An introductory course in differential geometry and the Atiyah-Singer index theorem

Paul Loya

BINGHAMTON UNIVERSITY, BINGHAMTON NY 13902-6000
E-mail address: paul@math.binghamton.edu

CHAPTER 1

What is the Atiyah-Singer index theorem?

1.1. The index of a linear map

To describe the index theorem we of course need to know what "index" means.

SECTION OBJECTIVES: THE STUDENT WILL BE ABLE TO ...

- define the index of a linear map.
- compute the index for "simple" linear maps.
- interpret the index analytically and topologically.

1.1.1. Definition and examples of the index. Let V and W be complex vector spaces, not necessarily finite-dimensional, and let $L : V \to W$ be a linear map. Recall that the **kernel** (or **null space**) of L is defined as

$$\ker L := \{v \in V \mid Lv = 0\}.$$

The **cokernel** of L is by definition W modulo the image of L:

$$\operatorname{coker} L := W/\operatorname{Im} L;$$

the represents the space "missed by L" (not in the image of L). Note that both $\ker L$ and $\operatorname{coker} L$ are vector spaces. If both $\ker L$ and $\operatorname{coker} L$ are finite-dimensional, then we say that L is **Fredholm**. In a very rough sense, this means that L is "almost bijective" in the sense that L is "almost injective" because L is injective except on the finite-dimensional part $\ker L$, and L is "almost surjective" because L only a finite-dimensional part of W is missed by L. In particular, note that if L is an isomorphism of the vector spaces V and W, then both the kernel and cokernel of L are zero, so L is Fredholm.

When L is Fredholm, the dimensions of the vector space $\ker L$ and $\operatorname{coker} L$ are integers, and the **index** of L is defined as the difference:

$$\boxed{\operatorname{ind} L := \dim \ker L - \dim \operatorname{coker} L \in \mathbb{Z}.}$$

In particular, the index of an isomorphism is zero. Here are some more examples.

Example 1.1. If V is finite-dimensional, observe that $\ker L \subset V$ is automatically finite-dimensional. Also, if W is finite-dimensional, then $\operatorname{coker} L = W/\operatorname{Im} L$ is automatically finite-dimensional. Thus, if both V and W are finite-dimensional, then $\ker L$ and $\operatorname{coker} L$ are finite-dimensional for any given linear map $L : V \to W$. In particular, *any* linear map between finite-dimensional vector spaces is Fredholm; in Theorem ? we compute the index of such a map.

Example 1.2. Before actually looking at Theorem ?, consider the example

$$L = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} : \mathbb{C}^2 \to \mathbb{C}^2.$$

One can check that

$$\ker L = \operatorname{span}\left\{\begin{pmatrix}-2\\1\end{pmatrix}\right\} \quad \text{and} \quad \operatorname{Im} L = \operatorname{span}\left\{\begin{pmatrix}1\\2\end{pmatrix}\right\}.$$

Thus, $\dim\ker L = 1$ and $\dim\operatorname{coker} L = \dim(\mathbb{C}^2/\operatorname{Im} L) = 1$, so $\operatorname{ind} L = 1 - 1 = 0$.

Consider now $L = \operatorname{Id} : \mathbb{C}^2 \to \mathbb{C}^2$. Then L is Fredholm with $\ker L = 0$ and $\operatorname{coker} L = \mathbb{C}^2/\operatorname{Im} L = 0$. Thus, $\operatorname{ind} L = 0 - 0 = 0$.

Lastly, consider $L = 0 : \mathbb{C}^2 \to \mathbb{C}^2$ (the zero map). Then L is Fredholm with $\ker L = \mathbb{C}^2$ and $\operatorname{coker} L = \mathbb{C}^2/\operatorname{Im} L = \mathbb{C}^2$, so $\operatorname{ind} L = 2 - 2 = 0$. Thus, we've seen three maps on $\mathbb{C}^2$, each with different kernels and cokernels, yet all have index zero. According to Theorem ?, any linear map $L : \mathbb{C}^2 \to \mathbb{C}^2$ will have index 0.

Example 1.3. In infinite-dimensions, linear maps may or may not be Fredholm. Consider $V = W = C^\infty(\mathbb{R})$, where $C^\infty(\mathbb{R})$ denotes the vector space of smooth, or infinitely differentiable, functions $f : \mathbb{R} \to \mathbb{C}$. Consider the linear map

$$L : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$$

given by $L(f) = \sin x \cdot f$; this map multiplies functions by sine. The map L is certainly a perfectly good (and nontrivial) linear map. Note that $\ker L = 0$. However, L is not Fredholm because, for example, the functions

$$1, x, x^2, x^3, \ldots,$$

define linearly independent elements of $C^\infty(\mathbb{R})/\operatorname{Im} L$ as you can check. Hence, $\operatorname{coker} L = C^\infty(\mathbb{R})/\operatorname{Im} L$ is infinite-dimensional and therefore L is not Fredholm.

Example 1.4. Although not every linear map between infinite-dimensional vector spaces is Fredholm there are many important examples. In this book the *most important* examples are first order differential operators, exactly the operators that appear in the Atiyah-Singer index theorem! Let us consider again $V = W = C^\infty(\mathbb{R})$. Consider the differentiation map

$$L = \frac{d}{dx} : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}),$$

that is, $L(f) = f'$. Note that

$$\ker L = \{f \in C^\infty(\mathbb{R}) \mid f' = 0\} = \{\text{ constant functions }\} = \mathbb{C}.$$

Thus, $\dim\ker L = 1$. We claim that L is surjective. Indeed, given $g \in C^\infty(\mathbb{R})$, define $f(x) = \int_0^x g(t)\,dt$. Then $Lf = g$, so L is surjective and $\dim\operatorname{coker} L = 0$. Hence, L is Fredholm and $\operatorname{ind} L = 1 - 0 = 1$.

Example 1.5. One more example. Let us consider $V = W = C^\infty(\mathbb{S}^1)$, where $C^\infty(\mathbb{S}^1)$ denotes the set of smooth functions on the unit circle $\mathbb{S}^1$, which are just smooth functions $f : \mathbb{R} \to \mathbb{C}$ that are periodic with period 2π. Denote by θ the variable on $\mathbb{S}^1$ and consider the differentiation map

$$L = \frac{d}{d\theta} : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$$

given by $L(f) = f'$. As before, we have

$$\ker L = \{f \in C^\infty(\mathbb{S}^1) \mid f' = 0\} = \{\text{ constant functions }\} = \mathbb{C}.$$

Thus, $\dim\ker L = 1$. In Problem 1 you will prove that $\dim\operatorname{coker} L = 1$ also. Thus, L is Fredholm and $\operatorname{ind} L = 1 - 1 = 0$.

Later we shall see that the numbers 1 and 0 in the previous two examples are exactly the Euler Characteristics of $\mathbb{R}$ and $\mathbb{S}^1$, respectively!

1.1.2. The index the analytical. Before explaining the topological aspect of the index we first explain how the index is an "analytical object". Let $L : V \to W$ be a Fredholm linear map between two complex vector spaces and consider the linear equation

$$Lv = w \quad \text{for } v \in V \text{ and } w \in W. \tag{1.1}$$

Let $v_1, \ldots, v_k$, where $k = \dim \ker L$, be a basis for $\ker L$. Then assuming that (1.1) holds for a v and w, by linearity, for arbitrary constants $a_1, \ldots, a_k \in \mathbb{C}$ we also have

$$L(v + a_1 v_1 + \cdots + a_k v_k) = w.$$

Thus, given that (1.1) holds for a v and w, in order to get a *unique* solution to (1.1), we need to put constraints on the $k = \dim \ker L$ constants $a_1, \ldots, a_k$. Hence we can think of

(1.2) $\dim \ker L$ as the number of constraints needed for uniqueness in (1.1).

Let $w_1, \ldots, w_\ell \in W$ such that $[w_1], \ldots, [w_\ell]$, is a basis for $\operatorname{coker} L = W/\operatorname{Im} L$, where $\ell = \dim \operatorname{coker} L$ and $[\ ]$ denotes equivalence class. Therefore, given any $w \in W$, we can write

$$[w] = b_1[w_1] + \cdots + b_\ell[w_\ell]$$

for some constants $b_1, \ldots, b_\ell$. Now observe that (1.1) just means that $w \in \operatorname{Im} W$, which is equivalent to $[w] = 0$. Hence, given $w \in W$, there exists a solution $v \in V$ to the equation (1.1) if and only if $b_1 = 0, b_2 = 0, \ldots, b_\ell = 0$. In other words, to get *existence* to (1.1), we need to constraint the $\ell = \dim \operatorname{coker} L$ constants $b_1, \ldots, b_\ell$ to vanish. Hence we can think of

(1.3) $\dim \operatorname{coker} L$ as the number of constraints needed for existence in (1.1).

