Differential Topology **Spring 2023**

Foundations of Manifolds

PROFESSOR DANIEL ALLCOCK ABHISHEK SHIVKUMAR

Manifolds and Maps

The basic object in differential topology is a smooth manifold, as opposed to C^k, topological, real-analytic, complex analytic (etc.) manifolds. We'll start with a Calc III-esque study of surfaces in space, curves in $\mathbb{R}^{2+}$, before moving onto abstract manifolds without respect to any specific embedding. Viewing manifolds as embedded in $\mathbb{R}^n$ eases some psychological difficulties but many manifolds do not come with natural embeddings into $\mathbb{R}^n$. The natural first example of such a manifold is $\mathbb{RP}^2 = S^2/\{\pm 1\}$.

This one is not a real definition yet. Need to unpack terms; also Hausdorff-ness, second countability are missing.

Definition 1.1.1: Manifolds

Let M be a topological space with an open cover U_α and homeomorphisms $\varphi_\alpha : U_\alpha \to V_\alpha \subseteq \mathbb{R}^n$. When U_α and U_β overlap, we require that

$$\varphi_\beta(U_\alpha \cap U_\beta) \xrightarrow{\varphi_\alpha \circ \varphi_\beta^{-1}} \varphi_\alpha(U_\alpha \cap U_\beta)$$

is a diffeomorphism, that is, a bijective map between two subsets of $\mathbb{R}^n$ that is C^∞ with C^∞ inverse.

The symbolic complexity of this definition (in contrast to its intuitive simplicity) is essentially tied to the fact that we have to bootstrap the definition of a diffeomorphism from the one place where we know what it means: open subsets of $\mathbb{R}^n$.

Example 1.1.2

The standard example of a map that is bijective and smooth but not a diffeomorphism is $f(x) = x^3$ from $\mathbb{R}$ to itself, since $f^{-1}(y) = \sqrt[3]{y}$ is not even differentiable at 0 (so f is not even a C^1 homeomorphism).

C^∞ functions require some more discussion. For functions from $\mathbb{R}$ to itself, the standard definition suffices, e.g, f is differentiable, f' is differentiable, etc. For $n \geq 2$, C^∞ requires more than just existence of partial derivatives to all orders.

I think all that needs to be done here is add "and are continuous" to every "partial derivatives exist" to make the definition work out. The exposition here gets a little confusing.

Example 1.1.3

Consider

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

One can check that f has partial derivatives of all orders because it vanishes identically along the x and y axes. However, f is not even continuous. Along the line $y = mx$, $f(x,y) = \frac{mx^2}{(1+m^2)x^2} = \frac{m}{1+m^2}$, so along each line through the origin, f takes a different, constant value (and therefore the limit at 0 is undefined). So the existence of partials of all orders is an artifact of the choice of coordinate system with which we described f.

Definition 1.1.4: Differentiability

Let $f : \mathbb{R}^m \to \mathbb{R}^n$, we say that f is differentiable at $x \in \mathbb{R}^m$ if there exists a linear function $\lambda : \mathbb{R}^m \to \mathbb{R}^n$ such that

$$\frac{f(x+\Delta x) - (f(x) + \lambda(\Delta x))}{\|\Delta x\|} \to 0$$

as $\|\Delta x\| \to 0$. If such a λ exists, then it is unique, and it is called the derivative of f at x.

Definition 1.1.5: C^k

f is differentiable on an open subset U of $\mathbb{R}^n$ if f is differentiable at each $x \in U$. Consider the map $(x \mapsto \lambda_x) : U \to \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ from points of U to linear maps (essentially matrices); we say f is C^1 if this assignment is continuous. Since $\mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n) \cong \mathbb{R}^{mn}$, we can now ask if the assignment $(x \mapsto \lambda_x)$ is differentiable, in which case f is C^2, etc.

All the following definition is doing is using the natural scaling operation on vector spaces to express the local linearity condition somewhat more cleanly, and without as much formal baggage.

Definition 1.1.6

Suppose $f : V \to W$ is a function between real vector spaces, and suppose that (perhaps after translation) $f(0) = 0$. Then f is differentiable at 0 if there exists a linear map $\lambda : V \to W$ such that for any bounded neighborhood A of 0 in V and any neighborhood B of 0 in W, for all $\epsilon > 0$, there exists $t_0 > 0$ such that $f(tv) - \lambda(tv) \in t(\epsilon B)$ for all $0 < t < t_0$, all $v \in A$.

Fixing A, B as above, and $g : V \to W$, we say that g vanishes faster than linearly at 0 if for all $\epsilon > 0$, there exists $t_0 > 0$ such that $g(tA) \subseteq t(\epsilon B)$ for all $0 < t < t_0$. Then f is differentiable with derivative λ when $f - \lambda$ vanishes faster than linearly at 0.

Lemma 1.1.7

If λ as above exists, it is unique, and is called the derivative of f at 0.

Notation for the derivative varies: Df, df, Tf all appear in the literature.

Theorem 1.1.8: Chain Rule

If we have a sequence $V \xrightarrow{f} W \xrightarrow{g} X$ with f and g differentiable at $v_0 \in V$ and $f(v_0) \in W$ resp., then so is their composition at v_0, with derivative $D_{f(v_0)}g \circ D_{v_0}f$.

Corollary 1.1.9

If f, g as above are C^k on some open set in V, and g is C^k on the image of that open set (under f), then $g \circ f$ is C^k on it as well.

One can show this by induction on k and the chain rule.

Theorem 1.1.10: Inverse Function Theorem

If $f : V \to W$ is differentiable at v_0, and $D_{v_0}f : V \to W$ is an isomorphism, then f is bijective from some neighborhood of v_0 in V to some neighborhood of $f(v_0)$ in W, and f^{-1} is differentiable as well. Since $f^{-1} \circ f$ is the identity, the chain rule tells us that $D_{f(v_0)}f^{-1} = (D_{v_0}f)^{-1}$, and thus f is actually a diffeomorphism on some neighborhood of v_0.

It is traditional to skip this proof, although apparently there's a very clean proof (or at least clean exposition) by Terry Tao on MO.

We are now ready to formally define manifolds (again?).

Definition 1.1.11: Smooth Manifolds

A smooth manifold M is a Hausdorff topological space equipped with charts $(U_\alpha, \varphi_\alpha : U_\alpha \xrightarrow{\sim} V_\alpha)$ where V_α is an open set in a real vector space, such that the U_α cover M and the charts agree on overlaps, i.e,

$$\varphi_\beta(U_\alpha \cap U_\beta) \xrightarrow{\varphi_\alpha \circ \varphi_\beta^{-1}} \varphi_\alpha(U_\alpha \cap U_\beta)$$

is C^∞ with everywhere invertible derivative. We also require that each component of M has a countable basis for its topology (e.g, is second countable).

There is now a problem, that our manifolds depend explicitly on given charts, so there are many different (redundant) representatives of what we would think of as the same manifold. We will resolve this by insisting that all charts compatible with given charts are included in our set of charts, e.g, by taking a maximal atlas.

Found it mildly confusing that we didn't say that the φ_α themselves were C^∞, just their compositions on overlaps, but this is because, as above, we have to bootstrap what it means to be C^∞ from $\mathbb{R}^m \to \mathbb{R}^n$ functions, which is what $\varphi_\beta^{-1} \circ \varphi_\alpha$ is. Saying that φ_α itself is C^∞ doesn't actually mean anything at this stage.

Many authors require that M itself is second countable, which only rules out the case where M has uncountably many components. Some natural (for some definition of natural) manifolds do have uncountably many components: consider the 2-torus foliated by lines of irrational slope. Every such line is dense in T^2. One can define a topology on T^2 (the "leaf topology") that makes each leaf into a copy of $\mathbb{R}$ and all leaves separate components.

As an example for why Hausdorffness should be imposed, consider the "line with two origins," the topological space obtained by gluing two copies of $\mathbb{R}$ along $\mathbb{R} \setminus \{0\}$, which is not Hausdorff since open neighborhoods of both origins will always intersect.

The two technical criteria, Hausdorffness and that every component has a countable basis, can be expressed together as paracompactness.

As far as I know, paracompactness does not usually include Hausdorffness.

To actually specify a manifold M, you choose an underlying space and a few charts satisfying the overlap conditions, and then add in all possible charts that are compatible with the given charts. One must verify that any two charts obtained this way are compatible with each other, e.g, that "the" maximal atlas is well-defined, i.e unique.

Here Daniel introduces the following slogan, which we should not regard as actually true: "Any concept you can define for open subsets in real vector spaces that is invariant under diffeomorphisms makes sense for manifolds as well."

However, it is quite rare to actually define a manifold with charts given the huge amount of data to track. When we need to actually use local charts, we will generally just choose them on the fly rather than having them all defined at the outset. Choosing local charts around a point p for a manifold M amounts to a choice of chart around p, and functions $x_1, \cdots, x_n : M \to \mathbb{R}$ which, when taken together as a tuple, give a diffeomorphism from the chart to $\mathbb{R}^n$.

For example, a manifold M with charts $(U_\alpha, \varphi_\alpha)$, $f : M \to \mathbb{R}$ is smooth if for all $x \in M$, there exists a neighborhood U_α of x in the charts such that $f \circ \varphi_\alpha^{-1}$ is smooth as a function from some open neighborhood in $\mathbb{R}^n$ to $\mathbb{R}$.

Skipping some examples here of explicit definitions by charts, open subsets of $\mathbb{R}^n$, spheres via stereographic projection from the poles, etc.

Definition 1.1.12: Smooth Functions

A function $f : M \to \mathbb{R}$ is smooth if for all $p \in M$, there exists a chart $(U_\alpha, \varphi_\alpha)$ containing p s.t for all $U_\beta \ni p$,

$$f \circ \varphi_\beta^{-1} = (f \circ \varphi_\alpha^{-1}) \circ (\varphi_\alpha \circ \varphi_\beta^{-1})$$

where $f \circ \varphi_\alpha^{-1}$ is smooth by assumption, as is $\varphi_\alpha \circ \varphi_\beta^{-1}$ by the compatibility criterion for charts.

More examples of explicit charts that I'm not backfilling since I missed the lecture anyway. Summary is that you might need actual charts if you're a general relativist.

All this is to say that $\mathbb{R}$-valued functions on manifolds "make sense," despite the oppressive verbosity that God has required to make formal "makes sense." We play a similar game to define smooth functions $f : M \to N$ between manifolds; in particular, to sketch the main idea, given a chart $(U_\alpha, \varphi_\alpha)$ about $p \in M$, and a chart (V_β, ψ_β) about $f(p) \in N$, the function we want to consider is $\psi_\beta \circ f \circ \varphi_\alpha^{-1} : \mathbb{R}^m \to \mathbb{R}^n$, from which we can bootstrap smoothness as above from the established definition on $\mathbb{R}^n$.

