Differential Geometry

Jacob Denson

February 14, 2016

Table Of Contents

Part I

Manifold Theory

Chapter 1

Topological Considerations

In a mathematician's heaven, all objects would be linear, and finite dimensional at that. Unfortunately, we live in the real world. When a physicist describes the motion of a robot's arm, rigidity forces the joints to move along spherical curves, never linear. When an algebraic geometer studies the solution set of the equation $X^2 + Y^3 - 5$, he must analyze a shape which bends and curves, never straight. Differential geometry gives the mathematician tools to cheat – we work with shapes that, though non-linear, are *locally linear*. The challenge, of course, is to figure out how to put all the locally linear properties together into a nice, global form.

1.1 Manifolds

Topology attempts to describe the properties of space invariant under actions which stretch and squash continuously. Differential geometry extends this description to spatial properties constant when space is stretched and squashed, but not 'bent' in some form. Four centuries of calculus have established differentiability in the nice cartesian spaces $\mathbf{R}^n$. A basic environment to extend the notions of 'bentness' should then be to consider spaces which are locally similar to $\mathbf{R}^n$. These are the **topological manifolds**, a Hausdorff space which is *locally euclidean*. In detail, at every point p on a manifold, there exists a neighbourhood U of p, and a non-negative integer n such that U is homeomorphic to $\mathbf{R}^n$. We take $\mathbf{R}^0 = \{0\}$. Non-Hausdorff manifolds are far and few between, and occur naturally only when constructing unnatural paradoxes. We only consider Hausdorff manifolds in this volume. The category of all manifolds with continuous morphisms will be denoted **Man**.

Example. *$\mathbf{R}^n$ is a manifold. Any ball in Euclidean space is also a manifold; in these examples, we may simply take the entire space as the neighbourhood of each point, since a ball in $\mathbf{R}^n$ is homeomorphic to $\mathbf{R}^n$.*

Example. *Let $f : \mathbf{R}^n \to \mathbf{R}^m$ be a continuous function, and consider the graph*

$$\Gamma(f) = \{(x, f(x)) : x \in \mathbf{R}^n\}$$

Then $\Gamma(f)$ is a manifold, since it is homeomorphic to $\mathbf{R}^n$.

The above examples are easily extended to show that any topological space homeomorphic to a manifold is also a manifold! This is a bad omen, for we wish to discuss properties invariant under differentiability, which should not necessarily be invariant under homeomorphism! Clearly, we must additional structure to a manifold, discussed in the next chapter. For now, lets enjoy some topological delights.

Example. *Consider the circle $S^1 = \{x \in \mathbf{R}^2 : \|x\| = 1\}$. For any proper subset U of S^1, an* **angle function** *is a continuous function $\theta : U \to \mathbf{R}$ such that $e^{i\theta(x)} = x$ for all $x \in U$. This restriction immediately implies θ is an embedding, with inverse $\theta^{-1}(t) = e^{it}$. Angle functions exist on any proper subset U of S^1, and therefore cover S^1, which is shown to be a 1-manifold.*

The circle is different to $\mathbf{R}^n$ in the sense that we cannot put coordinates over the whole space at once, we must analyze the circle piece by piece to determine the structure on the whole space. This is the main trick to manifold theory – a manifold might be a big nasty object globally, but locally, the shape is pleasant.

Example. *There is a less obvious coordinate system on the circle which permits an easy generalization to higher dimensional space. We will project the open subset $S^n - \{(1, 0, \dots, 0)\}$ onto the hyperplane $\{-1\} \times \mathbf{R}^{n-1}$ by taking the intersection of this plane with the line that passes through the projected point x and $z = (1, 0, \dots, 0)$. The set of points that lie on the line connecting z and a point x on the sphere is*

$$\{z + [\lambda x + (1 - \lambda)z] : \lambda \in \mathbf{R}\} = \{[(2 - \lambda)z + \lambda x, \lambda x_2, \dots, \lambda x_n] : \lambda \in \mathbf{R}\}$$

To find the intersection on the hyperplane, we set the first coordinate equal to -1, and find the projection

$$f(x_1,\dots,x_n) = \frac{2}{1-x_1}(x_2,\dots,x_n)$$

In similar manner, by computing intersections of points on the hyperplane with the sphere, we obtain a much more nasty formula for the inverse function f^{-1},

$$\begin{aligned} f^{-1}(y_2,\dots,y_n) &= \left(1 - \frac{8}{4+\sum_{k=2}^n y_k^2}, \frac{4y_2}{4+\sum_{k=2}^n y_k^2}, \dots, \frac{4y_n}{4+\sum_{k=2}^n y_k^2}\right) \\ &= \frac{1}{4+\sum_{k=2}^n y_k^2}\left(\sum_{k=2}^n y_k^2 - 4, 4y_2, \dots, 4y_n\right) \end{aligned}$$

This truly is the inverse, and the functions are both continuous. If we project from the point $(-1,0,\dots,0)$ to $\{1\}\times\mathbf{R}^{n-1}$, then the homeomorphism defined on $S^1 - \{(-1,0,\dots,0)\}$ is calculated to be

$$g(x_1,\dots,x_n) = \frac{1}{1+x_1}(x_2,\dots,x_n)$$

$$g^{-1}(y_2,\dots,y_n) = \frac{1}{4+\sum_{k=2}^n y_k^2}\left(4 - \left(\sum_{k=2}^n y_k^2\right), 4y_2, \dots, 4y_n\right)$$

And we have covered S^n with homeomorphisms; the space is a manifold.

Any open subset of $\mathbf{R}^n$ is a manifold; around any point, we may take U to be an open ball, and any open ball is homeomorphic to $\mathbf{R}^n$. In fact, any open subset of a manifold, with the subspace topology, is also a manifold, known as an **open submanifold**. When we analyze manifolds, it is convenient to consider not only homeomorphisms onto $\mathbf{R}^n$, but also maps onto open submanifolds of $\mathbf{R}^n$. We will call a homeomorphism $x: U \to V$, where V is an open subset of $\mathbf{R}^n$ a **chart**, and denote it (x,U). The letters x,y and z are often used for charts, so that it is easy to confuse coordinates $(x^1,x^2,\dots,x^n)$ in $\mathbf{R}^n$ with coordinates $(x^1(p),x^2(p),\dots,x^n(p))$ on a manifold.

Example. *Consider the set $M_n(\mathbf{R})$ of $n\times n$ matrices with entries in the real numbers. We can identify $M_n(\mathbf{R})$ with the space $\mathbf{R}^{n\times n}$, and therefore $M_n(\mathbf{R})$ is a topological manifold of dimension n^2. The determinant map* $\det: M_n(\mathbf{R}) \to \mathbf{R}$

can be viewed as a polynomial in the entries of the matrix, so the function is continuous, and

$$GL_n(\mathbf{R}) = \det^{-1}(\mathbf{R} - \{0\})$$

hence $GL_n(\mathbf{R})$ is an open submanifold of $M_n(\mathbf{R})$.

Example. *Let $M(n,m;k)$ be the set of n by m matrices of rank k. For any $X_0 \in M(n,m;k)$, there are permutation matrices P and Q such that*

$$PX_0Q = \begin{pmatrix} A_0 & B_0 \\ C_0 & D_0 \end{pmatrix}$$

where A_0 is an invertible k by k matrix. The map

$$L : X \mapsto PXQ = \begin{pmatrix} A(X) & B(X) \\ C(X) & D(X) \end{pmatrix}$$

is a rank-preserving linear endomorphism on $M(n,m)$, and locally around X_0, the resulting $A(X)$ are invertible. Fix X, and consider the rank-preserving linear map

$$T : Y \mapsto \begin{pmatrix} I_k & 0 \\ -C(X)A^{-1}(X) & I_{n-k} \end{pmatrix} Y$$

Notice that

$$(T \circ L)(X) = \begin{pmatrix} A(X) & B(X) \\ 0 & -C(X)A^{-1}(X)B(X) + D(X) \end{pmatrix}$$

It follows that X is rank k if and only if $D(X) = C(X)A^{-1}(X)B(X)$. We find that locally around $T(X_0)$ we may specify an element of $M(n,m;k)$ via a $k \times k$ invertible A, a $k \times (n-k)$ matrix B, and a $k \times (m-k)$ matrix C. But this is the same anywhere in $M(n,m;k)$, because the linear map $X \mapsto PXQ$ just swaps coordinates. Thus $M(n,m;k)$ is a

$$k^2 + k(n-k) + k(m-k) = k(n+m-k)$$

dimensional manifold.

$M(n,m;k)$ is an interesting manifold, since it embeds itself in a space in such a way that it is almost linear. The directions the space travels in relative to $M(n,m)$ are always along axis of the space. The best way to see this is to consider $M(2,1;1)$, which consists of vectors in $\mathbf{R}^2$ of the form

$(a, 0)$ and $(0, b)$, with $a, b \neq 0$. This is the x and y axis, with the origin removed, and is a 1 manifold.

Many proofs about manifolds use a reliable trick. First, we conjure forth local homeomorphisms to $\mathbf{R}^n$. Then we transport nice properties of $\mathbf{R}^n$ across the homeomorphism, thereby inducing the properties on the manifold. In fact, the general philosophy of manifold theory is that most properties of $\mathbf{R}^n$ will carry across to arbitrary spaces that look locally like $\mathbf{R}^n$ – we can perform linear algebra on spaces that are not really linear!

Theorem 1.1. *Every manifold is locally compact.*

Proof. Let x be an arbitrary point on a manifold, with an open neighbourhood U homeomorphic to $\mathbf{R}^n$ by a map $f : U \to \mathbf{R}^n$. Take any ball B around $f(x)$, whose closure $\overline{B}$ is compact. Since compactness is topologically invariant, $f^{-1}(\overline{B})$ is a compact neighbourhood of x. $\square$

The same method shows that every manifold is locally path-connected, and thus locally connected. The next problem requires more foresight on the reader, though the basic technique used is exactly the same.

Theorem 1.2. *A connected manifold is path-connected.*

Proof. Let x be a point on a connected manifold M, and consider the set U of all points in M path connected to x. Local path connectedness shows U is open. Suppose y is a limit point of P. Take some neighbourhood V of y homeomorphic to $\mathbf{R}^n$. Then V contains a point $p \in U$, which is path connected to x and y, since $V \cong \mathbf{R}^n$ is path connected. But then x is connected to y. Since U is non-empty, $U = M$. $\square$

Since every manifold is locally connected, any manifold can be split up into the disjoint sum of its connected components. It is therefore interesting to prove theorems about connected manifolds, since any manifold can be built up as a disjoint union of connected manifolds.

