Math 856 Differential Topology

Ever-expanding course notes

Differential topology is about introducing concepts and methods from calculus to the realm of topological spaces. That is, we wish to use notions of differentiation and integration in a topological setting. There are (at least) two reasons for doing so. The first is, essentially, waste not want not; lots of people have put in a lot of effort into developing the tools of analysis, why shouldn't topologists want to take advantage of all of that body of work? Any tool that we can bring to bear to better understand topological spaces helps us, well, understand topological spaces better. The other reason is that by figuring out how to introduce analysis into topology, we will have extended the range of applicability of these concepts. Experience has also shown that the topological point of view can, in hindsight, provide a more natural setting for many problems of analysis. It can also provide a natural framework for explaining some of its results; Stokes' Theorem is perhaps the first and most well-known, but certainly not the only such result that we may encounter in our study. As with nearly any branches of mathematics, once you figure out how to reconcile the immediate difficulties in introducing one subject to another (analysis to topology, or topology to analysis?), you discover innumerable ways in which they open up new avenues of exploration, and neither subject is ever the same again. The goal of this course is to explore the ways in which we can bring analysis and topology together, and some of the ways in which anaysis helps to illuminate the study of topology.

Our first task is to determine *which* topological spaces we can reasonably introduce such concepts and methods to.

Manifolds: A basic principle in topology is that a topological space is explored through its continuous functions/continuous maps, both in and out of the space. Calculus as we usually encounter it applies to functions between Euclidean spaces $\mathbb{R}^n$. We have derivatives, partial derivatives, integrals, and multipe integrals, and many variations, depending upon what domain or range/codomain we choose for our functions. So if we want to be able to introduce the idea of a "differentiable" map, the simplest tack to take is to look at topological spaces which "behave" like Euclidean spaces. Differentiability is a local property; a (partial) derivative of a function at a point (much less whether you have one, i.e., are differentiable) depends only on the values of the function near that point. Of course the notion of "local" is in some sense what a topology on a space is designed to describe; open neighborhoods of a point x are precisely the sets describing which points are "near" x. So on a most basic level, the topological spaces most naturally to introduce calculus to are those in which the the points have open neighborhoods which "look like" the spaces that we know how to do calculus on, namely, Euclidean spaces. This motivates our first definition.

A *topological manifold* M of dimension n is a Hausdorff, second countable space with the property that for every $x \in M$ there is an open neighborhood U of x which is homeomorphic to $\mathbb{R}^n$.

The shorthand for the last property is that M is *locally Euclidean*. The other two properties, Hausdorffness and second countability, are designed, really, to make the topologists job easier. One occasionally encounters situations in which a locally Euclidean space is either not Hausdorff or not second countable, but they are very much the exception rather than the rule. And being able to assume both conditions when someone starts tossing the term "manifold" around certainly make proving theorems a lot easier. Surely this isn't the first time that you have encountered hypotheses being imposed for the purpose of making theorems easier to prove? At any rate, any subset of a Euclidean space is both Hausdorff and second countable in the subspace topology; most (all?) manifolds we will meet can (with effort) be interpreted as such a subspace. A manifold of dimension n will be called an n-manifold for short. It is not at all clear from the definition, but it is the case that the n of n-manifold is a homeomorphism invariant. At a given point $x \in M$, this follows from a result called the Invariance of Domain; which says that if $U \subseteq \mathbb{R}^n$ is open, and $f : U \to \mathbb{R}^n$ is continuous and one-to-one, then $f(U) \subseteq \mathbb{R}^n$ is also open. (The cleanest proof uses homology theory, and can usually be found in any decent algebraic topology text.) It is a direct consequence that no open set $U \subseteq \mathbb{R}^n$ can be homeomorphic to an open subset of $\mathbb{R}^m$ for $m \neq n$ (and so can't be homeomorphic to $\mathbb{R}^m$, either). This also means that in a connected manifold, every point has neighborhoods locally homeomorphic to the same $\mathbb{R}^n$; this can be verified by the usual trick of showing that for a fixed n, the points with neighborhoods homeomorphic to $\mathbb{R}^n$ is open (and therefore also closed!).

Examples: Some standard examples: Euclidean space $\mathbb{R}^n$ itself. Spheres S^n = the points at unit distance in $\mathbb{R}^{n+1}$; given a point, $x \in S^n$ at least one of its coordinates x_i is non-zero. Then the set of points $y \in S^n$ whose i-th coordinate has the same sign as x form a locally Euclidean neighborhood of x; the homeo to the unit ball in $\mathbb{R}^n$ is given by projection onto the other coordinates. Cartesian products of manifolds are manifolds; take the Cartesian product of neighborhood in each as your local models. Open subsets of manifolds are manifolds. These basic building blocks already let you build a wide variety of examples.

Once we are confortable with the setting, manifolds, into which we will ultimately introduce differentiability, we are left with actually *doing* it. It turns out that in order to do so in a meaningful way, we have to introduce additional "structure"; simply having a topological manifold won't be enough.

Smooth functions: On the face of it, once we have a space M which locally "looks like" Euclidean space, we can seemingly define differentiability at a point for any function $f : M \to \mathbb{R}$. Given a point $x \in M$, we have, by definition, a neighborhood U of X and a homeomorphism $h : U \to \mathbb{R}^n$. This is, at least, enough to *describe* a function for which differentiability makes sense, namely the composition of h^{-1} with the restriction of f to U; $f \circ h^{-1} : \mathbb{R}^n \to \mathbb{R}$. So as a first approximation, we could say that f is differentiable at x if $f \circ h^{-1}$ is differentiable at $h(x)$.

There is only one problem with this. Surely if we are generalizing the notion of diferentiability to more general spaces we don't want to *change* what functions $g : \mathbb{R}^n \to \mathbb{R}$ we wish to consider to be differentable. But, technically, our first attempt at a definition just did. Consider the function $f(x) = |x|$, which we are all, presumably, willing to agree is not differentiable at 0. But we can treat the domain $\mathbb{R} = M$ as a 1-dimensional manifold, where $U = \mathbb{R}$ and the homeomorphism $h(x) = x^{1/3}$ serves as the proof for each $x \in M$ that M is locally Euclidean. But then in testing whether or not f is differentiable at 0, we can just check that $f \circ h^{-1}(x) = |x^3|$ is in fact differentiable at 0. Which it is; the derivative is 0.

Charts: What went wrong? Nothing. Unless you don't want to change the notion of differentiability... The point is, our definition of differentiability mentions both a neighborhood U of x (which won't, in the end, really affect things) and a specific homeomorphism $h : U \to \mathbb{R}^n$. The function f and the point x wasn't enough to define differentiability; we also needed a *chart* (U, h), that is, a specific description for *how* to identify a neighborhood of x with $\mathbb{R}^n$. And whether or not we decide f is differentiable depends on which chart we pick. (In our example above, if we chose the identity map to define our chart, we would have decided that $f(x) = |x|$ is not differentible at 0.) So, in order to *unambiguously* decide if a function is differentiable, we need to restrict which pairs (h, U) we are willing to allow ourselves to use as charts. This special collection of charts is the extra structure that we need.

What is the basic idea? We wish to find some way to ensure that if one of two charts (U, h) and (V, k) with $x \in U, V$ tells us that f is differentiable at x, then the other chart *must* do so, as well. That is, we wish to guarantee that $f \circ h^{-1}$ is differentiable at $h(x)$ iff $f \circ k^{-1}$ is differentiable at $k(x)$. And how to do this? The Chain Rule to the rescue! The thing which connects $f \circ h^{-1}$ to $f \circ k^{-1}$ is a *transition map* $k \circ h^{-1}$; $f \circ h^{-1} = (f \circ k^{-1}) \circ (k \circ h^{-1})$. This equality holds on $h(U \cap V)$, which is the image under a homeo of an open subset of U containing x, so is an open subset of $\mathbb{R}^n$ contining $h(x)$. And if $k \circ h^{-1}$ is differentiable, then Chain Rule tell us that $f \circ k^{-1}$ differentiable at $k(x)$ implies $f \circ h^{-1}$ differentiable at $h(x)$. The reverse implication follows from knowing that $h \circ k^{-1}$ is differentiable.

