Lecture notes on Morse homology (with an eye towards Floer theory and pseudoholomorphic curves)

Michael Hutchings

December 15, 2002

Abstract

These are informal lecture notes for a topics course that was taught at UC Berkeley in Fall 2002. Floer theory (for which Morse homology is a prototype) and pseudoholomorphic curves and their applications to low dimensional and symplectic topology are currently the subject of a lot of active and exciting research. The basic goal of this course was to introduce some of the fundamental ideas which should prepare and inspire one to understand what workers in this field are doing and why, and perhaps even begin new research in this area. We gave an introduction to some of the technical machinery which is needed, while referring to other sources for details of the analysis. We explored some of the frontiers of (at least the author's) knowledge.

The first part of the course covered Morse theory as a prototype for Floer theory. Unfortunately (but not too surprisingly), I only had time to write detailed notes for this part of the course. The second part of the course gave an introduction to pseudoholomorphic curves in symplectic manifolds, and the third part of the course gave a (sometimes quite sketchy) discussion of Floer theory. The last chapter of these notes gives a brief outline of these last two parts of the course, with references to some starting points for further reading on these topics.

I thank all of the participants of the course for their enthusiasm and comments and questions.

Contents

1 Introductory remarks on Morse theory

We begin with a brief overview of Morse theory in order to introduce what we will be doing and how it fits into the bigger picture. Like many "introductions" to mathematical works, this is easier to understand if you already know some of what follows — and we sometimes use terminology which we will not begin defining until the next section — so you might want to refer back to it later.

Let X be a finite dimensional compact smooth manifold, and $f : X \to \mathbb{R}$ a smooth function. In Morse theory, one often but not always assumes that the critical points of f are nondegenerate, and relates the topology of X to the critical points of X.

There are two basic approaches to Morse theory.

1.1 The classical approach: attaching handles

The classical approach [46] is to define

$$X_a := \{x \in X \mid f(x) \leq a\},$$

where $a \in \mathbb{R}$ is not a critical value of f, and study how the topology of X_a changes as a increases. One can show that if there are no critical values in the interval $[a, b]$, then X_a is diffeomorphic to X_b. If, say, $f^{-1}[a, b]$ contains a single critical point of index i, then up to diffeomorphism, X_b is obtained from X_a by attaching an i-handle. This has three important applications:

1. This leads to the Morse inequalities, which are lower bounds for the numbers of critical points of f of each index in terms of the ranks of the homology groups of X. Namely, if all critical points of f are nondegenerate, if c_i is the number of critical points of index i, and b_i is the rank of $H_i(X)$, then

$$c_i - c_{i-1} + c_{i-2} - \cdots + (-1)^i c_0 \geq b_i - b_{i-1} + b_{i-2} - \cdots + (-1)^i b_0 \quad (1)$$

   for all i. Without the assumption of nondegenerate critical points, there are techniques such as Lusternik-Shnirelman theory to establish weaker inequalities for arbitrary f, and there is also Morse-Bott theory for functions f with "nondegenerate critical submanifolds" which we will discuss later.

2. One can generalize this to certain functionals on certain infinite dimensional manifolds, particularly the energy functional E on the free loop space $\mathscr{L}X := \{\gamma : S^1 \to X\}$ of a Riemannian manifold X, defined by

$$E(\gamma) := \int_{S^1} |\gamma'(t)|^2 \, dt,$$

   whose critical points are the closed geodesics. This leads to existence theorems for closed geodesics, such as the famous result that for any metric on S^2 there exist at least three closed geodesics. One can also turn this around and determine the topology of the loop space of a manifold whose geodesics one understands; this approach was used in the original proof of Bott periodicity [8]. There are also relations between curvature and Morse theory of geodesics which lead to relations between curvature and topology.

3. Understanding finite dimensional smooth manifolds in terms of attaching handles is the basis for constructive methods for proving that manifolds are diffeomorphic.

(a) This lies at the heart of the h-cobordism theorem and the proof of topological Poincare conjecture in dimensions greater than four [47]. For example one step is the Smale cancellation lemma, which asserts that if f has critical points p and q of index i and $i+1$ with $f(p) < f(q)$, if there are no critical values in the interval $(f(p), f(q))$, and if the attaching sphere for the handle corresponding to q goes exactly once over the handle corresponding to p, then one can modify f to cancel the critical points p and q. If one can cancel all critical points except the minimum and maximum of f, then X must be homeomorphic to a sphere.

(b) There is also the Kirby calculus [37, 38, 26] which is used to explicitly describe three and four-dimensional smooth manifolds and show directly that different manifolds are diffeomorphic.

1.2 A newer approach: gradient flow lines

In these lectures we will focus on a second, newer approach to Morse theory. In this approach one introduces an auxiliary Riemannian metric g on X. One then considers the negative[1] gradient vector field of f with respect to g, which we denote by V. One then looks at *flow lines* of the vector field V which start at one critical point and end at another. If the metric is generic, then there are finitely many gradient flow lines from a critical point of index i to a critical point of index $i-1$. One then defines a chain complex $C_*^{\mathrm{Morse}}(f,g)$ over $\mathbb{Z}$, the Morse complex, whose chain group C_i is generated by the critical points of index i, and whose differential counts gradient flow lines between critical points of index difference one. A fundamental result is that the homology of this chain complex is canonically isomorphic to the singular homology of X. Roughly speaking, the isomorphism from Morse homology to singular homology sends a critical point to its descending manifold; this has appeared in various forms in many papers[2]. One easily deduces the Morse

[1]The negative gradient, as opposed to the positive gradient, fits in better with the classical approach above, but sometimes leads to annoying signs.

[2]The Morse complex has a confusing history. An essentially equivalent complex was described in Milnor's book on the h-cobordism theorem [47], but not in the language of gradient flow lines; and there were earlier suggestions by Thom [67] and Smale [64]. But in the form described above, the Morse complex was introduced to a large audience by Witten [70], who obtained it (over $\mathbb{R}$) from a radical approach inspired by supersymmetry, as a limit of deformed Hodge theory in which the de Rham differential d is replaced by

inequalities from this: there have to be enough critical points to generate the homology.

The significance of the language of gradient flow lines is that, as realized by Floer [16, 18, 17], it extends to important infinite dimensional cases where the classical approach is useless[3]. These are cases where the critical points have infinite index, so that passing through a critical point does not change the topology of the manifold X_a. However sometimes the index *difference* between two critical points is still finite, in that one can make sense of "gradient flow lines" between two critical points, and these form a finite dimensional moduli space. One can then define an analogue of Morse homology, called Floer homology.

The relation of the Floer homology to the topology of the infinite dimensional manifold X is somewhat unclear[4]. However Floer homology is typically associated to some finite dimensional manifold, e.g. as its loop space, or to some more complicated finite dimensional object, and the Floer homology has topological significance for the finite dimensional object. There are many interesting examples of Floer theory, but in order not to stray too far from our current topic of Morse homology, we will save these for later.

In the finite dimensional case, it is possible to describe topological notions other than just homology, such as Reidemeister torsion and the Leray-Serre spectral sequence, in terms of gradient flow lines, and these then have Floer theoretic analogues. There are also new constructions in Floer theory, such as the "quantum product" in symplectic Floer theory [51], which do not[5] have analogues in classical topology.

$e^{-tf}de^{tf}$ and $t \to \infty$. This is a remarkable way to establish the isomorphism between Morse homology and singular homology (over $\mathbb{R}$), and was made rigorous by Helffer and Sjostrand. See [9] for a nice survey.

[3] The book [56] gives a detailed technical treatment of Morse homology with an eye towards Floer-theoretic generalizations.

[4] This matter is discussed in [13]. Also, for some versions of Floer theory, the analogy with Morse theory of a function on a space begins to break down. For example symplectic field theory is like the Morse theory of the symplectic action functional on the loop space, except that there can be several loops which fuse and separate in a "gradient flow line" or pseudoholomorphic curve.

[5] It is possible to construct the cup product in finite dimensional Morse theory in a way which relates to the quantum product much like the way that Feynman diagrams relate to string theory [7]; but this is not a direct translation as in the preceding sentence, where one type of "gradient flow line" is replaced by another.

1.3 Comparison of the two approaches

To summarize, let us briefly describe how the Floer-theoretic approach compares with the three basic applications of the classical approach from §1.1.

1. Both approaches establish the Morse inequalities, and while the newer proof of the Morse inequalities seems more elegant, the two proofs have roughly the same content.

2. Roughly speaking, from an analytic point of view, the classical approach extends to infinite dimensional settings in which the gradient flow equation is *parabolic*, while the Floer-theoretic approach extends to cases where the gradient flow equation is *elliptic*.

3. Floer homology is generally used as an invariant to tell spaces apart. However it is very intereting to ask if it can lead to constructive results. For example, is there a Smale cancellation lemma in Floer theory? This question has been considered by Fukaya [22] and in a different form by Taubes [65].

2 The definition of Morse homology

2.1 Morse functions

Let X be a smooth (finite dimensional) manifold, say closed for now, and $f : X \to \mathbb{R}$ a smooth function.

A **critical point** of f is a point $p \in X$ such that $df_p = 0 : T_pX \to \mathbb{R}$. We let $\mathrm{Crit}(f)$ denote the set of critical points of f.

If p is a critical point, we define the **Hessian** $H(f,p) : T_pX \to T_p^*X$ as follows. Let ∇ be any connection on TX, and if $v \in T_pX$, define

$$H(f,p)(v) := \nabla_v(df).$$

This does not depend on the choice of connection ∇ because df vanishes at p and the difference between any two connections is a tensor[6]. If $x_1, \ldots, x_n$ are

[6]Here is another way to see why if s is a section of a vector bundle $E \to X$ and $s(x) = 0$ then the derivative $\nabla s : T_xX \to E_x$ is well defined, as this is an important point which we will need later. Let us write $E = \{(x,e) \mid e \in E_x\}$, and let $\pi : E \to X$ denote the projection. Let $\Gamma = \{(x, s(x))\}$ denote the graph of s. Then at any point $(x, s(x)) \in \Gamma$,

local coordinates for X near p, then with respect to the bases $\{\partial/\partial x_i\}$ and $\{dx_i\}$ for T_pX and T_p^*X, the Hessian is given by the matrix $(\partial^2 f/\partial x_i \partial x_j)$. Since this matrix is symmetric, if we use a Riemannian metric to identify $T_pX \simeq T_p^*X$, the Hessian becomes a symmetric bilinear form on T_pX, or a self-adjoint map $T_pX \to T_pX$.

The critical point p is **nondegenerate** if the Hessian does not have zero as an eigenvalue. In this case we define the **Morse index** $\operatorname{ind}(p)$ to be the number of negative eigenvalues of the Hessian.

It is easy to see that a nondegenerate critical point is isolated. Moreover, although we will not really use this, the **Morse lemma** asserts that if p is a nondegenerate critical point of index i, then there exist local coordinates $x_1, \ldots, x_n$ for X near p such that

$$f = f(p) - x_1^2 - \cdots - x_i^2 + x_{i+1}^2 + \cdots + x_n^2.$$

The function f is **Morse** if all of its critical points are nondegenerate. One can show that a generic smooth function on X is Morse. We will later do a systematic study of how to precisely formulate and prove such genericity statements.

2.2 The gradient flow

Let g be a metric on X, and let V denote the negative gradient of f with respect to g. The flow of the vector field V defines a one-parameter group of diffeomorphisms $\Psi_s : X \to X$ for $s \in \mathbb{R}$ with $\Psi_0 = \mathrm{id}$ and $d\Psi_s/dt = V$.