Submersions

Theorem 1.2.1: Implicit Function Theorem

Suppose $f : M \to N$ is differentiable at $x \in M$. Then if the linear map $T_x M \to T_{f(x)} N$ is surjective (i.e if f is a *submersion* at x), then $f^{-1}(f(x))$ is a manifold near x. Moreover, we can choose local coordinates x_i around x and $f(x)$ such that f is the canonical surjection: $f(x_1, \cdots, x_m) = (x_1, \cdots, x_n, 0, \cdots, 0)$ with $m \geq n$.

The phrasing of the implicit function theorem in the lecture is a little confusing and doesn't fully match the statement in Guillemin-Pollack (where it is called the local submersion theorem), but hopefully the general idea is correct.

We can derive this result from the earlier Theorem 1.1.10 (the inverse function theorem) by doing some linear algebra to massage submersions (surjective differentials) and immersions (injective differentials) into local bijections.

First, let us consider the case of submersions, and suppose that $D_u f$ is surjective for $f : U \to V$. Then the complement of $\ker D_u f$ maps isomorphically onto $T_{f(u)} V$. Write $T_u U = \ker D_u f \oplus W$, where by moderate abuse of notation we can write $W = T_{f(u)} V$ and that f is the identity on W.

Is there anything funky going on with choosing a complement? Implicitly we're picking an inner product.

From this map we can construct a new map $c : U \to T_u U$ (perhaps after shrinking U to a coordinate chart where the above splitting of the tangent space can be extended) given by $(k \in K, v \in V) \mapsto k + f(v)$; clearly c is an isomorphism, so by the inverse function theorem, c is a diffeomorphism near u.

This whole bit is reconstructed from other people's notes and is pretty incomprehensible to me; the analogous statement below for immersions is more comprehensible. Also omitted an example showing that $\mathrm{SL}_n \mathbb{R}$ is a smooth manifold using the implicit function theorem applied to det.

We will apply the implicit function theorem to prove $O(n)$ (the group of orthogonal $n \times n$ matrices) is a smooth manifold. Recall that if M is a symmetric $n \times n$ matrix, then it represents a symmetric bilinear form $(x, y) \mapsto x^T M y$. $\mathrm{GL}_n \mathbb{R}$ acts on $\mathbb{R}^n$ by left multiplication, so $\mathrm{GL}_n \mathbb{R}$ acts on the set S of symmetric bilinear forms by $(M \cdot g)(x, y) := M(gx, gy)$ which, in matrices, is the map $(M, g) \mapsto g^T M g$ for g symmetric.

The orthogonal group of M are simply the elements of GL_n preserving the inner product. In matrices, $O(M) = \{g \in \mathrm{GL}_n \mathbb{R} : g^T M G = M\}$. in particular, there exists a function $\mathrm{GL}_n \mathbb{R} \to S$ given by $g \mapsto g^T M g$, and $O(M)$ is the preimage of $\{M\}$ under this map. Since $\mathrm{GL}_n \mathbb{R} \to S$ is smooth, we expect the preimage of M (which is $O(M)$) to be a manifold, which will hold if the derivative of the map is surjective at every point of $O(M)$. Let's compute the derivative at $g = I$. Let ϵ be an $n \times n$ matrix very close (in the standard norm) to zero. Then the image of $(I + \epsilon)$ in S is

$$(I + \epsilon)^T M (I + \epsilon) = M + \epsilon^T M + M\epsilon + \epsilon^T M \epsilon$$

Setting the quadratically vanishing term on the right to zero, so $\epsilon^T M + M\epsilon$ is the derivative at I. Now, taking M to be the identity (so that $O(M) = O(n)$), the derivative is simply $\epsilon \mapsto \epsilon^T + \epsilon$; we want to show that this surjects to S, i.e, if it is true that every symmetric matrix can be written in this form. Let P be a symmetric matrix, and note that $P = \frac{1}{2}P^T + \frac{1}{2}P$, from which surjectivity follows.

Thus we can conclude that the orbit map $\mathrm{GL}_n \mathbb{R} \to S$ of the standard inner product $I \in S$ has surjective derivative $M_n\mathbb{R} \to S$ at the identity in $\mathrm{GL}_n \mathbb{R}$, so $O(n)$ is a manifold near $I \in \mathrm{GL}_n \mathbb{R}$. Smoothness everywhere in $O(n)$ follows by the homogeneity trick; if $g \in O(n)$, then multiplication by g^{-1} identifies a neighborhood of $g \in \mathrm{GL}_n \mathbb{R}$ with a neighborhood of I in $\mathrm{GL}_n \mathbb{R}$ and preserves $O(n)$

This is a standard proof pattern for Lie groups; use the fact that the group multiplication operation is smooth to translate open sets around.

Immersions

Recall that $f : M \to N$ has derivative injective at $x \in M$, then f is called an immersion at x. Any embedding (e.g $\mathbb{R}^n \hookrightarrow \mathbb{R}^{m+n}$) is an immersion.

Proposition 1.3.1

Every immersion is locally equivalent to the immersion $\mathbb{R}^m \hookrightarrow \mathbb{R}^{n \geq m}$.

By abuse of notation, we shrink M and N to small coordinate charts without changing their names.

PROOF : Let $f : M \to N$ be an immersion. Choosing charts, may suppose M is a neighborhood of 0 in $\mathbb{R}^m$, $x = 0$, N is a neighborhood of 0 in $\mathbb{R}^n$, $f(0) = 0$. We assume that $D_0 f : \mathbb{R}^m \to \mathbb{R}^n$ is injective; by following f by an element of $\mathrm{GL}_n \mathbb{R}$, may suppose that $D_0 f$ is in fact just the inclusion $\mathbb{R}^m \hookrightarrow \mathbb{R}^n$. Now we use the inverse function theorem: consider $M \times \mathbb{R}^{n-m} \xrightarrow{F} \mathbb{R}^n$ with

$$F(p, x_{m+1}, \cdots, x_n) = f(p) + (0, \cdots, 0, x_{m+1}, \cdots, x_n)$$

Now $D_0 F = D_0 f + \mathrm{id}_{\mathbb{R}^{n-m}}$, which is surjective. Thus F must be a diffeomorphism in a small neighborhood of 0, and identifying a neighborhood of $f(x) \in N$ with $\mathbb{R}^n$ via F, you get that $f : M \to N$ is given in coordinates by

$$f(p) = (p, 0, \cdots, 0) \in \mathbb{R}^m \times \mathbb{R}^{n-m}$$ ■

We can say a lot about submersions and immersions, but for $f : M \to N$ neither a submersion nor immersion, the local structure can be quite a bit more complicated.

As an aside, he says that singularity theory is the branch of math devoted to finding "normal forms" of degenerate functions, but that we will usually try to stick with nicer functions. He gave a few examples I couldn't really describe since they were mostly drawings. One of them showed a map from $\mathbb{R}^2$ to a surface in $\mathbb{R}^3$ where there's a fold in the sheet, and he draws a curve along that fold that when projected down to $\mathbb{R}^2$ is cuspidal, e.g, something rhyming with a blowup. This is called the elementary catastrophe, as in catastrophe theory (which, fun fact, Dalí was a fan of).

Example 1.3.2

Let $f(x_1, \cdots, x_n) = \sum_i x_i^2$ which "looks" like a paraboloid. The derivative at 0 is the zero map; however, for $g(x_1, \cdots, x_n) = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2$, the derivative at the origin is again the zero map, even though the two functions locally look nothing alike (f is a local minimum at 0, g is some kind of complicated saddle).

Tangent Spaces

Recall that a manifold M is covered by charts $(U_\alpha, \varphi_\alpha)$ with codomain V_α an open set in a vector space, with C^∞ (hence locally diffeomorphic, since $\varphi_\alpha \circ \varphi_\beta^{-1}$ and $\varphi_\beta \circ \varphi_\alpha^{-1}$ are inverse diffeomorphisms) transition maps. So we can transfer any concept invariant under diffeomorphisms from vector spaces to manifolds. As we've already seen, one such concept is the notion of $f : M \to \mathbb{R}$ being smooth, which is bootstrapped from charts, since if f is smooth at a point in one chart, it is smooth in all charts containing that point since the transition maps are diffeomorphisms. The same process for defining smoothness of functions $f : M \to N$ also makes sense, by choosing charts around $m \in M$ and $f(m) \in N$ and inspecting the only composition that gives a function between (an open subset of) $\mathbb{R}^m$ and $\mathbb{R}^n$.

Lots of nice pictures here that I unfortunately cannot draw. They're basically the natural pictures one would draw, with commutative diagrams of disks as open sets.

We will similarly define the tangent space by bootstrapping. The tangent space at $x \in M$ is a vector space T_xM attached to the point x. If $f : M \to N$, then the derivative $df : T_xM \to T_{f(y)}N$ has a natural pushforward action on tangent spaces. We have seen already that if df_x is surjective (i.e, if f is a submersion at x), then there are coordinates for which f (near x) is the "canonical" projection, and similarly when df_x is injective (f immersive at x) there are coordinates where f is the "canonical" inclusion of a linear space.

We already know what the tangent space T_xM is for M an open subset of a vector space, namely V itself, so, as above, given a chart $(U_\alpha, \varphi_\alpha)$ on M with $x \in U_\alpha$, we can just define $T_xM := T_{\varphi_\alpha(x)}V_\alpha$. Now we just need to check that this definition doesn't depend on the choice of chart.

Here Daniel defines a tangent vector as the collection of all possible representatives in chart tangent spaces, and runs the same argument to show they are consistent. This seems like a largely philosophical difference to me, and I think I prefer to think in terms of "Pick a chart; your choice didn't matter." Maybe I'm wrong and there's really some essential difference I'm missing here.

Note that for $x \in U_\alpha \cap U_\beta$, the diffeomorphism $\varphi_\beta \circ \varphi_\alpha^{-1} : V_\alpha \to V_\beta$ gives us a canonical isomorphism $d(\varphi_\beta \circ \varphi_\alpha^{-1})_{\varphi(x)}$ of tangent spaces, whereby we may use the compatibility conditions on transition maps to pass between different choices of representatives for a given tangent vector in different chart tangent spaces.

Now suppose $f : M \to M'$. We want $df_x : T_xM \to T_{f(x)}M'$ to be defined similarly, so letting $x \in U_\alpha \cap U_\beta \subseteq M$, $f(x) \in U'_\gamma \cap U'_\delta \subseteq M'$, we define df_x as the *family* of maps $d(\varphi'_\gamma \circ f \circ \varphi_\alpha^{-1})_{\varphi_\alpha(x)} : V_\alpha \to V'_\gamma$ for all charts U_α around x, all charts U'_γ around $f(x)$. We already know that this map is defined, so we just have to check compatibility:

$$\varphi'_\delta \circ f \circ \varphi_\beta^{-1} = (\varphi'_\delta \circ \varphi'^{-1}_\gamma) \circ \varphi'_\gamma \circ f \circ \varphi_\alpha^{-1} \circ (\varphi_\alpha \circ \varphi_\beta^{-1})$$

So, this construction identifies a tuple $(v_\alpha \in V_\alpha) \in T_xM$ satisfying compatibility to a tuple

$$d(\varphi_\beta \circ f \circ \varphi_\alpha^{-1})_{\varphi_\alpha(x)}(v_\alpha) \in T_{f(x)}M'$$

Then the submersion and immersion theorems follow formally from the vector space versions, which are much easier.