Atlases: This leads us to our basic construction. A $C^{(k)}$ *atlas* $\mathcal{A}$ on a topological manifold M is a collection (U_i, h_i) of charts on M so that (1) $\bigcup U_i = M$ and (2) for every i, j with $U_i \cap U_j \neq \emptyset$, $h_i \circ h_j^{-1} : h_i(U_i \cap U_j) \to h_j(U_i \cap U_j)$ is $C^{(k)}$, that is, has continuous partial derivatives through order k. Note that notationally, by reversing the roles of i and j, we are also insisting that $h_j \circ h_i^{-1}$ be $C^{(k)}$. Given a $C^{(k)}$ atlas $\mathcal{A}$ on a manifold M, we can then unambiguously define differentiatible functions, or $C^{(m)}$ functions for any $m \leq k$, $f : M \to \mathbb{R}$, by requiring that $f \circ h_i^{-1} : h(U_i) \to \mathbb{R}$ is $C^{(m)}$, for every i. More generally, given atlases on manifolds M, N, we can define a map $f : M \to N$ to be *differentiable* by requiring that $k_j \circ f \circ h_i^{-1}$ is differentiable for every k_j in the atlas for N and h_i in the atlas for M.

It will be useful to introduce some notation at this point, so that we don't have to keep writing " $h \circ k^{-1}$ is $C^{(k)}$ "; we will say that h and k are " $C^{(k)}$-related " if $h \circ k^{-1}$ and $k \circ h^{-1}$ are both $C^{(k)}$.

Smooth structures: A $C^{(k)}$ atlas is enough to be able to define $C^{(m)}$ functions for $m \leq k$, but from a philosophical (and functional) point of view, some atlases are better than others. If $f : M \to \mathbb{R}$ is a $C^{(m)}$ function and (h, U) is a chart on M, and $V \subseteq U$ is open, then $f \circ (h|_V)^{-1} : h(V) \to \mathbb{R}$, as the restriction of $f \circ h^{-1}$, is $C^{(m)}$. In fact, $h|_V$ is $C(k)$-related to every chart on M (if we started with a $C^{(k)}$ atlas), and so it doesn't hurt to add $h|_V$ to our atlas; it won't alter what functions we will call $C^{(m)}$. But it *might* actually help! We re all no doubt familiar with ϵ-δ arguments where we keep shrinking δ (effectively, shrinking the neighborhood of some point x) in order to make better things happen. The same will be true here; we will want to shink the domains of charts in order to make good things happen. It would be nice if such domains were already part of our atlas. So, we do the natural thing; just toss them in. And while we're at it, we might as well toss in everything that we can for free (without changing what we'll call a smooth map). This turns out to be everything which is $C^{(k)}$-related to *everything* already in our atlas. This is also the *largest* $C^{(k)}$ atlas which contains our original atlas. Such an atlas is called a *maximal* atlas.

A $C^{(k)}$ **structure** on a manifold M, $0 \leq k \leq \infty$, is a maximal $C^{(k)}$ atlas on M. M, together with a $C^{(k)}$ structure, will be called a $C^{(k)}$ manifold. A $C^{(0)}$ manifold is "just" a manifold; a $C^{(0)}$ structure is a collection of homeomorphisms from the sets of an open cover of M to $\mathbb{R}^n$ (that the transition maps are $C^{(0)}$, i.e., continuous, is automatic). In general we will content ourselves to study $C^{(\infty)}$ structures on manifolds, but it is important to know that there are other possible choices. (When an author never needs anything beyond a second derivative, they will often talk only about $C^{(2)}$ manifolds, for example. It is a fact (see, e.g., Hirsch, *Differential Topology*, p.51) that for every $1 \leq r \leq s \leq \infty$, a $C^{(r)}$ structure $\mathcal{A}$ on a manifold M *contains* a $C^{(s)}$ structure $\mathcal{B} \subseteq \mathcal{A}$; that is, $\mathcal{A}$ contains an atlas which is $C^{(s)}$-compatible. But we will likely not use this result.)

Examples: Our standard examples of manifolds above also provide some standard examples of smooth manifolds; one merely needs to verify that the charts that we built are $C^{(\infty)}$-related, so that the have an atlas, and then wave our magic wand to 'build' the corresponding maximal atlas. Restriction to an open set and Cartesian product both preserve smoothness, so we have several general approaches to building smooth manifolds at our fingertips.

Diffeomorphisms: Just as in topology we have a notion, homeomorphism, which allows us to treat two spaces as essentially the "same", there is a corresponding notion of same in the smooth setting. Two $C^{(k)}$ manifolds $(M,\mathcal{A}),(N,\mathcal{B})$ are *diffeomorphic* if there is a $C^{(k)}$ bijection $f: M \to N$ with $C^{(k)}$ inverse. Just as with a homeomorphism, a diffeomorphism induces a bijection between charts of M and N, via $h: U \to \mathbb{R}^n$, for $U \subseteq M$, is taken to $h \circ f^{-1}: f(U) \to \mathbb{R}^n$. Because f^{-1} is $C^{(k)}$, this map is $C^{(k)}$, hence is in the (maximal) atlas $\mathcal{B}$.

Some history: Just as in the "standard" definition of topology, the field of differential topology can be most succintly described as the study of the properties of smooth manifolds that are invariant under diffeomorphism (i.e., are defined in terms of the smooth structure). You will have learned in the homework that a given manifold can have many different smooth structures, meaning that the atlases defining them are distinct. But in many cases these atlases can still define the 'same' smooth structure, that is, they are diffeomorphic. In particular, up to diffeomorphism, $\mathbb{R}$,$\mathbb{R}^2$,$\mathbb{R}^3$, and $\mathbb{R}^n$ for $n \geq 5$ all have unique differentiable structure. (Except for $\mathbb{R}$, these are all fairly difficult results to establish!) It was a major breakthrough of the mid-1980's that $\mathbb{R}^4$ was discovered to have more than one smooth structure; it in fact has uncountably many non-diffeomorphic smooth structures. In fact, there are uncountably many open subsets of standard $\mathbb{R}^4$, each homeomorphic to $\mathbb{R}^4$, but (using the smooth structures it inherits from standard $\mathbb{R}^4$) *none* of them diffeomorphic to one another! If this isn't wierd enough for you, consider that, since $\mathbb{R}^5$ has only one smooth structure, up to diffeo, if you take these 'exotic' $\mathbb{R}^4$'s and cross them with $\mathbb{R}$ (with the standard structure), you obtain smooth manifolds, *all* of which are diffeomorphic to standard $\mathbb{R}^5$ (and hence to one another)!

Every 2-manifold has a unique smooth structure up to diffeo (Rado, 1920s?); the same is true for 3-manifolds, as well (Moise, 1950's). But there actually exist 4-manifolds which posess *no* smooth structure. This was first discovered as a result of work of Freedman and Donaldson (for which both received the Fields Medal in 1986). Freedman showed that simply-connected (meaning every map of a circle into M extends to a map of a disk) topological 4-manifolds were determined up to homeo by their 'intersection pairing on second homology' (whatever that is), and further, every unimodular symmetric bilinear pairing has a corresponding manifold. This, by the way, implies the topological 4-dimensional Poincaré conjecture. Donaldson, on the other hand, showed that for simply-connected *smooth* 4-manifolds, certain intersection pairings could not arise (if the pairing is positive definite, then it is diagonalizable). His work essentially involved PDE's on 4-manifolds. In particular, the pairing known as "E8" could not occur. So the 4-manifold "E8", which Freedman's work shows exists, has no smooth structure. Similar examples can be found for all higher dimensions, as well.

