M392C NOTES: MORSE THEORY

ARUN DEBRAY
SEPTEMBER 5, 2018

These notes were taken in UT Austin's M392C (Morse Theory) class in Fall 2018, taught by Dan Freed. I live-TEXed them using `vim`, so there may be typos; please send questions, comments, complaints, and corrections to a.debray@math.utexas.edu. Any mistakes in the notes are my own.

Contents

Lecture 1.

Critical points and critical values: 8/29/18

"The victim was a topologist." (nervous laughter)

In this course, manifolds are smooth unless assumed otherwise.

Morse theory is the study of what critical points of a smooth function can tell you about the topology of its domain manifold.

Definition 1.1. Let $f\colon M \to \mathbb{R}$ be a smooth function.

- A $p \in M$ is a *critical point* if $\mathrm{d}f|_p = 0$.
- A $c \in \mathbb{R}$ is a *critical value* if there's a critical point $p \in M$ with $f(p) = c$.

The set of critical points of f is denoted $\mathrm{Crit}(f)$.

Example 1.2. Consider the standard embedding of a torus T^2 in $\mathbb{R}^3$ and let $f\colon T^2 \to \mathbb{R}$ be the x-coordinate. Then there are four critical points: the minimum and maximum, and two saddle points. These all have dfifferent images, so there are four critical values. ⊲

If M is compact, so is $f(M)$, and therefore f has a maximum and a minimum: at least two critical points. (If M is noncompact, this might not be true: the identity function $\mathbb{R} \to \mathbb{R}$ has no critical points.) In the 1920s, Morse studied how the theory of critical points on M relates to its topology.

Example 1.3. On S^2, there's a function with precisely two critical points (embed $S^2 \subset \mathbb{R}^3$ in the usual way; then f is the z-coordinate). There is no function with fewer, since it must have a minimum and a maximum. ⊲

What about other surfaces? Is there a function on T^2 or $\mathbb{RP}^2$ with only two critical points?

Well, that was a loaded question – we'll prove early on in the course that the answer is no.

Theorem 1.4. *Let M be a compact n-manifold and $f\colon M \to \mathbb{R}$ be a smooth function with exactly two nondegenerate critical points. Then M is homeomorphic to a sphere.*

So, it "is" a sphere. But some things depend on what your definition of "is" is — Milnor constructed *exotic 7-spheres*, which are homeomorphic but not diffeomorphic to the usual S^7, and Kervaire had already produced topological 10-manifolds with no smooth structure. Freedman later constructed topological 4-manifolds with no smooth structure. In lower dimensions there are no issues: smooth structures exist and are unique in the usual sense. In dimension 4, there are some topological manifolds with a countably infinite number of distinct smooth structures. One of the most important open problems in geometric topology is to determine whether there are multiple smooth structures on S^4, and how many there are if so.

Morse studied the critical point theory for the energy functional on the based loop space ΩM of M, which is an infinite-dimensional manifold. This produced results such as the following.

Theorem 1.5 (Morse)**.** *For any $p, q \in S^n$ and any Riemannian metric on S^n, there are infinitely many geodesics from p to q.*

And you can go backwards, using critical points to study the differential topology of ΩM. Bott and Samelson extended this to study the loop spaces of symmetric spaces, and used this to prove a very important theorem.

Theorem 1.6 (Bott periodicity)**.** *Let $\mathrm{U} \coloneqq \varinjlim_{n\to\infty} \mathrm{U}_n$, which is called the infinite unitary group.*[1] *Then*

$$\pi_q \mathrm{U} \cong \begin{cases} \mathbb{Z}, & q \text{ even} \\ 0, & q \text{ odd}. \end{cases}$$

This theorem is at the foundation of a great deal of homotopy theory.

The traditional course in Morse theory (e.g. following Milnor) walks through these in a streamlined way. These days, one uses the critical-point data of a Morse function on M to build a CW structure (which recovers the homotopy theory of M), or better, a handlebody decomposition of M (which gives its smooth structure). We could also study Smale's approach to Morse theory, which has the flavor of dynamical systems, studying gradient flow and the stable and unstable manifolds. This leads to an infinite-dimensional version due to Floer, and its consequences in geometric topology, and to its dual perspective due to Witten, which we probably won't have time to cover. Our course could also get into applications to symplectic and complex geometry.

Milnor's Morse theory book is a classic, and we'll use it at the beginning. There's a more recent book by Nicolescu, which in addition to the standard stuff has a lot of examples and some nonstandard topics; we'll also use it. There will be additional references.

