t) = (t^2, t^3)$. It is injective everywhere, but it is not an immersion, because the rank is zero at the origin (because the derivative is $(2t, 3t^2)$ which is zero for $t = 0$). See Figure 5. ★

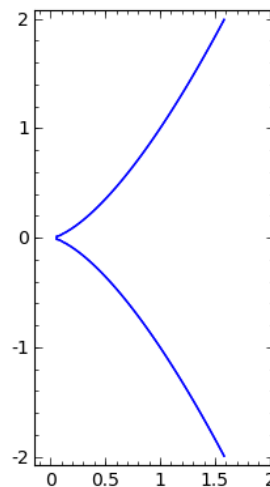


Figure 5: A cusp.

Example 5.3. Let $\alpha : \mathbb{R} \rightarrow \mathbb{R}^2$ be parametrized by $(t^2 - 1, (t^2 - 1)t)$. It is a nodal curve, and its derivative is $\alpha'(t) = (2t, 3t^2 - 1)$, which is nowhere zero. But it has self-intersections, and we do not want to allow that. See Figure 6. ★

Example 5.4. Another example is given by embedding a curve into the torus with irrational slope: $\alpha(t) = (e^{i\sqrt{2}t}, e^{i\sqrt{2}t})$. ★

This is an immersion, but the image is *dense* in $S^1 \times S^1$! ★

These problems “usually” vanish with the following definition:

Definition 5.5. An *embedding* is a map that is both an immersion and a homeomorphism onto its image. ■

We call the image of an embedding a *submanifold*.

Remark. *This is a sensible definition: it uses both the topological and the differentiable structure of the manifolds, whereas the first definition used only the differentiable structure. A similar situation occurs in algebraic geometry, where a closed immersion of schemes is defined not just as a closed inclusion*

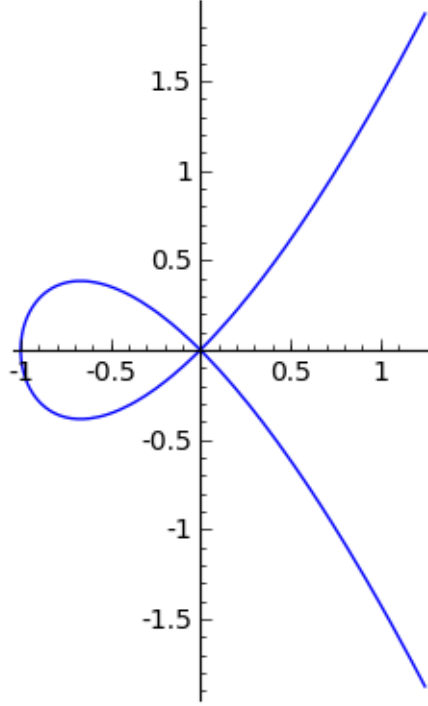


Figure 6: A nodal curve.

of the underlying topological spaces, but as a pair $(f, f^\#)$, where the first is a map of topological spaces, and the second is surjective map of sheaves $f^\# : f^* \mathcal{O}_Y \rightarrow \mathcal{O}_X$.

Here's an important result:

Proposition 5.6. *If $f : M \rightarrow N$ has constant rank k in a neighbourhood of $f^{-1}(y)$, then $f^{-1}(y)$ is a closed submanifold of M .*

Proof. To give $f^{-1}(y)$ the structure of a manifold, it is enough to give charts and glueing maps. Since f has rank k everywhere in $f^{-1}(y)$, there exists charts $(x, U) \subset M^n$ and $(y, V) \subset N^m$ such that $y \circ f \circ x^{-1} = (x_1, \dots, x_n) \mapsto (x_1, \dots, x_k, 0, \dots, 0)$.

We may assume that $y(y) = 0$, so that $x(U \cap f^{-1}(y)) = \{x_1 = x_2 = \dots = x_k = 0\}$. This is an $n - k$ -dimensional closed subspace of U . \square

Example 5.7. Consider Figure 7. If we define a map $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ by $(x, y) \mapsto x^2 - y^2$, we get a function whose graph has a saddle point. As long

as we're looking at $F^{-1}(a)$ for a non-zero, the inverse image is a hyperbola, which is a smooth (disconnected) manifold. However, when $a = 0$, the fiber (i.e. the contour line) is a pair of double lines, which is not even a manifold. The reason is that the rank of F is zero at $(0,0)$. ★

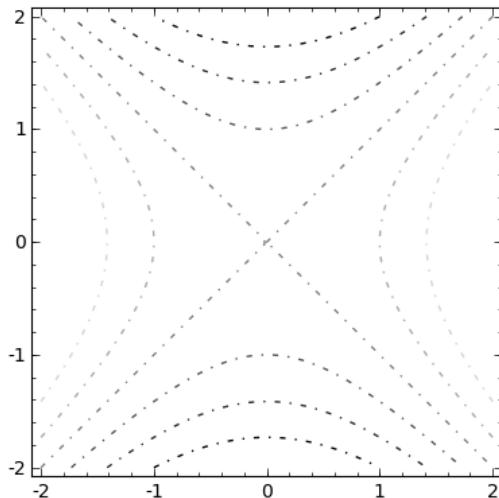


Figure 7: The countour lines of the function $F(x, y) = x^2 - y^2$.

Example 5.8. Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be the map $f(x, y) = x^2 + y^2 - 1$. It has constant rank 1 as long as we're away from the origin. Then the theorem says that $f^{-1}(0) = \{x^2 + y^2 = 1\} \approx S^1$ is a closed submanifold av \mathbb{R}^2 . ★

6 Bump functions and partitions of unity

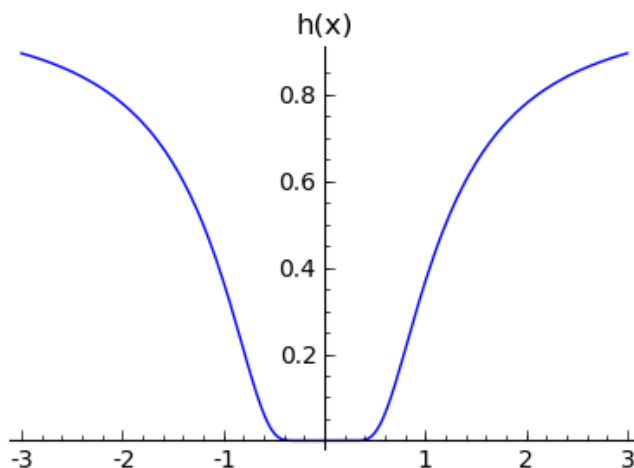
What makes differential geometry “taste” differently than complex or algebraic geometry is the existence of “bump functions”.

6.1 Useful functions

Recall that the *support* of a function $f : X \rightarrow \mathbb{R}$ is the closure of the set $\{x \mid f(x) \neq 0\}$.

Example 6.1. The first function is an example of a non-zero function whose Taylor series around any point is zero (thus it is a non-analytical function).

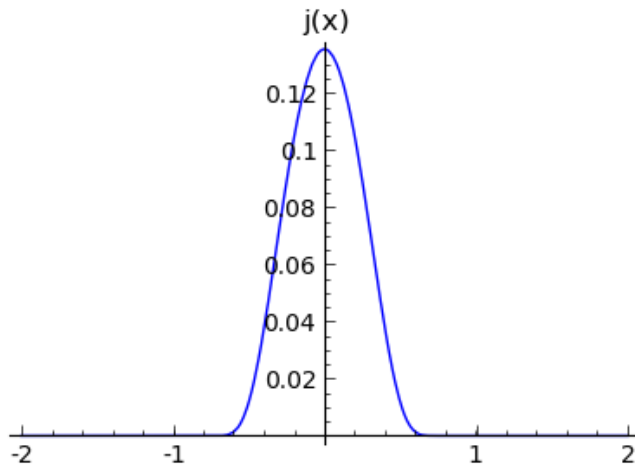
$$h(x) = \begin{cases} e^{-\frac{1}{x^2}} & x \neq 0 \\ 0 & x = 0. \end{cases}$$



★

Example 6.2. This is an example of a function whose support is $[-1, 1]$, but is zero everywhere else.

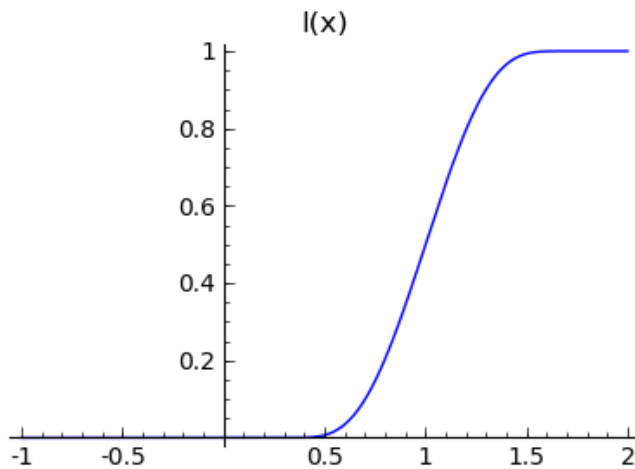
$$j(x) = \begin{cases} e^{-(x-1)^{-2}} \cdot e^{-(x+1)^{-2}} & x \in (-1, 1) \\ 0 & x \notin (-1, 1) \end{cases}$$



Note that this function is, roughly, the same as $h(x-1) \cdot h(x+1)$ inside $(-1, 1)$. By composing with a linear change of coordinates, we get a function which is positive on $(0, \delta)$ and 0 elsewhere. ★

Example 6.3. There is a function $l : \mathbb{R} \rightarrow \mathbb{R}$ which is zero for $x \leq 0$, strictly increasing on $(0, \delta)$, and equal to 1 for $x \geq \delta$:

$$l(x) = \frac{\int_0^x k(x) dx}{\int_0^\delta k(x) dx}.$$



★

Example 6.4. By mirroring the function $l(x)$ from the previous example around $x = \delta + 1$, we get a function $f(x)$ which is constantly equal to 1 on $(\delta, \delta + 1)$, and has support $[0, 2\delta + 1]$. By affine transformations, the function can be made to be 1 on any bounded interval K and support on any interval containing K . ★

Example 6.5. There is a function $g : \mathbb{R}^n \rightarrow \mathbb{R}$ which is positive on the open square $(-\epsilon, \epsilon) \times \cdots \times (-\epsilon, \epsilon)$ and zero elsewhere:

$$g(x) = \prod_{i=1}^n j(x^i/\epsilon).$$

★

Example 6.6. Generalizing Example 6.4, by defining $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ as

$$\phi(x) = \prod_{i=1}^n f(x^i),$$

we get a function which is constantly equal to one on a closed, bounded square, and has support in a slightly larger square. ★

The last example can be generalized to general manifolds:

Proposition 6.7. *Let M be a smooth manifold and K a compact subset of M and U an open subset containing K . Then there exists a smooth function $\beta : M \rightarrow [0, 1]$ that is constant equal to 1 on K and compact support contained in U .*

Proof. We first do the case $M = \mathbb{R}^n$. In this case, K is compact, so it is closed and bounded. For each $p \in K$, let U_p be an open square of radius ϵ_p centered at p and contained in U . The set of all these U_p is an open cover of K , and since K is compact, we can choose finitely many such p . By translation, the function in the last example, can be made such that f_p is positive in the interior of U_p and zero outside (so its support is \bar{U}_p), and constant 1 on $K \cap U_p$. Now define the following function:

$$\beta(x) = 1 - \prod_p (1 - f_p(x)),$$

where the product ranges over those finitely many p needed to cover K . Then, if $x \in K$, x is contained in one of the $U_p \cap K$, hence $\beta(x) = 1$. The support is clearly bounded, hence compact.