On the other hand, there are manifolds which have 'too many' smooth structures, i.e., admit multiple structures which are not diffeomorphic to one another. $\mathbb{R}^4$ is the most famous example these days, but it turns out that most spheres have this property, as well. In the late 1950's John ('Jack') Milnor showed that S^7 has more than one smooth structure; it was later shown that it has exactly 28 non-diffeomorphic structures. S^{31} has more than 16 million! And in case you think these structures are really wierd things that you are never likely to meet, the 28 structures in S^7 arise on the links of singularities of algebraic surfaces. Specifically, the intersection of the solutions (in $\mathbb{C}^5$) to the equation

$$a^2+b^2+c^2+d^3+e^{6k-1}=0$$

with a small sphere centered at the origin, for $k=1,\ldots,28$, gives all 28 exotic 7-spheres (source: the Wikipedia entry for 'exotic sphere'). While wandering the web, I found an assertion (by Ron Stern) that 'all known 4-manifolds have infinitely many distinct smooth structures', but I am not sure how to interpret that... A result of Kirby and Siebenmann from the 1960's says that, except possibly in dimension 4 (unless Ron Stern's statement deals with it?), every smooth n-manifold M^n has the same number of non-diffeomorphic smooth structures as S^n does. So *every* smooth 7-manifold has 28 distinct smooth structures, up to diffeomorphism...

Aside from being interesting and suprising facts, proved by really bright people, these kinds of results can have useful consequences. Since all 2- and 3-manifolds M have unique smooth structures, when somebody hands us such a manifold M, we can *assume* they have also handed us a smooth structure (even if they didn't); it comes for free. Even more, we don't need to worry about which smooth structure we might have picked; if you and I happen to have picked different ones to work with, any result you might find with yours can be translated into a result about mine, because we know that there is a diffeomorphism between them (we just might not know what it *is*...). And if we are trying to set up some problem or do some computation, we can choose the most convenient coordinate system (i.e., atlas) that we want (tailored to the functions involved, perhaps), to carry out our work; we know, again, that we can translate our results into any other coordinate system, since they all describe the 'same' smooth structure. The fact that this isn't true in dimensions 4 and above (except, technically, that the Kirby-Siebenmann results implies, for example, that all smooth 12-manifolds have unique smooth structures?) makes life in higher dimensions much more interesting, in this regard!

Manifolds with boundary: Our definitions so far do not allow for things like the unit interval $I = [0, 1]$ to be a manifold, much less a smooth one. And, semantically at least, they never will be; they will be *manifolds with boundary.* A manifold with boundary is a Hausdorff, second countable space in which every point has a neighborhood homeomorphic to *either* $\mathbb{R}^n$ *or* the upper half space $\mathbb{H}^n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_n \geq 0\}$. Points which do not have a neighborhood homeomorphic to $\mathbb{R}^n$ are called boundary points; the union of them is the boundary of M, denoted ∂M. It is a consequence of Invariance of Domain that if $x \in M$ has neighborhood homeomorphic to $\mathbb{H}^n$ and the image of x has last coordinate 0, then $x \in \partial M$; that is, *every* chart is homeomorphic to $\mathbb{H}^n$, not $\mathbb{R}^n$, and always sends x to the boundary. Putting a smooth structure on a manifold with boundary involves some extra requirements as well, motivated, mostly, by what analysts have found to be reasonable in dealing with regions with boundary in Euclidean space. That is, a map $\mathbb{H}^n \to \mathbb{R}$ is $C^{(k)}$ if it can be *extended* to a $C^{(k)}$ map $\mathbb{R}^{n-1} \times (-\epsilon, \infty) \to \mathbb{R}$ on an open neighborhood of $\mathbb{H}^n$. Then we can adopt the exact same definition of an atlas and smooth structure, using this augmented definition of smooth to test the compatibility of the charts; that is, for any point x on the boundary, the maps $h \circ k^{-1}$ and $k \circ h^{-1}$ must extend to smooth maps on open neighborhoods of $k(x)$ and $h(x)$ respectively.

Smooth maps: So what do you do when you have a smooth structure? Start building smooth maps! We know how to identify a smooth map $f : M^n \to N^m$; we must have $h \circ f \circ k^{-1} : K(V) \to h(U)$ smooth for every pair of charts on M and N. Note that it is enough, though, to verify this for charts in a pair of atlases contained in the smooth structures for M and N; the compatibility of every other chart in our smooth structure with those of the atlases will guarantee smoothness of $h \circ f \circ k^{-1}$ over the entire maximal atlas. (Note also that this does *not* contradict what you've shown in one of your homework problems!) So, for example, to verify that some function $f : S^5 \to S^8$ (using the standard smooth structures!) is smooth, it suffices to use an atlas consisting to two charts on each (the stereographic projections from the poles), so smoothness can be verified by examining only 4 functions from $\mathbb{R}^5$ to $\mathbb{R}^8$. Actually verifying that such functions *are* smooth we are going to mostly leave to the same slightly fuzzy realm one encounters in calculus: if it is built up out of functions that we "know" are smooth, then it is smooth wherever it is defined.

One thing that can help us in things is to recognize that smoothness is local. This is just like in topology, where continuity is local; if $f : M \to N$ is a map such that for every $x \in X$ there is a chart (h, U) for M with $x \in U$ and a chart (k, V) for N with $f(x) \in V$, and $h \circ f \circ k^{-1}$ is smooth (where it is defined), then f is smooth. This is simply because the h's and the k's form atlases for M and N, respectively. But if you turn it around it can be thought of as a prescription for building a smooth function, by patching together smooth functions defined on open sets; if $\mathcal{O}$ is an open cover of M, and for each $U \in \mathcal{O}$ we have a smooth map $f_U : U \to N$ such that $f_U = f_V$ on $U \cap V$ for every $U, V \in \mathcal{O}$, then the map $f : M \to N$ defined by '$f(x) = f_U(x)$ if $x \in U$' is smooth. This is the direct analogue of the Gluing Lemma from topology. Of course, in topology, one more often wants to glue together maps defined on *closed* sets, rather than open sets; it is less messy. But in the smooth setting things aren't nearly so nice; on $\mathbb{R}$ the function $f(x) = |x|$ can be obtained by gluing together two smooth functions, but it is not smooth (using the standard smooth structures!) Question: are there *other* smooth structures on $\mathbb{R}$ for which f *is* smooth?

Basic properties: We also have many of the standard results. The composition of two smooth maps is smooth; this is essentially just because the corresponding result is true for maps between Euclidean spaces. The sum, difference, and product of two smooth maps $M \to \mathbb{R}$ are all smooth; again, this is basically because this is true for maps from $\mathbb{R}^n$ to $\mathbb{R}$. And the quotient is smooth so long as the denomenator is never zero. And a map into a Cartesian product of smooth manifolds (using the product smooth structure) is smooth iff the map into each factor is smooth (i.e., the composition with projection onto each factor is smooth). This last fact you have probably already had to use, since to decide on the smoothness of $h \circ f \circ k^{-1} : K(V) \to h(U) \subseteq \mathbb{R}^m$, you had to look at each of the m coordinate functions (projecting onto each coordinate factor $\mathbb{R}$). But some things *don't* work; for example the maximum $\max\{f, g\}$ of two smooth functions (mapping to $\mathbb{R}$) *need not* be smooth; $h(x) = |x|$, for example, can be defined as the maximum of the functions $f(x) = x$ and $g(x) = -x$.

