Differential Geometry

Jacob Denson

October 24, 2018

Table Of Contents

Chapter 1

Topological Considerations

In some form of mathematical heaven, all objects would exist in the linear realm, where problems are easy to solve. 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 curves bound to a sphere, forced never to move in a linear fashion. When an algebraic geometer studies the solution set of the equation $X^2 + Y^3 = 5$ in the plane, he must analyze a shape which bends and curves, never straight. Differential geometry gives the mathematician tools to cheat: Most of the time, the shapes we study may not be non-linear, but are at least *locally linear*, so around each point we can recover the methods used in the linear case. The challenge is to figure out how to put all the locally linear properties together into a nice, global form.

1.1   Manifolds and Atlases

Topology attempts to describe the properties of space invariant under continuous stretching and squashing. Differential geometry extends this description to spatial properties constant when space is stretched and squashed, but not 'bent'. Four centuries of calculus have established differentiability in the nice cartesian spaces $\mathbf{R}^n$. A basic environment to extend the notions of differentiability to topological spaces are those which are locally similar to $\mathbf{R}^n$. A **topological manifold** is a Hausdorff space M such that at every point $p \in M$, there exists a neighbourhood U of p, and a non-negative integer $n \geqslant 0$ such that U is homeomorphic to $\mathbf{R}^n$.

Remark. It is convenient in the theory to let $n = 0$ be a permissible dimension of space for a manifold to be homeomorphic to, where $\mathbf{R}^0 = \{0\}$ is a point. This makes sense, because most techniques involve the use of linear algebra, and $\mathbf{R}^0$ is just the canonical example of a zero dimensional vector space. We assume our manifolds are Hausdorff because non-Hausdorff manifolds are far and few between in applications of manifold theory to other areas of mathematics, and make the theory less managable.

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 $\mathbf{R}^n$, by projection down onto the axis of the domain.

The above examples exemplify the fact that any topological space homeomorphic to a topological manifold is also a topological manifold. This is a bad omen, because we want to discuss properties of space which remain invariant under differentiable maps, and differentiability should certainly be a stronger concept than homeomorphism on more general topological spaces. This indicates we must add additional structure to a manifold, so we may distinguish the class of smooth maps from the class of continuous ones. We introduce this structure in the next chapter.

Geometrically a homeomorphism $x : U \to \mathbf{R}^n$ from an open set U on a manifold can be seen as a way of assigning coordinates to points on the manifold, because we associate with each geometric point $p \in U$ a sequence of numbers $x^1(p), \ldots, x^n(p)$. One way to see the study of manifolds is as an extension of analytic geometry to non-planar topological systems. Indeed, the technologies developed have immediate applications to projective, hyperbolic, and elliptic geometries, and the language of manifolds has become the common language of almost all modern day geometers.

Example. *Consider the circle $S^1 = \{x \in \mathbf{R}^2 : |x| = 1\}$. For any proper open 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 of U in $\mathbf{R}$, with continuous 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.

As a manifold, the circle is distinct from $\mathbf{R}^n$ because we cannot put coordinates over the whole space at once; instead, 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 just Euclidean space.

Example. *The method of stereographic projection provides another way of coming up with a system of coordinates on the circle, which generalizes to higher dimensional variants of the circle. We will project the open subset $S^1 - \{(1,0)\}$ to the line $\{-1\} \times \mathbf{R}$, by taking the intersection of the line between p and $(1,0)$ and the line $\{-1\} \times \mathbf{R}$. A formula for the projection from $S^1 - \{1,0\}$ to the line $\{-1\} \times \mathbf{R}$ is given by the continuous function*

$$f(x,y) = \frac{2y}{1-x}$$

Another calculation shows the inverse is given by

$$f^{-1}(y) = \left(1 - \frac{8}{4+y^2}, \frac{4y}{4+y^2}\right)$$

Similar calculations show that the on the n-dimensional spheres S^n, we ahve a projection map from $S^n - \{(1,0,\ldots,0)\}$ onto $\{-1\} \times \mathbf{R}^n$ by the same process, and one calculates that the formula for the projection and it's inverse is given by

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

$$f^{-1}(y_2,\ldots,y_n) = \left(1 - \frac{8}{4+\|y\|^2}, \frac{4y_2}{4+\|y\|^2}, \ldots, \frac{4y_n}{4+\|y\|^2}\right)$$

$$= \frac{1}{4+\|y\|^2}\left(\|y\|^2 - 4, 4y_2, \ldots, 4y_n\right)$$

If we project from the point $(-1,0,\ldots,0)$ to $\{1\} \times \mathbf{R}^{n-1}$ instead of $(1,0,\ldots,0)$, then the homeomorphism defined on $S^1 - \{(-1,0,\ldots,0)\}$ is calculated to be

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

$$g^{-1}(y_1,\ldots,y_{n-1}) = \frac{1}{4 + \|y\|^2} \left(4 - \|y\|^2, 4y_2, \ldots, 4y_n\right)$$

And we have covered S^n with homeomorphisms. We conclude that the space is a manifold.

Example. *Any open subset of $\mathbf{R}^n$ is a manifold; around any point, we may take an open ball as a neighbourhood, 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 subsets of $\mathbf{R}^n$. We will call a homeomorphism $x : U \to V$, where V is an open subset of Euclidean space a **chart**, and denote it by (x, U). The letters x, y and z are often used for charts, to make it easy to confuse coordinates $(x^1, x^2, \ldots, x^n) \in \mathbf{R}^n$ with coordinates $(x^1(p), x^2(p), \ldots, x^n(p))$ on a manifold. It is often very useful to trick your brain into the coordinate way of thinking, and the only disadvantage is that things can become somewhat confusing working on subsets of Euclidean space.

Example. *Consider the set $M(n)$ of $n \times n$ matrices with entries in the real numbers. We can identify $M(n)$ with $\mathbf{R}^{n \times n}$ by considering only the coefficients of the matrix, and this tells us $M(n)$ is a topological manifold of dimension n^2 (and more generally, the set $M(n, m)$ of $n \times m$ real matrices has dimension nm). The determinant map $\det : M(n) \to \mathbf{R}$ can be viewed as a polynomial in the entries of the matrix, so the function is continuous, and so the general linear group $GL(n)$ consisting of invertible matrices is an open submanifold of $M(n)$, since $GL(n) = \det^{-1}(\mathbf{R} - \{0\})$. An important fact about $GL(n)$ is that the operations of multiplying matrices is also continuous. This is easily seen, since multiplication is just a polynomial in the entires of the matrices. By Cramer's rule, the operation mapping M to M^{-1} is also continuous and differentiable, because it is a rational function of the matrix elements which doesn't have any singularities in $GL(n)$.*

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

$$PMQ = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where $A \in GL(k)$. The endomorphism L on $M(n, m)$ defined by

$$L(N) = PNQ = \begin{pmatrix} A(N) & B(N) \\ C(N) & D(N) \end{pmatrix}$$

is rank preserving, and by continuity, the $A(N)$ is still invertible for N in a suitably small neighbourhood of M. The matrix

$$\begin{pmatrix} I_k & 0 \\ -C(N)A^{-1}(N) & I_{n-k} \end{pmatrix}$$

is invertible, and so $L(N)$ has the same rank as

$$\begin{pmatrix} I_k & 0 \\ -C(N)A^{-1}(N) & I_{n-k} \end{pmatrix} \begin{pmatrix} A(N) & B(N) \\ C(N) & D(N) \end{pmatrix}$$
$$= \begin{pmatrix} A(N) & B(N) \\ 0 & D(N) - C(N)A^{-1}(N)B(N) \end{pmatrix}$$

and this matrix has rank k if and only if $D(N) = C(N)A^{-1}(N)B(N)$. Thus, locally around M we may specify an element of $M(n, m; k)$ by $A \in GL(k)$, $B \in M(k, n-k)$, and $C \in M(k, m-k)$, so we conclude $M(n, m; k)$ is a

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

dimensional manifold. As a special case, this verifies that if $k = \min(m, n)$, then the set $M(n, m; k)$ of matrices of full rank is actually an open subset of $M(n, m)$.

The manifold $M(n, m; k)$ is interesting, because its canonical embedding in Euclidean space is essentially just expressed along canonical coordinates obtained from the coefficients of the matrix entries by forgetting a few coefficients, yet the space cannot be globally linear because of the rank k requirement. 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 union of the x and y axis, with the origin removed, and is a 1 manifold. At any point you can travel left and right, or up and down, just not too far!

1.2 Constructing Abstract Manifolds

Sometimes, you may wish to place a topological structure on a set, with the purpose of making it into a manifold, without having any topological structure to begin with. In most cases, the set is contained in a bigger space with some topological structure, and you can just surplant the relative topology on the subspace. But in certain cases (most famously, the construction of moduli spaces in algebraic geometry), there is nothing to start working with. The next theorem provides a useful strategy for placing a topology on a space and simultaneously showing it is a manifold. We will later show it has a simpler extension to placing a topology and a structure of a differentiable manifold onto a space.

Theorem 1.1. *Let M be a set covered by a family U_α of sets, together with bijections x_α from U_α onto open subsets of euclidean space, such that*

- *For any α, β, $x_\alpha(U_\alpha \cap U_\beta)$ is open, and $x_\alpha \circ x_\beta^{-1}$ is continuous.*

- *If $a, b \in M$, then either there is U_α with $a, b \in U_\alpha$, or disjoint $U_\alpha, U_\beta = \varnothing$ such that $a \in U_\alpha$, $b \in U_\beta$.*

Then there is a unique topological structure on M making it into a Hausdorff manifold, such that each (x_α, U_α) is a chart.

Proof. Consider the topological structure on M generated by all sets of the form $x_\alpha^{-1}(V)$, where V is an open subset of $x_\alpha(U_\alpha)$. These sets form a base for the topology by the first condition of the theorem; if $x_\alpha(U)$ and $x_\beta(V)$ are open, then

$$x_\alpha(U \cap V) = x_\alpha(U) \cap (x_\alpha \circ x_\beta^{-1})(x_\beta(U_\alpha \cap V))$$

is open, since it is the intersection of open sets. This also shows that if $V \subset U_\alpha$ is open in our topology, then $x_\alpha(V)$ is open, so that each x_α is a homeomorphism, and the space is Hausdorff since points in a single U_α may be separated by Euclidean open sets, and points in disjoint U_α, U_β may surely be separated. We conclude that M is a manifold under this new topological structure. $\square$

Remark. If the family of charts is countable, the manifold we obtain will be second countable. More generally, we may also assume the charts above are maps into manifolds rather than open subsets of Euclidean space, since, locally, these identifications are essentially the same thing.

Example. *Any finite dimensional vector space V over the real numbers has a natural vector space structure. We associate to each basis $B = (v_1, \ldots, v_n)$ the map $f_B : V \to \mathbf{R}^n$ with $f(\sum a^i v_i) = \sum a^i e_i$. If $B' = (w_1, \ldots, w_n)$ is another basis, then there are b^j such that $v_i = \sum b_i^j w_j$, and so*

$$(f_{B'} \circ f_B^{-1})(a) = \sum a^i b_i^j e_j$$

and this map is clearly continous in the values a. Of course, the theorem above isn't really required, since each basis induces a map on the entire space, but our calculations here show that all bases give a homeomorphism. This provides another explanation as to why $M(n)$ is a manifold.

Example (Grassmannian). *Our first example using this construction is the assignment of a topological structure to the space $G(k,n)$ of k dimensional subspaces of $\mathbf{R}^n$, which makes clear the intuitive idea of two subspaces being 'close to one another'. Using the construction in the above theorem, given some subspace V of $\mathbf{R}^n$ of dimension $n - k$, we will let A_V consist of all k dimensional subspaces of $\mathbf{R}^n$ which intersect V trivially. We will show that the A_V can be made into a compatible covering of $G(k,n)$ which makes the space into a $k(n-k)$ dimensional manifold. We shall heavily rely on the fact that, given any k dimensional subspace V of $\mathbf{R}^n$, there is a (not necessarily unique) $n - k$ dimensional complementary subspace V' such that $\mathbf{R}^n = V \oplus V'$.*

It is easy to see that the $A_{V'}$ cover $\mathbf{R}^n$, because $V \in A_{V'}$ for any complementary V'. Given that $\mathbf{R}^n = V \oplus V'$, we will show that $A_{V'}$ corresponds to the set $L(V, V')$ of linear maps from V into V'. To do this, we take a map $T : V \to V'$ to the subspace $V_T = \{v + Tv : v \in V\}$. Denote the projection from $\mathbf{R}^n$ to V by π, and the projection onto V' by ψ. It is clear that the correspondence is injective, for if $V_T = V_S$, then for every $v \in V$ there is $w \in V$ such that $v + Tv = w + Sw$, and then the direct sum decomposition gives $v = w$ and $Tv = Sw = Sv$, so $T = S$. Conversely if $W \in A_{V'}$, then π is an isomorphism between W and V, and $W = V_{\psi \circ \pi^{-1}}$. Thus the correspondence is surjective. Since $L(V, V')$ is homeomorphic to $k(n-k)$ dimensional Euclidean space, it remains to show this correspondence is suitably regular to give sufficient topological structure to $G(k,n)$.

To show we get a Hausdorff topological structure, let V and W be two subspaces of $\mathbf{R}^n$ of dimension k, and suppose $V \cap W$ has dimension m. Then $V + W$ is a space of dimension $2k - m$, and somce choice of $(V + W)'$ has dimension $n - (2k - m) = (n - k) - (k - m)$. Since $(V + W)'$ intersects V and W trivially,

it now suffices to prove that $V + W$ has a subspace of dimension $k - m$ that intersects V and W trivially. Note that both

$$X = V' \cap (V + W) \quad Y = W' \cap (V \oplus W)$$

have dimension $k - m$. If we consider any isomorphism $T : X \to Y$, then the space X_T is $k - m$ dimensional and intersects both V and W trivially. Thus $X_T \oplus (V + W)'$ is the required subspace.

A more difficult task is to show that the charts are compatible with one another. First, we have to show that the set $A_{V'} \cap A_{W'}$ of spaces U which intersect V' and W' trivially is an open subset of linear transformations. Let π, ψ, ν and η denote the projections onto V, V', W, and W' respectively. The image of $A_V \cap A_W$ in $L(V, V')$ consists of maps $T : V \to V'$ such that $\nu \circ (1 + T)$ is an isomorphism. The association $T \mapsto \pi \circ (1 + T)$ is continuous, which shows that the set of all such isomorphisms is open, since this is the inverse image of the maps $\pi \circ (1 + T)$ which have nonzero determinant.

Now how does $A_{V'} \cap A_{W'}$ transform between $L(V, V')$ and $L(W, W')$? Consider a subspace $U \in A_{V'} \cap A_{W'}$ mapped to a transformation $T : V \to V'$, and a transformation $L : W \to W'$. Since U intersects V' and W' trivially, the projection maps $\pi|_U$ and $\nu|_U$ are isomorphisms, and $L = \eta \circ (\nu|_U)^{-1}$. Note that the map $(1 + T) : V \to U$ is an isomorphism, and so

$$L = (\eta|_U) \circ (1 + T) \circ (1 + T)^{-1} \circ (\nu|_U)^{-1} = [\eta \circ (1 + T)] \circ [\nu \circ (1 + T)]^{-1}$$

a formula now expressed independently of U, which holds over all choices of T. The map

$$T \mapsto [\eta \circ (1 + T)] \circ [\nu \circ (1 + T)]^{-1}$$

is a composition of continuous maps, hence the charts are consistant, and this concludes the construction that $G(k, n)$ is a topological manifold. If V is a finite dimensional vector space, then the construction above can be essentially copied to show the space $G_k(V)$ of k dimensional subspaces of V is also a manifold, also called a Grassmanian. A special case is the space $G_1(V)$ of lines through the origin in V, known as the projectivization of V and denoted by $\mathbf{P}(V)$.

1.3   Basic Properties of Manifolds

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 1.2. *Every manifold is locally compact.*

Proof. Let p be an arbitrary point on a manifold, with a chart (x, U). If we consider any closed ball around $x(p)$ in $x(U)$, the inverse image will be compact, since compactness is preserved under homeomorphisms. Thus p has a compact neighbourhood. $\square$

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 1.3. *A connected manifold is path-connected.*

Proof. Let p be a point on a connected manifold M, and consider the set U of all points in M path connected to p. Local path connectedness shows U is open. Suppose q is a limit point of U. Take some path-connected chart (x, V) around q. Then V contains some $r \in U$, which is path connected to p, and we can then patch this path with a path from r to q obtained in the chart to obtain a path from p to q. Thus U is open, closed, and non-empty, so $U = M$. $\square$

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. *$GL(n)$ is a disconnected manifold, since $\det(GL(n))$ is disconnected. Because of the last theorem, we can identify the connected components of $GL(n)$ by identifying the path connected components. To do this, we shall construct paths in $GL(n)$ reducing invertible matrices to certain canonical forms. In our proof, we shall use the fact that $GL_n(\mathbf{R})$ can be described as tuples of n linearly independent vectors in $\mathbf{R}^n$, which simplifies notation. If $v_1, \ldots, v_n$ are linearly independent, consider adding one vector to another.*

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

These vectors are path connected in $GL_n(\mathbf{R})$ by the path

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

Similarily, we may subtract rows from one another. Next, consider multiplying a row by a scalar $\gamma > 0$,

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

These matrices are path connected by

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

We cannot perform this technique if $\gamma < 0$, because then $(1-t) + t\gamma = 0$ for some choice of t, and the resulting vectors become linearly dependent. 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 determines at least two connected components. The same reasoning 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 $GL(n)$ consists of two homeomorphic path-connected components: the matrices with determinant greater than zero, and the component with determinant less than zero, reflecting the linear transformations which preserve orientation, and the ones which reverse orientation. This is quite a different situation from the space $GL(n, \mathbf{C})$ of invertible matrices over the complex numbers, which is connected for any n.

Another useful fact of a manifold, if the dimenison of the manifold is the same, is that all the local parts of space look identical to one another. We say a topological space X is **homogenous** if, for every two points $p, q \in X$, there is a homeomorphism mapping p to q.

Theorem 1.4. *Every connected manifold is homogenous.*

Proof. It is easy to show that $\mathbf{R}^n$ is homogenous. This theorem relies on the fact that we can extend a 'local homeomorphism' to a global homeomorphism. To see how this is done, consider the closed unit disk $\mathbf{D}^n$ in $\mathbf{R}^n$. Suppose that, for any two points p, q in the interior of the unit disk, there is a homeomorphism f of *the entire Euclidean space* $\mathbf{R}^n$ mapping p to q, such that $f(x) = x$ for all points outside of the unit ball. Given two points p, q on a general manifold M contained within a common chart (x, U), we can choose some closed disk K which p and q both lie in, find a homeomorphism $f : U \to U$ mapping p to q, and fixing all points outside the unit disk, and then f may be extended to a homeomorphism on M by fixing all points outside the chart. f continuous to be continuous, and a homeomorphism, because it is continuous on two overlapping open sets U and $M - K$, and these cover M. Thus all that remains is to prove the lemma for the closed disk, and this isn't too difficult. We define a smooth bump function ψ which is equal to one on the line segment between p and q, and vanishes outside of $\mathbf{D}^n$. We then consider the smooth vector field $v(x) = \psi(x)(q - p)$. This generates a family of evolution maps φ_t, which are homeomorphisms of $\mathbf{R}^n$, which fix all points outside of the unit disk because the vector field vanishes there, and for some t, $\varphi_t(p) = q$, which is the homeomorphism we wanted. $\square$

1.4   Construction in the Category of Manifolds

The category **Man** of topological manifolds has useful constructions which are good to know as tools for constructing more manifolds from simpler ones. The coproduct (disjoint-union) of two manifolds is easily shown to be a manifold. Here are some more constructions.

Theorem 1.5. *If M and N are manifolds, then the product space $M \times N$ is also a manifold*

Proof. Products of homeomorphisms onto open subsets of $\mathbf{R}^n$ suffice. $\square$

Example. *Since S^1 is a one dimensional manifold, we obtain a two dimensional manifold $\mathbf{T}^2 = S^1 \times S^1$, the torus. More generally, the n torus $\mathbf{T}^n = S^1 \times \cdots \times S^1$ is an n manifold.*

Example (Surfaces of Revolution)**.** *If M is a 1-manifold in the subplane of* $\mathbf{R}^3$ *defined by*

$$X = \{(x, y, 0) : x, y > 0\}$$

Then the space obtained by rotating M around the y axis is a 2-manifold, known as a surface of revolution, and homeomorphic to $M \times S^1$.

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, because we can push sufficiently small charts on M into charts on N. This is satisfied for many of the useful examples which occur in the earlier parts of the quotient theory of manifolds.

Example (The Möbius Strip)**.** *Consider the quotient space M obtained from* $[-\infty, \infty] \times (-1, 1)$ *by identifying* (x, y) *with* $(x + n, (-1)^n y)$, *for* $n \in \mathbf{Z}$. *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 two dimensional manifold, denoted* $\mathbf{RP}^2$ *and known as real projective space. In general* $\mathbf{RP}^n$ *is created by identifying opposite points on* S^n. *Another way to describe the space is as* $\mathbf{R}^{n+1} - \{0\}$, *where x and* λx *are identified, for* $\lambda \neq 0$. *To obtain explicit charts on* $\mathbf{RP}^n$, *define* $x : S^n \to \mathbf{R}^n$ *by*

$$x(a_1, \ldots, a_{n+1}) = \frac{1}{a_i}(a_1, \ldots, \widehat{a}_i, \ldots, a_{n+1})$$

This map is continuous whenever $a_i \neq 0$. *For all points p,* $x(p) = x(-p)$, *so the chart descends to a map on* $\mathbf{RP}^n$ *instead. It even has a continuous inverse*

$$x^{-1}(b_1, \ldots, b_n) = [b_1, \ldots, 1, \ldots, b_n]$$

Since our maps cover the space, $\mathbf{RP}^n$ *is an n dimensional manifold.*

Example. *A similar space can be formed in the case of complex scalars, by viewing $\mathbf{R}^{2n}$ as $\mathbf{C}^n$. We then identify z with λz, for $\lambda \in \mathbf{C} - \{0\}$ to form the quotient space $\mathbf{CP}^n$ from $\mathbf{C}^{n+1} - \{0\}$. For $n = 1$, $\mathbf{PC}^1$ is just the Riemann sphere, a 2 manifold. In general, $\mathbf{CP}^n$ is a 2n dimensional real manifold, or n 'complex' dimensions. We may consider the chart*

$$x(z_1,\ldots,z_{n+1}) = \frac{1}{z_i}(z_1,\ldots,\widehat{z_i},\ldots,z_{n+1})$$

with inverse

$$x^{-1}(w_1,\ldots,w_n) = [w_1,\ldots,1,\ldots,w_n]$$

where we consider complex multiplication instead of real multiplication. The space formed is known as complex projective space.

Even though $\mathbf{RP}^2$ 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 $\mathbf{RP}^2$, since our eye cannot see the full circumpherence of vision, but these topological problems occur in the algorithms involved in patching portions of vision across the entire circumpherence.

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 is unique up to homeomorphism, but this requires some tough topology to show for general manifolds. An example application is the construction of the n-holed torus $T\#T\#\ldots\#T\#T$, which is a surface embeddable in $\mathbf{R}^3$.

1.5 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 1.6 (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. In an appendix to this chapter, we shall prove the theorem based on the weaker assumption of the Jordan curve theorem, which is 'more intuitive' than invariance of domain.

Lemma 1.7. *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 $\pi : \mathbf{R}^n \to \mathbf{R}^m$

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

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**. As shorthand, we let M^n denote a manifold M which is n dimensional.

Corollary 1.8. *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 $n = m$. $\square$

Theorem 1.9. *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.

1.6 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 1.10. *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).

16

Lemma 1.11 $(1) \rightarrow (2)$. *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 $U_{x_1}, U_{x_2}, \ldots, U_{x_n}$. 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. $\qquad\square$

Lemma 1.12 $(2) \rightarrow (3)$. *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. $\qquad\square$

Lemma 1.13 $(3) \rightarrow (1)$. *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 $r : X \rightarrow \mathbf{R}$ is continuous:

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

17

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. $\qquad \square$

Lemma 1.14 $(4) \to (1)$. *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 $U_{1,1}, U_{1,2}, \ldots, U_{1,n_1}$. 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 $U_{2,1}, \ldots, U_{2,n_2}$, obtaining

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

18

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. $\qquad\square$

Lemma 1.15 $(1) \to (4)$. *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. Similarly, $C_2 \cup \overline{U_1}$ is contained in a precompact neighbourhood U_2 with compact closure. We find $U_1 \subset U_2 \subset \ldots$, 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 $U_1, \ldots, U_n$, and we let

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

be a collection of refined open sets which cover V_1. Do the same for each V_k, obtaining $\mathcal{V}_2, \mathcal{V}_3, \ldots$, 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. $\qquad\square$

1.7   Boundaries

There is an addition family of 'manifolds with sides' which often occurs in the theory of differential geometry. A **manifold with boundary** is a space also containing points that are *locally bounded*. Points in such a space must either be locally homeomorphic to some $\mathbf{R}^n$, or be a point on the 'boundary' of the manifold. That is, a point x lies on the **boundary** of the manifold if it has a neighbourhood homeomorphic to a 'halfspace' $\mathbf{H}^n = \{(x_1, \ldots, x_n) \in \mathbf{R}^n : x_1 \geqslant 0\}$. If M is a manifold with boundary, then we denote the set of points on its boundary by ∂M. This is well defined by the invariance of domain theorem, in the sense that a point on a manifold with boundary *either* has a neighbourhood homeomorphic to some $\mathbf{R}^n$, or to some $\mathbf{H}^m$, but not both. The non boundary points are known as the interior, denoted M°.

Theorem 1.16. *If M^n is a manifold with boundary, then ∂M, considered as a subspace of M, is a manifold (without boundary) of dimension $n-1$.*

Proof. Let x be a point in ∂M, and let U be a neighbourhood homeomorphic to $\mathbf{H}^n$ by a map $f : U \to \mathbf{H}^n$. Consider the points in U that map to the boundary plane under f,

$$V = \{y \in U : f(y) = (0, x_2, \ldots, x_n)\}$$

We contend that $V = U \cap \partial M$, so that V is a neighbourhood of x in the relative topology, and since V is homeomorphic to $\mathbf{R}^{n-1}$, this will show that ∂M is an $n-1$ dimensional manifold. It is easy to see that $V \subset U \cap \partial M$. Conversely, the other points in $U - V$ have a neighbourhood homeomorphic to $\mathbf{R}^n$, so that they do not lie in ∂M. Thus $U - V \subset M^\circ \cap U$, and this completes the proof. $\qquad\square$

Example. $\mathbf{H}^n$ *is the easiest example of a manifold with boundary. It's boundary consists of $\{0\} \times \mathbf{R}^{n-1}$, which is an $n-1$ manifold. Another manifold with boundary is the unit disc $D^n = \{x \in \mathbf{R}^n : \|x\| \leqslant 1\}$. We have already shown that the discs boundary, $\partial D^n = S^{n-1}$, is an $n-1$ manifold.*

We will see later that considering a manifold boundary as a disjoint structure is very useful. Stoke's theorem provides a direct application. But for now, let's move onto introducing calculus into the picture.

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 'non-differentiable' charts. 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. 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 abstract!

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**. The coordinate charts in the atlas of a smooth manifold are referred to as smooth, or curvilinear. In the literature, each map $y \circ x^{-1}$ is known as a **transition map**, so that we say 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.

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. Since differentiability is a *local* condition on Euclidean spaces, and manifolds are locally Euclidean spaces, the smooth structure, which preserves the differentiability across all local charts, allows us to define differentiability on arbitrary manifolds.

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. $\qquad\square$

This implies, in particular, that we may specify a smooth structure by just giving a family of C^∞ related transition maps covering a manifold, which is in almost all cases easier to specify.

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 $\mathcal{A}$.*

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. $\qquad\square$

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 smooth structure by considering the generating atlas containing only the identity map $id_{\mathbf{R}^n}$. This defines a smooth structure on $\mathbf{R}^n$, such that*

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.

We note that the transition maps given to define all the topological manifolds in the previous chapter are all C^∞ related to one another. It is left as an exercise to check this is true. Thus all the topological manifolds in the previous chapter are also differentiable manifolds. We make sure to note that there do exist some topological manifolds which have *no* differentiable structure – the creases in the manifold cannot be evened out. But these occur in certain pathological situations that we won't stumble upon in these introductory notes.

Example. *On $\mathbf{R}^2$, we have polar coordinates $(\theta, \mathbf{R}^2 - \{0\})$ defined 'in inverse coordinates' by the map $\theta^{-1}(r, u) = re^{iu}$. This chart is locally injective and has full rank at every point, so that θ^{-1} is actually invertible, and θ is C^∞, by the inverse function theorem. 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 to work with.*

Example. *The differentiable structure on S^n is defined by the stereographic projection maps. We may also define this structure on S^1 by the angle functions. If (θ, U) and (ψ, V) are angle functions, then $\theta \circ \psi^{-1}$ is just a translation by a multiple of 2π on each connected component of $\psi(U \cap V)$, hence C^∞. Indeed, if*

$$\psi^{-1}(t) = e^{it} = \theta^{-1}(t')$$

then $t = t' + 2\pi n$ for some unique integer n. Define $f(t) = n$, giving us a map $f : U \cap V \to \mathbf{Z}$. This map is continuous because if $t_i \to t$, then

$$(\theta \circ \psi^{-1})(t_i) \to (\theta \circ \psi^{-1})(t)$$

so $t_i + 2\pi f(t_i) \to t + 2\pi f(t)$, hence $f(t_i) \to f(t)$. The continuity of f implies that f is constant on every connected component of $U \cap V$.

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.*

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 $(-\infty, \infty) \times (0, 1)$.*

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. Later, we will show that every manifold which lies in Euclidean space has a natural smooth structure, however, and this will be taken as the canonical smooth structure in any of these cases.*

The C^∞ maps are not the only classes of maps occuring in calculus, and analogously, C^∞ manifolds are not the only types of manifolds. More generally, we can consider a C^k atlas, whose transition maps are C^k, and

form a C^k manifold. We can even specialize the family of C^∞ manifolds further by requiring transition maps are (real) analytic, in which case we obtain the family of C^ω manifolds, also known as analytic manifolds. If the dimension of the manifold is even, we can identify our charts into $\mathbf{R}^{2n}$ with $\mathbf{C}^n$, and if we require transition maps to be holomorphic, we obtain the family of complex manifolds. We will not discuss these types of manifolds in detail, but they occur in applications and in more general contexts.