Example. *$GL_n(\mathbf{R})$ is a disconnected manifold, since $\det(GL_n(\mathbf{R}))$ is disconnected. By Corollary (1.3) we should be able to identify the path connected components. We shall construct paths which represent operations by elementary matrices, thereby reducing a matrix to a canonical form by a series of paths.*

Let $v_1, \ldots, v_n$ be arbitrary row vectors in $\mathbf{R}^n$. Consider adding a row to another row in a matrix,

$$(v_1, \ldots, v_p, \ldots, v_q, \ldots, v_n)^t \mapsto (v_1, \ldots, v_p + v_q, \ldots, v_q, \ldots, v_n)^t$$

Every pair of matrices of this form are path connected in $GL(n)$ to its image above by the map

$$t \mapsto (v_1, \ldots, v_p + tv_q, \ldots, v_q, \ldots, v_n)$$

Subtracting rows is similarly a path-connected operation. Next, consider multiplying a row by a scalar $\gamma > 0$,

$$(v_1, \ldots, v_p, \ldots, v_n) \mapsto (v_1, \ldots, \gamma v_p, \ldots, v_n)$$

A path that connects the two matrices is defined by

$$t \mapsto \begin{pmatrix} v_1 & \ldots & [1 + t(\gamma - 1)]v_p & \ldots & v_n \end{pmatrix}^t$$

This path only remains in $GL_n(\mathbf{R})$ if $\gamma > 0$. We should not expect to find a path when $\gamma < 0$, since multiplying by a negative number reverses the sign of the determinant, and we know from the continuity of the determinant that the sign of the determinant separates into at least two connected components. The same reason shows we can't necessarily swap two rows. Fortunately, we don't need these operations – we may use the path-connected elementary matrices to reduce any matrix to a canonical form. A modification of the Gauss Jordan elimination algorithm (left to the reader as a simple exercise) shows all matrices can be path-reduced to a matrix of the form

$$\begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & & 1 & 0 \\ 0 & 0 & \ldots & 0 & \pm 1 \end{pmatrix}$$

One matrix has determinant greater than zero, the other has determinant less than zero. Thus $GL_n(\mathbf{R})$ consists of two homeomorphic path-connected components: the matrices with determinant greater than zero, and the component with determinant less than zero. This is quite a different situation from $GL_n(\mathbf{C})$, which is always connected.

1.2 Products and Quotients

The set of manifolds form a category. It would be useful therefore to find common constructions which work in general categories. The coproduct (disjoint-union) of two manifolds is easily shown to be a manifold. Here are some more constructions.

Example (Manifold Products and the Torus). *If M and N are manifolds, then the product $M \times N$ is also a manifold – we simply take products of homeomorphisms on the space. Since S^1 is a 1-manifold, we obtain a 2-manifold $T = S^1 \times S^1$, the Torus. More generally, the n-torus $S^1 \times S^1 \times \cdots \times S^1$ is an n-manifold.*

Most quotient spaces of manifolds will not be manifolds. Nonetheless, under some restrictions, the quotient space will be a manifold. It shall suffice that if $f : M \to N$ is a locally injective open surjective map, and M is a manifold, then N is a manifold.

Example (The Möbius Strip). *Consider the quotient space M obtained from $[-\infty,\infty]\times(-1,1)$ by identifying (x,y) with $(x+n,(-1)^n y)$. Then the projection is open and locally injective, so M is a manifold, known as the Möbius strip. By throwing away points, we find M can also be obtained from the product space $[-1,1] \times (-1,1)$ by identifying $(-1,x)$ with $(1,-x)$, for each $x \in (0,1)$. It only has one edge, even though it exists in three dimensional space, and if you have a paper copy at hand, try cutting it down the middle!*

Example (Projective Space). *Consider the quotient space of S^2 obtained by identifying opposite sides of the sphere: glue each point x to $-x$. The projection is locally injective, so the space is a 2-manifold, denoted $\mathbf{P}^2$ and known as projective space. In general $\mathbf{P}^n$ is created by identifying opposite points on $\mathbf{S}^n$. To obtain explicit homeomorphisms on $\mathbf{P}^n$, define a chart x on $\mathbf{S}^n$ by*

$$x(a_1,\dots,a_n) = \frac{1}{a_i}(a_1,\dots,a_{i-1},a_{i+1},\dots,a_n)$$

This map is continuous everywhere but where $a_i = 0$. For all points p, $f(p) = f(-p)$, so we may define the map on $\mathbf{P}^n$ instead, and this map will be continuous since the projection map is open. It even has a continuous inverse, defined by

$$x^{-1}(b_1,\dots,b_{n-1}) = \left[\frac{1}{\sqrt{\sum b_i^2 + 1}}(b_1,b_2,\dots,1,\dots,b_n) \right]$$

Since our maps cover the space, $\mathbf{P}^n$ is a manifold.

Even though $\mathbf{P}^2$ is locally trivial, the space is very strange globally, and cannot be embedded in $\mathbf{R}^3$. Nonetheless, the geometry our eyes percieve is modelled very accurately by the spherical construction of projective space. We don't see the really weird part of $\mathbf{P}^2$, since our eye cannot see a full circumpherence of vision.

Example (Gluing Surfaces). *Let M and N be connected n-manifolds. We shall define the connected sum $M\#N$ of the two manifolds. There are two sets B_1 and B_2 in M and N respectively, both homeomorphic to the closed unit ball in $\mathbf{R}^n$. Then there is a homeomorphism $h : \partial B_1 \to \partial B_2$, and we may define the connected sum as*

$$M\#N = (M - B_1^\circ) \cup_h (N - B_2^\circ)$$

The topological structure formed can be shown unique up to homeomorphism, but this is non-trivial to prove. The n-holed torus $T\#T\#\ldots\#T\#T$ is an example of such a structure.

1.3 Euclidean Neighbourhoods are Open

In these notes, we consider a neighbourhood as in the French school, as any subset containing an open set, regardless of whether it is open or not. Nonetheless, let M be a manifold, and take a point p with neighbourhood U homeomorphic to $\mathbf{R}^n$, lets say, by some continuous function $f : U \to \mathbf{R}^n$. Then U contains an open set V, and $f(V)$ is open in $\mathbf{R}^n$, so that $f(V)$ contains an open ball W around $f(x)$. But then W is homeomorphic to $\mathbf{R}^n$, and $f^{-1}(W)$ is a neighbourhood of x open in V (and therefore open in M) homeomorphic to $\mathbf{R}^n$. This complicated discussion stipulates that we may always choose open neighbourhoods in the definition in a manifold. Remarkably, it turns out that all neighbourhoods homeomorphic to $\mathbf{R}^n$ *must* be open; to prove this, we require an advanced theorem of algebraic topology.

Theorem 1.3 (Invariance of Domain). *If $f : U \to \mathbf{R}^n$ is a continuous, injective function, where U is an open subset of $\mathbf{R}^n$, then $f(U)$ is open, so that f is a homeomorphism.*

A domain is a connected open set, and this theorem shows that the property of being a domain is invariant under continuous, injective maps from $\mathbf{R}^n$ to itself. In multivariate calculus, the inverse function theorem shows this for differentiable mappings with non-trivial Jacobian matrices across its domain; invariance of domain stipulates that the theorem in fact holds for any such continuous map f on an open domain. The theorem can be proven in an excursion in some basic algebraic topology (homology theory, to be *exact*). In an appendix to this chapter, we shall prove the theorem based on the weaker assumption of the Jordan curve theorem.

Lemma 1.4. *If $U \subset \mathbf{R}^n$ and $V \subset \mathbf{R}^m$ are open, then $U \cong V$ implies $n = m$.*

Proof. If $n < m$, consider the projection $\pi : \mathbf{R}^n \to \mathbf{R}^m$

$$\pi(x_1, \dots, x_n) = (x_1, \dots, x_n, 0, \dots, 0)$$

Clearly no subset of $\pi(\mathbf{R}^n)$ is open. But if $f : V \to U$ is a homeomorphism, then $\pi \circ f : V \to \mathbf{R}^m$ is continuous and injective, so $\pi(V) \subset \pi(\mathbf{R}^n)$ is open by invariance of domain. □

The **dimension** of a point on a manifold is the dimension of the euclidean space which is locally homeomorphic to a neighbourhood of the point. When a manifold is connected, one can show simply that the dimension across the entire space is invariant, and we may call this the **dimension of the manifold**. An n-dimensional manifold M is often denoted M^n.

Corollary 1.5. *The dimension of a point on a manifold is unique.*

Proof. Let U and V be two non-disjoint neighbourhoods of a point homeomorphic to $\mathbf{R}^n$ and $\mathbf{R}^m$ by $f : U \to \mathbf{R}^n$ and $g : V \to \mathbf{R}^m$. Then $U \cap V$ is also open, and homeomorphic to open sets of $\mathbf{R}^n$ and $\mathbf{R}^m$. We conclude $n = m$. □

Theorem 1.6. *Any subset of a manifold locally homeomorphic to Euclidean space is open in the original topology.*

Proof. Let M be a manifold, and $U \subset M$ homeomorphic to $\mathbf{R}^n$ by a function f. Let $x \in U$ be arbitrary. There is an open neighbourhood V of x that is homeomorphic into $\mathbf{R}^n$ by a function g. Since V is open in M, $U \cap V$ is open in U, so $f(U \cap V)$ is open in $\mathbf{R}^n$. We obtain a one-to-one continuous function from $f(U \cap V)$ to $g(U \cap V)$ by the function $g \circ f^{-1}$. It follows by invariance of domain that $g(U \cap V)$ is open in $\mathbf{R}^n$, so $U \cap V$ is open in

V, and, because V is open in M, $U \cap V$ is open in M. In a complicated manner, we have shown that around every point in U there is an open neighbourhood contained in U, so U itself must be open. $\square$

Really, this theorem is just a generalized invariance of domain for arbitrary manifolds – since the concept of a manifold is so intertwined with Euclidean space, it is no surprise we need the theorem for $\mathbf{R}^n$ before we can prove the theorem here.

1.4 Equivalence of Regularity Properties

Many important results in differentiable geometry require spaces with more stringent properties than those that are merely Hausdorff. At times, we will want to restrict ourselves to topological manifolds with these properties. Fortunately, most of these properties are equivalent.

Theorem 1.7. *For any manifold, the following properties are equivalent:*

(1) Every component of the manifold is σ-Compact.