Technically, partial derivatives are taken with respect to coordinate charts, not variables. but if $h : U \to \mathbb{R}^n$ is a chart, and $f : M \to \mathbb{R}$ is a map, then if we adopt the notation that $h(z) = (x^1(z), \dots, x^n(z)) \in \mathbb{R}^n$, we will adopt the notation that

$$\frac{\partial}{\partial x^i}(f)(z) = \frac{\partial}{\partial x_i}(f \circ h^{-1})(h(z))$$

That is, we formally are taking the partial derivative of f with respect to the coordinate functions of the chart h. Therefore, a function does not really have a 'value' of a partial derivative at a point; it has such a value *with respect to* a given coordinate chart. If we change charts around a point, to (k, V), $k = (y^1, \dots, y^n)$ the Chain Rule tells us how to relate the two sets of partial derivatives; it works out to the familiar

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

(once you chase it through the notation). One thing that this formula immediately tells us is that if $\partial f/\partial x^i = 0$ at z for every i, using one coordinate chart, then $\partial f/\partial y^i = 0$ at z for every i, for any other chart; a linear combination of 0's is still 0. So the notion of a *critical point*, as in mulrivariate calculus, is still well-defined in our more general setting. And the usual proof that local max's and min's are critical points carries over as well (apply any proof you've ever seen to $f \circ h^{-1}$ for any chart around your local extremum). So we can, for example, formulate and solve max-min problems on smooth manifolds!

Partitions of unity: There will be many situations in the material to come where we will want to assemble information obtained locally into a single smooth map $f : M \to N$. To do so, we will introduce the notion of a *partition of unity*; this is a way of writing the function $f(x) = 1$ as a sum of smooth functions $1 = \sum g_i$. Of course, $f(x) = 1$ works; constant functions are smooth. But we will want each summand function to be 'local'; that is, zero outside of a small open set (think: the domain of a chart). For example, in order to define the integral of a function $f : M^n \to \mathbb{R}$, the idea will be to define it first over the domain of a chart (h, U) (as the integral in $\mathbb{R}^n$ of the function $f \circ h^{-1}$, essentially), and then define the integral of f as the sum of the integrals of $f \cdot g_i$ (which are 'really' integrals over $\mathbb{R}^n$), since $f = f \cdot 1 \sum f \cdot g_i$. Building such a collection of functions will take a bit more work, and will be the first time we really invoke the Hausdorffness and second countability conditions built into our original definition. Put slightly differently, the idea is that we want to make the *support* of each function, $\text{supp}(g_i) = \text{cl}(\{x \in M : g_i(x) \neq 0\})$, to be small.

Given an open cover $\mathcal{O}$ of a space M, a *locally-finite refinement* of $\mathcal{O}$ is an open cover $\mathcal{P}$ of M so that for every $P \in \mathcal{P}$ there is an $O \in \mathcal{O}$ so that $P \subseteq O$ (that's the refinement part), and for every $x \in M$ there is an open neighborhood $x \in U$ of x so that $U \cap P = \emptyset$ for all but finitely many $P \in \mathcal{P}$ (that's the locally finite part). A space M is *paracompact* if every open cover $\mathcal{O}$ has a locally finite refinement $\mathcal{P}$ such that $\overline{P}$ is compact for every $P \in \mathcal{P}$. Such P are called *precompact*.

The main result we are aiming at is:

If M is a smooth manifold and $\mathcal{O} = \{u_\alpha : \alpha \in I\}$ is an open cover of M, then there is a partition of unity $\{g_i : i \in I\}$ so that $\text{supp}(g_i) \subseteq U_i$, and for every $x \in M$, there is a neighborhood V of x so that only finitely many $\text{supp}(g_i)$ intersect V.

The statement that $\text{supp}(g_i) \subseteq U_i$ is referred to as having a partition of unity *subordinate* to the open cover. The last property of the statement allows us to make sense of adding the functions together; we don't need the convergence of some infinite series, since around every point all but finitely many of the functions take the value zero. We say that the supports of the functions are *locally finite*, for short.

The proof of the existence of partitions of unity essentially comes in two parts. The first part asserts that any open cover $\mathcal{O}$ of M has a *locally finite refinement*, that is, a locally finite cover $\mathcal{O}'$ so that for every $O \in \mathcal{O}$ there is an $O' \in \mathcal{O}'$ with $O' \subseteq O$ (i.e., the refinement has "smaller" sets). This property is known as *paracompactness*. In particular, we will show that the refinement can be built out of the domains for coordinate charts of M. The second part uses the refinement by coordinate charts to build the partition of unity, by building a collection of smooth "bump" functions supported on each coordinate chart.

Paracompactness: To prove paracompactness, start with an open cover $\mathcal{O}$ of M, and a countable basis $\mathcal{B}$ for the topology on M. First we need a locally finite open cover to help guide our steps. Every point $x \in M$ is in the domain of some chart $h : U \to \mathbb{R}^n$ (with, we can arrange, image containing $B(h(x), 2)$). The open sets $U_x = h^{-1}(B(h(x), 1))$ cover M, and have compact closure; and for each there is a basis element B_x with $x \in B_x \subseteq U_x$. Since $\mathcal{B}$ is countable, there are countably many x_i so that the B_{x_i}, and therefore the U_{x_i}, cover M. Call these sets $U_i, I \in \mathbb{N}$, and let $C_i = \overline{U_i}$. By construction, C_i is compact, so for any finite set $I \in \mathbb{N}$, $C_I = \bigcup_I C_i$ is compact. Let E_I denote $\bigcup_I U_i$. Set $I_1 = \{1\}$, then since the U_i cover M and therefore C_1, there are finitely many i with union J_2 so that $C_{I_1} \subseteq \bigcup_{J_2} U_i$, and set $I_2 = J_2 \cup \{2\}$. Inductively, we continue to build finite sets J_n so that $C_{I_{n-1}} \subseteq \bigcup_{J_n} U_i$ and $I_n = J_n \cup \{n\}$. Then $M = \bigcup_n C_{I_n} = \bigcup_n E_{I_n}$, $C_{I_{n-1}} \subseteq E_{I_n}$ C_{I_n} is compact and E_{I_n} is open. Then the sets $K_n = E_{I_n} \setminus C_{I_{n-2}}$ are open, have union M, have compact closure (contained in C_{I_n}), and are locally finite. To demonstrate the last assertion, for any $x \in M$, $x \in E_{I_n}$ for some n; assume n is minimal. Then $x \in U_j$ for some

$j \in I_n$, and since $U_j \subseteq E_{I_n} \subseteq E_{I_k}$ for all $k \geq n$, $U_j \cap K_r = \emptyset$ for all $r \geq n+2$. In fact, since $\overline{K_r} \subseteq C_{I_n} \setminus E_{I_{n-1}}$, only K_{r-1}, K_r, and K_{r+1} meet $\overline{K_r}$.

Now start again. We have our open cover $\mathcal{O}$, and the locally finite cover $\{K_n\}$ by precompact open sets. For every point $x \in M$ we can, by local finiteness, find an open neighborhood W_x so that if $x \in K_n$ then $W_x \subseteq K_n$; start with a neighborhood meeting only fnitely many of them, and then intersect it with each of them as well. Taking a further intersection with an element of $\mathcal{O}$ containing x, we can also assume that $W_x \subseteq U \in \mathcal{O}$. Then we may assume by intersecting with the domain of a chart that there is a chart $h : W_x \to \mathbb{R}^n$ sending W_x to an open neghborhood of $h(x)$. Rescaling h on the codomain side and shrinking the domain, we can asuume that $h(W_x) = B(h(x), 2)$, and so $V_x = h^{-1}(B(h(x), 1))$ is a neighborhood of x with compact closure, contained in an element of $\mathcal{O}$, and contained in every K_n that it meets. W_x satisfies all of these properties except possibly the compact closure.