2.2 Smooth Functions on a Manifold

The set of all real-valued differentiable maps defined on a manifold M form the algebra $C^\infty(M)$, contained with the algebra $C(M)$. Note that a continuous map $f : M \to N$ induces an algebra homomorphism $f^\# : C(N) \to C(M)$ defined by $f^\#(g) = g \circ f$. If f and g are C^∞, then $g \circ f$ is C^∞, and so we may restrict $f^\#$ to a map from $C^\infty(N)$ to $C^\infty(M)$. We see therefore that the 'map' C^∞ defines a contravariant functor from the category of differential manifolds to the category of algebras.

Lemma 2.4. *A continuous map $f : M \to N$ between manifold is smooth if and only if $f^\#(C^\infty(N)) \subset C^\infty(M)$.*

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 $y^i \circ f = f^\#(y^i)$ is differentiable, so that $y^i \circ f \circ x^{-1}$ is differentiable. But this implies $y \circ f \circ x^{-1}$ is differentiable, so f is differentiable. $\square$

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. Given $g = f^{-1}$, we note that $(g \circ f)^\# = f^\# \circ g^\#$ is just the identity map. Thus if f is a diffeomorphism, Then $f^\# \circ g^\#$ is the identity, so that $f^\#$ is invertible, and hence an isomorphism. Conversely, if $f^\#$ is a bijection between $C^\infty(N)$ and $C^\infty(M)$, then $f^\#(C^\infty(N)) = C^\infty(M) \subset f^\#(C^\infty(M))$, so f is differentiable, and $g^\#(C^\infty(M)) = C^\infty(N) \subset C^\infty(N)$, so g is also differentiable, hence f is a diffeomorphism. $\square$

Remark. More generally, f is C^k if and only if $f^\#(C^k(N)) \subset C^k(M)$, and a C^k diffeomorphism if and only if f maps $C^k(N)$ isomorphically to $C^k(M)$.

Suppose we know $C^\infty(M)$ for all manifolds 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. We can actually define a C^∞ manifold in a completely algebraic way. Given some topological space X, suppose we have a map associating each open subset U of X with a subalgebra $C^\infty(U) \subset C(U)$ containing the constant 1 function, such that

- For each $V \subset U$, if $f \in C^\infty(U)$ then $f|_V \in C^\infty(V)$.

- If $V = \bigcup U_\alpha$, and $f : V \to \mathbf{R}$ satisfies $f|_{U_\alpha} \in C^\infty(U_\alpha)$ for each α, then $f \in C^\infty(V)$.

- Every point $x \in X$ has a neighbourhood U with $x_1, \ldots, x_n \in C^\infty(U)$ such that $x = (x_1, \ldots, x_n)$ is a homeomorphism of U with an open subset of $\mathbf{R}^n$, and $f \in C^\infty(U)$ if and only if $f \circ x^{-1}$ is a C^∞ function.

Then there is a unique C^∞ structure on X such that the C^∞ real-valued maps on any open set U are precisely the elements of $C^\infty(U)$. These facts constitute the foundation of the algebraic viewpoint of manifold theory, which attempts to uncover the nature of manifolds solely by analyzing the commutative algebra $C^\infty(M)$. You can actually get pretty far with this approach: Nestruev's book "Smooth Manifolds and Observables" attempts to introduce differential geometry solely in this manner, but we prefer to introduce the geometric and algebraic approach simultaneously for maximal insight.

2.3   Partial Derivatives

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 $f \in C^1(M)$. We have no conventional coordinates to consider partial derivatives on, but if we fix some chart $x : U \to \mathbf{R}^n$ on M, we obtain a differentiable map $f \circ x^{-1}$, and we define, for a point $p \in U$,

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

Geometrically, this is the change in the function f when we trace the function along the coordinate lines from the map x; literally, if we define a curve $c(t) = (f \circ x^{-1})(x(p) + te_k)$, then

$$c'(0) = \frac{\partial f}{\partial x^k}\bigg|_p$$

Sometimes we use the notation $\partial_{x^k} f$ for the partial derivative in the k'th direction, which simplifies the notation in heavy calculations.

Theorem 2.6. *If x and y are coordinate systems at a point p, and $f \in C^1(M)$, then*

$$\frac{\partial f}{\partial x^i}\bigg|_p = \sum_j \frac{\partial y^j}{\partial x^i}\bigg|_p \frac{\partial f}{\partial y^j}\bigg|_p$$

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

$$\frac{\partial f}{\partial x_i}\bigg|_p = D_i(f \circ x^{-1})(x(p)) = D_i((f \circ y^{-1}) \circ (y \circ x^{-1}))(x(p))$$

$$= \sum D_j(f \circ y^{-1})(y(p)) D_i(y_j \circ x^{-1})(x(p))$$

$$= \sum \frac{\partial f}{\partial y_j}\bigg|_p \frac{\partial y_j}{\partial x_i}\bigg|_p$$

$\square$

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

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

(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

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

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

All the nastiness cancels out by use of the trigonometric identities, and we find

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

Partial derivatives also satisfy a nice 'derivation' property, which we leave to the reader to calculate. It is essentially the product rule for partial derivatives. We shall see later that this property characterizes the partial derivative maps in the space of linear maps on $C^\infty(M)$, which relates to the 'intrinsic' tangent space of a manifold.

29

Lemma 2.7. *If f and g are differentiable maps from a manifold M to* **R**, *then*

$$\frac{\partial f g}{\partial x^i}\bigg|_p = f(p)\,\frac{\partial g}{\partial x^i}\bigg|_p + g(p)\,\frac{\partial f}{\partial x^i}\bigg|_p$$

The partial derivatives provide immediate quantities to measure the rate of change of a real-valued differentiable function on a manifold. Next, we will consider a 'coordinate-independent' way to measure this rate of change for any two functions, by the introduction of 'infinitisimals' on manifolds.

2.4   The Differential Map

The coordinate operators allow us to extend differentiation to functions in $C^\infty(M)$, for any manifold M. However, defining the derivative of a differentiable map $f : M \to N$ between arbitrary manifolds is more tricky. We could define the derivative coordinatewise at a point p by considering the derivative $D(y \circ f \circ x^{-1})(x(p))$, for some 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. The trouble here is that there is no 'definite' space the operator is defined over – as we change the coordinate systems, the space changes. To obtain a 'universal' coordinate independent map, we need to form a 'universal' space which represents all coordinates at the same time, upon which the derivative operators are invariant. This space is known as the tangent bundle. There are deep and elegant constructions for this bundle, but for now, we only require the absolute basics.

Vectors v in $\mathbf{R}^n$ are often pictured as starting at the origin, and ending at the point with the same coordinates as v. But it is often convenient to picture these vectors as starting at a different beginning point than the origin. We introduce the notation v_p, where p is a point in Euclidean space, and v is a vector, to be the vector beginning at p and ending at $p + v$. One way to think of these values are as points p in $\mathbf{R}^n$ with an added 'second order' infinitisimal shift in the direction v. If U is an open subset of $\mathbf{R}^n$, we can set

$$TU = \{v_p : p \in U, v \in \mathbf{R}^n\}$$

to be the **tangent bundle** of U, consisting of all second order approximations of points in U. At each point $p \in U$, the **fibre** at p, denoted TU_p con-

sists of all tangent vectors beginning at p, which can be made into a vector space by defining the operations $v_p + w_p = (v + w)_p$ and $\lambda v_p = (\lambda v)_p$.

There are many complex and elegant ways to form the tangent bundle on a differentiable manifold. We will eventually discuss them in time, but we now construct the space in the most quick and dirty way. On a differentiable manifold M, we locally specify the tangent bundle in charts, and then patch them together into a reasonable structure on the entire manifold. For each point p, we let the fibre TM_p consist of equivalence classes of tuples of the form $(x, v)_p$, where $v \in \mathbf{R}^n$, and (x, U) is a chart with $p \in U$, where we identify $(x, v)_p$ and $(y, w)_p$ if

$$w = D(y \circ x^{-1})(p)(v)$$

In other words, if w looks like v stretched by a change in coordinates. Putting this altogether gives the required tangent vectors $[x, v]_p$. The fibres TM_p are still vector spaces by the operations $[x, v]_p + [x, w]_p = [x, v + w]_p$, and $\lambda[x, v]_p = [x, \lambda v]_p$.

If $f : U \to V$ maps an open subset $U \subset \mathbf{R}^n$ to $V \subset \mathbf{R}^m$, and is differentiable, then we can consider the derivatives $Df(p)$ at $p \in U$, which describe how the function acts locally around p. Since a differentation indicates that the space is 'locally linear', then the derivative should act on tangent vectors by the actual linear map. In other words, we can put all the derivatives together to define the **covariant derivative** of f, denoted f_*, by

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

Often, f_* is also denoted by df, and the map restricted to the domain of the fibre at p and the fibre at $f(p)$ is denoted df_p. The chain rule for derivative can be succinctly represented by the fact that if $g : V \to W$ is a differentiable function, then $(g \circ f)_* = g_* \circ f$. Now for a differentiable map $f : M \to N$, we can define the covariant derivative $f_* : TM \to TN$ by mapping $[x, v]_p$ to $[y, w]_{f(p)}$, where $w = D(y \circ f \circ x^{-1})(p)(v)$. This is easily shown to be independent of the coordinates chosen. Thus we obtain linear maps df_p from TM_p to $TN_{f(p)}$. We say that f has **rank** k at a point p if df_p has rank k. The next section will show the rank gives essentially all the local information about a differentiable map.

2.5 The Rank Theorems

The rank theorems provide the existence of coordinates on a manifold which simplify how the maps operate immensely. They are essentially an extension of the Euclidean inverse and implicit function theorems.

Theorem 2.8. *If f is rank k at a point p, there is a coordinate system x at p and y at $f(p)$ such that for $1 \leqslant i \leqslant k$,*

$$y^i \circ f \circ x^{-1}(a_1, \ldots, a_n) = a_i$$

or

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

where the asterixes denotes arbitrary functions.

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

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

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 $1 \leqslant i \leqslant k$,

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

and so we have found the required coordinate system. □

Corollary 2.9. *If f is rank k in a neighbourhood of a point p, then we may choose coordinate systems x and y such that*

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

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

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

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. $\qquad\square$

Corollary 2.10. *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, just the coordinates of x. $\qquad\square$

Corollary 2.11. *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, then it is rank n on a neighbourhood, since the set of full rank matrices is open. 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, and

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

and we have found the chart required. $\qquad\square$

2.6 Immersions, Submersions, and Covers

The extremely regular charts we constructed above are purely local, but we can use global topological properties to infer properties of an entire map. Call a map $f : M \to N$ an **immersion** if df_p is injective for all points p, and a **submersion** if df_p is always surjective. In terms of the rank discussion we have been having, if f is an immersion if the rank of f at $p \in M$ is the same as the dimension of M at p, and a submersion if the rank of f is the same as the rank of N at $f(p)$.

Theorem 2.12. *Let $f : M \to N$ be a map between smooth manifolds. Then*

    (a) If f is injective and has locally constant rank, then f is an immersion.

    (b) If f is surjective, has constant rank, M is second-countable, and N has the same dimension throughout, then f is a submersion.

Proof. If p is arbitrary, if the dimension of N is m at $f(p)$, and if f has rank $k < m$ at p, then there are coordinate systems x and y such that

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

which clearly isn't injective if $k < n$. This proves (a). To prove (b), consider a cover of f by charts (x_α, U_α), where $f(U_\alpha) \subset V_\alpha$ for some chart (y, V_α) on N. Since M is second-countable, it is also Lindelöf, and we may therefore assume the cover is countable. Let N have dimension m, and f has rank $k < n$. The coordinate systems can be chosen so that

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

This shows us that $(y \circ f \circ x_\alpha^{-1})(x(U_\alpha))$ is nowhere dense in $y(V_\alpha)$, and thus, $f(U_\alpha)$ is nowhere dense in N. It then follows by the Baire category theorem that $\bigcup f(U_\alpha) = f(M)$ is nowhere dense on N, which implies f is not surjective, a contradiction. $\qquad\square$

Variants of this proof can be adapted to various other cases, where f may not have constant rank, and N has non-constant dimension. We avoid discussing them in detail, because they are complicated to state, but obvious to adapt to any particular situation by working out the details there.

Theorem 2.13. *If $f : M^n \to N^m$ is a submersion, then f is an open map.*

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

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

This map is open, showing f is locally open, and thus open on its entire domain, since openness is a local property. $\square$

Corollary 2.14. *A submersion from a compact manifold to a connected manifold with the same dimensions is surjective.*

Proof. If M is compact, and $f : M \to N$ an immersion, then $f(M)$ is compact, hence closed, and since f is open, $f(M)$ is also an open subset of N. But then $f(M) = N$, since it is an open, closed, non-empty set. $\square$

Example. *Recall the Grassmanian space $G(k, n)$ of k dimensional subspaces of $\mathbf{R}^n$. There is another way to view the topology of $G(k, n)$ which is very useful for constructing continuous maps on the space. Consider the family $M(n, k; k)$ of full rank n by k matrices, which may be identified by separating columns with the family of linearly independant tuples $(v_1, \ldots, v_k)$ of vectors in $\mathbf{R}^n$. We obtain a surjective map $f : M(n, k; k) \to G(k, n)$ by mapping a matrix corresponding to the tuples $(v_1, \ldots, v_k)$ to $\mathrm{span}(v_1, \ldots, v_k)$. We claim that this map is a submersion, hence it is a continuous open map identifying the topological structure of $G(k, n)$ as a quotient space of $M(n, k; k)$, by identifying vector tuples which generate the same subspace. The utility of this is that continuous map g with domain $G(k, n)$ can be constructed as continuous maps with domain $M(n, k; k)$ which have the same value on vector tuples that generate the same subspace.*
     To obtain that the map is a submersion, we consider the open set

$$U = \left\{ \begin{pmatrix} A \\ B \end{pmatrix} : A \in GL_k(\mathbf{R}) \right\}$$

If we let $V = \mathrm{span}(e_1, \ldots, e_k)$, and $V' = \mathrm{span}(e_{k+1}, \ldots, e_n)$, then $f(U)$ is a subset of $A_{V'}$, and if we let $(y, A_{V'})$ be the standard coordinates corresponding to the set $A_{V'}$, identifying $L(V, V')$ with $M(k, n-k)$, then

$$(y \circ f) \begin{pmatrix} A \\ B \end{pmatrix} = BA^{-1}$$

which is the linear transformation mapping the i'th column of A to the i'th column of B. The map is now easily seen to be differentiable. For a fixed A,

the map $B \mapsto BA^{-1}$ is an invertible linear map, and therefore the map is a submersion everywhere, because we can always permute the coordinates so that a given matrix is in U (this corresponds to a diffeomorphism of $M(n,k;k)$), and when we take the span of this vectors, the way to get back to the original span is to unpermute the coordinates, so we can always assume our matrices have the nice form above. In other words, the group $GL_n(\mathbf{R})$ acts transitively on $M(n,k;k)$ by left multiplication, and transitively on $G(k,n)$ by the action $M \cdot V = \{Mv : v \in V\}$, and we find that f is a $GL_n(\mathbf{R})$ morphism, because $f(MN) = Mf(N)$.

2.7 Submanifolds

All this discussion of ranks and immersions goes to show that submanifolds defined by differentiable functions have a nice structure. If M is a manifold, and $N \subset M$ also has (under the relative topology) the structure of a differentiable manifold, then we say N is a **submanifold** of M if the embedding $i : N \to M$ is an immersion. If such a differentiable structure exists, it is unique, for if (x,U) is a chart in an atlas $\mathcal{A}$ on N, there is a chart $(\tilde{x},V)$ on M such that $(\tilde{x} \circ x^{-1})(a_1,\ldots,a_n) = (a_1,\ldots,a_n,0,\ldots,0)$, so that $(\tilde{x}^1,\ldots,\tilde{x}^n)$ is equal to x when restricted to N. But then if $\mathcal{B}$ is another atlas on $\mathcal{B}$, then any chart $(y,U) \in \mathcal{B}$ can be extended into a chart $\tilde{y}$ in M, and then $y \circ x^{-1} = \tilde{y} \circ \tilde{x}^{-1}$ is differentiable, so that $\mathcal{A} = \mathcal{B}$.

Example. *The biggest utility is when these manifolds lie in Euclidean space. Classically, a k-dimensional C^m manifold was a subset M of some Euclidean space $\mathbf{R}^n$, such that at ever $x \in M$, there is an open subset U of Euclidean space containing x, an open subset V of $\mathbf{R}^n$, and a C^m diffeomorphism $f : U \to V$ such that $f(U \cap M) = V \cap \mathbf{R}^k \times \{0\}$. All these diffeomorphisms, restricted to M, are C^m related to one another, because they are the restricted of diffeomorphisms on the whole of $\mathbf{R}^n$, and give a C^m differentiable structure onto M. Conversely, any submanifold of Euclidean space has this property because of the rank theorems, because if $i : M \to \mathbf{R}^n$ is an immersion, then we can extend any coordinate system on M to a coordinate system on an open subset of $\mathbf{R}^n$, such that M is flat in this coordinate system. The natural choice of a differentiable structure on any manifold lying in Euclidean space is this differentiable structure, and, as you can check, this is the differentiable structure we have chosen for all the manifolds we have used in past examples, which naturally can be embedded in Euclidean space.*

An important method of finding submanifolds of $\mathbf{R}^n$ is by specifying the submanifolds as a level set of Euclidean space. For instance, the sphere S^n can be defined as the level set $f^{-1}(1)$ of the map $f(x) = \|x\|$. Provided that ∇f does not vanish on the domain level set, the level set is always a manifold, and this fact can be generalized to maps between arbitrary manifolds.

Theorem 2.15. *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

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

for some charts (x, U) and (y, V), with y centred at q. This implies that if we adjust the last $n - k$ coordinates around p, the resulting point will still be in $f^{-1}(q)$, so that $x(U \cap f^{-1}(q)) = \{(a_1, \ldots, a_k)\} \times \mathbf{R}^{n-k}$. This means that $(x^{k+1}, \ldots, x^n)$, restricted to $f^{-1}(q)$, is a homeomorphism onto an open subset of $\mathbf{R}^{n-k}$. We can take all possible $\tilde{x} = (x^{k+1}, \ldots, x^n)$, over all possible points $p \in f^{-1}(q)$, as an atlas, because if $\tilde{y} = (y^{k+1}, \ldots, y^n)$ is also formed by this process, then $\tilde{y} \circ \tilde{x}^{-1}$ is equal to the restriction of $y \circ x^{-1}$ to $\{(a_1, \ldots, a_k)\} \times \mathbf{R}^{n-k}$, hence differentiable. Thus $f^{-1}(q)$ is a closed $n - k$ submanifold of M. $\qquad\square$

Example. *If M is a 1-submanifold of $\mathbf{R}^2$, specified as the level set of a function $f : \mathbf{R}^2 \to \mathbf{R}$,*

$$M = \{(x, y) \in \mathbf{R}^2 : f(x, y) = 0\}$$

then the surface of revolution about the z-axis related to M can be identified as the space

$$N = \left\{ (x, y, z) \in \mathbf{R}^3 : f\left(\sqrt{x^2 + y^2}, z \right) = 0 \right\}$$

It is a surface, because the differential of the defining level set mapping has rank 1 at any point in a neighbourhood of N (because we can use the chain rule, and the fact that the map $(x, y, z) \mapsto (\sqrt{x^2 + y^2}, z)$ has full rank at every point).

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

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

where $M(n,n)$ is identified with $(\mathbf{R}^n)^n$. Then for any $(v_1,\dots,v_n) \in GL_n(\mathbf{R})$,

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

So det *has rank 1 at every point (full rank), and $SL(n)$ is a (closed) submanifold of $GL(n)$ dimension $n^2 - 1$.*

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

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

This implies that M lies on $(S^{n-1})^n \subset \mathbf{R}^{n^2}$, so $O(n)$ is closed and bounded, and therefore compact. Consider the diffeomorphism $R_A : B \mapsto BA$ of $GL(n)$, for a fixed $A \in GL(n)$. We also have $\psi \circ R_A = \psi$, for $A \in O(n)$. 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

$$\left.\frac{\partial \psi^{ij}}{\partial M^{lk}}\right|_M = \begin{cases} 2M_{ik} & i = j = l \\ M_{ik} & j = l \\ M_{jk} & i = l \\ 0 & \text{elsewise} \end{cases}$$

In particular, at the identity,

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

It follows that $\psi_*(TGL(n)_I)$ is the space of all symmetric matrices in $TGL(n)_I$, which has dimension $n(n+1)/2$. Thus ψ has constant rank $n(n+1)/2$, and we find the space of orthogonal matrices has dimension

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

as a differentiable manifold.

Example. *Every orthogonal matrix has determinant ± 1. The special orthogonal group $SO(n)$ is the set of orthogonal matrices of determinant one, and is an open submanifold of $O(n)$. It's actually a connected component. To see this, consider the obvious (Lie-group) action of $SO(n)$ on S^{n-1}. For $n \geqslant 2$, it is easy to see that the action is transitive: if $v_1 \in S^{n-1}$ is given, we can extend v_1 to an orthonormal basis $(v_1, \ldots, v_n)$; possibly permuting this basis, the basis can be assumed oriented, and then the orthogonal matrix mapping e_i to v_i is in $SO(n)$. For a fixed $p \in S^{n-1}$, the action $f(M) = Mp$ is surjective. If N is fixed, $g(M) = NM$, and $h(p) = Np$, then we have a commutative diagram The vertical maps are diffeomorphisms, hence, taking differentials of all maps in the diagram, we conclude that $h_* \circ f_* = f_* \circ g_*$, and since h_* and g_* are rank preserving, we conclude that the rank of f_* at any M is the same as it's rank at NM (this argument works for any 'differentiable group', and any 'differentiable action', a notion we will define precisely later when we introduce Lie groups). Because $SO(n)$ is second countable, we conclude that f is a submersion, hence it is open. In the one dimensional case, the map isn't surjective, but in this case $SO(1)$ and S^0 are discrete spaces, so the map $M \mapsto Mp$ is trivially open. In any dimension $\geqslant 2$, the stabilizer of the unit vector e_n is the subgroup of matrices of the form*

$$\begin{pmatrix} X & 0 \\ 0 & 1 \end{pmatrix}$$

where $X \in SO(n-1)$, so the stabilizer is actually diffeomorphic to $SO(n-1)$. Using the orbit stabilizer theorem, we find that for $n \geqslant 2$, S^{n-1} is homeomorphic to $SO(n)/SO(n-1)$. Using induction, we may assume that $SO(n-1)$ is connected. But for $n \geqslant 2$, S^{n-1} is connected. The theorem then follows from the general fact that if H is a closed subgroup of G, and H and G/H are connected, then G is connected.

Remark. In the lower dimensional matrix groups there are standard coordinate systems which are often employed. On $SO(2) = \mathbf{T}$, the standard coordinate system is obtained by taking angles. In $SO(3)$ It is easy to prove

using the spectral theorem that any rotation R is given by an oriented rotation about the plane perpendicular to a vector x with $Rx = x$. Thus a rotation is given by picking a line $l \in \mathbf{RP}^2$ and an angle $\theta \in [0, 2\pi]$. More concretely, we can use Euler angles. It turns out that in three dimensions, every rotation can be given first by a rotation about the z axis, known as *yaw*, *pitch*, which is a further rotation about the y axis, and *roll*, which is rotation about the x axis. That is, we can describe almost every rotation by three angles ϕ, θ, and ψ, which describes the rotation

$$\begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The singularities of this coordinate system occur at rotations which cause gimbal locks.

Similar results about defining manifolds can be obtained for manifolds with boundary.

Theorem 2.16. *If $f \in C^\infty(M)$, where M is a manifold with boundary, and $df_p \neq 0$ for any p with $f(p) = a$ or $f(p) = b$, where $a, b \in \mathbf{R}$, then $f^{-1}([a,b])$ is a submanifold of M with boundary with the same dimension as the original space.*

Proof. Clearly $f^{-1}((a,b))$ is a open subset of M, so it is a smooth manifold with boundary. The boundary of this set is contained within $f^{-1}[a,b]$, and it actually equal to this, since $df_p \neq 0$ for any p with $f(p) = a$ or $f(p) = b$ (because this implies the function is changing near this point). The result then follows from the fact that if M is a subset of a manifold N, and ∂M is an imbedded submanifold with boundary, then $M \cup \partial M$ is a manifold with boundary. $\square$

Like the smooth immersions discussed here, a **topological immersion**, which is a continuous map $f : X \to Y$ between topological spaces which is locally an embedding. Certainly a smooth immersion is a topological immersion, but a smooth topological immersion may not be a smooth immersion (indeed, consider the map $t \mapsto t^3$). Injective smooth immersions are mostly well behaved, apart from the odd inconsistency of dropping differentiable maps to subdomains. Let $g : M \to N$ be a differentiable function with $g(M)$ contained in the image $f(N')$ of an immersed submanifold N'.

If f is globally injective, we may consider the function $g : M \to N'$. A suitable question to ask is whether this function is differentiable. In most cases, the answer is yes.

Example. *Immerse $(-\pi, \pi)$ in $\mathbf{R}^2$ via the lemniscate map*

$$f(t) = (\sin 2t, \sin t)$$

The map $g : (-\pi, \pi) \to (-\pi, \pi)$ defined by

$$g(x) = \begin{cases} \pi + x & : x < 0 \\ 0 & : x = 0 \\ x - \pi & : x > 0 \end{cases}$$

is not even continuous, yet $f \circ g$ is differentiable.

Continuity is all that can go wrong in this situation.

Theorem 2.17. *Let $f : M^n \to N^m$ be an injective immersion of M in N, and suppose $g : N' \to N$ is differentiable, and $g(N') \subset N$. 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 $p \in N'$. 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. $\qquad\square$

Thus we see that, for submanifolds whose topological structure is induced by the relative topology on the subspace, we can always descend to differentiable maps when needed. These submanifolds are known as **imbedded manifolds** (an immersion that is also an embedding).

Theorem 2.18. *The following are sufficient to conclude that an injective immersion $f : M \to N$ is an imbedding.*

 (a) f is open or closed.

 (b) f is proper (the inverse image of compact sets are compact).

(c) M is compact.

(d) The dimension of M is equal to the dimension of N.

Proof. An injective open or closed continuous map is a topological embedding, so (a) is fairly trivial. If M is compact or f is proper, f is closed, and if the dimension of M is equal to the dimension of N, f is open (by looking at the coordinate charts), so (b), (c) and (d) follow from (a). □

The idea of a submersive map $f : M \to N$ is very closely related to its family of smooth sections $s : N \to M$, maps such that $f \circ s = \mathrm{id}_N$. More precisely, a submersive map is closely related to its local smooth sections over neighbourhoods of N.

Theorem 2.19. *A smooth map $f : M \to N$ is a submersion if and only if every $p \in M$ is in the image of some local section.*

Proof. If f is a submersion, then there are local coordinates x and y around every point where $(y \circ f \circ x^{-1})(a^1,\ldots,a^n) = (a^1,\ldots,a^m)$. Given any point $p_0 \in M$, consider a point $q_0 \in N$ with $y^1(q_0) = x^1(p_0),\ldots,y^m(q_0) = x^m(p_0)$, and define the map $s(q) = x^{-1}(y(q),x^{m+1}(p),\ldots,x^n(p))$, then s is the required section. Conversely, if $s : U \to M$ is a local section, then $\mathrm{id} = f_* \circ s_*$, and therefore the rank of f_* at p must be at least the rank of id, so it has full rank. □

2.8 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 E has measure zero if, for any ε, there is a cover of E by open sets $U_1, U_2, \ldots$ with

$$\sum |U_k| < \varepsilon$$

Suppose that μ is a measure on a differentiable manifold M behaving locally like the Lebesgue measure, in the sense that for any chart (x, U), the pushforward measure $x_*(\mu)$ is absolutely continuous with respect to the Lebesgue measure. If $\mu(E) = 0$, then $x_*(\mu)(x(E)) = 0$ for every chart x, and conversely, assuming M is second countable, if this is true for every chart, then $\mu(E) = 0$. Thus it makes sense, without any measure on the manifold, to say a set E has **measure zero** if it can be covered by countably many charts, where the set has measure zero in each chart.

Lemma 2.20. *If $f : M^n \to N^n$ is C^1, and $E \subset M$ has measure zero, then $f(E)$ has measure zero.*

Proof. Begin with the case $M = N = \mathbf{R}^n$. Let E be a compact rectangle in $\mathbf{R}^n$, and suppose that $\|\partial_i f^j\|_{L^\infty(E)} \leqslant K$ for all i and j. Then for $x, y \in E$, by tracing the path from x to y along the coordinate lines,

$$|f(x) - f(y)| \leqslant \sum_{i=1}^n |f^i(x) - f^i(y)| \leqslant K \sum_{i=1}^n \sum_{j=i}^n |x^j - y^j| \leqslant Kn^2|x - y|$$

If E has measure zero, then, by the σ compactness of $\mathbf{R}^n$, we may assume E is compact, so that E is contained in some rectangle, and suppose on this rectangle that the partial derivative bound holds. If E is covered by rectangles $U_1, U_2 \ldots$ with $\sum |U_i| \leqslant \varepsilon$, then each U_i is mapped into a ball of radius $Kn^2 \mathrm{diam}(U_i)$, whose volume agrees with $|U_i|$ up to a constant C depending on K and n, and so $f(E)$ is covered by balls with total measure $\leqslant C\varepsilon$, and we may let $\varepsilon \to 0$. Now in general on any manifold, let E be covered by charts $(x_1, U_1), (x_2, U_2), \ldots$ with $x_k(E \cap U_k)$ a set of measure zero, in such a way that there are charts $(y_1, V_1), (y_2, V_2), \ldots$ with $U_k \subset f^{-1}(V_k)$ (we may take an initial covering of E by charts, and then duplicate the charts, suitably small to be contained in the V_k). Then each $(y_k \circ f \circ x^{-1})(x(E \cap U_k)) = y_k(f(E \cap U_k))$ has measure zero, and so $f(E) = \bigcup f(E \cap U_k)$ has measure zero. $\square$

If the rank of a map $f : \mathbf{R}^n \to \mathbf{R}^m$ has rank $< m$, then intuitively, it maps small neighbourhoods of points to sets close to hyperplanes, which have measure zero. We will show this idea shows the image of the set of low rank points always has measure zero on a manifold. Call a point p on M **criticial** for a map $f : M \to N$ if f has rank less than the dimension of p at p, and $f(p)$ is known as a critical value.