In view of (1.2) and (1.3), we can think of

$\operatorname{ind} L = \dim \ker L - \dim \operatorname{coker} L$ as the net number of constraints needed for existence and uniqueness in (1.1).

Thus, we can consider $\operatorname{ind} L$ as "analytic" because existence and uniqueness of solutions to (differential) equations is a topic usually covered in an analysis course.

1.1.3. The index is topological. We now explain how the index is "topological". To do so, consider the following theorem.

THEOREM 1.1. *Let V and W be finite dimensional vector spaces. Then any linear map $L : V \to W$ is Fredholm, and*

$$\operatorname{ind} L \;=\; \dim V - \dim W. \tag{1.4}$$

PROOF. Observe that if $w_1, \ldots, w_k$ are a basis for $\operatorname{Im} W$, then we can complete this set to a basis $w_1, \ldots, w_k, u_1, \ldots, u_\ell$ of W where $\ell = \dim W - \dim \operatorname{Im} L$. It follows that the equivalence classes $[u_1], \ldots, [u_\ell]$ are a basis of $W/\operatorname{Im} L$, so

$$\dim W/\operatorname{Im} L = \ell = \dim W - \dim \operatorname{Im} L.$$

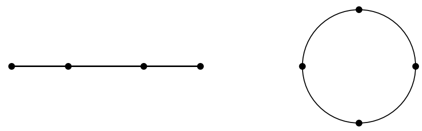

FIGURE 1.1. Computing the Euler Characteristic of $\mathbb{R}$ and $\mathbb{S}^1$.

By the "dimension theorem" or "rank theorem" of linear algebra, we have $\dim V = \dim\ker L + \dim\operatorname{Im} L$. Hence,

$$\begin{aligned}\operatorname{ind} L = \dim\ker L - \dim W/\operatorname{Im} L &= \dim\ker L - (\dim W - \dim\operatorname{Im} L)\\ &= \dim\ker L + \dim\operatorname{Im} L - \dim W\\ &= \dim V - \dim W.\end{aligned}$$

Thus, $\operatorname{ind} L = \dim V - \dim W$, which is a constant, regardless of L. □

This theorem is trivial to prove, but profound to think about: Note that the left-hand side of (1.4) is an *analytic* object (in accordance with our previous discussion), while the right-hand side of (1.4) is completely *topological*, involving the main topological aspect of vector spaces, their dimensions. Thus, we can very loosely interpret (1.4) as giving an equality8

"Analysis = Topology".

Another way the index is topological is that $\operatorname{ind} L$ is an *invariant* of L in the sense that it doesn't change under (particular) deformations of L. We can see this statement explicitly in Theorem 1.1 since the right-hand side of (1.4) is independent of L; see Problem 2 for another example. Compare this invariance with the corresponding situation in topology: Topological aspects of spaces (eg. connectivity, compactness, etc.) are invariant under deformations (homeomorphisms).

Another topological aspect of the index can be seen in Examples 1.4 and 1.5. Let us recall the **Euler characteristic**. Take a pen and mark as many dots on the circle as you wish, then count the number dots (also called vertices) and the number of segments (also called edges) between adjacent dots, then subtract the numbers. This difference is by definition the Euler characteristic of the circle, $\chi(\mathbb{S}^1)$, and it does not depend on the number of dots you mark. In Figure (1.1) we have four dots and four segments, so

$$\chi(\mathbb{S}^1) = 4 - 4 = 0.$$

The Euler characteristic is topological in the sense that any space homeomorphic to the circle also has Euler characteristic zero. Note that the index of $\frac{d}{d\theta}$ in Example 1.5 was also 0!

To define the Euler characteristic of $\mathbb{R}$ we do the same process: draw dots on $\mathbb{R}$ and count vertices and edges. However, we throw away the noncompact edges. For example, in Figure 1.1 we have four vertices and three edges after throwing away the noncompact edges, so

$$\chi(\mathbb{R}) = 4 - 3 = 1.$$

This is exactly the index of the first order differential operator $\frac{d}{dx}$ in Example 1.4!

In conclusion, the index of first order differential operators (at least for the examples we studied) is equal to a topological invariant of the underlying space! This is essentially the Atiyah-Singer index theorem, which is valid for more general manifolds and first order operators called Dirac operators. To give a rough idea what the Atiyah-Singer theorem says, let V and W be certain smooth functions on a space called a vector bundle. (More precisely, they are sections of a vector bundle over a manifold; all this will be discussed in Chapter 2.) We shall study operators like $\frac{d}{dx}$ and $\frac{d}{d\theta}$ already mentioned, and we shall consider such an operator $L : V \to W$ called a Dirac operator, named after the physicist Paul Dirac who first studied them; this will be done in Chapter 3. The Atiyah-Singer index formula states that

$$\operatorname{ind} L = \text{ an expression involving specific topological invariants of the space.}$$

The topological invariants on the right are called characteristic classes. Thus, the Atiyah-Singer index formula is a sense an equality:

$$\text{“Analysis = Topology”}.$$

EXERCISES 1.1.

1. In the notation of Example 2, define
$$C^\infty(\mathbb{S}^1) \ni f \ \mapsto \ \int_0^{2\pi} f(\theta)\, d\theta \in \mathbb{C},$$
Show that this map is zero on $\operatorname{Im} L = \operatorname{Im} \frac{d}{d\theta}$ and hence defines a map from $\operatorname{coker} L \to \mathbb{C}$. Prove that this map defines an isomorphism $\operatorname{coker} L \cong \mathbb{C}$. Hence $\dim \operatorname{coker} L = 1$.
2. In the notation of Example 1.4, define
$$L_{a,b} = a\frac{d}{dx} + b,$$
where $a, b \in \mathbb{C}$. Require $a \neq 0$ in order to keep $L_{a,b}$ a first-order differential operator. Show that this map is Fredholm and $\operatorname{ind} L_{a,b} = 1$ for all $a, b \in \mathbb{C}$ with $a \neq 0$. This gives an example showing that the index is stable under deformations.
3. In this exercise we see that indices of linear maps between fixed infinite-dimensional vector spaces can be arbitrary, in stark contrast to finite-dimensions. Let $V = W = C^\infty(\mathbb{R})$.
   (i) Find a linear map $L : V \to W$ that has index 2. Suggestion: Consider $\frac{d^2}{dx^2}$.
   (ii) Given any positive integer k, find a linear map $L : V \to W$ with index k.
   (iii) Given any negative integer k, find a linear map $L : V \to W$ with index k.

1.2. Heat kernel proof of the linear algebra index theorem

The proof of Atiyah-Singer index formula that we will use is known as the “heat kernel proof”. To illustrate the main ingredients that go into proving the Atiyah-Singer theorem, we will prove Theorem 1.1 using this heat kernel method.[1] Before proceeding, keep Problem 1 at the end of this section in mind!

SECTION OBJECTIVES: THE STUDENT WILL BE ABLE TO ...

- understand the “heat kernel” proof of Theorem 1.1.

[1]Of course, the linear algebra proof of Theorem 1.1 was very easy, and the proof we now give is complicated in comparison, but the heat kernel proof generalizes to infinite dimensions while the proof of Theorem 1.1 does not.

1.2.1. The adjoint matrix. Let V and W be finite-dimensional vector spaces and let $L : V \to W$ be a linear map. For concreteness, we choose any basis of V and W to identify them with products of $\mathbb{C}$. Thus, we can assume $V = \mathbb{C}^n$ and $W = \mathbb{C}^m$ and we can identify L with an $m \times n$ complex matrix. Let $L^* = \overline{L}^t$ denote the adjoint, or conjugate transpose, matrix of L; note that L^* is an $n \times m$ matrix. Properties of the adjoint include

$$(AB)^* = B^* A^* \quad \text{and} \quad (A^*)^* = A \tag{1.5}$$

for any matrices A and B (such that AB is defined). Focusing on L, the adjoint matrix has the property

$$L^* w \cdot v = w \cdot Lv \quad \text{for all } w \in \mathbb{C}^m \text{ and } v \in \mathbb{C}^n, \tag{1.6}$$

where the dot $\cdot$ denotes the standard dot, or inner, product (on $\mathbb{C}^n$ for the left-hand side of (1.6) and on $\mathbb{C}^m$ for the right-hand side of (1.6)). The adjoint matrix also has the property

$$\ker L^* \cong \operatorname{coker} L, \tag{1.7}$$

which you are to prove in Problem 2. Finally, note that since L is $m \times n$ and L^* is $n \times m$, L^*L is an $n \times n$ matrix and LL^* is an $m \times m$ matrix.

