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 $k = 1, \dots, 28$, 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 $\mathbb{H}^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_n \geq 0\}$. 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. The boundary ∂M of a smooth manifold M^n is a smooth manifold of dimension $n - 1$; the restrictions of the charts to $\partial\mathbb{H}^n = \mathbb{R}^{n-1} \times \{0\}$ form a smooth atlas, since the original charts are smooth; we just ignore any partial derivatives in the n-th direction.

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, $\mathrm{supp}(g_i) = \mathrm{cl}(\{x \in M : g_i(x) \neq 0\})$, 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 $\mathrm{supp}(g_i) \subseteq U_i$, and for every $x \in M$, there is a neighborhood V of x so that only finitely many $\mathrm{supp}(g_i)$ intersect V.

The statement that $\mathrm{supp}(g_i) \subseteq U_i$ 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 $(D_vf)(x)$ 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: $D_v(fg) = gD_vf + fD_vg$ (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_aM. 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 $i_* : T_aU \to T_aM$; 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 $X_1f = X_2f$ 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 $h = (x^1, \ldots, x^n)$. 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 $k = (y^1, \ldots, y^m)$. 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 $\sum_j [\sum_i v^i \partial(y^j \circ F)/\partial x^i]\partial/dely^j$.

In particular, if we set $F =$Id $=$ the identity function, we can recover a change of variables formula for tangent vectors: given two charts $h = (x^1, \ldots, x^n)$ and $k = (y^1, \ldots, y^n)$ 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), $h = (x^1, \ldots, x^n)$ for M, where $H(X_a) = (h(a), (X_a(x^1), \ldots, X_a(x^n))$. Given a second chart (k, V) , $k = (y^1, \ldots, y^n)$ for M, the transition function for TM, using the change of variables formula, is given by $H \circ K^{-1}(b, (v^1, \ldots, v^n)) = (h \circ k^{-1}(b), (\sum_i v^i(\partial x^1/\partial y^i), \ldots, \sum_i v^i(\partial x^n/\partial y^i)))$. 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.

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 that 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. A natural question to ask is "what does the Lie bracket really measure?". On the face of it, it measures the extent to which two vector feilds fail to commute. In fact, the fact that our coordinate vetor fields $\partial/\partial x^i$ do commute ("mixed partials are equal") has a converse; If $X_1, \ldots X_k$ are vector fields defined on some open set V so that the $X_i(a) \in T_aU$ are linearly independent at every a, and $[X_i, X_j] = 0$ on V for every i, j, then around every point there is a coordinate system (h, U) so that in these local coordinates $X_i = \partial/\partial x_i$ for $i = 1, \ldots, k$. [We may prove this at some future point.]

But what is the significance of the direction that $[X, Y]$ points, when it is non-zero?

To answer this, we need to introduce integral curves. And to do that we need to discuss pushforwards of vector fields.

Pushforwards and $\mathcal{F}$-relatedness: We have seen that a smooth map $f : M \to N$ induces a map $f_* : TM \to TN$ which is linear on each T_aM. So if we wish to push a vector field X on M to a vector field on N, the natural thing to do would be so set $(f_*X)(b) = f_*(X_a)$ where $f(a) = b$. There are, of course, two problems with this; f might not be onto, so there is no a in the preimage of b, and f might not be 1-to-1, so there might be more than one a in the preimage, and the various values of $f_*(X_a) \in T_bN$ might not agree. The concept of $\mathcal{F}$-relatedness essentially declares that the second problem does not occur: two vector fields X on M and Y on N are $\mathcal{F}$-related if $f_*(X_a) = Y_{f(a)}$ for all $a \in M$. Note that since a vector field is determined by its values on smooth (e.g., coordinate) functions, whether or not two vector fields are $\mathcal{F}$-related can also be determined by testing values on functions: X and Y are $\mathcal{F}$-related $\Leftrightarrow$ $X(g \circ f) = Y(g) \circ f$ for every smooth map $g : N \to \mathbb{R}$. Given a vector field X on M, there is of course no reason to expect there to be a vector field on N $\mathcal{F}$-related to X. Nor given Y on N should there be a vector field on M $\mathcal{F}$-related to Y. (Think: the constant function f.) But in one special case, which we use again and again, we can expect such good things:

If $f : M \to N$ is a diffeomorphism, then for every $X \in \mathcal{T}(M)$ there is a unique $Y \in \mathcal{T}(N)$ $\mathcal{F}$-related to X. This is mostly a matter of figuring out what Y has to be; we want $f_*(X_a) = Y_{f(a)}$, so define $Y_b = f_*(X_{f^{-1}(b)})$, which is well-defined since f has an inverse. Uniqueness of Y is immediate; plug $b = f(a)$ (so $a = f^{-1}(b)$) into the definition of $\mathcal{F}$-relatedness. All that is left is to show that Y is a smooth vector field. This can be done using the local coordinates $h \circ F^{-1}$ for some coordinate chart h on M about a; then Y has the exact same local formula as X does.

The point to this really, is that we want to be able to think of (the restriction of) a vector field on M as "really" being a vector field on $\mathbb{R}^n$, using a local coordinate chart. Since a chart $h : U \to \mathbb{R}^n$ is a diffeomorphism onto its image, we can use pushforwards to work with our vector fields as if they were in $\mathbb{R}^n$. This is really just a fancy way of saying that we are working in local coordinates. so you should mostly treat this as a warning that in general there is no such thing as the pushforward of a vector field?

Integral curves: Given a smooth curve $\gamma : I \to M$, the derivative $\gamma'(t)$ can be defined as $\gamma_*(d/dt) \in T_{\gamma(t)}M$; this is how our notion of velocity vector for a curve translates into the language of derivations. γ is an *integral curve* for a tangent vector field X on M if $\gamma'(t) = X(\gamma(t))$ for all t in the domain of γ. It is a consequence of the fundamental existence and uniqueness theorem for solutions to ODEs that for every smooth vector field X and $a \in M$ there is an integral curve γ for X with $\gamma(0) = a$; any two such integral curves agree on a neighborhood of 0. The idea is to choose a chart (h, U) about a; using the fact that $TU = TM|_U \cong T\mathbb{R}^n$ we can translate the problem of finding an integral curve in U to finding integral curves for a vector field Y on $\mathbb{R}^n$; set $Y_{h(x)} = h_*(X_x) \in T_{h(x)}\mathbb{R}^n$. This gives us a vector field on $\mathbb{R}^n$, which, in the standard coordinates, is just a collection of smooth functions $Y(y) = (f_1(y), \dots, f_n(y))$. We find an integral curve $\eta = (\eta_1, \dots, \eta_n)$ satisfying $\eta'(t) = Y(\eta(t)$ by solving each of the differential equations $\eta_i'(t) = f_i(\eta_i(t)$, which, since the f_i are smooth, each have a unique solution with $\eta_i(0) = x^i(a)$. Pushing this curve back to M using h^{-1} gives an integral curve to X on M. Uniqueness implies that any solution we find using another coordinate chart will agree with this one (push it over with the transition function and it will give "another" solution in our original coordinate system; uniqueness implies that they are actually the same).

Given two vector fields X, Y, we therefore have two sets of integral curves, γ_x tangent to X and η_x tangent to Y. If we choose ϵ small enough, we can, starting with $a \in M$, follow the integral curve to X through a from $t = 0$ to $t = \epsilon$ arriving at $b \in M$, then follow η_b from $b = \eta_b(0)$ to $c = \eta_b(\epsilon)$, then γ_c to $d = \gamma_c(-\epsilon)$, then eta_d to $e = \eta_d(-\epsilon)$. Think of $e = \theta(\epsilon)$ as a curve. This curve, it turns out, is smooth, $\theta^{|}prime(0) = 0$, and $\theta''(0) = 2[X, Y]_a$! So $[X, Y]$ points in the direction that, following the integral curves to X and Y around a "square", we infinitesimally fail to close up.

Vector bundles: The tangent space is our first example of a *vector bundle*. The idea behind a k-bundle E over a space M is that we have a k-dimensional vector space "over" each point of M, which locally fit together as a Cartesian product. More precisely, 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

$$\begin{tikzcd} p^{-1}(U) \arrow[r, "h_U"] \arrow[d, "p"'] & U \times \mathbb{R}^k \arrow[d, "\mathrm{pr}_1"] \\ U \arrow[r, "\mathrm{Id}"] & U \end{tikzcd}$$

commutes. Moreover, for each $x \in M$, the *fiber* $p^{-1}(x) = E_x$ has the structure of a vector space of dimension k so that, for each U with $x \in U$, the restriction of h_U maps E_x to $\{x\} \times \mathbb{R}^k$ by a linear isomorphism.

That is, locally the bundle map looks like the projection onto the first coordinate of a product with the vector space $\mathbb{R}^k$. This is precisely what we have for the tangent space. The basic idea is that each point of M has a vector space $\mathbb{R}^k$ attached to it, so that this assignment is 'locally trivial' in the sense of the commutative diagram. In the case of the tangent space (which we will now start calling the *tangent bundle*), our coordinate charts for TM are the local trivializations we need: $H : TU \to h(U) \times \mathbb{R}^n$ was given by $H(a, X_a) = (h(a), (X_a(X^1), \ldots, X_a(x^n))$, where $h = (x^1, \ldots, x^n)$.

The tangent space of a smooth manifold is just one such family of examples of vector bundles. The *trivial bundle* $M \times \mathbb{R}^k$, often denoted $\epsilon^k(M)$ or simply ϵ^k, with projection map pr_1, is another example. It is called trivial because it has a *global* trivialization. The (open) Möbius band is the total space of a (non-trivial) 1-bundle, or line bundle, over the circle.

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. So it will benefit us to discover some general facts about these objects. First, for any subset $A \subseteq B$, $E_A = p^{-1}(A)$ and the restriction map $p|_A : E_A \to A$ for a k-bundle, as well, the restrictions of the local trivializations for E demonstrate this.

A vector bundle $p : E \to M$ is *smooth* if E and M are smooth manifolds, p is a smooth map, and the bundle admits smooth local trivializations. We will, of course, be interested primarily in smooth bundles. Smooth bundles share several features in common with the tangent bundle. For the tangent space, the transition functions between two local trivializations had the form $t : h(U\cap V)\times\mathbb{R}^n \to k(U\cap V)\times\mathbb{R}^n$ is $(k\circ h^{-1}, D(k\circ h^{-1})$. As a tangent bundle, we supress the first function and think of TU as diffeomorphic to $U\times\mathbb{R}^n$ instead of $h(U)\times\mathbb{R}^n\subseteq\mathbb{R}^{2n}$. So the map on the first coordinate is the identity. In general, for a smooth vector bundle, the transition functions are smooth maps $(U\cap V)\times\mathbb{R}^k \to (U\cap V)\times\mathbb{R}^k$ of the form $t(a,v) = (a,\tau_a(v))$ where $\tau_a \in GL(n,\mathbb{R})$. The map $a\mapsto\tau_a$ is a smooth map from $U\cap V$ to $GL(n,\mathbb{R})$, since the $(i,j)-th$ entry of this map is given by $a \mapsto \mathrm{pr}_i\tau_a(\partial/\partial x^j)$, which is the composition of smooth functions.

A *section* of a vector bundle $p : E \to M$ is a function $s : M \to E$ satisfying $p\circ s = Id$. That is, it is a choice of vector at a for every $a \in M$. A continuous section is a continuous such function; a smooth section is a smooth such function. All bundles have continuous sections; choosing the 0-vector at each point we have the 0-section. If the bundle is smooth, so is this section. Not all bundles have a *nowhere zero section* $s : M \to E$ with $s(p) \neq 0$ for all p; the Möbius band, thought of as a line bundle over the circle, has no such section. Proof: if it did, it would be the trivial bundle (see below: global sections), but it isn't (see below: orientations!).

This statement foreshadows the fact that trivial bundles can be detected using sections. If $p : E \to M$ is a trivial k-bundle, then there is a fiber-preserving homeomorphism $H : E \to M\times\mathbb{R}^k$. If we let e^i represent the i-th coordinate vector, with 1 in the i-th coordinate and 0's elsewhere, then $X_i : x \mapsto h^{-1}(x,e^i)$ is a section of the bundle E, and for every $p \in M$ $x_1(x),\dots,X_k(x)$ form a basis for the vector space $p^{-1}(x) = E_x$. Such a collection of sections is called a *global frame* for the bundle (as in "frame of reference for doing computations"?). Because vector bundles are locally trivial, a similar construction will build *local frames* for us. Trivial bundles therefore have global frames; but the converse is also true: a global frame provides a trivialization of the bundle, by defining $H : E \to M\times\mathbb{R}^k$ by $H(z) = (p(z),(v^1,\dots,v^k))$, where $z = \sum_i v^i X_i(p(z))$. This is well-defined since the $X_i(p(z))$ form a basis for $p^{-1}(p(z))$. Showing it is a bijection is straighforward, and that it is continuous follows from the local triviality of the bundle (*). The inverse is continuous since the sections X_i are. If the bundle is smooth, the same argument establishes that the maps are smooth. A smooth manifold M^n is said to be *parallelizable* if its tangent bundle is fiber-preserving diffeomorphic to the trivial n-plane bundle. by the above, this is equivalent to the existence of n vector fields which are linearly independent at each point.

(*) Continuity of our map *does* follow from local triviality, but the route is more complicated than I first thought. To show that H is continuous/smooth, it suffice to show that it is locally so. Given a local trivialization $h_U : p^{-1}(U) \to U\times\mathbb{R}^n$, we can write $h_U(X_i(a)) = (a,(s_{1i}(a),\dots s_{ni}(a)))$. The functions s_{ji} are continuous/smooth, since h_U and the sections are, and since the sections form a basis, the matrix $A = (s_{ji}(a))$ is invertible for all a. Using Cramer's Rule, we can see that the entries of the inverse matrix B are continuous/smooth functions of a (the inverse is built out of determinants of minors, each of which is smooth on its inputs). S allows us to "build" h_U^{-1} in a way that is compatible with our definition of H; $H\circ h_U^{-1}(a,(v_1,\dots v_k)) = H(\sum_i v_i(h_U^{-1}(0,\dots,1,\dots,0)) = H(\sum_i v_i(\sum_j X_j(a)S_{ji}(a))) = (a,(\sum_i v_iS_{1i}(a),\dots,\sum_i v_iS_{ni}(a)))$, which is smooth. Since h_U is a diffeomorphism, H is smooth.

It turns out that there are a lot more parallelizable manifolds than we have any right to expect. For example, all Lie groups are parallelizable. A Lie group is a group which is also a smooth manifold, such that the multiplication snf inversion maps are smooth. As exmples, the familiar linear groups, $GL(n,\mathbb{R}), O(n), SO(n), U(n)$, etc, are Lie groups. Building vector fields to demonstrate the triviality of the tangent bundle is not too tough; you take your favorite basis for the tangent space at the identity, and then, using the map $g : x \mapsto gx$, we define vector fields by the map $g \mapsto g_*(X_e)$ for our basis vectors X_e at e. The fact that g is a diffeomorphism says that we get a basis at each point $g \in G$, and a little work shows that the vector fields that we have built are smooth.