Theorem 2.21 (Sard). *If $f : M^n \to N^n$ is C^1, and M is second countable, then the set of critical values is a set of measure zero in N.*

Proof. We begin with the case $N = \mathbf{R}^n$, and M an open subset of $\mathbf{R}^n$. Let C be a closed cube in M with side lengths l, and let B denote the critical values in C. Then f is uniformly continuously differentiable on C, and so for any ε, for large enough N, we can divide C into N^n cubes S_k of side length l/n, such that for $x, y \in S_k$,

$$|f(y) - f(x) - Df(x)(y - x)| < \varepsilon|y - x| \leqslant \varepsilon \frac{l\sqrt{n}}{N}$$

43

and we also know that f is Lipschitz on C, so there is M such that

$$|f(y) - f(x)| < M|y - x| < \frac{lM\sqrt{n}}{N}$$

If B intersects S_k, then there exists $x \in S_k$ such that $Df(x)(y - x)$ lies in an $n-1$ dimensional subspace of $\mathbf{R}^n$, and so the two inequalities about imply that f maps S_k into a 'n dimensional cylinder', formed by taking a ball of radius $lM\sqrt{n}/N$ in $n-1$ dimensional space as the base of the cylinder, and a height of length $2\varepsilon l\sqrt{n}/N$. This shape has volume equal, up to a constant depending on n, by

$$\left(\frac{2\varepsilon l\sqrt{n}}{N}\right)\left(\frac{lM\sqrt{n}}{N}\right)^{n-1} = 2\varepsilon\frac{l^n n^{n/2}M^{n-1}}{N^n}$$

But we have N^n cubes, which implies the set $f(B \cap C)$ can have volume at most

$$2\varepsilon l^n n^{n/2}M^{n-1}$$

and we may then let $\varepsilon \to 0$ to conclude $f(B \cap C)$ is a set of measure zero. But since $\mathbf{R}^n$ is the union of countably many cubes C, we conclude that B itself has measure zero. The general case for second countable manifolds follows from taking a countable set of coordinate charts covering M, and then applying Sard's theorem we have just proved in the Euclidean case. $\square$

Sard's theorem has a more general form which is very useful in differential topology, but we won't use it here.

Theorem 2.22 (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 is a set of measure zero in N.*

Our case occurs when $n = m$. It is easy to prove the case where $m > n$. The tricky case is where $m < n$.

Theorem 2.23. *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. $\square$

44

2.9 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 $h(x_1, \ldots, x_n) = g(x_1)g(x_2) \ldots g(x_n)$ which is positive on $(-1, 1)^n$, and zero elsewhere.

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

Theorem 2.24. *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. $\qquad \square$

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. 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 2.25 (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. $\qquad \square$

Theorem 2.26. *Any cover on a paracompact manifold induces a subordinate partition of unity.*

Proof. 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. $\qquad\square$

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 2.27. *In a σ-compact manifold M, there exists a smooth, real-valued function f 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 $x \notin B_1, \ldots, B_n$, 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 $\overline{B_1} \cup \cdots \cup \overline{B_n}$, and is therefore compact. $\qquad\square$

Other existence proofs also follow naturally.

Lemma 2.28. *If A is a closed subset of a paracompact manifold M, there is a differentiable function $f : M \to [0,1]$ with $f^{-1}(0) = A$.*

Proof. Let us begin proving this for $M = \mathbf{R}^n$. Let $\{U_i\}$ be a countable cover of $\mathbf{R}^n - A$ by open unit balls. For each U_i, pick a smooth $f_i : \mathbf{R}^n \to [0,1]$ positive on U_i, and equal to zero on $\mathbf{R}^n - U_i$. Define

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

Each α_j is well defined, because 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 if $k \geqslant n$,

$$\frac{1}{\alpha_k 2^k} \frac{\partial^n f_k}{\partial x_{i_1} \ldots \partial x_{i_n}}$$

so the sums of all partial derivatives converge uniformly, and $f^{-1}(0) = C$ because the function is positive for some coefficient everywhere else.

To address the general case, let M be paracompact, and find a cover of coordinate balls $\{B_\alpha\}$ for $M - A$. Then we may find a partition of unity $\{\psi_\alpha\}$ for this family, and find smooth $f_\alpha : M \to \mathbf{R}$ with $f_\alpha^{-1}(0) = A \cap B_\alpha$. But then $f = \sum \psi_\alpha f_\alpha$ is smooth (by local finiteness), and satisfies $f^{-1}(0) = A$. $\qquad\square$

Corollary 2.29. *If A and B are closed on a paracompact manifold M, then there is a function $h : M \to [0,1]$ with $h^{-1}(0) = A$, $h^{-1}(1) = B$.*

Proof. Modifying the function obtained in the last theorem, we can find $f : M \to [0,1]$ with $f^{-1}(1) = B$. Denote $f^{-1}(0)$ by C. If $g : M \to [0,1/2]$ satisfies $g^{-1}(0) = A$, and if $\psi : M \to [0,1]$ is a bump function on B, vanishing on $A \cup C$, then

$$h = \psi f + (1 - \psi) g$$

satisfies $h^{-1}(1) = B$, because for $p \notin B$, $f(p), g(p) < 1$. Now certainly $h(p) = 0$ for $p \in C$. If $p \in A - C$, $\psi(p) = 0$, so $h(p) = g(p) > 0$, and if $p \notin A \cup C$, $f(p), g(p) > 0$, so $h(p) > 0$. Thus $h^{-1}(0) = A$, and we have shown h is the needed function in the theorem. $\qquad\square$

2.10   Differentiable Manifolds with Boundary

Recall that we may extend differentiability on subsets of Euclidean space in the following way. In general, a map $f : A \to B$ between arbitrary subsets of Euclidean space is differentiable if f can be extended to a differentiable map on an open subset of Euclidean space containing A.

Theorem 2.30. *If $f : \mathbf{H}^n \to \mathbf{R}$ is locally differentiable (every point has a neighbourhood on which f can be locally extended to be differentiable), then f is differentiable in the sense defined above.*

Proof. Let $\{U_\alpha\}$ be a open cover of $\mathbf{H}^n$ in $\mathbf{R}^n$ with smooth functions $g_\alpha : U \to \mathbf{R}$ agreeing with f where the two are jointly defined. Consider a partition of unity $\{\Phi_\alpha\}$ subordinate to $\{U_\alpha\}$. Define $g = \sum g_k \Phi_k$, defined on $\bigcup U_\alpha$, a open extension of $\mathbf{H}^n$. Each pair g_k and Φ_k are differentiable, so the multiplication of the two is differentiable. Since these functions are locally finite, we also have g differentiable across its domain. If $p \in \mathbf{H}^n$, then $g_k(p) = f(p)$. Thus

$$g(p) = \sum g_k(p)\Phi_k(p) = \sum f(p)\Phi_k(p) = f(p)$$

since the Φ_k sum up to one. Thus $g|_{\mathbf{H}^n} = f$, and we have extended f to be differentiable. $\qquad\square$

This allows us to define a notion of differentiable structure for a manifold with boundary. We can extend the notion of two charts being C^∞ consistent, because we have extended differentiability on non open subsets of $\mathbf{H}^n$ to a non-local criterion. Similarily, a map $f : M \to N$ can also be considered differentiable if $y \circ f \circ x^{-1} : x(U) \to y(V)$ can be extended to be a differentiable function on an open subset of Euclidean space. Thus manifolds with boundary have effectively the same theory as manifolds without boundary.

Example (Differentiable structures on the boundary of a manifold)**.** *Given a differentiable manifold with boundary M, we can assign a unique differentiable structure to ∂M such that the inclusion map is an imbedding. We can generate it from the atlas consisting of the restriction of charts on M to the boundary, projected into an $(n-1)$ dimensional subspace of $\mathbf{R}^n$. The transition maps are easily verified to be C^∞.*

One issue with manifolds with boundary is that the rank theorem does not hold. The problem is that, at the boundary, we are restricted in how we reapply coordinates – we can only consider smooth automorphisms of $\mathbf{H}^n$ rather than $\mathbf{R}^n$. However, for immersions f of manifolds with boundary in manifolds without, we do have coordinate charts for which

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

because in the theorem we currently have, we need not change the coordinates on the manifold with boundary, only on the manifold without. We obtain similar results for submersions of manifolds without boundary in manifolds with boundary.

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 reconciling with the use of infinitisimals, which are formalized in the tangent bundle structure. One can describe modern differential geometry as studying manifolds with additional structure ascribed to the tangent bundle, and so a further analysis of the tangent bundle we constructed in the last chapter is essential.

3.1   A Smooth Structure on the Tangent Bundle

Recall that if M is a manifold, then the tangent bundle TM is defined as equivalence classes of vectors in various charts on the manifold, where $[x, v]_p$ is identified with $[y, w]_p$ if

$$w = D(y \circ x^{-1})(x(p))(v)$$

We view these vectors as eminating out from the manifold at every point. We then have a projection map $\pi : TM \to M$ which projects a vector to the point it is emitted from. On each of the 'tangent spaces' $\pi^{-1}(p)$, often denoted M_p, we have a vector space structure defined by $\lambda[x, v]_p + \gamma[x, w]_p = [x, \lambda v + \gamma w]_p$ (the linearity of derivatives makes this well defined). One

of the nicest parts about hte tangent bundle is that we can express the collection of all derivatives of a differentiable map $f : M \to N$ in all coordinate systems as a single mathematical object, by setting $f_* : TM \to TN$ by setting, if p is contained in a chart x and $f(p)$ is contained in a chart y,

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

for a fixed p, the map f_* is linear, since it is effectively the original derivative at p in disguise. Another nice feature, which we haven't already discussed, is that if $g : N \to L$ is another differentiable map, then we can consider $g_* \circ f_*$. The chain rule can then be expressed as the elegant formula $(g \circ f)_* = g_* \circ f_*$.

The Euclidean space $\mathbf{R}^n$ will provide the prototypical example of a tangent space. Here we can identify $T\mathbf{R}^n$ with $\mathbf{R}^n \times \mathbf{R}^n$, where we identify (p, v) with $[\mathrm{id}, v]_p$. Then $T\mathbf{R}^n$ possesses both topological and smooth structure, which agrees with our intuitions about how tangent vectors change over time. If $f : \mathbf{R}^n \to \mathbf{R}^m$ is smooth, then $f_* : T\mathbf{R}^n \to T\mathbf{R}^m$ is also a smooth map between the tangent spaces, now viewed as smooth manifolds. Given any manifold M, we would like to give TM a smooth structure which enables us to talk about continuous and smooth choices of vectors over space.

The most geometric way to get this structure to the tangent bundle is if we assume our manifold M occurs as a smooth submanifold of $\mathbf{R}^N$, for some N. Then the inclusion map $i : M \to \mathbf{R}^N$ is an imbedding, and so we have map $i_* : TM \to T\mathbf{R}^N$, which is injective. This identifies the vectors in the tangent bundle of M with subsets of $\mathbf{R}^N \times \mathbf{R}^N$, and we can use this to give TM a smooth structure, such that i_* is an imbedding. It is only slightly technical to show this smooth structure does not depend on the particular embedding of M in Euclidean space, so, assuming our manifold already occurs in Euclidean space, the problems of giving the tangent bundle smooth structure is solved.

This construction is probably the best way to view the smooth structure of hte tangent space geometrically, since Euclidean space is visualizable, at least in low dimensions. Since all manifolds can be imbedded in some dimension of Euclidean space, this is certainly a sufficient method to define the tangent space on all manifolds. Nonetheless, it is not entirely elegant because the imbedding is not unique, so there are many different candidates for the bundle (though all the candidates will essentially be equivalent), and it is not clear which imbedding will make the tangent bundle easiest to work with.

An alternative way to obtain the smooth structure is to use the coordinate charts. Given any chart (x, U) on the manifold, the map $x : U \to x(U)$ is a smooth map, and as such we can consider $x_* : TU \to Tx(U)$. We claim that the x_* can be chosen as a consistant family of charts on TM. Recall the way we construct abstract smooth manifolds from chapter 2. If (y, V) is another coordinate chart on the manifold, then $x_*(TV \cap TU) = x(U \cap V) \times \mathbf{R}^n$ is an open subset of Euclidean space, and by definition,

$$(y \circ x^{-1})_*(v_p) = (D(y \circ x^{-1})(p)(v)))_{(y \circ x^{-1})(p)}$$

Since $y \circ x^{-1}$ is C^∞, not only does the map change over points smoothly, but the choice of $D(y \circ x^{-1})$ changes smoothly as p ranges, which means the entire map, viewed from $\mathbf{R}^n \times \mathbf{R}^n$, to itself, is smooth. Thus these map combine to give a smooth structure on TM, and we have abstractly define what the smooth structure on the tangent bundle is. The coordinates x_* into $\mathbf{R}^{2n}$ on TM are often denoted by $(x, \dot{x})$, in the sense that $\dot{x}^i([x, v]_p) = v^i$.

We're not finished with our tangent bundle discussion yet. There are many different perspectives with which we can view the tangent bundle of a manifold, and all are useful at one point or another. The unity in description is best achieved with the concept of a vector bundle.

3.2   Vector Bundles

The abstract definition of a vector space eminating from points on a topological space comes from the theory of vector bundles. A **vector bundle** ξ over a topological space X is a topological space E together with a continuous, surjective projection $\pi : E \to X$ such that for each $x \in X$, the set $E_x = \pi^{-1}(\{x\})$ has a vector space structure, and such that we have local triviality; for each point $x \in X$, there is a neighbourhood U and a homeomorphism $\phi : \pi^{-1}(U) \to U \times \mathbf{R}^n$ which is a linear isomorphism on each fibre X_x, for $x \in U$. If X is connected, the n chosen here is unique, and ξ is known as an n-dimensional bundle. If X and E are both differentiable manifolds, and the projection maps π and triviality maps ϕ are differentiable, then we shall call this kind of bundle a **smooth vector bundle**.

Example. *For any topological space X, we have trivial n dimensional vector bundles $\varepsilon^n(X)$, where the extension space is $X \times \mathbf{R}^n$, and the projection map is the obvious one.*

Example. *The Möbius strip forms the extension space of a vector bundle over the circle. Indeed, the circle can be topologically identified with $\mathbf{T} = \mathbf{R}/\mathbf{Z}$, and the Möbius strip can also be defined as the quotient of the extension space $\mathbf{R}^2$ of $\varepsilon^1(\mathbf{R})$ where (x, v) is identified with $(x + n, (-1)^n v)$.*

Example. *If $\xi = \pi : E \to X$ and $\psi = v : F \to Y$ are two bundles, then we can define a bundle $\xi \times \psi$ in terms of the product of the projections $\pi \times v : E \times F \to X \times Y$.*

If X and Y are spaces with bundles $\xi = \pi : E \to X$ and $\psi = v : F \to Y$, then a **bundle map** is a pair of continuous maps $f : E \to F$ and $f_\sharp : E \to F$, where each $f_p^\sharp = (f^\sharp)|_{X_p}$ is linear from X_p to $Y_{f(p)}$, and the diagram below commutes An isomorphism in the category of bundle maps is known as an **equivalence** of bundles. If X and Y are the same space, it is often assumed that f *must* be the identity map, so $f_\sharp$ maps $\pi^{-1}(p)$ to $v^{-1}(p)$ linearly for each point p in the base space. Since f can be easily obtained from $f^\sharp$, we can describe a bundle map solely by the map $f^\#$ between bundles, or more particularly by the linear maps $f_p^\# : X_p \to Y_p$, which can be viewed as a continuous parameterization of linear maps over p. Note that now we have defined an equivalence of bundles, we can work backwards and *define* the local triviality conditions of bundles by saying that every vector bundle is locally equivalent to the trivial bundle over suitably small neighbourhoods.

Theorem 3.1. *Let ξ and ψ be bundles over the same base space X, then a bundle map $f : \xi \to \psi$ which is an isomorphism on each fibre is a bundle equivalence.*

Proof. Let $\xi = \pi : E \to X$ and $\psi = v : F \to X$. The map $f : E \to F$ is clearly injective, so all that remains is to check the inverse is also continuous. But locally at any point $p \in X$, we can consider a neighbourhood U for which there are trivializations $\phi : \pi^{-1}(U) \to U \times \mathbf{R}^n$ and $\psi : v^{-1}(U) \to U \times \mathbf{R}^n$, and then $g = \psi \circ f \circ \phi^{-1}$ is a bundle map from $U \times \mathbf{R}^n$ to itself which is an isomorphism on each fibre. Since being a homeomorphism is a local condition, it now suffices to prove that a bundle map from a trivial bundle to itself which is an isomorphism on each fibre is an equivalence. Note that the map g induces a continuous choice of linear isomorphisms $M : U \to GL(n)$ such that $g(p, v) = (p, M(p)v)$. Clearly such $M(p)$ must exist, and be unique, and if we let $h(p, v) = v$ be the projection of a point in the

bundle to it's vector part, then $M_{ij}(p) = e_j \cdot h(g(p,v))$ is continuous. But then $g^{-1}(p,v) = (p, M^{-1}(p)v)$, and inversion is a continuous operation on $GL(n)$. Of course, since inversion of matrices is also smooth, it follows that a smooth bundle map which is an isomorphism on each fibre is a smooth bundle equivalence. $\qquad\square$

3.3   The Space of Derivations

The algebraists found another characterization of the tangent bundle, which, though more elegant, is much more abstract then the definition by coordinates. Note that on $\mathbf{R}^n$, the space of vectors $v \in T\mathbf{R}_p^n$ can be identified with the space of directional derivatives, functionals D_v on $C^\infty(M)$ defined by

$$D_v(f)(p) = \lim_{h \to 0} \frac{f(p + hv) - f(p)}{h}$$

This map satisfies the product rule

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

The idea of the algebraic tangent bundle is to identify tangent vectors with certain functionals on $C^\infty(M)$. This is compatible with the interpretation of tangent vectors as velocities over the manifold. If we are moving across a surface, then the velocity we are travelling at determines how the surface below us changes, and the measurement of this change can be identified in the velocity we are travelling at.

Since $C^\infty(M)$ is a vector space, we may consider the dual space $C^\infty(M)^*$, which is a monstrous vector space consisting of all linear functionals from $C^\infty(M)$ to $\mathbf{R}$. A derivation at a point $p \in M$ is a linear functional $l \in C^\infty(M)^*$ satisfying $l(fg) = f(p)l(g) + g(p)l(f)$. A C^∞ map $f : M \to N$ induces a linear map $f^* : C^\infty(N) \to C^\infty(M)$, defined by $f^*(g) = g \circ f$, which further induces a map $f_* : C^\infty(M)^* \to C^\infty(N)^*$, 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 $f(p)$. We can directly calculate that $(f_* \circ g_*) = (f \circ g)_*$, so that if f is a diffeomorphism, f_* is an isomorphism. We shall soon find that we can identify the space of derivations at a point p as the space of tangent vectors at p. The identification will match the f_* defined here with the f_* defined on the tangent space, providing an elegant algebraic definition of the covariant derivative.

The differential operators

$$\frac{\partial}{\partial x^1}\bigg|_p, \ldots, \frac{\partial}{\partial x^n}\bigg|_p$$

are all derivations. We will show, in fact, that these operators span the space of all derivations, so that the space is n dimensional, and can be identified with the tangent bundle under the map

$$\sum v^i \frac{\partial}{\partial x^i}\bigg|_p \mapsto [x, v]_p$$

This means the set of all derivations *is* the tangent bundle, up to a change in notation. In other words, a 'tangent' on the manifold can be thought of a way of measuring change, by moving along the manifold at a certain velocity.

Lemma 3.2. *If $f \in C^\infty(M)$ is constant, $l(f) = 0$ for any derivation l.*

Proof. By scaling, assume without loss of generality that $f = 1$ for all p. Then
$$l(f) = l(f^2) = l(f) + l(f) = 2l(f)$$
We then just subtract $l(f)$ from both sides of the equation. $\square$

Lemma 3.3. *If l is a derivation, 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. $\square$

We will show this space, though lying in a very high dimensional vector space, is actually very low dimensional.

Lemma 3.4. *If $f \in C^\infty(\mathbf{R}^n)$ and $f(0) = 0$, then there exists functions $f_i \in C^\infty(\mathbf{R}^n)$, such that $f(x) = \sum f_i(x)x^i$, and $f_i(0) = \partial_i f(0)$.*

Proof. Define $g(t) = f(tx)$. Then $g'(t) = \sum x^i \partial_i f(tx)$

$$f(x) = \int_0^1 g'(t) \, dt = \sum x^i \int_0^1 \partial_i f(tx) = \sum x^i f_i(x)$$

and each f_i is verified to be C^∞, with $f_i(0) = \partial_i f(0)$. $\qquad\square$

Theorem 3.5. *The space of derivations at the origin on $\mathbf{R}^n$ is n dimensional, with basis*

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

Proof. Let $f \in C^\infty(\mathbf{R}^n)$. Then $l(f) = l(f - f_0)$, and so we can write $f - f_0 = \sum x^i f_i$, and then

$$l(f - f_0) = \sum l(x^i)\frac{\partial f}{\partial x^i}(0)$$

Thus

$$l = \sum l(x^i)\frac{\partial}{\partial x^i}$$

The independence of each partial derivative derivation is left to the reader.
$\qquad\square$

We can extend this theorem to arbitrary smooth manifolds.

Lemma 3.6. *If l is a derivation at p, 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

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

hence $l(f) = 0$. $\qquad\square$

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. This implies that we can identify the space of derivations at $p \in U$ in $C^\infty(U)$ with $C^\infty(M)$, and we find that derivations really act on the **germ** of functions defined in a neighbourhood of p. When we are working with analytic or complex manifolds, we no longer have bump functions to work with, so we must begin by working on the space of derivations defined on germs of analytic or holomorphic functions from the very beginning of the theory.

Theorem 3.7. *The space of derivations at p is n-dimensional, and if (x, U) is a chart centered at p, then a basis for the space are the partial derivatives*

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

Proof. By restriction, we can identify derivations at p on M with derivations at p on U, which can be identified by diffeomorphism with the derivations at the origin on $x(U) \subset \mathbf{R}^n$, and therefore with the space of derivations at the origin in $\mathbf{R}^n$. This identification maps the partial derivatives at p with respect to the coordinates x onto the partial derivatives at $x(p)$, and this is all that is needed to verify the proof. $\qquad\square$

If we collect all derivations at all points on a manifold together, we obtain a structure corresponding exactly 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, and alternate between the two notations depending on which is convenient. For instance, we will often speak of a tangent vector operating on functions, or of a derivation as an element of the tangent space. Ultimately, they are the same mathematical objects, viewed from two different lenses.

Remark. We can enlarge the space $C^\infty(M)$ of real-valued smooth functions to the complex vector space of complex-valued smooth functions on M. The linear functionals l in the complex dual space $C^\infty(M)^*$ such that $l(f) \in \mathbf{R}$ is f is a real-valued smooth function can then be identified with the original dual space. In the same way, the functionals l such that $l(f)$ is

purely imaginary when f is real-valued and smooth can also be identified with the original dual space. If l is any complex linear functional, and we define

$$l_1(u + iv) = \text{Re}(l(u)) + i\text{Re}(l(u))$$

$$l_2(u + iv) = \text{Im}(l(u)) + i\text{Re}(l(u))$$

Then l_1 and l_2 are complex linear, $l = l_1 + il_2$, and l_1 and l_2 are the unique such functions such that $l_1(f)$ is real and $l_2(f)$ is purely imaginary when f is real-valued. If d is a derivation on the space of real-valued functions, and if we define $d(u + iv) = du + idv$, then d is also a derivation on complex-valued functions by a trivial calculation. On the other hand, if d is a complex-valued derivation, then

$$d_1(uv) = \text{Re}(u(p)dv + v(p)du) = u(p)d_1v + v(p)d_1u$$

so d_1, and, left as an exercise, d_2, are easily seen to be a derivations. Thus every complex-valued derivation can be split into two real-valued derivations, and so it follows that every complex-valued derivations at p can be written as

$$\sum a_i \left.\frac{\partial}{\partial x_i}\right|_p$$

where $a_i \in \mathbf{C}$. This is sometimes more natural to work with than the original tangent space over the real numbers, and is a kind of 'complexification' of the original tangent space.

If V is the subspace of $C^\infty(M)$ consisting of all functions f with $f(p) = 0$, and W is the subspace of V generated by all products of functions in V, then the lemma above shows that any derivation on $C^\infty(M)$ vanishes on W. Conversely, if there is a functional $l : V \to \mathbf{R}$ which vanishes on W, then it can be extended to a unique derivation on $C^\infty(M)$, by defining, for arbitrary $f \in C^\infty(M)$,

$$l'(f) = l(f - f(p))$$

an equation which must hold for any derivation which extends l. It then follows that l' is a linear operator, and l' annihilates constant functions, from which it then follows that

$$\begin{aligned}
l'(fg) &= l(fg - f(p)g(p)) \\
&= l([f - f(p)][g - g(p)]) + f(p)l'(g) + g(p)l'(f) - 2f(p)g(p)l(1) \\
&= f(p)l'(g) + g(p)l'(f)
\end{aligned}$$

so derivations on $C^\infty(M)$ can be identified with $(V/W)^*$, which we have verified to be finite dimensional, so V/W is finite dimensional. On the other hand, we will show that V/W is infinite dimensional if V is the subspace of $C^1(\mathbf{R})$ functions f with $f(0) = 0$ and W the subspace of products. For each $0 < \varepsilon < 1$, let

$$f_\varepsilon(x) = \begin{cases} x^{1+\varepsilon} : x \geqslant 0 \\ 0 \qquad\qquad x \leqslant 0 \end{cases}$$

Then the f_ε are linearly independent in V/W. If gh is the product of C^1 functions vanishing at the origin, we can write $g(x) = xg_1(x)$ and $h(x) = xh_1(x)$, where g_1 and h_1 are continuous with $g_1(0) = g'(0)$, $h_1(0) = h'(0)$, and so every element of W is of the form $x^2 k(x)$ for some continuous function k. If, for all $x > 0$,

$$a_1 x^{1+\varepsilon_1} + a_2 x^{1+\varepsilon_2} + \cdots + a_N x^{1+\varepsilon_N} = x^2 k(x)$$

where $\varepsilon_1 < \cdots < \varepsilon_N$, and $a_k \neq 0$, then, dividing by $x^{1+\varepsilon_1}$, we conclude

$$a_1 + a_2 x^{\varepsilon_2 - \varepsilon_1} + \cdots + a_N x^{\varepsilon_N - \varepsilon_1} = x^2 k(x)$$

so $a_1 = 0$, evaluating both sides of the equation in the limit as $x \downarrow 0$, a contradiction proving that V/W is infinite dimensional. Thus we cannot define the tangent space of a C^1 manifold as the space of derivations on $C^1(M)$, and we shouldn't expect the derivations on C^k manifolds to behave any better.

3.4   Curves as Vectors

There is a third important view of the tangent bundle on a manifold, which is perhaps the most geometrically visual. To construct M_p, we consider curves $c : (a, b) \to M$ which pass through p at some time t. Given our previous construction of the tangent bundle, we can consider the tangent to the curve $c_*(1_t)$ in M_p, which represents the speed of the curve passing through the point; any tangent vector can be put in this form for some curve c. Classically, $c_*(1_t)$ is often denoted in Leibnitzian fashion as

$$\frac{dc}{dt}$$

where the additional parameter t is obscured. We could have defined M_p by these curves, provided we identify c and c' if $dc/dt = dc'/dt$. Of course, without the original tangent bundle to work with, stating this isn't so elegant. We have to fix some coordinate system (x, U) at p, and identify two curves c and c' with $c(t) = p$, $c(t') = p$ if $(x \circ c)'(t) = (x \circ c')(t')$. Our new tangent bundle is obviously equivalent to our original tangent bundle. This is the closest intrinsic way to visualize the tangent vectors on an manifold. If a manifold does not lie in $\mathbf{R}^n$, it isn't really fair to see tangent vectors as vectors lying 'off' the manifold, because the tangent vectors don't really 'point' to anything. For instance, in Einsteinian physics, we consider the universe as a 4-dimensional manifold, and seeing the tangent bundle as vectors lying 'outside of the universe' seems particularly strange. But the directions we can travel *are* visualizable from inside the manifold, and curves describe the way these curves travel.

3.5   Sections and Vector Fields

A **section** on a vector bundle $\xi = \pi : E \to X$ is a continuous map $f : X \to E$ for which $\pi \circ f$ is the identity map. One can think of a section as a continuous choice of a vector assigned to each point in space. Because of the local triviality property, the space of continuous sections forms a module over the space of real-valued continuous functions on X. To begin with, sections provide an interesting way to study the topological behaviour of a vector bundle. A **frame** on an n dimensional vector bundle $\pi : E \to B$ is a family of sections $s_1, \dots, s_n$ such that $\{s_1(p), \dots, s_n(p)\}$ is a basis at each point.

Theorem 3.8. *There exists a frame $s_1, \dots, s_n : U \to \pi^{-1}(U)$ on a set U if and only if there exists a trivialization $\phi : \pi^{-1}(U) \to U \times \mathbf{R}^n$.*

Proof. If $\phi : \pi^{-1}(U) \to U \times \mathbf{R}^n$ is a trivialization, we can define a *local frame* $s_1, \dots, s_n$ on U by setting $s_k(p) = \phi^{-1}(p, e_k)$. On the other hand, let s_k be a local frame defined on a common domain U, upon which there is a local trivialization $\psi : \pi^{-1}(V) \to V \times \mathbf{R}^n$ for some $V \subset U$, then the maps $\psi \circ s_k$ can be seen as a series of maps from V to $\mathbf{R}^n$, which can be stacked together, row by row, to form a single map $F : V \to GL(n)$. If we define a map $\phi : \pi^{-1}(U) \to U \times \mathbf{R}^n$ by $\phi(\sum a^i s_i(p)) = (p, a)$, then we find $(\psi \circ \phi^{-1})(p, v) = (p, F(p)(v))$, which is a homeomorphism since it is

a bundle equivalence. But by patching together over a covering of U by neighbourhoods V, we conclude that ϕ is an equivalence everywhere. $\square$

Sections also provide easy ways to form of **subbundles** $F \subset E$ of bundles $\pi : E \to B$, which are subsets which also form vector bundles in the sense that F has the same vector space structure as E (F_p is a subspace of E_p at each point), and under the same topology F has local trivializations.

Theorem 3.9. *A subset $F \subset E$ is an m dimensional subbundle of E if and only if there are m linearly independant sections $s_1, \ldots, s_m$ into F in a neighbourhood of each point in the base space B.*

Proof. The trivialization gives us the linearly independant sections, and conversely, a modification of our argument as to why frames give local trivializations show F has a local trivialization. Instead of arguments about $GL(n)$, we need to use arguments about constant rank matrices instead, which we leave to the reader to modify from our previous discussion as to why $M(n, m; k)$ is a manifold. $\square$

Example. *Let $f : \xi \to \psi$ be a bundle map between two vector bundles, such that the map $p \mapsto rank(f_p)$ is locally constant on the base space of ξ, then $Ker(f)$, the union of all the kernels of f_p on the fibres of ξ, forms a subbundle of ξ, whose dimension at each point p is the different in dimensions between the dimension of f at p and the rank of f_p. By locality, it suffices to prove this for bundles maps $f : \varepsilon^n(X) \to \varepsilon^m(Y)$ on trivial bundles with a constant rank k. Notice that in this case f corresponds to a choice of a continuous function from X to Y, as well as a continuous choice $M : X \to M(n, m; k)$ such that $f(p, v) = (f(p), M(p)(v))$. Slight modifications to our arguments about $M(n, m; k)$ in the first chapter guarantee that there exists a continuous choice of matrix $N(p) \in GL(n)$ locally around each point p such that*

$$M(p)N(p) = \begin{pmatrix} A(p) & 0 \\ B(p) & 0 \end{pmatrix}$$

where $A \in GL(k)$, and $B \in M(n-k, k)$. But then the sections $s_j(p) = (p, N(p)(e_{k+j}))$, for $j \leqslant n-k$, are linearly independent and parameterize the kernel of f locally. It follows from the last theorem that $Ker(f)$ is a subbundle of ξ. Notice that these ideas can also be used to prove that the image of f in ψ is a subbundle, because if we choose a continuous $N(p) \in GL(n)$ such that

$$N(p)M(p) = \begin{pmatrix} A(p) & B(p) \\ 0 & 0 \end{pmatrix}$$

where $A(p) \in GL(k)$ and $B(p) \in M(n, n-k)$, then $s_j(p) = (p, N(p)M(p)(e_j))$ are linearly independent and consistute a frame for the image of f.