Now for each n, the sets V_x with $x \in \overline{K_n}$ form an open cover of the compact set $\overline{K_n}$, so they have a finite subcover $\mathcal{P}_n = \{V_{x_1,k]n}, \dots, V_{x_{m_n},kn}\}$; we assume that each has non-empty intersection with $\overline{K_n}$ (otherwise we throw it away). Set $\mathcal{R}_n = \{W_{x_1,n}, \dots, Wx_{m_n}, n\}$. The collectionS $\mathcal{P} = \bigcup_n \mathcal{P}_n$ and $\mathcal{R} = \bigcup_n \mathcal{R}_n$ both form open covers of M, are refinements of $\mathcal{O}$, and, we now show, are locally finite. It is enough to show this for $\mathcal{R}$, since these sets are larger. We show that, in fact, each set in $\mathcal{R}$ meets only finitely many others, so each demonstrates local finiteness for every point in it. But each $W = W_{x_i,n}$ intersects, and is therefore contained in, some K_m. It therefore meets only $\overline{K_{m-1}}, \overline{K_m}$, or $\overline{K_{m+1}}$. Any other element W' of $\mathcal{R}$ meeting W meets, and therefore is contained in, one of these three sets. So the only $\overline{K_r}$ it could meet would be one of $\overline{K_{m-2}}$ through $\overline{K_{m+2}}$. W' is therefore a member of one of $\mathcal{R}_{m-2}$ through $\mathcal{R}_{m+2}$; it doesn't meet any of the other sets $\overline{K_r}$. Therefore, it is one of the finitely many elements of these five sets. So W meets only finitely many of the other elements of $\mathcal{R}$.

The partitioning of 1: Now that we know how to build a locally finite cover by (images of) charts $(h_i, h_i^{-1}(B(x_i, 2))$ for which $h_i^{-1}(B(x_i, 1))$ also cover and $h_i^{-1}(\overline{B(x_i, 1)})$ is compact, we turn to bulding a partition of unity with supported on these sets. We start with the fact that the function

$$f(x) = e^{-1/x} \text{ if } x > 0 \text{ ; } = 0 \text{ if } x \leq 0$$

is C^∞. This follows from the fact that the n-th derivative of $e^{-1/x}$ is $f_n(x) = p_n(x)e^{-1/x}/x^{2n}$ for some polynomial $p_n(x)$, which can be established by induction on n. The function has (one-sided) limit 0 at $x = 0$, which can be established by repeated use of L'Hôpital's Rule (to show that $e^{-1/x}/x^{2n}$ has limit 0). Together these imply that f has continuous derivatives of'all orders. Note that since $-1/x < 0$ for $x > 0$, $0 \leq f(x) < 1$ for all x. Now define $g(x) = f(2-x)/(f(2-x) + f(x-1))$; this function is smooth, since the denomentaor is always positive (one term is 0 only for $x \geq 2$ and the other is zero only for $x \leq 1$), takes values between 0 and 1, is one precisely when $f(x-1) = 0$, i.e., $x \leq 1$, and is 0 precisely when $f(2-x) = 0$, i.e., $x \geq 2$. Then the function $G : \mathbb{R}^n \to \mathbb{R}$ defined by $G(y) = g(||y - x_0||^2)$ is smooth (it's the composition of smooth functions), is 1 on $B(x_0, 1)$ and has support contained in $B(x_0, 2)$. Taking our charts h_i built above, the function $h_i \circ G$ extends (by taking the value 0) to a smooth function $f_i : M \to \mathbb{R}$ which is 1 on $h_i^{-1}(B(x_i\, 1))$ and has support in $h_i^{-1}(B(x_i\, 2))$. Since every point has a neighborhood which lies in only finitely many of the $h_i^{-1}(B(x_i\,2))$, the sum $F = \sum f_i$ is locally a finite sum and so is a smooth function on M. Since the $h_i^{-1}(B(x_i\,1))$ cover M, it is everywhere non-zero. So each of the functions $F_i = f_i/F$ is smooth, their supports = the supports of the f_i are locally finite, and their sum (which is locally a finite sum) is 1. that is, they form a smooth partition of unity subordinate to the cover $h_i^{-1}(B(x_i\,2))$, which is a refinement of our original cover $\mathcal{O}$. So they form a smooth partition of unity subordinqate to $\mathcal{O}$.

Density of smooth functions: Now that we have a partition of unity, what do we do with it? One immediate application of partitions of unity is: for every continuous function $f : M \to \mathbb{R}$ and $\epsilon > 0$, there is a smooth function $g : M \to \mathbb{R}$ with $|f(x)-g(x)| < \epsilon$ for all $x \in M$. The proof consists of looking at the open cover $\{f^{-1}(f(x)-\epsilon, f(x)+\epsilon)\}$, and choose a partition of unity g_i subordinate to this cover. For each g_i pick a point x_i with $\text{supp}(g_i) \subseteq f^{-1}(f(x_i)-\epsilon, f(x_i)+\epsilon)$. Then the function $g(y) = \sum f(x_i)g_i(y)$ is smooth (since the $f(x_i)$ are constants, so this is a locally finite sum of smooth functions), and $|f(y)-g(y)| = |\sum g_i(y)(f(y)-f(x_i))| \leq \sum g_i(y)|f(y)-f(x_i)| < \sum g_i(y)\epsilon = \epsilon$, since either $g_i(y) = 0$, or $g_i(y) > 0$, so $y \in f^{-1}(f(x_i)-\epsilon, f(x_i)+\epsilon)$, so $f(y) \in (f(x_i)-\epsilon, f(x_i)+\epsilon)$, so $|f(y)-f(x_i)| < \epsilon$.

Partitions of unity can also be used to build *bump functions*; given a closed set C of M^n and an open set U with $C \subseteq U$, we can build a smooth function $f : M \to \mathbb{R}$ which is 1 on C and has support contained in U. The idea is simply to take the open cover $\{U, M \setminus C\}$ and build a smooth partition of unity ψ_i, ϕ_j subordinate to it. with $\text{supp}(\psi_i) \subseteq U$ and $\text{supp}(\phi_j) \subseteq M \setminus C$ for every i and j. Then set $\psi = \sum_i \psi_i$ and $\phi = \sum_j \phi_j$; by local finiteness, both are smooth functions. Since $\psi(x)+\phi(x) = 1$ for all x and $\phi(x) = 0$ outside of $M \setminus C$ (since all summands are), i.e., for $x \in C$, we have $\psi(x) = 1$ for $x \in C$; since $\psi(x) = 0$ outside of U, we have $\psi(x) = 0$ for $x \notin U$, as desired. (This last statement does not quite say that $\text{supp}(\psi) \subseteq U$; but this can be remedied by using a slightly smaller open set V in place of U, with $C \subseteq V \subseteq \overline{V} \subseteq U$, which exists by the normality of M.)

Embedding in $\mathbb{R}^n$: Another immediate application (of our proof, really) is that if M^n is a smooth manifold, then there is a smooth embedding (that is, a topological embedding that is a smooth map) of M into $\mathbb{R}^n$ for some N. Right now we will prove this for compact M; later we will show it for all M. To build the embedding, cover M by finitely many coordinate charts (h_i, U_i), $i = 1, \ldots, k$, so that $B(x_i, 2) \subseteq h_i(U_i)$ and the $h_i^{-1}(B(x_i, 1))$ cover M. Then taking a smooth bump function g_i that is 1 on $h_i^{-1}(B(x_i, 1))$ and supported on U_i, we can build the smooth functions $f_i = g_i \cdot h_i : M \to \mathbb{R}^n$; Then the smooth function $F : M \to \mathbb{R}^{nk} = (\mathbb{R}^n)^k$ given by $F(x) = (f_1(x), \ldots, f_k(x))$ is 1-to-1; mapping from a compact space to a Hausdorff one, it is a homeomorphism onto its image. In a sense which we will eventually make precise, the smooth structure on M is induced from the map F and the smooth structure on $\mathbb{R}^{nk}$, making this a smooth embedding.