1.2.2. The heat operator. The first step in proving Theorem 1.1 is to define the "heat operators" of L^*L and LL^*:

$$e^{-tL^*L}, \; e^{-tLL^*}.$$

Here t is a real variable representing time, e^{-tL^*L} is an $n \times n$ matrix (a linear map on $\mathbb{C}^n$), and e^{-tLL^*} is an $m \times m$ matrix (a linear map on $\mathbb{C}^m$). We now discuss these operators more in depth. For concreteness, in the discussion we will focus on e^{-tL^*L} although similar statements hold for e^{-tLL^*}.

The heat operator e^{-tL^*L} is commonly found in ordinary differential equations texts, cf. [**3**], where it is shown that given $u_0 \in \mathbb{C}^n$, $u(t) := e^{-tL^*L}u_0$ is the unique solution to the initial value problem:

$$\Big(\frac{d}{dt} + L^*L\Big)u(t) = 0; \quad u(0) = u_0.$$

The equation $\Big(\frac{d}{dt} + L^*L\Big)u(t) = 0$ looks like the "heat equation" from mathematical physics that describes the propagation of heat in a material. Therefore it's not unusual to call e^{-tL^*L} a heat operator.

There are two ways to define the heat operator. The first way is by direct exponentiation:

$$e^{-tL^*L} := \sum_{k=0}^{\infty} \frac{(-t)^k}{k!}(L^*L)^k.$$

One can show that this sum is convergent within the $n \times n$ matrices. The second way to define the heat operator is via similar matrices. Note that L^*L is a self-adjoint or Hermitian matrix, which means that it's own adjoint: Using the properties in (1.5) we see that

$$(L^*L)^* = L^*(L^*)^* = L^*L.$$

Being self-adjoint, L^*L is similar to a diagonal matrix with entries given by the eigenvalues of L^*L (this is sometimes part of the "spectral theorem" for self-adjoint matrices). In other words, we can write

$$L^*L = U \begin{bmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_p & \\ & & & \mathbf{0} \end{bmatrix} U^{-1}, \tag{1.8}$$

for some matrix U, where the λ_i's represent the non-zero eigenvalues of L^*L, and where $\mathbf{0}$ represents the 0 matrix. Note that the dimension of the 0 matrix is equal to $\dim\ker L^*L$. We claim that $\dim\ker L^*L = \dim\ker L$. Indeed, if $u \in \ker L^*L$, then from property (1.6) we have

$$0 = 0 \cdot u = (L^*Lu) \cdot u = L^*(Lu) \cdot u = Lu \cdot Lu = \|Lu\|^2,$$

where $\|Lu\|$ denotes the norm of Lu. Thus, $Lu = 0$ and so, $\ker L^*L \subset \ker L$. Since $\ker L \subseteq \ker L^*L$, it follows that $\ker L^*L = \ker L$, and so their dimensions are also equal. Now plugging (1.8) into the series $e^{-tL^*L} := \sum_{k=0}^{\infty} \frac{(-t)^k}{k!}(L^*L)^k$, one can show that

$$e^{-tL^*L} = U \begin{bmatrix} e^{-t\lambda_1} & & & \\ & \ddots & & \\ & & e^{-t\lambda_p} & \\ & & & \mathrm{Id}_{\dim\ker L} \end{bmatrix} U^{-1}, \tag{1.9}$$

where $\mathrm{Id}_{\dim\ker L}$ is the identity matrix of dimension $\dim\ker L^*L = \dim\ker L$. If you're not conformable with the series definition of e^{-tL^*L}, you can take (1.9) as the definition instead.

Switching L and L^*, we can also get an expression

$$e^{-tLL^*} = U \begin{bmatrix} e^{-t\lambda'_1} & & & \\ & \ddots & & \\ & & e^{-t\lambda'_q} & \\ & & & \mathrm{Id}_{\dim\ker L^*} \end{bmatrix} U^{-1},$$

where $\lambda'_1, \ldots, \lambda'_q$ are the non-zero eigenvalues of L^*. Now it is an important fact that L^*L and LL^* have exactly the same non-zero eigenvalues. To verify this, let u be a non-zero eigenvector of L^*L corresponding to some λ_i. Then observe that

$$(LL^*)(Lu) = L(L^*Lu) = L(\lambda_i u) = \lambda_i(Lu).$$

Note that $Lu \neq 0$ since $L^*Lu = \lambda_i u \neq 0$. Thus, Lu is an eigenvector of LL^* with eigenvalue λ_i. Hence, we've shown that L defines a map from the eigenspace of L^*L with eigenvalue λ_i into the eigenspace of LL^* with eigenvalue λ_i. One can check that L^*/λ_i is the inverse to this map. In conclusion, we've shows that if λ_i is a nonzero eigenvalue of L^*L, then λ_i is also an eigenvalue of LL^* and the corresponding eigenspaces are isomorphic. One can also so the reverse: if λ'_i is a nonzero eigenvalue of LL^*, then λ'_i is also an eigenvalue of L^*L and the corresponding eigenspaces are isomorphic. This shows that the nonzero eigenvalues of L^*L and LL^* are the same. In particular, after reordering if necessary, we can write

$$e^{-tLL^*} = U \begin{bmatrix} e^{-t\lambda_1} & & & \\ & \ddots & & \\ & & e^{-t\lambda_p} & \\ & & & \mathrm{Id}_{\dim\ker L^*} \end{bmatrix} U^{-1}. \tag{1.10}$$

1.2.3. The McKean-Singer trick. Recall that the trace of a square matrix is just the sum of the diagonal entries of the matrix. So, for example, if $A = [a_{ij}]$ is a square matrix, then

$$\operatorname{Tr} A := \sum_i a_{ii}$$

The important property of the trace is that it is commutative in the sense that for square matrices A and B of the same dimension,

$$\operatorname{Tr} AB = \operatorname{Tr} BA.$$

In particular, the trace is invariant under conjugation:

$$\operatorname{Tr}(BAB^{-1}) = \operatorname{Tr}(B^{-1}BA) = \operatorname{Tr}(\mathrm{Id}A) = \operatorname{Tr}(A)$$

for any square matrix A and invertible matrix B of the same dimension as A.

Following McKean and Singer [**2**], we consider the function of t:

$$h(t) := \operatorname{Tr}(e^{-tL^*L}) - \operatorname{Tr}(e^{-tLL^*}).$$

This function has some amazing properties.

First, notice that by (1.9) and (1.10) and the conjugation invariant of the trace, we have

$$\begin{aligned} h(t) &= \sum_j e^{-t\lambda_j} + \operatorname{Tr}(\mathrm{Id}_{\dim\ker L}) - \sum_j e^{-t\lambda_j} - \operatorname{Tr}(\mathrm{Id}_{\dim\ker L^*}) \\ &= \sum_j e^{-t\lambda_j} + \dim\ker L - \sum_j e^{-t\lambda_j} - \dim\ker L^* \\ &= \dim\ker L - \dim\ker L^* \\ &= \dim\ker L - \dim\operatorname{coker} L, \end{aligned}$$

where we used (1.7). Thus, $h(t) \equiv \operatorname{ind} L$ for all t.

Second, let us choose a particular t: Since e^{-tL^*L} and e^{-tLL^*} are the identity matrices on $\mathbb{C}^n$ and $\mathbb{C}^m$ respectively at $t = 0$ (see the formulas (1.9) and (1.10)), we have

$$\operatorname{ind} L = h(0) = \operatorname{Tr}(\mathrm{Id}_n) - \operatorname{Tr}(\mathrm{Id}_m) = n - m = \dim V - \dim W.$$

Thus,

$$\operatorname{ind} L = \dim V - \dim W,$$

and Theorem 1.1 is proved.

Exercises 1.2.

1. Dig out your linear algebra book and review properties of linear algebra you weren't familiar with in Section 1.2. These properties are fundamental and should be known!
2. Prove the equality in (1.7): $\ker L^* \cong \operatorname{coker} L$. Suggestion: Observe that $\mathbb{C}^m = \operatorname{Im} L \oplus (\operatorname{Im} L)^\perp$ where $(\operatorname{Im} L)^\perp$ is the orthogonal complement of $\operatorname{Im} L$ in $\mathbb{C}^m$. Thus, $\operatorname{coker} L = \mathbb{C}^m / \operatorname{Im} L \cong (\operatorname{Im} L)^\perp$. Now prove that $\ker L^* = (\operatorname{Im} L)^\perp$.

CHAPTER 2

Manifolds and Riemannian geometry

2.1. Smooth manifolds

Intuitively a smooth manifold is a set which near each point "looks like" an open subset of Euclidean space; in this section we make this intuitive notion precise.

SECTION OBJECTIVES: THE STUDENT WILL BE ABLE TO ...