Example. *If $f : \xi \to \psi$ is a bundle map injective as a map over base spaces, and the rank of f is locally constant, then we can also make the union $\mathrm{coker}(f)$ of the cokernels of the f_p, formed by the union of the cokernels of each f_q (well defined because $f^{-1}(q)$ consists of at most one point) and given the quotient topology, into a vector bundle over the base space of ψ. To see this, it suffices to note the image of f is a subbundle of the extension space of ψ, so we may choose a frame $s_1, \ldots, s_k$ locally mapping into the image. This frame may then be extended locally to a frame on entire fibres of ψ, and these added sections, once projected into the quotient space, constitute a trivialization of the cokernel bundle.*

Example. *If $\xi = \pi : E \to X$ is a bundle, and $f : Y \to X$ is continuous, we can define an **induced bundle** $f^*(\xi)$ on X, by letting the extended space be the subspace*

$$F = \{(x, v) \in X \times E : f(x) = \pi(v)\}$$

If $s_1, \ldots, s_n : U \to E$ is a local frame for ξ, then the maps $t_k(x) = (x, (s_k \circ f)(x))$ is a local frame on $f^*(\xi)$, showing $f^*(\xi)$ is locally trivial. If $\psi = v : F_0 \to Y$ is another bundle over Y with a bundle map $f : \psi \to \xi$ which is an isomorphism on each fibre, then ψ is equivalent to $f^*(\pi)$ by the map $g(v) = (x, (s_k \circ f)(x))$. It follows from this that if $g : Z \to Y$ is another continuous map, then $g^*(f^*(\pi))$ is equivalent to $(f \circ g)^*(\pi)$.

Example. *If ξ and ψ are bundles over the same base space X, then we can define a new bundle $\xi \oplus \psi$ over X called the **Whitney sum** of the bundles, which using our previous exposition can be most easily described as $\Delta^*(\pi \times \xi)$, where $\Delta : X \to X \times X$ is the diagonal map $\Delta(x) = (x, x)$, so it is a vector bundle.*

The main sections we see in differential geometry are sections of the tangent bundle on the manifold, and we call these sections **vector fields**. On a smooth vector bundle, we can consider smooth sections, and the smooth vector fields are the vector fields we will really care about. On tangent bundles, we denote sections by lower case letters like s or t, whereas a vector field is often denoted by capital letters like X, Y, or Z, and the value of a vector field X at a point is X_p. Vector fields are so important that we often denote a single vector in the tangent bundle as X, as well, which could get confusing if we forget to distinguish the two, but often vector

fields act just like vectors in the same was that real valued functions 'act' like numbers, so this doesn't cause a problem. Vector fields form a vector space. Locally, around a chart (x, U), we may express the vector field in terms of the basis

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

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 $X(f) \in C^\infty(M)$ by setting $X(f)(p) = X_p(f)$. Since X_p is a derivation at p, which acts on functions in $C^\infty(M)$. What's more, we have the elegant equation

$$X(fg) = f X(g) + g X(f)$$

which expresses the pointwise derivation property globally. In general, a **derivation** is a map $F : A \to A$ on an algebra such that

$$F(ab) = a F(b) + b F(a)$$

If $F : C^\infty(M) \to C^\infty(M)$ is any derivation, then we may then define a vector field X such that X_p is the unique vector satisfying

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

and it is easily verified that $X_p \in M_p$, and that X itself is a smooth vector field. This vector field is the unique field which generates a derivation on $C^\infty(M)$, 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 on $C^\infty(M)$ are in one to one correspondence.

3.6   Tangents on Manifolds with Boundary

Given a manifold M with boundary, it is easy to define the tangent bundle on the interior of the manifold, but it is less clear what the fibres of the

bundle should look like on the boundary; should they have the same dimension as on the interior, or one dimension less? We shall find it is most convenient, suprisingly, to make the tangent space at the boundary as the dimension of the space on the interior.

The safest option is to look at the space of derivations at points on the boundary of the manifold, which are defined exactly the same as on any manifold without boundary. The locality properties apply, and so it suffices to determine the space V of derivations on $\mathbf{H}^n$. Surely we have the partial derivatives

$$\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^{n-1}}$$

So the space is at least $n-1$ dimensional. But perhaps suprisingly, the partial derivative operator $\frac{\partial}{\partial x^n}$ is also still well defined – to calculate it, we take some $f \in C^\infty(M)$, and extend to some differentiable $\tilde{f} : \mathbf{R}^n \to \mathbf{R}$, and then take partial derivatives. The value we obtain is independent of extension, because we can also calculate the partial derivative as the limit

$$\lim_{t \to 0^+} (f \circ x^{-1})(te_n)$$

Arguing the same way as in our calculation of the space of derivations in $\mathbf{R}^n$, we may take Taylor series to determine that these partial derivatives span the space of all derivations. Thus, viewing TM as the space of derivations over every point, we find that M_p is n dimensional at the boundary.

An additional feature of the tangent bundle on the boundary of a manifold is we can identify some vectors as 'outward' pointing, 'inward' pointing, and parallel to the manifold at the boundary. For a chart (x, U) at the boundary, we can write any derivation at $p \in U$ as

$$\sum a_i \frac{\partial}{\partial x^i}$$

Recalling that $\mathbf{H}^n = \{x : x_n \geqslant 0\}$, we say the derivation is outward pointing if $a_n < 0$, and inward pointing if $a_n > 0$, and parallel to the manifold if $a_n = 0$. This is well defined even when we vary coordinate systems, because if $y \circ x^{-1} : U \to V$ is a diffeomorphism, where $U, V \subset \mathbf{H}^n$, then by invariance of domain we know that

$$y \circ x^{-1}(t_1, \dots, t_{n-1}, 0) = (f^1(t_1, \dots, t_{n-1}), \dots, f^{n-1}(t_1, \dots, t_{n-1}), 0)$$

so that for $p \in \partial M$,

$$\frac{\partial y^n}{\partial x^i}\bigg|_p = \begin{cases} 0 & : i \neq n \\ \text{positive} & : i = n \end{cases}$$

where the fact for $i = n$ follows because y^n is always non-negative, and $(y^n \circ x^{-1})(x(p) + te^n) \to 0$ as $t \downarrow 0$, so that the function $t \mapsto (y^n \circ x^{-1})(p + te_n)$ must be an increasing function in a suitably small half neighbourhood of the origin. Thus we find that if a vector v can be written

$$\sum a_i \frac{\partial}{\partial x^i} = \sum b_j \frac{\partial}{\partial y^j}$$

then

$$b_n = \sum a_k \frac{\partial y^n}{\partial x_k} = a_n \frac{\partial y^n}{\partial x^n}$$

and we see that the sign of b_n and a_n agree.

3.7   Universality of the Tangent Bundle

For a map $f : M \to N$, the map $f_* : TM \to TN$ is meant to be a sufficient generalization of the derivative operator on Euclidean space. The fact that this is a 'universal' generalization can in fact be proved, once we introduce a categorical viewpoint. Note that the association of M with TM and f with f_* is a *functor* from the category of C^∞ manifolds to the category of smooth vector bundles, such that

- The bundle $T\mathbf{R}^n$ is isomorphic to $\varepsilon^n(\mathbf{R}^n)$ by a trivialization t_n such that for any map $f : \mathbf{R}^n \to \mathbf{R}^m$, the diagram commutes, where old f_* is defined by $(p, v) \mapsto (f(p), Df(p)(v))$.

- If $U \subset M$ is an open submanifold, there are equivalences

$$u_{U,M} : TU \to (TM|_U)$$

  such that if $i : U \to M$ is the inclusion map, then commutes, and for any differentiable $f : M \to N$, we have a diagram

The first bullet says that the functor, restricted to Euclidean spaces, is naturally equivalent to the functor which associates $\mathbf{R}^n$ with $\varepsilon^n(\mathbf{R}^n)$ and f :

$\mathbf{R}^n \to \mathbf{R}^m$ with the old definition of the differential. The second says that the family of maps $u_{U,M}$ is a natural transformation between the tangent functor and the restriction of the functor to open submanifolds.

We shall verify that any functor satisfying these properties is unique up to a natural equivalence. We shall first explore the properties of functors satisfying the properties above. So until necessary, we will let $M \to TM$ stand for such a functor. These properties will be obvious for the standard tangent functor, but as long as we don't use any facts about our constructed tangent bundle, the theorem will remain true for any of the other functors.

Given a chart (x, U) on a manifold M, we have a chain of isomorphisms

$$(TM)|_U \xleftarrow{u_{U,M}} TU \xrightarrow{x_*} Tx(U) \xrightarrow{u_{x(U),\mathbf{R}^n}} (T\mathbf{R}^n)|_{x(U)} \xrightarrow{t_n|_{x(U)}} \varepsilon^n(x(U))$$

Define $\alpha_x : (TM)|_U \to \varepsilon^n(x(U))$ to be the chain of compositions. We shall try and understand the properties of α_x independent of the properties of our tangent bundle, because if we have some other functor $M \mapsto T'M$, with other equivalences t'_n and $u'_{U,M}$, then we have $\beta_x : (T'M)|_U \to \varepsilon^n(x(U))$ defined exactly as α_x is defined, and we can consider $\beta_x^{-1} \circ \alpha_x : (TM)|_U \to (T'M)|_U$. Provided that this map is independent of x, we can put the maps together for all x, and obtain an equivalence between TM and $T'M$.

Lemma 3.10. *If $V \subset U$ is open, and $y = x|_V$, then $\alpha_y = (\alpha_x)|_V$.*

Proof. Denote the inclusion maps by $i : U \to M$, $\iota : V \to M$, $j : V \to U$, and $k : y(V) \to x(U)$. Then consider the diagram We verify commutativity by verifying the commutativity of each square. The first square follows by breaking the diagram into triangles. These subtriangles commutes by the universal properties of the functor, and these gives us the commutativity of the square. The second square follows by functoriality, because $x \circ j = k \circ y$. The third square commutes for the same reason the first square commutes, and the last square obviously commutes. $\qquad \square$

Lemma 3.11. *If $A \subset \mathbf{R}^n$ and $B \subset \mathbf{R}^m$ are open, and $f : A \to B$ is C^∞, then the diagram commutes.*

Proof. First, assume that f can be extended to a map $g : \mathbf{R}^n \to \mathbf{R}^m$. Denoting inclusions by $i : A \to \mathbf{R}^n$ and $j : B \to \mathbf{R}^m$. Then we have a huge diagram, whose subtriangles and subsquares all obviously commute. and

this diagram contains the other diagram as a subdiagram, since we know that old g_* and old f_* agree on A (since A is an open subset), so that the theorem is proved in this case.

In general, we might not be able to extend the entire map f to all of $\mathbf{R}^n$, but we can at least find a function g for each $p \in A$ such that g agrees with f on a neighbourhood $A' \subset A$ of p. Then we have a commutative diagram where $i : A' \to A$ is the inclusion map. Every subshape but the top left and bottom rectangle obviously commutes. First, note that the bottom rectangle is just an instance of this theorem, but where we can extend our map to all of $\mathbf{R}^n$, so we have already argued it's commutativity. To see that the top left square commutes, we extend it to a larger diagram. Defining $j : A \to \mathbf{R}^n$ to be the inclusion, we have a diagram To prove that $u_{A,\mathbf{R}^n} \circ i_* = u_{A',\mathbf{R}^n}$, it suffices to prove that $j_* \circ i_* = u_{A',\mathbf{R}^n}$, because j_* is equal to $u_{A,\mathbf{R}^n}$ when viewed as functions without a specified codomain. But $j \circ i$ is just the inclusion of A' in $\mathbf{R}^n$, so this fact is obvious. We have verified enough commutativity to conclude that the composition

$$ TA \xrightarrow{f_*} TB \xrightarrow{u_{B,\mathbf{R}^m}} (T\mathbf{R}^m)|_B \xrightarrow{t_m|_B} \varepsilon^m(\mathbf{R}^m)|_B $$

is equal to the composition

$$ TA \xrightarrow{u_{A,\mathbf{R}^n}} (T\mathbf{R}^n)|_A \xrightarrow{t_n|_A} \varepsilon^n(\mathbf{R}^n)|_A \xrightarrow{\text{old } g_*} \varepsilon^m(\mathbf{R}^m)|_B $$

on $TA|_{A'}$, and since old g_* is equal to old f_* on $(TA)|_{A'}$, we have verified the theorem over A'. Since A' was arbitrary, we obtain commutativity over all of A. $\qquad\square$

Lemma 3.12. *If $(x, U), (y, V)$ are two coordinates charts with $p \in U \cap V$, then $\beta_y^{-1} \circ \alpha_y$ and $\beta_x^{-1} \circ \alpha_x$ agree at p.*

Proof. We may assume $U = V$, because our first lemma shows the theorem is true in general otherwise. Assuming this is true, we consider the diagram The triangle obviously commutes, and the rectangle commutes by the second lemma. This implies that $\alpha_y = \text{old } (y \circ x^{-1})_* \circ \alpha_x$, and $\beta_y = \text{old}(y \circ x^{-1})_* \circ \beta_x$. The desired result follows immediately. $\qquad\square$

Putting together all $\beta_x^{-1} \circ \alpha_x : (TM)|_U \to (T'M)|_U$, we now have a well defined equivalence e_M from TM to $T'M$. Now we need only prove that this is in fact a natural equivalence – that is, for any $f : M \to N$,

$e_N \circ f_* = f_\sharp \circ e_M$. Let's prove this first for $f : \mathbf{R}^n \to \mathbf{R}^m$. First, note that $e_{\mathbf{R}^n} = (t_n')^{-1} \circ t_n$, because most of the maps involved in the construction of the equivalence are trivial. The properties of the tangent map imply that the squares and triangles of the diagram commute, and this shows the naturality equation holds. Second, note that $e_M \circ i_* = i_\sharp \circ e_U$, where $i : U \to M$ is the inclusion of an open submanifold of M, where (x, U) is a chart, because e_M is essentially formed by putting together all $i_\sharp \circ e_M \circ i_*^{-1}$. Similarily, we find that if (x, U) is a chart, then $e_{x(U)} \circ x_* = x_\sharp \circ e_U$. Putting all this together, we prove that, given $f : M \to N$, $e_N \circ f_* = f_\sharp \circ e_M$ holds at a neighbourhood of each point. Fixing some point p, we choose a chart (x, U) containing the point, and a chart (y, V) containing $f(p)$, such that $f(U) \subset V$. We may assume $x(U) = \mathbf{R}^n$ and $y(V) = \mathbf{R}^m$. If $g = y \circ f \circ x^{-1}$, then Then we have justified that all the squares in the diagram but the middle one commute, as do the top and bottom triangles, and the rectangle as a whole (in the sense that $e_{\mathbf{R}^m} \circ g_* = g_\sharp \circ e_{\mathbf{R}^n}$). But a final diagram chase justifies that $f_\sharp \circ e_M$ is equal to $e_N \circ f_*$ on $TM|_U$. Since (x, U) was arbitrary, we have shown that the property is a natural equivalence in full. Thus

Theorem 3.13. *There is a unique functor from the category of C^∞ manifolds to the category of vector bundles which satisfies the bullet point properties we denoted at the beginning of the section.*

The beauty of this categorical proof is that it depends on very little of the structure of C^∞ manifolds. The proof easily extends to C^k manifolds, and even to C^ω manifolds. It also shows uniqueness of association to C^∞ vector bundles, and effectively settles the question of how well the tangent bundle represents the linear property of differentiable maps on spaces.

3.8   Orientation

The key idea of differential geometry is that classical geometric concepts (which do not 'really' hold ground rigorously) can be given a rigorous standing when reinterpreted as some structure on the tangent bundle. These are normally just simple extensions of linear algebraic constructions, applied over each fibre of the tangent bundle. The first, easiest concept to introduce, is orientation. In a *real* vector space V (in complex vector spaces no such problem arises), each tuple $(v_1, \ldots, v_n)$ gives rise to a

linear isomorphism $T : \mathbf{R}^n \to V$ such that $T(e_i) = v_i$. Given another basis $(w_1,\ldots,w_n)$, we obtain another isomorphism $S : \mathbf{R}^n \to V$ with $S(e_i) = v_i$, and therefore a linear operator $T \circ S^{-1} : \mathbf{R}^n \to \mathbf{R}^n$. The determinant of this operator is non-zero and therefore positive or negative. We say these tuples are *equally oriented* if the determinant of this operator is positive, and *oppositely oriented* if the determinant is negative. This divides the bases of V into two equivalence classes, and an **orientation** for V is a choice of one of these classes. The equivalence class of some basis $(v_1,\ldots,v_n)$ shall be denoted $[v_1,\ldots,v_n]$, so that if V is an oriented vector space with fixed orientation μ, then $(v_1,\ldots,v_n)$ is oriented if and only if $[v_1,\ldots,v_n] = \mu$. Given two vector spaces V and W with a fixed orientation, a linear isomorphism $T : V \to W$ is orientation preserving if T maps oriented bases $(v_1,\ldots,v_n)$ in V to an oriented bases $(Tv_1,\ldots,Tv_n)$ in W. Either an isomorphism is orientation preserving or orientation reversing – it maps oriented bases $(v_1,\ldots,v_n)$ to unoriented bases $(Tv_1,\ldots,Tv_n)$.

The key reason that orientation exists is that the determinant of an operator is always positive or always negative. We see the same phenomenon occur when considering equivalences $f : \varepsilon^n(X) \to \varepsilon^n(X)$, which can be written

$$f(p,v) = \left(p, \sum a_{ij}(p)v^i e_j\right)$$

where the $a_{ij} : X \to \mathbf{R}$ change continuously, and since the matrix (a_{ij}) is always invertible, it's determinant is either always positive or always negative. Thus we may define an orientation to the bundle $\varepsilon^n(X)$ as a choice of orientation on each fibre, such that every equivalence of the form is orientation preserving on all fibres, or orientation reversing. Our reasoning shows that there are only two choices of orientation on $\varepsilon^n(X)$.

Now on an arbitrary vector bundle, defining such a vector bundle may be impossible, because we have to patch local equivalences up across the whole space. An orientation for an arbitrary bundle $\pi : E \to B$ is a choice of orientation μ_p on each fibre B_p, such that any local trivialization $t : \pi^{-1}(U) \to U \times \mathbf{R}^n$ around a connected set U is either orientation preserving on all fibres, or orientation reversing on all fibres. To verify that a particular orientation is 'consistant', we need only verify it for a set of connected open sets which cover the bundle, because trivializations are already orientation reversing or preserving when they are locally trivialized, as we just calculated noted. A bundle is called orientable if it has an orientation, and a differentiable manifold is called orientable if its tangent

bundle is orientable. An oriented manifold is a manifold with a fixed orientation on its tangent bundle, and if $f : M \to N$ is a local diffeomorphism between oriented manifolds of the same dimension, we say f is orientation preserving if $f_* : TM \to TN$ is orientation preserving on each fibre.

Example. *Since $T\mathbf{R}^n$ is equivalent to $\varepsilon^n(\mathbf{R}^n)$, Euclidean space is an orientable manifold, with a canonical orientation induced by the canonical choice of basis, the unit vectors $\mu_p = [(e_1)_p, \dots, (e_n)_p]$. If U is an open subset of $\mathbf{R}^n$, then TU can be embedded in $T\mathbf{R}^n$, and we can define an orientation μ_p in the same way. More generally, if $\pi : E \to B$ is an orientable bundle, and if $F \subset E$, then the bundle $\pi|_F$ is also orientable.*

Example. *The spheres S^n are all orientable – to obtain the orientation, we note that, viewing TS^n as a subset of $\varepsilon^{n+1}(\mathbf{R}^{n+1})$, we see that $p_p \notin TS^n$ for any $p \in S^n$. We may then define an orientation on S^n by letting $[v_1, \dots, v_n]$ be an oriented basis at p if $[p, v_1, \dots, v_n]$ is oriented in $\mathbf{R}^n$. More generally, if M is a manifold with boundary with orientation μ, then ∂M has a natural orientation by saying a basis $(v_1, \dots, v_{n-1})$ of ∂M_p is orientable if $[w, v_1, \dots, v_{n-1}] = \mu$ for any outward pointing vector $w \in M_p$. If we consider the orientation on $\mathbf{R}^{n-1}$ as a subset of $\mathbf{H}^n$, we see that we obtain $(-1)^n$ times the standard orientation. The reason for this choice will become clear when we talk about integration on manifolds, in which orientation plays a key role.*

Example. *For any manifold M, TM is always orientable as a differentiable manifold. We calculate the transition map between two charts induced by the coordinate maps (x, U) and (y, V) on M to be*

$$(y^1, \dots, y^n, \dot{y}^1, \dots, \dot{y}^n) \circ (x^1, \dots, x^n, \dot{x}^1, \dots, \dot{x}^n)^{-1}[p, v]$$

$$= \left[(y \circ x^{-1})(p), \sum_{i,j} v^j \left.\frac{\partial y^i}{\partial x^j}\right|_p e_i \right]$$

and so the matrix of partial derivatives is

$$M = \begin{pmatrix} \left(\frac{\partial y^i}{\partial x^j}\right) & 0 \\ X & \left(\frac{\partial y^i}{\partial x^j}\right) \end{pmatrix} = \begin{pmatrix} N & 0 \\ X & N \end{pmatrix}$$

Essentially, we have found that

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

Now

$$\begin{pmatrix} N^{-1} & 0 \\ 0 & I_n \end{pmatrix} \begin{pmatrix} N & 0 \\ X & N \end{pmatrix} \begin{pmatrix} I_n & 0 \\ 0 & N^{-1} \end{pmatrix} = \begin{pmatrix} I_n & 0 \\ X & I_n \end{pmatrix}$$

Taking determinants on both sides, we find

$$\frac{\det(M)}{\det(N)^2} = 1$$

and so $\det(M) = \det(N)^2 > 0$. *This implies that the family of charts* $(x, \dot{x})$ *have consistant orientations, and so together they give us an orientation on* $T(TM)$. *Here we have used the fact that if we have a family of chart* $\{(x_\alpha, U_\alpha)\}$ *covering a manifold M, and*

$$\det \left(\frac{\partial x_\beta^i}{\partial x_\alpha^j} \right)$$

is always positive, then the orientations induced from the $(x_\alpha)_*$ *combine to give an orientation on* TM.

We can view orientation, in another language, through the tool of frames, left to the reader to prove. A choice μ_p of orientation on a bundle $\pi : E \to B$ is consistant if and only if every local frame is either always orientation preserving or orientation reversing on each fibre (we assume the frame has some particular ordering). This is just a restatement of the duality between local frames and trivializations.

Example. *If* $\pi : E \to B$ *is a bundle with an orientation μ, and $f : X \to E$, then $f^*(\pi) = \xi : F \to X$ is orientable, which we can obtain by making the bundle map $g : F \to E$ orientation preserving on each fibre. Of course, $f^*(\pi)$ might be orientable even if π is not orientable. But, in the special case where we consider $\pi^*(\pi)$, which is a bundle over E, then $\pi^*(\pi)$ is orientable if and only if π is, because we can view the bundle π as the restriction of the bundle $\pi^*(\pi)$ over the set $\{v \in E : v \text{ is a zero vector}\}$, which is isomorphic to B.*

Example. *If* $\pi : E \to B$ *and $\xi : F \to B$ are two vector bundles over the same space which are both orientable, then the Whitney sum $\pi \oplus \xi : K \to B$ is also orientable, because the bundle can be covered by local trivializations which are just products of trivializations on each factor, and so we can easily define an orientation here by saying that if $s_k : U \to E$, $t_k : U \to F$ combine to give a local frame of $\pi \oplus \xi$, then this frame is oriented if either both s_k and t_k are*

oriented in each bundle, or if bundle is not oriented. On the other hand, if π is orientable, but ξ is non-orientable, then $\pi \oplus \xi$ is non-orientable. If $s_k : U \to F$ is a frame, then, after perhaps shrinking the neighbourhood, we may find an oriented frame $t_k : U \to E$, and then the frames combine to give a frame of $\pi \oplus \xi$. We could say s_k is oriented if the frame on $\pi \oplus \xi$ is oriented. However, if π is a bundle, then $\pi \oplus \pi$ is always orientable, because for any finite dimensional vector space V, $V \oplus V$ has a natural orientation such that, if $(v_1, \ldots, v_n)$ is any basis of V, then $(v_1, 0), \ldots, (v_n, 0), (0, v_1), \ldots, (0, v_n)$ is an oriented basis of $V \oplus V$. This allows us to give a natural orientation to $\pi \oplus \pi$ on each fibre, and it is not difficult to check this orientation is consistant in trivializations.

Another interesting perspective, provides an introduction to the generalization of vector bundles to more general 'fiber spaces', locally trivial supplementary spaces attached to topological spaces. Consider the discrete space $\{-1, 1\}$. Given a finite vector space V of dimension n, an orientation corresponds to choosing an element of $\pi_0(W)$, where

$$W(V) = \{v \in V^n : v \text{ is an ordered basis of } V\}$$

As a quotient space, W is homeomorphic to $\{-1, 1\}$. Given an n dimensional bundle $\xi = \pi : E \to X$, consider the space $O(\xi)$ obtained as the quotient of the n-fold Whitney sum $\xi \oplus \cdots \oplus \xi$ at each fibre E_p into $W(E_p)$. The projections $\pi : E^n \to X$ maps to the quotient, so we still have a continuous projection $\nu : O(\xi)$. On the trivial bundles $\varepsilon^n(X)$, $O(\varepsilon^n(X)) = O(X \times \mathbf{R}^n)$ is homeomorphic to $X \times W(\mathbf{R}^n)$, which again is homeomorphic to $X \times \{-1, 1\}$. Since ξ is locally trivial, $O(\xi)$ should also be 'locally' trivialized to $U \times \{-1, 1\}$ in certain neighbourhoods of every point. In general, given any two topological spaces X and Y, a fibre bundle ξ is a surjective map $\pi : E \to X$, such that around each point $p \in X$, there is a neighbourhood U and a homeomorphism $\phi : \pi^{-1}(U) \to U \times Y$ which respects the base point (from this, we see easily that each E_p is homeomorphic to Y). Thus a fibre bundle is a generalization of a vector bundle to be spaces which locally look like products, but without any local structure. Thus we see that $O(\xi)$ is just a fibre bundle where $Y = \{-1, 1\}$. As with vector bundles, we can consider sections of fibre bundles.

Theorem 3.14. *If X is connected, then ξ is orientable if and only if there exists a global section $s : X \to O(\xi)$, and then $O(\xi)$ is homeomorphic to the disjoint union of two copies of X, corresponding to the two choices of orientation on ξ.*

Example. *We have already seen that $O(\varepsilon^n(X))$ is homeomorphic to $X \times \{-1, 1\}$, so this provides an argument as to why $\varepsilon^n(X)$ is orientable.*

Example. *If $\xi = \pi : M \to S^1$ is the Möbius bundle over S^1, then $O(\xi)$ is homeomorphic to S^1, obtained by tracing a unit tangent vector along two revolutions of the Möbius strip. This is essentially how we argued that ξ was not orientable, and that it wasn't equivalent to the trivial bundle over S^1.*

3.9   Whitney's Embedding Theorem

We now use our results about tangent spaces and differentiability to show that all compact manifolds can be embedded in low dimensional space. This is a easy version of the general Whitney's embedding theorem, which shows this result is true for all manifolds.

Lemma 3.15. *Any compact manifold is imbeddable in some Euclidean space.*

Proof. Let M be a compact manifold. Then there is a finite set of charts $(x_1, U_1), \ldots, (x_n, U_n)$ covering M. Since the U_k is locally finite, by the shrinking lemma we may choose open sets $V_k \subset U_k$ with $\overline{V_k} \subset U_k$ which cover M. Consider a partition of unity ψ_k equal to 1 on V_k, and subordinate to U_k, and define

$$f(p) = (\psi_1(p)x_1(p), \ldots, \psi_n(p)x_n(p), \psi_1(p), \ldots, \psi_n(p)) = (f_1(p), \ldots, f_n(p))$$

then f is an immersion of M, because if $p \in V_k$, then in a neighbourhood of p we have $f_k(p) = x_k(p)$, so the partial derivative matrix

$$\left(\frac{\partial f^i}{\partial x_k^j} \right)$$

contains the identity as a submatrix, so the map has full rank. The map is also injective, because if $p \in V_m$, and $\psi_k(p)x_k(p) = \psi_k(q)x_k(q)$ and $\psi_k(p) = \psi_k(q)$ for some q, then $\psi_m(q) = \psi_m(p) = 1$, so $q \in U_m$, and $x_m(p) = x_m(q)$, so $p = q$. Since M is compact, the immersion is an imbedding, so M is imbeddable in Euclidean space. □

This theorem already has an application, for we can now use it to show all compact manifolds are embeddable in a fairly low dimensional Euclidean space.

Theorem 3.16 (Whitney's Embedding Theorem). *A compact submanifold M^n of $\mathbf{R}^N$ can be embedded in $\mathbf{R}^{2n+1}$.*

Proof. We say a *chord* of M is a vector in $\mathbf{R}^N$ of the form $p - q$, for distinct $p, q \in M$. If U is the open subset of $(p, q) \in M \times M$ with $p \neq q$, then the map $f : U \to S^{N-1}$ given by

$$f(p, q) = \frac{p - q}{|p - q|}$$

is C^∞, and since $2n < N - 1$, $f(U)$ has measure zero since all values are critical. Similarily, if V is the open set of vectors $v \in TM \subset T\mathbf{R}^N$ with $|v| \neq 0$, then we have a map $g : V \to S^{N-1}$ defined by $g(v_p) = v/|v|$, and again, since $2n < N - 1$, $g(V)$ has measure zero. It follows that $f(U) \cup g(V)$ has measure zero, hence $S^{N-1} - f(U) - g(V)$ is non-empty, and so there exists a vector v not parallel to any chord in $\mathbf{R}^N$, and not parallel to any tangent space. It follows that the projection of M onto a hyperplane perpendicular to v is an injective immersion, and since M is compact, it is an imbedding. Continuing this process, we may imbed M into $\mathbf{R}^{2n+1}$. $\qquad\square$

We have proven that all compact manifolds can be embedded in some Euclidean space, so that all compact n dimensional manifolds can be embedded in $\mathbf{R}^{2n+1}$. For non-compact manifolds, this argument only shows that these manifolds can be *immersed* in $\mathbf{R}^{2n+1}$, not imbedded.

