Lecture Notes on Differential Geometry

Jay Wilkins

August 19, 2008

Contents

Preface

This is a collection of lecture notes on differential geometry, focusing primarily on Riemannian geometry. They were compiled over a span of several years from varying sources. Consequently, they are, at present, somewhat uneven. A large portion of the material was put together while doing the research for my Master's Degree at Clemson University, under Dr. James Peterson. About half of the material came from a course in Riemannian geometry, taught by Dr. Conrad Plaut at the University of Tennessee. The rest of the material is pieced together from my own personal notes. I owe Dr. Peterson and Dr. Plaut a great deal for their support and help in my pursuit of this subject.

Since these are lecture notes and not a textbook, the reader is expected to have knowledge of basic linear algebra, particularly of linear transformations between finite dimensional vector spaces, change of basis, etc, as well as a strong grasp of real analysis on Euclidean spaces. Specifically, the use of the differential calculus is of primary importance, including knowledge of the Inverse and Implicit Function Theorems and the Rank Theorem. While these results will be essential to the material here, their proofs will not. So the reader willing to take them on faith can find them in [Ba], [Bo], or [Sp,2]. In addition, these notes take a decidedly topological point of view, so a certain degree of knowledge in that area will be assumed. The necessary results are not difficult or deep, and come primarily from the concepts of connectedness, compactness, separation, and continuity. At some point, we will use the notion of paracompactness, but the results in this section will be given with proof, since they are not standard. The standard topological reference for this material will be [Mu]. I understand that this may deter some aspiring differential geometers who do not yet have the necessary background in topology. The fact, however, is that if one wishes to pursue this subject, the topology must be learned at some point, as it is an indispensable tool in geometry. Moreover, the prevalent use of topology makes the proofs much more rigorous and formal.

These notes begin with the most basic fundamentals of topological manifolds, proceeding to smooth manifolds and their properties. After the definitions are presented, several examples of smooth manifolds are given. This list, and this collection of notes in general, is still growing, so the notes will be updated periodically. My hope is to eventually include in these notes a catalog of useful and interesting and examples of Riemannian manifolds. Local structures, such as the tangent space, are discussed in great detail before any global results are pursued. In particular, I give detailed and independent constructions of the tangent and cotangent spaces, showing, only afterward, that there is a natural duality between the two. I then proceed to global structures such as vector fields, the tangent bundle, tensor fields, metric tensors, etc. In particular, there are detailed constructions of the Riemannian and Lorentzian metric tensors, as well as the standard proof that a Riemannian manifold is an inner metric space.

After this metric space structure is developed, the more technical constructions of differential geometry are presented, which, of course, leads into the concept of curvature. The study of curvature and how it affects various other properties of a manifold lies at the heart of modern differential geometry. These notes are still being added as time permits me.

As a final remark, I should inform the reader ahead of time that my choice of presentation, notation, and organization, as it is for all mathematicians, is influenced by the work of those mathematicians whose research and textbooks I have personally found to be most accessible. In particular, the fundamental definitions given here of smooth manifolds, smooth mappings, the tangent and cotangent spaces, etc., are adapted from the work of S.S. Chern and William Boothby (see [Bo] and [Ch] in the bibliography). The later concepts and constructions, including geodesics, convexity, curvature, Jacobi fields, etc, are heavily influenced by Barret O'Neill's book [ON,2] and the course notes of Dr. Plaut. I have added my own touches in various places, and I have arranged the material according to my own taste, which, of course, may not be the preference of others. Such is the nature, I'm afraid, of the subject of Differential Geometry. I have, however, been careful to unify the notation and terminology. So, even though the material is drawn from different sources, it is presented here in these notes in a uniform way.

1 Manifolds and Smooth Functions

We begin by introducing the formal concept of a manifold. Roughly speaking, a manifold is a topological space that is locally Euclidean. More formally, we have the following.

Definition 1.1 *A* **topological manifold** *is a second-countable Hausdorff space M such that for every $p \in M$, there is a neighborhood, U, of p, a natural number m, and a corresponding map $\varphi_U : U \to \mathbb{R}^m$ with the property that φ_U is a homeomorphism of U onto an open subset of $\mathbb{R}^m$.*

First, there are some remarks on notation and terminology. As we have done here, we will not reference the specific topology on M unless it is necessary. The existence of the topology is the primary concern and will usually, but not always, be assumed. Second, as is common in topology, we will typically refer to a topological manifold merely as a *manifold*. We will introduce a different class of manifolds later on, so, occasionally, we will use the full name to distinguish topological manifolds from others.

The sets U are called **coordinate neighborhoods**, and their corresponding mappings φ_U are called **coordinate mappings**. Each pair (φ_U, U) is called a **coordinate chart**, or a **coordinate system**, and the collection of all such pairs is called an **atlas**. Note that the collection of coordinate neighborhoods forms an open covering of M. We will occasionally use the word "covering" when referring to an atlas, even though technically the elements of an atlas are pairs consisting of neighborhoods and corresponding mappings.

Two charts are equal if and only if their domains and mappings are identical. We will usually drop the subscript U from the map φ_U, since we will almost always reference a chart in the form (φ, U), from which it should be clear that φ is the coordinate mapping corresponding to the coordinate neighborhood U. For brevity, we will often simply call the pair (φ, U) a **chart around** p or **at** p to denote the fact that (φ, U) is a coordinate chart with $p \in U$.

The reason for this terminology is that these charts can be used to establish local coordinate systems around any point $p \in M$. If $p \in M$, then there is some chart (φ, U) such that $p \in U$ and $\varphi(U)$ is a neighborhood of $\varphi(p)$ in $\mathbb{R}^m$ for some m. (We take *neighborhood* here to mean an open set, and not, as some define it, as a set *containing* an open set.) We can define the coordinates of $q \in U$ to be the Euclidean coordinates of $\varphi(q) \in \varphi(U)$. It is in this respect that we describe a manifold as locally Euclidean. Note also that if (φ, U) is any chart around p, then the restriction of φ to any open subset, V, of U gives us a new coordinate chart, namely $(\varphi|_V, V)$, since this restriction gives us a homeomorphism of V onto $\varphi(V) \subset \varphi(U)$. More generally, if V is open in M and $U \cap V \neq \emptyset$, then $U \cap V$ is open in U and $(\varphi|_{U\cap V}, U \cap V)$ is a coordinate chart.

The **dimension** of a manifold, M, at a point $p \in M$, is defined to be the dimension of the Euclidean space containing $\varphi(U)$, where (φ, U) is a chart such that $p \in U$. This definition brings us to our first instance of *coordinate independence*. The point p may lie in several different coordinate neighborhoods, so in order to show that dimension is a well-defined concept, we must show that each one is homeomorphic to a subset of the same Euclidean space. So, suppose there are two coordinate charts, (φ, U) and (ψ, V) such that $p \in U \cap V$, and suppose $\varphi(U) \subset \mathbb{R}^n$ and $\psi(V) \subset \mathbb{R}^m$. Then $U \cap V$ is a neighborhood of p on which we have two coordinate mappings, φ and ψ. Moreover, $\psi(U \cap V)$ and $\varphi(U \cap V)$ are open subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. The mapping $\varphi \circ \psi^{-1} : \psi(U \cap V) \to \varphi(U \cap V)$ is a homeomorphism, as it is a composition of homeomorphisms. That is, $\varphi \circ \psi^{-1}$ maps an open subset of $\mathbb{R}^m$ homeomorphically to an open subset of $\mathbb{R}^n$. Two open subsets of Euclidean spaces cannot be homeomorphic unless the Euclidean spaces are of the same dimension. (This result is known as *Invariance of Domain.*) Hence, we must have $m = n$, and the dimension of a manifold is well-defined, meaning that it does not depend on the particular coordinate system chosen around a point. While local coordinate systems are useful for many purposes, a fundamental goal of differential geometry is to derive and express results that are independent of any particular coordinate system.

The dimension of a manifold is only defined locally, and a topological manifold, as we have defined it here, need not be of constant dimension. For example, the subspace $M = \{(x, y) \in \mathbb{R}^2 \ : \ 0 < x < 1\}$ is an open subset of $\mathbb{R}^2$. It is also a manifold, since given any point $(x, y) \in M$, we can let M, itself, be a coordinate neighborhood, and the coordinate map, then, is simply the inclusion map of M into $\mathbb{R}^2$. This also shows that M is of constant dimension, since every point in M has a neighborhood homeomorphic to an open set in $\mathbb{R}^2$. Similarly, the subspace $N = M \cup \{(x, y) \in \mathbb{R}^2 : x = 2\}$ is also a manifold, though not of constant dimension. Every point in M has a neighborhood homeomorphic to an open set in $\mathbb{R}^2$, while every point in $\{(x, y) \in \mathbb{R}^2 : x = 2\}$ has a neighborhood (in the subspace topology) homeomorphic to an open interval in $\mathbb{R}$.

In geometry, physics, and other applications, most manifolds are of constant dimension, and those that are not, like N above, are usually pathological and only serve as counter-examples. Indeed, in this example, N was disconnected. A connected manifold, M, must be of constant dimension. To see this, note that we can define $f : M \to \mathbb{R}$ by letting $f(p)$ be the dimension of M at p. Since we've shown the dimension to be a well-defined concept, f is well-defined. Moreover, f is continuous, for if $p \in M$, there is some coordinate neighborhood, U, around p, and f is constant, and therefore continuous, on U. Since the coordinate neighborhoods cover M, f is continuous. It now follows that $f(M)$ is a connected subspace of $\mathbb{R}$. But f takes values only in $\mathbb{N}$, and the only connected subspaces of $\mathbb{N}$ are the one-point sets. So, f, and the dimension of M, must be constant.

While we will not universally require that our manifolds be connected, *we will assume from here on that they are of constant dimension.* Consequently, we will use the following notational conventions when referring to manifolds. When we denote a manifold by capital letters M, N, etc, this is meant to imply that these manifolds are of constant dimension m, n, etc respectively. Occasionally, it will be convenient to use the same capital letter to denote multiple manifolds. In this case, we will always number them and denote the dimension with a superscript. For example, M_1^m and M_2^n will denote manifolds of dimension m and n, respectively.

We will now prove a series of topological properties of manifolds which will be useful in later arguments.

Lemma 1.1 *A manifold, M, is locally compact.*

<u>Proof</u> Since M is Hausdorff, it suffices to show that for any point, $p \in M$, and any neighborhood, V, of p, there is a neighborhood, W, of p, so that $\overline{W}$ is compact and $\overline{W} \subset V$.

So, let p be any point in M, and let V be any neighborhood of p. There is a coordinate chart (φ, U) such that $p \in U$, and, as we have seen, by restricting φ to an appropriate intersection, we can assume that $U \subset V$. Since $\varphi(U)$ is an open subset of $\mathbb{R}^m$ containing $\varphi(p) = x$, there is some $\epsilon > 0$ such that $\overline{B_\epsilon(x)} \subset \varphi(U)$. Now, $\overline{B_\epsilon(x)}$ is compact both in $\mathbb{R}^m$ and in the subspace $\varphi(U)$. Thus, $\varphi^{-1}(\overline{B_\epsilon(x)})$ is compact in the subspace $U \subset M$. It is also compact in M, for if $\{G_\lambda\}_{\lambda \in \Lambda}$ is an open cover of $\varphi^{-1}(\overline{B_\epsilon(x)})$ by sets open in M, then the collection $\{G_\lambda \cap U\}_{\lambda \in \Lambda}$ is an open cover of $\varphi^{-1}(\overline{B_\epsilon(x)})$ by sets open in U. This latter collection has, then, a finite subcover, say $\{G_k \cap U\}_{k=1}^s$, implying that the collection $\{G_k\}_{k=1}^s$ is a finite subcover from the original collection.

Now, let $W = \varphi^{-1}(B_\epsilon(x))$. Let $\overline{W}$ and $cl(W)$ denote the closures of W in M and U, respectively. Note that the closure of $B_\epsilon(x)$ in $\varphi(U)$ is just $\overline{B_\epsilon(x)}$. Since φ is a homeomorphism, we have $\varphi^{-1}(\overline{B_\epsilon(x)}) = cl(\varphi^{-1}(B_\epsilon(x))) = cl(W)$. Thus, $cl(W)$ is compact in M, and so is also closed in M. It follows that $\overline{W} \subset cl(W)$. Conversely, we have $cl(W) = \overline{W} \cap U$, implying that $cl(W) \subset \overline{W}$. Thus, $cl(W) = \overline{W}$, and $\overline{W}$ is compact in M. Finally, since $cl(W) \subset U \subset V$, the result follows.

QED

Theorem 1.2 *A manifold, M, is normal, separable, metrizable, and is a Lindelöf space. Furthermore, it is locally path connected (therefore locally connected), and M is connected if and only if any two points in M can be joined by a continuous curve. That is, M is connected if and only if it is path connected.*

Proof Every regular space with a countable basis is normal, while every locally compact Hausdorff space is regular. Thus, by the previous lemma and the second countability hypothesis in Definition 1.1, M is normal.

Now, every second countable space is both separable and Lindelöf, but we will prove these less obvious results directly. By hypothesis, there is a countable basis $\{B_n\}_{n\geq 1}$ for M. For each n, let p_n be a point in B_n. Define $D = \{p_n : p_n \in U_n, n \geq 1\}$. Given any nonempty open set $U \subset M$, let q be in U. Then there is some basis set, B_n, such that $q \in B_n \subset U$. It follows that $p_n \in U \Rightarrow U \cap D \neq \emptyset$. Hence, D intersects every open set in M, implying that D is a countable dense subset of M. Thus, M is separable.

Next, let S be an open covering of M. For each $p \in M$, there is an open set, $U_p \in S$ such that $p \in U_p$. So, there is a basis set, B_{n_p}, such that $p \in B_{n_p} \subset U_p$. That is, to each p, we associate some natural number, n_p. For example, we can choose the smallest index n such that $p \in B_n \subset U_p$. Let $I = \{n_p : p \in M\}$. Then I is countable. For each $k \in I$, choose a set $U_k \in S$ such that $B_k \subset U_k$. This is possible by our construction. The collection $\{U_k : k \in I\}$ covers M and is countable. Thus, S has a countable subcover, implying that M is a Lindelöf space.

Since M is second countable and normal, Urysohn's metrization theorem implies that M is metrizable.

Being locally Euclidean, every point in M has a neighborhood that his homeomorphic to an open subset of $\mathbb{R}^m$. The open sets in $\mathbb{R}^m$ are locally path connected, implying that M has this same property.

Finally, it is a standard topological result that connected and locally path connected implies path connected.

QED

Next, we prove that a manifold can be covered by a countable collection of compact sets. A space with this property is called σ-compact.

Theorem 1.3 *A manifold, M, is σ-compact.*

Proof If $p \in M$, there is a basis set, B_k, such that $p \in B_k$. Applying the local compactness, there is a neighborhood, V, of p such that $\overline{V}$ is compact and $p \in V \subset \overline{V} \subset B_k$. There is a basis set, B_j, such that $p \in B_j \subset V$. It follows that $\overline{B_j} \subset \overline{V}$. Moreover, since $\overline{V}$ is compact, $\overline{B_j}$ is compact. Thus, given $p \in M$, we have found a basis element, B_j, such that $p \in \overline{B_j}$ and $\overline{B_j}$ is compact. Let B_c be the set of of the closures of all such basis elements. Then B_c is a countable compact covering of M.

QED

Thus far, we have only worked with manifolds as topological spaces with certain special properties. In order to use calculus to study geometrical properties, we need to add one more condition to our definition of a manifold. As a preliminary remark on terminology, recall that C^∞ describes a function or mapping that has continuous derivatives of all orders. We will, from this point on, use the words *differentiable* and *smooth* as synonyms for C^∞, and we will interchange them without hesitation. Varying degrees of differentiability are required throughout differential geometry. Thus, to avoid having to continually state what degree is necessary, it will be assumed that all constructions for which the description "differentiable" makes sense are, in fact, C^∞.

Definition 1.2 *A* **differentiable** *(or* **smooth**, *or C^∞)* **manifold** *is a topological manifold, M, with an atlas, denoted here by $\mathscr{A}$, that satisfies the following additional properties.*

i) For any charts (φ, U) and (ψ, V) in $\mathscr{A}$ such that $U \cap V \neq \emptyset$, the map $\varphi \circ \psi^{-1} : \psi(U \cap V) \to \varphi(U \cap V)$ is C^∞, meaning that its component functions have continuous partial derivatives of all orders. Such pairs of charts are said to be ***C^∞-compatible****.*

ii) The atlas is maximal in the sense that any chart (φ, U) on M that is C^∞ compatible with all charts in $\mathscr{A}$ is also a member of $\mathscr{A}$. (We will assume that charts whose domains do not overlap are automatically C^∞-compatible, so that checking this condition reduces to looking at charts that do overlap.)

Technically, we should refer to the pair $(M, \mathscr{A})$ as the smooth manifold, since the existence of the atlas and its properties are the essential parts of this definition. However, as is common in mathematics, when no confusion will arise, we will omit explicit mention of the atlas, $\mathscr{A}$. Smooth manifolds will be the our primary objects of study, so some explanatory remarks are in order.

Remark 1.2.1 An atlas satisfying these properties is said to be a *differentiable*, or *smooth, structure on* M, and the atlas, $\mathscr{A}$, described above is called a **maximal atlas** to distinguish it from an atlas. An atlas, as defined after definition 1.1, may contain only a few charts, perhaps finitely many, since it is only required to cover the manifold. A maximal atlas will, in general, contain many more charts, and can be considered as an "extension" of the atlas. (See below.)

Remark 1.2.2 Condition *i* in this definition has a familiar geometric interpretation. If two charts (φ, U) and (ψ, V) overlap ("overlap" meaning that the domains intersect), then the map $\varphi \circ \psi^{-1} : \psi(U \cap V) \to \varphi(U \cap V)$ is just a change of coordinates from one coordinate system to another. The requirement here is that this transition from one system to another be smooth. Think, for example, of the transition from rectangular to cylindrical or spherical coordinates in $\mathbb{R}^3$.

Remark 1.2.3 Condition *ii* is not as transparent, but it turns out to be very useful. Most importantly for our purposes, it provides us with a great deal of freedom in choosing our coordinate charts for specific purposes.

Given a topological manifold, there really are no conditions imposed on the atlas except that it does, in fact, cover the space. The atlas may, indeed, be quite sparse in the sense that there may be just enough coordinate charts to cover the space. Thus, given a particular point $p \in M$, there will, in general, be little freedom in choosing a coordinate chart at p. The maximality condition in the previous definition is much stronger, though. Given any point, p, in a smooth manifold, we will see that there are always infinitely many distinct coordinate charts at p that are C^∞-compatible with all of the charts in the maximal atlas. Some will occasionally be more useful than others, thus allowing us to choose, in a given situation, a coordinate system that suits our needs. For instance, given a point $p \in M$, we will always be able to choose, if it is helpful, a chart (φ, U) at p such that $\varphi(p) = 0 \in \mathbb{R}^m$. The drawback is that this maximal atlas will, consequently, be uncountable, and it will be impossible to write down all of the coordinate charts in the covering and check the compatibility of each pair. So, one might think that showing that a space is a smooth manifold is an impossible task. It turns out, though, that it is not necessary to check the compatibility of each pair of charts. As the following theorem will show, all we need to do is find an atlas for M (not necessarily maximal) such that all of the charts in *that particular atlas* are C^∞-compatible. Then this next result does the rest of the work for us.

Theorem 1.4 *Let* M *be a topological manifold with an atlas,* $\{(\varphi, U), (\psi, V), \dots\}$*, such that any two charts in this atlas are* C^∞*-compatible. Then there is a unique* *maximal atlas, $\mathscr{A}$, on M containing these coordinate charts, thus making $(M, \mathscr{A})$ a smooth manifold.*

Proof   We define our differentiable structure, $\mathscr{A}$, to be the collection of all charts on M that are C^∞ compatible with all charts in the given atlas. Then $\mathscr{A}$ is nonempty by hypothesis and is an atlas for M. Suppose (ρ, W) and (θ, Z) are two charts in $\mathscr{A}$ such that $W \cap Z \neq \emptyset$. If at least one of them is in the original coordinate covering, then they are C^∞-compatible by hypothesis. So, we assume that they are not. The maps $\rho \circ \theta^{-1}$ and $\theta \circ \rho^{-1}$ are well defined homeomorphisms between open sets in $\mathbb{R}^m$, so we need only show that these maps are smooth. Let $x = \rho(p)$ be an arbitrary point of $\rho(W \cap Z)$. Then there is some chart (φ, U) in the original coordinate covering such that $p \in U$. So, $N = U \cap Z \cap W$ is a neighborhood of p and $\theta(N)$ is a neighborhood of x. On $\rho(N)$, we have $\theta \circ \rho^{-1} = \theta \circ \varphi^{-1} \circ \varphi \circ \rho^{-1}$. But $\theta \circ \varphi^{-1}$ and $\varphi \circ \rho^{-1}$ are C^∞ by hypothesis. Hence, their composition is also C^∞ on $\rho(N)$. Since p was arbitrary, it follows that for every point, q, in $\rho(W \cap Z)$, there is a neighborhood of q on which $\theta \circ \rho^{-1}$ is C^∞, implying that this map is C^∞. A similar argument shows that $\rho \circ \theta^{-1}$ is C^∞ on $\psi(W \cap Z)$. Hence, condition (1) of the definition of smooth manifold is satisfied. As for property (2), it is satisfied by our definition of $\mathscr{A}$.

QED

So, to prove that a topological space is a smooth manifold, we need only find an atlas in which all of the charts are C^∞-compatible, and then apply this theorem to extend that covering to a maximal atlas, thus giving us a smooth manifold. Whenever we construct a smooth manifold, we will usually just construct the initial atlas and show that the charts are C^∞-compatible. We will typically not even mention the application of this theorem, but it will be implicit in all such examples.

Example 1.1 *Euclidean Space*

Any Euclidean space, $\mathbb{R}^m$, is a smooth manifold. The single chart $(\boldsymbol{i}, \mathbb{R}^m)$, where $\boldsymbol{i}$ is the identity map $\boldsymbol{i}\colon \mathbb{R}^m \to \mathbb{R}^m$, gives us an atlas, and this map is trivially C^∞-compatible with itself. Applying Theorem 1.4 gives us a smooth structure on $\mathbb{R}^m$. This is referred to as the canonical smooth structure on $\mathbb{R}^m$.

Example 1.2 *Finite Dimensional Vector Spaces*

If V is a vector space of dimension n, then there is an isomorphism $\alpha : V \to \mathbb{R}^n$. This isomorphism allows us to define a topology on V making it homeomorphic to $\mathbb{R}^n$, and, thus, making α a homeomorphism. The single chart (α, V) gives us an atlas just as in the previous example. Theorem 1.4 extends this to a smooth structure on V.

Example 1.3 *Graph of a smooth function $f : U \to \mathbb{R}$, where U is open in $\mathbb{R}^2$*

In multivariable calculus, the graph of a smooth function $f : U \to \mathbb{R}$, viewed as a subset of $\mathbb{R}^3$, is known to induce a surface. As manifolds are meant to be generalizations of surfaces, it stands to reason that surfaces should certainly be manifolds. Suppose U is an open set in $\mathbb{R}^2$, and $f : U \to \mathbb{R}$ is a smooth function. Let $M = \{(x, y, f(x, y)) \in \mathbb{R}^3 \ : \ (x, y) \in U\}$. We will show that M is a manifold. First, since M is a subspace of a Euclidean space, M is Hausdorff and second countable. Define a map $\varphi : M \to U$ by $\varphi(x, y, f(x, y)) = (x, y)$. This map is clearly surjective, and it is injective, since $\varphi(x_1, y_1, f(x_1, y_1)) = \varphi(x_2, y_2, f(x_2, y_2)) \Rightarrow x_1 = x_2$ and $y_1 = y_2$. This implies that $f(x_1, y_1) = f(x_2, y_2)$. Moreover, φ is smooth as its component functions are smooth. The inverse mapping, $\varphi^{-1} : U \to M$, is given by $\varphi^{-1}(x, y) = (x, y, f(x, y))$. Since f is smooth by hypothesis, so is φ^{-1}. Thus, the single chart (φ, M) covers M and is C^∞-compatible with itself. Hence, it can be extended to a smooth structure on M.

Example 1.4 *S^1: the 1-dimensional sphere, or circle*

Let $S^1 = \{(x, y) \in \mathbb{R}^2 \ : \ x^2 + y^2 = 1\}$. This is, of course, the unit circle in $\mathbb{R}^2$. It is a subspace of $\mathbb{R}^2$, so it is Hausdorff and second countable. It is a curve, or a "1-dimensional surface," but it cannot be represented as the graph of a single function. Define the following mappings.

1. Let $U_1 = \{(x, y) \in S^1 \ : \ y > 0\}$. Define $\varphi_1 : U_1 \to (-1, 1)$ by $\varphi_1(x, y) = x$.
2. Let $U_2 = \{(x, y) \in S^1 \ : \ y < 0\}$. Define $\varphi_2 : U_2 \to (-1, 1)$ by $\varphi_2(x, y) = x$.
3. Let $U_3 = \{(x, y) \in S^1 \ : \ x > 0\}$. Define $\varphi_3 : U_3 \to (-1, 1)$ by $\varphi_3(x, y) = y$.
4. Let $U_4 = \{(x, y) \in S^1 \ : \ x < 0\}$. Define $\varphi_4 : U_4 \to (-1, 1)$ by $\varphi_4(x, y) = y$.

Noting that U_1 can be represented as the set of all points in S^1 of the form $(x, \sqrt{1-x^2})$ with $x \in (-1,1)$, we see that ${\varphi_1}^{-1}(x) = (x, \sqrt{1-x^2})$. The inverses of the other maps are defined similarly. We are simply projecting hemispheres of the circle onto an open interval. Given their respective domains, the maps φ_i, $1 \leq i \leq 4$ are easily seen to be homeomorphisms. The charts (φ_i, U_i) are also all C^∞-compatible. To see this, consider the case $\varphi_1 \circ {\varphi_3}^{-1} : \varphi_3(U_1 \cap U_3) \to \varphi_1(U_1 \cap U_3)$. $U_1 \cap U_3$ is the set of all points, (x, y), on the circle such that $x > 0$ and $y > 0$. We can represent such a point as $(x, \sqrt{1-x^2})$ or $(\sqrt{1-y^2}, y)$. So, if y is a point in $\varphi_3(U_1 \cap U_3)$, then $0 < y < 1$, and ${\varphi_3}^{-1}(y) = (\sqrt{1-y^2}, y)$. Thus, $\varphi_1 \circ {\varphi_3}^{-1}(y) = \varphi_1(\sqrt{1-y^2}, y) = \sqrt{1-y^2}$. Since $0 < y < 1$, this map is C^∞. The other possible compositions are computed in a similar way, and are also C^∞. So, S^1, with the atlas induced by the charts (φ_i, U_i), is a smooth manifold.

In a similar way, we can show that the set $S^n = \{(x_1, ..., x_{n+1}) \in \mathbb{R}^{n+1} \ : \ {x_1}^2 + \cdots + {x_{n+1}}^2 = 1\}$, called the n-dimensional sphere, is a smooth manifold. As in the 1-dimensional case, we simply divide S^n into $2(n+1)$ hemispheres (two for each direction or dimension in $\mathbb{R}^{n+1}$) and project to the appropriate coordinates. This will be referred to as the *standard coordinate atlas on* S^n. However, it is possible to define, on any sphere, a smooth coordinate covering consisting of just two coordinate charts. This is called the *stereographic projection*, and it has certain advantages over the standard coordinate covering that will become apparent later on.

Example 1.5 *S^n in stereographic coordinates*

Let $\mathbf{n}$ and $\mathbf{s}$ denote the points $(0, \ldots, 0, 1)$ and $(0, \ldots, 0, -1)$, respectively, on $S^n \subset \mathbb{R}^{n+1}$. These are the north and south poles of S^n. Given any point $p \in S^n - \{\mathbf{n}\}$, the line passing through p and $\mathbf{n}$ intersects the hyperplane $x_{n+1} = 0$, which is homeomorphic to $\mathbb{R}^n$, in exactly one point. We define a coordinate chart by sending p to that point. Define $\varphi : S^n - \{\mathbf{n}\} \to \mathbb{R}^n$ by

$$\varphi(x_1, \ldots, x_n, x_{n+1}) = \frac{1}{1 - x_{n+1}}(x_1, \ldots, x_n).$$

The inverse map is given by

$$\varphi^{-1}(y_1, \ldots, y_n) = \left(\frac{2y_1}{1 + \|y\|^2}, \ldots, \frac{2y_n}{1 + \|y\|^2}, \frac{-1 + \|y\|^2}{1 + \|y\|^2} \right).$$

Similarly, we define $\psi : S^n - \{\mathbf{s}\} \to \mathbb{R}^n$ by

$$\psi(x_1, \ldots, x_n, x_{n+1}) = \frac{1}{1 + x_{n+1}}(x_1, \ldots, x_n).$$

It is easy to check that φ and ψ are both homeomorphisms onto $\mathbb{R}^n$, thus making $\{(\varphi, S^n - \{\mathbf{n}\}),\ (\psi, S^n - \{\mathbf{s}\})\}$ an atlas for S^n. Another computation shows that the compositions $\varphi \circ \psi^{-1}$ and $\psi \circ \varphi^{-1}$ are smooth, thus inducing a smooth structure on S^n. Note that each chart covers all of S^n except for a single point.

Example 1.6 *Open submanifolds of a manifold, M*

Let M be a smooth manifold, and let O be an open subspace of M. Then O is both Hausdorff and second countable. Suppose the maximal atlas of M is given by $\{(\varphi, U), (\psi, V), \ldots\}$. Define an atlas on O by $\{(\varphi|_{U\cap O}, U \cap O), (\psi|_{V\cap O}, V \cap O), \ldots\}$. The restriction of a homeomorphism to an open subset of its domain is still a homeomorphism, so this covering clearly makes O a topological manifold. Moreover, if (φ, U) and (ψ, V) are two charts such that $(U \cap O) \cap (V \cap O) = U \cap V \cap O \neq \emptyset$, then $\varphi \circ \psi^{-1} : \varphi(U \cap V \cap O) \to \psi(U \cap V \cap O)$ is C^∞, as it is just the restriction of the C^∞ map $\varphi \circ \psi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$ to an open subset of $\varphi(U \cap V)$. Thus, any two charts in this atlas on O are C^∞-compatible, and we can conclude that O, with the induced maximal atlas, is a smooth manifold. We call it a ***smooth open submanifold***, or just an ***open submanifold***, of M. It follows that every coordinate neighborhood, U, is an open submanifold of M. Note that open submanifolds necessarily have the same dimension as the ambient, or "parent," manifold. This will not be true for more general submanifolds that will be discussed later.

Example 1.7 *Product Manifolds*

Suppose M and N are smooth manifolds. Then the products $A \times B$, where A and B are open in M and N, respectively, form a basis for the topology on $M \times N$. This makes $M \times N$ Hausdorff and second countable. Given any point $(p, q) \in M \times N$, there are charts (φ, U) on M and (ψ, V) on N at p and q, respectively. We form a chart $(\theta, U \times V)$ on $M \times N$ by letting $\theta(r, s) = (\varphi(r), \psi(s))$ for $(r, s) \in U \times V$. Doing this for each point in $M \times N$ gives us an atlas, making $M \times N$ a topological manifold. It is easy to check that any two of these charts which overlap are C^∞-compatible. (It just reduces to compatibility of charts on M and N, individually.) Hence, this atlas extends to a smooth structure, making $M \times N$ a smooth manifold. This is referred to as the ***standard product manifold structure***. Whenever we form product manifolds, it will be assumed, unless specifically stated otherwise, that this is the smooth structure we place on them.

The previous examples have all been either standard or trivial in their construction. The next example will show how manifolds can arise from more complicated and abstract structures. This example, itself, however, is considered standard. Manifolds may have a much more complex structure.

Example 1.8 *Real Projective n-space,* $\mathbb{R}P^n$

This example follows a construction given in Boothby. Let $X = \mathbb{R}^{n+1} - \{0\}$. Define an equivalence relation, $\sim$, on X as follows. For $x, y \in X$, $x \sim y$ if and only if there exists a real number, $\lambda \neq 0$, such that $x = \lambda y$. Denote the equivalence class containing x by $[x]$. Let $\mathbb{P}^n = \{[x] : x \in X\}$. Then $\mathbb{P}^n$ is the quotient space $\mathbb{X}/\sim$, and we assume it is endowed with the usual quotient space topology. This means that $O \subset \mathbb{P}^n$ is open if and only if $\pi^{-1}(O)$ is open in X, where $\pi : X \to \mathbb{P}^n$ is the usual quotient map, taking $x \in X$ to the equivalence class $[x]$. This automatically makes π continuous. Note that, geometrically speaking, we can identify $\mathbb{P}^n$ with the set of lines in R^{n+1} passing through the origin. We will show that $\mathbb{P}^n$ is a smooth n-dimensional manifold. The proof is direct, following our definition of a smooth manifold.

Note that, for $[x] = [(x^1, x^2, ..., x^{n+1})] \in \mathbb{P}^n$, if $x^i \neq 0$ for some $i = 1, 2, ..., n+1$, then $[(x^1, ..., x^i, ..., x^{n+1})] = [(x^1/x^i, ..., x^{i-1}/x^i, 1, x^{i+1}/x^i, ..., x^{n+1}/x^i)]$. Also, for $[x] \in \mathbb{P}^n$, if any element $(x^1, ..., x^{n+1}) \in [x]$ satisfies $x^i = 0$, then every element of $[x]$ has 0 i^{th} component. So, we can speak of an equivalence class $[x] \in \mathbb{P}^n$ as having a zero or nonzero i^{th} component, $[x]_i$.

Let $U \subset \mathbb{P}^n$ be given. Since $\pi^{-1}(U) = \bigcup_{[x]\in U}[x]$, it follows that a set $U \subset \mathbb{P}^n$ is open if and only if the union of all the equivalence classes in U is open in X. Define sets U_i, $i = 1, ..., n+1$, by $U_i = \{[x] \in \mathbb{P}^n : [x]_i \neq 0\}$. If $[x] \in \mathbb{P}^n$, then at least one of the components of $[x]$ must be nonzero. So, the sets U_i cover $\mathbb{P}^n$. They are also open. Consider $\pi^{-1}(U_i) = \bigcup_{[x]\in U_i}[x]$. This is just the union of all the equivalence classes that have a nonzero i^{th} component. Let $x = (x^1, ..., x^i, ..., x^{n+1})$ be in $\pi^{-1}(U_i)$. Then $x^i \neq 0$. Choose $\epsilon > 0$ such that $0 < \epsilon < \frac{|x^i|}{2}$. The set $\{y \in \mathbb{R}^{n+1} : \|x - y\| < \epsilon\} = B_\epsilon(x)$ is an open ball centered at x. If $y \in B_\epsilon(x)$, then $|x_i - y_i| \leq \|x - y\| \Rightarrow |x_i - y_i| < \epsilon \Rightarrow -\frac{|x_i|}{2} < y_i - x_i < \frac{|x_i|}{2}$. It follows that we must have $y_i \neq 0$. Thus $B_\epsilon(x) \subset \pi^{-1}(U_i)$, implying that $\pi^{-1}(U_i)$ is open in X. Hence, each set U_i is open. We will use this collection of sets to form the initial atlas.

For each $i = 1, ..., n+1$, define the map $\varphi_i : U_i \to \mathbb{R}^n$ by

$$\varphi_i([x]) = \varphi_i\Big(\big[(x^1, ..., x^i, ..., x^{n+1})\big]\Big) = \left(\frac{x^1}{x^i}, \frac{x^2}{x^i}, ..., \frac{x^{i-1}}{x^i}, \frac{x^{i+1}}{x^i}, ..., \frac{x^{n+1}}{x^i}\right)$$

Note that φ_i is well-defined, for if $x = (x^1, ..., x^{n+1})$, and if $y = (y^1, ..., y^{n+1})$ is in $[x]$, then $y = \lambda x$ for some $\lambda \neq 0$. Thus,

$$\begin{aligned} \Big(\frac{y^1}{y^i}, ..., \frac{y^{i-1}}{y^i}, \frac{y^{i+1}}{y^i}, ..., \frac{y^{n+1}}{y^i}\Big) &= \Big(\frac{\lambda x^1}{\lambda x^i}, ..., \frac{\lambda x^{i-1}}{\lambda x^i}, \frac{\lambda x^{i+1}}{\lambda x^i}, ..., \frac{\lambda x^{n+1}}{\lambda x^i}\Big) \\ &= \Big(\frac{x^1}{x^i}, ..., \frac{x^{i-1}}{x^i}, \frac{x^{i+1}}{x^i}, ..., \frac{x^{n+1}}{x^i}\Big). \end{aligned}$$

Now, if $\varphi_i([x]) = \varphi_i([y])$ for $[x], [y] \in U_i$, then for any elements $x \in [x]$ and $y \in [y]$, we have

$$\begin{aligned} &\left(\frac{x^1}{x^i}, ..., \frac{x^{i-1}}{x^i}, \frac{x^{i+1}}{x^i}, ..., \frac{x^{n+1}}{x^i}\right) = \left(\frac{y^1}{y^i}, ..., \frac{y^{i-1}}{y^i}, \frac{y^{i+1}}{y^i}, ..., \frac{y^{n+1}}{y^i}\right) \\ \Rightarrow \quad &\big(x^1, ..., x^{i-1}, x^{i+1}, ..., x^{n+1}\big) = \frac{x^i}{y^i}\big(y^1, ..., y^{i-1}, y^{i+1}, ..., y^{n+1}\big) \\ \Rightarrow \quad &\big(x^1, ..., x^{i-1}, x^i, x^{i+1}, ..., x^{n+1}\big) = \frac{x^i}{y^i}\big(y^1, ..., y^{i-1}, y^i, y^{i+1}, ..., y^{n+1}\big) \end{aligned}$$

from which it follows that $x = \frac{x^i}{y^i}y$. So $[x] = [y]$, and each φ_i is injective. If $y \in \mathbb{R}^n$, consider the element in X given by $\tilde{y} = (y^1, ..., y^{i-1}, 1, y^i, ..., y^n)$. That is, we put a 1 in the i^{th} position and shift the remaining components up one index value. If $i = n+1$, we simply attach a 1 on the end of y. Then $[\tilde{y}] \in U_i$ and $\varphi_i([\tilde{y}]) = (y^1, ..., y^{i-1}, y^i, ..., y^n) = y$. So, each φ_i is also surjective.

To show that each φ_i is continuous, it suffices to show that ${\varphi_i}^{-1}(B_\epsilon(y))$ is open in $\mathbb{P}^n$ for arbitrary ϵ and $y \in \mathbb{R}^n$. Hence, we want to show that $\pi^{-1}({\varphi_i}^{-1}(B_\epsilon(y)))$ is open in X. But $\pi^{-1}({\varphi_i}^{-1}(B_\epsilon(y))) = (\varphi_i \circ \pi)^{-1}(B_\epsilon(y))$, so, if the map $\varphi_i \circ \pi : \pi^{-1}(U_i) \to \mathbb{R}^n$ is continuous, this will prove the result. Now, $\pi^{-1}(U_i)$ consists of all those points in X such that $x^i \neq 0$. This set is open in X. If x is such a point, then $\pi(x) = [x]$ and

$$\varphi_i(\pi(x)) = \left(\frac{x^1}{x^i}, ..., \frac{x^{i-1}}{x^i}, \frac{x^{i+1}}{x^i}, ..., \frac{x^{n+1}}{x^i}\right).$$

Hence, since x^i never vanishes for any point $x \in \pi^{-1}(U_i)$, we see that each of the component functions of this map is continuous. It follows that $\varphi_i \circ \pi$ is continuous on $\pi^{-1}(U_i)$, implying that φ_i is continuous on U_i.

Since φ_i is a bijection, it has a well-defined inverse. Explicitly, the inverse mapping ${\varphi_i}^{-1} : \mathbb{R}^n \to U_i$ satisfies

$$
{\varphi_i}^{-1}(y_1, ..., y_n) = [y^1, ..., y^{i-1}, 1, y^i, ..., y^n].
$$

To see that ${\varphi_i}^{-1}$ is continuous, define a map $f : \mathbb{R}^n \to \mathbb{R}^{n+1}$ by $f(y^1, ..., y^n) = (y^1, ..., y^{i-1}, 1, y^i, ..., y^n)$. This map is continuous and its range lies in $\pi^{-1}(U_i)$. Moreover, ${\varphi_i}^{-1} = \pi \circ f$. Thus, ${\varphi_i}^{-1}$ is continuous, and each φ_i is a homeomorphism.

To show that $\mathbb{P}^n$ is a manifold, we still must show that $\mathbb{P}^n$ is Hausdorff and second countable. These results are not trivial, as topological properties are not generally well-preserved under the formation of quotient spaces.

First, we claim that π is an open map. For $t \in \mathbb{R}$, $t \neq 0$, consider $\alpha_t : X \to X$ defined by $\alpha_t(x) = tx$. This is just a scaling map, so it is a homeomorphism with ${\alpha_t}^{-1} = \alpha_{\frac{1}{t}}$. If $U \subset X$ is open, then $\alpha_t(U)$ is open in X also. Moreover, if $y \in [U] := \bigcup_{x \in U}[x]$, then $y \in [x]$ for some $x \in U$, implying that $y = tx$ for some $t \neq 0$. So, $y \in \alpha_t(U)$, and we have

$$
[U] \subset \bigcup_{\substack{t \in \mathbb{R} \\ t \neq 0}} \alpha_t(U).
$$

Conversely, if $y \in \bigcup_{t \in \mathbb{R}, t \neq 0} \alpha_t(U)$, then $y \in \alpha_t(U)$ for some t. So, $y = tx$ for some $x \in U$, implying that $y \sim x \Rightarrow y \in [U]$. Thus, we have

$$
[U] = \bigcup_{\substack{t \in \mathbb{R} \\ t \neq 0}} \alpha_t(U)
$$

Since each $\alpha_t(U)$ is open, $[U]$ is open in X. But it is easily seen that $\pi^{-1}(\pi(U)) = [U]$. Thus, by the definition of the quotient topology, $\pi(U)$ is open in $\mathbb{P}^n$.

This shows that $\mathbb{P}^n$ is second countable. If $\{B_n\}_{n \geq 1}$ is a countable basis for X, then $\{\pi(B_n)\}_{n \geq 1}$ is a countable basis for $\mathbb{P}^n$. Given any open set $O \subset \mathbb{P}^n$ and any $[x] \in O$, $\pi^{-1}(O)$ is open in X and contains x. So, there is some B_n such that $x \in B_n \subset \pi^{-1}(O)$, implying that $[x] = \pi(x) \in \pi(B_n) \subset \pi(\pi^{-1}(O)) \subset O$. Therefore, $\{\pi(B_n)\}$ is, indeed, a basis.

To prove the Hausdorff property, we will digress momentarily to prove the following lemma.

Lemma 1.5 *Let $\sim$ be an equivalence relation on a topological space X, and assume that the quotient map $\pi : X \to X/\sim$ is an open map. Then $R = \{(x, y) \in X \times X : x \sim y\}$ is a closed subset of $X \times X$ (with the standard product topology) if and only if $X/\sim$ is Haudsdorff.*

Proof First, suppose $X/\sim$ is Hausdorff. Suppose $(x, y) \notin R$. Then $x \nsim y$, which implies that $\pi(x) \neq \pi(y)$. Hence, there are disjoint open subsets, U and V, of $X/\sim$ such that $\pi(x) \in U$ and $\pi(y) \in V$. Then $\pi^{-1}(U)$ and $\pi^{-1}(V)$ are open sets in X containing x and y, respectively. So, the set $\pi^{-1}(U) \times \pi^{-1}(V)$ is an open set in $X \times X$ that contains (x, y). Suppose this set intersects R. Then it must contain some element (x', y') such that $x' \sim y'$. This implies that $\pi(x') = \pi(y')$. But (x', y') is also in $\pi^{-1}(U) \times \pi^{-1}(V)$, implying that $\pi(x') \in U$ and $\pi(y') \in V$. It follows that U and V have a nonempty intersection, contradicting the fact that they are disjoint. Thus, the set $\pi^{-1}(U) \times \pi^{-1}(V)$ does not intersect R. Therefore, for any $(x, y) \in R^c$, there is an open set containing (x, y) that does not intersect R. So, R^c is open, implying that R is closed.

Conversely, suppose R is closed. Since any point in $X/\sim$ is the image of some element under the map π, we can consider two elements of $X/\sim$ as $\pi(x)$ and $\pi(y)$. Suppose $\pi(x)$ and $\pi(y)$ are two distinct points of $X/\sim$. Then we cannot have $x \sim y$. Thus, the element (x, y) must be in R^c, which is open. Hence, there are open sets U and V in X such that $(x, y) \in U \times V$ and the set $U \times V$ does not intersect R. Consider the sets $\pi(U)$ and $\pi(V)$ in $X/\sim$. If there was an element, say $[z]$, in both of these sets, then we would have $x \sim z$ and $y \sim z$, implying that $x \sim y$. Hence, this contradiction shows that $\pi(U)$ and $\pi(V)$ are disjoint. The previous lemma shows that they are open, and we have $\pi(x) \in \pi(U)$ and $\pi(y) \in \pi(V)$. Since these two elements were arbitrary, this shows that $X/\sim$ is Hausdorff.

QED

We now apply this result to prove the Haudsorff condition. Define a function $f : X \times X \to \mathbb{R}$ by

$$f(x, y) = f(x^1, ..., x^{n+1}, y^1, ..., y^{n+1}) = \sum_{i,j=1}^{n} (x^i y^j - x^j y^i)^2.$$

Then f is continuous on $X \times X$. We can also show that $f(x, y) = 0 \Leftrightarrow y = tx$ for some $t \neq 0$. If $y = tx$, then it is clear that f vanishes. Conversely, suppose $f(x, y) = 0$. Then, viewing x and y as vectors in $\mathbb{R}^{n+1}$, this implies

$$
\begin{aligned}
& \sum_{i=1}^{n}\sum_{j=1}^{n}(x^i)^2(y^j)^2 - 2\sum_{i=1}^{n}\sum_{j=1}^{n}x^iy^ix^jy^j + \sum_{i=1}^{n}\sum_{j=1}^{n}(x^j)^2(y^j)^2 = 0 \\
\Rightarrow\ & \Big(\sum_{i=1}^{n}(x^i)^2\Big)\Big(\sum_{j=1}^{n}(y^j)^2\Big) - 2\Big(\sum_{i=1}^{n}x^iy^i\Big)\Big(\sum_{j=1}^{n}x^jy^j\Big) + \Big(\sum_{i=1}^{n}(x^i)^2\Big)\Big(\sum_{j=1}^{n}(y^j)^2\Big) = 0 \\
\Rightarrow\ & \|x\|^2\|y\|^2 = \langle x, y\rangle^2 \\
\Rightarrow\ & \|x\|\|y\| = \langle x, y\rangle,
\end{aligned}
$$

from which it follows that x and y must be linearly dependent. That is, $y = tx$ for some $t \neq 0$. Thus, $f^{-1}(0)$ consists of all those pairs, (x, y), in $X \times X$ such that $y = tx$ for some $t \neq 0$. Hence, $f^{-1}(0) = R = \{(x, y) \ : \ x \sim y\}$ is closed, and Lemma 1.6 implies that $\mathbb{R}^n$ is Hausdorff. So, $\mathbb{P}^n$ is a manifold. All that remains is to show that the maps φ_i are C^∞-compatible. This is nothing more than an elaborate computation.

Suppose $U_i \cap U_j \neq \emptyset$ and $x \in \varphi_i(U_i \cap U_j) \subset \mathbb{R}^n$. Then $x = \varphi_i([y])$ for some $[y] \in U_i \cap U_j$. This $[y]$ must, then, have nonzero i and j components. We can assume without loss of generality that $i < j$. The map φ_i will remove the i^{th} component of $[y]$ and shift the j^{th} component to the $(j-1)^{st}$ position. Hence, x has a nonzero $(j-1)^{st}$ component, and it follows that

$$\varphi_i{}^{-1}(x) = [x^1, ..., x^{i-1}, 1, x^i, ..., x^{j-1}, ..., x^n]$$

where the component x^{j-1} is nonzero and in the j^{th} position of the equivalence class. We then have

$$
\begin{aligned}
\varphi_j(\varphi_i{}^{-1}(x)) &= \varphi_j\big([x^1, ..., x^{i-1}, 1, x^i, ..., x^{j-1}, ..., x^n]\big) \\
&= \left(\frac{x^1}{x^{j-1}}, ..., \frac{x^{i-1}}{x^{j-1}}, \frac{1}{x^{j-1}}, \frac{x^i}{x^{j-1}}, ..., \frac{x^{j-2}}{x^{j-1}}, \frac{x^j}{x^{j-1}}, ..., \frac{x^n}{x^{j-1}}\right).
\end{aligned}
$$

This map is defined on $\varphi(U_i \cap U_j)$, which consists only of points whose $(j-1)^{st}$ component is nonzero. It is, therefore, C^∞. The indices i and j were arbitrary, so this shows that the coordinate charts (φ_i, U_i) are all C^∞-compatible. So, $\mathbb{P}^n$ is a smooth n-dimensional manifold. In the cases $n = 1$ and $n = 2$, we call $\mathbb{P}^n$ the *Projective Line* and the *Projective Plane*, respectively.

We should remark that the extensive effort that went in to showing the projective space was, in fact, a smooth manifold, is not always required. Indeed, in many cases, there are quicker and more indirect ways of showing that a space is a manifold. As we progress and further develop our theory, these methods will arise naturally. However, there are occasions where the direct method is the only viable choice, so the reader should have some experience with it.

Smooth Functions and Mappings

We will now introduce the concepts of smooth, or differentiable, functions on a smooth manifold, M. As M is a topological space, we have a well-defined notion of what it means for a function to be continuous. However, the methods of the differential calculus are the backbone of differential geometry. So, we need some means of differentiating functions on M. It is the local Euclidean structure of a manifold that makes this possible.

Definition 1.3 *Let f be a real-valued function defined on an open subset, W, of a smooth manifold, M. We say that f is C^∞ on W if for every $p \in W$ there is a coordinate chart (φ, U) at p such that the function $f \circ \varphi^{-1} : \varphi(U \cap W) \to \mathbb{R}$ is C^∞ on $\varphi(U \cap W)$. We call the function $f \circ \varphi^{-1} : \varphi(U \cap W) \to \mathbb{R}$ the* ***coordinate representation*** *of f on $U \cap W$. We let $\mathscr{F}(M)$ denote the set of smooth functions on M.*

Remark In this definition, it is important to remember that $\varphi(U \cap W)$ is an open subset of $\mathbb{R}^m$, so we know from basic Euclidean calculus what it means for $f \circ \varphi^{-1} : \varphi(U \cap W) \to \mathbb{R}$ to be differentiable.

