

Chapter 2

Manifolds

In this chapter we give the ‘patchwork’ definition of manifolds. Manifolds are geometrically nice spaces and a natural generalisation of \mathbb{R}^n . The most common way to define a manifold is as a special type of topological space, namely a ‘second-countable Hausdorff locally-euclidean topological space with an atlas’¹. Because students usually encounter differential geometry before abstract topology, the lecturer then gives a speed run of all the definitions in topology. I think this approach is better suited to the second time you encounter manifolds. Then you already know a little bit about what makes manifolds nice, and you can appreciate the interesting but weird topological spaces that need to be excluded from the definition. The standard approach in effect defines a manifold by saying what a manifold isn’t. In the approach below we avoid defining general topological spaces and instead use concrete gluing construction of open sets of \mathbb{R}^n . After this construction we will still need to impose certain conditions, so topology cannot be avoided completely, but hopefully they are suitable for a new-comer to manifolds.

Before we dive into theory, we define a concept that you probably know but have never had a word for. A *partial function* from X to Y is a function from a subset $S \subset X$ to Y . In the context of partial functions, a function with $S = X$ is called a *total function*. Many common functions are really partial functions: $\frac{1}{x}$ and \sqrt{x} are partial functions from \mathbb{R} to \mathbb{R} , with S being respectively $\mathbb{R} \setminus \{0\}$ and $[0, \infty)$. There doesn’t seem to be good standard terminology to talk about X and S , though S is often called the natural domain. Let’s call X the *source* of the function and S the *domain*, with the symbols $\text{src } f = X$ and $\text{dom } f = S$. You are no doubt familiar with the difference between the *codomain* $\text{codom } f = Y$ (also called the target) and the *image* $\text{img } f = f[S]$ (also called the range). We will use $f : X \dashrightarrow Y$ for partial functions (harpoon arrow), in contrast to $f : X \rightarrow Y$ for total functions.

Many students in Analysis I are confused about the relationship between injective and surjective, and those students are correct to be confused. Just as surjective means that the image is equal to the codomain, total means that the domain is equal to the source; they are the true counterparts to one another. In fact a partial function that is injective has an inverse partial function. If $f : X \dashrightarrow Y$ is injective then

$$f^{-1}(y) = x \text{ if } y = f(x)$$

is a perfectly valid definition of a partial function $f^{-1} : Y \dashrightarrow X$ with $\text{dom } f^{-1} = \text{img } f$ and

¹Sources differ: some use second-countable, others use the ‘Lindelöf’ property, which is equivalent in this context, and others use the more general ‘paracompact’

$\text{img } f^{-1} = \text{dom } f$.

Besides inverses, many of the usual definitions for functions carry over with sensible modifications. For example, partial functions from X to Y are equal if they have the same domains and are equal on all inputs. Likewise the composition of two partial functions $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ is a partial function $g \circ f : X \rightarrow Z$, but the domain of $g \circ f$ will be smaller than of the domain of f if the image of f lies partly outside the domain of g . We should also think about the ‘empty’ partial function. According to the set theory definition of a function, there is exactly one function from the empty set to any other set. This is usually considered a quirk in the definition. But the composition of partial functions where $\text{img } f \cap \text{dom } g = \emptyset$ results in the empty function, so we cannot ignore it.

Remark 2.1. In this course, functions and partial functions will always have open sets as their sources and their domains will be open subsets, unless specifically noted otherwise. This is needed so that derivatives can be defined.

2.1 Manifolds

Fix a dimension n . To keep track of the pieces we will glue together, let us introduce an index set \mathcal{I} . In our examples this will usually be a finite set, but we do not make this assumption generally. For every $i \in \mathcal{I}$ let U_i be an open subset of \mathbb{R}^n . U_i is called a *chart*. These are the pieces we will glue together.

For two charts, we describe how to glue them together using a partial function. A gluing function $\varphi_{ij} : U_i \rightarrow U_j$ is a partial function that is a homeomorphism from its domain to its image. This means it is a homeomorphism (bijective, continuous, continuous inverse) between open subsets $V_i \subset U_i$ and $V_j \subset U_j$. The idea is that the point $x \in V_i$ is glued to $\varphi_{ij}(x) \in V_j$. Note the order of the subscripts: “from i to j ”. The points of $U_i \setminus V_i$ are not glued to any points of U_j . Other names for gluing functions include ‘transition functions’, ‘change of coordinate functions’, and ‘overlap functions’. We allow here the possibility that $V_i = V_j = \emptyset$ and $\varphi_{ij} : \emptyset \rightarrow \emptyset$ is the empty function, this represents the situation that U_i and U_j are not glued together at all. To avoid having too many named sets, we will mostly use $\text{dom } \varphi_{ij}$ instead of V_i .

This information tells how to glue the pieces together, but how should we represent the completed glued object? First we define the *disjoint union*

$$\coprod_{i \in \mathcal{I}} U_i = \{(i, x) \mid i \in \mathcal{I}, x \in U_i\},$$

which is a set of pairs. We think of this as saying that even if a point is common to both U_i and U_j , in the disjoint union we consider it as two separate points. For example, if $U_1 = (-1, 1)$ and $U_2 = (0, 2)$ then the normal union is $U_1 \cup U_2 = (-1, 2)$ but the disjoint union is two intervals. We often do not write the index i if it is clear, and even when it is not clear we tend to write it as a subscript. Continuing the example, $U_1 \coprod U_2$ has two points that might both be called 0.5, namely $0.5_1 \in U_1$ and $0.5_2 \in U_2$. Formally these points should be written $(1, 0.5)$ and $(2, 0.5)$ respectively.

We want to create an equivalence relation on the disjoint union of all the charts such that $x \in U_i \sim y \in U_j$ iff $\varphi_{ij}(x) = y$. If this is to be an equivalence relation, the set of gluing

functions is required to have certain properties. To get reflexivity of \sim , we need $\varphi_{ii} = \text{id}_{U_i}$. Symmetry of the relation holds if and only if $\varphi_{ji} = \varphi_{ij}^{-1}$. These are simple enough, but expressing the condition for transitivity is more difficult. The usual way to express the transitivity condition is that $y \sim x, x \sim z \Rightarrow y \sim z$, but if we have symmetry then this is equivalent to $x \sim y, x \sim z \Rightarrow y \sim z$. That means for all $x \in \text{dom } \varphi_{ij} \cap \text{dom } \varphi_{ik}$ we need $y = \varphi_{ij}(x) \in \text{dom } \varphi_{jk}$ and $z = \varphi_{ik}(x) = \varphi_{jk}(\varphi_{ij}(x))$. In the language of partial functions

$$\varphi_{ik} = \varphi_{jk} \circ \varphi_{ij}.$$

If a set of gluing functions have these three properties, and thus defines an equivalence relation, we say that fulfill the *cocycle conditions*.

Definition 2.2. An atlas is a tuple $\mathcal{A} = (n, \mathcal{I}, \{U_i\}_{i \in \mathcal{I}}, \{\varphi_{ij}\}_{i, j \in \mathcal{I}})$, where $\{\varphi_{ij}\}_{i, j \in \mathcal{I}}$ fulfills the cocycle conditions. We have included n in our definition of atlas to make all the charts have the same dimension (some authors allow spaces with a mix of dimensions). We call $M = \coprod U_i / \sim$ the glued space.² Each ‘point’ in M is an equivalence class of points in different U_i .

Example 2.3 (Euclidean Space). Take any open subset $U \subset \mathbb{R}^n$. We can construct the trivial atlas for U as follows. Let the index set $\mathcal{I} = \{0\}$ and $U_0 = U$ the only chart. Then $\varphi_{00} = \text{id}_U$ is a gluing function. The cocycle condition is fulfilled, so $(n, \mathcal{I}, \{U\}, \{\varphi_{00}\})$ is an atlas. The corresponding equivalence relation \sim is the weakest equivalence relation on U , namely x is equivalent to itself but no other points. Therefore we say $M = U$.