(2) Every component of the manifold is second countable.

(3) The manifold is metrizable.

(4) The manifold is paracompact (so every compact manifold is metrizable).

Lemma 1.8 (1) $\to$ (2)**.** *Every σ-compact, locally second countable space is globally second countable.*

Proof. Let X be a locally second countable space, equal to the union of compact sets $\bigcup_{i=1}^{\infty} A_i$. For each x, there is an open neighbourhood U_x with a countable base $\mathcal{C}_x$. If, for some A_i, we consider the set of U_x for $x \in A_i$, we obtain a cover, which therefore must have a finite subcover $U_{x_1}, U_{x_2}, \dots, U_{x_n}$. Taking $\bigcup_{i=1}^{n} \mathcal{C}_{x_i}$, we obtain a countable base $\mathcal{C}_i$ for all points in a neighbourhood of A_i. Then, taking the union $\bigcup_{i=1}^{\infty} \mathcal{C}_i$, we obtain a countable base for X. $\square$

Lemma 1.9 (2) $\to$ (3)**.** *If a manifold is second countable, then it is metrizable.*

Proof. This is a disguised form the Urysohn metrization theorem, proved in a standard course in general topology. If you do not have the background, you will have to have faith that this lemma holds. All we need show here is that a second countable manifold is regular, and this follows because every locally compact Hausdorff space is Tychonoff. □

Lemma 1.10 (3) → (1)**.** *Every connected, locally compact metrizable space is σ-compact.*

Proof. Consider any connected, locally compact metric space (X,d). For each x in X, let

$$r(x) = \frac{\sup\{r \in \mathbf{R} : \overline{B}_r(x) \text{ is compact}\}}{2}$$

Since X is locally compact, this function is well defined and positive for all x. If $r(x) = \infty$ for any x, then $\{\overline{B}_n(x) : n \in \mathbf{Z}\}$ is a countable cover of the space by compact sets. Otherwise, $r(x)$ is finite for every x. Suppose that

$$d(x,y) + r' < 2r(x)$$

By the triangle inequality, this tells us that $\overline{B}_{r'}(y)$ is a closed subset of $\overline{B}_{r(x)}(x)$, which is hence compact. This shows that, when $d(x,y) < r(x)$,

$$r(y) \geqslant r(x) - \frac{d(x,y)}{2}$$

Put more succinctly, this equation tells us that the function $r : X \to \mathbf{R}$ is continuous:

$$|r(x) - r(y)| < \frac{d(x,y)}{2}$$

This has an important corollary. Consider a compact set A, and let

$$A' = \bigcup_{x \in A} \overline{B}_{r(x)}(x)$$

We claim that A' is also compact. Consider some sequence $\{x_i\}$ in A', and let $\{a_i\}$ be elements of A for which $x_i \in \overline{B}_{r(a_i)}(a_i)$. Since A is compact, we may assume $\{a_i\}$ converges to some a. When $d(a_i,a) < r(a)/2$,

$$r(a_i) < r(a) + r(a)/4$$

and so

$$d(a, x_i) \leqslant d(a, a_i) + d(a_i, x_i) < r(a)/2 + [r(a) + r(a)/4] = 7r(a)/4$$

Since we chose $r(a)$ to be half the supremum of compact sets, the sequence x_k will eventually end up in the compact ball $B_{3r(a)/4}(a)$, and hence will converge.

If A is a compact set, we will let A' be the compact set constructed above. Let A_0 consist of an arbitrary point x_0 is X, and inductively, define $A_{k+1} = A'_k$, and $A = \bigcup_{i=0}^{\infty} A_k$. Then A is the union of countably many compact sets. A is obviously open. If x is a limit point of A, then there is some sequence $\{x_i\}$ in A which converges to x, so $r(x_i) \to r(x)$. If $|r(x_i) - r(x)| < \varepsilon$, and also $d(x_i, x) < r(x) - \varepsilon$, then x is contained in $B_{r(x_i)}(x_i)$, and hence if x_i is in A_k, then x is in A_{k+1}. Thus A is non-empty and clopen, so $X = A = \bigcup A_k$ is σ-compact. □

Lemma 1.11 (4) $\to$ (1)**.** *A connected, locally compact, paracompact space is σ compact.*

Proof. Consider a locally-finite cover $\mathcal{C}$ of precompact neighbourhoods in a space X. Fix $x \in X$. Then x intersects finitely many elements of $\mathcal{C}$, which we may label $U_{1,1}, U_{1,2}, \dots, U_{1,n_1}$. Then

$$U_1 = \overline{U_{1,1}} \cup \overline{U_{1,2}} \cup \cdots \cup \overline{U_{1,n_1}}$$

intersects only finitely more elements of $\mathcal{C}$, since the set is compact, and we need only add finitely more open sets $U_{2,1}, \dots, U_{2,n_2}$, obtaining

$$U_2 = \overline{U_{2,1}} \cup \cdots \cup \overline{U_{2,n_2}}$$

Continuing inductively, we find an increasing sequence of compact neighbourhoods. Then $U = \bigcup U_i$ is open because a neighbourhood of $y \in U_k$ is contained in U_{k+1}. If y is a limit point of U, take a neighbourhood $V \in \mathcal{C}$, which must intersect some U_k. Then $y \in U_{k+1}$, so U is closed. We conclude $X = U$ is σ compact. □

Lemma 1.12 (1) $\to$ (4)**.** *A σ compact, locally compact Hausdorff space is paracompact.*

Proof. Let $X = \bigcup C_i$ be a locally compact, σ-compact space. Since C_1 is compact, it is contained in an open precompact neighbourhood U_1. Similarily, $C_2 \cup \overline{U_1}$ is contained in a precompact neighbourhood U_2 with compact closure. We find $U_1 \subset U_2 \subset \dots$, each with compact closure, and which cover the entire space. Now let $\mathcal{U}$ be an arbitrary open cover of X. Each $V_k = U_k - \overline{U_{k-2}}$ (letting $U_{-2} = U_{-1} = U_0 = \varnothing$) is open, and its closure $\overline{V_k}$ is a closed subset of compact space, hence compact. Since $\mathcal{U}$ covers $\overline{V_k}$, it has a finite subcover $U_1, \dots, U_n$, and we let

$$\mathcal{V}_1 = (U_1 \cap V_1), (U_2 \cap V_1), \dots, (U_n \cap V_1)$$

be a collection of refined open sets which cover V_1. Do the same for each V_k, obtaining $\mathcal{V}_2, \mathcal{V}_3, \dots$, and consider $\mathcal{V} = \bigcup \mathcal{V}_i$. Surely this is a cover of X, and each point is contained only in some $\mathcal{V}_k$ and $\mathcal{V}_{k+1}$, so this refined cover is locally finite. $\square$

Chapter 2

Differentiable Structures

As a topological space, we know when a map between manifolds is continuous, but when is a map differentiable? What we seek is a definition abstract enough to work on any manifold, yet possessing the same properties of differentiable functions on $\mathbf{R}^n$.

2.1 Defining Differentiability

Let us be given a map $f : M \to N$ between manifolds. Given a correspondence $b = f(a)$, a reasonable inquiry would be to consider two charts (x, U) and (y, V), where U is a neighbourhood of a and V is a neighbourhood of b. We obtain a map $y \circ f \circ x^{-1}$, defined between open subsets of Euclidean space. We have 'expressed f in coordinates'. f shall then be differentiable at a if $y \circ f \circ x^{-1}$ is differentiable at $x(a)$. Unfortunately, this idea is doomed to fail, for we can hardly expect that the statement holds for all charts when it holds for a pair of them.

Example. *Consider the chart $y : \mathbf{R} \to \mathbf{R}$, $y(t) = t^3$, and let x be the identity chart. If $f(x) = \sin(x)$, then $x \circ f \circ x^{-1} = f$ is differentiable, yet*

$$(y \circ f \circ y^{-1})(t) = \sin(\sqrt[3]{t})^3$$

is not differentiable at the origin.

If we are to stick with this definition, we either need to define differentiability in terms of the charts g and h used, or identify additional structure to manifolds. The latter option is clearly more elegant. Our method will be to identify charts which are 'correct', and ignore unnatural constructions like k. Two charts (x, U) and (y, V) are $\mathbf{C}^\infty$ **related**, if either U and V are disjoint, or

$$y \circ x^{-1} : x(U \cap V) \to y(U \cap V)$$

$$x \circ y^{-1} : y(U \cap V) \to x(U \cap V)$$

are C^∞ functions (diffeomorphisms, to be particular). One can see a chart as laying a blanket down onto a manifold. Two charts are C^∞ related if, when we lay them down over each other, they contain no creases! The fact that manifolds do not have a particular preference for coordinates is both a help and a hindrance. On one side, it forces us to come up with elegant, coordinate free approaches to geometry. On the other end, these coordinate free approaches can also be incredibly unnatural!

A **smooth** or $\mathbf{C}^\infty$ **atlas** for a manifold is a family of C^∞ charts whose domains cover the entire manifold. A maximal atlas is called a **smooth structure** on a manifold, and a manifold together with a smooth structure is called a **smooth** or **differentiable manifold**. In the literature, each map $y \circ x^{-1}$ is known as a **transition map**. An atlas for a manifold has C^∞ transition maps. From now on, when we mention a chart on a differentiable manifold, we implicitly assume the chart is the member of the smooth structure of the manifold. The category of smooth manifolds is denoted $\mathbf{Man}^\infty$.

A $f : M \to N$ be a map between two smooth manifolds. f is differentiable at $p \in M$ if it is continuous at p, and if for some chart $x : U \to \mathbf{R}^n$ whose domain contains p, and for some chart $y : V \to \mathbf{R}^m$ whose domain contains $f(p)$, the map $y \circ f \circ x^{-1} : x(f^{-1}(V) \cap U) \to \mathbf{R}^m$ is differentiable at $x(p)$. f itself is **differentiable** if it is differentiable at every point on its domain, or correspondingly, if $y \circ f \circ x^{-1}$ is differentiable for any two charts x and y.

It is uncomfortable to construct a maximal atlas explicitly on a manifold. Fortunately, we do not need to specify every single valid chart in our manifold.