Orientations: Perhaps one of the most persistent properties of bundles (and manifolds) that one comes across is *orientability.* It is also often a rather elusive concept to pin down. (Consequently, there are a large number of different, though functionally equivalent, ways to define it; not unlike tangent vectors...) We define orientability for bundles over general topological spaces; a smooth manifold M is then called orientable if its tangent bundle is orientable.

The basic idea is that a bundle $p : E \to B$ is orientable if you can assign an "orientation" to each fiber $p^{-1}(a)$ which are compatible with a collection of local trivializations which cover E. An orientation of a vector space V of dimension n is a generalization of the "right-hand rule" that we learn in calculus, except we don't discriminate against left-handed people; recall that a *frame* is any ordered basis $\mathcal{B} = (e_1, \ldots, e_n)$ for V. The standard frame for $\mathbb{R}^n$, for example, would be the coordinate axis vectors in their natural order. Any pair of bases defines an isomorphism $\varphi : V \to V$ which takes one ordered basis $\mathcal{B}$ to the other $\mathcal{B}'$ (and extends linearly); writing this as a matrix (using $\mathcal{B}$, say, in both domain and range), we essentially get the change of basis matrix M, expressing the $\mathcal{B}'$ in terms of $\mathcal{B}$. This matrix is non-singular, since both $\mathcal{B}$ and $\mathcal{B}'$ are bases, and so has non-zero determinant. We say that $\mathcal{B}$ and $\mathcal{B}'$ determine the *same orientation* on V if $\det(M) > 0$; since $\det(MN)=\det(M)\det(N)$, "determine the same orientation" is an equivalence relation on ordered bases. An equivalence class is called an *orientation* on V; there are two of them. Given a frame $\mathcal{B}$ representing an orientation on V and an isomorphism $h : V \to W$, there is an induced orientation on W, with representative $h(\mathcal{B}) = (h(v_1), \ldots h(v_n))$, so that h carries the orientation on V to the orientation on W.

A bundle $p : E \to B$ is *orientable* if every fiber $p^{-1}(a)$ can be given an orientation so that there is an open cover $\mathcal{O}$ of B and a collection of local trivializations $h_U : p^{-1}(U) \to U \times \mathbb{R}^n$ so that for every U and every $a, b \in U$, the induced orientations on $h(a) \times \mathbb{R}^n$ and $h(b) \times \mathbb{R}^n$ (by translation in the U-direction) are the same. Note that, logically, the orientations of the fibers comes first! It is in fact the case that if some cover $\mathcal{O}$ of a path-connected space B establishes that our bundle is orientable, then any cover $\mathcal{O}'$ does; choose two points in an element $U' \in \mathcal{O}'$, a path between them, and a finite open cover of the path by elements U_i of $\mathcal{O}$. Then we can compare the trivializations given by the U_i and U'; along the path we see a continuous family of linear isomorphisms (from $U_i \times \mathbb{R}^n$ to $U' \times \mathbb{R}^n$); applying the determinant we get a continuously changing family of non-zero numbers. They therefore all have the same sign, so have the same sign at either end of the curve. Therefore the maps we get from one end to the other from both trivializations are either both orientation-preserving or both orientation-reversing; since the one from $\mathcal{O}$

is orientation-preserving, so is the one from U'. The point to this is that orientability is a property of the bundle, not of the local trivializations we choose. (Although you need to be a little careful about how you interpret that statement...)

For example, trivial bundles are orientable; the global trivialization $E \to B \times \mathbb{R}^n$ will serve as our cover (i.e., $\mathcal{O} = \{B\}$), and we give each fiber the orientation induced by the inverse of the trivialization map (restricted to each fiber). Therefore, a non-orientable bundle cannot be trivial. Orientability turns out to be a crucial condition in many of the theorems we will encounter as we go forward, especially when we get to learning how to integrate over manifolds.

How do you *prove* that a bundle is not orientable? Well, the quickest way is to show that there is a loop $\gamma : [0,1] \to B$ in B, $\gamma(0) = \gamma(1) = a$, so that in trying to "transport" an orientation over a around γ and back to a, you return with the *opposite* orientation. And what do we mean by transport? Just what we were saying up above about comparing trivializations; over small segments we stay inside of a local trivialization chart, so we can use the orientation at one endpoint to determine what the compatible orientation at the other endpoint needs to be. In this way we can pull ourselves along the loop; if we return to our starting point with the opposite orientation, then one of the charts we used to get around the loop must have demonstrated that the orientations we thought we had at those endpoints were incompatible (no matter what they might have been). So there is no orientation at those points which will be compatible with all of the charts, so the bundle cannot be orientable.

This line of reasoning also demonstrates that we can use charts to *create* the orientations at each point; given a local trivialization and an orientation for a fiber over one point in the trivialization chart, insisting upon compatibility determines what the orientations on every other fiber must be. What we require is that on the overlaps between two charts, the orientations they impose from the chosen points in each chart to the points of the overlap are the same; this amounts to saying that for the transition functions $h_U \circ h_V^{-1} : (U \cap V) \times \mathbb{R}^n \to (U \cap V) \times \mathbb{R}^n$, the linear ismorphisms at each point preserve orientation, i.e., have positive determinant. Conversely, given an open cover by trivializations with that property, we can build a compatible family of orientations on each fiber by starting with an orientation on a single fiber and transporting it around as we have described above.

Boundary orientations: The above discussion works fine for manifolds with boundary. So the concept of an orientable manifold with boundary make sense. If M is orientable, then we can use an orientation on M to build an orientation for $T(\partial M)$ as follows: one of the homework problems talks about inward pointing tangent vectors at ∂M; the same is true for outward pointing vectors, the definition is independent of coordinate chart chosen. If we cover ∂M by coordinate charts (h_i, U_i), and take $M \setminus \partial M$ to complete an open cover of M, we can find a partition of unity subordinate to this cover. Choosing a trivialization of each TU_i and a vector field X_i over each U_i which points outward at the boundary, then $X = \sum_i g_i X_i$ is a vector field on M which points outward at every point of ∂M. If M is orientatble, then we know how to choose an orientation on each fiber that is locally compatible. Look at the orientations for M built along ∂M. We can compare them to the local orientations we can build *for* ∂M (i.e., choose local frames tangent to ∂M) with the vector field X appended at the end; this is a frame for M since X is not in the span of the frame for ∂M. The *induced* orientation for ∂M consists of the local frames for which this augmented frame agrees with the orientation of M at the points of ∂M. Checking that local compatibility holds amounts to the fact that being outward pointing is independent of chart. [The use of outward pointing vectors to build the induced orientation follows historical precedence, since the great integration theorems are always formulated in terms of "outwardly-pointing normals" at the boundary.]

Bundle maps: Given a smooth map $f : M \to N$, we have seen how to construct a smooth map $f_* : TM \to TN$ by pushing forward tangent vectors. Now that TM and TN have more structure - they are vector bundles - we can see that f_* also has more structure; it is a *bundle map*. In general, a bundle map between to vector bundles $p : P \to M$ and $q : Q \to N$ is a pair of maps $f : M \to N$ and $F : P \to Q$ so that F *covers* f, that is, $q \circ F = f \circ p$ (so F sends fibers P_x to fibers $Q_{f(x)}$), and F restricts to a linear map on fibers $F : P_x \to Q_{f(x)}$. A bundle map is continuous or smooth if both maps are. A bundle isomorphism is a bundle map f, F where f is an isomorphism (homeo- or diffeo-), and F is an isomorphism on each fiber. A global trivialization $E \to M \times \mathbb{R}^k$ is therefore "just" a bundle isomorphism from a bundle to the trivial bundle. If $M = N$ and $f =$Id, then F is called a *bundle map over M*.

Once you have maps, you of course have the notion of two objects being the "same": A bundle isomorphism is a bundle map which is a homeomorphism/diffeomorphism on the bases spaces and an isomorphism on every fiber. So saying, for example, that a bundle is trivial really means that it is bundle isomorphic to the trivial bundle over the same base space.

Given a vector bundle $p : E \to B$ and a map $f : X \to B$, there is a procedure for building a bundle $q : Y \to X$ and a bundle map $F : Y \to E$ over f, called the *pullback*. The total space Y consists of pairs $(x, e) \in X \times E$ satisfying $f(x) = p(e)$; $q : Y \to X$ is projection onto the first coordinate. $Y_x = q^{-1}(x)$ is "really" $E_{f(x)}$, that is, for a fixed $x \in X$ the $(x, e) \in Y$ are literally the $e \in E$ sitting over $f(x)$. To build local trivializations, given $a \in X$ choose an open neighborhood U of a so that $f(U)$ lies inside the domain V of some locally trivial neighborhood about $f(a)$. Then the map $h_U : q^{-1}(U) \to U \times \mathbb{R}^n$ given by $h_U(x, e) = (x, r(e))$, where $h_V(e) = (p(e), r(e))$. Since h_V is an isomorphism on each fiber, so is h_U; since r is continuous/smooth, so is h_U. The standard notation for the pullback Y in the literature is $f^*(E)$. It plays a crucial role in the study of classifying vector bundles up to isomorphism; it plays a less crucial role in the work we will be doing here. In the structure theory the basic idea is that every n-plane bundle is the pullback of a specific ("universal") bundle over a specific space, the Grassmanian of all n-planes through the origin in a sufficiently high dimensional Euclidean space (think: infinite-dimensional). The fiber of the bundle over a point <u>is</u> the n-plane. Understanding bundles up to isomorphism can then be turned into understanding maps into the Grassmanian up to homotopy; then the tools of algebraic topology can be turned onto the problem.

New bundles out of old ones: There are many ways to combine/transform vector bundles over the same base space B to create new vector bundles over B. Several of these will be of central importance to us as we continue to build new tools to study smooth manifolds.

Given two vector bundles $p : E \to B$ and $p' : E' \to B$, if we think of them as assignments of vector spaces E_a, E'_a to each point, any way that we can combine them to produce a new vector space can, with a little care, be stitched together locally to construct a new bundle. For example, we can take the direct sum of the two, producing the vector space $E''_a = E_a \oplus E'_a$ at each point. Their (disjoint) union over $a \in M$ we will call E''; the projection map to B is the obvious one. Local trivializations can be constructed out of trivializations $h_U : p^{-1}(U) \to U \times \mathbb{R}^k$ and $h'_U : (p')^{-1}(U) \to U \times \mathbb{R}^\ell$ (taking intersections to make the projections of the domains the same) by taking $h''_U : (p'')^{-1}(U) \to U \times \mathbb{R}^{k+\ell}$ to be $h''_U(e) = (p''(e), (H_U(e), H'_U(e)))$, where H_U, H'_U are the second coordinate functions for h_U, h'_U. A quick check show that these have all of the necessary properties, turning E'' into a $(k+\ell)$-dimensional vector bundle, called the *Whitney sum* of E and E', and denoted $E'' = E \oplus E'$. For example, $TS^n \oplus \epsilon^1(S^n) \cong \epsilon^{n+1}(S^n)$.

Given two bundles E, E' over B, with E'_a a subspace of E_a for every a, we can form the *quotient bundle* E/E' with fibers E_a/E'_a. The reader is invited to construct the appropriate local trivializations. This construction can be thought of in some sense as the inverse of the Whitney sum construction.

Given a vector space V, there is the dual vector space V^*=Hom$(V, \mathbb{R})$ of linear maps to $\mathbb{R}$; it has the same dimension as V (and, given a positive definite inner product on V, is isomorphic to V by the map $v \mapsto (w \mapsto \langle w, v\rangle))$. The duals to the vector spaces E_a over points can also be assembled into a "dual" vector bundle. The local trivializations are "dual" to the trivializations for V, that is, given $h_U = (p, H_U)$, we build h^*_U by setting $H^*_U(e) = \{w \mapsto \langle w, H_U(e)\rangle\}$, where $\{\cdot,\cdot\}$ is the standard inner product on $\mathbb{R}^k$. Again, these trivializations have the necessary properties; the resulting bundle is called the *dual* bundle to E, and is denoted E^*. When $E = TM$ is the tangent bundle to M, it is called the *cotangent* bundle to M, and denoted T^*M.

So what do cotangent vectors (tangent covectors?) look like? In local coordinates, T^*_aM has a basis dual to the coordinate basis $\partial/delx^1, \dots, \partial/\partial x^n$ for T_aM, which we denote dx^i; that is, $dx^i(\partial/\partial x^j) = \delta_{ij}$. So in local coordinates, a tangent covector field would be given by $\sum f_i(a)dx^i$. Does this notation look familiar? A covector field would be a choice of covector at each point, that is, a function from tangent vectors to $\mathbb{R}$ at each point. If we think more globally, a covector field ω would take a vector field, a choice of tangent vector at each point, to a number at each point. That is, it would take a vector field on M to a function on M. And one thing that can do that is a function (suitably interpreted). That is, instead of thinking of the expression Xf as starting with a vector field and feeding it a function, we can start with a function and feed it a vector field! To denote this change in roles, we introduce the *differential* of a (smooth) function f, denoted df . This is a covector, wherever f is defined, whose value is given by $df(X_a) = X_a(f)$. Since the expression $X_a(f)$ is linear in X_a, this is a covector. To see what this looks like in local coordinates, note that $df(\partial/\partial x^i) = \partial f/\partial x^i$, and $(\sum_j f_j dx^j)(\partial/\partial x^i) = f_i$, so $df = \sum_i \partial f/\partial x^i dx^i$. This <u>is</u> the standard notation for differentials in multivariate calculus.

Some properties of differentials, which can be verified either in coordinates or straight from the definition:

(a) $d(\alpha f + \beta g) = \alpha df + \beta dg$ for functions f, g and constants α, β.

(b) $d(fg) = f(dg) + g(df)$

(c) $d(f/g) = [g(df) - f(dg)]/g^2$ wherever $g \neq 0$

(d) $d(\text{constant}) = 0$. Conversely $df = 0$ on a connected open set implies that f is constant on that set. (This is just a reinterpretation of one of your homework problems...)

(e) If $f : M \to \mathbb{R}$ and $h : \mathbb{R} \to \mathbb{R}$ is smooth, then $h \circ f$ is smooth and $d(h \circ f) = (h' \circ f)(df)$.

Can every covector field ω be expressed as $\omega = df$ for some f (perhaps just locally)? Unfortunately (or fortunately) no, if $\omega = \sum f_i dx^i = df$ then $f_i = \partial f/\partial x^i$ for each i, and so $\partial f_i/\partial x^j = \partial f_j/\partial x^i$ for every i and j (since mixed partials are equal). Which of course doesn't always happen.