Even though this definition uses local coordinate systems and requires the existence of *only one* such chart at the point p, (when, as we know, there will be many charts at any given point) the smoothness of a function is, in fact, a well-defined coordinate-independent property. Given $p \in W$, if there is one chart (φ, U) that satisfies the above condition, then any chart at p will. For if (ψ, V) is another chart at p, then $f \circ \varphi^{-1} : \varphi(U \cap V \cap W) \to \mathbb{R}$ is C^∞, and the map $\varphi \circ \psi^{-1} : \psi(U \cap V \cap W) \to$

$\varphi(U \cap V \cap W)$ is C^∞. Thus, by the chain rule in Euclidean space (note that we have not yet developed a chain rule on manifolds), the function $f \circ \psi^{-1} = (f \circ \varphi^{-1}) \circ (\varphi \circ \psi^{-1})$ is C^∞ on $\psi(U \cap V \cap W)$, a neighborhood of $\psi(p)$. Thus, the definition does not depend on the particular coordinate representation of f.

Given this coordinate independence, a more compact way of saying all of this would be the following. ***A function*** $f : W \to \mathbb{R}$***, where*** $W \subset M$ ***is open, is smooth on*** W ***if and only if for any chart*** (φ, U) ***such that*** $U \cap W \neq \emptyset$***, the function*** $f \circ \varphi^{-1} : \varphi(U \cap W) \to \mathbb{R}$ ***is smooth.*** Moreover, as we should expect, differentiability of f at p implies continuity of f at p. We have $f = (f \circ \varphi^{-1}) \circ \varphi$ on $\varphi(U \cap W)$. Being a homeomorphism, φ is continuous at p, and $f \circ \varphi^{-1}$ is continuous at $\varphi(p)$, being differentiable there. Thus, smoothness implies continuity.

As a note on the algebraic structure of $\mathscr{F}(M)$, we should point out that it follows almost directly from the definition that $\mathscr{F}(M)$ is a commutative algebra over $\mathbb{R}$ with a multiplicative identity.

This definition of smooth real-valued functions leads us to the definition of a smooth mapping between manifolds. As a point of terminology, we will reserve the term *function* for real-valued functions on manifolds, while we will use the term *map* to refer to mappings between general manifolds.

Definition 1.4 *Let* M *and* N *be smooth manifolds, and let* $F : M \to N$ *be a mapping of* M *into* N*. We say that* F *is a smooth mapping if for every* $p \in M$ *there exist coordinate charts* (φ, U) *and* (ψ, V) *such that* $p \in U$, $F(p) \in V$, $F(U) \subset V$, *and the mapping* $\psi \circ F \circ \varphi^{-1} : \varphi(U) \to \psi(V)$ *is* C^∞*. We call* $\psi \circ F \circ \varphi^{-1}$ *the* ***coordinate representation*** *of* F *with respect to the charts* (φ, U) *and* (ψ, V)*.*

As with the previous definition, this definition of smoothness is coordinate independent and does not depend on the specific charts we choose. If (β, U') and (λ, V') are two other coordinate systems around p and $F(p)$, respectively, then $\lambda \circ F \circ \beta^{-1} = \lambda \circ \psi^{-1} \circ \psi \circ F \circ \varphi^{-1} \circ \varphi \circ \beta^{-1}$ on $\beta(U \cap U')$. Since the maps $\lambda \circ \psi^{-1}$, $\psi \circ F \circ \varphi^{-1}$, and $\varphi \circ \beta^{-1}$ are C^∞, so is $\lambda \circ F \circ \beta^{-1}$, by applying the Euclidean chain rule. Thus, $F : M \to N$ ***is*** C^∞ ***if and only if every possible coordinate representation of*** F ***is*** C^∞***.*** Moreover, it follows from this definition that smooth maps are continuous.

Note, also, that we have not bothered to restrict F to an open subset of M in this definition, as we did in the definition of a smooth function. Recalling Example 1.5, an open subset of M is an open submanifold of M, and so is a manifold, itself. So, if we have a mapping $F : W \to N$, where W is an open subset of M, we can still apply this definition by simply replacing M with W. Definition 1.4, therefore, includes all of the cases in which F may be defined on only some open subset of M.

There is a consequence of these definitions that we will use frequently and usually without comment. Each coordinate mapping $\varphi : U \to \varphi(U)$ is a smooth function from the coordinate neighborhood (and manifold), U, to the open set $\varphi(U) \subset \mathbb{R}^m$. This just follows from the C^∞ compatibility of the charts in the smooth structure on M. If p is any point in U and (ψ, V) is any coordinate chart around p (possibly (φ, U) itself), then $\varphi \circ \psi^{-1} : \psi(U \cap V) \to \varphi(U \cap V)$ is C^∞ because the charts are C^∞-compatible. Thus, φ is smooth. A similar argument shows that φ^{-1} is smooth as a map from the open set $\varphi(U)$ onto U. Hence, the coordinate maps on a smooth manifold are ***diffeomorphisms***.

Definition 1.5 *Let M and N be smooth manifolds. A map $F : M \to N$ is a* ***diffeomorphism*** *if it has an inverse $F^{-1} : N \to M$ and if both F and F^{-1} are smooth. If such a map exists, we say that M and N are* ***diffeomorphic***

Just as homeomorphisms preserve all topological properties between two topological spaces, diffeomorphisms preserve all properties that can be expressed in terms of the topological *and* smooth structures of manifolds. This concept is the foundation of the subject of *Differential Topology.* Clearly every diffeomorphism is a homeomorphism. The converse is not true, as can be seen by the mapping $t \mapsto t^{1/3}$ from $\mathbb{R}$ onto $\mathbb{R}$.

We now prove a series of useful lemmas concerning smooth maps on manifolds. The first is a version of the chain rule. We will strengthen this result in the next section.

Lemma 1.6 *Let $F : M \to N$ and $G : N \to L$ be smooth maps between smooth manifolds. Then $G \circ F$ is smooth.*

Proof Let $p \in M$ be given. Let $q = F(p) \in N$, and let $r = G(q) \in L$. By hypothesis, there are charts (φ, U) and (ψ, V) at q and r, respectively, such that $G(U) \subset V$ and $\psi \circ G \circ \varphi^{-1} : \varphi(U) \to \psi(V)$ is C^∞. Choose a chart (ξ, W) at p, and, using the continuity of F, shrink W so that $F(W) \subset U$. Since F is smooth, $\varphi \circ F \circ \xi^{-1} : \xi(W) \to \varphi(U)$ is smooth. Thus, by the Euclidean chain rule, the map $\psi \circ G \circ F \circ \xi^{-1} = \psi \circ G \circ \varphi^{-1} \circ \varphi \circ F \circ \xi^{-1}$ is C^∞ on $\xi(W)$. Since we also have $(G \circ F)(W) \subset V$, this shows that $G \circ F$ is smooth.

QED

Note that this version of the chain rule did not use the notion of the differential of a map, which is the manner in which the Euclidean chain rule is usually stated. We are assuming the Euclidean chain rule in this case. Later on, we will define a corresponding notion of the differential of a map and prove a chain rule in that context as well.

Lemma 1.7 *Let M be a smooth manifold, and let $W \subset M$ be open. Suppose $F : W \to \mathbb{R}^m$ is a diffeomorphism from W onto some open subset of $\mathbb{R}^m$. Then (F, W) is a coordinate chart on M.*

<u>Proof</u> Being a diffeomorphism, F is necessarily a homeomorphism. So (F, W) is a coordinate chart in the topological sense. Let (φ, U) be any chart in the maximal atlas on M such that $U \cap W \neq \emptyset$. Since F is smooth, $F \circ \varphi^{-1} : \varphi(U \cap W) \to F(U \cap W)$ is also smooth by the previous lemma (or simply by the definition of a smooth map on the open submanifold W). Likewise, the differentiability of F^{-1} implies that $\varphi \circ F^{-1} : F(U \cap W) \to \varphi(U \cap W)$ is C^∞. Since (φ, U) was arbitrary, it follows that (F, W) is C^∞-compatible with every overlapping chart on M. Hence, (F, W) is contained in the maximal smooth structure on M.

QED

As an example of the utility of this Lemma, consider any chart (φ, U) on a smooth manifold M, and let p be a fixed point in U. We know that φ is a diffeomorphism of U onto the open set $\varphi(U) \subset \mathbb{R}^m$, which contains $\varphi(p)$. Let $T : \mathbb{R}^m \to \mathbb{R}^m$ be the translation map, $T(x) = x - \varphi(p)$. It is easy to verify that translations are diffeomorphisms of Euclidean spaces. Since the restriction of a diffeomorphism to an open subset of its domain is, again, a diffeomorphism, it follows that the map $T \circ \varphi : U \to T(\varphi(U))$ is a diffeomorphism. By the preceding lemma, we can conclude that $(T \circ \varphi, U)$ is a coordinate chart on M. Moreover, we see that $(T \circ \varphi)(p) = 0$. Thus, we may choose, whenever convenient, a chart that maps any given point of M to $0 \in \mathbb{R}^m$.

In this same way, and utilizing the same lemma, we can always find charts with very desirable properties. We can assume, for instance, that a given chart maps its corresponding coordinate neighborhood to an open ball centered at any point we choose. We can also permute coordinates if we need to, since permutation of coordinates in $\mathbb{R}^m$ is a diffeomorphism.

Next, we prove two very important lemmas regarding the existence of certain kinds of smooth functions on manifolds. These are smooth analogues of well-known results in general topology concerning the existence of continuous functions on certain topological spaces. (Compare, for example, to Urysohn's Lemma.) Proving the theorems for manifolds follows easily once one has established them on ordinary Euclidean space, so we will prove those results first.

Lemma *Suppose $A, K \subset \mathbb{R}^m$, with A closed, K compact, and $A \cap K = \emptyset$. Then there is a smooth function $f : \mathbb{R}^m \to [0, 1]$ such that $f = 0$ on A and $f = 1$ on K.*

Proof The function $h(t) = e^{-1/t^2}$, extended to $(-\infty, 0]$ by setting $h(t) = 0$ for $t \leq 0$, is C^∞. Define $g : \mathbb{R}^m \to \mathbb{R}$ by

$$g(x) = \frac{h(\epsilon^2 - \|x\|^2)}{h(\epsilon^2 - \|x\|^2) + h(\|x\|^2 - \epsilon^2/4)}.$$

It is straightforward to check that g is C^∞ on $\mathbb{R}^m$, positive on $B_\epsilon(0)$, 1 on $\overline{B_{\epsilon/2}(0)}$, and 0 outside $B_\epsilon(0)$. Thus, for $a \in \mathbb{R}^m$, $\tilde{g}(x) = g(x-a)$ is C^∞ on $\mathbb{R}^m$, 1 on $\overline{B_{\epsilon/2}(a)}$, and 0 outside $B_\epsilon(a)$. Let $B_\epsilon(a_i)$, $i = 1, \ldots, k$, be a finite collection of balls in $\mathbb{R}^m - F$ such that $K \subset \cup_i B_{\epsilon/2}(a_i)$. For each i, let $\tilde{g}_i$ be as above. Then the function $f : \mathbb{R}^m \to [0, 1]$ defined by

$$f(x) = 1 - \Pi_{i=1}^k (1 - \tilde{g}_i(x))$$

satisfies the desired conclusions.

QED

Corollary *Suppose $U \subset \mathbb{R}^m$ is open and $f : U \to \mathbb{R}$ is C^∞. Let p be a point in U. There is a neighborhood, V, of p, with $V \subset U$, and a C^∞ function $g : \mathbb{R}^m \to \mathbb{R}$ such that $g = f$ on V and $g = 0$ outside of U.*

Proof Since $\mathbb{R}^m$ is regular and locally compact, choose neighborhoods V and W of p such that $p \in V \subset \bar{V} \subset W \subset \bar{W} \subset U$ with $\bar{V}$ compact. By the lemma, there is a smooth function $h : \mathbb{R}^m \to \mathbb{R}$ such that $h = 1$ on $\bar{V}$ and $h = 0$ outside W. Define g by letting $g = fh$ on U and letting $g = 0$ on $\mathbb{R}^m - U$. Then g is smooth and satisfies the desired properties.

QED

Lemma 1.8 *Let M be a smooth manifold. Suppose $A \subset M$ is closed, $K \subset M$ is compact, and $A \cap K = \emptyset$. Then there is a smooth function, $f : M \to [0, 1]$ such that $f = 1$ on K and $f = 0$ on A.*

Proof Given $p \in K$, let (φ, U) be a chart at p such that $U \cap A = \emptyset$, $\varphi(p) = 0$, and $\varphi(U) = B_r(0) \subset \mathbb{R}^m$. We will first show that there exist neighborhoods of p, W and V, such that $p \in W \subset V \subset \bar{V} \subset U$, with $\bar{V}$ compact in M, and a smooth function $g : M \to [0, 1]$ such that $g = 1$ on W, g is strictly positive on V, and $g = 0$ outside of V.

Choose $\epsilon < r$, so that $B_{\epsilon/2}(0) \subset B_\epsilon(0) \subset \overline{B_\epsilon(0)} \subset B_r(0)$. Let $W = \varphi^{-1}(B_{\epsilon/2}(0))$ and $V = \varphi^{-1}(B_\epsilon(0))$. From the Lemma above, there is a C^∞ function $\tilde{g} : \mathbb{R}^m \to [0,1]$ such that $\tilde{g} = 1$ on $\overline{B_{\epsilon/2}(0)}$, $\tilde{g}$ is positive on $B_\epsilon(0)$, and $\tilde{g} = 0$ outside $B_\epsilon(0)$.

It is clear that $W \subset V$. The closure of $B_\epsilon(0)$ in $B_r(0)$ is the same as the closure of $B_\epsilon(0)$ in $\mathbb{R}^m$. Thus, $\varphi^{-1}(\overline{B_\epsilon(0)}) = cl_U(\varphi^{-1}(B_\epsilon(0))) = cl_U(V)$, where cl_U denotes the closure of a set in U. (We use an overhead bar to denote closure in the whole space, M.) Moreover, $\overline{B_\epsilon(0)}$ is compact in $B_r(0)$, so $\varphi^{-1}(\overline{B_\epsilon(0)})$ is compact in U. That is, $cl_U(V)$ is compact in U. Following reasoning similar to that used in Lemma 1.1, we find that $cl_U(V)$ is compact in M and that $cl_U(V) = \bar{V}$. So $\bar{V}$ is compact in M and $\bar{V} = cl_U(V) \subset U$.

Define $g : M \to [0,1]$ by

$$g(q) = \begin{cases} \tilde{g}(\varphi(q)), & q \in U \\ 0, & q \in M - U. \end{cases}$$

Note that $q \in W \Rightarrow g(q) = 1$, $q \in V \Rightarrow g(q) > 0$, and $q \in U - V \Rightarrow g(q) = 0$. Thus, $g = 1$ on W, $g > 0$ on V, and $g = 0$ outside V. On the open set U, g agrees with the C^∞ function $\tilde{g} \circ \varphi$. On the open set $M - \bar{V}$, g is identically 0. So g is smooth on two open sets which, together, cover all of M. Hence, g is smooth.

Now, for each $p \in K$, choose a chart (φ_p, U_p) so that $U_p \cap A = \emptyset$, and let W_p, V_p, g_p be the corresponding neighborhoods and function as determined above, so that $p \in W_p \subset V_p \subset \bar{V}_p \subset U_p$, g_p maps M smoothly into $[0,1]$, $g_p = 1$ on W_p, $g_p > 0$ on V_p, and $g_p = 0$ oustide of V_p. The collection $\{W_p\}_{p \in K}$ covers K, so there is a finite collection of points, $\{p_1, \ldots, p_n\}$, such that $K \subset \bigcup_{i=1}^n W_{p_i}$. Define $f : M \to [0,1]$ by

$$f(q) = 1 - \Pi_{i=1}^n (1 - g_{p_i}(q)).$$

Clearly f is smooth. If $q \in M$, then $0 \leq g_{p_i}(q) \leq 1$ for each i, from which it follows that $0 \leq f(q) \leq 1$. So f does, indeed, map into $[0,1]$. If $q \in K$, then $q \in W_{p_i}$ for some $i \in \{1, \ldots, n\}$, so $1 - g_{p_i}(q) = 0$ for this i and $f(q) = 1$. That is, $f = 1$ on K. Finally, suppose $q \in A$. Then q cannot be in $\bigcup_{p \in K} U_p$ since $U_p \cap A = \emptyset$ for all $p \in K$. So, $q \notin V_{p_i}$ for all $i = 1, \ldots, n$, implying that $g_{p_i}(q) = 0$ for all such i. Thus, $f(q) = 0$, and we see that $f = 0$ on A. So, f is the desired function.

QED

Corollary 1.9 *Let $U \subset M$ be open, and let $f : U \to \mathbb{R}$ be smooth. Let p be a fixed point in U. Then there is a neighborhood, V, of p, with $V \subset U$, and a smooth function $g : M \to \mathbb{R}$ such that $g = f$ on V and $g = 0$ outside U.*

Proof As in the corollary above, we use the local compactness and the regularity of M to obtain neighborhoods V and W of p such that $\bar{V}$ is compact and $p \in V \subset \bar{V} \subset W \subset \bar{W} \subset U$. Apply the previous Lemma (with $K = \bar{V}$ and $A = M - W$) to get a smooth function $h : M \to [0,1]$ such that $h = 1$ on $\bar{V}$ and $h = 0$ on $M - W$. Define $g : M \to [0,1]$ by letting $g = fh$ on U and letting $g = 0$ on $M - U$. Then $q \in V \subset \bar{V} \Rightarrow h(q) = 1 \Rightarrow g(q) = f(q)$, so $g = f$ on V. Clearly g is smooth on U. On $M - \bar{W}$, $g = 0$ because h vanishes outside of W. So g is smooth on $M - \bar{W}$. Since U and $M - \bar{W}$ cover M, g is smooth on M.

QED

Corollary 1.10 *Let M be a smooth manifold. Given any point $p \in M$ and any neighborhood, U, of p, there is a smooth function $f : M \to [0,1]$ (called a* **bump function***) such that f has compact support, supp$(f) \subset U$, and f is identically 1 on some neighborhood of p contained in U.*

Proof Choose neighborhoods V and W of p such that $p \in V \subset \bar{V} \subset W \subset \bar{W} \subset U$ and $\bar{W}$ is compact (thus making $\bar{V}$ compact). Applying Lemma 1.8, we obtain a smooth function $f : M \to [0,1]$ such that $f = 1$ on $\bar{V}$ and $f = 0$ outside W. Then *supp*$(f) \subset \bar{W}$, so f has compact support and *supp*$(f) \subset U$. Moreover, f is identically 1 on V.

QED

Of these results, Corollary 1.10 and the existence of bump functions will prove most useful.

As a final note on smooth functions and maps, we will consider more carefully their coordinate representations. In the following chapters, when we discuss tangent vectors and derivatives of maps, we will find a way to represent the derivatives of these smooth maps. For the present, however, we will give some simple examples that will be useful later on.

Let I be an interval in $\mathbb{R}$ (closed or open, perhaps infinite), and let $\gamma : I \to M$ be a smooth map. This is a curve in the manifold M. If $p \in \gamma(I)$ and (φ, U) is a coordinate system around p, then the coordinate representation of γ on $I \cap \gamma^{-1}(U)$ is given by

$$(\varphi \circ \gamma)(t) = (x^1(t), x^2(t), ..., x^m(t))$$

where each x^i is a real-valued coordinate function on $I \cap \gamma^{-1}(U)$. That is, if π^i is the i^{th} coordinate projection on $\mathbb{R}^m$, then each x^i is defined by $x^i = \pi^i \circ \varphi \circ \gamma$. The smoothness of these maps implies that each x^i is smooth on the open set $I \cap \gamma^{-1}(U)$. This coordinate representation is just a usual curve in $\mathbb{R}^m$, and it has derivative $((x^1)'(t), ..., (x^m)'(t))$. We will see later that this is a coordinate representation of the tangent vector to the the curve γ.

Next, consider a point, $p \in M$, and suppose there are two charts, (φ, U) and (ψ, V), such that $p \in U \cap V$. Then we have two different means of representing points near p by local coordinates. Each coordinate mapping defines m coordinate functions from M to $\mathbb{R}$. That is, we can represent the map φ on U by $\varphi(q) = (x^1(q), ..., x^m(q))$, where each x^i maps U into $\mathbb{R}$. We refer to x^i as the i^{th} coordinate function of φ. If π^i is the i^{th} coordinate projection in $\mathbb{R}^m$, then $x^i = \pi^i \circ \varphi$. It is easily seen that the smoothness of φ implies the smoothness of the functions x^i. Likewise, there are smooth functions y^j such that $\psi(q) = (y^1(q), ..., y^m(q))$ for $q \in V$.

Now, a point $q \in U \cap V$ will have coordinates $\varphi(q) \in \varphi(U \cap V)$ and $\psi(q) \in \psi(U \cap V)$. Suppose we want to change from one coordinate system to another, say from ψ-coordinates to φ-coordinates. We need only apply to $\psi(q)$ the map $\varphi \circ \psi^{-1}$, which will give us $\varphi(q)$. Likewise, we change the other way by applying to $\varphi(q)$ the map $\psi \circ \varphi^{-1}$. In terms of the coordinate functions, these mappings indicate that we can think of the functions x^i as each being functions of the y^j's, and vice versa. In short, we have, via the mapping $\varphi \circ \psi^{-1}$, that

$$(y^1, ..., y^m) \longmapsto \big(x^1(y^1, ..., y^m), ..., x^m(y^1, ..., y^m)\big),$$

and, via the map $\psi \circ \varphi^{-1}$, we have

$$(x^1, ..., x^m) \longmapsto \big(y^1(x^1, ..., x^m), ..., y^m(x^1, ..., x^m)\big).$$

Moreover, since the manifold is smooth, the coordinate transformations are C^∞. Thus, each of the functions $x^i(y^1, \ldots, y^m)$ is smooth, as is each of the functions $y^j(x^1, \ldots, x^m)$.

As a final remark on notation, whenever we utilize a specific chart (φ, U) and we need to specifically use the coordinate system, we will often simply write $\varphi = (x^1, \ldots, x^m)$ to indicate that the x^i are the coordinate functions of φ.

2 The Tangent and Cotangent Spaces

2.1 The Tangent Space

The tangent space is the first additional structure we will add to our smooth manifolds. The motivation for this concept is clear. In classical differential geometry, the notion of a tangent plane to a surface was a well-defined concept, and it could be identified with a 2-dimensional subspace of $\mathbb{R}^3$, thus giving us the idea of the tangent space. However, in that context, the tangent space at a particular point was defined by means of tangent vectors to curves in the surface passing through that point. It is possible to extend this idea to manifolds, and we will do so. We will proceed, though, with a definition of the tangent space that is completely equivalent, though not quite as familiar. For the remainder of this section, p is a fixed point in the smooth manifold, M.

Let $C^\infty(p)$ be the set of all real-valued functions that are defined and smooth on some neighborhood of p. Thus,

$$C^\infty(p) = \{f : U_f \to \mathbb{R} \,\big|\, U_f \ is \ a \ neighborhood \ of \ p, \ f \ is \ C^\infty \ on \ U_f\}.$$

Note that we can always assume that a neighborhood, U_f, associated with a function, f, lies within a coordinate neighborhood.

Define a relation $\sim$ on $C^\infty(p)$ by $f \sim g \Leftrightarrow$ there exists some neighborhood V of p on which $f = g$. It is easy to see that this is an equivalence relation on $C^\infty(p)$, and we will denote the equivalence class containing f by $\tilde{f}$. Denote the quotient space $C^\infty(p)/\sim$ by $\mathscr{F}_p$.

Lemma 2.1 *$\mathscr{F}_p$ is a commutative algebra with a multiplicative identity*

<u>Proof</u> For $\tilde{f}, \tilde{g} \in \mathscr{F}_p$, define $\tilde{f} + \tilde{g}$ to be the equivalence class of the function $f + g$, where $f + g$ is defined on $N_f \cap N_g$. That is, we define $\tilde{f} + \tilde{g} = \widetilde{f + g}$. We need to show that this is independent of the choice of representatives from $\tilde{f}$ and $\tilde{g}$. Suppose $f_1, f_2 \in \tilde{f}$ and $g_1, g_2 \in \tilde{g}$. Then there are neighborhoods, N_f and N_g, of p such that $f_1 = f_2$ on N_f and $g_1 = g_2$ on N_g. Then, for $q \in N_f \cap N_g$, we have $f_1(q) = f_2(q)$ and $g_1(q) = g_2(q)$, implying that $f_1(q) + g_1(q) = f_2(q) + g_2(q)$. So, $f_1 + g_1 \sim f_2 + g_2 \Rightarrow \widetilde{f_1 + g_1} = \widetilde{f_2 + g_2}$. Hence, the addition operation is well-defined. Thus, we have a well-defined notion of vector addition in $\mathscr{F}_p$.

A similar argument shows that scalar multiplication is well-defined. The vector space properties are easily established from here. For example, the zero element, $\tilde{0}$, is the set of all functions in $C^\infty(p)$ that vanish in some neighborhood of p.

Finally, define $\tilde{f}\tilde{g} = \widetilde{fg}$. As with addition, this is a well-defined multiplication operation. Associativity and commutativity of multiplication follow easily, and distributivity of multiplication over addition is also clear. The multiplicative identity is just the set of all functions that are equal to the constant, 1, on some neighborhood of p.

QED

The reason for defining the equivalence classes above is one of convenience. We will generally only be concerned with derivatives of smooth functions at p, and functions that are equal on some neighborhood of p will have the same derivatives. This brings us to our definition of the tangent space at p.

Definition 2.1 *A* **tangent vector** *at p is a map, $X_p : \mathscr{F}_p \to \mathbb{R}$ such that, for any $\tilde{f}, \tilde{g} \in \mathscr{F}_p$ and $\alpha, \beta \in \mathbb{R}$,*

i) $X_p(\alpha\tilde{f} + \beta\tilde{g}) = \alpha X_p(\tilde{f}) + \beta X_p(\tilde{g})$ *(linearity)*

ii) $X_p(\tilde{f}\tilde{g}) = f(p)X_p(\tilde{g}) + X_p(\tilde{f})g(p)$ *(Leibniz property).*

Note that these are the defining characteristics of derivative operator. In general, a map on an algebra of functions satisfying properties *i* and *ii* is called a derivation. So, a tangent vector is just a derivation on $\mathscr{F}_p$. The set of all tangent vectors at p is called the *tangent space* at p, and is denoted by T_pM. Since a tangent vector is a local object, we will often denote an element of T_pM by X_p, to indicate that X_p is a tangent vector at p.

Clearly, T_pM is nonempty, as the zero derivation that maps all functions to 0 will be in this set. However, if it is not clear that there are nontrivial derivations, it will be as we progress and construct explicit examples. Indeed, we will construct a basis for T_pM with respect to a given coordinate chart.

First, however, note that T_pM is made into a real vector space in the obvious way. For X_p and Y_p in T_pM, and any $\tilde{f} \in \mathscr{F}_p$, $\alpha \in \mathbb{R}$, we define

$$\begin{aligned} (X_p + Y_p)(\tilde{f}) &= X_p(\tilde{f}) + Y_p(\tilde{f}) \\ (\alpha X_p)(\tilde{f}) &= \alpha X_p(\tilde{f}). \end{aligned}$$

Of course, one must show that these definitions do, in fact, produce tangent vectors, but that is an easy exercise.

Note that the existence and linear structure of T_pM is coordinate independent. We have not used any local coordinate system around p to construct anything so far. This shows that the tangent space and its linear structure are well-defined concepts, independent of any choice of coordinates. However, to construct a basis and represent tangent vectors in explicit form, a coordinate system is necessary.

Let (φ, U) be a coordinate system around p. In Euclidean spaces, we can identify the standard basis vectors with the partial derivative operators that give the directional derivatives in the direction of the coordinate axes. So, it is natural for us to ask whether we can construct a basis for our tangent space using the same idea. In fact, we can. For $i = 1, 2, ..., m$ define mappings $\partial_i : \mathscr{F}_p \to \mathbb{R}$ by

$$\partial_i(\tilde{f}) = \frac{\partial}{\partial x^i}(f \circ \varphi^{-1})\Big|_{\varphi(p)}.$$

Note that this is a well-defined mapping, since any $f_1, f_2 \in \tilde{f}$ are equal on some neighborhood of p, implying that $f_1 \circ \varphi^{-1}$ and $f_2 \circ \varphi^{-1}$ are equal on some neighborhood of $\varphi(p)$. Hence, ∂_i just maps $\tilde{f}$ to the i^{th} partial derivative of this coordinate representation of f at $\varphi(p)$.

Lemma 2.2 *Each ∂_i is a tangent vector. That is, each ∂_i is a derivation on $\mathscr{F}_p$.*

Proof Let $\tilde{f}$ and $\tilde{g}$ be in $\mathscr{F}_p$. Then

$$\partial_i(\tilde{f} + \tilde{g}) = \partial_i(\widetilde{f+g}) = \frac{\partial}{\partial x^i}\big((f+g) \circ \varphi^{-1}\big)\Big|_{\varphi(p)}.$$

But $(f+g) \circ \varphi^{-1} = f \circ \varphi^{-1} + g \circ \varphi^{-1}$, so

$$\begin{aligned} \partial_i(\tilde{f} + \tilde{g}) &= \frac{\partial}{\partial x^i}\big(f \circ \varphi^{-1} + g \circ \varphi^{-1}\big)\Big|_{\varphi(p)} \\ &= \frac{\partial}{\partial x^i}\big(f \circ \varphi^{-1}\big)\Big|_{\varphi(p)} + \frac{\partial}{\partial x^i}\big(g \circ \varphi^{-1}\big)\Big|_{\varphi(p)} \\ &= \partial_i(\tilde{f}) + \partial_i(\tilde{g}). \end{aligned}$$

Now, if $\alpha \in \mathbb{R}$, then we have

$$\begin{aligned}
\partial_i(\alpha\tilde{f}) &= \partial_i(\widetilde{\alpha f}) \\
&= \frac{\partial}{\partial x^i}\big((\alpha f)\circ\varphi^{-1}\big)\Big|_{\varphi(p)} \\
&= \frac{\partial}{\partial x^i}\big(\alpha(f\circ\varphi^{-1})\big)\Big|_{\varphi(p)} \\
&= \alpha\frac{\partial}{\partial x^i}\big(f\circ\varphi^{-1}\big)\Big|_{\varphi(p)} \\
&= \alpha\partial_i(\tilde{f}).
\end{aligned}$$

Finally, we have

$$\begin{aligned}
\partial_i(\tilde{f}\tilde{g}) &= \partial_i(\widetilde{fg}) \\
&= \frac{\partial}{\partial x^i}\big((fg)\circ\varphi^{-1}\big)\Big|_{\varphi(p)} \\
&= \frac{\partial}{\partial x^i}\big((f\circ\varphi^{-1})(g\circ\varphi^{-1})\big)\Big|_{\varphi(p)} \\
&= (f\circ\varphi^{-1})(\varphi(p))\frac{\partial}{\partial x^i}(g\circ\varphi^{-1})\Big|_{\varphi(p)} + (g\circ\varphi^{-1})(\varphi(p))\frac{\partial}{\partial x^i}(f\circ\varphi^{-1})\Big|_{\varphi(p)} \\
&= f(p)\partial_i(\tilde{g}) + g(p)\partial_i(\tilde{f}).
\end{aligned}$$

Thus, each E_i is tangent vector.

QED

Before we continue in showing that the set $\{\partial_i : 1 \leq i \leq m\}$ forms a basis for T_pM, we should say something about our coordinate representations. We have been using the partial derivative notation with respect to the variable x^i. There is a specific meaning behind this. These are the coordinate functions representing φ. As in the previous section, if π^i is the i^{th} coordinate projection from $\mathbb{R}^m$ to $\mathbb{R}$, then, for $q \in U$, $x^i(q) = \pi^i(\varphi(q))$. Note that these coordinate functions are smooth real-valued functions on U. Our tangent vectors, ∂_i, then, are just the partial derivatives of $f \circ \varphi^{-1}$ with respect to these coordinate functions.

Now, we will prove that the tangent vectors $\{\partial_i\}$ are linearly independent. Let 0_p denote the zero element of T_pM. That is, $0_p(\tilde{f}) = 0$ for all $\tilde{f} \in \mathscr{F}_p$. Suppose there are scalars $c_1, ..., c_m$ such that

$$\sum_{i=1}^{m} c_i \partial_i = 0_p$$

Consider the element $\sum_{i=1}^{m} c_i\partial_i$ acting on $\tilde{x}^k$, the equivalence class containing the k^{th} coordinate function. We have

$$\begin{aligned}
\sum_{i=1}^{m} c_i \partial_i(\tilde{x}^k) &= \sum_{i=1}^{m} c_i \frac{\partial}{\partial x^i}\left(x^k \circ \varphi^{-1}\right)\Big|_{\varphi(p)} \\
&= \sum_{i=1}^{m} c_i \frac{\partial}{\partial x^i}\left(\pi^k \circ \varphi \circ \varphi^{-1}\right)\Big|_{\varphi(p)} \\
&= \sum_{i=1}^{m} c_i \frac{\partial}{\partial x^i}\pi^k\Big|_{\varphi(p)} \\
&= \sum_{i=1}^{m} c_i \delta_i^k \\
&= c_k \\
&= 0.
\end{aligned}$$

This shows that $c_k = 0$. Applying this sum to each k, in turn, shows that $c_1 = c_2 = \cdots = c_m = 0$. Hence, these tangent vectors are linearly independent. To prove that they span T_pM, we need a preliminary result from multivariable calculus. This is just a consequence of the multivariable mean value theorem.

Lemma 2.3 *Let f be a real-valued function that is differentiable on some neighborhood, U, of $a \in \mathbb{R}^m$, and suppose that U is such that the straight line from a to x, for any $x \in U$, is contained in U. Such a neighborhood is called* **star-shaped about** *a. Then there are functions $g_k : U \to \mathbb{R}$, for $k = 1, 2, ..., m$, such that*

$$f(x) = f(a) + \sum_{k=1}^{m}(x^k - a^k)g_k(x) \qquad \textit{and} \qquad g_k(a) = \frac{\partial f}{\partial x^k}\Big|_a.$$

Proof For each $x \in U$, define $h_x : [0,1] \to \mathbb{R}$ by $h_x(t) = f(a + t(x-a))$. Then we have

$$\frac{dh_x(t)}{dt} = \sum_{k=1}^{m}(x^k - a^k)D_k f(a + t(x-a)).$$

By the fundamental theorem of calculus, we know that

$$\int_0^1 h'_x(t)dt = h_x(1) - h_x(0) = f(x) - f(a),$$

so it follows that

$$f(x) - f(a) = \sum_{k=1}^{m}(x^k - a^k)\int_0^1 D_k f(a + t(x-a))dt \qquad\qquad \forall\, x \in U.$$

Letting $g_k(x)$ be the integral in the above expression, we have $g_k(a) = D_k f(a)$ and

$$f(x) = f(a) + \sum_{k=1}^{m}(x^k - a^k)g_k(x).$$

QED

To use this result, we must see how to translate it to the manifold. Suppose f is a C^∞ function defined on some neighborhood, V, of p. Assume, by shrinking V if necessary, that $V \subset U$ and that V is the image under φ^{-1} of an open ball at p (or any open set in $\varphi(U)$ that is star-shaped about $\varphi(p)$). This implies that $f \circ \varphi^{-1}$ is a C^∞ function on $\varphi(V) \subset \varphi(U)$. We can apply this lemma to conclude that there are functions $g_k : \varphi(V) \to \mathbb{R}$, such that, for all $x \in \varphi(V)$,

$$(f \circ \varphi^{-1})(x) = (f \circ \varphi^{-1})(\varphi(p)) + \sum_{k=1}^{m}\big(x^k - [\varphi(p)]^k\big)g_k(x)$$

where $[\varphi(p)]^k$ is the k^{th} coordinate of $\varphi(p) \in \mathbb{R}^m$ and

$$g_k(\varphi(p)) = \frac{\partial}{\partial x^k}\big(f \circ \varphi^{-1}\big)\Big|_{\varphi(p)}.$$

Given $q \in V$, with $x = \varphi(q)$, the above equality implies that

$$f(q) = f(p) + \sum_{k=1}^{m}\big((\pi^k \circ \varphi)(q) - (\pi^k \circ \varphi)(p)\big)g_k\big(\varphi(q)\big). \qquad (1)$$

Since this last equality holds on V, we see that f and the function on the right-hand side belong to the same equivalence class. Remember that p is fixed in this discussion, so $f(p)$ and $(\pi^k \circ \varphi)(p)$ are constants in the right-hand side.

Lemma 2.4 *The elements* $\{\partial_i \,:\, 1 \leq i \leq m\}$ *span* T_pM.

<u>Proof</u> Let X_p be an arbitrary tangent vector, and let $\tilde{f}$ be in $\mathscr{F}_p$. By the representation formula we just derived, there is some neighborhood, V, of p, contained in U, and there are functions g_k, $k = 1, \ldots, m$, such that (1) holds for all $q \in V$. It follows that the function on the right-hand side of (1) represents the same equivalence class as f. Thus, we have

$$\begin{aligned}
X_p(\tilde{f}) &= X_p\left(\widetilde{f(p)} + \sum_{k=1}^{m}\Big[\widetilde{(\pi^k \circ \varphi)} - \widetilde{(\pi^k \circ \varphi)(p)}\Big]\widetilde{(g_k \circ \varphi)}\right) \\
&= X_p\big(\widetilde{f(p)}\big) + \sum_{k=1}^{m} X_p\Big[\Big(\widetilde{(\pi^k \circ \varphi)} - \widetilde{(\pi^k \circ \varphi)(p)}\Big)\widetilde{(g_k \circ \varphi)}\Big] \\
&= X_p\big(\widetilde{f(p)}\big) + \sum_{k=1}^{m} X_p\Big(\widetilde{(\pi^k \circ \varphi)} - \widetilde{(\pi^k \circ \varphi)(p)}\Big) g_k(\varphi(p)) \\
&\qquad + \sum_{k=1}^{m}\Big((\pi^k \circ \varphi)(p) - (\pi^k \circ \varphi)(p)\Big) X_p\big(\widetilde{g_k \circ \varphi}\big) \\
&= X_p\big(\widetilde{f(p)}\big) + \sum_{k=1}^{m} X_p\big(\widetilde{(\pi^k \circ \varphi)} - \widetilde{(\pi^k \circ \varphi)(p)}\big) g_k(\varphi(p)).
\end{aligned}$$

It is easy to see that a tangent vector (or any derivation) must map to 0 any element of $\mathscr{F}_p$ that is constant on a neighborhood of p. Since p is fixed in this argument, we get

$$X_p(\tilde{f}) = \sum_{k=1}^{m} X_p(\widetilde{\pi^k \circ \varphi})\partial_k(\tilde{f}).$$

Since $\tilde{f} \in \mathscr{F}_p$ was arbitrary, we can write this as the operator equation

$$X_p = \sum_{k=1}^{m} X_p(\tilde{x}^k)\partial_k,$$

and this shows that the set $\{\partial_i \,:\, 1 \leq i \leq m\}$ spans T_pM. Moreover, it also shows that the components of a tangent vector, X_p, with respect to the basis $\{\partial_i\}$ are the values obtained by X_p acting on the equivalence classes of the coordinate functions representing φ.

QED

As a consequence of this result, we see that the dimension of the tangent space, T_pM, is the same as that of the manifold at p. It should be emphasized again, however, that the basis we have constructed depends upon the coordinate system (φ, U). If we had used another chart, (ψ, V), such that $p \in V$, then we would still have obtained a basis for an m-dimensional vector space, but that basis would, in general, be different. We call the basis $\{\partial_i \ : \ 1 \leq i \leq m\}$ obtained at p from a particular chart the ***standard coordinate frame***, or simply the ***coordinate frame***, induced by the coordinate system.

As another consequence of this theorem, it becomes clear that we can consider tangent vectors as acting on individual functions instead of equivalence classes of functions. This will make our notation a great deal more convenient. To see why this follows, suppose $f \sim g$ in $C^\infty(p)$. Then, since f and g agree on some neighborhood of p, we see from the representation of X_p in the coordinate frame that

$$X_p(\tilde{f}) = \sum_{k=1}^{m} X_p(\tilde{x}^k)\frac{\partial}{\partial x^k}(f \circ \varphi^{-1})\Big|_{\varphi(p)} = \sum_{k=1}^{m} X_p(\tilde{x}^k)\frac{\partial}{\partial x^k}(g \circ \varphi^{-1})\Big|_{\varphi(p)} = X_p(\tilde{g}).$$

In other words, the value $X(\tilde{f})$ can be computed using any function in the equivalence class $\tilde{f}$. Thus, we can actually formally define $X_p(f)$, for $f \in C^\infty(p)$, to be the value $X_p(\tilde{f})$. The linearity and Leibniz properties still hold in this context since $\alpha X_p(f) + \beta X_p(g) = \alpha X_p(\tilde{f}) + \beta X_p(\tilde{g}) = X_p(\alpha\tilde{f} + \beta\tilde{g}) = X_p(\widetilde{\alpha f + \beta g}) = X_p(\alpha f + \beta g)$ and $X_p(fg) = X_p(\widetilde{fg}) = f(p)X_p(\tilde{g}) + g(p)X_p(\tilde{f}) = f(p)X_p(g) + g(p)X_p(f)$.

First, since we have shown that the action of a tangent vector on $\tilde{f}$ is independent of the function, f, we choose from $\tilde{f}$, we will omit the tilde in our notation whenever convenient. For example, the operator equation we derived for the representation of a tangent vector with respect to a particular coordinate system can be more succinctly written as

$$X_p = \sum_{k=1}^{m} X_p(x^k)E_k,$$

where it is implied that x^k actually represents the equivalence class $\tilde{x}^k$. This alternate notation will not be used exclusively. In particular, when we discuss the cotangent space later on, we will use the tilde again, since referring to the entire equivalence class will be necessary. However, it should be clear from the context which particular notation is appropriate. In general, all of the mappings and operators we define will be independent of any particular element chosen from an equivalence class. So, there will be no difference in the results regardless of which notation we use. This is really just a matter of convenient notation, which is very useful in differential geometry. The technical nature of the field is such that one's results can quickly become mired in what Elie Cartan called the "debauch of indices." Thus, expressions are simplified, both visually and mathematically, whenever possible.

Second, the following transformation theorem, along with its analogue in the case of the cotangent space, will require that we take partial derivatives of the change of variable maps. Recall that if (φ, U) and (ψ, V) are two coordinate systems such that $p \in U \cap V$, then we can change from φ-coordinates to ψ-coordinates via the map $\psi \circ \varphi^{-1}$, as well as from ψ-coordinates to φ-coordinates via the inverse map, $\varphi \circ \psi^{-1}$. Each of these is a map from $\mathbb{R}^m$ into $\mathbb{R}^m$. As before, we will denote the coordinate functions of φ by x^i, meaning that, for $p \in U$, we have $x^i(p) = \pi^i(\varphi(p))$, where π^i is the i^{th} coordinate projection in $\mathbb{R}^m$. Likewise, we will denote the coordinate functions of ψ by y^i. Hence, in forming the composition $\varphi \circ \psi^{-1}$, we see that each coordinate map, x^i, becomes a function of the coordinates $y^1, \ldots, y^m$, since $\varphi \circ \psi^{-1}$ maps points described by ψ-coordinates to points described by φ-coordinates. In other words, we can represent this map in terms of its input and output coordinates by the expression

$$\varphi \circ \psi^{-1}(y^1, \ldots, y^m) = (x^1(y^1, \ldots, y^m), \ldots, x^m(y^1, \ldots, y^m)).$$

Similarly, in forming the composition $\psi \circ \varphi^{-1}$, we see that each coordinate map, y^i, becomes a function of the coordinates $x^1, \ldots, x^m$, since $\psi \circ \varphi^{-1}$ maps points described by φ-coordinates to points described by ψ-coordinates. That is, we obtain a representation of the form

$$\psi \circ \varphi^{-1}(x^1, \ldots, x^m) = (y^1(x^1, \ldots, x^m), \ldots, y^m(x^1, \ldots, x^m)).$$

Now, suppose we wish to compute the i^{th} partial derivative of the j^{th} component of the map $\varphi \circ \psi^{-1}$ at the point $\psi(p)$. The correct, coordinate-free notation, of course, would be to write this partial derivative as

$$D_i\big(\varphi \circ \psi^{-1}\big)^j\Big|_{\psi(p)},$$

where D_i represents the general partial derivative operator with respect to the i^{th} variable of the function in question, and the superscript j indicates the j^{th} component function of the map $\varphi \circ \psi^{-1}$. Note that this is equivalent to the classically motivated expression

$$\frac{\partial\left(\pi^j \circ \varphi \circ \psi^{-1}\right)}{\partial y^i}\Big|_{\psi(p)},$$

since composing π^j with $\varphi \circ \psi^{-1}$ just gives us the j^{th} component function of this map. The classical partial derivative operator, $\partial/\partial y^i$, simply indicates that the derivative is to be taken with respect to the i^{th} variable. Since the composition operation is associative, consider composing π^j with φ in this expression, and leaving ψ^{-1} as is. We know that $\pi^j \circ \varphi$ is just x^j. So, the previous expression is further equivalent to

$$\frac{\partial\left(x^j \circ \psi^{-1}\right)}{\partial y^i}\Big|_{\psi(p)}.$$

So, in a very real sense, we are simply computing the i^{th} partial derivative, with respect to the ψ-coordinate system, of the j^{th} φ-coordinate. These particular partial derivatives are often simply written in the abbreviated form

$$\frac{\partial x^j}{\partial y^i}\Big|_{\psi(p)},$$

since the fact that we are composing x^j with ψ^{-1} is implied by the fact that we are taking the partial derivative with respect to one of the ψ-coordinates. The point of all of this is to alleviate any confusion that might arise later on. This abbreviated form is just a notational device. It is very elegant and convenient, but it should always be taken to represent the formal derivative $D_i(\varphi \circ \psi^{-1})^j|_{\psi(p)}$. In the exact same fashion, we will use the abbreviated expression

$$\frac{\partial y^j}{\partial x^i}\Big|_{|\varphi(p)}$$

to represent the formal partial derivative $D_i(\psi \circ \varphi^{-1})^j|_{\varphi(p)}$, which is the i^{th} partial derivative of the j^{th} component of the map $\psi \circ \varphi^{-1}$ at $\varphi(p)$. With these conventions in mind, we can finally present the following result.

Theorem 2.5 *Suppose $p \in U \cap V$, where (φ, U) and (ψ, V) are two coordinate systems around p. Let $\varphi = (x^1, \ldots, x^m)$ and let $\psi = (y^1, \ldots, y^m)$ denote the coordinate functions making up the coordinate representation of φ, and let $\{y^i\}$ denote the coordinates of ψ. Let $\{E_i^x\}$ and $\{E_i^y\}$ denote the coordinate frames induced by (φ, U) and (ψ, V), respectively, on T_pM. Then*

$$E_i^x = \sum_{k=1}^{m} \frac{\partial y^k}{\partial x^i}\Big|_{\varphi(p)} E_k^y \qquad \textit{and} \qquad E_j^y = \sum_{l=1}^{m} \frac{\partial x^l}{\partial y^j}\Big|_{\psi(p)} E_l^x.$$

If $X_p = \sum \alpha^i E_i^x = \sum \beta^j E_j^y$ *is a tangent vector at* p*, then*

$$\alpha^i = \sum_{j=1}^{m} \beta^j \frac{\partial x^i}{\partial y^j}\Big|_{\psi(p)} \qquad \textit{and} \qquad \beta^j = \sum_{i=1}^{m} \alpha^i \frac{\partial y^j}{\partial x^i}\Big|_{\varphi(p)}.$$

Proof The proof is just an application of change of coordinates at p. By our previous lemma, we know that $E_i^x = \sum_k E_i^x(y^k) E_k^y$. By our definition of E_i^x, and using our abbreviated notation, we have

$$\begin{aligned} E_i^x(y^k) &= \frac{\partial}{\partial x^i}\big(y^k \circ \varphi^{-1}\big)\Big|_{\varphi(p)} \\ &= \frac{\partial\big(\pi^k \circ \psi \circ \varphi^{-1}\big)}{\partial x^i}\Big|_{\varphi(p)} \\ &= D_i\big(\psi \circ \varphi^{-1}\big)^k\Big|_{\varphi(p)} \\ &= \frac{\partial y^k}{\partial x^i}\Big|_{\varphi(p)}. \end{aligned}$$

It follows that

$$E_i^x = \sum_{k=1}^{m} \frac{\partial y^k}{\partial x^i}\Big|_{\varphi(p)} E_k^y,$$

and the second formula follows similarly. Once those formulas are established, given X_p, we know that its i^{th} component, α_i, with respect to the basis $\{E_i^x\}$, is given by $X_p(x^i)$. So, we have

$$\begin{aligned} \alpha^i &= X_p(x^i) \\ &= \sum_{j=1}^{m} \beta^j E_j^y(x^i) \\ &= \sum_{j=1}^{m} \beta^j \frac{\partial x^i}{\partial y^j}\Big|_{\psi(p)}, \end{aligned}$$

and this verifies the third formula. The last one follows in the same manner.

QED

There is an interesting consequence of this last lemma. Recall from elementary linear algebra that, if $\mathscr{B}_1 = \{e_1, \ldots, e_m\}$ and $\mathscr{B}_2 = \{f_1, \ldots, f_m\}$ are two bases for an m-dimensional vector space, V, then we can easily construct a change of basis matrix, $P_{1,2}$. Given any vector, $v \in V$, if $[v]_1$ represents the vector of the components of v with respect to $\mathscr{B}_1$, then $P_{1,2}(v)$ will give us $[v]_2$, the vector of the components of v with respect to $\mathscr{B}_2$. The i^{th} column of $P_{1,2}$ is given by $[e_i]_2$, the vector of the components of e_i with respect to $\mathscr{B}_2$. Likewise, the inverse of this matrix, denoted $P_{2,1}$, maps $[v]_2$ to $[v]_1$.

Now, consider applying this idea to the results given in the previous lemma. If X_p is an arbitrary tangent vector in T_pM, with representations $X_p = \sum \alpha^i E_i^x = \sum \beta^j E_j^y$, then we can transform between the two component vectors $[X_p]_\varphi = [\alpha^1 \cdots \alpha^m]^T$ and $[X_p]_\psi = [\beta^1 \cdots \beta^m]^T$. According to the lemma, the matrix that maps $[X_p]_\varphi$ to $[X_p]_\psi$ is given by

$$\begin{bmatrix} \frac{\partial y^1}{\partial x^1} & \cdots & \frac{\partial y^1}{\partial x^m} \\ \vdots & \ddots & \vdots \\ \frac{\partial y^m}{\partial x^1} & \cdots & \frac{\partial y^m}{\partial x^m} \end{bmatrix},$$

where the partial derivatives are evaluated at $\varphi(p)$. Thus, the change of basis matrix that maps tangent vectors in span$\{E_i^x\}$ to tangent vectors in span$\{E_j^y\}$ is just the Jacobian of the change of coordinate mapping $\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$ evaluated at $\varphi(p)$. Likewise, the change of basis matrix that maps tangent vectors in span$\{E_j^y\}$ to tangent vectors in span$\{E_i^x\}$ is the inverse of that matrix, which is nothing more than the Jacobian of the change of coordinate mapping $\varphi \circ \psi^{-1} : \psi(U \cap V) \to \varphi(U \cap V)$ evaluated at $\psi(p)$. These Jacobians are often given in shorthand notation by

$$D(\psi \circ \varphi^{-1})(p) \qquad \textit{and} \qquad D(\varphi \circ \psi^{-1})(p).$$

3.2 The Cotangent Space

We have seen that the tangent space T_pM has the same dimension as the manifold, which we assume to be finite. Hence, after choosing a basis, the tangent space is isomorphic to $\mathbb{R}^m$. Consequently, it has a well-defined dual space of the same dimension. This dual space is called the *cotangent space*, and it is the set of all linear functionals on T_pM. We denote the cotangent space by T_p^*M. It is a perplexing but common feature of most differential geometry texts to leave the discussion of the cotangent space after just these few statements, if it is even mentioned at all. It is instructive, however, to actually construct the cotangent space independently, as we did the tangent space, and then produce a pairing between the two showing that they are dual spaces of each other. Given what we have constructed so far, we can show that the cotangent space at a point p is actually a quotient space of $\mathscr{F}_p$.