If p is a critical point, we define the **descending manifold**

$$\mathscr{D}(p) := \left\{ x \in X \,\middle|\, \lim_{s \to -\infty} \Psi_s(x) = p \right\}$$

and the **ascending manifold**

$$\mathscr{A}(p) := \left\{ x \in X \,\middle|\, \lim_{s \to +\infty} \Psi_s(x) = p \right\}.$$

the map $\pi_* : T_{(x,s(x))}\Gamma \to T_xX$ is an isomorphism, because $\pi \circ s = \mathrm{id}_X$. If $s(x) = 0$, we define

$$\nabla s : T_xX \overset{(\pi_*)^{-1}}{\longrightarrow} T_{(x,0)}\Gamma \subset T_{(x,0)}E = T_xX \oplus E_x \longrightarrow E_x.$$

The key point is that there is a canonical identification $T_{(x,0)}E = T_xX \oplus E_x$ because T_xX includes into $T_{(x,0)}E$ as the tangent space to the zero section. A connection ∇ is an extension of this (satisfying some restrictions) to an identification $T_{(x,e)}E \simeq T_xX \oplus E_x$ for all $(x,e) \in E$, but such an identification is not canonical except when $e = 0$.

(These are sometimes also called the "unstable manifold" and "stable manifold", respectively, of the flow V.)

If p is a nondegenerate critical point, then $\mathscr{D}(p)$ is an embedded open disc in X with dimension

$$\dim \mathscr{D}(p) = \operatorname{ind}(p).$$

In fact, the tangent space $T_p\mathscr{D}(p) \subset T_pX$ is just the negative eigenspace of the Hessian $H(f,p)$. Likewise, $\mathscr{A}(p)$ is an embedded open disc with the complementary dimension

$$\dim \mathscr{A}(p) = \dim(X) - \operatorname{ind}(p).$$

We refer to [1] for the proof[7].

We assume for the rest of this section that the pair (f,g) is **Morse-Smale**: namely, f is Morse and for every pair of critical points p and q, the descending manifold $\mathscr{D}(p)$ is transverse to the ascending manifold $\mathscr{A}(q)$. We will see later (maybe) that this condition holds generically.

If p and q are critical points, a **flow line** from p to q is a path $\gamma : \mathbb{R} \to X$ with $\gamma'(s) = V(\gamma(s))$ and $\lim_{s\to-\infty} \gamma(s) = p$ and $\lim_{s\to+\infty} \gamma(s) = q$. Note that $\mathbb{R}$ acts on the set of flow lines from p to q by precomposition with translations of $\mathbb{R}$. We let $\mathscr{M}(p,q)$ denote the moduli space of flow lines from p to q, modulo translation. We can identify

$$\mathscr{M}(p,q) = \mathscr{D}(p) \cap \mathscr{A}(q)/\mathbb{R},$$

where $\mathbb{R}$ acts on X by the flow $\{\Psi_s\}$. In particular, the Morse-Smale condition implies that $\mathscr{M}(p,q)$ is naturally a manifold with

$$\dim \mathscr{M}(p,q) = \operatorname{ind}(p) - \operatorname{ind}(q) - 1 \tag{2}$$

(except in the case $p = q$, when the $\mathbb{R}$ action is trivial, where $\dim \mathscr{M}(p,p) = 0$).

When $p \neq q$, we orient $\mathscr{M}(p,q)$ as follows[8]. For each critical point p, choose an orientation of the descending manifold $\mathscr{D}(p)$. At any point in the

[7] This is more or less obvious if one chooses the metric near the critical points to be Euclidean in a coordinate chart given by the Morse lemma. This assumption is sometimes made in the literature in order to simplify various technical arguments. However this condition is not generic, as the eigenvalues of the Hessian are all distinct for a generic metric.

[8] This convention follows [54]. There are other ways to do this which are more abstract and possibly nicer but also more difficult to work with. We we will see a very slightly more elegant version when we study Morse-Bott theory.

image of γ, we have an isomorphism, canonical at the level of orientations[9],

$$\begin{aligned} T\mathscr{D}(p) &\simeq T(\mathscr{D}(p)\cap\mathscr{A}(q))\oplus(TX/T\mathscr{A}(q)) \\ &\simeq T_\gamma\mathscr{M}(p,q)\oplus T\gamma\oplus T_q\mathscr{D}(q). \end{aligned} \tag{3}$$

The isomorphism in the first line comes from the Morse-Smale transversality assumption; the isomorphism $(\mathscr{D}(p)\cap\mathscr{A}(q))\simeq T_\gamma\mathscr{M}(p,q)\oplus T\gamma$ holds by (2), and the isomorphism $TX/T\mathscr{A}(q)\simeq T_q\mathscr{D}(q)$ is obtained by translating the subspace $T_q\mathscr{D}(q)\subset T_qX$ along γ while keeping it complementary to $T\mathscr{A}(q)$. We orient $\mathscr{M}(p,q)$ so that the isomorphism (3) is orientation-preserving.

2.3 Compactification by broken flow lines

When $\mathrm{ind}(p)-\mathrm{ind}(q)=1$, the moduli space $\mathscr{M}(p,q)$ has dimension zero, and we would like to count the points in it. For this purpose we need to know that $\mathscr{M}(p,q)$ is compact. This follows from the following general fact.

Recall that a **smooth manifold with corners** is a second countable Hausdorff space[10] such that each point has a neighborhood with a chosen homeomorphism with $\mathbb{R}^{n-k}\times[0,\infty)^k$ for some k, and the transition maps are smooth.

Theorem 2.1 *If X is closed and (f,g) is Morse-Smale, then for any two critical points p,q, the moduli space $\mathscr{M}(p,q)$ has a natural compactification to a smooth manifold with corners $\overline{\mathscr{M}(p,q)}$ whose codimension k stratum is*

$$\overline{\mathscr{M}(p,q)}_k=\bigcup_{\substack{r_1,\ldots,r_k\in\mathrm{Crit}(f)\\ p,r_1,\ldots,r_k,q\ \mathit{distinct}}}\mathscr{M}(p,r_1)\times\mathscr{M}(r_1,r_2)\times\cdots\times\mathscr{M}(r_{k-1},r_k)\times\mathscr{M}(r_k,q).$$

When $k=1$, as oriented manifolds[11] we have

$$\partial\overline{\mathscr{M}(p,q)}=\bigcup_{\substack{r\in\mathrm{Crit}(f)\\ p,q,r\ \mathit{distinct}}}(-1)^{\mathrm{ind}(p)+\mathrm{ind}(r)+1}\mathscr{M}(p,r)\times\mathscr{M}(r,q).$$

[9] That is, there are lots of choices involved in defining this isomorphism, but any two isomorphisms that result will differ by an automorphism of positive determinant.

[10] The "second countable" and "Hausdorff" conditions are the same conditions one makes in defining an ordinary manifold in order to rule out the long line and other strange beasts.

[11] For now we will omit the calculation of signs like this. Generally the fact that the sign behaves in a uniform way is more important than what the actual sign is. For example if equation (4) held with a global minus sign then we would still get $\partial^2=0$ below. The paper [19] describes a general procedure for showing that "coherent orientations" exist, where the signs behave in a sufficiently uniform way to give $\partial^2=0$ etc.

For example, if $\operatorname{ind}(p) = i$ and $\operatorname{ind}(q) = i - 1$, then $\mathscr{M}(p,q)$ is compact. If $\operatorname{ind}(q) = i - 2$, then $\mathscr{M}(p,q)$ has a compactifaction $\overline{\mathscr{M}(p,q)}$ which is a compact 1-manifold with boundary

$$\partial\overline{\mathscr{M}(p,q)} = \bigcup_{r \in \operatorname{Crit}_{i-1}(f)} \mathscr{M}(p,r) \times \mathscr{M}(r,q). \tag{4}$$

Note that by (2), a critical point r can arise here only if its index is $i-1$; because $\mathscr{M}(p,r) \neq \emptyset$ and $p \neq r$ implies that $\operatorname{ind}(r) \leq i-1$, while $\mathscr{M}(r,q) \neq \emptyset$ and $r \neq q$ implies $\operatorname{ind}(r) \geq i-1$.

Theorem 2.1 and many variants and infinite dimensional generalizations thereof comprise the technical cornerstone of Floer theory. The proof has two main parts. The first part is a compactness result asserting that any sequence of flow lines in $\mathscr{M}(p,q)$ has a subsequence that converges in an appropriate sense to a "broken flow line" in $\overline{\mathscr{M}(p,q)}_k$ for some $k \geq 0$. The second part is a "gluing theorem" which asserts that any broken flow line in $\overline{\mathscr{M}(p,q)}_k$ can be perturbed to an honest flow line in $\mathscr{M}(p,q)$, and these perturbations are parametrized by $(R,\infty)^k$, such that taking one of these gluing parameters to infinity corresponds to breaking the flow line at one of the k intermediate critical points $r_1,\ldots,r_k$. One also has to check that the orientations work out. We will go into more details of some of this later.

The basic idea to remember is that in favorable cases, one can compactify moduli spaces of flow lines into compact manifolds with corners by adding in suitably "broken" flow lines. (In unfavorable cases, there are issues such as "bubbling" which make compactification more complicated.)

2.4 The chain complex

We define the **Morse complex** $(C_*^{\text{Morse}}(f,g), \partial^{\text{Morse}})$ as follows. Let $\operatorname{Crit}_i(f)$ denote the set of index i critical points of f. The chain module C_i is the free $\mathbb{Z}$-module generated by this finite set:

$$C_i^{\text{Morse}}(f,g) := \mathbb{Z}\operatorname{Crit}_i(f).$$

The differential $\partial^{\text{Morse}} : C_i \to C_{i-1}$ counts gradient flow lines. That is, if $p \in \operatorname{Crit}_i(f)$, then

$$\partial^{\text{Morse}}(p) := \sum_{q \in \operatorname{Crit}_{i-1}(f)} \#\mathscr{M}(p,q) \cdot q.$$

Here $\#\mathscr{M}(p,q) \in \mathbb{Z}$ denotes the number of points in $\mathscr{M}(p,q)$, counted with the signs given by the orientation on $\mathscr{M}(p,q)$.

Lemma 2.2 $(\partial^{\text{Morse}})^2 = 0$.

Proof. This follows immediately from (4), because the boundary of a compact oriented 1-manifold has zero points counted with sign. More precisely, if $p \in \text{Crit}_i(f)$ and $q \in \text{Crit}_{i-2}(f)$, then

$$\begin{aligned}
\left\langle (\partial^{\text{Morse}})^2 p, q \right\rangle &= \sum_{r \in \text{Crit}_{i-1}(f)} \left\langle \partial^{\text{Morse}} p, r \right\rangle \left\langle \partial^{\text{Morse}} r, q \right\rangle \\
&= \# \bigcup_{r \in \text{Crit}_{i-1}(f)} \mathscr{M}(p,r) \times \mathscr{M}(r,q) \\
&= \# \partial \overline{\mathscr{M}(p,q)} \\
&= 0.
\end{aligned}$$

□

We define the **Morse homology** $H_*^{\text{Morse}}(f,g)$ to be the homology of the chain complex $(C_*^{\text{Morse}}(f,g), \partial^{\text{Morse}})$.

Example 2.3 Let $X = T^2$, let f be the height function for an embedding of T^2 into $\mathbb{R}^3$ in which the torus is "standing on end", and let g be the metric induced by the Euclidean metric. The height function f is Morse and there are four critical points: one minimum of index 0, two saddles of index 1, and one maximum of index 2. The pair (f,g) is not Morse-Smale, because there are two flow lines from the upper saddle to the lower saddle. However these will disappear if we perturb g slightly. Then $\partial^{\text{Morse}} = 0$, because for each saddle, there are two flow lines from the maximum which have opposite signs and cancel, and two flow lines to the minimum which also have opposite signs and cancel. Therefore $H_2^{\text{Morse}} \simeq \mathbb{Z}$, $H_1^{\text{Morse}} \simeq \mathbb{Z}^2$, and $H_0^{\text{Morse}} \simeq \mathbb{Z}$.

Example 2.4 Suppose f is a Morse function on S^2 with two maxima x_1, x_2, one saddle y, and one minimum z. Then for any metric g, the pair (f,g) is Morse-Smale, and for suitable orientation choices we have

$$\begin{gathered}
\partial^{\text{Morse}}(x_1) = -\partial^{\text{Morse}}(x_2) = y, \\
\partial^{\text{Morse}}(y) = 0.
\end{gathered}$$

Therefore $H_2^{\text{Morse}} \simeq \mathbb{Z}$, $H_1^{\text{Morse}} = 0$, and $H_0^{\text{Morse}} \simeq \mathbb{Z}$.