Missed this lecture because I like sleep more than manifolds apparently. Notes taken from Isaac.

Aside on Global Embeddings

Theorem 1.4.1

Any compact manifold M embeds in some $\mathbb{R}^n$, i.e, there exists an injective immersion $M \to \mathbb{R}^n$ that is homeomorphic onto its image.

PROOF : The proof will use bump functions. For all $x \in M$, open neighborhoods $K \subseteq \overline{K} \subseteq U$ of x, there exists a C^∞ function $f : M \to \mathbb{R}$ that is identically 1 on $\overline{K}$ and 0 outside U.

We didn't really prove the fact about bump functions, but we did a homework problem about a function related to $e^{-\frac{1}{x^2}}$ that is a bump function on concentric open balls. There's probably some argument to organize (and renormalize) these on the arbitrary union of open balls. Maybe compactness comes into play.

For all $x \in M$, there exists a chart $(U_\alpha, \varphi_\alpha)$ around x. Choose a smaller closed ball (by the definition of a neighborhood) $\overline{K}_x$ within U_α and let $f_x : M \to \mathbb{R}$ be a bump function that is 1 on $\overline{K}_x$ and 0 outside U_α.

Urysohn's lemma gives an at least continuous bump function from the fact that manifolds are normal (in fact, stronger adjectives hold). Wonder if it's possible to smooth these.

Let $\varphi_\alpha : U_\alpha \to \mathbb{R}^m$, and take the associated map $U_\alpha \to \mathbb{R}^{m+1}$ given by $y \mapsto (f_x(y) \cdot \varphi_\alpha(y), f_x(y))$ which agrees (after restricting to the first m coordinates) with φ_α on $\overline{K}_x$. Extending this construction to all of M by making the function vanish on the complement of U_α clearly gives a C^∞ function F_x.

Having done this around every point of M, pass to a finite subcover (via compactness) centered around $x_1, \cdots, x_k$ and the corresponding maps can be concatenated together to a C^∞ map $F : M \to \mathbb{R}^{k(m+1)}$ which is an embedding.

The derivative is injective everywhere since F_{x_i} is just a coordinate chart around x_i shifted to a hyperplane in $\mathbb{R}^{m+1}$. F is also injective since if $F(y) = F(z)$, then, specifically, $F_{x_i}(y) = F_{x_i}(z)$ for some i, so x and y are in the same U_α for one of the x_i, and $f_{x_i}(y) = f_{x_i}(z) = 1$, so

$$\varphi_\alpha(y) f_{x_i}(y) = \varphi_\alpha(z) f_{x_i}(z) \implies \varphi_\alpha(y) = \varphi_\alpha(z) \implies y = z$$

where the last implication is from the fact that φ_α is a chart. F is homeomorphic onto its image since a continuous bijection from a compact space to a Hausdorff space is a homeomorphism. ■

In the post proof remarks, it's not clear to me whether these claims refer to strengthenings of this proof or just general statements of this form. The former seems almost certainly false since it's impossible to get bounds on the size of a finite subcover in general.

Note that the embedding obtained by running this proof is almost certainly not optimal, dimensionally. M being compact was not necessary, second countability would have sufficed, and one can show that $n = 2m + 1$ is possible. Whitney showed that you can take $n = 2m$, and that this is optimal since $\mathbb{RP}^2$ doesn't embed in $\mathbb{R}^3$.

Tangent Vectors, Redux

If a particle is moving on a manifold, the tangent vectors at a point are the possible directions for the particle to move. The actual trajectory of the particle is a smooth curve $\gamma : R \to M$, which should always have well-defined tangent vectors. Based on this idea, one can (roughly) define tangent vectors and tangent spaces by looking at the tangent vectors of all possible curves on the manifold in question.

This is the physicist's definition, and the one I'm most comfortable with.

Definition 1.4.2

Let $\gamma, \delta : (-\epsilon, \epsilon) \to M$ satisfy $\gamma(0) = \delta(0) = x$. If

$$\frac{d}{dt}\bigg|_{t=0} (f \circ \gamma(t)) = \frac{d}{dt}\bigg|_{t=0} (f \circ \delta(t))$$

for all smooth f from a neighborhood of x to $\mathbb{R}$, we say that γ and δ are tangent vectors to each other.

We need to check that these vectors form a vector space. Given $v \in T_x M$, represented by $\gamma : (-\epsilon, \epsilon) \to M$, λv is represented by $\gamma(\lambda \cdot -)$, e.g, the same image curve, but going λ times as "fast" via the parameterization.

If we have γ, δ, two curves at x with tangent classes γ', δ', then $\gamma' + \delta'$ is defined as follows: pick a chart (U, φ) around x and (by abuse of notation) regard γ and δ as functions from $\mathbb{R}$ to $\mathbb{R}^n$, and define $\gamma' + \delta'$ to be the tangency class of the curve $t \mapsto \gamma(t) + \delta(t)$. There is some chart independence of this definition to check, which is left as an exercise.

Linear Approximations as Covectors

Definition 1.4.3: Cotangent Spaces

The dual of the tangent space $T_x M$ is called the *cotangent space* $T_x^* M$, and its elements are called *covectors*

It is dangerous and incorrect to think of covectors as essentially the same as vectors. For example, the gradient from multivariable calculus is not a vector, it is a covector. In particular, we are allowed to think of gradients as vectors due to the metric (Riemannian) structure on $\mathbb{R}^n$ that allows us to convert vectors to covectors and vice versa. The intrinsic definition of a gradient, in contrast, makes sense for any manifold M (as we will see), not necessarily possessing a natural metric or Riemannian structure.

I omit some discussion here of visualizing covectors on vector spaces via level sets since it was largely a series of pictures I can't convey well here.

The *differential* of a smooth function on (say) $\mathbb{R}^n$ is a linear function $df_x : T_x\mathbb{R}^n \to \mathbb{R}$ (i.e a covector) to be defined, and the gradient vector in $\mathbb{R}^n$ is given by the dot product isomorphism $T_x\mathbb{R}^n \cong T_x^*\mathbb{R}^n$. If we choose different charts this identification will still exist, but the transition maps will typically not be a Euclidean isometry, so ∇f will not be preserved, e.g, the two versions of ∇f are not equal.

Definition 1.4.4

Suppose f is a C^∞ $\mathbb{R}$-valued function defined on a neighborhood of $x \in M$. Then df_x is the linear function $T_xM \to \mathbb{R}$ defined as follows: for all curves γ through x, the derivative of $f \circ \gamma$ is the usual derivative of the function $(-\epsilon, \epsilon) \to \mathbb{R}$, so

$$df_x : \{\text{curves } \gamma \text{ through } x\} \to \mathbb{R}$$

given by $\gamma \mapsto \frac{d}{dt}\Big|_{t=0} f \circ \gamma$. One then checks that curves in the same tangency class get the same number, so that df_x is well-defined as a map from T_xM to $\mathbb{R}$. One also checks that this is a linear function and that $df_x = D_xf : T_xM \to T_{f(x)}\mathbb{R} = \mathbb{R}$ in the notation we sometimes used above.

An even more intrinsic way to get T_xM and T_x^*M is as follows: for $x \in M$, consider the ring of all $\mathbb{R}$-valued C^∞ functions defined on some neighborhood of x. This has an ideal $\mathfrak{m}_x$ of functions vanishing at x.

$\mathfrak{m}_x$ is only finitely generated as an ideal of $\mathcal{O}(U)$, the ring of functions on some open set U containing x, not as an $\mathbb{R}$-vector space.

Lemma 1.4.5: (Hadamard)

$\mathfrak{m}_x$ is generated as an ideal by the coordinate functions $x_1, \cdots, x_n$ in some chart.

Proof : We can show the more general statement that for any function f defined on a neighborhood of $x \in \mathbb{R}^n$, there exist C^∞ functions $g_1, \cdots, g_n$ nonvanishing at x such that $f(y) = f(x) + \sum_{i=1}^n y_i g_i(y)$ for all y close enough to x.

If this holds, then the functions vanishing at x locally have the form $\sum_{i=1}^n y_i g_i(y)$ which lies in the ideal generated by the y_i.

Let's take $x = 0$, $f(x) = 0$, then for all y close to 0,

$$f(y) = \int_0^1 \frac{d(t \mapsto f(ty))}{dt} dt = \int_0^1 \sum_{i=1}^n \frac{\partial f}{\partial y_i}\bigg|_{ty} \frac{dy_i t}{dt}\bigg|_t dt$$

where the first equality is just the fundamental theorem of calculus and the second is the chain rule. $\frac{dy_i t}{dt}\Big|_t = y_i$, so the above becomes

$$\sum_{i=1}^n y_i \int_0^1 \frac{\partial f}{\partial y_i}(ty_1, \cdots, ty_n) dt$$

and these are the C^∞ functions we want. ■

We can then define T_x^*M as $\mathfrak{m}_x/\mathfrak{m}_x^2$. Intuitively, whatever the definition of a derivative is, it should roughly obey some relation like

$$f(x + \Delta x) = f(x) + df_x(\Delta x) + O((\Delta x)^2)$$

so df_x is a linear function that vanishes at x, and we only care about the linear degree of vanishing. This says essentially that df_x is an element of $\mathfrak{m}_x/\mathfrak{m}_x^2$ represented by $f - f(x)$. $\mathfrak{m}_x/\mathfrak{m}_x^2$ is automatically a vector space, consisting of linear approximations to smooth functions vanishing at x.

Defining T_x^*M this way, we can also recover T_xM by dualizing. The advantage of this definition is the total avoidance of charts and coordinates, along with the fact that $\mathfrak{m}_x/\mathfrak{m}_x^2$ generalizes more nicely to (for example) schemes.

Note that the integrals in the above proof can in fact be ill-defined if (say) the chosen open neighborhood is non-convex (so the path of the integral leaves the neighborhood), so what we wrote last time really only shows that there exists a smaller neighborhood on which the formulas hold.

Why is shrinking the neighborhood a problem? Is it just an aesthetic problem?

The notion that will help us clarify this proof is the notion of germs:

Definition 1.4.6: Germs

If M is a manifold, $x \in M$, a *germ* of the C^∞ functions at x is an equivalence class of functions f defined on neighborhoods of x, where 2 such functions are equivalent if they agree identically on some neighborhood where they are defined.

It is clear that this is an equivalence relation, where transitivity follows by shrinking neighborhoods.

