MORSE HOMOLOGY

STUDENT GEOMETRY AND TOPOLOGY SEMINAR
FEBRUARY 26, 2015
MORGAN WEILER

Contents

Manifolds are closed. Fields are $\mathbb{Z}/2$.

I've used [1], [2], and [3], as well as things I've absorbed from the world. In particular I first heard about local slices from Katrin Wehrheim's 276 class (Regularization of Moduli Spaces) at UC Berkeley in Spring 2014.

1. Morse Functions

It's likely that you've seen the deRham complex: some algebra with (smooth) functions provides information about the topology of a manifold. What else can functions tell us about the topology of a manifold? In particular, it would be great to use functions which are somehow "abundant" or "standard."

Enter Morse functions. They are abundant and standard and provide handle and cell decompositions for manifolds, a hint at their topological power. They also provide a method for computing homology. The intuition from Morse homology in finite dimensions is the foundation for our understanding of many useful and beautiful modern invariants based on PDE solution spaces.

Let M be a (finite-dimensional, closed, smooth) manifold and $f \in C^\infty M$. Let $p \in M$ be a critical point of f, that is, $df_p = 0$. At a critical point we can define...

Definition 1. *The* ***Hessian*** *$H(f,p) : T_pM \to T_p^*M$ is the derivative of df at p.*

To understand $H(f,p)$ in a coordinate-free way, recall that $df : M \to T^*M$, so we have

$$\begin{tikzcd}
T_pM \arrow[r, "d(df)_p"] \arrow[dr, "H(f,p)"'] & T_{(p,0)}T^*M \arrow[r, leftrightarrow, "="] & T_pM \oplus T_p^*M \arrow[dl, "\pi_2"'] \\
& T_p^*M &
\end{tikzcd}$$

where $df_p = (p,0)$ because p is a critical point, and TT^*M splits canonically only at its zero section (it has a horizontal direction canonically only at its zero section).

Alternately, in local coordinates, $\mathrm{Hom}(T_pM, T_p^*M) = T_p^*M \otimes T_p^*M = (T_pM \otimes T_pM)^*$, so it makes sense to think of $H(f,p)$ as a bilinear form. We have $H(f,p) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)$ because $H(f,p)$ measures how df_p varies in the T_p^*M direction as p varies.

Definition 2. *The critical point p is* ***nondegenerate*** *if $H(f,p)$ is nondegenerate (in coordinates: if $H(f,p)$ has all nonzero eigenvalues).*

Note that nondegeneracy at p precisely means that df is transverse to the zero section of T^*M over M at p. In particular, nondegenerate critical points are isolated.

To a nondegenerate critical point we can associate a number.

Definition 3. *The* ***index*** *of a critical point p of f, denoted ind(p), is the dimension of the negative eigenspace of $H(f,p)$.*

Definition 4. *$f \in C^\infty M$ is a* ***Morse function*** *if all its critical points are nondegenerate.*

The index provides us with a grading on the chain complex generated by the critical points of a Morse function.

1.0.1. *Example.* Let $f : S^2 \to \mathbb{R}$ be the height function for the embedding of S^2 as the unit ball in $\mathbb{R}^3$, so in the coordinates (x,y), away from the $z=0$ plane,

$$f(x,y) = \pm\sqrt{1 - x^2 - y^2}$$

f is a Morse function, since at $p = (0,0,1)$ (and similarly at $(0,0,-1)$),

$$H(f,p) = \begin{pmatrix} \frac{-1+y^2}{1-x^2-y^2} & -\frac{xy}{1-x^2-y^2} \\ -\frac{xy}{1-x^2-y^2} & \frac{-1+x^2}{1-x^2-y^2} \end{pmatrix}(0,0) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

Morse Lemma. Morse functions are nice because there is always a chart around each critical point in which they equal their second derivatives (in such a chart one can always pick coordinates so that the coordinates are the eigenvectors of the Hessian):

Lemma 1. *Let p be a nondegenerate critical point of $f : M \to \mathbb{R}$. Then there is a neighborhood U of p and a diffeomorphism $\phi : (U, c) \to (\mathbb{R}^n, 0)$ such that*

$$f \circ \phi^{-1}(x_1, \ldots, x_n) = f(c) - \sum_{j=1}^{i} x_j^2 + \sum_{j=i+1}^{n} x_j^2$$

The number i appearing in the lemma is the index of p.

People like using this because it is possible to force all the PDEs to come to agree with the PDEs defining the gradient vector field of f in these coordinates (with the standard Euclidean metric). It's not necessary, however – the coordinates are useful mainly to make what follows more concrete.