Exercises for §2.

1. Let $\{\gamma_n\}$ be a sequence of flow lines from p to q, and let $\hat{\gamma} = (\hat{\gamma}_0, \ldots, \hat{\gamma}_k)$ be a **k-times broken flow line** from p to q; that is, there exist distinct critical points $r_0, \ldots, r_{k+1}$ with $r_0 = p$ and $r_{k+1} = q$ such that $\hat{\gamma}_i$ is a flow line from r_i to r_{i+1} for $i = 0, \ldots, k$. Let us say that $\lim_{n\to\infty}[\gamma_n] = [\hat{\gamma}]$ if for each n there exist real numbers $s_{n,0} < s_{n,1} < \cdots < s_{n,k}$ such that $\gamma_n(s_{n,i} + \cdot) \to \hat{\gamma}_i$ in C^∞ on compact sets.

   Show that any sequence of flow lines $\{\gamma_n\}$ from p to q has a subsequence which converges to some k-times broken flow line as above for some $k \geq 0$.

3 Morse homology is isomorphic to singular homology

We will now prove the following theorem, which is one of the most fundamental facts about finite-dimensional Morse theory.

Theorem 3.1 *If X is a closed smooth manifold and (f, g) is a Morse-Smale pair on X, then there is a canonical isomorphism*

$$H_*^{\text{Morse}}(f, g) \simeq H_*(X).$$

3.1 Outline of the proof

The idea of the proof of Theorem 3.1 is simple. We define a chain map $D : C_*^{\text{Morse}} \to C_*(X)$ by sending a critical point to its descending manifold. We define a map $A : C_*(X) \to C_*^{\text{Morse}}$ by taking a simplex, flowing it via V, and taking the sum of the critical points that it "hangs on". Then $A \circ D$ equals the identity on the chain level. On the other hand, $D \circ A$ is chain homotopic to the identity; the chain homotopy sends a singular chain to its entire forward orbit under the flow V.

To make this rigorous, we will use various compactifications by broken flow lines. But first, we need to decide what we mean by $C_*(X)$, and there are various approaches to handling the technicalities. Here we define $C_*(X)$ as follows. We say that an i-simplex $\sigma : \Delta_i \to X$ is **generic** if σ is smooth and each face of σ is transverse to the ascending manifolds of all the critical points of f. We let $C_i(X)$ denote the subspace of the set of all i-dimensional currents[12] on X generated by generic i-simplices. Standard arguments show that the homology of $C_i(X)$ so defined is canonically isomorphic to $H_*(X)$ as defined say by the Eilenberg-Steenrod axioms.

3.2 The chain map via compactified descending manifolds

To carry out the program outlined above, we start by defining a compactification of the descending manifold $\mathscr{D}(p)$ of a critical point p. The proof of the following proposition is similar to the proof of Theorem 2.1.

Proposition 3.2 *$\mathscr{D}(p)$ has a natural compactification to a smooth manifold with corners $\overline{\mathscr{D}(p)}$, whose codimension k stratum is*

$$\overline{\mathscr{D}(p)}_k = \bigcup_{\substack{q_1,\ldots,q_k\in\mathrm{Crit}(f)\\ p,q_1,\ldots,q_k \text{ distinct}}} \mathscr{M}(p,q_1)\times\mathscr{M}(q_1,q_2)\times\cdots\times\mathscr{M}(q_{k-1},q_k)\times\mathscr{D}(q_k).$$

When $k=1$, as oriented manifolds we have

$$\partial\overline{\mathscr{D}(p)} = \bigcup_{\substack{q\in\mathrm{Crit}(f)\\ p\neq q}} (-1)^{\mathrm{ind}(p)+\mathrm{ind}(q)+1}\mathscr{M}(p,q)\times\mathscr{D}(q).$$

The maps $\overline{\mathscr{D}(p)}_k\to X$ given by projecting to $\mathscr{D}(q_k)\subset X$ patch together to a smooth[13] *map*

$$e:\overline{\mathscr{D}(p)}\longrightarrow X$$

extending the inclusion $\mathscr{D}(p)\to X$.

Example 3.3 Define $f:[-1,1]^n\to\mathbb{R}$ by

$$f(x_1,\ldots,x_n) := \frac{1}{4}\sum_{i=1}^{n}(x_i+1)^2(x_i-1)^2$$

[12]The approach here is basically taken from [34], except that here we use currents instead of modding out by "degenerate singular chains". What we are doing here is different from the elegant treatment of Morse theory via currents in [29], which uses more general currents but makes additional assumptions on the gradient flow.

[13]I think this smoothness claim is OK but I need to check it, hopefully later.

and let g be the Euclidean metric. (If you like, include X into a closed n-manifold and extend f and g arbitrarily.) Then

$$-\nabla f = -\sum_{i=1}^{n}(x_i+1)x_i(x_i-1).$$

Thus f has a critical point of index k at the center of each k-face of the cube, and no other critical points. The descending manifold of a critical point is the interior of the corresponding face.

The compactified descending manifold of a critical point is diffeomorphic to a "fully truncated k-cube". If $k=2$, its boundary is an octagon. If $k=3$, its boundary is a polyhedron whose faces consist of 6 octagons, 12 quadrilaterals, and 8 hexagons.

Remark 3.4 One can show in general that $\overline{\mathscr{D}(p)}$ is homeomorphic to a closed ball, of course of dimension $\operatorname{ind}(p)$. Hence the compactified descending manifolds $\overline{\mathscr{D}(p)}$, together with the maps $e : \overline{\mathscr{D}(p)} \to X$, give X the structure of a CW-complex, with one i-cell for each critical point of index i. There are softer ways to see that a Morse function gives a CW-structure with one cell for each critical point; however the approach above shows that the metric gives a CW-structure more canonically.

Now the compact oriented manifold with corners $\overline{\mathscr{D}(p)}$ has a fundamental current $\left[\overline{\mathscr{D}(p)}\right]$, and we define

$$D(p) := e_*\left[\overline{\mathscr{D}(p)}\right].$$

Note that $D(p) \in C_*(X)$, because by the Morse-Smale assumption, we can compatibly triangulate all the descending manifolds using generic simplices by induction on the dimension.

Lemma 3.5 *D is a chain map: $\partial D = D\partial^{\mathrm{Morse}}$.*

Proof. Let $p \in \operatorname{Crit}_i(f)$. By Proposition 3.2 we have

$$\partial\overline{\mathscr{D}(p)} = \bigcup_{\substack{q\in\operatorname{Crit}(f)\\ p\neq q}} (-1)^{i+\operatorname{ind}(q)+1}\mathscr{M}(p,q)\times\mathscr{D}(q).$$

Therefore

$$\partial D(p) = \sum_{\substack{q\in\mathrm{Crit}(f)\\ p\neq q}} (-1)^{i+\mathrm{ind}(q)+1} e_* \left[\overline{\mathscr{M}(p,q)} \times \overline{\mathscr{D}(q)}\right] \in C_{i-1}(X).$$

Now if $\mathrm{ind}(q) > i-1$, then $\mathscr{M}(p,q)$ is empty by the Morse-Smale condition, while if $\mathrm{ind}(q) < i-1$, then the contribution on the right hand side is zero in $C_{i-1}(X)$, because e maps $\overline{\mathscr{M}(p,q)} \times \overline{\mathscr{D}(q)}$ to the support of $D(q)$, which is a current of dimension $\leq i-2$. Therefore

$$\begin{aligned}\partial D(p) &= \sum_{q\in\mathrm{Crit}_{i-1}(f)} \#\mathscr{M}(p,q)\cdot e_*\left[\overline{\mathscr{D}(q)}\right]\\ &= D\left(\partial^{\mathrm{Morse}}(p)\right).\end{aligned}$$

□

3.3 The left inverse chain map

If σ is a generic i-simplex and q is a critical point, let $\mathscr{M}(\sigma,q)$ denote the moduli space of gradient flow lines from σ to q, i.e. maps $\gamma : [0,\infty) \to X$ such that $\gamma(0) \in \sigma$ and $\gamma'(s) = V(\gamma(s))$ and $\lim_{s\to\infty}\gamma(s) = q$. As in (3), we have an isomorphism

$$T_{\gamma(0)}\sigma \simeq T_\gamma\mathscr{M}(\sigma,q) \oplus T_q\mathscr{D}(q),$$

and we orient $\mathscr{M}(\sigma,q)$ so that this isomorphism is orientation-preserving.

As in Theorem 2.1 and Proposition 3.2, $\mathscr{M}(\sigma,q)$ has a compactfication to a smooth manifold with corners $\overline{\mathscr{M}(\sigma,p)}$ whose codimension k stratum is

$$\overline{\mathscr{M}(\sigma,q)}_k = \bigcup_{j=0}^{k} \bigcup_{\substack{p_1,\ldots,p_j\in\mathrm{Crit}(f)\\ p_1,\ldots,p_j,q \text{ distinct}}} \mathscr{M}(\sigma_{k-j},p_1)\times\mathscr{M}(p_1,p_2)\times\cdots\times\mathscr{M}(p_{j-1},p_j)\times\mathscr{M}(p_j,q).$$

Here σ_j denotes the codimension j stratum of σ. When $k=1$, as oriented manifolds we have

$$\partial\overline{\mathscr{M}(\sigma,q)} = \mathscr{M}(\partial\sigma,q) \cup \bigcup_{\substack{p\in\mathrm{Crit}(f)\\ p\neq q}} (-1)^{i+\mathrm{ind}(q)}\mathscr{M}(\sigma,p)\times\mathscr{M}(p,q).$$

Clearly $\dim \mathscr{M}(\sigma, p) = i - \mathrm{ind}(p)$. By this and the compactness result, it makes sense to define

$$A(\sigma) := \sum_{p \in \mathrm{Crit}_i(f)} \#\mathscr{M}(\sigma, p) \cdot p.$$

Lemma 3.6 *A is a chain map: $A\partial = \partial^{\mathrm{Morse}} A$.*

Proof. This follows from the compactness result, since if $q \in \mathrm{Crit}_{i-1}(f)$, then

$$\begin{aligned}
\#\partial\overline{\mathscr{M}(\sigma, q)} &= \#\mathscr{M}(\partial\sigma, q) - \# \bigcup_{\substack{p \in \mathrm{Crit}(f) \\ p \neq q}} \mathscr{M}(\sigma, p) \times \mathscr{M}(p, q) \\
&= \#\mathscr{M}(\partial\sigma, q) - \# \bigcup_{p \in \mathrm{Crit}_i(f)} \mathscr{M}(\sigma, p) \times \mathscr{M}(p, q) \\
&= \langle A(\partial\sigma), q\rangle - \langle \partial^{\mathrm{Morse}} A(\sigma), q\rangle.
\end{aligned}$$

Here the second equality holds because of our transversality assumptions. □

Lemma 3.7 $A \circ D = \mathrm{id} : C_i^{\mathrm{Morse}} \to C_i^{\mathrm{Morse}}$.

Proof. If p is an index i critical point, then $\mathscr{M}(D(p), p)$ contains one point, the constant gradient flow line, oriented positively by our sign convention; while $\mathscr{M}(D(p), q)$ is empty if q is any other index i critical point, because $\mathscr{M}(p, q)$ is empty by the Morse-Smale condition. □

3.4 The chain homotopy via compactified forward orbits

If σ is a generic i-simplex, we define its **forward orbit** to be the set

$$\mathscr{F}(\sigma) := [0, \infty) \times \sigma$$

together with the map $e : \mathscr{F}(\sigma) \to X$ defined by

$$e(s, x) := \Psi_s(\sigma(x)).$$

The forward orbit has a natural compactification to a smooth manifold with corners $\overline{\mathscr{F}(\sigma)}$ whose codimension k stratum for $k > 2$ is

$$\overline{\mathscr{F}(\sigma)}_k = \mathscr{F}(\sigma_k) \cup \bigcup_{j=1}^{k} \bigcup_{\substack{r_1,\ldots,r_j \in \mathrm{Crit}(f) \\ r_1,\ldots,r_j \text{ distinct}}} \mathscr{M}(\sigma_{k-j}, r_1) \times \mathscr{M}(r_1, r_2) \times \cdots \times \mathscr{M}(r_{j-1}, r_j) \times \mathscr{D}(r_j).$$

When $k = 1$, as oriented manifolds we have

$$\partial\overline{\mathscr{F}(\sigma)} = -\sigma \cup -\mathscr{F}(\partial\sigma) \cup \bigcup_{r \in \mathrm{Crit}(f)} \mathscr{M}(\sigma, r) \times \mathscr{D}(r).$$

The map e extends over this compactification as a smooth map which projects to $\mathscr{D}(r_j) \subset X$.