Lemma 2.1. *Every atlas extends to a unique smooth structure.*

Proof. Let $\mathcal{A}$ be an atlas for a manifold M, and consider the set $\mathcal{A}'$, which is the union of all atlases containing $\mathcal{A}$. We shall show that $\mathcal{A}'$ is also an atlas, and therefore necessarily the unique maximal one. Let $x : U \to \mathbf{R}^n$

and $y : V \to \mathbf{R}^n$ be two charts in $\mathcal{A}'$ with non-disjoint domain, containing a point p. Let $z : W \to \mathbf{R}^n$ be a chart in $\mathcal{A}$ containing p. Then, on $U \cap V \cap W$, an open set containing p, we have

$$x \circ y^{-1} = (x \circ z^{-1}) \circ (z \circ y^{-1})$$

and by assumption, each component map is C^∞ on this domain, so $x \circ y^{-1}$ is smooth in a neighbourhood of p. The proof for $y \circ x^{-1}$ is exactly the same. Since the point p was arbitrary, we conclude that x and y are C^∞ related across their domains. □

Corollary 2.2. *If x is a chart defined on a differentiable manifold M, and is C^∞ related to each map in a generating atlas $\mathcal{A}$, then x is in the smooth structure generated by $\mathcal{A}$.*

The next few theorems are justified by the fact that C^∞ related charts play nicely with one another.

Lemma 2.3. *If a map f is differentiable at a point p in charts x and y, it is differentiable at p for any other charts containing p and q.*

Proof. Suppose $y \circ f \circ x^{-1}$ is differentiable at a point $x(p)$, and consider any other charts y' and x'. Then

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

On a smaller open neighbourhood than was considered. Nonetheless, since differentiability is a local concept, we need only prove the theorem for this map on a reduced domain. This follows since the component maps are differentiable. □

Example. *Let M be a manifold, and U an open submanifold. Define a differentiable structure on U consisting of all charts defined on M whose domain is a subset of U. This is a maximal atlas, and is the unique such structure such that*

1. *If $f : M \to N$ is differentiable, then $f|_U : U \to M$ is differentiable.*
2. *The inclusion map $i : U \to M$ is differentiable.*
3. *If $f : N \to M$ is differentiable, and $f(N) \subset U$, then $f : N \to U$ is differentiable.*

Example. *Consider the manifold $\mathbf{R}^n$, and define a generating atlas on $\mathbf{R}^n$ containing only the identity map $id_{\mathbf{R}^n}$. This defines a smooth structure on $\mathbf{R}^n$, which satisfies the following familiar properties:*

1. *x is a chart on $\mathbf{R}^n$ if and only if x and x^{-1} are C^∞.*
2. *A map $f : \mathbf{R}^n \to \mathbf{R}^m$ is differentiable in the sense of a manifold if and only if it is differentiable in the usual sense.*
3. *A map $f : M \to \mathbf{R}^n$ is differentiable if and only if each coordinate $f_i : M \to \mathbf{R}$ is differentiable.*
4. *A chart $x : U \to \mathbf{R}^n$ is a diffeomorphism from U to $x(U)$.*

Our definition has naturally extended calculus to arbitrary manifolds.

Example. *On $\mathbf{R}^2$, we have the polar coordinate system $(\theta, \mathbf{R}^2 - \{0\})$, defined by*

$$\theta^{-1}(r,u) = re^{iu}$$

This chart is injective and full rank, so that θ is C^∞. Thus the polar coordinate system truly is in the smooth structure generated by the identity. On $\mathbf{R}^3$, we have the spherical and cylindrical coordinate systems. We leave it to the reader that these are in fact C^∞ charts.

Example. *The differentiable structure on S^n is defined by the stereographic projection maps. On S^1, we may define the structure by angle functions. If (θ, U) and (ψ, V) are angle functions, then we may extend θ to $\tilde{\theta}$, defined on an open neighbourhood of U in $\mathbf{R}^2$, and defined by*

$$\tilde{\theta}^{-1}(r,t) = re^{it}$$

Then $\tilde{\theta}^{-1}$ is C^∞, and has full rank at every point, so the inverse if C^∞ by the inverse function theorem. It then follows that $\theta \circ \psi^{-1} = \tilde{\theta} \circ \psi^{-1}$ is C^∞. The smooth structure generated is the same as the one formed by stereographic projection.

Example. *Smooth structures on manifolds are not unique. Let $\mathbf{R}_1$ be the canonical smooth manifold on $\mathbf{R}$. Let $\mathbf{R}_2$ be the smooth structure on $\mathbf{R}$ generated by the map x, such that $x(t) = t^3$. Then $\mathbf{R}_1$ and $\mathbf{R}_2$ are diffeomorphic. Let $x : \mathbf{R}_2 \to \mathbf{R}_1$ be our diffeomorphism. It is surely bijective. Let y be a chart on $\mathbf{R}_2$. We must verify that $y = z \circ x$, where z is a chart on $\mathbf{R}_1$. We may show this by verifying that $y \circ x^{-1} = z$, and $x \circ y^{-1} = z^{-1}$ is C^∞ on $\mathbf{R}_1$. But this was exactly why y was a chart on $\mathbf{R}_2$ in the first place, hence the map is a diffeomorphism.*

2.2 The Function Space $C^\infty(M)$

The set of all real-valued differentiable maps defined on a manifold M is denoted $C^\infty(M)$. It is a subspace of the function space $C(M)$. First, note that a continuous map $f : M \to N$ induces a map $f^* : C(N) \to C(M)$ defined by $f^*(g) = g \circ f$. If f is differentiable, then we may consider f^* a map from $C^\infty(N)$ to $C^\infty(M)$ by restriction. Thus the 'map' C^∞ defines a contravariant functor from **Man** to **Vect**.

Lemma 2.4. *If $f^*(C^\infty(N)) \subset C^\infty(M)$, then f is smooth.*

Proof. Let (y, V) be a chart on N at a point q, and let (x, U) be a chart on M at $p \in f^{-1}(p)$. By assumption, each $y^i \circ f$ is differentiable, so that $y^i \circ f \circ x^{-1}$ is differentiable. But this implies $y \circ f \circ x^{-1}$ is differentiable. □

Theorem 2.5. *A homeomorphism $f : M \to N$ is a diffeomorphism if and only if f^* is an isomorphism between $C^\infty(N)$ and $C^\infty(M)$.*

Proof. Suppose f^* is an isomorphism. Then we know that f must be differentiable. One verifies that $(f^{-1})^* = (f^*)^{-1}$. The last lemma implies f^{-1} is also differentiable, so f is a diffeomorphism. The converse follows from the same argument, since if f is a diffeomorphism, then f^* maps $C^\infty(N)$ into $C^\infty(M)$, and $(f^{-1})^* = (f^*)^{-1}$ is an inverse, so f^* is an isomorphism. □

Suppose we know $C^\infty(M)$. Then we may recover the smooth structure on M, which is the set of diffeomorphisms from open subsets of M to open subsets of euclidean space. This is the foundation of the algebraic viewpoint of manifold theory, which attempts to uncover the nature of manifolds solely by analyzing the commutative algebra $C^\infty(M)$.

2.3 Partial Derivatives and Differentiable Rank

In calculus, when a function is differentiable, we obtained a derivative, a measure of a function's local change. On manifolds, determining an analogous object is difficult due to the coordinate invariant definition required. For now, we shall stick to structures corresponding to some particular set of coordinates. Consider a differentiable map f from a manifold M to the real numbers. We have no conventional coordinates to consider partial derivatives on, but consider some chart $x : U \to \mathbf{R}^n$ on M. We obtain a differentiable map $f \circ x^{-1}$. We define, for a point $p \in U$,

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

We are tracing the coordinates lines placed on U by the map x^{-1}; literally, if c is the curve defined by $c(t) = (f \circ x^{-1})(x(p) + tx_k)$, then

$$c'(0) = \left.\frac{\partial f}{\partial x_k}\right|_p$$

Our new definition of the partial derivative satisfies the familiar chain rule.

Theorem 2.6. *If x and y are coordinate systems at a point p, and $f : M \to \mathbf{R}$ is differentiable, then*

$$\left.\frac{\partial f}{\partial x_i}\right|_p = \sum \left.\frac{\partial y_j}{\partial x_i}\right|_p \left.\frac{\partial f}{\partial y_j}\right|_p$$

Proof. We just apply the chain rule in Euclidean space.

$$\begin{aligned}\left.\frac{\partial f}{\partial x_i}\right|_p = D_i(f \circ x^{-1})(x(p)) &= D_i((f \circ y^{-1}) \circ (y \circ x^{-1}))(x(p)) \\ &= \sum D_j(f \circ y^{-1})(y(p)) D_i(y_j \circ x^{-1})(x(p)) \\ &= \sum \left.\frac{\partial f}{\partial y_j}\right|_p \left.\frac{\partial y_j}{\partial x_i}\right|_p\end{aligned}$$

□

Example. *Let us compute the laplacian on $\mathbf{R}^2$ in polar coordinates.*

$$\triangle f = \frac{\partial f^2}{\partial x^2} + \frac{\partial f^2}{\partial y^2}$$

To do this, we need to relate partial differentives by the chain rule. If (r, θ) is the polar coordinate chart, and (x, y) the standard chart on $\mathbf{R}^2$, then

$$x = r\cos(\theta) \qquad y = r\sin(\theta)$$

(note that this is a relation between functions, and can be applied pointwise at any point on the charts). Thus the matrix of partial derivatives is

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}$$

We can invert this matrix to obtain the partial derivatives with respect to x and y. We have

$$\begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{pmatrix} = \frac{1}{r}\begin{pmatrix} r\cos(\theta) & r\sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\frac{\sin(\theta)}{r} & \frac{\cos(\theta)}{r} \end{pmatrix}$$

Now we apply the chain rule. We have

$$\frac{\partial}{\partial x} = \cos(\theta)\frac{\partial}{\partial r} - \frac{\sin(\theta)}{r}\frac{\partial}{\partial \theta} \qquad \frac{\partial}{\partial y} = \sin(\theta)\frac{\partial}{\partial r} + \frac{\cos(\theta)}{r}\frac{\partial}{\partial \theta}$$