1.0.2. *Example.* Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $(x, y) \mapsto \cos(x) + \cos(y)$. f descends to a function on $\mathbb{T}^2 = \mathbb{R}^2/(2\pi\mathbb{Z})^2$. Its critical points are $(0,0), (0,\pi), (\pi,0)$, and (π,π), mapping to $2, 0, 0$, and -2, with indices $2, 1, 1$, and 0, respectively.

1.0.3. *Example.* Stand a round torus on its side (so that there is a rotational symmetry about the z axis) and consider the height function.

1.0.4. *Example.* A dimpled sphere, with the height function. In general, the height function will work on any manifold embedded in Euclidean space, so long as it only intersects horizontal hyperplanes transversely.

Existence.

Proposition 1. *Let $M \subset \mathbb{R}^n$ be a submanifold. For almost every $q \in \mathbb{R}^n$, $f_q : M \to \mathbb{R}$ given by $x \mapsto ||x - q||^2$ is a Morse function.*

Use Sard's theorem. In fact, Sard's theorem is exploited systematically throughout the proofs of almost everything in the setup of Morse homology.

Genericness.

Theorem 1. *The set of Morse functions on a manifold M is dense open in $C^\infty M$.*

That is, for any $k \in \mathbb{Z}_{>0}$, a given function f can be uniformly approximated with all its derivatives on any compact set by a Morse function. The proof constructs a Morse function from a given function.

Topology. The sublevel sets $f^{-1}((-\infty, c])$ are diffeomorphic between critical values and change by handle attachments as c passes a critical value.

More precisely, choose a Riemannian metric g on M. Then use the flow of $\nabla_g f$ to retract $f^{-1}((-\infty, b])$ onto $f^{-1}((-\infty, a])$ for $a < b$ with no critical values in $[a, b]$.

For c a critical value, $f^{-1}((-\infty, c + \epsilon])$ differs from $f^{-1}((-\infty, c + \epsilon])$ by a set of handles $D^i \times D^{n-i}$, attached to $f^{-1}((-\infty, c - \epsilon])$ along $\partial D^i \times D^{n-i}$. As an example, think of f on S^3 with three critical points: you build S^3 as a ball, a torus, a ball, then S^3.

From this handle decomposition we get the Morse inequalities $\#\{p \in \mathrm{Crit}(f) \mid \mathrm{ind}(p) = i\} \geq b_i$.

The "handle attachment" view of the topology visible from a Morse function generalizes to infinite-dimensional cases when the gradient flow equation is parabolic. For example, critical points of the energy functional on the free loop space of a Riemannian manifold are closed geodesics. In this case it is possible to use the handle attachment perspective.

Apparently, one can understand how the topology changes when a critical value is passed. In particular it is possible to make sense of the "index."

What I will now describe is the finite-dimensional analogue of the approach used when the gradient flow equation is elliptic. Examples are the symplectic action functional on either the free loop space or the space of paths with Lagrangian boundary conditions, or the Chern-Simons functional on the space of flat connections modulo gauge equivalence.

In these cases we don't know what the resulting homology tells us about the infinite-dimensional manifold. However, the infinite-dimensional manifolds are obtained from finite-dimensional manifolds, and the homologies we get tell us something about the finite-dimensional manifolds. The main difficulty is that we cannot make sense of the index, and so we must try instead to understand an "index difference" between critical points. The index difference is a property of the flow lines between the critical points, and it is possible to put a compatible grading on the chain complex. But this is getting ahead of ourselves.

2. The Morse Chain Complex

Let $\mathrm{Crit}(f)$ be the set of critical points of f.

Generators.

$$C_i^{\text{Morse}}(f,g) = \sum_{\substack{p \in \text{Crit}(f) \\ \text{ind}(p)=i}} \mathbb{Z}\langle p \rangle$$

Differential. The differential of a critical point p is a sum over all critical points q of index one lower than p's, with weights given by a count (mod 2) of negative gradient flow lines from p to q.

But how do we know that this count is possible, that is, that the negative gradient flow lines from p of index i to q of index $i-1$ always form a finite set? Also note that I have swept under the rug the fact that the gradient depends on a choice of metric, so we need to know that it is possible to choose a metric so that there will be nice (existing, converging to the critical points, stable under perturbations of f and g) negative gradient flow lines between critical points.

Stable and Unstable Manifolds. Let $\Gamma(t,x)$ denote the flow of $-\nabla_g f$, where g is some Riemannian metric on M.