We now define $F : C_i(X) \to C_{i+1}(X)$ by

$$F(\sigma) := e_* \left[\overline{\mathscr{F}(\sigma)}\right].$$

Then the above compactification result implies that F is a chain homotopy between $D \circ A$ and the identity:

Lemma 3.8 $\partial F + F\partial = D \circ A - \mathrm{id}$.

Lemmas 3.5, 3.6, 3.7 and 3.8 complete the proof of Theorem 3.1.

3.5 Morse cobordisms and relative homology.

Theorem 3.1 has the following useful generalization. Let X be a compact smooth manifold with boundary, whose boundary is partitioned into two unions of connected components X_0 and X_1. A **Morse cobordism** is a smooth function $f : X \to [0, 1]$ such that $f^{-1}(i) = X_i$ for $i = 0, 1$, and all critical points of f are nondegenerate and in the interior of X.

Theorem 3.9 *Let $f : X \to [0, 1]$ be a Morse cobordism and let g be a metric on X such that (f, g) is Morse-Smale. Then there is a canonical isomorphism*

$$H_*^{\mathrm{Morse}}(f, g) \simeq H_*(X, X_0).$$

Exercises for §3.

1. Deduce the Morse inequalities (1) from Theorem 3.1.
2. Use Theorem 3.1 to prove the Künneth formula for closed manifolds.
3. Use Theorem 3.1 to prove the Poincar‘’e-Hopf index theorem: if X is a closed oriented smooth manifold, then $\int_X e(TX)$, i.e. the signed number of zeroes of a generic vector field, is equal to the Euler characteristic $\chi(X)$.
4. Use Theorem 3.1 to prove Poincaré duality for closed oriented manifolds.
5. Prove Theorem 3.9. Deduce Poincaré-Lefschetz duality.

4 *A priori* invariance of Morse homology

Let X be a closed smooth manifold and (f, g) a Morse-Smale pair. We will now give an *a priori* proof that the Morse homology $H_*^{\mathrm{Morse}}(f, g)$ is a topological invariant, i.e. it depends only on X and f and g. Of course we already know this as a corollary of Theorem 3.1. The point of this exercise is that it provides a model for proofs that various versions of Floer homology are topological invariants, where an interpretation in terms of a previously known topological invariant might not be available or possible[14].

A natural and enlightening strategy for the proof is "bifurcation analysis": one deforms the pair (f, g), studies explicitly how the chain complex changes, and checks that the homology stays the same, see [17, 41]. However, bifurcation analysis is technically difficult in general, and Floer discovered an elegant alternative approach [16] which uses the same ideas as the proof that $\partial^2 = 0$, and which we will now explain.

4.1 Continuation maps

Let (f_0, g_0) and (f_1, g_1) be two Morse-Smale pairs. Let (C_*^0, ∂_0) and (C_*^1, ∂_1) denote the corresponding Morse complexes. Let

$$\Gamma = \{(f_t, g_t) \mid t \in [0, 1]\}$$

be a path of functions and metrics from (f_0, g_0) to (f_1, g_1). Under a genericity assumption to be explained below, we define the **continuation map**

$$\Phi_\Gamma : C_*^0 \longrightarrow C_*^1$$

as follows.

Define a vector field V on $[0, 1] \times X$ by

$$V := (1 - t)t(1 + t)\partial_t + V_t, \tag{5}$$

where t denotes the $[0, 1]$ coordinate and V_t denotes the negative gradient of $f_t : X \to \mathbb{R}$ with respect to the metric g_t. The vector field V is sufficiently

[14] Perhaps on some other planet, Morse homology was discovered before any other form of homology. Then on that planet, this result proved that Morse homology is a powerful tool for distinguishing closed smooth manifolds; and whoever discovered this probably received that planet's analogue of the Fields medal.

well behaved that we can define its critical points, ascending and descending manifolds, and flow lines just as if it were the negative gradient of a Morse function, and the same transversality and compactness properties will hold[15]. The function $(t+1)^2(t-1)^2/4$ on $\mathbb{R}$ has a critical point of index 1 at $t=0$ and a critical point of index 0 at $t=1$ with no critical points in between. Thus

$$\mathrm{Crit}_i(V) = \{0\} \times \mathrm{Crit}_{i-1}(f_0) \bigcup \{1\} \times \mathrm{Crit}_i(f_1). \tag{6}$$

We say that the family Γ is **admissible** if the ascending and descending manifolds of the critical points of V intersect transversely. One can show that if (f_0, g_0) and (f_1, g_1) are Morse-Smale, then a generic homotopy Γ between them is admissible. This is a slight modification of the proof that a generic pair (f, g) is Morse-Smale. So assume from now on that Γ is admissible. Note that for an admissible Γ, there might (and often must) be some "bifurcation times" t for which the pair (f_t, g_t) is not Morse-Smale. Genericity of a family does not imply genericity of all the individual points in the family.

To continue, if P and Q are critical points of V, let $\mathscr{M}(P, Q)$ denote the moduli space of flow lines of V from P to Q, modulo the $\mathbb{R}$ action as usual. The orientation of $[0,1]$ and the orientations of the descending manifolds for (f_0, g_0) and (f_1, g_1) induce orientations of the descending manifolds for V and hence of the moduli spaces $\mathscr{M}(P, Q)$. Now if $p \in \mathrm{Crit}_i(f_0)$, we define

$$\Phi_\Gamma(p) := \sum_{q \in \mathrm{Crit}_i(f_1)} \#\mathscr{M}((0,p),(1,q)) \cdot q.$$

Lemma 4.1 *Φ_Γ is a chain map: $\partial_1 \Phi_\Gamma = \Phi_\Gamma \partial_0$.*

Proof. If $p \in \mathrm{Crit}_i(f_0)$ and $q \in \mathrm{Crit}_{i-1}(f_1)$, then the usual argument shows that $\mathscr{M}((0,p),(1,q))$ has a compactification to a compact oriented 1-manifold

[15] In the first draft of this lecture I defined V to be the negative gradient of the function $F : [0,1] \times X \to \mathbb{R}$ defined by $F(t,x) := \frac{1}{4}(t+1)^2(t-1)^2 + f_t(x)$, with respect to the metric G on $[0,1] \times X$ defined by $G(t,x) = dt^2 + g_t(x)$. But this doesn't work in the discussion below because I want the $[0,1]$ component of the vector field to be positive on $(0,1) \times X$. Thanks to Tamas Kalman who pointed out this mistake, and also suggested fixing it by multiplying the term $\frac{1}{4}(t+1)^2(t-1)^2$ in the definition of F by a large constant and assuming that f_t is independent of t for t close to 0 or 1. That would work fine here, and also makes concatenation of paths nicer. However the vector field (5) is the one I wanted in the first place because of generalizations that I have in mind in [33].

$\overline{\mathscr{M}((0,p),(1,q))}$ with boundary

$$\partial\overline{\mathscr{M}((0,p),(1,q))} = \bigcup_{r\in \mathrm{Crit_i(f_1)}} \mathscr{M}((0,p),(1,r)) \times \mathscr{M}((1,r),(1,q))$$
$$\cup \bigcup_{r\in \mathrm{Crit_{i-1}(f_0)}} \mathscr{M}((0,p),(0,r)) \times \mathscr{M}((0,r),(1,q)).$$

If $\mathscr{M}_0$ and $\mathscr{M}_1$ denote the moduli spaces for (f_0, g_0) and (f_1, g_1), then as oriented manifolds we have

$$\mathscr{M}((0,p),(0,r)) = (-1)^{\mathrm{ind}(p)+\mathrm{ind}(r)}\mathscr{M}_0(p,r),$$
$$\mathscr{M}((1,r),(1,q)) = \mathscr{M}_1(r,q).$$

The lemma follows immediately[16]. □

Thus Φ_Γ induces a map

$$(\Phi_\Gamma)_* : H_*^{\mathrm{Morse}}(f_0, g_0) \to H_*^{\mathrm{Morse}}(f_1, g_1). \tag{7}$$

4.2 Chain homotopies

Now let Γ and Γ' be two different generic paths with the same endpoints (f_0, g_0) and (f_1, g_1). Let Φ and Φ' denote the corresponding continuation maps.

Lemma 4.2 *A generic homotopy between the paths Γ and Γ' induces a chain homotopy*

$$K : C_*^0 \longrightarrow C_{*+1}^1,$$
$$\partial_1 K + K\partial_0 = \Phi - \Phi'.$$

Proof. We regard the homotopy as a family $\{(f_d, g_d) \mid d \in D\}$, where D is a **digon** (a closed 2-manifold with corners with two edges and two vertices). Let $\hat{g}$ be a metric on D such that the edges have length 1. Let $\hat{f} : D \to \mathbb{R}$ be a function with an index 2 critical point at one vertex and an index 0 critical point at the other vertex and no other critical points, such that the negative

[16]Another way to say this is that the Morse differential ∂ for the vector field V is well-defined and still satisfies $\partial^2 = 0$. With respect to the decomposition of $\mathrm{Crit}(V)$ given by (6), we have $\partial = \begin{pmatrix} -\partial_0 & 0 \\ \Phi_\Gamma & \partial_1 \end{pmatrix}$ so $\partial^2 = \begin{pmatrix} \partial_0^2 & 0 \\ -\Phi_\Gamma\partial_0 + \partial_1\Phi_\Gamma & \partial_1^2 \end{pmatrix}$.

gradient of $\hat{f}$ with respect to $\hat{g}$ is tangent to the edges and agrees with the negative gradient of $(t+1)^2(t-1)^2/4$ there. Let $\hat{V}$ be the negative gradient of $\hat{f}$ with respect to $\hat{g}$. We then define a vector field V on $D \times X$ by

$$V := \hat{V} + V_d$$

where V_d denotes the negative gradient of f_d with respect to g_d. The map K then counts flow lines of the vector field V. We omit the verification of the chain homotopy equation. $\square$

This proves that the map $(\Phi_\Gamma)_*$ in (7) depends only on the homotopy class of Γ. In fact, since the space of paths Γ here is contractible[17], this implies that the map $(\Phi_\Gamma)_*$ does not depend on anything[18]. We now want to prove that it is an isomorphism. If Γ_1 is an admissible path from (f_0, g_0) to (f_1, g_1), and Γ_2 is an admissible path from (f_1, g_1) to (f_2, g_2), let $\Gamma_2 * \Gamma_1$ denote the concatenation of these two paths, reparametrized to be smooth and perturbed if necessary to be admissible.

Lemma 4.3 *$\Phi_{\Gamma_2 * \Gamma_1}$ is chain homotopic to $\Phi_{\Gamma_2} \circ \Phi_{\Gamma_1}$.*

Proof. This is similar to the proof of the preceding lemma except that we use a triangle instead of a digon. $\square$

The preceding lemma, together with Exercise 1 below, imply that for any two Morse-Smale pairs (f_0, g_0) and (f_1, g_1), there is a canonical[19] isomorphism $H_*^{\mathrm{Morse}}(f_0, g_0) \simeq H_*^{\mathrm{Morse}}(f_1, g_1)$.

Exercises for §4.

1. Show that if $\Gamma = \{(f_t, g_t)\}$ is a *constant* family with (f_t, g_t) Morse-Smale, then Γ is admissible and $\Phi_\Gamma = \mathrm{id}$.