Example 2.4 (Polar Coordinates). Next we consider an example with two charts: $\mathcal{I} = \{1, 2\}$. Consider the plane $U_1 = \mathbb{R}^2$ and the half-strip $U_2 = (0, \infty) \times (-\pi, \pi) \subset \mathbb{R}^2$. To give an atlas it is sufficient to describe φ_{21} , because the cocycle condition requires $\varphi_{11} = \text{id}_{U_1}$ and $\varphi_{22} = \text{id}_{U_2}$, as well as $\varphi_{12} = \varphi_{21}^{-1}$. Set

$$\varphi_{21} : U_2 \rightarrow U_1, \quad (r, \theta) \mapsto (r \cos \theta, r \sin \theta).$$

This is a homeomorphism from U_2 to $\text{img } \varphi_{21} = \mathbb{R}^2 \setminus \{(x, 0) \mid x \leq 0\}$. Therefore $\mathcal{A} = (2, \mathcal{I}, \{U_1, U_2\}, \{\varphi_{11}, \varphi_{12}, \varphi_{21}, \varphi_{22}\})$ is an atlas. This example shows why φ_{ij} is sometimes called a change of coordinates function.

Every point of U_2 is glued to some point of U_1 , but not every point of U_1 is glued to some point of U_2 . This leads to the points of the glued space M are of two types: either they are an equivalence class with two points $\{(r \cos \theta, r \sin \theta), (r, \theta)\}$ for $(r \cos \theta, r \sin \theta) \in U_1$ and $(r, \theta) \in U_2$, or they are an equivalence class with a single point $\{(x, 0)\}$ for $(x, 0) \in U_1$ with $x \leq 0$. Every equivalence class in M contains an element of U_1 , so we might say casually $M = U_1$.

Example 2.5 (Glued Circle). Consider $n = 1$, $\mathcal{I} = \{1, 2\}$ and $U_1 = (-1, 1)$, $U_2 = (-1, 1)$. As in the previous example, it is sufficient to give φ_{12} . Set $V_1 = U_1 \setminus \{0\}$ and $V_2 = U_2 \setminus \{0\}$ and give the gluing function $\varphi_{12} : V_1 \rightarrow V_2$ by the formula

$$\varphi_{12}(x) = \begin{cases} x + 1 & \text{for } x \in (-1, 0), \\ x - 1 & \text{for } x \in (0, 1). \end{cases}$$

²If you are familiar with topology, the quotient topology makes M a locally euclidean topological space.

For example, this tells us that we should glue $-0.5_1 \in U_1$ to $\varphi_{12}(-0.5) = 0.5_2 \in U_2$ and $0.3_1 \in U_1$ to $\varphi_{12}(0.3) = -0.7_2 \in U_2$. Here the glued space M is a circle, which you can see by cutting two strips of paper of the same length, drawing a number line from -1 to 1 on each of them, and then gluing as instructed. Every point of M is equivalent to a point of U_1 or to 0_2 .

Example 2.6. Consider everything the same as in Example 2.5, but this time give φ_{12} the formula

$$\varphi_{12}(x) = x, \quad x \in (-1, 0) \cup (0, 1).$$

The space M is called an interval with two origins. This is because every point of M is either $\{x_1, x_2\}$ for $x \neq 0$, or 0_1 , or 0_2 .

The interval with two origins might seem like a harmless curiosity, but in fact it is a weird topological space that we want to avoid. Let us say that a sequence p_n in M converges to a point p if the sequence $\phi_i(p_n)$ converges to $\phi_i(p)$ in the chart U_i . In Example 2.6 consider the sequence n^{-1} for $n \in \mathbb{N}^+$. We can view this sequence in U_1 or U_2 . In U_1 it has the limit 0_1 , but in U_2 it has the limit 0_2 . Therefore in M this sequence has two different limits!

There are **other ways** that this space is badly behaved. It is not a metric space, because the distance between the two origins is zero. So although every point has a neighbourhood that is homeomorphic to an open subset of euclidean space (locally euclidean), it is not really like euclidean space. We therefore want to exclude such pathological spaces. It turns out that the above example is the only way a gluing can cause non-unique limits.

Lemma 2.7 (Non-unique Limits). *A glued space M has non-unique limits if and only if there is a sequence $x_n \in \text{dom } \varphi_{12} \subset U_1$ such that x_n converges to a point $x \in U_1 \setminus \text{dom } \varphi_{12}$ and $\varphi_{12}(x_n)$ converges to a point $y \in U_2$, for some charts U_1, U_2 .*

Proof. Suppose first we have a sequence $x_n \in \text{dom } \varphi_{12} \subset U_1$ such that x_n converges to a point $x \in U_1 \setminus \text{dom } \varphi_{12}$ and $\varphi_{12}(x_n)$ converges to a point $y \in U_2$. By definition, x and y are both limits of the sequence x_n . But x is outside the domain of the gluing function, so by definition it is not glued to any point of U_2 . Therefore x and y are distinct points in M .

Conversely suppose a glued space M has a sequence with two distinct limits. In euclidean space limits are unique, so if there are two limits then they must come from two charts U_1 and U_2 . Let's use the notation $x_n, x \in U_1$ and $y_k, y \in U_2$ with $y_n = \varphi_{12}(x_n)$. Suppose x were in the domain of the gluing function. The gluing function by definition is a continuous function, so we would have $\varphi_{12}(x) = \lim \varphi_{12}(x_n) = \lim y_n = y$. But $\varphi_{12}(x) = y$ means exactly that $x \sim y$ and this contradicts our assumption that the limits are distinct in M . Therefore x is not in the domain of φ_{12} , but it is the limit of a sequence in the domain. We have shown that $x \in U_1 \setminus \text{dom } \varphi_{12}$. \square

Definition 2.8. *We say that an atlas has the unique limit property if it does not satisfy the condition of Lemma 2.7. That is, if for every $i, j \in \mathcal{I}$ and every sequence $x_n \in \text{dom } \varphi_{ij}$ that converges to some $x \in U_1 \setminus \text{dom } \varphi_{ij}$, the sequence $\varphi_{ij}(x_n)$ does not converge in U_2 .*

Observe why this is not an issue with Example 2.5. Consider the sequence $x_n = n^{-1}$ in U_1 . It has the limit 0_1 . On the other hand, the sequence $y_n = \varphi_{12}(x_n) = n^{-1} - 1$ converges to -1 in \mathbb{R} , but -1 is not in U_2 . Therefore this 'other' limit point is not in the space M .

The other way that gluing can produce a topologically bad space is if we glue too many charts together. We will not provide an example of this; interested students may search for the ‘**long line**’ or the ‘long ray’, which are standard examples of this phenomenon.

Definition 2.9. A manifold is an atlas with the unique limit property and such that the index set is countable.

Exercise 2.10. For students who know topology: show that this is equivalent to the standard definition of a manifold.

It is very common to talk about the glued space M as the manifold without explicitly stating the atlas. This is similar to talking about a vector space as the set of vectors, when in fact it is the operations of addition and scalar multiplication that make a vector space interesting.

There are different sorts of manifolds, based on additional conditions on the gluing functions. We will use the notation that a function is \mathcal{C}^ℓ when it is ℓ -times continuously differentiable. By convention, \mathcal{C}^0 means that the function is continuous, and \mathcal{C}^∞ means that the function is smooth. An atlas (or a manifold) is called \mathcal{C}^ℓ when all of the gluing functions are \mathcal{C}^ℓ . Probably the most common type of manifold that is studied, and the one we will study in this course, are *smooth* manifolds. Henceforth, when we say manifold we mean smooth manifold.

Because we have given a non-standard definition of manifolds, we should explain how this compares to the standard definition. We do this using the example of a sphere and stereographic projection, which seems to be the first non-trivial example in every book on manifolds.

Example 2.11 (Stereographic Projection). The circle is commonly defined as the set $\mathbb{S}^1 = \{p \in \mathbb{R}^2 \mid \|p\| = 1\}$, and we name the north pole $N = (0, 1)$ and the south pole $S = (0, -1)$. If one draws chooses a point p other than the north pole, and considers the line through p and N , there is a unique intersection point $\phi_N(p)$ on the x -axis. This is a bijective function $\phi_N : \mathbb{S}^1 \rightarrow \mathbb{R}$ called stereographic projection. A nice geometry exercise using similar triangles gives the formula

$$\phi_N : (p^1, p^2) \mapsto \frac{p^1}{1 - p^2}.$$

Likewise stereographic projection from the south pole is

$$\phi_S : \mathbb{S}^1 \setminus \{S\} \rightarrow \mathbb{R}, \quad (p^1, p^2) \mapsto \frac{p^1}{1 + p^2}.$$

Notice that these constructions are ill-defined when applied to the pole itself, because a single point does not determine a line, so it is naturally a partial function on the circle.