∽ · ∼

Let M be a manifold and $(x^1, \dots, x^n)$ be a local coordinate system (or, we're working on an open subset of affine n-space $\mathbb{A}^n$). One defines the first derivative using coordinates, but then finds that it's intrinsic: if $x = x(y)$ is a change of coordinates (so $x = x(y^1, \dots, y^n)$), then

$$\frac{\partial f}{\partial x^i}\,\mathrm{d}x^i = \frac{\partial f}{\partial y^\alpha}\frac{\partial y^\alpha}{\partial x^i}\frac{\partial x^i}{\partial y^\beta}\,\mathrm{d}y^\beta = \frac{\partial f}{\partial y^\alpha}\,\mathrm{d}y^\alpha, \tag{1.7}$$

and so this is usually just called $\mathrm{d}f$, and can even be defined intrinsically. For critical points we're also interested in second derivatives, but the second derivative isn't usually intrinsic:

$$\frac{\mathrm{d}^2 f}{\mathrm{d}y^2} = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\left(\frac{\mathrm{d}x}{\mathrm{d}y}\right) + \frac{\mathrm{d}f}{\mathrm{d}x}\frac{\mathrm{d}^2 x}{\mathrm{d}y^2}. \tag{1.8}$$

The second term depends on our choice of x, so it's nonintrinsic. In general one needs more data, such as a connection, to define intrinsic higher derivatives. But at a critical point, the second term vanishes, and the second derivative is intrinsic![2]

Definition 1.9. Let $f\colon M \to \mathbb{R}$ and $p \in \mathrm{Crit}(f)$. Then the *Hessian* of f at p is the function $\mathrm{Hess}_p(f)\colon T_pM \times T_pM \to \mathbb{R}$ sending $\xi_1, \xi_2 \mapsto \xi_1(\xi_2 f)(p)$, where we extend ξ_2 to a vector field near p.

Of course, one must check this is independent of the extension. Suppose η is a vector field vanishing at p. Then

$$\xi_1 \cdot (\eta f)(p) = \eta(\xi_1 f)(p) + [\eta, \xi] \cdot f(p) = 0 + 0 = 0, \tag{1.10}$$

so everything is good.

Lemma 1.11. *The Hessian is a symmetric bilinear form.*

[1] The map $\mathrm{U}_n \to \mathrm{U}_{n+1}$ sends $A \mapsto \left(\begin{smallmatrix} A & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

[2] This generalizes: if the first n derivatives vanish at x, the $(n+1)$st derivative is intrinsic.

Proof. Extend both ξ_1 and ξ_2 to vector fields in a neighborhood of p. Then

$$\xi_1 \cdot (\xi_2 f)(p) - \xi_2(\xi_1 f)(p) = [\xi_1, \xi_2] f(p) = 0. \tag{1.12}$$ ⊠

In order to study the Hessian, let's study bilinear forms more generally. Let V be a finite-dimensional real vector space and $B\colon V \times V \to \mathbb{R}$ be a symmetric bilinear form.

Definition 1.13. The *kernel* of B is the set K of $\xi \in V$ with $B(\xi, \eta) = 0$ for all η. If $K = 0$, we say B is *nondegenerate*.

Equivalently, B determines a map $b\colon V \to V^*$ sending $\xi \mapsto (\eta \mapsto B(\xi, \eta))$, and $K = \ker(b)$. Any symmetric bilinear form descends to a nondegeneratr form $\widetilde{B}\colon V/K \times V/K \to \mathbb{R}$.

Example 1.14.

(1) If B is *positive definite*, meaning $B(\xi, \xi) > 0$ for all $\xi \neq 0$, then B is an *inner product.*

(2) On $V = \mathbb{R}^3$, consider the nondegenerate and indefinite form

$$B((\xi^1, \xi^2, \xi^3), (\eta^1, \eta^2, \eta^3)) := \xi^1\eta^1 - \xi^2\eta^2 - \xi^3\eta^3. \tag{1.15}$$

The *null cone*, namely the subspace of ξ with $B(\xi, \xi) = 0$, is a cone opening in the x-direction. We can restrict B to the subspace $\{(x, 0, 0)\}$, where it becomes positive definite, or to the subspace $\{(0, y, z)\}$, where it's negative definite. ◀

However, we can't canonically define anything like *the* maximal positive or negative definite subspace — the only canonical subspace is the kernel. We can fix this by adding more structure.