So

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= \left(\cos(\theta)\frac{\partial}{\partial r} - \frac{\sin(\theta)}{r}\frac{\partial}{\partial \theta}\right)\left(\cos(\theta)\frac{\partial f}{\partial r} - \frac{\sin(\theta)}{r}\frac{\partial f}{\partial \theta}\right) \\ &= \cos^2(\theta)\frac{\partial^2 f}{\partial r^2} + \frac{\cos(\theta)\sin(\theta)}{r^2}\frac{\partial f}{\partial \theta} - \frac{\cos(\theta)\sin(\theta)}{r}\frac{\partial^2 f}{\partial r \partial \theta} \\ &\quad + \frac{\sin^2(\theta)}{r}\frac{\partial f}{\partial r} - \frac{\sin(\theta)\cos(\theta)}{r}\frac{\partial^2 f}{\partial \theta \partial r} + \frac{\sin(\theta)\cos(\theta)}{r^2}\frac{\partial f}{\partial \theta} + \frac{\sin^2(\theta)}{r^2}\frac{\partial^2 f}{\partial \theta^2}\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 f}{\partial y^2} &= \left(\sin(\theta)\frac{\partial}{\partial r} + \frac{\cos(\theta)}{r}\frac{\partial}{\partial \theta}\right)\left(\sin(\theta)\frac{\partial f}{\partial r} + \frac{\cos(\theta)}{r}\frac{\partial f}{\partial \theta}\right) \\ &= \sin^2(\theta)\frac{\partial^2 f}{\partial r^2} - \frac{\cos(\theta)\sin(\theta)}{r^2}\frac{\partial f}{\partial \theta} + \frac{\cos(\theta)\sin(\theta)}{r}\frac{\partial^2 f}{\partial r \partial \theta} \\ &\quad + \frac{\cos^2(\theta)}{r}\frac{\partial f}{\partial r} + \frac{\sin(\theta)\cos(\theta)}{r}\frac{\partial^2 f}{\partial \theta \partial r} - \frac{\sin(\theta)\cos(\theta)}{r^2}\frac{\partial f}{\partial \theta} + \frac{\cos^2(\theta)}{r^2}\frac{\partial^2 f}{\partial \theta^2}\end{aligned}$$

It follows that, by use of the trigonometric identities,

$$\triangle f = \frac{\partial f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2}$$

Partial derivatives also satisfy a nice ‘derivation’ property, which we lead to the reader to calculate.

Lemma 2.7. *For $f, g \in C^\infty(M)$,*

$$\left.\frac{\partial fg}{\partial x^i}\right|_p = f(p)\left.\frac{\partial g}{\partial x^i}\right|_p + g(p)\left.\frac{\partial f}{\partial x^i}\right|_p$$

Now we could define the full derivative of a function $f : M \to N$ at a point p to be $D(y \circ f \circ x^{-1})(x(p))$, for coordinate systems x and y. The trouble is that in this way we can only talk of properties of the derivative that are invariant under the coordinate systems chosen, since the linear operator D is not invariant of the coordinate system. Later on, we will be able to come up with a universal differential operator that contains all coordinate representations of D.

Theorem 2.8. *The rank of the matrix $D(y \circ f \circ x^{-1})(x(p))$ is the same regardless of which coordinate systems x and y are chosen.*

Proof. Let y' and x' be two more coordinate systems around p and $f(p)$.

$$\begin{aligned} D(y' \circ f \circ x'^{-1})(x'(p)) &= D(y' \circ y^{-1} \circ y \circ f \circ x^{-1} \circ x \circ x'^{-1})(x'(p)) \\ &= D(y' \circ y^{-1})(y(f(p))) \circ D(y \circ f \circ x^{-1})(x(p)) \\ &\quad \circ D(x \circ x'^{-1})(x'(p)) \end{aligned}$$

The first and last derivatives in the last composition are invertible linear operators, so the rank is invariant. □

The rank of the partial derivative matrix tells us the freedom of movement of the mapping around the point p. Using the rank will allow us to extend the inverse function theorem to arbitrary manifolds.

Theorem 2.9. *If $f : M^n \to N^m$ is rank k at a point p, there is a coordinate system x at p and y at $f(p)$ such that for $1 \leqslant i \leqslant k$,*

$$y_i \circ f \circ x^{-1}(a_1, \dots, a_n) = a_i$$

Proof. Let (x, U) and (y, V) be arbitrary coordinate systems around p and $f(p)$. By a permutation of the coordinates, we may, by arranging coordinates, guarantee that the matrix

$$\left(\left.\frac{\partial y_i \circ f}{\partial x_j}\right|_p \right)_{i,j=1}^k$$

is invertible. Define a map $z : U \cap f^{-1}(V) \to \mathbf{R}^n$ around p by $z_i = y_i \circ f$, for $1 \leqslant i \leqslant k$, and $z_i = x_i$ otherwise. The matrix

$$D(z \circ x^{-1})(p) = \begin{pmatrix} \left(\left. \frac{\partial y_i \circ f}{\partial x_j} \right|_p \right) & X \\ 0 & I \end{pmatrix}$$

is invertible, hence, by the inverse function theorem, $x \circ z^{-1}$ is a diffeomorphism in a neighbourhood of $z(p)$. It follows that z is a coordinate system at p, and for $1 \leqslant i \leqslant k$,

$$y_i \circ f \circ z^{-1}(a_1, \dots, a_n) = a_i$$

and we have found the right coordinate system. □

Corollary 2.10. *If f is rank k in a neighbourhood of a point p, then we may choose coordinate systems x and y such that*

$$y \circ f \circ x^{-1}(a_1, \dots, a_n) = (a_1, \dots, a_k, 0, \dots, 0)$$

Proof. Choose x and y from the theorem above. Then

$$D(y \circ f \circ x^{-1})(p) = \begin{pmatrix} I & 0 \\ X & \left(\frac{\partial y_i \circ f}{\partial x_j} \right) \end{pmatrix}$$

Since f is rank k, the matrix in the bottom right corner must vanish in a neighbourhood of p. Therefore, for $i > k$, $y_i \circ f \circ x^{-1}$ can be viewed only as a function of the first k coordinates. Define $z_i = y_i$, for $i < k$, and

$$z_i = y_i - (y_i \circ f)(y_1 \dots y_k)$$

We have an invertible change of coordinate matrix,

$$D(z \circ y^{-1}) = \begin{pmatrix} I & 0 \\ X & I \end{pmatrix}$$

So z is a coordinate system, and

$$z \circ f \circ x^{-1}(a_1, \dots, a_n) = (a_1, \dots, a_k, 0, \dots, 0)$$

we have constructed a coordinate system as needed. □

Corollary 2.11. *If $f : M^n \to N^m$ is rank m at p, then for any coordinate system y, there exists a coordinate system x such that*

$$y \circ f \circ x^{-1}(a_1, \dots, a_n) = (a_1, \dots, a_m)$$

Proof. In the proof of the theorem, we need not rearrange coordinates of y in the case that the matrix is rank m. □

Corollary 2.12. *If $f : M^n \to N^m$ is rank n at p, then for any coordinate system x, there exists a coordinate system y such that*

$$y \circ f \circ x^{-1}(a_1, \dots, a_n) = (a_1, \dots, a_n, 0, \dots, 0)$$

Proof. If f is rank n at a point, it is clearly rank n on a neighbourhood by the continuity of the determinant. Choose coordinate systems u and v such that $u \circ f \circ v^{-1}(a_1, \dots, a_n) = (a_1, \dots, a_n, 0, \dots, 0)$. Define a map λ on $\mathbf{R}^m$ by $\lambda(a_1, \dots, a_m) = (x \circ v^{-1}(a_1, \dots, a_n), a_{n+1}, \dots, a_m)$. Then λ is a diffeomorphism, hence $\lambda \circ y$ is a coordinate system, so

$$\begin{aligned}(\lambda \circ y) \circ f \circ x^{-1}(a_1, \dots, a_n) &= \lambda \circ (y \circ f \circ v^{-1}) \circ (v \circ x^{-1})(a_1, \dots, a_n) \\ &= \lambda(v \circ x^{-1}(a_1 \dots a_n), 0 \dots 0) \\ &= (a_1 \dots a_n, 0 \dots 0)\end{aligned}$$

and we have found the chart required. □

Theorem 2.13. *If $n \geqslant m$, and $f : M^n \to N^m$ has full rank at every point. Then f is open.*

Proof. If $p \in M$, pick a neighbourhood x around p and y around $f(p)$ such that

$$x \circ f \circ y^{-1}(a_1, \dots, a_n) = (a_1, \dots, a_m)$$

This map is open, showing f is locally open, and thus entirely open. □

All this goes to show that manifolds defined by a differentiable function still have a very nice structure.

Theorem 2.14. *If $f : M^n \to N^m$ has constant rank k in a neighbourhood of the points mapping to $q \in N$, then $f^{-1}(q)$ is a closed $n - k$ submanifold of M.*

Proof. If $f(p) = q$, and f has rank k at p, then we may write

$$y \circ f \circ x^{-1}(a_1, \dots, a_n) = (a_1, \dots, a_k, 0, \dots, 0)$$

This implies that if we adjust the last $n-k$ coordinates around p, that point will still map to p. These are the only coordinates we can vary, so that $(x_{k+1}, \dots, x_n)$ form a coordinate system around p in $f^{-1}(q)$. $\square$

The spheres S^n give us examples of how this theorem can be applied.

Example. *The special linear group $SL_n(\mathbf{R})$ is the set of invertible matrices with determinant one. Since the determinant is a multilinear function, we can find the determinant via the formula*

$$D(\det)(v_1, \dots, v_n)(w_1, \dots, w_n) = \sum_{k=1}^{n} \det(v_1, \dots, w_k, \dots, v_n)$$

Then in $GL_n(\mathbf{R})$,

$$D(\det)(v_1, \dots, v_n)(v_1, \dots, v_n) = n \det(v_1, \dots, v_n) \neq 0$$

So det *has full rank at every point, and $SL_n(\mathbf{R})$ is dimension $n^2 - 1$.*

Example. *The orthogonal group $O_n(\mathbf{R})$ is the set of matrices M such that $MM^t = I$. Then $O_n(\mathbf{R})$ is closed, for the map $\psi : M \mapsto MM^t$ is continuous, and $O_n(\mathbf{R}) = \psi^{-1}(I)$. ψ maps into the set of symmetric matrices, which is a subspace of $M_n(\mathbf{R})$ of dimension $n(n+1)/2$. If we take the i'th entry of MM^t, we obtain that*

$$v_{i1}^2 + v_{i2}^2 + \dots + v_{in}^2 = 1$$

This implies that M lies on a sphere in $\mathbf{R}^{n^2}$. Thus $O_n(\mathbf{R})$ is closed, bounded, and therefore compact. Consider the diffeomorphism $R_A : B \mapsto BA$ of $GL_n(\mathbf{R})$, for a fixed $A \in GL_n(\mathbf{R})$. We also have $\psi \circ R_A = \psi$, for $A \in O_n(\mathbf{R})$. We conclude that

$$D(\psi)(A) \circ D(R_A)(I) = D(\psi \circ R_A)(I) = D(\psi)(I)$$

Since R_A is a diffeomorphism, $D(\psi)(A)$ has the same rank as $D(\psi)(I)$. Let us find this rank. Explicitly, we may write the projections of ψ as

$$\psi^{ij}(M) = \sum_{k=1}^{n} M_{ik} M_{jk}$$

Then

$$\left.\frac{\partial \psi^{ij}}{\partial x^{kl}}\right|_M = \begin{cases} 2M_{il} & k = j = i \\ M_{jl} & k = j \\ M_{il} & k = i \\ 0 & elsewise \end{cases}$$

In particular, at the identity,

$$\left.\frac{\partial \psi^{ij}}{\partial x^{kl}}\right|_I = \begin{cases} 2 & k = l = j = i \\ 1 & k = j = l \\ 1 & k = i = l \\ 0 & elsewise \end{cases}$$

It follows that the range of the derivative at the identity is the space of all symmetric matrices, which has dimension $n(n+1)/2$. Since we are mapping into the space of all symmetric matrices, this map has full rank, and thus has the same rank in a neighbourhood. Using the theorem, we find the space of orthogonal matrices has dimension

$$n^2 - n(n+1)/2 = n(n-1)/2$$

Every orthogonal matrix has determinant ± 1. The special orthogonal group $SO_n(\mathbf{R})$ is the set of orthogonal matrices of determinant one, and is an open submanifold of $O_n(\mathbf{R})$.