Definition 5. *The* ***stable/ascending manifold*** *of a critical point p is*

$$A(p) = \left\{ x \in M \mid \lim_{t \to +\infty} \Gamma(t,x) = p \right\}$$

Similarly, the ***unstable/descending manifold*** *is*

$$D(p) = \left\{ x \in M \mid \lim_{t \to -\infty} \Gamma(t,x) = p \right\}$$

For nondegenerate p, $A(p)$ is an embedded open $n - \text{ind}(p)$ dimensional disk and $D(p)$ is an embedded open $\text{ind}(p)$ dimensional disk. Essentially it is because their tangent spaces at p have to be the positive and negative eigenspaces of the Hessian, respectively, and everything outside a small neighborhood of p where they can be modeled on these eigenspaces is traced out by the flow.

Morse-Smale Pairs.

Definition 6. *A pair (f,g) is* ***Morse-Smale*** *if for all critical points $p \neq q$, $D(p) \pitchfork A(q)$.*

Morse-Smale pairs "exist and are generic," I believe in the sense that given a Morse function it is possible to pick a metric so that (f,g) is Morse-Smale, and furthermore the choice of such g is open and dense. (Rather than in the sense that the choice of *pairs* (f,g) is open and dense in the choice of all pairs.)

2.0.5. *Example.* The pair in Example 1.0.3 is not Morse-Smale.

The Space of Flow Lines. Let $p, q \in \mathrm{Crit}(f)$.

Definition 7. *A **flow line** of $-\nabla_g f$ from p to q is a map $\gamma : \mathbb{R} \to M$ such that*

$$\dot{\gamma}(t) = -\nabla_g f(t) \text{ and } \lim_{t\to-\infty} \gamma(t) = p \text{ and } \lim_{t\to+\infty} \gamma(t) = q$$

$\mathbb{R}$ acts on flow lines by translation $c \cdot \gamma = \gamma(c + \cdot)$ (there are no other reparameterizations since the first derivative is fixed).

Definition 8. *Let*

$$\widehat{\mathcal{M}}_{(f,g)}(p, q) = \{\gamma : \mathbb{R} \to M \mid \gamma \text{ is a flow line from } p \text{ to } q\}$$

*The **moduli space of flow lines** from p to q is*

$$\mathcal{M}_{(f,g)}(p, q) = \widehat{\mathcal{M}}_{(f,g)}(p, q)/\mathbb{R}$$

Notice that the "evaluation" of $\widehat{\mathcal{M}}$ minus the endpoints p and q is $D(p) \cap A(q)$, that is,

$$\left\{x \in M \mid x = \gamma(t) \text{ for } \gamma \in \widehat{\mathcal{M}}\right\} \Big\backslash \{p, q\} = D(p) \cap A(q)$$

Index Difference One Flows. Assume $\mathrm{ind}(q) = \mathrm{ind}(p) - 1$. Since $D(p)$ and $A(q)$ intersect transversely, and they are of dimensions $\mathrm{ind}(p)$ and $n - \mathrm{ind}(p) + 1$ respectively, their intersection must be a manifold of dimension 1. If $x \in D(p) \cap A(q)$ then the forward and backward flow of $-\nabla_g f$ is also in $D(p) \cap A(q)$, so $D(p) \cap A(q)$ must precisely consist of a disjoint union of images of flow lines. Modding out by $\mathbb{R}$ translation gives a 0-manifold.

It remains to show that the 0-manifold is compact, in which case it will consist of a finite set of points, which can be counted for the coefficient of q in dp. Instead of modding by $\mathbb{R}$, simply consider the set of points $\gamma(0)$ for a representative of each equivalence class in $\mathcal{M}$. So we have a bijection between a set of points in M and elements of $\mathcal{M}$.

Because M is compact, this set is finite (good) or has an accumulation point (bad). In the latter case, the images of the flow lines must accumulate. If the accumulation point is in $D(p) \cap A(q)$ then $D(p) \cap A(q)$ is not a 1-manifold near the accumulation point. If it is not in $D(p) \cap A(q)$ then consider the flow of $-\nabla_g f$ through that point; it has to flow from p to q since one point in it is arbitrarily close to gradient flows between p and q.

3. It Really is a Chain Complex, or, Index Difference Two Flows

In other words, $d^2 = 0$ so we can take homology.

Notice now that we do *not* have $\widehat{\mathcal{M}} \cong D(p) \cap A(p)$. To see this, consider the height function on a dimpled sphere, Example 1.0.4. The moduli space of flow lines between either of the index two critical points and the index zero critical point should be an interval, but $D(p) \cap A(q)$ is a disk. In particular, the images of the two endpoints of $\mathcal{M}$ intersect nontrivially, but they still need to be thought of as different elements of the moduli space.