With this definition, we can define $\mathcal{O}_{M,x} = \varinjlim_{U \ni x} C^\infty(U)$ (the direct limit of C^∞ functions on neighborhoods of x) as the set of germs of smooth functions at x, which has a natural ring structure (as one can check), and $\mathfrak{m}_x$ the maximal ideal of $\mathcal{O}_{M,x}$ of germs vanishing at x.

Thus, we can restart Hadamard's lemma as follows: if $\in \mathcal{O}_{M,x}$ then there exist $g_i \in \mathcal{O}_{M,x}$ such that

$$f = f(x) + \sum_{i=1}^{n} x_i g_i$$

and the proof works as follows: suppose $[f] \in \mathcal{O}_{M,x}$, so there exists a neighborhood of U of x and some function f on U representing $[f]$. By shrinking U, we can suppose that it is a ball, without changing the germ $[f]$, and then we integrate along radial segments as before.

We in fact did not even avoid shrinking the neighborhood by using germs, but Daniel says that the notion of germs makes the "shrinking neighborhoods" bit more natural/baked in. This seems like a bit of the whole "making choices" vs "working with the moduli space of all possible choices" philosophy that keeps cropping up.

We now have three definitions of T_xM:

1. One vector space for each chart ($U_\alpha, \varphi_\alpha : U_\alpha \to V_\alpha$ containing x, all identified with each other via derivatives of transition maps.
2. Tangency classes of curves through x.

3. The dual vector space to $\mathfrak{m}_x/\mathfrak{m}_x^2$.

One would hope that these are all the same: that the first and second definitions are equivalent amounts to the statement that two curves through the same point in $\mathbb{R}^n$ are in the same tangency class iff their derivatives agree at 0. If some derivative $\dot{\gamma}_i \neq \dot{\delta}_i$, then the coordinate function x_i gives a different rate of change:

$$\frac{d(x_i \circ \gamma)}{dt} \neq \frac{d(x_i \circ \delta)}{dt}$$

so these will not be in the same tangency class. For the other direction, after subtracting off a linear function with row vector matrix $(\dot{\gamma}_1, \cdots, \dot{\gamma}_n) = (\dot{\delta}_1, \cdots, \dot{\delta}_n)$, it is enough to prove that $f \circ \gamma$ has 0 derivative for all f iff $x_i \circ \gamma$ has 0 derivative for all i.

Suppose $x_i \circ \gamma$ has 0 derivative for all i, and let f be given,

$$f(x) = f(0) + \sum_i x_i g_i(x)$$

for x near 0. Then

$$(f \circ \gamma)(t) = f(0) + \sum_i x_i(\gamma(t)) g_i(\gamma_1(t), \cdots, \gamma_n(t))$$

and the product rule gives the derivative of this with respect to t is equal to the sum of terms, all of which contain x_i or $\frac{dx_i}{dt}$ which are then equal to 0.

To see that the second and third formulations are equivalent, note that in $\mathbb{R}^n$, we have a coordinate system $(x_1, \cdots, x_n)$, and their images in $\mathfrak{m}_x/\mathfrak{m}_x^2$ are a basis for $T_x^*\mathbb{R}^n$ for $x = 0$, the origin. They span $\mathfrak{m}_x/\mathfrak{m}_x^2$ since any $f \in \mathfrak{m}_x$ can be written as

$$f = f(0) + \sum_i x_i g_i(x) = \sum_i x_i g_i(0) + \sum_i x_i (\text{functions vanishing at } 0)$$

by Hadamard's lemma, and the second sum on the right lies in $\mathfrak{m}_x^2$ and therefore vanishes in the quotient. The matrix of partial derivatives of the coordinates x_i is just the identity matrix, so the map $\mathfrak{m}_x/\mathfrak{m}_x^2 \to \mathbb{R}^n$ is onto and therefore the x_i form a basis as claimed.

I was confused here why we need to split up the $g_i(x)$ into constant terms and higher order terms since $\mathfrak{m}_x$ is an ideal over the stalk $\mathcal{O}_{\mathbb{R}^n,x}$ so the coefficients of the x_i are allowed to be arbitrary germs at x (not just constants), which to me seemed to imply that the x_i generate $\mathfrak{m}_x$ itself as an ideal (rather than its quotient by the square of itself), in concert with Hadamard's lemma. I guess we expand things out this way to get the second order terms explicitly and kill them off.

Note that we often refer to an element in some ring and its image in some quotient of that ring by the same symbol f, but here, the notation is df, which the element of $\mathfrak{m}_x/\mathfrak{m}_x^2$ represented by $f - f(x)$.

Was spiritually unwilling to attend a Zoom class, and the recording hasn't been posted yet, so this section of notes is adapted from Vincent Hoffmann's notes.
Note the subtlety that df_p and the dx_i all have independent meaning, but the $\frac{\partial f}{\partial x_i}$ do not, in the sense that you can only calculate the partials with respect to a full coordinate system, whereas picking a single coordinate function x_i and looking at the corresponding form dx_i is reasonable and allowed. The partials only have meaning when they are taken together.

Thus, by the above argument, we have that, if $x_1, \cdots, x_n$ are coordinates around a point p of a manifold M, then $dx_1, \cdots, dx_n$ form a basis for $\mathfrak{m}_p/\mathfrak{m}_p^2$. So for any smooth function f defined near p, df_p can be written as a linear combination of the dx_i,

$$df_p = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$$

The above relation holds in any coordinate system. In particular, if φ_α is the chart corresponding to the coordinates x_i, and φ_β is some other chart around p corresponding to coordinates y_i, then $\varphi_\beta \circ \varphi_\alpha^{-1}$ expresses the y_i in terms of the x_i (NB: not necessarily a linear combination), which gives rise to

$$dy_i = \sum_{j=1}^{n} \frac{\partial y_i}{\partial x_j} dx_j$$

Example 1.4.7

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given in polar coordinates, $f(r, \theta) = r^2$. Then

$$df = \frac{\partial f}{\partial r} dr + \frac{\partial f}{\partial \theta} d\theta = 2r dr$$

In rectangular coordinates, $f(x, y) = x^2 + y^2$, so we also have that

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = 2x dx + 2y dy$$

One can show that $rdr = xdx + ydy$ using the chain rule as above and the equality $r = \sqrt{x^2 + y^2}$, so these two expressions coincide as one would expect.

Example 1.4.8

If $\gamma(t) = (5, t)$ in rectangular coordinates, we normally write $\dot{\gamma}(t) = (0, 1)$. In fact, it is more natural to write $\dot{\gamma} = 0\frac{\partial}{\partial x} + 1\frac{\partial}{\partial y}$, e.g, as a partial differential operator, so that evaluating a form (say, df) on it is more straightforward in coordinates. For example, $\frac{d}{dt} f \circ \gamma$ is equal to the pairing of $df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$ with $\dot{\gamma} = \frac{\partial}{\partial y}$, which evaluates to $\frac{\partial f}{\partial y}$ since dx and $\frac{\partial}{\partial x}$ are dual by construction, and similarly for y.

Some discussion right at the end about what a vector field is, I assume it'll be covered in detail next time, so I've omitted it.

The Tangent Bundle

If M is a manifold, then $\cup_{x \in M} T_x M$ can be made into a manifold itself, $TM \to M$. A vector field on M is defined as a global section of this bundle, i.e, a map $M \to TM$ which when followed by $TM \to M$ is the identity.

To build TM as a manifold, we will work with explicit charts. For all charts $\varphi_\alpha : U_\alpha \to V_\alpha$ on M, consider the map

$$U_\alpha \times T_x M \ni (x, v) \mapsto (\varphi_\alpha(x), (d\varphi_\alpha)_x(v)) \in V_\alpha \times V_\alpha$$

These are the charts that make TM a manifold. One must check that the transition maps are smooth, which is essentially immediate since the transition map will be the tuple of a known smooth function (since M is a smooth manifold) and a linear map.

Explicitly, the transition map at $(x_\alpha, v_\alpha) \in U_\alpha \times V_\alpha$ is

$$(\varphi_\beta \circ \varphi_\alpha^{-1}(x_\alpha), (d\varphi_\beta)_x \circ (d\varphi_\alpha^{-1})_{x_\alpha}(v_\alpha))$$

which is smooth since $\varphi_\beta \circ \varphi_\alpha^{-1}$ is by assumption and so its derivative must be as well, and invertible for the same reason.

In these coordinates, a vector field means a function on U_α taking values in V_α that is smooth. In coordinates $x_1, \cdots, x_n$ on U_α, we can write our vector field (locally) as

$$v(x_1, \cdots, x_n) = \sum_{i=1}^{n} v_i(x_1, \cdots, x_n) \frac{\partial}{\partial x_i}$$

so local coordinates x_i give us natural coordinates $\frac{\partial}{\partial x_i}$ on the tangent space.

An important question, in practice, is given a smooth manifold M, does there exist a smooth nowhere vanishing vector field on M?

On S^1, for example, this is possible, and $TS^1 \cong S^1 \times \mathbb{R}$, i.e, the tangent bundle of S^1 is trivial. Note that even smooth vector fields can be fairly poorly behaved, with support the interior of a Cantor set, for example (by adding up bump functions), but in general, after some perturbation, any map can be assumed "good," here meaning transverse.

Sard and Whitney

Theorem 1.5.1: Sard

If $f : M \to N$ is a smooth map of manifolds, then the critical values of f have measure 0.

Note that the critical *values* lie in the codomain, and the critical *points* in the domain. So, for any constant map $M \to \mathbb{R}$, there is a single critical value (assuming $\dim M \geq 1$) but all of M are critical points.

One application of this is that, for almost every $y \in N$, $f^{-1}(y)$ is a manifold (possibly empty). Recall that measure 0 means for a set $X \subseteq \mathbb{R}^n$ that, for all $\epsilon > 0$, there exists a covering of X by boxes (products of intervals) whose sum of volumes is less than ϵ. For a subset of a general manifold, $X \subseteq M$, X has measure 0 if for all $\epsilon > 0$, there exist countably many charts $(U_\alpha, \varphi_\alpha)$ s.t $\varphi_\alpha(X \cap U_\alpha) \subseteq \mathbb{R}^n$ has measure 0 for all α.

So, to make sure the latter definition is sensible, we must show that if $U \subseteq \mathbb{R}^n$, $f : U \to U'$ a diffeomorphism to some other open set in $\mathbb{R}^n$, then $X \subseteq U$ has measure 0 iff $f(X)$ does. To see this, let $V \subseteq U$ be a compact subset, and chop V into countably many boxes V_i, each of which is necessarily compact and lies in U. If $f(V_i \cap X)$ has measure 0 for all i, then $f(V \cap X)$ has measure 0 as well. To see this, the key point is that for all V_i, there exists a bound on $\|df\|$ (in the operator norm sense), say

K, everywhere on V_i. Then given $\epsilon > 0$, choose a covering of $X \cap U_i$ by boxes that have total volume less than $\frac{\epsilon}{K^n}$. The f-images of these boxes will not necessarily be boxes, but can be then covered by boxes which will have volume at most ϵ.