Now to the general case. If K is contained in a single chart, we are done by the above. If not, K is contained in finitely many charts (U_i, x_i) , hence we can find compact sets K_1, \dots, K_k with $K \subset \bigcup_{i=1}^k K_i$, $K_i \subset U_i$ and $\bigcup_i U_i \subset U$. Let ϕ_i be identically 1 on K_i and zero on $M \setminus U_i$. Then define

$$\beta(x) = 1 - \prod_{i=1}^k (1 - \phi_i(x)).$$

□

6.2 Partition of unity

Partitions of unity is an extremely important tool, but to define it, we need some technical definitions. To motivate all this, we will note that when we are done, we will be able to prove that any compact manifold can be embedded into real Euclidean space \mathbb{R}^N for some large N .

Definition 6.8. We say that a family \mathcal{U} of open sets is an *open cover* of M if

$$\bigcup_{U \in \mathcal{U}} U = M.$$

■

Definition 6.9. We say that \mathcal{U}' is a *refinement* of \mathcal{U} if for all $U \in \mathcal{U}'$, there exists some $V \in \mathcal{U}$ with $U \subseteq V$. ■

Example 6.10. Cover \mathbb{R} with the single open set \mathbb{R} , so that $\mathcal{U} = \{\mathbb{R}\}$. Now consider the cover given by $\mathcal{U}' = \{(-\infty, 1), (-1, \infty)\}$. Then \mathcal{U}' is a refinement of \mathcal{U} .

Let $\mathcal{U}'' = \{(-2, 2), (0, \infty), (-\infty, 0)\}$. Then \mathcal{U}'' is *not* a refinement of \mathcal{U}' . ★

Definition 6.11. We say that an open cover \mathcal{U} is *locally finite* if for every $p \in M$, there are only finitely open sets U in \mathcal{U} with $\{p\} \cap U \neq \emptyset$. ■

Theorem 6.12. If \mathcal{U} is an open cover of a connected manifold M , then there exists a locally finite refinement \mathcal{U}' of \mathcal{U} .

Moreover, we can choose these \mathcal{U}' in such a way that $\mathcal{U}' \approx \mathbb{R}^n$ as differentiable manifolds.

Remark. We skip the proof. Now it's time to note that we are hiding details under carpets: for this theorem to be true, we must assume that M is σ -compact, meaning that we can write M as a countable union of compact

subsets. In particular, this is clearly true if M itself is compact. Note also that all subsets of \mathbb{R}^n have this property.

Theorem 6.13 (Shrinking Lemma). *Let \mathcal{U} be an open cover of M . Then it is possible to choose for each $U \in \mathcal{U}$ an open set U' with $\bar{U}' \subset U$ in such a way that the new collection $\{U'\}$ is also an open cover M .*

Remark. *In particular, this new open cover is a refinement of \mathcal{U} .*

Theorem 6.14 (Existence of partitions of unity). *Let \mathcal{U} be an open locally finite cover of a manifold M . Then there is a collection of C^∞ -functions $\varphi_U : M \rightarrow [0, 1]$ for each $U \in \mathcal{U}$ such that*

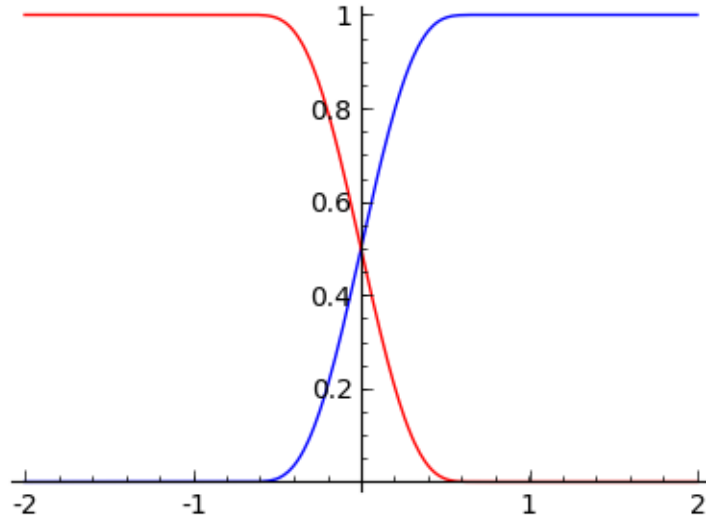
1. $\text{supp } \varphi_U \subset U$ for each U .
2. $\sum_{U \in \mathcal{U}} \varphi_U(p) = 1$ for all $p \in M$. (this makes sense because the cover is locally finite!)

Definition 6.15. A collection $\{\varphi_i : M \rightarrow [0, 1]\}$ is called a *partition of unity* if (1) the collection $\{p \mid \varphi_i(p) \neq 0\}$ is locally finite and (2) if $\sum_i \varphi_i(p) = 1$ for all $p \in M$.

If for each i there is an $U \in \mathcal{U}$ such that $\text{supp } \varphi_i \subset U$, then we say that the collection is *subordinate to \mathcal{U}* . ■

So the theorem says that given a locally finite cover \mathcal{U} of a manifold M , there exists a partition of unity subordinate to \mathcal{U} .

Example 6.16. Let $M = \mathbb{R}$ and cover M by $U_1 = (-1, \infty)$ and $U_2 = (-\infty, 1)$. Let $\varphi_1(x) = \frac{l(x+1)}{l(x+1)+l(-x-1)}$. And similarly, $\varphi_2(x) = \frac{l(-x+1)}{l(x+1)+l(-x+1)}$.



Then the partition of unity looks like the graph above. Notice that it sums to 1 everywhere. ★

Theorem 6.17. *Let M^n be a compact manifold. Then there exists an embedding $f : M \rightarrow \mathbb{R}^N$ for some N .*

Proof. By compactness we can choose a finite number of coordinate systems $\{(x_i, U_i)\}_{i=1, \dots, k}$, covering M . By the Shrinking Lemma 6.13, choose $U'_i \subseteq U_i$, and by Partition of Unity, choose functions $\psi_i : M \rightarrow [0, 1]$ that are 1 on \bar{U}'_i and have support contained in U_i . This can be done by Proposition 6.7.

Define $f : M \rightarrow \mathbb{R}^N$, where $N = k(n + 1)$, by

$$f = (\psi_1 \cdot x_1, \dots, \psi_k \cdot x_k, \psi_1, \dots, \psi_k).$$

This is an immersion: any point $p \in M$ is contained in some U'_i , and on U'_i , where $\psi_i = 1$, the $N \times n$ Jacobian matrix contains the identity matrix $\begin{pmatrix} \frac{\partial x_i^\alpha}{\partial x_i^\beta} \end{pmatrix}$.

The map is also one-one. For suppose that $f(p) = f(q)$. There is some i with $p \in U'_i$. But then $q \in U_i$ since the support of ψ_i is contained in that set. Moreover, $\psi_i \cdot x_i(p) = \psi_i x_i(q)$, so $p = q$, since x_i is 1-1. □

Remark. *Note that this embedding is highly non-canonical. It contains several layers of choices. The proof begins by choosing a finite open cover, and then choosing a refinement, and then choosing a partition of unity. None of these choices are natural in any sense.*

7 The tangent bundle

Think of a surface $S \subseteq \mathbb{R}^3$ and let p be a point on S . Then the tangent plane of S is a linear subspace of \mathbb{R}^3 . Thus, in the case of embedded manifold, it is easy to assign a “tangent space” to each point $p \in S$. However, for general manifolds, we are not given an embedding anywhere. What we want is some rule that assigns to each point $p \in S$ a vector space of the same dimension as S , called the “tangent space”. Also, we want this rule to reflect the global structure of S in a natural way.

More formally, we want a functor from manifolds to vector bundles, such that its restriction to chart domains, the result is just \mathbb{R}^n , as a vector space. To carry this out, we first formally introduce the “tangent space” of \mathbb{R}^n (or open subsets thereof).

We define $T(\mathbb{R}^n) := \mathbb{R}^n \times \mathbb{R}^n$, and we write its elements as (p, v) or v_p , where we think of the left factor as the *manifold* \mathbb{R}^n , and the right factor as the *vector space* \mathbb{R}^n . Thus, to each *point* p in \mathbb{R}^n , we attach a *vector* v_p .

Now we have defined the functor T on objects. Now we define it on morphisms, so let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a smooth function. Then we define $Tf : T(\mathbb{R}^n) \rightarrow T(\mathbb{R}^m)$ by $Tf(p, v) = (f(p), Df(p)(v))$. By the chain rule, this is really a functor, because if $g : \mathbb{R}^m \rightarrow \mathbb{R}^k$ is another smooth function, we have

$$\begin{aligned} T(g) \circ Tf(p, v) &= T(g)(f(p), Df(p)(v)) \\ &= (g(f(p)), Dg(f(p))(Df(p)(v))) \\ &= (g \circ f(p), D(g \circ f)(v)) \\ &= T(g \circ f). \end{aligned}$$

To generalize this construction to general manifolds, we need to notion of a *vector bundle*:

Definition 7.1. A *vector bundle* over M is a map $\pi : E \rightarrow M$, where E is a manifold, such that for each $p \in M$, there is an open neighbourhood U , and a homeomorphism $e_q : \pi^{-1}(U) \rightarrow U \times \mathbb{R}^n$ in such a way that

$$t|_{\pi^{-1}(\{q\})} : \pi^{-1}(\{q\}) \rightarrow \{q\} \times \mathbb{R}^n$$

is a vector space isomorphism for all $q \in U$.

A map of vector bundles is just a commutative diagram

$$\begin{array}{ccc} E_1 & \xrightarrow{\tilde{f}} & E_2 \\ \pi_1 \downarrow & & \downarrow \pi_2 \\ B_1 & \xrightarrow{f} & B_2 \end{array}$$

such that, when restricted to each fiber, the map $\tilde{f} : \pi_1^{-1}(p) \rightarrow \pi_2^{-1}(f(p))$ is a linear map. ■

Thus we have a category of vector bundles, which we shall denote by **VBundles**. We have a natural map $p : \mathbf{VBundles} \rightarrow \mathbf{Diff}$ that sends a vector bundle to its base. The fiber $p^{-1}(M)$ is a category, called the category of *vector bundles over M* , which we shall often denote by $\mathbf{VBundles}(M)$.

We call a vector bundle E *trivial* if it is isomorphic to $M \times \mathbb{R}^n$ (with the obvious maps).

This really generalizes the $T(\mathbb{R}^n)$ above: it is trivially a vector bundle over \mathbb{R}^n , and it is easy to see that it restricts to a sub-vector bundle for any open subset $U \subset \mathbb{R}^n$.

Requiring only that $T(U) = U \times \mathbb{R}^n$ for U an open subset of \mathbb{R}^n and functoriality, there is a unique extension of T from open subsets of \mathbb{R}^n to general manifolds.

Theorem 7.2. *Let M be a smooth manifold. Then there is a functor $T : \mathbf{Diff} \rightarrow \mathbf{VBundles}$ that to each manifold M associates a vector bundle over M in such a way that if (U, x) is a chart domain of M , we have $T(U) = U \times \mathbb{R}^n$.*

Proof. Here is a proof sketch. Cover M by open chart domains (x_i, U_i) . Each of these is isomorphic to \mathbb{R}^n , so we define $TM|_{U_i} := x_i(U_i) \times \mathbb{R}^n$, with the obvious projection map. On the intersections $U_i \cap U_j$ we have a map $x_j \circ x_i^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, and we define

$$TM|_{U_i \cap U_j} \rightarrow TM|_{U_i \cap U_j}$$

by

$$(u, v) \mapsto (x_j \circ x_i^{-1}(u), D(x_j \circ x_i^{-1})(u)(v)).$$

That this is well-defined on triple overlaps should follow from the chain rule, and all these glueing maps make TM into a manifold. Local triviality is clear by construction. □

We call the vector bundle $\pi : TM \rightarrow M$ the *tangent bundle* of M .