Tangent vectors push forward under a smooth map $f : M \to N$; The map $f_* : T_aM \to T_{f(a)}N$ was given by $(f_*(X_a))g = X_a(g \circ f)$. Covectors, on the other hand, pull back. In general, a linear map $L : V \to W$ induces a linear map $L^* : W *^{\to} V^*$, its *dual*, given by $L^*(\varphi)v = \varphi(Lv)$. Given a covector $\omega \in T^*_{f(a)}N$, we can define $f^*\omega \in T^*_aM$ by $f^*\omega(X_a) = \omega(f_*X_a)$. Unlike the pushforward, however, the pullback extends to cotangent *fields*, without any condition required. That is, if ω is a cotangent field on N, then we can define a cotangent field $f^*\omega$ on M by, well, it's the same formula! $f^*\omega(X_a) = \omega(f_*X_a)$; this is well-defined, since a point $a \in M$ has exactly one image in M under f; there is no ambiguity involved. Writing this in local coordinates, we find that if $\omega = \sum_i g_i dy^i$, then $f^*\omega = \sum_i f_i dx^i$ for some smooth functions f_i; to figure out which, we evaluate $f_j = (\sum f_i dx^i)(\partial/\partial x^j) = \omega(f_*(\partial/\partial x^j)) = \omega(\sum_i \partial(y^i \circ f)/\partial x^j]\partial/dely^i) = \sum_i g_i \partial(y^i \circ f)/\partial x^j$, so $f^*\omega = \sum_j(\sum_i g_i \partial(y^i \circ f)/\partial x^j)dx^j$.

Double duals: Given a vector space V, the dual V^* is of course a vector field. So it has its own dual $(V^*)^* = V^{**}$. But unlike V and V^*, we can define an isomorphism $\varphi : V \to V^{**}$ without choosing a basis first; given $v \in V$, we define $\varphi(v) = v^{**} \in V^{**} = \text{Hom}(V^*, \mathbb{R})$ by $v^{**}(f) = f(v)$. (This is the classic "make the variable the function and the function the variable" switch.) This map is linear (compare $(\alpha v + \beta w)^{**}(f)$ and $(\alpha v^{**} + \beta w^{**})(f)$), and injective: if $0 = v^{**}(f) = f(v)$ for all $f \in V^*$, then if we write $v = \sum a_i e_i$, then setting $f = \sum a_i e_i^*$ we have $f(v) = \sum a_i^2 = 0$, so $a_i = 0$ for all i and $v = 0$. If we cheat and only give a proof for finite dimensional vector spaces, then injective linear maps between vector spaces of the same dimension are surjective, and so φ is an isomorphism.

Since this isomorphism does not invoke a basis for its construction, we can use it on every fiber of a vector bundle to build an isomorphism $E \to E^{**}$ over the identity map on the base space. So the double dual E^{**} of any vector bundle E is isomorphic to E.

Inner products: We've seen that the double dual of a bundle is isomorphic to the bundle. But what about the dual itself? To answer that, we need to introduce something which we have so far avoided introducing into a vector bundle: an inner product. An inner product on a vector space is a bilinear function $(v, w) \mapsto \{v, w\}$ which is symmetric ($\{v, w\} = \{w, v\}$) and *non-degenerate*: for every $v \neq 0$ there is a w with $\{v, w\} \neq 0$. The inner product is *positive definite* if $\{v, v\} > 0$ for every $v \neq 0$. Note that positive definiteness immediately implies non-degeneracy (take $w = v$). An positive definite inner product allows us to define the norm of a vector: $||v|| = \sqrt{\{v, v\}}$ satisfies the triangle inequality. We can also define angles using the law of cosines $\cos A = \{v, wl\}/(||v|| \cdot ||w||)$. Perhaps most importantly, we can define orthogality: $\{v, w\} = 0$.

All this can be carried over to vector bundles. A *Riemannian metric* on a vector bundle (p, E, B) is a choice of positive definite inner product $\phi_b(\cdot, \cdot) = \{\cdot, \cdot\}_b$ for every fiber $b \in B$, which varies continuously, in the sense that for any pair of sections $v_1(b), v_2(b)$ of E the function $\{v_1(b), v_2(b)\}_b$ is continuous. If the bundle is smooth, we can also require that the assignment of inner product vary smoothly. (A *pseudo-Riemannian metric* is a choice of non-degenerate inner product.) A Riemannian metric on the tangent bundle TM is referred to as a Riemannian metric on M. And M equipped with a Riemannian metric is called a *Riemannian manifold.* Riemannian metrics allow to introduce angles (and therefore orthogonality) into differential topology.

Provided we can show that bundles have Riemannian metrics! But, it turns out, they do. Certainly they have Riemannian metrics locally; a trivial bundle $U \times \mathbb{R}^n$ has an Riemannian metric obtained by choosing any (e.g., the standard) positive definite inner product on $\mathbb{R}^n$ and using it on every fiber. We can then stitch these locall-defined inner product together using (what else?) a partition of unity. Given a partition of unity g_i subordinate to an open cover $\mathcal{O}$ of the base manifold M by locally trivializable sets, and a positive definite inner product ϕ_U (as above) on each $p^{-1}(U)$, since a finite positive linear combination of positive definite inner products is positive definite (and an inner product!), we can define an inner product on each fiber E_b by $\{v, w\}_b = \sum_i g_i(b)\phi_{U_i}(v, w)$ (where $\phi_{U_i}(v, w) = 0$ if $b \notin U_i$). This defines a Riemannian metric on E, which is smooth if E is smooth. (The locally defined inner products incorporate the (smooth) trivializations into their definition, to transport a given inner product to each fiber of $p^{-1}(U_i)$.)

The dual bundle again: The existence of a Riemannian metric allows us to answer the question of when E is isomorphic to E^*, at least in the case of smooth vector bundles over smooth manifolds. The answer is, they always are! The point is that a Riemannian metric $\{\cdot, \cdot\}_b$ allows us to define a (bundle) map $\psi : E \to E^*$ by $\psi(e)(e') = \{e, e'\}_{p(e)}$. This map is certainly linear in each fiber, and by definition of a Riemannian metric it is smooth. It is also an isomorphism on each fiber, by the non-degenerateness of the metric; if $\psi(e)(e') = \{e, e'\}_{p(e)} = 0$ for every $e' \in E_{p(e)}$, then $e = 0$. So the map $E_b \to E_b^*$ is injective and linear, hence surjective, hence an isomorphism.

Riemannian metrics again: So why is an inner product on the tangent bundle called a metric? Because it can be used to build one. The inner product allows us to compute a norm of a tangent vector; in calculus, we learn that to compute the length of a (parametrized) curve, we integrate the length of its velocity vector. We can do the same thing in a Riemannian manifold M. Given a smooth curve $\gamma : [0, 1] \to M$, we define its length to be $L(\gamma) = \int_0^1 ||\gamma_*(d/dt)|_{\gamma(t)}||\ dt$. As in calculkus, we can show that this is independent of prarametrization; if $h : [0, 1] \to [0, 1]$ is a smooth increasing bijection, then $L(\gamma \circ h) = \int_0^1 ||(\gamma \circ h)_*(d/dt)|_{\gamma(h(t))}||\ dt = \int_0^1 ||\gamma_*(d/dt)|_{\gamma(h(t))}h'(t)||\ dt = \int_0^1 ||\gamma_*(d/dt)|_{\gamma(h(t))}||h'(t)\ dt = \int_0^1 ||\gamma_*(d/dt)|_{\gamma(u)}||\ du = L(\gamma)$, using the u-substitution $u = h(t)$. So, essentially, the length of a curve is a function its image, not how we traverse it.

Once we can define the length of a curve, we can define a metric on M by choosing the "shortest" curve between two points; $d(a, b) = \inf\{L(\gamma) : \gamma$ a piecewise smooth curve from a to $b\}$. By using piecewise smooth curves, verifying the triangle inequality requires no work; take curves close to the infimum for $d(a,b)$ and $d(b,c)$ and concatenate them to get a curve from a to c whose length is ϵ above the sum, and let ϵ goto 0. Symmetry is also quick; use the curve $\overline{\gamma}(t) = \gamma(1-t)$ to compare $d(b,a)$ to $d(a,b)$. The constant path demonstrates that $d(a,a) = 0$. The only property of a metric that takes real work is showing that $d(a,b) = 0$ implies $a = b$.

The idea behind this is to push our computations into $\mathbb{R}^n$ usng a chart. That is, given a chart (h,U) about a and $k = h^{-1}$, we can put a Riemannian metric on $\mathbb{R}^n = h(U)$ by defining $\langle v,w\rangle_b = \langle k_*v, k_*w\rangle_{k(b)}$. At every point b this is an inner product, defining a norm whose unit sphere, in the usual inner product, is an ellipse with minor axis $\alpha(b)$ and major axis $\beta(b)$. These functions are smooth and positive, so on a compact neighborhood V of $h(a)$ α has a minimum α_0 and β has a maximum β_0. Then the length $L(\gamma)$ of any curve that stays in V, from the point of view of the pushed forward inner product, is bounded between α_0 and β_0 times the length computed from the point of view of the ordinary inner product on $\mathbb{R}^n$. In particular, really short curves must also be really short from the point of view of the ordinary inner product, and so stay inside of V. And, finally, if $d(a,b) = 0$, we can find arbitrarily short curves, which are therefore arbitrarily short from the ordinary Euclidean metric point of view, and so $a = b$. So the metric $d(a,b)$ is a metric on M.

Even more, the metric topology on M from the Riemannian metric is the same as the manifold topology on M. This also follows from the basic inequality above; it's enough to show that $d_M(a,b) \leq c \cdot d_E(a,b)$ for some c, where d_E is the ordinary Euclidean metric on $\mathbb{R}^n$ and d_M is the transported metric from M; then standard facts about metrics imply that the d_E metric topology is finer than the d_M topology. The reverse inequality gives a coarser topology.

Immersions, submersions, and embeddings

The inverse function theorem: We are going to take a break from our march towards ever more interesting constructions of vector bundles out of the tangent bundle to discuss something which is also at the heart of calculus, namely linear approximation. That is, to what extent does the "derivative" F_* of a smooth function $F : M \to N$ approximate the function? On the face of it this is a non-question, because unlike in calculus the values of F_*, tangent vectors, live in different places than the values of F. So the one can't approximate the other. But some things still hold true. In particular, there is a result from (advanced) calculus, the Inverse Function Theorem, which says (in our current terminology) if F_* is invertible at a point, then F is invertible in a neighborhood of that point. (Lee's text has a proof of this, in Chapter 7.) This is a fundamental result, which we will often use. In the case of smooth manifolds, the proof amounts to imposing local coordinates (h,U) and (k,V) about a and $F(a)$, noticing that the hypothesis still holds in $\mathbb{R}^n$ (the Jacobian of $k \circ F \circ h^{-1}$ is invertible; that's what F_* being invertible means!), and invoking the conclusion of the $\mathbb{R}^n$ version to give our conclusion ($k \circ F \circ h^{-1}$ invertible implies F is, since h and k are invertible).

A smooth map $F : M^n \to N^m$ is an *immersion* if $F_* : T_aM \to T_{f(a)}N$ is an injective linear map for every $a \in M$. (Note that thsi requires $n \leq m$.) F is a submersion if F_* is surjective at every point. (We need $n \geq m$.) A 1-to-1 immersion which is a homeomorphism onto its image is an *embedding.* The cheapest way to insure that an injective immersion is an embedding is to have the domain M compact, since N is by hypothesis Hausdorff. The homeo onto image is not a vacuous condition, as the map $f : (-2\pi, 2\pi) \to \mathbb{R}^2$ given by $f(t) = (t, 0)$ on the left half of the interval, and $= (\cos t, \sin t)$ on the right half shows.

One way to interpret the inverse function theorem is that if F_* is an isomorphism at a point a then (since F is a diffeomorphism in a neighborhood of a) there are charts (h, U) about a and (k, V) about $b = F(a)$, namely $k = h \circ F^{-1}$, so that in local coordinates, $k \circ F \circ h^{-1} = h \circ F^{-1} \circ F \circ h^{-1} = Id$ is the identity function. That is, there is a local model for the map F in which it looks like (and therefore behaves like) the identity. The key here is that $F_*(a) : T_aM \to T_{F(a)}N$ is a map of rank n between n-dimensional vector spaces, and so it is automatically true that there is a neighborhood of a for which this is true; locally, the rank of F_* can only go up (it's the size of the largest invertible $k \times k$ minor), so in this case it must stay n. In the case of an immersion, the same is true; since F_* is injective, its rank equals $n =$dim(M), and can't get higher, so locally it remains n. The same is true of a submersion; the rank of F_* equals $m =$dim(N), and again, can't get higher. So all of these maps fit the pattern of: the rank of F_* is constant in a neighborhood of a. And it is for such a situation, it turns out, that we can build the most meaningful local models. The general statement is: if $F : M^n \to N^m$ is smooth, and for some $a \in M$ there is a neighborhood U of a for which $F_*(b) : T_bM \to T_{f(b)}N$ has constant rank k (note that then $k \leq n, m$), then there are charts about a and $F(a)$ so that, in local coordinates, F looks like $(x_1, \dots, x_n) \mapsto (x_1 \dots, x_k, 0, \dots, 0)$.