Let Γ_p be the set of all smooth curves in M through p. By a suitable translation, we can assume that each curve maps $0 \in \mathbb{R}$ to p and is defined on an interval of the form $(-\epsilon, \epsilon)$. So, precisely, we have

$$\Gamma_p = \{\gamma : (-\epsilon, \epsilon) \to M \mid \gamma(0) = p,\ \gamma\, is\, C^\infty\}.$$

Note that ϵ depends on γ in this definition. Now, define a map $T : \mathscr{F}_p \times \Gamma_p \to \mathbb{R}$ by

$$T(\tilde{f}, \gamma) = \frac{d}{dt}\big(f \circ \gamma\big)\Big|_{t=0}.$$

This map is well-defined, for if $f_1, f_2 \in \tilde{f}$, then $f_1 = f_2$ on some neighborhood, N, of p, implying that $f_1 = f_2$ on $N \cap \gamma(-\epsilon, \epsilon)$. T is also linear in the first variable, since if $\tilde{f}, \tilde{g} \in \mathscr{F}_p$ and $\alpha, \beta \in \mathbb{R}$, then

$$\begin{aligned}
T(\alpha\tilde{f} + \beta\tilde{g}, \gamma) &= T(\widetilde{\alpha f + \beta g}, \gamma) \\
&= \frac{d}{dt}\big((\alpha f + \beta g) \circ \gamma\big)\Big|_{t=0} \\
&= \frac{d}{dt}\big(\alpha f \circ \gamma + \beta g \circ \gamma\big)\Big|_{t=0} \\
&= \alpha\frac{d}{dt}\big(f \circ \gamma\big)\Big|_{t=0} + \beta\frac{d}{dt}\big(g \circ \gamma\big)\Big|_{t=0} \\
&= \alpha T(\tilde{f}, \gamma) + \beta T(\tilde{g}, \gamma).
\end{aligned}$$

Next, define a set $\mathscr{H}_p$ by

$$\mathscr{H}_p = \{\tilde{f} \in \mathscr{F}_p \ :\ T(\tilde{f}, \gamma) = 0\ \ \forall\, \gamma \in \Gamma_p\}.$$

Since T is linear in the first variable, it follows that $\mathscr{H}_p$ is a linear subspace of $\mathscr{F}_p$. In fact, we can completely characterize those elements of $\mathscr{F}_p$ that are in $\mathscr{H}_p$.

Lemma 2.6 *Let $\tilde{f}$ be in $\mathscr{F}_p$. Then $\tilde{f} \in \mathscr{H}_p$ if and only if for any coordinate chart, (φ, U), with $p \in U$, we have*

$$\frac{\partial}{\partial x^i}\big(f \circ \varphi^{-1}\big)\Big|_{\varphi(p)} = 0$$

for all $i = 1, 2, ..., m$. In short, $\tilde{f} \in \mathscr{H}_p$ if and only if all coordinate partial derivatives of f vanish at p.

<u>Proof</u> Let (φ, U) be any coordinate system around p, and let γ be any curve in Γ_p. We can write the coordinate representation of γ as

$$(\varphi \circ \gamma)(t) = (x^1(t), ..., x^m(t)),\ t \in (\epsilon, \epsilon).$$

Then for any $f \in \tilde{f}$, we have

$$\begin{aligned}
T(\tilde{f}, \gamma) &= \frac{d}{dt}\big(f \circ \gamma\big)\Big|_{t=0} \\
&= \frac{d}{dt}\big(f \circ \varphi^{-1} \circ \varphi \circ \gamma\big)\Big|_{t=0} \\
&= \frac{d}{dt}\big(f \circ \varphi^{-1}\big)\big(x^1(t), ..., x^m(t)\big)\Big|_{t=0} \\
&= \sum_{i=1}^{m} \frac{\partial}{\partial x^i}\big(f \circ \varphi^{-1}\big)\Big|_{\varphi(p)} \frac{dx^i}{dt}\Big|_{t=0}. \qquad (2)
\end{aligned}$$

So, if all the partial derivatives of $f \circ \varphi^{-1}$ vanish at p, we must have $T(\tilde{f}, \gamma) = 0$. Conversely, suppose $\tilde{f} \in \mathscr{H}_p$, so that $T(\tilde{f}, \gamma) = 0$ for all $\gamma \in \Gamma_p$. For each $i = 1, 2, ..., m$, define a particular curve in Γ_p as follows. Let $\lambda_i : \mathbb{R} \to \mathbb{R}^m$ be given by $\lambda_i(t) = ((\varphi(p))^1, ..., (\varphi(p))^{i-1}, (\varphi(p))^i + t, (\varphi(p))^{i+1}, ..., (\varphi(p))^m)$, where $(\varphi(p))^k$ is the k^{th} component of $\varphi(p) \in \mathbb{R}^m$. Then define $\gamma_i : \mathbb{R} \to M$ by $\gamma_i = \varphi^{-1} \circ \lambda_i$. Each curve, γ_i, is smooth and satisfies $\gamma_i(0) = p$. Then the coordinate representation, $\varphi \circ \gamma_i = (x^1(t), ..., x^m(t))$, satisfies

$$\frac{dx^k}{dt}\Big|_{t=0} = \begin{cases} 1 & k = i \\ 0 & k \neq i \end{cases}.$$

Thus, applying this to equation (2), we see that

$$T(\tilde{f}, \gamma_i) = 0 \Rightarrow \frac{\partial}{\partial x^i}\big(f \circ \varphi^{-1}\big)\Big|_{\varphi(p)} = 0.$$

Applying this to each $i = 1, 2, ..., m$ shows that $\frac{\partial}{\partial x^i}(f \circ \varphi^{-1}) = 0$ at $\varphi(p)$ for all i. Moreover, the coordinate system (φ, U) was arbitrary. So, this proves the result.

QED

We are now ready to define the cotangent space. Recall that, for a vector space, V, and a subspace $W \subset V$, the quotient space V/W is defined as follows. We define an equivalence relation on V by $x \approx y$ if and only if $x - y \in W$. Then V/W is the set of all equivalence classes under this relation. It is a linear space with the operations $[x] + [y] = [x + y]$ and $\alpha[x] = [\alpha x]$. The zero element of V/W is the entire subspace W. Using this concept, we have the following.

Definition 2.2 *The quotient space $\mathscr{F}_p/\mathscr{H}_p$ is called the* **cotangent space** *and is denoted by T_p^*M. The $\mathscr{H}_p$ equivalence class of $\tilde{f}$ is denoted $d\tilde{f}_p$, and is called a* **cotangent vector** *at p.*

So, we see right away that T_p^*M is a linear space, as it is the quotient space defined by a subspace of $\mathscr{F}_p$. Moreover, by the general structure of a quotient space, we know that for any $\tilde{f}, \tilde{g} \in \mathscr{F}_p$ and $\alpha \in \mathbb{R}$, we have $d\tilde{f}_p + d\tilde{g}_p = d(\widetilde{f+g})_p$ and $\alpha d\tilde{f}_p = d(\widetilde{\alpha f})_p$.

As a remark on notation, we will usually denote a cotangent vector simply by $d\tilde{f}$, if it is clear from the context to which specific point we are referring. Since all of our discussions in this chapter are focused on the single point $p \in M$, we will use this notation.

The definition of the cotangent space may not be as immediately transparent as that of the tangent space. A moment of thought, though, will convince us that two elements $\tilde{f}, \tilde{g} \in \mathscr{F}_p$ are in the same $\mathscr{H}_p$ equivalence class if and only if each of the coordinate partial derivatives of f, in terms of any coordinate system around p, equals the corresponding coordinate partial derivative of g. This characterization does not, of course, depend on which functions $f \in \tilde{f}$ and $g \in \tilde{g}$ we might choose to compute such partial derivatives. Hence, a cotangent vector, $d\tilde{f}$, can be thought of as a set of smooth functions defined in a neighborhood of p that all have identical coordinate partial derivatives at p.

Our next goal is to show that the dimension of T_p^*M is m. To do this, we will proceed as in the case of the tangent space, choosing a local coordinate system and constructing a basis. Our construction thus far has not made use of any coordinate system, so the existence and linear structure of T_p^*M is well-defined and independent of any coordinate system around p.

To illustrate, first, the natural choice for a basis, consider the following simple example. Let $M = \mathbb{R}^2$, and let $p = (0,0)$. Suppose two functions, f and g, defined and smooth on respective neighborhoods of p, have the same coordinate partial derivatives at the origin. Then we would equate these functions as a cotangent vector. In essence, their gradients, ∇f and ∇g, are equal at p. Any other function whose gradient equalled these would also be identified with f and g. Hence, we can think of the cotangent vector $d\tilde{f}$ as simply defining a vector in $\mathbb{R}^2$ emanating from the origin, namely the vector defined by the gradient of any function identified with f. Conversely, given a vector in $\mathbb{R}^2$, is straightforward to construct a smooth function having this vector as its gradient. Hence, we can actually identify the cotangent space with $\mathbb{R}^2$. This space is spanned by the standard basis vectors $e_1 = (1, 0)$ and $e_2 = (0, 1)$. So, to think of these vectors as a basis for the cotangent space at p, we need to consider what functions have these vectors as gradients. The simple choice, of course, is the set of coordinate functions $f_1(x, y) = x$ and $f_2(x, y) = y$. That is, the coordinate functions are the natural choice for a basis for the cotangent space. This is the idea we will pursue.

Let (φ, U) be a coordinate system around p. Then we have local coordinates on U defined by $\varphi(q) = (x^1(q), ..., x^m(q))$ for $q \in U$. The functions $x^k : U \to \mathbb{R}$ are the coordinate functions on U. As before, we will also use the fact that each x^i can be obtained by composing the i^{th} component projection on $\mathbb{R}^m$, π^i, with φ. That is, $x^i = \pi^i \circ \varphi$. Each coordinate function x^i is also smooth, so it is an element of $C^\infty(p)$. We want to show that $\{d\tilde{x}^i : 1 \leq i \leq m\}$ is a basis for T_p^*M induced by the coordinate system (φ, U).

We will begin by proving the following technical lemma, which will be useful in our basis construction.

Lemma 2.7 *Suppose $f_1, ..., f_k$ are in $C^\infty(p)$, and let $F(y^1, ..., y^k)$ be a C^∞ real-valued function defined in a neighborhood of $(f_1(p), ..., f_k(p))$. Then $f = F(f_1, ..., f_k)$ is in $C^\infty(p)$ and*

$$d\tilde{f} = \sum_{i=1}^{k} \frac{\partial F}{\partial f_k}\big(f_1(p), ..., f_k(p)\big) d\tilde{f}_i.$$

Proof If U_j is the neighborhood of of p on which f_j is smooth, then all the functions f_j are smooth on $N = \cap_{j=1}^{k} U_j$. It follows that $f = F(f_1, ..., f_k)$ is smooth on N. So, $f \in C^\infty(p)$, and we can consider $\tilde{f} \in F_p$ and $d\tilde{f} \in T_p^*M$. Let γ be any curve in Γ_p. Then, using the linearity of the map T in its first variable, we have

$$
\begin{aligned}
T(\tilde{f},\gamma) &= \frac{d}{dt}\big(f\circ\gamma\big)\Big|_{t=0} \\
&= \frac{d}{dt}\Big(F\big(f_1\circ\gamma,...,f_k\circ\gamma\big)\Big)\Big|_{t=0} \\
&= \sum_{i=1}^{k}\frac{\partial F}{\partial f_i}\big(f_1(p),...,f_k(p)\big)\frac{d}{dt}\big(f_i\circ\gamma\big)\Big|_{t=0} \\
&= \sum_{i=1}^{k}\frac{\partial F}{\partial f_i}\big(f_1(p),...,f_k(p)\big)T(\tilde{f}_i,\gamma) \\
&= T\left(\sum_{i=1}^{k}\frac{\partial F}{\partial f_i}\big(f_1(p),...,f_k(p)\big)\tilde{f}_i,\gamma\right).
\end{aligned}
$$

This implies that

$$
T\left(\tilde{f}-\sum_{i=1}^{k}\frac{\partial F}{\partial f_i}\big(f_1(p),...,f_k(p)\big)\tilde{f}_i,\gamma\right)=0.
$$

Since γ was arbitrary, this implies that

$$
\tilde{f}-\sum_{i=1}^{k}\frac{\partial F}{\partial f_i}\big(f_1(p),...,f_k(p)\big)\tilde{f}_i\in\mathscr{H}_p.
$$

Applying the linear structure of the cotangent space, it follows that

$$
d\tilde{f}=\sum_{i=1}^{k}\frac{\partial F}{\partial f_i}\big(f_1(p),...,f_k(p)\big)d\tilde{f}_i.
$$

QED

Now, to show that the cotangent vectors $d\tilde{x}^i$ form a basis for T_p^*M, we first need to know that they form a set of m distinct elements. That is, we want to know that if $i \neq j$, then $d\tilde{x}^i \neq d\tilde{x}^j$. This result is obvious in the Euclidean case, but it may not be in an arbitrary coordinate system. So, suppose i and j are indices such that $1 \leq i < j \leq m$. If $\tilde{x}^i = \tilde{x}^j$, then there is a neighborhood, N, of p on which $x^i = x^j$. Thus, every point $(y^1, ..., y^m) \in \varphi(N)$ has the same i^{th} and j^{th} coordinates, implying that there is an open ball, $B_\epsilon(\varphi(p))$, such that every point in this ball, including $\varphi(p)$, has the same i^{th} and j^{th} coordinates. Consider, however, the point $y = \varphi(p) + \frac{\epsilon}{2}e_i$, where e_i is the i^{th} standard basis vector in $\mathbb{R}^m$. Then this point is in the ball $B_\epsilon(\varphi(p))$, but its i^{th} and j^{th} coordinates are not the same, as we have shifted one by a positive amount while leaving the other fixed. Hence we cannot have $\tilde{x}^i = \tilde{x}^j$ when $i \neq j$. This shows that the elements $\{\tilde{x}^i \in F_p \,:\, 1 \leq i \leq m\}$ are all distinct.

Next, with indices i and j as before, suppose $d\tilde{x}^i = d\tilde{x}^j$. Then for any $k = 1, ..., m$, we have

$$\frac{\partial}{\partial x^k}\big(x^i \circ \varphi^{-1}\big)\Big|_{\varphi(p)} = \frac{\partial}{\partial x^k}\big(x^j \circ \varphi^{-1}\big)\Big|_{\varphi(p)}.$$

But $x^i = \pi^i \circ \varphi$, so $x^i \circ \varphi^{-1} = \pi^i$, and likewise for x^j. Thus, letting $k = i$, the previous equation implies that

$$\frac{\partial}{\partial x^i}\pi^i\Big|_{\varphi(p)} = \frac{\partial}{\partial x^i}\pi^j\Big|_{\varphi(p)},$$

from which it follows that $1 = 0$, as $i \neq j$. The contradiction shows that, if $i \neq j$, then $d\tilde{x}^i \neq d\tilde{x}^j$. Hence, the set $\{d\tilde{x}^i \,:\, 1 \leq i \leq m\}$ does, indeed, contain m distinct elements. Thus, we can prove the following result.

Theorem 2.8 *The set $\{d\tilde{x}^i \,:\, 1 \leq i \leq m\}$ forms a basis for T_p^*M with respect to the coordinate system (φ, U).*

Proof The fact that this set spans T_p^*M follows from Lemma 3.7. We let $k = m$ and let the functions $f_1, ..., f_m$ be the coordinate functions $x^1, ..., x^m$. Then, for any element $d\tilde{f} \in T_p^*M$, we let F be $f \circ \varphi^{-1}$. Then the function $g = F(x^1, ..., x^m)$ satisfies $g = f \circ \varphi^{-1} \circ \varphi = f$. Hence, $d\tilde{g} = d\tilde{f}$, and the lemma implies that

$$\begin{aligned} d\tilde{f} &= \sum_{i=1}^{m} \frac{\partial F}{\partial x^i}\big(x^1(p), ..., x^m(p)\big)d\tilde{x}^i \\ &= \sum_{i=1}^{m} \frac{\partial\big(f \circ \varphi^{-1}\big)}{\partial x^i}\Big|_{\varphi(p)} d\tilde{x}^i. \end{aligned}$$

So, the set $\{d\tilde{x}^i \,:\, 1 \leq i \leq m\}$ spans T_p^*M. Now, let $d\tilde{0}$ denote the zero element of T_p^*M, and suppose there exist scalars $c_1, c_2, ..., c_m$ such that

$$\sum_{i=1}^{m} c_i d\tilde{x}^i = d\Big(\sum_{i=1}^{m} c_i \tilde{x}^i\Big) = d\tilde{0}.$$

Then, using the fact that T_p^*M is a quotient space, it follows that

$$\sum_{i=1}^m c_i \tilde{x}^i \in \mathscr{H}_p.$$

So, for any $\gamma \in \Gamma_p$, we have

$$\begin{aligned} T\Big(\sum_{i=1}^m c_i\tilde{x}^i, \gamma\Big) &= \sum_{i=1}^m c_i T(\tilde{x}^i, \gamma) \\ &= \sum_{i=1}^m c_i \frac{d}{dt}\big(x^i \circ \gamma\big)\Big|_{t=0} \\ &= 0. \end{aligned}$$

Now, for each $i = 1, ..., m$, define curves γ_i as in the proof of Lemma 3.6. Then

$$\begin{aligned} T\Big(\sum_{i=1}^m c_i\tilde{x}^i, \gamma_j\Big) &= \sum_{i=1}^m c_i \frac{d}{dt}\big(x^i \circ \varphi^{-1} \circ \varphi \circ \gamma_j\big)\Big|_{t=0} \\ &= \sum_{i=1}^m c_i \frac{d}{dt}\big(x^i \circ \varphi^{-1} \circ \lambda_j\big)\Big|_{t=0} \\ &= \sum_{i=1}^m c_i \frac{d}{dt}\pi^i\big((\varphi(p))^1, ..., (\varphi(p))^{j-1}, (\varphi(p))^j + t, (\varphi(p))^{j+1}, ..., (\varphi(p))^m\big)\Big|_{t=0} \\ &= \sum_{i=1}^m c_i\delta_j^i \\ &= c_j. \end{aligned}$$

Thus, since T vanishes for any curve γ, it follows that $c_j = 0$. Letting $j = 1, ..., m$, in turn, shows that $c_1 = c_2 = \cdots = c_m = 0$. Hence, the set $\{d\tilde{x}^i \;:\; 1 \leq i \leq m\}$ is linearly independent and forms a basis for T_p^*M with respect to the coordinate system (φ, U).

QED

Note that the components of the cotangent vector $d\tilde{f}$ with respect to this basis are the partial derivatives

$$\frac{\partial\big(f \circ \varphi^{-1}\big)}{\partial x^i}\Big|_{\varphi(p)} \qquad i = 1, 2, \ldots, m.$$

These are just the partial derivatives of the coordinate representation of f with respect to the coordinate system (φ, U). This gives us another way to think of a cotangent vector. A cotangent vector $d\tilde{f}$ is essentially just a coordinate gradient of f at p. It is even more interesting to note that the i^{th} component of the cotangent vector, $d\tilde{f}$, is just the value obtained by the i^{th} tangent basis vector, E_i, acting on $\tilde{f}$. That is, it follows directly from the definition of E_i that

$$\frac{\partial\big(f \circ \varphi^{-1}\big)}{\partial x^i}\Big|_{\varphi(p)} = E_i(\tilde{f}).$$

Hence, we can represent a cotangent vector with respect to the standard cotangent basis by

$$d\tilde{f} = \sum_{i=1}^{m} E_i(\tilde{f}) d\tilde{x}^i.$$

Of course, $E_i(\tilde{f})$ is independent of the choice of $f \in d\tilde{f}$, so we usually write this expansion as

$$d\tilde{f} = \sum_{i=1}^{m} E_i(f) d\tilde{x}^i.$$

It should be pointed out, as an interesting consequence of Lemma 3.7, that we can define a Leibniz rule for cotangent vectors just as we have for tangent vectors. In the lemma, just let F be defined by $F(y^1, y^2) = y^1 y^2$. Then, for any two functions $f_1, f_2 \in C^\infty(p)$, we have $f = F(f_1, f_2) = f_1 f_2$. It follows that

$$\begin{aligned} d(\tilde{f}_1 \tilde{f}_2) &= d(\widetilde{f_1 f_2}) \\ &= d\tilde{f} \\ &= \frac{\partial F}{\partial f_1}\big(f_1(p), f_2(p)\big) d\tilde{f}_1 + \frac{\partial F}{\partial f_2}\big(f_1(p), f_2(p)\big) d\tilde{f}_2 \\ &= \tilde{f}_2(p) d\tilde{f}_1 + \tilde{f}_1(p) d\tilde{f}_2. \end{aligned}$$

That is, for $\tilde{f}, \tilde{g} \in \mathscr{F}_p$, we have $d(\tilde{f}\tilde{g}) = \tilde{f}(p) d\tilde{g} + \tilde{g}(p) d\tilde{f}$.

We can also derive a change of basis formula for the representation of a cotangent vector as we did with the tangent vectors. Let (φ, U) and (ψ, V) be two coordinate systems around p. We will denote φ-coordinates by x^i and ψ-coordinates by y^i. Thus, the standard bases for T_p^*M with respect to these coordinate systems are, respectively, $\{d\tilde{x}^i\}$ and $\{d\tilde{y}^i\}$. Similarly, the standard bases for T_pM with respect to these coordinate systems will be denoted, respectively, by $\{E_i^x\}$ and $\{E_i^y\}$. Recalling how we can express the components of a cotangent vector in terms of the action of the tangent vectors, it follows that

$$d\tilde{x}^i = \sum_{j=1}^{m} E_j^y(x^i)d\tilde{y}^j.$$

So, if the representation of an arbitrary cotangent vectgor, $d\tilde{f}$, with respect to the coordinate system (φ, U) is $d\tilde{f} = \sum_i E_i^x(f)d\tilde{x}^i$, then it follows that

$$\begin{aligned} d\tilde{f} &= \sum_{i=1}^{m} E_i^x(f) \sum_{j=1}^{m} E_j^y(x^i)d\tilde{y}^j \\ &= \sum_{j=1}^{m} \sum_{i=1}^{m} E_j^y(x^i)E_i^x(f)d\tilde{y}^j \\ &= \sum_{j=1}^{m} \left(\sum_{i=1}^{m} E_j^y(x^i)E_i^x(f) \right) d\tilde{y}^j. \end{aligned}$$

But $E_j^y(x^i)$ is just the j^{th} partial derivative of $x^i \circ \psi^{-1}$ at $\psi(p)$, for which we have an abbreviated notation. Using this notation, the previous equality reduces to

$$d\tilde{f} = \sum_{j=1}^{m} \left(\sum_{i=1}^{m} \frac{\partial x^i}{\partial y^j}\Big|_{\psi(p)} E_i^x(f) \right) d\tilde{y}^j.$$

But we also know that $d\tilde{f} = \sum_j E_j^y(f)d\tilde{y}^j$, so we must have

$$E_j^y(f) = \sum_{i=1}^{m} \frac{\partial x^i}{\partial y^j}\Big|_{\psi(p)} E_i^x(f).$$

This is the transformation law for the components of the cotangent vector. In terms of matrices, with all components considered, this equation takes the form

$$\begin{bmatrix} E_1^y(f) \\ \vdots \\ E_m^y(f) \end{bmatrix} = \begin{bmatrix} \frac{\partial x^1}{\partial y^1} & \cdots & \frac{\partial x^m}{\partial y^1} \\ \vdots & \ddots & \vdots \\ \frac{\partial x^1}{\partial y^m} & \cdots & \frac{\partial x^m}{\partial y^m} \end{bmatrix} \begin{bmatrix} E_1^x(f) \\ \vdots \\ E_m^x(f) \end{bmatrix},$$

where the partial derivatives in the matrix are evaluated at $\psi(p)$. Note that this matrix is the transpose of the Jacobian of the map $\varphi \circ \psi^{-1}$ at $\psi(p)$. Thus, the change of basis transformation that maps the φ-components of a cotangent vector to its ψ-components is given in matrix form by the transpose of $D(\varphi \circ \psi^{-1})$.

3.3 The Duality Between T_pM and T_p^*M

We have independently constructed both the tangent space and the cotangent space. Based on our constructions and the bases we derived for each space, it seems intuitive that there is some connection between the two. In fact, as a consequence of the way we have defined these spaces, the pairing between the two is quite natural. A tangent vector is an operator, acting on smooth functions defined at p. A cotangent vector is an equivalence class of functions that all have identical coordinate derivatives. The natural pairing, then, should be the action of a tangent vector on the functions that make up a cotangent vector. More precisely, we have the following.

Define a map $\omega : T_pM \times T_p^*M \to \mathbb{R}$ by

$$\omega\big(X, d\tilde{f}\big) = X(f).$$

Theorem 2.9 *The mapping ω is well-defined, meaning that it does not depend on the particular $f \in d\tilde{f}$ we choose to compute it. Moreover, it is bilinear, and, if (φ, U) is any coordinate system around p, inducing the bases $\{E_i\}$ and $\{d\tilde{x}^i\}$ for T_pM and T_p^*M, respectively, then we have $\omega(E_i, dx^j) = E_i(d\tilde{x}^j) = \delta_i^j$.*

Proof First, the mapping is clearly linear in the second variable, as any $X \in T_pM$ is a linear operator. Likewise, using the linear structure of the tangent space, we have $\omega(X_1 + X_2, d\tilde{f}) = (X_1 + X_2)(f) = X_1(f) + X_1(f)$. Scalar multiplication holds similarly, showing that ω is bilinear.

To show that the mapping is well-defined, we use local coordinate systems. Let (φ, U) be a coordinate system around p. Suppose X is a tangent vector,$d\tilde{f}$ a cotangent vector, and let f_1 and f_2 be any two functions in $d\tilde{f}$. Then $X = \sum_i \alpha_i E_i$. So,

$$
\begin{aligned}
X(f_1) &= \Big(\sum_{i=1}^{m} \alpha_i E_i\Big)(f_1) \\
&= \sum_{i=1}^{m} \alpha_i E_i(f_1) \\
&= \sum_{i=1}^{m} \alpha_i \frac{\partial}{\partial x^i}\big(f_1 \circ \varphi^{-1}\big)\Big|_{\varphi(p)} \\
&= \sum_{i=1}^{m} \alpha_i \frac{\partial}{\partial x^i}\big(f_2 \circ \varphi^{-1}\big)\Big|_{\varphi(p)} \\
&= \sum_{i=1}^{m} \alpha_i E_i(f_2) \\
&= \Big(\sum_{i=1}^{m} \alpha_i E_i\Big)(f_2) \\
&= X(f_2)
\end{aligned}
$$

Finally, since the choice of function $f \in d\tilde{f}$ does not matter, we have

$$
\begin{aligned}
E_i(d\tilde{x}^j) &= \frac{\partial}{\partial x^i}\big(x^j \circ \varphi^{-1}\big)\Big|_{\varphi(p)} \\
&= \frac{\partial}{\partial x^i}\big(\pi^j \circ \varphi \circ \varphi^{-1}\big)\Big|_{\varphi(p)} \\
&= \frac{\partial \pi^j}{\partial x^i}\Big|_{\varphi(p)} \\
&= \delta_i^j.
\end{aligned}
$$

QED

Thus, since the dual of a given vector space is unique up to isomorphism, we can conclude that the cotangent space is the dual of the tangent space. Moreover, given a coordinate system, (φ, U), around p, this shows that $\{d\tilde{x}^i : 1 \leq i \leq m\}$ is the dual basis to $\{E_i : 1 \leq i \leq m\}$. In fact, since the tangent and cotangent spaces are of finite dimension, when a coordinate system is chosen, the mapping, ω, is just a bilinear form on $\mathbb{R}^m$. Given a tangent vector X_p and a cotangent vector $d\tilde{f}_p$, we call the real number $\omega(X_p, d\tilde{f})$ the *directional derivative* of f at p in the direction of X_p with respect to the chosen coordinate system. To actually compute this value, we would have to choose a coordinate system around p. Our construction shows that the directional derivative is independent of any choice of coordinates. However, one can work out the details with two different coordinate systems and show it directly also. It is merely a change of basis coordinates. Mappings of this sort are used frequently in mathematics and physics. They are called *tensors*,and we will study such mappings in detail in Chapter 5.

We use the mapping ω to define the action of cotangent vectors on tangent vectors. Recall T_p^*M is the dual space of T_pM, so it must consist of linear functionals on T_pM. This mapping gives us the means for defining how a cotangent vector operates on a tangent vector. For $X_p \in T_pM$ and $d\tilde{f} \in T_p^*M$, we define $d\tilde{f}(X_p)$ to be $\omega(X_p, d\tilde{f}) = X_p(f)$. This is a well-defined linear functional on T_pM.

As a final note on the map, ω, there is a common and intuitive notational convenience employed in most differential geometry texts. Once bases are chosen for T_pM and T_P^*M, both spaces are isomorphic to $\mathbb{R}^m$, inducing an identification between tangent and cotangent vectors. Hence, the map ω behaves just like the usual inner product on a Euclidean space. Consequently, given $X_p \in T_pM$ and $d\tilde{f} \in T_p^*M$, the map, $\omega : T_pM \times T_p^*M \to \mathbb{R}$, is typically denoted by $\langle X_p, d\tilde{f}\rangle$. Thus, for example, the directional derivative of f at p in the direction of X_p is given by the real number $\langle X_p, d\tilde{f}\rangle$. Henceforth, we will use this notation whenever the duality map is required.

To conclude this section on the duality of the tangent and cotangent spaces, we point out an expected consequence of the formula $d(\tilde{f}\tilde{g}) = \tilde{f}(p)d\tilde{g} + \tilde{g}(p)d\tilde{f}$. The directional derivative also satisfies a Leibniz rule. If X_p is a tangent vector, then the action of X_p on a the product $\tilde{f}\tilde{g}$ is given by

$$\begin{aligned}\langle X_p, d(\tilde{f}\tilde{g})\rangle &= \langle X_p, \tilde{f}(p)d\tilde{g} + \tilde{g}(p)d\tilde{f}\rangle \\ &= \tilde{f}(p)\langle X_p, d\tilde{g}\rangle + \tilde{g}(p)\langle X_p, d\tilde{f}\rangle.\end{aligned}$$

Since $d(\tilde{f}\tilde{g}) = d(\widetilde{fg})$, we can rewrite $\langle X_p, d(\tilde{f}\tilde{g})\rangle$ using the definition of the directional derivative as $X_p(fg)$. Hence, we have

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

3.4 The Differential of a Map

As an important application of the tangent and cotangent spaces, we will introduce the notion of the differential of a map, a concept we will use repeatedly. The differential of a map is the natural generalization of the classical derivative of a function $f : \mathbb{R}^n \to \mathbb{R}^m$. The derivative of such a function is usually defined to be a linear map $Df : \mathbb{R}^n \to \mathbb{R}^m$. We want to translate this concept to functions mapping manifolds to manifolds. To do so, it is necessary to look at the classical definition in a new way.

We know that the Euclidean spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ are manifolds. It is easy to see that the tangent space of a Euclidean space at any point, p, is simply another copy of that Euclidean space. That is, for any $p \in \mathbb{R}^n$, we have $T_p\mathbb{R}^n = \mathbb{R}^n$. This is a consequence of the unique structure of the Euclidean spaces. We can think of any vector in a Euclidean space as being anchored at the origin, and, in this way, we identify all of the tangent spaces of $\mathbb{R}^n$, at any point, with $\mathbb{R}^n$ itself. So, a more precise way of defining the derivative of a map $f : \mathbb{R}^n \to \mathbb{R}^m$ at $p \in \mathbb{R}^n$ would be as a linear map from $T_p\mathbb{R}^n$ to $T_{f(p)}\mathbb{R}^m$. This distinction is not made in classical analysis for the reason we just gave. These tangent spaces are nothing more than copies of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. This is not the case for manifolds with arbitrary coordinate systems. Hence, we must define the derivative of a map in terms of the tangent spaces.

We begin by first defining the *cotangent map*.

Definition 2.3 *Let M and N be smooth manifolds, and let $F : M \to N$ be a smooth map. For $p \in M$ and $q = F(p) \in N$, the* **cotangent map** *induced by F at q is the map $F^* : T_q^*N \to T_p^*M$ defined by $F^*(d\tilde{f}_q) = d(\widetilde{f \circ F})_p$.*

Note that the direction of this mapping is opposite the direction of F.

The cotangent map is linear, for if $d\tilde{f}_q, d\tilde{g}_q \in T_q^*N$ and $\alpha, \beta \in \mathbb{R}$, then

$$\begin{aligned}
F^*(\alpha d\tilde{f}_q + \beta d\tilde{g}_q) &= F^*(d(\alpha \widetilde{f + \beta} g)_q) \\
&= d((\alpha f \widetilde{+ \beta} g) \circ F)_p \\
&= d(\alpha \widetilde{f \circ F})_p + d(\widetilde{\beta g \circ F})_p \\
&= \alpha d(\widetilde{f \circ F})_p + \beta d(\widetilde{g \circ F})_p \\
&= \alpha F^*(d\tilde{f}_q) + \beta F^*(d\tilde{g}_q).
\end{aligned}$$

We now use this map to define the differential of F.

Definition 2.4 *Let M, N, and F be as before. The* **differential***, or* **tangent map***, induced by F at p is the map $F_*: T_pM \to T_qN$ defined by $F_*(X_p)(f) = X_p(F^*(d\tilde{f}_q))$ for $f \in C^\infty(q)$.*

This definition does make sense. Since f is a smooth function in a neighborhood of q, there is a cotangent vector $d\tilde{f}_q$. The map F^* will map this cotangent vector to a cotangent vector at $p \in M$. A cotangent vector is nothing more than an equivalence class of functions all having the same coordinate partial derivatives. Hence, X_p can act on any function in the cotangent vector $F^*(d\tilde{f}_q)$ and produce the same value. We will usually denote the tangent map by DF instead of F_* and refer to it as the differential. The reason for the '*' notation is to distinguish between the tangent and cotangent maps, but we will rarely have occasion to refer to the cotangent map beyond this. So, to emphasize that the differential is a generalization of the classical derivative, we will use the familiar nomenclature and notation.

As expected, the differential is linear also. If $X_p, {X_p}' \in T_pM$ and $\alpha, \beta \in \mathbb{R}$, then, for any $f \in C^\infty(q)$, we have

$$\begin{aligned} D(\alpha X_p + \beta {X_p}')(f) &= \big(\alpha X_p + \beta {X_p}'\big)\big(F^*(d\tilde{f}_q)\big) \\ &= \alpha X_p\big(F^*(d\tilde{f}_q)\big) + \beta {X_p}'\big(F^*(d\tilde{f}_q)\big) \\ &= \alpha D(X_p)(f) + \beta D({X_p}')(f). \end{aligned}$$

So, just as in the classical case, the differential of a map, $F: M \to N$, establishes a homomorphism between the tangent spaces at p and at $q = F(p)$.

Theorem 2.10 *Let $F: M \to N$ be a smooth map between smooth manifolds. For $p \in M$, let $q = F(p)$. Then, if F is a diffeomorphism of some neighborhood of p onto some neighborhood of q, then the differential $DF: T_pM \to T_qN$ is an isomorphism. Also, if $G: N \to L$ is another smooth map between manifolds, then $D(G \circ F) = DG \circ DF$ (the chain rule).*

Proof   We prove the chain rule first, and then the first result will follow from that.

$G \circ F$ is a map from M to L. Let p be in M and let $q = F(p)$ and $r = G(F(p))$. Suppose X_p is a tangent vector at p and $f \in C^\infty(r)$. Then, by definition we have

$$D(G \circ F)(X_p)(f) = X_p\big((G \circ F)^*(d\tilde{f}_r)\big).$$

Now, $(G \circ F)^*(d\tilde{f}_r) = d(\widetilde{f \circ G \circ F})_p$, so

$$D(G \circ F)(X_p)(f) = X_p\Big(d\big(\widetilde{f \circ G \circ F}\big)_p\Big).$$

On the other hand, we have, for any function $h \in C^\infty(q)$,

$$\begin{aligned} DF(X_p)(h) &= X_p\big(F^*(d\tilde{h}_q)\big) \\ &= X_p\big(d(\widetilde{h \circ F})_p\big), \end{aligned}$$

from which it follows that

$$\begin{aligned} DG\big(DF(X_p)\big)(f) &= DF(X_p)\big(G^*(d\tilde{f}_r)\big) \\ &= DF(X_p)\big(d(\widetilde{f \circ G})_q\big) \\ &= DF(X_p)(f \circ G) \\ &= X_p\Big(F^*\big(d(\widetilde{f \circ G})_q\big)\Big) \\ &= X_p\big(d(\widetilde{f \circ G \circ F})_p\big). \end{aligned}$$

Thus, $(DG \circ DF)(X_p) = D(G \circ F)(X_p)$, and this proves the chain rule.

Now, suppose $F : M \to N$ is a diffeomorphism of a neighborhood, V, of p onto a neighborhood, W, of q. Then $F^{-1} : N \to M$ is also a well-defined smooth map between manifolds. Let Y_q be a tangent vector in T_qN. Then $DF^{-1}(Y_q)$ is a tangent vector, X_p, in T_pM, and the chain rule implies that $DF(X_p) = DF(DF^{-1}(Y_p)) = D(F \circ F^{-1})(Y_q) = Y_q$. Hence, DF is surjective. If $DF(X_{p_1}) = DF(X_{p_2})$, then $DF^{-1}(DF(X_{p_1}) = DF^{-1}(DF(X_{p_2})$, implying that

$$D(F^{-1} \circ F)(X_{p_1}) = D(F^{-1} \circ F)(X_{p_2}) \Rightarrow X_{p_1} = X_{p_2}.$$

So DF is injective. Hence, it is an isomorphism. Note that this also requires the dimensions of M at p and N at q to be equal.

QED

In terms of local coordinates, we can establish a familiar expression for the differential of a map.

Theorem 2.11 *Let $F : M \to N$ be a smooth map between manifolds, and let p be a point in M with $q = F(p)$. Suppose (φ, U) and (ψ, V) are coordinate systems around p and q respectively, with respective coordinate functions denoted by $\{x^i\}$ and $\{y^j\}$.*

Let $\{E_i^x\}$ and $\{E_j^y\}$ be the bases induced by each coordinate system for T_pM and T_qN, respectively. Then, for each $i = 1, \ldots, m$,

$$DF(E_i^x) = \sum_{j=1}^{n} \left(\frac{\partial F^j}{\partial x^i} \right) \Bigg|_{\varphi(p)} E_j^y,$$

where F^j denotes the j^{th} coordinate function of the coordinate representation of F, $\psi \circ F \circ \varphi^{-1}$. Moreover, if $X_p = \sum_i \alpha_i E_i^x$ and $DF(X_p) = \sum_j \beta_j E_j^y$, then

$$\beta_j = \sum_{i=1}^{m} \alpha_i \left(\frac{\partial F^j}{\partial x^i} \right) \Bigg|_{\varphi(p)}.$$

Proof The proof of the first formula is just a computation. The components of a tangent vector with respect to a given coordinate system are given by the values obtained by that tangent vector acting on the coordinate functions. That is, the j^{th} component of $DF(E_i^x)$ is given by $DF(E_i^x)(y^j)$. Thus, we have

$$\begin{aligned} DF(E_i^x)(y^j) &= E_i^x\big(F^*(d\tilde{y}^j_q)\big) \\ &= E_i^x\big(d(\widetilde{y^j \circ F})_p\big) \\ &= E_i^x(y^j \circ F) \\ &= \frac{\partial}{\partial x^i}\big(y^j \circ F \circ \varphi^{-1}\big). \end{aligned}$$

But $y^j \circ F \circ \varphi^{-1}$ is just the j^{th} coordinate function of the coordinate representation, $\psi \circ F \circ \varphi^{-1}$, of F. Denoting this component function by F^j, we get

$$DF(E_i^x)(y^j) = \frac{\partial F^j}{\partial x^i},$$

and this is the j^{th} component of $DF(E_i^x)$. Hence,

$$DF(E_i^x) = \sum_{j=1}^{n} \left(\frac{\partial F^j}{\partial x^i} \right) \Bigg|_{\varphi(p)} E_j^y.$$

Now, suppose $X_p = \sum_i \alpha_i E_i^x$ and $DF(X_p) = \sum_j \beta_j E_j^y$. Then, we just use the linearity of the differential to obtain

$$
\begin{aligned}
DF(X_p) &= DF\Big(\sum_{i=1}^{m} \alpha_i E_i^x\Big) \\
&= \sum_{i=1}^{m} \alpha_i DF(E_i^x) \\
&= \sum_{i=1}^{m} \alpha_i \sum_{j=1}^{n} \Big(\frac{\partial F^j}{\partial x^i}\Big)\Big|_{\varphi(p)} E_j^y \\
&= \sum_{j=1}^{n} \left(\sum_{i=1}^{m} \alpha_i \Big(\frac{\partial F^j}{\partial x^i}\Big)\Big|_{\varphi(p)} \right) E_j^y.
\end{aligned}
$$

So, if β_j is the j^{th} component of $DF(X_p)$, then this shows that

$$
\beta_j = \sum_{i=1}^{m} \alpha_i \Big(\frac{\partial F^j}{\partial x^i}\Big)\Big|_{\varphi(p)}.
$$

QED

Notice that this theorem states that the local coordinate representation of DF is just the Jacobian of the coordinate representation $\psi \circ F \circ \varphi^{-1}$. Moreover, for the basis vector E_i^x, the components of the image $DF(E_i^x)$ are just the entries of the i^{th} column of this Jacobian matrix.

As an application of the differential concept, we will discuss the tangent vectors of curves in the manifold M. This is a notion we will need later on when we construct a metric on M in terms of the arclengths of curves.

Let $\gamma : (a, b) \to M$ be a smooth map between the submanifold $(a, b) \subset \mathbb{R}$ and M. Looking at (a, b) as a manifold, we see that its tangent space is 1-dimensional. So, once we choose a basis, the tangent space $T_{t_0}(a, b)$ at any point $t_0 \in (a, b)$ is just the set of real numbers, thought of as a vector space. This tangent space is spanned by the tangent vector E_{t_0} defined by

$$
E_{t_0}(f) = \frac{df}{dt}\Big|_{t_0},
$$

for any smooth function, f, defined on some neighborhood of t_0. Hence, E_{t_0} is just the ordinary derivative operator on $\mathbb{R}$ evaluated at t_0. Let us denote E_{t_0} by the more intuitive symbol d/dt, where the point t_0 at which the derivative is evaluated is assumed to be known. Suppose $\gamma(t_0) = p$. The differential of γ at t_0, which we will denote $D_{t_0}\gamma$, is a linear map from $T_{t_0}(a,b)$ to T_pM. For any function $f \in C^\infty(p)$, we have

$$D_{t_0}\gamma\Big(\frac{d}{dt}\Big)f = \frac{d}{dt}\big(f \circ \gamma\big)\Big|_{t_0}.$$

This is simply the directional derivative of f in the direction of the curve γ. More formally, this is the image under $D_{t_0}\gamma$ of d/dt, and we will call this tangent vector in T_pM the *tangent vector to the curve* γ at $p = \gamma(t_0)$. Since the curve, γ, is smooth, we can let t vary over (a,b), and we obtain a tangent vector to γ at every point $p \in \gamma(a,b)$. The field of tangent vectors we obtain is called the *tangent vector field* of γ.

Now, suppose (φ, U) is a coordinate system around $p = \gamma(t_0)$. The local coordinate representation of γ on $\gamma(a,b) \cap U$ is given by $\varphi \circ \gamma(t) = (x^1 \circ \gamma(t), \ldots, x^m \circ \gamma(t))$. Let $\{E_i\}$ denote the basis induced by this coordinate system at T_pM. The i^{th} component of $D_{t_0}\gamma$ with respect to this basis is the action of this tangent vector on x^i. Thus, this component is given by

$$D_{t_0}\gamma\Big(\frac{d}{dt}\Big)x^i = \frac{d}{dt}\big(x^i \circ \gamma\big)\Big|_{t_0},$$

which is just the ordinary derivative of the i^{th} component function of the coordinate representation of γ. If, for brevity of notation, we denote this ordinary derivative, evaluated at t_0, by

$$\frac{d}{dt}\big(x^i \circ \gamma\big)\Big|_{t_0} = \dot{x}^i(t_0),$$

then we have

$$D_{t_0}\gamma\Big(\frac{d}{dt}\Big) = \sum_{i=1}^{m} \dot{x}^i(t_0)E_i. \tag{3}$$

This is the local coordinate representation of the tangent vector to the curve, γ at $p = \gamma(t_0)$.

Remarks on Notation and Interpretation We have two remarks to make on notation before we move on. The first one concerns representation of cotangent vectors. Henceforth, for simplicity of notation, we will denote cotangent vectors simply by df, or, if we need to reference the point at which the cotangent space exists, df_p. The transition from $d\tilde{f}$ should not cause any confusion. We have shown in our constructions that all of our definitions and results are independent of which

$f \in \tilde{f}$ or $d\tilde{f}$ we might choose. Indeed, any two functions $f_1, f_2 \in \tilde{f}$ will be equal near p, so they will certainly be in the same $\mathscr{H}_p$-equivalence class $d\tilde{f}$. Moreover, as we have also shown, any computations we might perform with cotangent (or tangent) vectors are independent of the choice of function from the equivalence class $d\tilde{f}$. Thus, referencing a cotangent vector as df will simply refer to the cotangent vector induced by the equivalence class $\tilde{f}$.

Secondly, the notation we have used to represent the standard basis vectors in T_pM is common, but not standard. We have used E_{p_i} to denote the tangent vector at p giving the partial derivative of a smooth function at p with respect to its i^{th} coordinate. We will continue with our current notation for tangent vectors throughout. However, there is a more intuitively pleasing notation that is used more often. We give a description of this notation in the appendix. We have chosen not to use it here because the notation is such that one tends to lose sight of the fact that we are seeking coordinate independent results. Hence, when learning the material, we feel that the general operator form is more appropriate.

Finally, we want to point out that there are many different constructions and interpretations of the tangent space. Undoubtedly, the interpretation of a tangent vector as an operator on smooth functions is the most concise and useful when doing analysis on a manifold. However, the geometric intuition of these operators is lacking. There is another means of constructing the tangent space in terms of the tangent vectors to curves passing through points on the manifold. This definition is entirely equivalent to the one we have given here, and many think it is more geometrically pleasing. To give the reader a different, but just as valid, point of view, we have placed a description of the tangent space using curves in the appendix.

Having developed the local structure of a manifold in great detail, we will turn, in the next chapter, to a study of the global structure of manifolds.

3 Mapping Theorems and Submanifolds

in progress

4 Vector Fields and the Tangent Bundle

4.1 Classical Vector Fields

All of our constructions in the previous chapter were local in character. They were all restricted to a single point, p, in the manifold. This can only take us so far, though. Thus, we will, from here on, be extracting global consequences of these constructions.

The first construction that should be discussed in any branch of global analysis is that of a vector field. A motivation for the study of vector fields is not necessary. They are essential in numerous mathematical and physical disciplines, including geometry, analysis, and dynamics. A manifold and its tangent bundle, which we will define soon, is the most natural setting for a general study of vector fields. The tangent bundle is essentially the union of all the tangent spaces over a manifold. Vector fields are simply fields of tangent vectors over a manifold, so they exist, naturally, in the collection of tangent spaces. The tangent bundle has a great deal of structure, however, and this provides an elegant setting for the study of vector fields. Hence, we will devote this entire chapter to a thorough description of the tangent bundle and its place as the natural setting for vector fields.

To further motivate the usefulness of the tangent bundle, however, we will begin by giving the classical definition of a smooth vector field on a manifold. It will be seen that this definition and characterization of smoothness is awkward. The modern definitions utilizing the tangent bundle are much more concise.

Definition 4.1 *A* **tangent vector field***, X, on a smooth manifold, M, is an assignment of a tangent vector, $X_p \in T_pM$, to each $p \in M$. That is, at each p, $X(p)$ is an element of T_pM. We will denote vector fields by letters X, Y, etc, and we will denote their vectors at particular points, p, by X_p or $X(p)$.*

We will usually simply refer to a tangent vector field as a vector field. Note that a vector field, X, defines a field of operators that act on smooth functions defined on the manifold. Let $C^\infty(M)$ denote the collection of smooth functions defined on M. For any $f \in C^\infty(M)$, define a real-valued function on M by

$$(Xf)(p) = X_p(f).$$

That is, the value of Xf at p is the value of the tangent vector, X_p, acting on f.

Defintion 4.2 *Let X be a vector field on M. We say that X is a* **smooth vector field** *if for any $f \in C^\infty(M)$, we have $Xf \in C^\infty(M)$.*

As a consequence of this definition, it follows that a smooth vector field defines an operator from $C^\infty(M)$ to itself. That is, we can think of X as mapping $f \in C^\infty(M)$ to $Xf \in C^\infty(M)$. Moreover, by the properties of tangent vectors, this operator satisfies $X(\alpha f + \beta g) = \alpha Xf + \beta Xg$ and $X(fg) = f \cdot Xg + g \cdot Xf$, where $f, g \in C^\infty(M)$ and $\alpha, \beta \in \mathbb{R}$.

Observe that, restricted to a particular coordinate neighborhood, U, with corresponding coordinate mapping, φ, we can express a vector field, X, in local coordinates. For each $p \in U$, this coordinate system induces the standard coordinate frame, $\{E_{p_i}\}$. The vector field on U whose tangent vector at each $p \in U$ is E_{p_i} is smooth, since E_{p_i} is just a partial derivative operator acting on smooth functions. Hence, we have a *standard coordinate basis field* on U, denoted by $\{E_i\}_{i=1}^m$, which is a set of vector fields on U such that at each $p \in U$, the set $\{E_{p_i}\}$ is the standard coordinate frame at for T_pM. Moreover, each vector field in this standard basis field is smooth. Hence, at each $p \in U$, there are scalars $\alpha_1, \dots, \alpha_m$ such that we can express $X(p)$ in the form

$$X(p) = \sum_{i=1}^{m} \alpha_i E_{p_i}.$$

Letting p vary over all of U, we obtain m real-valued functions, $\alpha_i : U \to \mathbb{R}$, $i = 1, \dots, m$, such that $\alpha_i(p)$ is the i^{th} coordinate of $X(p)$ with respect to the basis $\{E_{p_i}\}$. Hence, we can represent the vector field, X, on U by

$$X = \sum_{i=1}^{m} \alpha_i E_i,$$

where we interpret each α_i as a real-valued function on U and each E_i as the vector field on U such that $E_i(p) = E_{p_i}$. Using this local representation, we can give another characterization of smoothness based on local coordinates.