Tangent vectors: In multivariable calculus, a prominent place is taken up by vectors, underlying many constructions and techniques. Tangent vectors, directional derivatives, gradients, and vector fields appear throughout the subject. Our next task is to introduce this technology into smooth manifolds. It turns out there are about as many ways to approach the concept of tangent vector as there were early researchers in the field. But in a way which we will make fairly precise, all are really the same. We will introduce (at least) two of them, since they both have their own advantages in different situations.

In $\mathbb{R}^n$, the notion of a direction is expressed by a vector v based at a point. This leads to the notion of the directional derivative $D_v f$; the rate of change of f in the direction of v. One way to approach (tangent) vectors for manifolds is to borrow directional derivatives, making a defnition out of the properties which they have in multivariable calculus. This will be one point of view we will take.

Velocity vectors: Borrowing vectors directly will work (with a little effort); but we can reformulate them more directly, in terms of things that we can borrow more directly, namely smooth functions. Specifically, in $\mathbb{R}^n$ a vector v at x describes a direction by way of the curve $\gamma(t) = x + tv$; v is the derivative of γ at $t = 0$. We can translate this picture to a smooth manifold using charts; given a chart (h, U) around x, $\eta = h^{-1} \circ \gamma$, defined on a small interval around 0, is a smooth curve $\eta : (-\epsilon, \epsilon) \to M$. It's derivative at 0, using the coordinate chart h, is v. But of course this result is dependent upon the chart chosen, both to define it and to evaluate it. But the idea of a smooth curve *isn't*. So instead we make our definition based on them. A tangent vector at a point x will "be' the derivative, at $t = 0$, of a smooth curve with value x at $t = 0$. But different curves can have the same derivative, so we need to introduce an equivalence relation to make a formal definition.

A tangent vector at $x \in M$ is an equivalence class of smooth curves $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = x$. Two such curves γ, η are equivalent if for some chart (h, U) about x we have $(h \circ \gamma)'(0) = (h \circ \eta)'(0)$.

Informally, we tend to think of this as saying that two curves are equivalent if they have the same velocity vector at $t = 0$! Note that the equivalence is independent of coordinate chart chosen; if (k, V) is another chart (for convenience, let us suppose that $h(x) = k(x) = 0$), then $(k \circ \gamma)'(0) = [D(k \circ h^{-1})(h(x))](h \circ \eta)'(0)$ (where this is matrix multiplication), so $(h \circ \gamma)'(0) = (h \circ \eta)'(0)$ implies $(k \circ \gamma)'(0) = (k \circ \eta)'(0)$.

But now we see how we *ought to* to relate tangent vectors from the point of view of different charts; we use the total derivative map $D(k \circ h^{-1})(h(x))$. And we can use this to get rid of the smooth curves! Writing $k \circ \gamma)'(0) = w$ and $h \circ \eta)'(0) = v$, what is important is that $w = D(k \circ h^{-1})(h(x))v$, which needs no mention of curves at all. So we can define a tangent vector at a point $x \in M$ as an equivalence class of triples (x, h, v), where h is a chart whose domain contains x, and v is a vector based at $h(x) \in \mathbb{R}^n$. Another tangent vector (y,k,w) is equivalent if $y = x$ and $w = k \circ \gamma)'(0) = w$. We will let $[h, v]_x$ denote the equivalence class. This construction illustrates a basic theme that runs throughout the development of differential topology: To introduce an object from calculus, all we need to do, really, is figure out how the object would transform when we change our point of view by using a different chart around at point, and incorporate that into the definition, in the form of an equivalence relation. The point, really, is that so long as we are working locally, we can essentially pretend that it *is* the familiar object from calculus; it is only when we start looking at how the object behaves as we wander around the manifold that we need to remember how they transform as we need to keep changing coordinate charts, as our point of view keeps shifting.

The set of tangent vectors $[h, v]_x$ at a point form a vector space, the *tangent space*, TM_x or T_xM, at the point x. The union $\bigcup T_xM = TM$ is the tangent space of M. We could keep talking about this, exploring it various properties from this point of view, but let us back up and start again using the directional derivative point of view.

Derivations: Given a vector v based at $z \in \mathbb{R}^n$, it allows us to define the directional derivative $(D_v f)(x)$ of any differentiable function whose domain contains a neighborhood about z. That is, we have an operator D_v from smooth functions to $\mathbb{R}$. This operator is linear, and satisfies a Leibnitz rule: $D_v(fg) = gD_v f + fD_v g$ (this last is because D_v is 'really' the gradient dotted with v, so it is a linear combination of the partial derivatives, and the partial derivatives satisfy the Leibnitz rule). Such an operator is called a *derivation*. But such a concept makes sense anywhere that the notion of 'differentiable function' makes sense, e.g., on a smooth manifold. Since D_v takes too long to write, and D_v is really *replacing* the notion of the vector v, we will write $D_v = X$ in general.

For $a \in M^n$ a smooth manifold, a *derivation at* a is a map $X : C^\infty(M) \to \mathbb{R}$ satisfying $X(\alpha f + \beta g) = \alpha X(f) + \beta X(g)$ and $X(fg) = f(a)X(g) + g(a)X(f)$, for $\alpha, \beta \in \mathbb{R}$ and $f, g \in C^\infty(M)$.

The set of all derivations at a will be denoted $T_a M$. These will be our 'tangent vectors' at a. On the face of it, this definition is hopelessly abstract, but we can get a better handle on it by outlining some of its basic properties.

Lemma: If $c \in C^\infty(M)$ is constant, $c(x) = c$, then $X(c) = 0$. If $f(a) = g(a) = 0$, then $X(fg) = 0$. If $f = g$ on a neighborhood of a, then $X(f) = X(g)$.

Proof: First, $X(1) = x(1 \cdot 1) = 1 \cdot X(1) + 1 \cdots X(1) = 2 \cdot X(1)$, so $X(1) = (2-1)X(1) = 0$. Then $X(c) = cX(1) = c \cdot 0 = 0$. For the second, $X(fg) = f(a)X(g) + g(a)X(f) = 0 + 0 = 0$. And finally, choose a small chart (h, U) around a, with domain contained in the neighborhood on which f and g agree, and build a bump function B supported on U and equal to 1 on the closure of a small ball $h^{-1}(B(b, \epsilon))$. Then $fB = gB$ on all of M; inside of U they agree by hypothesis, and outside of U they are both 0. So $X(fB) = X(gB)$; but then $0 = X(fB) - X(gB) = (f(a)X(B) + B(a)X(f)) - (g(a)X(B) + B(a)X(g)) = g(a)X(B) + X(f) - g(a)X(B) - X(g) = X(f) - X(g)$, so $X(f) = X(g)$.

So the derivation X is really 'local'; it depends only on the values of f near a. Which would indicate that we ought to be able to understand them better using charts! Which we will do. But first, a little more theory. A chart can be thought of as a C^∞ map from a neighborhood in M to the standard smooth structure on $\mathbb{R}^n$. So understanding how derivations behave under smooth maps will help us understand derivations.