The inverses of these functions,

$$\phi_N^{-1} : \mathbb{R} \rightarrow \mathbb{S}^1 \setminus \{N\}, \quad x \mapsto \frac{1}{x^2 + 1}(2x, x^2 - 1),$$

and

$$\phi_S^{-1} : \mathbb{R} \rightarrow \mathbb{S}^1 \setminus \{S\}, \quad y \mapsto \frac{1}{y^2 + 1}(2y, 1 - y^2),$$

are regular parametrisations of (parts of) the circle in the sense of Section 1.1. So given a point p on the sphere, we can apply ϕ_N to get a point x in \mathbb{R} , called its coordinate with

respect to ϕ_N , and putting the coordinate into the parameterisation ϕ_N^{-1} gives back the point p . If we know the coordinate x with respect to ϕ_N then $y = \phi_S \circ \phi_N^{-1}(x)$ is the coordinate with respect to ϕ_S . For this reason $\phi_S \circ \phi_N^{-1} : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R} \setminus \{0\}$ is the change of coordinates function. In this example, one can calculate

$$\phi_S \circ \phi_N^{-1}(x) = \phi_S \left(\frac{2x}{x^2 + 1}, \frac{x^2 - 1}{x^2 + 1} \right) = \frac{\frac{2x}{x^2 + 1}}{1 + \frac{x^2 - 1}{x^2 + 1}} = \frac{2x}{x^2 + 1 + x^2 - 1} = \frac{1}{x}.$$

Notice that it is not possible to cover the circle by a single regular parameterisation (parameterisations must be injective). Thus to deal with spaces generally we must consider multiple overlapping parameterisations and change of coordinate functions.

In this example we have presented the standard approach to manifolds through coordinate functions and parameterisations. Let us connect it now with our definition. The index set is $\mathcal{I} = \{N, S\}$. The image of the coordinate functions are the charts, $U_N = \mathbb{R}$ and $U_S = \mathbb{R}$. And the gluing function from $U_N \setminus \{0\}$ to $U_S \setminus \{0\}$ is $\varphi_{NS}(x) = x^{-1}$. This data forms an atlas.

But what does this atlas have to do with the circle? We indeed see that the equivalence classes of points are $S = \{0_N\}$, $N = \{0_S\}$, or $\{x, y\}$ for $x \in U_N \setminus \{0\}$ and $y = x^{-1} \in U_S \setminus \{0\}$. That is, there are in one-to-one correspondence with the points of \mathbb{S}^1 . The glued space M is a space that has all the same points as \mathbb{S}^1 and the same topology, but it is not a subset of a euclidean space. Whatever we can describe on M must be intrinsic to the circle, independent of its relationship to \mathbb{R}^2 .

Exercise 2.12 (Stereographic Projection). Repeat the above construction for the n -sphere $\mathbb{S}^n = \{p \in \mathbb{R}^{n+1} \mid \|p\| = 1\}$. We choose³ our coordinates on \mathbb{R}^{n+1} such that the north pole is $N = (0, \dots, 0, 1)$ and the south pole is $S = (0, \dots, 0, -1)$. Show that stereographic projection have the formulas

$$\begin{aligned} \phi_N : \mathbb{S}^n \setminus \{N\} &\rightarrow \mathbb{R}^n, & p &\mapsto \frac{1}{1 - p^{n+1}}(p^1, \dots, p^n), \\ \phi_S : \mathbb{S}^n \setminus \{S\} &\rightarrow \mathbb{R}^n, & p &\mapsto \frac{1}{1 + p^{n+1}}(p^1, \dots, p^n), \end{aligned}$$

and inverse stereographic projection is

$$\begin{aligned} \phi_N^{-1} : \mathbb{R}^n &\rightarrow \mathbb{S}^n \setminus \{N\}, & x &\mapsto \frac{1}{\|x\|^2 + 1}(2x^1, \dots, 2x^n, \|x\|^2 - 1), \\ \phi_S^{-1} : \mathbb{R}^n &\rightarrow \mathbb{S}^n \setminus \{S\}, & y &\mapsto \frac{1}{\|y\|^2 + 1}(2y^1, \dots, 2y^n, 1 - \|y\|^2). \end{aligned}$$

Finally, show that the transition function is

$$\varphi_{NS}(x) = \phi_S \circ \phi_N^{-1}(x) = \|x\|^{-2}x.$$

In the standard definition of a manifold you begin with M and coordinate functions $\phi_i : M \rightarrow \mathbb{R}^n$. The procedure in Example 2.11 to go from this data to our definition of a manifold is fully general: define the charts to be $U_i = \text{img } \phi_i$ and the gluing functions as $\varphi_{ij} = \phi_j \circ \phi_i^{-1}$.

³Another common convention is to zero-index the components the have the ± 1 in the zeroth component

Conversely, if you begin with a manifold in our sense then M is the glued space $\coprod U_i / \sim$. There is of course the canonical projection $\pi_M : \coprod U_i \rightarrow M$ that sends every element to its equivalence class, but this is too coarse. We call the restriction $\Phi_i := \pi_M|_{U_i} : U_i \rightarrow M$ a parameterisation of $\pi_M[U_i] \subset M$. It sends a point of U_i to its equivalence class in M . We really should think of this as a parameterisation because an ordinary set U_i in euclidean space is describing part of a complicated object M . In the other direction $\phi_i := (\Phi_i)^{-1} : \pi_M[U_i] \rightarrow U_i \subset \mathbb{R}^n$ is the coordinate function. It sends an equivalence class to its representative that lies in U_i . We usually write ϕ_i^{-1} for the parameterisation rather than Φ_i , as it is unnecessary to have two symbols. Thus the two approaches to manifolds give equivalent information.

Let us summarise our terminology and the relations between the objects. Charts are open sets of \mathbb{R}^n . An atlas is set of charts and gluing functions with the cocycle property. A gluing function can also be called a change of coordinate function or a transition function. The information of an atlas allows us to glue the charts together to get a manifold M . Functions from the charts to the manifold are called parametrisations and functions from the manifold to the charts are coordinates. The composition of a parameterisation and a coordinate is a transition function. One of the drawbacks of differential geometry being an old and widely practised field is that notation and terminology has been around for a long time and is not completely standardised. Different authors use the word chart to describe ϕ_i, ϕ_i^{-1}, U_i , or $\phi_i^{-1}[U_i]$.

2.2 Functions

Next we want to define functions between manifolds. On one hand, there is nothing to do. If we have manifolds M and \tilde{M} , they are sets, and a function $f : M \rightarrow \tilde{M}$ is defined in the normal way. But manifolds are more than sets of points, they have atlases. Let

$$\mathcal{A} = (n, \mathcal{I}, \{U_i\}_{i \in \mathcal{I}}, \{\varphi_{i,j}\}_{i,j \in \mathcal{I}}), \quad \tilde{\mathcal{A}} = (\tilde{n}, \tilde{\mathcal{I}}, \{\tilde{U}_i\}_{i \in \tilde{\mathcal{I}}}, \{\tilde{\varphi}_{i,j}\}_{i,j \in \tilde{\mathcal{I}}})$$

be atlases for M and \tilde{M} respectively. Then we can look at f ‘in charts’. This means we look at the partial functions

$$f_{ik} := \tilde{\phi}_k \circ f \circ \phi_i^{-1} : U_i \rightarrow \tilde{U}_k \text{ for } i \in \mathcal{I} \text{ and } k \in \tilde{\mathcal{I}}.$$

These are functions between euclidean spaces, so we can ask whether they are \mathcal{C}^ℓ .