2. Find counterexamples with $X = S^1$ to each of the following statements.

   (a) Suppose (f_t, g_t) is Morse-Smale for all $t \in [0,1]$, so that there is a canonical identification $\mathrm{Crit}(f_0) \simeq \mathrm{Crit}(f_1)$. Then the family $\Gamma =$

[17] Note that the space of metrics on a manifold is contractible, because one can contract all metrics to a given one by averaging.

[18] It is important to note that in Floer theory, there are often different homotopy classes of paths connecting two objects, and sometimes the induced maps on Floer homology can distinguish them, see [60].

[19] In Floer theory the analogous isomorphism might not be canonical, see the preceding footnote.

$\{(f_t, g_t)\}$ is admissible, and Φ_Γ is given by the canonical identification above.

(b) $\Phi_{\Gamma_2 * \Gamma_1} = \Phi_{\Gamma_2} \circ \Phi_{\Gamma_1}$ at the chain level.

3. Show that the diagram

$$\begin{array}{ccc} H_*^{\mathrm{Morse}}(f_0, g_0) & \longrightarrow & H_*^{\mathrm{Morse}}(f_1, g_1) \\ \big\downarrow & & \big\downarrow \\ H_*(X) & = & H_*(X) \end{array}$$

commutes, where the top arrow is the continuation isomorphism, and the vertical arrows are the isomorphisms given by Theorem 3.1.

5 Genericity and transversality

We now explain at least some of how to prove statements such as "a generic function is Morse". We begin with a general definition of "generic".

Definition 5.1 Let X be a topological space and let $P(x)$ be a statement for each $x \in X$ (which might be true or false). We say that $P(x)$ is true for **generic** $x \in X$ if the set $\{x \in X \mid P(x)\} \subset X$ contains a countable intersection of open dense sets.

This is a reasonable definition of "generic", for example because the Baire category theorem asserts that if X is a complete metric space then a countable intersection of open dense sets in X is itself dense.

5.1 The Sard-Smale theorem

The basic strategy for proving genericity statements is encapsulated in Theorem 5.4 below[20]. It requires the Sard-Smale theorem, an infinite dimensional generalization of Sard's theorem. We first recall the following definition.

Definition 5.2 Let V and W be Banach spaces. A bounded linear operator $F : V \to W$ is **Fredholm** if:

[20]This theorem is distilled out of [44], which provides tons of details regarding a lot of the analysis we will be discussing (and there will be even more details in the second edition).

- F has closed range, i.e. $F(V)$ is a closed subspace of W.
- $\dim \operatorname{Ker}(F) < \infty$.
- $\dim \operatorname{Coker}(F) < \infty$.

If F is Fredholm we define the **index**

$$\operatorname{ind}(F) := \dim \operatorname{Ker}(F) - \dim \operatorname{Coker}(F).$$

The index is a locally constant function on the space of Fredholm operators with the norm topology.

Theorem 5.3 (Sard-Smale) *Let X and Y be separable*[21] *Banach manifolds*[22]*. Let $f : X \to Y$ be a C^k map*[23] *such that $df_x : T_xX \to T_{f(x)}Y$ is Fredholm of index l for all $x \in X$. Assume $k \geq 1$ and $k \geq l + 1$.*

Then a generic $y \in Y$ is a regular value of f, i.e. df_x is onto for all $x \in f^{-1}(y)$, so $f^{-1}(y)$ is naturally a manifold[24] *of dimension l.*

The idea of the proof is to use the Fredholm assumption to locally reduce to Sard's theorem in finite dimensions, and to use the separability assumption to get a countable intersection of open dense sets.

We use the Sard-Smale theorem as follows. Suppose we have an equation of the form $\psi(y, z) = 0$, and we want to show that for generic $y \in Y$, the set of z such that $\psi(y, z) = 0$ is "cut out transversely".

[21] A topological space is **separable** if it contains a countable dense set.

[22] A **Banach manifold** is defined just like a smooth manifold except that it is locally modelled on a Banach space rather than $\mathbb{R}^n$.

[23] If V and W are Banach spaces then a function $f : V \to W$ is **differentiable** at $p \in V$ if there exists a bounded linear map $df_p : V \to W$ such that

$$\lim_{v \to 0} \frac{||f(p + v) - f(p) - df_p(v)||}{||v||} = 0.$$

If such a df_p exists then it is necessarily unique. If f is differentiable everywhere then df is a map $V \to \operatorname{Hom}(V, W)$ and one can similarly talk about the derivative of df, etc.

[24] The implicit function implies that if $f : X \to Y$ is a C^k map between Banach manifolds and if y is a regular value of f, then $f^{-1}(y)$ is a C^k submanifold of X, with $T_xf^{-1}(y) = \operatorname{Ker}(df_x)$. The proof is a nice application of the contraction mapping theorem, and if you haven't seen this before you should learn it because it's cool. This kind of analysis is needed for gluing theorems in Floer theory.

Theorem 5.4 (useful) *Let Y, Z be separable Banach manifolds, $E \to Y \times Z$ a Banach space bundle, and $\psi : Y \times Z \to E$ a smooth section. Suppose that for all $(y,z) \in \psi^{-1}(0)$, the following hold:*

(a) The differential $\nabla\psi_{(y,z)} : T_{(y,z)}(Y \times Z) \to E_{(y,z)}$ is surjective.

(b) The restricted differential $\nabla\psi_{(y,z)} : T_z Z \to E_{(y,z)}$ is Fredholm of index l.

Then for generic $y \in Y$, the set $\{z \in Z | \psi(y,z) = 0\}$ is an l-dimensional submanifold of Z (and moreover at each point in this set, $\nabla\psi$ is surjective on the tangent space to Z).

Proof. Hypothesis (a) and the implicit function theorem imply that $\psi^{-1}(0)$ is a Banach manifold. Now let $\pi : \psi^{-1}(0) \to Y$ be the projection.

Claim: For each $(y,z) \in \psi^{-1}(0)$, the projection $d\pi : T_{(y,z)}\psi^{-1}(0) \to T_y Y$ is Fredholm.

Proof of claim: The finite dimensional kernel, finite dimensional cokernel, and closed range properties follow from the corresponding properties for the restricted differential in (b). First, we have a tautological equality

$$\operatorname{Ker}\left(d\pi : T_{(y,z)}\psi^{-1}(0) \to T_y Y\right) = \operatorname{Ker}\left(\nabla\psi : T_z Z \to E_{(y,z)}\right). \tag{8}$$

Furthermore $\nabla\psi : T_y Y \to E_{(y,z)}$ induces an injection on cokernels which by (a) is in fact an isomorphism,

$$\nabla\psi : \operatorname{Coker}\left(d\pi : T_{(y,z)}\psi^{-1}(0) \to T_y Y\right) \xrightarrow{\simeq} \operatorname{Coker}\left(\nabla\psi : T_z Z \to E_{(y,z)}\right). \tag{9}$$

Finally, $d\pi : T_{(y,z)}\psi^{-1}(0) \to T_y Y$ has closed range because

$$d\pi\left(T_{(y,z)}\psi^{-1}(0)\right) = \{y \in T_y Y \mid \nabla\psi(y,0) \in \nabla\psi(T_z Z)\},$$

$\nabla\psi(T_z Z)$ is closed, and the inverse image of a closed set under a continuous map is closed.

The claim and the Sard-Smale theorem imply that a generic $y \in Y$ is a regular value of $\pi : \psi^{-1}(0) \to Y$. For such a y, the set $\{z \in Z \mid \psi(y,z) = 0\}$ is then a submanifold of Z by the implicit function theorem; by (8) this submanifold has dimension l, and by (9), for for each (y,z) in this submanifold, the restricted differential $\nabla\psi_{(y,z)} : T_z Z \to E_{(y,z)}$ is surjective. $\square$

5.2 Generic functions are Morse

Here is a simple example of the application of Theorem 5.4.

Proposition 5.5 *Let Z be a closed smooth manifold and let $k \geq 2$ be an integer. Then a generic C^k function $f : Z \to \mathbb{R}$ is Morse.*

Proof. Let $Y = C^k(Z, \mathbb{R})$, and let $E \to Y \times Z$ be the pullback of the cotangent bundle $TZ \to Z$ via the projection $Y \times Z \to Z$, so that $E_{(f,z)} = T_z Z$. Define a section ψ of E by $\psi(f, z) = df(z)$. Suppose $(f, z) \in \psi^{-1}(0)$. If f_1 is another C^k function on Z and $v \in T_z Z$ then

$$\nabla\psi_{(f,z)}(f_1, v) = df_1(v) + \nabla_v(df).$$

Theorem 5.4 is applicable because: (a) clearly $\nabla\psi_{(f,z)} : T(Y \times Z) \to T^*zZ$ is surjective, since $df_1(v)$ can be arbitrary; (b) the restricted differential

$$\nabla\psi_{(f,z)} : T_z Z \to T^*zZ \tag{10}$$

is automatically Fredholm since it maps between finite dimensional vector spaces.

So for generic f, for each $z \in Z$ such that $\psi(f, z) = 0$, i.e. for each critical point z of f, the restricted differential (10) is surjective. But now we recognize that the operator (10) is just the Hessian, and if it is surjective then the critical point is nondegenerate. □

This argument does not work for C^∞ functions because $C^\infty(Z, \mathbb{R})$ is not a Banach space. However there is a general technique for passing from C^k-genericity to C^∞-genericity. We refer the reader to [44] for the details.

5.3 Spectral flow

Our next goal is to show that if f is a Morse function, then for a generic metric g, the pair (f, g) is Morse-Smale. Before doing so, we need to introduce an important principle. The discussion here is based on the paper [53], which does much more stuff in much more detail.

Let $\mathscr{H}$ be a Hilbert space and let $\{A_s \mid s \in \mathbb{R}\}$ be a continuous family of operators on $\mathscr{H}$. The operators A_s may be unbounded. We assume that A_s converges in the norm topology to invertible self-adjoint operators $A^\pm$ as

$s \to \pm\infty$. If the family $\{A_s\}$ is reasonable[25], then one can make sense of the **spectral flow**

$$\mathrm{SF}\{A_s\} \in \mathbb{Z},$$

which intuitively is the number of eigenvalues of A_s which cross from negative to positive as s goes from $-\infty$ to $+\infty$. If $\mathscr{H}$ is finite dimensional then no additional assumptions are needed for the family to be "reasonable" and the spectral flow is simply the dimension of the positive eigenspace of A^+ minus the dimension of the positive eigenspace of A^-.

We now consider the operator[26]

$$\partial_s - A_s : L^2_1(\mathbb{R}, \mathscr{H}) \to L^2(\mathbb{R}, \mathscr{H}). \tag{11}$$

A precise statement and proof of the following principle is given in [53].

Principle 5.6 *If $\{A_s\}$ is a reasonable family of operators as above, then $\partial_s - A_s$ is Fredholm, and*

$$\operatorname{ind}(\partial_s - A_s) = -\,\mathrm{SF}\{A_s\}.$$

Example 5.7 Here is a sketch of some of the proof when $\mathscr{H}$ is finite dimensional.

Let us first try to understand the kernel of $\partial_s - A$. For each $h \in \mathscr{H}$, by the fundamental theorem of ODE's, there exists[27] a unique differentiable function $f_h : \mathbb{R} \to \mathscr{H}$ solving the equation

$$(\partial_s - A_s) f_h(s) = 0, \quad f_h(0) = h.$$

Now this function may or may not be in L^2_1. To analyze this, we define subspaces

$$\mathscr{H}^+ := \left\{ h \in \mathscr{H} \mid \lim_{s \to +\infty} f_h(s) = 0 \right\},$$
$$\mathscr{H}^- := \left\{ h \in \mathscr{H} \mid \lim_{s \to -\infty} f_h(s) = 0 \right\}.$$

[25] One set of sufficient technical assumptions is given in [53].

[26] Recall that if $p \geq 1$ and k is a nonnegative integer then the **Sobolev space** L^p_k is the completion of the space of smooth functions f, such that f and its first k derivatives are in L^p, with respect to the sum of the L^p norms of f and its first k derivatives.