Lemma 4.1 *A vector field, X, is smooth if and only if for every coordinate chart, (φ, U), on M, the coordinate representation of X restricted to U is given by $\sum_i \alpha_i E_i$ where each α_i is a smooth function on U.*

Proof Suppose, first, that for every coordinate chart, (φ, U), on M, we have $X|_U = \sum_i \alpha_i E_i$ where each α_i is a smooth function on U. Let f be in $C^\infty(M)$. Then $f|_U$ is a smooth function on U. Hence,

$$\begin{aligned} Xf|_U(p) &= \sum_{i=1}^{m} \alpha_i(p) E_{p_i}(f) \\ &= \sum_{i=1}^{m} \alpha_i(p) \frac{\partial}{\partial x^i} \big(f \circ \varphi^{-1}\big)\Big|_{\varphi(p)} \end{aligned}$$

But, since f is smooth, the function

$$p \mapsto \frac{\partial}{\partial x^i} \big(f \circ \varphi^{-1}\big)\Big|_{\varphi(p)}$$

is smooth on U. Hence, $Xf|_U$ is a finite sum of smooth functions on U, implying that $Xf|_U$ is a smooth function on U. The coordinate chart, (φ, U), was arbitrary, though, so this implies that Xf is a smooth function on a collection of open sets covering M, namely the collection of coordinate neighborhoods. Hence, Xf must be smooth on M. Since $f \in C^\infty(M)$ was arbitrary, this shows that X is a smooth vector field.

Conversely, suppose X is smooth according to the definition, and let (φ, U) be any coordinate chart. Then X restricted to U is a smooth vector field on U. We can express $X|_U$ as

$$X|_U = \sum_{i=1}^{m} \alpha_i E_i,$$

where each α_i is a real-valued function on U and $E_i(p) = E_{p_i}$. The coordinate functions $x^j : U \to \mathbb{R}$ are smooth functions on U, so, by hypothesis, $X|_U(x^j) : U \to \mathbb{R}$ is a smooth function for each $j = 1, \ldots, m$. But, for any $p \in U$, we have

$$\begin{aligned} X|_U(x^j)(p) &= \sum_{i=1}^{m} \alpha_i(p) E_{p_i}(x^j) \\ &= \sum_{i=1}^{m} \alpha_i(p) \delta^i_j \\ &= \alpha_j(p). \end{aligned}$$

Thus, $X|_U(x^j) = \alpha_j$ for each $j = 1, \ldots, m$, implying that each α_j is a smooth function on U. **QED**

Characterizing the smoothness of vector fields in terms of local coordinates is computationally useful in certain circumstances, but it is clearly not the most efficient or concise means of doing so. The characterization involving smooth functions, while not coordinate dependent, still relies on an external set of objects to verify smoothness. Moreover, we must show that a vector field, X, induces a smooth function, Xf, for *every* $f \in C^\infty(M)$, which may be difficult to do analytically. After we construct the tangent bundle, we will give more elegant restatements of the previous definitions, showing that the calculus of vector fields in the tangent bundle is very efficient.

4.2 The Tangent Bundle

As a set, the tangent bundle is just the union of all the tangent spaces over a manifold, M. Formally, we have the following.

Definition 4.3 *The* **tangent bundle** *of M, denoted TM, is defined by*

$$TM = \bigcup_{p \in M} T_pM = \{X \,:\, X \in T_pM \;\; for \;\; some \;\; p \in M\}.$$

We will define a topology on TM, making it a second countable Hausdorff space. Even more surprisingly, we will define a coordinate covering for TM, making it a $2m$-dimensional manifold. This dimension makes sense intuitively. We will usually denote an element of the tangent bundle by (p, X), where $X \in T_pM$. This will save us from continually having to refer to the tangent space in which a tangent vector lies. Thus, given this representation, we need $2m$ coordinates to distinguish an element of TM, m for the point, p, and m for the basis coordinates of $X \in T_pM$.

Let $\tau : TM \to M$ be the surjective projection mapping such that $\tau(p, X) = p$. So, if U is any subset of M, $\tau^{-1}(U) = \cup_{p\in U} T_pM$. That is, $\tau^{-1}(U)$ is the union of all the tangent spaces T_pM for $p \in U$.

For each chart, (φ, U) in the coordinate covering of M, we can think of the mapping φ as a smooth mapping between the manifolds U and $\varphi(U) \subset \mathbb{R}^m$. By the way we have defined our smooth manifolds, this map is actually a diffeomorphism. Hence, the differential $D_p\varphi : T_pM \to \mathbb{R}^m$, where $D_p\varphi$ indicates that this is the differential of φ at p, is an isomorphism. Likewise, the mapping $\varphi^{-1} : \varphi(U) \to U$ is a diffeomorphism, and, for $x \in \varphi(U)$, the differential $D_x\varphi^{-1} : \mathbb{R}^m \to T_{\varphi^{-1}(x)}M$ is an isomorphism. We will use these facts in the construction of both the topology and the coordinate covering on TM.

Define a mapping $T\varphi : \tau^{-1}(U) \to \varphi(U) \times \mathbb{R}^m$ by

$$T\varphi(p, X) = \big(\varphi(p), D_p\varphi(X)\big).$$

If $T\varphi(p, X_p) = T\varphi(q, X_q)$, then $(\varphi(p), D_p\varphi(X_p)) = (\varphi(q), D_q\varphi(X_q)) \Rightarrow \varphi(p) = \varphi(q)$ and $D_p\varphi(X_p) = D_q\varphi(X_q)$. But φ is injective, so this implies $p = q$. This, in turn, implies that $D_p\varphi = D_q\varphi$, so $(p, X_p) = (q, X_q)$, and $T\varphi$ is injective. If $(a, b) \in \varphi(U) \times \mathbb{R}^m$, then $T\varphi(\varphi^{-1}(a), D_a\varphi^{-1}(b)) = (\varphi(\varphi^{-1}(a)), D_{\varphi^{-1}(a)}\varphi(D_a\varphi^{-1}(b))) = (a, D_a(\varphi \circ \varphi^{-1})(b)) = (a, b)$. Hence, $T\varphi$ is surjective, and we see that this map is a bijection of $\tau^{-1}(U)$ onto $\varphi(U) \times \mathbb{R}^m$.

Now, let $\mathbb{T} = \{W \subset TM$: for each chart (φ, U) on M, $T\varphi(W \cap \tau^{-1}(U))$ is open in $\mathbb{R}^m \times \mathbb{R}^m\}$.

Theorem 4.2 *$\mathbb{T}$ is a topology on TM.*

<u>Proof</u> The empty set, $\emptyset$, is in $\mathbb{T}$, because, for any chart (φ, U), we have $\emptyset \cap \tau^{-1}(U) = \emptyset$ and $T\varphi(\emptyset) = \emptyset$. TM is in $\mathbb{T}$ because $TM \cap \tau^{-1}(U) = \tau^{-1}(U) \Rightarrow T\varphi(\tau^{-1}(U)) = \varphi(U) \times \mathbb{R}^m$, which is open in $\mathbb{R}^m \times \mathbb{R}^m$. Now, let $\{W_\alpha\}_{\alpha \in A}$ be an arbitrary collection of sets in $\mathbb{T}$ indexed by some set A. Then, for any chart, (φ, U), we have

$$\begin{aligned} T\varphi\Big(\big(\bigcup_\alpha W_\alpha\big) \cap \tau^{-1}(U)\Big) &= T\varphi\Big(\bigcup_\alpha \big[W_\alpha \cap \tau^{-1}(U)\big]\Big) \\ &= \bigcup_\alpha T\varphi\Big(W_\alpha \cap \tau^{-1}(U)\Big) \\ &\in \mathbb{T}. \end{aligned}$$

Finally, if $\{W_i\}_{i=1}^n$ is a finite collection of sets in $\mathbb{T}$, then

$$\begin{aligned} T\varphi\Big(\bigcap_{i=1}^n W_i \cap \tau^{-1}(U)\Big) &= \bigcap_{i=1}^n T\varphi\Big(W_i \cap \tau^{-1}(U)\Big) \\ &\in \mathbb{T}. \end{aligned}$$

So, $\mathbb{T}$ is a topology on TM.

QED

Defining the coordinate covering on TM is fairly straightforward. The following theorem does most of the technical work.

Theorem 4.3 *If (φ, U) is a coordinate chart on M, then $\tau^{-1}(U)$ is open in TM, $T\varphi : \tau^{-1}(U) \to \varphi(U) \times \mathbb{R}^m$ is a homeomorphism, and, for any other chart (ψ, V) on M, the map $T\varphi \circ (T\psi)^{-1} : T\psi(\tau^{-1}(U \cap V)) \to T\varphi(\tau^{-1}(U \cap V))$ is C^∞.*

Proof To show that $\tau^{-1}(U)$ is open, we need to show that for any chart, (ψ, V), the set $T\psi(\tau^{-1}(U) \cap \tau^{-1}(V))$ is open. But

$$T\psi\big(\tau^{-1}(U) \cap \tau^{-1}(V)\big) = T\psi\big(\tau^{-1}(U \cap V)\big) = \psi(U \cap V) \times \mathbb{R}^m,$$

which is open in $\mathbb{R}^m \times \mathbb{R}^m$. So, $\tau^{-1}(U)$ is open for any chart (φ, U).

Next, we show that for any two charts, (φ, U) and (ψ, V), the map $T\varphi \circ (T\psi)^{-1}$ is C^∞. Note that the inverse of the map, $T\psi$, is the map $(T\psi)^{-1} : \psi(V) \times \mathbb{R}^m \to \tau^{-1}(V)$ defined explicitly by

$$\big(T\psi\big)^{-1}(a^1 \ldots, a^m, b^1, \ldots, b^m) = \big(\psi^{-1}(a), D_a\psi^{-1}(b)\big),$$

for $(a, b) = (a^1, \ldots, a^m, b^1, \ldots, b^m) \in \psi(V) \times \mathbb{R}^m$. So, the map $T\varphi \circ (T\psi)^{-1}$ will map $T\psi(\tau^{-1}(U \cap V))$ to $T\varphi(\tau^{-1}(U \cap V))$. Let us denote the coordinate functions of φ and ψ by x^i and y^j, respectively. Then, for $(a, b) \in T\psi(\tau^{-1}(U \cap V))$, we have

$$\begin{aligned} T\varphi \circ \big(T\psi\big)^{-1}(a, b) &= T\varphi\Big(\psi^{-1}(a), D_a\psi^{-1}(b)\Big) \\ &= \Big((\varphi \circ \psi^{-1})(a), D_{\psi^{-1}(a)}\varphi\big(D_a\psi^{-1}(b)\big)\Big) \\ &= \Big((\varphi \circ \psi^{-1})(a), D_a(\varphi \circ \psi^{-1})(b)\Big) \\ &= \left((\varphi \circ \psi^{-1})(a), \sum_{i=1}^m \frac{\partial x^1}{\partial y^i} b^i, \sum_{i=1}^m \frac{\partial x^2}{\partial y^i} b^i, \ldots, \sum_{i=1}^m \frac{\partial x^m}{\partial y^i} b^i, \right) \end{aligned}$$

where the last equality follows from the fact that $D_a(\varphi \circ \psi^{-1})(b)$ is just the differential of the change of coordinates map, $\varphi \circ \psi^{-1}$. It is implied that these partial derivatives are to be evaluated at $a \in \psi(U \cap V)$. Now, the map $\varphi \circ \psi^{-1}$ is C^∞ because M is a smooth manifold. By this same fact, the partial derivatives $\frac{\partial x^i}{\partial y^j}$ are smooth functions. Hence, each component function of this map is smooth, and we can conclude that $T\varphi \circ (T\psi)^{-1}$ is smooth. It follows by almost the same argument that $T\psi \circ (T\varphi)^{-1}$ is C^∞ also. Hence, since these charts were arbitrary, it follows that all possible compositions of the form $T\varphi \circ (T\psi)^{-1}$ are smooth. In fact, since each map, $T\varphi$, is bijective, this shows that the maps $T\varphi \circ (T\psi)^{-1}$ are diffeomorphisms.

Finally, we will show that $T\varphi : \tau^{-1}(U) \to \varphi(U) \times \mathbb{R}^m$ is a homeomorphism for any chart (φ, U). Suppose W is open in $\tau^{-1}(U)$. Since $\tau^{-1}(U)$ is open in TM, it follows that W is also open in TM, for there is some open set, $\Omega \subset TM$, such that $W = \Omega \cap \tau^{-1}(U)$. So, for any chart, (ψ, V), the set $T\psi(W \cap \tau^{-1}(V))$ is open in $\mathbb{R}^m \times \mathbb{R}^m$. Hence, $T\varphi(W \cap \tau^{-1}(U)) = T\varphi(W)$ is open in $\varphi(U) \times \mathbb{R}^m$. So, $T\varphi$ maps open sets to open sets.

Conversely, suppose $Z \subset \varphi(U) \times \mathbb{R}^m$ is open. We must show that $T\psi(\tau^{-1}(V) \cap (T\varphi)^{-1}(Z))$ is open for any chart (ψ, V). We first note that

$$\begin{aligned} T\varphi\Big(\tau^{-1}(V) \cap \big(T\varphi\big)^{-1}(Z)\Big) &= Z \cap \Big(T\varphi\big(\tau^{-1}(V)\big)\Big) \\ &= Z \cap \Big(T\varphi\big(\tau^{-1}(V) \cap \tau^{-1}(U)\big)\Big), \end{aligned}$$

where this last line follows from the fact that $T\varphi$ is only defined on $\tau^{-1}(U)$. So, the set $T\varphi(\tau^{-1}(V))$ is empty unless $\tau^{-1}(V)$ intersects $\tau^{-1}(U)$, and if it does intersect $\tau^{-1}(U)$, then we are only concerned with the intersection. Hence, we have

$$\begin{aligned} T\varphi\Big(\tau^{-1}(V) \cap \big(T\varphi\big)^{-1}(Z)\Big) &= Z \cap \Big(T\varphi\big(\tau^{-1}(U \cap V)\big)\Big) \\ &= Z \cap \Big(\varphi(U \cap V) \times \mathbb{R}^m\Big), \end{aligned}$$

which implies that

$$\begin{aligned} & \tau^{-1}(V) \cap \big(T\varphi\big)^{-1}(Z) = \big(T\varphi\big)^{-1}\Big(Z \cap \big(\varphi(U \cap V) \times \mathbb{R}^m\big)\Big) \\ \Rightarrow \quad & T\psi\Big(\tau^{-1}(V) \cap \big(T\varphi\big)^{-1}(Z)\Big) = T\psi \circ \big(T\varphi\big)^{-1}\Big(Z \cap \big(\varphi(U \cap V) \times \mathbb{R}^m\big)\Big). \end{aligned}$$

Since $T\psi \circ (T\varphi)^{-1}$ is a diffeomorphism, the right side of this last equality is an open set. Hence, the set $T\psi(\tau^{-1}(V) \cap (T\varphi)^{-1}(Z))$ is open, and $T\varphi$ is a homeomorphism.

QED

Corollary 4.4 $\tau : TM \to M$ *is continuous.*

Proof Let $\Omega \subset M$ be open, and let (φ, U) be any chart on M. We have $T\varphi(\tau^{-1}(\Omega) \cap \tau^{-1}(U)) = T\varphi(\tau^{-1}(\Omega \cap U))$. If $\Omega \cap U = \emptyset$, then $T\varphi(\tau^{-1}(\Omega \cap U)) = \emptyset$, which is open. If $\Omega \cap U \neq \emptyset$, then $\Omega \cap U$ is an open subset of U, so $T\varphi(\tau^{-1}(\Omega \cap U)) = \varphi(\Omega \cap U) \times \mathbb{R}^m$, which is open.

QED

We can now prove that TM is a smooth manifold. We have already done the necessary work in constructing the coordinate covering. The most difficult part about proving that TM is a manifold is proving that it has the correct topological structure. This will be the bulk of the following theorem.

Theorem 4.5 *The tangent bundle, TM, is a smooth manifold of dimension $2m$ with coordinate covering $A = \{(T\varphi, \tau^{-1}(U)) : (\varphi, U)$ is a chart on $M\}$.*

Proof For any $(p, X) \in TM$, we must have $p \in U$ for some chart (φ, U), so $(p, X) \in \tau^{-1}(U)$. Moreover, we have shown that each $T\varphi : \tau^{-1}(U) \to \varphi(U) \times \mathbb{R}^m$ is a homeomorphism. So, A is a coordinate covering of TM. We have also shown that, for any two charts, (φ, U) and (ψ, V), the composition $T\varphi \circ (T\psi)^{-1}$ is C^∞. Hence, TM will be a smooth manifold if we can show that it is both Hausdorff and second countable. We will prove the Hausdorff condition first.

Let (p, X_p) and (q, X_q) be any two distinct points of TM. If $p = q$, then we must have $X_p \neq X_q$, which implies that $D_p\varphi(X_p) \neq D_p\varphi(X_q)$. Since $\mathbb{R}^m$ is Hausdorff, there are open balls, B_1 and B_2, around $D_p\varphi(X_p)$ and $D_p\varphi(X_q)$, respectively, such that $B_1 \cap B_2 = \emptyset$. Then $(T\varphi)^{-1}(\varphi(U) \times B_1)$ and $(T\varphi)^{-1}(\varphi(U) \times B_2)$ are disjoint open sets containing (p, X_p) and (p, X_q) respectively. If $p \neq q$, we will look at two mutually exclusive possibilities.

i) Suppose there exists a single coordinate chart, (φ, U), such that $p, q \in U$. Since M is Hausdorff, we can find open sets Ω_p and Ω_q such that $p \in \Omega_p \subset U$, $q \in \Omega_q \subset U$, and $\Omega_p \cap \Omega_q = \emptyset$. Then $\varphi(\Omega_p)$ and $\varphi(\Omega_q)$ are disjoint open subsets of $\varphi(U)$. It follows that $\varphi(\Omega_p) \times \mathbb{R}^m$ and $\varphi(\Omega_q) \times \mathbb{R}^m$ are disjoint open sets containing $(\varphi(p), D_p\varphi(X_p))$ and $(\varphi(q), D_q\varphi(X_q))$ respectively. Since $T\varphi$ is a homeomorphism, it follows that $(T\varphi)^{-1}(\varphi(\Omega_p) \times \mathbb{R}^m)$ and $(T\varphi)^{-1}(\varphi(\Omega_q) \times \mathbb{R}^m)$ are disjoint open sets in TM such that $(p, X_p) \in (T\varphi)^{-1}(\varphi(\Omega_p) \times \mathbb{R}^m)$ and $(q, X_q) \in (T\varphi)^{-1}(\varphi(\Omega_q) \times \mathbb{R}^m)$.

ii) Now, suppose there is no single coordinate neighborhood containing p and q. Let (φ, U) and (ψ, V) be charts on M such that $p \in U$ and $q \in V$. There are open sets $\Omega_p \subset U$ and $\Omega_q \subset V$ such that $p \in \Omega_p$, $q \in \Omega_q$, and $\Omega_p \cap \Omega_q = \emptyset$. Since τ is continuous, the sets $\tau^{-1}(\Omega_p)$ and $\tau^{-1}(\Omega_q)$ are disjoint open sets in TM containing (p, X_p) and (q, X_q), respectively.

Hence, TM is Hausdorff.

Finally, we will construct a countable basis for TM. Define a collection, S, of subsets of TM as follows. The Euclidean space $\mathbb{R}^m$ is second countable, so there is a countable collection of open balls, $\{B_i\}_{i=1}^{\infty}$, that form a basis for $\mathbb{R}^m$. Moreover, for any chart, (φ, U), in the coordinate covering of M, the set $\varphi(U)$ is an open subspace of $\mathbb{R}^m$, so there is a countable collection of open balls, $\{B_{U_i}\}_{i=1}^{\infty}$, that form a basis for $\varphi(U)$. Hence, all the sets of the form $B_{U_i} \times B_j$, $i, j \geq 1$, form a basis for the subspace $\varphi(U) \times \mathbb{R}^m \subset \mathbb{R}^m \times \mathbb{R}^m$. We define S as

$$S = \{(T\varphi)^{-1}(B_{U_i} \times B_j) \,:\, (\varphi, U) \textit{ is a chart on } M,\ i, j \geq 1\}.$$

That is, S is the set of all images of the basis elements of $\varphi(U) \times \mathbb{R}^m$ under $(T\varphi)^{-1}$ for all charts, (φ, U), on M. The basis, $\{B_{U_i} \times B_j\}_{i,j\geq 1}$, for $\varphi(U) \times \mathbb{R}^m$ is countable. Moreover, the coordinate covering of M is countable. Thus, S is a countable collection of open subsets of TM. If $(p, X) \in TM$, then we must have $p \in U$ for some chart (φ, U). Hence, $(p, X) \in \tau^{-1}(U) = (T\varphi)^{-1}(\varphi(U) \times \mathbb{R}^m)$. There are sets $B_{U_i} \subset \varphi(U)$ and $B_j \subset \mathbb{R}^m$ such that $(\varphi(p), D_p\varphi(X)) \in B_{U_i} \times B_j$, implying that $(p, X) \in (T\varphi)^{-1}(B_{U_i} \times B_j)$. So, S covers TM. Now, let W be any open set in TM, and let (p, X) be any point in W. Let (φ, U) be a chart such that $p \in U$. Then $(p, X) \in W \cap \tau^{-1}(U)$, so this set is nonempty and open. Hence, the set $T\varphi(W \cap \tau^{-1}(U))$ is open in $\varphi(U) \times \mathbb{R}^m$ and contains $T\varphi(p, X) = (\varphi(p), D_p\varphi(X))$. Thus, there are sets $B_{U_i} \subset \varphi(U)$ and $B_j \subset \mathbb{R}^m$ such that $B_{U_i} \times B_j \subset T\varphi(W \cap \tau^{-1}(U))$ and $T\varphi(p, X) = (\varphi(p), D_p\varphi(X)) \in B_{U_i} \times B_j$. It follows that $(p, X) \in (T\varphi)^{-1}(B_{U_j} \times B_j) \subset W \cap \tau^{-1}(U) \subset W$. Since $(p, X) \in W$ was arbitrary, this implies that we can express any open set, $W \subset TM$, as a union of elements from S. Hence, S must be a basis, and we can conlcude that TM is second countable.

QED

Thus, the tangent bundle of a smooth manifold is, itself, a smooth manifold. This is a particularly useful property in the study of dynamics. In classical mechanics, the *configuration space*, Ω, of a dynamical system is the set of all possible positions of a particle or system of particles, and it was usually assumed to be a subset of some Euclidean space, say $\mathbb{R}^m$. The motion of these particles induced velocity fields, or tangent vector fields on the set Ω, where the velocity vectors existed in $\mathbb{R}^m$. The *state space* of the system, then, was defined to be $\Omega \times \mathbb{R}^m$. As mechanical theories progressed, however, and it was seen that configuration spaces were more naturally described as manifolds, the notion of the state space had to change along with it. The problem with this new formulation of the state space, though, was that, for a manifold, M, the state space was not as simple as the set $M \times \mathbb{R}^m$. The possibility of curved manifolds meant that, if we wanted some kind of smooth structure on the state space, we needed a more sophisticated means of describing the collection of all the tangent spaces over a manifold. Thus, the tangent bundle was developed, and, in modern terminology, the state space of a dynamical system is described as $M \times TM$, where M is the manifold representing the configuration space.

It is tempting to simply think that, once a basis for each tangent space is chosen and each tangent space is, consequently, isomorphic to $\mathbb{R}^m$, that the tangent bundle of a manifold can simply be described as $M \times \mathbb{R}^m$. This is not true. Manifolds whose tangent bundles are of this structure are said to have a *trivial bundle structure*. The possibility of a curved manifold presents us with the problem that the tangent bundle cannot be globally described by a single Cartesian product like this. Instead, such products must be pasted together over the manifold to create the tangent bundle.

In purely geometric and topological terms, the tangent bundle is a typical example of a more general class of structures called **vector bundles**. A complete description of vector bundles requires an extensive digression into differential topology, so we will only briefly mention them here. A vector bundle of rank k over a manifold, M, is a $4-tuple$, (E, M, V, τ), where E and M are smooth manifolds, $\tau : E \to M$ is a smooth surjective map, and V is a k-dimensional vector space, with the following properties.

i) For each $p \in M$, $\tau^{-1}(p) = V$. V is called the typical fiber of the vector bundle.

ii) If (φ, U) is a coordinate chart on M, then there is a corresponding map, $\tilde{\varphi}$, such that $\tilde{\varphi} : \tau^{-1}(U) \to U \times V$ is a homeomorphism. For any single point, $p \in U$, $\tilde{\varphi} : \tau^{-1}(p) \to p \times V$ is a vector space isomorphism.

iii) If (φ, U) and (ψ, W) are two charts on M such that $U \cap W \neq \emptyset$, then for any $p \in U \cap W$, the map $G_{UW_p} = \tilde{\psi} \circ \tilde{\varphi}^{-1} : V \to V$ is an automorphism and depends smoothly on p.

So, loosely speaking, a vector bundle is just a collection of vector spaces parametrized by points of some base space (a manifold in our case), such that there is some notion of transferring smoothly from one fiber to another. In the case of the tangent bundle, the typical fiber at p is the tangent space, T_pM, which is m-dimensional. So, we can identify the typical fiber with $\mathbb{R}^m$. The mappings $\tilde{\varphi}$ are just the mappings $T\varphi$, and these mappings induce the smooth structure on TM. Restricted to just a single point, $p \in U$, where (φ, U) is a chart on M, then $T\varphi(p, T_pM) = \varphi(p) \times \mathbb{R}^m$. Since $T\varphi$ is a homeomorphism, and the differential $D_p\varphi$ is linear, this induces an isomorphism between T_pM and $\varphi(p) \times \mathbb{R}^m$, as we have shown.

The tangent bundle is the most widely-used example of a vector bundle. As one might expect, we could, using almost the same arguments as above, construct the **cotangent bundle** as well. We will not do this, as the construction is quite similar to what we have already done. Moreover, the cotangent bundle is used primarily in the discipline of *global analysis*, including the theory of measure and integration on manifolds. This is an important field of study in its own right, but it will not play the role in our study of geometry that the tangent bundle will. Hence, our construction of the tangent bundle will suffice for the remainder of our studies.

Having completed this construction, it will be, as we stated earlier, instructive to revisit our notion of a smooth vector field on a manifold, and reformulate those ideas in a more concise fashion.

4.3 Vector Fields Revisited

Previously, we defined a vector field as an assignment of a tangent vector, $X \in T_pM$, to each $p \in M$. The term 'assignment' is an awkward means of defining a mathematical object. Note, however, that after we have constructed the tangent bundle, a vector field is nothing more than a mapping from the manifold to the tangent bundle. That is, a vector field, X, maps $p \in M$ to an element $X(p) \in T_pM \subset TM$. The smoothness characterization is also more elegant.

Definition 4.4 *A smooth vector field, X, on a smooth manifold, M, is a C^∞ mapping $X : M \to TM$.*

It is not quite obvious that this definition is equivalent to the previous one, so we will prove this.

Theorem 4.6 *Let X be a smooth vector field in the sense of definition 4.2, and let X_p be the tangent vector at p defined by this vector field. Then the mapping from M to TM defined by $X(p) = X_p$ is a smooth mapping. Conversely, if $X : M \to TM$ is a smooth map, then, for any smooth function $f \in C^\infty(M)$, we have $Xf \in C^\infty(M)$.*

Proof Let $X : M \to TM$ be a smooth map. We will actually show that the alternate characterization of smoothness given in Lemma 4.1 holds. Since this is equvialent to our original definition, this will prove the first implication.

Let (φ, U) be any chart on M, and let $\{E_i\}_{i=1}^m$ be the standard coordinate basis field induced by this chart on U. For a particular point, $p \in U$, we will denote the basis of T_pM by $\{E_{p_i}\}$, so that the vector field E_i satisfies $E_i(p) = E_{p_i}$. Now, we can observe the restriction of X to U, which must be smooth. For $p \in U$, we have $X(p) \in T_pM = \text{span}\{E_{p_i}\}$, so there are scalars $\alpha_i, \ldots, \alpha_m$ such that $X(p) = \sum_i \alpha_i E_{p_i}$.

Letting p vary over U, we obtain functions $\alpha_i : U \to \mathbb{R}$ such that the restriction of X to U can be represented as

$$X|_U = \sum_{i=1}^{m} \alpha_i E_i.$$

We need to show that each α_i is a C^∞ function on U. By the definition of a smooth map between manifolds, the smoothness of $X : M \to TM$ implies that the coordinate representation $T\varphi \circ X \circ \varphi^{-1} : \varphi(U) \to T\varphi(\tau^{-1}(U))$ is a smooth mapping. We will denote the φ-coordinate functions by x^i. For a point $x = (x^1, \ldots, x^m) \in \varphi(U)$ with $p = \varphi^{-1}(x)$, we have

$$\begin{aligned} T\varphi \circ X \circ \varphi^{-1}(x) &= T\varphi\big(X\big(\varphi^{-1}(x)\big)\big) \\ &= T\varphi\left(\varphi^{-1}(x), \sum_{i=1}^{m} \alpha_i\big(\varphi^{-1}(x)\big) E_{\varphi^{-1}(x)_i}\right), \end{aligned}$$

where we have simply represented the point $X(\varphi^{-1}(x)) \in TM$ in our usual form. Thus, we have

$$\begin{aligned} T\varphi \circ X \circ \varphi^{-1}(x) &= \left(\varphi(\varphi^{-1}(x)), D_p\varphi\Big(\sum_{i=1}^{m} \alpha_i(p) E_{p_i}\Big)\right) \\ &= \left(x, \sum_{i=1}^{m} \alpha_i(p) D_p\varphi(E_{p_i})\right). \end{aligned}$$

Now, the differential of φ at p maps E_{p_i} to the partial derivative operator $\frac{\partial}{\partial x^i}$, which can be identified with the i^{th} standard basis vector e_i in $\mathbb{R}^m$. Hence,

$$\begin{aligned} T\varphi \circ X \circ \varphi^{-1}(x) &= \Big(x, \sum_{i=1}^{m} \alpha_i(p) e_i\Big) \\ &= \big(x, \alpha_1(p), \alpha_2(p), \ldots, \alpha_m(p)\big) \\ &= \big(x, \alpha_1(\varphi^{-1}(x)), \ldots, \alpha_m(\varphi^{-1}(x))\big). \end{aligned}$$

Hence, we can express this map as $T\varphi \circ X \circ \varphi^{-1} = (\mathbf{id}, \boldsymbol{\alpha_1 \circ \varphi^{-1}}, \ldots, \boldsymbol{\alpha_m \circ \varphi^{-1}})$, where **id** represents the identity function. Since this is a C^∞ function, each of the real-valued functions $\alpha_i \circ \varphi^{-1}$ is C^∞. So, by definition, each of the functions α_i is smooth on U. This proves that the new characterization implies the old one.

Now, suppose X is a smooth vector field in the sense of definition 4.2 and Lemma 4.1. We have the induced mapping $X(p) = X_p$. Let p be any point on M, and let (φ, U) be a coordinate chart on M such that $p \in U$. Then $(T\varphi, \tau^{-1}(U))$ is a coordinate system around $X(p)$. The local representation of X on U is

$$X|_U = \sum_{i=1}^{m} \alpha_i E_i,$$

where each α_i is a smooth function on U. Going through the same process we just completed, we can show that, for $x \in \varphi(U)$,

$$T\varphi \circ X \circ \varphi^{-1}(x) = \big(x, \alpha_1(\varphi^{-1}(x)), \ldots, \alpha_m(\varphi^{-1}(x))\big).$$

Since each α_i is smooth by hypothesis, this is a C^∞ map. Finally, let $(T\psi, \tau^{-1}(V))$ be any coordinate chart containing $(p, X(p)$, and consider the coordinate representation $T\psi \circ X \circ \varphi^{-1}$. We must have $p \in U \cap V$, so $T\psi \circ X \circ \varphi^{-1}$ maps $\varphi(U \cap V)$ to $T\psi(\tau^{-1}(U \cap V))$. Moreover, we also see that

$$T\psi \circ X \circ \varphi^{-1} = T\psi \circ (T\varphi)^{-1} \circ T\varphi \circ X \circ \varphi^{-1},$$

which is C^∞ on $U \cap V$. But the point $p \in M$ was arbitrary, as were the charts (φ, U) and (ψ, V). So, we can conclude that X is a C^∞ mapping from M to TM. This shows that the old definition implies the new one.

QED

Hence, we lose nothing with our new definitions. In fact, we gain a more elegant and concise means of working with vector fields on a manifold. Moreover, this shows that the tangent bundle is the natural setting in which to develop the calculus of vector fields. Particularly in dynamics, since momentum and velocity fields are naturally expressed by vector fields, the natural state space of a dynamical system is given by the product manifold $M \times TM$.

In the next chapter, we will discuss tensors on manifolds. Like vector fields, tensors are also most naturally described within vector bundles over manifolds. While we will not use this approach, we simply want to point out the utility of the general concept of the vector bundle, of which the tangent bundle is just one typical example.

Note to self: Show that the projection map $\pi : TM \to M$ is smooth and open.

5 Tensor Analysis on Manifolds

5.1 Tensors Over a Vector Space

Tensor analysis is a deep field of study in its own right. It is an indispensable tool, however, in modern geometry and physics. In the spirit of Riemann's ideas on the nature of space, geometry, and the foundations of physics, tensors allow us to express both geometrical and physical ideas in a coordinate free manner. The classical notion of a vector in Euclidean space was dependent upon a particular coordinate system. So, results based on these ideas were, in turn, coordinate dependent. To free themselves from this restriction, and to develop laws of geometry and physics that hold regardless of any particular coordinate system, mathematicians needed a class of objects that behaved like vectors but could be referenced and manipulated free of any restrictions imposed by coordinates. Elie Cartan, along with his student, Tullio Levi-Cevita, was the first to address this need with his development of the tensor calculus.

One could fill volumes with results on tensor analysis by itself. Instead, we will focus on results that are particularly useful in geometry and physics, and we will specialize our results to the analysis of tensors on manifolds. We will begin by discussing tensors on an arbitrary finite dimensional vector space, V, pointing out results that will be useful for our purposes. Later on, we will transfer these results to the case where the vector space in question is the tangent space of a manifold.

Definition 5.1 *Let V be a vector space of dimension n, and let V^* denote its dual space. An $(r, s)-$type tensor, σ, over V is a multilinear mapping*

$$\sigma : \underbrace{V^* \times \cdots \times V^*}_{r\ copies} \times \underbrace{V \times \cdots \times V}_{s\ copies} \to \mathbb{R}.$$

As a note on this definition, tensors can, in general, map to any field $\mathbb{F}$. We will restrict ourselves to real-valued tensors, though.

Let σ_1, σ_2, and σ be (r, s)-type tensors over V, and let α be in $\mathbb{R}$. If $v_1^*, \ldots, v_r^*$ are dual vectors in V^*, and if $v_1, \ldots, v_s$ are vectors in V, then we can define pointwise addition and scalar multiplication of tensors by the relations

$$(\sigma_1 + \sigma_2)(v_1^*, \ldots, v_r^*, v_1, \ldots, v_s) = \sigma_1(v_1^*, \ldots, v_r^*, v_1, \ldots, v_s) + \sigma_2(v_1^*, \ldots, v_r^*, v_1, \ldots, v_s)$$

$$(\alpha\sigma)(v_1^*, \ldots, v_r^*, v_1, \ldots, v_s) = \alpha\sigma(v_1^*, \ldots, v_r^*, v_1, \ldots, v_s).$$

In this way, the set of (r, s)-type tensors over V becomes a vector space. We denote this vector space by

$$\underbrace{V \otimes \cdots \otimes V}_{r \text{ copies}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{s \text{ copies}},$$

and we call this space the *tensor product* of the vector spaces $V, \ldots, V, V^*, \ldots, V^*$. For brevity of notation, we will usually denote this vector space with the shorthand notation V_s^r. The number r is called the *contravariant order* of σ, and s is called the *covariant order*. An $(r, 0)$-type tensor is called a *contravariant tensor* of order r, and a $(0, s)$-type tensor is called a *covariant tensor* of order s. As a final remark on notation, observe that the domain of σ is defined to be $V^* \times \cdots \times V^* \times V \times \cdots \times V$, while the vector space containing these tensors is denoted by $V \otimes \cdots \otimes V \otimes V^* \otimes \cdots \otimes V^*$. The purpose for reversing the order of the spaces V and V^* in the vector space notation is to emphasize the vectors that are performing the action, as opposed to those that are being acted upon. For example, the first input variable of σ is a dual vector from V^*. Since σ is a multilinear map, the notation $V \otimes \cdots \otimes V \otimes V^* \otimes \cdots \otimes V^*$ is used to signify that the functional acting on this particular variable is a vector from V. Likewise, the $(r+s)^{th}$ input variable of σ is a vector from V. Hence, the vector space notation emphasizes the fact that the functional acting on this particular variable is a dual vector from V^*.

We can also define the product of two tensors. If σ_1 and σ_2 are (r, s) and (k, l)-type tensors, respectively, then we define their **tensor product**, $\sigma_1 \otimes \sigma_2$, to be the $(r + k, s + l)$-type tensor given by

$$(\sigma_1 \otimes \sigma_2)(v_1^*, \ldots, v_r^*, v_{r+1}^*, \ldots, v_{r+k}^*, v_1, \ldots, v_s, v_{s+1}, \ldots, v_{s+l}) =$$

$$\sigma_1(v_1^*, \ldots, v_r^*, v_1, \ldots, v_s)\sigma_2(v_{r+1}^*, \ldots, v_{r+k}^*, v_{s+1}, \ldots, v_{s+l}).$$

Moreover, the tensor product is associative. The proof of the associativity property is not difficult. It is just a tedious computation, so we will only illustrate associativity in the case $\sigma_1, \sigma_2, \sigma_3 \in V_1^1$. Let v^{*1}, v^{*2}, v^{*3} be dual vectors and let v_1, v_2, v_3 be vectors. Then

$$\begin{aligned}
\sigma_1 \otimes (\sigma_2 \otimes \sigma_3)(v^{*1}, v^{*2}, v^{*3}, v_1, v_2, v_3) &= \sigma_1(v^{*1}, v_1)(\sigma_2 \otimes \sigma_3)(v^{*2}, v^{*3}, v_2, v_3) \\
&= \sigma_1(v^{*1}, v_1)\sigma_2(v^{*2}, v_2)\sigma_3(v^{*3}, v_3) \\
&= (\sigma_1(v^{*1}, v_1)\sigma_2(v^{*2}, v_2))\sigma_3(v^{*3}, v_3) \\
&= (\sigma_1 \otimes \sigma_2)(v^{*1}, v^{*2}, v_1, v_2)\sigma_3(v^{*3}, v_3) \\
&= (\sigma_1 \otimes \sigma_2) \otimes \sigma_3(v^{*1}, v^{*2}, v^{*3}, v_1, v_2, v_3).
\end{aligned}$$

Hence, we can give meaning to expressions like $\sigma_1 \otimes \sigma_2 \otimes \sigma_3$, or any tensor product of a finite number of tensors.

As an interesting consequence of our definitions, note that the set of $(0, 1)$-type tensors is just the dual space of V. That is, $V_1^0 = V^*$. Likewise, since V is finite dimensional and can be identified with its double dual, V^{**}, we see that the set of $(1, 0)$-type tensors is just V itself, since each vector $v \in V$ can be naturally identified with a linear functional on V^*. Thus, $V_0^1 = V$.

Before going on, we will give some examples of common tensors.

Example 5.1 An inner product, $\langle , \rangle$, on a real linear space, V, is a $(0, 2)$-type tensor, or a covariant tensor of order 2, over V, as it is real-valued and linear in each variable. This tensor has the additional properties of symmetry, $\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle$ for $v_1, v_2 \in V$, and positive definiteness, $\langle v, v \rangle \geq 0$ for $v \in V$ and equality holds if and only if $v = 0$. ■

Example 5.2 The determinant function, $det : (\mathbb{R}^n)^n \to \mathbb{R}$, thought of as acting on the cartesian product $(\mathbb{R}^n)^n$, is a $(0, n)$-type tensor, or a covariant tensor of order n, over $\mathbb{R}^n$, since the determinant is linear in each row of a matrix. If we identify $\mathbb{R}^n$ with row vectors and the dual space $\mathbb{R}^{n*}$ with column vectors, we can also think of the determinant function as an $(n, 0)$-type tensor, since it is linear in each column as well. ■

Example 5.3 For a vector space, V, let $T : V \to V$ be a linear transformation. Define a $(1, 1)$-type tensor, σ, as follows. For $v \in V, w^* \in V^*$, let $\sigma(w^*, v) = w^*(T(v))$. This a $(1, 1)$ tensor over V. Note that this is nothing more than a bilinear form on V, determined by the linear transformation, T. If $V = \mathbb{R}^n$, then, assuming we have chosen a basis, we can represent T by a matrix, A, and we can identify V with V^* (this identification is only possible after a basis has been chosen). Hence, we simply think of w^* as its corresponding element $w \in V$, so we have $\sigma(w^*, v) = w^T A v$. ■

The following theorem tells us that we can use a basis of V and its dual basis to construct a basis for V_s^r.

Theorem 5.1 *Let V be an m-dimensional vector space. Let $\{e_1, \ldots, e_m\}$ be a basis for V, and let $\{e^{*1}, \ldots, e^{*m}\}$ be its corresponding dual basis, so that $e^{*i}(e_j) = \delta^i_j$. Then the set*

$$e_{i_1} \otimes \cdots e_{i_r} \otimes e^{*j_1} \otimes \cdots \otimes e^{*j_s}, \qquad 1 \leq i_1, \ldots, i_r, j_1, \ldots, j_s \leq m$$

forms a basis for V^r_s. In other words, the m^{r+s} possible tensor products, consisting of all possible permuations (with repititions allowed) of size r from the set $\{e_1, \ldots, e_m\}$ and all possible permutations (with repeititions allowed) of size s from the set $\{e^{*1}, \ldots, e^{*m}\}$, form a basis for the set V^r_s.

Proof Suppose there are scalars, $c^{i_1 \ldots i_r}_{j_1 \ldots j_s}$, such that

$$\sum_{\substack{i_1, \ldots, i_r, \\ j_1, \ldots, j_s = 1}}^{m} c^{i_1 \ldots i_r}_{j_1 \ldots j_s} e_{i_1} \otimes \cdots e_{i_r} \otimes e^{*j_1} \otimes \cdots \otimes e^{*j_s} = 0_T,$$

where 0_T indicates the zero element of the space V^r_s. Then, for fixed indices $1 \leq k_1, \ldots, k_s, l_1, \ldots, l_r \leq m$, we can evaluate this tensor at $(e^{*l_1}, \ldots, e^{*l_r}, e_{k_1}, \ldots, e_{k_s})$ to obtain

$$\sum_{\substack{i_1, \ldots, i_r, \\ j_1, \ldots, j_s = 1}}^{m} c^{i_1 \ldots i_r}_{j_1 \ldots j_s} e_{i_1} \otimes \cdots e_{i_r} \otimes e^{*j_1} \otimes \cdots \otimes e^{*j_s} (e^{*l_1}, \ldots, e^{*l_r}, e_{k_1}, \ldots, e_{k_s}) = 0.$$

But this implies

$$\begin{aligned} 0 &= \sum_{\substack{i_1, \ldots, i_r, \\ j_1, \ldots, j_s = 1}}^{m} c^{i_1 \ldots i_r}_{j_1 \ldots j_s} e_{i_1}(e^{*l_1}) \cdots e_{i_r}(e^{*l_r}) e^{*j_1}(e_{k_1}) \cdots e^{*j_s}(e_{k_s}) \\ &= \sum_{\substack{i_1, \ldots, i_r, \\ j_1, \ldots, j_s = 1}}^{m} c^{i_1 \ldots i_r}_{j_1 \ldots j_s} \delta^{l_1}_{i_1} \cdots \delta^{l_r}_{i_r} \delta^{k_1}_{j_1} \cdots \delta^{k_s}_{j_s} \\ &= c^{l_1 \ldots l_r}_{k_1 \ldots k_s}. \end{aligned}$$

Thus, evaluating this tensor at each of the elements, $e_{i_1} \otimes \cdots e_{i_r} \otimes e^{*j_1} \otimes \cdots \otimes e^{*j_s}$, in turn, shows that each of the scalars must be zero. So, these tensors are linearly independent.

Now, let σ be an arbitrary (r, s)-type tensor. Let $v^{*1}, \ldots, v^{*r}$ be r dual vectors, and let $v_1, \ldots, v_s$ be s vectors. Each dual vector, v^{*i} can be expressed in terms of the dual basis as

$$v^{*i} = \sum_{k=1}^{m} c_k^i e^{*k}$$

for some scalars c_k^i. Likewise, each vector v_i can be expressed in the form

$$v_i = \sum_{l=1}^{m} c_i^l e_l$$

for scalars c_i^l. Now, we just use the multilinearity of σ to obtain

$$\begin{aligned} \sigma(v^{*1}, \ldots, v^{*r}, v_1, \ldots, v_s) &= \sigma\left(\sum_{k_1=1}^{m} c_{k_1}^1 e^{*k_1}, \ldots, \sum_{k_r=1}^{m} c_{k_r}^r e^{*k_r}, \sum_{l_1=1}^{m} c_1^{l_1} e_{l_1}, \ldots, \sum_{l_s=1}^{m} c_s^{l_s} e_{l_s}\right) \\ &= \sum_{\substack{k_1,\ldots,k_r \\ l_1,\ldots,l_s=1}}^{m} c_{k_1}^1 \cdots c_{k_r}^r c_1^{l_1} \cdots c_s^{l_s} \sigma\big(e^{*k_1}, \ldots, e^{*k_r}, e_{l_1}, \ldots, e_{l_s}\big). \end{aligned}$$

But, for $i = 1, \ldots, r$, the scalar $c_{k_n}^i$, for $1 \leq k_n \leq m$, is given by

$$\begin{aligned} e_{k_n}(v^{*i}) &= e_{k_n}\Big(\sum_{k_i=1}^{m} c_{k_i}^i e^{*k_i}\Big) \\ &= \sum_{k_i=1}^{m} c_{k_i}^i e_{k_n}(e^{*k_i}) \\ &= \sum_{k_i=1}^{m} c_{k_i}^i \delta_{k_n}^{k_i} \\ &= c_{k_n}^i. \end{aligned}$$

Similarly, for $i = 1, \ldots, s$, the scalar $c_i^{l_n}$, for $1 \leq l_n \leq m$, is given by $e^{*l_n}(v_i)$. Thus, we have

$$\sigma(v^{*1}, \ldots, v^{*r}, v_1, \ldots, v_s)$$

$$
\begin{aligned}
&= \sum_{\substack{k_1,\ldots,k_r\\ l_1,\ldots,l_s=1}}^{m} \sigma\big(e^{*k_1},\ldots,e^{*k_r},e_{l_1},\ldots,e_{l_s}\big)e_{k_1}(v^{*1})\cdots e_{k_r}(v^{*r})e^{*l_1}(v_1)\cdots e^{*l_s}(v_s)\\
&= \sum_{\substack{k_1,\ldots,k_r\\ l_1,\ldots,l_s=1}}^{m} \sigma\big(e^{*k_1},...,e^{*k_r},e_{l_1},...,e_{l_s}\big)e_{k_1}\otimes\cdots\otimes e_{k_r}\otimes e^{*l_1}\otimes\cdots\otimes e^{*l_s}(v^{*1},...,v^{*r},v_1,...,v_s).
\end{aligned}
$$

Hence, these elements span V_s^r. Moreover, this shows that the dimension of V_s^r is m^{r+s}, and the components of a tensor, σ, with respect to this basis are the values obtained by σ acting on the basis elements.

QED

The last result we will need from general tensor analysis concerns the formula describing how the components of a tensor change when the basis of the underlying space, V, changes.

Theorem 5.2 *Let σ be an (r,s)-type tensor over an m-dimensional vector space, V. Let $\{e_i\}_{i=1}^m$ and $\{f_i\}_{i=1}^m$ be two bases for V, and let $\{e^{*i}\}$ and $\{f^{*i}\}$ be their corresponding dual bases of V^*. Suppose σ has representations*

$$
\begin{aligned}
\sigma &= \sum_{\substack{i_1,\ldots,i_r\\ j_1,\ldots,j_s=1}}^{m} c^{i_1\cdots i_r}_{j_1\cdots j_s} e_{i_1}\otimes\cdots\otimes e_{i_r}\otimes e^{*j_1}\otimes\cdots\otimes e^{*j_s}\\
\sigma &= \sum_{\substack{k_1,\ldots,k_r\\ l_1,\ldots,l_s=1}}^{m} d^{k_1\cdots k_r}_{l_1\cdots l_s} f_{k_1}\otimes\cdots\otimes f_{k_r}\otimes f^{*l_1}\otimes\cdots\otimes f^{*l_s}
\end{aligned}
$$

with respect the two bases induced on V_s^r. Further, suppose that we have, for each $i = 1,\ldots,m$,

$$
\begin{aligned}
e_i &= \sum_{k=1}^{m} \alpha_i^k f_k\\
e^{*i} &= \sum_{k=1}^{m} \beta_k^i f^{*k}
\end{aligned}
$$

for scalars $\{\alpha_i^k\}$ *and* $\{\beta_k^i\}$. *Then the components of* σ *transform according to the formula*

$$c_{j_1\cdots j_s}^{i_1\cdots i_r} = \sum_{\substack{k_1,\ldots,k_r\\ l_1,\ldots,l_s=1}}^{m} \beta_{k_1}^{i_1}\cdots\beta_{k_r}^{i_r}\alpha_{j_1}^{l_1}\cdots\alpha_{j_s}^{l_s} d_{l_1\cdots l_s}^{k_1\cdots k_r}.$$

Proof   The component $c_{j_1\cdots j_s}^{i_1\cdots i_r}$ is given by $\sigma(e^{*i_1},\ldots,e^{*i_r},e_{j_1},\ldots,e_{j_s})$, and the component $d_{l_1\cdots l_s}^{k_1\cdots k_r}$ is given by $\sigma(f^{*k_1},\ldots,f^{*k_r},f_{l_1},\ldots,f_{l_s})$. Thus, we see that

$$\begin{aligned} c_{j_1\ldots j_s}^{i_1\cdots i_r} &= \sigma\left(\sum_{k_1=1}^{m}\beta_{k_1}^{i_1}f^{*k_1},\ldots,\sum_{k_r=1}^{m}\beta_{k_r}^{i_r}f^{*k_r},\sum_{l_1=1}^{m}\alpha_{j_l}^{l_1}f_{l_1},\ldots,\sum_{l_s=1}^{m}\alpha_{j_s}^{l_s}f_{l_s}\right)\\ &= \sum_{\substack{k_1,\ldots,k_r\\ l_1,\ldots,l_s=1}}^{m}\beta_{k_1}^{i_1}\cdots\beta_{k_r}^{i_r}\alpha_{j_1}^{l_1}\cdots\alpha_{j_s}^{l_s}\sigma(f^{*k_1},\ldots,f^{*k_r},f_{l_1},\ldots,f_{l_s}),\end{aligned}$$

which implies that

$$c_{j_1\cdots j_s}^{i_r\cdots i_r} = \sum_{\substack{k_1,\ldots,k_r\\ l_1,\ldots,l_s=1}}^{m} \beta_{k_1}^{i_1}\cdots\beta_{k_r}^{i_r}\alpha_{j_1}^{l_1}\cdots\alpha_{j_s}^{l_s} d_{l_1\cdots l_s}^{k_1\cdots k_r}.$$

QED

Now let us consider what these results looks like when the underlying vector space is the tangent space, T_pM, at $p \in M$. Let (φ, U) be a coordinate system around p, and let $\{E_i\}$ and $\{dx^i\}$ denote the bases of T_pM and T_p^*M, respectively. Then, these bases induce a basis for $(T_pM)_s^r$ given by the set

$$\{E_{i_1}\otimes\cdots\otimes E_{i_r}\otimes dx^{j_1}\otimes\cdots\otimes dx^{j_s} \,:\, 1\le i_1,\ldots,i_r,j_1,\ldots,j_s\le m\}.$$

Thus, an (r,s)-type tensor over $(T_pM)_s^r$ has representation

$$\sigma = \sum_{\substack{i_1,\ldots,i_r\\ j_1,\ldots,j_s=1}}^{m} c_{j_1\ldots j_s}^{i_1\ldots i_r} E_{i_1}\otimes\cdots\otimes E_{i_r}\otimes dx^{j_1}\otimes\cdots\otimes dx^{j_s},$$

where $c_{j_1\cdots j_s}^{i_1\cdots i_r} = \sigma(dx^{i_1},\ldots,dx^{i_r},dx^{j_1},\ldots,dx^{j_s})$.