Definition 2.13. *A function $f : M \rightarrow \tilde{M}$ is \mathcal{C}^ℓ at a point $p \in M$ if there is a chart $U_i \ni \phi_i(p)$ and chart $\tilde{U}_k \ni \tilde{\phi}_j(f(p))$ such that f_{ik} is \mathcal{C}^ℓ , and it obeys a further technical condition.⁴ We say that f is \mathcal{C}^ℓ on M if it is \mathcal{C}^ℓ at every point of M .*

The definition of \mathcal{C}^ℓ uses a chart. But we know that a point may belong to more than one chart. This opens the possibility that f is \mathcal{C}^ℓ at p according to one chart, but not \mathcal{C}^ℓ according to another chart. However, because of the relation

$$\begin{aligned} f_{jl} &= \tilde{\phi}_l \circ f \circ \phi_j^{-1} = (\tilde{\phi}_l \circ \tilde{\phi}_k^{-1} \circ \tilde{\phi}_k) \circ f \circ (\phi_i^{-1} \circ \phi_i \circ \phi_j^{-1}) \\ &= \tilde{\varphi}_{k,l} \circ (\tilde{\phi}_k \circ f \circ \phi_i^{-1}) \circ \varphi_{j,i} = \tilde{\varphi}_{k,l} \circ f_{ik} \circ \varphi_{j,i}, \end{aligned}$$

the definition does not depend on the which charts are used.

Example 2.14 (Euclidean Space). We have already seen that open subsets of euclidean space are manifolds in a particularly simple way, namely the coordinates and parametrisations are the identity function. Therefore a function in charts is the same thing as a function. This shows that a function between euclidean spaces is \mathcal{C}^ℓ according to the manifold definition if and only if it is \mathcal{C}^ℓ according to the ordinary definition.

Example 2.15 (Stereographic Projection). Consider the circle \mathbb{S}^1 and the function $f : \mathbb{S}^1 \rightarrow \mathbb{R}$ given by $f(p^1, p^2) = p^2$. This is the height function. We can look at this function in charts

$$f_{N0} : U_N \rightarrow \mathbb{R} \quad f_{N0}(x) = \tilde{\phi}_0 \circ f \circ \phi_N^{-1}(x) = \text{id}_{\mathbb{R}} \circ f \left(\frac{2x}{x^2+1}, \frac{x^2-1}{x^2+1} \right) = \frac{x^2-1}{x^2+1},$$

and likewise

$$f_{S0} : U_S \rightarrow \mathbb{R} \quad f_{S0}(y) = \frac{1-y^2}{y^2+1}.$$

It may help to understand this if we calculate it for a few points. Consider the south pole $S = (0, 0, -1)$, which has a height $f(0, 0, -1) = -1$. In ϕ_N -coordinates the south pole is $0 \in U_N$ and $f_{N0}(0) = -1$ as expected. Now consider the point $p = (0.65, 0.76) \in \mathbb{S}^2$. Clearly

⁴It is necessary to require that $\phi_i(p)$ has an open neighbourhood $U \subset U_i$ such that $f[U] \subset \tilde{U}_j$. If you omit this condition, it is possible to make an example where f is continuous in every chart but is not continuous on M as a whole. See Lee ‘Introduction to Smooth Manifolds’ Problem 2-1 for an example.

is has $f(p) = 0.76$. In ϕ_N -coordinates it is $2.7 \in U_N$ and in ϕ_S -coordinates it is $0.37 \in U_S$. So then we compute

$$f_{N0}(2.7) = \frac{7.3 - 1}{7.3 + 1} = 0.76, \quad f_{S0}(0.37) = \frac{1 - 0.137}{0.137 + 1} = 0.76.$$

In conclusion, f in charts is nothing other than a manipulation of the formula for f to use coordinates; it gives the same result.

Perhaps not unexpectedly, this is a smooth function in the sense of manifolds because both f_{N0} and f_{S0} are smooth functions in the usual sense.

Example 2.16 (Glued Circle). We can give a function f that *embeds* M into euclidean space \mathbb{R}^2 . It is easier to write the formulas in charts. We define

$$\begin{aligned} f_{11}(x) &= (\cos \pi x, \sin \pi x), \\ f_{21}(y) &= (\cos \pi(y + 1), \sin \pi(y + 1)). \end{aligned}$$

We can see that this is well-defined in the following way. Take a point $x \in \text{dom } \varphi_{12}$. Then

$$\begin{aligned} f_{21}(\varphi_{12}(x)) &= (\cos \pi(\varphi_{12}(x) + 1), \sin \pi(\varphi_{12}(x) + 1)) \\ &= \begin{cases} (\cos \pi(x + 1 + 1), \sin \pi(x + 1 + 1)) & \text{for } x \in (-1, 0) \\ (\cos \pi(x - 1 + 1), \sin \pi(x - 1 + 1)) & \text{for } x \in (0, 1) \end{cases} \\ &= \begin{cases} (\cos(\pi x + 2\pi), \sin(\pi x + 2\pi)) & \text{for } x \in (-1, 0) \\ (\cos \pi x, \sin \pi x) & \text{for } x \in (0, 1) \end{cases} \\ &= (\cos \pi x, \sin \pi x) \\ &= f_{11}(x). \end{aligned}$$

What we have shown is that if $x \in U_1 \sim y \in U_2$ then $f_{11}(x) = f_{21}(y)$. This means that it doesn't matter which chart you use, the result is the same. In other words we have defined a function f that doesn't depend on charts, f is defined on M .

This example shows us how to embed Example 2.5 into \mathbb{R}^2 to get what we would normally think of as the circle. But please keep in mind that manifolds are defined as the gluing of charts; they are defined as an abstract space that does not need live in a bigger space. There are many ways to embed the circle into euclidean space. Even when we define a manifold starting with a subset of euclidean space, we leave the embedding behind.

2.3 Vectors

Our definition of manifolds makes it easy to define (tangent) vectors and vector fields. At any point $x \in U_1 \subset \mathbb{R}^n$ we have the tangent vectors $\{(v^1, \dots, v^n) \in \mathbb{R}^n\}$ and a vector field on U_1 is a function $X : U_1 \rightarrow \mathbb{R}^n$. To make this into a definition on a manifold, however, we need a way to make this independent of the chart. Alternatively, we need a way to compare vectors that are defined using different charts. There are essentially two equivalent ways to do this: using curves and using directional derivatives. We first give an example using curves as motivation.

Example 2.17 (Polar Coordinates). Consider the plane \mathbb{R}^2 with cartesian coordinates (x^1, x^2) , but also polar coordinates $(y^1, y^2) = (r, \theta)$, as in Example 2.4. The transition function from polar to cartesian is $\varphi(r, \theta) = (r \cos \theta, r \sin \theta)$. We consider vectors at the point p with $(r, \theta) = (1/\sqrt{2}, \pi/4)$, which is equivalent to $(x^1, x^2) = (0.5, 0.5)$.

Let $\alpha(t) = (t + 1/\sqrt{2}, \pi/4)$, a curve in U_2 with $\alpha(0) = p$. In the U_2 chart it is a horizontal line with tangent vector $\alpha'(0) = (1, 0)$. But we can also view this curve in the U_1 chart. It has the formula

$$\beta(t) = \varphi_{21}(\alpha(t)) = \left((t + \frac{1}{\sqrt{2}}) \cos \frac{\pi}{4}, (t + \frac{1}{\sqrt{2}}) \sin \frac{\pi}{4} \right) = \left(\frac{1}{\sqrt{2}}t + \frac{1}{2}, \frac{1}{\sqrt{2}}t + \frac{1}{2} \right).$$

At p it has the tangent vector $\beta'(0) = (1/\sqrt{2}, 1/\sqrt{2})$.

But these two charts represent the same points of manifold; the curves α and β have the same points under the gluing equivalence relationship. Likewise we should think of the two tangent vectors, $v = \alpha'(0) = (1, 0)$ at $(1/\sqrt{2}, \pi/4)$ in the U_2 chart and $w = \beta'(0) = (1/\sqrt{2}, 1/\sqrt{2})$ at $(0.5, 0.5)$ in the U_1 chart, as equivalent.

Both methods have the same setup. Let $U_1, U_2 \subset \mathbb{R}^n$ be two charts and $\varphi : U_1 \rightarrow U_2$ the transition (we leave off the subscripts for this explanation to simplify notation). Let x be a point in $\text{dom } \varphi$, $y = \varphi(x)$, and $v = (v^1, \dots, v^n)$ be a vector on U_1 and $w = (w^1, \dots, w^n)$ be a vector on U_2 .