Seems like the bound isn't good enough since you'll gain some volume by re-covering the image with boxes.

Theorem 1.5.2: Whitney's Theorem

Let M be a compact n-manifold, then $M \hookrightarrow \mathbb{R}^{2n+1}$.

PROOF : The idea is as follows: first embed M in some huge $\mathbb{R}^N$ (we showed how to do this far above; this is where we apply compactness). Then, linearly project onto some hyperplane. Sard's theorem shows that this will work with probability 1.

Suppose $N > 2n+1$, then the hyperplanes to project onto are classified by $\mathbb{RP}^{N-1} = \mathbb{RP}^{\geq 2n}$, and consider $(M \times M \setminus \Delta) \xrightarrow{L} \mathbb{RP}^{N-1}$ given by the direction of the line $(m, m') \mapsto \vec{mm'}$ (note that this is not well-defined on the diagonal). Then Sard's theorem implies that the image of $(M \times M) \setminus \Delta \to \mathbb{RP}^{N-1}$ has a regular value (since the critical values have measure zero). Since $\dim((M \times M) \setminus \Delta) = 2n < N-1$, the derivative cannot possibly be surjective anywhere, so every point of $(M \times M) \setminus \Delta$ is a critical point, and the critical *values* are the whole image of L. However, the critical values are known to be measure 0, and therefore not all of $\mathbb{RP}^{N-1}$. Let y be a point not in the image of L, i.e a line in $\mathbb{R}^N$.

Then projection along lines parallel to y (i.e the projection onto the hyperplane orthogonal to y in the standard inner product) is a one-to-one map from M to $\mathbb{R}^{N-1}$. Since y is not in the image of L, no line parallel to y can contain two points of M. This shows that we have a set-level embedding of M in $\mathbb{R}^{N-1}$ and (therefore, inducting down) into $\mathbb{R}^{2n+1}$. It remains to show that the derivative is injective too.

Intuitively, this amounts to showing that we can choose the projection map π to separate "distinct infinitesimally close points," not just separate distinct points. Another way to think of this is that π should separate tangent directions, or never project a tangent space to 0. Imagine a curve with a vertical tangent vector being projected down to the x-axis; this is the situation we are trying to avoid.

To that end, we will work with TM instead of M; the above map $i : M \hookrightarrow \mathbb{R}^N$ induces a map $TM \hookrightarrow T\mathbb{R}^N = \mathbb{R}^{2N}$, which is an embedding since i is. Then we can project to $\mathbb{R}^N$ by the map $TM \ni (x, v) \mapsto i(x) + (D_x i)(v)$.

If we choose the parallelization class y of lines to project along which does not lie in any of the tangent spaces regarded as subsets of $\mathbb{R}^N$, then π remains injective on tangent spaces and therefore embeds M in $\mathbb{R}^{N-1}$. To see that we can pick such a y, we employ Sard's theorem again.

Consider $TM\backslash\{\mathbf{0}\} \to \mathbb{RP}^{N-1}$ where $\mathbf{0}$ denotes the 0-section, taking (x, v) to the line through x in the v direction, i.e, the line corresponding to $(D_x i)(v)$. Argue as before: since $\dim TM\backslash\{\mathbf{0}\} < \dim \mathbb{RP}^{N-1}$, again critical values are the whole image, so there exists $y \in \mathbb{RP}^{N-1}$ not in the image. This is exactly to say that y is transverse to every one of the tangent spaces $T_x M \subseteq \mathbb{R}^N$. Note that we actually need to choose y to satisfy this criteria and the above criteria simultaneously at each step of the downward induction. ∎

Differential Topology **Spring 2023**

Beyond the Basics

Professor Daniel Allcock Abhishek Shivkumar

Morse Functions

No good reason to have a chapter break here, but really the whole class will be "Foundations of Manifolds," so here's as good a spot as any.

Definition 2.1.1

Let $f : M \to \mathbb{R}$ be a smooth function. $x \in M$ is a *nondegenerate critical point* of f if $df|_x = 0$ and the Hessian of f at x is nonsingular.

Recall that the Hessian of a function is (in coordinates) its matrix of second partial derivatives. Alternatively, it can be described as a symmetric bilinear form as follows: let γ represent a tangency class at x, then the second derivative of $f \circ \gamma$ at $t = 0$ gives a function H from T_xM to $\mathbb{R}$. Thinking of H as a "norm" (in quotes because it is allowed to be negative), we may obtain a corresponding symmetric bilinear form via the standard construction:

$$\langle v, w \rangle_H = v^T H w$$

By linear algebra, there exists a basis in which a symmetric bilinear form can be written as a digaonal matrix with only ± 1 and 0 on the diagonal (0s only if the form is degenerate), in which case the corresponding form can be written as $Q(x_1, \cdots, x_n) = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2$ (assuming nondegeneracy).

Lemma 2.1.2: Morse

If x is a nondegenerate critical point of a function $f : M \to \mathbb{R}$, then there exist local coordinates at x such that $f = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2$.

Morally, this is just expressing the idea that any (nondegenerate) critical point corresponds to a cap, a cup, or some kind of complicated saddle. As an example, one can picture the "top" of a sphere as being well-approximated by a paraboloid, and the inside corner of a torus being well-approximated by a saddle.

There were several examples here in class that I'm omitting because they're not super useful without the accompanying pictures. See Milnor's Morse Theory for examples.

Note also that a function being Morse (i.e having nondegenerate critical points) is not an arduous restriction; in various contexts one can prove that Morse functions are dense.

Lemma 2.1.3

Let $M \subseteq \mathbb{R}^n$ be a submanifold, and $f : M \to \mathbb{R}$ any smooth function. Then, for a linear function λ from $\mathbb{R}^n$ to $\mathbb{R}$, the set of such λ such that $f+\lambda$ is *not* Morse has measure zero, i.e, the linear perturbations of f that are Morse have full measure.

As an example, consider $f(x) = x^3$ from $\mathbb{R}$ to $\mathbb{R}$, which has a degenerate critical point at $x = 0$. Perturbing f by a positive multiple of x leads to no critical points ($f'(x) = 3x^2 + a = 0$ has no solutions) and by a negative multiple of x leads to two nondegenerate critical points ($f'(x) = 3x^2 - a = 0 \iff x = \pm\sqrt{\frac{a}{3}}$).

Towards this theorem, we have the following lemma:

Lemma 2.1.4

$f : M \to \mathbb{R}$ is Morse, that is, each critical point is nondegenerate (i.e the Hessian at the point has nonzero determinant), iff the partials of f form local coordinates around each critical point of f (even though they must, by definition, vanish *at* the critical point).

PROOF : Since this is a question about each critical point of f, we can immediately localize to a single chart and assume that M is an open subset of $\mathbb{R}^n$ (and that the critical point we are interested in is $0 \in M$). Assuming the Hessian is nondegenerate, f is locally given by $f = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2$, so

$$\left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n}\right) = (2x_1, \cdots, 2x_k, -2x_{k+1}, \cdots, -2x_n)$$

For x_i near 0, these are clearly local coordinates.

For the converse, suppose the partials of f are local coordinates at 0, i.e, $df : x \mapsto \left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n}\right)$ is invertible, so, its differential is invertible at 0. The differential of df is precisely the Hessian, from which the result follows. ■

This immediately implies the following:

Lemma 2.1.5

A nondegenerate critical point p of $f : M \to \mathbb{R}$ is necessarily isolated.

Degenerate critical points can evidently be non-isolated; take the zero map for example.

PROOF : $p \mapsto 0$ under the isomorphism $x \mapsto \left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n}\right)$ and by above lemma, every point near p maps to a nonzero tuple of partial derivatives, from which it follows that the critical point is isolated. ■

Then, we are ready to prove one formulation of the result that Morse functions are dense (as above):

PROOF : First assume that $M \subseteq \mathbb{R}^k$ is an open subset, and consider $g : M \to \mathbb{R}^k$ given by

$$g(m) = \left(\frac{\partial g}{\partial x_1}(m), \cdots, \frac{\partial g}{\partial x_k}(m) \right)$$

Then the derivative of g is a $k \times k$ matrix whose entries are the partials of the components of g, i.e, the Hessian of f at $m \in M$. Consider $a \in \mathbb{R}^k$ a regular value of g; then we claim that $f - \lambda_a$ is Morse (where $\lambda_a : \mathbb{R}^k \to \mathbb{R}$ is the linear map given by taking the dot product with a). To see this, note that

$$d(f - \lambda_a) = df - a = \left(\frac{\partial f}{\partial x_1} - a_1, \cdots, \frac{\partial f}{\partial x_k} - a_k \right)$$

If this vanishes at some point p, then, by definition, $g(p) = a$. But then Dg is an isomorphism since a is a regular value, and so the Hessian is nondegenerate. So for all critical points of $f - \lambda_a$, the Hessian of f is nondegenerate, so $f - \lambda_a$. To finish, we apply Sard's theorem to show that regular values a have full measure.

In the general case, cover M by open sets U for which some $m = \dim M$ (not necessarily equal to k) many of the k standard linear coordinates form a coordinate system on that chart. By second countability we can take countably many of these open sets to cover M, so it is enough to show that the theorem holds for a given open set U; this is because on each open set, we get a measure zero set of the space of linear functions that are "bad," i.e, s.t $f - \lambda$ is not Morse, and the countable union of measure zero sets is still measure zero, so on all of M we will still have full measure of choices of linear perturbations.

To see that the theorem holds for such an open set U, replace f by $f \circ \pi^{-1}$: where $\pi : U \to \mathbb{R}^m$ is the coordinate projection map, and apply the first half of this proof, i.e, that linear perturbations in m of the k coordinates of f will give a Morse function. For simplicity, assume that we can take these to be the first m coordinates

Note that this argument applies just as well to $f + b_{m+1}x_{m+1} + \cdots + b_k x_k$, so for a.e $(a_1, \cdots, a_m) \in \mathbb{R}^m$, $(f + b_{m+1}x_{m+1} + \cdots + b_k x_k) + (a_1 x_1 + \cdots + a_m x_m)$ is Morse. Thus, for each such tuple b, the set of a which make the perturbation non-Morse is measure zero, so it follows by Fubini's theorem for product measures that the whole set of k-dimensional linear perturbations that are not Morse has measure zero. ■

The basic idea: partitions of unity extend results from compact manifolds to all manifolds by giving us a nice framework for patching together local constructions.

Partitions of Unity

Definition 2.2.1: Partition of Unity

A *partition of unity* on a manifold M is a family of pairs (U_α, f_α) where U_α is open in M, and $f_\alpha : M \to [0,1]$ is smooth and supported in U_α, such that every $x \in M$ has a neighborhood meeting only finitely many U_α, and $\sum_\alpha f_\alpha$ is identically equal to 1; this sum is always finite since only finitely many f_α are nonzero at each point in x, so this sum makes sense.