Even more interesting is the change of basis formula in this case. Let (φ, U) and (ψ, V) be two coordinate systems such that $p \in U \cap V$, and let us denote the coordinate functions of φ and ψ by $\{x^i\}$ and $\{y^i\}$, respectively. Let $\{E_i^x\}$ and $\{dx^i\}$ be the bases for T_pM and T_p^*M induced by φ, with $\{E_i^y\}$ and $\{dy^i\}$ the corresponding bases induced by ψ. Further, suppose, for each $i = 1, \ldots, m$,

$$\begin{aligned} E_i^x &= \sum_{k=1}^{m} \alpha_i^k E_k^y \\ dx^i &= \sum_{k=1}^{m} \beta_k^i dy^k \end{aligned}$$

for scalars $\{\alpha_i^k\}$ and $\{\beta_k^i\}$. Then, if

$$\begin{aligned} \sigma &= \sum_{\substack{i_1,\ldots,i_r \\ j_1,\ldots,j_s=1}}^{m} a_{j_1\cdots j_s}^{i_1\cdots i_r} E_{i_1}^x \otimes \cdots \otimes E_{i_r}^x \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s} \\ \sigma &= \sum_{\substack{k_1,\ldots,k_r \\ l_1,\ldots,l_s=1}}^{m} b_{l_1\cdots l_s}^{k_1\cdots k_r} E_{k_1}^y \otimes \cdots \otimes E_{k_r}^y \otimes dx^{l_1} \otimes \cdots \otimes dx^{l_s} \end{aligned}$$

are the representations of σ with respect to the coordinate systems (φ, U) and (ψ, V), respectively, we have

$$a_{j_1\cdots j_s}^{i_r\cdots i_r} = \sum_{\substack{k_1,\ldots,k_r \\ l_1,\ldots,l_s=1}}^{m} \beta_{k_1}^{i_1} \cdots \beta_{k_r}^{i_r} \alpha_{j_1}^{l_1} \cdots \alpha_{j_s}^{l_s} b_{l_1\cdots l_s}^{k_1\cdots k_r}.$$

But, for $1 \le n \le r$ and any $1 \le i_n, k_n \le m$, the scalar $\beta_{k_n}^{i_n}$ is given by $E_{k_n}^y(dx^{i_n})$. Likewise, for $1 \le n \le s$ and any $1 \le j_n, l_n \le m$, the scalar $\alpha_{j_n}^{l_n}$ is given by $dy^{l_n}(E_{j_n}^x)$. Thus, the previous equality becomes

$$\begin{aligned} a_{j_1\cdots j_s}^{i_r\cdots i_r} &= \sum_{\substack{k_1,\ldots,k_r \\ l_1,\ldots,l_s=1}}^{m} E_{k_1}^y(dx^{i_1}) \cdots E_{k_r}^y(dx^{i_r}) dy^{l_1}(E_{j_1}^x) \cdots dy^{l_s}(E_{j_s}^x) b_{l_1\cdots l_s}^{k_1\cdots k_r} \\ &= \sum_{\substack{k_1,\ldots,k_r \\ l_1,\ldots,l_s=1}}^{m} \frac{\partial x^{i_1}}{\partial y^{k_1}}\Big|_{\psi(p)} \cdots \frac{\partial x^{i_r}}{\partial y^{k_r}}\Big|_{\psi(p)} \frac{\partial y^{l_1}}{\partial x^{j_1}}\Big|_{\varphi(p)} \cdots \frac{\partial y^{l_s}}{\partial x^{j_s}}\Big|_{\varphi(p)} b_{l_1\cdots l_s}^{k_1\cdots k_r}, \end{aligned}$$

where we have used the fact that the action of a cotangent vector, dy^{l_n}, on a tangent vector, $E^x_{j_n}$, is given by the pairing $\langle E^x_{j_n}, dy^{l_n}\rangle = E^x_{j_n}(dy^{l_n})$. In classical tensor analysis, this relation was used to define tensors. The definition was later refined, however, when it was discovered that tensors whose transformation formula was defined by the partial derivatives of the coordinate functions were simply special cases of a larger class of multilinear mappings.

5.2 Tensor Fields On a Manifold

Just as we first defined tangent vectors at a point p, and then extended this idea to the notion of a vector field on M, we will extend our idea of a tensor over T_pM to a field of tensors over M. At each point, $p \in M$, this tensor field will induce an (r,s)-type tensor over T_pM, a real-valued mapping on r cotangent vectors and s tangent vectors at p. As p varies over M, we have field of tensors acting on r fields of cotangent vectors and s fields of tangent vectors. We already have a well-defined notion of a field of tangent vectors, or a vector field. To complete our definition of a tensor field, we need the definition of a *cotangent vector field*. We will not dwell on this partiuclar topic very much, for we will not use it again after this. We merely need it to define a smooth tensor field.

Definition 5.2 *A* **smooth cotangent vector field***, or* **covector field***, ω, is an assignment of a cotangent vector, df_p, to each point $p \in M$, such that, for any smooth vector field, X, on M, the function $\omega(X) : M \to \mathbb{R}$ defined by $\omega(X)(p) = df_p(X_p) = X_p(f)$, is C^∞ on M.*

For notational purposes, we will denote a covector field by ω, and its cotangent vector at a particular point, p, by $\omega_p = df_p$. Note that, on a given coordinate system, (φ, U), a covector field, ω, can be represented in terms of the coordinate frame induced for the cotangent spaces at points $p \in U$. This coordinate frame is denoted $\{dx^i\}_{i=1}^m$, where x^i is the i^{th} coordinate function of the mapping φ. That is, for any $i = 1, \ldots, m$, dx^i is the covector field on U such that, for any $p \in U$, we denote $dx^i(p)$ by dx^i_p. Hence, there are real-valued functions, β_i, on U such that $\omega|_U = \sum_i \beta_i dx^i$.

As with the tangent vectors, it would be more elegant to construct the cotangent bundle, T^*M, and define a covector field as a mapping $\omega : M \to T^*M$. We will not construct the cotangent bundle here, though, as it would take us too far afield from our stated purpose. We will, however, use, without proof, a result for covector fields analogous to Lemma 4.1. As expected from that result, it simply states that a covector field, ω, is smooth if and only if its component functions on every coordinate neighborhood are smooth. The proof of this result is almost exactly like that of Lemma 4.1, so we do not lose anything significant by omitting the proof here.

Definition 5.3 *A smooth, (r, s)-type* **tensor field***, σ, on a smooth manifold, M, is an assignment of an (r, s)-type tensor, σ_p, to each point $p \in M$, such that for any smooth covector fields $\omega_1, \ldots, \omega_r$ and any smooth vector fields $X_1, \ldots, X_s$, the function $\sigma(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s) : M \to \mathbb{R}$ defined by $\sigma(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s)(p) = \sigma_p(\omega_{1_p}, \ldots, \omega_{r_p}, X_{1_p}, \ldots, X_{s_p})$ is C^∞.*

This characterization of smoothness parallels that given in definition 4.1 for vector fields.

Now, let (φ, U) be a coordinate system on M. Let $\{E_i\}$ and $\{dx^i\}$ denote the coordinate frames for T_pM and T_p^*M, respectively, that are induced by this coordinate chart. Thus, notationally speaking, for any $i = 1, \ldots, m$ and any $q \in U$, the vector field, E_i, and covector field, dx^i, satisfy $E_i(q) = E_{q_i}$ and $dx^i(q) = dx^i_q$. Suppose that with respect to these coordinate frames, the covector fields ω_i have representation

$$\omega_i = \sum_{k=1}^{m} \beta^i_k dx^k \tag{4}$$

and the vector fields X_i have representation

$$X_j = \sum_{l=1}^{m} \alpha^l_j E_l. \tag{5}$$

Then, for all $1 \le i \le r$, $1 \le j \le s$, and $1 \le k \le m$, the functions β^i_k and α^k_j are smooth functions on U. Moreover, on the coordinate neighborhood, U, the expression $\sigma(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s)$, for an (r, s)-type tensor field, σ, takes the form

$$\sigma\left(\sum_{k_1=1}^{m} \beta^1_{k_1} dx^{k_1}, \ldots, \sum_{k_r=1}^{m} \beta^r_{k_r} dx^{k_r}, \sum_{l_1=1}^{m} \alpha^{l_1}_1 E_{l_1}, \ldots, \sum_{l_s=1}^{m} \alpha^{l_s}_s E_{l_s}\right),$$

which yields, using the multilinearity of σ,

$\sigma(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s)$

$$= \sum_{\substack{k_1,\dots,k_r \\ l_1,\dots,l_s=1}}^{m} \beta^1_{k_1}\cdots\beta^r_{k_r}\alpha^{l_1}_1\cdots\alpha^{l_s}_s\sigma\big(dx^{k_1},\dots,dx^{k_r},E_{l_1},\dots,E_{l_s}\big)$$

$$= \sum_{\substack{k_1,\dots,k_r \\ l_1,\dots,l_s=1}}^{m} E_{k_1}(\omega_1)\cdots E_{k_r}(\omega_r)dx^{l_1}(X_1)\cdots dx^{l_s}(X_s)\sigma\big(dx^{k_1},\dots,dx^{k_r},E_{l_1},\dots,E_{l_s}\big)$$

$$= \sum_{\substack{k_1,\dots,k_r \\ l_1,\dots,l_s=1}}^{m} \sigma\big(dx^{k_1},...,dx^{k_r},E_{l_1},...,E_{l_s}\big)E_{k_1}\otimes\cdots\otimes E_{k_r}\otimes dx^{l_1}\otimes\cdots\otimes dx^{l_s}(\omega_1,...,\omega_r,X_1,...,X_s).$$

This shows that the (r,s)-type tensor fields $E_{k_1}\otimes\cdots\otimes E_{k_r}\otimes dx^{l_1}\otimes\cdots\otimes dx^{l_s}$ span the set of (r,s)-type tensor fields on the coordinate neighborhood U. Thus, for any (r,s)-type tensor field, σ, on U,

$$\sigma = \sum_{\substack{k_1,\dots,k_r \\ l_1,\dots,l_s=1}}^{m} \sigma\big(dx^{k_1},...,dx^{k_r},E_{l_1},...,E_{l_s}\big)E_{k_1}\otimes\cdots\otimes E_{k_r}\otimes dx^{l_1}\otimes\cdots\otimes dx^{l_s}.$$

Now, if σ is smooth according to definition 5.3, then $\sigma|_U$ is smooth on this coordinate neighborhood. Hence, the local representation given above is smooth. If we evaluate σ at the covector fields $dx^{i_1},\dots,dx^{i_r}$ and the vector fields $E_{j_1},\dots,E_{j_s}$, all of which are smooth, we simply obtain the component of σ given by

$$\sigma(dx^{i_1},\dots,dx^{i_r},E_{j_1},\dots,E_{j_s}),$$

and this must be a C^∞ real-valued function on U, in the sense that the mapping $p\mapsto\sigma_p(dx^{i_1}_p,\dots,dx^{i_r}_p,E_{p_{j_1}},\dots,E_{p_{j_s}})$, for $p\in U$, is smooth. Applying this to each component in turn shows that if σ is smooth according to the definition, then each of the component functions on an arbitrary coordinate neighborhood is also smooth. Conversely, suppose the local coordinate representations of σ on any coordinate neighborhood are smooth. Let $\omega_1,\dots,\omega_r$ and $X_1,\dots,X_s$ be smooth covector and vector fields, respectively, on M. Then, we obtain smooth covector and vector fields by restricting these to U. Suppose the restrictions of these covector and vector fields have local coordinate representations given by relations (4) and (5). It follows that

$\sigma|_U\big(\omega_1|_U,\dots,\omega_r|_U,X_1|_U,\dots,X_s|_U\big)$

$$= \sigma|_U\left(\sum_{k_1=1}^{m}\beta^1_{k_1}dx^{k_1}, ..., \sum_{k_r=1}^{m}\beta^r_{k_r}dx^{k_r}, \sum_{l_1=1}^{m}\alpha^{l_1}_1 E_{l_1}, ..., \sum_{l_s=1}^{m}\alpha^{l_1}_s E_{l_s}\right)$$
$$= \sum_{\substack{k_1,...,k_r\\ l_1,...,l_s=1}}^{m} \beta^1_{k_1}\cdots\beta^r_{k_r}\alpha^{l_1}_1\cdots\alpha^{l_s}_s\sigma|_U(dx^{k_1}, ..., dx^{k_r}, E_{l_1}, ..., E_{l_s}).$$

But the functions $\sigma|_U(dx^{k_1}, \ldots, dx^{k_r}, E_{l_1}, \ldots, E_{l_s})$, for any indices $1 \leq k_1,...,k_r, l_1, ..., l_s \leq m$, are smooth on U by hypothesis, as are the functions $\beta^i_{k_i}$ and $\alpha^{l_j}_j$ for any $1 \leq i \leq r$ and $1 \leq j \leq s$. Hence, the function $\sigma|_U$ is smooth on U. Since the chart (φ, U) was arbitrary, it follows that σ is smooth on a collection of open sets covering M. Thus, σ is smooth on M, and we have shown the following.

Theorem 5.3 *Let σ be an (r, s)-type tensor field on a smooth manifold, M. Then σ is smooth if and only if for any coordinate chart, (φ, U), on M, the local coordinate representation of σ on U is given by*

$$\sigma|_U = \sum_{\substack{k_1,...,k_r\\ l_1,...,l_s}}^{m} \sigma|_U(dx^{k_1}, ..., dx^{k_r}, E_{l_1}, ..., E_{l_s})E_{k_1}\otimes\cdots\otimes E_{k_r}\otimes dx^{l_1}\otimes\cdots\otimes dx^{l_s}$$

where the component functions $\sigma|_U(dx^{k_1}, \ldots, dx^{k_r}, E_{l_1}, \ldots, E_{l_s})$ are smooth functions on U.

Having derived a series of results on tensors and tensor fields in the most general setting, we will now focus on a particular type of tensor. These particular tensors are an important aspect of both geometry and theoretical physics, and they will be the focus of the remainder of our studies.

5.3 Metric Tensors

In order to study geometrical properties of a space or physical phenomena within a space, we need some means of measurement. Such a construction is *not* intrinsic to the space itself. There will, in fact, be many means of taking measurements within any particular space. For theoretical purposes, however, this does not pose a problem. We simply want to know that such a means does exist. That is, we need some kind of metric structure on our space. The construction of this metric structure requires a specific type of tensor, namely a $(0,2)$-type tensor, or a second order covariant tensor. At a point, p, on a manifold, M, such a tensor will be a real-valued, bilinear operator on T_pM.

The motivation for this is not difficult to see if we simply look at our past experience. Take, for example, the problem of measuring the length of a curve in three dimensional space. If the curve is given by $\gamma(t) = (x^1(t), x^2(t), x^3(t))$, for $t \in (a,b)$, then the arclength, L, is defined to be

$$\begin{aligned} L &= \int_a^b \sqrt{\left(\frac{dx^1}{dt}\right)^2 + \left(\frac{dx^2}{dt}\right)^2 + \left(\frac{dx^3}{dt}\right)^2}\, dt \\ &= \int_a^b \sqrt{\langle \dot{\gamma}(t), \dot{\gamma}(t) \rangle} dt. \end{aligned}$$

Thus, our classical definition of arclength requires the notion of an inner product on $\mathbb{R}^3$, which, as we have seen, is just a $(0,2)$-type tensor on $\mathbb{R}^3$. Other classical geometrical quantities, such as surface area, curvature, etc, can also be formulated using inner products. Hence, it makes sense to base our construction of a means of measurement on these particular types of tensors.

Of course, an inner product, in addition to being bilinear, is symmetric and positive definite. So, we will require a bit more of our second order covariant tensors than just their bilinearity.

Definition 5.4 *A second order covariant tensor, ρ, over a vector space, V, is said to be* **symmetric** *if, for every $v_1, v_2 \in V$, we have $\rho(v_1, v_2) = \rho(v_2, v_1)$. If σ is a second order covariant tensor field over a smooth manifold, M, we say σ is symmetric if, for every point, $p \in M$, the tensor σ_p is symmetric over T_pM.*

Definition 5.5 *A* **metric tensor***, G, on a smooth manifold, M, is a smooth, symmetric, second order covariant tensor field over M. At a point, $p \in M$, we will denote the $(0,2)$-type tensor induced by this field by G_p.*

Note that we have not required G to be positive definite. This restriction will come later. Also, we should technically refer to G as a *metric tensor field*, but this is to be implied, as we want G to be defined on each tangent space of M. Many authors abuse the nomenclature even more, referring to this tensor field as a *metric*. We will not resort to this, as we will show later on that the metric tensor is used to define an actual metric on the manifold. Throughout the remainder of this chapter, we will derive consequences of this definition, assuming that such a tensor field exists. This is not at all obvious. We defer this existence proof, which is lengthy but instructive, until Chapter 6.

Intuitively, a metric tensor just defines a symmetric bilinear form on each tangent space of M, such that the transition from one point to another is smooth. Consider the local coordinate representation of G on a coordinate chart (φ, U). As usual, let $\{E_i\}$ and $\{dx^i\}$ denote the coordinate frames for the tangent and cotangent spaces, respectively, on U. Then, from our general results, we can express the metric tensor $G|_U$ in the form

$$G = \sum_{i,j=1}^{m} g_{ij} dx^i \otimes dx^j,$$

where each g_{ij} is a smooth real-valued function on U. For any point, $p \in U$, this actually defines a bilinear form in the classical sense. If X_1 and X_2 are smooth vector fields on M, then they have local coordinate representations on U given by

$$X_1 = \sum_{k=1}^{m} \alpha_1^k E_k \tag{6}$$

$$X_2 = \sum_{l=2}^{m} \alpha_2^l E_l. \tag{7}$$

For any point, $p \in U$, we have

$$\begin{aligned}
G_p(X_{1_p}, X_{2_p}) &= \sum_{i,j=1}^{m} g_{ij}(p) dx_p^i \otimes dx_p^j \left(\sum_{k=1}^{m} \alpha_1^k(p) E_{p_k}, \sum_{l=1}^{m} \alpha_2^l(p) E_{p_l} \right) \\
&= \sum_{i,j=1}^{m} g_{ij}(p) dx_p^i \left(\sum_{k=1}^{m} \alpha_1^k(p) E_{p_k} \right) dx_p^j \left(\sum_{l=1}^{m} \alpha_2^l(p) E_{p_l} \right) \\
&= \sum_{i,j=1}^{m} g_{ij}(p) \left(\sum_{k=1}^{m} \alpha_1^k(p) dx_p^i(E_{p_k}) \right) \left(\sum_{l=1}^{m} \alpha_2^l(p) dx_p^j(E_{p_l}) \right) \\
&= \sum_{i,j=1}^{m} g_{ij}(p) \sum_{k,l=1}^{m} \alpha_1^k(p) \alpha_2^l(p) \delta_k^i \delta_l^j \\
&= \sum_{i,j=1}^{m} g_{ij}(p) \alpha_1^i(p) \alpha_2^j(p)
\end{aligned}$$

Let $[G_p]$ be the $m \times m$ matrix whose entries are defined by $[G_p]_{ij} = g_{ij}(p)$, and let $v(p)$ and $w(p)$ be the $m \times 1$ vectors whose components are given, respectively, by $(v(p))_j = \alpha_2^j$ and $(w(p))_i = \alpha_1^i$. Then we have

$$G_p(X_{1_p}, X_{2_p}) = v(p)^T [G_p] w(p).$$

Letting p vary over U, we have the local coordinate representation of G on U, given in this matrix form by

$$\begin{aligned} G|_U\big(X_1|_U, X_2|_U\big) &= \sum_{i,j=1}^{m} g_{ij}\alpha_1^i \alpha_2^j \\ &= v^T[G]w, \end{aligned}$$

where X_1 and X_2 are arbitrary smooth vector fields on M with local representations given by relations (6) and (7), $[G]$ is the matrix whose entries are the smooth functions $g_{ij} : U \to \mathbb{R}$, and v and w are the vector valued functions on U whose images at a point $p \in U$ are the components of X_{1_p} and X_{2_p} respectively.

This local representation in terms of matrices will be important in Chapter 7 when we discuss Lorentzian metric tensors. However, our first goal will be to construct a particular type of metric tensor that is positive definite at each point in the following sense.

Definition 5.6 *A second order covariant tensor, ρ, over a vector space, V, is said to be* **positive definite** *over V if $\rho(v, v) \geq 0$ for all $v \in V$ and equality holds if and only if $v = 0$. A metric tensor, G, on a smooth manifold is said to be positive definite if, for each point, $p \in M$, the tensor G_p is positive definite over T_pM.*

This brings us the to the definition that lies at the foundation of differential geometry.

Definition 5.7 *Let M be a smooth manifold. A metric tensor, G, on M is a* **Riemannian Metric Tensor** *if it is positive definite. If this is the case, we call M a* **Riemannian manifold***.*

Recalling that the tangent space of $\mathbb{R}^m$ is just $\mathbb{R}^m$, itself, we see right away that there is a Riemannian metric tensor on the manifold $M = \mathbb{R}^m$. There is only one coordinate chart, namely the identity map $\mathbf{id} : \mathbb{R}^{\mathbf{m}} \to \mathbb{R}^{\mathbf{m}}$, so we simply define G by $G_p(x, y) = \langle x, y \rangle$, where p is any point in $\mathbb{R}^m$ and x and y are any two vectors in

$\mathbb{R}^m$. Thus, on manifolds that are *globally Euclidean*, the typical Riemannian metric tensor is simply the usual inner product on the space.

The whole of Chapter 6 is devoted to proving that a Riemannian metric tensor exists on any smooth manifold. Part of this proof will require transferring, locally at least, the structure of the Euclidean metric tensor, $\langle , \rangle$, to M. That is, we will use the coordinate charts, (φ, U), to transfer the Euclidean metric tensor on $\varphi(U)$ to $U \subset M$. To facilitate this, we will need the following results.

Definition 5.8 *Let $F : M \to N$ be a smooth map between smooth manifolds. We say that F is an* **immersion** *if for each $p \in M$, the rank of F at p equals m, the dimension of M.*

The rank of a mapping $F : M \to N$ at $p \in M$ is defined as one would expect from classical analysis. It is the rank of the Jacobian matrix of the coordinate representation $\psi \circ F \circ \varphi^{-1}$, for coordinate systems (ψ, V) and (φ, U) around $F(p)$ and p, respectively.

Theorem 5.4 *Let $F : M \to N$ be an immersion, and suppose G is a Riemannian metric tensor on N. Define a tensor field, F^*G, on M by $(F^*G)_p(X_{1_p}, X_{2_p}) = G_{F(p)}(D_pF(X_{1_p}), D_pF(X_{2_p}))$, for each $p \in M$ and any smooth vector fields, X_1 and X_2 on M. Then F^*G is a Riemannian metric tensor on M.*

Proof We first show that F^*G is smooth. We will do this by using the local coordinate characterization.

Let p be any point in M, and let $q = F(p)$. Let (φ, U) and (ψ, V) be coordinate systems around p and q, respectively. Denote the coordinate functions of φ and ψ by x^i and y^i respectively. On V, we can represent G locally in the form

$$G = \sum_{i,j=1}^{m} g_{ij} dy^i \otimes dy^j.$$

Let X_1 and X_2 be smooth vector fields on M, and consider their representations on U, given by relations (6) and (7) above. Then

$$
\begin{aligned}
(F^*G)_p(X_{1_p}, X_{2_p}) &= G_q(D_pF(X_{1_p}), D_pF(X_{2_p})) \\
&= G_q\left(D_pF\left(\sum_{k=1}^{m} \alpha_1^k(p)E^x_{p_k}\right), D_pF\left(\sum_{l=1}^{m} \alpha_2^l(p)E^x_{p_l}\right)\right) \\
&= G_q\left(\sum_{k=1}^{m} \alpha_1^k(p)D_pF(E^x_{p_k}), \sum_{l=1}^{m} \alpha_2^l(p)D_pF(E^x_{p_l})\right) \\
&= \sum_{k,l=1}^{m} \alpha_1^k(p)\alpha_2^l(p)G_q\big(D_pF(E^x_{p_k}), D_pF(E^x_{p_l})\big) \\
&= \sum_{k,l=1}^{m} \alpha_1^k(p)\alpha_2^l(p)G_q\left(\sum_{i=1}^{n} \frac{\partial y^i}{\partial x^k}\Big|_{\varphi(p)} E^y_{q_i}, \sum_{j=1}^{n} \frac{\partial y^j}{\partial x^l}\Big|_{\varphi(p)} E^y_{q_j}\right) \\
&= \sum_{k,l=1}^{m} \alpha_1^k(p)\alpha_2^l(p) \sum_{i,j=1}^{n} \frac{\partial y^i}{\partial x^k}\Big|_{\varphi(p)} \frac{\partial y^j}{\partial x^l}\Big|_{\varphi(p)} G_q(E^y_{q_i}, E^y_{q_j}) \\
&= \sum_{k,l=1}^{m} \sum_{i,j=1}^{n} \frac{\partial y^i}{\partial x^k}\Big|_{\varphi(p)} \frac{\partial y^j}{\partial x^l}\Big|_{\varphi(p)} G_q(E^y_{q_i}, E^y_{q_j}) dx^k_p \otimes dx^l_p(X_{1_p}, X_{2_p})
\end{aligned}
$$

But this actually holds for any point in U, so we have

$$
(F^*G)|_U(X_1|_U, X_2|_U) = \sum_{k,l=1}^{m} \sum_{i,j=1}^{n} \frac{\partial y^i}{\partial x^k} \frac{\partial y^j}{\partial x^l} G(E^y_i, E^y_j) dx^k \otimes dx^l(X_1|_U, X_2|_U),
$$

where the partial derivatives are to be evaluated at the point $\varphi(p)$ in question. This shows that the component functions of F^*G on U are given by

$$
\sum_{i,j=1}^{n} \frac{\partial y^i}{\partial x^k} \frac{\partial y^j}{\partial x^l} G(E^y_i, E^y_j).
$$

Since F is a smooth mapping, the mappings $p \mapsto \frac{\partial y^i}{\partial x^k}|_{\varphi(p)}$, for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$ are smooth. Moreover, since G is smooth by hypothesis, its components, $G(E^y_i, E^y_j)$, are smooth on V, implying that the mappings $p \mapsto G_q(E^y_{q_i}, E^y_{q_j})$, for $p \in U$ and $q = F(p)$, are smooth. Hence, the components of F^*G are smooth on U. The chart, (φ, U), was arbitrary, so, by Theorem 5.3, we can conlcude that F^*G is smooth on M.

To see that F^*G is symmetric, let p be any point in M, and consider $(F^*G)_p$. If X_p and X'_p are any two tangent vectors at p, then the symmetry of G implies

$$\begin{aligned}(F^*G)_p(X_p, X'_p) &= G_{F(p)}\big(D_pF(X_p), D_pF(X'_p)\big)\\ &= G_{F(p)}\big(D_pF(X'_p), D_pF(X_p)\big)\\ &= (F^*G)_p(X'_p, X_p).\end{aligned}$$

Finally, to see that F^*G is positive definite, let p be any point in M, and let X_p be any tangent vector at p. Then

$$(F^*G)_p(X_p, X_p) = G_{F(p)}\big(D_pF(X_p), D_pF(X_p)\big).$$

But G is positive definite, so $(F^*G)_p(X_p, X_p) \geq 0$ for all $X_p \in T_pM$. Moreover, if $X_p = 0$, then $D_pF(X_p) = 0 \Rightarrow G_{F(p)}(D_pF(X_p), D_pF(X_p)) = 0 \Rightarrow (F^*G)(X_p, X_p) = 0$. Conversely, suppose $(F^*G)_p(X_p, X_p) = 0$. Then, since G is positive definite, we have $G_{F(p)}(D_pF(X_p), D_pF(X_p)) = 0 \Rightarrow D_pF(X_p) = 0$. But $F : M \to N$ is an immersion, so the rank of F at p is m. This implies that the dimension of $D_pF(T_pM)$ is m. Hence, D_pF is injective, and it follows that $D_pF(X_p) = 0 \Rightarrow X_p = 0$.

QED

Like vector fields, tensor fields over a smooth manifold are more naturally described in the context of vector bundles. We know that the set of (r, s)-type tensors over a finite dimensional vector space is, itself, a vector space of finite dimension. So, as we used the tangent space at each point $p \in M$ as the typical fiber of the tangent bundle, we can take the set of all (r, s)-type tensors over T_pM, for each $p \in M$, to be the typical fiber of another vector bundle, called the (r, s)-**type tensor bundle**. This vector bundle is usually denoted T^r_s. Using a proof that is not much different from the tangent bundle case, one can show that T^r_s is a smooth manifold of dimension $m + m^{r+s}$. A smooth (r, s)-type tensor field, σ, is then defined to be a smooth map $\sigma : M \to T^s_r$. It is simply a map that defines an (r, s)-type tensor at each point $p \in M$ such that the transition between points p and q in M is smooth. We have, of course, not followed this method of construction, as the necessary excursion into differential topology would take us too far off course. It is important to point out, however, that all of the material in this and the preceding chapter can be generalized considerably. An interested reader is referred to the works by Chern and Steenrod.

6 Riemannian Manifolds

In this chapter, we will prove that a Riemannian metric tensor always exists on a smooth manifold. The proof is constructive, but only to a certain extent. We will explicitly construct a smooth, symmetric, positive definite tensor field, but its specific form will not be evident from our proof. This is a consequence of the fact that the proof is largely topological in nature, but not completely. This is a significant point, and, given our goal of adequately redefining the mathematical notion of a space, we would be remiss if we did not comment on it.

We will construct the Riemannian metric tensor almost entirely from the topological properties of a manifold. We will begin by proving the existence of a *partition of unity* on a manifold, a tool which is extremely useful in topology and global analysis. A partition of unity is essentially just a collection of functions on a topological space satisfying certain properties. Such a collection always exists on a manifold, whether it is differentiable or not. Typically, partitions of unity in general topological studies consist only of continuous functions. With smooth manifolds, however, we will need various degrees of differentiability. Thus, we will require the functions in our partition of unity to be smooth. We will have to rely on the smooth structure of the manifold to aid us in this respect. The actual construction of such a set of functions is lengthy, and it will require a series of preliminary definitions and technical lemmas. After the existence of the partition of unity is established, we then use this to define a tensor field on M, concluding by showing that this tensor field is a Riemannian metric tensor.

The significance of this process is that we must venture beyond the mere topological properties of the manifold to define a means of measurement. The topology of the manifold takes us a long way in the construction process, as will be evident during our progression. However, in order to completely develop a well-defined metric tensor on the manifold, we must appeal to the differentiable structure of the manifold. This structure is not intrinsic. As we discussed in chapter 1, it may not even be unique. Hence, the very nature of our construction process shows that there can be no preferred means of measurment, no preferred coordinate system, and no preferred point of view in making observations concerning a space in either geometry or physics. Yet, having a means of measurement is crucial to all of science. We cannot make observations, collect data, make and test hypotheses, or conduct experiments without some means of measuring various properties of the phenomena we observe. In this respect, the mere act of making observations and taking measurements becomes a scientific discipline of its own. By what means do we make such observations? What information can we draw from our measurements? What further hypotheses are valid based on these observations? With questions like these, it is clear that we must be assured of at least the existence of some means of measurement on an arbitrary space.

6.1 Partitions of Unity

Throughout this section, and the rest of this chapter, we will assume that M is a smooth manifold. This section will rely heavily on the topology of the manifold. Consequently, it will be abstract in nature, and intuition may be difficult. The results we derive will be vital, however, to our further work. Thus, we will be very detailed in our proofs.

We will begin with a series of definitions. The concepts we discuss here are all common to general topology, but we will specialize our results slightly to the concept of a manifold.

Definition 6.1 *An open covering, $\{\mathscr{O}\}_{\alpha\in\mathscr{A}}$, of a manifold M is said to be* **locally finite** *if each $p \in M$ has a neighborhood, U, that intersects only a finite number of the sets, $\mathscr{O}_\alpha$.*

Definition 6.2 *If $\{A_\alpha\}$ and $\{B_\beta\}$ are two coverings of M, then we say that $\{A_\alpha\}$ is a* **refinement** *of $\{B_\beta\}$ if each $A_\alpha \subset B_\beta$ for some B_β.*

Definition 6.3 *A C^∞* **partition of unity** *on M is a collection of C^∞ functions, $\{f_\alpha\}$ on M such that*

i) $f_\alpha \geq 0$ *on M for all α;*

ii) the collection $\{supp(f_\alpha)\}$ forms a locally finite covering of M, where $supp(f_\alpha)$ denotes the support of f_α, or the closure of the set on which f_α is nonzero;

iii) $\sum_\alpha f_\alpha(p) = 1$ *for each $p \in M$.*

A partition of unity, $\{f_\alpha\}$, is said to be subordinate to an open covering $\{A_\gamma\}$ if, for each α, there is some set A_γ such that $supp(f_\alpha) \subset A_\gamma$.

Note that the sum in property *iii* of the partition of unity is always well-defined. For any $p \in M$, property *ii* implies that there is a neighborhood, U, of p such that only finitely many functions, say $\{f_{\alpha_i}\}_{i=1}^k$, are nonzero on U. Hence, this sum is always finite. Moreover, for any neighborhood, V, of p having this property, it follows that any function, f_α, whose value at p is nonzero must be included in this finite collection. Hence, the sum is always finite and must equal 1 regardless of what neighborhood we choose around p to intersect only finitely many of the sets $supp(f_\alpha)$.

We will prove several lemmas leading up to the existence theorem for partitions of unity. The first is the most lengthy and the most technical, and it establishes the existence of countable locally finite refinement of the coordinate covering of M. In general topology, this result is proved in a more general setting. A topological space with the property that every open covering has a locally finite refinement is called a **paracompact** space. All metrizable spaces are paracompact, and, in more generality, so are all locally compact second countable Hausdorff spaces. So, we could appeal to these results. However, for the sake of a complete and self-contained exposition, we would like to tie all of our results together using a common thread. Thus, we will follow the method of Boothby and prove our own version of this result, staying true to our chosen notation and conventions.

Lemma 6.1 *Let $\{A_\alpha\}$ be an open covering of M. Then there exists a countable locally finite refinement of $\{A_\alpha\}$, denoted $\{(\varphi_i, U_i)\}_{i\geq 1}$, consisting of coordinate charts and satisfying the following.*

i) For each $i \geq 1$, $\varphi_i(U_i)$ is an open ball of finite radius in $\mathbb{R}^m$. That is, for each i, there is a point $a_i \in \mathbb{R}^m$ and a real number $\epsilon_i > 0$ such that $\varphi_i(U_i) = B_{\epsilon_i}(a_i)$.

ii) For each $i \geq 1$, the sets $V_i = \varphi_i^{-1}(B_{\epsilon_i/4}(a_i)) \subset U_i$ also form a locally finite countable cover of M.

<u>Proof</u> By our results in chapter 2, there is a countable covering of M by open sets, $\{\mathscr{O}_i\}_{i\geq 1}$, such that $\overline{\mathscr{O}_i}$ is compact for all $i \geq 1$. For each $i \geq 1$, define

$$P_i = \bigcup_{1\leq k\leq i} \overline{\mathscr{O}_k}.$$

Then each P_i is compact and satisfies $P_i \subset P_{i+1}$. Moreover, $\cup_i P_i = M$.

Now, we will inductively construct another sequence of compact sets, $\{Q_i\}_{i\geq 1}$, such that $P_i \subset Q_i$. Let $Q_1 = P_1$. Then Q_1 is obviously compact, and $P_1 \subset Q_1$. Suppose we have constructed $Q_1, \ldots, Q_i$ for some $i > 1$, and suppose that all of these sets are compact and satisfy $P_i \subset Q_i$. Then the set $Q_i \cup P_{i+1}$ is a compact subset of M. Hence, there are finitely many sets $\mathscr{O}_j$, $1 \leq j \leq s$, such that

$$Q_i \cup P_{i+1} \subset \bigcup_{j=1}^{s} \mathscr{O}_j.$$

Define Q_{i+1} by

$$Q_{i+1} = \bigcup_{j=1}^{s} \overline{\mathscr{O}_j}.$$

Then Q_{i+1} is compact and we have $P_{i+1} \subset \cup_{j=1}^{s} \mathscr{O}_j \subset Q_{i+1}$. Hence, by induction, we have constructed our desired seqeunce, $\{Q_i\}$. Moreover, note that

$$Q_i \subset \bigcup_{j=1}^{s} \mathscr{O}_j \subset int(Q_{i+1}) \subset Q_{i+1},$$

where $int(Q_{i+1})$ denotes the interior of Q_{i+1}. Finally, since the collection $\{P_i\}$ covers M, for any $p \in M$ there is a set P_k such that $p \in P_k$. It follows that $p \in P_k \subset Q_k$, implying that the collection $\{Q_i\}$ also covers M.

Next, define $Q_{-1} = Q_0 = \emptyset$. We will define two more collections of sets from the collection $\{Q_i\}$. Define $L_i = Q_i - int(Q_{i-1})$ and $K_i = int(Q_{i+1}) - Q_{i-2}$ for $i \geq 1$. Then each L_i is compact, each K_i is open, and $L_i \subset K_i$ for all i.

For any $i \geq 1$, consider the set K_i. Let p be any point in K_i. Since $\{A_\alpha\}$ is an open covering of M by hypothesis, we have $p \in A_\alpha$ for some α. Likewise, there is some coordinate chart (φ, U) such that $p \in U$. Let $U_p = K_i \cap A_\alpha \cap U$. Then U_p is a neighborhood of p, and we have $U_p \subset K_i$, $U_p \subset A_\alpha$, and $(\varphi|_{U_p}, U_p)$ is a coordinate system around p. By shrinking U_p, if necessary, we can assume without loss of generality that $\varphi(U_p)$ is an open ball of finite radius in $\mathbb{R}^m$ centered at $\varphi(p)$. (If $\varphi(U_p)$ is not an open ball, it will contain an open ball, B, centered at $\varphi(p)$, so we can simply let U_p be the set $\varphi^{-1}(B) \subset U_p$.) Let ϵ_p denote the radius of this open ball, and let $a_p = \varphi(p)$, so that $\varphi(U_p) = B_{\epsilon_p}(\varphi(p))$. We then define $V_p = \varphi^{-1}(B_{\epsilon_p/4}(a_p)$, so that $V_p \subset U_p$.

Since $p \in K_i$ was arbitrary, we can follow this same construction for every $p \in K_i$. Doing so, we obtain two collections of sets associated with K_i, $\mathscr{U}_i = \{U_p(a_p, \epsilon_p) : p \in K_i\}$ and $\mathscr{V}_i = \{V_p(a_p, \epsilon_p) : p \in K_i\}$. Here we use the notation $U_p(a_p, \epsilon_p)$ and $V_p(a_p, \epsilon_p)$ to indicate that there is some coordinate map, φ, defined on some coordinate neighborhood containing U_p, such that $\varphi(U_p) = B_{\epsilon_p}(a_p)$ and $V_p = \varphi^{-1}(B_{\epsilon_p/4}(a_p))$. We do not indicate the coordinate map, since the same map may correspond to more than one point $p \in K_i$. We only need the fact that such a map exists.

The notation and indexing is clearly cumbersome here, so, to sum up what we have constructed so far, we note that $\mathscr{U}_i$ and $\mathscr{V}_i$ are collections of open sets satisfying

1. each $U_p(a_p, \epsilon_p) \subset A_\alpha$ for some α,
2. each $U_p(a_p, \epsilon_p) \subset K_i$,

3. for each $p \in K_i$, there is a coordinate mapping, φ, such that $(\varphi, U_p(a_p, \epsilon_p))$ is a coordinate system around p,

4. $\varphi(U_p) = B_{\epsilon_p}(a_p)$ for each p,

5. $V_p = \varphi^{-1}(B_{\epsilon_p/4}(a_p)) \subset U_p$.

Each V_p is a subset of K_i, and there is one containing each point $p \in K_i$. So, $K_i = \cup_{p \in K_i} V_p$. Likewise, it also follows that $K_i = \cup_{p \in K_i} U_p$. Since L_i is a compact subset of K_i, there is a finite collection of sets in $\mathscr{V}_i$, say $\{V_1^i(a_1^i, \epsilon_1^i), V_2^i(a_2^i, \epsilon_2^i), \ldots, V_{r_i}^i(a_{r_i}^i, \epsilon_{r_i}^i)\}$, such that

$$L_i \subset \bigcup_{\lambda=1}^{r_i} V_\lambda^i(a_\lambda^i, \epsilon_\lambda^i) \subset K_i.$$

As before, we use the notation $V_\lambda^i(a_\lambda^i, \epsilon_\lambda^i)$ to indicate the fact that, for each $\lambda = 1, \ldots, r_i$, there is a coordinate mapping φ, a point a_λ^i, and a positive real number ϵ_λ^i such that $V_\lambda^i(a_\lambda^i, \epsilon_\lambda^i) = \varphi^{-1}(B_{\epsilon_\lambda^i/4}(a_\lambda^i))$. If we consider the corresponding collection of sets $\{U_1^i(a_1^i, \epsilon_1^i), U_2^i(a_2^i, \epsilon_2^i), \ldots, U_{r_i}^i(a_{r_i}^i, \epsilon_{r_i}^i)\}$, where, as before, there is a coordinate map φ, for each λ, such that $\varphi(U_\lambda^i(a_\lambda^i, \epsilon_\lambda^i)) = B_{\epsilon_\lambda^i}(a_\lambda^i)$, then we further have

$$L_i \subset \bigcup_{\lambda=1}^{r_i} V_\lambda^i(a_\lambda^i, \epsilon_\lambda^i) \subset \bigcup_{\lambda=1}^{r_i} U_\lambda^i(a_\lambda^i, \epsilon_\lambda^i) \subset K_i.$$

We next note that $\cup_{i \geq 1} L_i = M$. To see this, suppose $p \in M$. Then $p \in Q_i$ for at least one $i \geq 1$. Let k be the smallest natural number such that $p \in int(Q_k)$. Then $p \in Q_k \Rightarrow p \in Q_k - int(Q_{k-1}) = L_k$.

Now, let $\mathscr{B}_U^i = \{U_1^i(a_1^i, \epsilon_1^i), U_2^i(a_2^i, \epsilon_2^i), \ldots, U_{r_i}^i(a_{r_i}^i, \epsilon_{r_i}^i)\}$, and let $\mathscr{B}_V^i = \{V_1^i(a_1^i, \epsilon_1^i), V_2^i(a_2^i, \epsilon_2^i), \ldots, V_{r_i}^i(a_{r_i}^i, \epsilon_{r_i}^i)\}$ for each $i \geq 1$. Then define

$$\mathscr{B}_{\mathscr{U}} = \bigcup_{i=1}^{\infty} \mathscr{B}_U^i$$
$$\mathscr{B}_{\mathscr{V}} = \bigcup_{i=1}^{\infty} \mathscr{B}_V^i.$$

Both $\mathscr{B}_{\mathscr{U}}$ and $\mathscr{B}_{\mathscr{V}}$ are countable unions of finite collections of sets. Thus, they are countable collections of open sets. Moreover, since the collection $\{L_i\}$ covers M, the collections $\mathscr{B}_{\mathscr{U}}$ and $\mathscr{B}_{\mathscr{V}}$ do also. Since each $U_\lambda^i(a_\lambda^i, \epsilon_\lambda^i) \subset A_\alpha$ for some α, it follows that $\mathscr{B}_{\mathscr{U}}$ is a refinement of $\{A_\alpha\}$, and $\mathscr{B}_{\mathscr{V}}$ is a refinement of both $\{A_\alpha\}$ and $\mathscr{B}_{\mathscr{U}}$. In addition, since each $U^i_\lambda(a^i_\lambda, \epsilon^i_\lambda)$ is a subset of a coordinate chart, we see that $(\varphi^i_\lambda, U^i_\lambda(a^i_\lambda, \epsilon^i_\lambda))$ is a coordinate chart, where φ^i_λ is a coordinate mapping corresponding to some coordinate neighborhood containing $U^i_\lambda(a^i_\lambda, \epsilon^i_\lambda)$. That is, by pairing each set in $\mathscr{B}_{\mathscr{U}}$ with a coordinate mapping defined on it, we see that $\mathscr{B}_{\mathscr{U}}$ contains coordinate charts. Thus, all we have left to show is that the collections $\mathscr{B}_{\mathscr{U}}$ and $\mathscr{B}_{\mathscr{V}}$ are locally finite.

Let p be any point in M. Then $p \in int(Q_i) \subset Q_i$ for some $i \geq 1$. For any $k \geq i+2$, we have $K_k = int(Q_{k+1}) - Q_{k-2} \subset int(Q_{k+1}) - Q_i$, because $Q_i \subset Q_{k-2}$. It follows that $K_k \cap Q_i = \emptyset$ for $k \geq i+2$. Now, consider $V^k_\lambda(a^k_\lambda, \epsilon^k_\lambda)$ for $k \geq i+2$ and $1 \leq \lambda \leq r_k$. That is, we consider sets in the collection $\mathscr{B}^k_V$ for $k \geq i+2$. If $p \in V^k_\lambda(a^k_\lambda, \epsilon^k_\lambda) \cap int(Q_i)$ for $k \geq i+2$, $1 \leq \lambda \leq r_k$, then $p \in Q_i$ because $int(Q_i) \subset Q_i$, and $p \in K_k$ because

$$p \in V^k_\lambda(a^k_\lambda, \epsilon^k_\lambda) \subset \bigcup_{\lambda=1}^{r_k} V^k_\lambda(a^k_\lambda, \epsilon^k_\lambda) \subset K_k.$$

Thus, $[V^k_\lambda(a^k_\lambda, \epsilon^k_\lambda) \cap int(Q_i)] \subset [K_k \cap Q_i] = \emptyset$ for all $k \geq i+2$, $1 \leq \lambda \leq r_k$. In other words, none of the sets $V^k_\lambda(a^k_\lambda, \epsilon^k_\lambda)$ for $k \geq i+2$ and $1 \leq \lambda \leq r_k$ intersect $int(Q_i)$. The only other sets left in $\mathscr{B}_{\mathscr{V}}$ are those in the collections $\mathscr{B}^k_V$ for $k = 1, 2, \ldots, i+1$, each of which is finite. Hence, only a finite number of sets $V^i_\lambda(a^i_\lambda, \epsilon^i_\lambda)$ can intersect $int(Q_i)$. Since $p \in M$ was arbitrary, we can conlude that $\mathscr{B}_{\mathscr{V}}$ is locally finite.

An exactly parallel argument shows that $\mathscr{B}_{\mathscr{U}}$ is also locally finite. So, finally, by reindexing the collections $\mathscr{B}_{\mathscr{U}}$ and $\mathscr{B}_{\mathscr{V}}$, we conclude that $\mathscr{B}_{\mathscr{U}}$ is a countable collection of coordinate charts, (φ_i, U_i), such that, for each $i \geq 1$, there is a point $a_i \in \mathbb{R}^m$ and a positive real number ϵ_i satisfying $\varphi_i(U_i) = B_{\epsilon_i}(a_i)$. Likewise, the collection $\mathscr{B}_{\mathscr{V}}$ is a corresponding collection of sets, V_i, such that, for each $i \geq 1$, $V_i = \varphi_i^{-1}(B_{\epsilon_i/4}(a_i))$. This proves the result.

QED

Any manifold must, by definition, have a covering of coordinate neighborhoods. Thus, applying this lemma, we can always conclude that any manifold, M, smooth or not, has a locally finite coordinate covering $\{(\varphi_i, U_i)\}_{i \geq 1}$, such that, for each i, $\varphi_i(U_i) = B_{\epsilon_i}(a_i)$ for some point $a_i \in \mathbb{R}^m$ and some $\epsilon_i > 0$, $V_i = \varphi_i^{-1}(B_{\epsilon_i/4}(a_i)) \subset U_i$, and such that this covering refines the original coordinate covering. For brevity of notation, we will denote this coordinate covering by $\{(\varphi_i, U_i, V_i)\}_{i \geq 1}$, and we call it a **regular covering**. We also point out another consequence of the existence of such a covering that we will use later on. For any $i \geq 1$, let $W_i = \varphi_i^{-1}(B_{\epsilon_i/2}(a_i))$. Then, since $\overline{B_{\epsilon_i/4}(a_i)} \subset B_{\epsilon_i/2}(a_i) \subset \overline{B_{\epsilon_i/2}(a_i)} \subset B_{\epsilon_i}(a_i)$, we have

$$V_i \subset \overline{V_i} \subset W_i \subset \overline{W_i} \subset U_i \qquad\qquad \forall\, i \geq 1.$$

Now, we need another lemma, this one concerning the existence of smooth functions. We will use this result to define the functions that make up our partition of unity.

Lemma 6.2 *Let $\epsilon > 0$ be given, and let $a \in \mathbb{R}^m$ be arbitrary. Then there is a C^∞ function, $f : \mathbb{R}^m \to [0,1]$, such that $f > 0$ on $B_\epsilon(a)$, $f = 1$ on $\overline{B_{\frac{\epsilon}{2}}(a)}$, and $f = 0$ outside $B_\epsilon(a)$.*

Proof   The function, $h : \mathbb{R} \to \mathbb{R}$, defined by

$$h(t) = \begin{cases} e^{-1/t} & t > 0 \\ 0 & t \leq 0 \end{cases}$$

is C^∞ on $\mathbb{R}$. Define $f : \mathbb{R}^m \to \mathbb{R}$ by

$$f(x) = \frac{h(\epsilon^2 - \|x-a\|^2)}{h(\epsilon^2 - \|x-a\|^2) + h(\|x-a\|^2 - \frac{1}{4}\epsilon^2)}.$$

If $\|x-a\| \geq \epsilon$, then $\|x-a\|^2 \geq \epsilon^2 \Rightarrow \epsilon^2 - \|x-a\|^2 \leq 0$ and $f(x) = 0$. So, f vanishes outside $B_\epsilon(a)$. If $\|x-a\| < \epsilon$, then $\epsilon^2 - \|x-a\|^2 > 0 \Rightarrow f(x) > 0$, so f is strictly positive on $B_\epsilon(a)$. If $0 \leq \|x-a\| \leq \frac{\epsilon}{2}$, then $0 \leq \|x-a\|^2 \leq \frac{\epsilon^2}{4} \Rightarrow \|x-a\|^2 - \frac{\epsilon^2}{4} \leq 0$ and

$$f(x) = \frac{h(\epsilon^2 - \|x-a\|^2)}{h(\epsilon^2 - \|x-a\|^2)} = 1.$$

Thus, f is C^∞ on $\mathbb{R}^m$ and satisfies the stated conditions.