Curve method: Consider the curve $\alpha(t) = x + vt \in U_1$. This curve has $\alpha(0) = x$ and $\alpha'(0) = v$. Using the transition, we also have a curve $\beta = \varphi \circ \alpha$ in U_2 with $\beta(0) = \varphi(x) = y$. The idea is that v in the first chart is transformed to $w = \beta'(0)$ in the second chart. Using the chain rule

$$w = \beta'(0) = (J_x \varphi) \alpha'(0) = (J_x \varphi) v,$$

where $J_x \varphi$ is the matrix of partial derivatives of φ , also called the Jacobian matrix, evaluated at the point x .

Directional derivative method: Take any function $h : U_2 \rightarrow \mathbb{R}$. We can use the transition function to write h as a function on U_1 , at least near x , namely $h \circ \varphi$. We compute the derivative in the direction of v at the point x and apply the chain rule

$$\sum_{j=1}^n v^j \frac{\partial (h \circ \varphi)}{\partial x^j}(x) = \sum_{i,j=1}^n v^j \frac{\partial h}{\partial y^i}(\varphi(x)) \frac{\partial y^i}{\partial x^j}(x) = \sum_{i,j=1}^n v^j \frac{\partial h}{\partial y^i}(y) (J_x \varphi)_{ij} = \sum_{i=1}^n (J_x \varphi v)_i \frac{\partial h}{\partial y^i}(y).$$

We see that the derivative of $h \circ \varphi$ at x in the direction of v is equal to the derivative of h at y in the direction $w = J_x \varphi v$. From both of these methods we get the same answer:

Definition 2.18. The vector v at $x \in U_1$ is equivalent to the vector w at $y \in U_2$ when $w = J_x \varphi v$. This formula is called the change of coordinates for vectors. We denote the set of equivalence classes of vectors at a point $p \in M$ by $T_p M$, called the tangent space to M at p .

Because of the second method, it is common to write a tangent vector as $\sum_{i=1}^n v^i \frac{\partial}{\partial x^i} \Big|_p$. This notation has two advantages. First that the change of coordinates is built into the notation via the chain rule, as above. And second, if we leave out the point at which we should evaluate the derivative then $\frac{\partial}{\partial x^i}$ is a vector field. In fact it gives a basis of the vector fields. Every vector field on U_1 can be written as a function

$$X : x \mapsto \sum_{i=1}^n v^i(x) \frac{\partial}{\partial x^i}$$

for functions $v^i : U_1 \rightarrow \mathbb{R}$. Because vector fields thought of this way can be evaluated at a point to give a tangent vector as well as act on a function, the notation $X(x)$ is potentially ambiguous. We will use $X|_p$ for evaluation and $X(f)$ for action on a function.

Example 2.19 (Stereographic Projection). Let's see how the coordinate vector fields transform between the two charts of stereographic projection on \mathbb{S}^1 . In particular $y = \varphi_{NS}(x) = x^{-1}$ has the Jacobian matrix

$$J\varphi = \left(\frac{\partial y}{\partial x} \right) = (-x^{-2}).$$

Therefore

$$\frac{\partial}{\partial x} = \frac{\partial y}{\partial x} \frac{\partial}{\partial y} = -x^{-2} \frac{\partial}{\partial y} = -y^2 \frac{\partial}{\partial y}.$$

We should note that this equality, which is really equivalence of vectors as per Definition 2.18, holds on the overlap between the two charts, namely away from the north and south poles. In particular, the vector $\frac{\partial}{\partial x}$ is not defined at the north pole, this expression only has meaning on the chart U_N . If we were to define a vector field by the formula

$$X = \begin{cases} \frac{\partial}{\partial x} & \text{for } x \in U_N, \\ -y^2 \frac{\partial}{\partial y} & \text{for } y \in U_S \end{cases}$$

then this is a well-defined vector field on all of \mathbb{S}^1 , because it gives a vector at every point and on overlaps the two cases give equivalent vectors. Observe that $X|_{y=0}$ is the zero vector.

On the other hand, if we consider the vector field $x^3 \frac{\partial}{\partial x}$ on U_1 then it corresponds to the vector field

$$x^3 \frac{\partial}{\partial x} = (y^{-3}) \left(-y^2 \frac{\partial}{\partial y} \right) = -y^{-1} \frac{\partial}{\partial y}$$

on $U_2 \setminus \{y = 0\}$. This has no continuous extension to all of \mathbb{S}^1 .

The main lesson of this example is that what looks like the 'same vector' at two different points in one chart look like completely different vectors in another chart. From the above example,

$\frac{\partial}{\partial x}\Big|_{x=1}$ and $\frac{\partial}{\partial x}\Big|_{x=2}$ look the same in the U_N chart, but in the U_S chart these two vectors are

$$-\frac{\partial}{\partial y}\Big|_{y=1} \quad \text{and} \quad -\frac{1}{4}\frac{\partial}{\partial y}\Big|_{y=0.5}$$

respectively. This is because the Jacobian $J_x\varphi$ has a dependence on x , on the point in the manifold. Hence the equivalence relation is different at different points. The consequence is that there is no easy way to identify tangent vectors at different points of a manifold. We will examine how this problem can be overcome in the next chapter, which concerns ‘connections’ between different points of a manifold.

Example 2.20 (Stereographic Projection). Let us continue Example 2.19 to see that equivalent vectors give the same directional derivatives. Consider how the vector field

$$X = \begin{cases} \frac{\partial}{\partial x} & \text{for } x \in U_N, \\ -y^2 \frac{\partial}{\partial y} & \text{for } y \in U_S \end{cases}$$

acts on the height function f from Example 2.15. In the chart U_N we have

$$(Xf)_N = \frac{\partial f_N}{\partial x} = \frac{\partial}{\partial x} \left(\frac{x^2 - 1}{x^2 + 1} \right) = \frac{2x(x^2 + 1) - 2x(x^2 - 1)}{(x^2 + 1)^2} = \frac{4x}{(x^2 + 1)^2}.$$

Likewise in the chart U_S we have

$$(Xf)_S = -y^2 \frac{\partial f_S}{\partial y} = -y^2 \frac{\partial}{\partial y} \left(\frac{1 - y^2}{y^2 + 1} \right) = -y^2 \left(-\frac{4y}{(y^2 + 1)^2} \right) = \frac{4y^3}{(y^2 + 1)^2}$$

So in the U_N chart for the point $x = 0.5$ we have that the function f is changing in the direction X by $(Xf)_N(0.5) = 1.28$. And in the U_S chart for the corresponding point $y = 2$ we have that the function f is changing in the direction X by $(Xf)_S(2) = 1.28$. The directional derivatives in the two charts agree. This is a demonstration that everything is all well-defined.

Example 2.21 (Glued Circle). Because the Jacobian of the transition function is just the constant 1, the following is a well-defined vector field:

$$X = \begin{cases} \frac{\partial}{\partial x} & \text{for } x \in U_1, \\ \frac{\partial}{\partial y} & \text{for } y \in U_2 \end{cases}$$

We can act this vector field on the function $f : \mathbb{S}^1 \rightarrow \mathbb{R}$ defined in charts as

$$f_{11}(x) = \cos \pi x, \quad f_{21}(y) = \cos \pi(y + 1),$$

which is just the first component of the function from Example 2.16. We get

$$(Xf)_{11}(x) = \frac{\partial}{\partial x} \cos \pi x = -\pi \sin \pi x, \quad (Xf)_{21}(y) = \frac{\partial}{\partial y} \cos \pi(y + 1) = -\pi \sin \pi(y + 1).$$

This is a well-defined function on \mathbb{S}^1 , which we know because it is $-\pi$ multiplied by the second component from Example 2.16.

These two examples illustrate that a vector field X applied in a chart to a function f_i is another function $(Xf)_i$. The fact that we used the chain rule to define equivalence of vectors is precisely the condition to ensure that the resulting functions in charts piece together to give a well-defined function Xf on the whole manifold.