Chapter 4

The Cotangent Bundle

We now begin the real task of differentiable geometry, which is to equip the tangent space of a differentiable manifold with enough structure to begin doing some actual geometry. Normally anything we can do to a vector space, we can do to a vector bundle, and we will use this to obtain lots of different objects to work with. The essential idea to this is provided by the following general theorem. A covariant endofunctor F on the category of finite dimensional vector spaces associates with any two vector spaces V and W a map from $\mathrm{Hom}(V, W)$ to $\mathrm{Hom}(F(V), F(W))$, and F is called a *continuous* functor if the map is continuous for each V and W.

Theorem 4.1. *If F is a continuous endofunctor, and $\xi = \pi_0 : E \to B$ is any vector bundle, then there is a bundle $F(\xi) = \pi_1 : E' \to B$ for which $E'_p = F(E_p)$, and such that every trivialization $t : \pi^{-1}(U) \to U \times \mathbf{R}^n$ corresponds to a trivialization $F(t) : \pi_1^{-1}(U) \to U \times F(\mathbf{R}^n)$.*

Proof. Given ξ, define $E' = \bigcup F(E_p)$, with $\pi_1 : E' \to B$ projecting $F(E_p)$ onto p. A trivialization t corresponds to a set of linear maps $t_p : E_p \to \mathbf{R}^n$, for each p in some open set U, and we define s as the map corresponding to the set of maps $F(t_p) : F(E_p) \to F(\mathbf{R}^n)$. We declare a topology on $F(\xi)$ by letting the maps $F(t)$ be trivializations. If t and s are two trivializations on a common open set U, then $F(t) \circ F(s)^{-1} = F(t \circ s^{-1})$ corresponds to a map from $U \times F(\mathbf{R}^n)$ to itself, which is a homeomorphism if and only if the map $p \mapsto F((t \circ s^{-1})_p)$ from U to $GL(F(\mathbf{R}^n))$ is continuous. But we know that the map $t \mapsto (t \circ s^{-1})_p$ is continuous, and so the continuity of the corresponding map is also continuous because the functor is continuous. By choosing some isomorphism of $F(\mathbf{R}^n)$ with Euclidean space once and

for all, we obtain a vector bundle $F(\xi)$. If $f : \xi \to \eta$ is a bundle map, then it is given by maps $f_p : E_p \to F_p$, for each p, and we can obtain a map $F(f) : F(\xi) \to F(\eta)$ by taking the union of the $F(f_p)$. The continuity of the functor again guarantees that $F(f)$ remains continuous, and is therefore a bundle map. $\qquad\square$

Thus a continuous functor on the category of vector spaces extends uniquely to a functor on the category of vector bundles which acts as the original functor on the fibres of bundle maps. If we also consider smooth functors instead of continuous functors, the same arguments show that we obtain a functor on the category of smooth vector bundles. A similar process can be carried out when F is a functor of multiple variables, or is contravariant, and we will take these generalizations for granted in the sequel.

4.1   The Dual of a Vector Bundle

The first object we can equip the tangent bundle is with cotangent vectors, which often operate as differentials in certain geometric arguments. Recall that if V is a vector space, then there is a natural vector space V^* associated with it, known as the *dual space* of V, which is the space of all linear functionals $f : V \to \mathbf{R}$. If $f : V \to W$ is a linear map, we can define another linear map, the dual $f^* : W^* \to V^*$, by setting $f^*(g) = g \circ f$. Thus the dual can be viewed as a *contravariant* functor in the category of vector spaces. If V is finite dimensional, and has some basis $(v_1, \ldots, v_N)$, then we can construct a *dual basis* $(v_1^*, \ldots, v_N^*)$ on V^* defined by $v_n^*(v_m) = \delta_{nm}$. This implies the functor $V \mapsto V^*$ is smooth, for if a map f is given in some pair of bases by a matrix M, then f^* will be given by the matrix M^T with respect to the dual bases, and the map $M \mapsto M^T$ is smooth. Thus, given any vector bundle ξ, we can associate with it the *dual bundle* ξ^*, whose elements consist of functionals over a particular fibre.

For any finite dimensional vector space V, V^* is isomorphic to V. But this is an artificial fact emerging from the fact that V^* has the same dimension as V, and there is no 'canonical' way to identify V with it's dual. This implies that ξ^* is not always equivalent to ξ, though, since V^{**} is naturally isomorphic to V, we can always identify ξ^{**} with ξ. This example can be generalized, and the proof of the equivalence is no more complicated.

Theorem 4.2. *Let F and G be continuous functors on finite dimensional vector spaces naturally isomorphic to one another. Then the extensions of F and G to vector bundles are naturally isomorphic to one another.*

Proof. Let η be a natural isomorphism between F and G. Given $\xi = \pi : E \to B$, we can define a map $\eta_\xi : F(\xi) \to G(\xi)$, which is an isomorphism on each fibre, by putting together the equivalences $\eta_{F(E_p)} : F(E_p) \to G(E_p)$, for each point p. The only tricky part is to verify this map is continuous. For any trivialization t of ξ on $U \subset B$, we have a commutative diagram

$$
\begin{array}{ccccccc}
 & & & \eta_{F(\varepsilon^n(U))} & & & \\
U \times F(\mathbf{R}^n) & \xleftarrow{\;F(t)\;} & F(\pi)^{-1}(U) & \xrightarrow{\;\eta_\xi\;} & G(\pi)^{-1})(U) & \xrightarrow{\;G(t)\;} & U \times G(\mathbf{R}^n)
\end{array}
$$

The diagram above implies that in order to prove η_ξ is continuous, it suffices to prove that $\eta_{F(\varepsilon^n(U))}$ is continuous for each set U. But this is just

$$\eta_{F(\varepsilon^n(U))}(v_p) = \eta_{F(\mathbf{R}^n)}(v)_p$$

which is obviously continuous, because it is a constant choice of a linear map. Thus η truly does extend to a bundle equivalence, and the naturality is clear. $\qquad\square$

The most important case of the dual bundle occurs when we study the tangent bundle TM, in which case the result bundle T^*M is known as the **cotangent bundle** of M. It is a smooth vector bundle, so we can consider smooth sections, which we call *covector fields*. These covectors fields are rarely geometrically viewable. Instead, covector fields *act* on vector fields; given a covector field ω and a vector field X, we can define a function $\omega(X)$ by $\omega(X)(p) = \omega(p)(X(p))$. If ω and X are both smooth fields, then $\omega(X)$ will also be smooth.

Example. *For any smooth real-valued function $f : M \to \mathbf{R}$, we can define a covector field $df : M \to T^*M$, such that $df(p)(v) = v(f)$. The covector field df is known as the differential of f. If X is a vector field, then $df(X) = X(f)$ is a smooth function.*

If (x, U) is a coordinate system on M, then the x^i are smooth, and so we can consider the local differentials dx^i. Now we calculate

$$dx^i(p)\left(\left.\frac{\partial}{\partial x^j}\right|_p\right) = \left.\frac{\partial x^i}{\partial x^j}\right|_p = \delta^i_j$$

78

And so the $dx^i(p)$ are precisely the dual basis corresponding to the vectors $\partial/\partial x^i$. This means every covector field ω can be locally written as

$$\omega = \sum \omega \left(\frac{\partial}{\partial x_i} \right) dx^i = \sum \omega_i dx^i$$

and ω is continuous/smooth if on this chart if and only if the ω_i are smooth/continuous. In particular, since

$$df \left(\frac{\partial}{\partial x^i} \right) = \frac{\partial f}{\partial x^i}$$

We obtain the classical formula

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

which holds throughout the domain of the coordinate system. If $\sum \omega_i dx^i = \sum \eta_j dy^j$, then

$$\sum \eta_j dy^j = \sum \eta_j \frac{\partial y^j}{\partial x^i} dx^i$$

so we find

$$\omega_i = \sum \eta_j \frac{\partial y^j}{\partial x^i}$$

A smooth covector field used to be seen as an assignment of n functions to each coordinate system on a manifold, which are related to one another by the equation above, complicated by the resulting topology of the manifold. The advantage of the modern approach using the tangent bundle is that these complications are summarized in the topology of TM rather than in the relations between the various coordinate systems. Just like the fact that a coordinate system q induces a coordinate system $(q, \dot{q})$ on TM, it also induces a coordinate system (q, p) on T^*M, such that $p_i(\sum a_j dq^j) = a_j$.

Example. *Let M be the configuration space of a dynamical system with local generalized coordinates $q_1, \ldots, q_n$, with a time independant Lagrangian $L(q, \dot{q})$. Thus we can think of L as a real valued function on TM. We now consider the transition of Lagrangian mechanics to Hamiltonian mechanics. The generalized momenta are defined as*

$$p_i = \frac{\partial L}{\partial \dot{q}^i}$$

In classical mechanics, assuming that $\det(\partial p_i / \partial \dot{q}^j) \neq 0$, the coordinates (q, p) are taken as an alternate coordinate system on TM. One problem with this is that the coordinates p are no longer geometrically significant. Furthermore, if we change from coordinates q_0 to q_1, we obtain two different generalized momenta p^0 and p^1, and we then find that

$$p_i^0 = \frac{\partial L}{\partial \dot{q}_0^i} = \sum \frac{\partial L}{\partial \dot{q}_1^j} \frac{\partial \dot{q}_1^j}{\partial \dot{q}_0^i} = \sum p_j^1 \frac{\partial \dot{q}_1^j}{\partial \dot{q}_0^i}$$

Thus the momenta behave more like covectors than vectors, and so we find that we can put all the maps p together to come up with a bundle map $p : TM \rightarrow T^*M$, setting $p(q, \dot{q}) = \sum p_i(q, \dot{q}) dq^i(q)$, which is an isomorphism because of the determinant criterion. The space T^*M is therefore known as the phase space of the dynamical system. Because of the nonvanishing determinant condition, p is surjective. The Lagrangian can in most cases in mechanics be expressed as $L(q, \dot{q}) = T(q, \dot{q}) - V(q)$, where T is a kinetic energy, and V a potential energy independent velocity. Often T is a positive definite symmetric quadratic form

$$T(q, \dot{q}) = \frac{1}{2} \sum g_{ij}(q) \dot{q}^i \dot{q}^j$$

where the g_{ij} are known as masses, or at least can be approximated by one. Then

$$p_k = \frac{\partial T}{\partial \dot{q}^k} = \sum g_{ik}(q) \dot{q}^i$$

The operation $2T$ is a two tensor, that gives M the structure of a Riemannian manifold. And now we see that the momentum is just the covariant version of the velocity vector, under the musical isomorphism. Thus velocity and momentum are duals of one another with respect to the kinetic energy tensor.

There is a canonical covariant vector field on T^*M (a section into T^*T^*M!) known as the Poincare one form λ, given by $\lambda(q, p) = \sum p_i dq^i$. It is not a covariant vector field on M itself, even though it is defined over dq^i, and not over dp_i, because the coefficients p_i are not functions of M. Essentially, we just embed T^*M in $T^*(T^*M)$ in a canonical way. The coordinate independant definition of λ is given as follows. Given a one form $\alpha \in T^*M$, since the map $\pi : T^*M \rightarrow M$ is differentiable, we can consider $\pi^* \alpha \in T^*(T^*M)$, and this is $\lambda(\alpha)$.

Let $\phi(x, u)$ be an equation on $\mathbf{R}^n \times \mathbf{R}^m$ homogenous of degree two in the variable u, i.e. $\phi(x, \lambda u) = \lambda^2 \phi(x, u)$. This implies

$$\langle \nabla_u \phi(x, u), u \rangle = \lim_{\lambda \to 0} \frac{\phi(x, (\lambda+1)u) - \phi(x, u)}{\lambda} = \phi(x, u) \lim_{\lambda \to 0} \frac{(\lambda+1)^2 - 1}{\lambda}$$

$$= 2\phi(x, u)$$

If we define $\rho_i = \partial_{u^i}(\phi)$, then $2\phi = \sum \rho_i u^i$. Thus taking differentials gives $2d\phi = \sum \rho_i du^i + u^i d\rho_i$. But we also have $d\phi = \sum \rho_i du^i + \partial_{x^i} \phi dx^i$, so subtracting these equations gives $d\phi = u^i d\rho_i - (\partial_{x^i} \phi) dx^i$. But we also have

$$d\phi = \frac{\partial \psi}{\partial x^i} dx^i + \frac{\partial \psi}{\partial \rho_i} d\rho_i$$

so by uniqueness of the expansion we conclude

$$u^i = \frac{\partial \psi}{\partial \rho_i} \qquad -\frac{\partial \phi}{\partial x^i} = \frac{\partial \psi}{\partial x^i}$$

These equations are very useful in mechanics, where $\phi(x, u) = \sum a_{ij} u^i u^j$ is a measure of kinetic energy, and $\rho_i = \sum a_{ij} u^j$ is some form of generalized momentum.

Classical differential geometers were not afraid to describe df as the 'infinitisimal change' with respect to f as the x_i change. Eventually, it was realized that an infinitisimal change can be described as an action on infinitisimal lengths, or tangent vectors. Indeed, since

$$f(x + v) = f(x) + \sum v^i \frac{\partial f}{\partial x^i} + o(|v|^2)$$

to a first order, we find $f(x + v) - f(x) = df(v)(x)$. Thus the functions df and dx^i become functions, and the classical notation was preserved so the usual formulae hold.

If $f : M \to N$ is smooth, then we know we obtain a differential map $f_* : TM \to TN$. For a fixed $p \in M$, f_* acts as a linear map $(f_*)_p : M_p \to N_{f(p)}$. We may then consider the dual map $(f_*)_p^* : N_{f(p)}^* \to M_p^*$. The fact that f is not injective prevents up from putting all these maps together to form a map $f^* : T^*N \to T^*M$. On the other hand, given a covector field ω on N, one can define a covector field $f^*\omega$ on M by setting $(f^*\omega)(p)(v) =$

$\omega(f(p))(f_*v)$, which is a complicated notation for describing the pushforward map. This is known as the *pullback field* corresponding to ω. This makes covector fields operate quite differently to vector fields, since there is no universal way to *pushforward* a vector field X on M to f_*X on N.

4.2   The Method of Lagrangian Multipliers

The theory we have developed is enough to obtain an incredibly practical theorem, which has usage almost everywhere in applied mathematics, and has theoretical usage as well. Let M be a differentiable manifold, and let N be a C^∞ submanifold. Given a function $f \in C^\infty(M)$, we may wish to consider maximizing f over N. One method of finding the maximum over f is to find a cover of N by coordinate charts (x_α, U_α), and then find points where

$$\frac{\partial f}{\partial x_\alpha^i} = 0$$

for all i, which are candidates for extrema. In terms of our notation, if $h = f|_N$, then candidates for extrema are points p for which $dh(p) = 0$. This may prove impractical, especially when the equations for the partial derivatives are cumbersome to solve. The method of Lagrangian multipliers provides an alternate method of proof, applying when N is specified as the level set of some specifying function $g : M \to M'$, where $N = g^{-1}(p)$. We begin by trying to understand the tangent and cotangent bundle of N, in terms of the function g.

Theorem 4.3. *If $N = g^{-1}(p)$ is a differentiable manifold, where g has locally constant rank on N, then the tangent bundle TN can be defined as the set of vectors in $TM|_N$ for which $g_*(v) = 0$.*

Proof. If $g : M \to M'$ is rank k at p, then choose some coordinate system (x, U) on M and (y, V) centred at $f(p)$ such that

$$(y \circ g \circ x^{-1})(t^1, \ldots, t^n) = (t^1, \ldots, t^k, 0, \ldots, 0)$$

Then $(x^{n-k+1}, \ldots, x^n)$ is a coordinate system on N, and

$$g_*\left(\sum a^i \frac{\partial}{\partial x^i}\right) = \sum_{i,j} a^i \frac{\partial y^j \circ g}{\partial x^i} \frac{\partial}{\partial y^j} = \sum_{i=1}^{k} a^i \frac{\partial}{\partial y^i}$$

and so $g_*\left(\sum a^i \frac{\partial}{\partial x^i}\right) = 0$ if and only if $a^i = 0$ for $i = 1,\ldots,k$, in which case $\sum a^i \frac{\partial}{\partial x^i}$ is a vector in TN. $\square$

Example. *This was essentially the way we defined tangent vectors on S^n, as vectors v_p such that $\langle p, v \rangle = 0$. This is the correct definition, because the manifold can be defined by the equation $\|p\|^2 = \langle p, p \rangle = 1$, and if $f(v) = \langle v, v \rangle$ defines the inner product, then*

$$\left.\frac{\partial f}{\partial x^i}\right|_p = 2p^i$$

so a vector v_p is in TS^n if and only if $2\sum v^i p^i = 2\langle v, p \rangle = 0$.

Now consider the special case where $g : M \to \mathbf{R}^n$. Then we can consider the covector fields dg^i, and the vectors in TN are exactly those vectors for which $dg^i(v) = 0$, since dg^i is obtained from g^i_* by 'forgetting the base point'. That is

$$g_*(v_p) = (dg^1(v),\ldots,dg^n(v))_{g(p)}$$

The kernel of the reduction $M_p^* \to N_p^*$ is a space of dimension k, because $\sum a_i dx^i(p)$ annihilates all vectors in N_p if and only if $a_i = 0$ for all $i > k$, and the $dg^i(p)$ span a space of dimension k, because

$$dg^i(p) = \sum \frac{\partial g^i}{\partial x^j} dx^j(p)$$

and $D(g \circ x^{-1})$ has rank k. This implies that the dg^i are actually a spanning set of the kernel. Now if p maximizes f on N, then $df(p)(v) = 0$ for all vectors $v \in N_p$, implying $df(p)$ is in the kernel of the reduction, and implying the existence of λ_i such that

$$df(p) = \sum \lambda_i dg^i(p)$$

This gives us a different system of equations to solve, with an additional set of unknowns λ_i. If $g : M \to \mathbf{R}$, then we need only find λ such that $df(p) = \lambda dg(p)$.

Example. *Consider finding extrema of the function $f(x,y) = 5x - 3y$, subject to the constraint $x^2 + y^2 = 136$. Then the constraint function is $g(x,y) = x^2 + y^2$, and*

$$df = 5dx - 3dy \qquad dg = 2xdx + 2ydy$$

83

Extrema occur at points (x,y) such that

$$5dx - 3dy = 2\lambda x dx + 2\lambda y dy$$

Which implies that $2\lambda x = 5$, $2\lambda y = -3$, and in order for the point (x,y) to occur in the constraint region, we require $25/4\lambda^2 + 9/4\lambda^2 = 17/2\lambda^2 = 136$, so $1/16 = \lambda^2$, and so $\lambda = \pm 1/4$. If $\lambda = 1/4$, then we have an extrema $(10, -6)$ and $(-10, 6)$. Since the constraint region is compact, we must have a maxima and minima, and since the function is non-constant, one of these points must be the maixmum, and one the minimum. Since $f(10, -6) = 68$, $f(-10, 6) = -68$, $(10, -6)$ is the maxima of f, and $(-10, 6)$ the minima.

Example. *Let $T : \mathbf{R}^n \to \mathbf{R}^n$ be a self-adjoint operator, and consider maximizing $\langle Tx, x \rangle$ over S^{n-1}. If $f(x) = \langle Tx, x \rangle$, then*

$$f(x) = \sum_{i,j} T_{ij} x_i x_j$$

so we may take differentials,

$$df(x) = \sum_{i=1}^{n} \left(\sum_{j \neq i} (T_{ij} + T_{ji}) x_j + 2 T_{ii} x_i \right) dx^i = 2 \sum_{i,j} T_{ij} x_j dx^i$$

and the constraint region is defined by the function $g(v) = \sum x_i^2$, so $dg(x) = 2 \sum x_i dx^i$. Since S^{n-1} is compact, extrema exist, and at this extremum point x there is λ such that

$$\sum_{i,j} T_{ij} x_j dx^i = \lambda \sum x_i dx^i$$

and so $Tx = \lambda x$, implying that x is an eigenvector of T, with eigenvalue λ. If V is the orthogonal complement of the span of x, then $T(V) \subset V$, because if $\langle w, v \rangle = 0$, then $\langle Tw, v \rangle = \langle w, Tv \rangle = \lambda \langle w, v \rangle = 0$, and $T : V \to V$ is still self-adjoint. This implies that if V is non-trivial, we may find another eigenvector in V. Continuing this process, we find a sequence of orthogonal eigenvectors $v_1, \ldots, v_n$ for T, which diagonalizes T.

The fact that all orthogonal matrices can be diagonalized gives us a general decomposition results for matrices in $GL(n)$, which is analogous to the fact that a nonzero complex number can be written as rz where $r > 0$ and $|z| = 1$. In general, the result is known as the polar decomposition theorem.

Theorem 4.4. *Any invertible matrix M can be uniquely decomposed as $M_1 M_2$, where M_1 is orthogonal and M_2 is positive definite. In fact, $GL(n)$ is homeomorphic to $O(n) \times \mathbf{R}^{n(n+1)/2}$, where we identify $\mathbf{R}^{n(n+1)/2}$ with the space of self adjoint matrices.*

Proof. The uniqueness is easy, because if $M = M_1 M_2 = M_1' M_2'$, then

$$(M_2')^2 = M_2'(M_1')^{-1} M_1' M_2' = M^t M = M_2 M_1^{-1} M_1 M_2 = M_2^2$$

But since M_2 and M_2' are positive definite, this fact implies $M_2 = M_2'$, and therefore that $M_1 = M_1'$. Existence is a bit more tricky. Given $M \in GL(n)$, MM^t is self adjoint and positive definite, so that there is a positive definite matrix N such that $MM^t = N^2$ (If MM^t was diagonal, then N would be easy to find, because we could take square roots along the diagonal, but in general we can just diagonalize M and then take square roots). Then $N^{-1}M$ is orthogonal, because N^{-1} is also positive definite, so

$$(N^{-1}M)^t (N^{-1}M) = M^t N^{-2} M = M^t (MM^t)^{-1} M = I$$

and so $M = N(N^{-1}M)$.

Now we claim the association $M \mapsto (M_1, M_2)$ is continuous. Since $M_2 = M_1^{-1} M$, it suffices to prove that the map $M \mapsto M_1$ is continuous. Suppose that $M^1, M^2, \ldots$ is a sequence in $GL(n)$ converging to $GL(n)$. Then since $O(n)$ is compact, some subsequence $M_1^{i_k}, M_1^{i_2}, \ldots$ converges to an orthogonal matrix N. But now $M_2^{i_k} = (M_1^{i_k})^{-1} M^{i_k}$ converges as well, to $N^{-1}M$, and the space of self adjoint matrices is closed so $N^{-1}M$ is self adjoint. Since $M^{i_k} = M_1^{i_k} M^{i_k}$ converges to $N(N^{-1}M)$, it follows that $N = M_1$ and $N^{-1}M = M_2$, hence $M^{i_k} \to M_1$. But now we conclude that in general $M_1^k \to M_1$ because this argument can be adapted to show every subsequence contains a further subsequence converging to M_1. Since the map $(M_1, M_2) \to M_1 M_2$ is continuous, this verifies the decomposition is actually a homeomorphism. $\square$

4.3   Tensors

We now consider tensors, which are a generalization of covectors as linear functionals to multilinear functionals. Recall that if V is a vector space, then we let $T^n(V)$ denote the space of maps $f : V^n \to \mathbf{R}$ which are *multilinear*, or linear in each variable. Of course, $T^1(V)$ is just our ordinary dual

space V^*. If T is an n tensor, and S is an m tensor, then the *tensor product* $(T \otimes S)(v, w) = T(v)S(w)$ is an $n + m$ tensor, and the tensor product is an associative, distributive operation which turns the space $\bigoplus T^n(V)$ into a graded algebra (by convention, if we let $T^0(V)$ denote the real numbers, and $\lambda \otimes T = \lambda T$, then this algebra has an identity). In particular, if $f_1, \ldots, f_n$ is a basis for V^*, then a basis for $T^m(V)$ is given by the elements $f_{i_1} \otimes f_{i_2} \otimes \cdots \otimes f_{i_m}$, with $i_1, \ldots, i_m \in [n]$, so $T^m(V)$ has dimension n^m.

Given a linear map $f : V \to W$, we can define a map $f^* : T^n(W) \to T^n(V)$ by setting $f^*(\omega)(v_1, \ldots, v_n) = \omega(f(v_1), \ldots, f(v_n))$. Thus we obtain a continuous tensor functor, so for each vector bundle ξ, we can find the bundle $T^n(\xi)$ of n tensors over ξ. In particular, we obtain the bundle $T^n(TM)$ of tensors over the tangent bundle, and since a basis at M_p is given by $dx^1(p), \ldots, dx^n(p)$ for some coordinate system x around p, a basis for $T^m(TM)$ is given by $dx^{i_1}(p) \otimes \cdots \otimes dx^{i_m}(p)$. Thus an m tensor field ω on M can be locally written as

$$\omega = \sum f_{i_1 \ldots i_m} dx^{i_1} \otimes \cdots \otimes dx^{i_m}$$

and ω is continuous/smooth if and only if the functions $f_{i_1 \ldots i_m}$ are. If

$$\sum f_{i_1 \ldots i_m} dx^{i_1} \otimes \cdots \otimes dx^{i_m} = \sum g_{j_1 \ldots j_m} dy^{j_1} \otimes \cdots \otimes dy^{j_m}$$

Then we have the horrendous conversion formula

$$f_{i_1 \ldots i_m} = \sum g_{j_1 \ldots j_m} \frac{\partial y^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial y^{j_m}}{\partial x^{i_m}}$$

where the products here are just ordinary products of functions.

Though not as often occuring as covariant tensors, contravariant tensor fields are defined similarily, as multilinear functionals operating on V^*, and the space of n contravariant tensors is denoted $T_n(V)$. $T_1(V)$ is just V^{**}, and is naturally isomorphic to V, and we can take tensor products to obtain all elements of $T_n(V)$, so if V has a basis $v_1, \ldots, v_n$, then a basis for $T_m(V)$ is given by $v_{i_1} \otimes \cdots \otimes v_{i_m}$. Given $f : V \to W$, we obtain $f^* : T_n(V) \to T_n(W)$ in the obvious way, and this functor gives us $T_n(\xi)$ for any vector bundle ξ. Putting covariant and contravariant tensors gives us the space $T_n^m(V)$ of mixed tensors, multilinear functions over m copies of V and n copies of V^*, and we can therefore form $T^m(\xi)$. Over the tangent bundle

TM, we can form the space $T_n^m(TM)$, and vector fields over these sets can be locally written as

$$\sum f_{i_1\dots i_n}^{j_1\dots j_m} dx^{i_1} \otimes dx^{i_n} \otimes \frac{\partial}{\partial j_1} \otimes \cdots \otimes \frac{\partial}{\partial j_m}$$

which transforms in the coordinates in the obvious way. We call an element of $T_n^m(V)$ an (m,n) tensor.

There are all kinds of algebraic tricks one can do on mixed tensors. Consider the map $V \mapsto \mathrm{End}(V)$, where if $f : V \to W$ is an isomorphism, then we can define $f_* : \mathrm{End}(V) \to \mathrm{End}(W)$ by $f_*(T) = f \circ T \circ f^{-1}$. Then the map $V \mapsto \mathrm{End}(V)$ defines a functor from the category of vector spaces with morphisms restricted to isomorphisms to itself, and this is enough to make $\mathrm{End}(\xi)$ into a bundle for each bundle ξ, though now only *equivalences* $f : \xi \to \phi$ extend to maps $f_* : \mathrm{End}(\xi) \to \mathrm{End}(\phi)$. Now the mixed tensor space $T_1^1(V)$ is isomorphic to $\mathrm{End}(V)$, where $S : V \to V$ corresponds to the tensor $\omega_S(v,\lambda) = \lambda(S(v))$. If $f : V \to W$ is an isomorphism, and $S : W \to W$ is an endomorphism, then

$$\omega_{f_*S}(v,\lambda) = \lambda((f_*S)(v)) = \lambda((f \circ S \circ f^{-1})(v)) = (f^*\omega_S)(v,\lambda)$$

This means that $\mathrm{End}(\xi)$ is naturally equivalent to $T_1^1(\xi)$, and so any operation we can perform on $\mathrm{End}(V)$ can be transfered naturally to $T_1^1(V)$. For instance, we can define the *trace*, or *contraction* of a tensor $\omega \in T_1^1(V)$ to be the trace of the endomorphism S corresponding to ω. If we fix a basis, and write $\omega = \sum a_i^j e_i^* \otimes e_j$, then $\mathrm{trace}(\omega) = \sum a_i^i$. In particular, a smooth vector field on $T_1^1(TM)$ can be contracted pointwise to produce a smooth function. Similarly, we can always contract a bottom and top index on an (n,m) tensor to form a $(n-1,m-1)$ tensor. We note that unless a Riemannian metric is fixed, contravariant indices and covariant indices work very differently. For instance, the bilinear forms have two covariant indices, whereas matrices have one contravariant and one covariant index. If we have a $(1,1)$ tensor a_i^j, we can consider contravariant and covariant eigenvectors, those vectors v^i such that $a_i^j v^i = \lambda v^j$, or those covectors b_j such that $b_j a_i^j = \lambda b_i$. It does not make sense to talk about the eigenvectors and eigenvalues of $(0,2)$ or $(2,0)$ tensors, independently of some Riemannian metric. However, given some covariant 2 tensor ω and Riemannian metric, we can consider eigenvector fields X such that $\omega(X,Y) = \lambda\langle X,Y\rangle$, which in coordinates, if ω has coordinates A_{ij}, is equivalent to solving $A_{ij}X^i = \lambda g_{jk}X^k$.

4.4 The Derivational Viewpoint

We previously identified the smooth vector fields on a manifold M with the derivations on $C^\infty(M)$. We now identify tensor fields as certain operators on vectors fields. Given an n tensor field ω and n vector fields $X_1, \ldots, X_n$, we can define a smooth function

$$\omega(X_1, \ldots, X_n)(p) = \omega(p)(X_1(p), \ldots, X_n(p))$$

and ω operates as a multilinear map on $\Gamma(TM)^n$, where the map is linear not only over real numbers, but also C^∞ functions, since the map is 'defined pointwise'. We find that all such multilinear maps are just tensor fields in disguise.

