An Introduction to Differential Geometry

Eugenen Lerman

PREFACE

Note by the transcriber

This is a transcription of "Introduction to Differential Geomtry" by Eugene Lerman.

The transcriber is currently binding the notes of Brian Conrad on differential geometry. His notes are detailed, but unfortunately they are nothing more than a colleciton of seemingly sporadic topics: in other words, they are not well-organised. So, I looked for a compact and well-organised lecture notes on differential geometry.

First, I tried "Notes on Differential Geometry" by Hicks, but they have a crucial flaw: namely he assigns the same symbol X to both tangent vectors and vector fields, which makes a confusing and frustrating read.

Then I found Lerman's notes and decided to transcribe them in the hope that his notes and Conrad's notes can coplement each other.

For the record, Lerman has not provided preface to his notes.

The appendix contains the material omitted in the original notes: namely,

1. the proof of the inverse function theorem,

2. the proof of Sard's theorem.

Contents

Chapter 1

Introduction: why manifolds?

There are many different ways to formulate mathematically the notion of a 'space' that occurs in different branches of science and engineering. For instance one can talk about the space of configurations of a physical system. This, of course, requires a decision as to the level of details one is trying to model. For example, we can regard the configuration space of a system consisting of a sun and a planet as $\mathbb{R}^3 \times \mathbb{R}^3$. We use three real numbers to describe the position of the center of mass of the sun and three real numbers to describe the position of the center of mass of the planet. In this model we assume that the sun and the planet are simply two points in space. We also allow collisions. If we exclude collisions (but still allow the sun and the planet to come arbitrarily close to each other), the con
guration space is then

$$Q = \{(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 | x \neq y\}.$$

Here is another idealised example: the configuration space of a penny tumbling through the air. Fix a frame of reference. We will need a triple of real numbers to describe the position of the penny's center of gravity and three orthonormal vectors to describe the orientation of the penny. Thus the con
guration space in question is

$$Q = \mathbb{R}^3 \times \mathrm{O}(3),$$

where $\mathrm{O}(3)$ denotes the set of 3×3 orthogonal matrices (recall that an $n \times n$ matrix is orthogonal if (and only if) its columns form an orthonormal basis of $\mathbb{R}^n$).[1]

Exercise 1.0.1. What is the configuration space of a penny rolling on a plane?
Manifolds constitute a particular way to formalise the notion of a configuration space. These are the spaces that "locally look like $\mathbb{R}^n$." The reason we will limit

[1]Strictly speaking the configuration space is $\mathbb{R}^3 \times \mathrm{SO}(3)$, where $\mathrm{SO}(3)$ denotes the set of orthogonal matrices with positive determinant. Why?

ourselves to manifolds is that they are particularly suitable for generalising the ideas of calculus — differentiation and integration. We will see that the two examples of configuration spaces given above: $Q = \{(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 | x \neq y\}$ and $Q = \mathbb{R}^3 \times \mathrm{O}(3)$ are, indeed, manifolds.

Remark 1.0.2. There are, of course, many other notions of a "space." In linear algebra one studies vector spaces and maps between them. In algebraic geometry one studies spaces of solutions of polynomial equations which give rise to the notion of an algebraic variety. In metric topology/geometry one studies metric spaces, spaces with a notion of a distance. In point set topology and in algebraic topology one talks about topological spaces. In analysis one may study the space of solutions of a partial differential equation. In geometry and topology one may be forced to study spaces that have singularities such as orbifolds and stratified spaces. Before we can discuss orbifolds and more complicated spaces we should first come to terms with manifolds which are smooth.

Chapter 2

Smooth manifolds

2.1 Digression: smooth maps from open subsets of $\mathbb{R}^n$ to $\mathbb{R}^m$

We start out by recalling the definition of a differentiable map.

Definition 2.1.1. Let $U \subset \mathbb{R}^n$ be an open subset. A map $f: U \to \mathbb{R}^m$ is *differentiable* at a point $x \in U$ if there is a linear map $L: \mathbb{R}^n \to \mathbb{R}^m$ so that

$$\lim_{h \to 0} \frac{1}{\|h\|}(f(x+h) - f(x) - Lh) = 0.$$

It is not hard to show that if such a map L exists, it is unique. The linear map L is variously called the *derivative* of f at x, the *differential* of f at x, ... and is denoted by df_x or by Df_x or by $Df(x)$ or by a similar notation. Moreover, the matrix corresponding to L with respect to the standard basis of $\mathbb{R}^n$ and $\mathbb{R}^m$ is the so called *Jacobian matrix*. That is, if $f = (f_1, \dots, f_m)$ then

$$Df_x = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix}$$

Definition 2.1.2. Let $U \subset \mathbb{R}^n$ be an open subset. A map $f: U \to \mathbb{R}^m$ is *smooth* (or C^∞) on the set U if all partial derivatives of f to all orders exist at all points of U.

Here is a more "sophisticated" version of the the definition above. Suppose $f: U \to \mathbb{R}^m$ is differentiable at all points of U. Then we have a map $g(x) := Df_x: U \to \mathbb{R}^{nm}$. We can require that g is differentiable as a map from U to $\mathbb{R}^{nm}$. The derivative of g is a map from U to a bigger vector space $\mathbb{R}^N$ for an appropriate N. We can require that this map is differentiable and so on... In other words, if all derivatives of $f: U \to \mathbb{R}^n$ exist and are differentiable we say that f is *smooth*.

2.2 Definitions and examples of manifolds

A smooth manifold is a generalisation of a smooth surface in $\mathbb{R}^3$. A smooth surface in $S \subset \mathbb{R}^3$ has local parameterisations: for every point $p \in S$ there is an open set $V \subset \mathbb{R}^3$ with $p \in V$ and a map $x \colon U \to S \cap V$ (where $U \subset \mathbb{R}^2$ is an open set) such that

(1) x is C^∞. That is $x(u_1, u_2) = (x_1(u_1, u_2, x_2(u_1, u_2), x_3(u_1, u_2))$ and each $x_i(u_1, u_2), 1 \le i \le 3$ is an infinitely differentiable function of $u = (u_1, u_2) \in U$;

(2) x is 1-1 (injective) and onto.

The map x is a *local parameterisation* of S.

Example 2.2.1. The two sphere

$$S^2 = \{s \in \mathbb{R}^3 \ \|x\| = 1\}$$

is a smooth surface. In fact, if $p = (p_1, p_2, p_3) \in S^2$ and $p_3 > 0$ take $V = \{x \in \mathbb{R}^3 | x_3 > 0\}$, $U = \{(u_1, u_2) | \|u\| < 1\}$ and a local parameterisation $x \colon U \to S^2 \cap V$ to be $x(u_1, u_2) = (u_1, u_2, \sqrt{1 - u_1^2 - u_2^2})$. It's easy to check that this x is 1-1, onto and C^∞. If $p_3 < 0$ take the local parameterisation $x(u) = (u_1, u_2, -\sqrt{1 - u_1^2 - u_2^2})$. If $p_3 = 0$ then either p_1 or p_2 is non-zero (or both) and there are formulasefor local parameterisations similar to the ones above.

Note that if S is a smooth surface and $x_\alpha \colon \mathbb{R}^2 \supset U_\alpha \to S$ and $x_\beta \colon \mathbb{R}^2 \supset U \to S$ are two local parameterisations with

$$W_{\alpha\beta} := x_\alpha(U_\alpha) \cap x_\beta(U_\beta) \neq \varnothing$$

then

$$x_\beta^{-1} \circ x_\alpha \colon x_\alpha^{-1}(W_{\alpha\beta}) \to x_\beta^{-1}(W_{\alpha\beta}) \subset \mathbb{R}^2$$

is C^∞.

This motivates

Definition 2.2.2 (of a C^∞ manifold, first approximation, not quite right)**.** A C^∞ manifold of dimension m is a set M and a family of injective maps $\{x_\alpha \colon U_\alpha \to M\}$ where $U \subset \mathbb{R}^m$ are open sets, such that

(1) $\cup x_\alpha(U_\alpha) = M$;

(2) if for some pair of indices α and β, the set $W_\alpha := x_\alpha(U_\alpha) \cap x_\beta(U_\beta) \neq \varnothing$ then $x_\alpha^{-1}(W_{\alpha\beta}), x_\beta^{-1}(W_\beta)$ are open in $\mathbb{R}^m$ and

$$x_\beta^{-1} \circ x_\alpha \colon x_\alpha^{-1}(W_{\alpha\beta}) \to x_\beta^{-1}(W_{\alpha\beta})$$

are C^∞.

One thing that is wrong with this definition is that there is no topology specified on M. The other is that instead of parameterisations one usually works with charts that go the other way. Namely

Definition 2.2.3 (Chart)**.** Let X be a topological space. An $\mathbb{R}^n$ *(coordinate) chart* on X is a homeomorphism $\phi\colon X \supset U \to U' \subset \mathbb{R}^n$.

Notation 2.2.4. We will often write $\phi\colon U \to \mathbb{R}^n$ or even (U, ϕ) for a coordinate chart $\phi\colon X supset U \to U' \subset \mathbb{R}^n$. Note that since ϕ takes values in $\mathbb{R}^n$, it is an n-tuple of functions $\phi = (x_1, \ldots, x_n)$ for some functions $x_i\colon U \to \mathbb{R}$, the *coordinate functions on* U associated to the coordinate chart $\phi\colon U \to \mathbb{R}^n$.

Notation 2.2.5. When dealing with charts it will be convenient to to adopt the notation where the standard coordinate functions on $\mathbb{R}^n$ are denote by $r_i, 1 \le i \le n$. That is, r_i assigns to a point $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ the number a_i. If $\phi\colon U \to \mathbb{R}^n$ is a chart then

$$x_i = r_i \circ \phi\colon U \to \mathbb{R}$$

are the coordinate functions on U.

Definition 2.2.6 (Atlas)**.** A C^∞ *atlas* on a topological space X is a collection of charts $\{\phi_\alpha\colon U_\alpha \to U'_\alpha\}$ (with all U''s being open subsets of one fixed $\mathbb{R}^n$ such that

(1) $\{U_\alpha\}$ is an open cover of X,[1] and

(2) If $U_\alpha \cap U_\beta \neq \varnothing$, then $\phi_\beta \circ \phi_\alpha\colon \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$ is C^∞ as a map from an open subset of $\mathbb{R}^n$ to $\mathbb{R}^n$. That is, changes of coordinates are smooth.

Example 2.2.7. The identity map $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = x$ is the standard chart on $\mathbb{R}$. The set $\{(f, \mathbb{R})\}$ consisting of one chart is an atlas on $\mathbb{R}$. The map $g\colon \mathbb{R} \to \mathbb{R}$, $g(x) = x^3$ is also a chart on $\mathbb{R}$; it defines a different atlas on $\mathbb{R}$.

Here is a third atlas on $\mathbb{R}$. For each integer $n \in \mathbb{Z}$, $\phi_n\colon (n, n+2) \to \mathbb{R}$, $\phi_n(x) = x$ is a chart. The set $\{(\phi_n, (n, n+2)\}$ is an atlas on $\mathbb{R}$.

Definition 2.2.8. We say that *two atlases are equivalent* if their union is also an atlas.

The definition above amounts to: an atlas $\{x_\alpha\colon U_\alpha \to U'_\alpha\}$ is equivalent to an atlas $\{y_\beta\colon V_\beta \to V'_\beta\}$ if for any indices α, β with $U_\alpha \cap V_\beta \neq \varnothing$ the map $x_\alpha \circ y_\beta^{-1}\colon y_\beta(U_\alpha \cap V_\beta) \to x_\alpha(U_\alpha \cap V_\beta)$ is smooth. One can easily verify that this is indeed an equivalence relation

Exercise 2.2.9. Convince yourself that the first and the third atlases in Example 2.2.7 are equivalent. Show that the first and the second example of atlases are not equivalent.

Definition 2.2.10 (Manifold)**.** An *n-dimensional* (C^∞) *manifold* is a topological space M together with an equivalence class of C^∞ atlases.

Notation 2.2.11. We will denote the manifold and the underlying topological space by the same letter, with the equivalence class of atlases usually understood.

[1] That is, each $U_\alpha \subset X$ is open and $\cup_\alpha U_\alpha = X$

Example 2.2.12. Let $M = \mathbb{R}^n$. We cover M by one open set and take the identity map as our chart. This is the *standard* manifold structure on $\mathbb{R}^n$.

Example 2.2.13. Let $M = \mathbb{C}^n$. Again we cover $\mathbb{C}^n$ by one open set $U = \mathbb{C}^n$ and take the identity map as our coordinate chart map $\phi\colon \mathbb{C}^n \to \mathbb{R}^{2n}$ which is given by

$$\phi(z_1, \dots, z_n) = (\Re z_1, \Im z_1, \dots).$$

Example 2.2.14. If M is a manifold, and $V \subset M$ is an open subset, then V is naturally a manifold. Check this!

Example 2.2.15. The set $\mathrm{M}_n(\mathbb{R})$ of $n \times n$ matrices with real coeffcients is a manifold, since it is $\mathbb{R}^{n^2}$. The subset $\mathrm{GL}(n, \mathbb{R}) \subset \mathrm{M}_n(\mathbb{R})$ of invertible matrices is an open subset: a matrix A is invertible if and only if its determinant is non-zero and determinant $\det\colon \mathrm{M}_n(\mathbb{R}) \to \mathbb{R}$ is a polynomial map, hence continuous. Hence the subset $\{A \in \mathrm{M}_n(\mathbb{R})\} | \det A \neq 0\}$ is open. So by the previous example, $\mathrm{GL}(n, \mathbb{R})$ is a manifold.

Example 2.2.16. The two-sphere $S^2 := \{x \in \mathbb{R}^3 | \|x\|^2 = 1\}$ is a manifold. To see this, we give S^2 the subspace topology that it inherits as a subset of $\mathbb{R}^3$. Next we define charts. To do this, let

$$U_i^+ = \{x = (x_1, x_2, x_3) \in S^2 : x_i > 0\}$$

and

$$U_i^- = \{x = (x_1, x_2, x_3) \in S^2 : x_i < 0\},$$

$i = 1, 2, 3$ (6 charts altogether) which gives us an open cover of S^2. Define $\phi_1^\pm(x) = (x_2, x_3), \phi_2^\pm(x) = (x_1, x_3)$, and $\phi_3^\pm(x) = (x_1, x_2)$. We need to verify that changes of coordinates are smooth. Consider, for example, $\phi_2^+ \circ (\phi_1^+)^{-1}(u_1, u_2) = (\sqrt{1 - u_1^2 - u_2^2}, u_2)$, which is smooth in its region of definition. The other compositions yield similar results. It follows that S^2 is indeed a manifold.

Example 2.2.17. Now we consider a slightly more interesting example of a manifold, the real projective space $\mathbb{R}P^{n-1}$ which is, by definition, the space of lines through the origin in $\mathbb{R}^n$. To give $\mathbb{R}P^{n-1}$ a topology, we think of it as the set of equivalence classes of nonzero vectors in $\mathbb{R}^n$. That is,

$$\mathbb{R}P^{n-1} = (\mathbb{R}^n \setminus \{0\}) / \sim$$

where two non-zero vectors v and v' are equivalent ($v \sim v'$) if and only if there is a constant $\lambda \neq 0$ such that $v = \lambda v'$. Note that this is an equivalence relation. We then have a surjective map

$$\pi\colon \mathbb{R}^n \setminus \{0\} \to \mathbb{R}P^{n-1}, \quad \pi(v) = [v],$$

where $[v]$ denotes the equivalence class of v ($[v]$ is the line through v).

We put on $\mathbb{R}P^{n-1}$ the quotient topology: $U \subset \mathbb{R}P^{n-1}$ is open if and only if $\pi^{-1}(U)$ is open in $\mathbb{R}^n \setminus \{0\}$. I leave it to the reader to check that this topology is Hausdorff.