- define what a smooth manifold is.
- verify that a set with an atlas is a manifold.

2.1.1. Smooth functions. We begin by discussing smooth functions. Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open. A function $f : \mathcal{U} \to \mathbb{R}$ is said to be **smooth** or C^∞ if all partial derivatives of all orders of f exist everywhere. In other words, if $x_1, \ldots, x_n$ are the standard coordinates on $\mathbb{R}^n$, then we require all first partials to exist:

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n},$$

all second partials to exist:

$$\frac{\partial^2 f}{\partial x_1^2}, \frac{\partial^2 f}{\partial x_1 \partial x_2}, \ldots, \frac{\partial^2 f}{\partial x_i \partial x_j},$$

all third, fourth, etc., to exist. We use the same terminology for complex-valued functions, which we shall use a lot later on; thus, $f : \mathcal{U} \to \mathbb{C}$ is smooth or C^∞ if all partial derivatives of all orders of f exist.[1] However, we shall concentrate on real-valued functions for the first section in this chapter.

Example 2.1. Any polynomial $p : \mathcal{U} \to \mathbb{R}$ is smooth. Here, a polynomial is just a finite sum of the form

$$p(x) = a + \sum a_i x_i + \sum a_{ij}\, x_i x_j + \cdots + \sum a_{ij\cdots k}\, x_i x_j \cdots x_k,$$

where the coefficients $a, a_i, \ldots$ are real numbers.

The next examples are useful for constructing partitions of unity in Section 2.2.

Example 2.2. Here is an interesting example of a smooth function. Define

$$f(t) = \begin{cases} e^{-1/t} & \text{if } t > 0, \\ 0 & \text{if } t \le 0. \end{cases}$$

Notice that $f(t)$ is infinitely differentiable for $t \neq 0$, so we just have to check that $f(t)$ is infinitely differentiable at $t = 0$. To this end, observe that

$$\frac{f(t) - f(0)}{t} = \begin{cases} \frac{e^{-1/t}}{t} & \text{if } t > 0, \\ 0 & \text{if } t \le 0. \end{cases}$$

[1]This is equivalent to the smoothness of the real and imaginary parts of f.

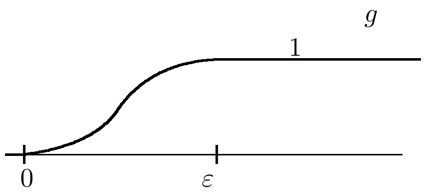

FIGURE 2.1. The function $g(t)$.

By L'Hospital's rule (putting $\tau = 1/t$),

$$\lim_{t \to 0^+} \frac{e^{-1/t}}{t} = \lim_{\tau \to +\infty} \frac{e^{-\tau}}{1/\tau} = \lim_{\tau \to +\infty} \frac{\tau}{e^{\tau}} = \lim_{\tau \to +\infty} \frac{1}{e^{\tau}} = 0.$$

Therefore,

$$\lim_{t \to 0^+} \frac{f(t) - f(0)}{t} = 0 = \lim_{t \to 0^-} \frac{f(t) - f(0)}{t}$$

and it follows that $f(t)$ is differentiable at 0 with derivative zero. Combining this fact with the derivatives of $f(t)$ for $t \neq 0$, we obtain

$$f'(t) = \begin{cases} e^{-1/t}/t^2 & \text{if } t > 0, \\ 0 & \text{if } t \leq 0. \end{cases} \tag{2.1}$$

Continuing with L'Hospital, one can check that all derivatives of $f(t)$ exist at $t = 0$ and equal zero. It is instructive to draw the graph of $f(t)$; it looks something like the graph in Figure 2.1 but $f(t)$ is only asymptotically equal to 1 as $t \to \infty$.

Example 2.3. Let $\varepsilon > 0$. Then with $f(t)$ the nonnegative function in the previous example, consider the function

$$g(t) = \frac{f(t)}{f(t) + f(\varepsilon - t)}.$$

Observe that for any $t \in \mathbb{R}$, either t or $\varepsilon - t$ is positive,[2] thus either $f(t)$ or $f(\varepsilon - t)$ is positive. In particular, the denominator in $g(t)$ is a positive function, so $g(t)$ is well-defined for all $t \in \mathbb{R}$ and (by the quotient rule) is smooth. Moreover, since $f(t) = 0$ for $t \leq 0$, we also have $g(t) = 0$ for $t \leq 0$, and for $t \geq \varepsilon$ we have $\varepsilon - t \leq 0$, so

$$g(t) = \frac{f(t)}{f(t) + f(\varepsilon - t)} = \frac{f(t)}{f(t) + 0} = \frac{f(t)}{f(t)} = 1 \quad \text{for } t \geq \varepsilon;$$

putting these facts together we get the graph of $g(t)$ found in Figure 2.1. Note that in the graph, g is nondecreasing. We can rigourously prove this by computing $g'(t)$:

$$g'(t) = \frac{f'(t)\, f(\varepsilon - t) + f(t)\, f'(\varepsilon - t)}{(f(t) + f(\varepsilon - t))^2}.$$

By the formula (2.1) we have $f'(t) \geq 0$ for all t, so $g'(t) \geq 0$ for all t and hence g is nondecreasing.

Example 2.4. Let $0 < \varepsilon < a$ and define

$$\varphi(t) = g(a - |t|);$$

see Figure 2.2. Then φ is a smooth nonnegative function on $\mathbb{R}$ that vanishes outside

[2]Positive means "> 0", which others may call "strictly positive". Negative means "< 0".

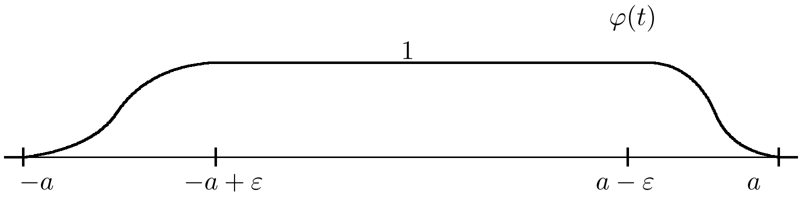

FIGURE 2.2. The function $\varphi(t)$.

the interval $[-a, a]$ and is 1 on $[-a+\varepsilon, a+\varepsilon]$. There are similar functions on $\mathbb{R}^n$: Just define

$$\psi(x) = \varphi(\|x\|), \qquad \text{where } \|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Then ψ is a smooth nonnegative function on $\mathbb{R}^n$ that vanishes outside the ball $\{x \in \mathbb{R}^n \mid \|x\| \le a\}$ and is 1 on the ball $\{x \in \mathbb{R}^n \mid \|x\| \le a - \varepsilon\}$.

Let $\mathcal{U} \subseteq \mathbb{R}^n$ and $\mathcal{V} \subseteq \mathbb{R}^m$ be open and let $f : \mathcal{U} \to \mathcal{V}$. Then writing f in terms of its coordinate functions, we have

$$f(x) = (f_1(x), f_2(x), \dots, f_m(x)),$$

where each $f_j : \mathcal{U} \to \mathbb{R}$. We say that f is smooth or C^∞ if each function $f_j : \mathcal{U} \to \mathbb{R}$ is smooth. We say that $f : \mathcal{U} \to \mathcal{V}$ is a **diffeomorphism**, in which case we say that $\mathcal{U}$ is **diffeomorphic** to $\mathcal{V}$, if f is smooth, a bijection, and $f^{-1} : \mathcal{V} \to \mathcal{U}$ is also smooth.

Example 2.5. The function $f(x) = x^3 : \mathbb{R} \to \mathbb{R}$ is a smooth and a bijection, but $f^{-1}(x) = x^{1/3} : \mathbb{R} \to \mathbb{R}$ is not smooth; hence f is not a diffeomorphism (although it is a homeomorphism).

Example 2.6. We claim that any open ball in $\mathbb{R}^n$ is diffeomorphic to $\mathbb{R}^n$. To see this, let $\mathbb{B}_r(p) = \{x \in \mathbb{R}^n \mid \|x - p\| < r\}$ be an open ball in $\mathbb{R}^n$. Then the map

$$f : \mathbb{B}_r(p) \to \mathbb{R}^n,$$

defined by

$$f(x) := \frac{x - p}{r^2 - \|x - p\|^2}\ , \quad \text{where } \|y\|^2 = y_1^2 + \cdots + y_n^2,$$

is smooth. After some algebra, one can prove that the inverse map is given by

$$f^{-1}(x) = p + \frac{2r^2\, x}{1 + \sqrt{1 + 4r^2\, \|x\|^2}},$$

which is also smooth. Hence $\mathbb{B}_r(p)$ is diffeomorphic to $\mathbb{R}^n$.