Corollary 2.15. *An immersion between n-manifolds whose domain is compact and connected is onto.*

A function $f : M \to N$ has **full rank** at a point $p \in M$ if the rank of f at p is the same as the dimension of M at p. f is an **immersion** if it has full rank at every point. An immersion need not be one-to-one, but it is always locally one-to-one. At any point $f(p) \in N$, we may choose a coordinate system (y, U) around $f(p)$ such that, for some open set V, $((y_1, \dots, y_n), V)$ is a coordinate system around p, and

$$f(V) = \{q \in U : y_{n+1}(q) = y_{n+2}(q) = \dots = y_m(q) = 0\}$$

This follows from theorems in the previous section. Immersions are mostly well behaved, apart from the odd inconsistency. Let $g : P \to N$ be a differentiable function with $g(P) \subset f(M)$. If f is globally one-to-one, we may define $g : P \to M$. A suitable question to ask is whether this function is differentiable. In most cases, the answer is yes.

Example. *Immerse the non-negative real numbers in S^2 via the map f defined by*

$$f(x) = e^{2\pi(1-e^{-x})i}$$

whose image is $S^2 - \{1\}$. Define $g : (-1,1) \to S^2$ by $g(x) = e^{-ix}$. Then g is C^∞ onto S^2, yet is not even continuous when considered a map into the real numbers via the immersion defined above.

Continuity is all that can go wrong in this situation.

Theorem 2.16. *Let $f : M^n \to N^m$ be an injective immersion of M in N, and suppose $g : P \to N$ is differentiable, and $g(P) \subset N$. If g is continuous considered as a map into M, then g is differentiable considered as a map into M.*

Proof. Consider an arbitrary point $p \in P$. There is a coordinate system $((y_1, \ldots, y_m), U)$ around $g(p)$ such that $(y_1, \ldots, y_n)$ is a coordinate system around $f^{-1}(g(p))$. Since g is continuous into M, there is a coordinate system x around p which maps into $U \cap f(M)$. $y \circ g \circ x^{-1}$ is differentiable, so each $y_i \circ g \circ x^{-1}$, which we have constructed a coordinate system at $f^{-1}(g(p))$ at, is differentiable. Thus g is differentiable mapping into M. □

None of this can happen if M is just a subspace of N. In this case, we say that M is embedded in N, and call M a **submanifold**. If M is closed in N, M is called a **closed submanifold**.

2.4 Manifolds and Measure

The Lebesgue measure on $\mathbf{R}^n$ gives us a nice way to calculate the volume of various sorts of objects. A subset A has measure zero if, for any ε, there is a cover of A by open sets $B_1, B_2, \ldots$ with

$$\sum \lambda(B_i) < \varepsilon$$

Manifolds may not necessarily have a measure like Lebesgue's, but we can still define a subset $A \subset M$ to be of **measure zero** if $A = \bigcup C_i$ where $C_i \subset U_i$, and there are charts (x_i, U_i) in which $x_i(C_i)$ has measure zero.

We shall show that there are few problems with differentiable points on manifolds. Call a point p on M **critical** for a map $f : M \to N$ if f has rank less than the dimension of p at p. Call $f(p)$ a critical value. Recall Sard's theorem, that if $f : \mathbf{R}^n \to \mathbf{R}^m$ is C^1, then the critical values form a set of measure zero in $\mathbf{R}^m$. It follows by localization that the critical values of every C^1 map between manifolds form a set of measure zero. Sard's theorem can actually be generalized to the following:

Theorem 2.17 (Sard)**.** *If $f : M^n \to N^m$ is C^k, and $k \geqslant n - m$, with $k \geqslant 1$ then the set of critical values from a set of measure zero in N.*

We proved the case $n = m$. It is easy to prove the case where $m > n$. The tricky case is where $m < n$.

Theorem 2.18. *If $f : M^n \to N^m$ is C^1, M^n is connected, and $n < m$, then $f(M)$ has measure zero in N^m.*

Proof. Consider the map $g : M \times \mathbf{R}^{m-n} \to N$, defined by $g(p,x) = f(p)$. Then all values of g are critical, and so $g(M \times \mathbf{R}^{m-n}) = f(M)$ has measure zero. □

A Peano curve is a continuous map from $[0,1]$ onto $[0,1] \times [0,1]$. Such a curve cannot be differentiable, by the theorem above. Fortunately, the measure zero properties hold also for rectifiable curves, ones for which

$$\sum_{k=1}^{n} d(c(t_k), c(t_{k+1})) \leqslant M$$

For all points $t_1 < t_2 < \cdots < t_n$. We might as well consider M the supremum of such numbers. Pick $t_1 < \cdots < t_n$ be such that

$$M - \varepsilon \leqslant \sum_{k=1}^{n} d(c(t_k), c(t_{k+1})) \leqslant M$$

Then all points of $c([0,1])$ are within ε of the line segments between the $c(t_k)$, which therefore in contained within an volume of $CM\varepsilon^n$, for some constant C. Taking $\varepsilon \to 0$, we conclude the image of the curve has measure zero.

2.5 The C^∞ Category

We now show that the topological category of manifolds naturally restricts to the category of differentiable manifolds.

Example (Differentiable Product)**.** *If M and N are differentiable manifolds, we may consider an atlas on $M \times N$ with the differentiable structure generated by all maps $x \times y$, where x is a chart on M and y is a chart on N. From this definition, we have the property that $(f,g) : X \to M \times N$ is differentiable if and only if $f : X \to M$ and $g : X \to N$ are differentiable. This is the unique differentiable structure on $M \times N$ which has this property. To see this, suppose that $(M \times N)_2$ is the manifold with a different differential structure, such that the projection $\pi_1 : (M \times N)_2 \to M$ and $\pi_2 : (M \times N)_2 \to N$ are differentiable for some arbitrary differential structure on $M \times N$. It then follows that the identify map $(\pi_1, \pi_2) : (M \times N)_2 \to M \times N$ is differentiable. We may perform the same action to conclude that the identity is a diffeomorphism, and thus maps charts onto charts.*

Example (Differentiable Quotients and $\mathbf{P}^n$)**.** *If N is a quotient space of a differentiable manifold M whose projection $\pi : M \to N$ is locally injective, then we may ascribe a differentiable structure to it. We take all charts $x : U \to \mathbf{R}^n$ on M such that U is homeomorphic to $\pi(U)$ by π. We may then push the chart onto N, and all the charts placed down on N will be C^∞ related. As a covering, this can be extended to a maximal atlas. In fact, this is the unique structure on N which causes $f : N \to L$ to be differentiable if and only if $f \circ \pi$ is differentiable. It allows us to consider $\mathbf{P}^n$ a differentiable manifold, taking the differentiable structure on S^n, as does the Möbius strip, taking the projection from $(-\infty, \infty) \times (0,1)$.*

2.6 Defining C^∞ functions: Partitions of Unity

The use of C^∞ functions relies on the fact that manifolds possess them in plenty. The following theorem gives us our first plethora. First, we detail some C^∞ functions on $\mathbf{R}^n$.

1. The map $f : \mathbf{R} \to \mathbf{R}$, defined by

$$g(t) = \begin{cases} e^{-t} & : t > 0 \\ 0 & : \text{elsewhere} \end{cases}$$

is C^∞. We have $0 < f(t) < 1$ on $(0,\infty)$, and $f^{(n)}(0) = 0$ for all n.

2. The C^∞ map $g(t) = f(t-1)f(t+1)$ is positive on $(-1,1)$, and zero everywhere else. Similarily, for any ε, there is a map g_ε which is positive on $(-\varepsilon,\varepsilon)$ and zero elsewhere.

3. The map
$$l(t) = \begin{cases} \left(\int_{-\varepsilon}^{t} g_\varepsilon\right) / \left(\int_{-\varepsilon}^{\varepsilon} g_\varepsilon\right) & : t \in (0,\infty) \\ 0 & : \text{elsewise} \end{cases}$$
is C^∞, is zero for negative t, increasing on $(0,\varepsilon)$, and one on $[\varepsilon,\infty)$.

4. There is a differentiable map $h : \mathbf{R}^n \to \mathbf{R}$ defined by $h(x_1,\dots,x_n) = g(x_1)g(x_2)\dots g(x_n)$ which is positive on $(-1,1)^n$, and zero elsewhere.

With these nice functions in hand, we may form them on arbitrary manifolds.

Theorem 2.19. *If M is a differentiable manifold, and C is a compact set contained in an open set U, then there is a differentiable function $f : M \to \mathbf{R}$ such that $f(x) = 1$ for x in C, and whose support $\overline{\{x \in M : f(x) \neq 0\}}$ is contained entirely within U.*

Proof. For each point p in C, consider a chart (x,V) around p, with $\overline{V} \subset U$, and $x(V)$ containing the open unit square $(-1,1)^n$ in $\mathbf{R}^n$. We may clearly select a finite subset of these charts (x_k,V_k) such that the $x_k^{-1}((-1,1)^n)$ cover C. We may define a map $f_k : V_k \to \mathbf{R}$ equal to $h \circ x_k$, where h is defined above. It clearly remains C^∞ if we extend it to be zero outside of V_k. Then $\sum f_k$ is positive on C, and whose closure is contained within $\bigcup \overline{V_k} \subset U$. Since C is compact, and the function is continuous, $\sum f_k$ is bounded below by ε on C. Taking $f = l \circ (\sum f_k)$, where l is defined above, we obtain the map needed. □

To enable us to define C^∞ functions whose support lie beyond this range, we need to consider a technique to extend C^∞ functions defined locally to manifolds across the entire domain.