If we have a function between two manifolds $f : M \rightarrow \tilde{M}$, the *tangent map* or *pushforward* of f at p is a function between the tangent spaces $T_p f : T_p M \rightarrow T_p \tilde{M}$. We will define it in two ways. The easiest definition is using the curve method of vectors. If we have a vector $v \in T_x M$ then in a chart U_i the curve $\alpha(t) = x + vt$ is a representative of this vector. Observe that $f \circ \alpha : (a, b) \rightarrow \tilde{M}$ is a curve in \tilde{M} . We define $T_p f(v)$ to be the tangent vector of the curve $f \circ \alpha$ at $t = 0$.

We also give a practical formula for calculating the tangent map. The curve $f \circ \alpha$ must lie in some chart \tilde{U}_j of \tilde{M} . In charts we have $\alpha_i : (a, b) \rightarrow U_i$ and $f_{ij} : U_i \rightarrow \tilde{M}_j$. So the tangent vector of the curve is

$$\left. \frac{d}{dt} f_{ij}(\alpha(t)) \right|_{t=0} = (J_{\alpha(t)} f_{ij}) \alpha'(t) \Big|_{t=0} = (J_{\alpha(0)} f_{ij}) \alpha'(0) = (J_x f_{ij}) v,$$

using the chain rule and the fact that $\alpha(0) = x$, $\alpha'(0) = v$. Thus we see in charts that the tangent map is the Jacobian of f_{ij} .

We could have also given a definition of the pushforward based on the directional derivative idea. If $w = T_p f(v) \in T_{f(p)} \tilde{M}$ is the pushforward of a vector $v \in T_p M$, then we can ask how w acts on a function $g : \tilde{M} \rightarrow \mathbb{R}$. The observation is that $g \circ f : M \rightarrow \mathbb{R}$ and

$$w(g) = T_p f(v)(g) = v(g \circ f).$$

Example 2.22 (Stereographic Projection). Consider the inclusion map $\iota : \mathbb{S}^1 \rightarrow \mathbb{R}^3$. As a formula it is just identity, but we consider it as a function between manifolds. We can ask how the vector $\partial_x|_0 \in T_S \mathbb{S}^1$, a coordinate vector of the U_N chart, is pushed forward by this map to a vector in $T_{(0,-1)} \mathbb{R}^2$, since $0 \in U_N = S \in \mathbb{S}^1$ is mapped to $\iota(S) = (0, -1)$. Let us calculate the tangent map in the chart U_N . As per above, this is Jacobian of $\iota_{N0} = \text{id}_{\mathbb{R}^3} \circ \iota \circ \phi_N^{-1} = \phi_N^{-1}$:

$$\begin{aligned} J_x \phi_N^{-1} &= \begin{pmatrix} \frac{\partial(\phi_N^{-1})^1}{\partial x} \\ \frac{\partial(\phi_N^{-1})^2}{\partial x} \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x} \left(\frac{2x}{x^2+1} \right) \\ \frac{\partial}{\partial x} \left(\frac{x^2-1}{x^2+1} \right) \end{pmatrix} \\ &= \frac{2}{(x^2+1)^2} \begin{pmatrix} -x^2+1 \\ 2x \end{pmatrix} \end{aligned}$$

and at $x = 0$ we have

$$J_0 \phi_N^{-1} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

In coordinates, the vector $\partial_x|_0$ is just 1, so its pushforward in coordinates is

$$(T_S \iota)(\partial_x|_0) = (J_0 \phi_N^{-1})(1) = \begin{pmatrix} 2 \\ 0 \end{pmatrix} (1) = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2 \frac{\partial}{\partial p^1} \Big|_{(0,-1)} + 0 \frac{\partial}{\partial p^2} \Big|_{(0,-1)}.$$

| Note that this vector is tangent to \mathbb{S}^1 in the usual sense.

An important observation should be made about the previous example. We were careful to stress that the tangent map was applied to a vector at a particular point and not a vector field. That is because the tangent map does not, as students sometimes assume, take vector fields to vector fields. There are two ways it fails to do this. We could easily generalise this example to ∂_x at any point of U_N , which would then give a vector on \mathbb{R}^2 . But not every point of \mathbb{R}^2 gets a vector. Only the points in the image of ϕ_N^{-1} get a vector. Therefore the result is not a vector field on \mathbb{R}^2 .

The other way that the tangent map can fail to transform a vector field into another vector field, which does not apply to the above example, is if the function is not injective. In that case some points in the image get multiple vectors from the tangent map. Since a vector field is a certain type of function, and functions must have exactly one output, the pushforward of a vector field may not be a function.

Exercise 2.23 (Stereographic Projection). Verify the following formula for the tangent map of $\iota : \mathbb{S}^2 \rightarrow \mathbb{R}^3$ in the U_N chart.

$$J_x(\phi_N^{-1}) = \frac{2}{(\|x\|^2 + 1)^2} \begin{pmatrix} -(x^1)^2 + (x^2)^2 + 1 & -2x^1x^2 \\ -2x^1x^2 & (x^1)^2 - (x^2)^2 + 1 \\ 2x^1 & 2x^2 \end{pmatrix}.$$

We end this section by highlighting a distinction that we have elided until now. A tangent vector is an intrinsic concept on a manifold, it does not require the manifold to be immersed in euclidean space. And yet in the previous chapter we repeatedly talked about tangent vectors to surfaces as a vector in \mathbb{R}^3 that was tangent in the geometrical sense. The connections between these two ideas is exactly the idea of immersion and pushforward.

Definition 2.24. A function $f : M \rightarrow \tilde{M}$ is an immersion at $p \in M$ if $T_p f$ is injective.

We commented on this at the end of Example 2.22. Another possible example would be Example 2.16. We saw that a manifold can be mapped into a euclidean space by a function. The pushforward of its tangent vectors by this function are exactly the tangent vectors in the intuitive picture. Thus for an immersion, it is possible to consider the elements of $T_p M$ as certain elements of $T_p \tilde{M}$, and this is unambiguous because by definition $T_p f$ is injective.

2.4 Vector Bundles

This section will not be used later in the script. It exists just to flex how natural the construction of the tangent bundle is in our approach to manifolds.

Above we have defined the tangent space T_pM to a manifold M at a point $p \in M$ and discussed how the vectors of T_pM and T_qM should be thought of as distinct, even if p and q are in the same chart. But we also worked with examples of vector fields, whose values are vectors at different points. We can reconcile this tension by putting all the vectors of a manifold together into a new manifold:

Definition 2.25. Let M be a manifold with atlas $\mathcal{A} = (n, \mathcal{I}, \{U_i\}_{i \in \mathcal{I}}, \{\varphi_{ij}\}_{i,j \in \mathcal{I}})$. We construct a new atlas $T\mathcal{A} = (2n, \mathcal{I}, \{TU_i\}_{i \in \mathcal{I}}, \{T\varphi_{ij}\}_{i,j \in \mathcal{I}})$, where $TU_i = U_i \times \mathbb{R}^n$ and

$$T\varphi_{ij} : \text{dom } \varphi_{ij} \times \mathbb{R}^n \rightarrow \text{img } \varphi_{ij} \times \mathbb{R}^n, \quad (x, v) \mapsto (\varphi_{ij}(x), (J_x \varphi_{ij})v).$$

The corresponding manifold is called the tangent bundle TM . There is a function $\pi : TM \rightarrow M$ in charts by

$$\pi_{ii}(x, v) = x.$$

We tend to think of the tangent bundle of M as all the vectors of M , with the understanding that vectors at different points are distinct from one another. The function π is called the canonical projection of the bundle. Intuitively it takes a tangent vector to its base point. Hence $\pi^{-1}[\{p\}]$ are all the vectors which live at the same point, in other words T_pM .

The tangent bundle allows us to speak and reason formally about tangent vectors as a whole. We can define a vector field on M as a function X from M to TM with the property that $\pi \circ X = \text{id}_M$. This property says that for every point $p \in M$ the vector $X|_p$ must have the base point $\pi(X|_p) = p$, which is exactly what a vector field is.

Exercise 2.26. Check that the tangent bundle is a manifold. In particular you must check the unique limit condition.

Example 2.27 (Euclidean Space). Since the trivial atlas for an open subset $U \subset \mathbb{R}^n$ has only one chart $U_0 = U$, so too does its tangent bundle $TU_0 = U \times \mathbb{R}^n$. Therefore the tangent bundle is just the cartesian product. The only difference between TU and usual way we think about vectors in euclidean space is that the base point of the vector is important for the tangent bundle.