With the preliminary material, we are now ready to discuss manifolds.

2.1.2. Coordinate patches. Throughout this section, a nonnegative integer n shall be fixed. We begin with a *set* M; as of this moment M is just a set and has no topology but we'll put a topology on M later. Recall that intuitively a manifold is a set which near each point "looks like" an open subset of Euclidean space. To make this precise we define a coordinate patch. A **coordinate patch** (or **chart** or **system**) is a function $F : \mathcal{U} \to \mathcal{V}$ where

$$\mathcal{U} \subseteq M, \quad \mathcal{V} \subseteq \mathbb{R}^n \text{ is open}, \quad \text{and} \quad F : \mathcal{U} \to \mathcal{V} \text{ is a bijection.}$$

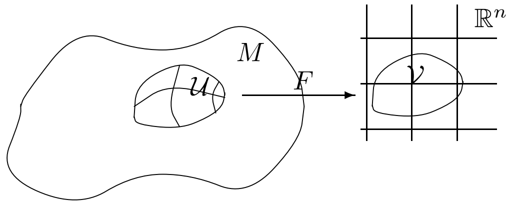

FIGURE 2.3. A patch on a set M. The coordinates on $\mathcal{V}$ give coordinates on $\mathcal{U}$.

See Figure 2.3 is a pictorial representation of this concept. The integer n is *always* fixed for any given M. The idea of a patch is that F "identifies" the set $\mathcal{U} \subseteq M$ with the open subset $\mathcal{V} \subseteq \mathbb{R}^n$ and in this sense M "looks like" an open subset of Euclidean space. We sometimes denote the patch by $(F, \mathcal{U}, \mathcal{V})$ to emphasize the function and sets $\mathcal{U} \subseteq M$ and $\mathcal{V} \subseteq \mathbb{R}^n$. Lastly, we remark that $(F, \mathcal{U}, \mathcal{V})$ is called a *coordinate* patch because $\mathcal{V}$ has Cartesian coordinates from $\mathbb{R}^n$ (represented by the lines on the right in Figure 2.3) to describe the location of its points, so under the "identification" of $\mathcal{U}$ with $\mathcal{V}$, the set $\mathcal{U} \subseteq M$ inherits the coordinates (represented by the curved lines on the left in Figure 2.3).

Example 2.7. Let $M = \mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1\}$, the unit circle. Let

$$\mathcal{U} = \mathbb{S}^1 \setminus \{(0,1)\}, \quad \mathcal{V} = \mathbb{R},$$

and define $F : \mathcal{U} \to \mathcal{V}$ via **stereographic projection** from the point $(0, 1)$. This means that given a point $(x, y) \in \mathbb{S}^1$, we draw the line from $(0, 1)$ through (x, y) and define $F(x, y)$ as the point where the line intersects the x-axis. See the left-hand picture in Figure 2.4. Elementary algebra shows that

$$F(x, y) = \frac{x}{1 - y}.$$

Since F is certainly a bijection (which should be checked), $(F, \mathcal{U}, \mathcal{V})$ is a coordinate patch on $\mathbb{S}^1$. Coordinate patches are certainly not unique and we can define another one, say $(G, \widetilde{\mathcal{U}}, \widetilde{\mathcal{V}})$ by stereographic projecting from, for example, the point $(0, -1)$; see Figure 2.4. Explicitly,

$$\widetilde{\mathcal{U}} = \mathbb{S}^1 \setminus \{(0, -1)\}, \quad \widetilde{\mathcal{V}} = \mathbb{R},$$

and after some algebra,

$$G(x, y) = \frac{x}{1 + y}.$$

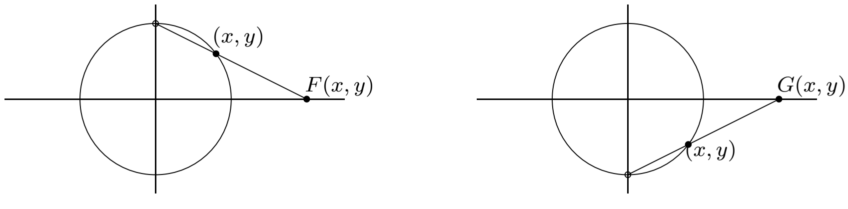

FIGURE 2.4. Stereographic projections from the north and south poles of $\mathbb{S}^1$.

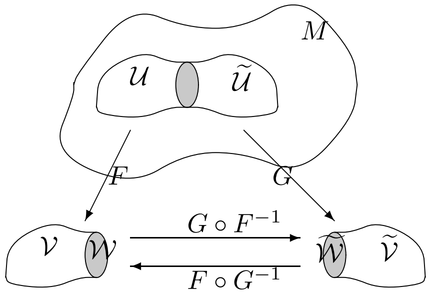

FIGURE 2.5. Compatibility of charts. The shaded "oval" in M is the intersection $\mathcal{U} \cap \widetilde{\mathcal{U}}$.

As we've seen, coordinate patches may not be unique and in particular, it is possible to have coordinate patches that overlap, as they do in the previous example. Our goal is to define *smooth* (C^∞) manifolds so we need to require that when two coordinate patches overlap, they do so in a "smooth fashion." We make this precise as follows. Let $(F, \mathcal{U}, \mathcal{V})$ and $(G, \widetilde{\mathcal{U}}, \widetilde{\mathcal{V}})$ be two coordinate patches on a set M. We say that these patches are (C^∞-)**compatible** if the following three conditions are satisfied:

$$\mathcal{W} := F(\mathcal{U} \cap \widetilde{\mathcal{U}}) \subseteq \mathbb{R}^n \text{ is open}, \quad \widetilde{\mathcal{W}} := G(\mathcal{U} \cap \widetilde{\mathcal{U}}) \subseteq \mathbb{R}^n \text{ is open},$$

and

$$G \circ F^{-1} : \mathcal{W} \to \widetilde{\mathcal{W}}$$

is a diffeomorphism. Make sure you study Figure 2.5 until you understand exactly what's going on here. Note that $\mathcal{W}$ and $\widetilde{\mathcal{W}}$ are open subsets of $\mathbb{R}^n$ so we know what it means for $G \circ F^{-1} : \mathcal{W} \to \widetilde{\mathcal{W}}$ to be a diffeomorphism. Also note that since $(G \circ F^{-1})^{-1} = F \circ G^{-1}$, instead requiring $G \circ F^{-1} : \mathcal{W} \to \widetilde{\mathcal{W}}$ to be a diffeomorphism we could instead require $F \circ G^{-1} : \widetilde{\mathcal{W}} \to \mathcal{W}$ to be a diffeomorphism.

Example 2.8. Let's return to our example $M = \mathbb{S}^1$ with the two patches

$$\mathcal{U} = \mathbb{S}^1 \setminus \{(0,1)\}, \quad \mathcal{V} = \mathbb{R}, \quad F(x,y) = \frac{x}{1-y},$$

and

$$\widetilde{\mathcal{U}} = \mathbb{S}^1 \setminus \{(0,-1)\}, \quad \widetilde{\mathcal{V}} = \mathbb{R}, \quad G(x,y) = \frac{x}{1+y}.$$

Then, see Figure 2.6, $\mathcal{U} \cap \widetilde{\mathcal{U}} = \mathbb{S}^1 \setminus \{(0,1),(0,-1)\}$, and

$$F(\mathcal{U} \cap \widetilde{\mathcal{U}}) = (-\infty, 0) \cup (0, \infty)\ , \quad G(\mathcal{U} \cap \widetilde{\mathcal{U}}) = (-\infty, 0) \cup (0, \infty),$$

which are open. Moreover, observe that

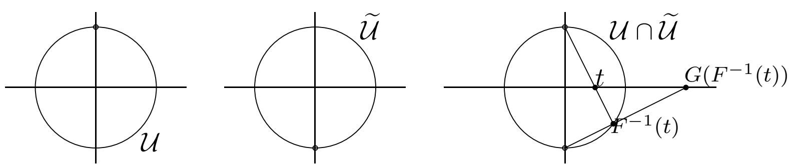

FIGURE 2.6. Overlapping coordinate charts.

$$F(x,y) = \frac{x}{1-y} = \frac{x(1+y)}{1-y^2} = \frac{x(1+y)}{x^2} \quad (\text{since } x^2+y^2=1)$$
$$= \frac{1+y}{x} = \frac{1}{G(x,y)},$$

that is, $F(x,y) = \frac{1}{G(x,y)}$ are just reciprocals! Hence,

$$G \circ F^{-1} : (-\infty,0) \cup (0,\infty) \to (-\infty,0) \cup (0,\infty)$$

is simply

$$G \circ F^{-1}(t) = \frac{1}{t},$$

which is certainly a diffeomorphism. Thus, $(F,\mathcal{U},\mathcal{V})$ and $(G,\widetilde{\mathcal{U}},\widetilde{\mathcal{V}})$ are compatible.