Definition. A partition of unity on a manifold M is a family of C^∞ functions $\{\phi_i : i \in I\}$, and such that the following two properties hold:

1. The supports of the functions forms a locally finite set.
2. For each point $p \in M$, the finite sum $\sum_{i \in I} \phi_i(p)$ is equal to 1.

If $\{U_i\}$ is an open cover of M, then a partition of unity is subordinate to this cover if it also satisfies (3):

3. The closure of each function is contained in some element of the cover.

It is finally our chance to use the topological 'niceness' established in the previous chapter.

Lemma 2.20 (The Shrinking Lemma)**.** *If M is a paracompact manifold, and $\{U_i\}$ is an open cover, then there exists a refined cover $\{V_i\}$ such that for each $i \in I$ there exists i' such that $\overline{V_i} \subset U_{i'}$.*

Proof. Without loss of generality, we may assume $\{U_i\}$ is locally finite, and M is connected. Then M is also σ-compact, $M = \bigcup C_i$. Since C_i is compact, and each $p \in C_i$ locally intersects only finitely many U_i, then C_i intersects only finitely many U_i. Therefore $\bigcup C_i$ intersects only countably many U_i, and thus our locally finite cover is countable. Consider an ordering $\{U_1, U_2, \dots\}$ of $\{U_i\}$. Let C_1 be the closed set $U_1 - (U_2 \cup U_3 \cup \dots)$. Let V_1 be an open set with $C_1 \subset V_1 \subset \overline{V_1} \subset U_1$. Inductively, let C_k be the closed set $U_k - (V_1 \cup \dots \cup V_{k-1} \cup U_{k+1} \cup \dots)$, and define V_k to be an open set with $C_k \subset V_k \subset \overline{V_k} \subset U_k$. Then $\{V_1, V_2 \dots\}$ is the desired refinement. □

Theorem 2.21. *Any cover on a paracompact manifold induces a subordinate partition of unity.*

Proof. c Let $\{U_i\}$ be an open cover of a paracompact manifold M. Without loss of generality, we may consider $\{U_i\}$ locally finite. Suppose that each U_i has compact closure. Choose $\{V_i\}$ satisfying the shrinking lemma. Apply Theorem (2.13) to $\overline{V_i}$ to obtain functions ψ_i that are 1 on $\overline{V_i}$ and zero outside of U_i. These functions are locally finite, and thus we may define

ϕ_i by

$$\phi_i(p) = \frac{\psi_i(p)}{\sum_j \psi_j(p)}$$

Then ϕ_i is the partition of unity we desire.

This theorem holds for any $\{U_i\}$ provided Theorem (2.13) holds on any closed set, rather than just a compact one. Let C be a closed subset of a manifold, contained in an open subset U, and for each $p \in C$, choose an open set U_p with compact closure contained in U. For each $p \in C^c$, choose an open subset V_p contained in C^c with compact closure. Then our previous compact case applies to this cover, and we obtain a subordinate partition of unity $\{\zeta_i\}$. Define f on M by

$$f(p) = \sum_{\overline{\zeta_i} \subset U_p} \zeta_i(p)$$

Then the support of f is contained within U, and $f = 1$ on C. $\square$

Partitions of unity allow us to extend local results on a manifold to global results. The utility of these partitions is half the reason that some mathematicians mandate that manifolds must be paracompact – otherwise many nice results are lost.

Theorem 2.22. *In a σ-compact manifold, there exists a smooth function $f : M \to \mathbf{R}$ such that $f^{-1}((-\infty, t])$ is compact for each t.*

Proof. Let M be a σ-compact manifold with $M = \bigcup B_i$, Where $\overline{B_i}$ is compact, B_i is diffeomorphic to a ball, and the B_i are a locally finite cover. Consider a partition of unity $\{\psi_i\}$ subordinate to $\{B_i\}$, and take the sum

$$f(x) = \sum k\psi_k(x)$$

Then f is smooth, since locally it is the finite sum of smooth functions. If $x \notin B_1, \dots, B_n$, then

$$f(x) = \sum_{k=1}^{\infty} k\psi_k(x) = \sum_{k=n+1}^{\infty} k\psi_k(x) \geqslant (n+1) \sum_{k=n+1}^{\infty} \psi_k(x) = (n+1)$$

Therefore if $f(x) < n$, x is in some B_i. Thus $f^{-1}((-\infty, n])$ is a closed subset of $\overline{B_1} \cup \dots \cup \overline{B_n}$, and is therefore compact. $\square$

Other existence proofs also follow naturally.

Lemma 2.23. *If C is closed on a paracompact manifold M, there is a differentiable function $f : M \to [0,1]$ with $f^{-1}(0) = C$.*

Proof. Let $\{U_i\}$ be a countable cover of $M - C$, with each U_i homeomorphic to the open unit ball in $\mathbf{R}^m$ by some map x. For each U_i, pick f_i be greater than zero on U_i, and equal to zero on $M - U_i$. Define

$$\alpha_j = \min\left\{ \left\| \frac{\partial^n f_i}{\partial x_{i_1} \dots \partial x_{i_n}} \right\|_\infty : i \leqslant j, n \leqslant j \right\}$$

Each α_j is well defined since f_i is C^∞ and tends to zero as we leave U_i. Define

$$f = \sum_{k=1}^{\infty} \frac{f_k}{\alpha_k 2^k}$$

Then f is differentiable, since all partial derivatives locally uniformly converge, and $f^{-1}(0) = C$. □

Corollary 2.24. *If C and D are closed on a paracompact manifold M, then there is a function $f : M \to [0,1]$ with $f^{-1}(0) = C$, $f^{-1}(1) = D$.*

Proof. Let $f^{-1}(1) = C$, and $g^{-1}(0) = D$. Then if ψ is a bump function which is one on C, and zero on D, then $\psi f + (1 - f)g$ is zero on D, and 1 on C. □

Chapter 3

The Tangent Bundle

Historically, calculus was the subject of infinitisimals, differentiable functions which are 'infinitisimally linear'. It took over 200 years to make precise the analytical notions defining the field; in the process, infinitisimals vanished from sight, replaced by linear approximations, epsilons and deltas. On manifolds, we cannot discuss global linear approximations, since the space is not globally linear. Thus we must return to using infinitisimals, which lie on a structure called the tangent bundle.

3.1 Vector Bundles

Geometrically, $\mathbf{R}^n$ is just a system of points in space. Identification of these points with coordinates gives us numerical facts about the space we live in. Algebraically, $\mathbf{R}^n$ is a system of arrows about an origin point, which can be added, subtracted, and scaled. But why do these arrows have to start at the origin? A differentiable curve $c : (a,b) \to \mathbf{R}^n$ has a tangent vector $c'(x)$, canonically pictured as eminating from the point $c(x)$. The idea of a vector space eminating from each point on a manifold will pave the way to all future endeavors. We call such a space a vector bundle, a pair (E, B) of spaces, called the **total** and **base** space respectively, with a continuous function π mapping E onto B, such that at each point $p \in B$, the **fibre space** $\pi^{-1}(p)$ forms a real vector space. We denote this vector space by B_p. A vector bundle must also be locally trivial. For every point $p \in B$, there is a neighbourhood U of p, a space $\mathbf{R}^k$, and a homeomorphism ϕ mapping $\pi^{-1}(U)$ onto $U \times \mathbf{R}^k$, linear on each fibre. Every connected component of B has a unique dimension k for which (2) holds. A vector bundle is **k-dimensional** if the dimension k is invariant across all of B. If E and B are manifolds, and the mappings ϕ and E differentiable, we call the vector bundle **differentiable**. Collectively, we shall denote a vector bundle by (E, B, π).

Example. *Consider any topological space U, and let $\varepsilon^n(U) = U \times \mathbf{R}^n$. Denote $(p, v) \in \varepsilon^n(U)$ by v_p, and define π by $v_p \mapsto p$. If we define $v_p + w_p = (v+w)_p$, and $c(v_p) = (cv)_p$, we form a vector bundle on U, known as the* **trivial bundle**. *We may picture $\varepsilon^1(\mathbf{R})$ as the plane, and looking at each $\mathbf{R}_p$ will tell you why we call such a space a fibre. We shall denote $\varepsilon^n(\mathbf{R}^n) = T\mathbf{R}^n$.*

Example. *The Möbius strip M can also be define to be $\mathbf{R} \times \mathbf{R}$, with (x, y) identified with $(x+n, (-1)^n y)$. We can also see S^1 as R with each integer identified. Then the projection $\pi(x, y) = x$ induces a projection $\pi : M \to S^1$. (M, S^1, π) is then a tangent bundle, if we ascribe a vector space structure to M by defining*

$$[(x,y)] + [(x,z)] = [(x, y+z)] \qquad c[(x,y)] = [(x, cy)]$$

This bundle is known as the **Möbius bundle** *on S^1.*

If (E, A, π) and (F, B, ϕ) are two vector bundles, a **bundle map** between the two is a pair of continuous functions $(f, f_\sharp)$ with $f : A \to B$ and $f_\sharp : E \to F$, where $f_\sharp$ maps A_p linearly into $B_{f(p)}$, for any $p \in A$. We shall denote $(f_\sharp)|_{E_p}$ as $(f_\sharp)_p$, so that each $(f_\sharp)_p$ as linear.

$$\begin{tikzcd} E \arrow[r, "f_\sharp"] \arrow[d, "\pi"] & F \arrow[d, "\phi"] \\ A \arrow[r, "f"] & B \end{tikzcd}$$

Such a map makes the diagram above commute. An isomorphism in the category of bundle maps is known as an equivalence. Here, bundle maps are introduced to extend the notion of a derivative.