Theorem 4.5. *If $\Omega : \Gamma(TM)^n \to C^\infty(M)$ is a $C^\infty(M)$ multilinear map, there is a unique smooth n tensor field ω on M such that $\Omega(X_1, \ldots, X_n) = \omega(X_1, \ldots, X_n)$.*

Proof. We first note that if $X_k = Y_k$ in a neighbourhood of p, then

$$\Omega(X_1, \ldots, X_n)(p) = \Omega(Y_1, \ldots, Y_n)(p)$$

To see this, we find a smooth bump functions $f_1, \ldots, f_n$ equal to one in a neighbourhood of p such that $f_k X_k = f_k Y_k$, and then

$$\begin{aligned}
\Omega(X_1, \ldots, X_n)(p) &= f_1(p) \ldots f_n(p) \Omega(X_1, \ldots, X_n)(p) \\
&= \Omega(f_1 X_1, \ldots, f_n X_n)(p) \\
&= \Omega(f_1 Y_1, \ldots, f_n Y_n)(p) \\
&= f_1(p) \ldots f_n(p) \Omega(Y_1, \ldots, Y_n)(p) \\
&= \Omega(Y_1, \ldots, Y_n)(p)
\end{aligned}$$

Thus Ω is locally defined. To show that $\Omega(X_1, \ldots, X_n)(p)$ depends only on $X_1(p), \ldots, X_n(p)$, it suffices to show that $\Omega(X_1, \ldots, X_n)(p) = 0$ if $X_k(p) = 0$, for some k. If (x, U) is a coordinate system around p, then we can write

$$X_k = \sum b_i \frac{\partial}{\partial x^i}$$

But then

$$\Omega(X_1, \ldots, X_k, \ldots, X_n)(p) = \sum b_i(p) \Omega\left(X_1, \ldots, \frac{\partial}{\partial x_i}\right) = 0$$

which shows that Ω is defined pointwise. We can now set

$$\omega(p)(v_1,\ldots,v_n) = \Omega(X_1,\ldots,X_n)(p)$$

where $X_k(p) = v_k$, and we have verified this is well defined. ω is smooth because

$$\omega = \sum \omega_{i_1\ldots i_n} dx^{i_1} \otimes \cdots \otimes dx^{i_n}$$

then $\omega_{i_1\ldots i_n} = \Omega(dx^{i_1},\ldots,dx^{i_n})$ is a smooth function. $\qquad\square$

Because of this, we do not distinguish tensor fields from their corresponding operators on vector fields. Similar results hold for contravariant tensors, which operate on covector fields.

4.5   The Classical Viewpoint

The classical viewpoint of vectors in the tangent bundle, and tensors in general, is obsessed with the coefficients associated to these objects and the way they transform with respect to coordinate systems. We previous mentioned that for a coordinate chart (x, U) an (n, m) tensor field ω may be written on U as

$$\omega = \sum_{\substack{i_1,\ldots,i_n \\ j_1,\ldots,j_m}} a_{i_1,\ldots,i_n}^{j_1,\ldots,j_m} dx^{i_1} \otimes \cdots \otimes dx^{i_n} \otimes \frac{\partial}{\partial x^{j_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{j_m}}$$

If (y, V) is another coordinate system, then we also have

$$\omega = \sum_{\substack{i_1,\ldots,i_n \\ j_1,\ldots,j_m}} b_{i_1,\ldots,i_n}^{j_1,\ldots,j_m} dy^{i_1} \otimes \cdots \otimes dy^{i_n} \otimes \frac{\partial}{\partial y^{j_1}} \otimes \cdots \otimes \frac{\partial}{\partial y^{j_m}}$$

and on $U \cap V$, we have the monstrous transformation rule

$$a_{i_1,\ldots,i_n}^{j_1,\ldots,j_m} = \sum_{\substack{\alpha_1,\ldots,\alpha_n \\ \beta_1,\ldots,\beta_m}} b_{\alpha_1,\ldots,\alpha_n}^{\beta_1,\ldots,\beta_m} \frac{\partial y^{\alpha_1}}{\partial x^{i_1}} \cdots \frac{\partial y^{\alpha_n}}{\partial x^{i_n}} \frac{\partial x^{j_1}}{\partial y^{\beta_1}} \cdots \frac{\partial x^{j_m}}{\partial y^{\beta_m}}$$

Classically, one *defines* a (n, m) tensor on a k dimensional manifold as an assignment of k^{n+m} functions to each chart (x, U), defined on U, such that the transformation above is satisfied on $U \cap V$, where (y, V) is another coordinate system. An *invariant* if a $(0, 0)$ tensor, or a smooth function on the entire manifold. Here's an example.

Example. *A classical differential geometer will tells you that the Kronecker delta function δ_i^j is a tensor. What they mean by this is, if we assign the same n^2 functions δ_i^j to every coordinate system, then for any two coordinate systems x and y,*

$$\delta_i^j = \frac{\partial x^j}{\partial x^i} = \sum \frac{\partial x^j}{\partial y^\alpha} \frac{\partial y^\alpha}{\partial x^i} = \sum \delta_\alpha^\beta \frac{\partial x^j}{\partial y^\beta} \frac{\partial y^\alpha}{\partial x^i}$$

so, mystically, these coordinate systems really do define a smooth tensor field A over $T_1^1(TM)$. To figure out what A really is, we calculate

$$A(X, \omega) = \sum \omega_i X^i = \omega(X)$$

so A is just the tensor which operates as evaluation of a functional on a vector field. In terms of the identification of $End(TM)$ with $T_1^1(TM)$, A just corresponds to the identity map.

It is a useful skill to be able to translate the classical language into modern language, and vice versa, and be able to prove things in both settings. Thus in the following, we state theorems in their classical context, prove them by classical means, then reexpress them into modern theorems and prove them in this context.

Theorem 4.6. *If $\sum v_i \mu^i$ is an invariant, for any covector field $\sum v_i dx^i$, then $\sum \mu^i \frac{\partial}{\partial x_i}$ is a well defined vector field.*

Proof. For any covector field $\sum v_i dx^i = \sum (v')_i dy^i$, we have

$$\sum \mu^i v_i = \sum \mu'^i v_i'$$

and we have the relation $(v')_i = \sum v_j \frac{\partial x^j}{\partial y_i}$. If we choose $v_\alpha = 1$, and $v_i = 0$ otherwise, we find

$$\mu^\alpha = \sum \mu^i v_i = \sum_i \mu'^i v_i' = \sum_{i,j} \mu'^i \frac{\partial x^j}{\partial y_i} v_j = \sum_i (\mu')^i \frac{\partial x^\alpha}{\partial y_i}$$

so

$$\sum_i \mu^i \frac{\partial}{\partial x^i} = \sum_{i,j,k} (\mu')^k \frac{\partial x^i}{\partial y_k} \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j} = \sum_{j,k} (\mu')^k \frac{\partial y^j}{\partial y^k} \frac{\partial}{\partial y^j} = \sum_k (\mu')^k \frac{\partial}{\partial y^k}$$

is a well defined vector field, and is uniquely determined since we can determine the value of μ^i at any point x by choosing a smooth vector field v such that $v_i(x) = 1$, $v_j(x) = 0$, and then $\mu^i(x) = \sum \mu^i(x)v_i(x)$. In modern terms, the functions v_i correspond to a $C^\infty(M)$ linear map $T\mu = \sum \mu^i v_i$ from $\Gamma(T^*M)$ to $C^\infty(M)$, which therefore corresponds to a unique smooth vector field v, and the v_i are obviously the coordinates of v. $\qquad\square$

Chapter 5

Lie Derivatives and Differential Equations

So far, we've managed to construct a manifold, and a coordinate-independant way to measure the rate of change of functions between manifolds. However, a problem results when we wish to measure the rate of change of a rate of change. But it's difficult to measure the rate of change of vector field X, because we cannot compare X_p and X_q for $p \neq q$, as the elements lie in different vector spaces. We can of course switch to a coordinate system for close enough p and q, associating with each X a function $F(x) : \mathbf{R}^n \to \mathbf{R}^n$ we can differentiate, but this doesn't really apply to a global perspective in any way, and one cannot patch together these maps so that the derivative of a vector field is a vector field. On their own, we shouldn't expect there to be a coordinate independent way to measure the rate of change of these vector fields, because if we can choose coordinate arbitrarily, then the rate of change of X can take the form of any derivative we like. Thus we must add additional structure to this scenario. In this case, a direction with which to differentiate. We will define the *Lie derivative* of a vector field X in the direction of another vector field Y to be the vector field $L_X(Y)$. To do this, we must use the theory of differential equations to represent vector fields as motions around the manifold.

Theorem 5.1. *If $f : \mathbf{R}^n \to \mathbf{R}^n$ is a locally Lipschitz function, then at any point $p \in \mathbf{R}^n$, there is a suitably small ε and a unique curve $x : (-\varepsilon, \varepsilon) \to \mathbf{R}^n$ with $x(0) = p$, and $x' = f \circ x$, and if we extend x to have the maximum possible interval domain, then either x is defined everywhere, or x is defined on some*

interval (a, b), and $\lim_{t \to a} |x(t)| = \lim_{t \to b} |x(t)| = \infty$.

The theorem has the following interpretation in the theory of differentiable manifolds. Given any differentiable vector field $X : M \to TM$, and $p \in M$, there is a suitably small ε and a unique differentiable curve $x : (-\varepsilon, \varepsilon) \to M$ such that $x'(t) = X_{x(t)}$. We shall call a curve satisfying this equation an **integral curve** for X. this results obviously follows because we can switch to any particular coordinate system. The unique curves imply that we can define a function on some open subset of $M \times \mathbf{R}$ by $\phi_t(p) = x(t)$, where x is the unique integral curve for X with $x(0) = p$, where we assume such an x exists in the definition of the function. If $\phi_{t+h}(p)$, and $\phi_t(\phi_h(p))$ are all well defined, then $\phi_t(\phi_h(p)) = \phi_{t+h}(p)$ because if $x : (-\varepsilon_0, \varepsilon_0) \to M$ is an integral curve satisfying $x(0) = p$, and $y : (-\varepsilon_1, \varepsilon_1) \to M$ is an integral curve with $y(0) = x(h)$, then the uniqueness theorem tells us that $y(u) = x(h + u)$ where these functions are defined, because $z(u) = x(h + u)$ is also an integral curve with $z(0) = x(h)$. ϕ is defined on an open submanifold of $M \times \mathbf{R}$, and it is a very difficult theorem of functional analysis to show that if X is a C^k vector field, then ϕ is C^k. We will take this for granted in the sequel. Since $\phi_{-t} = \phi_t^{-1}$ on a small enough neighbourhood of every domain, we conclude that each ϕ_t is 'almost' a diffeomorphism, and is certainly a local diffeomorphism. For compact manifolds, we can obtain a global family of diffeomorphisms.

Theorem 5.2. *If X is a vector field with compact support, then ϕ is defined everywhere.*

Proof. If K is the support of X, then ϕ_t is defined on K^c for all t, because if $p \notin K$, then $x(p) = p$ is an integral curve of X. If we cover K by finitely many coordinate charts $(x_1, U_1), \ldots, (x_n, U_n)$, such that ϕ_t is defined for $|t| < \varepsilon_i$ on U_i, then we have define ϕ_t on M for all $|t| < \min(\varepsilon_1, \ldots, \varepsilon_n)$, and we can then define ϕ for arbitrarily large real numbers by considering the composition $\phi_t^n = \phi(nt)$. $\square$

The reason we can solve all nonzero vector fields are locally the same up to a change in coordinates.

Theorem 5.3. *If X is a differentiable vector field on a manifold M with $X_p \neq 0$, then there is a coordinate chart x around p such that in a neighbourhood of p, $X_p = \partial_1$.*

Proof. It obviously suffices to prove this for $p = 0$, and $M = \mathbf{R}^n$. We may also start with a coordinate chart such that $X_0 = \partial_1$. Consider the induced diffeomorphism ϕ from the vector field X, and take the map $\alpha : (-\varepsilon, \varepsilon) \times U \to \mathbf{R}^n$, where $U \subset \mathbf{R}^{n-1}$ is a neighbourhood defined by $\alpha(t, x) = \phi_t(0, x)$. Clearly the map is C^∞, and if $f : x(U) \to \mathbf{R}$ is arbitrary, then

$$\alpha_* \left(\frac{\partial}{\partial x^i}(0, x) \right) (f) = \frac{\partial (f \circ \alpha)}{\partial x^i}(0, x) = \frac{\partial f}{\partial x_i}(x)$$

Thus $\alpha_* \partial_i(0, x) = \partial_i(0, x)$. We also calculate

$$\alpha_* \left(\frac{\partial}{\partial t}(0, x) \right) (f) = \lim_{h \to 0} \frac{f(\phi_h(x)) - f(x)}{h} = X_x(f)$$

so $\alpha_*(\partial_t(0, x)) = X_x$. We conclude that α_* is nonsingular in a neighbourhood of $(0, x)$, and so locally α is invertible, and the inverse, which we call y, is a coordinate system. But we know that

$$\frac{\partial}{\partial y^1} = \alpha_* \left(\frac{\partial}{\partial t} \right) = X_x$$

which completes the proof. $\qquad \square$

Note that here we have used the fact that

$$X_p f = \lim_{h \to 0} \frac{f(\phi_h(p)) - f(p)}{h}$$

This limit suggests that the translation maps ϕ_h can be used to measure the rate of change along the vector X_p. As foreshadowed at the beginning of the chapter, we can use this process to measure the rate of change of the many objects defined on a differentiable manifold. First, we switch notations, and denote $X(f)$ by $L_X(f)$, and call it the **Lie derivative** of f along X. Using this notation, we can also consider the derivatives of covariant vector fields ω, obtaining a new covariant vector field

$$(L_X \omega)(p) = \lim_{t \to 0} \frac{(\phi_t^* \omega)(p) - \omega(p)}{t}$$

So that $(L_X \omega)(\phi_t(p)) = \omega(p) + t(L_X \omega)(p) + o(t)$. Similarily, and most importantly, we consider the derivatives of other vector fields, defined by

$$(L_X Y)(p) = \lim_{h \to 0} \frac{Y_p - ((\phi_h)_* Y)_p}{h}$$

94

Where the order of terms is switched because $((\phi_h)_*Y)_p$ is really looking at the point on the vector field at $Y_{\phi_{-h}}$, and therefore backwards in time rather than forwards. We can define $(\phi_h^*Y)_p = ((\phi_{-h})_*Y)_{\phi_h(p)}$, and then we can interchange the equation, but we'll stick with this definition.

Theorem 5.4. *If $f, \omega\ X$, and Y are given, then*

   (i) $L_X(f\omega) = (L_Xf)\omega + f(L_X\omega)$.

   (ii) $L_X(fY) = (L_Xf)X + f(L_XY)$.

   (iii) $L_X(\omega(Y)) = \omega(L_XY) + \omega(L_XY)$.

Proof. We write

$$
\begin{aligned}
L_X(f\omega)(p)(v) &= \lim_{h\to 0} \frac{\phi_h^*(f\omega)(p)(v) - (f\omega)(p)(v)}{h} \\
&= \lim_{h\to 0} \frac{f(\phi_h(p))\omega(\phi_h(p))((\phi_h)_*(v)) - f(p)\omega(p)(v)}{h} \\
&= (FG)'(0)
\end{aligned}
$$

Where $F(t) = f(\phi_t(p))$, and $G(t) = \omega(\phi_t(p))((\phi_t)_*(v))$. Since $F(0)' = (L_Xf)(p)$, and $G(0)' = (L_X\omega)(p)(v)$, we find that

$$
L_X(f\omega)(p)(v) = f(p)(L_X\omega)(p)(v) + (L_Xf)(p)\omega(p)(v)
$$

and this completes the proof of (i). The propositions (ii) and (iii) are proved in essentially the same way. $\square$

The Lie derivative is obviously linear, and we can obtain a formula for the derivatives in the coordinates. First, let's compute the Lie derivative for a covariant vector field. If

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

Then

$$(L_X dx^i)(p)\left(\frac{\partial}{\partial x^j}\right) = \lim_{h\to 0} \frac{(\phi_h^* dx^i)(p)\left(\frac{\partial}{\partial x^j}\right) - dx^i(p)\left(\frac{\partial}{\partial x^j}\right)}{h}$$

$$= \lim_{h\to 0} \frac{dx^i(\phi_h(p))\left((\phi_h)_*\frac{\partial}{\partial x^j}\right) - \delta_i^j}{h}$$

$$= \lim_{h\to 0} \frac{1}{h}\left(\left.\frac{\partial x^i \circ \phi_h}{x^j}\right|_{\phi_h(p)} - \delta_i^j\right)$$

$$= \left.\frac{\partial^2 (x^i \circ \phi_h)}{\partial h \partial x^j}\right|_p = \left.\frac{\partial a^i}{\partial x^j}\right|_p$$

Thus if $\omega = \sum b_i dx^i$, then

$$L_X \omega = \sum \left(b_i \frac{\partial a^i}{\partial x^j} + a_i \frac{\partial b_i}{\partial x^j}\right) dx^j$$

To calculate $(L_X Y)$ in coordinates, let $X = \sum a^i \frac{\partial}{\partial x^i}$, and $Y = \sum b^i \frac{\partial}{\partial x^i}$. Using (iii) from the last proposition, we find

$$0 = (L_X \delta_i^j) = L_X \left(dx^i \left(\frac{\partial}{\partial x^j}\right)\right)$$

$$= dx^i \left(L_X \frac{\partial}{\partial x^j}\right) + (L_X dx^i)\left(\frac{\partial}{\partial x^j}\right)$$

$$= dx^i \left(L_X \frac{\partial}{\partial x^j}\right) + \frac{\partial a^i}{\partial x^j}$$

Hence $\left(L_X \frac{\partial}{\partial x^j}\right) = \sum -\frac{\partial a^i}{\partial x^j}\frac{\partial}{\partial x^i}$, and so

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

We can use this theorem to find a much simpler expression for $L_X Y$. Note that if f is a function, then $Y(f)$ is also a function, so we can consider $X(Y(f))$.

Theorem 5.5. $(L_X Y)(f) = X(Y(f)) - Y(X(f))$.

96

Proof. If $X = \sum a_i \frac{\partial}{\partial x^i}$, and $Y = \sum b_i \frac{\partial}{\partial x^i}$, then

$$
\begin{aligned}
X(Y(f)) - Y(X(f)) &= \sum a_i \frac{\partial Y(f)}{\partial x^i} - b_i \frac{\partial X(f)}{\partial x^i} \\
&= \sum a_i \frac{\partial b_j}{\partial x_i} \frac{\partial f}{\partial x_j} - a_i b_j \frac{\partial^2 f}{\partial x_i \partial x_j} - b_i \frac{\partial a_j}{\partial x_i} \frac{\partial f}{\partial x_j} + b_i a_j \frac{\partial^2 f}{\partial x_i \partial x_j} \\
&= \sum \left(a_i \frac{\partial b_j}{\partial x_i} - b_i \frac{\partial a_j}{\partial x_i} \right) \frac{\partial f}{\partial x_j} \\
&= (L_X Y)(f)
\end{aligned}
$$

$\square$

Chapter 6

Forms and Integration

The change of variables formula for integration says that if y is sufficiently smooth, injective function, then

$$\int_A (g \circ y)|\det(Dy)| = \int_{y^{-1}(A)} g$$

We now wish to define integration of certain objects on manifolds. Since we have no natural choice of a measure on a manifold, the objects ω we integrate cannot be functions, this object must contain additional information which substitutes for the existence of a measure. This information will be carried by how the function changes with respect to coordinates. If, in a certain coordinate system (x, U), the object ω is associated with a certain integrable function ω_x, we would like to define

$$\int_U \omega = \int_{x(U)} \omega_x$$

In order for this definition to make sense, in another coordinate system y, the associated function ω_y would then have the satisfy

$$(\omega_y \circ y \circ x^{-1})|\det(D(y \circ x^{-1}))| = \omega_x$$

The objects associated with this transformation law will be known as *absolute tensors*. However, if our manifold M has a fixed orientation, then we can restrict ourselves to oriented charts, and we could then consider the more general objects for which

$$(\omega_y \circ y \circ x^{-1})\det(D(y \circ x^{-1})) = \omega_x$$

This transformation law leads natural to the discussion of alternating tensors, also known as differential forms. Of course, it is also natural to integrate over submanifolds of your manifold. In this case, we need objects which satisfy the transformation law above for any coordinate system on any submanifold of space. These objects will be distinct from the objects we can integrate on the entirety of the manifold, and more generally, on an n dimensional manifold we will have objects corresponding to integrations over submanifolds of each dimension $m \leqslant n$. For the beginning parts of this chapter, we introduce these objects abstractly, then return to integration once the theory has been fully established. But keep in mind that the main reason for the introduction of these objects is that they have a transformation law that allows us to integrate in a coordinate independant manner.

6.1   Alternating Tensors

An n tensor T on a vector space V is called **alternating** if $T(v_1,\ldots,v_n) = 0$ whenever $v_i = v_j$. There is another description. For $i \neq j$,

$$0 = T(v_1,\ldots,v_i + v_j,\ldots,v_i + v_j,\ldots,v_n)$$
$$= T(v_1,\ldots,v_i,\ldots,v_i,\ldots,v_n) + T(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_n)$$
$$+ T(v_1,\ldots,v_j,\ldots,v_i,\ldots,v_n) + T(v_1,\ldots,v_j,\ldots,v_j,\ldots,v_n)$$
$$= T(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_n) + T(v_1,\ldots,v_j,\ldots,v_i,\ldots,v_n)$$

Rearranging, we find that

$$T(v_1,\ldots,v_j,\ldots,v_i,\ldots,v_n) = T(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_n)$$

A tensor satisfying this equation is known as *skew symmetric*. Of course, over the real numbers, these two definitions are equivalent. More generally, if $\sigma \in S_n$, and we define the right group action

$$(T\sigma)(v_1,\ldots,v_n) = T(v_{\sigma(1)},\ldots,v_{\sigma(n)})$$

on the space of tensors, then an alternating tensor is precisely one for which $T\sigma = \text{sgn}(\sigma)T$ for all permutations σ. We will let $\Omega^k(V)$ denote the vector space of all alternating tensors on V. If $f : V \to W$ is linear,

then f^* maps $\Omega^k(W)$ into $\Omega^k(V)$, since then if $v_i = v_j$, $f(v_i) = f(v_j)$, so $(f^*T)(v_1,\ldots,v_n) = T(f(v_1),\ldots,f(v_n)) = 0$.

There is a natural way to convert any tensor into an alternating tensor. If T is an n tensor, we define

$$\text{Alt}(T) = \frac{1}{n!} \sum_{\sigma \in S_n} \text{sgn}(\sigma) T\sigma$$

It then follows that

$$\text{Alt}(T)\tau = \frac{1}{n!} \sum_{\sigma \in S_n} \text{sgn}(\sigma)(T\sigma\tau) = \frac{\text{sgn}(\tau)}{n!} \sum_{\sigma \in S_n} \text{sgn}(\sigma\tau)(T\sigma\tau) = \text{sgn}(\tau)\text{Alt}(T)$$

so $\text{Alt}(T)$ is an alternating tensor. The term $1/n!$ is included to normalize the alternator, so that Alt is a projection of $T^n(V)$ onto $\Omega^n(V)$. The analogous operation to the tensor product on $T^n(V)$ is the *wedge product* of alternating n and m tensors ω and η, defined to be

$$\omega \wedge \eta = \frac{(n+m)!}{n!m!}\text{Alt}(\omega \otimes \eta)$$

The operation is bilinear, and *anticommutative*, in the sense that

$$\eta \wedge \omega = (-1)^{nm}(\omega \wedge \eta)$$

This is most easy to see if $\eta = dx^{i_1} \otimes \cdots \otimes dx^{i_n}$ and $\omega = dx^{j_1} \otimes \cdots \otimes dx^{j_m}$, the most tricky part to prove being that the permutation swapping the first n coordinates with the latter m coordinates has sign $(-1)^{nm}$ (provable by induction). The product also satisfies

$$f^*(\omega \wedge \eta) = f^*(\omega) \wedge f^*(\eta)$$

The operation is associative. In particular, if $\omega_1,\ldots,\omega_n$ are k_n tensors, then

$$\omega_1 \wedge \cdots \wedge \omega_n = \binom{k_1 + \cdots + k_n}{k_1 \; k_2 \; \ldots \; k_n}\text{Alt}(\omega_1 \otimes \cdots \otimes \omega_n)$$

Thus the wedge product turns $\bigoplus \Omega^k(V)$ into a graded algebra, which has an identity if we let $\Omega^0(V)$ denote the real numbers.

The normalization factor $(n+m)!/n!m!$ included in the definition of the wedge product is included so that if $e_1, \ldots, e_n$ is a basis of the space, then

$$(e_1^* \wedge \cdots \wedge e_n^*)(e_1, \ldots, e_n) = \frac{n!}{n!} \sum_{\sigma \in S_n} (e_1^* \otimes \cdots \otimes e_n^*)\sigma(e_1, \ldots, e_n) = 1$$

More generally,

$$(e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*) \left(\sum a_1^i e_i, \ldots, \sum a_k^i e_i \right) = \det \left(\{ k \text{ by } k \text{ minor of } (a_j^i) \} \right)$$

To simplify notation, if a basis $\{\phi_1, \ldots, \phi_n\}$ for the dual space is fixed and $I = (i_1, \ldots, i_k)$ is a multi-index, then ϕ_I will denote the wedge product of the functionals ϕ_k.

Corollary 6.1. *If $\omega_1, \ldots, \omega_k$ are linear functionals, then $\omega_1 \wedge \cdots \wedge \omega_k \neq 0$ precisely when $\omega_1, \ldots, \omega_k$ are linearly independant.*

Proof. If $\omega_1, \ldots, \omega_k$ are linearly independant, we can extend them to a basis of the space of all linear functionals, and if $e_1, \ldots, e_k$ form part of the dual basis, then

$$(\omega_1 \wedge \cdots \wedge \omega_k)(e_1, \ldots, e_k) = 1$$

Thus $\omega_1 \wedge \cdots \wedge \omega_k \neq 0$. Conversely, if ω_1 is in the span of $\omega_2, \ldots, \omega_k$, then expanding out the product by linearity, we find the wedge product must vanish. $\square$

Because Alt is surjective, we know that $\Omega^k(V)$ is spanned by the ϕ_I. But because the wedge product is anticommutative, the space is still spanned by the subfamily of indices $I = (i_1, \ldots, i_n)$ when $i_1 < i_2 < \cdots < i_n$. This forms a basis for the space of all alternating tensors, because we know $e_I \neq 0$ for any such I, and for distinct families of indices the alternator operates over completely different tensor indices. This means that $\Omega^k(V)$ is $\binom{n}{k}$ dimensional. In particular, $\Omega^n(V)$ is 1 dimensional, which has the following consequence.

Corollary 6.2. *If $e_1, \ldots, e_n$ is a basis for V, $\omega \in \Omega^n(V)$, then*

$$\omega \left(\sum a_{1j} e_j, \ldots, \sum a_{nj} e_j \right) = \det((a_{ij}))\omega(e_1, \ldots, e_n)$$

Proof. The determinant of the coefficients is clearly an n alternating tensor, and so disagrees with the left hand side of the equation only by a constant multiple. This multiple is precisely $\omega(e_1, \ldots, e_n)$. $\square$

Because of the corollary above, if $e_1,\ldots,e_n$ is a basis for V, then for a nonzero alternating n tensor ω, $\omega(e_1,\ldots,e_n) \neq 0$. This gives an interesting connection between alternating tensors and orientations on vector spaces. If ω is an n alternating tensor on V, then we can define an orientation μ on V by letting $(e_1,\ldots,e_n) \in V$ precisely when $\omega(e_1,\ldots,e_n) > 0$. Conversely, every orientation is induced by some alternating tensor – i.e. the determinant tensor given by an oriented basis.

6.2 Differential Forms

We now apply the theory of alternating tensors to manifolds. Given a manifold M, we can now construct the subbundle $\Omega^k(TM)$ of $T^k(TM)$ consisting of alternating tensors on TM. A section of this bundle is known as a **differential form**, the space of all such sections forming a $C^\infty(M)$ module. By taking wedge products locally, we can take the direct sum of bundles $\Omega(TM) = \bigoplus \Omega^k(TM)$ into a bundle of algebras, and the sections of this space form a $C^\infty(M)$ algebra. Every differential form can locally be written uniquely as $\sum a_I dx^I$, where I ranges over monotonic multi-indexes. Notice that any element of $\Omega^n(TM)$ is written in coordinates in terms of a single function $f\,dx^1 \wedge \cdots \wedge dx^n$, and if

$$f\,dx^1 \wedge \cdots \wedge dx^n = g\,dy^1 \wedge \cdots \wedge dy^n$$

Then since $dx^i = \frac{\partial x^i}{\partial y^j} dy^j$, our characterization of n forms on vector spaces guarantees

$$g = \det\left(\frac{\partial y^i}{\partial x^j}\right) f$$

Thus differential forms of order n are precisely the coordinate invariant integrable objects we desired at the beginning of the chapter. We will find that forms of order $< n$ also satisfy the ability to be integrated, but over low dimensional submanifolds rather than the entire space.

Theorem 6.3. *A manifold M^n is orientable precisely when $\Omega^n(TM)$ has a nowhere vanishing section.*

Proof. If ω is a nowhere zero differential n form, we can define an orientation μ by letting $(e_1,\ldots,e_n) \in \mu_p$ if $\omega(p)(e_1,\ldots,e_n) > 0$. In coordinates (x,U), $\omega = f\,dx^1 \wedge \cdots \wedge dx^n$, and we must either have $f > 0$ everywhere, or $f < 0$

everywhere, since ω doesn't vanish. This shows the choice of orientation is locally trivial. Conversely, if M^n is orientable, we can cover M by a family of orientated coordinate charts (x_α, U_α). If φ_α is a partition of unity subordinate to these charts, and we choose a nonvanishing differential form ω_α on U_α, then we can consider the form

$$\omega = \sum \varphi_\alpha \omega_\alpha$$

and then $\omega(p) \neq 0$ naywhere, since for a given oriented basis $e_1, \ldots, e_n$ of M_p, $\omega_\alpha(p)(e_1, \ldots, e_n) > 0$, hence $\omega(p)(e_1, \ldots, e_n) > 0$. $\qquad \square$

6.3 The Differential Operator

A section of $\Omega^0(TM)$, or a differential form of order 0, is just a smooth function on M. If f is such a function, we defined its differential as

$$df = \sum \frac{\partial f}{\partial x^\alpha} dx^\alpha$$

Notice that we can now say that 'the differential of a zero form is a one form'. In this section we extend this definition, defining an operator d taking m forms to $m+1$ forms. If $\omega = \sum a_I dx^I$, then we define

$$d\omega = \sum da_I \wedge dx^I = \sum \frac{\partial a_I}{\partial x^\alpha} dx^\alpha \wedge dx^I$$

The linearity of the one dimensional differential immediately extends to this generalized differential. Secondly, if $\omega = \sum a_I dx^I$ is an n form, and $v = \sum b_J dx^J$ is an m form, then

$$\begin{aligned} d(\omega \wedge v) &= \sum d(a_I b_J) \wedge dx^I \wedge dx^J \\ &= \sum a_I db_J \wedge dx^I \wedge dx^J + b_J da_I \wedge dx^I \wedge dx^J \\ &= (-1)^n \omega \wedge dv + d\omega \wedge v \end{aligned}$$

Thirdly, because mixed partials are equal, we find $d^2\omega = 0$, since if $\omega = \sum a_I dx^I$, then

$$d(d\omega) = d\left(\frac{\partial a_I}{\partial x^\alpha} dx^\alpha \wedge dx^I \right) = \frac{\partial^2 a_I}{\partial x^\beta \partial x^\alpha} (dx^\beta \wedge dx^\alpha) \wedge dx^I$$

and the coefficient corresponding to $\partial^2 a_I / \partial x^\beta \partial x^\alpha$ cancels out with the coefficient corresponding to $\partial^2 a_I / \partial x^\alpha \wedge x^\beta$.