2.1.3. Atlases. A manifold is basically a bunch of patches patched together. To make this precise, we define an **atlas** $\mathcal{A}$ on the set M as a collection of patches

$$\mathcal{A} = \{(F_\alpha, \mathcal{U}_\alpha, \mathcal{V}_\alpha)\}$$

such that the collection $\{\mathcal{U}_\alpha\}$ is a cover of M (that is, $M = \bigcup_\alpha \mathcal{U}_\alpha$) and any two patches in $\mathcal{A}$ are compatible.

Example 2.9. Back to our example $M = \mathbb{S}^1$ with the two compatible patches $(F,\mathcal{U},\mathcal{V})$ and $(G,\widetilde{\mathcal{U}},\widetilde{\mathcal{V}})$ considered in the previous two examples, since $\mathbb{S}^1 = \mathcal{U} \cup \widetilde{\mathcal{U}}$, $\mathcal{A} = \{(F,\mathcal{U},\mathcal{V}), (G,\widetilde{\mathcal{U}},\widetilde{\mathcal{V}})\}$ is an atlas for $\mathbb{S}^1$.

Suppose that $\mathcal{A}$ is an atlas on M. Then we can define a topology on M by declaring a set $\mathcal{W} \subseteq M$ to be open if and only if given any coordinate patch $(F,\mathcal{U},\mathcal{V}) \in \mathcal{A}$, the set $F(\mathcal{W} \cap \mathcal{U}) \subseteq \mathbb{R}^n$ is an open subset of $\mathbb{R}^n$. This is easily checked to define a topology on M (that is, M and $\varnothing$ are open, and open sets are closed under arbitrary unions and finite intersections). Thus, any set with an atlas has a natural topology induced by the coordinate patches.

Example 2.10. For the example $M = \mathbb{S}^1$ with atlas $\mathcal{A} = \{(F,\mathcal{U},\mathcal{V}), (G,\widetilde{\mathcal{U}},\widetilde{\mathcal{V}})\}$ we encourage you to check that the topology on $\mathbb{S}^1$ induced by $\mathcal{A}$ is exactly the standard topology, that is, the relative topology induced from $\mathbb{R}^2$.

As seen in the following proposition, coordinate patches make M locally "look like" subsets of $\mathbb{R}^n$ in the stronger sense of being homeomorphic instead of just being in bijective correspondence.

Proposition 2.1. *Let $(F,\mathcal{U},\mathcal{V})$ be a coordinate patch in an atlas on a set M (with topology induced by the atlas). Then $F : \mathcal{U} \to \mathcal{V}$ is a homeomorphism from $\mathcal{U} \subseteq M$ onto $\mathcal{V} \subseteq \mathbb{R}^n$. Simply put, coordinate patches are homeomorphic to open subsets of $\mathbb{R}^n$.*

Proof. $F : \mathcal{U} \to \mathcal{V}$ is a bijection so we just have to prove that F is a continuous open map. To prove continuity, let $\mathcal{W} \subseteq \mathcal{V}$ be open. To show that $F^{-1}(\mathcal{W})$ is open, by definition of the topology on M, we just need to show that $G(F^{-1}(\mathcal{W}))$ is open in $\mathbb{R}^n$ for any coordinate patch G on M. But by compatibility, $G \circ F^{-1}$ is a diffeomorphism so in particular is a homeomorphism, and hence $G(F^{-1}(\mathcal{W})) = (G \circ F^{-1})(\mathcal{W})$ is open. To see that F is open, let $\mathcal{W} \subseteq \mathcal{U}$ be open; we need to prove that $F(\mathcal{U})$ is open. But by definition of the topology on M, $G(\mathcal{U})$ is open for any coordinate patch G on M; in particular, $F(\mathcal{U})$ is open. Thus, F is a homeomorphism. $\square$

Returning to the atlas $\{(F,\mathcal{U},\mathcal{V}),(G,\widetilde{\mathcal{U}},\widetilde{\mathcal{V}})\}$ on the circle $\mathbb{S}^1$ in Example 2.9, we remark that there are many other coordinate patches we could have chosen to form an atlas; for example, stereographic projection from the point $(1,0)$ onto the y-axis and from $(-1,0)$ onto the y-axis. One can check that these new coordinate patches are compatible with each other and with $(F,\mathcal{U},\mathcal{V})$ and $(G,\widetilde{\mathcal{U}},\widetilde{\mathcal{V}})$. In order to make sure we are not giving preference of one atlas over another one even though the charts in the atlases are perfectly compatible with each other, e.g. choosing stereographic projection from $(0,\pm 1)$ over $(\pm 1,0)$, we always take one "maximal" atlas. An atlas $\mathcal{A}$ is **maximal** if it contains all coordinate patches that are compatible with every element of $\mathcal{A}$; that is, if $(F,\mathcal{U},\mathcal{V})$ is a coordinate patch on M and it is compatible with every element of $\mathcal{A}$, then in fact $(F,\mathcal{U},\mathcal{V})$ is in the atlas $\mathcal{A}$. A maximal atlas is also called a **smooth structure** or C^∞ **structure** on M. For example, the atlas in Example 2.9 is *not* maximal as we already observed. However, we can make any atlas maximal as follows: Given any atlas $\mathcal{A}$ on a set M, we can define

$$\widehat{\mathcal{A}} := \{(F,\mathcal{U},\mathcal{V})\,|\,(F,\mathcal{U},\mathcal{V}) \text{ is compatible with every element of } \mathcal{A}\}.$$

You can readily check that $\widehat{\mathcal{A}}$ is indeed maximal; $\widehat{\mathcal{A}}$ is sometimes called the "completion of $\mathcal{A}$". Because of the compatibility condition, the topology induced by $\widehat{\mathcal{A}}$ is exactly the same as the topology induced by $\mathcal{A}$. With a maximal atlas we can improve Proposition 2.1 as follows.

PROPOSITION 2.2. *Let M be a space with topology induced by a maximal atlas. Then given any open set $\mathcal{W} \subseteq M$ and point $p \in \mathcal{W}$ there is a coordinate patch $F : \mathcal{U} \to \mathbb{R}^n$ in the atlas with $p \in \mathcal{U}$, $\mathcal{U} \subseteq \mathcal{W}$, and $F(p) = 0$. Simply put, in any neighborhood of a point $p \in M$ there is a coordinate patch (in the atlas) homeomorphic to $\mathbb{R}^n$ that identifies $p \in M$ with the origin $0 \in \mathbb{R}^n$.*

PROOF. Given any point $p \in \mathcal{W}$, by definition of atlas we can choose a coordinate patch $F : \mathcal{U}' \to \mathcal{V}'$ in the atlas with $p \in \mathcal{U}'$. Let $q = F(p) \in F(\mathcal{W} \cap \mathcal{U}')$. Since $F(\mathcal{W} \cap \mathcal{U}')$ is open in $\mathbb{R}^n$ (because $\mathcal{W}$ is open in M), there is an open ball $\mathbb{B}_r(q) = \{x \in \mathbb{R}^n \,|\, \|x - q\| < r\} \subseteq F(\mathcal{W} \cap \mathcal{U}')$. Let $\mathcal{U} := F^{-1}(\mathbb{B}_r(q)) \subseteq \mathcal{U}'$ and let

$$G : \mathcal{U} \to \mathbb{B}_r(q)$$

be the restriction of F to $\mathcal{U} \subseteq \mathcal{U}'$. Then G, being a restriction of a coordinate patch, is also coordinate patch. In Example 2.6 we found a diffeomorphism $f : \mathbb{B}_r(q) \to \mathbb{R}^n$ that takes q to 0. Then

$$H : \mathcal{U} \to \mathbb{R}^n \quad \text{defined by} \quad H := f \circ G$$

is a coordinate patch that takes p to 0. Moreover, using that F is compatible with every element of the atlas and that f is a diffeomorphism, it's easy to check that H is compatible with every element of the atlas. Hence by maximality, H is in the atlas. By Proposition 2.1, H defines a homeomorphism of $\mathcal{U}$ onto $\mathbb{R}^n$. □

2.1.4. Smooth manifolds.

Recall that a topological space X is **Hausdorff** means that given any distinct points $p, q \in X$ there are disjoint open sets $\mathcal{U}, \mathcal{V} \subseteq X$ such that $p \in \mathcal{U}$ and $q \in \mathcal{V}$. For example, $\mathbb{R}^n$ is Hausdorff.