Lemma 1.16. *Let $N, N' \subset V$ be maximal subspaces of V on which B is negative definite. Then* $\dim N = \dim N'$.

This is called the *index* of B.

Proof. Since N and N' don't intersect K, we can pass to V/K, and therefore assume without loss of generality that B is nondegenerate. Assume $\dim N' < \dim N$; then, $V = N \oplus N^\perp$. Let $\pi\colon V \twoheadrightarrow N$ be a projection onto N, which has kernel $N^\perp$. Then $\pi(N')$ is a proper subspace of N. Let $\eta \in N$ be a nonzero vector with $B(\eta, \pi(N')) = 0$. Then $B(\eta, N') = 0$, and so $B(\xi + \eta, \xi + \eta) < 0$ for all $\xi \in N'$, and therefore N' isn't maximal. ⊠

Applying the same proof to $-N$, there's a maximal dimension of a positive-definite subspace P. So B determines three numbers, $\dim K$ (the *nullity*), $\lambda := \dim N$ (the *index*), and $\rho := \dim P$. This doesn't have a name, but the *signature* is $\rho - \lambda$. In Morse theory we'll be particularly concerned with the index.

Proposition 1.17. *There exists a basis of V, $e_1, \dots, e_\lambda, e_{\lambda+1}, \dots, e_{\lambda+\rho}, e_{\lambda+\rho+1}, \dots, e_n$, such that*

$$B(e_i, e_j) = 0, \qquad i \neq j, B(e_i, e_i) = \begin{cases} -2, & 1 \leq i \leq \lambda, \\ 2, & \lambda + 1 \leq i \leq \lambda + \rho \\ 0, & \textit{otherwise.} \end{cases} \tag{1.18}$$

Proof. We have the kernel $K \subset V$, and can choose a complement V' for it; then $B|_{V'}$ is nondegenerate. Let $N \subset V'$ be a maximal negative definite subspace, and $N^\perp$ be its orthogonal complement with respect to $B|_{V'}$. Then $V = N \oplus N^\perp \oplus K$, and we can choose these bases in each subspace. ⊠

Remark 1.19. If we choose an inner product $\langle -, - \rangle$ on V and define $T\colon V \to V$ by

$$B(\xi_1, \xi_2) = \langle \xi_1, T\xi_2 \rangle \tag{1.20}$$

for all $\xi_1, \xi_2 \in V$, then T is symmetric and therefore diagonalizable. ◀

With the linear algebra interlude over, let's get back to topology. The Hessian is a very useful invariant, e.g. defining the curvature of embedded hypersurfaces in $\mathbb{R}^n$.

Definition 1.21. Let $f\colon M \to \mathbb{R}$ be smooth.

(1) A $p \in \mathrm{Crit}(f)$ is *nondegenerate* if $\mathrm{Hess}_p(f)$ is nondegenerate.

(2) If every critical function is nondegenerate, f is called a *Morse function.*

Example 1.22. For example, on the torus as above, the y-coordinate is a Morse function. But the z-coordinate is not Morse: there's a whole circle of maxima, and another one of minima, and therefore the Hessians on these circles cannot be nondegenerate. ◀

Example 1.23. For another example, consider $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$. This isn't Morse: it has one critical point, which is degenerate. Unlike the previous example, this is a degenerate critical point which is isolated. ◀

Example 1.24. Let V be a finite-dimensional inner product space over $\mathbb{R}$ or $\mathbb{C}$,[3] and let $T\colon V \to V$ be a symmetric linear operator with distinct eigenvalues (i.e. its eigenspaces are one-dimensional). Then $\mathbb{P}(V)$, the set of lines through the origin (i.e. one-dimensional subspaces) in V is a closed manifold. Define $f\colon \mathbb{P}(V) \to \mathbb{R}$ by

$$L \longmapsto \frac{\langle \xi, T\xi\rangle}{\langle \xi, \xi\rangle}, \qquad \xi \in L \setminus 0. \tag{1.25}$$

It's a course exercise to show the critical points of f are the eigenlines of T, and to compute their Hessians and their indices.

It may be useful to know that there's a canonical identification $T_L\mathbb{P}(V) \cong \mathrm{Hom}(L, V/L)$. This also generalizes to Grassmannians. ◀

The next thing we'll study is a canonical local coordinate system around a critical point of a Morse function (the Morse lemma). It's a bit bizarre to build coordinates out of nothing, so we'll start with an arbitrary coordinate system and deform it. We will employ a very general tool to do this, namely flows of vector fields. This may be review if you like differential geometry.