Pushforwards: Given a smooth map $F : M^n \to N^m$ and $a \in M$, we can 'push forward' a derivation X at a to a derivation at $F(a)$, which we will call $F_*(X)$; we define $F_*(X)(f) = X(f \circ F)$. It is a straighforward calculation, using the fact that $(f \cdot g) \circ F = (f \circ F) \cdot (g \circ F)$ that $F_*(X)$ is a derivation at $F(a)$ [don't forget linearity!]. The following facts are also pretty straighforward:

Lemma: $(F \circ G)_* = F_* \circ G_*$. $\mathrm{Id}_* = \mathrm{Id}$. If F is a diffeomorphism, then F_* is an isomorphism for every $a \in M$.

With these, we can go explore derivations using charts, and get a better understanding of them. First, because the definition of a derivation is really local, if $U \subseteq M$ is open and $a \in U$, then the inclusion map $i : U \to M$ induces an isomorphism $i_* : T_a U \to T_a M$

; choosing a bump function g which is 1 on neighborhood of a and supported on U, we can extend any smooth function f on U to a function gf on M which equals f on a neighborhood of a. Then comparing a derivation X on U and the derivation i_*X on M, we have $i_*X(gf) = X(gf \circ i) = X(f)$, since $gf \circ i$ and f agree on a neighborhood of a. So if $i_*X_1 = i_*X_2$, then $X_1 f = X_2 f$ for every $f : U \to \mathbb{R}$, so $X_1 = X_2$ and i_* is injective. To show surjectivity, given $X \in T_aM$, define $Y \in T_aU$ by $Y(f) = X(gf)$. A computation shows that Y is a derivation at a. Then for any $f \in C^\infty(M)$, $i_*(Y)(f) = Y(f \circ i) = X(g(f \circ i)) = X(f)$, again, since $g(f \circ i)$ and f gree near a. So $i_*Y = X$ and i_* is onto. SO i_* is an isomorphism.

But now a chart (h, U) is a diffeomorphism $U \to \mathbb{R}^n$ for the standard smooth structure on $\mathbb{R}^n$, so T_aU is isomorphic to $T_{h(a)}\mathbb{R}^n$. So to understand the tangent space at a point, we may assume that $M = \mathbb{R}^n$! But we can build a collection of derivations in $\mathbb{R}^n$; the directional derivatives $D_v = \sum_i v^i(\partial/\partial x^i)$ for $v = (v^1, \dots, v^n)$. The last piece of the puzzle is:

Lemma: The map $I : v \mapsto D_v$ is an isomorphism.

To prove it, look at the coordinate functions $f_j : x = (x^1, \dots, x^n) \mapsto x^j$ and note that $D_v(x^j) = \sum_i v^i(\partial x^j/\partial x^i) = v^j$, so $D_v = D_w$ implies $v = w$, and I is injective. For surjectivity, given a derivation X on $\mathbb{R}^n$ at a, let $v^i = X(f_i)$, and $v = (v^1, \dots, v^n)$. We show $X = D_v$; given $f \in C^\infty(\mathbb{R}^n)$, we can expand f as a power series centered at $a = (a^1, \dots, a^n)$.

$$f(x) = f(a) + \sum_i \frac{\partial f}{\partial x^i}(a)(x^i - a^i) + \sum_i g_i(x)(x^i - a^i)$$

where $g_i(a) = 0$, by a result of advanced calculus. Since $f(a)$ is constant and $g_i(x)$ and $x^i - a^i$ are both 0 at a, $X(f) = \sum_i f_{x^i}(a)v^i = D_v(f)$, so $X = D_v$. Finally, from calculus we know that I is linear. So I is an isomorphism.

The derivations $X_i = \partial/\partial x^i$ form a basis for $T_a\mathbb{R}^n$. Carrying these back to M via a chart $h : U \to \mathbb{R}^n$, or rather the map $h_* : T_aU \to T_{h(a)}\mathbb{R}^n$. But to get back to M, we use $(h^{-1})_*$; this map carries the basis X_i to $(h^{-1})_*X_i$, where $((h^{-1})_*X_i)f = X_i(h^{-1} \circ f) = \partial(h^{-1} \circ f)/\partial x^i$. Which should sound familiar! This is what we denoted $\partial f/\partial x^i$, where $h(y) = (x^1(y), \dots x^n(y))$. So the derivatives with respect to the coordinate functions of our chart h, $\partial/\partial x^i$, form a basis for T_aM.

The three approaches to the concept of a tangent vector can all be brought together by decribing the basis vectors for T_aM from each point of view. For derivations, they are the differentiation operators $\partial/\partial x^i$ for the coordinate functions. For curves, they are the derivatives at a of the functions $t \mapsto h^{-1}(h(a) + t(0, \dots, 1, \dots, 0))$. For vectors, they are the equivalence classes $[h, (0, \dots, 1, \dots, 0)]_a$. Each of these descriptions are local, usng the chart h.

In particular, given a smooth curve $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = a$, the map $f \mapsto (f \circ \gamma)'(0)$ is a derivation at a; this follows from the fact that $(fg) \circ \gamma = (f \circ \gamma)(g \circ \gamma)$. Chasing this through both descriptions will verify that this assignment is an isomorphism between our equivalence classes of curves and the space of derivations at a.

Computations in local coordinates: Tangent vectors as derivations are defined globally, but are typically worked with locally. Given $a \in M$ and a chart (h, U) around a, a derivation X can be written $X = \sum_i v^i \partial/\partial x^i$ (evaluated at $h(a)$), where $h = (x^1, \dots, x^n)$. Given a map $F : M^n \to N^m$ we can express the pushforward in local coordinates: let (k, V) be a chart around $b = F(a)$, with $k = (y^1, \dots, y^m)$. We know we can write $F_*X = \sum_j w^j \partial/\partial y^j$ for some w^j; the task is to compute w^j. But $w^k = (\sum_j w^j \partial/\partial y^j)(y^k)$, since $\partial y^k/\partial y^j = \delta_{kj}$. So in order to compute w^j we need to compute $w^j = F_*X(y^j) = X(y^j \circ F) = \sum_i v^i \partial(y^j \circ F)/\partial x^i$. The constants $\partial(y^j \circ F)/\partial x^i$ form the matrix a partial derivative of F, in local coordinates, and so $\sum_i v^i \partial/delx^i$ is carried to $\sum_j [\sum_i v^i \partial(y^k \circ F)/\partial x^i] \partial/dely^j$.

In particular, if we set F =Id = the identity function, we can recover a change of variables formula for tangent vectors: given two charts $h = (x^1, \dots, x^n)$ and $k = (y^1, \dots, y^n)$ about $a \in M$, we have $\partial/\partial x^i = \sum_j (\partial y^k/\partial x^i) \partial/dely^j$, which is, of course, the exact same change of variables formula we had for tangent vectors as smooth curves and as vectors. This formula extends to T_aM by linearity. This formula allows us to translate computations when we switch perspectives by using a different chart.

As an example, let us examine the tangent vectors to S^2 using a variety of standard charts on S^2. We have the standard projection coordinates $h_1 : (x^1, x^2, x^3) \mapsto (x^1, x^2)$, etc., with inverse $(x^1, x^2) \mapsto (x^1, x^2, \sqrt{(1 - (x^1)^2 - (x^2)^2})$. There are the stereographic coordinates $k_1 : (y^1, y^2, y^3) \mapsto (y^1, y^2)/(1 - y^3)$, etc.. , with inverse $(y^1, y^2) \mapsto (2y^1, 2y^2, |y|^2 - 1)/(|y|^2 + 1)$. We also have spherical coordinates, which we are all probably more familiar writing the inverse for: $\ell_1^{-1} : (\theta, \varphi) \mapsto (\cos\theta \sin\varphi, \sin\theta \sin\varphi, \cos\varphi)$. The inverse of the inverse is $\ell_1 : (z^1, z^2, z^3) \mapsto (\arctan(z^2/z^1), \arccos(z^3))$. So, for example, on the upper hemisphere, we can compute the change of coordinates formula from sphereical to projection coordinates as the total derivative of the map $(\theta, \varphi) \mapsto (\cos\theta \sin\varphi, \sin\theta \sin\varphi)$, which is the matrix $(-sin\theta \sin\varphi, \cos\theta \sin\varphi; \cos\theta \cos\varphi, \sin\theta \cos\varphi)$.