Example 2.2.2

Suppose we have $a : M \to \mathbb{R}$. Cover M by open sets V_β and choose perturbations a_β of a on each of them. Choose a partition of unity whose open sets refine the V_β, i.e, a cover U_α such that for each α, $U_\alpha \subseteq V_\beta$ for some β. Then, for all α, consider $f_\alpha \cdot a_\beta$ for $U_\alpha \subseteq V_\beta$, which is a smooth function supported on V_β and can be smoothly extended by 0 to all of M.

Then, we claim that the sum $\sum_\alpha f_\alpha \cdot a_\beta$ makes sense, where if U_α lies in more than one V_β, we just pick one (implicitly using the axiom of choice here).

For existence of a partition of unity, it suffices to treat the case that M is connected. In this case, M being second countable (as a manifold) is equivalent to being σ-compact, i.e, the countable union of compact sets. Choose a sequence $K_0, K_1, \cdots$ compact with $K_0 \subseteq K_1 \subseteq \cdots$, and $\cup_i K_i = M$. Each point has a basis of neighborhoods, on each of which we can define smooth functions whose support is contained in the neighborhood via bump functions.

Note that the existence of bump functions is a feature of the smooth and continuous case, and fails in the analytic case (an analytic manifold being one whose transition maps have local Taylor expansions), since the bump function $e^{-\frac{1}{x^2}}$ is smooth but not analytic since its Taylor expansion is the zero function. The distinction between smooth and analytic disappears in the complex case (with smooth replaced by holomorphic).

Cover K_0 by finitely many sets, and $K_n \setminus \mathrm{Int}(K_{n-1})$ by finitely many open sets chosen to miss K_{n-2}, and define functions on these which are identically 1 on smaller open sets, which also cover $K_n \setminus \mathrm{Int}(K_{n-1})$. These open sets are locally finite by construction because each $x \in M$ lies in the interior of some K_n, and is outside all open sets used for stages $n+2$ and beyond. So $\sum_\alpha f_\alpha$ makes sense, and is identically ≥ 1. Define $g_\alpha(x) = \frac{f_\alpha(x)}{\sum_\beta f_\beta(x)}$ which is identically 1, so we are done.

The majority of this proof was a picture that I unfortunately cannot reproduce here.

Theorem 2.2.3: Whitney, again

Every σ-compact manifold M of dimension m embeds in $\mathbb{R}^{2m+1}$.

I didn't fully follow this proof at the time and it was largely a sketch, some of which is preserved here. It doesn't seem to match any extant proofs of the Whitney embedding theorem that I can find.

PROOF : Choose a union of open sets $U_i \subseteq U_{i+1}$ exhausting M, where each U_i is constructed from U_{i-1} by adjoining the domain of a single chart (V, φ) (we can do this by the assumption of σ-compactness, i.e, M is the union of countably many compact subsets; in fact all manifolds are σ-compact from second countability) Suppose that U_i is already embedded in $\mathbb{R}^{2m+1}$ (since we may assume that U_i is an honest coordinate chart for $i = 1$, which must embed in $\mathbb{R}^m \hookrightarrow \mathbb{R}^{2m+1}$, and induct up), and U_{i+1} is constructed by adjoining (V, φ). Extend φ to a function on all of M by setting it to 0 outside of V, and multiply by a bump function defined on a smaller open subset of V (so that φ remains smooth).

Recall that a map is *proper* if preimages of compact sets are compact.

In this manner, we will construct a sequence of functions $f_i : M \to \mathbb{R}^{2n+1}$ which are all compactly supported, but are embeddings on open sets U_i that exhaust M. We will put them all together to embed M. Use a partition of unity $f_1, \cdots$, and consider $g(x) = \sum_n n f_n$. If x lies in the support of $f_k, \cdots, f_l$, and not in any others, then $k \leq f(x) \leq l$. One can show that g is a proper function. ■

Manifolds with Boundary

To allow manifolds with boundary, we just modify our notion of chart to allow $(U_\alpha, \varphi_\alpha)$ where U_α is an open set of H^n, the closed upper half of $\mathbb{R}^n$ (i.e with ≥ 0 first coordinate).

Example 2.3.1

$[0, 1]$ is a closed manifold with two charts; $[0, 1)$ and $(0, 1]$ both map homeomorphically to H^1 via $\tan\left(\frac{\pi}{2}x\right)$ and $\tan\left(\frac{\pi}{2}(1-x)\right)$ respectively.

Example 2.3.2

The closed n-ball is a manifold with boundary with charts that can be given by extending the stereographic projection.

The compatibility condition is that the transition functions are smooth but one needs to be careful since we are now dealing with half spaces, which have a boundary. In particular, smoothness means that $\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$ is the restriction of of a smooth map defined on an open set in $\mathbb{R}^n$ whose intersection with H^n is $\varphi_\alpha(U_\alpha \cap U_\beta)$.

Don't think I understand the intrinsic meaning of this over a variant of this definition. Just that we can say (imprecisely) that the charts are "smooth at the boundary"?

Tangent vectors can be defined as before. Note that the tangent space at the boundary will not match the dimension of the boundary, but the dimension of the manifold itself. In the equivalence classes of vectors in charts picture, this should be clear since $T_x H^n = \mathbb{R}^n$ at each point x (including in the boundary). In the equivalence classes of curves picture, this is even more clear, since there can always be a curve towards the boundary trying to "escape" the manifold, which recovers our tangent direction that *a priori* maybe “should” be missing.

Proposition 2.3.3

If M is a manifold with boundary, then ∂M is a manifold without boundary in a natural way.

I think natural here should just mean that there is a unique manifold structure on the boundary compatible with the manifold structure on M. The essential point is that boundary points have to map to the boundary of H^n, which I don't think we spent much time on.

PROOF : The charts at a point p of ∂M identify a neighborhood of p with a neighborhood in the boundary of H^n, with boundary points mapping to boundary points. So a chart φ takes $p \in U$ to $\varphi(p) \in \partial H^n = \mathbb{R}^{n-1}$, and $\varphi|_{\partial M}$ is itself a chart valued in $\partial H^n = \mathbb{R}^{n-1}$. The compatibility condition between charts arising in this way is automatically satisfied since they are satisfied by the original charts. ■

Many of our results for manifolds without boundary will hold for manifolds with boundary.

Theorem 2.3.4: Transversality (with ∂)

Let $f : X \to Y$ with X a manifold with boundary, Y a manifold without boundary, and $Z \subseteq Y$ a submanifold. Then $f^{-1}(Z)$ is a manifold with boundary if f is transverse to Z and $\partial f := f|_{\partial X}$ is transverse to Z.

Don't see why x_n (the height function) ought to be independent of the other coordinates. Why should the f^*y_i be supported only in ∂X? This proof is incomplete since I don't really follow that step.

PROOF : We will examine $f^{-1}(Z)$ in the neighborhood of a single point $x \in f^{-1}(Z) \cap \partial X$, since if $x \in X \setminus \partial X$, then $f^{-1}(Z)$ is a manifold near x by the ordinary submersion theorem. By transversality of f to Z, there exists a partial set of local coordinates $y_1, \cdots, y_k$ which define Z locally (i.e by $y_i = 0$) and have pullbacks to X with linearly independent differentials, so we may extend the $f^*y_i(x) := y_i(f(x))$ to a larger set of functions on a neighborhood of x in H^n (after identification with a coordinate chart around x), which serve as local coordinates in ∂X. Then x_n, the height function (above the boundary) on H_n, is linearly independent of all coordinates so far introduced, so the implicit function theorem implies that f^*y_i, x_j, x_n are local coordiantes around p. In these coordinates, the preimage of Z is given by $f^*y_1, \cdots, f^*y_k = 0$. The corresponding locus in H^n is clearly a manifold with boundary given by $x_n = 0$. ■

Theorem 2.3.5: Sard (with ∂)

If X is a manifold with boundary, $f : X \to Y$, Y without boundary, then for almost all $y \in Y$, $f^{-1}(y)$ is a manifold with boundary.

Just use the regular Sard's Theorem twice: once on the manifold and once on the boundary.

PROOF : f can only fail to be transverse to $\{y\}$ for a measure zero set of y, and for only a measure zero set of y can ∂f fail to be transverse to $\{y\}$. Union of these sets is measure zero. ■

Theorem 2.3.6

A compact connected smooth manifold possibly with boundary is diffeomorphic to S^1 or $[0,1]$.

Only a proof sketch here.

PROOF : M embeds in $\mathbb{R}^3$ by Whitney's theorem, so we can use this embedding to write a Morse function f on M. The critical points of f are isolated. Consider the following two families of open sets: small non-overlapping open neighborhoods of the critical points of f and the components of the complement of the set of critical points. The latter components are each diffeomorphic to an open interval via f (which is locally a diffeomorphism at non-critical points). Thus, taking the two families together, we have covered M by open intervals that are disjoint except for neighborhoods of the critical points. Then one argues that these either glue together to a circle or an interval. ■

Corollary 2.3.7

If M is a compact manifold with boundary, then there does not exist a smooth retraction $M \xrightarrow{f} \partial M$.

PROOF : Choose a regular value p of f, the preimage is a manifold with boundary. The number of boundary points of this manifold is 1 since p itself lies in the preimage, and no other boundary points can map to p since f is a retraction. But there is no manifold with a single boundary point by the above classification. ■

Corollary 2.3.8: Brouwer

If $f : D^n \to D^n$ is smooth, then it has a fixed point.

PROOF : Suppose f has no fixed points, and define a retraction $r(x)$ from D^n to $\partial D^n = S^{n-1}$ given by the intersection of the boundary with the (oriented) ray from $f(x)$ to x. This map is clearly smooth, and is the identity on the boundary by construction. But no such smooth retraction can exist, so we have a contradiction. ■

Corollary 2.3.9: Continuous Brouwer

The above result holds with "smooth" replaced by "continuous."

SKETCH : The idea is to just perturb f slightly to make it smooth; for example, by the Stone-Weierstraß theorem, we can approximate any continuous function (with some adjectives) uniformly and arbitrarily well by polynomials, so we obtain $g : D^n \to \mathbb{R}^n$ whose image is close enough to the disk that g has fixed points. By rescaling the image, we can obtain a smooth map from D^n to itself with no fixed points, and argue as above. ■

Theorem 2.3.10

Suppose $F : X \times S \to Y$ is a submersion, with X potentially having boundary, S and Y without boundary. Suppose $Z \subseteq Y$ is a submanifold. Then for almost every $s \in S$, $f_s : X \to Y$ and ∂f_s are transverse to Z.

Slogan: submersions are generic.