[27] Actually, since $\mathscr{H}$ is not compact, the basic existence theorem for ODE's only gives us a short-time solution defined for $s \in (-\delta, \delta)$ for some $\delta > 0$. But in the present situation the short-time solution can be continued for all time because we have a uniform upper bound on the eigenvalues of A so that the solution cannot escape to infinity in finite time.

Then there is an isomorphism

$$\begin{aligned} \mathscr{H}^+ \cap \mathscr{H}^- &\xrightarrow{\simeq} \operatorname{Ker}(\partial_s - A_s), \\ h &\longmapsto f_h. \end{aligned}$$

Namely, one can show that if $h \in \mathscr{H}^+ \cap \mathscr{H}^-$, then f_h and hence its first derivative decay exponentially as $s \to \pm\infty$, so $f_h \in L^2_1$. Conversely, if $f \in \operatorname{Ker}(\partial_s - A_s)$ then $f = f_h$ for some h, and we must have $h \in \mathscr{H}^+ \cap \mathscr{H}^-$, or else one can show that f blows up exponentially as f approaches one end of $\mathbb{R}$ or the other so that $f \notin L^2_1$.

Furthermore, if $E^-(A^+)$ denotes the negative eigenspace of A^+, then (it takes some thought to justify this) we have an isomorphism

$$\begin{aligned} \mathscr{H}^+ &\xrightarrow{\simeq} E^-(A^+), \\ h &\longmapsto |h| \lim_{s \to +\infty} \frac{f_h(s)}{|f_h(s)|}. \end{aligned} \tag{12}$$

Similarly, $\mathscr{H}^-$ is isomorphic to $E^+(A^-)$, the positive eigenspace of A^-.

It is shown in [53] that $\partial_s - A_s$ has closed range.

If we believe this, then the cokernel of $\partial_s - A_s$ is just the kernel of its formal adjoint, i.e. the kernel of $\partial_s + A_s^*$. More specifically, we claim that there is an isomorphism

$$\begin{aligned} \operatorname{Ker}(\partial_s + A_s^*) &\xrightarrow{\simeq} (\mathscr{H}^+)^\perp \cap (\mathscr{H}^-)^\perp, \\ \tilde{f} &\longmapsto \tilde{f}(0). \end{aligned} \tag{13}$$

To see that this map is well-defined, suppose $\tilde{f} \in \operatorname{Ker}(\partial_s + A_s^*)$ and let $h \in \mathscr{H}^\pm$. Then

$$\begin{aligned} \partial_s \langle \tilde{f}, f_h \rangle &= \langle \partial_s \tilde{f}, f_h \rangle + \langle \tilde{f}, \partial_s f_h \rangle \\ &= \langle -A^* \tilde{f}, f_h \rangle + \langle \tilde{f}, A f_h \rangle \\ &= 0. \end{aligned}$$

On the other hand since $\lim_{s \to \pm\infty} \tilde{f}(s) = 0$ we have

$$\lim_{s \to \pm\infty} \langle \tilde{f}(s), f_h(s) \rangle = 0.$$

Hence $\langle \tilde{f}(s), h \rangle = 0$. Now the map (13) is injective by the uniqueness of solutions to ODE's, and it is surjective by an argument similar to the (omitted) proof of (12).

Therefore

$$\begin{aligned}
\operatorname{ind}(\partial_s - A_s) &= \dim(\mathscr{H}^+ \cap \mathscr{H}^-) - \dim((\mathscr{H}^+)^\perp \cap (\mathscr{H}^-)^\perp) \\
&= \dim(\mathscr{H}^+ \cap \mathscr{H}^-) + \dim \operatorname{span}(\mathscr{H}^+, \mathscr{H}^-) - \dim(\mathscr{H}) \\
&= \dim(\mathscr{H}^+) + \dim(\mathscr{H}^-) - \dim(\mathscr{H}) \\
&= \dim(E^-(A^+)) + \dim(E^+(A^-)) - \dim(\mathscr{H}) \\
&= -\dim(E^+(A^+)) + \dim(E^+(A^-)) \\
&= \mathrm{SF}\{A_s\}.
\end{aligned}$$

5.4 Morse-Smale transversality for generic metrics

Proposition 5.8 *Let X be a closed smooth manifold, let k be a positive integer, and let $f : X \to \mathbb{R}$ be a C^{k+1} Morse function on X. Then for a generic C^k metric on X, the pair (f, g) is Morse-Smale.*

Proof. We proceed in three steps.

Step 1 (setup): Fix distinct critical points p, q of f. Let Y be the space of C^k metrics on X; this is a C^∞ Banach manifold. Let Z be the space of locally L^2_1 (in particular continuous[28]) maps $\gamma : \mathbb{R} \to X$ such that:

- $\lim_{s\to -\infty} \gamma(s) = p$, and for $R << 0$, so that $\gamma(-\infty, R]$ is contained in a coordinate chart centered at p, the restriction of γ to $(-\infty, R]$, viewed as a map to $\mathbb{R}^n$ via the coordinate chart, is L^2_1.
- $\lim_{s\to +\infty} \gamma(s) = q$, and γ is analogously L^2_1 on $[R, \infty)$ for $R >> 0$.

Note that Z is a C^∞ Banach manifold[29] with $T_\gamma Z = L^2_1(\gamma^* TX)$, where L^2_1 is defined with respect to the metric on $\gamma^* TX$ obtained by pulling back a fixed metric on X. We define a Banach space bundle $E \to Y \times Z$ by

$$E_{(g,\gamma)} := L^2(\gamma^* TX).$$

[28] The Sobolev embedding theorem asserts that for functions defined on an n-dimensional manifold, there is an embedding $L^p_k \to L^{p'}_{k'}$ whenever $k > k'$ and $k - n/p > k' - n/p'$, which moreover is a compact embedding when the domain is compact. (The number $k - n/p$ is the "conformal weight" which measures how the L^p_k norm on $\mathbb{R}^n$ behaves under scaling of $\mathbb{R}^n$.) So on a 1-manifold, $L^2_1 \subset L^\infty_0 = C^0$, because $1 - 1/2 > 0$.

[29] One can define a coordinate chart for Z around each smooth γ using the exponential map associated to some fixed smooth metric on X.

We define a section ψ of E by

$$\psi(g,\gamma)(s) := \gamma'(s) - V(\gamma(s))$$

where V denotes the negative gradient of f with respect to g as usual[30]. Thus $\psi(g,\gamma) = 0$ if and only if γ is a C^{k+1} negative gradient flow line of f from p to q with respect to g.

Step 2 (applying Theorem 5.4): We claim now that the hypotheses of Theorem 5.4 are satisfied. If $\psi(g,\gamma) = 0$ then

$$\nabla\psi(\dot{g},\dot{\gamma}) = \nabla_{\gamma'}\dot{\gamma} - \nabla_{\dot{\gamma}}V - \dot{V}$$

where on the right side, ∇ is the Levi-Civita connection[31] on $TX \to X$ associated to some fixed smooth metric on X, and $\dot{V}$ denotes the derivative of V with respect to $\dot{g}$.

(a) We claim that $\nabla\psi$ is surjective. To see this suppose that $w \in L^2(\gamma^*TX)$ is orthogonal to the image of $\nabla\psi$. Then for any $\dot{g}$ we have

$$\int \langle \dot{V}, w\rangle ds = 0.$$

Now at any given point in the image of γ, it is an exercise in linear algebra to check that there exists $\dot{g}$ such that $\dot{V} = W$. Since γ is a flow line between distinct critical points, γ is injective, so if we choose $\dot{g}$ supported near that point then we conclude that w is zero there. Hence $w = 0$.

---

[30] Note that the second term in $\psi(g,\gamma)$ is really in L^2, because for instance the restriction of γ to $(-\infty, R]$, viewed as a map to $\mathbb{R}^n$ where the critical point p corresponds to zero, is L^2, and near the critical point we have an estimate $|V(x)| \le c|x|$.

[31] To clarify this calculation: we can extend $\gamma : \mathbb{R} \to X$ to a map $\tilde{\gamma} : \mathbb{R} \times [-1,1] \to X$ with $\tilde{\gamma}(s,0) = \gamma(s)$ and

$$\frac{\partial\tilde{\gamma}(s,t)}{\partial t}\Big|_{s=0} = \dot{\gamma}.$$

Then

$$\begin{aligned}\nabla\psi(0,\dot{\gamma}) &= \nabla_t\left(\frac{\partial\tilde{\gamma}}{\partial s} - V\right)\\ &= \nabla_s\frac{\partial\tilde{\gamma}}{\partial t} - \nabla_t V\end{aligned}$$

where in the second line we have used the torsion-free condition. (Of course $\nabla\psi$ is independent of the connection we choose on X, but the torsion-free condition allows us to write it in this nice way.)

(b) We claim that the restricted differential

$$\dot{\gamma} \longmapsto \nabla_{\gamma'}\dot{\gamma} - \nabla_{\dot{\gamma}} V \tag{14}$$

is Fredholm. To see this, we can choose a trivialization of γ^*TX which is parallel with respect to our chosen connection on TX. Then in this trivialization, the operator (14) has the form (11), where $\mathscr{H} = \mathbb{R}^n$ and A_s is the covariant derivative $\nabla V(\gamma(s)) : T_{\gamma(s)}X \to T_{\gamma(s)}X$ in this trivialization[32]. Now we observe that $\lim_{s\to-\infty} A_s = -H(f,p)$ and $\lim_{s\to+\infty} = -H(f,q)$. Since these are self-adjoint and invertible, Principle 5.6 applies to prove the Fredholm property.

In conclusion, Theorem 5.4 implies that for generic g, the operator (14) is surjective for every flow line γ.

Step 3 (recovering the Morse-Smale condition): To complete the proof, we need to show that surjectivity of (14) implies the Morse-Smale transversality condition. This basically follows from the discussion in Example 5.7. We observe that if γ is a flow line from p to q, then

$$\mathscr{H}^+ = T_{\gamma(0)}\mathscr{D}(p), \quad \mathscr{H}^- = T_{\gamma(0)}\mathscr{A}(q).$$

Since the operator (11) is surjective, its cokernel $(\mathscr{H}^+)^\perp \cap (\mathscr{H}^-)^\perp$ is zero, so $\mathscr{D}(p)$ and $\mathscr{A}(q)$ intersect transversely at $\gamma(0)$. $\square$

Note that Example 5.7 shows that the index of (11) here is $\operatorname{ind}(p) - \operatorname{ind}(q)$, which agrees with our earlier calculation that the moduli space of flow lines (before modding out by the $\mathbb{R}$-action) has dimension $\dim(\mathscr{D}(p) \cap \mathscr{A}(q)) = \operatorname{ind}(p) - \operatorname{ind}(q)$.

Exercises for §5.

1. Verify the isomorphism (9). (This really is tautological if you work through all the notation.)

2. (a) Give a complete proof of Principle 5.6 when $\dim(\mathscr{H}) = 1$.

   (b) Suppose $\dim(\mathscr{H}) = 1$ and $A_s = 0$ for $s > s_0$. Explain why the operator (11) fails to be Fredholm.

[32] Note that the metric compatibility of the connection insures that the metric on γ^*TX induces a well defined metric on $\mathscr{H}$ so that the spaces L^2_1 and L^2 in (11) agree with $L^2_1(\gamma^*TX)$ and $L^2(\gamma^*TX)$.

3. Let V be a vector field on an n-dimensional smooth manifold X. Let us define a **closed orbit** of V to be an embedding[33] $\gamma : S^1 \to X$ such that $\gamma'(s) = \lambda V(\gamma(s))$ for some constant $\lambda > 0$. Let us say that γ is "nondegenerate" if the linearized return map[34] does not have 1 as an eigenvalue. Show that if k is a positive integer, then for a generic C^k vector field, all closed orbits are nondegenerate.

4. It was asserted in §4 that if (f_0, g_0) and (f_1, g_1) are Morse-Smale, then a generic homotopy Γ between them is admissible. Prove this.