Charts here are given as follows: for each $1 \leq i \leq n$, let

$$U_i = \{[x_1, \dots, x_n] \in \mathbb{R}P^{n-1} : x_i \neq 0\}$$

and define

$$\phi_i \colon U_i \to \mathbb{R}^{n-1}$$

by

$$[x_1, \dots, x_n] \mapsto \left(\frac{x_1}{x_i}, \cdots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \cdots \frac{x_n}{x_i} \right).$$

Note that the inverse ϕ_i^{-1} is given by

$$\phi_i^{-1} \colon (x_1, \cdots, x_{n-1}) \mapsto [x_1, \cdots, x_{i-1}, 1, \cdots, x_n].$$

We must check that the change of coordinates maps are smooth. If $j < i$, then on the interesection $U_i \cap U_j$

$$\phi_j \circ \phi_i^{-1}(u_1, \cdots, u_{n-1}) = \phi_j(u_1, \cdots, u_{i-1}, 1, \cdots, u_n) = \left(\frac{u_1}{u_j}, \cdots, \frac{u_{i-1}}{u_j}, \frac{1}{u_j}, \cdots, \frac{u_n}{u_j} \right)$$

which is smooth. Other computations are similar (and are left to the reader).

Example 2.2.18. Define the complex projective space $\mathbb{C}P_{n-1}$ to be the set of complex lines through the origin in $\mathbb{C}^n$ and prove that it is a manifold.

Example 2.2.19. If M and N are manifolds, show that $M \times N$ is also naturally a manifold.

Example 2.2.20. Let V be a finite-dimensional vector space over $\mathbb{R}$. Then V is a manifold: a choice of basis $v_1, \dots, v_n$ ($n = \dim V$) of V defines a linear bijection $\sigma \colon \mathbb{R}^n \to V$, $\sigma(r_1, \dots, r_n) = \sum r_i v_i$. Define a topology on V by requiring that σ is a homeomorphism (that is, $U \subset V$ is open $\Leftrightarrow \sigma^{-1}(U) \subset \mathbb{R}^n$ is open). Check that this is indeed a Hausdorff second countable topology. Define $\sigma^{-1} \colon V \to \mathbb{R}^n$ to be a chart and $\{\sigma^{-1} \colon V \to \mathbb{R}^n\}$ to be an atlas (one chart!). Prove that a different choice of basis of V defines the same topology and an equivalent atlas.

Example 2.2.21. Let M be a manifold. Show that for each point $x \in M$ there is a coordinate chart $\phi \colon U \to \mathbb{R}^n$ with $x \in U$ such that $\phi(x) = 04and\phi(U)$ is $B_1(0)$, the ball of radius 1 centered at 0.

Remark 2.2.22. In Definition 2.2.10 we have made no assumption on the topology of our manifolds. It is standard to assume that the manifolds are Hausdorff. Otherwise all sorts of pathologies turn up. Another set of standard assumptions guarantees the existence of partitions of unity (see §2.4 below). For this the simplest assumption to make is that the manifold in question is *second countable*. However, this assumption is too stringent and paracompactness is much more reasonable. All of this will be discussed later on.

2.3 Maps of manifolds

In the Bourbakist view every area of mathematics has its collection of objects and its collection of maps between objects (or, more generally, morphisms). (*Note by the transcriber*: That is the view of the categorist S. Mac Lane rather than of Bourbaki.)
While it is enjoyable to make fun of Bourbaki and Bourbakists, there is some merit to this point of view. A map $f\colon M \to N$ between two manifolds is smooth if it is continuous and is smooth in coordinates. More precisely we have:

Definition 2.3.1 (smooth map)**.** Let M and N be two smooth manifolds with atlases $\{(U_\alpha, \phi_\alpha)\}$ and $\{(V_\beta, \psi_\beta)\}$, respectively. A continuous map $f\colon M \to N$ is a *smooth map* (or a *morphism* of C^∞ manifolds) if for all α and β with

$$f^{-1}(V_\beta) \cap U_\alpha \neq \varnothing,$$

the composition

$$\psi_\beta \circ f \circ \phi_\alpha^{-1}\colon \phi_\alpha(U_\alpha \cap f^{-1}(V_\beta)) \to \psi(V_\beta)$$

is C^∞.

We will write $C^\infty(M, N)$ to denote the set of all smooth maps from M to N. Note that this definition does not depend on which atlases on M and N we choose [check this]. Also note a special case of this deinition is that of a smooth function on a manifold, which is a map from M to $\mathbb{R}$. To wit

Definition 2.3.2. A function $f\colon M \to \mathbb{R}$ is *smooth* if f is continuous and if for all coordinate charts $\{U_\alpha, \phi_\alpha)\}$, $f \circ \phi_\alpha^{-1} \to \mathbb{R}$ is C^∞. It's consistent with the previous definition: we think of the real line $\mathbb{R}$ as a manifold with the standard coordinate chart $\mathrm{id}_{\mathbb{R}}\colon \mathbb{R} \to \mathbb{R}$. We denote the collection of all smooth functions on a manifold M by $C^\infty(M) = C^\infty(M, \mathbb{R})$.

Exercise 2.3.3. Let M be a manifold. Check that $C^\infty(M)$ is a vector space over the reals under the standard addition of functions and multiplication by scalars. Is it finite dimensional?

Exercise 2.3.4. Let M be a manifold. Check that a constant function on a manifold M is smooth.

Here are some examples of smooth maps.

Example 2.3.5. Take $M = \mathbb{R}^n \setminus \{0\}$, and let $N = \mathbb{R}P^{n-1}$. Let $\pi\colon \mathbb{R}^n \setminus \{0\} \to \mathbb{R}P^{n-1}$ be the projection $\pi(v) = [v]$. I claim that π is a smooth map. Let's check it.

The atlas on M is given by one chart — the inclusion ϕ of M into $\mathbb{R}^n$. The charts on $\mathbb{R}P^{n-1}$ are the same as last time. Note that $\pi^{-1}(U_i) = \{v \in \mathbb{R}^n \setminus \{0\} : v_i \neq 0\}$. To see that π is smooth, we need to check that $\phi_i \circ \pi \circ \phi^{-1}\colon \pi^{-1}()U_i) \to \mathbb{R}^{n-1}$ is C^∞. But note that

$$(\phi_i \circ \pi \circ \phi^{-1})(v) = \phi_i(\pi(v)) = \phi_i([v]) = \left(\frac{v_1}{v_i}, \cdots \frac{v_n}{v_i}\right).$$

Example 2.3.6. Let $M = \mathbb{R}$ with the coordinate chart $\phi(x) = x^3$. Let $N = \mathbb{R}$ with the coordinate chart $\psi(x) = x$. Let $f\colon M \to N$ be the map $x \mapsto x^3$. Is f a C^∞ map?

$$\psi \circ f \circ \phi^{-1})(x) = \psi \circ f(x^{1/3}) = \psi(x) = x,$$

which is smooth. So f is smooth.

Now let us see if the map $h\colon M \to N$, $h(x) = x$ is smooth. We have $\psi \circ f \circ \phi^{-1})(x) = x^{1/3}$, which is not differentiable at 0. So h is not smooth. Finally note that $f^{-1}\colon N \to M$ is smooth:

$$\phi \circ \psi^{-1}(x) = (x^{1/3})^3 = x.$$

Example 2.3.7. Constant functions are smooth maps of manifolds.

The appropriate notion of "isomorphism" in differential geometry is the following one:

Definition 2.3.8 (Diffeomorphism)**.** A C^∞ map $f\colon M \to N$ between two smooth manifolds is a diffeomorphism if f is a homeomorphism and both f and f^{-1} are C^∞ maps.

Example 2.3.9. The map $f\colon M \to N$ of Example 2.3.6 is a diffeomorphism.

Exercise 2.3.10. If M and N are manifolds, prove that $M \times N$ is diffeomorphic to $N \times M$.

Exercise 2.3.11. Show that the composition of smooth maps is smooth.

Exercise 2.3.12. Let $L_A\colon \mathrm{GL}(n,\mathbb{R}) \to \mathrm{GL}(n,\mathbb{R})$ be left multiplication by $A \in \mathrm{GL}(n,\mathbb{R})$. Prove that L_A is a diffeomorphism. [Recall that $\mathrm{GL}(n,\mathbb{R}) \subset \mathbb{R}^{n^2}$ is the set of all invertible $n \times n$ matrices and that it is open in $\mathbb{R}^{n^2}$.]

2.4 Partitions of unity

In this subsection we define partitions of unity (that is, writing the constant function 1 as a sum of bump functions with certain properties) and prove the existence of a partition of unity subordinate to a cover on a second countable manifold. The existence of such partitions of unity is very useful. The proof of the existence of the partition of unity is not terribly useful and should be skipped on the first (and second) reading. The reason for this advice is that the proof is technical and the techniques will never be used again in this course. We start with a string of definitions.

Definition 2.4.1 (second countable)**.** A topological space X is *second countable* if there is a countable collection of open subsets $\{U_i\}$ of X such that any open set in X is the union of some collection of U_i's. In other words, the topology of X has a countable basis.

Example 2.4.2. The real line $\mathbb{R}$ with the standard topology is second countable: the collection $\{U_i\}$ is consists of open intervals (a, b) where a and b are rational numbers.

Similarly $\mathbb{R}^n$ is second countable: the collection $\{U_i\}$ consists of open balls $B_r(x)$ of rational radius r centered at points x with rational coordinates.

Remark 2.4.3. Any (topological) subspace of a second countable space is second countable [prove it]. Hence any manifold that can be realised as a subspace of some $\mathbb{R}^n$ has to be second countable.

The condition of second countability is much more than necessary for the existence of the partition of unity. One can get away with assuming only paracompactness. Here, for the record, is its definition. It takes a paragraph to state because we have to define a few more things first.

Definition 2.4.4. Let M be a topological space. A collection $\{U_\alpha\}$ of subsets of M is a *cover* of a subset $W \subset M$ if $W \cup U_\alpha$. It is an *open cover* if each $\{U_\alpha\}$ is open. A *refinement* $\{V_\beta\}$ of a cover $\{U_\alpha\}$ is a cover such that for each index β there is an index $\alpha = \alpha\beta$ with $V_\beta \subset U_\alpha$.

A collection of subsets $\{U_\alpha\}$ of subsets of M is *locally finite* if for every point $m \in M$ there is a neighbourhood W of M with $W \cap U_\alpha \neq \varnothing$ for only finitely many α.

Example 2.4.5. The cover $\{(n, n+2)\}_{n\in\mathbb{Z}}$ is a locally finite cover of $\mathbb{R}$. The cover $\{[-\frac{1}{n}, \frac{1}{n}]\}$ is a cover of $(-1, 1)$ which is *not* locally finite — there is a problem at 0.

Definition 2.4.6 (paracompactness)**.** A topological space is *paracompact* if every open cover has a locally finite refinement.

Example 2.4.7. Any compact space is paracompact. We will see shortly that second countable Hausdorff manifolds are paracompact.

Definition 2.4.8 (support)**.** The *support* $\operatorname{supp} f$ of a continuous function $f\colon X \to \mathbb{R}$ is the closure of the set of points where f is non-zero:

$$\operatorname{supp} f = \overline{\{x \in X : f(x) \neq 0\}}.$$

Definition 2.4.9 (Partition of Unity)**.** Let $\{U_\alpha\}$ be an open cover of a manifold M. A *partition of unity subordinate to the cover* $\{U_\alpha\}$ is a collection of smooth functions $\{\rho_\beta\colon M \to [0,1]\}$ such that:

(1) For each index βthere is an index α with $\operatorname{supp}(\rho_\beta) \subset U_\alpha$.

(2) For each point $m \in M$, there is a neighbourhood W of m such that $\rho_\beta\_W \neq 0$ for only finitely many β. That is, the collection of supports $\{\operatorname{supp} \rho_\beta\}$ is locally finite.

(3) $\sum_\beta \rho_\beta = 1$.

Remark 2.4.10. Note that we need condition (2) to make sense of the sum in (33): by (2), for each point $m \in M$ the sum $\sum \rho_\beta(m)$ is actually a finite sum. So there are no problems with convergence.

Theorem 2.4.11. *Let M be a second countable Hausdorff manifold. Then every open cover of M has a partition of unity subordinate to it.*

Proof. (You should not read this proof the first time around.)
Step1. We first construct a collection $\{X_k\}_{k=1}^{\infty}$ of open subsets of M such that their closures $\overline{X}_k$ are compact, $\overline{X}_k \subset X_{k+1}$ and $M = \cup_{k=1}^{\infty} X_k$. Since M is second countable, there is a countable basis of the topology of M. Out of this collection of open sets choose those that have compact closure and denote them by $W_1, W_2, \ldots$. We claim that that they cover M: $M = \sup W_i$. Indeed, a point $x \in M$ has a neighbourhood homeomorphic to an open subset of $\mathbb{R}^n$ ($n = \dim M$, of course). For any point y in an open set $U \subset \mathbb{R}^n$ there is a closed ball $\overline{B}_r(y)$ centered at y with $\overline{B}_r(y) \subset U$. Closed balls in $\mathbb{R}^n$ are compact. Hence every point $x \in M$ has a neighbourhood $U(x)$ whose closure $\overline{U(x)}$ is compact. Now $U(x)$ is a union of a certain number of elements of the countable basis of the topology of M. The closure of each of these elements is compact. Therefore $x \in W_i$ for some index i. This proves that $M = \cup W_i$.

Let $X_1 = W_1$. The whole collection $\{W_i\}_{i=1}^{\infty}$ covers $\overline{X}_1$. Since $\overline{X}_1$ is compact, $\overline{X}_1 = W_{i_1} \cup W_{i_2} \cup \cdots \cup W_{i_p}$ for some $i_1 < i_2 < \cdots < i_p$. Let $X_2 = W_{i_1} \cup W_{i_2} \cup \cdots \cup W_{i_p}$. Then $\overline{X}_2$ is compact $\cdots$. Continuing in this manner we get the desired collection $\{X_k\}_{k=1}^{\infty}$.
Step2. We construct three open countable covers $\{V_{\beta,1}\}$, $\{V_{\beta,2}\}$, $\{V_{\beta,3}\}$ with $\{V_{\beta,1}\} \subset \{V_{\beta,2}\} \subset \{V_{\beta,3}\}$, $\cup_\beta \{V_{\beta,1}\} = M$ and $\{V_{\beta,3}\}$ is locally finite and subordinate to $\{U_\alpha\}$, the cover we started out with. Note that this will prove that any Hausdorff second countable manifold is paracompact, as promised.