Example 2.3.11

Suppose $f_0 : X \to \mathbb{R}^n$ is smooth. Then, we can find a family S of perturbations of f_0 s.t $F : X \times S \to \mathbb{R}^n$ satisfies the submersion hypothesis. For example, we can take S to be the unit ball, and define $F(x,s) = f_0(x) + s$.

Proof : Since F is a submersion, $W = F^{-1}(Z)$ is a submanifold of $X \times S$ (with boundary). We want s s.t $X \times \{s\}$ is transverse to W; we claim that if s is a regular value of $\pi : X \times S \to S$ (the projection map to S) *restricted* to W and ∂W, then f_s and ∂f_s (respectively) are transverse to Z. Sard's theorem implies that such s have full measure, so the theorem will follow from this claim.

To see the claim, suppose that s is a regular value of $\pi|_W$ and $\pi|_{\partial W}$. Let $f_s(x) = z$ for some $x \in X$. Since F is transverse to Z by assumption,

$$dF_{(x,s)} T_{(x,s)}(X \times S) + T_z Z = T_z Y$$

Concretely, this means that for any vector $a \in T_z Y$, there exists a vector $b \in T_{(x,s)}(X \times S)$ s.t $dF_{(x,s)}(b) - a \in T_z Z$. Since a is arbitrary, for transversality of f_s, we want to show that there exists $v \in T_x X$ s.t $d(f_s)_x(v) - a \in T_z Z$. Since $T_{(x,s)}(X \times S) = T_x X \times T_s S$, so $b = (w, e)$ with $w \in T_x X$ and $e \in T_s S$. If $e = 0$, we are done, since $dF_{(x,s)}(w, 0) = d(f_s)_x(w)$. If $e \neq 0$, we may use $d\pi$ to kill off the vector e. Since s is a regular value of $\pi|_W$, $d\pi_{(x,s)}$ maps $T_{(x,s)}(W)$ onto $T_s S$, so for $e \in T_s S$, there exists $(u, e) \in T_{(x,s)}(W)$ mapping to e. But then $dF_{(x,s)}(u, e) \in T_z Z$ since $F|_W : W \to Z$ should map TW to TZ. Then set $v = w - u \in T_x X$, and we have

$$d(f_s)_x(v) - a = dF_{(x,s)}[(w,e) - (u,e)] - a = [dF_{(x,s)}(w,e) - a] - dF_{(x,s)}(u,e)$$

$dF_{(x,s)}(w, e) - a \in T_z Z$ by assumption, and $dF_{(x,s)}(u, e) \in T_z Z$ by construction, so $d(f_s)_x(v) - a \in T_z Z$ as desired. The same argument applies for ∂f_s when s is a regular value of $\partial \pi$; in fact, this case is just the result for the case of a boundaryless manifold. ■

Missed another lecture, notes transcribed again from Vincent Hoffmann's notes.

The above result will nearly suffice to prove that, for any smooth map between manifolds, and a boundaryless submanifold of the target, there exists a smooth map homotopic to it that is transverse to the given submanifold. First, a lemma:

Lemma 2.3.12

If $f : X \to Y$ is smooth, Y boundaryless (so X potentially with boundary), $Z \subseteq Y$ a submanifold, $U, V \subseteq X$ open sets whose closures are disjoint, then there exists a deformation $F : X \times S \to Y$ of f, s.t $F(x, s) = f(x)$ for all $x \in U$, all s, and $F|_V$, $\partial F|_V$ are transverse to Z.

PROOF : We induct on the number of charts on Y (we can take Y to have finitely many charts by first embedding in $\mathbb{R}^{2n+1}$ and looking at the $2n+1$ projection maps to coordinate axes as local coordinates). The base case is when Y is contained in $\mathbb{R}^n$ as an open subset. By the smooth Urysohn lemma, we can take a smooth function $\tau : X \to [0, 1]$ which is 0 on $\overline{U}$ and 1 on $\overline{V}$. Let $r : X \to \mathbb{R}$ be given by the distance from $f(x)$ to the boundary of Y in $\mathbb{R}^n$ (this is well-defined since it is a real-valued function on the small S^n around $f(x)$, and S^n is compact so the function achieves a minimum value). Clearly, r is smooth.

Let S be the unit open ball in $\mathbb{R}^n$, and define $F : X \times S \to Y$ given by $F(x, s) = f(x) + sr(x)\tau(x)$ (rescaling r by a constant factor less than one half so that F actually lands in Y). $f_s = f$ on U since $\tau|_U = 0$, and F is a submersion on $\overline{V} \times S$ since for each x, $F|_{\{x\}\times S}$ is an embedding of S into Y, so F satisfies the desiderata.

Induction is a little sketchy, a lot left to check.

For the inductive step, suppose Y is covered by charts $W_1, \cdots, W_n$. By induction, there exists $F : X \times S \to Y$ satisfying the requirements on $W_1 \cup \cdots \cup W_{n-1}$. Now, we essentially repeat the base case: take F in place of f, and perturb F on $W_n \times S$ to get a function $G : (X \times S) \times T \to Y$ that is transverse to Z. Moreover, choose G to disagree with F only on $\overline{W_n} \cap \overline{U}$, and regard $S \times T$ as the parameter space. Then G is our desired function. ■

Corollary 2.3.13

Suppose $f : X \to Y$ is smooth, Y boundaryless, Z a submanifold of Y. Suppose f is transverse to Z at each x in some closed $K \subseteq X$, and ∂f is transverse to Z for all $x \in K \cap \partial X$. Then there exists a homotopy f_s of f s.t f_1 is transverse to Z, ∂f_1 is transverse to Z, and f_1 can be taken to be "arbitrarily close" to f.

Unclear to me in what sense f_1 can be taken to be arbitrarily close to f.

PROOF : Apply the lemma to an open set $U' \subseteq \overline{U'} \subseteq K$ and $X \setminus K$; now we have a map $F : X \times S \to Y$ s.t $F(x, s) = f(x)$ for all $x \in \overline{U'}$, F is transverse to Z on $X \setminus K$, and F is transverse to Z on U' since f was. By choosing a small enough perturbation, we can conclude transversality on K as well, since f was transverse there, and picking a path in the ball S, we have the desired homotopy to a function transverse to Z everywhere. ■

Tubular Neighborhood Theorem

For an embedded curve in (say) $\mathbb{R}^3$, each point along the curve has a transverse disk that does not intersect the curve, so it is intuitively easy to build a tubular neighborhood of the curve by gluing together these "normal" disks to build a space which looks like the curve $\times$ a disk locally. All the tubular neighborhood theorem tells us is that this is true generally.

Note that we can only ask that the neighborhood we build looks like a product *locally*, as in the definition of fiber bundles, since (for example) on a Möbius band, the transverse intervals to an embedded equatorial curve will reverse orientation as the curve goes along the band, and in fact, the tubular neighborhood built this way will just be a smaller copy of the Möbius band.

For our discussion, we will need a notion of orthogonality; for $Z \subseteq \mathbb{R}^n$ this is clear, since the normal space to $z \in Z$ (denoted N_zZ) can be taken as the subspace of $T_z\mathbb{R}^n$ which is orthogonal to T_zZ. Implicitly we are using the standard inner product on $\mathbb{R}^n$.

For $Z \subseteq \mathbb{R}^n$, there is a smooth map from $NZ := \cup_{z\in Z} N_zZ$ to $\mathbb{R}^n$ given by $(z, v) \mapsto z + v$, identifying normal vectors with actual nearby points to z in a natural way. Of course, one must check that NZ is actually a manifold, which we will show by identifying it as a subset of $T\mathbb{R}^n$ in a natural way.

First, we choose linear coordinates $x_1, \cdots, x_n$ so that a given point z is the origin, and T_zZ is defined by the equations $x_1 = \cdots = x_k = 0$, i.e, T_zZ is spanned by the unit vectors in the remaining $n - k + 1$ directions, and N_zZ is spanned by the tangent vectors $\partial_1, \cdots, \partial_k$. Thus, the defining equations for $Z \subseteq \mathbb{R}^n$ are $0 = f_i = x_i + \cdots$ where the higher order terms are omitted, so for z' close enough to z, $T_{z'}Z$ is given by evaluating $df_1, \cdots, df_k$ at z'. Thus, we can pick a neighborhood U of $z \in Z$ and we have the map $U \times N_zZ \to NZ$ given by $(z', v) \mapsto (z', \mathrm{proj}_{N_{z'}Z} v)$ where $\mathrm{proj}_{N_{z'}Z} v$ is just the projection of v to $N_{z'}Z$ (regarded as a subspace of $\mathbb{R}^n$). The derivative of this map at z is the identity map $T_zY \to T_zY$, hence the map is a local diffeomorphism, and therefore these maps as z varies over Z give a manifold structure to NZ.

Theorem 2.4.1

If Z is a submanifold of $\mathbb{R}^n$, then the map $(z, v) \mapsto z + v$ from NZ to $\mathbb{R}^n$ is a diffeomorphism from a neighborhood of the 0-section of NZ to a neighborhood of Z in $\mathbb{R}^n$.

Note that for an ambient space Y not equal to $\mathbb{R}^n$, building a normal bundle requires a few more steps. For one thing, we no longer have a given notion of orthogonality, so $NZ \subseteq TY$ requires more work. We also won't necessarily have a notion of addition on Y, so $NZ \to Y$ requires some work as well.

One approach is to avoid orthogonality altogether, and define $N_z Z$ as $T_z Y / T_z Z$. Another approach is to introduce orthogonality by choosing a Riemannian metric (a smoothly varying family of symmetric bilinear forms on the tangent spaces) on Y (via an embedding in some $\mathbb{R}^N$ or a partition of unity argument) and proceed as before, defining $N_z Z$ as the orthogonal complement of $T_z Z$ in $T_z Y$. One can show that these definitions give the "same" normal bundles and tubular neighborhoods, and that the choice of metric does not matter.

To define a map $NZ \to Y$, however, a Riemannian metric is required (the abstract formulation does not work), and establish the properties of the exponential map from TY to Y which is given by extending tangent directions to Riemannian geodesics (thereby replacing the addition on $\mathbb{R}^n$ with a local notion of addition). With this discussion in mind, we are ready to prove the above (limited) version of the theorem:

That adding z to the disk $N_z Z$ gives a submanifold that intersects Z transversely doesn't seem obvious to me. This is mostly a sketch.