5. Prove some genericity statement which you have always wanted to rigorously justify.

6 Morse-Bott theory

The definition of Morse homology that we have given requires that the pair (f, g) be generic, so that the moduli spaces of gradient flow lines are cut out transversely. However for purposes of computation it is often easier to explicitly understand the gradient flow lines of a particular example in a nongeneric case, e.g. when there is symmetry. Morse-Bott theory is an extension of Morse theory to certain cases where the critical points of f are not isolated[35].

6.1 Morse-Bott functions

Definition 6.1 Let X be a closed smooth (finite dimensional) manifold. A function $f : X \to \mathbb{R}$ is **Morse-Bott** if:

(a) The set $\mathrm{Crit}(f)$ of critical points of f is a union of submanifolds of X.

(b) If S is a critical submanifold then for any $p \in S$, the kernel of the Hessian $\nabla df(p) : T_pX \to T_p^*X$ consists only of T_pS, so that for any

[33]Warning: in the literature "closed orbits" are sometimes not required to be embedded.

[34]Let p be a point in the image of γ and let $D \subset X$ be a small $(n-1)$-disc transverse to γ. The **return map** $\phi : D \to D$ takes a point in D and follows its trajectory under V until it hits D again. This is a well-defined diffeomorphism from a small neighborhood of p in D to another small neighborhood of p in D. The eigenvalues of the **linearized return map** $d\phi_p : T_pD \to T_pD$ do not depend on the choice of p or D.

[35]The previous chapter had way too many footnotes. So we won't have any footnotes in this chapter (except of course for this one).

metric on X, the Hessian restricts to an *invertible* self-adjoint map on the normal bundle,

$$H(f,p) : N_pS \to N_pS. \tag{15}$$

If S is a critical submanifold, its index is most naturally regarded as an *interval* $[i_-(S), i_+(S)]$, where $i_-(S)$ is the dimension of the negative eigenspace of the restricted Hessian (15), and $i_+(S) = i_-(S) + \dim(S)$.

A simple example of a Morse-Bott function is the height function on a torus lying on its side. There are two critical submanifolds: a circle of minima of index $[0,1]$, and a circle of maxima of index $[1,2]$.

6.2 The chain complex: first version

Fix a Morse-Bott function f on X. Let g be a generic metric on X and let V be the negative gradient of f with respect to g. We now want to define a chain complex counting flow lines of the vector field V. The treatment here is based on [21], which explains more details, although we are treating orientations and chains differently.

6.2.1 Moduli spaces of flow lines

If S_1, S_2 are two critical submanifolds, a flow line from S_1 to S_2 is a path $\gamma : \mathbb{R} \to X$ such that $\gamma'(s) = V(\gamma(S))$ and $\lim_{s\to-\infty} \gamma(s) \in S_1$ and $\lim_{s\to+\infty} \gamma(s) \in S_2$. Let $\mathscr{M}(S_1, S_2)$ denote the moduli space of flow lines from S_1 to S_2, modulo the $\mathbb{R}$-action by reparametrization as usual. For a generic metric g, the descending manifold of S_1 and the ascending manifold of S_2 will intersect transversely so that

$$\dim \mathscr{M}(S_1, S_2) = i_+(S_1) - i_-(S_2) - 1. \tag{16}$$

(On the other hand, for generic $p_i \in S_i$, the moduli space of flow lines from p_1 to p_2 has dimension $i_-(S_1) - i_+(S_2) - 1$.) There are natural **endpoint maps**

$$e_+ : \mathscr{M}(S_1, S_2) \to S_1, \quad e_- : \mathscr{M}(S_1, S_2) \to S_2$$

sending a flow line γ to $\lim_{s\to-\infty} \gamma(s)$ and $\lim_{s\to+\infty} \gamma(s)$ respectively.

Before continuing, recall that if A, B, C are sets with given maps $i : A \to C$ and $j : B \to C$, then the **fiber product** is defined by

$$A \times_C B := \{(a,b) \mid i(a) = j(b)\} \subset A \times B.$$

If A, B, C are manifolds and the maps i and j are transverse to each other, then $A \times_C B$ is a manifold with

$$\dim(A \times_C B) = \dim(A) + \dim(B) - \dim(C).$$

For a generic metric, the moduli space $\mathscr{M}(S_1, S_2)$ has a compactification to a manifold with corners $\overline{\mathscr{M}(S_1, S_2)}$, whose boundary (codimension one stratum) is the fiber product

$$\partial\overline{\mathscr{M}(S_1, S_2)} = \bigcup_{S'} \mathscr{M}(S_1, S') \times_{S'} \mathscr{M}(S', S_2).$$

Here the union is over all critical submanifolds S' distinct from S_1 and S_2. (The property of the metric required here is that $e_- : \mathscr{M}(S_1, S') \to S'$ is transverse to $e_+ : \mathscr{M}(S', S_2) \to S'$, together with an inductively defined generalization of this which ensures that all iterated fiber products of moduli spaces of flow lines between critical submanifolds are cut out transversely. This holds for a generic metric. Some papers make stronger assumptions, such as that $e_- : \mathscr{M}(S_1, S_2) \to S_1$ is a submersion; while this holds for some important examples and makes certain technicalities nicer, there are many Morse-Bott functions, even on surfaces, for which no metric exists satisfying this assumption.)

6.2.2 Slightly incorrect definition of the chain complex

The rough idea of the chain complex is to define the chain group

$$C_k := \bigoplus_S C_{k-i_-(S)}(S)$$

and the differential

$$D\sigma := \partial\sigma + \sum_{S' \neq S} e_- \left[\sigma \times_S \overline{\mathscr{M}(S, S')}\right],$$

where ∂ is the ordinary differential on singular chains. However this isn't quite right; in order to get the signs to work out one has to modify this a little. We will now be a little more careful and give a correct definition.

6.2.3 Orientations

The signs in Morse-Bott theory are a bit subtle, because the moduli space $\mathscr{M}(S_1, S_2)$ might not be orientable, even when S_1, S_2, and X are all orientable. (It is not hard to cook up an example where S_1 and S_2 are circles and $\mathscr{M}(S_1, S_2)$ is a Klein bottle.) However we can still orient it locally given some choices. More generally, let σ be a generic simplex in S_1 and define

$$\overline{\mathscr{M}(\sigma, S_2)} := \sigma \times_{S_1} \overline{\mathscr{M}(S_1, S_2)}. \tag{17}$$

On the open stratum, if $\gamma \in \mathscr{M}(\sigma, S_2)$ represents a flow line from p_1 to p_2, then we have a natural isomorphism (up to automorphisms of positive determinant)

$$T_{p_1}\sigma \oplus T_{p_1}\mathscr{D}(p_1) \simeq T_\gamma \mathscr{M}(\sigma, S_2) \oplus T\gamma \oplus T_{p_2}\mathscr{D}(p_2). \tag{18}$$

Hence orientations of σ, $\mathscr{D}(p_1)$, and $\mathscr{D}(p_2)$ determine a local orientation of $\mathscr{M}(\sigma, S_2)$.

It is then natural to introduce chains on the critical submanifolds with *twisted coefficients*, so that they have local orientations of the descending manifolds built into them. Namely, there is a locally constant sheaf $\mathscr{O}$ on S, whose stalk at a point $p \in S$ is isomorphic to $\mathbb{Z}$, where an orientation of $\mathscr{D}(p)$ determines such an isomorphism with $\mathbb{Z}$, and the opposite orientation determines the opposite isomorphism. If $i_-(S) > 1$, then one can equivalently describe the stalk at p as

$$\mathscr{O}_p = H_{i_-(S)-1}(\mathscr{D}(p) \setminus p) \simeq \mathbb{Z}.$$

We let $C_*^{\mathrm{sing}}(S, \mathscr{O})$ denote the space of singular chains with coefficients in $\mathscr{O}$. More concretely, $C_*^{\mathrm{sing}}(S, \mathscr{O})$ is the $\mathbb{Z}$-module generated by pairs (σ, o), where σ is a simplex in S and o is a continuously varying orientation of $T\mathscr{D}(p)$ for each p in the image of σ, modulo the relation

$$(\sigma, -o) = -(\sigma, o). \tag{19}$$

For technical reasons as in §3, we actually want to consider only a subspace of currents (with coefficients in $\mathscr{O}$) spanned by pairs (σ, o) where σ is suitably generic. We let $C_*(S, \mathscr{O})$ denote the resulting chain complex. (A simplex σ is suitably generic if it is smooth and if each face of σ is transverse to e_+ of all moduli spaces of flow lines between critical submanifolds and all iterated fiber products thereof.)

6.2.4 The chain complex

We now define a chain complex as follows. The k^{th} chain group is

$$C_k^{\mathrm{Bott}} := \bigoplus_S C_{k-i_-(S)}(S, \mathscr{O}).$$

We define $D : C_k^{\mathrm{Bott}} \to C_{k-1}^{\mathrm{Bott}}$ as follows. If $\sigma \in C_*(S, \mathscr{O})$ is a generic simplex with locally oriented descending manifolds, and if $S' \neq S$, then we have a well-defined current

$$e_- \left[\overline{\mathscr{M}(\sigma, S')}\right] \in C_*(S', \mathscr{O}).$$

Thanks to (18) and (19), we have just enough orientation data for this to be well-defined. Furthermore if $\dim(\sigma) = k - i_-(S)$ then

$$\begin{aligned}\dim(\mathscr{M}(\sigma, S')) &= (k - i_-(S)) + (i_+(S) - i_-(S') - 1) - \dim(S) \\ &= k - 1 - i_-(S').\end{aligned}$$

So it makes sense to define

$$D\sigma := \partial\sigma + \sum_{S' \neq S} e_- \left[\overline{\mathscr{M}(\sigma, S')}\right].$$

Lemma 6.2 *$D^2 = 0$.*

Proof. We omit the signs. For the proof we use the fiber product interpretation (17). We note that

$$\partial\left(\sigma \times_S \overline{\mathscr{M}(S, S')}\right) = \partial\sigma \times_S \overline{\mathscr{M}(S, S')} \bigcup \sigma \times_S \partial\overline{\mathscr{M}(S, S')}.$$

We then have

$$\begin{aligned} D^2\sigma = \partial^2\sigma + \partial \sum_{S' \neq S} e_- \left[\sigma \times_S \overline{\mathscr{M}(S, S')}\right] + \sum_{S' \neq S} e_- \left[\partial\sigma \times_S \overline{\mathscr{M}(S, S')}\right] \\ + \sum_{S'' \neq S' \neq S} e_- \left[\left(\sigma \times_S \overline{\mathscr{M}(S, S')}\right) \times_{S'} \overline{\mathscr{M}(S', S'')}\right]. \end{aligned}$$

The first term is zero, the sum of the second and third terms is

$$\sum_{S' \neq S} e_- \left[\sigma \times_S \partial\overline{\mathscr{M}(S, S')}\right]$$

(up to sign), and this equals the fourth term. □

We define the **Morse-Bott homology** $H_*^{\mathrm{Bott}}(f, g)$ to be the homology of the chain complex (C_*^{Bott}, D).

Example 6.3 Consider again our example of a Morse-Bott function on the torus with two critical submanifolds, one a circle S_0 of minima and the other a circle S_1 of maxima. Then

$$C_*^{\text{Bott}} = C_*(S_0, \mathscr{O}) \oplus C_*(S_1, \mathscr{O})[1].$$

Here the notation [1] indicates that the grading is shifted upward by 1. In this example all simplices in the critical submanifolds are generic.

(a) If we choose a symmetric metric, then for each point in S_1 there are two flow lines to the same point in S_0. Then the differential is given simply by

$$D((\sigma_0, o), (\sigma_1, o)) = ((\partial\sigma_0, o), (\partial\sigma_1, o)).$$

Note that $D(0, (\sigma_1, o))$ has no component in $C_*(S_0)$, because $\mathscr{M}(\sigma_1, S_0)$ consists of two copies of σ_1 which contribute with opposite signs. Hence

$$H_*^{\text{Bott}} = H_*(S_0, \mathscr{O}) \oplus H_*(S_1, \mathscr{O})[1]. \tag{20}$$

Since the orientation sheaf $\mathscr{O}$ is trivial here, $H_*(S_i, \mathscr{O}) \simeq H_*(S^1)$.