Definition 1.26. Suppose ξ is a vector field on M. A curve $\gamma\colon (a,b) \to M$ is an *integral curve* of ξ if for $t \in (a,b)$, $\dot\gamma(t) = \xi|_{\gamma(t)}$.

Theorem 1.27. *Integral curves exist: for all $p \in M$, there exists an $\varepsilon > 0$ and an integral curve $\gamma\colon (-\varepsilon,\varepsilon) \to M$ for ξ with $\gamma(0) = p$.*

This is a geometric reskinning of existence of solutions to ODEs, as well as smooth dependence on initial data (whose proof is trickier). If you don't know the proof, you should go read it!

We can also allow ξ to depend on t with a trick: consider the vector field $\frac{\partial}{\partial t} + \xi_t$ on $(a,b) \times M$. By the theorem, integral curves exist, and since this vector field projects onto $\frac{\partial}{\partial t}$ on (a,b), the integral curve we get projects onto the integral curve for $\frac{\partial}{\partial t}$. So what we've constructed is exactly the graph of γ. In ODE, this is known as the non-autonomous case.

We'd like to do this everywhere on a manifold at once.

Definition 1.28. A *flow* is a function $\varphi\colon (a,b) \times M \to M$ such that $\varphi(t,\text{–})\colon M \to M$ is a diffeomorphism.

We'd like to say that vector fields give rise to flows. Certainly, we can differentiate flows, to obtain a time-dependent vector field $\frac{\mathrm{d}\varphi}{\mathrm{d}t} = \xi_t$.

Example 1.29. For a quick example of nonexistence of flow for all time, consider $\xi = \frac{\partial}{\partial t}$ on $\mathbb{R} \setminus \{0\}$. You can't flow from a negative number forever, since you'll run into a hole. Now maybe you think this is the problem, but there's not so much difference with just $\mathbb{R}$ and the vector fields $t\frac{\partial}{\partial t}$ or $t^2\frac{\partial}{\partial t}$, where you will reach infinity in finite time. ◀

One of the issues with global-time existence of flow is that the metric might not be complete. But it's not the only obstruction, as we saw above.

Theorem 1.30. *Let ξ_t be a family of vector fields for $t \in (t_-, t_+)$, where $t_- < 0$ and $t_+ > 0$.*

1. *Given a $p \in M$, there are neighborhoods of p $U' \subset U$ and an $\varepsilon > 0$ such that there's a flow $\varphi\colon (-\varepsilon,\varepsilon) \times U' \to U$ with $\frac{\mathrm{d}\varphi}{\mathrm{d}t} = \xi_t$.*
2. *If M has a complete Riemannian metric and there's a $C > 0$ in which $|\xi_t| \le C$, then the flow is global: we can replace $(-\varepsilon,\varepsilon)$ with (t_-, t_+).*

A compact manifold is complete in any Riemannian metric, so for ξ arbitrary, global flows exist.

[3] With a little more work, we can make this work over the quaternions.

Remark 1.31. If ξ is *static*, i.e. independent of t, then $t \mapsto \varphi_t$ is a *one-parameter group*, i.e. $\varphi_{t_1+t_2} = \varphi_{t_1} \circ \varphi_{t_2}$. ◀

Example 1.32. Let M be a Riemannian manifold and $f\colon M \to \mathbb{R}$ be smooth. Define its *gradient vector field* by

$$\mathrm{d}f|_p(\eta) \coloneqq \langle \eta, \mathrm{grad}_p f\rangle \tag{1.33}$$

for all $\eta \in T_pM$. ◀

Let's (try to) flow by $-\operatorname{grad} f$.

Definition 1.34. Let $\omega \in \Omega^*(M)$ and ξ be a vector field with local flow φ generated by ξ. The *Lie derivative* is

$$\mathcal{L}_\xi \omega \coloneqq \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \varphi_t^* \omega,$$

which is also a differential form, homogeneous of degree k if ω is.

Theorem 1.35 (H. Cartan)**.** *$\mathcal{L}_\xi\omega = (\mathrm{d}\iota_\xi + \iota_\xi \mathrm{d})\omega$. Here ι_ξ denotes contracting with ξ.*

With this in our pockets, let's turn to the Morse lemma.