So we'll have to do something else.

d^2p is a count of critical points r whose indices are two lower than that of p. The weights are given by a count of "broken trajectories," those which pass first through a critical point q with $\mathrm{ind}(q) = \mathrm{ind}(p) - 1$, then secondly flow from q to r with $\mathrm{ind}(r) = \mathrm{ind}(q) - 1$. If we can show that there are always an even number of such trajectories then, mod 2, the count is zero.

The method is to find a compact 1-manifold $\overline{\mathcal{M}_{(f,g)}(p,r)}$ for $\mathrm{ind}(r) = \mathrm{ind}(p) - 2$, containing $\mathcal{M}_{(f,g)}(p,r)$ (in this case as its interior, though in infinite-dimensional cases $\mathcal{M}$ may not be dense in $\overline{\mathcal{M}}$), whose boundary is the set of broken trajectories. Since all compact 1-manifolds have an even number of boundary components, there are an even number of broken trajectories.

Manifold, Dimension. Local slices give maps between $\mathcal{M}_{(f,g)}(p,r)$ and $D(p) \cap A(r)/\mathbb{R}$ on interiors, providing the manifold structure of the moduli space and the dimension of two by transversality of the intersection. This gets harder in the infinite-dimensional setting because we have to translate all of these theorems into theorems about maps with Fredholm derivatives.

Compactness. Done by showing that a sequence of flow lines converges in $\mathcal{M}_{(f,g)}(p,r)$ to either another in $\mathcal{M}_{(f,g)}(p,r)$ or to a broken flow line in $\mathcal{M}_{(f,g)}(p,q) \times \mathcal{M}_{(f,g)}(q,r)$. This is nontrivial, and different references have different was to do it, each of which generalize or not to infinite dimensions.

The Boundary is the Brokens. Otherwise known as gluing. To include the brokens in the boundary, take elements of $\widehat{\mathcal{M}}_{(f,g)}(p,q)$ and $\widehat{\mathcal{M}}_{(f,g)}(q,r)$ and glue them together to get an element of $\widehat{\mathcal{M}}_{(f,g)}(p,r)$.

The process goes like so: use cutoff functions and a single gluing parameter (so that the dimension is one) to get a smooth curve which is close to being a gradient flow. Use a Newton iteration on the linearization of the gradient flow to make it actually a gradient flow.

4. Independence of (f,g)

Directly. Choose a homotopy between (f_0, g_0) and (f_1, g_1). There is an open-dense set of homotopies for which the following construction will work, but for each t in the homotopy it is not necessarily the case that (f_t, g_t) is Morse-Smale. (It is true, however, that at only finitely many times is it possible for (f_t, g_t) to *not* be Morse-Smale.)

Choose a smooth function $\beta : [0,1] \to \mathbb{R}$ which looks like an upside-down parabola, zero at endpoints, nonzero derivative at endpoints. Let $V(s,x) = \beta(s)\partial_t - \nabla_{g_t} f_t$ and define a map on the chains by setting the coefficient of q_1 with index the same as the index of p_0 in the image of p_0 to be the count of flow lines of V between p_0 and q_1. Showing this gives an isomorphism on homology is similar to showing that $d^2 = 0$.

Morse Homology is Cellular Homology. Put a cell structure on M by making the descending manifolds the cells (everything is in a descending manifold of some critical point since M is compact so f has a maximum value). From (generalized to higher amounts of breaking) compactness, the boundary of a descending manifold is either a critical point of lower index or a broken flow line made up of flow lines in the descending manifolds of lower index critical points.

In this formulation the Morse and cellular homologies are the same. The map on chains is the map between a critical point and its descending manifold. The map on the differential comes from the fact that the boundary of an i cell in the $i-1$ skeleton is given by the images of the twice-broken flow lines. It's important here to realize that the distinction between boundary components of $\overline{\mathcal{M}_{(f,g)}(p,r)}$ is determined entirely by $\mathcal{M}_{(f,g)}(q,r)$ for $\mathrm{ind}(q) = \mathrm{ind}(p) - 1$.

We again get independence of the choice of (f,g) from this equivalence. We also get the Morse inequalities again.

(It is possible to do all of this on manifolds with boundary rather than closed manifolds. The homology depends on the choice splitting of ∂X into ingoing and outgoing ends of a cobordism, because that choice influences the choice of Morse function (up to equivalence depending only on the splitting of ∂X).)

References

[1] Audin, Michéle and Mihai Damian, trans. Reinie Erné. *Morse Theory and Floer Homology*, Springer-Verlag, London, 2014.

[2] Hutchings, Michael. "Lecture notes on Morse homology (with an eye towards Floer theory and pseudoholomorphic curves)." Fall 2002, UC Berkeley.

[3] Hutchings, Michael. My notes from Math 242, Spring 2014, UC Berkeley.