QED

Now we can prove the existence of a smooth partition of unity on M.

Theorem 6.3 *Let M be a smooth manifold, and let $\{(\varphi_i, U_i, V_i)\}_{i=1}^\infty$ be a regular covering of M. Then there exists a C^∞ partition of unity, $\{f_i\}_{i=1}^\infty$, subordinate to $\{U_i\}$.*

Proof   As we noted before, our regular covering gives us, for each i, the relation

$$V_i \subset \overline{V_i} \subset W_i \subset \overline{W_i} \subset U_i,$$

where $U_i = \varphi_i^{-1}(B_{\epsilon_i}(a_i))$, $W_i = \varphi_i^{-1}(B_{\epsilon_i/2}(a_i))$, and $V_i = \varphi_i^{-1}(B_{\epsilon_i/4}(a_i))$ for some $a_i \in \mathbb{R}^m$ and $\epsilon_i > 0$. There is a C^∞ function, $c_i : \mathbb{R}^m \to \mathbb{R}$, such that $c_i = 1$ on $\overline{B_{\epsilon_i/4}(a_i)}$, $c_i = 0$ outside $B_{\epsilon_i/2}(a_i)$, and $c_i > 0$ on $B_{\epsilon_i/2}(0)$.

Now, for each $i \geq 1$, define $g_i : M \to [0,1]$ by

$$g_i(p) = \begin{cases} (c_i \circ \varphi_i)(p) & p \in U_i \\ 0 & p \in M - U_i \end{cases}$$

Each g_i is identically 0 on $U_i - \varphi_i^{-1}(B_{\epsilon_i/2}(a_i))$, and it is identically 0 on $M - U_i$. Hence, each g_i is C^∞ on U_i and vanishes outside a compact proper subset of U_i. Thus, each g_i is C^∞ on M. Moreover, any given g_i is only nonzero when $\varphi_i(p) \in B_{\epsilon_i/2}(a_i)$, or when $p \in \varphi_i^{-1}(B_{\epsilon_i/2}(a_i)) = W_i$. Thus, it follows that $supp(g_i) \subset \overline{W_i} \subset U_i$. Consider the family of sets $\{supp(g_i)\}_{i=1}^\infty$. Since the family $\{U_i\}_{i\geq 1}$ is locally finite, and since there is exactly one set $supp(g_i)$ corresponding to each U_i, this family is locally finite. Moreover, if $p \in M$, then $p \in V_k$ for some $k \geq 1$, implying that $g_k(p) = 1$. That is, at least one of the functions g_i is positive at p, so $\{supp(g)\}_{i\geq 1}$ covers M. Hence, $\{supp(g_i)\}_{i\geq 1}$ forms a locally finite covering of M.

Now, define $S : M \to \mathbb{R}$ by $S(p) = \sum_i g_i(p)$. This sum is well-defined. For any $p \in M$, there is a neighborhood, N_p, of p such that N_p intersects $supp(g_i)$ for only finitely many indices, say $\{supp(g_{i_k})\}_{k=1}^n$. Consider $S|_{N_p}$. If $i \neq i_k$ for any $k = 1, \ldots, n$, then $supp(g_i) \cap N_p = \emptyset$, so that $g_i(q) = 0$ for all $q \in N_p$. Hence, on N_p,

$$S(q) = \sum_{i=1}^n g_{i_k}(q).$$

Thus, S is C^∞ on N_p, as it is a finite sum of smooth functions. Since we can repeat this reasoning for every $p \in M$, it follows that S is C^∞ on a collection of open sets covering M. Hence, S is smooth on M.

Now, define $f_i : M \to \mathbb{R}$ by $f_i = \frac{g_i}{S}$. Each f_i is clearly nonnegative. Consider $supp(f_i) = supp(\frac{g_i}{S})$. Since $S(p) > 0$ for all p, the support of f_i is the same as $supp(g_i)$. Thus, $\{supp(f_i)\}_{i\geq 1}$ forms a locally finite covering of M, and $supp(f_i) \subset U_i$ for all i.

Finally, let p be any point in M. Then

$$\sum_i f_i(p) = \sum_i \frac{g_i(p)}{S(p)} = \frac{S(p)}{S(p)} = 1.$$

Hence, $\{f_i\}_{i\geq 1}$ is a partition of unity subordinate to $\{U_i\}$.

QED

6.2 The Riemannian Metric Tensor

The whole of this section is devoted to the Riemannian metric tensor. This can be thought of as a culmination of all that we have done to this point. The existence proof is not difficult to follow. Most of the technical work has been done in Chapter 5 and section 6.1. All we need to do here is put the pieces together. After verifying that a Riemannian metric always exists, we will give some examples of Riemannian manifolds, showing the various ways in which a metric tensor might be constructed in actual practice.

Typically, we will denote a Riemannian metric tensor by a capital G. However, as we will use several different tensor fields in this proof, it will be notationally convenient to use capital greek letters. This will be done only in this proof. Afterwards, we will return to denoting Riemannian metric tensors by G.

Theorem 6.4 *Every smooth manifold, M, admits a Riemannian metric tensor.*

Proof Let $\{(\varphi_i, U_i, V_i)\}_{i\geq 1}$ be a regular covering of M, and let $\{f_i\}_{i\geq 1}$ be a corresponding partition of unity subordinate to the covering $\{U_i\}$.

For any $i \geq 1$, the map $\varphi_i : U_i \to B_{\epsilon_i}(a_i)$ is a diffeomorphism between the submanifold $U_i \subset M$ and the submanifold $B_{\epsilon_i}(a_i)$. Let Ψ denote the usual Euclidean inner product on $\mathbb{R}^m$ restricted to the submanifold $B_{\epsilon_i}(a_i)$. The Euclidean inner product is a smooth, symmetric, positive definite metric tensor. Hence, by Theorem 5.4, $\Theta_i = \varphi_i^*\Psi$ defines a smooth, symmetric, positive definite tensor field on U_i. This gives us a Riemannian metric tensor on U_i for each $i \geq 1$.

By the way we have constructed our partition of unity, each f_i is strictly positive on $V_i = \varphi_i^{-1}(B_{\epsilon_i/4}(a_i))$ and $supp(f_i) \subset \varphi_i^{-1}(\overline{B_{\epsilon_i/2}(a_i)}) = \overline{W_i} \subset U_i$. That is, f_i is nonzero only on a proper compact subset of U_i. So, if we temporarily consider f_i defined only on U_i, then $f_i\Theta_i$ is a smooth second order covariant tensor field on U_i, is a Riemannian metric tensor on V_i, and vanishes outside $W_i = \varphi_i^{-1}(B_{\epsilon_i/2}(a_i))$. (It is easy to see that multiplying a Riemannian metric tensor by a smooth positive function just gives us another Riemannian metric tensor.) So, we can extend $f_i\Theta_i$ over all of M by defining it to be identically 0 outside U_i.

For each $i \geq 1$, let $\Lambda_i = f_i\Theta_i$, where we mean the extension of $f_i\Theta_i$ over all of M by defining it to be zero outside of U_i. It follows that Λ_i is a smooth, symmetric, second order covariant tensor field on M that vanishes outside $W_i = \varphi_i^{-1}(B_{\epsilon_i/2}(a_i))$, but is also positive definite, and, therefore, a Riemannian metric tensor, on $V_i = \varphi_i^{-1}(B_{\epsilon_i/4}(a_i))$.

Now, define $G = \sum_i \Lambda_i$, or more precisely, at any $p \in M$, and for any $X_p, Y_p \in T_pM$, let

$$G_p(X_p, Y_p) = \sum_{i=1}^{\infty} \Lambda_{i,p}(X_p, Y_p).$$

Keep in mind that, when discussing global structure, this expression is preferred over the sum

$$G_p(X_p, Y_p) = \sum_{i=1}^{\infty} f_i(p)\Theta_{i,p}(X_p, Y_p),$$

since each Λ_i is defined on all of M, while it does not make sense to refer to $\Theta_{i,p}$ if $p \notin U_i$. However, at each $p \in M$, a finite variant of the latter expression is always valid, as only a finite number of the functions, f_i, are nonzero at p.

We will show that G is a Riemannian metric tensor by first showing that it is well-defined and symmetric. We will then prove smoothness, concluding by showing that G is positive definite.

First, let p be in M, and let X_p, Y_p be two tangent vectors at p. Since $\{supp(f_i)\}$ forms a locally finite covering of M, there is a neighborhood, N, of p such that only a finite number of functions, say $\{f_{i_j}\}_{j=1}^{l}$, have support so that $supp(f_{i_j}) \cap N \neq \emptyset$. This also implies that, on N, we can express G as

$$G_q(X_q, Y_q) = \sum_{j=1}^{l} \Lambda_{i_j,q}(X_q, Y_q), \qquad q \in N,$$

since the only indices that would possibly produce nonzero terms are the indices i_j for $j = 1, \ldots, l$. We can assume, then, that the set $\{f_{i_j}\}_{j=1}^{l}$ is maximal in the sense that, if $supp(f_i) \cap N \neq \emptyset$, then $f_i \in \{f_{i_j}\}_{j=1}^{l}$. So, in particular, we have

$$G_p(X_p, Y_p) = \sum_{j=1}^{l} \Lambda_{i_j,p}(X_p, Y_p). \tag{8}$$

Now, the set of functions, $\{f_r\}$, such that $p \in supp(f_r)$ must be a subset of $\{f_{i_j}\}_{j=1}^{l}$. It could be all of these functions, but there could also be a function f_i such that $supp(f_i) \cap N \neq \emptyset$ and $p \notin supp(f_i)$. Any function f_{i_j}, for $1 \leq i_j \leq l$, such that $p \notin supp(f_{i_j})$ will not affect the sum defining G_p because $f_{i_j}(p) = 0$. Thus, at the point, p, we can reduce this sum as follows. If $\{f_r\}_{r=1}^{s}$ denotes the set of functions such that $p \in supp(f_r)$, then equation 7 reduces to

$$G_p(X_p, Y_p) = \sum_{r=1}^{s} f_r(p)\Theta_{r,p}(X_p, Y_p). \tag{9}$$

Note that it is valid to express this sum in terms of the tensor fields, Θ_i, since, for each index $r = 1, \ldots, s$, p is in $supp(f_r)$, implying that $p \in U_r$.

Now, suppose that N' is another neighborhood of p such that only a finite number of functions, say $\{f_{i_k}\}_{k=1}^n$, satisfy $supp(f_{i_k}) \cap N' \neq \emptyset$. Suppose also that this set is maximal as before. Then

$$G_p(X_p, Y_p) = \sum_{k=1}^{n} \Lambda_{i_k,p}(X_p, Y_p).$$

Using the same reasoning as before, the set $\{f_r\}_{r=1}^s$ must be a subset of $\{f_{i_k}\}_{k=1}^n$. Thus, this sum, like the previous one, reduces to equation 9. Hence, $G_p(X_p, Y_p)$ is independent of the finite expansion we choose. Since p, X_p, and Y_p were arbitrary, this shows that G is well-defined.

Next, at any point $p \in M$, using this finite expansion in a neighborhood of p, and the symmetry of the tensor fields Λ_i, we have

$$\begin{aligned} G_p(X_p, Y_p) &= \sum_{j=1}^{l} \Lambda_{i_j,p}(X_p, Y_p) \\ &= \sum_{j=1}^{l} \Lambda_{i_j,p}(Y_p, X_p) \\ &= G_p(Y_p, X_p). \end{aligned}$$

Hence, G is symmetric.

Now, to prove that G is smooth, consider again the neighborhood, N, of p, on which G can be expressed as the finite sum

$$G = \sum_{j=1}^{l} \Lambda_{i_j}.$$

Each Λ_{i_j} is C^∞. Hence, on N, G is a finite sum of smooth, second order covariant tensor fields. Let X and Y be two smooth vector fields defined on M, and consider their restrictions to N. Then, for $q \in N$, we have

$$\begin{aligned} G(X,Y)(q) &= G_q(X_q,Y_q) \\ &= \sum_{j=1}^{l} \Lambda_{i_j,q}(X_q,Y_q) \\ &= \sum_{j=1}^{l} \Big(\Lambda_{i_j}(X,Y)\Big)(q). \end{aligned}$$

Since Λ_{i_j} is smooth on N for each j, it follows that the function $G(X,Y)$ is C^∞ on N. Hence, for every $p \in M$, there is an open set containing p on which $G(X,Y)$ is a smooth function. It follows that $G(X,Y)$ is C^∞ on M. Since the vector fields X and Y were arbitrary, this shows that G is a smooth second order covariant tensor field on M.

Finally, we will show that G is positive definite. If $p \in M$, suppose, for $X_p \in T_pM$, we have $G_p(X_p,X_p) = 0$. Choosing a neighborhood, N, of p as before, it follows that

$$G_p(X_p,X_p) = \sum_{j=1}^{l} \Lambda_{i_j,p}(X_p,X_p).$$

Now, $p \in V_k$ for some $k \geq 1$. Taking intersections if necessary, we can assume that $N \subset V_k$. Thus, $f_k(p) > 0$, and f_k must be in the collection $\{f_{i_j}\}_{j=1}^{l}$, since $supp(f_k) \cap N \neq \emptyset$. In fact, f_k must also be in the collection of functions whose support actually contains p, since $f_k(p) > 0$, and, since $p \in N \subset V_k$, we know that Λ_k is positive definite on N. Since $G_p(X_p,X_p)$ is a sum of nonnegative terms, it can only be zero if each term in the sum equals zero. One of these terms, however, corresponds to the index k. Thus, we must have $\Lambda_{k,p}(X_p,X_p) = f_k(p)\Theta_k(X_p,X_p) = 0 \Rightarrow \Theta_k(X_p,X_p) = 0$. But Θ_k is positive definite, so this implies that $X_p = 0$. Hence, G is positive definite, and this shows that G is a Riemannian metric tensor on M.

QED

While our proof has been constructive, its usefulness in application is somewhat limited. In actual practice, Riemannian metric tensors are not usually constructed using partitions of unity, as the collections of functions are often difficult to explicitly use and, in any case, are not unique. Instead, Theorem 5.4 is used in most applications to induce metric tensors on manifolds.

If $F : M \to N$ is a smooth mapping between smooth manifolds, and if F is an immersion, then we say that $F(M)$ is an **immersed submanifold** of N. Note that an immersion must be locally injective since the rank of F at any point $p \in M$ is m. This implies, of course, that the differential DF at each point p is an injective map between T_pM and $T_{F(p)}N$. This requires, then, that m be less than or equal to n. So, the idea of immersing a manifold, M, into another manifold, N, is similar to the notion of imbedding one linear space into another. However, even if F is an immersion, it need not be globally injective. A smooth map $F : M \to N$ such that F is an immersion and F is a homeomorphism is called an **imbedding**, and we say that $F(M)$ is an **imbedded submanifold** of N.

Most manifolds that occur in applications are imbedded or immersed submanifolds of another manifold which usually has a simpler structure. In fact, it has been shown several times, using a variety of methods, that any smooth manifold of dimension m can be smoothly (not just homeomorphically) imbedded as a closed submanifold of $\mathbb{R}^{2m+1}$. The dimension of the Euclidean space can be less than $2m+1$ in special cases, but this is the best one can do in general. We have not focused on such results here, as our goal has been to develop a theory of spaces without any reference to an ambient space or outlying global structure. Nevertheless, in applications, this is a useful result. Given a smooth manifold, M, we may be able to imbed or immerse M into a Euclidean space, $\mathbb{R}^n$. Using Theorem 5.4, this allows us to induce a Riemannian metric tensor on M by transferring the Euclidean metric struture of $\mathbb{R}^n$ to M.

Example 6.1 *Parametrized surfaces in* $\mathbb{R}^3$

If M and N are smooth manifolds, we say that the map $F : \Omega \subset M \to N$ is a **parametrization** of $\Omega \subset M$ in N if Ω is an open submanifold of M, F is a diffeomorphism, and $F(\Omega)$ is an open submanifold of N. The most common examples of parametrizations are mappings from two-dimensional manifolds into $\mathbb{R}^3$. Let M be a smooth two-dimensional manifold, and let (φ, U) be a coordinate chart on M. Denote the coordinate functions of φ by u^1 and u^2, and suppose $r : U \to \mathbb{R}^3$ is a parametrization. The Euclidean metric tensor on $\mathbb{R}^3$ is just the usual inner product, which, in tensor notation is just $G = \sum_{i=1}^{3} dx^i \otimes dx^i$. If X and Y are two smooth vector fields on M, and we consider their restrictions to U, then, applying Theorem 5.4, we have a Riemannian metric tensor, r^*G, on U defined for $p \in U$ by

$$(r^*G)_p(X_p, Y_p) = G_{r(p)}(D_pr(X_p), D_pr(Y_p)). \tag{10}$$

Now, if we think of r as a function of its local coordinates on U, we can represent r by $r(u^1, u^2) = (x^1(u^1, u^2), x^2(u^1, u^2), x^3(u^1, u^2))$. This is nothing more than the local coordinate representation of r at any point $p \in U$. Thus, the differential of r at $p \in U$ takes the form

$$Dr(u^1, u^2) = \begin{bmatrix} D_1x^1 & D_2x^1 \\ D_1x^2 & D_2x^2 \\ D_1x^3 & D_2x^3 \end{bmatrix},$$

where we have used the shorthand notation D_ix^j for the partial derivative $\frac{\partial x^j}{\partial u^i}$. Since, for any $p \in U$, the basis vectors for the tangent space, T_pM, are just the partial derivative operators with respect to the coordinate functions u^1 and u^2, let us denote the basis field induced on U by the more intuitive form

$$\left\{\frac{\partial}{\partial u^1}, \frac{\partial}{\partial u^2}\right\}.$$

Likewise, we can represent the basis of $T_{r(p)}\mathbb{R}^3$ by

$$\left\{\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right\}.$$

Then we can represent $X|_U$ and $Y|_U$ by

$$\begin{aligned} X &= \alpha^1\frac{\partial}{\partial u^1} + \alpha^2\frac{\partial}{\partial u^2} \\ Y &= \beta^1\frac{\partial}{\partial u^1} + \beta^2\frac{\partial}{\partial u^2}. \end{aligned}$$

Then, at any point $p \in U$, equation 10 takes the form

$$\begin{aligned}
(r^*G)(X_p, Y_p) &= G_{r(p)}\big(D_p r(X_p), D_p r(Y_p)\big) \\
&= \sum_{k=1}^{3} dx^k \otimes dx^k \big(D_p r(X_p), D_p r(Y_p)\big) \\
&= \sum_{k=1}^{3} dx^k \left(D_p r\left(\sum_{i=1}^{2} \alpha^i \frac{\partial}{\partial u^i}\right)\right) dx^k \left(D_p r\left(\sum_{j=1}^{2} \beta^j \frac{\partial}{\partial u^j}\right)\right) \\
&= \sum_{k=1}^{3}\sum_{i=1}^{2} \alpha^i dx^k \left(D_p r\left(\frac{\partial}{\partial u^i}\right)\right) \sum_{j=1}^{2} \beta^j dx^k \left(D_p r\left(\frac{\partial}{\partial u^j}\right)\right) \\
&= \sum_{k=1}^{3}\sum_{i,j=1}^{2} \alpha^i \beta^j dx^k \left(\sum_{s=1}^{3} \frac{\partial x^s}{\partial u^i}\frac{\partial}{\partial x^s}\right) dx^k \left(\sum_{t=1}^{3} \frac{\partial x^t}{\partial u^j}\frac{\partial}{\partial x^t}\right) \\
&= \sum_{i,j=1}^{2}\sum_{k=1}^{3} \alpha^i \beta^j \frac{\partial x^k}{\partial u^i}\frac{\partial x^k}{\partial u^j}.
\end{aligned}$$

But, α^i and β^j are just $du^i(X_p)$ and $du^j(Y_p)$, respectively, so we have

$$(r^*G)(X_p, Y_p) = \sum_{i,j=1}^{2}\sum_{k=1}^{3} \frac{\partial x^k}{\partial u^i}\frac{\partial x^k}{\partial u^j} du^i \otimes du^j (X_p, Y_p),$$

or, in operator form,

$$r^*G = \sum_{i,j=1}^{2}\left(\sum_{k=1}^{3} \frac{\partial x^k}{\partial u^i}\frac{\partial x^k}{\partial u^j}\right) du^i \otimes du^j. \tag{11}$$

Note, however, that, for fixed i and j, the coefficient of $du^i \otimes du^j$, given by

$$\sum_{k=1}^{3} \frac{\partial x^k}{\partial u^i}\frac{\partial x^k}{\partial u^j},$$

is just the usual dot product of the i^{th} and j^{th} columns of $D_p r$. This differential is just a matrix of real numbers, namely the values of the partial derivatives of the coordinate functions x^i with respect to the coordinates u^j at $\varphi(p)$. Hence, it makes sense to use the term dot product, if we think of the columns of the matrix as vectors in $\mathbb{R}^3$. If we denote the differential, $D_p r$ by the abbreviated form $D_p r = [r_1 \;\; r_2]$, where r_i denotes the column vector whose entries are $D_i x^j$, for $j = 1, 2, 3$, then we can rewrite equation 11 as

$$r^*G = \sum_{i,j=1}^{2} \langle r_i, r_j \rangle du^i \otimes du^j, \tag{12}$$

which is the familar expression for the **first fundamental form** of a surface. In classical differential geometry, the first fundamental form of a surface is just the inner product that is defined on the tangent plane at each point. This is nothing more than the metric tensor field over M. In other words, the Riemannian metric tensor on a two dimensional manifold (i.e. a surface) imbedded into $\mathbb{R}^3$ is the same thing as the first fundamental form of the surface. ■

Example 6.2 *The metric tensor on S^2 in spherical coordinates*

The previous example allows us to construct Riemannian metric tensors on all of the classical surfaces in differential geometry. Consider S^2, the two dimensional sphere, as a surface imbedded in $\mathbb{R}^3$, and consider the family of charts on S^2 given by the following.

1. Let $U_1 = \{(x, y, z) \in S^2 : x > 0\}$. Define $\psi_1 : U_1 \to (-\frac{\pi}{2}, \frac{\pi}{2}) \times (0, \pi)$ by $\psi_1(x, y, z) = (\tan^{-1}(\frac{y}{x}), \cos^{-1} z)$ with ${\psi_1}^{-1}(\theta, \phi) = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi)$.
2. Let $U_2 = \{(x, y, z) \in S^2 : x < 0\}$. Define $\psi_2 : U_2 \to (\frac{\pi}{2}, \frac{3\pi}{2}) \times (0, \pi)$ by $\psi_2(x, y, z) = (\pi + \tan^{-1}(\frac{y}{x}), \cos^{-1} z)$ with ${\psi_2}^{-1}(\theta, \phi) = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi)$.
3. Let $U_3 = \{(x, y, z) \in S^2 : y > 0\}$. Define $\psi_3 : U_3 \to (0, \pi) \times (0, \pi)$ by $\psi_3(x, y, z) = (\cot^{-1}(\frac{x}{y}), \cos^{-1} z)$ with ${\psi_3}^{-1}(\theta, \phi) = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi)$.
4. Let $U_4 = \{(x, y, z) \in S^2 : y < 0\}$. Define $\psi_4 : U_4 \to (\pi, 2\pi) \times (0, \pi)$ by $\psi_4(x, y, z) = (\pi + \cot^{-1}(\frac{x}{y}), \cos^{-1} z)$ with ${\psi_4}^{-1}(\theta, \phi) = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi)$.

These charts do not cover S^2, as the north and south poles are not in any of the sets U_i. The poles, however, are degenerate points in the standard spherical coordinate system, since they are not uniquely defined in terms of the parameters θ and ϕ. Nevertheless, these charts are C^∞ compatible with all other admissible charts one can define on S^2. Moreover, these charts uniquely associate to every point in $S^2 - \{(0, 0, 1), (0, 0, -1)\}$ a point in $[0, 2\pi) \times (0, \pi)$. Consequently, we can consider the mapping $s : (0, 2\pi) \times (0, \pi) \to \mathbb{R}^3$ defined by

$$s(\theta, \phi) = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi).$$

This is the coordinate representation of a parametrization of an open submanifold of S^2 into $\mathbb{R}^3$. The differential of this map at (θ, ϕ) is the linear map given by

$$Ds(\theta, \phi) = \begin{bmatrix} -\sin\phi\sin\theta & \cos\phi\cos\theta \\ \sin\phi\cos\theta & \cos\phi\sin\theta \\ 0 & -\sin\phi \end{bmatrix},$$

and this map has rank 2 for all (θ, ϕ) in $(0, 2\pi) \times (0, \pi)$. Hence, s is an immersion. In fact, since we have restricted the domain, it is an imbedding. Using the result derived in the previous example, we see that the metric tensor on S^2 in the spherical coordinate system, which we will denote by G, is given by

$$\begin{aligned} G &= (\sin^2\phi\sin^2\theta + \sin^2\phi\cos^2\theta)d\theta \otimes d\theta \\ &\quad + 2(-\sin\phi\sin\theta\cos\phi\cos\theta + \sin\phi\sin\theta\cos\phi\cos\theta)d\theta \otimes d\phi \\ &\quad + (\cos^2\phi\cos^2\theta + \cos^2\phi\sin^2\theta + \sin^2\phi)d\phi \otimes d\phi \\ &= \sin^2\phi d\theta \otimes d\theta + d\phi \otimes d\phi. \end{aligned}$$

This is the first fundamental form given for the sphere in classical geometry. ■

Example 6.3 *The infinite cylinder in cylindrical coordinates*

Consider the set $C^2 = \{(x, y, z) \in \mathbb{R}^3 \ : \ x^2 + y^2 = 1\}$. This is the infinite cylinder of radius 1 centered around the z-axis in $\mathbb{R}^3$. As a manifold, one can easily define C^2 to be the product manifold $S^1 \times \mathbb{R}$. The standard family of charts, though, in cylindrical coordinates is given by the following.

1. Let $U_1 = \{(x, y, z) \in C^2 \ : \ x > 0\}$ and let $\varphi_1 : U_1 \to (-\frac{\pi}{2}, \frac{\pi}{2}) \times \mathbb{R}$ be defined by $\varphi_1(x, y, z) = (\tan^{-1}(\frac{y}{x}), z)$.

2. Let $U_2 = \{(x, y, z) \in C^2 \ : \ x < 0\}$ and let $\varphi_2 : U_2 \to (\frac{\pi}{2}, \frac{3\pi}{2}) \times \mathbb{R}$ be defined by $\varphi_2(x, y, z) = (\pi + \tan^{-1}(\frac{y}{x}), z)$.

3. Let $U_3 = \{(x, y, z) \in C^2 \ : \ y > 0\}$ and let $\varphi_3 : U_3 \to (0, \pi) \times \mathbb{R}$ be defined by $\varphi_3(x, y, z) = (\cot^{-1}(\frac{x}{y}), z)$.

4. Let $U_4 = \{(x, y, z) \in C^2 \ : \ y < 0\}$ and let $\varphi_4 : U_4 \to (\pi, 2\pi) \times \mathbb{R}$ be defined by $\varphi_4(x, y, z) = (\pi + \cot^{-1}(\frac{x}{y}), z)$.

The map $c : (0, 2\pi) \times \mathbb{R} \to \mathbb{R}^3$ defined by $c(\theta, z) = (\cos\theta, \sin\theta, z)$ is a parametrization of an open submanifold of C^2, and, by symmetry, we could use this same map to cover any portion of C^2. The differential of this map at a point (θ, z) is given by

$$Dc(\theta, z) = \begin{bmatrix} -\sin\theta & 0 \\ \cos\theta & 0 \\ 0 & 1 \end{bmatrix}.$$

The rank of this mapping is 2 for all $(\theta, z) \in (0, 2\pi) \times \mathbb{R}$. Hence, this is an immersion. Using our derivation in Example 6.1, we see that the standard Riemannian metric tensor on C^2, denoted by G, is given by

$$\begin{aligned} G &= (\sin^2\theta + \cos^2\theta) d\theta \otimes d\theta + dz \otimes dz \\ &= d\theta \otimes d\theta + dz \otimes dz. \end{aligned}$$

This is the first fundamental form of the cylinder in classical geometry. ■

6.3 The Riemannian Metric

Having constructed the Riemannian metric tensor, we will now apply this structure to construct an actual metric on our smooth manifold. We know that manifolds are always metrizable from our results in chapter 2. However, any metric space will always have infinitely many possible metrics defined on it. Thus, the question naturally arises as to whether there is a metric that is more reasonable than others with respect to the study of geometry and physics. As we have said several times already, no means of measurement can be preferred over another. Nevertheless, falling back on our mathematical and physical intuition, there is a metric that seems to arise "naturally" in these contexts.

Consider measuring the distance between two points on the 2-dimensional sphere, S^2, where we picture this surface as an imbedded subset of $\mathbb{R}^3$. We would not measure the distance between these two points by drawing a straight (in the Euclidean sense) line between them. This line does not even lie on the space we are considering. Instead, we would measure the length of curves on the sphere connecting these two points. Consequently, it seems natural to *define* the distance between these points to be the length of the curve of shortest arclength connecting them. Note that we used the word 'define' here for a significant reason. The notion of distance is not something that is provided for us *a priori*, as was thought in Euclid's geometry. We must define what we mean by distance in terms of the metric we choose.

This is the method we will discuss in this section. We will define a metric on M in terms of the lengths of curves between two points. Furthermore, we will show that the topology induced by this metric agrees with the original manifold topology, lending more credibility to our choice of metric as being natural in some sense. There are two restrictions on this construction we must point out, though.

First, we have not yet required our manifolds to be connected. We must do so to define this metric. If M is not connected, there will be points on M that cannot be connected by any curve. This is necessary for our definition. Second, in our example of the sphere, we mentioned the curve of shortest arclength connecting two points on the surface. Such a curve need not exist, but the arclengths will always be bounded below by 0. Thus, our definition will be in terms of the infimum of the arclengths of all curves connecting the two points. There need not be any curve that actually attains this infimum. In fact, there is a famous result in differential geometry called the Hopf-Rinow Theorem that gives necessary and sufficient conditions for a manifold to have the property there is always a curve between any two points that attains the minimum arclength of curves between them. We will not prove this result here, but we will mention it again in our conclusion.

To begin our construction of the metric, let M be a Riemannian manifold with Riemannian metric tensor, G. A curve connecting two points, p and q, is a smooth map $\gamma : [a, b] \to M$ such that $\gamma(a) = p$ and $\gamma(p) = b$. The domain of γ could, of course be an open or half open interval as well. We will simply use closed intervals throughout for consistency of notation. Recalling our previous discussion of the tangent vectors of a curve, we know that the tangent vector of γ at $p = \gamma(t_0)$, for some $t_0 \in [a, b]$, is given by $D_{t_0}\gamma(d/dt)$. (If $t_0 = a$ or b, we will assume that γ can be smoothly extended so that the tangent vectors at the endpoints may be well-defined.) We will abuse our notation slightly and denote this tangent vector by the more intuitive form

$$D_{t_0}\gamma\Big(\frac{d}{dt}\Big) = \frac{d\gamma}{dt},$$

and the particular point to which we are referring should be clear from the context. We define the *length of the curve*, γ, to be the real number

$$L = \int_a^b \left(G\Big(\frac{d\gamma}{dt}, \frac{d\gamma}{dt}\Big) \right)^{1/2} dt.$$

First, note that this is nothing more than an integral of a real-valued function on $[a, b]$. For each $t \in [a, b]$, $d\gamma/dt$ is just a tangent vector in $T_{\gamma(t)}M$. The metric tensor, G, then maps $T_pM \times T_pM$ to $\mathbb{R}$. Moreover, since G and γ are both smooth, this is a smooth function, so its integral is defined just as in the classical sense of the Riemann integral. Taking the square root is always valid, as G is positive definite. Hence, this integral does make sense. Observe, also, that this is just a generalization of the classical definition of arclength involving the inner product of the tangent vectors with themselves.

Next, we want to point out that the value of this integral is independent of the parametrization of γ. This just follows from the chain rule as we have defined it on manifolds. Suppose $\alpha : [c,d] \to [a,b]$ is a smooth injective map. Then we can think of γ as being a function on $[c,d]$ of the form $\gamma \circ \alpha(s)$, for $s \in [c,d]$. By the chain rule, we have, for any $f \in C^\infty(p)$,

$$\begin{aligned} D_{s_0}(\gamma \circ \alpha)\Big(\frac{d}{ds}\Big) f &= \frac{d}{ds}\big(f \circ \gamma \circ \alpha\big)\Big|_{s_0} \\ &= \frac{d}{dt}\big(f \circ \gamma\big)\Big|_{\alpha(s_0)} \frac{d\alpha}{ds}\Big|_{s_0} \\ &= \frac{d\alpha}{ds}\Big|_{s_0} D_{\alpha(s_0)}\gamma\Big(\frac{d}{dt}\Big) f. \end{aligned}$$

If we write this equality in operator form, for a general point $s \in [c,d]$, and using our current notation, we have the familiar form

$$\frac{d\gamma}{ds} = \frac{d\gamma}{dt}\frac{d\alpha}{ds}.$$

Now, considering the length of γ expressed as a function of s, we have

$$
\begin{aligned}
\int_c^d \left(G\Big(\frac{d\gamma}{ds}, \frac{d\gamma}{ds}\Big)\right)^{1/2} ds &= \int_c^d \left(G\Big(\frac{d\gamma}{dt}\frac{d\alpha}{ds}, \frac{d\gamma}{dt}\frac{d\alpha}{ds}\Big)\right)^{1/2} ds \\
&= \int_c^d \left(\Big(\frac{d\alpha}{ds}\Big)^2 G\Big(\frac{d\gamma}{dt}, \frac{d\gamma}{dt}\Big)\right)^{1/2} ds \\
&= \int_a^b \left(G\Big(\frac{d\gamma}{dt}, \frac{d\gamma}{dt}\Big)\right)^{1/2} \frac{d\alpha}{ds}\frac{d\alpha^{-1}}{dt} dt \\
&= \int_a^b \left(G\Big(\frac{d\gamma}{dt}, \frac{d\gamma}{dt}\Big)\right)^{1/2} \frac{d(\alpha \circ \alpha^{-1})}{dt} dt \\
&= \int_a^b \left(G\Big(\frac{d\gamma}{dt}, \frac{d\gamma}{dt}\Big)\right)^{1/2} dt.
\end{aligned}
$$

This shows that the length of the curve is a true geometric property, meaning that it depends only on the set of points making up the curve and not the particular parametrization.

This independence of parametrization allows us to define the *arclength function* for the curve γ. For $t \in [a, b]$, we see that the arclength of γ between $\gamma(a)$ and $\gamma(t)$ is given by

$$
L(t) = \int_a^t \left(G\Big(\frac{d\gamma}{du}, \frac{d\gamma}{du}\Big)\right)^{1/2} du.
$$

As in classical differential geometry, this is a continuous strictly increasing function of t. Hence, it defines a new paramter, s, called the arclength parameter. If we let $s(t) = L(t)$, then we can, in principle, at least, change between the parameter t and the arclength parameter s. By the definition of $s(t)$, we see that

$$
\Big(\frac{ds}{dt}\Big)^2 = G\Big(\frac{d\gamma}{dt}, \frac{d\gamma}{dt}\Big).
$$

This is just one expression for the *element of arclength* defined by the metric tensor G.

Next, consider what the arclength integral looks like in a local coordinate system. That is, suppose that $\gamma([a, b])$ lies within a single coordinate system (φ, U). Let $\{E_i\}$ denote the basis field induced by this coordinate system on the tangent spaces over U, and denote the coordinate functions of φ by x^i. Then, for any $p = \gamma(t) \in U \cap \gamma([a,b])$, we can represent the tangent vector $d\gamma/dt$ at p as we discussed in Chapter 2 by the expression

$$\frac{d\gamma}{dt} = \sum_{i=1}^{m} \dot{x}^i(t) E_{p_i},$$

where $\dot{x}^i(t)$ simply represents the ordinary derivative of the coordinate function $x^i \circ \gamma$ at t. Hence, for any point $p = \gamma(t) \in U \cap \gamma([a,b])$, we have

$$\begin{aligned} G_p\Big(\frac{d\gamma}{dt}, \frac{d\gamma}{dt}\Big) &= G_p\left(\sum_{i=1}^{m} \dot{x}^i(t) E_{p_i}, \sum_{j=1}^{m} \dot{x}^j(t) E_{p_j}\right) \\ &= \sum_{i,j=1}^{m} \dot{x}^i(t)\dot{x}^j(t) G_p\Big(E_{p_i}, E_{p_j}\Big) \\ &= \sum_{i,j=1}^{m} g_{ij}(p)\dot{x}^i(t)\dot{x}^j(t). \end{aligned}$$

Thus, the arclength integral, in this coordinate system, can be expressed as

$$s(t) = \int_a^t \left(\sum_{i,j=1}^{m} g_{ij}(\gamma(t)) \frac{d(x^i \circ \gamma)}{dt} \frac{d(x^j \circ \gamma)}{dt} \right)^{1/2} dt. \qquad (13)$$

Now, the metric tensor components, g_{ij}, are real-valued functions on U and, thus, on $U \cap \gamma([a,b])$. However, because the coordinate map, φ, is a homeomorphism, we can uniquely associate each point $\gamma(t)$ with its coordinates $\varphi \circ \gamma(t) = (x^1(t), \ldots, x^m(t))$. In this way, we can think of g_{ij} as being a function of the local coordinates of the curve. That is, we will implicitly identify g_{ij} with its coordinate representation $g_{ij} \circ \varphi^{-1}$. Consequently, we can rewrite the previous equation in a form that will be useful to us, namely

$$s(t) = \int_a^t \left(\sum_{i,j=1}^{m} g_{ij}(x(t)) \frac{d(x^i \circ \gamma)}{dt} \frac{d(x^j \circ \gamma)}{dt} \right)^{1/2} dt, \qquad (14)$$

where $g_{ij}(x(t))$ denotes g_{ij} as a function of all the coordinate functions $x^i(t)$.

Now, from here on, we will assume that M is connected. Let p and q be two fixed points in M. Let $\Gamma_{p,q}$ be the set of all piecewise smooth curves connecting p and q.

Thus, $\Gamma_{p,q}$ is the set of curves, $\gamma : [a, b] \to M$ such that $\gamma(a) = p$, $\gamma(b) = q$, and there is a partition of $[a, b]$, $a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$, such that $\gamma(t_i) = \gamma(t_{i+1})$ for $1 \leq i \leq n-2$ and the curve $\gamma : [t_{i-1}, t_i] \to M$ is smooth for $1 \leq i \leq n$. For a curve, $\gamma \in \Gamma_{p,q}$, let L_γ denote the arclength of the curve as we have defined it. Define a function $d : M \times M \to \mathbb{R}$ by

$$d(p, q) = \inf\{L_\gamma \ : \ \gamma \in \Gamma_{p,q}\}.$$

First, note that the function d is well-defined. For any curve, there will be, at most, a finite number of points where the integrand of the arclength integral is not differentiable. It will always be continuous, though, so the integral is well-defined. Moreover, since G is positive definite, the integrand in the arclength function is always nonnegative, so the integral is bounded below by zero. Hence the infimum always exists and is nonegative. The symmetry and the triangle inequality are easily verified. The positive definiteness is more difficult, and will take some work.

Every curve connecting p and q can be reparametrized, reversing the orientation, as a curve connecting q and p. Thus, there is a one-to-one correspondence between curves in $\Gamma_{p,q}$ and curves in $\Gamma_{q,p}$. Moreover, simply reversing the orientation of a curve is just a reparametrization, and we have shown that the arclength is invariant under reparametrizations. Thus, the sets $\Gamma_{p,q}$ and $\Gamma_{q,p}$ contain the same real numbers and, thus, have the same infimum. It follows, then, that $d(p, q) = d(q, p)$, and d is symmetric.

Now, suppose p, q, and r are three distinct points in M. Suppose we have $\gamma_1 \in \Gamma_{p,q}$ and $\gamma_2 \in \Gamma_{q,r}$, where γ_1 is defined on $[a, b]$ and γ_2 is defined on $[c, d]$. We can always find a suitable translation so that γ_2 is defined on $[b, b + (d - c)]$. Then the curve $\gamma : [a, b + (d - c)] \to M$ defined by

$$\gamma(t) = \begin{cases} \gamma_1(t) & a \leq t \leq b \\ \gamma_2(t) & b \leq t \leq b + (d - c) \end{cases}$$

is piecewise smooth and satisfies $\gamma(a) = p$, $\gamma(b) = \gamma_1(b) = q = \gamma_2(b)$, and $\gamma(b + (d - c)) = r$. Thus, $\gamma \in \Gamma_{p,r}$, and, by the summation properties of the Riemann integral, we have $L_\gamma = L_{\gamma_1} + L_{\gamma_2}$. It follows that $d(p, r) \leq L_\gamma \Rightarrow d(p, r) \leq L_{\gamma_1} + L_{\gamma_2}$. But $\gamma_1 \in \Gamma_{p,q}$ and $\gamma_2 \in \Gamma_{q,r}$ were arbitrary, so this implies that $d(p, r) \leq \inf\{L_{\gamma_1} + L_{\gamma_2} \ : \ \gamma_1 \in \Gamma_{p,q},\ \gamma_2 \in \Gamma_{q,r}\}$.

Now, let $u_1 = \inf\{L_\gamma \ : \ \gamma \in \Gamma_{p,q}\}$, $u_2 = \inf\{L_\gamma \ : \ \gamma \in \Gamma_{q,r}\}$, and let $\epsilon > 0$ be given. Then there is some $\gamma_1 \in \Gamma_{p,q}$ and some $\gamma_2 \in \Gamma_{q,r}$ such that $u_1 \leq L_{\gamma_1} < u_1 + \frac{\epsilon}{2}$ and $u_2 \leq L_{\gamma_2} < u_2 + \frac{\epsilon}{2}$. Thus, it follows that $u_1 + u_2 \leq L_{\gamma_1} + L_{\gamma_2} < u_1 + u_2 + \epsilon$. Since ϵ was arbitrary, this shows that $\inf\{L_{\gamma_1} + L_{\gamma_2} \ : \ \gamma_1 \in \Gamma_{p,q},\ \gamma_2 \in \Gamma_{q,r}\} = \inf\{L_\gamma \ : \ \gamma \in \Gamma_{p,q}\} + \inf\{L_\gamma \ : \ \gamma \in \Gamma_{q,r}\} = d(p, q) + d(q, r)$. Hence, we see that

$$d(p,r) \leq d(p,q) + d(q,r),$$

and d satisfies the triangle inequality.

Next, suppose $p = q$. Then p and q, obviously, lie in some coordinate system, (φ, U) and $\varphi(p) = \varphi(q)$. Let $\epsilon > 0$ be given. There is an open ball of radius, $\delta > 0$, centered at $\varphi(p)$ and contained in $\varphi(U)$, and we can assume without loss of generality that $\delta < \epsilon$. Denote this ball by $B_\delta(\varphi(p))$ and let $V = \varphi^{-1}(B_\delta(\varphi(p)))$. On the coordinate neighborhood, U, the components of the metric tensor, G, are smooth functions $g_{ij} : U \to \mathbb{R}$, for $1 \leq i, j \leq m$. Hence, the functions $g_{ij} \circ \varphi^{-1}$ are smooth functions on $\varphi(U) \subset \mathbb{R}^m$. Since Euclidean spaces are normal spaces, there is a neighborhood, W, of $\varphi(p)$ such that $\overline{W}$ is compact and $B_\delta(\varphi(p)) \subset W \subset \overline{W} \subset \varphi(U)$. Hence, the functions $g_{ij} \circ \varphi^{-1}$ are smooth on the compact set $\overline{W}$. So, for any i, j, there is a positive real number, M_{ij}, such that $|g_{ij} \circ \varphi^{-1}(x)| \leq M_{ij}$ for all $x \in \overline{W}$. Let $M = \max_{i,j} M_{ij}$, and assume also that $M > 1$. Now, assuming, temporarily, that the dimension of M is at least two, define a curve, λ in $\varphi(U) \subset \mathbb{R}^m$ as follows. Let r be a positive real number such that $0 < r < \min\{\frac{\delta}{4\pi\sqrt{M}}, \frac{\delta}{\sqrt{2}}\}$, and define λ by

$$\lambda(t) = \big(r\cos t - r + (\varphi(p))^1, r\sin t + (\varphi(p))^2, (\varphi(p))^3, \ldots, (\varphi(p))^m\big), \qquad t \in [0, 2\pi].$$

Then, for any $t \in [0, 2\pi]$, we have

$$\begin{aligned}
\|\lambda(t) - \varphi(p)\| &= \sqrt{(r(\cos t - 1))^2 + (r\sin t)^2} \\
&= \sqrt{r^2(\cos^2 t + \sin^2 t - 2\cos^2 t + 1)} \\
&= r\sqrt{2 - 2\cos^2 t} \\
&= r\sqrt{2}|\sin t| \\
&\leq r\sqrt{2} \\
&< \delta,
\end{aligned}$$

which implies that $\lambda(t) \in B_\delta(\varphi(p))$. Hence, the curve λ lies inside this ball, and we can define a curve in $V \subset U$ by $\gamma(t) = \varphi^{-1} \circ \lambda(t)$. The length of this curve can be calculated using its local coordinate representation on U by

$$\begin{aligned}
L_\gamma &= \int_0^{2\pi}\left(\sum_{i,j=1}^{m} g_{ij}(\gamma(t))\frac{d(x^i\circ\gamma)}{dt}\frac{d(x^j\circ\gamma)}{dt}\right)^{1/2}dt\\
&= \int_0^{2\pi}\left(\sum_{i,j=1}^{m} (g_{ij}\circ\varphi^{-1})(\lambda(t))\frac{d(x^i\circ\varphi^{-1}\circ\lambda)}{dt}\frac{d(x^j\circ\varphi^{-1}\circ\lambda)}{dt}\right)^{1/2}dt\\
&= \int_0^{2\pi}\left(\sum_{i,j=1}^{m} (g_{ij}\circ\varphi^{-1})(\lambda(t))\frac{d\lambda^i}{dt}\frac{d\lambda^j}{dt}\right)^{1/2}dt.
\end{aligned}$$

But, we also have

$$\begin{aligned}
\sum_{i,j=1}^{m}(g_{ij}\circ\varphi^{-1})(\lambda(t))\frac{d\lambda^i}{dt}\frac{d\lambda^j}{dt} &\le \sum_{i,j=1}^{m}|(g_{ij}\circ\varphi^{-1})(\lambda(t))|\left|\frac{d\lambda^i}{dt}\right|\left|\frac{d\lambda^j}{dt}\right|\\
&\le M\sum_{i,j=1}^{m}\left|\frac{d\lambda^i}{dt}\right|\left|\frac{d\lambda^j}{dt}\right|\\
&\le M\Big(r^2\sin^2 t+2r^2\sin t\cos t+r^2\cos^2 t\Big),
\end{aligned}$$

and this implies that

$$\begin{aligned}
L_\gamma &\le \int_0^{2\pi}\Big(Mr^2(\sin^2 t+2\sin t\cos t+\cos^2 t)\Big)^{1/2}dt\\
&\le r\sqrt{M}\int_0^{2\pi}(\sin^2 t+2\sin t\cos t+\cos^2 t)^{1/2}dt\\
&\le r\sqrt{M}\int_0^{2\pi}|\sin t+\cos t|dt\\
&\le 2r\sqrt{M}\int_0^{2\pi}dt\\
&\le 4\pi r\sqrt{M}.
\end{aligned}$$

Hence, we see that $L_\gamma < \delta < \epsilon$. Since ϵ was arbitrary, it follows that we can find curves from p to p of arbitrarily small arclength, implying that $d(p,p)$ must equal 0.

Now, suppose M is 1-dimensional, and let p be a point in M, with some coordinate system, (φ, U), around p. If G is the Riemannian metric tensor on M, then, on U,

G takes the simple form $G_p(X_p, Y_p) = g(p)dx \otimes dx$, where we denote the single coordinate function of φ by x. Now, $\varphi(U)$ is an open subset of $\mathbb{R}$ containing $\varphi(p)$. As we have done, we can assume that $\varphi(p) = 0$. So, we can find a closed interval, I, that is contained in $\varphi(U)$, centered at $0 = \varphi(p)$, and is symmetric about 0. Hence, the the function $g \circ \varphi^{-1} : \varphi(U) \to \mathbb{R}$ is bounded by some positive real number M on I, and we can assume that $M > 1$. Given $\epsilon > 0$, let δ be a positive real number such that $\delta < \epsilon$ and the interval $(-\delta, \delta)$ is contained in I. Then the function $g \circ \varphi^{-1}$ is bounded by M on $(-\delta, \delta)$. Let r be a positive real number such that $r < \frac{\delta}{2\pi\sqrt{M}}$. Define the function $\lambda : [0, r] \to \mathbb{R}$ by $\lambda(t) = r\sin(\frac{2\pi}{r}t)$. Then λ is actually a curve in $\mathbb{R}$ connecting $\lambda(0) = 0 = \varphi(p)$ and $\lambda(r) = 0 = \varphi(p)$. Moreover, for any t, $\lambda(t) \leq r$, so this curve lies in the interval $(-\delta, \delta) \subset I$. So, we can define a curve in $\varphi^{-1}(I) \subset U$ by $\gamma(t) = (\varphi^{-1} \circ \lambda)(t)$. The arclength of this curve is given by

$$\begin{aligned}
L_\gamma &= \int_0^r \left((g \circ \gamma)(t) \frac{d\lambda}{dt} \frac{d\lambda}{dt} \right)^{1/2} dt \\
&= \int_0^r \left((g \circ \varphi^{-1} \circ \lambda)(t) \left(\frac{d\lambda}{dt} \right)^2 \right)^{1/2} dt \\
&\leq \sqrt{M} \int_0^r \left| 2\pi \cos\left(\frac{2\pi}{r} t \right) \right| dt \\
&\leq 2\sqrt{M}\pi \int_0^r \left| \cos\left(\frac{2\pi}{r} t \right) \right| dt \\
&\leq 2\pi\sqrt{M} r.
\end{aligned}$$

Thus, we have $L_\gamma < \delta < \epsilon$. Since ϵ was arbitrary, we can find curves from p to p of arbitrarily small arclength. So, in all cases, we must have $d(p, p) = 0$.