The proof of this consists, first, of choosing any two charts h and k about a and $F(a)$ and noting that $k \circ F \circ h^{-1}$ has constant rank k in a neighborhood of $h(a)$, reducing the problem to a question of maps between open subsets of Euclidean spaces. Then we invoke the corresponding theorem for Euclidean spaces, which is, really, a result from analysis. But since it follows fairly directly from the inverse function theorem, we present the proof; it also requires us to build charts on Euclidean spaces, so it really still belongs in the realm of differential topology. To prove this, note that the hypothesis implies that the matrix of partial derivatives of $F = (f_1, \dots f_m) : \mathbb{R}^n \to \mathbb{R}^m$ has an invertible $k \times k$ minor, which, by permuting the coordinates of domain and range, we may assume is the uper left corner of the matrix. (Think of these permutations as the first level of a series of coordinate charts we build around F to make the map successively nicer.) We now view F as a function $F(x, y) = (Q(x, y), R(x, y))$, $F : \mathbb{R}^k \times \mathbb{R}^{n-k} \to \mathbb{R}^k \times \mathbb{R}^{m-k}$. We may assume, by translating a to $(0, 0)$ and adding a constant, that $F(0, 0) = (0, 0)$. Our hypothesis tells us that the matrix $M(b)$ of partial derivatives of Q with respect to the first k variables is non-singular, for b in a neighborhood of $(0, 0)$. If we consider the map $G : \mathbb{R}^n \to \mathbb{R}^n$ given by $G(x, y) = (Q(x, y), y)$, then the matrix of partial derivatives of G at $(0, 0)$ is a block matrix with the partials of Q at top and a block of 0s followed by the Identity matrix at bottom. This matrix is non-singular precisely where $M(b)$ is; in particular it is non-singular at $(0, 0)$. So G is a diffeomorphism in a neighborhood of $(0, 0)$, and using it as a chart in the domain of F, and writing its inverse G^{-1} as $G^{-1}(x, y) = (A(x, y), B(x, y))$, we find that $(x, y) = G(A(x, y), B(x, y)) = (Q(A(x, y), B(x, y)), B(x, y))$, so $y = B(x, y)$ and $G^{-1}(x, y) = (A(x, y), y)$. Consequently, $(x, y) = G(A(x, y), y) = (Q(A(x, y), y), y)$, so $x = Q(A(x, y), y)$, and so $F \circ G^{-1}(x, y) = F(A(x, y), y) = (Q(A(x, y), y), R(A(x, y), y)) = (x, R(A(x, y), y)) = (x, S(x, y))$ where S is the obvious function. The matrix of partial derivative of this function is also a block matrix, with the identity matrix and 0s on the top and the partials of S on the bottom. Since the composition with the diffeomorphism G^{-1} cannot change the rank of the derivative, this matrix has rank K in a neighborhood of $(0,0)$, and therefore the lower right corner of the matrix must be identically 0 in this neighborhood. But by the mean value theorem this implies that S is constant on the $n-k$-dimensional planes where the first k variables are held consant, i.e., $S(x,y) = S(x)$ is a function only of the first k variables. This allow us to build a chart on $\mathbb{R}^m = \mathbb{R}^k \times \mathbb{R}^{m-k}$ to finish our transformation; The map $H : (x,z) \mapsto (x, z - S(x))$, on a neighborhood of $(0,0)$, gives a smooth (since S is smooth) function for which $H \circ F \circ G^{-1}(x,y) = H(x, S(x)) = (x,0)$, which is projection onto the first k coordinates. Note that H is a diffeomorphism (on the neighborhood of $(0,0)$ where S is a function of x alone); its inverse is $H^{-1}(x,z) = (x, z + S(x))$. Therefore we have built charts G and H on $\mathbb{R}^n$ and $\mathbb{R}^m$ so that in these local coordinates F is projection onto the first k coordinates, as desired.

This immediately implies that for an immersion there are charts so that in local coordinates the map looks like $x \mapsto (x,0)$. For a submersion, there are charts so that the map locally looks like $(x,y) \mapsto x$.

The implicit function theorem: Another useful result that follows from the inverse function theorem is the implicit function theorem: If W is an open subset of $\mathbb{R}^n \times \mathbb{R}^k$ and $F = (F_1, \dots, F_k) : W \to \mathbb{R}^k$ is a smooth map, $(x,y) \mapsto F(x,y)$, and the matrix of partial derivatives $(\partial f_i / \partial y_j(a,b))$ is non-singular, then there are neighborhoods U of a and V of b and a smooth map $G : U \to V$ so that $F^{-1}(F(a,b)) \cap (U \times V)$ is the graph of G. That is, $(x,y) \in U \times V$ satisfies $F(x,y) = F(a,b) \Leftrightarrow \ y = G(x)$. that is, under the non-singularity condition, a point inverse of F looks, locally, like the graph of a smooth function.

The proof is, again, a matter of building the right charts. On the domain side, we use the map $H : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n \times \mathbb{R}^k$ given by $H(x,y) = (x, F(x,y))$. At (a,b) the derivative is an isomorphism, so H is a diffeomorphism in a neighborhood of (a,b), and arguing just as before, we find that we can express the inverse as $H^{-1}(x,y) = (x, E(x,y))$ for a smooth map E. Setting $F(a,b) = c$, we define $G(x) = E(x,c)$, which is a smooth function defined in a neighborhood of a. From $(x,c) = H(H^{-1}(x,c)) = (x, F(x, E(x,c))) = (x, F(x, G(x)))$, we have $F(x, G(x)) = c = F(a,b)$, so $y = G(x) \Rightarrow F(x,y) = F(a,b)$. For the opposite implication, $F(x,y) = c$ implies that $H(x,y) = (x,c)$, so $(x,y) = H^{-1}(x,c) = (x, E(x,c)) = (x, G(x))$, so $y = G(x)$.

Some useful things follow fairly quickly from our Rank Theorem. For example, a submersion is an open map, since in local coordinates it looks like throwing away the last few coordinates, which takes a small open cube to a small open cube. Therefore, locally, it takes an open set (thought of as a union of small open cubes) to a union of open cubes, i.e., an open set. A surjective submersion $F : M \to N$, therefore, is a quotient map, because surjective, continuous, open maps are quotient maps. [Proof: $F(F^{-1}(V)) = V$, since F is surjective; so if $F^{-1}(V)$ is open, so is V.] Submersions behave, locally, like bundle maps; for every $y_0 = F(x_0) \in F(X)$, there is a neighborhood V of y_0 and a smooth "section" $s : V \to M$ with $s(y_0) = x_0$ and $F \circ s = \text{Id}$. We simply choose charts $(h,u),(k,V)$ as in the theorem and take $s(x) = h^{-1}(k(x), 0, \dots, 0)$.

Even more, surjective submersions satisfy a smooth analogue of the "fundamental lemmas" of quotient spaces. First, if $F : M \to N$ is a surjective submersion and $g : N \to P$ is a map, then g is smooth $\Leftrightarrow$ $g \circ F$ is. $\Rightarrow$ is immediate; for the opposite implication, at $a \in M$ we choose charts on M and N as in the theorem and any chart around $g(F(a))$, then in local coordinates the composition looks like $(x_1, \dots, x_n) \mapsto (x_1, \dots x_m) \mapsto (g_1(x_1, \dots, x_m), \dots, g_r(x_1, \dots, x_m))$, which has continuous partial derivatives to all degrees for all choices of variables $x_1, \dots, x_n$, so the g_i certainly do when only choosing among $x_1, \dots, x_m$. With these charts, therefore, g is seen to be C^∞.

Second, if $F : M \to N$ is a surjective submersion and $g : M \to P$ is a smooth map such that $F(x) = F(y)$ implies $g(x) = g(y)$, then there is an induced map $G : N \to P$ defined by $G(z) = g(x)$ where $F(x) = z$. Since F is a quotient map, G is continuous; but even more, G is smooth. This is because $G \circ F = g$ is smooth. Put in words, a smooth map to P from the domain of a surjective submersion F which is constant on point inverses of F induces a smooth map from the quotient to P. As in point-set topology, this provides a far less painful way to build smooth maps out of the quotients of smooth manifolds.

Embedded submanifolds: So far we haven't really talked about the usefulness of immersions, just submersions. The languge of immersions allow us to talk meaningfully of a manifold obtaining a smooth structure from an embedding into (for example) Euclidean space $\mathbb{R}^N$, which was probably really the original way to think about smooth manifolds; they look locally, like the graph of a smooth function $\mathbb{R}^n \to \mathbb{R}^{N-n}$ (which, being an immersion, really locally looks like the graph of the 0-function). We start with a definition. A subset $S \subseteq M^n$ of a smooth manifold M^n is called an *embedded submanifold of dimension* k if for every $a \in S$ there is a chart $h(U)$ of M about a so that $h(S \cap U) = \mathbb{R}^k \times \{0\} \subseteq \mathbb{R}^n$. The terminology is that $h(S \cap U)$ is called a k-slice of U, and h is a slice chart for S in M. The complementary dimension, $n - k$, is called the *codimension* of S in M.

Note that, as a subset of M, in the subspace topology S is Hausdorff and second countable; the k-slices also demonstrate that S is locally Euclidean (of dimension k). But even more, we can use the slice charts to assemble a smooth structure on S; composing $s|S \cap U$ with the projection to the first k cooridinates of $\mathbb{R}^n$ form a collection of charts on S. The inverse of this chart is the inclusion of $\mathbb{R}^k$ into $\mathbb{R}^n$ as $\mathbb{R}^k \times \{0\}$, followed by the inverse of s, and so the transition function for two such charts consists of inclusion of $\mathbb{R}^k$ as $\mathbb{R}^k \times \{0\}$, followed by the transition function of the two charts on M, followed by projection, all of which are smooth, so the charts on S form a smooth atlas. They therefore define a smooth structure on S, which makes the inclusion map $\iota : S \to M$ smooth.

Furthermore, this is the <u>only</u> smooth structure on S making ι smooth. To see this, suppose that (k, V) is a chart (not necessarily in our smooth structure) for which $\iota \circ k^{-1} : k(U) \to M$ is smooth; we show that (k, V) actually <u>is</u> in our smooth structure, that is, it is C^∞-related to all of our slice charts (well, a collection of slice charts that cover S, anyway). Given $a \in V$ and a slice chart (h, U) for M about a, we may assume (since we are investigating a local condition (smoothness)) that $\iota(V) \subseteq U$, and we have $h \circ \iota \circ k^{-1} : \mathbb{R}^k \to V \to U \to \mathbb{R}^n$ is smooth, by hypothesis. Since h is a slice chart, $h|_S : S \cap U \to \mathbb{R}^k \times \{0\}$ is a chart on S, and $h|_S \circ \iota \circ k^{-1} : \mathbb{R}^k \to V \to U \to \mathbb{R}^k$ is also smooth, since $h \circ \iota \circ k^{-1}$ really maps into $\mathbb{R}^k \times \{0\}$, anyway, so it has the same matrix of partial derivatives (ignoring the 0 columns from mapping into $\mathbb{R}^k \times \{0\}$). But this map is "really" just $h|_S \circ k^{-1}$, since the middle ι really only serves to redefine the codomain of

k^{-1}, which we no longer really care about. Since this map is smooth, we have half of our C^∞-relatedness. To prove the other half, we wish to use the Inverse Function Theorem! It suffices to show that $h|_S \circ k^{-1}$ has invertible Jacobian matrix at $k(a)$; then, locally, $h|_S \circ k^{-1}$ is invertible and has smooth inverse. And for this, it suffices to show that $(h|_S \circ k^{-1})_*$ is injective at $k(a)$. But this map is the same as (where pr_k is the projection onto the first k coordinates of $\mathbb{R}^n$) $(\text{pr}_k \circ h \circ k^{-1})_* = (\text{pr}_k)_* \circ h_* \circ k_*^{-1} = \text{pr}_k \circ h_* \circ k_*^{-1}$, and h_* and k_*^{-1} are isomorphisms (being charts), while pr_k is injective <u>on</u> <u>the</u> <u>image</u> <u>of</u> $h_* \circ k_*^{-1}$, which lies in $\mathbb{R}^k \times \{0\}$. So the derivative map is injctive, sohwing that $h|_S \circ k^{-1}$ is a diffeomorphism, as desired.

With this smooth structure on S, S becomes a smooth k-manifold in its own right, and (since we are using the subspace topology) $\iota : S \to M$ is then a smooth embedding, that is, an immesrion which is a homeomorphism onto its image. So embedded submanifolds are images of smooth embeddings. Using the Rank Theorem, the converse is readily established; the image of a smooth embedding is an embedded submanifold. This is because given an embedding $f : M^n \to N^m$ and setting $S = f(M)$, given $a \in S$ we let $b = f^{-1}(a)$, so $f(b) = a$ and by the rank theorem there are charts (h, U) around b and (k, V) around a so that $k \circ f \circ h^{-1}$ is the inclusion into the first n coordinates of $\mathbb{R}^m$. (WOLOG, $f(U) \subseteq V$.) In particular, $k(f(M) \cap V) = k(S \cap V)$ lies in $\mathbb{R}^n \times \{0\}$. To finish, we need only show that this slice $\mathbb{R}^n \times \{0\}$ contains an open neighborhood of $k(a)$; but this follows from Invariance of Domain, since $\text{pr}_n \circ k \circ f \circ h^{-1}$ is a continuous injective map from the open set $h(U) \subseteq \mathbb{R}^n$ to $\mathbb{R}^n$. So (on a possibly smaller neighborhhood than V) $k(f(M) \cap V)$ <u>equals</u> the entire slice.

The tangent space: An embedded submanifold S of a smooth manifold M is itself a smooth manifold, so it has a tangent bundle. For $a \in S$, we can identify T_aS with a subspace of the vector space T_aM, allowing us to treat TS as a subbundle of $TM|_S$. (Did we talk about subbundles? A subbundle F of E is a bundle such that F_a is a subspace of E_a for every $a \in B$. If we want it to be a smooth subbundle, we insist that locally F is the span of a collection of smooth sections of E.) Using the inclusion map $\iota : S \to M$, which the above discussion shows is a smooth immersion, we can make this identification concrete, identifying T_aS with $\iota_*(T_aS) \subseteq T_aM$. From the point of view of derivations, a tangent vector X of T_aS is identified with the tangent vector $\widetilde{X}$ in T_aM satisfying $\widetilde{X}f = X(f|_S)$ for $f \in C^\infty(M)$. Turning this around, an element $X \in T_aM$ (for $a \in S$) is <u>really</u> an element of T_aS (that is, is in the image of ι_*) if, working in the local coordinates where $h(U \cap S)$ is the subspace with all but the first k coordinates equal to 0, $X = \sum_{i=1}^k a_i \partial/\partial x^i$, that is, $X(x^j) = 0$ for $j > k$. But this in turn can be characterized as saying that if $f \in C^\infty(M)$ is constant on S, then $Xf = 0$, and since $Xc = 0$ for all constant functions, this is the same as saying that $Xf = 0$ for any smooth function that is 0 on S. So we have a characterization of the tangent vectors to S, they are the tangent vectors to M which are 0 on smooth functions that vanish on S.

Given a submanifold $S \subseteq M$, we now have TS as a subbundle of $TM|_S$. Together with a Riemannian metric on M, we can create a complementary subbundle NS of $TM|_S$ as, for $a \in S$, the set of vectors in TM_a orthogonal to every vector in TS_a. If we choose a slice chart for S about a, and take the standard basis vectors, the first k in TS_a, and apply Gram-Schmidt to them, a recent homework problem states that the local frame will be carried to a local frame; the first k still span TS_a, and the remainder span NS_a. This gives us a smooth local frame of NS, so NS is smooth. Since $TS_a \oplus NS_A = TM_a$ at every point, the normal bundle NS, as it is called, allows us to split $TM|_S$ as $TS \oplus NS$. For example, applying this to the standard embedding of S^n in $\mathbb{R}^{n+1}$, and noting that the normal bundle has a nowhere-zero (outward pointing) section, so $NS^n = \varepsilon^1$, we find that $TS^n \oplus \varepsilon^1 \cong \varepsilon^{n+1}$. (It is a fact that since all Riemannian metrics are the "same", different choices of Riemannian metric yield isomorphic normal bundles; two such metrics can be joined by a path of metrics, and "nearby" metrics build isomorphic bundles.)