Lemma 1.36 (Morse lemma)**.** *Let $f\colon M \to \mathbb{R}$ be smooth and p be a nondegenerate critical point of f of index λ. Then there exist local coordinates $x^1, \dots, x^n$ near p with $x^i(p) = 0$ and*

$$f(x^1, \dots, x^n) = f(p) - ((x^1)^2 + \cdots (x^\lambda)^2) + ((x^{\lambda+1})^2 + \cdots + (x^n)^2).$$

The proof employs a technique of Moser. Moser used this to provide a nice proof of Darboux's theorem, that symplectic manifolds all look like affine space locally.

Lemma 1.37. *Let $U \subset \mathbb{R}^n$ be a star-shaped open set with respect to the origin and $g\colon U \to \mathbb{R}$ be such that $g(0) = 0$. Then there exist $g_i\colon U \to \mathbb{R}$ with $g(x) = x^i g_i(x)$.*

Proof. Well, just let

$$g_i(x) = \int_0^1 \frac{\partial g}{\partial x^i}(tx)\,\mathrm{d}t. \tag{1.38}$$ ⊠

Proof of Lemma 1.36. Choose local coordinates $x^1, \dots, x^n$ such that

$$\frac{1}{2}\mathrm{Hess}_p(f) = \left(-\left(\mathrm{d}x^1 \otimes \mathrm{d}x^1 + \cdots + \mathrm{d}x^\lambda \otimes \mathrm{d}x^\lambda\right) + \left(\mathrm{d}x^{\lambda+1} \otimes \mathrm{d}x^{\lambda+1} + \cdots \mathrm{d}x^n \otimes \mathrm{d}x^n\right)\right)_p. \tag{1.39}$$

Since we're only asking for this at p, we can start with any coordinate system and then apply Lemma 1.37. Set

$$h(x) \coloneqq f(p) - ((x^1)^2 + \cdots (x^\lambda)^2) + ((x^{\lambda+1})^2 + \cdots + (x^n)^2) - f(x). \tag{1.40}$$

We're hoping for this to be zero. Also set

$$\alpha_t \coloneqq (1-t)\Bigg(\underbrace{-\left(x^1\,\mathrm{d}x^1 + \cdots + x^\lambda\,\mathrm{d}x^\lambda\right) + \left(x^{\lambda+1}\,\mathrm{d}x^{\lambda+1} + \cdots x^n\,\mathrm{d}x^n\right)}_{\alpha_0}\Bigg) + t\,\mathrm{d}f, \tag{1.41}$$

for $t \in [0,1]$. We claim that in a neighborhood of $x = 0$, we can find a vector field ξ_t such that $\iota_{\xi_t}\alpha_t = h$; in particular, h does not depend on t; and such that $\xi_t(p = 0) = 0$. We'll then use this to move the coordinates; at p everything looks right, so we'll use this to move the coordinates elsewhere.

Assuming the claim, let φ_t be the local flow generated by ξ_t, which exists at least in a neighborhood of U. Then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^*\alpha_t &= \varphi_t^*\mathcal{L}_{\xi_t}\alpha_t + \varphi_t^*\left(\frac{\mathrm{d}}{\mathrm{d}t}\alpha_t\right)\\ &= \varphi_t^*(\mathrm{d}\iota_{\xi_t}\alpha_t + \iota_{\xi_t}\,\mathrm{d}\alpha_t - \mathrm{d}h).\end{aligned}$$

Since α_t is exact,

$$= \varphi_t^*(\varphi_t^*\,\mathrm{d}(\iota_{\xi_t}\alpha_t - h)) = 0.$$

Therefore $\varphi_1^*(\mathrm{d}f) = \varphi_1^*\alpha_1 = \varphi_0^*\alpha_0 = \alpha_0$. In particular, φ_1 is a local diffeomorphism fixing $p = 0$, and it pulls $\mathrm{d}f$ back to d of something quadratic. Therefore $\varphi_1^* f$ is quadratic, and has the desired form.

Now we need to prove the claim. Observe $\alpha_t(0) = 0$ and $h(0) = 0$. Then write

$$\alpha_t(x) = A_{ij}(t,x)x^j \,\mathrm{d}x^i$$
$$h(x) = h_j(x)x^j$$
$$\xi_t = \xi^k(t,x)\frac{\partial}{\partial x^k},$$

so $\iota_{\xi_t}\alpha_t h$ is equivalent to

$$A_{ij}(t,x)x^j\xi^i(t,x) = h_j(x)x^j, \tag{1.42}$$

which is implied by

$$A_{ij}(t,x)\xi^j(t,x) = h_j(x). \tag{1.43}$$

Since $(A_{ij}(t,0))$ is invertible, we can solve this in some neighborhood of $x = 0$ uniform in t (it remains invertible in that neighborhood). ⊠

Lecture 2.