Example. *The derivative of a map at a point is a linear approximation. We can collect these maps together to form a bundle map: if $f \in C^\infty(\mathbf{R}^n, \mathbf{R}^m)$, then we may construct $f_* : T\mathbf{R}^n \to T\mathbf{R}^m$, defined by*

$$f_*(v_p) = [Df(p)(v)]_{f(p)}$$

This is the 'infinitisimal action' of f. We have $(id_{\mathbf{R}^n})_* = id_{T\mathbf{R}^n}$. The chain rule in calculus implies that

$$\begin{aligned}(g \circ f)_*(v_p) &= [D(g \circ f)(v)]_{(g \circ f)(p)} = [Dg(f(v)) \circ Df(v)]_{g(f(p))}\\ &= g_*([Df(v)]_{f(p)}) = (g_* \circ f_*)(v_p)\end{aligned}$$

Giving us the nice formula $(g \circ f)_* = g_* \circ f_*$. Similar results hold for open submanifolds $U \subset \mathbf{R}^n$, giving us the tangent spaces TU.

Our current aim is to define, for each manifold M, a bundle TM extending off the space, such that if $f : M \to N$ is differentiable, then it induces a map $f_* : TM \to TN$, which possesses the properties of infinitisimal differentiation. There are many candidates, and all are useful at some point or another. The only structure we have on a manifold is its atlas $\mathcal{A}$, so we might as well make use of this to define the tangent bundle. We shall define TM_p, for each p separately, and then put a topological structure on the disjoint union of the TM_p. Let TM_p shall consist of an equivalence class of the set of all tuples (x, v), where $x : U \to \mathbf{R}^n$ is a chart in a neighbourhood of p, v is a vector in $\mathbf{R}^n$. We shall identify (x, v) and (y, w) if $(x \circ y)_*(v_{x(p)}) = w_{y(p)}$. Then TM_p is seen to be n dimensional, since every element can be identified with an element of $\{x\} \times \mathbf{R}^n$, for each chart x around p. Given a map $f : M \to N$ differentiable at p, we define $f_* : TM \to TN$, defined by

$$f_*([x, v]_p) = [y, D(y \circ f \circ x^{-1})(x(p))(v)]_{f(p)}$$

f_* is well defined by the equivalence relation imposed on the tangent bundle. It is obviously a linear map. We obtain a tangent bundle structure if we let each x_* be a homeomorphism and local trivialization of the space. We still have the identity $(f \circ g)_* = f_* \circ g_*$, and $(\mathrm{id}_M)_* = \mathrm{id}_{TM}$. In fact, our new definition of the tangent bundle agrees with the previously defined definition on $T\mathbf{R}^n$, which naturally injects corresponds to $\{[\mathrm{id}_{\mathbf{R}^n}, v]_p : v \in \mathbf{R}^n\}$.

The tangent bundle TM has a natural smooth structure induced by the smooth structure on M. If (x, U) is a chart, then we will let $x_* : TM \to Tx(U)$ be a chart on TM. These charts are C^∞ related, for

$$[y_* \circ (x_*)^{-1}](v_w) = y_*([x, v]_{x^{-1}(w)}) = \left[D(y \circ x^{-1})(w)(v)\right]_{(y \circ x^{-1})(w)}$$

and the total derivative is just the sum and product of the partial derivatives of $y \circ x^{-1}$, which we know to be C^∞.

3.2 The Space of Derivations

The algebraists found another characterization of the tangent bundle, which is an elegant, though much more abstract introduction to the bundle. Since $C^k(M)$ is a vector space, we may consider the dual space $C^k(M)^*$, which consists of all linear maps from $C^k(M)$ to $\mathbf{R}$. Let

$$\mathrm{Der}_p^k(M) = \{l \in C^k(M)^* : l(fg) = f(p)l(g) + g(p)l(f)\}$$

be the space of **derivations at** p in $C^k(M)$. A C^k map $f : M \to N$ induces a linear map $f^* : C^k(N) \to C^k(M)$, defined by

$$f^*(g) = g \circ f$$

which induces a map $f_* : C^k(M)^* \to C^k(N)^*$, defined by

$$[f_* l](g) = l(f^*(g))$$

If l is a derivation at p, then

$$\begin{aligned} f_*(gh) = l(f^*(gh)) &= l((gh) \circ f) = l((g \circ f)(h \circ f)) \\ &= g(f(p))l(h \circ f) + h(f(p))l(g \circ f) \\ &= g(f(p))f_*(h) + h(f(p))f_*(g) \end{aligned}$$

so $f_*(l)$ is a derivation at $f(p)$, and $f_* : \mathrm{Der}_p^k(M) \to \mathrm{Der}_{f(p)}^k(M)$. We can directly calculate that $(f_* \circ g_*) = (f \circ g)_*$, so that if f is a diffeomorphism, f_* is an isomorphism. The differential operators

$$\left.\frac{\partial}{\partial x^1}\right|_p, \dots, \left.\frac{\partial}{\partial x^n}\right|_p$$

are all derivations. We will show, in fact, that these operators span the space of all derivations.

Lemma 3.1. *If $f \in C^k(M)$ is a constant map, and $l \in Der_p^k(M)$, then $l(f) = 0$.*

Proof. By scaling, assume without loss of generality that $f(p) = 1$ for all p. Then

$$l(f) = l(f^2) = l(f) + l(f) = 2l(f)$$

which verifies triviality. □

Lemma 3.2. *If $l \in Der_p^k(M)$, and $f(p) = g(p) = 0$, then $l(fg) = 0$.*

Proof.

$$l(fg) = f(p)l(g) + g(p)l(f) = 0 + 0 = 0$$

We verified the proof by direct calculation. ☐

Theorem 3.3. *$Der_0^\infty(\mathbf{R}^n)$ is n dimensional, with basis*

$$\left.\frac{\partial}{\partial x^1}\right|_0, \dots, \left.\frac{\partial}{\partial x^n}\right|_0$$

Proof. Let $f \in C^\infty(\mathbf{R}^n)$. Using Taylor's theorem, we may write

$$f(x) = f(0) + \sum_{k=1}^n x^i \left.\frac{\partial f}{\partial x^i}\right|_0 + \sum_{i=1,j=1}^n \frac{x^i x^j}{2} g_{ij}(x)$$

where $g_{ij} \in C^\infty(\mathbf{R}^n)$. If l is an arbitrary derivation at zero,

$$\begin{aligned} l(f) &= l(f(0)) + \sum_{k=1}^n \left.\frac{\partial f}{\partial x^i}\right|_0 l(x^i) + \sum_{i=1,j=1}^n l(x^i x^j g_{ij}) \\ &= \sum_{k=1}^n \left.\frac{\partial f}{\partial x^i}\right|_0 l(x^i) \end{aligned}$$

Since the higher order terms of the series are the product of the functions x^i and $x^j g_{ij}$, which are both zero at 0. But this implies that l is determined by its values on $l(x^i)$

$$l = \sum_{k=1}^n l(x^i) \left.\frac{\partial}{\partial x^i}\right|_0$$

☐

We can extend this theorem to arbitrary smooth manifolds.

Lemma 3.4. *If $l \in Deriv_p^k(M)$, and f and g are equal in a neighbourhood of p, then $l(f) = l(g)$.*

Proof. We shall prove that if $f = 0$ in a neighbourhood U of p, then $l(f) = 0$. Consider a bump function $\psi \in C^\infty(M)$ such that $\psi = 1$ at p, and $\psi = 0$ outside of U. Then $\psi f = 0$, and

$$0 = l(0) = l(\psi f) = \psi(p)l(f) + f(p)l(\psi) = l(f)$$

☐

Corollary 3.5. *If U is an open neighbourhood of p, then $Der_p^k(M) \cong Der_p^k(U)$.*

If f is only defined in a neighbourhood U of p, we may still compute a well-defined value $l(f)$. Consider a function $\psi = 1$ in $V \subset U$, and equal to zero outside of U. Then $\psi f \in C^\infty(M)$, and $l(\psi f)$ is invariant of the bump function chosen, by the last lemma. Thus $\mathrm{Der}_p^k(M)$ acts of the **germ** of functions defined in a neighbourhood of p.

Theorem 3.6. *$Der_p^\infty(M)$ is n-dimensional, and if (x, U) is a chart centered at p, then a basis for $Der_p^k(M)$ is*

$$\left.\frac{\partial}{\partial x^1}\right|_p, \dots, \left.\frac{\partial}{\partial x^n}\right|_p$$

Proof. $\mathrm{Der}_p^k(M) \cong \mathrm{Der}_p^k(U) \cong \mathrm{Der}_p^k(x(U))$. The inverse image of this congruence maps the partial derivatives at zero to the partial derivatives at p. □

If we consider the disjoint union $\mathrm{Der}_p^\infty(M)$ as we vary $p \in M$, we obtain a structure isomorphic to TM. The correspondence is

$$[x, v]_p \mapsto \sum_{k=1}^n v_i \left.\frac{\partial}{\partial x^i}\right|_p$$

which induces a topology (and smooth structure) on derivations making the correspondence a homeomorphism (diffeomorphism). We will rarely distinguish between the two sets. We will speak of a tangent vector operating on functions, or of a derivation in the tangent space.

3.3 Vector Fields

A **section** on a vector bundle (E, B, π) is a continuous map $f : B \to E$ for which $\pi \circ f = \mathrm{id}_B$. We want to consider **vector fields**, which are sections from a manifold to its tangent bundle. We denote a vector field by capital letters like X, Y, and Z, and denote the value of a vector field X at a point p by X_p. Vector fields form a vector space. Locally, around a chart (x, U), we may express

$$X_p = \sum a^i(p) \left.\frac{\partial}{\partial x^i}\right|_p$$

This vector field is differentiable or continuous if and only if the functions a^i are differentiable or continuous. The space of all differentiable vector fields is itself a vector space, an algebra over $C^\infty(M)$. If X is a vector space, and $f \in C^\infty(M)$, then we define a new function $X(f) \in C^\infty(M)$, defined by

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

Since X_p is a derivation on $C^\infty(M)$. In general, we call a map $F : A \to A$ a derivation if $F(ab) = aF(b) + bF(a)$. If $F : C^\infty(M) \to C^\infty(M)$ is any derivation, and if f and g agree in a neighbourhood of p, then $F(f)(p) = F(g)(p)$, by the same trick for derivations at a point. We may then define a vector field X by the equation

$$X_p(f) = F(f)(p)$$

This vector field is unique, for if locally,

$$X = \sum a^i \frac{\partial}{\partial x^i}$$

Then $a^i(p) = X_p(x^i) = F(x^i)(p)$. The derivation corresponding to X is F, so C^∞ vector fields and derivations are in one to one correspondence.

Bibliography

[1] Michael Spivak, *A Concise Introduction to Differential Geometry: Vol. One*

[2] James Lee, *An Introduction to Smooth Manifolds*

[3] Paul Halmos, *Naive Set Theory*

[4] Wikipedia, *Lie Groups*