103

Example. *One nice thing about the differential is it generalizes the differential operators found in three dimensional vector calculus. If f is a function, then*

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz$$

carries exactly the same information as the gradient ∇f. If $\omega = f_1 dx + f_2 dy + f_3 dz$ is a one form, then

$$d\omega = \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right) dx \wedge dy + \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z} \right) dy \wedge dz + \left(\frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x} \right) dz \wedge dx$$

carries the same imformation as the curl of the vector field (f_1, f_2, f_3). If $\omega = f_1 dy \wedge dz + f_2 dz \wedge dx + f_3 dx \wedge dy$ is a two form, then

$$d\omega = \frac{\partial f_1}{\partial z} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}$$

carries the same information as the divergence of a vector. The fact that $d^2 \omega = 0$ for any form ω implies the classical equations

$$\nabla \cdot (\nabla \times f) = \nabla \times \nabla(f) = 0$$

we shall find that a further study of differential forms give generalizations of the integral theorems of vector calculus: Green's theorem, Gauss's theorem, and Stoke's theorem.

Example. *If the values a^i form the component of a contravariant tensor field with respect to coordinates x, then the values $\partial_j a^i$ do not form the components of a tensor, since if b^i are components with respect to y, then*

$$b^i = \frac{\partial y^i}{\partial x^j} a^j$$

and so

$$\partial_j b^i = \partial_j \left(\frac{\partial y^i}{\partial x^k} a^k \right) = \frac{\partial y^i}{\partial x^k} \partial_j a^k + \frac{\partial^2 y^i}{\partial x_j x_k} a^k$$

This shows that contravariant vector fields cannot be differentiated in the same way that we can with exterior forms. Note also that if a_i are components of a covariant tensor, then $\partial_j a_i - \partial_i a_j$ are components of a covariant tensor, which is the differential.

Theorem 6.4. *If d' is another linear map taking m forms to $m + 1$ forms, satisfying $d(\omega \wedge v) = d\omega \wedge v + (-1)^n \omega \wedge dv$, and $(d')^2 = 0$, with $d'f = df$ when f is a function, then $d' = d$.*

Proof. It suffices to prove this if $\omega = f\,dx^I$. We calculate

$$d'(f\,dx^I) = d'(f \wedge dx^I) = d'f \wedge dx^I - f \wedge d'(dx^I)$$

and since $d'f = df$, we now only need to prove $d'(dx^I) = d(dx^I)$, but these two terms both vanish. $\qquad\square$

This theorem implies that the d operator is coordinate independant, and thus extends to globally defined differential forms. Another way to see this is to give a coordinate independant definition of the operator.

Chapter 7

Lie Groups

A **Lie group** is a group with a differentiable action. That is, it is a group with a differentiable structure such that multiplication and inversion are differentiable functions on the space. It suffices to verify that the map $(x, y) \mapsto xy^{-1}$ is differentiable. We have actually already considered many of the important Lie groups. The simplest Lie groups are abelian, like the group $\mathbf{R}^n$ under addition, the circle group $S^1 = \mathbf{R}/\mathbf{Z}$, and more generally, the torus groups $\mathbf{T}^n = \mathbf{R}^n/\mathbf{Z}^n$. The main noncommutative exmamples occur as matrix groups, like $GL_n(\mathbf{R})$, $SL_n(\mathbf{R})$, $O_n(\mathbf{R})$ and $SU_n(\mathbf{R})$.

Any subgroup of a Lie group which is also a submanifold is a Lie group, because the group structure on the subgroup is just the restriction of the group structure on the entire group, which is differentiable. This is essentially the argument we used to show that $SL_n(\mathbf{R})$, $O_n(\mathbf{R})$, and $SU_n(\mathbf{R})$ are Lie groups. We could have also used the fact that S^1 is a subgroup of the multiplicative group of non-zero complex numbers to show it was a Lie group. S^3 is a Lie group, because it is a subgroup of the Lie group of quaternions, consisting of elements of norm one. More generally, we define a **Lie subgroup** of a Lie group to be a subgroup, with some C^∞ structure making the operations of the subgroup differentiable, and such that the C^∞ structure makes the inclusion of the subgroup an immersion. As an example of a subgroup that is not an imbedded submanifold, consider the set of points $(x, cx) \in \mathbf{T}^2$, where c is an irrational number.

Lie groups are incredibly useful in geometry, because we often want to consider some symmetries which occur in a problem, which often turn out to be a group with some Lie structure. As an example, suppose we are discussing the metric structure of $\mathbf{R}^n$. In this situation, it is natural to

discuss the Euclidean group E_n, which is the collection of all isometries of $\mathbf{R}^n$. The easiest case to analyze is the group E_1. For each $x \in \mathbf{R}$, and $z \in \{-1, 1\}$, define $T_{xz}(t) = x + zt$. Every $T \in E_1$ can be written as T_{xz} for some x and some z. To see this, let $x = T(0)$. Since T is an isometry, either $T(1) = x + 1$ or $T(1) = x - 1$, because $|T(1) - x| = 1$. If $T(1) = x + 1$, then $T(y) = x + y$, because $x + y$ is the only number satisfying

$$|T(y) - x| = |y| \quad |T(y) - (x+1)| = |y - 1|$$

Similarily, if $T(1) = y - 1$, then $T(x) = y - x$. Since

$$T_{x_0 z_0} \circ T_{x_1 z_1} = T_{(x_0 + z_0 x_1)(z_0 z_1)}$$

and so E_1 can be described as the *semidirect product* of the multiplicative group $\{-1, 1\}$ and $\mathbf{R}$ under the representation $\rho : \{-1, 1\} \to \mathrm{Aut}(\mathbf{R})$ defined by $\rho(x)(y) = -y$. For E_2, we note that if $T : \mathbf{C} \to \mathbf{C}$ is an isometry such that $T(0) = 0$, and $T(1) = 1$, then $|T(i)| = 1$ and $|T(i) - 1| = \sqrt{2}$, so either

- $T(i) = 1$, which implies $T(z) = z$ for all $z \in \mathbf{C}$, because z is uniquely specified by the values $|z|$, $|z - 1|$, and $|z - i|$.

- $T(i) = -1$, which implies $T(z) = \bar{z}$ for all $z \in \mathbf{C}$, because $\bar{z}$ is the unique point with $|\bar{z}| = |z|$, $|\bar{z} - 1| = |z - 1|$, and $|\bar{z} - (-i)| = |z - i|$.

If $z \in \mathbf{C}$, and $w \in \mathbf{T}$, we define T_{zw} to be the isometry $T(u) = z + wu$. If $T \in E_2$ is arbitrary, and if $T(0) = u$, $T(1) = w$, then $T_{uw}^{-1} \circ T$ maps 0 to 0, and 1 to 1, hence either $T = T_{uw}$ or $T = \overline{T_{uw}}$. Note that $\overline{T_{uw}}(z) = T_{\overline{uw}}(\bar{z})$, so that the set of all T_{uw} is a normal subgroup of E_2. Since the group of all T_{uw} is isomorphic to the semidirect product $\mathbf{C} \rtimes \mathbf{T}$, because $T_{u_0 w_0} \circ T_{u_1 w_1} = T_{(u_0 + w_0 u_1)(w_0 w_1)}$, and therefore E_2 is isomorphic to the semidirect product $(\mathbf{C} \rtimes \mathbf{T}) \rtimes \{-1, 1\}$ with multiplication law

$$(z_0, w_0, t_0)(z_1, w_1, t_1) = \begin{cases} (z_0 + \overline{w_0 z_1}, w_0 \overline{w_1}, t_0 t_1) & t_0 = -1 \\ (z_0 + w_0 z_1, w_0 w_1, t_0 t_1) & t_0 = 1 \end{cases}$$

The fact that $\{-1, 1\}$ is isomorphic to O_1, and $\{-1, 1\} \rtimes \mathbf{T}$ is isomorphic to O_2 hints at a more general fact about the structure of the Euclidean groups, but we need to know some structure of the metric of $\mathbf{R}^n$ first. Say a point x lies *between* two points y and z if $x = \lambda y + (1 - \lambda)z$ for $0 \leqslant \lambda \leqslant 1$.

This holds if and only if $\|y-x\|+\|x-z\| = \|y-z\|$, because if $x = \lambda y+(1-\lambda)z$, then

$$\|y - x\| + \|x - z\| = \|(1 - \lambda)y - (1 - \lambda)z\| + \|\lambda y - \lambda z\| = \|y - z\|$$

and the Cauchy Schwartz inequality implies that this inequality occurs only when $y - x = \lambda(x - z)$ for some $\lambda > 0$, in which case we find

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

We say x, y, z are colinear if one point lies between the other pair of points. This occurs if and only if $y - x$ and $z - x$ are linearly dependant, because if $\lambda(y - x) = (z - x)$, then

- If $\lambda < 0$, then $\|y - x\| + \|x - z\| = \|y - z\|$, so x lies between y and z.

- If $0 < \lambda < 1$, then z lies between x and y, because $z = \lambda y + (1 - \lambda)x$.

- If $\lambda > 1$, then y lies between x and z, because

$$y = \frac{1}{\lambda}z + \frac{\lambda - 1}{\lambda}x$$

This tells us that an isometry maps straight lines to straight lines, because betweenness and colinearity are purely metric conditions, hence preserved by an isometry, and a line can be described by a set of points such that any triple of points is colinear. Furthermore, an isometry maps planes to planes, because a plane can be described as the smallest set containing a triple of non-colinear points x, y, z, and also containing the line generated by any points in the set. We claim that this plane is the set of points

$$\{x + \lambda(y - x) + \gamma(z - x) : \lambda, \gamma \in \mathbf{R}\}$$

If $x+\lambda_0(y-x)+\gamma_0(z-x)$ and $x+\lambda_1(y-x)+\gamma_1(z-x)$ are two points in this plane, then the set of points on the line between these two points is exactly $x + (\lambda_0 + t\lambda_1)(y - x) + (\gamma_0 + t\gamma_1)(z - x)$, as t ranges over all real numbers, and these points all lie in the set above. Conversely, if X is any colinearily closed set containing x, y, and z, then $x + t_0(y - x)$ and $x + t_1(z - x)$ are elements of x, for all $t \in \mathbf{R}$, and therefore

$$x + t_0(y - x) + t_2(t_1(z - x) - t_0(y - x)) = x + t_0(1 - t_2)(y - x) + t_2 t_1(z - x)$$

108

are also points in X, for all $t_0, t_1, t_2 \in \mathbf{R}$. This implies that the set of all points $x + \lambda(y - x) + \gamma(z - x)$ are contained in X. Since colinearity is a metric notion, an isometry maps planes to planes.

Now suppose $T : \mathbf{R}^n \to \mathbf{R}^n$ is an isometry, with $T(0) = 0$. Our discussion implies that T maps lines through the origin to lines through the origin. Thus $T(cx) = cT(x)$, because $cT(x)$ is the only point on the line through the origin and x which lies at a distance $|c|\|x\|$ from the origin and a distance $|c - 1|\|x\|$ from x. Similarly, for a fixed $x, y \in \mathbf{R}^n$, if we assume that T maps the plane generated by x and y to itself, then since $T(0) = 0$ we find $T(x + y) = T(x) + T(y)$. Otherwise, we consider a linear isometry S which projects the plane generated by $T(x)$ and $T(y)$ to the plane generated by x and y, and then it follows that $(S \circ T)(x + y) = S(Tx + Ty)$, hence $T(x + y) = Tx + Ty$, because S is a bijection. It follows that $T(0) = 0$ holds if and only if T is an element of the orthogonal group of isometric linear transformations O_n. If $T \in E_n$ is any linear transformation, and if $T(0) = x$, then the isometry $T_x^{-1} \circ T$ maps zero to zero, hence $T_x^{-1} \circ T \in O_n$, and we find that we can write any Euclidean transformation as a rotation and a translation, and by normality we find the Euclidean group is actually the semidirect product of O_n and $\mathbf{R}^n$, since if $M, N \in O_n$,

$$(T_x \circ M) \circ (T_y \circ N) = T_{x + My} \circ MN$$

E_n can be given the structure of a Lie group if we take the topology corresponding to $\mathbf{R}^n \times O_n$, in which case

$$(x, M)(y, N)^{-1} = (x, M)(-Ny, N^{-1}) = (x - MNx, MN^{-1})$$

which is differentiable, since the multiplication map on O_n is differentiable, and the map $x - MNx$ is differentiable since the action of M_n on $\mathbf{R}^n$ defined by $(M, x) \mapsto Mx$ is differentiable.

The left and right translation maps $L_x(y) = xy$ and $R_x(y) = yx$ are diffeomorphisms on any Lie group G, so they induce bundle equivalences $(L_x)_* : TG \to TG$ and $(R_x)_* : TG \to TG$. We say a vector field X is **left-invariant** if $(L_x)_* X = L_x \circ X$ for all $x \in G$, i.e. if $(L_x)_*(X_y) = X_{xy}$. It suffices to show that $(L_x)_*(X_e) = X_x$, because then

$$(L_x)_*(X_y) = (L_x \circ L_y)_*(X_e) = (L_{xy})_*(X_e) = X_{xy}$$

Given any $v \in G_e$, we can define a unique left invariant vector field X with $X_e = v$ by setting $X_x = (L_x)_*(v)$.

Theorem 7.1. *Any left-invariant vector field is automatically C^∞.*

Proof. We need only verify that the map $X_p = (L_p)_*(v)$ is C^∞ for any $v \in G_e$, and it suffices to prove this in a neighbourhood of the origin. Let (x, U) be a chart around a neighbourhood of the origin. Let $V \subset U$ be a neighbourhood chosen such that $ab^{-1} \in U$ if $a, b \in U$. Then

$$Xx_i$$

$\square$

Corollary 7.2. *A Lie group always has trivial tangent bundle.*

Since a Lie group is differentiable, we should be able to 'linearly approximate' the group multiplication action.

To remedy this fact, we are required to introduce some complicated machinery, which can be skimmed at a first reading. First, the elementary theory of Lie groups tells us that every tangent vector at the identity can be extended to a left-invariant smooth vector field on the Lie group. Given a vector $X_e \in \mathfrak{g}$, we can consider it as a base-point of a left-invariant vector field X, and use the field to generate a unique curve $\phi : \mathbf{R} \to G$ satisfying $\phi_0 = e$ and $d\phi_t/dt = X_{\phi_t}$. It turns out that $\phi_{t+u} = \phi_t \phi_u$, so that ϕ is actually a *homomorphism* from $\mathbf{R}$ to G (a 'one-parameter subgroup of G'), and we let the **exponential map** from $\mathfrak{g}$ to G by letting $e^X = \phi_1$. It turns out that $e^{tX} = \phi_t$, so the exponential map models all curves emerging from infinitisimals at the identity.

Example. *Over the multiplicative group $\mathbf{C}^\times$ of non-zero complex numbers, we can identify the tangent bundle $T\mathbf{C}^\times$ with $\mathbf{C}^\times \times \mathbf{C}$, and the left-invariant vector fields take the form $X_z = wz$ for some $w \in \mathbf{C}$. This tells us that the exponential on this space is the unique solution to the differential equation*

$$\frac{dz}{dt} = wz$$

and this is just the standard exponential map $z(t) = e^{tw}$, so that the exponential on $\mathbf{C}^\times$ is just the normal exponential function. Since the exponential descends consistently to Lie subgroups, this tells us that the exponential map on the multiplicative group $\mathbf{R}^\times$ is just the standard exponential as well.

Example. *Over the group $GL_n(\mathbf{R})$, we can identify the tangent space at each point $M \in GL_n(\mathbf{R})$ with the space $M_n(\mathbf{R})$ of all $n \times n$ matrices. The left-invariant vector fields on $GL_n(\mathbf{R})$ are of the form $X_M = NM$ for some $N \in M_n(\mathbf{R})$, and the unique solution to the system of linear equations*

$$\frac{dM}{dt} = NM$$

is the matrix exponential

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

hence the exponential of Lie groups generalizes many versions of the exponential defined in analysis.

Example. *Over the additive group of real numbers* $\mathbf{R}$, $T\mathbf{R}$ *is just the trivial tangent bundle* $\mathbf{R} \times \mathbf{R}$, *so elements of the tangent space at the identity can be identified with real numbers, and the left-invariant vector fields are just the vector fields with a constant velocity. Recalling the unique solutions to the differential equation*

$$\frac{dx}{dt} = c$$

We find that the for a tangent vector $t \in \mathbf{R}$ *at the identity,* $e^t = t$ *is just the identity map. More generally, the exponential related to the additive group* $\mathbf{R}^n$ *is just the identity map under a suitable idenfication of the tangent bundle.*

Some elementary facts about the exponential, proved in an elementary introduction to differentiable manifolds are that

- For any tangent vector X,

$$\left. \frac{de^{tX}}{dt} \right|_{t=0} = X$$

  so as t ranges over all real-numbers, e^{tX} is just the one-parameter subgroup of G corresponding to the map ϕ used to construct the exponential. Thus $e^{(t+u)X} = e^{tX} e^{uX}$.

- If $\phi : G \to H$ is a homomorphism, then $e^{\phi_*(X)} = \phi(e^X)$.

- $\exp : \mathfrak{g} \to G$ is a diffeomorphism in a neighbourhood of the identity. It is not always surjective, even if G is connected, but it is surjective for the group $GL_n(\mathbf{C})$, or any compact, connected Lie group.

Returning to our discussion of constructing homomorphisms of Lie groups from operations on the tangent space, we see that $e^{\phi_*(X)} = \phi(e^X)$ gives a much stronger condition on the linear map ϕ_* restricting which linear maps can be differentials of homomorphisms. In particular, since the image of exp always contains a neighbourhood of the identity, given any linear map ϕ_* such that $\phi_*(X) = \phi_*(Y)$ if $e^X = e^Y$, we can define a map ϕ on a neighbourhood of the identity of G by letting $\phi(e^X) = e^{\phi_*(X)}$. Provided that $\phi(gh) = \phi(g)\phi(h)$ where defined, we can try to extend ϕ to a homomorphism on the entire space by using the fact that a neighbourhood of the identity in the Lie group generates the entire space. The sufficient condition for this method to work is that the Lie group is simply connected,

and this is not too much of a problem because we can always swap a connected Lie group G with its simply connected cover $\tilde{G}$, and the underlying space of infinitisimals will be the same.

Thus we are left with determining the conditions on ϕ_* such that $\phi(gh) = \phi(g)\phi(h)$. This is where the Lie bracket enters the picture. If X and Y are arbitrary smooth vector fields on G, then they induce integral curves $\phi : \mathbf{R} \times G \to G$ and $\psi : \mathbf{R} \times G \to G$ respectively, and the Lie bracket operation $[X, Y]$ (turning the space of smooth vector fields into an infinite dimensional Lie algebra) defines a smooth vector field with

$$[X, Y]_p = (1/2)\left.\frac{d^2(\psi_{-t} \circ \phi_{-t} \circ \psi_t \circ \phi_t)}{dt}\right|_p$$

which effectively means that for any C^∞ function f,

$$[X, Y]_p(f) = (1/2)\left.\frac{d^2(f \circ \psi_{-t} \circ \phi_{-t} \circ \psi_t \circ \phi_t)}{dt}\right|_p$$

The Lie bracket essentially measures the commutivity of X and Y.

It turns out that if X and Y are both left-invariant vector fields on a Lie group G, then $[X, Y]$ is also a left-invariant vector field, hence the Lie bracket descends to an operation on the tangent space $\mathfrak{g}$, where if $X, Y \in \mathfrak{g}$, then $[X, Y] \in \mathfrak{g}$ satisfies

$$[X, Y](f) = (1/2)\frac{d^2}{dt^2}f(e^{tX}e^{tY}e^{-tX}e^{-tY})$$

Viewing $\mathfrak{g}$ as the space of curves through the origin identified up to first order, this implies that for any two curves λ, γ,

$$f(\lambda_t \gamma_t \lambda_t^{-1} \gamma_t^{-1}) = f(e) + t^2[\lambda, \gamma](f) + o(t^3)$$

Thus the Lie bracket expresses the second order coefficients of conjugation on the Lie group. Surprisingly, the second order terms are sufficient to characterize the Lie group operation. First note that if $\phi : G \to H$ is a group homomorphism, then in two different ways, we calculate that

$$f(\phi(\lambda_t \gamma_t \lambda_t^{-1} \gamma_t^{-1})) = f(e) + t^2[\lambda, \gamma](f \circ \phi) + o(t^3)$$
$$= f(e) + t^2(\phi_*[\lambda, \gamma])(f) + o(t^3)$$
$$f(\phi(\lambda_t \gamma_t \lambda_t^{-1} \gamma_t^{-1})) = f((\phi \circ \lambda)_t(\phi \circ \gamma)_t(\phi \circ \lambda)_{-t}(\phi \circ \gamma)_{-t})$$
$$= f(e) + t^2[\phi_*(\lambda), \phi_*(\gamma)](f) + o(t^3)$$

113

It follows that $\phi_*[\lambda, \gamma] = [\phi_*(\lambda), \phi_*(\gamma)]$, so all linear maps on infinitisimals induced by homomorphisms of groups preserve the Lie bracket operation. We shall find that we can 'almost' always find a Lie group homomorphism from a linear map on the space of infinitisimals preserving the Lie bracket.

Since exp is locally a diffeomorphism in a neighbourhood of both identities, we can find a differentiable inverse log, and the Baker-Hausdorff formula implies that there is a neighbourhood of the origin in $\mathfrak{g}$ and universal constants a_α such that

$$\log(e^X e^Y) = X + Y + \sum_{|\alpha|>1} a_\alpha (X, Y)^\alpha$$

where α are multi-indexes, and if $\alpha = (\alpha_1, \ldots, \alpha_n)$, then

$$(X, Y)^\alpha = \underbrace{[X, [X, [\ldots, [X,}_{\alpha_1 \text{ times}} \underbrace{[Y, [\ldots, [Y,}_{\alpha_2 \text{ times}} \ldots]]]]]]]]$$

A nasty formula which implies that given a linear map $\phi_* : \mathfrak{g} \to \mathfrak{h}$, the sufficient condition for the map $\phi : G \to H$ defined by $\phi(e^X) = e^{\phi_*(X)}$ to satisfy $\phi(gh) = \phi(g)\phi(h)$ where defined, a sufficient condition for this to hold is that

$$\log(e^{\phi_*(X)} e^{\phi_*(Y)}) = \phi_*(\log(e^X e^Y))$$

and in terms of the Baker-Hausdorff formula, we find that this means

$$\phi_*(X) + \phi_*(Y) + \sum a_\alpha (\phi_*(X), \phi_*(Y))^\alpha = \phi_*(X) + \phi_*(Y) + \sum a_\alpha \phi_*((X, Y)^\alpha)$$

It can be proved by induction that for any map ϕ_* with $\phi_*[X, Y] = [\phi_* X, \phi_* Y]$, $\phi_*((X, Y)^\alpha) = (\phi_*(X), \phi_*(Y))^\alpha$, so a sufficient condition for the homomorphism condition $\phi(gh) = \phi(g)\phi(h)$ is that ϕ_* preserves the Lie bracket.

Now that this property holds, if U is the neighbourhood upon which the homomorphism property holds, then any $g \in G$ can be written as $h_1 h_2 \ldots h_n$ for some $h_i \in U$, and if ϕ can be extended to a homomorphism on all of G, then we can let $\phi(g) = \phi(h_1) \ldots \phi(h_n)$, and provided this is well defined, ϕ will be a homomorphism that is differentiable at the identity, hence differentiable everywhere, and ϕ_* is the differential of ϕ. The only problem is that this extension of ϕ won't necessarily be well defined everywhere, and in order for this to work, we will need to switch to studying simply connected Lie groups.

Chapter 8

Riemannian Manifolds

On $\mathbf{R}^n$, an inner product enables us to discuss distances and angles. In 1854, Bernard Riemann figured out how to generalize this concept to a smooth manifold. We have exploited almost every tool of finite dimension linear algebra on manifolds, except for the theory of bilinear forms. Recall that a symmetric bilinear form β on a real vector space V is a bilinear map satisfying $\beta(v,w) = \beta(w,v)$. This form is called positive-definite if $\beta(v,v) > 0$ for $v \neq 0$, and *nondegenerate* if, for every $v \neq 0$, there is w with $\beta(v,w) \neq 0$ (every positive-definite form is automatically nondegenerate). Given a non-degenerate form, we can define a map $v \mapsto v$ from V to V^* by defining $v(w) = \beta(v,w)$. This map is injective precisely when β is non-degenerate, and thus naturally identifies V with V^* if V is finite dimensional. Given a basis $e_1, \ldots, e_n$ for V, there are coefficients $g_{ij} = \beta(e_i, e_j)$ such that $\beta(a^i e_i, b^j e_j) = a^i b^j g_{ij}$. A **Riemannian metric** on a manifold M is a smooth assignment of a positive-definite bilinear form $\langle \cdot, \cdot \rangle_p$ on the tangent spaces $T_p M$, for each point $p \in M$. This smoothness is characterized either by talking in the language of smooth tensor fields, so that a Riemannian metric is a smooth $(0,2)$ tensor field, in more basic terms, for any smooth vector fields X and Y, $\langle X, Y \rangle$ is a smooth function, or through coordinates, using a coordinate system x to locally write

$$\langle \cdot, \cdot \rangle_p = \sum a_{ij}(p) dx^i \otimes dx^j$$

and then the field is smooth if the a_{ij} are smooth. If the choice bilinear form is merely non-degenerate, the object is called a **psuedo-Riemannian metric**. A **(psuedo) Riemannian manifold** is just a smooth manifold with a (psuedo) Riemannian metric.

Example. $\mathbf{R}^n$ *is a Riemannian manifold, with the Riemannian tensor field*

$$\sum dx^i \otimes dx^i = \sum (dx^i)^2$$

Often, the field is just denoting by δ.

Example. *A surface in $\mathbf{R}^3$ inherits a Riemannian metric from $\mathbf{R}^3$ by constricting the restricted inner product on each tangent space to the surface. More generally, a metric is induced on submanifolds of a Riemannian manifold; if we have an immersion $i : M \to N$, and N has a Riemannian metric $\langle \cdot, \cdot \rangle_N$, then we can define a Riemannian metric by the equation*

$$\langle X, Y \rangle_M = \langle i_* X, i_* Y \rangle_N$$

so submanifolds of Riemannian manifolds are naturally given the structure of a Riemannian submanifold. Note that since every smooth manifold can be immersed in Euclidean space, every smooth manifold has a Riemannian metric (alternatively, we can construct a Riemannian metric by locally defining a positive definite form, and then extending it using a partition of unity).

Example. *The hyperbolic plane $\mathbf{H}^2$ can be modelled as the upper half plane, where the metric is given by $g = \delta/y^2$, or*

$$\langle X, Y \rangle_{(x,y)} = \frac{X^1 Y^1 + X^2 Y^2}{y^2}$$

This metric pushes the x axis 'off to infinity' by stretching the distance between points near the axis. This is known as the Poincare model of the hyperbolic plane. More generally, we can define the higher dimensional hyperbolic spaces $\mathbf{H}^n = \{(x,y) : x \in \mathbf{R}^{n-1}, y > 0\}$ with a metric δ/y^2.

An **isometry** between two Riemannian manifolds is a diffeomorphism which pullsback one metric on one manifold to a metric on the other. The study of Riemannian geometry can be considered the study of properties of Riemannian manifolds which are invariant under isometries.

Psuedo-Riemannian metrics most often occur in the physical theories employing differential geometry. Since the eigenvalues corresponding to the symmetric operator change smoothly, and cannot be zero by nondegeneracy, the number of negative and positive eigenvalues do not change across a connected psuedo-Riemannian manifold. The number of negative eigenvalues for a pseudometric is known as the index. An index 0 metric is a Riemannian metric, and an index 1 metric is known as a Lorentz metric.

Example. *The classical example of a Lorentz metric occurs in special relativity, known as the Minkowski metric on $\mathbf{R}^n \times \mathbf{R}$, which is given by*

$$\eta(X, Y) = X^1 Y^1 + \cdots + X^n Y^n - X^{n+1} Y^{n+1}$$

In general relativity, we replace $\mathbf{R}^n \times \mathbf{R}$ by a general $n+1$ dimensional manifold, and the Minkowski metric by an arbitrary Lorentz metric, satisfying certain physical equations known as the Einstein equations, which control the geometry of space. The Lorentz transformations are precisely the isometries of the Riemannian manifold with respect to the Lorentz metric.

8.1   The Gradient Vector

Since a form enables us to identify a vector space V with its dual, a metric enables us to identify TM and $(TM)^*$ in a natural way. Given a scalar valued function $f : M \to \mathbf{R}$, we obtain a section df of $(TM)^*$. By taking the dual, we obtain the **gradient** ∇f, which is a section of TM, defined by

$$\langle \nabla f, X \rangle = df(X)$$

In local coordinates, by taking $X = \partial/\partial x^j$, this equation says

$$dx^i(\nabla f) g_{ij} = \frac{\partial f}{\partial x^j}$$

If we let g^{ij} denote the components of the inverse of g_{ij}, so $g^{ij} g_{jk} = g_{ij} g^{jk} = \delta^i_k$, then

$$\nabla f = \frac{\partial f}{\partial x^j} g^{ji} \frac{\partial}{\partial x^i}$$

Note that then

$$|\nabla f|^2 = \langle \nabla f, \nabla f \rangle = df(\nabla f) = \frac{\partial f}{\partial x^i} g^{ij} \frac{\partial f}{\partial x^j}$$

Over Euclidean space, this is precisely the definition of the gradient.