Example 2.28 (Glued Circle). We saw in Example 2.21 that there was a vector field X on \mathbb{S}^1 that was never zero. Because $T_p\mathbb{S}^1$ is one dimensional, every vector in it must be a scalar multiple of $X|_p$. This means we can define a function from T_pM to \mathbb{R}

$$v \mapsto a, \text{ where } v = aX|_p.$$

Allowing p to change gives us a function from $T\mathbb{S}^1$ to $\mathbb{S}^1 \times \mathbb{R}$. Thus the tangent bundle of the circle is also a product. Tangent bundles that are products are called *trivial*.

Example 2.29. In this example we want to show a non-trivial tangent bundle, but it's actually somewhat difficult to prove non-triviality. Instead we will try to convey the idea. The tangent bundle of \mathbb{S}^2 is non-trivial. If it were trivial, that would mean there would exist a smooth bijective correspondence between $T\mathbb{S}^2$ and $\mathbb{S}^2 \times \mathbb{R}^2$. We could use this correspondence to write any vector field X on \mathbb{S}^2 as $(p, \tilde{X}(p))$ with $\tilde{X}(p) \in \mathbb{R}^2$. Conversely there would exist a vector field with the form $(p, (1, 0))$. This vector field is never zero.

But it is a theorem that every vector field on the sphere has at least one zero. This is called the **hairy ball theorem** or the hedgehog theorem. There are many intuitive interpretations of this theorem; one is that there must be somewhere on earth where the wind is not blowing. Try imagining different wind maps, and perhaps you will convince yourself of the truth of this theorem. This is one reason that the tangent bundle of \mathbb{S}^2 cannot be trivial.

The tangent map takes its nicest form when expressed with the tangent bundle. We can collect together all the tangent maps $T_p f : T_p M \rightarrow T_p N$ into a single map $Tf : TM \rightarrow TN$. Using the formula in charts in terms of the Jacobian we get

$$(Tf)_{ij}(x, v) = \left(f_{ij}(x), (J_x f_{ij})v \right) \in V_j \times \mathbb{R}^n = TV_j.$$

This formula shows that the tangent map is the generalisation of the Jacobian, and all local properties of the Jacobian carry over to the tangent map.

Exercise 2.30. Prove that the above formula for the tangent map is well-defined (in different charts) by using the chain rule for Jacobians: for $y = \varphi_{M,i,j}(x)$ and $z = f_{j,k}(y)$

$$J_x(\varphi_{N,k,l} \circ f_{j,k} \circ \varphi_{M,i,j}) = (J_z \varphi_{N,k,l})(J_y f_{j,k})(J_x \varphi_{M,i,j}).$$

Exercise 2.31 (Glued Circle). Calculate the tangent map for f from Example 2.16. Use this to give an alternative description of the tangent bundle $T\mathbb{S}^1$ by locating it as a set inside $T\mathbb{R}^2 = \mathbb{R}^2 \times \mathbb{R}^2$.

2.5 Summation Convention

As you have seen, when working with vectors in charts, there are many Σ summations, but all of them are just from $i = 1$ to n . There is a convention, called the Einstein summation convention or the Einstein rule, that allows us to omit all these sigmas.

Definition 2.32 (Einstein Summation). *We apply the following notational rule: when an index occurs twice in a term, once as an upper index and once as a lower index, we sum that index over its range. For the purposes of this rule $\frac{\partial}{\partial x^i} = \partial_i$ counts as a lower index.*

In this notation, the chain rule reads

$$\frac{\partial}{\partial x^1} = \frac{\partial y^1}{\partial x^1} \frac{\partial}{\partial y^1} + \frac{\partial y^2}{\partial x^1} \frac{\partial}{\partial y^2} = \sum_{i=1}^2 \frac{\partial y^i}{\partial x^1} \frac{\partial}{\partial y^i} = \frac{\partial y^i}{\partial x^1} \frac{\partial}{\partial y^i}.$$

A vector field X that has coefficients X^i with respect to a chart can neatly be written as $X = X^i \partial_i$, and its expression with respect to a second chart y might be written using the equivalences of vectors as

$$X = X^i \frac{\partial}{\partial x^i} = X^i \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j}.$$

This convention is useful for even ordinary linear algebra. If you have a matrix A we would normally write the entries A_{ij} . However if we instead write it A_j^i with a mix of upper and lower indices then we can write matrix multiplication in the following way

$$Av = \left(A_j^i \right) \left(v^j \right) = \sum_{j=1}^n A_j^i v^j = A_j^i v^j.$$

Likewise matrix multiplication AB would be $(AB)_k^i = A_j^i B_k^j$. The saying to turn matrix algebra into index algebra with summation convention is “Upper indices go up to down; lower indices go left to right”.

We know that bilinear functions can also be represented as a matrix; in usual matrix notation:

$$g(v, w) = v^T \left(g_{ij} \right) w = \sum_{i,j=1}^n v^i g_{ij} w^j.$$

In summation convention we write this as $g_{ij} v^i w^j$, with the vector components having upper indices, which forces g to have two lower indices. This is actually an advantage of index notation over matrix notation. Though linear transforms and bilinear transforms can both be represented as a matrix, those matrix representations behave differently under change of basis. We can see that they are different in index notation, but not in matrix notation.

Some authors do not have the condition that you need one upper and one lower index, and allow summation over any repeated index. I do not like this, because it makes it impossible to talk about the diagonal elements g_{ii} . If you need to sum over two indices on the same level, you can of course use a summation sign. Alternatively you can use an identity matrix, whose entries are usually called δ_{ij} . For example

$$\text{tr} \left(g_{ij} \right) = \sum_{i=1}^n g_{ii} = \delta^{i,j} g_{ij}.$$

Finally, some useful rules of calculation are

$$\frac{\partial x^i}{\partial y^j} \frac{\partial y^j}{\partial x^k} = \delta_k^i \quad \text{and} \quad v^i \delta_i^j = v^j.$$

2.6 The Lie Bracket

We have seen examples of how vectors are a type of directional derivative of functions and that a vector field applied to a function gives another function. A natural question is to ask whether we can also differentiate a vector field using a vector field. The answer is yes: in fact vector fields can differentiate many objects on manifolds using a process called the Lie derivative \mathcal{L}_X , but the process complicated and we won't explain it here. We show with the following example why the 'obvious' way to differentiate vector fields doesn't work.

Example 2.33 (Polar Coordinates). You might guess that you can differentiate a vector field just by differentiating its coefficient functions:

$$Y(X) \stackrel{?}{:=} \sum_{i=1}^n Y(X^i) \frac{\partial}{\partial x^i},$$

but this doesn't produce a well-defined vector field, because in different charts it produces inequivalent vectors.

Let us give an example. Consider the vector fields

$$X = x^1 \frac{\partial}{\partial x^1} + x^2 \frac{\partial}{\partial x^2}, \quad Y = x^2 \frac{\partial}{\partial x^1} - x^1 \frac{\partial}{\partial x^2},$$

on $U_1 = \mathbb{R}^2$. For example, if we act Y on X with this definition, in U_1 we have

$$\left[x^2 \frac{\partial x^1}{\partial x^1} - x^1 \frac{\partial x^1}{\partial x^2} \right] \frac{\partial}{\partial x^1} + \left[x^2 \frac{\partial x^2}{\partial x^1} - x^1 \frac{\partial x^2}{\partial x^2} \right] \frac{\partial}{\partial x^2} = x^2 \frac{\partial}{\partial x^1} - x^1 \frac{\partial}{\partial x^2} = Y.$$

But in U_2 , polar coordinates, we have

$$\begin{aligned} r \frac{\partial}{\partial r} &= r \frac{\partial x^i}{\partial r} \frac{\partial}{\partial x^i} = r \cos \theta \frac{\partial}{\partial x^1} + r \sin \theta \frac{\partial}{\partial x^2} = X, \\ -\frac{\partial}{\partial \theta} &= -\frac{\partial x^i}{\partial \theta} \frac{\partial}{\partial x^i} = r \sin \theta \frac{\partial}{\partial x^1} - r \cos \theta \frac{\partial}{\partial x^2} = Y. \end{aligned}$$

So the same calculation in polar coordinates gives

$$\left[-\frac{\partial r}{\partial \theta} \right] \frac{\partial}{\partial r} = 0.$$

Hence this operation depends on which coordinates one uses. It is not well-defined on the manifold.