Fix an index k. For each point $z \in \overline{X}_k \setminus X_{k-1}$ choose an open set $V_{z,3}$ such that $V_{z,3} \subset U_\alpha$ for some α, $V_{z,3} \subset X_{k+1}$ and $V_{z,3} \cap \overline{X}_{k-1} = \varnothing$. Additionally we require that there is a coordinate chart ψ_z mapping $V_{z,3}$ homeomorphically onto

$$B_3(0) := \{x \in \mathbb{R}^n | \|x\| < 3\}.$$

Let $V_{z,i} = psi_z^{-1} B_i(0)$ for $i = 1, 2$. The open sets $V_{z,1}$ cover the compact set $\overline{X}_k \setminus X_{k-1}$ (and are contained in $X_{k+1} \setminus \overline{X}_{k-2}$). Therefore, for each k, there is a *finite* collection of $V_{z,1}$'s covering $\overline{X}_k \setminus X_{k-1}$. Take all of these finite collections. We get a cover $\{V_{\beta,1}\}$ of M. Similarly we get two more covers: $\{V_{\beta,2}\}$ and $\{V_{\beta,3}\}$. Note that by construction they are locally finite and are subordinate to $\{U_\alpha\}$: for each β there is $\alpha(\beta)$ with $V_{\beta,i} \subset U_{\alpha(\beta)}$.
Step3. Now we construct a partition of unity. The function

$$f(t) = \begin{cases} \exp^{-\frac{1}{t}}, & \text{if} \quad t > 0 \\ 0, & \text{if} \quad t \leq 0 \end{cases}$$

is smooth on all of $\mathbb{R}$ [this fact is not entirely trivial]. Hence

$$\widetilde{f}(t) = \begin{cases} \exp^{-11-t}, & \text{if} \quad t < 1 \\ 0, & \text{if} \quad t \geq 1 \end{cases}$$

is smooth on all of $\mathbb{R}$. Therefore $h\colon \mathbb{R}^n \to [0, \infty)$ given by

$$h(x) = \widetilde{f}(\|x\|^2/4)$$

is also smooth. Note that $h(x) > 0$ for all $x \in B_2(0)$ and $h(x) = 0$ for all $x \notin B_2(0)$. Therefore, for each index β,

$$g_\beta(x) = \begin{cases} h(\psi_\beta)(x), & \text{if} \quad x \in V_{\beta,3} \\ 0, & \text{if} \quad x \notin V_{\beta,3}, \end{cases}$$

where $\psi_\beta \colon V_{\beta,3} \to B_3(0)$ is the corresponding coordinate chart, is a smooth function on M. Moreover, $g_\beta(x) > 0$ for $x \in V_{\beta,1}$. Since the cover $\{V_{\beta,3}\}$ is locally finite, the sum

$$G(x) = \sum_\beta g_\beta(x)$$

makes sense [converges for each $x4$] and defines a smooth function on M. Since $\{V_{/beta,1}\}$ covers M, $G(x) > 0$ for all $x \in M$. Let

$$\rho_\beta(x) = g_\beta(x)/G(x).$$

Then $1 \geq \rho_\beta(x) \geq 0$, $\sum \rho_\beta = 1$ and $\operatorname{supp} \rho_\beta \subset V_{\beta,3} \subset U_{\alpha(\beta)}$. Thus the collection $\{\rho_\beta\}$ is the desired partition of 1. □

Corollary 2.4.12. *Let M be a second countable Hausdorff manifold and $\{U_i\}_{i=1}^\infty$ a countable open cover. Then there is a partition of unity $\{\rho_i\}$ with supp $\rho_i \subset U_i$.*

Proof. By Theorem 2.4.11 there is a partition of unity $\{\tau_\beta\}$ with $\operatorname{supp} \tau_\beta \subset U_i$ for some $i = i(\beta)$. Let

$$I(i) = \{\beta | \operatorname{supp} \tau_\beta \subset U_i \quad \text{and} \quad \operatorname{supp} \tau_\beta \not\subset U_j \quad \text{for} \quad j < i\}.$$

Define

$$\rho_i = \sum_{\beta \in I(i)} \tau_\beta.$$

The collection $\{\rho_i\}$ is the desired partition of 1. □

Proposition 2.4.13. *Suppose that M is a second countable Hausdorff manifold, $K \subset M$ a closed subset and $U \subset M$ an open set with $K \subset U$. Then there is a smooth function $f \colon M \to [0,1]$ such that*

(1) $f|_K \equiv 1$ *and*

(2) $\operatorname{supp}(f) \subset U$.

Proof. Let $U_1 = U$ and $U_2 = M \setminus K$. By Corollary 2.4.12 there exists smooth functions $\rho_1, rho_2 \colon M \to [0,1]$ with $\operatorname{supp} \rho_i \subset U_i$ and $\rho_1 + \rho_2 = 1$. Since $\operatorname{supp} \rho_2 \subset M \setminus K$, $\rho_2|_K \equiv 0$. Hence $\rho_1|_K \equiv 1$. Now let $f = \rho_1$. □

Corollary 2.4.14. *Let M be a (second countable Hausdorff) manifold. uchor any point $x \in M$ and any neighbourhood U of x in M there is a smooth function $f \colon M \to \mathbb{R}$ such that*

(1) $f \equiv 1$ *on a neighbourhood* V *of* x *contained in* U *and*

(2) $\mathrm{supp}(f) \subset U$.

Proof. Exercise. You can use the proposition above. Alternatively prove it directly first in the case where $M = \mathbb{R}^n$ and then use a coordinate chart around x to prove it for arbitrary M. Is the condition that M is second countable really necessary? □

Chapter 3

Tangent vectors and tangent spaces

3.1 Tangent vectors and tangent spaces

We learn in physics that a vector is an arrow sticking out of a point in space and that a vector field assigns an arrow to each point in space. When we learn linear algebra, we are told to forget this point of view: all vectors are sticking out of one point — the origin. For the purposes of differential geometry the physics point of view is correct after all: all our vectors are anchored at various points in space.

There is another issue we need to deal with. If $S \subset \mathbb{R}^3$ is a smooth convex surface, one can imagine that for every point $p \in S$ there is a two-plane T_pS touching S at that point, a plane *tangent* to S at p. (It is not entirely clear that such a plane is unique, but that's another story.) A vector tangent to S at p would be an arrow anchored at p and lying in T_pS. This raises a problem: our manifolds are defined abstractly and not as subsets of some $\mathbb{R}^n$. So what would a tangent plane be in this case? and what vector space would it lie in?

There is another issue we need to deal with. If $S \subset \mathbb{R}^3$ is a smooth convex surface, one can imagine that for every point $p \in S$ there is a two-plane T_pS touching S at that point, a plane *tangent* to S at p. (It is not entirely clear that such a plane is unique, but that's another story.) A vector tangent to S at p would be an arrow anchored at p and lying in T_pS. This raises a problem: our manifolds are defined abstractly and not as subsets of some $\mathbb{R}^n$. So what would a tangent plane be in this case? and what vector space would it lie in?

The solution is to think of vectors as directional derivatives. A directional derivative of a function on $\mathbb{R}^n$ depends on two things: a direction and the point at which the function is being differentiated. For a smooth function $f \in C^\infty(\mathbb{R}^n)$, we write

$$D_v f(p) = \frac{d}{dt}|_0 f(p + tv)$$

for the directional derivative of f at a point $p \in \mathbb{R}^n$ in the direction $v \in \mathbb{R}^n$. Observe that

(1) the directional derivatives are *linear*: for any $f, g \in C^\infty(\mathbb{R}^n)$ and any $\lambda, \mu \in \mathbb{R}$

$$D_v(\lambda f + \mu g)(p) = \lambda D_v f(p) + \mu D_v g(p);$$

(2) the directional derivatives have a *derivation* property:

$$D_v(fg)(p) = f(p) D_v g(p) + D_v f(p) g(p).$$

This motivates the following definition:

Definition 3.1.1 (Tangent vector)**.** Let M be a manifold and $a \in M$ a point. A *tangent vector* to M at a is an $\mathbb{R}$-*linear* map $v\colon C^\infty(M) \to \mathbb{R}$ such that

$$v(fg) = f(a)v(g) + g(a)v(f) \tag{3.1.2}$$

for all functions $f, g \in C^\infty(M)$.

Linear maps $C^\infty(M) \to \mathbb{R}$ satisfying (3.1.2) are also said to have a *derivation* property and are called *derivations* (into $\mathbb{R}$).

Definition 3.1.3 (Tangent space)**.** The *tangent space* $T_a M$ to a manifold M at a point a is the collection of all tangent vectors to M at a.

Exercise 3.1.4. The tangent space $T_a M$ is a *vector space over the reals.* [That's why the elements of the tangent space are called "vectors"!] That is, if $v, w \in T_a M$ and $\lambda, \mu \in \mathbb{R}$ then the linear map $\lambda v + \mu w\colon C^\infty(M) \to \mathbb{R}$ is a derivation.

Note that by our definition every direction derivative at a point $p \in \mathbb{R}^n$ is a tangent vector at p to $\mathbb{R}^n$. This begs a question: are there tangent vectors that are not directional derivatives? The answer is no, tangent vectors to points of $\mathbb{R}^n$ are directional derivatives and that's all there is to it:

Proposition 3.1.5. *Let $w \in T_a\mathbb{R}^n$ be a tangent vector. That is, suppose $w\colon C^\infty(\mathbb{R}^n) \to \mathbb{R}$ is a linear map satisfying (3.1.2). Then*

$$w(f) = D_v f(a)$$

for some $v \in \mathbb{R}^n$. The same result holds with $\mathbb{R}^n$ replaced by some open ball $B_r(a)$.

To prove the proposition we first "recall" a version of Taylor's theorem.

Lemma 3.1.6. *Let f be a smooth function on $\mathbb{R}^n$. Fix a point $a \in \mathbb{R}^n$. Then for any $x \in \mathbb{R}^n$*

$$f(x) = f(a) + \sum (x_i - a_i) h_i(x) \tag{3.1.7}$$

where $h_i(x)$ are smooth functions with

$$h_i(a) = \frac{\partial f}{\partial x_i}(a).$$

Proof. Suppose first that $a = 0$. Then, by the fundamental theorem of calculus and chain rule,

$$f(x) - f(0) = \int_0^1 \frac{d}{dt} f(tx)dt = \int_0^1 \left(\sum x_i \frac{\partial f}{\partial x_i}(tx) \right) dt = \sum x_i \int_0^1 \frac{\partial f}{\partial x_i}(tx)dt.$$

Let $h_i(x) = \int_0^1 \frac{\partial f}{\partial x_i}(tx)dt$. These are the desired functions. If $a \neq 0$ apply the previous argument to $\overline{f}(x) = f(x - a)$. □

Remark 3.1.8. If f is a smooth function on an open ball $B_r(a)$ then (3.1.7) still holds at all $x \in B_r(a)$, except now $h_i \in C^\infty(B_r(a))$. The proof is exactly the same.

Before proving the proposition we need one more simple lemma.

Lemma 3.1.9. *Let M be a manifold and $w \in T_a M$ a tangent vector. Then for any constant function c we have $w(c) = 0$.*

Proof. Apply the tangent vector w to the constant function 1:

$$w(1) = w(1 \cdot 1) = 1w(1) + w(1)1 = 2w(1). \Rightarrow w(1) = 0.$$

Since w is linear, for any constant function $c = c \cdot 1$

$$w(c) = w(c \cdot 1) = cw(1) = 0.$$

□

of Proposition 3.1.5. By Lemma 3.1.6, $f(x) = f(a) + \sum (x_i - a_i)h_i(x)$. Hence

$$\begin{aligned} w(f) &= w(f(a)) + \sum (w(x_i - a_i)h_i(a) + (a_i - a_i)w(h_i)) \\ &= 0 + \sum w(x_i)h_i(a) + 0 = \sum w(x_i) \frac{\partial f}{\partial x_i}(a). \end{aligned}$$

Therefore $w = D_v f(a)$, where $v = (w(x_1), \ldots, w(x_n))$.

We leave the ball version of the proof as an exercise. □

Remark 3.1.10. The proof above actually shows that the derivations $\{\frac{\partial}{\partial x_i}|_a\}$ form a basis of $T_a\mathbb{R}^n$.

For arbitrary manifolds a choice of coordinates near a point also defines a basis of the tangent space at the point. To express this precisely it will be convenient to slightly change our notation. To this end, denote the points of $\mathbb{R}^n$ by $r = (r_1, \ldots, r_n)$. We also think of r_i as a function that assign to a point its i-th coordinate. If $\phi \colon U \to \mathbb{R}^n$ is a coordinate chart on a manifold M, then $\phi = (r_1 \circ \phi, \ldots, r_n \circ \phi)$. We then think of $x_i = r_i \phi$ as coordinate functions on U.

The coordinates de

ne tangent vectors at points of U: for any $a \in U$ and any $f \in C^\infty(M)$ we define $\frac{\partial}{\partial x_i}|_a$ by

$$\frac{\partial}{\partial x_i}|_a(f) := \frac{\partial}{\partial r_i}|_{\phi(a)}(f \circ \phi^{-1}).$$

It is easy to see that these are, indeed, tangent vectors. It should come as no surprise that they form a basis of the tangent space T_aM. After all, manifolds locally look like $\mathbb{R}^n$ and in $\mathbb{R}^n$ the partial derivatives do form bases of tangent spaces. Now let's prove this. We first observe that tangent vectors are local.

Lemma 3.1.11. *Let M be a manifold and $v \in T_aM$ a tangent vector. Then for any two functions $f, g \in C^\infty(M)$ with $f = g$ in a neighbourhood U of a, we have*

$$v(f) = v(g).$$

In particular, if h is constant on a neighbourhood U of a, then $v(h) = 0$ (cf. Lemma 3.1.9).

Proof. As $v\colon C^\infty(M) \to \mathbb{R}$ is $\mathbb{R}$-linear, it is enough to show that $v(f - g) = 0$. Chose a smooth bump function $\rho\colon M \to [0,1]$ with $\operatorname{supp} \rho \subset U$ which is identically 1 on a neighborhood V of a. We then have that $\rho \cdot (f - g) = 0$ on all of M by construction. Furthermore, because v is linear, $v(0) = 0$, hence

$$0 = v(\rho(f \circ g)) = v(\rho)(f \circ g)(a) + \rho(a)v(f \circ g) = v(f \circ g).$$

□

What's the point of the lemma, aside from its esthetic appeal? If $\phi = (x?1, \dots, x_n)\colon U \to \mathbb{R}^n$ is a coordinate chart on a manifold M and $v \in T_aM$ is a tangent vector at some point $a \in U$, then we cannot apply v to a coordinate function x_i. The function x_i is only defined on U; it is not a smooth function on all of M. However, there is a way around this problem. Pick a smooth bump function $\rho\colon M \to [0,1]$ with $\operatorname{supp} \rho \subset U$ which is identically 1 on some neighbourhood of a. Then $x_i\rho$ is a smooth function on M and so $v(x_i\rho)$ does make sense. Moreover, this number *does not depend on the choice of the bump function*: if $\tau\colon M \to [0,1]$ is another choice of a bump function with the same properties, then $x_i\rho = x_i\tau$ on some (perhaps smaller) neighborhood of a. Therefore, by the preceding lemma, $v(x_i\rho) = v(x_i\tau)$. We therefore *define*

$$v(x_i) := v(x_i\rho)$$

for some choice of the bump function ρ. Similarly, if $h \in C^\infty(U)$ we define

$$v(h) := v(h\rho)$$

for some (any) choice of the appropriate bump function ρ.

Lemma 3.1.12. *If $\phi = (x_1, /dotsc, x_n)\colon U \to \mathbb{R}^n$ is a coordinate chart on a manifold M and $v \in T_aM$ is a tangent vector at some point $a \in U$. Then*

$$v = \sum_i v(x_i) \frac{\partial}{\partial x_i}|_a. \tag{3.1.13}$$

Moreover, the vectors $\{\frac{\partial}{\partial x_i}\ {}_a\}$ *form a basis of* T_aM.