Example. *On the sphere S^2, we have spherical coordinates $x = r\cos\phi\cos\theta$, $y = r\sin\phi\cos\theta$, and $z = r\sin\theta$. We then have*

$$dx = \cos\phi\cos\theta\, dr - r\sin\phi\cos\theta\, d\phi - r\cos\phi\sin\theta\, d\theta$$

$$dy = \sin\phi\cos\theta dr + r\cos\phi\cos\theta d\phi - r\sin\phi\sin\theta d\theta$$

$$dz = \sin\theta dr + r\cos\theta d\theta$$

and so the metric is

$$g = dx^2 + dy^2 + dz^2 = dr^2 + (r^2\cos^2\theta)d\phi^2 + (r^2)d\theta^2$$

Thus the basis is orthogonal, but nor orthonormal, with

$$\left|\frac{\partial}{\partial r}\right| = 1 \qquad \left|\frac{\partial}{\partial\phi}\right| = r|\cos\theta| \qquad \left|\frac{\partial}{\partial\theta}\right| = r$$

If we normalize these vectors, letting

$$\frac{\widehat{\partial}}{\partial\phi} = \frac{1}{r\cos\theta}\frac{\partial}{\partial\phi} \qquad \frac{\widehat{\partial}}{\partial\theta} = \frac{1}{r}\frac{\partial}{\partial\theta}$$

then

$$\nabla f = \frac{\partial f}{\partial r} + \frac{1}{r\cos\theta}\frac{\partial f}{\partial\phi}\frac{\widehat{\partial}}{\partial\phi} + \frac{1}{r}\frac{\partial f}{\partial\theta}\frac{\widehat{\partial}}{\partial\theta}$$

These are known as the physical components of the gradient with respect to spherical coordinates.

Example. *In Minkowski space, we have a time coordinate t, and three spatial coordinates x, y, and z. If t is chosen as the 'zeroeth' coordinate, the Minkowski metric is chosen such that*

$$g_{ij} = \begin{cases} 1 & i = j > 0 \\ -c^2 & i = j = 0 \\ 0 & otherwise \end{cases}$$

Then if f is a function on Minkowski space, we find

$$\nabla f = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z} - \frac{1}{c^2}\frac{\partial f}{\partial t}$$

The gradient ∇f has the same interpretation as in Euclidean space as the direction of steepest ascent on the manifold. If $v \in M_p$ is any vector with $|v| = 1$, then the Schwarz inequality implies

$$|v(f)| = |df(v)| = |\langle v, \nabla f\rangle| \leqslant |\nabla f|$$

If f is real valued, then the tangent space of the level sets are precisely the vectors orthogonal to the gradient vector.

8.2 Unit Tangent Bundle

If M is a Riemannian manifold, we can obtain an interesting submanifold of TM by considering the subspace of *unit* vectors v with $|v| = 1$. The equation $g(v, v) = 1$ is given in local coordinates by

$$\sum_{ij} \dot{x}^i g_{ij} \dot{x}^j = 1$$

If we let $f(v) = g(v, v)$. Then

$$\frac{\partial f}{\partial \dot{x}^i} = 2 \sum_j g_{ij} \dot{x}^j$$

Since g is non-degenerate, these partial derivatives cannot all simulatenously vanish unless $\dot{x} = 0$. Thus the space of v with $|v| = 1$ is a submanifold of TM of dimension $2n - 1$, if M has dimension n.

Example. *If w is a unit tangent vector to S^2 at some point $v \in S^2$, then $(v, w, v \times w)$ is a right handed orthonormal coordinate frame. Thus we find that the space of unit tangent vectors is diffeomorphic to $SO(3)$.*

8.3 The Volume Form

In $\mathbf{R}^n$, the length of a smooth curve c parameterized on $[a, b]$ is given by

$$\int_a^b |c'(t)| \, dt$$

In a Riemannian manifold, we can measure the lengths of tangent vectors v by the equation $|v|^2 = \langle v, v \rangle$, and so we can also measure the lengths of curves on Riemannian manifolds. We also have the existence of a **volume forms**, which enable one to integrate smooth functions on the manifold in a natural way.

8.4 Connections

On $\mathbf{R}^n$, given two vector fields X and Y, we can define the 'derivative' of the vector field Y along X to be the vector field

$$(D_X Y)_p = \lim_{t \to 0} \frac{Y_{p+tX} - Y_p}{t}$$

On a manifold, there is no canonical way to differentiate vector fields; the Lie derivative $L_X Y$ isn't a staisfying definition because it measures how 'independent' two vector fields are, rather than measuring the change of Y along X. The main problem is that we are unable to 'connect' tangent spaces on a manifold close to one another, like we can in $\mathbf{R}^n$. On a Riemannian manifold, there is a canonical way to connect the fibres of a tangent bundle at points close to one another, and this is best explained through the concept of a connection.

Recall the definition of a smooth vector bundle $\pi : E \to M$, and smooth sections $s : M \to E$. We denote the $C^\infty(M)$ module of all smooth sections by $\Gamma(M)$. A **connection** is a map $\nabla : \Gamma(TM) \times \Gamma(E) \to \Gamma(E)$ which maps $(X, s) \to \nabla_X(s)$, which is $C^\infty(M)$ linear in X, $\mathbf{R}$ linear in s, and satisfies the product rule

$$\nabla_X(fs) = X(f)s + f\nabla_X(s)$$

for any $f \in C^\infty(M)$, so that ∇_X operates 'like a derivative' on s. An **affine connection** is a connection where $E = TB$.

Example. *The definition*

$$(\nabla_X Y)_p = D_X(Y)_p = \lim_{t \to 0} \frac{Y_{p+tX} - Y_p}{t}$$

is an affine connection on $\mathbf{R}^n$.

Example. *The Lie derivative is not an affine connection on a manifold, because it isn't* $C^\infty(M)$ *linear in the first variable.*

The fact that ∇ is $C^\infty(M)$ linear in the 1st variable tells us that for a fixed section s, the map $X \mapsto \nabla_X s$ acts 'like a tensor' in X. An argument analogous to the tensor characterization lemma reveals that $(\nabla_X s)(p)$ depends only on the value X_p of X at p, not the entire vector field. On the other hand, for the sections s we only have $\mathbf{R}$ linearity, so $\nabla_X s$ is allowed to depend on more of the behaviour of s. A bump function type argument, analogous to the proof that derivations on $C^\infty(M)$ are local, shows that $(\nabla_X s)(p)$ depends only on the behaviour of s in a neighbourhood of p. The introduction of Christoffel symbols will show that, if fact, $(\nabla_X s)(p)$ depends only on the behaviour of s on a curve tangent to X.

Suppose that we consider a trivialization U from $\pi^{-1}(U) \to U \times \mathbf{R}^n$, some local frame $s_1, \ldots, s_n$ on U, and a coordinate chart (x, U). We know

that we can define $\nabla_X(s_k)$, even though the sections s_k are not defined globally, because ∇_X depends only on the behaviour of s_k locally. The **Christoffel symbols** with respect to this setup are functions $\Gamma^\alpha_{i\beta}$ such that

$$\nabla_{e_i}(s_\beta) = \sum \Gamma^\alpha_{i\beta} s_\alpha$$

where $e_i = \partial/\partial_{x_i}$. Then if $X = \sum a^i e_i$, and $s = \sum b^\beta s_\beta$

$$\begin{aligned} \nabla_X(s) = \sum \nabla_X(b^\beta s_\beta) &= \sum X(b^\beta) s_\beta + b^\beta \nabla_X(s_\beta) \\ &= \sum X(b^\beta) s_\beta + b^\beta a^i \nabla_{e_i}(s_\beta) \\ &= \sum X(b^\beta) s_\beta + b^\beta a^i \Gamma^\alpha_{i\beta} s_\alpha \\ &= \sum \left(X(b^\alpha) + b^\beta a^i \Gamma^\alpha_{i\beta} \right) s_\alpha \end{aligned}$$

The Christoffel symbols and the a^i at p do not depend on s, and $b^\beta(p)$ depends only on the values of s at p. $X(b^\alpha)$ depends only on the behaviour of s tangent to X, and therefore $\nabla_X(s)$ depends only on the bahviour of s locally on a curve tangent to X.

There are infinitely many degrees of freedom for defining an affine connection on a manifold. Any smooth family of Christoffel symbols gives rise to a connection on $\mathbf{R}^n$, and using a partition of unity type decomposition, one can patch together local connections to a global connection. The only thing to note here is that linear combinations of connections need not be a connection, but convex combinations are.

To introduce the natural choice of connection on a Riemannian manifold, we need to discuss how parallel transports arise from a particular connection. It turns out that if we have a connection on a smooth bundle $\pi : E \to M$, and γ is a curve on M from a point p to a point q, then there arises a linear transformation $P_p^q : E_p \to E_q$ arising from the curve γ, called a parallel transport. Conversely, a family of parallel transports give rise to a connection.

8.5   Riemannian Submanifolds

Recall that if M is a Riemannian manifold with metric g, and N is an immersed submanifold with immersion $i : M \to N$, then we can give $i^*(TM)$

a metric structure by the pullback metric $i^*(g)$. We can of course embed TN in $i^*(TM)$, and obtain a Riemannian metric structure on TN by restricting $i^*(g)$, but we can also define another interesting bundle from this embedding. We define the **normal bundle** of N with respect to the embedding to be the orthogonal complement of TN in $i^*(TM)$.

Theorem 8.1. *The Levi-Civita connection ∇ relative to the induced metric on N is given by*

$$\nabla_X^N Y = P(\nabla_X^M Y)$$

For $X, Y \in \Gamma(TM)$, where P is the orthogonal projection of $i^*(TM)$ onto TN.

Proof. It is straightforward to check ∇^N defines an affine connection on TN. It suffices to sheck that the connection is metric compatible and torsion free. We find $\nabla_X^N Y - \nabla_Y^N X = P(\nabla_X^M Y - \nabla_Y^M X) = P([X, Y]) = [X, Y]$, since we know that if X and Y lie in a subbundle of the tangent bundle, then $[X, Y]$ also lies in this subbundle. To check metric compatibility, we find

$$X_p(i^*g)(Y_p, Z_p) = X_p g(i_* Y_p, i_* Z_p) = g(\nabla_X^M(i_* Y_p), Z_p) + g(Y, \nabla_X^M Z) = g(\nabla_X^N Y, Z) + g(Y, \nabla_X^N Z)$$

and the uniqueness of the Levi-Cevita connection gives our result. $\qquad\square$

The **second fundamental form** of a pair $N \subset M$ of Riemannian manifold is defined to be $\mathbf{II}(X, Y) = Q(\nabla_X^M Y)$, where Q is the orthogonal projection onto $TN^\perp$. It is $C^\infty(M)$ bilinear and symmetric, and so corresponds to a smooth covariant two tensor field on $i^*(TM)$. The equation

$$\nabla_X^M Y = \nabla_X^N Y + \mathbf{II}(X, Y)$$

The manifold N is **totally geodesic** if $\mathbf{II} = 0$. The **mean curvature vector** H is the trace of $\mathbf{II}$, and is the gradient of the volume function on the manifold. A manifold is **minimal** if the mean curvature vector vanishes. Given a unit vector $\nu \in N_p M$ normal to M, set $A^\nu : T_p M \to T_p M$ by $A^\nu(X) = -Q(\nabla_X^M \nu)$, known as the **shape operator**. Then A^ν is self-adjoint, and $\langle \mathbf{II}(X, Y), \nu \rangle = \langle A^\nu(X), Y \rangle$.

The Gauss equation says that

$$\langle R^N(X, Y)Z, W \rangle = \langle R^M(X, Y)Z, W \rangle + \langle \mathbf{II}(X, W), \mathbf{I}(Y, Z) \rangle - \langle \mathbf{II}(X, Z), \mathbf{II}(Y, W) \rangle$$

8.6 Jacobi Fields

Let γ be a geodesic. Then γ is locally minimizing if and only if γ has no conjugate points. Recall that $I(V, V)$ is the index form

$$I(V, V) = \int_0^t |\nabla_{\partial_t} V|^2 - \langle R(V, \gamma')\gamma', V \rangle$$

A pair of **conjugate points** along a geodesic if there is a nonvanishing Jacobi field between the two points.

Theorem 8.2. *If γ has no conjugate points then $I(\cdot, \cdot)$ is positive definite.*

We proved this theorem this morning, in the lecture I missed. This lecture we focus on the converse. This means that past a conjugate point pair, geodesics can fail to be locally minimizing.

Theorem 8.3 (Index Inequality). *Let γ be a geodesic with no conjugate points, and let J be a Jacobi field along γ. If V is a vector field along γ with $V(0) = J(0)$ and $V(l) = J(l)$ then $I(J, J) \leqslant I(V, V)$, with equality if and only if $J = V$.*

Proof. Since γ contains no conjugate points, $I(V, V) = 0$ for all V with $V(0) = V(l) = 0$, with $I(V, V) = 0$ if and only if $V = 0$. In particular, $I(J - V, J - V) = I(J, J - V) - I(V, J - V) = I(V, J - V) \geqslant 0$, since $I(J, J - V) = 0$. But then $I(J, J) = I(J, V)$, which shows $I(J, J) \leqslant I(V, V)$. $\qquad\square$

Corollary 8.4. *A geodesic containing an interior point which is conjugate to $\gamma(0)$ is not locally minimizing.*

Proof. Let $\gamma(t_0)$ be the first conjugate point to $\gamma(0)$. Then there is a nonzero Jacobi field J on $\gamma_{[0,t_0]}$ with $J(0) = J(t_0)$, which can be extended to a vector field X on all of γ by making the vector field vanish elsewhere. Then $I(X, X) = 0$ on $[0, t_0]$. Note X is not smooth at t_0, since $(\nabla_{\partial_t} J)(t_0) = 0$, because this would imply J is zero everywhere. However, if δ is sufficiently smlal, then γ contains no pairs of conjugate points around t_0, so there exists a local Jacobi field W with $W(t_0 - \delta) = J(t_0 - \delta)$ and $W(t_0 + \delta) = 0$. The index inequality implies $I(V, V) < I(X, X)$ on $[t_0 - \delta, t_0 + \delta]$. But $X = V$ outside of the $[t_0 - \delta, t_0 + \delta]$, so $I(V, V) < I(X, X) = 0$ on $[0, l]$, and so there is a variation of γ decreasing the length. $\qquad\square$

Remark. A quantitative generalization of this theorem is given by the Morse index theorem. The Morse index is the maximal dimension of a subspace of variation fields on which I is negative definite. We just proved that if there is a conjugate point, the index is at least one. The morse index theorem says the index is finite and equals the # of interior cojugate points, counted with multiplicity.

8.7   Nonpositive Curvature

Recall that give $p \in M$ and a 2-plane $\Pi \subset T_pM$, we define the *sectional curvature*

$$K(\Pi) = \frac{\langle R(V,W)W,V \rangle}{|V|^2|W|^2 - \langle V,W \rangle^2}$$

We set $K_M > k$ if $K(\Pi) > k$ for all 2 planes $\Pi \subset T_pM$.

Theorem 8.5. *If $K_M \leqslant 0$ then M has no conjugate points.*

Proof. Let γ be a unit speed geodesic on the manifold, and let V be a normal vector field along γ with $V(0) = V(l) = 0$. Then

$$\frac{d^2 L(\gamma_s)}{ds^2}\bigg|_{s=0} = \int_0^l \left[|\nabla_{\partial_t} V|^2 - \langle R(V,\gamma',V) \rangle \right] \geqslant \int_0^l \left| \nabla_{\partial_t} V \right|^2 \, dt > 0$$

except when $\nabla_{\partial_t} V = 0$, but this only occurs if $V = 0$ (uniqueness of initial conditions). $\square$

Theorem 8.6 (Cartan-Hadamard)**.** *If M is a complete Riemannian manifold, and $K_M \leqslant 0$, then the exponential is a covering map.*

Proof. It suffices to note that if M is a complete Riemannian manifold, and F is a local isometry, thne F is a covering map. We know that since $K_M \leqslant 0$, M has no conjugate points, which is the set of points where the differential of the exponential map is non-invertible. $\square$

We shall find that if K_M is stictly negative, then geodesics are unique.

8.8 March 22nd

Now we address problems related to positive curvature. Recall that the Ricci tensor is a symmetric quadratic form which is obtained by contracting the Riemann curvature tensor, i.e.

$$\mathrm{Ric}(X,Y) = \sum_{n=1}^{N} \langle R(X,e_n)e_n, Y \rangle$$

It is the directional average of sectional curvatures. We say $\mathrm{Ric}_M \geqslant K$ if $\mathrm{Ric}_M(X,X) \geqslant K$ for all unit vectors X.

Theorem 8.7 (Bonnet-Myers). *If (M,g) is complete, and $\mathrm{Ric}_M \geqslant (n-1)K > 0$, then ever geodesic of length $\geqslant \pi/\sqrt{K}$ contains conjugate points. THus the diameter of the manifold is less than or equal to $\pi/\sqrt{K}$.*

Proof. Let $\gamma : [0,l] \to M$ be a unit length geodesic. Suppose $L(\gamma) = l > \pi/\sqrt{K}$. We will now show γ is not locally minimizing. Choose an orthonormal basis $\{E_1,\dots,E_n\}$ of $T_{\gamma(0)}M$, and extend the basis by parallle transport. Set $V_i = \varphi E_i$ where $\varphi(0) = \varphi(l) = 0$. We calculate

$$\sum I(V_i,V_i) = -\sum \int_0^l \langle \nabla_{\partial t}\nabla_{\partial t}V_i, V_i \rangle + \langle R(V_i,\varphi')\varphi', V_i \rangle\, dt$$

$$= -\int_0^l (n-1)\varphi''\varphi + \varphi^2\mathrm{Ric}(\varphi',\varphi')$$

$$\leqslant -(n-1)\int_0^l (\varphi + k\varphi)\varphi\, dt$$

Setting $\varphi(t) = \sin(\pi t/l)$, which is the first laplacian eigenfunction on $[0,l]$, this causes the integral to be bounded by

$$-(n-1)\int_0^l [-(\pi/l)^2 + k]\varphi^2 < 0$$

provided $l > \pi/\sqrt{K}$. $\qquad\square$

Corollary 8.8. *Let (M,g) be a complete manifold, with $\mathrm{Ric}_M \geqslant (n-1)k >?$. Then M is compact with a finite fundamental group.*

125

Proof. Every point in the manifold can be connected to be a geodesic of length at most $\pi/\sqrt{K}$, so it follows that the exponential map on the closure of a ball maps surjectively onto the whole space, hence the image is compact. To obtain that the image is a finite fundamental group, the universal cover of this manifold also satisfies the complete curvature condition, hence it is compact, and so the number of sheets is finite, hence the fundamental group is finite. $\square$

Example. *SInce the fundamental group of $S^1 \times S^2$ is **Z**, which is infinite, for every metric there is a plane with zero curvature. The Hopf conjecture is whether $S^2 \times S^2$ has a product metric with $K \geqslant 0$.*

Theorem 8.9 (Synge)**.** *A compact orientable even dimensional manifold with positive sectional curvature*

Proof. Suppose M is not simply connected. Every nontrivial free homotopy class has a minimizing geodesic $\square$

Chapter 9

Moving Frames

Let's start by considering frames in $\mathbf{R}^3$. Let e_1, e_2, e_3 be a right handed smooth orthonormal frame. Since we are working in Euclidean space, we can think of the e_n as vector fields. Because e_1, e_2, e_3 are a basis, we can write

$$e_i' = \omega_{i1}e_1 + \omega_{i2}e_2 + \omega_{i3}e_3$$

Since $e_i \cdot e_j = \delta_{ij}$, we can differentiate on both sides to conclude that $e_i' \cdot e_j + e_i \cdot e_j' = 0$. This implies

$$\omega_{i1} = e_i' \cdot e_1 = -e_1' \cdot e_i = -\omega_{1i}$$

Thus the matrix $\Omega = (\omega_{ij})$ is skew symmetric. In particular, $e_i' \cdot e_i = 0$. If we consider the one forms $\sigma_k = dx(e_k)dx + dy(e_k)dy + dz(e_k)dz$, then

Chapter 10

De Rham Cohomology

We now want to ask how many non-exact closed k forms there are on a manifold M. If ω is non exact, then $\omega + d\eta$ is non-exact, so it is natural to consider these forms as equivalent when counting the non-exact forms. This is also natural from the point of integration. If N is a closed submanifold of M, then Stoke's theorem implies that

$$\int_N \omega + d\eta = \int_N \omega$$

so the 'induced action' of the form on closed submanifolds is the same for closed forms which differ by an exact differential. If we let $Z^k(M) \subset \Omega^k(M)$ denote the class of closed k forms, and $B^k(M)$ the class of exact k forms, then we can form the quotient space $H^k(M) = Z^k(M)/B^k(M)$, known as the **De Rham Cohomology**. This is exactly what we want, since the dimension of this space counts the number of different non-exact closed k forms on space. De Rham showed that this cohomology is equivalent to the cohomology over the reals, and therefore that the spaces $H^k(M)$ are actually a topological invariant of M, rather than just a differential invariant.

Example. *The Poincaré lemma on $\mathbf{R}^n$ says that every closed form is exact. It follows that every $H^k(\mathbf{R}^n)$ is trivial for $k > 0$. Thus there are no interesting ways to integrate closed submanifolds on Euclidean space.*

Example. *The space $B^0(M)$ is trivial, so $H^0(M)$ is really just the space of all smooth real-valued functions f on M such that $df = 0$. Another way of describing this is the space of all functions on M which vanish on each component of M, so the dimension of $H^0(M)$ gives the number of components of M.*

Example. *If M is a compact and oriented n manifold with orientation μ, then there is a form ω such that $\omega(x)(v_1,\ldots,v_n) > 0$ if and only if $(v_1,\ldots,v_n) \in \mu_x$. Because of this, we find that*

$$\int_M \omega > 0$$

This implies that $\omega \notin B^k(M)$, since $\int_M d\eta = 0$ for all η. Thus $H^n(M)$ is non-trivial.

10.1   Integration in Polar Coordinates

Aside from these trivial remarks, our knowledge of integration is really not enough to compute the cohomology of many other manifolds. To get further, we require some knowledge of integration on spheres. On the sphere S^{n-1} in $\mathbf{R}^n$, there is a natural choice of orientation form, the volume form

$$\sigma(x)(v_1,\ldots,v_{n-1}) = \det(x,v_1,\ldots,v_{n-1})$$

In coordinates, we have

$$\sigma(x) = \sum (-1)^{k-1} x^i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n$$

The natural retraction map $r : \mathbf{R}^n - \{0\} \to S^{n-1}$ with section $i : S^{n-1} \to \mathbf{R}^n - \{0\}$ thereby induces an $n-1$ form $r^*\sigma$ on $\mathbf{R}^n - \{0\}$. It is closed, since $d\sigma$ is closed, but not exact, since if $r^*\sigma = d\eta$, then

$$\sigma = (r \circ i)^*(\sigma) = i^*(r^*\sigma) = i^*(d\eta) = d(i^*\eta)$$

We shall use it to compute in polar coordinates.

If we extend σ to all of $\mathbf{R}^n - \{0\}$ by the same formula as on S^{n-1}, then $(r^*\sigma)(x) = \sigma(x)/|x|^n$. To see this, we note that the tangent space $\mathbf{R}_x^n$ for $x \neq 0$ is spanned by the tangent vectors to the sphere of radius $|x|$ and x_x. So it suffices to show that the two $n-1$ forms operate on a family of $n-1$ vectors, each of this form. The vector x_x is the tangent vector to the straight line $c(t) = (1 + t)x$ from the origin to the point x, and since $(r \circ c)(t) = r(x)$ is constant, we conclude that $r_*(x_x) = 0$, and so $(r^*\sigma)(x)(v_1,\ldots,v_{n-1})$ vanishes if any of the v_n are of the form x_x. On the other hand, since

$$\sigma(x)(v_1,\ldots,v_{n-1}) = \det(x,v_1,\ldots,v_{n-1})$$

129

if any of the v_n are x_x, the determinant vanishes. Thus we now need only show that the two forms operate the same on tangent vectors to the sphere of radius $|x|$. For such a vector v, we have $r_*(v_x) = v_{r(x)}/|x|$, because if $c(t)$ is a curve on the sphere of radius x, then $(c \circ r)(t) = c(t)/|x|$, and thus $(c \circ r)'(t) = c'(t)/|x|$. This proves the equality in this case, since

$$(r^*\sigma)(x)(v_1,\ldots,v_n) = \sigma(x/|x|)(v_1/|x|,\ldots,v_n/|x|) = |x|^{-n}\det(x,v_1,\ldots,v_n)$$

As a corollary of this computation, we find that we can integrate by polar coordinates.

Corollary 10.1. *If f is a real valued function on the unit ball B in $\mathbf{R}^n$, and we set $g : S^{n-1} \to \mathbf{R}$ by $g(x) = \int_0^1 t^{n-1}f(tx)\,dt$, then*

$$\int_B f = \int_{S^{n-1}} g\sigma$$

Proof. Consider the projections $\pi_1 : S^{n-1} \times [0,1] \to S^{n-1}$ and $\pi_2 : S^{n-1} \times [0,1] \to [0,1]$. For any form ω on S^{n-1} and η on $[0,1]$, we can consider the form $\pi_1^*\omega \wedge \pi_2^*\eta$, which we denote by $\omega \wedge \eta$. If y is a coordinate system on S^{n-1}, then there exists h such that $\sigma = h\,dy$. If we consider the coordinate system $y = y \circ \pi_1$ and $t = \pi_2$ on $S^{n-1} \times [0,1]$, then $\sigma \wedge dt = (h \circ \pi_1)dy \wedge dt$. Thus if $k(x,t) = t^{n-1}f(tx)$, then an easy arguments by partitions of unity establishes that

$$\int_{S^{n-1}} g\sigma = (-1)^{n-1}\int_{S^{n-1} \times [0,1]} k\sigma \wedge dt$$

Now we consider the standard diffeomorphism from $\phi : B - \{0\} \to S^{n-1} \times (0,1]$, and calculate

$$\begin{aligned}
\phi^*(\sigma \wedge dt) &= \phi^*(\pi_1^*\sigma \wedge \pi_2^*dt) \\
&= (\pi_1 \circ \phi)^*(\sigma) \wedge (\pi_2 \circ \phi)^*(dt) \\
&= r^*\sigma \wedge v^*
\end{aligned}$$

$\square$

10.2   Compact De Rham Cohomology

We would like to reduce our computations to coordinate neighbourhoods, but in order to do this, we must introduce another quotient class of forms.

We restrict our knowledge to the class $\Omega_c^k(M)$ of differential forms with compact support, with the corresponding spaces $Z_c^k(M)$ and $B_c^k(M)$, which is the class of forms which is the differential of a form *with compact support*. We therefore get an induced cohomology $H_c^k(M)$, known as the **De Rham cohomology with compact support**. Of course, if M is compact, these two cohomologies are exactly the same.

Example. *The space $B_c^k(M)$ is not the same as the class of all exact differentials with compact support. On $\mathbf{R}^n$, consider a compactly supported function f with $f(x) \geqslant 0$, and $f(x_0) > 0$ at some point x_0. Now the form $\omega = f\, dx^1 \wedge \cdots \wedge dx^n$ is exact, but it is not a differential of an $n-1$ form η with compact support, because if this were true, then by Stoke's theorem*

$$0 < \int_{\mathbf{R}^n} f(x)\, dx = \int_{\mathbf{R}^n} d\eta = 0$$

Thus $H_c^n(\mathbf{R}^n)$ is non trivial, even though $H^n(\mathbf{R}^n)$ is trivial. A similar argument shows that if M is any oriented n manifold, then $H_c^n(M)$ is nontrivial.

In fact, we now prove $H_c^n(M)$ is always one dimensinoal.

Theorem 10.2. *If M is oriented and connected, the map*

$$\omega \mapsto \int_M \omega$$

induces an isomorphism of $H_c^n(M)$ to $\mathbf{R}$.

Proof. First, take $M = \mathbf{R}$. If ω is a closed one form on $\mathbf{R}$, there exists a function f such that $df = \omega$. If ω vanishes outside of $[a, b]$, then f must be constant on $(-\infty, a]$ and $[b, \infty)$. Since

$$f(b) - f(a) = \int_a^b \omega = \int_{\mathbf{R}} \omega$$

If $\int_{\mathbf{R}} \omega = 0$, then $f(b) = f(a)$, and so if $g(x) = f(x) - f(a)$, then g has compact support and $dg = \omega$. Thus we have proved the theorem in this special case. Now we prove that if the theorem is true for all $n-1$ manifolds, then it ust be true for $\mathbf{R}^n$. Consider a closed n form ω with compact support on

$\mathbf{R}^n$. There exists an $n-1$ form η such that $d\eta = \omega$, and if $\omega = f\,dx$, we can actually define η by the formula

$$\eta(x) = \sum_{k=1}^{n} (-1)^{k-1} \left(\int_0^1 t^{n-1} f(tx)\,dt \right) x^i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n$$

which follows because

$$\frac{\partial}{\partial x^i} \left(x^i \int_0^1 t^{n-1} f(tx)\,dt \right) = \int_0^1 t^{n-1} f(tx)\,dt + x^i \int_0^1 t^{n-1} \frac{\partial f}{\partial x^i}(tx)\,dt$$

and so

$$d\eta(x) = \left(n \int_0^1 t^{n-1} f(tx)\,dt + \sum_{i=1}^{n} x^i \int_0^1 t^{n-1} \frac{\partial f}{\partial x^i}(tx)\,dt \right) dx$$

If $F(t) = t^n |x|^n f(tx)$, then

$$F'(x) = |x|^n \left(n t^{n-1} f(tx) + t^n \sum_{k=1}^{n} \frac{\partial f}{\partial x^i} x^i \right)$$

and the fundamental theorem of calculus implies

$$d\eta(x) = [F(1) - F(0)]\,dx = f(x)\,dx = \omega(x)$$

Now $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

Example. *If M is a compact and oriented n manifold, then $H^n(M)$ is nontrivial. To see this, we note that if the orientation of M is μ, M has a non-vanishing n form ω defined on the whole space, such that $\omega(v_1, \ldots, v_n) > 0$ if and only if $(v_1, \ldots, v_n) \in \mu$. It follows that*

$$\int_M \omega > 0$$

which proves that ω must be non-zero in $H^n(M)$, since exact forms must integrate to zero over the entire manifold.

Chapter 11

Differential Topology Notes

A **connection** d_A on a bundle E is an **R** linear map $d_A : \Omega^0(M, E) \to \Omega^1(M, E)$ satisfies $d_A(fs) = df \wedge s + f d_A s$ for all $f \in C^\infty(M)$, $s \in \Gamma(E)$. Note that the support of $d_A s$ must be contained in the support of s. The operator d_A is local, and for local frames $s^1, \ldots, s^n$, we have $d_A s^i = \sum a^i_j s^j$. This is the connected matrix, or matrix valued one forms, or gauge fields.

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*