Level sets: So where do we get embedded submanifolds, anyway? One place is from smooth maps of constant rank! For example, the function $F : \mathbb{R}^{n+1} \to \mathbb{R}$ given by $F(x_1, \dots, x_{n+1}) = x_1^2 + \cdots + x_{n+1}^2$ has rank 1 everywhere except at the origin. Consequently, near any point except the origin, there are coordinate charts which make F look like projection onto the first coordinate. So in those charts, a point inverse, say, $S^n = F^{-1}(1)$, looks like the points with first coordinate equal to some constant c. Permuting the coordinates and composing with a translation, we can make this looking like the last coordinate equals 0, i.e, we have S^n looking like the embedded submanifold it is, of codimension 1. We refer to point inverses as *level sets*, as people often do outside of point-set topology, especially in analytical cricles. This works in complete generality: if $F : M^n \to N^m$ is a smooth map of constant rank, then for every $c \in N$, the level set $F^{-1}(c)$ is a (closed: F is continuous and $\{c\} \subseteq N$ is closed) embedded submanifold of M of codimension k, hence of dimension $n - k$. The proof is exactly the same. A particular instance is when F is a submersion; then level sets are closed submanifolds of codimension m, hence dimension $n - m$. Note that insisting on constant rank k is a bit of overkill; what we really need is that F has constant rank k in a neighborhood of $F^{-1}(c)$, in order to apply the rank theorem. This specializes to the the submersion case to require only that F_* is a surjection at every point of the level set; the statement for nearby points is automatic. This prompts us to make a definition; a point $a \in M$ is a *regular point* for F if F_* is a surjection at a; otherwise, a is a *critical point.* A point $b \in N$ is a *regular value* for F if every $a \in F^{-1}(b)$ is a regular point; otherwise it is a *critical value.* Then the inverse image of every regular value is a closed submanifold of M.

In the particular case of functions $F : M \to \mathbb{R}$, we have a particularly succinct way of saying these things; $a \in M$ is a regular point of F iff $dF_a \neq 0$. (The text leaves this as an exercise; so will we.)

Examples to illustrate these concepts are all around; every closed orientable surface is the level set corresponding to a regular value of a smooth map from $\mathbb{R}^3$ to $\mathbb{R}$ (it just takes a little work to build it... One approach; using the standard Riemannian metric on $\mathbb{R}^3$, we can find the vector in $T\mathbb{R}^3$ normal to $T\Sigma$; these form a 1-dimensional bundle $N\Sigma$, which, since we can build a nowhere-zero (outward pointing) section, is trivial. $N\Sigma \cong \Sigma \times \mathbb{R}$ can be embedded in $\mathbb{R}^3$ with Σ as 0-section; the map which crushes each $\Sigma \times \{t\}$ to a point is a map to $\mathbb{R}$ with 0 as a regular value. [Send the stuff outside of $N(\Sigma)$ anywhere you want, away from 0.]). In general, write down a smooth map F from $\mathbb{R}^n$ to $\mathbb{R}^k$; nearly every point in $\mathbb{R}^k$ will turn out to be a regular value for F. (We will shortly make this statement precise.)

Induced smooth maps: Since the smooth structure on an embedded submanifold is "derived" from he ambient manifold M, certain typical operations work well. For example, if $F : M \to N$ is smooth and $S \subseteq M$ is an embedded submanifold, then $F|_S : S \to N$ is smooth. This is because this map is "really" the composition of the map F and the (smooth) inclusion map ι. On the other end, if $F : M \to N$ is smooth and $T^k \subseteq N^m$ is an embedded submanifold with $F(M) \to T$, then we can think of F as a map from M to T, and the map $F : M \to T$ is smooth; it is continuous, since T has the subspace topology, so $U \subseteq T$ is open implies that $U = V \cap T$ for some $V \subseteq N$ open, and then $F^{-1}(U) = F^{-1}(V)$, since if $F(x) \in V$ then $F(x) \in F(M) \cap V \subseteq T \cap V = U$, so $F^{-1}(U)$ is open in M. To see that $F : M \to T$ is, moreover, smooth, we note that, in slice coordinates for T, F looks locally like a smooth map whose last $m - k$ coordinates are 0. Lopping off those zeroes gives a local representation of F as a map into T, since the first k coordinates for the chart on N are the chart on T, and this map is still smooth. So as a map into T, F is smooth.

Vector and covector fields behave well, too. Covector fields are simplest; since $\iota : S \to M$ is smooth, any covector field ω on M pulls back to a covector field $\iota^*\omega$ on S. Our characterization of tangent vectors to S as tangent vectors to M which vanish on smooth functions zero on S carries over to vector fields as well; a vector field X on M is tangent to S if $X_a \in \iota_* T_a S$ for every $a \in S$, and this is true if, for every $a \in S$, $X_a f = 0$ for every $f \in C^\infty(M)$ with $f|_S \equiv 0$. In this case, $X_a = \iota_* Y_a$ for some $Y_a \in T_a S$, and since ι_* is injective at every point, Y_a is uniquely defined. The Y_a's together define a smooth vector field Y on S, which is therefore (by definition) ι-related to X. To see that Y is in fact smooth, we evaluate it on a collection of coordinate functions for charts covering S, specifically, slice charts. In slice charts, for $i \leq k$, $X(x^i) = \iota_* Y(x^i) = Y(x^i \circ \iota) = Y(x^i)$, where the first x^i is a coordinate function on M and the last is a coordinate function on S. Since $X(x^i)$ is smooth (as a function of n variables), so is $Y(x^i)$ (as a function of the first k). Since these functions allow us to describe Y in local coordinates $Y = \sum Y(x^i)\partial/\partial x^i$, Y is smooth.

Our next goal is really to formulate the notion of integration of smooth manifolds. To do so, we first need to build up the concepts necessary to let us describe what the objects are that we will be integrating. These objects are best described in the language of tensors.

Tensor products: On one level, tensor products are just another method for combining two vector spaces to build another vector space. Ultimately, we will be interested in the k-fold tensor product of the dual vector space V^*, as a mechanism for building the higher-dimensional analogues of covectors and covector fields.

Given vector spaces V, W, we first buld the free vector space on $V \times W$, that is, the vector space $\mathbb{R}(V \times W)$ with basis all pairs (v, w). Then we take the quotient by the linear subspace spanned by the following:

$(v_1 + v_2, w) - (v_1, w) - (v_2, w)$, for $v_1, v_2 \in V$, $w \in W$
$(v, w_1 + w_2) - (v, w_1) - (v, w_2)$, for $v \in V$, $w_1, w_2 \in W$
$c(v, w) - (cv, w)$, for $c \in \mathbb{R}$, $v \in V$, $w \in W$
$c(v, w) - (v, cw)$, for $c \in \mathbb{R}$, $v \in V$, $w \in W$

The quotient vector space is denoted $V \otimes W$, and the image of (v, w) is denoted $v \otimes w$. Given bases $\{v_i\}$ for V and $\{w_j\}$ for W, it is a quick matter of using the facts that

$(v_1 + v_2) \otimes w - v_1 \otimes w - v_2 \otimes w = 0$ in $V \otimes W$, etc., to show that every element $v \otimes w$ (and therefore, taking linear combinations, every element of $V \otimes W$) can be expressed as a linear combination of the $v_i \otimes w_j$; just write each of v and w in terms of their bases and expand out the sums. That is, after all, what the four conditions above are designed to let you do. A slightly more involved argument establishes that the elements $v_i \otimes w_j$ form a *basis* for $V \otimes W$, rather than just a spanning set. To see this,we first note that this method of constructing $V \otimes W$ makes immediate the fact that the tensor product has the universal property:

Given a vector space Z and a bilinear map $F : V \times W \to Z$ (that is, a map which is linear on each separate variable), there is a unique linear map $L : V \otimes W \to Z$ such that $F = L \circ \pi$, where $\pi : V \times W \to V \otimes W$ is the map sending (v, w) to $v \otimes w$. L is built by extending F to $\mathbb{R}(V \times W)$ by making the image of a linear combination the combination of their images, and noting that by bilinearity the extension is 0 on the relations we mod out by, and so descends to a well-defined map from $V \otimes W$ to Z. This map is uniquely determined by the fact that $L(v_i \otimes w_j) = F(v_i, w_j)$, and linearity is built-in by the fact that our extension to $\mathbb{R}(V \times W)$ was by linearity.

Applying this to the duals V^*, W^* of two vector spaces, we can build a bilinear map from $V^* \times W^*$ to the vector space $\mathcal{B}(V, W)$ of bilinear maps $F : V \times W \to \mathbb{R}$ by sending (φ, ϑ) to $(v, w) \mapsto \varphi(v)\vartheta(w)$. This map descends to a linear map $V^* \otimes W^* \to \mathcal{B}(V, W)$ sending $\varphi \otimes \vartheta$ to the map above. If $\{v_i\}, \{w_j\}$ form bases for V and W, then their dual bases are $\{v_i^*\}, \{w_j^*\}$, and $v_i^* \otimes w_j^*$ is sent to the map which is 1 on (v_i, w_j) and 0 on all other pairs. This shows that the images of the $v_i^* \otimes w_j^*$ are linearly independent in $\mathcal{B}(V, W)$; the coefficient of any $L(v_i^* \otimes w_j^*)$ in a linear combination can be extracted by evaluating the combination on (v_i, w_j). But this implies that the $v_i^* \otimes w_j^*$ are then linearly independent in $V^* \otimes W^*$, as well. But these also, by the argument above, span $V^* \otimes W^*$, so they form a basis. Applying this to $V^{**} \cong V$ and $W^{**} \cong W$ shows that $v_i \otimes w_j$ form a basis for $V \otimes W$, as well. Consequently, the dimension of the tensor product of two vector spaces is the product of their dimensions.

Iterating this process allows us to define the k-fold tensor product $V_1 \otimes \cdots \otimes V_n$ of vector spaces; technically, this requires showing that the order in which we pair things is unimportant, that is, $(V \otimes W) \otimes Z$ is naturally isomorphic to $V \otimes (W \otimes Z)$. Such an isomorphism can be constructed from the universal property, by starting, for $z \in Z$ with the bilinear map $v \times W \to V \otimes (W \otimes Z)$ given by $(v, w) \mapsto v \otimes (w \otimes z)$, which gives a linear map $V \otimes W \to V \otimes (W \otimes Z)$. These can be assembled into a bilinear map $(V \otimes W) \times Z \to V \otimes (W \otimes Z)$, giving a linear map $(V \otimes W) \otimes Z \to V \otimes (W \otimes Z)$. This map is readily shown to be surjective (its image contains a basis), so for dimension reasons it is an isomorphism. It is also exactly what you think it is; it sends $(v \otimes w) \otimes z$ to $v \otimes (w \otimes z)$. Alternatively, we could simply rebuild the theory without parentheses, defining multilinear maps $V_1 \times \cdots \times V_n \to Z$ and constructing $V_1 \otimes \cdots \otimes V_n$ to satisfy the corresponding universal property.

Our primary object of study will be the k-fold tensor product of the dual V^* of a vector space V (think: $V = TM_a$). Elements of this vector space will be called *covariant tensors of rank k* (for reasons of historical accident), and the space will be denoted $T^k(V)$. This tensor product has another life (as hinted at above), as the vector space of multilinear maps from the k-fold Cartesian product of V with it self, to $\mathbb{R}$. The isomorphism, as above, is defined by sending $v_1^* \otimes \cdots \otimes v_k^*$ to the map $(w_1, \ldots, w_k) \mapsto v_1^*(w_1) \cdots v_k^*(w_k)$, and extending linearly. This is the map that the (undeveloped here) universal property would construct out of the multilinear map sending $(v_1^*, \ldots, v_k^*)$ to the map above. It will sometimes be more convenient to build elements of $T^k(V)$ by building such multilinear maps, and then passing through this isomorphism. Though we won't encounter it much, the k-fold tensor product of V with itself is the vector space of *contravariant tensors of rank* k, and denoted $T_k(V)$. Finally, the tensor product $T^k(V) \otimes T_\ell(V)$ is the space of *mixed tensors of type* (k, ℓ), and is denoted $T^k_\ell(V)$.

Tensors on manifolds: We now have a new means of combining vectors spaces to create a new one. So, of course, we want to turn this into a way to build new bundles. In particular, taking the k-fold tensor product of T^*M_a at every $a \in M$, we can build a bundle which we will denote $T^k(TM)$, or $T^k(M)$ for short (note that in this new notation $T^1(M) = T^*M$), called the *bundle of covariant tensors of rank k*. The bundles of contravariant tensors, and mixed tensors, are built analogously. The bundle structure, for example on $T^k(M)$, can be described by explicitly building local trivializations, which we have seen amounts to building local frames; but given a chart $h = (x^1, \ldots, x^n)$ about $a \in M$, the covariant tensors $dx^{i_1} \otimes \cdots \otimes dx^{i_k}$ form a basis for $T^k(TM_a)$, and so give local frames about each point, and therefore local trivializations. A smooth structure can be imposed on $T^k(M)$ by declaring these trivializations to be an atlas; it is a tedious exercise to check that they are C^∞-related.

A section of the bundle $T^k(M)$ is called a covariant tensor field of rank k on M, or just a tensor field (when he rest is clear from context). The (vector) space of smooth section of $T^k(M)$ is denoted $\mathcal{T}^k(M)$. Analogous contravariant and mixed tensor fields can be similarly defined. As we saw with vector fields, a section $\sigma : M \to T^k(M)$ can be tested for smoothness by writing $\sigma(x)$ as a linear combination of our local frames, and insisting that the coeficients be smooth functions of x. Using the interpretation of a tensor field as a choice of multilinear function $TM_a \times \cdots \times TM_a \to \mathbb{R}$ at each point, we can also test for smoothness by feeding σ a k-tuple of vector fields $X_1, \ldots X_k$, and insisting that the function $x \mapsto \sigma_x(X_1(x), \ldots, X_k(x))$ is always smooth. (Since σ is multilinear, this can be checked by feeding it the coordinate vectors fields $\partial/\partial x^i$ in all possible combinations.) Given $\sigma \in \mathcal{T}^k(M)$ and $\tau \in \mathcal{T}^\ell(M)$, we can build their tensor product $\sigma \otimes \tau \in \mathcal{T}^{k+\ell}(M)$ by building the tensor product in each fiber (supressing the necessary parentheses).