Proof. We evaluate both sides of (3.1.13) on a function $f \in C^\infty(M)$. It is no loss of generality to assume that $\phi(U)$ is a ball and that $\phi(a) = 0$. By Lemma 3.1.6,

$$(f \circ \phi^{-1})(r) = (f \circ \phi^{-1})(0) + \sum r_i h_i(r)$$

where $h_i(0) = \frac{\partial}{\partial r_i}(f \circ \phi^{-1})|_0$. Thus,

$$f(x) = f(a) + \sum x_i \cdot f_i(x),$$

where

$$f_i(a) = \frac{\partial}{\partial r_i}(f \circ \phi^{-1})(0) = \frac{\partial}{\partial x_i}|_a(f),$$

for all $x \in U$. Hence, for any $v \in T_aM$, we have

$$\begin{aligned} v(f) &= v(f(a) + \sum x_i f_i) \\ &= \sum x_i(a) v(f_i) + \sum v(x_i) f_i(a) \\ &= \sum v(x_i) f_i(a) \\ &= \sum v(x_i) \frac{\partial}{\partial x_i}|_a(f). \end{aligned}$$

This shows that $\{\frac{\partial}{\partial x_i}\ {}_a\}$ span T_aM. To check linear independence observe that

$$\frac{\partial}{\partial x_i}|_a(x_j) = \delta_{ij},$$

where δ_{ij} denotes the Kronecker delta function: it's 1 if $i = j$ and zero otherwise. □

Remark 3.1.14. We have seen in the preceding discussion that for any $p \in \mathbb{R}^n$ the tangent space $T_p\mathbb{R}^n$ is isomorphic to $\mathbb{R}^n$. Explicitly the isomorphism is give by taking a vector $v \in \mathbb{R}^n$ to the directional derivative at p in the direction of v:

$$\mathbb{R}^n \xrightarrow{\simeq} T_p\mathbb{R}^n \quad v \mapsto D_v(\cdot)(p).$$

In particular

$$\mathbb{R} \xrightarrow{\simeq} T_a\mathbb{R} \quad s \mapsto s\frac{d}{dr}|_a.$$

3.2 Digression: vector spaces and their duals

Given two (fiite dimensional) vector spaces V and W we denote the set of all linear maps from V to W by $\mathrm{Hom}(V, W)$. It is a vector space: any linear combination of two linear maps is again a linear map. Of special interest is the vector space $V^\vee := \mathrm{Hom}(V, \mathbb{R})$ of linear maps from a vector space V to $\mathbb{R}$, the so called *dual* vector space. If $\{v_i\}_{i=1}^n$ is a basis of V, the *dual basis* is a basis $\{v_i^\vee\}$ of $V^\vee$ defined by

$$v_i^\vee(v_j) = \delta_{ij}$$

for all $1 \leq i, j \leq n$. This is indeed a basis. If $\ell \in V^\vee$ is an arbitrary functional, then

$$\ell = \sum \ell(v_i) v_i^\vee$$

because both sides of the formula above agree on the basis vectors v_j (I am tacitly using the fact that if two linear maps $\mu, \nu \colon V \to \mathbb{R}$ agree on basis vectors, then they agree). It follows that $\dim V^\vee = \dim V$. Finally observe that for any vector $u \in V$,

$$u = \sum v_i^\vee(u) v_i.$$

Why is the formula above true? Apply $v_j^\vee$ to both sides.

Exercise 3.2.1. Show that a choice of basis of vector spaces V and W identifies $\mathrm{Hom}(V, W)$ with a space of matrices. Conclude that $\dim \mathrm{Hom}(V, W) = \dim V \cdot \dim W$.

3.3 Differentials

Definition 3.3.1. Let $f \colon M \to N$ be a smooth map of manifolds and $a \in M$ a point. The *differential* of f at a is the linear map

$$df_a \colon T_aM \to T_{f(a)}N \quad (df_a(v))(h) = v(h \circ f)$$

for all $v \in T_aM$ and all $h \in C^\infty(N)$.

Exercise 3.3.2. Check that the de

nition above makes sense. That is, given $v \in T_aM$, check that the map

$$C^\infty(N) \to /\mathbb{R}, \quad h \mapsto v(h \circ f)$$

is a linear map satisfying (3.1.2).

We will check shortly that in the case of a smooth map $f \colon \mathbb{R}^n \to \mathbb{R}^m$, $df_a = Df_a$ under the natural identification $T_a\mathbb{R}^n \simeq \mathbb{R}^n$.

We next sort out what the definition of a differential amounts to in the case where $f \colon M \to \mathbb{R}$ is a smooth *function* (in other words the target manifold $N = \mathbb{R}$). By definition 3.3.1, df_a is a map from T_aM to $T_{f(a)}\mathbb{R} \simeq \mathbb{R}$. That is, if we compose df_a with the isomorphism $T_{f(a)}\mathbb{R} \xrightarrow{\simeq} \mathbb{R}$ (see Remark 3.1.14), we get a linear map

$$\underline{df_a} \colon T_aM \to \mathbb{R}$$

By definition, $\underline{df_a}$ is an element of the dual vector space $T_a^\vee M := \mathrm{Hom}(T_aM, \mathbb{R})$. I claim that

Proposition 3.3.3. *The linear map $\underline{df_a}$ is given by*

$$\underline{df_a}(v) = v(f). \tag{3.3.4}$$

for any tangent vector $v \in T_aM$.

Proof. Let $r\colon \mathbb{R} \to \mathbb{R}$ denote the identity map. We think of it as the standard coordinates on $\mathbb{R}$. Then for every point $x \in \mathbb{R}$ the vector $\frac{d}{dr}|_x$ is a basis vector of $T_x\mathbb{R}$, which gives us an isomorphism

$$T_x\mathbb{R} \to \mathbb{R}, \quad t\frac{d}{dr}|_x \mapsto t.$$

The map above has a "coordinate free" description as well. It is:

$$T_x\mathbb{R} \ni v \mapsto v(r).$$

Therefore

$$\underline{df_a}(v) = (df_a(v))(r) = v(r \circ f) = v(f).$$

□

Remark 3.3.5. It is customary not to distinguish between df_a and $\underline{df_a}$. Thus, in the case of $f \in C^\infty(M)$, the differential df_a denotes both the linear map $df_a\colon T_aM \to T_{f(a)}\mathbb{R}$ and the linear functional $df_a\colon T_aM \to \mathbb{R}$. In other words, from now on we drop the notation $\underline{df_a}$ and write (3.3.4) as

$$df_a(v) = v(f). \tag{3.3.6}$$

for all $f \in C^\infty(M)$, $a \in M$, $v \in T_aM$.

Definition 3.3.7. The vector space

$$T_a^\vee M := \mathrm{Hom}(T_aM, \mathbb{R})$$

is called the *cotangent* space of M at a.

The new concept of the differential allows us to re-interpret the formula (3.1.13). Recall that a choice of coordinates $\phi = (x_1, \dots, x_n)\colon U \to \mathbb{R}^n$ on a manifold M gives rise to basis $\{\frac{\partial}{\partial x_i}|_a\}$ of T_aM for any point $ainU$. We claim that $\{(dx_i)_a\}$ form the dual basis of the cotangent space $T_a^\vee M$. Indeed, by (3.3.6),

$$(dx_j)_a\left(\frac{\partial}{\partial x_i}|_a\right) = \frac{\partial}{\partial x_i}|_a(x_j) = \delta_{ij}.$$

Since for $v \in T_aM$ we have $v(x_i) = (dx_i)_a(v)$, (3.1.13) becomes

$$v = \sum (dx_i)_a(v)\frac{\partial}{\partial x_i}|_a. \tag{3.3.8}$$

Let $f = (f_1, \dots, f_m)\colon \mathbb{R}^n \to \mathbb{R}^m$ be a smooth map. We are now in the position to compare $dfa\colon T_a\mathbb{R}^n \to T_{f(a)}\mathbb{R}^m$ with $Df_a\colon \mathbb{R}^n \to \mathbb{R}^m$. Let $r_1, \dots, r_n$ denote the standard coordinates on $\mathbb{R}^n$ and $s_1, \ldots, s_m$ the standard coordinates on $\mathbb{R}^m$. Using (3.3.8) we compute:

$$\begin{aligned}(ds_i)f(a)(df_a(\frac{\partial}{\partial r_j}|_a)) = (df_a(\frac{\partial}{\partial r_j}|_a))(s_i) &= \frac{\partial}{\partial r_j}|_a(s_i \circ f)\\ &= \frac{\partial}{\partial r_j}|_a(f_i)\\ &= \frac{\partial f_i}{\partial r_j}(a)\end{aligned}$$

Thus the matrix of the linear map $df_a\colon T_a\mathbb{R}^n \to T_{f(a)}\mathbb{R}^m$ with respect to the basis $\{\frac{\partial}{\partial r_j}|_a\}$ and $\{\frac{\partial}{\partial s_i}|_{f(a)}\}$ is the Jacobian matrix of Df_a.

It is worth singling out another special case of the definition of a differential of a map: $M = \mathbb{R}$. In this case $f\colon \mathbb{R} \to N$ is a smooth curve. We define the tangent vector to f at $t \in \mathbb{R}$ to be

$$f'(t) := df_t\left(\frac{d}{dr}|_t\right).$$

. Note that by definition $f'(t)$ is a tangent vector in $T_{f(t)}N$, the tangent space to N at $f(t)$.

Exercise 3.3.9. Let M be a manifold, $p \in M$ a point and $v \in T_pM$ a tangent vector at the point p. Show that there is a curve $\gamma\colon I \to M$ (where I is an open interval containing 0) with $\gamma(0) = p$ and $\gamma'(0) = v$.

We next observe that the chain rule holds for the differentials of smooth maps.

Theorem 3.3.10 (Chain Rule). *If $F\colon X \to Y$ and $H\colon Y \to Z$ are smooth maps of manifolds, then*

$$d(H \circ F)_a = dH_{F(a)} \circ dF_a$$

for any point $a \in X$.

Proof. Fix $a \in X$, $v \in T_aX$, and $f \in C^\infty(Z)$. Then

$$\begin{aligned}(d(H \circ F)_a(v))(f) &= v(f \circ (H \circ F))\\ &= v((f \circ H) \circ F)\\ &= (dF_a(v))(f \circ H)\\ &= (dH_{F(a)}(dF_a(v)))(f).\end{aligned}$$

□

Remark 3.3.11. Theorem 3.3.10 and Exercise 3.3.9 give us a useful way of computing differentials $df_a\colon T_aM \to T_{f(a)}N$. By the exercise, for any $v \in T_aM$ we can find a curve $\gamma\colon I \to M$ with $\gamma(0) = a$ and $\gamma'(0) = v$. Then, by the chain rule,

$$df_a(v) = df_a(\gamma'(0)) = df_a(d\gamma(\frac{d}{dr}|_0)) = d(f \circ \gamma)_0(\frac{d}{dr}|_0) = (f \circ \gamma)'(0).$$

Exercise 3.3.12. Prove that if $F\colon M \to N$ is a diffeomorphism then the differential $dF_a\colon T_aM \to T_{F(a)}N$ is an isomorphism.

Exercise 3.3.13. Let M and N be manifolds. Prove that for any $(a,b) \in M \times N$ the tangent space $T_{(a,b)}(M \times N)$ is isomorphic to $T_aM \times T_bN$.

Exercise 3.3.14. Suppose that $\gamma\colon \mathbb{R} \to \mathbb{R}^n$, $\gamma(t) = (\gamma_1(t), \dots, \gamma_n(t))$ is a smooth curve. Show that

$$d\gamma(\frac{d}{dt}) = \sum_i \gamma_i'(t)\frac{\partial}{\partial r_i},$$

where $\gamma'(t)$ are ordinary derivatives.

3.4 The tangent bundle

Definition 3.4.1 (provisional)**.** The tangent bundle TM of a manifold M is (as a set)

$$TM = \coprod_{a\in M} T_aM.$$

Note that there is a natural projection (the tangent bundle projection)

$$\pi\colon TM \to M$$

which sends a tangent vector $v \in T_aM$ to the corresponding point a of M.

We want to show that the tangent bundle TM itself is a manifold in a natural way and the projection map $\pi\colon TM \to M$ is smooth. Strictly speaking, we first should specify a topology on TM. However, our strategy will be different. We will first find candidates for coordinate charts on the tangent bundle TM. They will be constructed out of coordinate charts on M. We will check that the change of these candidate coordinates on TM is smooth. We will then use these candidate coordinates to manufacture a topology on TM.

Let $\phi = (x_1, \dots, x_n)\colon U \to \mathbb{R}^n$ be a coordinate chart on M. Out of it we construct a chart on TU. The first n functions come for free: we take the functions $x_1 \circ \pi, \dots, x_n \circ \pi$. Another set of n functions come for free also: by (3.3.8), given a vector $v \in T_aU$,

$$v = \sum (dx_i)_a(v)\frac{\partial}{\partial x_i}|_a.$$

Hence, abusing the notation a bit, we get maps

$$dx_i\colon TU \to \mathbb{R}, \quad TU \ni v \mapsto (dx_i)_a(v), \quad \text{where} \quad a = \pi(v).$$

Thus we define a candidate coordinate chart

$$\widetilde{\phi} := (x_1 \circ \pi, \dots, x_n \circ \pi, dx_1, \dots, dx_n)\colon TU \to \mathbb{R}^n \times \mathbb{R}^n$$

by

$$\widetilde{\phi}(v) = (x_1(\pi(v)), \dots, x_n(\pi(v)), (dx_1)_{\pi(v)}(v), \dots, (dx_n)_{\pi(v)}(v)).$$

If $\{U_\alpha, \phi_\alpha)\}$ is an atlas on M, we get a candidate atlas $\{(TU_\alpha, \widetilde{\phi}_\alpha)\}$ on TM. To see why this could possibly be an atlas, we need to check that the change of coordinates in this new purported atlas is smooth.

To this end pick two coordinate charts $(U, \phi = (x_1, \ldots, x_n))$ and $(V, \psi = (y_1, \ldots, y_n))$ on M with $U \cap V \neq \varnothing$. Then $T(U \cap V) = TU \cap TV \neq \varnothing$. Let

$$\widetilde{\phi} = (x_1, \ldots, x_n, dx_1, \ldots, dx_n) \colon TU \to \mathbb{R}^n \times \mathbb{R}^n$$

and

$$\widetilde{\psi} = (y_1, \ldots, y_n, dy_1, \ldots, dy_n) \colon TV \to \mathbb{R}^n \times \mathbb{R}^n$$

be the corresponding candidates charts on TM. Now let us compute the change of coordinates $\widetilde{\psi} \circ \widetilde{\phi}^{-1}$.

First, note that

$$\begin{aligned}&\widetilde{\phi}^{-1}(r_1, \ldots, r_n, u_1, \ldots, u_n)\\ &= \sum_i u_i \frac{\partial}{\partial x_i}|_{\phi^{-1}(r_1, \ldots, r_n, u_1, \ldots, u_n)} \in T_{\phi^{-1}(r_1, \ldots, r_n, u_1, \ldots, u_n)} M.\end{aligned}$$

So

$$\begin{aligned}&\widetilde{\psi}(\sum u_i \frac{\partial}{\partial x_i}|_{\phi^{-1}(r_1, \ldots, r_n)}) =\\ &(\psi(\phi^{-1}(r_1, \ldots, r_n)), dy_1(\sum_i u_i \frac{\partial}{\partial x_i}), \ldots, dy_n(\sum_i u_i \frac{\partial}{\partial x_i})).\end{aligned}$$

But

$$\begin{aligned}dy_j(\sum_i u_i \frac{\partial}{\partial x_i}) = \sum_i u_i (\frac{\partial}{\partial x_i}(y_j)) &= \sum_i \frac{\partial y_j}{\partial x_i} u_i\\ &= \sum_i \frac{\partial}{\partial r_i}(r_j(\psi \circ \phi^{-1})) u_i.\end{aligned}$$

Thus the change of the candidate coordinates is given by

$$\begin{aligned}\widetilde{\psi} \circ \widetilde{(\phi)}^{-1}(r_1, \ldots, r_n, u_1, \ldots, u_n) &= (\psi \circ \phi^{-1}(r), (\sum_i \frac{\partial y_1}{\partial x_i}(r) u_i, \ldots, \sum_i \frac{\partial y_n}{\partial x_i}(r) u_i))\\ &= \psi \circ \phi^1(r), \left(\frac{\partial y_j}{\partial x_i}(r)\right) \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}).\end{aligned} \tag{3.4.2}$$

where $r = (r_1, \ldots, r_n)$. Clearly $\widetilde{\psi} \circ \widetilde{\phi}^{-1}$ is smooth wherever it is defined. It remains to define a topology on TM so that the charts $\widetilde{\phi} \colon TU \to \phi(U) \times \mathbb{R}^n$ are homeomorphisms. We declare a subset $O \subset TM$ to be open if for any coordinate chart $\phi \colon U \to \mathbb{R}^n$ on M, the set $\widetilde{\phi}(O \cap TU) \subset \mathbb{R}^n \times \mathbb{R}^n$ is open.