Thus, the only remaining metric property to verify is the fact that $d(p, q) = 0 \Rightarrow p = q$. To do this, we will have to derive some preliminary results. We will assume the point p to be fixed, and let (φ, U) be a coordinate system around p. As we have done before, we can assume, by an appropriate translation if necessary, that $\varphi(p) = 0 \in \mathbb{R}^m$. Then $\varphi(U)$ is a neighborhood of $0 \in \mathbb{R}^m$. Let δ be a fixed, positive real number such that $\overline{B_\delta(0)} \subset \varphi(U)$. As usual, we will denote the coordinate functions of φ by x^i. So, we can express the metric tensor, G, on U by $G = \sum_{i,j} g_{ij} dx^i \otimes dx^j$, and we will think of the functions g_{ij} in terms of their coordinate representations. That is, g_{ij} will denote the function $g_{ij} \circ \varphi^{-1}$, the function of the local coordinates, $g_{ij}(x^1, \ldots, x^m)$.

Now, for each $r \leq \delta$, define $K_r \subset \mathbb{R}^m \times \mathbb{R}^m$ by $K_r = \{(x, \alpha) : \|x\| \leq r, \|\alpha\| = 1\}$. For each $r \leq \delta$, K_r is a compact subset of $\mathbb{R}^m \times \mathbb{R}^m$. The function $F_r : K_r \to \mathbb{R}$, defined by

$$F_r(x, \alpha) = \left(\sum_{i,j=1}^{m} g_{ij}(x) \alpha^i \alpha^j \right)^{1/2}$$

is continuous on K_r. So, on K_r, it assumes a maximum and minimum value, denoted by M_r and m_r, respectively. Moreover, if we fix $x \in \overline{B_r(0)}$, then we see that $F_r(x, \alpha)$ is just the square root of the value obtained by the metric tensor, G, acting on the tangent vector in $T_{\varphi^{-1}(x)}M$ whose components are the components of $\alpha \in \mathbb{R}^m$. That is, if $p = \varphi^{-1}(x)$, and if $X_p \in T_pM$ has coordinate representation $X_p = \sum_i \alpha^i E_{p_i}$, then

$$F_r(x, \alpha) = \Big(G_p(X_p, X_p) \Big)^{1/2}.$$

Since $\alpha \neq 0$, this quantity must be positive. Thus, for each $r \leq \delta$, the function F_r must be strictly positive on K_r. Since F_r attains its minimum value, it follows that $m_r > 0$ for all $r \leq \delta$. Also, note that F_r, for $r < \delta$, is just $F_\delta|_{K_r}$.

Now, it follows that for any $(x, \alpha) \in K_r$, we have

$$0 < m_\delta \leq m_r \leq \left(\sum_{i,j=1}^{m} g_{ij}(x) \alpha^i \alpha^j \right)^{1/2} \leq M_r \leq M_\delta.$$

If $(\beta^1, \ldots, \beta^m)$ is any nonzero vector in $\mathbb{R}^m$, and if $b = \|(\beta^1, \ldots, \beta^m)\| \neq 0$, then

$$\sum_{i=1}^{m} \left(\frac{\beta^i}{b} \right)^2 = 1.$$

So, for $x \in \overline{B_r(0)}$, we see that

$$0 < m_\delta \leq m_r \leq \left(\sum_{i,j=1}^{m} g_{ij}(x) \frac{\beta^i}{b} \frac{\beta^j}{b} \right)^{1/2} \leq M_r \leq M_\delta,$$

which implies that

$$0 < m_\delta b \leq m_r b \leq \left(\sum_{i,j=1}^{m} g_{ij}(x) \beta^i \beta^j \right)^{1/2} \leq M_r b \leq M_\delta b$$

for all $x \in \overline{B_r(0)}$ and $(\beta^1, \ldots, \beta^m) \in \mathbb{R}^m$.

Now, the curve in $\mathbb{R}^m$ that minimizes the distance between two points $x, y \in \mathbb{R}^m$ is the straight line connecting them. In other words, for $x, y \in \mathbb{R}^m$, we have $d(x, y) = \|x - y\|$. This seems intuitively obvious, but if we ignore our Euclidean intuition, take $\mathbb{R}^m$ to be a manifold, and define a metric on it using the arclength of curves, the result is not at all trivial. Nevertheless, the proof is straightforward, especially after having gone through our current work. We will assume this result here, so that we may focus more on the purpose at hand. For the sake of completeness, however, we have given a proof of this result in the appendix.

Next, we fix $r \leq \delta$. Let $\gamma : [c, d] \to M$ be a piecewise smooth curve between $p = \gamma(c)$ and $q = \gamma(d)$ such that $\gamma([c, d])$ lies entirely within $\varphi^{-1}(\overline{B_r(0)} \subset U$. The length of this curve is then given by

$$L_\gamma = \int_c^d \left(\sum_{i,j=1}^{m} g_{i,j}(x(t))\dot{x}^i(t)\dot{x}^j(t) \right)^{1/2},$$

where $\dot{x}^i(t)$ denotes the i^{th} component of the tangent vector $d\gamma/dt$ at t. Since $\gamma(t)$ is regular, the tangent vector is never 0. Thus, $\sum_i (\dot{x}^i)^2 \neq 0$. So, for a fixed $t \in [c, d]$, we can let the numbers $\dot{x}^i(t)$ act as the vector $(\beta^1, \ldots, \beta^m)$ above. This yields

$$0 < m_\delta \left(\sum_{i,j=1}^{m} \left(\dot{x}^i(t)\right)^2 \right)^{1/2} \leq m_r \left(\sum_{i,j=1}^{m} \left(\dot{x}^i(t)\right)^2 \right)^{1/2} \leq \left(\sum_{i,j=1}^{m} g_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t) \right)^{1/2}$$

and

$$\left(\sum_{i,j=1}^{m} g_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t) \right)^{1/2} \leq M_r \left(\sum_{i,j=1}^{m} \left(\dot{x}^i(t)\right)^2 \right)^{1/2} \leq M_\delta \left(\sum_{i,j=1}^{m} \left(\dot{x}^i(t)\right)^2 \right)^{1/2}.$$

These inequalities hold for each $t \in [c, d]$, so we have

$$0 < m_\delta \int_c^d \|\varphi \circ \gamma(t)\| dt \leq m_r \int_c^d \|\varphi \circ \gamma(t)\| dt \leq \int_c^d \left(\sum_{i,j=1}^{m} g_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t) \right)^{1/2} dt$$

and

$$\int_c^d \left(\sum_{i,j=1}^m g_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t) \right)^{1/2} dt \le M_r \int_c^d \|\varphi \circ \gamma(t)\| dt \le M_\delta \int_c^d \|\varphi \circ \gamma(t)\| dt.$$

But, since $\varphi \circ \gamma$ is a curve in $\mathbb{R}^m$ connecting $\varphi(\gamma(c)) = \varphi(p) = 0$ and $\varphi(\gamma(d)) = \varphi(q)$, it follows that

$$\int_c^d \|\varphi \circ \gamma(t)\| dt \ge \|\varphi(q) - \varphi(p)\| = \|\varphi(q)\|.$$

Thus, we have

$$0 < m_\delta \|\varphi(q)\| \le m_r \|\varphi(q)\| \le L \le M_r \int_c^d \|\varphi \circ \gamma(t)\| dt \le M_\delta \int_c^d \|\varphi \circ \gamma(t)\| dt. \quad (15)$$

This holds for any curve, γ, lying in $\varphi^{-1}(\overline{B_r(0)}$ connecting $p = \varphi^{-1}(0)$ and a point $q \in \varphi^{-1}(\overline{B_r(0)})$. Moreover, $r \le \delta$ was arbitrary, so we can repeat this construction for any $\overline{B_r(0)}$.

Now, suppose that q' is a point of M distinct from p. Then, for some $r \le \delta$, q' lies outside of $\varphi^{-1}(B_r(0))^*$. Let $\gamma : [0, c] \to M$ be a piecewise smooth curve connecting $p = \gamma(0) = \varphi^{-1}(0)$ and $q' = \gamma(c)$, and denote the length of this curve by L_γ. At least a portion of this curve must lie in U, so choose $\lambda \in (0, c]$ such that $\gamma(t)$, for $0 \le t \le \lambda$, lies in U but $\gamma(\lambda)$ is outside of $\varphi^{-1}(B_r(0))$. Such a value, λ, must exist, because $q' \notin \varphi^{-1}(B_r(0))$. Then $(\varphi \circ \gamma)(t) = (x^1(t), \ldots, x^m(t))$ is the local coordinate representation of γ for $0 \le t \le \lambda$. So, $\varphi \circ \gamma$ connects $0 \in \mathbb{R}^m$ and the point $\varphi(\gamma(\lambda)) \in \mathbb{R}^m$. Also, since $\gamma(\lambda) \notin \varphi^{-1}(B_r(0))$, we have $(\varphi \circ \gamma)(\lambda) \notin B_r(0)$.

Now, consider the function $t \mapsto \|\varphi \circ \gamma(t)\|$. This is a continuous function on the compact set $[0, \lambda]$, and $\|\varphi \circ \gamma(0)\| = 0$ and $\|\varphi \circ \gamma(\lambda)\| \ge r$. So, by the intermediate value theorem, there is some $t_0 \in (0, \lambda]$ such that $\|\varphi \circ \gamma(t_0)\| = r$. Moreover, we can assume that t_0 is the smallest such value, meaning that $\|\varphi \circ \gamma(t)\| < r$ for $t \in [0, t_0)$ and $\|\varphi \circ \gamma(t_0)\| = r$. Such an assumption is valid because $\|\varphi \circ \gamma\|^{-1}(r)$ is a closed subset of $[0, \lambda]$.

So, t_0 is the smallest value for which $\varphi \circ \gamma(t_0) \notin B_r(0)$, but $\varphi \circ \gamma(t_0) \in \overline{B_r(0)}$. Thus, $\gamma(t_0) \in \varphi^{-1}(\overline{B_r(0)})$, and $\gamma(t_0)$ is the first point of the curve γ that lies outside of $\varphi^{-1}(B_r(0))$. That is, $\gamma(t_0)$ is the first point of the curve such that $\|\varphi(\gamma(t_0))\| = r$. Let $q = \gamma(t_0)$. Let L'_γ denote the length of the curve $\gamma(t)$, $t \in [0, t_0]$. Then $L'_\gamma \le L_\gamma$. Moreover, the curve $\gamma(t)$, with t restricted to $[0, t_0]$, lies in $\varphi^{-1}(\overline{B_r(0)})$. So the inequalities in (15) hold for this curve, implying that

$$0 < m_\delta \|\varphi(q)\| \leq m_r \|\varphi(q)\| \leq L'_\gamma \leq L_\gamma$$

where $q = \gamma(t_0) \Rightarrow \|\varphi(q)\| = \|\varphi(\gamma(t_0))\| = r$. Thus $L_\gamma \geq L'_\gamma \geq m_\delta r > 0$. Now, the curve $\gamma : [0, c] \to M$, with length L_γ, and connecting $\gamma(0) = p$ and $\gamma(c) = q'$, was chosen arbitrarily. Moreover, the value of r does not depend on the particular curve. It only depends on q'. So, $m_\delta r$ is a lower bound for $\Gamma_{p,q'}$. Thus, $d(p, q') > 0$, and this shows that $p \neq q' \Rightarrow d(p, q') > 0$. Hence, d is positive definite on M, and we have shown the following.

Theorem 6.5 *Let M be a connected Riemannian manifold with metric tensor G. Then the function $d : M \times M \to \mathbb{R}$ defined by*

$$d(p, q) = \inf\{L_\gamma \ : \ \gamma \in \Gamma_{p,q}\}$$

defines a metric on M.

To conclude this chapter, we will show that this metric is natural in the sense that its induced topology is equivalent to the original manifold topology. Thus, in the case of Riemannian manifolds, there is a distinct connection between the original manifold structure and the smooth metric structure we add to it.

Theorem 6.6 *The topology induced on a Riemmanian manifold, M, by the metric d is equivalent to the original topology on M.*

<u>Proof</u> Let W be an open subset of M relative to the original topology, and let p be a point of W. Let (φ, U) be a coordinate system around p. Then $V = U \cap W$ is a neighborhood of p and $(\varphi|_V, V)$ is a coordinate system around p. As before, we can assume that $\varphi(p) = 0 \in \mathbb{R}^m$, so that $\varphi(V)$ is a neighborhood of $0 \in \mathbb{R}^m$. Also as before, let $\delta > 0$ be a fixed real number such that $V_\delta = \varphi^{-1}(\overline{B_\delta(0)}) \subset V$.

Choose $r \in (0, \delta)$ such that $V_r = \varphi^{-1}(B_r(0)) \subset V_\delta \subset V$. Recalling what we have constructed thus far, for $r \in (0, \delta]$, let m_δ denote the lower bound corresponding to the function $F_\delta(x, \alpha)$ on $K_r = \{(x, \alpha) \ : \ \|x\| \leq r, \ \|\alpha\| = 1\}$, so that we have $m_r \geq m_\delta$ for all $r \in (0, \delta]$. Now, let $\epsilon > 0$ be such that $\epsilon < m_\delta r$, and consider $S_\epsilon(p) = \{q \in M \ : \ d(p, q) < \epsilon\}$. Let q be any point in $S_\epsilon(p)$, so that $d(p, q) < m_\delta r$. We will show that $q \in V_r$ by contradiction.

Suppose $q \notin V_r$. Then q lies outside of $\varphi^{-1}(B_r(0))$. We can go back to our previous results (starting at * on the previous page) and repeat those steps verbatim to conclude that $d(p, q) \geq m_\delta r > 0$. This contradicts the fact that $d(p, q) < m_\delta r$, so $q \in S_\epsilon(p) \subset V_r \subset V$. Thus, $S_\epsilon(p) \subset V \subset W$. That is, for every $p \in W$, there is an open ball relative to the metric topology that is centered at p and contained in W. Hence, W is open with respect to the metric topology.

To prove the converse, it suffices to show that the open ball, $S_\epsilon(p)$, for arbitrary p and $\epsilon > 0$, is open with respect to the original manifold topology.

We still assume that (φ, U) is a coordinate system around $p \in M$ and $\varphi(p) = 0$. Let $\delta > 0$ be a fixed real number such that $\overline{B_\delta(0)} \subset \varphi(U)$. Choose a real number $r > 0$ such that $r < \min\{\delta, \frac{\epsilon}{M_\delta}\}$, where M_δ is the upper bound of the function F_δ that we have defined previously. Consider the set $V_r = \varphi^{-1}(B_r(0)) \subset \varphi^{-1}(B_\delta(0)) \subset U$. Suppose $q \in V_r$. Define a curve, γ, from p to q as follows. Let $(\beta^1, \ldots, \beta^m)$ be the coordinates of q. That is, $\varphi(q) = (\beta^1, \ldots, \beta^m)$. Define $\mu : [0,1] \to \mathbb{R}^m$ by $\mu(t) = (\beta^1 t, \ldots, \beta^m t)$. Note that $\mu(0) = 0 = \varphi(p)$ and $\mu(1) = \varphi(q)$. Thus, $(\varphi^{-1} \circ \mu)(0) \in V_r$ and $(\varphi^{-1} \circ \mu)(1) \in V_r$. Moreover, for any $t \in [0,1]$, we have

$$\|\mu(t)\| = \left(\sum_{i=1}^{m} (\beta^i t)^2\right)^{1/2} = t\|\varphi(q)\| \leq \|\varphi(q)\|.$$

Since $q \in V_r$, we know that $\varphi(q) \in \varphi(V_r) = B_r(0)$, so $\|\varphi(q)\| < r$, which implies that $\|\mu(t)\| < r$ for all t. So, $\mu(t) \in B_r(0)$ for all t, from which it follows that $(\varphi^{-1} \circ \mu)(t) \in V_r$ for all t. So, let $\gamma(t) = (\varphi^{-1} \circ \mu)(t)$ for $t \in [0,1]$. The length of this curve satisfies

$$\begin{aligned}
L_\gamma &= \int_0^1 \left(\sum_{i,j=1}^{m} g_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t)\right)^{1/2} dt \\
&= \int_0^1 \left(\sum_{i,j=1}^{m} g_{ij}(t\varphi(q))\beta^i\beta^j\right)^{1/2} dt \\
&\leq M_r \int_0^1 \left(\sum_{i=1}^{m} (\beta^i)^2\right)^{1/2} dt \\
&\leq M_\delta \int_0^1 \left(\sum_{i=1}^{m} (\beta^i)^2\right)^{1/2} dt.
\end{aligned}$$

It follows that

$$L_\gamma \leq M_\delta \|\varphi(q)\| < M_\delta r.$$

But $r < \frac{\epsilon}{M_\delta} \Rightarrow M_\delta r < \epsilon$, and we see that $L_\gamma < \epsilon$. Thus, if $d(p,q) \leq L_\gamma < \epsilon$, then $q \in S_\epsilon(p)$. Hence, $V_r = \varphi^{-1}(B_r(0)) \subset S_\epsilon(p)$, and V_r is an open set relative to the original topology.

QED

7 A Brief Excursion into Lorentzian Manifolds

7.1 Development of the Lorentzian Metric

While the Riemannian metric is the natural measure of distance to consider in geometry, classical mechanics, and most dynamical systems, it is not an appropriate measure in a large portion of theoretical physics. In prerelativity physics, space and time were not well understood. With Euclidean geometry motivating a great deal of physical theories, the space that makes up the universe, both globally and locally, was thought to be a three-dimensional Euclidean space, with time as an independent parameter. Such a model worked well for Newtonian mechanics involving the motion of particles very near the earth's surface and at speeds much less than the speed of light. However, efforts to adapt Newtonian mechanics in order to model phenomena such as light, electricity, magnetism, and cosmological mechanics failed.

One of the biggest wounds to the status of Newtonian mechanics was James Maxwell's remarkable theory of electrodynamics and magnetism. Immediately after its inception, though, Maxwell's theory was questioned and dismissed by staunch classical physicists, and it continued to be a source of controversy for years. Even a young Albert Einstein, some half a century after their introduction, questioned Maxwell's results. The reason for all of this controversy, while not evident at the time, was that acceptance of Maxwell's theories required a paradigm shift in theoretical physics as a whole. The problem was that Maxwell's theories were not consistent with Newtonian mechanics. In particular, the core of Maxwell's theory, his now famous fundamental equations of electromagnetism, were not invariant in form under the standard coordinate transformations used in Newtonian mechanics. Consequently, acceptance of Newtonian mechanics as a global dynamical theory prevented one from accepting Maxwell's theories. Yet, over and over, Maxwell's theories stood up to experiment, and the consequences predicted as a result of his theories continually turned out to be true. So ingrained, however, was the basis of Newtonian mechanics, that physicists sought more strongly to invalidate the new theory, rather than modify the old one.

This was not the only direction, though, from which Newton's classical theories were being threatened. While Maxwell's theories were pushing aside Newtonian mechanics at the microscopic level, it was becoming obvious that the classical theories failed macroscopically as well. Even though experimentation was still relatively primitive, the study of global dynamics on the universal scale was slowly revealing that Newtonian mechanics failed to accurately predict even the most basic phenomena. The motion in space of objects with masses, densities, and speeds far beyond anything observable on Earth was outside the scope of Newtonian mechanics. Many of the predictions made as a result of Newton's theories were in direct conflict with actual observed behavior. Thus, it was again clear that the principles of classical physics were not suitable for a global mechanical theory.

Hence, as the motivation grew for a reworking of theoretical physics, the mathematics involved progressed accordingly. Riemann, Cartan, Poincare, and others had developed and formalized the notions of a manifold, tensor, and metric, thus shedding new light on the mathematical notion of space. Extensions of these ideas led to modern notions of curvature and geodesics, or curves of minimal distance. Einstein used ideas from these early stages of modern differential geometry to develop his theory of special relativity, which placed Maxwell's theory of electromagnetism on consistent ground with the mechanics of light. He later developed his general theory of relativity, a theory of global mechanics that accounted for, on the macroscopic level, at least, the phenomenon of gravity, something Newtonian mechanics could not do. What made all of this possible was the introduction of a new type of metric, or, more accurately, metric tensor, that was more suitable to modeling in these extreme cases where speed, mass, and size were signifcantly greater than or less than that which was common to everyday phenomena.

This metric tensor, named the *Lorentzian metric tensor*, after the mathematician and physicist who first published results on its use, is now one of the cornerstones of theoretical physics. All of relativity theory relies on the postulate that space and time, as we know it, make up a four-dimensional manifold equipped with a Lorentzian metric. It has also found a place in quantum mechanics, as efforts are continually made to reconcile the seemingly disparate theories of relativity and mechanics on the microscopic scale. Consequently, along with the Riemannian metric tensor, the Lorentzian metric tensor is the most widely used metric structure on manifolds. It is very different from the Riemannian metric tensor, though, and, therefore, induces an entirely different array of geometric properties. Nevertheless, it is an indispensable tool in theoretical physics, and both metric tensors should be studied thoroughly in order to fully understand the physical and geometrical ramifications of using one over the other. The first question that naturally arises, however, is what conditions are necessary and sufficient for a Lorentzian metric tensor to exist on a manifold. If it is entirely different in nature from the Riemannian metric tensor, it is natural to expect that a different environment is required for its existence. Indeed, this is true, and we will spend the remainder of this chapter exploring exactly what manifolds admit Lorentzian metric tensors.

7.2 Nondegeneracy and the Index of a Metric Tensor

As we have been doing, we will assume throughout that M is a smooth manifold. Before we construct a Lorentzian metric tensor on M, we must provide some new definitions as well as redefine some of our previous notions in a more precise manner. Let Θ be a smooth, symmetric second order covariant tensor field on M. So, Θ defines a bilinear form on the tangent space at each point $p \in M$. If Θ was also positive definite, it would be a Riemannian metric tensor. However, for physical modeling purposes, we actually do *not* want a Lorentzian metric tensor to be *positive* definite, though we do want it to be *definite*. That is, at any point $p \in M$, we must be able to obtain negative numbers by evaluating $\Theta(X_p, X_p)$ for particular types of vectors, X_p, but we still require that $\Theta(X_p, Y_p) \neq 0$ unless the vectors are orthogonal (in a sense to be defined later) or at least one of them is the zero vector. This variant of "definiteness" is called **nondegeneracy**.

Definition 7.1 *Let V be a vector space of arbitrary dimension, and let $\alpha : V \times V \to \mathbb{R}$ be a real-valued, symmetric bilinear form on V. We say α is* **nondegenerate** *if, whenever $\alpha(v, w) = 0$ for all w in V, we must have $v = 0$. If Θ is a metric tensor on a smooth manifold, M, we say Θ is nondegenerate if the bilinear form, Θ_p, induced on the tangent space at each point $p \in M$ is nondegenerate.*

Note that, at each $p \in M$, the induced bilinear form, Θ_p, has all the properties of a Euclidean inner product on T_pM except positive definiteness. A positive definite bilinear form is clearly nondegenerate.

An important notion associated with an inner product, whether it is positive definite or not, is that of orthogonality. As is true in analysis in general, the notion orthogonality can be utilized in a number of ways and has far reaching consequences. In particular, it allows us to find a basis for the vector space in question with very desirable properties. The following lemma can be extended to infinite dimensional spaces, but, as we are working with finite dimensional manifolds, we will prove the result for just the finite case.

Lemma 7.1 *Let V be an n-dimensional vector space, and let $\alpha : V \times V \to \mathbb{R}$ be a real-valued, nondegenerate symmetric bilinear form on V. Then there is a basis, $\{e_1, \ldots, e_n\}$, for V such that $\alpha(e_i, e_j) = 0$ for $i \neq j$ and $\alpha(e_i, e_i) = \pm 1$ for all i.*

Proof For clarity, the quadratic form induced by α will be denoted by Q. That is, $Q : V \to \mathbb{R}$ is defined by $Q(v) = \alpha(v, v)$. The proof follows by induction on n, the dimension of V.

First, however, note the following. Since α is nondegenerate, there must be a pair of vectors, $v, w \in V$, such that $\alpha(v, w) \neq 0$. Consider $\alpha(v+w, v+w) = \alpha(v, v) + 2\alpha(v, w) + \alpha(w, w)$. If one of v or w satisfies $\alpha(v, v) \neq 0$ or $\alpha(w, w) \neq 0$, then we have a vector such that $Q(v) \neq 0$ (or $Q(w) \neq 0$). If neither v nor w satisfies this condition, then $Q(v) = Q(w) = 0$ and we have $Q(v+w) = \alpha(v+w, v+w) = Q(v) + 2\alpha(v, w) + Q(w) = 2\alpha(v, w) \neq 0$. Hence, because of nondegeneracy, there is always a vector, $v \in V$, such that $Q(v) \neq 0$.

Now, we continue with the induction proof. If $n = 1$, choose $v \in V$ such that $Q(v) \neq 0$. Define $e_1 = v/|Q(v)|^{1/2}$. Then

$$\begin{aligned} Q(e_1) &= Q\Big(\frac{v}{|Q(v)|^{1/2}}\Big) \\ &= \alpha\Big(\frac{v}{|Q(v)|^{1/2}}, \frac{v}{|Q(v)|^{1/2}}\Big) \\ &= \frac{Q(v)}{|Q(v)|} \\ &= \pm 1. \end{aligned}$$

The vector space, V, was arbitrary, so the result holds for all one-dimensional spaces.

Now suppose $n > 1$, and suppose that every real symmetric bilinear form on any vector space of dimension $< n$ has such a basis. Let V be a vector space of dimension n with a real symmetric bilinear form, α. Choose a vector, $v \in V$, such that $Q(v) = \alpha(v, v) \neq 0$. Let $e_n = v/|Q(v)|^{1/2}$, so that $Q(e_n) = \pm 1$.

Let $W = (\text{span}\{e_n\})^{\perp} = \{w \in V \ : \ \alpha(w, u) = 0 \, \forall \, u \in \text{span}\{e_n\} \, \}$. Then W is a subspace of V, and, since $e_n \notin W$, we must have $\dim W < \dim V = n$. In fact, the dimension of W must be $n-1$, for if we define the function $f : V \to \mathbb{R}$ by $f(w) = \alpha(w, e_n)$, then this is a linear transformation of *rank* 1, implying that the dimension of the kernel is $n-1$. Moreover, the kernel of f must equal W, for if $f(w) = 0$ and $u \in \text{span}\{e_n\}$, then $u = ce_n$ for some $c \in \mathbb{R}$ and $\alpha(w, u) = c\alpha(w, e_n) = cf(w) = 0 \Rightarrow w \in W$. Hence, W is an $n-1$-dimensional subspace of V.

Now, the restriction of α to W is clearly bilinear. To show it is nondegenerate, suppose there is some vector $w_1 \in W$ such that $\alpha(w_1, w) = 0$ for all $w \in W$. We know that $\alpha(w_1, e_n) = 0$, because $w_1 \in (\text{span}\{e_n\})^{\perp}$. Any vector, $u \in V$, must lie in W or $V - W$. If $u \in W$, then $\alpha(w_1, u) = 0$ by our assumption. If $u \in V - W$, then $u \in \text{span}\{e_n\}$, so $u = ce_n$ for some $c \in \mathbb{R}$. Thus, $\alpha(w_1, u) = c\alpha(w_1, e_n) = 0$. That is, $\alpha(w_1, u) = 0$ for all $u \in V$. But α is nondegenerate on V, so this implies $w_1 = 0$, and $\alpha|_W$ is also nondegenerate.

Hence, $\alpha|_W$ is a symmetric nondegenerate bilinear form on W. By our induction hypothesis, there is a basis $\{e_1, \ldots, e_{n-1}\}$ for W such that the restriction of α to W

satisfies $\alpha(e_i, e_j) = 0$ for $i \neq j$ and $\alpha(e_i, e_i) = \pm 1$ for all $i = 1, \ldots, n-1$. Consider the set $\{e_1, \ldots, e_{n-1}, e_n\} \subset V$. Suppose there are scalars $c_1, \ldots, c_n$ such that

$$\sum_{i=1}^{n} c_i e_i = 0.$$

Since each e_i, $i = 1, \ldots, n-1$ lies in W, it is orthogonal to e_n. Thus, we have

$$\begin{aligned} \alpha\left(\sum_{i=1}^{n} c_i e_i, e_n\right) &= \sum_{i=1}^{n} c_i \alpha(e_i, e_n) \\ &= c_n \alpha(e_n, e_n) \\ &= \pm c_n. \end{aligned}$$

So, we must have $c_n = 0$, implying that $\sum_{i=1}^{n-1} c_i e_i = 0$. But the vectors $\{e_1, \ldots, e_{n-1}\}$ form a basis, so they are linearly independent. Hence, we have $c_1 = c_2 = \cdots = c_{n-1} = c_n = 0$, and the vectors $\{e_1, \ldots, e_n\}$ are a linearly independent set of n vectors in an n-dimensional vector space. They, therefore, form a basis for V, and they satisfy the stated conditions.

QED

A basis satisfying the properties given in this lemma is called an **orthonormal basis** for V. It is important to remember that the orthonormality of a set of vectors depends upon the bilinear form in question. A set of vectors that form an orthnormal basis with respect to one bilinear form will, in general, not be orthonormal with respect to another bilinear form. In fact, it is possible to characterize a symmetric, nondegenerate bilinear form by the sign of its values on basis vectors.

Defintion 7.2 *Let $\alpha : V \times V \to \mathbb{R}$ be a symmetric nondegenerate bilinear form on a finite dimensional vector space, V, and let $\{e_1, \ldots, e_n\}$ be any orthonormal basis for V with respect to α. The* **index** *of α is the number of vectors in $\{e_1, \ldots, e_n\}$ such that $\alpha(e_i, e_i) = -1$.*

Of course, in order for this definition to make sense, it must be true that the index does not depend on the particular orthonormal basis chosen. This is a remarkably useful result in linear algebra.

Theorem 7.2 *Let V be an n-dimensional vector space with symmetric nondegenerate bilinear form α. The index of α is independent of the particular orthonormal basis chosen. In fact, the index of α is equal to the dimension of the highest dimensional subspace of V on which α is negative definite.*

Proof α is trivially negative definite on the subspace $\{0\}$. So, there are, indeed, subspaces of V on which α is negative definite. Since the possible dimensions of such a subspace are finite and bounded, there must be a subspace, W, of V of largest dimension on which α is negative definite. Note that W need not be unique, but we will show that its dimension equals the index of α. This will give us a basis-independent characterization of the index.

Let $\{e_1, \ldots, e_n\}$ be any orthonormal basis for V with respect to α. We can assume that this basis is ordered so that $\alpha(e_i, e_i) = -1$ for $1 \leq i \leq k$ and $\alpha(e_i, e_i) = 1$ for $k+1 \leq i \leq n$. Let $E = \text{span}\{e_1, \ldots, e_k\}$, and let v be a nonzero vector in E. Then there are scalars $v^1, \ldots, v^k$ such that $v = \sum_{i=1}^k v^i e_i$. We can think of v as a vector in $\text{span}\{e_1, \ldots, e_n\}$ by letting the scalars $v^{k+1}, \ldots, v^n$ equal zero. So, we have

$$\begin{aligned} Q(v) &= \alpha(v, v) \\ &= \sum_{i,j=1}^{k} v^i v^j \alpha(e_i, e_j) \\ &= \sum_{i=1}^{k} v^i \sum_{j=1}^{k} v^j \alpha(e_i, e_j) \\ &= \sum_{i=1}^{k} v^i \sum_{j=1}^{k} -v^j \delta^i_j \\ &= -\sum_{i=1}^{k} (v^i)^2. \end{aligned}$$

It follows that $Q(v) < 0$. Since $v \in E$ was arbitrary, this shows that α is negative definite on E. Thus, dimension $E \leq$ dimension W.

Now, define $\lambda : W \to E$ as follows. For $v \in W \subset V$, with basis representation $v = \sum_{i=1}^n v^i e_i$, define

$$\lambda(v) = \sum_{i=1}^{k} v^i e_i.$$

Suppose there is $v \in W$ be such that $v \neq 0$ and $\lambda(v) = 0$. Then

$$\lambda(v) = \sum_{i=1}^{k} v^i \lambda(e_i) = 0.$$

But, for $i = 1, \ldots, k$, we have $\lambda(e_i) = e_i \neq 0$, so $\lambda(v) = 0$ implies that $v^i = 0$ for $i = 1, \ldots, k$. Thus, $v = \sum_{i=k+1}^n v^i e_i$, and we have

$$\begin{aligned} Q(v) &= \alpha(v, v) \\ &= \sum_{i,j=k+1}^{n} v^i v^j G(e_i, e_j) \\ &= \sum_{i=k+1}^{n} v^i \sum_{j=k+1}^{n} v^j \delta^i_j \\ &= \sum_{i=k+1}^{n} (v^i)^2. \end{aligned}$$

Now, since $v \neq 0$ and $v^1 = \cdots = v^k = 0$, one of the scalars v^i, $k+1 \leq i \leq n$ must be nonzero. Thus, by this last equality, we must have $Q(v) > 0$. This contradicts the fact that α is negative definite on W. Hence, there can be no vector, $v \in W$, such that $\lambda(v) = 0$ with $v \neq 0$. That is, $\lambda(v) = 0 \Rightarrow v = 0$, which further implies that λ is an injective linear transformation. It follows that dimension $W \leq$ dimension E. So, we have dim W = dim $E = k$ = index(α). Nothing in our construction depended upon the particular orthonormal basis. W exists independently of any basis, and only depends on the bilinear form α. So, this shows that index(α) equals the dimension of the largest subspace of V on which α is negative definite.

QED

So, the index of a symmetric nondegenerate bilinear form only depends on the bilinear form itself.

Note that a Euclidean inner product has an index of 0. In fact, any positive definite bilinear form has an index of 0. Thus, if Θ is a Riemannian metric tensor on a smooth manifold, M, then, for any point $p \in M$, the induced bilinear form, Θ_p, on T_pM has and index of 0. This yields a new way to define a Riemannian metric tensor.

Definition 7.3 *Let Θ be a nondegenerate metric tensor on a smooth manifold, M, and let Θ_p denote the bilinear form induced on T_pM for any point $p \in M$. We say that Θ is a* **Riemannian metric tensor** *if Θ_p has index 0 for each $p \in M$. We say that Θ is a* **Lorentzian metric tensor** *if Θ_p has index 1 for each $p \in M$. In these cases, we call M a* **Riemannian manifold** *and a* **Lorentzian Manifold**, *respectively.*

Now, suppose M is a Lorentzian manifold with Lorentzian metric tensor, Θ. For any point $p \in M$, we can, by Lemma 7.1, find a basis, $\{e_1, \ldots, e_m\}$, for T_pM that is orthonormal with respect to Θ. Moreover, we can assume without loss of generality that this basis is ordered so that $\Theta(e_i, e_i) = 1$ for $1 \le i \le m-1$ and $\Theta(e_m, e_m) = -1$. Then, with respect to this basis, if $X_p = \sum_{i=1}^m c_i e_i$ is a tangent vector at p, we have

$$Q(X_p) = \Theta(X_p, X_p) = \sum_{i,j=1}^{m} c^i c^j \Theta(e_i, e_j) = \sum_{i=1}^{m-1} (c^i)^2 - (c^m)^2.$$

This is the canonical representation of the action of a Lorentzian metric tensor on a vector. Admittedly, this is atypical of a bilinear form that we relate to an inner product. As we said before, though, Lorentzian geometry is quite different from Riemannian geometry. It is, however, a fascinating subject in its own right. This simple variant of the usual "dot" product produces some very interesting geometrical properties of manifolds. Moreover, it is essential in theoretical physics and the modeling of phenomena on both the global and microscopic levels.

Some examples of Lorentzian manifolds are certainly in order, as such a new and unorthodox construction may easily produce skepticism among the first time reader. We will postpone these examples, however, until after the following section, in which we establish necessary and sufficient conditions for the existence of Lorentzian manifolds.

7.3 The Lorentzian Metric Tensor

In this section, we establish the existence of a Lorentzian metric on a suitable smooth manifold. Since, in general, positive definiteness is such a stringent condition to place on an operator, one might be inclined to think that Lorentzian metric tensors are actually more common than Riemannian metric tensors. Suprisingly, this is not the case. In fact, as our proof will show, a requirement for the existence of a Lorentzian metric tensor on M is that M also be a Riemannian manifold. Indeed, as we will discuss later, there are Riemannian manifolds that are not Lorentzian manifolds, but all Lorentzian manifolds are also Riemannian.

To start our construction, we need the following definition.

Definition 7.4 *Let M be a smooth manifold, and let $h \le m$ be fixed. Suppose there are smooth vector fields, $X_1, \ldots, X_h$, on M satisfying the following condition: for every point $p \in M$, there is a neighborhood, U, of p, such that, for every $q \in U$, the set $\{X_{1_q}, \ldots, X_{h_q}\}$ is linearly independent. Then for each $p \in M$, define $L^h(p) =$*

span$\{X_{1_p}, \ldots, X_{h_p}\}$. *We call* L^h *an* h-**dimensional smooth tangent subspace field***, or, more concisely, an* h-**dimensional smooth distribution** *on* M.

While this definition may seem complicated, all it says is that an h-dimensional distribution on M is an establishment of an h-dimensional subspace of T_pM at each $p \in M$, so that the transition from the subspace at p to the subspace at q near p is smooth. The smoothness condition is characterized by vector fields. If there is a set of h smooth vector fields on M, such that at each point $p \in M$, the induced tangent vectors at p are linearly indepdnent, then they establish an h-dimensional subspace of T_pM. We will not dwell too long on this topic, as we will only need a special case of it for our purposes. We should point out, however, that this is an indispensable tool in the theory of manifolds. For example, the famous Frobenius Theorem, which, in one form, establishes necessary and sufficient conditions for the solution of Pfaffian systems on manifolds, requires the notion of a smooth distribution.

In establishing a Lorentzian metric, we will need the notion of a 1-dimensional smooth distribution, L^1, on M. Such a distribution is also called a C^∞ **field of lines**, since it establishes a 1-dimensional subspace of T_pM for each $p \in M$. The smoothness characterization for L^1 is as follows. For each point $p \in M$, there must be a neighborhood, U, of p and a smooth vector field, X, on M such that, for $q \in U$, the subspace $L^1(q)$ (i.e. the "line" at q) is equal to the span of X_p. Note that a smooth, everywhere nonzero vector field, X, induces a smooth 1-dimensional distribution, and conversely.

Now we want to consider the local matrix representation of a metric tensor. This is another concept we will need in our construction of the Lorentzian metric tensor. We discussed this idea briefly in chapter 5, but we will need more detail here. Let (φ, U) be a coordinate neighborhood of M, and suppose G is a nondegenerate metric tensor on M. This discussion will not depend on whether G is positive definite or not, so we need not specify whether G is Riemannian, Lorentzian, or of some other form. Then we know that, on U, G can be expressed $G = \sum_{i,j=1}^m g_{ij} dx^i \otimes dx^j$, where $\{dx^i\}$ is the standard basis field for the cotangent spaces over U. For each i and j, the function g_{ij} is a smooth real-valued function on U. Note that the form of this sum is exactly the same as that which defines a bilinear form over a finite dimensional vector space vector space. In fact, if we fix $p \in U$, and we have tangent vectors $X_p = \sum_{k=1}^m \alpha^k(p) E_{p_k}$, $Y_p = \sum_{l=1}^m \beta^l(p) E_{p_l}$ in T_pM represented in terms of the local tangent basis field $\{E_i\}$, then we have

$$\begin{aligned} G_p(X_p, Y_p) &= \sum_{i,j=1}^{m} g_{ij}(p) dx^i \otimes dx^j (X_p, Y_p) \\ &= \sum_{i,j=1}^{m} g_{ij}(p) dx^i \left(\sum_{k=1}^{m} \alpha^k(p) E_{p_k} \right) dx^j \left(\sum_{l=1}^{m} \beta^l(p) E_{p_l} \right) \\ &= \sum_{i,j=1}^{m} g_{ij}(p) \alpha^i(p) \beta^j(p). \end{aligned}$$

Now, let $[X_p]$ and $[Y_p]$ denote, respectively, the coordinate vectors of X_p and Y_p with respect to the basis $\{E_{p_i}\}$. Likewise, let $[G_p]$ denote the matrix whose ij^{th} entry is $g_{ij}(p)$. Then this last equality can be expressed as the matrix equation $G_p(X_p, Y_p) = [X_p]^T [G_p][Y_p]$. By letting p vary over U, we obtain such a matrix expression for G over all of U. In doing so, we obtain a matrix function, $[G]$, defined on U, where $[G](p) = [G_p]$. The component functions defining the matrix function $[G]$ are just the functions $g_{ij} : U \to \mathbb{R}$. Since G is a metric tensor, these functions are smooth on U, implying that this matrix function is also smooth.

Extending this idea, we can think of matrices at a point $p \in M$ as operators on T_pM. This leads us to the following definition.

Definition 7.5 *A* **tangent operator field** *over M is an assignment of a linear operator $H_p : T_pM \to T_pM$ to each $p \in M$. Globally, we refer to the tangent operator field as H, and, for a point $p \in M$, we denote the specific operator at p by $H(p) = H_p$.*

Note that, for any $p \in M$ with coordinate system (φ, U) around p, we can obtain a matrix expression for H_p using the usual linear algebra methods. If $\{E_i\}$ is the standard tangent basis field over U, then H_p is just a linear transformation from T_pM into T_pM. If we assume that $\{E_{p_i}\}$ is the desired basis in both the domain and range, then we can compute the matrix representation of H_p with respect to these bases. Thus, the k^{th} column of H_p is the coordinate vector of $H_p(E_{p_k})$ with respect to the basis $\{E_{p_i}\}$. Suppose for each $k = 1, \ldots, m$, we have scalars $h_k^j(p)$, $1 \leq j \leq m$ such that $H_p(E_{p_k}) = \sum_{j=1}^{m} h_k^j(p) E_{p_j}$. Then the matrix expression for H_p, denoted $[H_p]$, is given by

$$[H_p] = \begin{bmatrix} h_1^1(p) & h_2^1(p) & \cdots & h_m^1(p) \\ h_1^2(p) & h_2^2(p) & \cdots & h_m^2(p) \\ \vdots & \vdots & \ddots & \vdots \\ h_1^m(p) & h_2^m(p) & \cdots & h_m^m(p) \end{bmatrix}$$

Letting p vary over U, we obtain m^2 functions $h^i_j : U \to \mathbb{R}$. These functions define a matrix function, $[H]$, on U, such that $[H](p) = [H_p]$. That is, at $p \in U$, the matrix $[H](p)$ is just the matrix representation of H_p with respect to the standard bases. The functions h^i_j, $1 \leq i, j \leq m$ are the local component functions of this matrix function. In this manner, we obtain a local coordinate expression for the tangent operator field, H, just as we have for other globally defined objects.

Now consider the following. If X is a vector field on M, then we can use H to define a new vector field, $H(X)$, on M in a natural way. For each $p \in M$, define $(H(X))(p) = (H(X))_p$ to be $H_p(X_p)$.

Definition 7.6 *We say that a tangent operator field, H, is smooth, or C^∞, if for every smooth vector field, X, on M, the vector field $H(X)$ is a smooth vector field on M.*

As with our other globally defined objects in which questions of smoothness arise, we can also characterize smoothness of tangent operator fields in terms of local coordinate representations.

Lemma 7.3 *Let H be a tangent operator field over M. Then H is smooth in the sense of Definition 7.6 if and only if, for every coordinate chart (φ, U) on M, the local coordinate matrix function $[H]$, defined on U, has smooth component functions.*

Proof Suppose H is smooth according to Definition 7.6, and let (φ, U) be any coordinate system on M. Let h^i_j, $1 \leq i, j \leq m$, denote the component functions that define $[H]$, as in our previous discussion. We want to show that these are smooth real-valued functions on U.

As usual, let $\{E_i\}$ denote the standard tangent basis field on U. For any $1 \leq j \leq m$, the vector field E_j is smooth on U. Thus, the vector field $H(E_j)$ is smooth on U by hypothesis. But by our definitions we have $H(E_j) = \sum_{i=1}^m h^i_j E_i$. Since a smooth vector field has smooth local component functions, it follows that the functions h^i_j, for $1 \leq i, j \leq m$, are smooth. Thus, the local coordinate matrix function, $[H]$, has smooth component functions. Since the chart (φ, U) was arbitrary, this proves the necessary condition.

Conversely, suppose that, for any chart (φ, U), on M, the local coordinate matrix function, $[H]$, representing H on U, has smooth component functions. Let X be any smooth vector field on M, and consider the vector field $H(X)$, restricted to U. Since the chart is arbitrary, it suffices to show that $H(X)$ has smooth component functions on U. Let the local coordinate representation of X on U be given by $X = \sum_{i=1}^m \alpha^i E_i$. Then each function $\alpha^i : U \to \mathbb{R}$ is smooth. For any fixed point $p \in U$, we have

$$
\begin{aligned}
H_p(X_p) &= H_p\left(\sum_{i=1}^{m} \alpha^i(p) E_{p_i}\right) \\
&= \sum_{i=1}^{m} \alpha^i(p) H_p(E_{p_i}) \\
&= \sum_{i=1}^{m} \alpha^i(p) \sum_{j=1}^{m} h_i^j(p) E_{p_j} \\
&= \sum_{j=1}^{m} \left[\sum_{i=1}^{m} \alpha^i(p) h_i^j(p)\right] E_{p_j}.
\end{aligned}
$$

So, letting p vary over U, we obtain the expression

$$
H(X)|_U = \sum_{j=1}^{m} \left[\sum_{i=1}^{m} \alpha^i h_i^j\right] E_j.
$$

Since the component functions h_i^j and α^i, for $1 \le i, j \le m$ are all smooth, it follows that $H(X)|_U$ has smooth component functions on U. This shows that the component functions of the vector field $H(X)$ on every local coordinate system are smooth. Hence, $H(X)$ is smooth.

QED

Now, having laid the groundwork sufficiently, we are ready to prove the main result of this chapter.

Theorem 7.4 *A smooth manifold M admits a Lorentzian metric tensor if and only if it admits a smooth* 1*-dimensional distribution.*

Proof First, suppose M admits a smooth 1-dimensional distribution, L^1. Since M is a smooth manifold, there is a Riemannian metric tensor on M, which we will denote by G. Let p be any point in M. The $L^1(p)$ is a 1-dimensional subspace of T_pM. Choose a unit vector, $v_p \in T_pM$, such that $\text{span}\{v_p\} = L^1(p)$. Note that v_p is unique up to a sign. Choose such a v_p for each $p \in M$. Define a second order covariant tensor field, $\widehat{G}$, on M by

$$
\widehat{G}_p(u_p, w_p) = G_p(u_p, w_p) - 2G_p(v_p, u_p)G_p(v_p, w_p),
$$

where, for each $p \in M$, we have $u_p, w_p \in T_pM$ and $L^1(p) = \text{span}\{v_p\}$. We first note that $\widehat{G}$ is symmetric, since, for any $p \in M$ and $u_p, w_p \in T_pM$, we have

$$\begin{aligned}
\widehat{G}_p(u_p, w_p) &= G_p(u_p, w_p) - 2G_p(v_p, u_p)G_p(v_p, w_p) \\
&= G_p(w_p, u_p) - 2G_p(v_p, w_p)G_p(w_p, u_p) \\
&= \widehat{G}_p(w_p, u_p).
\end{aligned}$$

Moreover, $\widehat{G}$ is well-defined, independent of the choice of sign of v_p, since

$$\begin{aligned}
\widehat{G}_p(u_p, w_p) &= G_p(u_p, w_p) - 2G_p(-v_p, u_p)G_p(-v_p, w_p) \\
&= G_p(u_p, w_p) - 2(-1)(-1)G_p(v_p, u_p)G_p(v_p, u_p) \\
&= G_p(u_p, w_p) - 2G_p(v_p, u_p)G_p(v_p, u_p).
\end{aligned}$$

Now, choose an orthonormal basis, $\{e_1, \ldots, e_n\}$, for T_pM relative to G, such that $e_1 = v_p$. Note that this is always possible since $v_p \neq 0$ (see Lemma 7.1). Then we see that

$$\begin{aligned}
\widehat{G}_p(e_i, e_j) &= G_p(e_i, e_j) - 2G_p(v_p, e_i)G_p(v_p, e_j) \\
&= \delta^i_j - 2G_p(e_1, e_i)G_p(e_1, e_j) \\
&= \delta^i_j - 2\delta^i_1\delta^j_1.
\end{aligned}$$

It follows that

$$\widehat{G}_p(e_i, e_j) = \begin{cases} -1 & i = j = 1 \\ 1 & i = j \neq 1 \\ 0 & i \neq j \end{cases}$$

Hence, we see that $\{e_1, \ldots, e_n\}$ is also an orthonormal basis for T_pM relative to $\widehat{G}$, and $\widehat{G}$ has index 1. Since $p \in M$ was arbitrary, it follows that $\widehat{G}$ is Lorentzian. So, all we have left to show is that $\widehat{G}$ is smooth.

Let (φ, U) be any coordinate chart on M. Choose a smooth vector field, X, on M such that for each $p \in M$, X_p is a unit vector relative to G and $L^1(p) = \text{span}\{X_p\}$. Such a vector field exists by hypothesis, since we are assuming the existence of a smooth 1-dimensional distribution. Consider the restriction of X to U, which we denote simply by X. We can then express X is is local coordinate form, where, for $q \in U$, we have

$$X(q) = \sum_{i=1}^{m} x^i(q) E_{q_i}, \qquad\qquad x^i : U \to \mathbb{R}.$$

The component functions of $\widehat{G}$ on U, denoted $\widehat{g_{ij}}$, are given by $\widehat{G}(E_i, E_j)$. So, at any point $p \in U$,

$$\begin{aligned}
\widehat{g_{ij}(p)} &= \widehat{G}_p(E_{p_i}, E_{p_j}) \\
&= G_p(E_{p_i}, E_{p_j}) - 2G_p(X_p, E_{p_i})G_p(X_p, E_{p_j}) \\
&= g_{ij}(p) - 2G_p\left(\sum_{k=1}^{m} x^k(p)E_{p_k}, E_{p_i}\right)G_p\left(\sum_{k=1}^{m} x^k(p)E_{p_k}, E_{p_j}\right) \\
&= g_{ij}(p) - 2\sum_{k,l=1}^{m} x^k(p)x^l(p)G_p(E_{p_k}, E_{p_i})G_p(E_{p_l}, E_{p_j}) \\
&= g_{ij}(p) - 2\sum_{k,l=1}^{m} x^k(p)x^l(p)g_{ki}(p)g_{lj}(p).
\end{aligned}$$

Thus, the component functions of $\widehat{G}$ on U are

$$\widehat{g_{ij}} = g_{ij} - 2\sum_{k,l=1}^{m} x^k x^l g_{ki} g_{lj}.$$

Since G is smooth, the functions g_{ij} are smooth on U for all i and j. Since X is a smooth vector field, the functions x^i are smooth on U. Hence, the component functions of $\widehat{G}$ are smooth on U. Finally, since the chart (φ, U) was arbitrary, we can conclude that $\widehat{G}$ is smooth on M. That is, $\widehat{G}$ is a smooth Lorentzian metric tensor.