Covariant tensors pull back under smooth maps: given $F : M \to N$ smooth, and $\sigma \in \mathcal{T}^k(N)$, we can define $F^*\sigma \in \mathcal{T}^k(M)$ by (thinking of it as a choice of multilinear function and using) the formula $(F^*\sigma)_a(X_a^1, \ldots X_a^k) = \sigma_{F(a)}(F_*X_a^1, \ldots F_*X_a^k)$. As with covector fields, this pullback operation is well-behaved, e.g., F^* is $\mathbb{R}$-linear (but over smooth functions we need to shift the domain: $F^*(f\sigma) = (f \circ F)\ F^*\sigma)$, $(F \circ G)^* = G^* \circ F^*$, and $F^*(\sigma \otimes \tau) = F^*\sigma \otimes F^*\tau$. These are all nearly immediate from the definition.

Covariant tensors (or rather, certain covariant tensors) will be the "correct" objects to integrate on manifolds and submanifolds. The reason for this is one of invariance. The basic idea on integration is going to be to play our standard trick; we do it by taking the thing to integrate (whatever it turns out to be) and multiply it by one, in the form of a partition of unity subordinate to a cover by charts. But integrating $\sum g_i(\text{blah})$ will be done by integrating each $g_i(\text{blah})$, and adding; and each $g_i(\text{blah})$ will be integrated by working in the range of the chart, that is, on an open subset of Euclidean space (where we think we know how to integrate). Proving that this make sense amounts to showing that the result is independent of choices, namely the cover by charts and the partition of unity. But by taking the product of the two partitions of unity, we have (blah) = $\sum_j G_j(\sum_i g_i(\text{blah}))$, which can be evaluated two ways; collecting i's and summing over the j's (using the charts for the j's), or vice versa. Independence, in the end, comes down to knowing that integrating $G_i g_j(\text{blah})$ over the chart from the i's gives the same answer as when integrating it over the chart from the j's. That is, we want the representation of (blah) in local coordinates to transform in such a way that integrating it gives the same answer over both charts. But a covariant tensor transforms "in the same way" as the coordinate charts (hence the unfortunate terminology), which is why, in the end, this will work....

We have, as it happens, already encountered tensor fields in this class. A Riemannian metric on M is a choice of bilinear map $TM_a \times TM_a \to \mathbb{R}$ (with other nice properties) at every point of M. But such a bilinear map now has a different life, as a covariant 2-tensor on TM_a. That is, a Riemannian metric is a certain kind of covariant 2-tensor field on M. In particular, it is a *symmetric* 2-tensor field. A covariant k-tensor σ over a vector space V is called symmetric if (in the language of multilinear maps) $\sigma(x_1, \dots, x_k) = \sigma(x_{\tau(1)}, \dots x_{\tau(k)})$ for any permutation τ. (Since the group of permutations is generated by transpositions, it is enough to check this for τ=transpositions.) In the language of a basis $v_{i_1}^* \otimes \cdots \otimes v_{i_k}^*$ of $T^k(V)$, we need the coefficient of $v_{i_1}^* \otimes \cdots \otimes v_{i_k}^*$ (in multi-language, the value of $\sigma(v_{i_1}, \dots, v_{i_k})$ to be the same as the coefficient of $v_{i_{\tau(1)}}^* \otimes \cdots \otimes v_{i_{\tau(k)}}^*$. A tensor field is symmetric if it is symmetric at every point. Since linear combinations of symmetric tensors are symmetric, the symmetric tensors form a vector subspace of $T^k(V)$, which we denote by $\Sigma^k(V)$. The argument above shows that the dimension of this subspace is equal to the number of orbits of the basis elements $v_{i_1}^* \otimes \cdots \otimes v_{i_k}^*$ under the action of the symmetric group S_k. For example, if dim $V = n$, then dim $\Sigma^2(V) = n(n+1)/2$, having basis $v_i^* \otimes v_i^*$ and $v_i^* \otimes v_j^* + v_j^* \otimes v_i^*$ for $i < j$. A similar count establishes that $\Sigma^3(V)$ has dimension $n + n(n-1) + n(n-1)(n-2)/6$ (counting those with all indices the same, exactly two the same, and all different).

A process of *symmetrization* gives a (projection) map Sym:$T^k(V) \to \Sigma^k(V)$, given by $\sigma \mapsto (1/k!) \sum_{\tau \in S_k} {}^\tau\sigma$, where ${}^\tau\sigma(x_1, \dots, x_k) = \sigma(x_{\tau(1)}, \dots x_{\tau(k)})$. A tensor is symmetric $\Leftrightarrow$ σ =Sym(σ) . Given two symmetric tensors σ and τ, their tensor product $\sigma \otimes \tau$ need not be symmetric (think of dx^1 and dx^2 (both symmetric) and $dx^1 \otimes dx^2$). But we can introduce the *symmetric product* ST of a symmetric k- and ℓ-tensor, giving the symmetric $(k+\ell)$-tensor defined (by way of multilinear maps) by $ST(x_1, \dots, x_{k+\ell}) = (1/(k+\ell)!) \sum_{\sigma \in S_{k+\ell}} S(x_{\sigma(1)}, \dots, x_{\sigma(k)}) T(x_{\sigma(k+1)}, \dots, x_{\sigma(k+\ell)})$. It is straightforward to verify that the symmetric product is commutative and linear in each argument.

Alternating tensors: But for the most part we don't want our tensors to be symmetric, we want the exact opposite. That is, we want anti-symmetric, or *alternating*, tensors, which satisfy ${}^\tau S = -S$ for every transposition of the inputs of S (here again, we are thinking of S as a multilinear function). It follows that ${}^{\tau}S = (\text{sgn}\tau)S$ for any permutation τ, where $\text{sgn}(\tau)$ is the sign of the permutation τ.

Why do we want to use alternating tensors? Because when we do integration in Euclidean space we create huge sums of the values of our function F at points in little boxes, times the *signed* volume of the little boxes. (This is why $\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$; we are really using the signed lengths of the intervals in our partition of $[a, b]$.) Alternating tensors, we shall see, are a way to capture the "signedness" of volume, as it appears in the formulation of multiple integrals. But before looking into <u>why</u> they work, let us look into <u>how</u> they work. The basic idea is the they behave like the determinant; thinking of an $n\times n$ matrix as an ordered collection of n column vectors $X_1, \ldots, X_n$, then $det(X_1, , \ldots, X_n)$ is an alternating, multilinear map from V^n to $\mathbb{R}$, and so defines an alternating covariant n-tensor on $\mathbb{R}^n$. The uphot to what follows is that, in a certain sense, this is the only covariant tensor around...

The collection of alternating covariant k-tensors for a sub-vector space of $T^k(V)$; the linear combination of alternating tensors is alternating, since we hit the combination with a permutation τ by hitting each term, and, by hypothesis, it amounts to just multiplying by $\text{sgn}(\tau)$, which we can factor out of the sum. We will denote the subspace of alternating k-tensors by $\Lambda^k(V) \subseteq T^k(V)$; $\Omega^k(V)$ is also common. Linear maps $L : T^k(V) \to T^k(W)$ also take alternating tensors to alternating tensors. A k-tensor S is alternating $\Leftrightarrow$ $S(x_1, \ldots, x_k) = 0$ whenever $x_i = x_j$ for some $i \neq j$. This is because the transposition of the two inputs introduces a change of sign without changing the value; the opposite implication comes from expanding out $0 = S(x_1, \ldots, x_i+x_j, \ldots, x_i+x_j, \ldots, x_k)$ using the multilinearity. Using a basis v_i^* for the underlying vector space, alternating requires that the coefficient of $v_{i_1}^*\otimes\cdots\otimes v_{i_k}^*$ in S determines the coeficient of $v_{i_{\tau(1)}}^*\otimes\cdots\otimes v_{i_{\tau(k)}}^*$ for any permutation τ; they are equal up to multiplication by $\text{sgn}\tau$. In particular, in any representation of an alternating tensor in the "standard basis", a term with a repeated basis element has coefficient 0. Consequently, the only terms that can have non-zero coefficient are those with all basis elements distinct, and there are as many elements in a basis for $\Lambda^k(V)$ as there are orbits of those basic tensors under the action of the symmetric group S_k. Consequently, the dimension of $\Lambda^k(V)$ (if V has dimension n) is n choose k; a basis consists of the signed sums over each orbit of a representative of each orbit, which we can choose, unambiguously, as $v_{i_1}^* \otimes \cdots \otimes v_{i_k}^*$ with the i_j strictly increasing. That is, the basis consists of elements $\sum_{\tau\in S_k}(\text{sgn}\tau)v_{i_{\tau(1)}}^* \otimes \cdots \otimes v_{i_{\tau(k)}}^*$ one for each fixed increasing sequence $i_1, \ldots, i_k$ from the sequence $1, \ldots, n$. The dimension of $\Lambda^k(V)$ is therefore n choose k. In particular, if $k > n$ the basis is empty; there are no (non-zero) alternating k-tensors when $k >\text{dim}(V)$. Significantly, when $k = n$, the space of alternating tensors has dimension 1; the basis vector is the alternating sum S of the tensors $v_{\tau(1)}^* \otimes \cdots \otimes v_{\tau(n)}^*$. This is, essentially, the determinant; if we write a collection of vectors w_i, $i = 1, \ldots, n$ as column vectors in terms of the v_i's, assembling them into a matrix then $S(w_1, \ldots w_n)$ is the determinant of that matrix. Similarly, the basis element given by the alternating sum of terms $v_{\tau(i_1)}^* \otimes \cdots \otimes v_{\tau(i_k)}^*$, summed over S_k, gives the alternating k-tensor T for which $T(w_1, \ldots w_k)$ is the determinant of the $k \times k$ "minor" built from the matrix of column vectors w_i by taking the rows $i_1, \ldots, i_k$ (in that order). So every alternating k-tensor is, essentially, a linear combination of determinants of $k \times k$ minors.

Just as with symmetric tensors, there is a projection map Alt: $T^k(V) \to \Lambda^k(V)$ (alternification?) given by $\mathrm{Alt}(S)(v_1, \ldots, v_k) = (1/k!) \sum_\sigma (\mathrm{sgn}\sigma) S(v_{\sigma(1)}, \ldots, v_{\sigma(k)})$. As before, $\mathrm{Alt}(S)$ is alternating, S is alternating $\Leftrightarrow$ $S = \mathrm{Alt}(S)$. Note, for example, that the basis for $\Lambda^k(V)$ consisted of the tensors $(k!)\mathrm{Alt}(v_{i_1}^* \otimes \cdots v_{i_k}^*)$ for $1 \leq i_1 < \cdots < i_k \leq n = \dim(V)$. And again, the tensor product of two alternating tensors typically isn't alternating, by we can just push it back into $\Lambda^{k+\ell}(V)$ to create a *wedge product* $\wedge : \Lambda^k(V) \otimes \Lambda^\ell(V) \to \Lambda^{k+\ell}(V)$ given by $S \wedge T = ((k+\ell)!/k!\ell!)\mathrm{Alt}(S \otimes T)$. The coefficient out front is to make this formula behave well when applied to a dual basis: if we write, for a multi-index $I = (i_1, \ldots, i_k)$, $v^I = \sum_{\tau \in S_k} (\mathrm{sgn}\tau) v_{i_{\tau(1)}}^* \otimes \cdots \otimes v_{i_{\tau(k)}}^*$, then $v^I \wedge v^J = v^{IJ}$, where IJ is the concatenation of I and J. (Note that $v^I = 0$ if the multi-index contains a repeated index, otherwise it is (up to sign) one of our standard basis elements.) The proof is tedious, and not very enlightening; it suffices to show that both sides represent the same multilinear function, for which it is enough to verify on all k-tuples of basis vectors that the v_i^* are dual to.

This wedge product is bilinear (since $\otimes$ and Alt are), associative (from $v^I \wedge (v^J \wedge v^K) = v^{IJK} = (v^I \wedge v^J) \wedge v^K$ and bilinearity), and is *anticommutative* : $S \wedge T = (-1)^{k\ell} T \wedge S$ (again, since this is true for our basis elements V^I). Further, $v_{i_1}^* \wedge \cdots \wedge v_{i_k}^* = v^I$, where I is the obvious multi-index. Finally, if $\omega^i \in V^*$ and $X_i \in V$, $i = 1, \ldots, k$, then
$(\omega^1 \wedge \cdots \wedge \omega^k)(x_1, \ldots, X_k) = \det(\omega^i(X_j))$.

Every alternating covariant k-tensor is therefore a linear combination of $v_{i_1}^* \wedge \cdots \wedge v_{i_k}^*$ for a dual basis to a basis v_i for V. (WOLOG, $i_1 < \cdots < i_k$.) Using the properties of the wedge product, it is a straightforward computation to determine a change of variables formula, to express an alternating tensor in terms of to different dual basis $w_1^*, \ldots, w_n^*$; since $v_i^* = \sum a_{ij} w_j^*$ for some matrix of coefficients (a_{ij}), $v_{i_1}^* \wedge \cdots \wedge v_{i_k}^* = (\sum a_{i_1 j} w_j^*) \wedge \cdot \wedge (\sum a_{i_k j} w_j^*) = \sum_{j_1, \ldots, j_k} a_{i_1 j_1} \cdots a_{i_k j_k} w_{j_1}^* \wedge \cdot \wedge w^* j_k$. The wedge product is 0 if there is a repeated j-index, and the sum can be partitioned into sums for each k-element subset of $\{1, \ldots, n\}$; writing each sum in terms of a single representative $w_{j_1}^* \wedge \cdot \wedge w^* j_k$ with $j_1 < \cdots < j_k$ introduces the sgn of the permutation τ taking the given ordering to this standard one, so $v_{i_1}^* \wedge \cdots \wedge v_{i_k}^* = \sum_{j_1 < \cdots < j_k} det(A_{j_1 \ldots j_k}) w_{j_1}^* \wedge \cdots \wedge w_{j_k}^*$, where $A_{j_1 \ldots j_k}$ is the obvious $k \times k$ minor. In particular, if $k = n = \dim V$, then $v_1^* \wedge \cdots \wedge v_n^* = \det(a_{ij}) w_1^* \wedge \cdots \wedge v_n^*$. This will be the fundamental fact which will allow us to integrate alternating tensors.

k-forms: An alternating k-tensor field, that is, a tensor field that is alternating at each point, is called a *k-form*. A smooth k-form can be expressed, in local coordinates $(x^1, \ldots, x^n)$, as sums $\sum_{i_1 < \cdots < i_k} a_{i_1 \ldots i_k}(x) dx^{i_1} \wedge \cdots dx^{i_k}$, where the $a_{i_1 \ldots i_k}(x)$ are smooth functions, since the tensors $dx^{i_1} \wedge \cdots dx^{i_k}$ form a basis at each point in the chart, and each is a smooth section. We can write out a change of coordinate chart formula for k-forms; this follows from the formula above applied at each point, using the fact that, in local coordinates, $dy^i = \sum_j (\partial y^i / \partial x^j) dx^j$. This is because $dx^j(\partial / \partial x^k) = \delta_{jk} =$, so both sides of this equation evaluate to the same thing on each $\partial / \partial x^k$. In particular, if $\dim M = n$, then on n-forms we have $dy^1 \wedge \cdots \wedge dy^n = \det(\partial y^i / \partial x^j) dx^1 \wedge \cdots \wedge dx^n$.