Proposition 3.4.3. *The collection of open sets on TM defined above does indeed form a topology. Moreover, if M is Hausdorff and second countable, so is TM.*

Proof. Left as an exercise for the reader. □

We conclude that if M is an n-dimensional Hausdorff second countable manifold then ts tangent bundle TM is a $2n$-dimensional Hausdorff second countable manifold. Moreover, each coordinate chart $(x_1, \dots, x_n)\colon U \to \mathbb{R}^n$ on M gives rise to a coordinate chart $(x_1 \circ \pi, \dots, x_n \circ \pi, dx_1, \dots, dx_n)\colon TU \to \mathbb{R}^{2n}$.

Remark 3.4.4. The following notation is suggestive: we write $(m, v) \in TM$ for $v \in T_m(M)$. Strictly speaking, it is redundant since $m = \pi(v)$.

Remark 3.4.5. It is customary to simply write $x_i\colon TU \to \mathbb{R}$ for $x_i \circ \pi\colon TU \to \mathbb{R}$.

Exercise 3.4.6. Prove that the map $\pi\colon (TM) \to M$ is smooth and that the differential $d\pi_v\colon T_v(TM) \to T_{\pi(v)}M$ is surjective for all tangent vectors $v \in TM$. Hint: do it in (convenient) coordinates.

3.5 The cotangent bundle

As a set, the cotangent bundle $T^*M = (TM)^\vee$ is the disjoint union of cotangent spaces:

$$T^*M = (TM)^\vee = \coprod_{s \in M} T_a^*M = \coprod_{a \in M} (T_aM)^\vee.$$

Note that there is a natural projection (the cotangent bundle projection)

$$\pi\colon T^*M \to M$$

which sends a cotangent vector (a covector for short) $\eta \in T_a^*M$ to the corresponding point a of M. We make the cotangent bundle T^*M into a manifold in more or less the same way we made the tangent bundle into a manifold. That is, we manufacture new coordinate charts on T^*M out of coordinate charts on M and check that the transition maps between the new coordinate charts are smooth.

So let $\phi = (x_1, \dots, x_n)\colon U \to \mathbb{R}^n$ be a coordinate chart on M. Then for each point $a \in U$ the covectors $\{(dx_i)_a\}$ form a basis of T_a^*M. The partials $\{\frac{\partial}{\partial x_i}|_a\}$ form the dual basis. Hence for any $\eta \in T_a^*M$,

$$\eta = \sum \eta(\frac{\partial}{\partial x_i}|_a)(dx_i)_a.$$

Therefore the partials $\{\frac{\partial}{\partial x_i}\}$ give us coordinate functions on T^*U:

$$\frac{\partial}{\partial x_i}\colon T^*U \to \mathbb{R}^n, \quad T^*U \ni \eta \mapsto \eta(\frac{\partial}{\partial x_i}|_a),$$

where $a = \eta(\pi)$. We now define the candidate coordinates

$$\overline{\phi}\colon T^*U \to \mathbb{R}^n \times \mathbb{R}^n$$

by

$$\overline{\phi} = (x_1 \circ \pi, \ldots, x_n \circ \pi, \frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}).$$

Note that

$$\overline{\phi}^{-1}(r_1, \ldots, r_n, w_1, \ldots, w_n) = \sum_{i=1}^n w_i (dx_i)_{\phi^{-1}(r)} \in T^*_{\phi^{-1}(r)} M,$$

where again we have abbreviated $(r_1, \ldots, r_n)$ as r. We now check the transition maps. Let $\psi = (y_1, \ldots, y_n)\colon V \to \mathbb{R}^n$ be a coordinate chart on M with $V \cap U \neq \varnothing$. Then

$$\begin{aligned}\overline{\psi} \circ \overline{phi}^{-1}(r_1, \ldots, r_n, w_1, \ldots, w_n) &= \overline{\psi}(\sum_{i=1}^n w_i (dx_i)_{\phi^{-1}(r)}) \\ &= ((\psi \circ \phi^{-1})(r), \frac{\partial}{\partial y_1}(\sum_{i=1}^n w_i dx_i), \ldots, \frac{\partial}{\partial y_n}(\sum_{i=1}^n w_i dx_i)) \\ &= ((\psi \circ \phi^{-1})(r), \sum_i w_i \frac{\partial x_i}{\partial y_1}, \ldots, \sum_i w_i \frac{\partial x_i}{\partial y_n}).\end{aligned}$$

We conclude that

$$\overline{\psi} \circ \overline{\phi}^{-1}(r_1, \ldots, r_n, w_1, \ldots, w_n) = ((\psi \circ \phi^{-1})(r), \left(\frac{\partial x_i}{\partial y_j}(r)\right) \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}), \qquad (3.5.1)$$

which is smooth. The rest of the argument proceeds as in the case of the tangent bundle.

Remark 3.5.2. Later on, when we look at the general vector bundles, it will be instructive to compare the formulae for the change of coordinates in the tangent and the cotangent bundles. In particular note that the matrices $\left(\frac{\partial y_j}{\partial x_i}(r)\right)$ and $\left(\frac{\partial x_i}{\partial y_j}(r)\right)$ are inverse transposes of each other.

3.6 Vector fields

A vector field X on a manifold M smoothly assigns to a point $a \in M$ a tangent vector $X(a) \in T_a M$.[1] What does "smoothly" mean? If X is a vector field in $\mathbb{R}^n$ then

$$X(a) = \sum f_i(a) \frac{\partial}{\partial r_i}|_a$$

[1] Sometimes this is also written X_a.

for certain functions $f_i(a) \in \mathbb{R}$ of the point $a \in \mathbb{R}^n$. So whatever we mean by "smooth" should amount to the functions f_i being smooth. This suggests one definition of a smooth vector field:

Definition 3.6.1. A vector field X on a manifold M is *smooth* if for any coordinate chart $\phi = (x_1, \dots, x_n) \colon U \to \mathbb{R}^n$ we have, for any point $a \in U$,

$$X(a) = \sum f_i(a) \frac{\partial}{\partial r_i}|_a \tag{3.6.2}$$

for some smooth functions $f_i \colon U \to \mathbb{R}$.

There is something a bit unsatisfactory about this definition: is it possible that the functions f_i in (3.6.2) are smooth for one choice of coordinates and not smooth for another choice? So we will use it as as starting point for a better one. Note that the functions f_i in (3.6.2) are smooth for one choice of coordinates and not smooth for another choice? So we will use it as as starting point for a better one. are given by:

$$f_i(a) = (dx_i)_a(X(a)),$$

for any $a \in U$. Thus Definition 3.6.1 simply says that the composite

$$(x_1, \dots, x_n, dx_1, \dots, dx_n) \circ X \colon U \to \mathbb{R}^n \times \mathbb{R}^n$$

is smooth. But this is the same thing as saying that the map $X \colon M \to TM$ is smooth. Not every map $Z \colon M \to TM$ is a vector field: we need to make sure that $Z(a) \in T_aM$. The condition is equivalent to

$$\pi(Z(a)) = a$$

for all aßM. Here, as before, $\pi \colon TM! \to M$ is the natural projection. This gives us a slightly more "sophisticated" definition of a vector field:

Definition 3.6.3. A (smooth) *vector field* X on a manifold M is a smooth map $X \colon M \to TM$ such that $\pi \circ = \mathrm{id}$.

There is yet another definition of a vector field, which is quite useful from some points of view:

Definition 3.6.4. A smooth *vector field* X on a manifold M is a linear map $X \colon C^\infty(M) \to C^\infty(M)$ such that

$$X(fg) = fX(g) + gX(f) \quad \text{for all} \quad f, g \in C^\infty(M). \tag{3.6.5}$$

Proposition 3.6.6. *Definitions 3.6.3 and 3.6.4 are equivalent.*

Proof. Exercise.

Here are a few hints. Given a vector field $X \colon M \to TM$ define a map $\widetilde{X}$ from $C^\infty(M)$ to functions on M by

$$(\widetilde{X}(f))(a) = X_a(f)$$

for all $f \in C^\infty(M)$ and all $a \in M$. Check that $\widetilde{X}(f)$ is a smooth function and that the map $\widetilde{X}$ so defined is a derivation. That is, show that (3.6.5) holds with X replaced by $\widetilde{X}$.

Conversely, given a map $\widetilde{X}\colon C^\infty(M) \to C^\infty(M)$ with the derivation property as above, define $X\colon M \to TM$ by

$$X_a(f) = (\widetilde{X}(f))(a)$$

for all $f \in C^\infty(M)$ and all $a \in M$. Check that X_a is indeed a tangent vector in T_aM and that the map $X\colon M \to TM$, $a \mapsto X_a$ is smooth in a. □

Remark 3.6.7. From now on we will not distinguish between the two definitions and will think of vector fields as either smooth maps $M \to TM$ satisfying certain conditions or as $\mathbb{R}$-linear maps $C^\infty(M) \to C^\infty(M)$ satisfying the appropriate conditions. We will make no notation distinction between the two ways of looking at vector fields. Thus $X(a)$ will stand for the value of a vector field at a point a if a is a point. On the other hand, if f is a smooth function, $X(f)$ will stand for a new smooth function, the "derivative" of f with respect to the vector field X.

Notation 3.6.8. There are several standard ways to denote the space of all smooth vector fields on a given manifold M. The two most common ones are $\Gamma(TM)$ [vector fields are sections of the tangent bundle, see below] and $\chi(M)$.

Remark 3.6.9. 1. The space of vector fields$\Gamma(TM)$ is a vector space over $\mathbb{R}$: if $X, Y \in \Gamma(TM)$ are (smooth) vector fields and $\lambda, \mu \in \mathbb{R}$ are scalars, then their linear combination $\lambda X + \mu Y$ is defined by

$$(\lambda X + \mu Y)(a) := \lambda X(a) + \mu Y(a)$$

for any $a \in M$. It is again a smooth vector field.

2. We can also multiply vector fields on M by smooth functions: if $X \in \Gamma(TM)$ and $f \in C^\infty(M)$ then fX is defined by

$$(fX)(a) := f(a)X(a)$$

for all $a \in M$.

A fancy way of describing 2 is to say that $\Gamma(TM)$ is a *module* over the ring of smooth functions $C^\infty(M)$. See if you can impress your date.

If $X, Y \in \Gamma(TM)$ are two vector fields on a manifold M then it is not true that the $\mathbb{R}$-linear map

$$C^\infty(M) \to C^\infty(M), \quad f \mapsto X(Y(f)).$$

is a vector field — it does not have the correct derivation property. For example, if $M = \mathbb{R}$ and $X = Y = d/dt$, then $X(Y(f)) = f''$ and

$$(fg)' = (f'g + fg')' = f''g + 2f'g' + fg'' \neq f'g + fg''.$$

However,

Lemma 3.6.10. *Let $X, Y \in \Gamma(TM)$ be two smooth vector fields on a manifold M. Then the map*

$$[X, Y]\colon C^\infty(M) \to C^\infty(M), \quad f \mapsto X(Y(f)) - Y(X(f)) \tag{3.6.11}$$

is a vector field.

Proof. Clearly the map $[X, Y]$ is $\mathbb{R}$-linear. We need to check that it has the correct derivation property. This is a mechanical computation. Pick two functions $f, g \in C^\infty(M)$. Then

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

□

Definition 3.6.12. The *Lie bracket* of two vector fields X and Y on a manifold M is the vector field $[X, Y]$ defined by (3.6.11).

We now quickly recall the definitions of bilinear and skew-symmetric bilinear maps, the point being that Lie bracket will turn out to be a skew-symmetric bilinear map.

Definition 3.6.13. Let V, U and W be three vector spaces over the reals. A map

$$b\colon V \times U \to W$$

is *bilinear* if it is ($\mathbb{R}$-) linear in each argument: for all $u_1, u_2 \in U$, $c_1, c_2 \in \mathbb{R}$ and all $v \in V$,

$$b(v, c_1u_1 + c_2u_2) = c_1b(v, u_1) + c_2b(v, u_2);$$

and for all $v_1, v_2 \in V$, $c_1, c_2 \in \mathbb{R}$ and all $u \in U$,

$$b(c_1v_1 + c_2v_2, u) = c_1b(v_1, u) + c_2b(v_2, u).$$

Definition 3.6.14. A bilinear map $b\colon U \times U \to V$ is *skew-symmetric* if

$$b(u_1, u_2) = -b(u_2, u_1)$$

for all $u_1, u_2 \in U$.

It is easy to see that the Lie bracket on a manifold M is $\mathbb{R}$-bilinear and skew-symmetric. Note that it is *not* $C^\infty(M)$-bilinear:

$$[X, hY] = X(h)Y + h[X, Y]$$

for any $X, Y \in \Gamma(TM)$, $h \in C^\infty(M)$ (prove this).

Somewhat surprisingly the Lie bracket has a kind of derivation property:

Lemma 3.6.15 (Jacobi identity)**.** *For any three vector fields $X, Y, Z \in \Gamma(TM)$ on a manifold M*

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

Here is how one sees this as a derivation property: for a vector field $X \in \Gamma(TM)$ define

$$L_X \colon \Gamma(TM) \to \Gamma(TM)$$

by

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

With this definition (3.6.15) becomes:

$$L_X([Y, Z]) = [L_X(Y), Z] + [Y, L_X(Z)].$$

of Lemma 3.6.15. This is another computation that's easier to do yourself than watch someone else doing it. To keep the notation from getting out of hand, we will drop parentheses. Thus $XYZf$ stands for $X(Y(Z(f)))$ etc. We pick a function $f \in C^\infty(M)$ and compute:

$$\begin{aligned}([[X, Y], Z] + [Y, [X, Z]])f &= [X, Y]Zf - Z[X, Y]f + Y[X, Z]f - [X, Z]Yf \\ &= XYZf - YXZf - ZXYf + ZYXf \\ &\quad + YXZf - YZXf - XZYf + ZXYf \\ &= XYZf + ZYXf - YZXf - XZYf \\ &= X(YZf - ZYf) + (ZY - YZ)Xf = [X, [Y, Z]]f.\end{aligned}$$

This proves the Jacobi identity. □

Equation (3.6.15) is called the Jacobi identity and is often written as

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

(it is equivalent to (3.6.15) by skew-symmetry of $[\cdot, \cdot]$.

Definition 3.6.16. A (real) *Lie algebra* is a vector space V over $\mathbb{R}$ (possibly infinite dimensional) together with a map $[\cdot, \cdot] \colon V \times V \to V$, a Lie bracket, such that

(1) $[\cdot, \cdot]$ is bilinear,

(2) $[\cdot, \cdot]$ is skew-symmetric, and

(3) $[\cdot, \cdot]$ satisfies the Jacobi identity: for all $v, u, w \in V$

$$[u, [v, w]] = [[u, v], w] + [v, [u, w]].$$

Example 3.6.17. We have proved that the space of vector fields $\Gamma(TM)$ on a manifold M forms a Lie algebra.

Example 3.6.18. $\mathbb{R}^3$ with the cross (vector) product is a Lie algebra.

Remark 3.6.19. The bracket on a Lie algebra can be thought of as a multiplication. Note that it is *not* associative in general because of the Jacobi identity.

The geometric meaning of the Lie brackets of vector fields will be discussed later.

Chapter 4

Submanifolds and the implicit function theorem

Given a smooth function $F\colon \mathbb{R}^m \to \mathbb{R}^n$ and a point $c \in \mathbb{R}^n$ the level set

$$F^{-1}(c) := \{x \in \mathbb{R}^m | F(x) = c\}$$

may or may not be a smooth manifold. For example, take $f(x, y) = x^2 - y^2$, a smooth function on $\mathbb{R}^2$. Then $f^{-1}(0)$ is the union of two lines: $y = \pm x$. It is not a manifold. However, for $c \neq 0$, $f^{-1}(c)$ is a union of two smooth curves, hence a 1 dimensional manifold. The goal of this section is to describe a sufficient condition for the level sets $F^{-1}(c)$ to be manifolds. We then generalise this to level sets of smooth maps between manifolds. The key technical result that makes it all possible is the inverse function theorem.