Example 7.3. Consider the circle S^1 . It is covered by two open sets $S^1 \setminus N$ and $S^1 \setminus S$. The map $\phi_1 : S^1 \setminus N$ is given by $(x, y) \mapsto \frac{x}{1-y}$, and the inverse map is given by $\phi_1^{-1}(x) = \left(\frac{2x}{x^2+1}, \frac{x^2-1}{x^2+1} \right)$. Similarly, $\phi_2(x, y) = \frac{x}{1+y}$, and the inverse is $\phi_2^{-1}(x) = \left(\frac{2x}{x^2+1}, -\frac{x^2-1}{x^2+1} \right)$. Thus $T(S^1)$ is covered by two open sets $U = T(R^1)$ and $V = T(R^1)$. The transition map ϕ_{UV} is given by $(x, v) \mapsto (\phi_2 \circ \phi_1^{-1}(x), D(\phi_2 \circ \phi_1^{-1})(x)(v))$, which we compute to be

$$(x, v) \mapsto \left(\frac{1}{x}, -\frac{v}{x^2} \right).$$

This is well-defined, since $x \in \mathbb{R} \setminus \{0\}$.

Then I claim that $TS^1 \approx S^1 \times \mathbb{R}^1$, that is, the tangent bundle of the circle is trivial. To do this, we cover the circle with the same open sets as above, so that S^1 is covered by the two open sets $U' \times \mathbb{R}$ and $V' \times \mathbb{R}$ with transition maps $\phi_{U'V'} : (x, v) \mapsto (\frac{1}{x}, v)$. This is truly a product, because the second factor does not interact with the first.

An isomorphism $TS^1 \approx S^1 \times \mathbb{R}^1$ is the same thing as maps from the charts from one manifold to the other that agrees on the overlaps. So we define, from $U \times \mathbb{R}$ to $U' \times \mathbb{R}$ a map given by $(x, v) \mapsto (x, v)$, and from $V \times \mathbb{R}$ to $V' \times \mathbb{R}$, a map given by $(x, v) \mapsto (x, -\frac{v}{x^2})$. Then the reader can check that the following diagram commutes, and so we have defined a map on manifolds, that is an isomorphism on charts and agrees on overlaps, hence an isomorphism of manifolds:

$$\begin{array}{ccc} U \times \mathbb{R} & \xrightarrow{\phi_{UV}} & V \times \mathbb{R} \\ (x,v) \mapsto (x,v) \downarrow & & \downarrow (x,v) \mapsto (x, -\frac{v}{x^2}) \\ U' \times \mathbb{R} & \xrightarrow{\phi_{U'V'}} & V' \times \mathbb{R} \end{array}$$

★

Example 7.4 (The tangent bundle of S^2 is not trivial). It is not hard to see that a n -bundle $E \rightarrow M$ is trivial if and only if it admits n everywhere linearly independent sections.

Then the non-triviality of TS^2 follows from the the Hairy Ball Theorem, which says that every vector field on the sphere must vanish at least somewhere. This in turn follows from the Poincaré-Hopf theorem, which says that

$$\sum_i \text{index}_{x_i}(X) = \chi,$$

where the left-hand-side is a certain sum ranging over the zeros of the vector field X , and the right-hand-side is the Euler characteristic of the sphere, which is two - so X must have at least one zero. ★

The construction of the tangent space here is the most intuitive one: locally it looks like $U \times \mathbb{R}^n$. However, there are other views one can take. One is to consider points of TM as equivalence classes of curves, as follows: let $[x, v]_p \in T_p M$. Then we say that this corresponds to a curve $x^{-1} \circ \gamma$ where γ is a curve in \mathbb{R}^n with $\gamma'(0) = v$ (here x is a coordinate chart). It is not so difficult to see for every vector v in $T_p M$, there is a curve with derivative v . Taking equivalence classes resolves uniqueness.

The other view is to consider elements of $T_p M$ to be *derivations*. To define these, let first \mathcal{O}_p consist of all functions $f : U \rightarrow \mathbb{R}$ where U is a neighbourhood of p , where we consider $f : U \rightarrow \mathbb{R}$ and $g : V \rightarrow \mathbb{R}$ to be equal if there is a neighbourhood $W \subset U \cap V$ such that $f|_W = g|_W$. More formally, \mathcal{O}_p is really the direct limit $\varinjlim \mathcal{O}(U)$, where $\mathcal{O}(U)$ is the sheaf of C^∞ functions $f : U \rightarrow \mathbb{R}$.

Then we claim that $T_p M$ can be considered as the set of *derivations* $\ell : \mathcal{O}_p \rightarrow \mathbb{R}$. Formally, a derivation at p is a function from \mathcal{O}_p to \mathbb{R} that is linear and satisfies the Leibniz rule

$$\ell(fg)(p) = f(p)\ell(g)(p) + g(p)\ell(f)(p).$$

To see that derivations actually land in \mathbb{R} , we must show the following lemma:

Lemma 7.5. *If $f : W \rightarrow \mathbb{R}$ and $g : W \rightarrow \mathbb{R}$ represent the same function in \mathcal{O}_p , and ℓ is a derivation, then $\ell(f) = \ell(g)$.*

Proof. Clearly we can assume $g = 0$. So we need to prove that if f is zero in a neighbourhood of p , then $\ell(f)$ is also zero in a neighbourhood of p .

Choose a C^∞ function $h : M \rightarrow \mathbb{R}$ with $h(p) = 1$ and support contained in $f^{-1}(\{0\})$. Such a function exists by Proposition 6.7. Let $q \in W$, where W is the smallest neighbourhood of p such that $f \equiv 0$ on W .

$$0 = \ell(0)(q) = \ell(fh)(q) = f(q)\ell(h)(q) + h(q)\ell(f)(q) = 0 + \ell(f)(q).$$

Thus $\ell(f)(q) = 0$ for all $q \in W$, hence $\ell(f) = 0$. □

Here's a lemma:

Lemma 7.6. Let $f \in \mathcal{O}(U)$ ¹ with $f(0) = 0$, where $U \subset \mathbb{R}^n$ and U is convex. Then there are C^∞ functions $g_i : U \rightarrow \mathbb{R}$ with

1. $f(x^1, \dots, x^n) = \sum_{i=1}^n x^i g_i(x^1, \dots, x^n)$ for all $x \in U$.
2. $g_i(0) = D_i(f)(0)$.

Proof. Define a new function $h_x : [0, 1] \rightarrow \mathbb{R}$ for each $x \in U$ as $h_x(t) = f(tx)$. This makes sense since U is assumed to be convex. Then

$$f(x) = f(x) - f(0) = \int_0^1 h'_x(t) dt = \int_0^1 \sum_{i=1}^n D_i f(tx) x^i dt.$$

Therefore we can let $g_i(x) = \int_0^1 D_i f(tx) dt$. □

Note that if (x, U) is a chart domain of M around p , then $\frac{\partial}{\partial x^i} \Big|_p$ is a derivation at p . In fact, the derivations $\frac{\partial}{\partial x^i} \Big|_p$ span the space of derivations at p .

Theorem 7.7. The set of derivations at $p \in M^n$ is an n -dimensional vector space. In fact, if (x, U) is a coordinate system around p , then

$$\frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p$$

span this vector space, and any derivation ℓ can be written

$$\ell = \sum_{i=1}^n \ell(x^i) \cdot \frac{\partial}{\partial x^i} \Big|_p.$$

Proof. We can clearly assume $M = \mathbb{R}^n$. Then using the Lemma, we see that any $\ell(f)$, where ℓ is a derivation, can be written as

$$\ell(f) = \sum_{i=1}^n \ell(x^i) \frac{\partial f}{\partial x^i}(0).$$

Thus they span the space of derivations, and they are clearly linearly independent. □

If (y, V) is another coordinate system, then we know that $\frac{\partial}{\partial y^i} \Big|_p = \sum_{j=1}^n \frac{\partial x^j}{\partial y^i} \Big|_p \frac{\partial}{\partial x^j} \Big|_p$, so the derivations transform exactly the same way as the charts of TM . So we identify them.

¹From now on, unless otherwise stated, the notation $\mathcal{O}(U)$ will denote the set of C^∞ -functions from U to \mathbb{R} .

7.1 Operations on vector fields

A *vector field* is a section of TM . They are often denoted by capital letters such as X, Y and Z . The vector $X(p) \in M_p$ is often denoted X_p . Thinking of TM as the set of derivations, we have

$$X(p) = \sum_{i=1}^n a^i(p) \frac{\partial}{\partial x^i} \Big|_p$$

for some continuous functions $a^i(p)$.

If X and Y are vector fields, then they can be added:

$$(X + Y)(p) = X(p) + Y(p).$$

Similarly, if $f : M \rightarrow \mathbb{R}$, we can define the vector field fX by

$$fX(p) = f(p)X(p).$$

Thus we see that the set of vector fields, $\Gamma(TM)$, is a module over the C^∞ -functions $C^\infty(M) = \Gamma(\mathcal{O}_M)$.

If $f : M \rightarrow \mathbb{R}$ is a function and X is a vector field, then we can define a new function $X(f) : M \rightarrow \mathbb{R}$ by letting X operate on f at each point:

$$X(f)(p) = X_p(f).$$

Example 7.8. Let $M = \mathbb{R}^3$ and let X be the vector field $X = x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y}$. Let f be the function $f(x, y, z) = x^2 + y^2 + z^2$. Then

$$X(f) = x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} = 2x^2 + 2y^2.$$

★

8 Tensors

8.1 Miscellany on the exterior algebra

The construction of the tangent bundle can be generalized to give other types of vector bundles: each fiber TM_p can be dualized to give T^*M_p , the dual vector space. These spaces glue to give the *cotangent bundle*, denoted by T^*M . It has a basis consisting of *differentials*.

Now each of the T^*M_p are finite-dimensional vector spaces, and as such, we can form the *exterior algebra* $\bigwedge T^*M_p$. This is the quotient of the tensor algebra on T^*M_p by the two-sided ideal generated by elements of the form $v_1 \otimes \cdots \otimes v_k$ with $v_i = v_{i+1}$ for some i . This is a graded algebra, so we write $\bigwedge T^*M_p = \bigoplus_{k=0}^n \bigwedge^k T^*M_p$. The sum is finite, since $\bigwedge^{n+1} T^*M_p = 0$, as is easily seen (here n is the dimension of M). Elements of $\bigwedge^p T^*M_p$ are called p -forms, and we will see later that they are what one integrates over.

Since we have a canonical isomorphism (for vector spaces V) $\bigwedge V^* \simeq (\bigwedge V)^*$, we think of the exterior algebra as parametrizing multilinear maps $T_pM \times \cdots \times T_pM \rightarrow \mathbb{R}$.

8.2 The cotangent bundle

Given a map $f : M \rightarrow \mathbb{R}$, we can produce a section of the cotangent bundle:

Definition 8.1. We define a section $df : M \rightarrow T^*M$ as follows: $df(p)$ should act on tangent vectors, so given a tangent vector $X_p \in T_pM$, we define

$$df(p)(X_p) := X_p(f)$$

■

To see how this looks like in a chart, see the Appendix.

The cotangent bundle have better functorial properties than the tangent bundle. In the case of the tangent bundle, there were no way to push tangent fields forward nor pull them back via a map $f : M \rightarrow N$. However, it is possible to pull back cotangent fields (this is sort of a miracle, as the cotangent bundle does not appear all that different from the tangent bundle). Given a map $f : M \rightarrow N$, and a section $\omega : N \rightarrow T^*N$, we can define a new section $f^*\omega : M \rightarrow T^*M$ as follows: we must say how it acts on vectors $X_p \in T_pM$:

$$f^*\omega(p)(X_p) := \omega(f(p))(f_*(X_p)).$$

The idea is clear: tangent vectors can be pushed forward, and this allows us to act on it by ω . To see how this works in coordinates, see the Appendix.