: 9/5/18

Last time, we proved the Morse lemma: if $f\colon M \to \mathbb{R}$ is a smooth function and $p \in M$ is a nondegenerate critical point, then there are local coordinates $x^1, \dots, x^n$ with $x(p) = 0$ and

$$f(x) = f(p) - ((x^1)^2 + \cdots + (x^\lambda))^2 + ((x^{\lambda+1})^2 + \cdots + (x^n)^2). \tag{2.1}$$

In this case we can define the Hessian; λ is its index, which is the maximal dimension d such that there's a d-dimensional subspace $N \subset T_pM$ on which the Hessian is negative definite.

Corollary 2.2. *A nondegenerate critical point is isolated.*

Recall that a smooth function is called Morse if all of its critical points are nondegenerate.

Corollary 2.3. *If f is a Morse function, then* $\mathrm{Crit}(f) \subset M$ *is discrete. If M is compact, then* $\mathrm{Crit}(f)$ *is finite.*

So Morse functions are really nice. But they're nontheless generic.

Theorem 2.4. *Let M be a smooth manifold.*

(1) *M admits a Morse function; in fact, Morse functions are dense in $C^\infty(M)$.*
(2) *M admits a proper Morse function.*[4]

To make precise the notion of density of Morse functions, we need to specify a topology on $C^\infty(M)$; that can be done, but we're not going to do it here. Proofs will be given in the next section.

Definition 2.5. Let $f\colon M \to \mathbb{R}$ be smooth and $a \in \mathbb{R}$. Then define $M^a \coloneqq f^{-1}((\infty, a])$, which is called a *sublevel set.*

See Figure 1 for examples of sublevel sets. Sublevel sets of M define a filtration of M indexed by $\mathbb{R}$.

The second fundamental theorem of Morse theory, which we'll do next time, is about handles and handlebodies, and that when you cross a critical point, the diffeomorphism type of the sublevel set changes precisely by adding a handle.

We probably should have already mentioned an important theorem from differential topology.

Theorem 2.6. *If a is a regular value, $f^{-1}(a) \subset M$ is a manifold, and M^a is a manifold with $\partial M^a = f^{-1}(a)$.*

Since a point is compact, and an interval is compact, choosing proper Morse functions allows us to get compact level sets for $f^{-1}(a)$. Moreover, the preimage of $[a, b]$ is a compact manifold with boundary $f^{-1}(a) \amalg f^{-1}(b)$ (here a and b should be regular values), i.e. a *bordism* from $f^{-1}(a)$ to $f^{-1}(b)$.

This perspective, involving handles and differential topology, is geometric, and is due to Smale in the 1960s or so. But there's another, homotopical approach, where one uses a Morse function to define a CW structure. This not only shows that all manifolds have CW structures, which is nice, but also is a gateway

[4] Recall that a *proper map* is a map $f\colon X \to Y$ such that the preimage of any compact set in Y is compact.

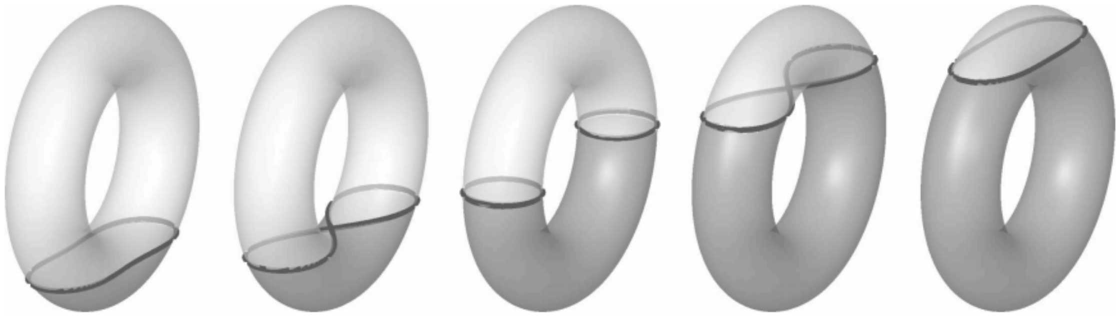

FIGURE 1. Sublevel sets for the standard height function on a torus. We can also get the empty 2-manifold $\varnothing^2$ for sublevel sets for a below the minimum, and T^2 for sublevel sets for a above the maximum.

to good calculations of homology and cohomology. The idea is to think of handle attachment by collapsing the "irrelevant" dimensions, so that instead of attaching a handle, you can attach a k-cell (depending on the index), and so on.