It turns out the correct way to differentiate vector fields is to take the commutator of the above guess.

Definition 2.34. *The Lie derivative of a vector field Y with respect to a vector field X , also called the Lie bracket, acts on a function f by*

$$(\mathcal{L}_X Y)(f) = [X, Y](f) := X(Y(f)) - Y(X(f)).$$

From this definition we can ask how to write it in charts:

$$\begin{aligned}
[X, Y](f) &= X(Y(f)) - Y(X(f)) = X\left(Y^j \frac{\partial f}{\partial x^j}\right) - Y\left(X^i \frac{\partial f}{\partial x^i}\right) \\
&= X^i \frac{\partial}{\partial x^i} \left(Y^j \frac{\partial f}{\partial x^j}\right) - Y^j \frac{\partial}{\partial x^j} \left(X^i \frac{\partial f}{\partial x^i}\right) \\
&= X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial f}{\partial x^j} + X^i Y^j \frac{\partial^2 f}{\partial x^i \partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \frac{\partial f}{\partial x^i} - Y^j X^i \frac{\partial^2 f}{\partial x^j \partial x^i} \\
&= \left[X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j}\right] \frac{\partial f}{\partial x^i} = [X(Y^i) - Y(X^i)] \frac{\partial f}{\partial x^i}.
\end{aligned}$$

So even though the definition of $[X, Y]$ appears to have second derivatives of f , these cancel out and what remains is indeed a (first order) vector field acting on f .

Example 2.35 (Polar Coordinates). Let us take the vector fields from the previous example and compute their Lie bracket. In the U_1 chart

$$\begin{aligned}
[X, Y] &= \left[\left(x^1 \frac{\partial}{\partial x^1} + x^2 \frac{\partial}{\partial x^2}\right)(x^2) - \left(x^2 \frac{\partial}{\partial x^1} - x^1 \frac{\partial}{\partial x^2}\right)(x^1)\right] \frac{\partial}{\partial x^1} \\
&\quad + \left[\left(x^1 \frac{\partial}{\partial x^1} + x^2 \frac{\partial}{\partial x^2}\right)(-x^1) - \left(x^2 \frac{\partial}{\partial x^1} - x^1 \frac{\partial}{\partial x^2}\right)(x^2)\right] \frac{\partial}{\partial x^2} \\
&= [x^2 - x^2] \frac{\partial}{\partial x^1} + [-x^1 + x^1] \frac{\partial}{\partial x^2} = 0.
\end{aligned}$$

and in the U_2 chart

$$\begin{aligned}
[X, Y] &= \left[\left(r \frac{\partial}{\partial r} + 0 \frac{\partial}{\partial \theta}\right)(0) - \left(0 \frac{\partial}{\partial r} - \frac{\partial}{\partial \theta}\right)(r)\right] \frac{\partial}{\partial r} \\
&\quad + \left[\left(r \frac{\partial}{\partial r} + 0 \frac{\partial}{\partial \theta}\right)(-1) - \left(0 \frac{\partial}{\partial r} - \frac{\partial}{\partial \theta}\right)(0)\right] \frac{\partial}{\partial \theta} = 0.
\end{aligned}$$

Whenever we have an expression in charts, it is good to check that in different charts it produces equivalent vectors. We do so in the following lemma. This is not strictly necessary, because the $[X, Y](f) = X(Y(f)) - Y(X(f))$ definition doesn't use charts, but good practice none-the-less.

Lemma 2.36. *The above chart expression gives a vector field that is independent of the choice of chart.*

Proof. If $X = X^i(x) \frac{\partial}{\partial x^i}$ and $Y = Y^i(x) \frac{\partial}{\partial x^i}$ are the expressions of two vector fields in the x -chart, then the expression of these fields in the y -chart is

$$X = X^j \frac{\partial y^i}{\partial x^j} \frac{\partial}{\partial y^i}, \quad Y = Y^j \frac{\partial y^i}{\partial x^j} \frac{\partial}{\partial y^i}.$$

In this chart, the i th component of $[X, Y]$ is

$$\begin{aligned}
X(\tilde{Y}^i) - Y(\tilde{X}^i) &= X^l \frac{\partial y^k}{\partial x^l} \frac{\partial}{\partial y^k} \left(Y^j \frac{\partial y^i}{\partial x^j} \right) - Y^j \frac{\partial y^k}{\partial x^j} \frac{\partial}{\partial y^k} \left(X^l \frac{\partial y^i}{\partial x^l} \right) \\
&= X^l \frac{\partial y^k}{\partial x^l} \frac{\partial}{\partial y^k} \left(Y^j \frac{\partial y^i}{\partial x^j} \right) - Y^j \frac{\partial y^k}{\partial x^j} \frac{\partial}{\partial y^k} \left(X^l \frac{\partial y^i}{\partial x^l} \right) \\
&= X^l \frac{\partial}{\partial x^l} \left(Y^j \frac{\partial y^i}{\partial x^j} \right) - Y^j \frac{\partial}{\partial x^j} \left(X^l \frac{\partial y^i}{\partial x^l} \right) \\
&= X^l \frac{\partial Y^j}{\partial x^l} \frac{\partial y^i}{\partial x^j} + X^l Y^j \frac{\partial^2 y^i}{\partial x^l \partial x^j} - Y^j \frac{\partial X^l}{\partial x^j} \frac{\partial y^i}{\partial x^l} - Y^j X^l \frac{\partial^2 y^i}{\partial x^j \partial x^l} \\
&= X^l \frac{\partial Y^j}{\partial x^l} \frac{\partial y^i}{\partial x^j} - Y^j \frac{\partial X^l}{\partial x^l} \frac{\partial y^i}{\partial x^j} = \left[X^l \frac{\partial Y^j}{\partial x^l} - Y^j \frac{\partial X^l}{\partial x^l} \right] \frac{\partial y^i}{\partial x^j}.
\end{aligned}$$

In total then, in the vector field $[X, Y]$ in the y -chart is

$$[X, Y] = \left[X^l \frac{\partial Y^j}{\partial x^l} - Y^j \frac{\partial X^l}{\partial x^l} \right] \frac{\partial y^i}{\partial x^j} \frac{\partial}{\partial y^i}.$$

But we see that this is equivalent to the expression in the x -chart. Therefore the definition produces a well-defined vector field. \square

There are of course many things that can be said about the Lie bracket. The first observation is that it is \mathbb{R} -bilinear and antisymmetric: $[aX + b\tilde{X}, Y] = a[X, Y] + b[\tilde{X}, Y]$ and $[X, Y] = -[Y, X]$. If you have two coordinate vector fields, then their Lie bracket is zero. Thus one interpretation of the Lie bracket is that it is a measurement of how far two vector fields are from being coordinate vector fields. The final property that we will give is a product rule: if f is a function, then

$$\begin{aligned}
[X, fY] &= X(f)Y + f[X, Y], \\
[fX, Y] &= -[Y, fX] = -Y(f)X - f[Y, X] = f[X, Y] - Y(f)X.
\end{aligned}$$

Example 2.37 (Euclidean Space). We give some euclidean examples. Consider the plane \mathbb{R}^2 . Let $X = \partial_1$ and $Y = \partial_2$. Then plugging in the definitions

$$[X, Y] = X(Y^i)\partial_i - Y(X^i)\partial_i = X(1)\partial_2 - Y(1)\partial_1 = 0.$$

Next consider $V = (1 + x^2)\partial_1$. Then

$$[V, Y] = V(1)\partial_2 - Y(1 + x^2)\partial_1 = -\partial_1.$$

Finally, set $W = x^2\partial_1 - x^1\partial_2$. then

$$[V, W] = V(x^2)\partial_1 - V(x^1)\partial_2 - W(1 + x^2)\partial_1 = 0 - (1 + x^2)\partial_2 - (0 - x^1)\partial_1 = -(1 + x^2)\partial_2 + x^1\partial_1.$$