4.1 The inverse function theorem and a few of its consequence

Theorem 4.1.1 (Inverse function theorem). *Let $U, U' \subset \mathbb{R}^n$, be open sets and $F\colon U \to U'$ a smooth map. Suppose for some point $a \in U$ the differential*

$$dF_a\colon \mathbb{R}^n \to \mathbb{R}^n$$

is invertible. Then there are open neighbourhoods U_0 of a in U and U_0' of $F(a)$ in U' such that

$$F\colon U_0 \to U_0'$$

is a diffeomorphism.

We will assume this result. (See Appendix : Chapter 5.1 for the proof.) It is not essential that U and U' are open subsets of $\mathbb{R}^n$ — any finite dimensional vector space will do. It is even true with $\mathbb{R}^n$ replaced by a Banach space. We now discuss various consequences of the inverse function theorem. The most famous one is the implicit function theorem. But first we prove the manifold version.

Proposition 4.1.2. *Let $f\colon N \to M$ be a smooth map of manifolds with $f(p) = q$ ($p \in N, q \in M$). Suppose*

$$df_p\colon T_pN \to T_qM$$

is an isomorphism (invertible linear map). There there are neighbourhoods U of $p \in N$, V of $q \in M$ so that

$$f_U\colon U \to V$$

is a diffeomorphism (invertible map with a smooth inverse).

Proof. Note first that if $\phi\colon U' \to \mathbb{R}^n$ is a coordinate chart on N then for any $z \in U'$ the map $d\phi_z\colon T_zN \to T_{\phi(z)}\mathbb{R}^n$ is an isomorphism (for instance if $\phi = (x_1, \ldots, x_n), d\phi_x(\frac{\partial}{\partial x_i}) = \frac{\partial}{\partial r_i})$.

So let $p \in U' \xrightarrow{\phi} \mathbb{R}^n$ and $q \in V' \xrightarrow{\psi} \mathbb{R}^m$ be two coordinate charts on M and N respectively. Then the diagram

$$\begin{array}{ccc} U' & \xrightarrow{\quad f \quad} & V' \\ \downarrow{\scriptstyle \phi} & & \downarrow{\scriptstyle \psi} \\ \phi(U') & \xrightarrow[(\psi\circ f\circ\phi^{-1})]{} & \psi(V') \end{array} \tag{4.1.3}$$

commutes: $\psi \circ f = (\psi \circ f \circ \phi^{-1}) \circ \phi$. Hence the diagram of differentials

$$\begin{array}{ccc} T_pN & \xrightarrow{\quad df_p \quad} & T_qM \\ \downarrow{\scriptstyle d\phi_p} & & \downarrow{\scriptstyle d\psi_q} \\ T_{\phi(p)}\phi(U') & \xrightarrow[d(\psi\circ f\circ\phi^{-1})_{\phi(p)}]{} & T_{\psi(q)}\psi(V') \end{array} \tag{4.1.4}$$

commutes as well. By the inverse function theorem, there are neighborhoods $\overline{U}$ of $\phi(p)$ and $\overline{V}$ of $\psi(q)$ such that

$$(\psi \circ f \circ \phi^{-1})|_{\overline{U}}\colon \overline{U} \to \overline{V}$$

is a diffeomorphism. Consequently,

$$f\colon \phi^{-1}(\overline{U}) \to \psi^{-1}(\overline{V})$$

is a diffeomorphism $\square$

Next we turn to the implicit function theorem, the vector space version.

Theorem 4.1.5 (Implicit function theorem)**.** *Let $F\colon \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^k$ be a smooth map, $(a,b) \in mathbbR^n \times \mathbb{R}^k$ a point and $c = F(a,b)$. Suppose that the restriction of the differential*

$$dF(a,b)|_{\{0\}\times\mathbb{R}^k}\colon \{0\} \times \mathbb{R}^k \times \mathbb{R}^k$$

is onto. Then there are neighborhoods U of $a2 \in \mathbb{R}^n$, W of (a,b) in $\mathbb{R}^n \times \mathbb{R}^k$ and a smooth map $g\colon U \to \mathbb{R}^k$ with $g(a) = b$ such that the

$$F^{-1}(c) \cap W = graph\{g\colon U \to \mathbb{R}^k\}.$$

That is, for $(x,y) \in W$

$$F(x,y) = c \Leftrightarrow y = g(x).$$

In other words the function g is implicitly defined by the equation $F(x, g(x)) = c$.

Proof. We write suggestively

$$\frac{\partial F}{\partial x}(a,b) \overset{def}{=} dF_{(a,b)|_{\mathbb{R}^n\times\{0\}}}, \quad \frac{\partial F}{\partial y}(a,b) \overset{def}{=} dF_{(a,b)|_{\{0\}\times\mathbb{R}^k}}.$$

Consider the smooth map $H\colon \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n \times \mathbb{R}^k$ defined by

$$H(x,y) = (x, F(x,y))$$

for all $(x,y) \in \mathbb{R}^n \times \mathbb{R}^k$. Then the differential of H at (a,b) is of the form

$$dH_{(a,b)} = \begin{bmatrix} I & 0 \\ \frac{\partial F}{\partial x}(a,b) & \frac{\partial f}{\partial y}(a,b) \end{bmatrix}$$

where $I\colon \mathbb{R}^n \to \mathbb{R}^n$ is the identity map. By assumption $\frac{\partial F}{\partial y}(a,b)$ is invertible. Hence $dH_{(a,b)}$ is invertible. By the inverse function theorem the function H is invertible on a neighborhood of (a,b). Let $G(u,v) = (G_1(u,v), G_2(u,v))$ denote its inverse, which is defined on a neighborhood of $H(a,b) = (a, F(a,b)) = (a,c)$. We may take this neighborhood to be of the form $U \times V$, with $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^k$ being open. Let $W = G(U \times V)$. Then

$$(u,v) = H(G(u,v)) = (G_1(u,v), F(G_1(u,v), G_2(u,v))$$

for all $(u,v) \in U \times V$. Hence $G_1(u,v) = u$. Therefore

$$F(u, G_2(u,v)) = v$$

for all $(u,v) \in U \times V$.

Conversely, if for any $(x,y) \in W$ we have $F(x,y) = v$ then

$$(x,y) = G(H(x,y)) = G(x, F(x,y)) = G(x,v) = (G_1(x,v), G_2(x,v))$$

and therefore $y = G_2(x,v)$. Define the function $g\colon U \to \mathbb{R}^k$ by

$$g(x) = G_2(x,c).$$

It is a smooth function and, by the above discussion,

$$F(x,y) = c \quad \Leftrightarrow \quad y = g(x)$$

for any $(x,y) \in W$. □

Remark 4.1.6. Here is a slightly different and ultimately more useful way to look at what we have proved. The argument above shows that there is a diffeomorphism

$$H\colon W \to U \times V$$

mapping bijectively the set

$$\{F = c\} \cap W := \{(x, y) \in W | F(x, y) = c\}$$

onto the set

$$H(W) \cap (\mathbb{R}^n \times \{c\})$$

This motivates the following definition.

Definition 4.1.7 (Submanifold)**.** Let M be an m-dimensional manifold. A subset $N \subset M$ is an n-dimensional *embedded submanifold* if for every point $q \in N$, there is a coordinate chart $\phi = (x_1, \dots, x_m)\colon U \to \mathbb{R}^m$ with $q \in U$ such that

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

That is, for all $a \in N \cap U$,

$$\phi(a) = (x_1(a), \dots, x_n(a), 0, \dots, 0).$$

Such charts are said to be *adapted to* N.

Example 4.1.8. The sphere S^2 is an embedded submanifold of $\mathbb{R}^3$. For example if $(x_1, x_2, x_3) \in S^2$ and $x_3 > 0$ then

$$\phi(x_1, x_2, x_3) = \left(x_1, x_2, x_3 - \sqrt{1 - x_1^2 - x_2^2}\right)$$

is a chart adapted to S^2 (and there are 5 more charts like this).

Thus the implicit function theorem says that, under certain conditions, portions of a level set of a map $F\colon \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^k$ are embedded submanifolds. Naturally the embedded submanifolds are manifolds in their own right.

Lemma 4.1.9. *If $N \subset M$ is an n-dimensional embedded submanifold of an m-dimensional manifold M then it is naturally an n-dimensional manifold in its own right, and the inclusion map $\iota\colon N \hookrightarrow M$, $\iota(a) = a$ is smooth.*

Proof. We make N into a topological space by giving it the subspace topology. If $\phi\colon U \to \mathbb{R}^m$ is a chart on M adapted to N, then

$$p \circ \phi|_N\colon N \cap U \to \phi(U) \cap \mathbb{R}^n$$

is a homeomorphism. Here $p\colon \mathbb{R}^m \to \mathbb{R}^n$ is the projection

$$p(x_1, \dots, x_n, \dots, x_m) = (x_1, \dots, x_n).$$

If $\psi \colon V \to \mathbb{R}^m$ is another chart adapted to N, then the map

$$\psi \circ \phi^{-1} \colon \phi(U \cap V) \to \psi(U \cap V), \quad \phi(U \cap V) \cap (\mathbb{R}^n \times \{0\}) \mapsto (U \cap V) \cap (\mathbb{R}^n \times \{0\})$$

is a diffeomorphism. Hence if $\{\phi_\alpha \colon U \to \mathbb{R}^m\}$ is a collection of charts on M adapted to N with $M = \cup U_\alpha$ then

$$\{p \circ \phi_\alpha|_{U_\alpha \cap N} \colon U_\alpha \cap N \to \mathbb{R}^n\}$$

is an atlas on N. Checking that the inclusion map ι is smooth is easy: in coordinates it's the inclusion

$$\mathbb{R}^n \to \mathbb{R}^m, \quad (r_1, \ldots, r_n) \mapsto (r_1, \ldots, r_n, 0, \ldots, 0)$$

□

We now generalise the implicit function theorem.

Proposition 4.1.10. *Let$F \colon \mathbb{R}^m \to \mathbb{R}^k$ be a smooth map and $c \in F(\mathbb{R}^m) \subset \mathbb{R}^k$ a point. Suppose that for all points $q \in F^{-12}(c)$ the differential*

$$dF_q \colon \mathbb{R}^m \to \mathbb{R}^k$$

is onto. Then the level set $F^{-1}(c)$ is a submanifold of $\mathbb{R}^m$ and (if $F^{-1}(c)$ is nonempty)

$$\dim F^{-1}(c) = m - k (= \dim \mathbb{R}^m - \dim \mathbb{R}^k).$$

Proof. Fix a point $q \in F^{-1}(c)$. Let $Z = \ker dF_p$. Let $X \subset \mathbb{R}^m$ be the vector space complement to Z so that

$$\mathbb{R}^m = Z \oplus X \simeq Z \times X.$$

We can thus think of a point $p \in \mathbb{R}^m$ as a pair $(z, x) \in Z \times X$. By assumption on dF_q and by construction of X, the restriction

$$dF_q|_X \colon X \to \mathbb{R}^k$$

is an isomorphism of vector spaces. We now proceed as in the proof of the implicit function theorem. Consider

$$H \colon Z \times X \to Z \times \mathbb{R}^k, \quad H(z, x) = (z, F(z, x)).$$

Write $\frac{\partial F}{\partial z}$ for $dF|_Z$ and $\frac{\partial F}{\partial x}$ for $dF|_X$.

Then the differential of H is of the form

$$dH(z, x) = \begin{bmatrix} I & 0 \\ \frac{\partial F}{\partial z} & \frac{\partial F}{\partial x} \end{bmatrix}.$$

By construction $\frac{\partial F}{\partial x}(q) \colon X \to \mathbb{R}^k$ is a bijection. Hence dH_q is a bijection. By the inverse function theorem there exist neighborhoods W of q in $\mathbb{R}^m$ and $U \times V$ of $H(q)$ in $Z \times \mathbb{R}^k$ such that $H \colon W \to U \times V$ is a diffeomorphism. Moreover, as in the proof of the implicit function theorem H maps bijectively $\{_{F=c\}\cap W}\}$ to $(U \times V) \cap (Z \times \{c\})$. Therefore $F^{-1}(c) = \{F = c\}$ is a submanifold of $\mathbb{R}^m$ of dimension $m - k$, i.e,

$$\dim Z = \dim \mathbb{R}^m - \dim \mathbb{R}^k.$$

□

Example 4.1.11. Consider $F : \mathbb{R}^n \to \mathbb{R}$, $F(x) = \sum x_i^2$. Then $dF_x = (2x_1, \dots, 2x_n)$. Hence dF_x is surjective for all nonzero x. In particular $F\ -1(1) = \{x \in \mathbb{R}^n | \sum x_i^2 = 1\}$ is a submanifold of $\mathbb{R}^n$ of dimension $n - 1$. This is, of course, the standard sphere of radius 1.

Definition 4.1.12 (Regular value). Suppose $f : M \to N$ is a smooth map of manifolds. A point $c \in N$ is a *regular value* of f if for all $x \in f^{-1}(c)$ the differential

$$df_x : T_x M \to T_c N$$

is surjective.

Note that Proposition 4.1.10 then simply states that non-empty preimages of a regular values of a map $F : mathbbR^m \to \mathbb{R}^k$ are submanifolds of $\mathbb{R}^m$.

Remark 4.1.13. Note that if $f^{-1}(c) = \varnothing$, then c is a regular value of f. It seems silly to construct a definition this way. The reason for the peculiar phrasing is that it makes easier to state Sard's theorem.

Theorem 4.1.14 (Sard's theorem). *Let $f : M \to N$ be a smooth map. Then the set of regular values of f is dense in M (and in fact its compliment has measure 0).*

Note that if $F : M \to N$ maps everything to one point $\{c\}$ then c is not a regular value (the differential of F is 0 everywhere), but $N \setminus \{c\}$ does consist of regular values. So Sard's theorem does hold for constant maps, except for the preimage of every regular value of a constant map is empty. It will take us too far afield to prove Sard's theorem here, so we will give a proof in the appendix (See §5.2.) On the other hand Proposition 4.1.10 nicely generalises to manifolds:

Theorem 4.1.15. *If c is a regular value of a smooth map of manifolds $f : M \to N$ and if $f^{-1}(c) \neq \varnothing$ then the level set $f^{-1}(c)$ is an embedded submanifold of M of dimension*

$$\dim f^{-1}(c) = \dim(M) - \dim(N).$$

Before we proceed with the proof of Theorem 4.1.15, we make an observation:

1. Let $\{\phi_\alpha : U_\alpha \to \mathbb{R}^m\}$ be an atlas on a manifold M. Suppose for some index β there is a diffeomorphism $\sigma : \phi_\beta(U_\beta) \to W \subset \mathbb{R}^m$ (W is some open set). Then

   (i) $\sigma \circ \phi_\beta : U_\beta \to \mathbb{R}^m$ is a chart on M,

(ii) this chart is copatible with the atlas $\{\phi_\alpha \colon y_\alpha \to \mathbb{R}^m\}$

These imply that

2. If Z is a submanifold of a manifold M and $H \colon M \to M'$ is a diffeomorphism, then $H(Z)$ is a submanifold of M'.

Proof. (*of Theorem 4.1.15*) It is enough to show that for every point a of $f^{-1}(c)$ there is a neighbourhood U of $f^{-1}(c)$ such that $U \cap f^{-1}(c)$ is a submanifold of U of dimension $m - n$.