Wikipedia has a good article on pullbacks: [http://en.wikipedia.org/wiki/Pullback_\(differential_geometry\)](http://en.wikipedia.org/wiki/Pullback_(differential_geometry)).

Example 8.2. Let $M = \mathbb{R}^3$. Let ω be the section of the cotangent bundle defined by $zdx - dz$. Let $X = y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}$. Then

$$\omega(X) = z.$$

Here we use only that dx^i and $\frac{\partial}{\partial x^i}$ are dual bases. ★

Example 8.3. Let $c : S^1 \rightarrow \mathbb{R}^2$ be given by $\theta \mapsto (\sin \theta, \cos \theta)$ and let $\omega = ydx - xdy$ be a differential form on \mathbb{R}^2 . Then $c^*\omega$ is given by

$$c^*\omega = \cos \theta \frac{\partial c^1}{\partial \theta} d\theta - \sin \theta \frac{\partial c^2}{\partial \theta} d\theta = (\cos^2 \theta + \sin^2 \theta) d\theta = d\theta.$$

★

8.3 General tensor fields

Proposition 8.4. *The smooth bilinear pairing $\Gamma(TM) \times \Gamma(T^*M) \rightarrow C^\infty(M)$ given by $(X, \omega) \mapsto \omega(X)$ defines an isomorphism $\Gamma(T^*M) \approx \Gamma(TM)^*$.*

Thus to give a tensor field is the same as giving a multilinear map from tangent vectors. This is really a theorem, and relies heavily on the existence of bump functions.

9 Vector fields and differential equations

Let $X : M \rightarrow TM$ be a vector field. Then one can ask: is there a curve $\rho : (-\epsilon, \epsilon) \rightarrow M$ with $\rho(0) = p \in M$ such that

$$\rho'(t) = X_{\rho(t)} \quad \forall t \in (-\epsilon, \epsilon)?$$

This is a local question, so we can assume that $M = \mathbb{R}^n$. Then a vector field X on \mathbb{R}^n is just a smooth function $f : V \rightarrow \mathbb{R}^n$, where V is some open neighbourhood, which we assume contains zero. Then the above equation just reads:

$$\rho'(t) = f(\rho(t)).$$

This is just an ordinary differential equation of order one. In the beginning of this section we will study the existence and uniqueness of solutions of these in \mathbb{R}^n .

Two examples will illuminate the kind of pathologies (or problems, if you like) that can arise:

Example 9.1. Set $n = 1$ and $f(y) = -y^2$. We seek a curve $c(t)$ with $c(0) = x \in \mathbb{R}$. Here f corresponds to the vector field X that assigns to every number y a vector pointing backward of length y^2 . We get the equation

$$-\frac{1}{c^2} \frac{dc}{dt} = 1.$$

Integrating both sides:

$$\begin{aligned} \int -\frac{1}{c^2} \frac{dc}{dt} dt &= \int 1 dt \\ \int -\frac{1}{c^2} dc &= \int 1 dt \\ \frac{1}{c} &= t + C, \end{aligned}$$

yielding either $c(t) = 1/(t + C)$ as a solution, for some constant C , or $c(t) = 0$ for all t , the latter alternative occurring if $c(0) = 0$. But if $c(0) \neq 0$, the solution is not defined for all t ! It does not extend outside $(-\frac{1}{x}, \frac{1}{x})$ in both directions. This is one of the problems that can occur. ★

Example 9.2. Again $n = 1$. Now let $f(y) = y^{2/3}$. Set $c(0) = 0$. Then we get the equation

$$\frac{dc}{dt} = c^{2/3}, \quad c(0) = 0.$$

But there are two solutions! The curve is not unique! ★

There is a theorem, however:

Theorem 9.3. *Let $V \subset \mathbb{R}^n$ be open and $f : V \rightarrow \mathbb{R}^n$. Let $x_0 \in V$ and $a > 0$ be such that $B_{2a}(x_0) \subset V$. If there are K, L such that*

1. $|f(x)| \leq L$ on $\overline{B_{2a}(x_0)}$, i.e. f is L -bounded.
2. $|f(x) - f(y)| \leq K|x - y|$ on $\overline{B_{2a}(x_0)}$, i.e. f is K -Lipschitz.

then choose $b > 0$ such that

3. $b \leq 1/L$ and
4. $b \leq 1/K$.

Then for each $x \in \overline{B_a(x_0)}$, there exists a unique curve $\alpha_x : (-b, b) \rightarrow U$ such that

$$\alpha'_x(t) = f(\alpha_x(t)) \text{ and } \alpha_x(0) = x.$$

Sketch of proof. Topologize the set of maps $Y = \{\alpha : (-b, b) \rightarrow \overline{B_{2a}(x_0)}\}$ by the sup metric. Then M is a complete metric space. One defines an operator $S : Y \rightarrow Y$ by

$$(S\alpha)(t) = x + \int_0^t f(\alpha(u))du.$$

Then one sees that if S has a fixed point, then it is a solution to our differential equation. One shows that S is a contraction, and this implies that there exists a unique fixed point solution by the contraction lemma of analysis.

Then there is some checking to do, and the proof is complete. \square

Write $\alpha_x(t)$ as $\alpha(t, x)$ to get a map

$$\begin{aligned} \alpha : (-b, b) \times \overline{B_a(x_0)} &\longrightarrow V \\ (t, x) &\longmapsto \alpha(t, x) = \alpha_x(t) \end{aligned}$$

satisfying $\alpha(0, x) = x$ and $\frac{d}{dt}\alpha(t, x) = f(\alpha(t, x))$. This map α is called the **local flow** for f in $(-b, b) \times B_a(x_0)$.

Suppose $y = \alpha_x(t_0)$ for some $t_0 \in (-b, b)$ (think of this as starting at x and following the flow for a time t_0). Then the reparametrized integral curve

$$t \longmapsto \beta(t) := \alpha_x(t + t_0)$$

satisfies $\beta'(t) = f(\alpha(t_0 + t)) = f(\beta(t))$, with $\beta(0) = \alpha_x(t_0) = y$. This means that β satisfies the conditions that uniquely determine α_y , so $\beta(t) = \alpha_y(t)$ for t near 0 on

$$(-b, b) \cap (-b - t_0, b - t_0).$$

This means that $\alpha_x(t + t_0) = \alpha_y(t)$ for t near zero. Thus:

Proposition 9.4. *For each $t \in (-b, b)$ we get a map*

$$\begin{aligned} \phi_t : B_a(x_0) &\longrightarrow V \\ x &\longmapsto \alpha(t, x) \end{aligned}$$

Such that $\phi_0(x) = x$ and $\phi_{s+t}(x) = \phi_s(\phi_t(x))$ for $s, t, s + t \in (-b, b)$. In particular, $\phi_{-t} = \phi_t^{-1}$, so each ϕ_t is a bijection.

Theorem 9.5. *The flow*

$$\alpha : (-b, b) \times B_a(x_0) \longrightarrow V$$

is continuous. Hence each ϕ_t is also continuous.

Sketch of proof. Let S denote the operator used in the previous theorem. Using a geometric series trick, one proves that

$$\sup_t |\alpha(t, x) - \alpha(t, y)| = \|\alpha_x - \alpha_y\| \leq \frac{1}{1 - bK} |x - y|,$$

where α_x is a solution of the differential equation starting at x and α_y a solution starting at y . This inequality implies continuity of α . \square

Proposition 9.6 (Spivak cites Lang). *If $f : V \rightarrow \mathbb{R}^n$ is C^∞ , then the flow is also C^∞ . Hence each ϕ_t is smooth.*

Proof. "Introduction to Differentiable Manifolds" by Serge Lang. \square

Here's a theorem:

Theorem 9.7. *Let X be a C^∞ vector field on M and let $p \in M$. Then there is an open set V containing p and an $\epsilon > 0$ such that there is a unique collection of diffeomorphisms $\phi_t : V \rightarrow \phi_t(V) \subset M$ for $|t| < \epsilon$ with the following properties:*

1. $\phi : (-\epsilon, \epsilon) \times V \rightarrow M$ defined by $\phi(t, p) = \phi_t(p)$ is C^∞ .

2. If $|s|, |t|, |s+t| < \epsilon$, and q and $\phi_t(q) \in V$, then

$$\phi_{s+t}(q) = \phi_s \circ \phi_t(q).$$

3. If $q \in V$, then X_q is the tangent vector at $t = 0$ of the curve $t \mapsto \phi_t(q)$.

If M is compact, or more generally, if X has compact support, then the local diffeomorphisms above can be extended to all of M :

Theorem 9.8. *If X has compact support, then there are diffeomorphisms $\phi_t : M \rightarrow M$ for all $t \in \mathbb{R}$, satisfying the above three properties.*

For proof, see Spivak. To get a sense of what this theorem is about, image X being a vector field on the sphere: let imagine it be horizontal “winds”. Then the ϕ_t “twist” the sphere more and more.

The map $t \mapsto \phi_t$ above is called a *1-parameter group of diffeomorphisms* and is *generated by X* . If $\text{Diff}(M)$ is the group of diffeomorphisms of M , then the above tells us that each vector field X gives us a map $\mathbb{R} \rightarrow \text{Diff}(M)$, given by $t \mapsto \phi_t$.

Theorem 9.9. *Let X be a smooth vector field on M with $X(p) \neq 0$. Then there is a coordinate system (x, U) around p such that*

$$X = \frac{\partial}{\partial x^1}.$$

9.1 Lie derivatives

Here are some definitions that are really different instances of the same definition:

$$L_X f = X(f) = \lim_{h \rightarrow 0} \frac{1}{h} [f \circ \phi_h - f]$$

$$L_X \omega = \lim_{h \rightarrow 0} \frac{1}{h} [(\phi_h^* \omega) - \omega]$$

$$L_X Y = \lim_{h \rightarrow 0} \frac{1}{h} [Y - (\phi_{h*} Y)_p]$$

Proposition 9.10. *Let X, Y_1, Y_2 be vector fields and ω_1, ω_2 be covector fields. Then*

$$1. \quad L_X(Y_1 + Y_2) = L_X Y_1 + L_X Y_2.$$

$$2. \quad L_X(\omega_1 + \omega_2) = L_X \omega_1 + L_X \omega_2.$$

$$3. L_X(fY) = Xf \cdot Y + f \cdot L_X Y.$$

$$4. L_X(f \cdot \omega) = Xf \cdot \omega + f \cdot L_X \omega.$$

If we let $[X, Y]$ denote the vector field given by $[X, Y]_p(f) = X_p(Yf) - Y_p(Xf)$, then one sees that

$$L_X Y = [X, Y].$$

Now it follows, almost trivially, that

$$L_X Y = -L_Y X$$

and

$$L_X X = 0.$$

We also have the Jacobi identity:

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

Recall: If we have a diffeomorphism, we can push forward vector fields by defining $(\alpha_* X)_q = \alpha_* X_{\alpha^{-1}(q)}$.

Lemma 9.11. *Let $\alpha : M \rightarrow N$ be a diffeomorphism and X a vector field on M which generates $\{\phi_t\}$. Then $\alpha_* X$ generated $\{\alpha \circ \phi_t \circ \alpha^{-1}\}$.*

Corollary 9.12. *If $\alpha : M \rightarrow M$, then $\alpha_* X = X$ if and only if $\phi_t \circ \alpha = \alpha \circ \phi_t$ for all t .*

Lemma 9.13. *Let X generate $\{\phi_t\}$ and Y generate $\{\psi_t\}$. Then $[X, Y] = 0$ if and only if $\phi_t \circ \psi_s = \psi_s \circ \phi_t$ for all s, t .*

Theorem 9.14. *If X_1, \dots, X_k are linearly independent C^∞ vector fields around p and $[X_i, X_j] = 0$ for all i, j , then there is a coordinate system (x, U) around p such that*

$$X_\alpha = \frac{\partial}{\partial x^i}$$

for $i = 1, \dots, n$.