But the simplest question you can ask is: if a and b are regular values with no critical values in $[a, b]$, how do M^a and M^b differ? The answer is, more or less, they don't.

Theorem 2.7. *Let $f\colon M \to \mathbb{R}$ be a smooth function and $a < b$ such that every $y \in [a, b]$ is regular for f. Assume $f^{-1}([a, b])$ is compact. Then,*

(1) *M^a and M^b are diffeomorphic.*
(2) *M^a is a deformation retract of M^b: in particular, inclusion $M^a \hookrightarrow M^b$ is a homotopy equivalence.*[5]

Again, we have a smooth manifold statement and a homotopical statement.

Proof. First, introduce a Riemannian metric on M. This additional data is necessary so that we can measure things (such as lengths and angles and so on). Riemannian metrics exist on all smooth manifolds; let's talk about why. An inner product on V is a positive definite bilinear pairing; these form a convex space in $\mathrm{Sym}^2 V^*$. In fact, it's a convex cone, because if $a > 0$ and g is an inner product, ag is also an inner product.

Now let M be a smooth manifold and $\mathfrak{U}$ be an atlas. Each open $U \in \mathfrak{U}$ is diffeomorphic to affine space, so we can introduce the standard Euclidean metric on it. We can then use a partition of unity to sum these metrics into a global one: because inner products form a convex space and the partition of unity is a locally finite convex combination, this works.

From the Riemannian metric, we obtain a vector field $\operatorname{grad} f$ with $\operatorname{grad}_p f = 0$ iff f is a critical point. This flows in the direction of increasing height; we want to push M^b down to M^a, so we'll flow along $-\operatorname{grad} f$. But we don't want to flow too much beyond that, so let's introduce a cutoff function $\rho\colon M \to \mathbb{R}^{\geq 0}$ such that

$$\rho(x) = \begin{cases} \dfrac{1}{\|\operatorname{grad} f\|^2}, & x \in f^{-1}([a, b]) \\ 0 & \text{outside } U, \end{cases} \tag{2.8}$$

where U is an open neighborhood of $\overline{f^{-1}([a, b])}$ whose closure is compact.

Set $\xi := -\rho \operatorname{grad} f$. Then ξ generates a global flow $\varphi_t\colon M \to M$. If $p \in M$,

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\varphi_t(p)) = \left\langle \operatorname{grad} f, \frac{\mathrm{d}\varphi_t(p)}{\mathrm{d}t} \right\rangle = -\rho \|\operatorname{grad} f\|^2. \tag{2.9}$$

In $f^{-1}([a, vb])$ this is just -1, and outside of U, this is the identity. In particular, $\varphi_{b-a}\colon M^b \to M^a$ is a diffeomorphism: its inverse is φ_{a-b}.

[5]Recall that given an inclusion $i\colon A \hookrightarrow X$, a map $r\colon X \to A$ is a *deformation retraction* if theres a homotopy $h\colon [0, 1] \times X \to X$ such that $h_0 = \mathrm{id}_X$ and $h_1 = i \circ r$, and such that $r \circ i = \mathrm{id}_A$.

For the second part, we can define the requisite homotopy $h\colon [0,1]\times M^b \to M^b$ by

$$h(t,p) := \begin{cases} p, & p \in M^a \\ \varphi_{t(f(p)-a)}, & p \in f^{-1}([a,b]). \end{cases} \tag{2.10}$$

⊠

Exercise 2.11. Let $M=\mathbb{R}$ and $f(x)=(\log x)^2$. Make the theorem explicit in this case.

Let $M=\mathrm{GL}_n(\mathbb{R})$ (resp., $\mathrm{GL}_n(\mathbb{C})$). Show that M deformation retracts onto O_n (resp. U_n). Make the theorem explicit for $f(A)=\mathrm{tr}(\log(A^*A))$.