PROOF : One can compute the derivative of this map and see that it is surjective at each $(z, 0)$. Clearly, the composition $Z \to \{0\text{-section}\} \subseteq NZ \to Y$ is the identity map on Z, and the map $N_z Z \to z + N_z Z \to Y$ given by adding z is a diffeomorphism onto a submanifold of complementary dimension and transverse to Z. Then, by the inverse function theorem, the map is a local diffeomorphism around any $(z, 0)$. This isn't quite enough to finish, since there is *a priori* the possibility that our neighborhoods get arbitrarily small to avoid self-intersection. To find a neighborhood of the 0-section on which the exponential map is a diffeomorphism, we need to shrink our neighborhoods at each $z \in Z$ which are small enough to miss Z (except at z). ■

Intersection Theory

Morally, what we want some machinery to deal with is the idea that the intersection of two submanifolds of a manifold X is a submanifold of expected codimension. One can think of two spheres in three space which may miss each other, though we can always perturb them to intersect in a circle. From our point of view, then, this intersection should be empty, so the circle obtained by intersecting should be trivial (since it bounds a disk in either S^2):

Example 2.5.1

Consider the 3-torus T^3 obtained by identifying opposite faces of a cube. Two orthogonal squares in the middle of T^3 can be thought of as a representation of the intersection of two ordinary tori in T^3, with intersection S^1. However, one can imagine sending a "feeler" out from one of the tori (drawn as a square inside a cube) so that the two tori intersect in two circles. The intuition we want to have is that the second circle is trivial, because it bounds a disk in the second torus.

A lot of the power of working $\pmod 2$ seems to come from repeated use of the fact that a compact one manifold has an even number of boundary points.

For a more formal development, we will restrict to submanifolds that have complementary codimension so that (after a perturbation) their intersection will be a discrete set. When this is a finite set, the size of this set is called the intersection number; for now, we will only discuss the intersection number mod 2, since this is all we can deal with as a topological invariant (for now), since, for example, the intersection of two lines in $\mathbb{R}^2$ is homotopic to the intersection of some cubic curve and a line, with 1 and 3 intersection points respectively (and generically).

Suppose $f : X \to Y$ is smooth, with X closed, and Y connected, and $Z \subseteq Y$ a submanifold. Then we define $I_2(f, Z)$ as the $\pmod 2$ number of points in $f'^{-1}(Z)$ where f' is a perturbation of f that is transverse to Z.

Proposition 2.5.2

If $f_0, f_1 : X \to Y$ are both transverse to Z as above, then $I_2(f_0, Z) \equiv I_2(f_1, Z) \pmod 2$.

PROOF : Regard $X \times I$ as a manifold with boundary, and $F : X \times I \to Y$ the homotopy from f_0 to f_1. Perturbing F if necessary, we may assume F is transverse to Z (via a homotopy that is trivial near $X \times \{0, 1\}$). Then $W := F^{-1}(Z)$ is a compact manifold with boundary whose boundary lies in $\partial X \times I = X \times \{0, 1\}$. Now, by dimension counts, W is a compact one-manifold with boundary, it is a union of circles and closed intervals, ∂W has an even number of points. We can decompose $\partial W = f_0^{-1}(Z) \cup f_1^{-1}(Z)$, so $I_2(f_0, Z) \equiv I_2(f_1, Z) \pmod 2$ as desired. ■

Lemma 2.5.3

Suppose that X is a compact manifold with boundary, $F : X \to Y$, Z a closed submanifold of Y. Then $I_2(\partial F, Z) = 0$.

PROOF : The idea is that given $f : \partial X \to Y$, if you can find a compact manifold X that it bounds, and extend f to this larger manifold X, then $I_2(f, Z) = 0$ (this is just a rephrasing). First perturb F so that F and ∂F are both transverse to Z. Then $F^{-1}(Z)$ is a compact one-dimensional manifold with boundary (by dimension counting), and therefore has an even number of boundary points. But $\|\partial F^{-1}(Z)\| = \|f^{-1}(Z)\| \equiv 0 \pmod 2$. ■

Orientations

An orientation on our manifolds is what will allow us to pass from (mod 2) intersection theory to "honest" intersection theory.

Recall that an orientation for a real vector space V is an equivalence class of ordered bases for V, where an equivalence between bases x_i and y_i is a linear transformation $x_i \mapsto y_i$ with positive determinant (i.e, a choice of connected component of $\mathrm{GL}_n \mathbb{R}$). Equivalently, the matrix whose columns are the y_i with respect to the basis of the x_i has positive determinant. Therefore, there are clearly two choices of orientation for a vector space V.

A more abstract approach is given by looking at the exterior algebra $\bigoplus_{k=0}^{\dim V} \bigwedge^k V$ where $\dim \bigwedge^k V = \binom{\dim V}{k}$. Recall that the exterior algebra is the quotient of the tensor algebra by all relations of the form $x_i \otimes x_j = -x_j \otimes x_i$ (with all other tensor factors implicitly left constant), so there is a natural S_k action (that descends to a $\mathbb{Z}/2$ action) on $\bigwedge^k V$ given by permuting factors. The key fact (for our purposes) is that the top exterior power $\bigwedge^{\dim V}$ is one dimensional, with basis $x_1 \wedge \cdots \wedge x_n$, with $y_1 \wedge \cdots \wedge y_n = \det M x_1 \wedge \cdots \wedge x_n$ where M is the linear transformation taking x_i to y_i, so an orientation of V is a choice of a component of $\bigwedge^{\dim V} V \setminus \{0\}$.

In class, Allcock defines the exterior power $\bigwedge^k V$ as a subspace of $V^{\otimes k}$ consisting of all "totally antisymmetric tensors," i.e, $x_1 \wedge x_2$ is literally equal to $x_1 \otimes x_2 - x_2 \otimes x_1$.

An orientation on a manifold is defined similarly, via the tangent spaces; an orientation of a manifold M at a point x is an orientation of $T_x M$. If there are local coordinates $x_1, \cdots, x_n$, then the orientation is given by the ordered basis $\frac{\partial}{\partial x_1}, \cdots, frac\partial\partial x_n$ as above. Note that we can then immediately extend this orientation to all points of the coordinate patch.

Definition 2.6.1: Orientation

An *orientation* on a manifold M means a smooth choice of orientation over all the tangent spaces $T_x M$, i.e, a smooth map $M \to \bigwedge^n TM = \det M$.

One can parse this in practice as either a coordinate system around each point that induces the given orientation, or via coordinate patches: two coordinate patches valued in $\mathbb{R}^m$ define the same orientation if their transition map has positive determinant.

NB: a vector subspace is *not* canonically oriented.

For the purposes of intersection theory, for a vector space V, a subspace A, and the quotient $Q = V/A$, an orientation on any two of these will give an orientation on the third. In the case treated in Guillemin and Pollack, when $V = Q \oplus A$, suppose $q_1, \cdots, q_l, a_1, \cdots, a_k$ are ordered bases for Q, A (where the order of summands is specified). Then the basis $a_1, \cdots, a_k, q_1, \cdots, q_l$ differs from the original basis for V by kl many sign changes, so if A and

Q are odd dimensional, then the order of summands matters, but if one of them is even dimensional, then it does not.

Mostly a long series of facts listed here.

If $Q = V/A$ (rather than a complement), then you can lift a basis $q_1, \cdots, q_l$ for Q to V and then run the same construction. If we have orientations on V and A, then each orientation on Q combines (as above) with the orientation on A to give an orientation on V, so we pick the orientation on Q that will agree with the orientation on V under this construction.

Now, if Y, Z are oriented manifolds, Z a submanifold of Y, then $N_Z(Y)$ is oriented, since $N_Z(Y) = TY|_Z/TZ$. Similarly, an orientation in the normal bundle and on Y gives an orientation on Z.

If $f : X \to Y$ with X having boundary, Y boundaryless, f and ∂f transverse to a submanifold Z of Y, then $W := f^{-1}(Z)$ acquires an orientation if each of X, Y, Z has one. Transversality implies that $df_x(T_xX)$ surjects onto T_yY/T_yZ where $y = f(x)$, which induces an isomorphism of N_xW with N_yZ via a complement of the subspace of T_xX which maps into T_yZ.

Thus, $T_xX = N_xW \oplus T_xW$ and $T_yY = N_yZ \oplus T_yZ$; in the first case, orientations on T_xX and N_xW give an orientation on T_xW, and in the second case, orientations on T_yY and T_xZ give an orientation on N_yZ.

Note that our definition of orientation is not well-defined for 0-manifolds, so, for compatibility reasons, we define an orientation on a 0-manifold as a formal symbol $+$ or $-$.

If X is oriented, then ∂X automatically obtains an orientation, via the "outward normal first" convention. Given an orientation on $T_x(\partial X)$, pick an ordered basis $(v_1, \cdots, v_n)$ representing this orientation, and prepend the outward normal vector to ∂X to obtain an ordered basis $(\hat{n}, v_1, \cdots, v_n)$ for T_xX. However, this is the wrong direction: we want an orientation on ∂X from an orientation on X, not the other way around. The resolution is as follows: pick a random orientation on ∂X and see if prepending an outward normal vector coincides with the given orientation on X, and if not, flip the orientation on ∂X.

The key point is that if X is a manifold without boundary, then $I \times X$ is a manifold with boundary (it is important that I is the first, not the second factor); then an orientation on $I \times X$ induces an orientation on $\partial(I \times X)$, i.e, on $\{0\} \times X$ and $\{1\} \times X$. These orientations are "opposite" in the sense that we can identify $\{0\} \times X$ and $\{1\} \times X$ in the obvious way and then compare their orientations, and the normal vectors at 0 and 1 are antiparallel. If we have an orientation on X, say, $v_1, \cdots, v_k$, then the orientation on $\{1\} \times X$ gives rise to a basis $(\rightarrow, v_1, \cdots, v_k)$ and the orientation on $\{0\} \times X$ gives rise to a basis $(\leftarrow, v_1, \cdots, v_k)$. These are opposite orientations.

I have no idea what the above example was meant to illustrate.

An orientation on a 1-manifold M diffeomorphic to $[0,1]$ is given by a smooth choice of nonzero vector at each point of M, so it is determined by its value at any point of M (since the vector at a single point determines a connected component of $\mathbb{R} \setminus \{0\}$). The orientation on ∂M induced by this therefore must have an *inward* normal vector at one of the boundary points. This illustrates the following:

Lemma 2.6.2

If M is any compact oriented 1-manifold, then the sum of the signs of its boundary is 0.

Now suppose $f : X \to Y$ is as above, $Z \subseteq Y$ a boundaryless submanifold, $f, \partial f$ transverse to Z. First, if $\dim X + \dim Z = \dim Y$, then f is an immersion. Chose any $z \in Z$, and $f^{-1}(z)$ is a finite set if X is compact, and orientations on X, Y, Z let you define the *degree* of f as follows: at each $x \in f^{-1}(z)$, you get an orientation on $N_z Z$ via df_x as above, then

$$\deg(f) := \sum_{x \in f^{-1}(z)} \operatorname{sign}(x)$$

where $\operatorname{sign}(x)$ is $+$ if the orientation transverse to Z pulled back to $T_x X$ coincides with the given orientation on X, and $-$ otherwise.

This definition requires X to be compact so the sum is finite, and requires the transversality hypothesis of f to Z. It is not clear that f is invariant under homotopy or that $\deg(f)$ does not depend on the choice of $z \in Z$.