We are now ready to define a smooth manifold, the most fundamental definition in this course:[3]

[3]Requiring a covering by countably many coordinate patches is equivalent to M being **second countable**, which means that M has a countable basis. Explicitly, there is a countable cover $\{\mathcal{U}_i\}$

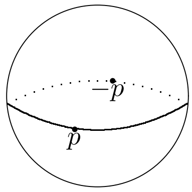

FIGURE 2.7. The points p and $-p$ are identified in real projective space.

A **(smooth) manifold** is a pair $(M, \mathcal{A})$ where M is a set and $\mathcal{A}$ is a maximal atlas on M such that M can be covered by countably many coordinate patches in $\mathcal{A}$ and the topology induced by the atlas is Hausdorff.

Most of the time we call M the manifold with the atlas implicit. The integer n in the patches (and fixed for a given M) is called the **dimension** of M, denoted $n = \dim M$, and we say that M is an n**-dimensional manifold**.

Example 2.11. Consider one last time $M = \mathbb{S}^1$ with atlas the completion of $\mathcal{A} = \{(F,\mathcal{U},\mathcal{V}),(G,\widetilde{\mathcal{U}},\widetilde{\mathcal{V}})\}$. Since we stated that the induced topology is just the relative topology from $\mathbb{R}^2$, $\mathbb{S}^1$ is Hausdorff (one can also check the Hausdorff condition using the definition of the topology induced by $\mathcal{A}$ without reference to the relative topology). Moreover, $\mathbb{S}^1$ is covered by $\mathcal{U}$ and $\widetilde{\mathcal{U}}$, hence $\mathbb{S}^1$ is a one-dimensional manifold! It took forever just to get to this point didn't it!

We remark that whenever we are constructing manifolds we look for atlases with as few elements as possible; only later we complete the atlas so that we are free then to use any other coordinates compatible with the original ones.

Example 2.12. In this example we work out in painful detail (two-dimensional) real projective space; this example shows that in general there is quite a bit to check when proving that something is a manifold, at least if we do everything by first principles! We define two-dimensional **real projective space** as the set of equivalence classes:

$$\mathbb{R}P^2 := \{[p] \mid \text{where } p \in \mathbb{S}^2 \text{ is equivalent to } -p \in \mathbb{S}^2\}.$$

In other words, as seen in Figure 2.7, we identify opposite points of the sphere: $p \in \mathbb{S}^2$ and $-p \in \mathbb{S}^2$ are identified. We can define a manifold structure on $\mathbb{R}P^2$ as follows. First, we let

$$\mathcal{U}_1 = \{[p] \mid p_1 \neq 0\}\ , \quad \mathcal{U}_2 = \{[p] \mid p_2 \neq 0\}\ , \quad \mathcal{U}_3 = \{[p] \mid p_3 \neq 0\},$$

where we write p as $p = (p_1, p_2, p_3)$, and $\mathcal{V}_1 = \mathcal{V}_2 = \mathcal{V}_3 = \mathbb{R}^2$. Second, we define functions

$$F : \mathcal{U}_1 \to \mathcal{V}_1\ , \quad G : \mathcal{U}_2 \to \mathcal{V}_2\ , \quad H : \mathcal{U}_3 \to \mathcal{V}_3\ ,$$

---

of open subsets of M having the property that any given open set $\mathcal{U} \subseteq M$ is a union of some $\mathcal{U}_i$'s. It is standard to define a manifold using the second countability condition, but we shall *never* explicitly use this condition; however, we shall repeatedly use the equivalent condition on having a countable coordinate cover. Thus, we err to the side of usefulness rather than convention ☺.

by

$$F\big([(x,y,z)]\big) := \left(\frac{y}{x}, \frac{z}{x}\right) \ , \quad G\big([(x,y,z)]\big) := \left(\frac{x}{y}, \frac{z}{y}\right) \ , \quad H\big([(x,y,z)]\big) := \left(\frac{x}{z}, \frac{y}{z}\right).$$

Note that $[(-x,-y,-z)] = [(x,y,z)]$, and

$$F\big([(-x,-y,-z)]\big) = \left(\frac{-y}{-x}, \frac{-z}{-x}\right) = \left(\frac{y}{x}, \frac{z}{x}\right) = F\big([(x,y,z)]\big),$$

so F is well-defined. For similar reasons, G and H are well-defined. Also notice that $\mathbb{R}P^2 = \mathcal{U}_1 \cup \mathcal{U}_2 \cup \mathcal{U}_3$, so F, G, H define an atlas for $\mathbb{R}P^2$ if we can show that F, G, H are pairwise compatible coordinate patches. Let's show that F is a coordinate patch; checking G and H are analogous.

To see that F is one-to-one, assume that $F\big([(x,y,z)]\big) = F\big([(x',y',z')]\big)$. Replacing (x,y,z) with $(-x,-y,-z)$ and (x',y',z') with $(-x',-y',-z')$ if necessary, we may assume that $x, x' > 0$. Now, $F\big([(x,y,z)]\big) = F\big([(x',y',z')]\big)$ means that

$$\left(\frac{y}{x}, \frac{z}{x}\right) = \left(\frac{y'}{x'}, \frac{z'}{x'}\right), \quad \text{that is,} \quad \frac{y}{x} = \frac{y'}{x'} \ , \quad \frac{z}{x} = \frac{z'}{x'}. \tag{2.2}$$

Since $(x,y,z), (x',y',z') \in \mathbb{S}^2$ we have

$$1 = x^2 + y^2 + z^2 \ , \quad 1 = (x')^2 + (y')^2 + (z')^2.$$

Dividing the first equation by x and the second by x' we see that

$$\frac{1}{x^2} = 1 + \left(\frac{y}{x}\right)^2 + \left(\frac{z}{x}\right)^2 \ , \quad \frac{1}{(x')^2} = 1 + \left(\frac{y'}{x'}\right)^2 + \left(\frac{z'}{x'}\right)^2.$$

In view of (2.2) (that $\frac{y}{x} = \frac{y'}{x'}$ and $\frac{z}{x} = \frac{z'}{x'}$), we conclude that $\frac{1}{x^2} = \frac{1}{(x')^2}$, which implies that $x = x'$ since $x, x' > 0$. Again using (2.2) and the fact that $x = x'$, we see that $y = y'$ and $z = z'$. Hence, $[(x,y,z)] = [(x',y',z')]$ and F is one-to-one.

To see that F is onto, let $(s,t) \in \mathbb{R}^2$. Define

$$p := \frac{(1,s,t)}{\rho}, \quad \text{where } \rho = \|(1,s,t)\| = \sqrt{1+s^2+t^2}.$$

Certainly $p \in \mathbb{S}^2$, $[p] \in \mathcal{U}_1$, and by definition of F,

$$F\big([p]\big) = \left(\frac{s/\rho}{1/\rho}, \frac{t/\rho}{1/\rho}\right) = (s,t). \tag{2.3}$$

Hence, F is onto. Thus, $F : \mathcal{U}_1 \to \mathcal{V}_1$ is a bijection, so is a coordinate patch.

We now show that F, G, and H are pairwise compatible. For concreteness, let's show that F and G are compatible. First, we leave you to verify that since

$$\mathcal{U}_1 \cap \mathcal{U}_2 = \big\{[(x,y,z)] \mid x \neq 0 \, , \, y \neq 0\big\},$$

we have

$$F(\mathcal{U}_1 \cap \mathcal{U}_2) = \mathbb{R}^2 \setminus \big(\{0\} \times \mathbb{R}\big) \ , \quad G(\mathcal{U}_1 \cap \mathcal{U}_2) = \mathbb{R}^2 \setminus \big(\{0\} \times \mathbb{R}\big),$$

which shows that both $F(\mathcal{U}_1 \cap \mathcal{U}_2)$ and $G(\mathcal{U}_1 \cap \mathcal{U}_2)$ are open. Second, we need to show that $G \circ F^{-1}$ is a diffeomorphism. To this end, let $(s,t) \in \mathbb{R}^2 \setminus \big(\{0\} \times \mathbb{R}\big)$. Then by (2.3), we have

$$F^{-1}(s,t) = [p] \ , \quad \text{where } p = \frac{(1,s,t)}{\rho} \text{ with } \rho = \|(1,s,t)\| = \sqrt{1+s^2+t^2}.$$

Hence, by definition of G,

$$(G \circ F^{-1})(s,t) = G\big([p]\big) = \left(\frac{1/\rho}{s/\rho}, \frac{t/\rho}{s/\rho}\right) = \left(\frac{1}{s}, \frac{t}{s}\right).$$

This is a smooth function on $\mathbb{R}^2 \setminus \big(\{0\} \times \mathbb{R}\big)$. Moreover, it's its own inverse:

$$(G \circ F^{-1})^{-1}(s,t) = \left(\frac{1}{s}, \frac{t}{s}\right).$$

Therefore, $G \circ F^{-1}$ is a diffeomorphism and so G and F are compatible.