The bundle of alternating k-tensors on M is denoted $\Lambda^k(M)$; the collection of smooth sections of $\Lambda^k(M)$, the k-forms, is denoted $\mathcal{A}^k(M)$. Given a smooth map $F : M \to N$, there is a pullback map $F^* : \mathcal{A}^k(N) \to \mathcal{A}^k(M)$; this is just the same pullback map on k-tensor fields, the point is that it takes alternating tensors to alternating tensors. (Thinking in terms of multilinear maps, the pullback pushes the k-vector of the domain forward, without mixing, so a transposition on the M-side is evaluated as a transposition on the N-side, where we pick up a $-$ sign.) Pullbacks behave well w.r.t. wedge product: $F^*(\omega \wedge \eta) = F^*\omega \wedge F^*\eta$, and, locally, since $F^*(dx^i) = d(x^i \circ F)$, we have $F^*(f\ dx^{i_1} \wedge \cdots \wedge dx^{i_k} = (f \circ F) d(x^{i_1} \circ F) \wedge \cdots \wedge d(x^{i_k} \circ F)$.

Orientations again: We have described an orientation on a manifold M^n as a collection of charts covering M so that the transition functions are orientation-preserving, that is, the matrix of partial derivatives has positive determinant at (one, hence) every point. Given such a collection of charts (h_i, U_i), we can build a nonwhere-zero n-form ω on M by choosing a partition of unity g_i subordinate to the cover $\{U_i\}$, defining $\omega_i = g_i dx^1 \wedge \cdots \wedge dx^n$ on U_i and zero outside of it, and setting $\omega = \sum \omega_i$. This is nowhere-zero because about each point, in terms of a fixed chart $\omega = \sum g_i(x)\ \det(\partial x^j / \partial y^k) dy^1 \wedge \cdots \wedge dy^n$, each of the determinants have the same sign, and at least one of the $g_i(x)$ is non-zero. Conversely, a nowhere-zero n-form ω determines an orientation by choosing local orientations $(v_1, \ldots, v_n)$ at each point x so that $\omega(v_1, \ldots, v_n) > 0$. (Note that since $\omega_x \in \Lambda^n(T_x M)$ is non-zero, it must be non-zero on any basis (otherwise it is 0 on all of them.) The fact that ω applied to any other basis in the same orientation will also be positive follows quickly from the properties outlined above.

Consequently, an orientation can (and often is) thought of, and introduced, as a nowhere-zero n-form, sometimes known as an *orientation form*. Given an orientation form ω, any other n-form η can be uniquely expressed as $\eta = f\omega$ for some $f \in C^\infty(M)$; $f(x) = \eta(v_1, \ldots, v_n)/\omega(v_1, \ldots, v_n)$ for any basis for $T_x M$. Therefore, for an orientable manifold, $\Lambda^n(M) \cong C^\infty(M)$.

Integration: We now have all of the pieces in place to make sense out of integration on smooth manifolds. It applies to oriented manifolds M^n, that is, orientable manifolds equiped with a specific choice of orientation. On any such manifold we can integrate an n-form ω with compact support, that is, $\omega_x = 0$ outside of a compact set $K \subseteq M$. (For example, if M is itself compact, then all forms have compact support.) The idea is to follow the outline we have already hinted at; cover M by charts (h_i, U_i), choose a partition of unity $\{g_i\}$ subordinate to the cover, and integrate ω over M by taking the sum of the integrals of $g_i\omega$ over each U_i, by integrating $\omega_i = (h_i^{-1})^*(g_i\omega)$ over $h_i(U_i) \subseteq \mathbb{R}^n$. Note that since supp$(g_i\omega)$ =supp$(g_i)\cap$supp$(\omega) \subseteq U_i$ is compact, ω_i has compact support as well. Of course, we also need to understand how to integrate n-forms over subsets $h_i(U_i)$; but in this case we can write $\omega_i = f\ dx^1 \wedge \cdots \wedge dx^n$ for some function f (since $dx^1 \wedge \cdots \wedge dx^n$ is an orientation form on $\mathbb{R}^n$; here the x^i are coordinate functions of the identity chart), and we set $\int_{h(U_i} f\ dx^1 \wedge \cdots \wedge dx^n = \int_{h(U_i} f\ dx^1 \cdots dx^n$ (the usual integral in $\mathbb{R}^n$). We are not quite done, though; we haven't actually used the orientation on M! The point is that letting ourselves use <u>any</u> cover by charts is trouble; if take a chart (h, U) and build a new one by composing h with a reflection $(x^1, x^2, x^3, \ldots, x^n) \mapsto (x^2, x^1, x^3, \ldots, x^n)$, building a chart (k, U), then $(k^{-1})^*(\omega)_{k(x)} = -(h^{-1})^*(\omega)_{h(x)}$, so the integrals are negatives of one another. So we need the orientation to help us choose a cover by charts <u>compatible</u> with our choice of orientation, to carry out the construction above. Since we focus only on n-forms with compact support, we also may insist that finitely many of the charts cover that support, and all others are disjoint from it; so our sum above is really a finite sum (the integrals over the chart disjoint from the support are all 0).

With this modification, we can show that the definition of $\int_M \omega$ given above is independent of the choices involved, that is, the choice of cover by charts compatible with the orientation, and the choice of partition of unity subordinate to that cover. To see this, suppose (k_j, V_j) is another cover by compatible charts, and G_j is a partition of unity subordinate to it. We wish to show that

$$\sum_i \int_{h_i(U_i)} (h_i^{-1})^*(g_i\omega) = \sum_i \int_{k_j(V_j)} (k_j^{-1})^*(G_j\omega).$$

But $\omega = \sum_i g_i\omega = \sum_i \sum_j g_iG_j\omega$, and

$$\int_{h(U_i)} (h_i^{-1})^*(g_i\omega) = \int_{h(U_i)} (h_i^{-1})^*(\sum_j g_iG_j\omega) = \int_{h(U_i)} \sum_j (h_i^{-1})^*(g_iG_j\omega)$$
$$= \sum_j \int_{h(U_i)} (h_i^{-1})^*(g_iG_j\omega) = \sum_j \int_{h_i(U_i\cap V_j)} (h_i^{-1})^*(g_iG_j\omega),$$

since the sum is a finite sum, and $g_iG_j\omega$ has support contained in $U_i \cap V_j$. So $\int_M \omega = \sum_i \sum_j \int_{h_i(U_i\cap V_j)} (h_i^{-1})^*(g_iG_j\omega)$, from one point of view, and there is a similar formulation involving the k_j. So to show that both give the same numerical answer, it is enough to show that $\int_{h_i(U_i\cap V_j)} (h_i^{-1})^*(g_iG_j\omega) = \int_{k_j(U_i\cap V_j)} (k_j^{-1})^*(g_iG_j\omega)$.

But if we write $\eta = (h_i^{-1})^*(g_iG_j\omega) = f dx^1 \wedge \cdots \wedge dx^n$, then since $k_j^{-1} = h_i^{-1} \circ (h_i \circ k_j^{-1})$, $(k_j^{-1})^*(g_iG_j\omega) = (h_i \circ k_j^{-1})^*\eta$. And of course $h_i(U_i \cap V_j) = (h_i \circ k_j^{-1})(k_j(U_i \cap V_j))$, where $h_i \circ k_j^{-1}$ is an orientation-preserving diffeomorphism between open subsets of $\mathbb{R}^n$. So it is enough to show that for $U, V \subseteq \mathbb{R}^n$ open and $F : U \to V$ an orientation-preserving diffeomorphism, we have $\int_V \eta = \int_U F^*\eta$. These integrals are expressed in terms of n-forms; to write them as "ordinary" integrals, if $\eta = f dx^1 \wedge \cdots \wedge dx^n$, then since $F^*(dx^i) = d(x^i \circ F) = d(F^i) = \sum_j (\partial F^i/\partial x^j)\ dx^j$, $F^*\eta = (f \circ F)\det(\partial F^i/\partial x^j)dx^1 \wedge \cdots \wedge dx^n$, so what we want is that $\int_V f\ dx^1 \cdots dx^n = \int_U (f \circ F)\det(\partial F^i/\partial x^j)\ dx^1 \cdots dx^n$. But since F is orientation-preserving the determinant is positive, and since the change of variables formula for multiple integrals (you can look it up!) is $\int_V f\ dx^1 \cdots dx^n = \int_U (f \circ F)|\det(\partial F^i/\partial x^j)|\ dx^1 \cdots dx^n$, we are done.

So by employing a parition of unity, we can formulate the integral of any n-form with compact support, over an oriented smooth n-manifold M. Since any choice of orientation form gives us an isomorphism $\mathcal{A}^n(M) \cong C^\infty(M)$, we can think of this as allowing us to integrate <u>functions</u> with compact support, provided we have some kind of "canonical" orientation form ω on M, as $\int_M f = \int_M f\omega$. One way to achieve this is to have (an orientation and) a particular choice of Riemannian metric on M, $< \cdot, \cdot >_x$. In this case, there is a unique orientation form on M for which $\omega_x(v^1, \ldots, v^n) = 1$ for every orthonormal basis in the orientation at T_xM; start with any orientation form and rescale it at each point using a single orthonormal basis at each point. Since any other orthonormal basis has an orthogonal change of basis matrix, with determinant 1, $\omega_x(v^1, \ldots, v^n) = 1$ holds for every orthonormal basis. It is called the *Riemannian volume form*, and is sometimes denoted dV_g (although it is not "d" of anything, in the sense about to be introduced!). So on a Riemannian manifold we can, in fact, allow ourselves to integrate functions.

What about all of those other differential forms? A k-form can be integrated over an oriented k-manifold; so a k-form ω can be integrated over oriented k-submanifolds $S \subseteq M$ of M, by pulling them back by the inclusion map ι: $\int_S \omega = \int_S \iota^*(\omega)$. Relating this back to

M would require some sort of canonical way to induce an orientation on S from M. One submanifold for which we do know how to do this is the boundary ∂M of M; we previously built the induced orientatin on ∂M from M. so an orientation on M allows us to integrate $(n-1)$-forms on ∂M.

Exterior derivatives: One of the main results on the integral is Stokes' Theorem, which relates the integrals of n-forms on M integrals of $(n-1)$-forms on ∂M. To formulate it, we need to build a map $d : \mathcal{A}^{n-1}(M) \to \mathcal{A}^n(M)$, the *exterior differential*. We build it more generally, $d : \mathcal{A}^k(M) \to \mathcal{A}^{k+1}(M)$, in terms of local coordinates, as $d(f dx^{i_1} \wedge \cdots \wedge dx^{i_k}) = df \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} = \sum_i (\partial f/\partial x^i)\, dx^i \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}$, and extending linearly. The fun part is showing that this definition is independent of the coordinate chart chosen! This can be done by a laborious, and ultimately unenlightening, computation, but there is also a slightly underhanded way to do this; we show that it is the unique function $d : \mathcal{A}^k(U) \to \mathcal{A}^{k+1}(U)$ (where U is our coordinate chart) that satisfies certain properties that <u>don't</u> <u>mention</u> the coordinates. In the process, we will uncover some very useful formulas. The first of these is that for a 0-form (i.e., a function f), this definition of "d" agrees with the one that we have already given, i.e., $df = d(f)$. This first property is, we should note, independent of coordinates. Other basic properties are:

$d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$ (this is immediate);

$d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^{\deg \omega_1} \omega_1 \wedge d\omega_2$. From the second property above, it is enough to verify this for $\omega_1 = f\ dx^I$ and $\omega_2 = g\ dx^J$. Then $\omega_1 \wedge \omega_2 = (fg) dx^I \wedge dx^J = (fg) dx^{IJ}$, and so $d(\omega_1 \wedge \omega_2) = d(fg) \wedge dx^{IJ} = (df \cdot g + f \cdot dg) \wedge dx^{IJ} = (df \wedge dx^I) \wedge (g\ dx^J) + f\ dg \wedge dx^I \wedge dx^J = (df \wedge dx^I) \wedge (g\ dx^J) + (-1)^{|I|}(f\ dx^I) \wedge (dg \wedge dx^J) = d\omega_1 \wedge \omega_2 + (-1)^{\deg \omega_1} \omega_1 \wedge d\omega_2$, as desired.

$d(d\omega) = 0$. Again, it is enough to verify this for $\omega = f\ dx^I$. Then $d\omega = \sum_i (\partial f/\partial x^i) dx^i \wedge dx^I$, and $d(d\omega) = \sum_j \sum_i (\partial^2 f/\partial x^i \partial x^j) dx^j \wedge dx^i \wedge dx^I$. But since mixed partials commute, but $dx^i \wedge dx^j = -dx^i \wedge dx^i$, the terms all cancel in pairs to sum to 0.

Then we can show that there is only one function D satisfying these four properties. By the second, it is enough to verify this for forms $f\ dx^I$. By the third, $D(f\ dx^I) = Df \wedge dx^I + f\ D(dx^I) = df \wedge dx^I + f\ D(dx^I)$, while we <u>want</u> this to be equal to $df \wedge dx^I$, so it is enough to show that $D(dx^I) = 0$. But $dx^I = dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ for some set of indices, which (since each x^{i_j} is 'just' a smooth function) equals $Dx^{i_1} \wedge \cdots \wedge Dx^{i_k}$, so we wish to have $D(Dx^{i_1} \wedge \cdots \wedge Dx^{i_k}) = 0$. But $D(Dx^{i_1} \wedge \cdots \wedge Dx^{i_k}) = D(Dx^{i_1}) \wedge Dx^{i_2} \wedge \cdots \wedge Dx^{i_k} - Dx^{i_1} \wedge D(Dx^{i_2} \cdots \wedge Dx^{i_k}) = 0 - 0 = 0$ by the fourth property and an inductive argument. So $d = D$.

So for each chart U, we have a unique function d_U satisfying these properties. From this it follows that there is a unique function $d : \mathcal{A}^k(M) \to \mathcal{A}^{k+1}(M)$ satisfying these same properties; we define $d\omega$, at a point $p \in M$, as $d_U \omega$ for any chart U containing p. If we were to choose another chart V, then (since the definition of d_U at p requires only knowledge of the form in a (or any) neighborhood of p (to define the partial derivatives), that $d_U = d_{U \cap V} = d_V$, so the result is independent of choice of chart. It follows directly from the local definition (now that we know it is independent of it) that for $F : M \to N$ smooth, $F^*(d\omega) = d(F^*\omega)$; for $\omega = f dx^{i_1} \wedge \cdots \wedge dx^{i_k}$, both are equal to $d(f \circ F) \wedge d(x^{i_1} \circ F) \wedge \cdots \wedge d(x^{i_k} \circ F)$.