Let $a \in f^{-1}(c)$ be a point. Let $\phi \colon U \to \mathbb{R}^m$ be a chart of M with $a \in U$ and $\psi \colon V \to \mathbb{R}^n$ be a chart on N with $c \in V$. Then

$$\psi \circ f \circ \phi^{-1} \colon \phi(U) \to V$$

is a smooth map. Moreover, by the chain rule,

$$*d(\psi \circ f \circ \phi{-}1)_0 = d\psi_c \circ df_a \circ d(\phi^{-1})_0.$$

Since $d\psi_c$ and $d\phi_a$ are isomorphisms and df_a is onto for any $a2 \in f^{-1}(c)$ by assumption,

$$d(\psi \circ f \circ \phi^{-1})_{\phi(a)} \colon T_{\phi(a)}\mathbb{R}^m \to T_{\psi(c)}\mathbb{R}^n$$

is onto for any $a \in f^{-1}(c) \cap U$. By Proposition 4.1.10 $(\psi \circ f \circ \phi^{-1})^{-1}(\psi(c)) = \phi(U \cap f^{-1}(c))$ is a submanifold of $\phi(U)$ of dimension $m - n$. Therefore $U \cap f^{-1}(c)$ is a submanifold of $U \subset M$ of dimension $m - n$. Since a is arbitrary, $f^{-1}(c)$ is a submanifold of M of the desired dimension. □

The next statement describes the tangent bundle of a regular level set $f^{-1}(c)$.

Corollary 4.1.16. *Suppose that c is a regular value of $f \colon M \to N$ and $f^{-1}(c) \neq \varnothing$. Then for all $a \in f^{-1}(c)$,*

$$T_a f^{-1}(c) = \ker(df_a).$$

Proof. Since

$$\dim T_a f^{-1}(c) = \dim f^{-1}(c) = \dim M - \dim N = \dim \ker df_a,$$

it is enough to prove that $T_a f^{-1}(a) \subset \ker df_a$. Let $v \in T_a f^{-1}(c)$ be a vector. By exercise 3.3.9 there is a curve $\gamma \colon I \to f^{-1}(c)$ (where I is the unit interval $[0,1]$) such that $\gamma(0) = a$ and $d\gamma\left(\frac{d}{dt}\right) = v$. Since $f \circ \gamma$ is a constant map, $d(f \circ \gamma)_0 = 0$. By the chain rule,

$$d(f \circ \gamma)_0 \left(\frac{d}{dt}\right) = df_{\gamma(0)}(d\gamma_0 \left(\frac{d}{dt}\right)) = df_a(v)$$

Therefore $T_a f^{-1}(c) \subset \ker df_a$ and we are done. □

Example 4.1.17. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be given by $f(x) = \sum x_i^2$. Then, as we have seen before, 1 is a regular value of f and $df_x = (2x_1, \ldots, 2x_n)$ for all $x \in \mathbb{R}^n$. Therefore, for any $x \in f^{-1}(1) = \mathbb{S}^{n-1}$ the tangent space $T_x\mathbb{S}^{n-1}$ is naturally isomorphic to $\ker\{v \mapsto \sum 2x_i v_i\}$, which is the $(n-1)$ dimensional hyperplane in $\mathbb{R}^n \simeq T_x\mathbb{R}^n$ orthogonal to the vector x.

Exercise 4.1.18. Show that $\mathrm{O}(n)$, the set of all $n \times n$ orthogonal matrices, is a submanifold of $\mathrm{GL}(n, \mathbb{R})$.
Hint: Consider the map $f\colon \mathrm{GL}(n,\mathbb{R}) \to \mathrm{Sym}(n,\mathbb{R})$ given by $A \mapsto AA^T$. Show that the identity matrix I is a regular value of f.

4.2 Transversality

We now have enough tools to do a bit of differential topology.

Definition 4.2.1 (Transversality)**.** A smooth map $F\colon M \to N$ of manifolds is *transverse* to a submanifold Z of N if for every $z \in Z$ and any $m \in F^{-1}(z)$, we have

$$T_z Z + dF_m(T_m M) = T_z N$$

Note that the sum is not necessarily a direct sum!.

Notation 4.2.2. We write $F \pitchfork Z$ if a map F is transverse to a submanifold Z.

Example 4.2.3. Let $N = \mathbb{R}^2$, $M = \mathbb{R}^3$, $Z = \mathbb{S}^2 \subset M$, the unit sphere. Let $f\colon N \to M$ is given by $f(x_1, x_2) = (x_1, x_2, 0)$. Then $f \pitchfork \mathbb{S}^2$.

Remark 4.2.4. A map $F\colon M \to N$ is transverse to submanifold Z consisting of one point c if and only if c is a regular value of F.

Example 4.2.5. Take $M = N = \mathbb{R}^2$. Consider $F\colon M \to N$ given by $F(x,y) = (x, x^2)$. Then F is transverse to $\{0\} \times \mathbb{R}$, but it is not transverse to $\mathbb{R} \times \{0\}$.

Theorem 4.2.6. *If a smooth map $F\colon M \to N$ of manifolds is transverse to a submanifold Z of N, then $F^{-1}(Z)$ is a submanifold of M. Moreover,*

$$T_a(F^{-1}(Z)) = (dF_a)^{-1}(T_{F(a)}Z),$$

for all $a \in F^{-1}(Z)$, and

$$\dim(M) - \dim(F^{-1}(Z)) = \dim(N) - \dim(Z).$$

Proof. We first consider a special case: assume that

$$N = \mathbb{R}^n, \quad Z = \mathbb{R}^k \times \{0\} \subset \mathbb{R}^k \times \mathbb{R}^{n-k} = \mathbb{R}^n.$$

Let $\pi\colon \mathbb{R}^k \times \mathbb{R}^{n-k} \to \mathbb{R}^{n-k}$ denote the canonical projection map. Then

$$\pi^{-1}(0) = \mathbb{R}^k \times \{0\} = Z,$$

hence

$$(\pi \circ F)^{-1}(0) = F^{-1}(Z).$$

Additionally, for all $a \in F^{-1}(Z)$

$$\begin{aligned} d(\pi \circ F)_a(T_aM) &= d\pi_{F(a)}(dF_a(T_aM)) = d\pi_{F(a)}(dF_a(T_aM) + T_{F(a)}Z) \\ &= d\pi_{F(a)}(\mathbb{R}^n) = \mathbb{R}^{n-k}, \end{aligned}$$

where for the second equality we used the fact that $d\pi_{F(a)}(T_{F(a)}Z) = 0$. Therefore 0 is a regular value of $\pi \circ F$ and consequently $(\pi \circ F) - -1(0) = F^{-1}(Z)$ is a submanifold of M. Moreover,

$$\begin{aligned} T_aF^{-1}(Z) &= T_a(\pi \circ F)^{-1}(0) = \ker d(\pi \circ F)_a = \ker\{d\pi_{F(a)} \circ dF_a\} \\ &= (dF_a)^{-1}(\ker d\pi_{F(a)}) = (dFa)^{-1}(T_{F(a)}Z). \end{aligned}$$

Finally, since $(d\pi \circ F)_a$ is surjective,

$$\dim F^{-1}(Z) = \dim(\ker(d\pi \circ F)_a) = \dim M - \dim \mathbb{R}^{n-k}.$$

Therefore

$$\dim M - \dim F^{-1}(Z) = \dim M - (\dim M - dim\mathbb{R}^{n-k}) = \dim N - \dim Z.$$

The general case follows from the following consideration. Since Z is an embedded submanifold for all $z \in Z$, there is a coordinate chart $\psi = (x_1, \dots, x_n)\colon N \to \mathbb{R}^n$ adapted to Z with $z \in V$. Hence $psi(Z) = \psi(V) \cap (\mathbb{R}^k \times \{0\})$. Now apply the previous argument to $\psi \circ F\colon F^{-1}(V) \to \mathbb{R}^n$ and $\psi(V) \cap (\mathbb{R}^k \times \{0\})$. □

Example 4.2.7. Consider two surfaces S_1 and S_2 in $\mathbb{R}^3$ such that $T_xS_1 \neq T_xS_2$ for every $x \in S_1 \cap S_2$. Then $T_xS_1 + T_xS_2 = \mathbb{R}^3$ for all $x \in S_1 \cap S_2$.

Let $F\colon S_1 \hookrightarrow \mathbb{R}^3$ be the inclusion map. Then $dF_x(T_xS_1) = T_xS_1$. Thus, F is transverse to S_2. By the theorem above $F^{-1}(S_2) = S_1 \cap S_2$ is a submanifold of S_1 of dimension 1. In other words, if two surfaces are nowhere tangent then they intersect in a collection of curves.

4.3 Embeddings, Immersions, and Rank

Definition 4.3.1 (Immersion). A smooth map of manifold $f\colon Z \to M$ is an *immersion* if its differential is injective at every point of Z.

Immersions need not be injective: consider the map $f\colon \mathbb{S}^1 \to \mathbb{S}^1$, $f(e^{i\theta}) = e^{2i\theta}$. It is a 2-1 map but its differential everywhere is a bijection.

Example 4.3.2. The inclusion map of a submanifold is a 1-1 immersion.

Definition 4.3.3 (Submersion). A map $f\colon M \to N$ between smooth manifolds is called a *submersion* if its differential at every point is surjective.

Exercise 4.3.4. Show that for any manifold M the canonical projection $\pi\colon TM \to M$ is a submersion — compute in the appropriate coordinates.

Exercise 4.3.5. Show that if $Z \subset M$ is an embedded submanifold, then $\pi^{-1}(Z) \subset TM$ is an embedded submanifold of the tangent bundle TM of M. Here again $\pi \colon TM \to M$ is the projection. Note that $\pi^{-1}(Z) = \cup_{a \in Z} T_a M$. It is often denoted by $TM|_Z$.

Definition 4.3.6 (Embedding)**.** A smooth map of manifold $f \colon Z \to M$ is an *embedding* if $f(Z) \subset M$ is an embedded submanifold and $f \colon Z \to f(Z)$ is a diffeomorphism.

This says, in particular, that every embedding is a 1-1 immersion. The converse is not true.

Chapter 5

Appendix

5.1 Inverse function theorem

This section is a slightly edited version of "Inverse Function Theorem" by Ethan Y. Jaffe. The reason that the transcriber adopted this one is two-fold:

1. It is written in the same spirit as "Baby-Rudin", and
2. it is self-contained.

Theorem 5.1.1 (Inverse Function Theorem)**.** *Let U be an open set in $\mathbb{R}^n$, and let $f\colon U \to \mathbb{R}^n$ be continuously differentiable, i.e., f is of class C^1. Suppose that $x_0 \in U$ and $Df(x_0)$ is invertible. Then there exists a smaller neighbourhood $V \ni x_0$ such that f is a homeomorphism onto its image. Furthermore, V may be taken small enough so that f^{-1} is also continuously differentiable, with its derivative satisfying $D(f^{-1})_y = (Df)^{-1}_{f^{-1}}(y)$. Moreover, if f is of class C^k, ($k \in \mathbb{N} \cup \{\infty\}$), then so is f^{-1}.*

The version of the proof presented here depends on a version of the Banach fixed point theorem with parameter, which we now state.

Theorem 5.1.2 (Banach Fixed Point Theorem)**.** *Let (X, d) be a complete metric space, and $T\colon X \to X$ be a contraction of factor $r < 1$, i.e. $d(Tx, Ty) \leq rd(x, y)$. Then T has a unique fixed point. Furthermore, if Λ is another metric space, and $T(\lambda)$ $\lambda \in \Lambda$ is a continuous family of contractions of factor r, that is,*

$$\lim_{\lambda \to \lambda_0} \sup_{x \in X} d(T(\lambda)x, T(\lambda_0)x) = 0$$

then the fixed points of $T(\lambda)$ are continuous of λ. Stated otherwise, if $x(\lambda)$ is the unique fixed point of $T(\lambda)$, then the map $\lambda \mapsto x(\lambda)$ is continuous.

Proof. First we show uniqueness. If $Tx = x$ and $Ty = y$, then

$$d(x, y) = d(Tx, Ty) \leq rd(x, y),$$

which is only possible if $d(x, y) = 0$, i.e., $x = y$.

Now for existence. Fix any $x_0 \in X$, and consider $y = \lim_{n\to\infty} T^n(x_0)$. If this exists, then

$$T(y) = T\left(\lim_{n\to\infty} T^n(x_0)\right) = \lim_{n\to\infty} T^{n+1}(x_0) = y,$$

since T is continuous. To prove convergence, notice that the sequence is Cauchy. Indeed, for any n it is easy to see inductively that

$$d(T^n(x_0), T^{n+1}(x_0)) \leq r^n d(x_0, T(x_0)).$$

By the triangle inequality, it follows that for $k \geq 1$

$$d(T^n(x_0), T^{n+k}(x_0)) \leq d(x_0, T(x_0)) \sum_{i=n}^{n+k-1} r^i \leq r^n \frac{d(x_0, T(x_0))}{1-r}$$

This upper bound is independent of k, so it follows that if $n, m \geq N$,

$$d(T^n(x_0), T^m(x_0)) \leq r^N \frac{d(x_0, T(x_0))}{1-r},$$

which shows that the sequence is Cauchy.

Now for the version with parameter. Observe that

$$\begin{aligned} d(x(\lambda), x(\lambda_0)) &= d(T(\lambda)x(\lambda), T(\lambda_0)x(\lambda_0)) \\ &\leq d(T(\lambda)x(\lambda), T(\lambda)x(\lambda_0)) + d(T(\lambda)x(\lambda_0), T(\lambda_0)x(\lambda_0)) \\ &\leq rd(x(\lambda), x(\lambda_0)) + d(T(\lambda)(x(\lambda_0)), T(\lambda_0)(x(\lambda_0))). \end{aligned}$$

Rearranging,

$$d(x(\lambda), x(\lambda_0)) \leq (1-r)^{-1} d(T(\lambda)x(\lambda_0), T(\lambda_0)x(\lambda_0)) \to 0$$

as $\lambda \to \lambda_0$ by continuity of the map $\lambda \mapsto T(\lambda)$. □

Now we prove the inverse function theorem.

Proof. (of Inverse Function Theorem) Translating and multiplying by a linear map, we may assume that $x_0 = 0$, $f(x_0) = 0$ and $Df_0 = \mathrm{id}$. Since f is continuously differentiable, Df_x remains close to Df_0 as matrices if x is close to 0. For $y \in \mathbb{R}^n$, with y close to 0, consider the map $T_y \colon x \mapsto x - f(x) + y$. Observe that a fixed point x of T_y is precisely an x for which $f(x) = y^2$.

The following part should be skipped for the first reading. Let us motivate the choice of this map. Let us use the notation $F(x) = f(x) - x$ of the sequel. Since $Df_0 = \mathrm{id}$, $F(x) \in o(1)$ can be thought of as a perturbation of the constant map 0, and hence $f = \mathrm{id} - F$ is a peturbation of the identity. We are seeking to solve $f(x) = y$, i.e. $(\mathrm{id} - F)(x) = y$.

One way to motivate the choice of the map T_y is to rearrange this equality into $\mathrm{id}(x) = F(x) + y = T_y(x)$, i.e., finding a fixed point for T_y. A more brute force approach, however, is to try to build a sequence of approximate solutions x_n to this equation, starting with $x_1 = y$, and iteratively improving the error. Explicitly, we may think of $F(x_n) + y = x_n + R_n$, where R_n is some error. To improve the error, we try to perturb by adding some z of size roughly R_n to x_n, and trying to solve $F(x_n + z) + y = x_n + z$. For such z,

$$F(x_n + z) = F(x_n) + DF_{x_n} z + o(|z|)$$

by definition of the derivative, and so we are trying to solve

$$F(x_n) + y + DF_{x_n} z + o(|z|) = x_n + z, \quad \text{i.e.} \quad R_n + DF_{x_n} z + o(|z|) = z.$$

Since x_n should be thought of as close to 0, DF_{x_n} is small, and thus $DF_{x_n} z + o(|z|)$ should be thought of as neglible compared to R_n if z is about size R_n. Thus $z = R_n$ is of size R_n and solves the previous equation modulo an error of

$$R_{n+1} = DF_{x_n} R_n + o(|R_n|),$$

which is of order smaller than R_n. Thus $x_{n+1} := x_n + z$ solves $F(x_{n+1}) + y = x_{n+1} + R_{n+1}$, and R_{n+1} is an improved error compared to R_n.