The tangent space: In any of its manifestations, we can assemble the tangent spaces T_aM at points a into a single tangent space $TM = \bigcup_a T_aM$. But the change of variable formula above can be turned into a prescription for putting a topology, and a smooth structure, on M. TM is locally homeomorphic to $\mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$, by charts $H : TU \to \mathbb{R}^{2n}$, induced by a chart (h, U), $h = (x^1, \dots, x^n)$ for M, where $H(X_a) = (h(a), (X_a(x^1), \dots, X_a(x^n))$. Given a second chart (k, V) , $k = (y^1, \dots, y^n)$ for M, the transition function for TM, using the change of variables formula, is given by $H \circ K^{-1}(b, (v^1, \dots, v^n)) = (h \circ k^{-1}(b), (\sum_i v^i(\partial x^1/\partial y^i), \dots, \sum_i v^i(\partial x^n/\partial y^i)))$. This function is smooth, since in the first n coordinates it is the transition function for M, and in the second n coordinates it is essentially built out of the partial derivatives of the coordinate functions, which are also smooth. The topology on TM is generated from these charts; a basis consists of images under the H^{-1} of a basis of open sets on $\mathbb{R}^{2n}$. Since a countable number of charts cover M and each of the charts provide a countable collection of sets to add to the basis, we have second countability. Hausdorffness proceeds similarly (two cases: in the same chart or never in the same chart). The natural map $p : TM \to M$ sending a derivation at a to a is smooth. This is our first example of a *vector bundle.* The idea behind a k-bundle E over a space M is that we have a map $p : E \to M$ so that for every $a \in M$ there is a neighborhood U of a for which $p^{-1}(U) \cong U \times \mathbb{R}^k$ via a map $h_U : p^{-1}(U) \to U \times \mathbb{R}^k$, so that the diagram

$$\begin{array}{ccc} p^{-1}(U) & \xrightarrow{h_U} & U \times \mathbb{R}^k \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle \mathrm{pr}_1} \\ U & \xrightarrow{\mathrm{Id}} & U \end{array}$$

commutes. This is precisely what we have for the tangent bundle. The basic idea is that each point of M has an $\mathbb{R}^k$ attached to it, so that this assignment is 'locally trivial' in the sense of the commutative diagram. The tangent space of a smooth manifold is just one such family of examples of vector bundles. The *trivial bundle* $M \times \mathbb{R}^k$, with projection map pr_1, is another example. The (open) Möbius band is the total space of a (non-trivial) 1-bundle over the circle. A smooth manifold M^n is said to be *parallelizable* if its tangent bundle is fiber-preserving diffeomorphic to the trivial n-plane bundle.

Vector fields: The tangent bundle gives us the framework to introduce another standard concept from advanced calculus. A vector field is a choice of vector at every point in a space. From the point of view of TM, a (tangent) vector field is a choice of element of T_aM for every $a \in M$. Put differently, a vector field X is a map $X : M \to TM$ so that $X(a) \in T_aM$ for every $a \in M$; i.e., $p \circ X : M \to TM \to M$ is the identity map. Generally, for a vector bundle $p : E \to M$, a map $s : M \to E$ with $p \circ s = \mathrm{Id}_M$ is called a *section* of the bundle. So a vector field on M is a section of the tangent bundle. The vector field is *smooth* if the section is a smooth map. The set of all smooth vector fields on M is denoted $\mathcal{T}(M)$.

Writing things in local coordinates (h, U), a vector field can be expressed as $X = \sum_i v^i \partial/\partial x^i$, where the v^i are functions from U to $\mathbb{R}$. X is smooth $\Leftrightarrow$ the functions v^i are smooth; this follows from the construction of the smooth structure on TM. Given a smooth vector field X and a smooth function $f : M \to \mathbb{R}$ the assignment $a \mapsto X_a f$ is a function. Writing things in local coordinates, $Xf = \sum v^i \partial f/\partial x^i$ is a smooth function. This point of view actually provides another characterization of smoothness: X is smooth $\Leftrightarrow$ Xf is smooth for every smooth map $f : M \to \mathbb{R}$. This is because in local coordinates the v^i can be recovered as Xx^i (or rather, the coordinate function x^i multiplied by a bump function for the chart), and x^i is smooth, so $Xx^i = v^i$ is smooth and X is smooth.

The fact tht Xf is a smooth function for smooth f and smooth vector field X means that we can use Xf as the function to feed another vector field Y, allowing us to define YX as $(YX)(f) = Y(Xf)$. But YX is not a vector field; it fails to be a derivtion at a point. We can compute

$YX(fg) = Y(f(Xg) + g(Xf)) = Y(f(Xg)) + Y(g(Xf)) = f(YX)g + (Yf)(Xg) + g(YX)f + (Yg)(Xf) = [f(YX)g + g(YX)f] + (Yf)(Xg) + (Yg)(Xf)$

and we have no reason to believe that the last two terms will cancel one another. But! Those last two terms are symmetric in X and Y, and $YX(fg)$ isn't. So if we compute $XY(fg)$, we will get the same two extra terms. So if we subtract these two expressions, they will cancel. That is, $(XY - YX)(fg) = f(XY - YX)g + g(XY - YX)f$, so $XY - YX$ is a derivation (a quick check shows that it is linear), and therefore defines a vector field, called the *Lie bracket* $[X, Y] = XY - YX$ of X and Y. A direct computation in local coordinates reveals that $[\sum_i v^i \partial/\partial x^i, \sum_i w^i \partial/\partial x^i] = \sum_i (\sum_j v^j (\partial w^i/\partial x^j) - w^j (\partial v^i/\partial x^j))\partial/\partial x^i$. So the Lie bracket of two smooth vector fields is a smooth vector field.

The Lie bracket satisfies several useful properties:

(a) it is $\mathbb{R}$-linear in each entry: $[aX + bY, Z] = a[X, Z] + b[Y, Z]$ for $a, b \in \mathbb{R}$, etc.

(b) it is antisymmetric: $[X, Y] = -[Y, X]$

(c) it satisfies the *Jacobi identity*: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ (the 0 vector field).

Each can be verified by evaluating each side on a smooth function f, checking that we get the same answer.

Vector bundles: The tangent bundle will not be the only 'bundle' over a smooth manifold M that we will find ourselves interested in. There will be many other linear spaces that we will wish to assign to each point $a \in M$, and assemble into a bundle.

Some random, interesting but useless(?), facts:

Every n-manifold can be covered by at most $n + 1$ charts. The minimum number of contractible open sets which cover M is called its *Lusternik-Schnirelmann category*, $LS(M)$. For example, $LS(S^n) = 2$ for every n. In also happens to be the minimal number of critical points that a function $f : M \to \mathbb{R}$ (whose critical points are discrete) can have.

$\mathbb{R}^4$ has uncountably many non-differomorphic smooth structures; but since $\mathbb{R}^5$ has only one, crossing exotic $\mathbb{R}^4$'s with $\mathbb{R}$ always gives *standard* $\mathbb{R}^5$.

Every 3-manifold M^3 is parallelizable; TM^3 is fiber-preserving diffeomorphic to $M^3 \times \mathbb{R}^3$.