10 Differential forms

Depending upon taste, there are several ways to define differential forms. This has already been mentioned in an earlier chapter. Perhaps this is the time to reflect upon the strange order of things in Spivak's book?

Suffice it to say that the only thing that is needed of the exterior algebra, regardless of the particular construction, is the anticommutativity property, mimicking determinants (or if you're so inclined: defining determinants).

10.1 Basic theorems

Theorem 10.1. *Let $f : M \rightarrow N$ be a map between n -manifolds of the same dimension n . Let (x, U) be a coordinate system around $p \in M$ and (y, V) be a coordinate system around $q = f(p) \in N$. Then*

$$f^*(g \, dy^1 \wedge \dots \wedge dy^n) = (g \circ f) \cdot \det \left(\frac{\partial y^i \circ f}{\partial x^j} \right) dx^1 \wedge \dots \wedge dx^n.$$

Theorem 10.2. *A manifold M is orientable if and only if there exists a non-vanishing form $\omega \in \Omega^n(M)$.*

10.2 The de Rham complex

If

$$\omega = \sum_I \omega_I dx^I.$$

Then we define

$$d\omega := \sum_I d\omega_I dx^I = \sum_I \sum_{\alpha=1}^n \frac{\partial \omega_I}{\partial x^\alpha} dx^\alpha \wedge dx^I.$$

Example 10.3. Let ω be the 2-form on \mathbb{R}^3 defined by $\omega = xdx \wedge dy - ydx \wedge dz$. Then

$$d\omega = dx \wedge dx \wedge dy - dy \wedge dx \wedge dz = dx \wedge dy \wedge dz.$$

★

Proposition 10.4. 1. $d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$.

2. *There is a Leibniz rule: if ω_1 is a k -form, then*

$$d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge d\omega_2.$$

3. $d^2 = 0$.

Further, if d' is another function $\Omega^k(M) \rightarrow \Omega^{k+1}(M)$ obeying these rules, then $d = d'$.

Corollary 10.5. *There is a unique operator from k -forms on M to $k+1$ -forms on M satisfying the above rules.*

Proof. The above proposition says that for each coordinate system (x, U) , we have a unique d_U defined. Given the form ω and $p \in M$, pick any U with $p \in U$ and define

$$d\omega(p) = d(\omega_U)(p).$$

This proves the claim. ²

□

Proposition 10.6. *If $f : M \rightarrow N$ is C^∞ and ω is a k -form on N , then*

$$f^*(d\omega) = d(f^*\omega).$$

10.3 The Poincaré Lemma

Recall that a manifold M is *smoothly contractible to a point* $p_0 \in M$ if there is a map $H : M \times [0, 1] \rightarrow M$ such that $H|_1 = \text{id}_M$ and $H|_0$ is the constant map mapping to p_0 .

Example 10.7. The canonical example is \mathbb{R}^n : a contraction to $\vec{0} \in \mathbb{R}^n$ is defined just by $t\vec{x}$. More generally, any convex region in \mathbb{R}^n is contractible. Even more generally, every star-shaped region in \mathbb{R}^n is contractible.

Even more generally, any region homotopic to a starshaped (hence convex) region is contractible. ★

The Poincaré Lemma states that if ω is a closed form on a contractible manifold M , then M is exact. In short, $H_{dR}^k(M) = 0$ for $k > 0$. In fact, we will prove something more general. To do that, we need some terminology:

For any $t \in [0, 1]$ we have natural maps $i_t : M \hookrightarrow M \times [0, 1]$. Thus for each t , we have pullback maps $i_t^* : \Omega_{M \times [0, 1]}^k \rightarrow \Omega_M^k$. By Proposition 10.6, this induces maps on cohomology. We will prove that they in fact define the *same* map on cohomology. In fact, they are chain homotopic. We will construct a map $I : \Omega_{M \times [0, 1]}^{k+1} \rightarrow \Omega_M^k$ for all k , such that $d(I\omega) + I(d\omega) = i_1^*\omega - i_0^*\omega$. By standard homological algebra, this implies that the maps on cohomology are equal.

²However, I don't see why. If U' is another coordinate chart, how does one see that $d_{U'}(\omega|_{U'})(p)$ gives the same number?

We have the following diagram:

$$\begin{array}{ccccc}
\Omega_M^{k-1} & \xrightarrow{d} & \Omega_M^k & & \\
\uparrow i_0^* & & \uparrow i_1^* & \swarrow I & \\
\Omega_{M \times [0,1]}^{k-1} & \xrightarrow{d} & \Omega_{M \times [0,1]}^k & \xrightarrow{d} & \Omega_{M \times [0,1]}^{k+1}
\end{array}$$

We want to show the relation $d(I\omega) + I(d\omega) = i_1^*\omega - i_0^*\omega$ for some I that will be defined now.

Recall that if V is a vector space that is the tensor product of two vector spaces A, B , then $\bigwedge^k V = \bigoplus_{i+j=k} \bigwedge^i A \otimes \bigwedge^j B$. Applying this to $\Omega_{M \times [0,1]}^k$ gives that we have a canonical decomposition as abelian groups

$$\Omega_{M \times [0,1]}^k \approx \Omega_M^k \oplus \Omega_{[0,1]}^1 \otimes \Omega_M^{k-1} \approx \Omega_M^k \oplus \Omega_M^{k-1},$$

where the last isomorphism is non-canonical, and can be seen after choosing bases. In particular, this says that any k -form $\omega \in \Omega_{M \times [0,1]}^k$ can be written uniquely (after choosing bases) as $\omega_1 + (dt \wedge \eta)$ for some $\eta \in \Omega_{M \times [0,1]}^{k-1}$.³

Having written ω in this way, we define I as follows:

$$I\omega(p)(X_1, \dots, X_{k-1}) := \int_0^1 \eta(p, t)(i_{t*}X_1, \dots, i_{t*}X_{k-1})dt.$$

In coordinates, this works as follows: Suppose (x, U) is a coordinate chart. Then we can write $\omega = \sum_I a^I dx^I + \sum_J a^J dt \wedge dx^J$ where I ranges over k -subsets of $\{1, \dots, n\}$, and J ranges over $(k-1)$ -subsets of $\{1, \dots, n\}$. Then $\eta = \sum_J a^J dx^J$. Thus

$$I\omega(p) = \sum_J \int_0^1 a^J(p, t) dt dx^J.$$

Here's the theorem:

Theorem 10.8. *For any k -form ω on $M \times [0, 1]$, we have*

$$i_1^*\omega - i_0^*\omega = d(I\omega) + I(d\omega).$$

Proof. Since all maps involved are linear, there are basically just two cases.

Case A is when $\omega = \omega_1$ (in the notation above), involving no dt -terms, say

³As it stands, this is not precise. There exists a canonical splitting, but not as stated. Both factors are still $C^\infty(M \times [0, 1])$ -modules, and this is not reflected in the notation.

$\omega = f dx^I$. In this case $I\omega = 0$, so we need to prove that $i_1^*\omega - i_0^*\omega = I(d\omega)$. The left hand side is just

$$i_1^*\omega(p) - i_0^*\omega(p) = f(p, 1)dx^I - f(p, 0)dx^I = (f(p, 1) - f(p, 0)) dx^I.$$

The right hand side is

$$\begin{aligned} I(d(f dx^I)) &= I \left(\left(\sum_i \frac{\partial f}{\partial x^i} dx^i + \frac{\partial f}{\partial t} dt \right) \wedge dx^I \right) \\ &= \left(\int_0^1 \frac{\partial f}{\partial t} dt \right) dx^I \end{aligned}$$

But these two expressions are equal by the fundamental theorem of calculus.

Now for **Case B**. Here $\omega = f dt \wedge dx^J$. In this case $i_t^*\omega = 0$ for all t , so we want to prove that $d(I\omega) = -I(d\omega)$:

$$\begin{aligned} d(I\omega)(p) &= d \left(\int_0^1 f(p, t) dt \wedge dx^J \right) \\ &= \sum_{i=1}^n \frac{\partial}{\partial x^i} \left(\int_0^1 f(p, t) dt \right) dx^i \wedge dx^J \\ I(d\omega)(p) &= I \left(\sum_{i=1}^n \frac{\partial f}{\partial x^i} dx^i \wedge dt \wedge dx^J \right) \\ &= I \left(- \sum_{i=1}^n \frac{\partial f}{\partial x^i} dt \wedge dx^i \wedge dx^J \right) \\ &= - \sum_{i=1}^n \left(\int_0^1 \frac{\partial f}{\partial x^i} dt \right) dx^i \wedge dx^J. \end{aligned}$$

Thus the proof is complete. \square

The Poincaré Lemma now follows as an easy corollary:

Lemma 10.9 (Poincaré⁴). *If M is contractible, then $H_{dR}^k(M) = 0$ for $k > 0$.*

Proof. Let ω be such that $d\omega = 0$. We are given a $H : M \times [0, 1] \rightarrow M$ as above. I.e. such that $H \circ i_0 = p_0$ (the constant map to p_0), and $H \circ i_1 = \text{id}_M$. Thus

$$\omega = \text{id}_M^* \omega = (H \circ i_1)^* \omega = i_1^*(H^* \omega)$$

⁴Though it is not clear if Poincaré actually proved it.

and also

$$0 = (H \circ i_0)^*(\omega) = i_0^*(H^*\omega)$$

This last pullback is zero because every form on a point-space vanishes. But also $d(H^*\omega) = H^*(d\omega) = H^*(0) = 0$, by Proposition 10.6. But then $\omega = i_1^*(H^*\omega) = d(I(H^*\omega))$ by the Theorem. Thus ω is a boundary, and so $H^k = 0$. \square

11 Integration

A map $c : [0, 1]^k \rightarrow M$ is called a *singular k -cube* in M . By convention, a singular 0-cube is just the assignment of a point in M . The inclusion map $I : [0, 1]^k \rightarrow \mathbb{R}^k$ will be called the *standard k -cube*. The idea is that we will integrate over “cubes”.

Given a k -form ω on $[0, 1]^k$ and a coordinate system x^1, \dots, x^k , ω can be uniquely written as

$$\omega = f dx^1 \wedge \dots \wedge dx^k.$$

For some function $f \in \Gamma([0, 1]^k, \mathcal{O}_{[0, 1]^k})$.

Definition 11.1. If ω is a k -form on $[0, 1]^k$, we define the *integral of ω* to be

$$\int_{[0, 1]^k} \omega := \int_{[0, 1]^k} f,$$

where the right-hand side denotes classical integration over subsets of \mathbb{R}^n . ■

This immediately generalizes to forms on M :

Definition 11.2. If ω is a k -form on M , and c is a singular k -cube in M , we define

$$\int_c \omega = \int_{[0, 1]^k} c^* \omega.$$

The right-hand side was defined in the previous proposition. ■

Remark. A k -cube is just a map $c : [0, 1]^k \rightarrow M$, and on the left manifold, we have a canonical set of coordinates. The values of $c^* \omega$ at points $p \in [0, 1]^k$ only depend upon this chart, and no charts of M are needed in the definition. Hence this is well-defined.