While this idea can be turned into a formal proof, one should just notice that

$$x_{n+1} = x_n + z = x_n + R_n = F(x_n) + y = T_y(x_n)$$

is just a fixed-point iteration, and so the formalisation of this proof may be abstracted away to an invokation of the fixed point theorem, anyway. The reader should also notice the similary of this approach to inverting the linear operator $\mathrm{id} - F$ on a Banach space, if $\|F\| < 1$, via the Neumann series $\sum_{k=0}^{\infty} F^k$; in this case, the choice of x_n above are precisely the partial sums of the Neumann series, truncated at $k = 2^{n-1}$, applied to y.

Returned from the paragragh to be skipped for the first reading. Let B_R denote the closed ball of radius $R > 0$ centred at 0. Note that B_R is a complete metric space. We will prove that if R is small enough, and y is small enough, T_y maps B_R to itself and is a contraction. We will use $\|x\|$ to denote the usual (ℓ^2) Euclidean norm on points, and for a linear map A, $\|A\|$ to denote the ℓ^2 operator norm.

Let us start by considering the map $F(x) = f(x) - x$. F is continuously differentiable with $DF_0 = 0$. Then for $R > 0$ small enough that $B_R \subseteq U$, and any two $x, x' \in B_R$,

$$\begin{aligned}
\|F(x) - F(x')\| &= \left\| \int_0^1 DF_{(x-x')t+x'} \cdot (x - x') dt \right\| \\
&\leq \int_0^1 \|DF_{(x-x')t+x'}\| \|x - x'\| dt \\
&\leq \left(\sup_{z \in B_R} \|DF_z\| \right) \|x - x'\|.
\end{aligned}$$

Since $DF_0 = 0$ and F is continuously differentiable, for all $0 < \varepsilon < 1$, if R is small enough, $(\sup_{z\in B_R}\|DF_z\|) \leq \varepsilon$. Fix such an ε.

Suppose $\|y\| \leq R(1-\varepsilon)$. Then we will show $T_y\colon B_R \to B_R$ is a contraction. Fix $x \in B_R$. Then we compute

$$\begin{aligned}\|T_y(x)\| &= \|x - f(x) + y\| \leq \|F(x)\| + \|y\| \\ &= \|F(x) - F(0)\| + \|y\| \\ &\leq \|x\| + R(1-\varepsilon) \leq R.\end{aligned}$$

Thus $T_y\colon B_R \to B_R$.

Now for the contraction. Fix $x, x' \in B_R$. Then we compute

$$\|T_y(x) - T_y(x')\| \leq \|F(x) - F(x')\| \leq \varepsilon\|x - x'\|.$$

By the fixed point theorem, T_y has a unique fixed point $x \in B_R$, i.e. if $\|y\|$ is small enough, there exists a unique solution x to $f(x) = y$ with $x \in B_R$. In other words, we have established the existence of $f^{-1}\colon B_{R(1-\varepsilon)} \to B_R$.

We still need to prove that f is a homeomorphism. In finite dimensions, we can appeal to the fact that a continuous bijection between compact subsets of $\mathbb{R}^n$ is a homeomorphism, and that is the end of the story.

For an extra bonus, we show that we do not need the assumption of finite dimensions, so we will use the version of the fixed point theorem with parameter. We just need to prove that f^{-1} is continuous, i.e. the fixed points of T_y are continuous in y. By the fixed point theorem, we just need to show that the map $y \to T_y$ is continuous, since they all have the same contractive factor ε. We easily compute for $y, y' \in B_{R(1-\varepsilon)}$.

$$\sup_{x\in B_R}\|T_y x - T_{y_0}x\| = \|y - y'\|,$$

which certainly tends to 0 as $y \to y'$. Thus f^{-1} is continuous. If $0 \in V \subseteq B_R$ is open, then restricting f to U, it follows that f is a homeomorphism onto its image, which we will call W. This completes the first part of the theorem

Now we need to show that f^{-1} is continuously differentiable. Shrinking V if necessary, we may assume that Df_x is nonsingular on V . Now we show that $f^{-1}\colon W \to V$ (which we know to be a homeomorphism) is differentiable on W, with derivative $(Df)^{-1}_{f^{-1}(y)}$. Since Df is non-singular and f^{-1} is continuous, this automatically shows that $(Df)^{-1}_{f^{-1}(y)}$ is continuous, and hence f^{-1} is continuously differentiable.

Fix $y_0 \in W$, and write $x_0 = f^{-1}(y_0)$, and for any $y \in W$ write $x = f^{-1}(y)$. Then since f is a homeomorphism

$$\begin{aligned}&\lim_{y\to y_0}\frac{f^{-1}(y) - f^{-1}(y_0) - (Df)^{-1}_{f^{-1}(y_0)}(y - y_0)}{\|y - y_0\|} \\ &= \lim_{x\to x_0}\frac{x - x_0 - (Df)_{x_0}(f(x) - f(x_0))}{\|f(x) - f(x_0)\|} \\ &= \lim_{x\to x_0} -Df^{-1}_{x_0}\left(\frac{f(x) - f(x_0) - Df_{x_0}(x - x_0)}{\|x - x_0\|}\right)\frac{\|x - x_0\|}{\|f(x) - f(x_0)|}.\end{aligned}$$

Since $Df_{x_0}^{-1}$ is a linear map, it is continuous, and so the first factor converges to 0 by definition of differentiability. The second factor is bounded above as $x \to x_0$. Indeed,

$$\begin{aligned}
&\liminf_{x \to x_0} \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} \\
&\geq \liminf_{x \to x_0} \left| \frac{\|Df_{x_0}(x - x_0)\|}{\|x - x_0\|} - \frac{\|f(x) - f(x_0) - Df_{x_0}(x - x_0)\|}{\|x - x_0\|} \right| \\
&= \liminf_{x \to x_0} \frac{\|Df_{x_0}(x - x_0)\|}{\|x - x_0\|} \geq c > 0,
\end{aligned}$$

since Df_{x_0} invertible means that there is some $c > 0$ for which $\|Df_{x_0}(x-x_0)\| \geq c\|x - x_0\|$. Putting these two things together means that

$$\lim_{y \to y_0} \frac{f^{-1}(y) - f^{-1}(y_0) - (Df)^{-1}_{f^{-1}(y_0)}(y - y_0)}{\|y - y_0\|} = 0,$$

i.e. f^{-1} is differentiable at y_0 with the desired derivative.

Lastly, we show that if f is C^k on V, then f^{-1} is C^k on W, without the need to shrink V. Because we do not shrink V, if we can show this is true for $k < \infty$, we automatically show it's true for $k = \infty$. First, observe that $\mathrm{GL}(n, \mathbb{R})$ is an open subset of $M_n(\mathbb{R}) \simeq \mathbb{R}^{n^2}$, and that the inversion map $I\colon \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$ is of class C^∞ (since it is just a rational function of the entries). If f is of class C^k, then the map $Df\colon V \to \mathrm{GL}(n, \mathbb{R})$ is of class C^{k-1}. Now, from the above, $D(f^{-1})\colon W \to \mathrm{GL}(n, \mathbb{R})$ is just

$$D(f^{-1}) = I \circ Df \circ f^{-1},$$

i.e., is the composition of three maps, the first of which is C^∞, and the second of which is C^{k-1}. This argument shows that that if f^{-1} is of class C^r for $r < k$, then $D(f^{-1})$ is of class C^r, too, so that f^{-1} is of class C^{r+1}. Starting with the case $r = 1$, which we know to be true, we obtain iteratively that f^{-1} is of class C^k, too. □

Remark 5.1.3. As mentioned in the motivation section, this proof easily extends to infinite dimensions, with the derivative replaced by the Fréchet derivative. Indeed, the only thing which needs changing is $\mathbb{R}^n$ to whichever Banach space X is in question, and changing the norms to the norms in the Banach spaces. The last part about f^{-1} inheriting the regularity of f does not quite carry through, as we have to manipulate maps that are k times continuously Fréchet differentiable like the following.

If $f\colon U \to W$ is a differentiable function at all points in an open subset U of V, it follows that its derivative

$$Df\colon U \to L(V, W)$$

is a function from U to the space $L(V, W)$ of all bounded linear operators from V to W. This function may also have a derivative, the second order derivative of f, which, by the definition of derivative, will be a map

$$D^2 f\colon U \to L(V, L(V, W)).$$

To make it easier to work with second-order derivatives, the space on the right-hand side is identified with the Banach space $L^2(V \times V, W)$ of all continuous bilinear maps from V to W. An element φ in $L(V, L(V, W))$ is thus identified with $\psi \in L^2(V \times V, W)$ such that for all x and y in V,

$$\varphi(x)(y) = \psi(x, y).$$

(Intuitively: a function φ linear in x with $\varphi(x)$ linear in y is the same as a bilinear function ψ in x and y).

One may differentiate

$$D^2 f\colon U \to L^2(V \times V, W)$$

again, to obtain the third order derivative, which at each point will be a trilinear map,

$$D^3 f\colon U \to L(V, L(V, L(V, W))) \simeq L^3(V \times V \times V, W)$$

and so on. The n-th derivative will be a function

$$D^n f\colon U \to L^n(V \times V \times \cdots \times V, W),$$

taking values in the Banach space of continuous multilinear maps in n arguments from V to W. Recursively, a function f is $n+1$ times differentiable on U if it is n times differentiable on U and for each x in U there exists a continuous multilinear map A of $n+1$ arguments such that the limit

$$\lim_{h_{n+1}\to 0} \frac{\|\Delta - A(h_1, h_2, \ldots, h_n, h_{n+1})\|}{\|h_{n+1}\|} = 0 \quad \text{where}$$
$$\Delta = D^n f(x + h_{n+1})(h_1, h_2, \ldots, h_n) - D^n f(x)(h_1, h_2, \ldots, h_n)$$

exists uniformly for $h_1, h_2, \ldots, h_n$ in bounded sets in V. In that case, A is the $(n+1)$st derivative of f at x.

Moreover, we may obviously identify a member of the space

$$L^n(V \times V \times \cdots \times V, W)$$

with a linear map

$$L(\bigotimes_{j=1}^{n} V_j, W)$$

through the identification

$$f(x_1, x_2, \ldots, x_n) = f(x_1 \otimes x_2 \otimes \cdots \otimes x_n)$$

thus viewing the derivative as a linear map.

5.2 Sard's theorem

This section is an excerpt from "Topology from the differentialble veiw point" by John Milnor. The reason for the seclection is quite simple: to the knowwledge of the transcriber, no other textbooks of manifolds and/or differential topology states and proves the theorem as completely as Milnor's book.

In general, it is too much to hope that the set of critical values of a smooth map be finite. But this set will be "small," in the sense indicated by the next theorem, which was proved by A. Sard in 1942 following earlier work by A. P. Morse.

Theorem 5.2.1. *Let $f\colon U \to \mathbb{R}^n$ be a smooth map, defined on an open set $U \subset \mathbb{R}^m$, and let*

$$C = \{x \in U \,|\, \operatorname{rank} df_x < n\}.$$

Then the image f(C) C Rn has Lebesgue measure zero, i.e, given any $\varepsilon > 0$, it is possible to cover $f(C)$ by a sequence of cubes in $\mathbb{R}^n$ having total n-dimensional volume less than ε.

Since a set of measure zero cannot contain any nonvacuous open set, it follows that the complement $\mathbb{R}^n \setminus f(C)$ must be everywhere dense in $\mathbb{R}^n$.

Note that f should be "reasonably" smooth for the proof which we will give later.

Before giving the proof, let us remark that We will be mainly interested in the case $m \geq n$. If $m < n$, then clearly $C = U$; hence the theorem says simply that $f(U)$ has measure zero.

More generally consider a smooth map $f\colon M \to N$, from a manifold of dimension m to a manifold of dimension n. Let C be the set of all $x \in M$ such that

$$df_x\colon TM_x \to TN_{f(x)}$$

has rank less than n (i.e. is not onto). Then C will be called the set of *critical points*, $f(C)$ the set of *critical values*, and the complement $N \setminus f(C)$ the set of *regular values* of f. (This agrees with our previous definitions in the case $m = n$.) Since M can be covered by a countable collection of neighbourhoods each diffeomorphic to an open subset of $\mathbb{R}^m$, we have:

Corollary 5.2.2 (A. B. Brown)**.** *The set of regular values of a smooth map $f\colon M \to N$ is everywhere dense in N.*

In order to exploit this corollary we will need the following (See Theorem 4.1.15):

Lemma 5.2.3. *If $f\colon M \to N$ is a smooth map between manifolds of dimension $m \geq n$, and if $y \in N$ is a regular value, then the set $f^{-1}(y) \subset M$ is a smooth manifold of dimension $Em - n$.*

Proof. Let $x \in f^{-1}(y)$. Since y is a regular value, the derivative df_x must map TM_x onto TN_y. The null space $\mathfrak{N} \subset TM_x$ of df_x will therefore he an $(m - n)$-dimensional vector space.

At this satege we take granted the Whitney imbedding theorem which states that any smooth manifold of dimension m can be embedded into $\mathbb{R}^k$ for some integer k (actually k can be assumed to be $2m+1$.) If $M \subset \mathbb{R}^k$, choose a linear map $L\colon \mathbb{R}^k to \mathbb{R}^{m-n}$ that is nonsingular on this subspace $\mathfrak{N} \subset TM_x \subset \mathbb{R}^k$. Now define

$$F\colon M \to N \times \mathbb{R}^{m-n}$$

by $F(\xi) = (f(\xi), L(\xi))$. The derivative dF_x is clearly given by the formula

$$dF_x(v) = (df_x(v), L(v)).$$

Thus dF_x is nonsingular. Hence F maps some neighbourhood U of x diffeomorphically onto a neighbourhood V of $(y, L(x))$. Note that $f^{-1}(y)$ corresponds, under F, to the hyperplane $y \times \mathbb{R}^{m-n}$. In fact F maps $f^{-1}(y) \cap U$ diffeomorphically onto $(y \times \mathbb{R}^{m-n}) \cap V$. This proves that $f^{-1}(y)$ is a smooth manifold of dimension $m - n$. □

As an example we can give an easy proof that the unit sphere $\mathbb{S}^{m-l}$ is a smooth manifold. Consider the function $f\colon \mathbb{R}^m \to \mathbb{R}$ defined by

$$f(x) = x_1^2 + x_2^2 + \cdots + x_m^2.$$

Any $y \neq 0$ is a regular value, and the smooth manifold $f^{-1}(1)$ is the unit sphere.

If M' is a manifold which is contained in M, it has already been noted that TM'_x is a subspace of TM_x for $x \in M'$. The orthogonal complement of TM'_x in TM_x is then a vector space of dimension $m - m'$ called *the space of normal vectors to* M' in M at x.

In particular let $M' = f^{-1}(y)$ for a regular value y of $f\colon M \to N$.

Lemma 5.2.4. *The null space of $df_x\colon TM_x \to TN_y$ is precisely equal to the tangent space $TM'_x \subset TM_x$ of the submanifold $M' = f^{-1}(y)$. Hence df_x maps the orthogonal complement of TM'_x isomorphically onto TN_v.*

Proof. From the diagram

$$\begin{array}{ccc} M' & \xrightarrow{i} & M \\ \downarrow & & \downarrow f \\ y & \longrightarrow & N \end{array}$$

we see that df_x maps the subspace $TM'_x \subset TM_x$ to zero. Counting dimensions we see that df_x maps the space of normal vectors to M' isomorphically onto TN_y. □