There is, in fact, a coordinate-independent definition of the exterior derivative. Given vectors at p, $v^1,\ldots,v^{k+1}$, pick locally-defined vector fields $X^1,\ldots,X^{k+1}$ equalling the v^i at p, and define $d\omega(v^1,\ldots,v^{k+1}) = d\omega(X^1,\ldots,X^{k+1}) = \sum_i X^i(\omega(X^1,\ldots,\widehat{X^i},\ldots,X^{k+1})) + \sum_{i<j}(-1)^{i+j}\omega([X^i,X^j],X^1,\ldots,\widehat{X^i},\ldots,\widehat{X^j},\ldots,X^{k+1})$, where $\widehat{X^i}$ means that X^i is omitted. The tough part with this approach is showing that the definition is independent of the extensions chosen (and showing that it defines a tensor!). We will not pursue this.

Stokes' Theorem: We now have the tools to state (and prove) Stokes' Theorem. It says: if ω is a compactly supported $(n-1)$-form on the oriented manifold M^n, then $\int_M d\omega = \int_{\partial M}\omega$. The proof, with a little care, is suprisingly brief. The point is that we can evaluate these integrals using any collection of compatible charts that we want, so we put a lot of effort into our choice of charts. The other point is that both sides of the equality are linear in ω (since the integral is "really" a sum of integrals over Euclidean space, where linearity holds). So we can focus on fairly "simple" forms, as well.

To build our collection of charts, the driving force is to make our ultimate integrals in Euclidean space as uncomplicated as possible. Each point $a \in M$ has a chart neighborhood (h_a, U_a) so that $h_a(U_a) \subseteq \mathbb{R}^n$ contains the closed cube $[0,1]^n$ with $h(a)$ in its interior (if $a \in$int(M)) or in the interior of the bottom face $[0,1]^{n-1}\times\{0\}$ (if $a \in \partial M$). Further, the interiors of the cubes cover int(M) and the interiors of the bottom faces cover ∂M (forming a compatible cover of ∂M by charts). Then choose a finite subcover of supp(ω) by these charts (or rather, the interiors of the inverse images of the half-open cubes $(0,1)^{n-1}\times[0,1)$, or $\times(0,1)$ if we are in the interior of M), (h_i, U_i), and a partition of unity $\{g_i\}$ subordinate to this cover (and the open set $M\backslash$supp(ω)). Note that the $g_i|\partial M$ form a partition of unity subordinate to the cover $h_i|_{\partial M}$ of our collection of charts for ∂M. To show $\int_M d\omega = \int_{\partial M}\omega$, what we show is that $\int_{U_i} d(g_i\omega) = \int_{U_i\cap\partial M} g_i\omega$ for every i. Stokes' Theorem will follow by summing each set of integrals separately, since $\sum_i\int_{U_i} d(g_i\omega) = \sum_i\int_M d(g_i\omega) = \sum_i\int_M dg_i\wedge\omega + g_i d\omega = \int_M d(\sum_i g_i)\wedge\omega + \int_M(\sum_i g_i)d\omega = \int_M d(1)\wedge\omega + \int_M d\omega = \int_M 0 + \int_M d\omega = \int_M d\omega$.

So now we can focus, essentially, on forms $g_i\omega = \eta$ supported on the interior (interpreted broadly) of $[0,1]^n$ in $\mathbb{R}^n$ or $\mathbb{H}^n$, since $\int_{U_i} d(\eta) = \int_{[0,1]^n}(h_i^{-1})^*(d\eta)$ and $\int_{U_i\cap\partial M}\eta = \int_{[0,1]^{n-1}\times\{0\}}(h_i^{-1})^*(\eta)$ if $U_i\cap\partial M \neq \emptyset$; it is (formally) equal to 0 if $U_i\cap\partial M = \emptyset$. But since $(h_i^{-1})^*(d\eta) = d((h_i^{-1})^*(\eta))$, we can pretend we are just working with "some" $(n-1)$-form, which we will go back to calling ω, on $\mathbb{R}^n$ or $\mathbb{H}^n$, depending on the circumstance. Stokes' Theorem then follows from the following two routine calculations:

If $\omega = f dx^1\wedge\cdots\widehat{dx^i}\cdots\wedge dx^n$ is an $(n-1)$-form with support contained in $(0,1)^n$, then $\int_{[0,1]^n} d\omega = 0$. This is because $d\omega = df\wedge dx^1\wedge\cdots\widehat{dx^i}\cdots\wedge dx^n = (\sum_j(\partial f/\partial x^j)dx^j)\wedge dx^1\wedge\cdots\widehat{dx^i}\cdots\wedge dx^n = (-1)^{i-1}(\partial f/\partial x^i)dx^1\wedge\cdots\wedge dx^n$, so $\int_{[0,1]^n} d\omega = \int_{[0,1]^n}(-1)^{i-1}(\partial f/\partial x^i)dx^1\cdots dx^n$. Integrating out the dx^i first gives us $f(x^1,\ldots,1,\ldots,x^n) - f(x^1,\ldots,0,\ldots,x^n)$, which since ω (i.e., f) is supported on the interior of the cube, is a difference of 0's, hence 0. So the remaining integrals are integrals of 0, so $\int_{[0,1]^n} d\omega = 0$.

If $\omega = f dx^1\cdots\widehat{dx^i}\cdots\wedge dx^n$ is an $(n-1)$-form with support contained in $(0,1)^{n-1}\times$

$[0,1)$, then $\int_{[0,1]^n} d\omega = \int_{[0,1]^{n-1}\times\{0\}} \omega$. In this case we need to be a trifle more careful to pay attention to how the orientation is chosen on ∂M; and, as it turns out, we did it wrong! We should have **pre**pended the outward-pointing vector field, not appended it. That is, translating into the language of forms, it is chosen so that $(-dx^n)\wedge\eta$ is the orientation form on M, since we prepend the outward-pointing normal, which in our standard upper half-space model, is *downward* pointing. That is, $(-1)^n dx^1\wedge\cdots\wedge dx^{n-1}$ is the induced orientation on $\partial\mathbb{H}^n$ from the standard orientation on $\mathbb{H}^n\subseteq\mathbb{R}^n$, since we have to push dx^n past $(n-1)$ dx^i's to properly rearrange terms. This sign will shortly be crucial!

Otherwise, it is really the exact same calculation as above, although there are two cases. When $i\neq n$, the calculation above goes through to show that $\int_{[0,1]^n} d\omega = 0$, while $\omega = f dx^1\cdots\widehat{dx^i}\wedge\cdots\wedge dx^n$, <u>as a form on ∂M</u>, is 0, since the function x^n is constant $x^n=0$ on ∂M, so $dx^n=0$. So $\int_{[0,1]^{n-1}\times\{0\}}\omega = 0$ in this case, as well. When $i=n$, however, when we integrate out dx^n above, we don't get 0, we get $-f(x^1,\dots,x^{n-1},0)$. So
$\int_{[0,1]^n} d\omega = (-1)^{n-1}\int_{[0,1]^{n-1}} -f(x^1,\dots,x^{n-1},0)dx^1\cdots dx^{n-1}$
$= (-1)^n\int_{[0,1]^{n-1}} f(x^1,\dots,x^{n-1},0)dx^1\cdots dx^{n-1}$. But as just saw (or rather, corrected), this is precisely the value of
$\int_{[0,1]^{n-1}\times\{0\}}\omega = \int_{[0,1]^{n-1}\times\{0\}} f dx^1\wedge\cdots\wedge dx^{n-1} = \int_{[0,1]^n\cap\partial\mathbb{H}^n} f dx^1\wedge\cdots\wedge dx^{n-1}$, when we take into account the correct choice of orientation on ∂M.

With these two calculations, Stokes' Theorem is proved.

Stokes' Theorem has many consequences; for example, applied (delicately) to Euclidean spaces, and to curves and surfaces in Euclidean spaces, it unifies (i.e., is really the same statement as) all of the great integration theorems of advanced calculus; the fundamantal theorem itself, as well as Green's, the Divergence, and (the classical) Stokes' Theorems. From our prespective, though, and to finish out the semester (?), we show how it gives some, essentially purely topological, results.

The first is that given a compact oriented manifold M^n with boundary, then there is no (smooth) retraction $r: M\to\partial M$, that is, no map with $r(x)=x$ for all $x\in\partial M$. For the proof, let ω be an orientation form for ∂M, giving the boundary orientation on ∂M. Then we should note that $\int_{\partial M}\omega > 0$, because in all charts compatible with the orientation $\int_U g_i\omega = \int_{h(U)} g_i(h^1)^*\omega$ is the integral of a non-negative function which is postive on an open set - $(h^1)^*\omega$ is an orientation form on $h(U)$ compatible with the standard one - and $\int_{\partial M}\omega$ is a sum of such integrals.

But then $r^*\omega = \eta$ is an $(n-1)$-form on M, with restriction to ∂M $\iota^*\eta = \omega$, since $r\circ\iota = \mathrm{Id}_{\partial M}$, and so by Stokes' Theorem, $\int_M d\eta = \int_{\partial M}\eta = \int_{\partial M}\omega > 0$. But! $d\eta = d(r^*\omega) = r^*(d\omega) = r^*(0) = 0$, since $d\omega\in\mathcal{A}^n(\partial M) = 0$, since ∂M <u>has</u> no non-trivial n-forms, being $(n-1)$-dimesional. So $\int_M d\eta = \int_M 0 = 0$, a contradiction, so no such retraction exists. (Note: This result is true for topological manifolds, as well, although this proof, of course, does not work in that more general case.)

Another result that we can prove using these techniques is the "Hairy Sphere Theorem": if n is even, then S^n (with the standard smooth structure, although that is unimportant) does not admit a nowhere-zero tangent vector field. (A straighforward calculation (in one of your homework problems!) shows that odd-dimesional spheres do have such fields.) The proof is a minor tour-de-force of the concepts that we have introduced, so it seems like a fitting ending point for the class!

So, suppose we have such a vector field X. Since "nowhere-zero" in invariant under small deformations, we may assume that X is smooth (choose smooth deformations in local coordinates, and patch together using a partition of unity). First, we think of S^n as the embedded submanifold $S = \{(x_1, \ldots, x_{n+1}) : x_1^2 + \cdots x_{n+1}^2 = 1\}$. Using the induced Riemannian metric from $\mathbb{R}^{n+1}$, we can divide $X(a)$ by its length, to obtain a smooth unit-length tangent vector field on S. With this embedding, an element $v \in T_a\mathbb{R}^{n+1}$ is tangent to S (i.e, is in T_aS) if $v \perp a$, thinking of a as a vector based at the origin. So a unit-length tangent vector field is really a smooth map $F : S \to S$ with $F(x) \perp x$ for all $x \in S$. Using such a map, we can smothly deform the identity map $I : S \to S$ to the antipodal map $A : S \to S$ given by $A(x) = -x$, through maps from S to S, by "pushing" x around S in the direction of $F(x)$. Symbolically, this map is given by $H(x,t) = \cos(\pi t)x + \sin(\pi t)F(x)$. (Taking the inner product of this formula with itself verifies that it always lies in S, since $< x, F(x) = 0.$).

S is orientable; it is the boundary of the unit ball in $\mathbb{R}^{n+1}$, which is orientable. Let ω be an orientation form on S. In general, given an oriented manifold (M, ω) and a diffeomorphism $G : M \to M$, G induces an isomorphsm of forms, and so $G^*\omega$ is another orientation form on M. We say that G is *orientation-preserving* if ω and $G^*\omega$ induce the same orientation on M, it is *orientation-reversing*, otherwise. Note that if G is orientation reversing, then $\int_M G^*\omega = -\int_M \omega$, where both integrals are computed using the orientation ω on M. In particular, $\int_M G^*\omega < 0$, while $\int_M \omega > 0$. Now the fundamental point to this line of reasoning is that if n is even, then the antipodal map $A : S \to S$ is orientation-reversing. Probably the cheapest way to see this is to note that A is the restriction to S of the composition of the $n+1$ reflections in $\mathbb{R}^{n+1}$ through the coordinate hyperplanes. This composition takes the standard orientation form ω_0 of $\mathbb{R}^{n+1}$ to $(-1)^{n+1}\omega_0$, so when n is even it reverses orientation. Since it preserve the outward-pointng direction on S, however, it also reverses orientation on S.

Given the orientation form ω on S, H induces an n-form $H^*\omega$ on $M = S \times [0,1]$. (Note that H is indeed a smooth map.) Call the 0- and 1-components of $\partial(S \times [0,1]$ S_0 and S_1. Now note that, in the induced boundary orientations that the S_i get from $S \times [0,1]$, the map $S_0 \to S_1$ given by $(x,0) \mapsto (x,1)$ is orientation reversing, because the outward-pointing normals point in opposite directions! The resulting change in signs when comparing integrals is important. Now we use Stokes' Theorem! $\int_M d(H^*\omega) = \int_{\partial M} H^*\omega$

Some facts, mostly difficult to prove, which our new vocabulary allows us to express:

(*) 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* $\mathbb{R}^5$.

(*) Tangent bundles can be constructed for topological manifolds; the trick is to come up with a characterization of the tangent bundle of a smooth manifold which doesn't mention derivatives! Topological tangent bundles can then be used to attack questions such as the existence of smooth structures; it amounts to being able to put a “linear” structure on the topological tangent bundle.

(*) The *span* of a vector bundle is the maximum number of linearly independent vector fields the the bundle can have. The span of a smooth manifold is the span of its tangent bundle. A manifold M has span $\geq 1 \Leftrightarrow$ its Euler characteristic $\chi(M) = 0$. Higher spans can be detected by increasingly sophisticated tools from algebraic topology.

(*) Every orientable 3-manifold M^3 is parallelizable; TM^3 is fiber-preserving diffeomorphic to $M^3 \times \mathbb{R}^3$ (Stiefel, 1936). Among spheres, only S^1, S^3 and S^7 are parallelizable (Adams 1962, Le Lionnais 1983).

(*) Every vector bundle over a contractible space X (i.e., there is a continuous map $H : X \times I \to X$ which restricted to $X \times \{0\}$ is the identity and to $X \times \{1\}$ is constant) is trivial. The usual proof consists of showing that the pullback bundle induced from a constant map is trivial, and the pullbacks induced by homotopic maps are isomorphic.

(*) Every vector bundle over S^3 is trivial (apparently, because $\pi_2(GL(3, \mathbb{R})) = 0$?).

(*) $\mathbb{R}P^{31}$ can be embedded in $\mathbb{R}^{54}$, immersed in $\mathbb{R}^{53}$, and cannot immerse in $\mathbb{R}^{52}$. It is unknown if it can embed in $\mathbb{R}^{53}$. (Source: http://www.lehigh.edu/∼dmd1/immtable)