The proof of the converse is not difficult. It is merely a tedious exercise in notation. So we will have to derive some preliminary results. Suppose $\widehat{G}$ is a smooth Lorentzian metric tensor on M. As before, let G denote the Riemannian metric tensor on M. Let p be any point in M. For each p, define maps $\tau_p : T_pM \to T_p^*M$ and $\widehat{\tau}_p : T_pM \to T_p^*M$ by $\tau_p(v) = G_p(v, *)$ and $\widehat{\tau}_p(v) = \widehat{G}_p(v, *)$, respectively. That is, for $v \in T_pM$, we define a linear functional, $\tau(v)$, on T_p^*M by $\tau(v)(w) = G_p(v, w)$, and likewise for $\widehat{\tau}$.

It is straightforward to show that both τ and $\widehat{\tau}$ are linear. We also want to show that τ and $\widehat{\tau}$ are bijections. Suppose $\tau(v_1) = \tau(v_2)$. Then, for all $w \in T_pM$, we have $G_p(v_1, w) = G_p(v_2, w) \Rightarrow G_p(v_1 - v_2, w) = 0$. Since G is nondegenerate, this implies that $v_1 = v_2$. Similarly, if $\widehat{\tau}(v_1) = \widehat{\tau}(v_2)$, then, for all $w \in T_pM$, we have $\widehat{G}_p(v_1, w) = \widehat{G}_p(v_2, w) \Rightarrow \widehat{G}_p(v_1 - v_2, w) = 0 \Rightarrow v_1 = v_2$. Thus, both τ and $\widehat{\tau}$ are injective. Consequently, since T_pM and T_p^*M both have dimension m, it follows that these maps are also both surjective. So, both τ and $\widehat{\tau}$ are isomorphisms from T_pM onto T_p^*M.

Let (φ, U) be a coordinate system on M, and let $\{E_i\}$ and $\{dx^i\}$ denote the standard basis fields for the tangent and cotangent spaces, respectively, over U. We next claim that, for each $p \in U$, the matrix representations of τ_p and $\widehat{\tau}_p$ with respect to the standard coordinate bases are, respectively, the matrices $[g_{ij}(p)]$ and $[\widehat{g_{ij}}(p)]$, where g_{ij} and $\widehat{g_{ij}}$ are the component functions of G and $\widehat{G}$, respectively, on U. To see this, we first point out that the j^{th} column of $[\tau_p]$ is the vector whose components are the coordinates of $\tau_p(E_{p_j})$ with respect to the basis $\{dx_p^i\}$. So, let w be any tangent vector. Then,

$$\begin{aligned} \tau_p(E_{p_j})(w) &= G\left(E_{p_j}, \sum_{i=1}^{m} w^i E_{p_i}\right) \\ &= \sum_{i=1}^{m} w^i G(E_{p_j}, E_{p_i}) \\ &= \sum_{i=1}^{m} G(E_{p_i}, E_{p_j}) dx_p^i(w). \end{aligned}$$

That is, $\tau_p(E_{p_j}) = \sum_{i=1}^{m} G(E_{p_i}, E_{p_j}) dx_p^j$, implying that the coordinates of $\tau_p(E_{p_j})$ with respect to the basis $\{dx_p^i\}$ are the numbers $G(E_{p_i}, E_{p_j}) = g_{ij}(p)$. Hence, the j^{th} column of $[\tau_p]$ is the vector with components $g_{ij}(p)$, and it follows that the matrix representation of τ_p with respect to the standard coordinate bases is the matrix $[g_{ij}(p)]$. It follows in exactly the same manner that the matrix representation of $\widehat{\tau}_p$ with respect to the standard coordinate bases is the matrix $[\widehat{g_{ij}}(p)]$. So, as p varies over U, we see that the representation of the field of maps, $\tau_p : T_pM \to T_p^*M$, $p \in U$, with respect to the standard basis fields, is the matrix of functions, $[g_{ij}]$. Likewise, the reprsentation of the field of maps $\widehat{\tau}_p : T_pM \to T_p^*M$, $p \in U$, is the matrix of functions $[\widehat{g_{ij}}]$. Moreover, by what we have shown, the maps τ_p and $\widehat{\tau}_p$ are invertible for each $p \in U$. We will need these inverses to complete the proof. For simplicity of notation, we will denote the component functions of the inverses, ${\tau_p}^{-1}$ and $\widehat{\tau}_p^{-1}$, by $g^{ij}(p)$ and $\widehat{g}^{ij}(p)$, respectively. Hence, the representations of the fields of maps, τ_p^{-1} and $\widehat{\tau}_p^{-1}$, for $p \in M$, are, respectively, the matrices of functions $[g^{ij}]$ and $[\widehat{g}^{ij}]$. Finally, note that, since the entries of an inverse matrix depend smoothly on the entries of the original matrix, it follows that the component functions of these inverses are also smooth.

Now, for each $p \in U$, define a map $H_p : T_pM \to T_pM$ by $H_p(v) = (\tau_p^{-1} \circ \widehat{\tau_p})(v)$. Then, for each $p \in U$, H_p is an automorphism of T_pM. Hence, we have obtained a tangent operator field, H, such that $H(p) = H_p$. Relative to any basis we choose for T_pM, it follows that the matrix representation of H_p is the product $[\tau_p]^{-1}[\widehat{\tau_p}]$. Thus relative to the standard basis fields on U, H has a local matrix representation whose ij^{th} entry is given by

$$[H]_{ij} = \sum_{k=1}^{m} g^{ik}\widehat{g_{kj}}.$$

Since the coordinate system (φ, U) has been arbitrary, this shows that H is a smooth tangent operator field, as its component functions on each chart are smooth.

We will construct a C^∞ field of lines on M as follows. We will show that at each $p \in M$, $[H_p]$ has exactly one negative eigenvalue, while the rest are strictly positive. It will follow that the eigenspace associated with the single negative eigenvalue will be one-dimensional for each $p \in M$. This will give us a field of lines on M. The fact that H is smooth will imply that our field of lines is also smooth.

Fix any $p \in M$. Choose an orthonormal basis for T_pM relative to G, denoted $\{e_1, \ldots, e_m\}$. Note that for such a basis, $[\tau_p]^{-1} = I$, so the matrix representation of H_p is just $[\widehat{\tau_p}]$. Using a common result from linear algebra on the simultaneous reduction of two symmetric matrices, there is a nonsingular matrix, R, such that $R^T R = I = [\tau_p]^{-1}$ and $[\widehat{\tau_p}] = R^T DR$, where D is a diagonal matrix. Since $R^T R = I$, we have $R^{-1} = R^T$. So, R is orthogonal, $R^T DR$ is the orthogonal diagonalization of $[\widehat{\tau_p}]$, and the elements on the diagonal of D are the eigenvalues of $[\widehat{\tau_p}]$. Now, if we consider another basis, $\{\widetilde{e_1}, \ldots, \widetilde{e_m}\}$, where

$$\widetilde{e_i} = \sum_{j=1}^{m} r_{ji} e_j,$$

and r_{ij} is the ij^{th} entry of R, then this is also an orthonormal basis for T_pM relative to G. Hence, relative to this basis, we still have $[\tau_p]^{-1} = I$. Moreover, since $[\widehat{\tau_p}] = R^T DR$, it follows that relative to this new basis, we have $[\widehat{\tau_p}] = D$. So, the matrix representation of H_p relative to this new basis is the diagonal matrix D. That is, $[H_p] = [\widehat{\tau_p}] = [\widehat{g_{ij}(p)}] = D$. This implies that the basis $\{\widetilde{e_1}, \ldots, \widetilde{e_m}\}$ is orthogonal with respect to the Lorentzian metric tensor, $\widehat{G}$, though not necessarily orthonormal. However, since $\widehat{G}$ is nondegenerate, none of the diagonal of elements of D are zero, so we can normalize this basis to obtain an orthonormal basis for T_pM relative to $\widehat{G}$.

Since the index is independent of the orthonormal basis, it follows that exactly one of these normalized basis vectors, say $\widetilde{e_k}/\|\widetilde{e_k}\|$, satisfies

$$\widehat{G}\Big(\frac{\widetilde{e_k}}{\|\widetilde{e_k}\|}, \frac{\widetilde{e_k}}{\|\widetilde{e_k}\|}\Big) = -1.$$

For this particular vector, it follows that $\widehat{G}(\widetilde{e_k}, \widetilde{e_k}) = -\|\widetilde{e_k}\|^2$, implying that $d_k = -\|\widetilde{e_k}\|^2 < 0$, where d_k is the k^{th} diagonal element of D. Hence, this one eigenvalue will be negative, while all the rest will be strictly positive. We can conclude, then, that the eigenspace corresponding to this negative eigenvalue is one-dimensional. Let v_p be an eigenvector corresponding to this negative eigenvalue. Then $\text{span}\{v_p\}$ is equal to this eigenspace.

Since the point $p \in M$ was arbitrary in this construction, we can find such an eigenspace for each $p \in M$. Define a 1-dimensional distribution, L^1, on M by $L^1(p) = \text{span}\{v_p\}$. Now, for a fixed point $p \in M$, this eigenspace is an eigenspace of the operator $H_p = H(p)$. But H is a smooth tangent operator field. If a field of symmetric matrices, A, depends smoothly on some domain, $\mathscr{D}$, and if, for each $x \in \mathscr{D}$, $A(x)$ has a unique negative eigenvalue with a corresponding 1-dimensional eigenspace, then there must exist a smooth scalar function $\lambda : \mathscr{D} \to \mathbb{R}$ and a smooth vector-valued function, v, such that for each x, $\lambda(x)$ is this negative eigenvalue and $v(x)$ is a corresponding eigenvector. Thus, our choice of eigenspace in defining L^1 must define a smooth distribution. Hence, L^1 is our C^∞ field of lines, and this completes the proof.

QED

This result is the most common means of identifying Lorentzian manifolds, although equivalent forms are used just as often. The next problem that naturally confronts us is determining which manifolds admit smooth 1-dimensional distributions. This is a difficult question, indeed, and is of deep importance in both algebraic and differential topology. For example, it is well-known, though difficult to prove, that the standard 2-dimensional sphere, S^2, does not admit a continuous, everwhere non-zero vector field. (You can't comb the hair on a tennis ball, as it is often phrased.) Thus, every smooth vector field on S^2 must have a zero. It follows that a C^∞ field of lines cannot exist on S^2, for if one did, there would have to be some smooth everywhere nonzero vector field defining the subspace at each point p. More generally, no even dimensional sphere, S^{2n}, admits a continuous, everywhere nonzero vector field. Consequently, no sphere of even dimension is a Lorentzian manifold. However, all spheres of odd dimension are Lorentzian manifolds.

In more general topological discussions, there is an equivalent reformulation of the preceding result involving the Euler characteristic of a manifold, denoted $\chi(M)$. A complete discussion of this number would require a significant digression into notions of curvature and various tensors defined on M. So, we will only point out here that the Euler characteristic of a manifold, in one interpretation, is a means of "measuring", in a specific sense, the set of singular points on a manifold. A common result equivalent to Theorem 7.4 is the following.

Theorem 7.5 *A simply connected smooth manifold, M, admits a smooth field of lines if and only if $\chi(M) = 0$.*

Hence, a simply connected smooth manifold, M, is Lorentzian if and only if its Euler characteristic vanishes.

To close this section, we present a couple of examples of Lorentzian manifolds. These examples are considered basic by modern theoretical physics. The most interesting examples of Lorentzian manifolds, however, require a good knowledge of theoretical mechanics, relativity, and the geometrical properties of general Lorentzian manifolds. This, of course, is a perfect topic for another work, but we will not pursue these questions here.

Example 7.1 *Minkowski Spacetime*

Minkowski space, denoted simply by M^4, is the natural setting for Einstein's special theory of relativity. As a set, M^4 contains exactly the same elements as $\mathbb{R}^4$. However, M^4 is not a Euclidean space, due to its unusual topology. A complete description and motivation for this topology requires at least a decent introduction to the physics of light, so we will only given a mathematical overview. Minkowski space is constructed like a metric space in classical analysis. We will use a quadratic form on $\mathbb{R}^4$ to define a collection of sets that give $\mathbb{R}^4$ a distinct topology. This quadratic form will turn out to be the quadratic form associated with the Lorentzian metric tensor.

Let $x = (x^1, x^2, x^3, x^4)$ denote a point of $\mathbb{R}^4$. We define a bilinear form, G, on $\mathbb{R}^4$ by $G(x, y) = x^1y^1 + x^2y^2 + x^3y^3 - x^4y^4$. The quadratic form associated with this bilinear form will be denoted by Q, and is defined by $Q(x) = (x^1)^2 + (x^2)^2 + (x^3)^2 - (x^4)^2$. Hence, this bilinear form has index 1. For each $x_0 \in \mathbb{R}^4$, we define the **time cone** at x_0, $C_T(x_0)$, by

$$C_T(x_0) = \{x \in \mathbb{R}^4 \ : \ Q(x - x_0) < 0\}.$$

Geometrically, $C_T(x_0)$ is the interior of a right circular cone with vertex at x_0. The two conic pieces open along opposite directions parallel to the $x^4 - axis$. If we choose an orientation along the $x_4 - axis$, so that we have *positive* and *negative* directions, then we can distinguish the two conic pieces as the **future time cone** and the **past time cone**. That is, the future time cone is the conic piece of $C_T(x_0)$ that opens in the positive x^4 direction, while the past time cone is the piece that opens in the negative x_4 direction. We denote the sets, respectively, by $C_T^+(x_0)$ and $C_T^-(x_0)$. The reason for this terminology is that, in special relativistic mechanics, the fourth dimension is usually the time dimension. So, if we think of x_0 as an event in space and time, then the time cone at x_0 distinguishes the past and future of this event. Note that the time cone, $C_T(x_0)$, does not include the point x_0. So, the sets $C_T^+(x_0)$ and $C_T^-(x_0)$ do not contain x_0 either.

Now, let $B_\epsilon(x_0)$ denote the standard open ball of radius ϵ centered at x_0 with respect to the usual Euclidean metric on $\mathbb{R}^4$. Let $C(x_0) = C_T^+(x_0) \cup C_T^-(x_0) \cup \{x\}$, which is just the time cone at x_0 with the point x_0 included. Finally, define $N_\epsilon(x_0) = C(x_0) \cap B_\epsilon(x_0)$. The set $N_\epsilon(x_0)$ is nothing more than the time cone at x_0, with the point x_0 included, intersected with the usual open ball of radius ϵ centered at x_0. Hence, it is a type of hourglass set, with a curved top and bottom, centered at x_0. Note that $N_\epsilon(x_0)$ is not open with respect to the usual Euclidean topology, since the point x_0 will actually be a boundary point with respect to this topology.

The collection of sets $\{N_\epsilon(x) \ : \ x \in \mathbb{R}^4, \ \epsilon > 0\}$ form a basis for a topology on $\mathbb{R}^4$. They clearly cover $\mathbb{R}^4$, since $N_\epsilon(x)$ contains x. Also, note that, for any $x \in \mathbb{R}^4$ and any $\epsilon > 0$, the sets $C_T^+(x) \cap B_\epsilon(x)$ and $C_T^-(x) \cap B_\epsilon(x)$ are open with respect to the usual Euclidean topology. So, if two sets $N_{\epsilon_1}(x_1)$ and $N_{\epsilon_2}(x_2)$ have a nonempty intersection, then this intersection must contain an open ball with respect to the Euclidean topology, within which we can always define a set of the form $N_\epsilon(x)$ for some x and some ϵ. Hence, these sets do, in fact, form a basis for a topology on $\mathbb{R}^4$. We denote $\mathbb{R}^4$ with this topology by M^4, and we call it Minkowski space, or Minkowski spactime, and we refer to the topology as the spacetime topology.

Every open set in the Euclidean topology is open with respect to the spacetime topology. As we have already noted, the converse is not true, since $N_\epsilon(x)$ is not open in the Euclidean topology. Thus, the spacetime topology is finer than the Euclidean topology. From this, it follows that M^4 is Hausdorff and second countable, so it is a 4-dimensional manifold. There is only one coordinate chart, namely the identity map taking each $x \in M^4$ to $x \in \mathbb{R}^4$. Since Q has index 1 at each point $x \in M^4$, it follows that M^4 is a Lorentzian manifold. ■

Example 7.2 *The Lorentzian Cylinder*

The standard 2-dimensional cylinder in $\mathbb{R}^3$ is denoted C^2. We will denote the

Lorentzian cylinder by C_L^2. If we imbed C^2 in $\mathbb{R}^3$ in the standard fashion, so that it is centered around the $z - axis$, then it is fairly straightforward to see that there is a smooth field of lines on C^2. We can define a smooth 1-dimensional distribution, L^1, by simply letting $L^1(x, y, z)$ be the line passing through (x, y, z) parallel to the $z - axis$. This line is spanned by the vector e_z, the standard unit vector, anchored at (x, y, z), in the direction of the positive $z - axis$.

As a manifold, C_L^2 is the same as C^2. We found in example 6.3 that the Riemannian metric induced on C^2 by imbedding it in $\mathbb{R}^3$ is given by $G = d\theta \otimes d\theta + dz \otimes dz$. We can define a Lorentzian metric tensor on C^2 by following the method used in the proof of Theorem 7.4. For any point $p \in C^2$, let u_p be e_z. Define a bilinear form, $\widehat{G}$, on C^2 by

$$\widehat{G}_p(v_p, w_p) = G_p(v_p, w_p) - 2G_p(e_z, v_p)G_p(e_z, w_p),$$

where v_p and w_p are arbitrary tangent vectors at p. Now, the tangent space T_pM is spanned by the two real-valued operators that give the partial derivatives of smooth functions with respect to the coordinates θ and z. We can naturally identify the latter with the vector e_z. Consequently, we see that $dz(e_z) = 1$ and $d\theta(e_z) = 0$. Then, by simply computing, we obtain

$$\begin{aligned} G_p(v_p, w_p) &= d\theta(v_p)d\theta(w_p) + dz(v_p)dz(w_p) \\ &\qquad - 2[d\theta(e_z)d\theta(v_p) + dz(e_z)dz(v_p)][d\theta(e_z)d\theta(w_p) + dz(e_z)dz(w_p)] \\ &= d\theta(v_p)d\theta(w_p) + dz(v_p)dz(w_p) - 2[dz(v_p)][dz(w_p)] \\ &= d\theta(v_p)d\theta(w_p) - dz(v_p)dz(w_p) \\ &= (d\theta \otimes d\theta - dz \otimes dz)(v_p, w_p). \end{aligned}$$

Thus, the Lorentzian cylinder, C_L^2, is the standard cylinder with Lorentzian metric tensor $d\theta \otimes d\theta - dz \otimes dz$. ■

Admittedly, these examples are elementary. We simply wanted to show how Lorentzian metric tensors might be constructed on smooth manifolds.

8 Connections and Covariant Differentiation

still in progress

Definition 1 *A **connection**, D, on a smooth manifold, M, is a mapping $D : \mathscr{X}(M) \times \mathscr{X}(M) \to \mathscr{X}(M)$ such that*

i) $D_V W$ *is* $\mathscr{F}(M)$*-linear in* V

ii) $D_V W$ *is* $\mathbb{R}$*-linear in* W

iii) $D_V(fW) = fD_V W + (Vf)W$ *for* $f \in \mathscr{F}(M)$.

$D_V W$ is called the **covariant derivative** of W with respect to V for D. Other connections will produce different derivatives. May also be referred to as an affine connection.

**Show tensorial property of D in V, so $(D_V W)_p$ only depends on V at p.

Lemma 1 *Let $X, Y \in \mathscr{X}(M)$. Suppose either $X = 0$ or $Y = 0$ on an open set $U \subset M$. If D is a connection on M, then $D_X Y = 0$ on U.*

Lemma 2 *Let $p \in M$ be given. If $X, Y \in \mathscr{X}(M)$ are such that $X_p = Y_p$, then for every vector field Z, $(D_X Z)_p = (D_Y Z)_p$. If we denote this uniquely determined vector by $D_{X_p} Z$, then the map from T_pM to T_pM defined by $X_p \mapsto D_{X_p} Z$ is linear for every fixed Z.*

This allows us to define the restriction of a connection to a local neighborhood, so we can give meaning to $D|_U$, for $U \subset M$.

For a coordinate chart (φ, U) with basis field $\{\partial_i\}$, we have

$$D_{\partial_i}(\partial_j) = \sum_k \Gamma_{ij}^k \partial_k.$$

Theorem 1 ***Existence of Covariant Derivative Along a Curve*** *Let M be a smooth manifold with a connection D. Let $\alpha : I \to M$ be a smooth curve in M. There exists a unique function $Z \mapsto Z' := DZ/dt$ from $\mathscr{X}(\alpha)$ to $\mathscr{X}(\alpha)$ such that*

i) $\frac{D}{dt}(aV + bW) = a\frac{DV}{dt} + b\frac{DW}{dt}$ *for* $a, b \in \mathbb{R}$ *and* $V, W \in \mathscr{X}(M)$.

ii) $\frac{D}{dt}(fV) = f\frac{DV}{dt} + f'V$ *for* $f \in \mathscr{F}(I)$.

iii) *If* $V \in \mathscr{X}(M)$ *and* V_α *denotes the restriction of* V *to the curve* $\alpha(I)$, *then* $V'_\alpha(t) = D_{\alpha'(t)}V$.

Locally, if $V = \sum_i V^i \partial_i$, then

$$\frac{DV}{dt} = \sum_k \left[\frac{dV^k}{dt} + \sum_{i,j} \frac{dx^i}{dt} V^j \Gamma^k_{ij} \right] \partial_k.$$

Definition 2 *V along α is **parallel** if $DV/dt = 0$ for all $t \in I$.*

Theorem 2 ***Parallel Transport*** *Let M be a smooth manifold with a connection D. Let $\alpha : I \to M$ be a smooth curve in M, $V_0 \in T_{\alpha(t_0)}M$, $t_0 \in I = (a,b)$. Then there exists a unique parallel vector field, V, along α such that $V(t_0) = V_0$. $V(t)$ is the **parallel transport** of $V(t_0)$ to $\alpha(t)$ along α.*

Notes: Prove by ODE theory. Parallel transport is path dependent. It is an isometry in the case of a (semi-)Riemannian manifold.

There are 2 other properties a connection on a manifold may satisfy. If $[V, W] = D_V W - D_W V$ for $V, W \in \mathscr{X}(M)$, then D is **symmetric** or **torsion free**. If M is a (semi-)Riemannian manifold with a metric tensor $\langle , \rangle$, then we say that D is compatible with the metric if for any smooth curve $\alpha : I \to M$ and any parallel vector fields V, W along α, we have $\langle V, W \rangle(t) =$constant.

Lemma 3 *TFAE*

i) *D is compatible with $\langle , \rangle$.*

ii) *For any $\alpha : I \to M$ and any $V, W \in \mathscr{X}(\alpha)$, we have*

$$\frac{d}{dt}\langle V, W \rangle = \left\langle \frac{DV}{dt}, W \right\rangle + \left\langle V, \frac{DW}{dt} \right\rangle \ t \in I$$

iii) $X\langle Y, Z \rangle = \langle D_X Y, Z \rangle + \langle Y, D_X Z \rangle$ *for* $X, Y, Z \in \mathscr{X}(M)$.

Theorem 3 *On a semi-Riemannian manifold, M, there exists a unique connection D such that D is symmetric and compatible with $\langle , \rangle$.*

Corollary 4 *On a semi-Riemannian manifold, M, on a coordinate system (φ, U), we have*

$$\Gamma^k_{ij} = \frac{1}{2} \sum_m g^{km} \Big(\frac{\partial g_{jm}}{\partial x^i} + \frac{\partial g_{im}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^m} \Big).$$

9 Geodesics, Normal Neighborhoods, and Hopf-Rinow

in progress

10 Curvature

*****Also consider sections on Jacobi fields, differential forms, local isometries, .******

Appendix

A.1 Proofs of Lemmas 2.6 and 2.7

We used two lemmas from general topology to construct the projective manifold in Chapter 2. The proofs of those lemmas are given here.

Definition 1 *An equivalence relation $\sim$ on a topological space X is said to be open if , for any open subset, $A \subset X$, the set $[A]$ is open in X, where $[A] = \bigcup_{x \in A}[x]$.*

Lemma 1 *An equivalence relation $\sim$ on X is open if and only if the quotient map, p, is an open mapping. When $\sim$ is open and X is second countable, then $X/\sim$ is second countable also.*

Proof First, suppose $\sim$ is open. To show that p is an open mapping, we must show that it maps open sets in X to open sets in $X/\sim$. By definition of the quotient topology, a set $\Omega \subset X/\sim$ is open if $p^{-1}(\Omega)$ is open. Let A be an open subset of X. Then the set $[A]$ is also open in X. Note, however, that $p^{-1}(p(A)) = [A]$. Thus, it follows that $p^{-1}(p(A))$ is open, which, in turn, implies that $p(A)$ is open in $X/\sim$. Since A was arbitrary, this shows that p is an open map.

Conversely, suppose p is an open map. Let A be any open subset of X. Then $p(A)$ is open in $X/\sim$. Since p is continuous, this implies that $p^{-1}(p(A))$ is open in X. Since $p^{-1}(p(A)) = [A]$, it follows that $[A]$ is open. So, $\sim$ is an open equivalence relation.

Finally, suppose $\sim$ is open, and suppose that X is second countable. Let $\{U_i\}_{i=1}^{\infty}$ be a countable basis of open sets for X. Then the collection $\{p(U_i)\}_{i=1}^{\infty}$ is a countable collection of open sets in $X/\sim$. Let W be an open subset of $X/\sim$. Then $p^{-1}(W)$ is an open subset of X, implying that $p^{-1}(W)$ can be expressed as a union of basis sets U_i. Suppose

$$p^{-1}(W) = \bigcup_{j=1}^{\infty} U_{i_j}.$$

Then $W = p(p^{-1}(W)) = p(\cup_{j \geq 1} U_{i_j}) = \cup_{j \geq 1} p(U_{i_j})$. Hence, W can be expressed as the union of sets in the collection $\{p(U_i)\}$. Since W was arbitrary, this shows that this collection forms a basis for $X/\sim$. Thus, the quotient space is second countable.

QED

Lemma 2 *Let $\sim$ be an open equivalence relation on X. Then $R = \{(x, y) \in X \times X : x \sim y\}$ is a closed subset of $X \times X$ (with the standard product topology) if and only if $X/\sim$ is Haudsdorff.*

<u>Proof</u> First, suppose $X/\sim$ is Hausdorff. We will show that R is closed by showing that its complement is open. Suppose $(x, y) \notin R$. Then $x \nsim y$, which implies that $p(x) \neq p(y)$. Hence, there are disjoint open subsets, U and V, of $X/\sim$ such that $p(x) \in U$ and $p(y) \in V$. Then $p^{-1}(U)$ and $p^{-1}(V)$ are open sets in X containing x and y, respectively. So, the set $p^{-1}(U) \times p^{-1}(V)$ is an open set in $X \times X$ that contains (x, y). Suppose this set intersects R. Then it must contain some element (x', y') such that $x' \sim y'$. This implies that $p(x') = p(y')$. But (x', y') is also in $p^{-1}(U) \times p^{-1}(V)$, implying that $p(x') \in U$ and $p(y') \in V$. It follows that U and V have a nonempty intersection, contradicting the fact that they are disjoint. Thus, the set $p^{-1}(U) \times p^{-1}(V)$ does not intersect R. Therefore, for any $(x, y) \in R^c$, there is an open set containing (x, y) that does not intersect R. So, R^c is open, implying that R is closed.

Conversely, suppose R is closed. Since any point in $X/\sim$ is the image of some element under the map p, we can consider two elements of $X/\sim$ as $p(x)$ and $p(y)$. Suppose $p(x)$ and $p(y)$ are two distinct points of $X/\sim$. Then we cannot have $x \sim y$. Thus, the element (x, y) must be in R^c, which is open. Hence, there are open sets U and V in X such that $(x, y) \in U \times V$ and the set $U \times V$ does not intersect R. Consider the sets $p(U)$ and $p(V)$ in $X/\sim$. If there was an element, say $[z]$, in both of these sets, then we would have $x \sim z$ and $y \sim z$, implying that $x \sim y$. Hence, this contradiction shows that $p(U)$ and $p(V)$ are disjoint. The previous lemma shows that they are open, and we have $p(x) \in p(U)$ and $p(y) \in p(V)$. Since these two elements were arbitrary, this shows that $X/\sim$ is Hausdorff.

QED

A.2 Additional Remarks on the Tangent Space

When constructing a basis for the tangent space, T_pM, at a point $p \in M$, we denoted the basis vectors induced by a particular coordinate system, (φ, U), by E_i, or E_{p_i}, where these operators are defined on smooth functions in a neighborhood of p by

$$E_i(f) = \frac{\partial(f \circ \varphi^{-1})}{\partial x^i}\Big|_{\varphi(p)}.$$

In this relation, x^i denotes the i^{th} coordinate function of the map φ. Thus, E_i is just the partial derivative operator at p with respect to the i^{th} coordinate. Consequently, it is common in differential geometry to denote this basis vector simply by the partial derivative operator symbol

$$\frac{\partial}{\partial x^i}.$$

We chose not to use this notation during our construction of the basis, since it explicitly refers to a coordinate system by using the coordinate function x^i. However, once the nature of the tangent vectors is understood, this representation is actually quite useful. Most notably, it tells us directly that a tangent vector is a partial derivative operator, whereas the operator E_i does not. Using this notation, the basis for the tangent space T_pM induced by the coordinate system (φ, U), where the coordinate functions of φ are denoted by x^i, can be written as

$$\left\{\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}\right\}.$$

Recall that, using our standard notation, the basis for T_p^*M is denoted $\{dx^i\}$. We have shown that this is the dual basis to the standard tangent basis. Thus, in our new notation, we have the relations

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

Having introduced this new notation, we can give another interpretation of a tangent vector that is distinct from, but equivalent to, our original one. Recall that in our construction of the cotangent space, we used the set Γ_p, the set of all smooth curves, $\gamma : I \to M$ such that $\gamma(0) = p$. We will define an equivalence relation, $\sim$, on Γ_p as follows. Let $\gamma_1 \sim \gamma_2$ if and only if there is a coordinate system, (φ, U) around p such that $(\varphi \circ \gamma_1)'(0) = (\varphi \circ \gamma_2)'(0)$. This is easily seen to be an equivalence relation, and it simply states that two curves are equivalent if their coordinate tangent vectors at p are the same. We will denote the equivalence class containing γ by $\tilde{\gamma}$.

We should first point out one consequence of our definition. If two curves are equivalent with respect to one chart (φ, U), then they are also equivalent with respect to any other chart containing p. To see this, suppose $\gamma_1 \sim \gamma_2$. Then there is a chart (φ, U) around p such that $(\varphi \circ \gamma_1)'(0) = (\varphi \circ \gamma_2)'(0)$. Let (ψ, V) be another coordinate system around p. Then

$$
\begin{aligned}
(\psi \circ \gamma_1)'(0) &= (\psi \circ \varphi^{-1} \circ \varphi \circ \gamma_1)'(0) \\
&= D_{\varphi(p)}(\psi \circ \varphi^{-1})(\varphi \circ \gamma_1)'(0) \\
&= D_{\varphi(p)}(\psi \circ \varphi^{-1})(\varphi \circ \gamma_2)'(0) \\
&= (\psi \circ \varphi^{-1} \circ \varphi \circ \gamma_2)'(0) \\
&= (\psi \circ \gamma_2)'(0).
\end{aligned}
$$

Thus, we need only verify that two curves are equivalent with respect to a single chart at p, and it will follow that they have the same coordinate derivative with respect to any coordinate system.

Consider the set $\Gamma_p/\sim$. We will continue to let T_pM denote the tangent space at $p \in M$ as we have already constructed it. So, elements of T_pM are still considered as operators on $\mathscr{F}_p$, the of smooth functions defined on a neighborhood of p. Our goal is to show that $\Gamma_p/\sim\, = T_pM$. More precisely, we should say that $\Gamma_p/\sim\, \cong T_pM$, as we will have to define an identification between equivalence classes of curves and derivations on $\mathscr{F}_p$.

We begin by uniquely associating to each $\tilde{\gamma} \in \Gamma_p$ a derivation on $\mathscr{F}_p$, or, simply, a tangent vector. For $\tilde{\gamma} \in \Gamma_p/\sim$, define an element $T_\gamma \in T_pM$ such that for $\tilde{f} \in \mathscr{F}_p$,

$$
T_\gamma(\tilde{f}) = \frac{d}{dt}(f \circ \gamma)\Big|_{t=0}.
$$

This definition is independent of the choice of curve $\gamma \in \tilde{\gamma}$ and of the choice of function $f \in \tilde{f}$, since any two functions in $\tilde{f}$ will be equal on a neighborhood of p and any two curves in $\tilde{\gamma}$ will have the same coordinate derivatives at p with respect to any chart. Thus, it is not ambiguous to denote the operator by T_γ, without the tilde. Of course, we still must show that our definition makes sense. That is, we must know that T_γ actually is a tangent vector. Each T_γ is a linear operator on $\mathscr{F}_p$, since for $\alpha, \beta \in \mathbb{R}$ and $\tilde{f}, \tilde{g} \in \mathscr{F}_p$, we have

$$
\begin{aligned}
T_\gamma(\alpha\tilde{f} + \beta\tilde{g}) &= T_\gamma(\widetilde{\alpha f + \beta g}) \\
&= \frac{d}{dt}\big((\alpha f + \beta g) \circ \gamma\big)\Big|_{t=0} \\
&= \frac{d}{dt}\big(\alpha(f \circ \gamma) + \beta(g \circ \gamma)\big)\Big|_{t=0} \\
&= \alpha\frac{d}{dt}(f \circ \gamma)|_{t=0} + \beta\frac{d}{dt}(g \circ \gamma)|_{t=0} \\
&= \alpha T_\gamma(\tilde{f}) + \beta T_\gamma(\tilde{g}).
\end{aligned}
$$

Likewise, T_γ satisfies the Leibniz property, since for $\tilde{f}, \tilde{g} \in \mathscr{F}_p$, we have

$$\begin{aligned} T_\gamma(\tilde{f}\tilde{g}) &= T_\gamma(\widetilde{fg}) \\ &= \frac{d}{dt}\big((fg)\circ\gamma\big)\Big|_{t=0} \\ &= \frac{d}{dt}\big((f\circ\gamma)(g\circ\gamma)\big)\Big|_{t=0} \\ &= f(p)\frac{d}{dt}\big(g\circ\gamma\big)\Big|_{t=0} + g(p)\frac{d}{dt}\big(f\circ\gamma\big)\Big|_{t=0} \\ &= f(p)T_\gamma(\tilde{g}) + g(p)T_\gamma(\tilde{f}). \end{aligned}$$

Thus, each T_γ is a tangent vector. Moreover, this association is unique. If $\tilde{\gamma}_1$ induces T_{γ_1} and $\tilde{\gamma}_2$ induces T_{γ_2} with $T_{\gamma_1} = T_{\gamma_2}$, then for any $\tilde{f} \in \mathscr{F}_p$ we have

$$\frac{d}{dt}(f\circ\gamma_1)\Big|_{t=0} = \frac{d}{dt}(f\circ\gamma_2)\Big|_{t=0}.$$

Let (φ, U) be a coordinate system around p. Since the above equality must hold for any smooth function f, we can let $f = x^i$, the i^{th} coordinate function of the mapping φ. Recall that this mapping is given by $\pi^i\circ\varphi$, where π^i is the i^{th} coordinate projection on $\mathbb{R}^m$. We then obtain

$$\begin{aligned} \frac{d}{dt}\big(x^i\circ\varphi^{-1}\circ\varphi\circ\gamma_1 - x^i\circ\varphi^{-1}\circ\varphi\circ\gamma_2\big)\Big|_{t=0} &= \frac{d}{dt}\big(\pi^i\circ\varphi\circ\gamma_1 - \pi^i\circ\varphi\circ\gamma_2\big)\Big|_{t=0} \\ &= \frac{d}{dt}\big(\pi^i\circ(\varphi\circ\gamma_1 - \varphi\circ\gamma_2)\big)\Big|_{t=0} \\ &= \dot{x}_1^i(0) - \dot{x}_2^i(0), \end{aligned}$$

where $\dot{x}_k^i(0)$ denotes the derivative if the i^{th} coordinate function of $\varphi\circ\gamma_k$ at $t = 0$. It follows that $\dot{x}_1^i(0) = \dot{x}_2^i(0)$ for each $i = 1, \ldots, m$. That is, we have $(\varphi\circ\gamma_1)'(0) = (\varphi\circ\gamma_2)'(0)$, implying that $\tilde{\gamma}_1 = \tilde{\gamma}_2$.

Now, let T_p^cM be the set defined by

$$T_p^cM = \{T_\gamma \,:\, \tilde{\gamma} \in \Gamma_P/\sim\}.$$

Then, by what we have shown, we know that $T_p^cM \subset T_pM$, and that the association between elements of T_p^cM and elements of T_pM is one to one. To conclude, we will show that, in fact, $T_p^cM = T_pM$, implying that each tangent vector can be identified with a unique equivalence class of curves. To do so, we will need to use local coordinate systems. Thus, we will also need to show the identification does not depend on the particular coordinate system chosen.

Let X be any tangent vector in T_pM, and let (φ, U) be any coordinate system around p. We must find a curve, $\gamma \in \Gamma_p$ such that $T_\gamma = X$. Our coordinate system induces the standard basis for T_pM, which, in our new notation, is denoted

$$\left\{\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^m}\right\}.$$

Then X has a representation with respect to this basis given by

$$X = \sum_{i=1}^{m} \alpha^i \frac{\partial}{\partial x^i},$$

for some set of scalars $\alpha^1, \dots, \alpha^m$. Define a curve $\lambda : \mathbb{R} \to \mathbb{R}^m$ by

$$\lambda(t) = \Big(\alpha^1 t + (\varphi(p))^1, \dots, \alpha^m t + (\varphi(p))^m\Big).$$

Then define a curve, γ, in M by $\gamma(t) = (\varphi^{-1} \circ \lambda)(t)$. This is a smooth curve in M and it satisfies $\gamma(0) = \varphi^{-1}(\lambda(0)) = \varphi^{-1}(\varphi(p)) = p$. So, γ defines an equivalence class $\tilde{\gamma} \in \Gamma_p/\sim$, and we have $(\varphi \circ \gamma)'(0) = (\alpha^1, \dots, \alpha^m)$. The operator T_γ induced by this equivalence class is given by

$$\begin{aligned}
T_\gamma(f) &= \frac{d}{dt}\big(f \circ \gamma\big)\Big|_{t=0} \\
&= \frac{d}{dt}\big(f \circ \varphi^{-1} \circ \varphi \circ \gamma\big)\Big|_{t=0} \\
&= D_{\varphi(p)}(f \circ \varphi^{-1})(\varphi \circ \gamma)'(0) \\
&= \sum_{i=1}^{m} \alpha^i \frac{\partial (f \circ \varphi^{-1})}{\partial x^i}\Big|_{\varphi(p)}.
\end{aligned}$$

Recalling the definition of our tangent basis vectors and our new notation, this last equality is simply

$$T_\gamma(f) = \sum_{i=1}^{m} \alpha^i \frac{\partial}{\partial x^i}(f),$$

or, in operator form,

$$T_\gamma = \sum_{i=1}^m \alpha^i \frac{\partial}{\partial x^i} = X.$$

Thus, every tangent vector $X \in T_pM$ can be associated with an equivalence class of curves in $\Gamma_p/\sim$. However, we used only one particular coordinate system in this proof. In order to show that the identification is natural, we must show that if X is associated to $\tilde{\gamma}_1$ under the coordinate system (φ, U) and to $\tilde{\gamma}_2$ under another coordinate system, say (ψ, V), then $\tilde{\gamma}_1 = \tilde{\gamma}_2$.

Let the coordinate functions of φ be denoted by x^i and let the coordinate functions of ψ be denoted by y^i. Further, let the bases for T_pM induced by these coordinate systems be denoted, respectively, by

$$\left\{\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}\right\} \qquad \textit{and} \qquad \left\{\frac{\partial}{\partial y^1}, \ldots, \frac{\partial}{\partial y^m}\right\}.$$

Suppose that with respect to these two bases, X has representations

$$X = \sum_{i=1}^m \alpha^i \frac{\partial}{\partial x^i} = \sum_{j=1}^m \beta^j \frac{\partial}{\partial y^j}.$$

We know that the association between tangent vectors and equivalence classes of curves under a particular coordinate system is one to one. Thus, with respect to the coordinate system (φ, U), we associate X with the equivalence class $\tilde{\gamma}_1$, where $\gamma_1 = \varphi^{-1} \circ \lambda_1$ and

$$\lambda_1(t) = \Big(\alpha^1 t + (\varphi(p))^1, \ldots, \alpha^m t + (\varphi(p))^m\Big).$$

Likewise, with respect to the coordinate system (ψ, V), we associate X with the equivalence class $\tilde{\gamma}_2$, where $\gamma_2 = \psi^{-1} \circ \lambda_2$ and

$$\lambda_2(t) = \Big(\beta^1 t + (\psi(p))^1, \ldots, \beta^m t + (\psi(p))^m\Big).$$

Since equivalence of two curves with respect to one chart implies equivalence with respect to all charts, we need only show that $(\varphi \circ \gamma_1)'(0) = (\varphi \circ \gamma_2)'(0)$. Now, we know that $(\varphi \circ \gamma_1)'(0) = (\alpha^1, \ldots, \alpha^m)$. So, consider $(\varphi \circ \gamma_2)'(0)$. This can be rewritten as

$$\begin{aligned}(\varphi \circ \gamma_2)'(0) &= (\varphi \circ \psi^{-1} \circ \lambda_2)'(0) \\ &= D_{\psi(p)}(\varphi \circ \psi^{-1})(\beta^1, \ldots, \beta^m),\end{aligned}$$

where $D_{\psi(p)}(\varphi\circ\psi^{-1})$ is the Jacobian of the change of coordinates map, $\varphi\circ\psi^{-1}$, evaluated at $\psi(p)$. We have already discussed the structure of this Jacobian and changes of coordinates in Chapter 2. The ij^{th} entry of this matrix is just the coordinate partial derivative $\partial x^i/\partial y^j$. Thus, we see that

$$(\varphi\circ\gamma_2)'(0) = \left(\sum_{i=1}^{m}\frac{\partial x^1}{\partial y^i}\beta^i, \ldots, \sum_{i=1}^{m}\frac{\partial x^m}{\partial y^i}\beta^i\right).$$

But, by Theorem 3.5, where we showed how components of a tangent vector transform under a change of coordinates, the sum

$$\sum_{i=1}^{m}\frac{\partial x^j}{\partial y^i}\beta^i$$

is equal to the component, α^j, of X. Thus, we see that $(\varphi\circ\gamma_2)'(0) = (\alpha^1, \ldots, \alpha^m) = (\varphi\circ\gamma_1)'(0)$. This implies that $\tilde{\gamma_1} = \tilde{\gamma_2}$, and our association of equivalence classes of curves with tangent vectors is not coordinate dependent.

So, we can uniquely associate to each tangent vector, X, an equivalence class of curves, $\tilde{\gamma}$, such that the operator T_γ satisfies $X(\tilde{f}) = T_\gamma(\tilde{f})$ for all $\tilde{f}\in\mathscr{F}_p$. Moreover, we see by our construction that the association has an intuitively geometric basis. The equivalence class of curves, $\tilde{\gamma}$, that we associate to $X\in T_pM$ are simply the curves whose tangent vectors denote the direction in which X gives us the directional derivative of a function f at p. Conversely, the tangent vector X gives us the directional derivative of a function f in the direction of the tangent vectors of the curves in $\tilde{\gamma}$ at p.

A.3 Addendum to Theorem 6.5

In Theorem 6.5, we showed that the function $d(p,q) = \inf\{L_\gamma : \gamma : [a,b]\to M, \gamma(a) = p, \gamma(b) = q, \gamma\ is\ C^\infty\}$ defined a metric on a smooth, connected manifold, M. Part of that proof relied on the fact that the value of this metric on the Euclidean manifold $\mathbb{R}^m$ is attained by the straight line connecting the two points in question. That is, in Euclidean manifolds, the shortest distance between two points is, in fact, a straight line. We used that result without proof there, so we wish to give a proof here.

This result was a fundamental postulate in Euclidean geometry, but, when thinking of $\mathbb{R}^m$ as a manifold, the result is not at all obvious. The result can be obtained as a simple problem in the Calculus of Variations, but we do not require any knowledge of this subject here. Hence, we will give an independent proof without referring to this topic.

Theorem *Let p and q be two points in $\mathbb{R}^m$, with the metric, d, defined as before. Then the value of $d(p,q)$ is attained by the straight line connecting p and q. That is, $d(p,q) = \|p-q\|$.*

Proof We can assume without loss of generality that the initial point is the origin, since arclength is invariant under translations. Let $\gamma : [0,1] \to \mathbb{R}^m$ be a curve in $\mathbb{R}^m$ such that $\gamma(0) = 0$ and $\gamma(1) = p$. Note that by a suitable reparametrization, we can assume that all curves connecting 0 and p are defined on $[0,1]$. We will assume that γ is piecewise smooth and regular. Moreover, we will make the assumption that $\|\gamma(t)\| > 0$ for all $t > 0$. That is, we do not want γ to circle back to its initial point before reaching the final point, p. If a curve does do so, that case can be reduced to this one.

Let $\lambda : [0,1] \to \mathbb{R}$ be a real-valued function such that $\lambda(t) = \|\gamma(t)\|$. Then, if $v(t)$ denotes a unit vector function on $[0,1]$, then we can write $\gamma(t) = \lambda(t)v(t)$. Since γ is piecewise smooth, so are the functions λ and v. We want to compute $\langle \gamma'(t), \gamma'(t)\rangle = \|\gamma'(t)\|^2$. First, note that

$$\gamma'(t) = \lambda'(t)v(t) + \lambda(t)v'(t).$$

Since, γ is only piecewise smooth, this derivative will not be defined for all t, but it will be defined almost everywhere. In fact, by our conditions on the curves connecting two points, the set of discontinuities of $\gamma'(t)$ will be finite, and γ' will be piecewise continuous. Hence, it will still make sense to integrate $\|\gamma'(t)\|$ using the basic Riemann integral. Thus, for almost every $t \in [0,1]$, we have

$$\begin{aligned}
\langle \gamma'(t), \gamma'(t)\rangle &= \langle \lambda'(t)v(t) + \lambda(t)v'(t), \lambda'(t)v(t) + \lambda(t)v'(t)\rangle \\
&= \langle \lambda'(t)v(t), \lambda'(t)v(t)\rangle + 2\langle \lambda'(t)v(t), \lambda(t)v'(t)\rangle + \langle \lambda(t)v'(t), \lambda(t)v'(t)\rangle \\
&= (\lambda'(t))^2\|v(t)\|^2 + 2\lambda'(t)\lambda(t)\langle v(t), v'(t)\rangle + (\lambda(t))^2\|v'(t)\|^2 \\
&= (\lambda'(t))^2 + (\lambda(t))^2\|v'(t)\|^2,
\end{aligned}$$

Now, the length of the curve, γ is given by

$$L_\gamma = \int_0^1 \|\gamma'(t)\|dt = \int_0^1 \sqrt{\langle \gamma'(t), \gamma'(t)\rangle}dt.$$

Hence, we see that

$$
\begin{aligned}
L_\gamma &= \int_0^1 \sqrt{(\lambda'(t))^2 + (\lambda(t))^2\|v'(t)\|^2}dt \\
&\geq \int_0^1 \sqrt{(\lambda'(t))^2}dt \\
&\geq \int_0^1 |\lambda'(t)|dt \\
&\geq \Big|\int_0^1 \lambda'(t)dt\Big| \\
&\geq \big|\lambda(1) - \lambda(0)\big| \\
&\geq \|\gamma(1)\| \quad (\Leftarrow \lambda(0) = \|\gamma(0)\| = 0) \\
&\geq \|\gamma(1) - \gamma(0)\| \\
&\geq \|p - 0\|.
\end{aligned}
$$

So, this shows that $L_\gamma \geq \|\gamma(1)-\gamma(0)\| = \|p-0\|$, and, since the curve γ was arbitrary, it follows that $\|p - 0\| \leq d(0, p)$. But the straight line connecting 0 and p is an admissible curve, and its arclength is just $\|p - 0\|$, so we also have $d(0, p) \leq \|p - 0\|$. Hence, we have $d(0, p) = \|p - 0\|$.

QED

BIBLIOGRAPHY AND FURTHER READING

[Ab] ABRAHAM, R., and MARSDEN, J. *Foundations of Mechanics*, second edition. Benjamin-Cummings, New York. (1979).

[Ba] BARTLE, R.G. *The Elements of Real Analysis*, second edition. Wiley, New York. (1964).

[Bo] BOOTHBY, W.M. *An Introduction to Differentiable Manifolds and Riemannian Geometry*, second edition. Academic Press, San Diego, California. (1975).

[Ch] CHERN, S.S., CHEN, W.H., and LAM, K.S. *Lectures on Differential Geometry*. World Scientific, Singapore. (1999).

[Cu] CURTIS, W.D., and MILLER, F.R. *Differential Manifolds and Theoretical Physics*. Academic Press, Orlando, Florida. (1985).

[Da] DARLING, R.W.R. *Differential Forms and Connections*. Cambridge University Press, Cambridge, United Kingdom. (1994).

[Di] DIEUDONNE, J. *Foundations of Modern Analysis*. Academic Press, New York. (1960).

[Do] DO CARMO, M.P.

[i] *Differential Geometry of Curves and Surfaces*. Prentice Hall, Englewood Cliffs, New Jersey. (1976).

[ii] *Riemannian Geometry*. Springer-Verlag, New York. (1992).

[Ga] GAUSS, C.F. *General Investigations of Curved Surfaces*. Raven Press, Hewlett, New York. (1965).

[Hil] HILBERT, D. *Foundations of Geometry*. (Translated by Leo Unger). Open Court Publishing Co., La Salle, Illinois. (1971).

[Hir] HIRSCH, M.W. *Differential Topology*. Springer-Verlag, New York. (1997).

[Ko] KOBAYASHI, S. and NOMIZU, K. *The Foundations of Differential Geometry*. Vols. I and II. Wiley, New York. (1966).

[Le] LEVI-CEVITA, T. *The Absolute Differential Calculus* (*CalculusofTensors*). (Translated by Marjorie Long). Dover Publications, New York. (1977).

[ON] O'NEILL, B.

[1] *Elementary Differential Geometry*. Academic Press, New York. (1966).

[2] *Semi-Riemannian Geometry* (publication details to be included)

[Mu] MUNKRES, J.R. *Topology, A First Course*. Prentice Hall, Inc., Englewood Cliffs, New Jersey. (1975).

[Na] NABER, G. *The Geometry of Minkowski Spacetime*. Springer-Verlag, New York. (1992).

[Sp] SPIVAK, M.

[1] *Calculus on Manifolds*. Perseus Press, New York. (1965).

[2] *A Comprehensive Introduction to Differential Geometry*, vols. I and II. Publish or Perish, Inc., Wilmington, Delaware. (1979).

[Ste] STEENROD, N. *The Topology of Fiber Bundles*, Princeton University Press, Princeton, New Jersey. (1999).

[Str] STRUIK, D.J. *Lectures on Classical Differential Geometry*, second edition. Dover, New York. (1988).

[Wa] WALD, R.M. *General Relativity*. The University of Chicago Press, Chicago, Illinois. (1994).

[Ze] ZEEMAN, E.C. "The Topology of Minkowski Space," Topology, 6(1967), 161-170.