Summarizing what we've done so far: We've defined three coordinate patches on $\mathbb{R}P^2$ that cover $\mathbb{R}P^2$ and shown that they are compatible. Thus, we've defined an atlas on $\mathbb{R}P^2$. Since this atlas contains a finite number of patches, $\mathbb{R}P^2$ is a manifold once we verify that the induced topology is Hausdorff.

To prove that $\mathbb{R}P^2$ is Hausdorff, let $[p], [q] \in \mathbb{R}P^2$ be distinct. If $[p]$ and $[q]$ happen to lie in a single coordinate patch, say $\mathcal{U}_1$ for example, then, since $\mathcal{U}_1 \cong \mathbb{R}^2$ and $\mathbb{R}^2$ is Hausdorff, we can certainly separate $[p]$ and $[q]$ by open sets. Thus, we may assume that $[p]$ and $[q]$ cannot lie in a single patch. By definition of $\mathcal{U}_1$, $\mathcal{U}_2$, and $\mathcal{U}_3$, this can happen only for the points $[(1,0,0)]$, $[(0,1,0)]$, and $[(0,0,1)]$. For concreteness let's assume that $[p] = [(1,0,0)]$ and $[q] = [(0,1,0)]$. Define

$$\mathcal{U} := \{[(x,y,z)] \,|\, x^2 > 1/2\} \quad , \quad \mathcal{V} := \{[(x,y,z)] \,|\, y^2 > 1/2\}.$$

Note that $[p] \in \mathcal{U}$, $[q] \in \mathcal{V}$. Also, since $x^2 + y^2 + z^2 = 1$ it is impossible for $x^2 > 1/2$ and $y^2 > 1/2$ to both hold, so $\mathcal{U} \cap \mathcal{V} = \varnothing$. Therefore, once we show that $\mathcal{U}$ and $\mathcal{V}$ are both open sets in $\mathbb{R}P^2$, we can conclude that $\mathbb{R}P^2$ is Hausdorff and finally we can conclude that $\mathbb{R}P^2$ is a manifold.

Let us prove that $\mathcal{U}$ is open, which by the definition of the topology on $\mathbb{R}P^2$ means that $F(\mathcal{U})$, $G(\mathcal{U})$, and $H(\mathcal{U})$ are each open. Let's take for example $H(\mathcal{U})$. By definition of H we have

$$H(\mathcal{U}) = \left\{ \left(\frac{x}{z}, \frac{y}{z}\right) \;\middle|\; x^2 > \frac{1}{2} \right\}.$$

Let $s = \frac{x}{z}$ and $t = \frac{y}{z}$. Then

$$x^2 + y^2 + z^2 = 1 \quad \Longrightarrow \quad \frac{1}{z^2} = 1 + \frac{x^2}{z^2} + \frac{y^2}{z^2} \quad \Longrightarrow \quad \frac{1}{z^2} = 1 + s^2 + t^2,$$

and

$$s = \frac{x}{z} \quad \Longrightarrow \quad x^2 = z^2\, s^2 \quad \Longrightarrow \quad x^2 = \frac{s^2}{1 + s^2 + t^2}.$$

Therefore,

$$H(\mathcal{U}) = \left\{ (s,t) \;\middle|\; \frac{s^2}{1 + s^2 + t^2} > \frac{1}{2} \right\}.$$

This is certainly open. Analogous arguments show that $F(\mathcal{U})$ and $G(\mathcal{U})$ are open.

EXERCISES 2.1.

1. Consider the unit n-sphere:

$$M = \mathbb{S}^n = \{(x_1, x_2, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} \,|\, x_1^2 + x_2^2 + \cdots + x_{n+1}^2 = 1\}.$$

Let $\mathcal{U} = \mathbb{S}^n \setminus \{(0,0,\ldots,0,1)\}$, $\widetilde{\mathcal{U}} = \mathbb{S}^n \setminus \{(0,0,\ldots,0,-1)\}$, $\mathcal{V} = \widetilde{\mathcal{V}} = \mathbb{R}^n$, and define $F : \mathcal{U} \to \mathcal{V}$ and $G : \widetilde{\mathcal{U}} \to \widetilde{\mathcal{V}}$ as the stereographic projections from the north and south poles, respectively, just like in Figure 2.4; here the vertical axis is the x_{n+1}-axis.

(i) Show that F and G are given by the formulas

$$F(x_1,\dots,x_{n+1}) = \frac{(x_1,\dots,x_n)}{1-x_{n+1}}\ , \quad G(x_1,\dots,x_{n+1}) = \frac{(x_1,\dots,x_n)}{1+x_{n+1}}$$

(ii) Prove that F and G define coordinate patches on $\mathbb{S}^n$.
(iii) Prove that F and G form an atlas on $\mathbb{S}^n$.
(iv) Prove that $\mathbb{S}^n$ is a manifold.

2. Consider real projective n-space:

$$M = \mathbb{R}P^n := \big\{ [p] \mid \text{where } p \in \mathbb{S}^n \text{ is equivalent to } -p \in \mathbb{S}^n \big\}.$$

For each $j = 1,\dots,n+1$, put

$$\mathcal{U}_j := \big\{[p] \mid p_j \neq 0\big\},$$

where we write p as $p = (p_1,\dots,p_{n+1})$. For each j, let $\mathcal{V}_j = \mathbb{R}^n$ and define $F_j : \mathcal{U}_j \to \mathcal{V}_j$ by

$$F_j\big([(x_1,\dots,x_{n+1})]\big) := \left(\frac{x_1}{x_j}, \frac{x_2}{x_j},\dots,\frac{x_{j-1}}{x_j},\frac{x_{j+1}}{x_j},\dots,\frac{x_{n+1}}{x_j}\right). \tag{2.4}$$

For example,

$$F_1\big([(x_1,\dots,x_{n+1})]\big) := \left(\frac{x_2}{x_1},\frac{x_3}{x_1},\dots,\frac{x_{n+1}}{x_1}\right).$$

and

$$F_2\big([(x_1,\dots,x_{n+1})]\big) := \left(\frac{x_1}{x_2},\frac{x_3}{x_2},\dots,\frac{x_{n+1}}{x_2}\right).$$

(i) Show that for each j, F_j is well-defined; that is, the definition (2.4) doesn't depend on the equivalence class of $(x_1,\dots,x_{n+1})$.
(ii) Prove that each F_j defines a coordinate patch on $\mathbb{R}P^n$.
(iii) Prove that $\mathcal{A} = \{F_1,\dots,F_{n+1}\}$ is an atlas on $\mathbb{R}P^n$. (If you wish to simplify things for yourself just prove that F_1 and F_2 are compatible.)
(iv) Prove that $\mathbb{R}P^n$ is a manifold.

3. If M is an n-dimensional manifold in $\mathbb{R}^n$ and M' is an n'-dimensional manifold, prove that $M \times M'$ is an $(n+n')$-dimensional manifold. In fact, prove that given any finite number of manifolds $M_1,\dots,M_k$ of dimensions $n_1,\dots,n_k$, respectively, the product $M_1 \times \cdots \times M_k$ is an $(n_1+\cdots+n_k)$-dimensional manifold. In particular, the n-**torus**,

$$\mathbb{T}^n := \mathbb{S}^1 \times \mathbb{S}^1 \times \cdots \times \mathbb{S}^1, \quad \text{a product of } n \text{ circles},$$

is an n-dimensional manifold.

4. Prove that a manifold can be covered by countably many coordinate patches, each of which is homeomorphic to $\mathbb{R}^n$.

5. Here is a simple example of a set with an atlas that is almost a one-dimensional manifold except for the Hausdorff condition. Let $0'$ denote an object that is not a real number and let $M = \mathbb{R} \cup \{0'\}$. Let $\mathcal{U} = \mathbb{R} \subseteq M$, $\widetilde{\mathcal{U}} = (-\infty,0)\cup(0,\infty)\cup\{0'\} \subseteq M$ and let $\mathcal{V} = \widetilde{\mathcal{V}} = \mathbb{R}$. Define

$$F : \mathcal{U} \to \mathcal{V} \ \text{ by } \ F(x) = x \ \text{ for all } x \in \mathcal{U},$$

$$G : \widetilde{\mathcal{U}} \to \widetilde{\mathcal{V}} \ \text{ by } \ G(x) = \begin{cases} x & \text{if } x \in (-\infty,0)\cup(0,\infty), \\ 0 & \text{if } x = 0'. \end{cases}$$

Show that $\mathcal{A} = \{F, G\}$ is an atlas on M (the "real line with two origins") but the induced topology is not Hausdorff.