∽ · ∼

Now we'll do a short review of some Riemannian geometry. Let A be an affine space modeled on a vector space V and $\eta\colon A\to V$ be a smooth function to some vector space. We can define the directional derivative in the direction of an $\eta \in V$ by

$$D_\xi \eta := \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \eta(p+t\xi). \tag{2.12}$$

If we're on a smooth manifold M, though, we can't make sense of $p+t\xi$. Instead, we'd like to choose a curve $\gamma\colon(-\varepsilon,\varepsilon)\to M$ with $\gamma(0)=p$ and $\dot\gamma(0)=\xi$, and use this to define the directional derivative. However, we then have a problem: as t varies, $\eta(\gamma(t))$ lives in different vector spaces, so we can't define their difference, which is important for taking the derivative. So we need to introduce more structure in order to define directional derivatives.

Definition 2.13. Let M be a smooth manifold. A *covariant derivative* on $TM\to M$, also called a *linear connection*, i a bilinear map $\nabla\colon \mathcal{X}(M)\times\mathcal{X}(M)\to\mathcal{X}(M)$ such that

1. (linearity over functions) if $f\in C^\infty(M)$, then $\nabla_{f\xi}\eta = f\nabla_\xi\eta$.
2. (Leibniz rule) if $g\in C^\infty(M)$, then $\nabla_\xi(g\eta)=(\xi\cdot g)\eta+g\nabla_\xi\eta$.

The first condition implies $\nabla_\xi\eta|_p$ depends only on $\xi|_p$, which expresses tensoriality.

Definition 2.14. ∇ is *torsion-free* if

$$\nabla_X Y - \nabla_Y X = [X,Y]. \tag{2.15}$$

If $\langle -,-\rangle$ is a Riemannian metric on M, then ∇ is *orthogonal* with repsect to g if

$$X\langle Y,Z\rangle = \langle\nabla_X Y, Z\rangle + \langle Y,\nabla_X Z\rangle. \tag{2.16}$$

Remarkably, these exist and are unique! This is a foundational theorem in Riemannian geometry.

Theorem 2.17. *For any Riemannian manifold (M,g), there's a unique torsion-free orthogonal connection on TM.*

This connection is called the *Levi-Civita connection.* It turns out this can be explicitly constructed with a straightedge and compass, though it would take a while.

Exercise 2.18. Prove Theorem 2.17 by explicitly writing a formula for $\langle\nabla_X Y, Z\rangle$ and using the torsion-free and orthogonal conditions to expand it out, hence defining $\nabla_X Y$.

There are lots of different ways to say the proof, but it's really a formula proof, and no synthetic proof exists. There are special classes of manifolds (e.g. Kähler manifolds) on which a synthetic proof exists.

If (M,g) is a Riemannian manifold and $N\hookrightarrow M$ is an immersed submanifold, then it inherits a Riemannian metric: a subspace of an inner product space gains an inner product by restriction, and doing this for all $T_pN\subset T_pM$ defines the metric on N. Moreover, if $X,Y\in\mathcal{X}(M)$ and $p\in N$, then $\nabla^M_X Y|_p\in T_pM$ need not be in T_pN. But $T_pM=T_pN\oplus\nu_p$, where ν_p is the normal bundle; to choose this splitting we needed to use the metric.

Using this, let $I\!I(X,Y)$ denote the component of $\nabla^M_X Y|_p$ in ν_p, where ∇^M denotes the Levi-Civita connection on M.

Lemma 2.19. *$I\!I(X,Y)$ is linear over functions in both of its arguments, and $I\!I(X,Y)=I\!I(Y,X)$; in particular, it's a symmetric bilinear form.*

The proof is a calculation. $II(X,Y)$ is called the *second fundamental form*.[6] Moreover, it expresses the difference between ∇^M and ∇^N.

Lemma 2.20. *The tangential component of $\nabla^M_X Y$ is $\nabla^N_X Y$.*

If Z is a normal vector field to N in M, we can define $II^Z(X,Y) \coloneqq \langle II(X,Y), Z\rangle$. Then II^Z is a symmetric bilinear form $T_pM \times T_pM \to \mathbb{R}$, and we know what the invariants of symmetric bilinear forms are. We can also define $S\colon T_pM \to T_pM$ by $\langle S(X), Y\rangle = II(X,Y)$. This is symmetric, so we can diagonalize, and therefore recover an orthonornal basis $e_1, \dots, e_m$ of T_pM (up to units and reordering) such that $Se_j = \lambda_j e_j$ for some $\lambda_j \in \mathbb{R}$. These λ_j are expressing the amount of curvature in various directions — unless they coincide (this is called an *umbilic point*). S is called the *shape operator*, as it determines the local shape of the surface.

Lecture 3.

: 9/5/18

[6]The "first fundamental form" is another word for the inner product on T_pN.