Remark. If $k = 0$, then sections of $\Omega^k(M)$ are just functions $M \rightarrow \mathbb{R}$. In particular, each 0-form is just a function f , and a 0-cube is just a map $c : \{0\} \rightarrow M$. In this case

$$\int_c \omega = \int_{\{0\}} c^* \omega = c(f(0)).$$

Now, let C^k be the free abelian group generated by *all* k -forms. Explicitly, the elements of C^k are finite sums $n_1 c_1 + \dots + n_r c_r$, where the n_i are real

numbers and the c_i are k -forms. We call the elements of C^k for k -chains. Then we define the integral of a k -chain c to be

$$\int_c \omega := \sum_{i=1}^r n_i \int_{c_i} \omega.$$

This way, we get for each k an “integration” pairing $\int : C^k \times \Omega^k(M) \rightarrow \mathbb{R}$, which is obviously bilinear.

Proposition 11.3. *Let $c : [0, 1]^n \rightarrow \mathbb{R}^n$ be a one-one singular n -cube with $\det c' \geq 0$ on $[0, 1]^n$. Let ω be the n -form*

$$\omega = f dx^1 \wedge \dots \wedge dx^n.$$

Then

$$\int_c \omega = \int_{c([0,1]^n)} f.$$

Proof.

$$\begin{aligned} \int_c \omega &= \int_{[0,1]^n} c^* \omega \\ &= \int_{[0,1]^n} (f \circ c)(\det c') dx^1 \wedge \dots \wedge dx^n \\ &= \int_{[0,1]^n} (f \circ c) |(\det c')| dx^1 \wedge \dots \wedge dx^n \\ &= \int_{c([0,1]^n)} f \end{aligned}$$

The first equality is by definition, the second is by a basic theorem about transformation of wedge algebras. The third is by assumption, and the fourth is by the change of variable formula in ordinary calculus (see for example Spivak’s “Calculus on Manifolds”, Theorem 3.13, page 67). \square

The corollary below says that the integral is independent of orientation-preserving reparametrizations of the k -cube. In particular, the map \int above is really a function from equivalence classes of k -cubes to \mathbb{R} .

Corollary 11.4. *Let $p : [0, 1]^k \rightarrow [0, 1]^k$ be one-one onto with $\det p' \geq 0$. Let c be a singular k -cube in M and let ω be a k -form on M . Then*

$$\int_c \omega = \int_{c \circ p} \omega.$$

Proof. We have

$$\begin{aligned}\int_{c \circ p} \omega &= \int_{[0,1]^k} (c \circ p)^* \omega = \int_{[0,1]^k} p^*(c^*(\omega)) \\ &= \int_{[0,1]^k} c^*(\omega) \\ &= \int_c \omega.\end{aligned}$$

The last equality is by definition, the one before that follows from the proposition, since p is onto. \square

11.1 The boundary of a chain

One should think of 1-chains as sums of “paths” on M . Then the boundary ∂c is just the difference $c(1) - c(0)$ in $C^0(M)$ ⁵. For a 2-form c , one should think of c as some “surface” lying on M , and ∂c as the oriented boundary of c .

Let $I^k : [0, 1]^k \rightarrow \mathbb{R}^n$ be the standard k -cube. From this, we can define several $(k-1)$ -cubes as follows: Suppose $x \in [0, 1]^{k-1}$. Then let

$$I_{(i,\alpha)}^k(x) = (x^1, \dots, x^{i-1}, \alpha, x^i, \dots, x^{k-1}).$$

Here $\alpha = 0, 1$ and $i = 1, \dots, k-1$. We call this the (i, α) -face of I^k . For a general k -cube c , we define the (i, α) -face to be $c_{i,\alpha} := c \circ I_{(i,\alpha)}^n$.

Definition 11.5. The *boundary* ∂c of a k -cube c is defined to be

$$\partial c := \sum_{i=1}^n \sum_{\alpha=0,1} (-1)^{i+\alpha} c_{(i,\alpha)}.$$

We extend to general k -chains $c \in C^k(M)$ by additivity.

If $k = 0$, we define $\partial c = \sum a_i$, where the a_i are the coefficients of c in C^k . \blacksquare

We have a complex!

Proposition 11.6. *If c is a k -chain in M , then $\partial(\partial(c)) = 0$. Or in jargon: $\partial^2 = 0$.*

⁵Recall that this abelian group is the group freely generated by the *points* on M . In fancy language, one obtains $C^0(M)$ by first applying the forgetful functor $F : \text{Diff} \rightarrow \text{Set}$, and then the “free abelian group functor” $\text{Set} \rightarrow \text{Ab}$.

Proof. Without loss of generality, we can assume that $i \leq j$. An easy computation gives that:

$$(I_{(i,\alpha)}^n)_{(j,\beta)}(x) = (x^1, \dots, x^{i-1}, \alpha, x^i, \dots, x^{j-1}, \beta, x^j, \dots, x^{n-2}),$$

where the α is in the i th position, and β is in the $j+1$ th position. Similarly:

$$(I_{(j+1,\beta)}^n)_{(i,\alpha)}(x) = (x^1, \dots, x^{i-1}, \alpha, x^i, \dots, x^{j-1}, \beta, x^j, \dots, x^{n-2}).$$

Thus $(I_{(i,\alpha)}^n)_{(j,\beta)} = (I_{(j+1,\beta)}^n)_{(i,\alpha)}$ for $i \leq j \leq n-1$. Thus, for any singular n -cube c , we have that $(c_{(i,\alpha)})_{(j,\beta)} = (c_{(j+1,\beta)})_{(i,\alpha)}$ for $i \leq j \leq n-1$. Now, by definition:

$$\begin{aligned} \partial\partial(c) &= \partial\left(\sum_{i=1}^n \sum_{\alpha} (-1)^{i+\alpha} c_{(i,\alpha)}\right) \\ &= \sum_{i=1}^n \sum_{\alpha} \sum_{\beta} \sum_{j=1}^{n-1} (-1)^{i+j+\alpha+\beta} (c_{(i,\alpha)})_{(j,\beta)}. \end{aligned}$$

But $(c_{(j+1,\beta)})_{(i,\alpha)}$ appear with opposite sign as $(c_{(i,\alpha)})_{(j,\beta)}$, so all terms cancel! \square

The homology groups gives the *singular smooth homology groups of M* (with coefficients in \mathbb{R}), and are denoted simply by $H_{\infty}^k(M, \mathbb{R})$. By applying $\text{Hom}(-, \mathbb{R})$ to the complex, we get the dual complex, and thus the *singular smooth cohomology groups of M* , denoted by $H_{\infty}^k(M, \mathbb{R})$.

We say that a chain is *closed* if $\partial c = 0$. A chain c is a *boundary* if $c = \partial b$ for some $b \in C^{k+1}(M)$.

Theorem 11.7 (Stoke's Theorem). *If ω is a $(k-1)$ -form on M and c is a k -chain in M , then*

$$\int_c d\omega = \int_{\partial c} \omega.$$

The integration map descends to homology:

Proposition 11.8. *We have a map:*

$$\int : H_k(M, \mathbb{R}) \otimes H_{dR}^k(M) \rightarrow \mathbb{R}.$$

Proof. We need to show that the map $(c, \omega) \mapsto \int_c \omega$ does not depend on the homology classes of c and ω . By bilinearity of the pairing, it is enough to

show that if either of c or ω are boundaries, then the pairing is zero. So assume $\partial c = 0$ and $d\omega = 0$, and first that $c = \partial c'$.

$$\int_c \omega = \int_{\partial c'} \omega = \int_{c'} d\omega = \int_{c'} 0 = 0.$$

Now assume $\omega = d\eta$. Then

$$\int_c \omega = \int_c d\eta = \int_{\partial c} \eta = 0$$

since $\partial c = 0$. □

11.2 Connection to the fundamental group

It is a fact of topology that $\pi_1(M)/[\pi_1(M), \pi_1(M)] \approx H^1(M)$. We will prove something weaker.

Proposition 11.9. *If $\pi_1(M)$ is finite, then $H^1(M) = 0$.*

For this, we need a lemma.

Lemma 11.10. *Suppose M is connected. Let $\omega \in Z^1(M)$ and suppose $[\omega] \neq 0$ in $H^1(M)$. Then there exists a loop $\gamma : S^1 \rightarrow M$ such that*

$$\omega(\gamma) := \int_{S^1} \gamma^* \omega \neq 0.$$

Proof. Assume on the contrary that $\omega(\gamma) = 0$ for all $\gamma : S^1 \rightarrow M$. Choose a base point $x_0 \in M$. We want to define f such that $df = \omega$.

Suppose $\gamma_1, \gamma_2 : [0, 1] \rightarrow M$ are two paths, both starting at x_0 and ending at x . Then $\gamma_1 \circ \gamma_2^{-1} : S^1 \rightarrow M$ is a loop. Then by the assumption:

$$\omega(\gamma_1 \circ \gamma_2^{-1}) = \int_0^1 \gamma_1^* \omega - \int_0^1 \gamma_2^* \omega = 0$$

This shows that the function

$$f(x) := \int_0^1 \gamma^* \omega$$

for any γ a path starting at x_0 , ending at x is a well-defined function.

Now we claim that $df = \omega$. For this, choose a chart (U, x) around x , and write $\omega = \sum_{i=1}^n a^i dx^i$. We can suppose that $x = 0$ in this chart. Write

$df(0)$ as $\sum_{i=1}^n \frac{\partial f}{\partial x^i} \Big|_0 dx^i$. Then we want to check that each of the terms are equal. Let us compute $\frac{\partial f}{\partial x^1} \Big|_0$:

$$\begin{aligned} \frac{\partial f}{\partial x^1} \Big|_0 &= \lim_{h \rightarrow 0} \frac{1}{h} (f(h, 0, \dots, 0) - f(0)) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \left(\int_0^1 \gamma^* \omega + \int_0^1 \tau_h^* \omega - \int_0^1 \gamma^* \omega \right) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \int_0^1 \tau_h^* \omega. \end{aligned}$$

Here τ_h is the path starting at 0 and ending at h , given simply by $ht\vec{e}_1$, where \vec{e}_1 is the first unit vector in \mathbb{R}^n . Then $\tau_h^* \omega = ha^1(ht, 0, \dots, 0)dx^1$. Thus the last integral is simply

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{1}{h} \int_0^1 \tau_h^* \omega &= \lim_{h \rightarrow 0} \frac{1}{h} \int_0^1 ha^1(ht, 0, \dots, 0)dx^1 = \lim_{h \rightarrow 0} \int_0^1 a^1(ht, 0, \dots, 0)dt \\ &= \int_0^1 a^1(0, \dots, 0)dt = a^1(0, \dots, 0). \end{aligned}$$

As wanted. But this means that $[\omega] = 0$ in $H^1(M)$, contrary to assumption. \square

Proof of the Proposition. If $\pi_1(M)$ is finite, then for all $[\gamma] \in \pi_1(M)$, there is an integer n such that $[\pi]^n = [0]$. This means that for the n -fold concatenation $\Gamma := \gamma \cdot \dots \cdot \gamma$, there is a homotopy $H : S^1 \times [0, 1] \rightarrow M$ such that $H|_0 = \Gamma$ and $H|_1 = k$, a constant function.

Now let $\omega \in \Omega^1(M)$ be a 1-form with $d\omega = 0$. Then by Stoke's theorem

$$\int_{S^1 \times [0, 1]} F^* d\omega = \int_{S^1} (\gamma \cdot \dots \cdot \gamma)^* \omega - \int_{S^1} k^* \omega = n \int_{S^1} \gamma^* \omega = 0.$$

Thus all curves integrate to zero on all forms, and then by the Lemma, we must have $H^1(M) = 0$. \square

A Definitions

Here $f : M \rightarrow N$ is a map. $\alpha : M \rightarrow N$ is a diffeomorphism. X is a vector field. $\{\phi_h\}$ is the 1-parameter subgroup generated by X .

$$\begin{aligned}
(f_*X_p)(g) &= X_p(g \circ f) \\
(f^*\omega)(p)(X_p) &= \omega(f(p))(f_*X_p) \\
(\alpha_*X)_p &= \alpha_*(X_{\alpha^{-1}(p)}) \\
L_X f &= Xf \\
L_X \omega &= \lim_{h \rightarrow 0} \frac{1}{h} [\phi_h^* \omega - \omega] \\
L_X Y &= \lim_{h \rightarrow 0} \frac{1}{h} [Y - \phi_{h*} Y] = [X, Y](p) \\
\phi^*(df)(Y) &= Y(f \circ \phi) \\
df(p)(X_p) &= X_p(f) \\
L_X df &= d(L_X f) = d(Xf)
\end{aligned}$$

B Formulas

B.1 Local description of (co-)vector fields

By local triviality, every vector field $X : M \rightarrow TM$ locally looks like $U \times \mathbb{R}^n$, where \mathbb{R}^n is the n -dimensional vector space spanned by $\{\frac{\partial}{\partial x^1}|_p, \dots, \frac{\partial}{\partial x^n}|_p\}$ for $p \in U$. In particular:

Proposition B.1. *Let $X : M \rightarrow TM$ be a section of the tangent bundle. Then in a local coordinate system $x : U \rightarrow \mathbb{R}^n$, X can be written as*

$$X(p) = \sum_{i=1}^n a^i(p) \frac{\partial}{\partial x^i} \Big|_p.$$

The a^i are C^∞ functions $U \rightarrow \mathbb{R}$.

Similarly, if $\omega : M \rightarrow T^*M$ is a section of the cotangent bundle, the fibers are spanned by the differentials $\{dx^1, \dots, dx^n\}$.

Proposition B.2. *Let $\omega : M \rightarrow T^*M$ be a section of the cotangent bundle. Then in a local coordinate system $x : U \rightarrow \mathbb{R}^n$, ω can be written as*

$$\omega(p) = \sum_{i=1}^n a^i(p) dx^i(p).$$

Vector fields act on functions $f : M \rightarrow \mathbb{R}$. In coordinates:

Proposition B.3. *If $f : M \rightarrow \mathbb{R}$ is a function, and U is an open trivializing neighbourhood of p , then*

$$X(f)(p) = \sum_{i=1}^n a^i(p) \frac{\partial(f \circ x^{-1})}{\partial x^i} \Big|_p.$$

Proof. This is just Proposition B.1 applied to a function f , keeping in mind that the $\frac{\partial}{\partial x^i}|_p$ are derivations. \square

If $f : M \rightarrow \mathbb{R}$ is a C^∞ map, then it induces a section of the cotangent bundle as follows: $df(p)(X) = X(f)$. In local coordinates we have the following:

Proposition B.4. *Let $f : M \rightarrow \mathbb{R}$ be a C^∞ map and (x, U) a chart. Then*

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x^i} dx^i$$

Proof. Let $X = \frac{\partial}{\partial x^i}|_p$ and use that the set of these are a dual basis for $\{dx_i(p)\}$. \square

B.2 Pullbacks and pushforwards

Given a map $f : M \rightarrow N$ and a point $p \in M$ and a vector $X_p \in T_p M$, we can *push forward* X_p to $f_* X_p$, simply by functoriality of f_* .

Given a map $f : M \rightarrow N$ and a covector field on N (i.e. a section $\omega : N \rightarrow T^* N$ of the cotangent bundle), there is a pullback covector field on M , denoted $f^* \omega$.

Definition B.5 (Pushforward of tangent vectors). If X_p is a tangent on M , then we can define a tangent vector on N by

$$f_*(X_p)(g) := X_p(g \circ f).$$

■

Definition B.6. Let $x : U \rightarrow \mathbb{R}^n$ be local coordinates on M and let $y : V \rightarrow \mathbb{R}^m$ be local coordinates on N . Let

$$X_p = \sum_{i=1}^n a^i \frac{\partial}{\partial x^i} \Big|_p.$$

Then

$$f_*(X_p)(g) = \sum_{j=1}^m \sum_{i=1}^n \frac{\partial y^j}{\partial x^i} \Big|_p \frac{\partial}{\partial y^j} \Big|_p.$$

■

Definition B.7 (Pullback of cotangent fields). We define how $f^* \omega$ acts on vectors $X_p \in T_p M$:

$$(f^* \omega)(p)(X_p) := w(f(p))(f_*(X_p)).$$

■

In coordinates this looks as follows:

Proposition B.8. Let $x : U \rightarrow \mathbb{R}^n$ be local coordinates on M and $y : V \rightarrow \mathbb{R}^m$ be local coordinates on N . Let ω be described locally by

$$\omega = \sum_{i=1}^m a^i dy^i.$$

Then $f^* \omega$ is described locally by

$$f^* \omega(p) = \sum_{i=1}^n \sum_{j=1}^m \frac{\partial (y^j \circ f)}{\partial x^i} \Big|_p a^j(f(p)) dx^i(p).$$

If we define $\langle -, - \rangle$ by $\langle \omega, X \rangle := \omega(X)$, we have:

Proposition B.9. *Pushforward and pullback are actually dual. That is,*

$$\langle f^*\omega, X \rangle = \langle \omega, f_*X \rangle.$$

Proof. This is actually the definition. □

B.3 Lie derivatives

The Lie derivative of a function $f : M \rightarrow \mathbb{R}$ along X is just $L_X f := Xf$. Thus, in coordinates, this is just $X(f)$, so the coordinate description is the content of Proposition B.3.

The Lie derivative of a covariant vector field is by definition

$$L_X \omega(p) := \lim_{h \rightarrow 0} \frac{1}{h} [(\phi_h^* \omega)(p) - \omega(p)].$$

Before we give the coordinate description of $L_X \omega$, we want to prove a useful lemma:

Lemma B.10. *We have $L_X df = d(L_X f)$.*

Proof. This is just definition-hunting. We let $L_X(df)$ act on a tangent vector Y .

$$\begin{aligned} L_X(df)(p)(Y_p) &= \lim_{h \rightarrow 0} \frac{1}{h} [(\phi_h^* df)(p) - df(p)](Y_p) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} [(\phi_h^* df)(p)(Y_p) - df(p)(Y_p)] \\ &= \lim_{h \rightarrow 0} \frac{1}{h} [Y_p(f \circ \phi_h)(p) - Y_p(f)(p)] \\ &= Y_p \lim_{h \rightarrow 0} \frac{1}{h} [(f \circ \phi_h)(p) - f(p)] \\ &= Y_p \cdot Xf(p) = d(Xf)(Y_p)(p). \end{aligned}$$

Dropping the Y_p and the p from the notation gives the result. □

Here's another useful lemma:

Lemma B.11.

$$\phi^*(df)(Y) = Y(f \circ \phi).$$

Proof. By definition $\phi^*(df)(p)(Y_p) = df(\phi(p))(\phi_* Y_p)$. Also $df(\phi(p))(\phi_* Y_p) = \phi_* Y_p(\phi)(f) = Y_p(f \circ \phi)$. Everything was just definitions. □

In coordinates we have the following:

Proposition B.12. *Let ω have local description as*

$$\omega(p) = \sum_{i=1}^n b^i(p) dx^i.$$

Then

$$L_X \omega = \sum_{i=1}^n \sum_{j=1}^n \left(a^j \frac{\partial(b^i \circ x^{-1})}{\partial x^j} dx^i + b^i \frac{\partial a^j}{\partial x^j} dx^j \right).$$

Proof. We first calculate $L_X dx^i$. From Proposition B.10 we have $L_X df = d(L_X f) = d(Xf)$, so that

$$d(L_X x^i) = d(Xx^i) = d\left(\sum_{j=1}^n a^j \frac{\partial x^i}{\partial x^j}\right) = da^i.$$

Thus, using Proposition B.4, we conclude that $d(L_X x^i) = \sum_{j=1}^n \frac{\partial a^i}{\partial x^j} dx^j$. Now, using the fact that $L_X(f \cdot \omega) = Xf \cdot \omega + f \cdot L_X \omega$, we can calculate $L_X(b^i dx^i)$:

$$\begin{aligned} L_X(b^i dx^i) &= X(b^i) \cdot dx^i + b^i \cdot L_X dx^i \\ &= \sum_{j=1}^n a^j \frac{\partial b^i}{\partial x^j} dx^i + \sum_{j=1}^n b^i \frac{\partial a^j}{\partial x^j} dx^j \\ &= \sum_{j=1}^n \left(a^j \frac{\partial b^i}{\partial x^j} dx^i + b^i \frac{\partial a^j}{\partial x^j} dx^j \right). \end{aligned}$$

The result follows by summing over all i . □

The Lie derivative of a vector field Y along a vector field X is defined as follows:

$$(L_X Y)(p) = \lim_{h \rightarrow 0} \frac{1}{h} [Y_p - (\phi_{h*} Y)_p].$$

We can define the *bracket* $[X, Y]$ by $[X, Y](p)(f) = X_p(Yf) - Y_p(Xf)$. Then we have:

Proposition B.13.

$$L_X Y = [X, Y]$$

Proof. Let $f : M \rightarrow \mathbb{R}$ be a function and let X generate $\{\phi_t\}$. Define

$$\alpha(t, h) := Y_{\phi_{-t}(p)}(f \circ \phi_h).$$

We differentiate in each direction. First, we have

$$\begin{aligned} \left. \frac{\partial \alpha}{\partial t} \right|_{(0,0)} &= \lim_{t \rightarrow 0} \frac{1}{t} [Y_{\phi_{-t}(p)}(f) - Y_p(f)] \\ &= -\lim_{t \rightarrow 0} \frac{1}{t} [Y_p(f) - Y_{\phi_{-t}(p)}(f)] = -X_p(Yf). \end{aligned}$$

On the other hand:

$$\begin{aligned} dd|_{\alpha} h(0, 0) &= \lim_{h \rightarrow 0} \frac{1}{h} [Y_p(f \circ \phi_h) - Y_p(f)] \\ &= Y_p(Xf). \end{aligned}$$

Now, set $c(h) = \alpha(h, h) = Y_{\phi_{-h}(p)}(f \circ \phi_h)$. Then, by the chain rule, the derivative $c'(0)$ is given by adding the two derivatives above. In addition:

$$\begin{aligned} -c'(0) &= \lim_{h \rightarrow 0} \frac{1}{h} [Y_p f - Y_{\phi_{-h}(p)}(f \circ \phi_h)] \\ &= \lim_{h \rightarrow 0} \frac{1}{h} [Y_p f - \phi_{h*} Y_{\phi_{-h}(p)}(f)] \\ &= \lim_{h \rightarrow 0} \frac{1}{h} [Y_p f - (\phi_h Y)_p(f)] = L_X Y(p)(f). \end{aligned}$$

Thus $L_X Y(p)(f) = [X, Y]_p(f)$. □

From the above proof, it is easy to derive a coordinate expression for $L_X Y = [X, Y]$.

Proposition B.14. Let $X = \sum_{i=1}^n a^i \frac{\partial}{\partial x^i}$ and $Y = \sum_{i=1}^n b^i \frac{\partial}{\partial x^i}$. Then

$$L_X Y = \sum_{i=1}^n \sum_{j=1}^n \left(a^j \frac{\partial b^i}{\partial x^j} - b^i \frac{\partial a^j}{\partial x^i} \right) \frac{\partial}{\partial x^i}.$$

Proof. This is just a matter of calculation:

$$\begin{aligned} L_X Y(f) &= [X, Y](f) = X(Yf) - Y(Xf) \\ &= X \left(\sum_{i=1}^n b^i \frac{\partial f}{\partial x^i} \right) - Y \left(\sum_{i=1}^n a^i \frac{\partial f}{\partial x^i} \right) \\ &= \sum_{i=1}^n \left(X \left(b^i \frac{\partial f}{\partial x^i} \right) - Y \left(a^i \frac{\partial f}{\partial x^i} \right) \right) \end{aligned}$$

Now, using the chain rule, we find that

$$X \left(b^i \frac{\partial f}{\partial x^i} \right) = \sum_{j=1}^n \left(a^j \frac{\partial f}{\partial x^i} \frac{\partial b^i}{\partial x^j} + a^j b^i \frac{\partial^2 f}{\partial x^i \partial x^j} \right).$$

The double derivatives cancel in the difference above, so that we end up with

$$L_X Y = \sum_{i=1}^n \sum_{j=1}^n \left(a^j \frac{\partial b^i}{\partial x^j} - b^i \frac{\partial a^j}{\partial x^i} \right) \frac{\partial}{\partial x^i}.$$

□

C A concrete exercise

This will be a solution to a mandatory assignment given in 2013.

Let $M(3)$ be the real vector space of 3×3 matrices. Let $\text{Sym}(3)$ be the subspace of symmetric matrices ($A^T = A$) and $\text{Skew}(3)$ be the subspace of skew-symmetric matrices. Note that $\text{Sym}(3) \simeq \mathbb{R}^6$ and $\text{Skew}(3) = \mathbb{R}^3$. Let $f : M(3) \rightarrow \text{Sym}(3)$ be the smooth map given by $f(A) = AA^T$. Let $O(3)$ be the orthogonal group, i.e. the set $\{A \in M(3) \mid AA^T = I\}$, and notice that it is the kernel of the map f .

Exercise 1. Show that $f_{*A} : T_A M(3) \rightarrow T_{f(A)} \text{Sym}(3)$ is given by

$$f_{*A}(C) = AC^T + CA^T.$$

♠

Solution 1. Consider the curve $c(t) = A + tC$ in $M(3)$ through A with tangent vector C . Then we want to compute $(f \circ c)_* \left(\frac{\partial}{\partial t} \Big|_0 \right)$. The composition $f \circ c$ is given by

$$(f \circ c)(t) = AA^T + tAC^T + tCA^T + t^2CC^T.$$

Now derivation with respect to t gives $(f \circ c)_{*A}(C) = AC^T + CA^T$. ♡

Exercise 2. Show that I is a regular value of f . ♠

Solution 2. Recall that p is a regular value if p the differential f_{*p} is surjective.

So let $D \in T_I \text{Sym}(3)$. Then $f_{*I}(C) = C^T + C$. Then $\frac{1}{2}D$ is mapped to D , since D is symmetric. ♡

Exercise 3. Show that $O(3)$ is a smooth submanifold of $M(3)$, and determine its dimension. ♠

Solution 3. This is immediate from Proposition 5.6. The dimension is easily seen to be $9 - 6 = 3$. This is intuitively clear because $O(3)$ is the set of orthogonal transformation of \mathbb{R}^3 . The orientation-preserving elements of $O(3)$ are given by rotation about a line. Choosing a line is a 2-dimensional choice (equivalent to picking a point in $\mathbb{P}_{\mathbb{R}}^2$), and rotation is a 1-dimensional choice. ♥

Now let $k : O(3) \hookrightarrow M(3)$ denote the inclusion map.

Exercise 4. Prove that the image of $k_{*A} : T_A O(3) \hookrightarrow T_A M(3) \approx M(3)$ is equal to the subspace $K := \{C \in M(3) \mid AC^T + CA^T = 0\}$. ♠

Solution 4. Note that the exercise is equivalent to the exactness of the sequence

$$0 \longrightarrow T_A O(3) \xrightarrow{k_{*A}} T_A M(3) \xrightarrow{f_{*A}} T_{f(A)} \text{Sym}(3) \longrightarrow 0$$

The only non-trivial part is exactness in the middle. Since f is constant on elements of $O(3)$, the composite $f_* \circ k_* = (f \circ k)_*$ is zero. Hence $\text{im } k_* \subseteq \ker f_*$.

Consider the map $\text{Skew}(3) \rightarrow K$ given by $S \mapsto SA^{-1,T}$ (A inverse, transposed). This is a bijection since A is invertible, hence K has dimension 3. Thus the dimensions are equal, and the sequence is exact. ♥

Thus we can identify $T_A O(3)$ with this subspace of $M(3)$. In particular, k_{*I} identifies $T_I O(3)$ with $\text{Skew}(3)$.

For each $C \in \text{Skew}(3)$, let $X(C) : O(3) \rightarrow TO(3)$ be the vector field given by $X(C)_A = AC$ in $T_A O(3)$, for each $A \in O(3)$.

Exercise 5. Show that $X = X(C)$ is left invariant, in the sense that $\ell_* X = X$ for each diffeomorphism $\ell : O(3) \rightarrow O(3)$ given by left multiplication with elements $B \in O(3)$. ♠

Solution 5. The map ℓ is the restriction of a linear map, hence its derivative is given simply by multiplication by B . Then, by definition

$$(\ell_* X)_D = B(B^{-1}DC) = DC$$

But also $X_D = DC$. ♥

Take as known the fact that the series

$$e^C = \sum_{n=0}^{\infty} \frac{C^n}{n!}$$

defines a smooth map $\exp : M(3) \rightarrow M(3)$, satisfying $e^{(s+t)C} = e^{sC}e^{tC}$ for $s, t \in \mathbb{R}$ and $C \in M(3)$.

Exercise 6. Prove that \exp restricts to a smooth map $\exp : \text{Skew}(3) \rightarrow O(3)$, and compute the derivative at $t = 0$ of the curve $t \mapsto \exp(tC)$ through I in $O(3)$. ♠

Solution 6. It is enough to prove that $\exp|_{\text{Skew}(3)}$ factors through the submanifold $O(3) \subset M(3)$. But this is easy: if $C \in \text{Skew}(3)$, then $\exp(C)\exp(C)^T = \exp(C + C^T) = \exp(0) = I$, so the image lands in $O(3)$. The derivative is easily calculated:

$$\lim_{t \rightarrow 0} \frac{\exp(tC) - I}{t} = \lim_{t \rightarrow 0} (I + tC + \frac{t^2 C^2}{2} + \dots - I) = C.$$

♡

Exercise 7. Let $C \in \text{Skew}(3)$ and let $X = X(C)$ be the associated left invariant vector field with $X_I = C$. Show that the 1-parameter group of diffeomorphisms $\{\phi_t\}_t$ generated by X is given by

$$\phi_t(A) = A \exp(tC)$$

for $t \in \mathbb{R}$ and $A \in O(3)$. ♠

Solution 7. Since such a 1-parameter group of diffeomorphisms is uniquely defined by the vector field, it is enough to check that the map $\phi_t(A)$ satisfies the conditions.

First of all, the map $\mathbb{R} \times O(3) \rightarrow O(3)$ given by $(t, A) \mapsto A \exp(tC)$ is smooth, since a product of smooth functions is smooth.

Secondly, by the additive properties of the exponential, we have

$$\phi_{s+t}(A) = A \exp((s+t)C) = A \exp(sC) \exp(tC) = \phi_s(A) \exp(tC) = \phi_t(\phi_s(A)).$$

Thirdly, the tangent vector of the curve $t \mapsto \phi_t(C)$ is by the previous exercise simply C . ♡

Exercise 8. Let $D \in \text{Skew}(3)$ and let $Y = X(D)$ be the associated left invariant vector field with $Y_I = D$. Show that

$$(\phi_t)_*(Y)_A = A \exp(tC)^{-1} D \exp(tC).$$

for all $t \in \mathbb{R}$ and $A \in O(3)$. ♠

Solution 8. Note that ϕ_t is given by right-multiplication by $\exp(tC)$, hence the same is true for ϕ_{t*} . Now, by definitions:

$$\begin{aligned} (\phi_t)_*(Y)_A &= \phi_{*\phi_t^{-1}(A)}(Y_{\phi_t^{-1}(A)}) \\ &= \phi_{*\phi_{-t}(A)}(A \exp(-tC)D) \\ &= A \exp(-tC)D \exp(tC). \end{aligned}$$

Now the result is clear by a property of the exponential map. ♡

Exercise 9. Prove that the Lie derivative $L_X Y$ of $Y = X(D)$ with respect to $X = X(C)$ satisfies

$$L_{X(C)}(X(D)) = X([C, D])$$

for $C, D \in \text{Skew}(3)$, where the commutator $[C, D] = CD - DC$ is computed in $\text{Skew}(3) \subset M(3)$. ♠

Solution 9. We first compute the Lie derivative $(L_X Y)_A$ for $A \in O(3)$. This is a straight-forward calculation:

$$\begin{aligned} (L_X Y)_A &= A \lim_{h \rightarrow 0} \frac{1}{h} [D - \exp(tC)^{-1} D \exp(tC)] \\ &= A \lim_{h \rightarrow 0} \frac{1}{h} [D - D \exp(tC) + D \exp(tC) - \exp(tC)^{-1} D \exp(tC)] \\ &= A \lim_{h \rightarrow 0} \frac{1}{h} [D(I - \exp(tC)) + (1 - \exp(tC)^{-1}) D \exp(tC)] \\ &= A \left(D \lim_{h \rightarrow 0} \frac{1}{h} [I - \exp(tC)] + \lim_{h \rightarrow 0} \frac{1}{h} [I - \exp(-tC)] D \right) \\ &= A(-DC + CD) = A(CD - DC) = A[C, D] = X([C, D])_A. \end{aligned}$$

The statement now follows. ♡

Let $\xi : O(3) \rightarrow \mathbb{R}^3$ be given by $\xi(A) = (a_{23}, a_{13}, a_{12})$ where $A = (a_{ij})_{ij}$, so that $\xi(I) = 0 = (0, 0, 0)$.

Exercise 10. Show that $\xi_{*I} : T_I O(3) \rightarrow T_0 \mathbb{R}^3 \approx \mathbb{R}^3$ is an isomorphism. Explain why there is an open neighbourhood $U \subset O(3)$ of I with $\xi(U) \subset \mathbb{R}^3$ an open neighbourhood of 0 such that $\xi|_U \rightarrow \xi(U)$ is a diffeomorphism. ♠

Solution 10. The map ξ is (the restriction of) a linear map, hence its derivative is given by the same map. In particular, the natural basis for $\text{Skew}(3)$ is mapped to the standard basis vectors of \mathbb{R}^3 , so ξ_* is an isomorphism at the identity.

The other claim follows from the inverse function theorem. ♡

Let $x = \xi|_U$ so that (x, U) is a smooth chart on $O(3)$ near I .

Exercise 11. Compute the images C, D and $E \in \text{Skew}(3)$ (under k_{*I}) of the vectors

$$\left. \frac{\partial}{\partial x^1} \right|_I, \left. \frac{\partial}{\partial x^2} \right|_I, \left. \frac{\partial}{\partial x^3} \right|_I \in T_I O(3)$$

♠

Solution 11. By the explanation in the previous exercise, we have that these tangent vectors are mapped to the matrices below, in the same order:

$$C = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, D = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, E = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

♡

Exercise 12. Compute the commutators $[C, D], [D, E]$ and $[E, C]$ in $\text{Skew}(3)$. Determine the Lie derivatives $L_{X(C)}X(D), L_{X(D)}X(E)$ and $L_{X(E)}X(C)$. ♠

Solution 12. This is just matrix multiplication. We get that $[C, D] = E$, $[D, E] = C$ and $[E, C] = D$. Hence it follows from the formula above ($L_{X(C)}X(D) = X([C, D])$), that $L_{X(C)}X(D) = X(E)$, $L_{X(D)}X(E) = X(C)$ and $L_{X(E)}X(C) = X(D)$. ♡

Remark. Given a basis for a Lie algebra (in this case the algebra generated by C, D, E), the expression of the commutator in terms of the other three elements determine the structure constants of the Lie algebra. They are the coefficients of $[C, D]$ (etc.) as a linear combination of C, D, E .