(b) If the metric on the torus is not symmetric then the two flow lines from a given point in S_1 may have different lower endpoints in S_0. But with a bit more work one can see that (20) still holds.

Example 6.4 Starting with the previous example, do surgery on a horizontal circle of the torus to obtain a Morse-Bott function on S^2 with a circle S_0 of minima, a circle S_1 of maxima, an isolated minimum m_0, and an isolated maximum m_2. In this example again, all simplices in the critical submanifolds are generic, and all orientation sheaves are trivial. Up to orientations, if p is any point in S_1 then we have

$$Dp = \pm m_0 \pm \phi(p)$$

where $\phi : S_1 \to S_0$ is a diffeomorphism. We also have

$$Dm_1 = \pm[S_0].$$

These are the only components of D that relate different critical submanifolds. It follows fairly readily that

$$H_*^{\text{Bott}} \simeq H_0(S_0, \mathscr{O}) \oplus H_1(S_1, \mathscr{O})[1].$$

6.2.5 The homology

Theorem 6.5 *If f_0 and f_1 are two Morse-Bott functions with generic metrics g_0 and g_1, then there is a canonical isomorphism*

$$H_*^{\mathrm{Bott}}(f_0, g_0)) \simeq H_*^{\mathrm{Bott}}(f_1, g_1).$$

Proof. This is an extension of the arguments in §4, defining Morse-Bott versions of the continuation maps and chain homotopies by analogy with the definition of the Morse-Bott differential. □

Corollary 6.6 *For any Morse-Bott function f_0 and generic metric g_0, there is a canonical isomorphism*

$$H_*^{\mathrm{Bott}}(f_0, g_0) \simeq H_*(X).$$

Proof. Let $f_1 = 0$ and let g_1 be any metric. Then by definition, $H_*^{\mathrm{Bott}}(f_1, g_1) = H_*(X)$. □

Remark 6.7 In particular any Morse function is Morse-Bott, and the Morse-Bott complex then agrees with the Morse complex, so this gives another proof of Theorem 3.1. This may make §3 appear retrospectively superfluous, but in fact the work needed to flesh out the details of the proof of Theorem 6.5 is similar to the work done in §3; and for natural choices of homotopies one can see that the two proofs of Theorem 3.1 have essentially the same content.

6.3 An example from symplectic geometry

We now present, following [6], a quick application of Corollary 6.6. This example requires some basic symplectic geometry as in [45].

Let (M, ω) be a closed symplectic manifold, and suppose there is Hamiltonian S^1 action on M with moment map $f : M \to \mathbb{R}$. Then the critical points of f are the fixed points of the action. It is known from symplectic geometry that f is a Morse-Bott function; the S^1 representation on the normal bundle to a critical submanifold has no trivial components and thus splits as a sum of 2-dimensional components. In particular, a critical submanifold is even dimensional, and its index is also even, namely twice the number of components on which the S^1 action has positive weights. The orientation sheaf $\mathscr{O}$ over a critical submanifold is naturally trivialized by the symplectic form.

We claim now that f is a **perfect Morse-Bott function**, i.e.

$$\begin{aligned} H_*(M) &= \bigoplus_S H_*(S, \mathscr{O})[i_-(S)] \\ &\simeq \bigoplus_S H_*(S)[i_-(S)]. \end{aligned}$$

Equivalently $D = \partial$, i.e. the Morse-Bott differential D sends a current in a critical submanifold to another current in the same critical submanifold.

The idea of the proof is simple. We need to choose a generic metric g which is also S^1-invariant, and this can be done (I think). Then if S_1, S_2 are two distinct critical submanifolds, S^1 acts nontrivially on $\mathscr{M}(S_1, S_2)$, while fixing S_1 and S_2. This means that the endpoint map

$$e_+ \times e_- : \mathscr{M}(S_1, S_2) \to S_1 \times S_2$$

factors through $\mathscr{M}(S_1, S_2)/S^1$, and so its image has dimension one less than expected. Hence if $\sigma \in C_{k-i_-(S_1)}(S_1)$ is a generic simplex, then $e_-(\sigma \times_{S_1} \mathscr{M}(S_1, S_2))$ is supported in a current of dimension $k - i_-(S_2) - 2$, and hence is zero when regarded as a current of dimension $k - i_-(S_2) - 1$.

In fact, general results of [4, 39] imply that f is *equivariantly* perfect, i.e. the S^1-equivariant cohomology of M is the sum of the equivariant cohomologies of the fixed point sets. For a treatment of equivariant cohomology via Morse-Bott theory, see [6].

6.4 The Morse-Bott spectral sequence(s)

We claimed that Morse-Bott theory would simplify computations, but it may appear that we have taken a step backward by replacing the finite dimensional Morse complex with the the infinite dimensional Morse-Bott complex. However it is possible to compute the homology of the Morse-Bott complex by first passing to the homology of the critical submanifolds, and then defining differentials on the homology of the critical submanifolds. To do this we need to use the spectral sequence associated to a filtered complex, see e.g. [10, 27].

6.4.1 The weakly self-indexing case

Let f be a Morse-Bott function and let g be a generic metric. The pair (f, g) is **weakly self-indexing** if $\mathscr{M}(S, S') = \emptyset$ whenever $i_-(S) < i_-(S')$. In this case i_- defines a filtration on the complex (C_*^{Bott}, D), namely

$$\mathscr{F}_i C_*^{\mathrm{Bott}} = \bigoplus_{i_-(S)\le i} C_*(S,\mathscr{O})[i_-(S)].$$

We then obtain a spectral sequence which converges to the Morse-Bott homology, with

$$E^1_{p,q} = \bigoplus_{i_-(S)=p} H_q(S,\mathscr{O}).$$

The first differential

$$\partial_1 : E^1_{p,q} \to E^1_{p-1,q}$$

is defined as follows. Given $\alpha \in H_q(S,\mathscr{O})$, we choose a cycle C representing it. For each S' with $i_-(S') = p-1$, the weakly self-indexing assumption implies that $\mathscr{M}(S,S')$ is a compact manifold with no boundary. Thus $e_+^{-1}(C) = C \times_S \mathscr{M}(S,S')$ is a cycle in $\mathscr{M}(S,S')$, and its pushforward by e_- is a cycle in S'. Then up to orientations,

$$\partial_1(\alpha) = \sum_{i_-(S')=p-1} \pm \left[e_-(C\times_S \mathscr{M}(S,S'))\right].$$

The higher differentials in the spectral sequence are more subtle. However they are given by a formula similar to the formula for ∂_1 in the simple case when there are no broken flow lines involved. If we are lucky the other differentials will vanish due to the bigrading on the spectral sequence so that we can compute the Morse-Bott homology by computing the homology of ∂_1.

6.4.2 The general case

Although the weakly self-indexing case is nice, there is always (at least when f is real-valued!) an obvious filtration given by f itself. Namely one can order the critical submanifolds as $S_1, S_2, \cdots$ with $f(S_i) \le f(S_j)$ for $i < j$. Then we have the filtration

$$\mathscr{F}_i C_*^{\mathrm{Bott}} = \bigoplus_{j\le i} C_*(S_j,\mathscr{O})[i_-(j)]$$

with an associated spectral sequence, whose E^1 term is the sum of the (twisted, grading-shifted) homologies of the critical submanifolds. For an application of this spectral sequence where the weakly self-indexing condition does not hold, see [66].

This spectral sequence is essentially what we used (without explicitly saying so) to work out Examples 6.3 and 6.4, which might now be worth revisiting.

6.5 Another Morse-Bott complex

We now sketch another approach to Morse-Bott theory which we learned about from [11]. The idea is as follows.

First, we can always perturb a Morse-Bott function f to obtain a Morse function. Explicitly, for each critical submanifold S_i, choose a Morse function $f_i : S_i \to \mathbb{R}$. We extend this to some smooth function $\widetilde{f}_i : X \to \mathbb{R}$. For $\epsilon \in \mathbb{R}$, define

$$f_\epsilon := f + \epsilon \sum_i \widetilde{f}_i : X \to \mathbb{R}.$$

If $\epsilon > 0$ is small, then f_ϵ is a Morse function and we have a one-to-one correspondence

$$\mathrm{Crit}(f_\epsilon) = \bigcup_i \mathrm{Crit}(f_i).$$

Moreover, if $p \in \mathrm{Crit}(f_i)$, then the index of the corresponding critical point of f_ϵ is $\mathrm{ind}(p) + i_-(S_i)$. In particular, the indices of the critical points of f_ϵ on S_i lie in the interval $[i_-(S_i), i_+(S_i)]$.

If g is a generic metric on X, then (f_ϵ, g) will be Morse-Smale. Now the key point is that we can read off the Morse differential $\partial_\epsilon^{\mathrm{Morse}}$ for (f_ϵ, g) from the Morse-Bott setup (f, g), without actually carrying out the perturbation. In this way we obtain a finite-dimensional complex from the Morse-Bott data.

Here is how it works. If p, q are critical points of f_i on the same S_i, then $\langle \partial_\epsilon^{\mathrm{Morse}} p, q\rangle$ is determined by our choice of f_i in a way which we already in principle understand. And more interestingly, if $p \in \mathrm{Crit}(f_i)$ and $q \in \mathrm{Crit}(f_j)$ with $i \neq j$, then for ϵ sufficiently small, up to sign we have

$$\left\langle \partial_\epsilon^{\mathrm{Morse}} p, q \right\rangle = \#\mathscr{M}(\mathscr{D}(p), \mathscr{A}(q)). \tag{21}$$

Here $\mathscr{D}(p)$ is the descending manifold of f_i in S_i, $\mathscr{A}(q)$ is the ascending manifold of f_j in S_j, and $\mathscr{M}(\mathscr{D}(p), \mathscr{A}(q))$ is the set of flow lines γ for (f, g) with $\lim_{s\to-\infty} \gamma(s) \in \mathscr{D}(p)$ and $\lim_{s\to+\infty} \gamma(s) \in \mathscr{A}(q)$.

We leave it to the reader to ponder why (21) might be true.

Exercises for §6.

1. Justify equation (16).
2. Justify equation (18).
3. Work out Example 6.3 explicitly (without using a spectral sequence).
4. Find an example of a Morse-Bott function such that for at least one of the critical submanifolds, the orientation sheaf $\mathcal{O}$ is nontrivial. Compute the Morse-Bott homology for your example.
5. Let $\pi : Z \to B$ be a fiber bundle of closed smooth manifolds. Let $f : B \to \mathbb{R}$ be a Morse function.
   (a) Show that $\pi^* f : Z \to \mathbb{R}$ is a Morse-Bott function. Show that for each critical submanifold the orientation sheaf $\mathcal{O}$ is trivial. Show that for any generic metric on Z, the pair (f, g) is weakly self-indexing.
   (b) Now that you are warmed up, see if you can show that the Morse-Bott spectral sequence for $\pi^* f$ using the i_- filtration agrees, from the E^2 term on, with the Leray-Serre spectral sequence for the fiber bundle $Z \to B$. (I have seen this last point asserted many times, but I have never seen the proof.)
6. If you haven't seen spectral sequences before, do some examples until you get the hang of it.

7 Morse theory for circle-valued functions and closed 1-forms

Many functionals that arise in Floer theory are not $\mathbb{R}$-valued but rather $\mathbb{R}/\mathbb{Z}$-valued. Thus it is important to understand Morse theory for such functions. In fact, if f is a real-valued or circle-valued function on X, then after a metric is chosen, the gradient flow depends only on the closed 1-form df. When f is a real-valued or circle-valued function, the cohomology class of df in $H^1(X; \mathbb{R})$ is zero or the image of an integral cohomology class, respectively[36];

[36] Recall that there is a natural bijection $[X, S^1] = H^1(X; \mathbb{Z})$, which sends a homotopy class of map $f : X \to S^1$ to the pullback by f of the fundamental class in $H^1(S^1; \mathbb{Z})$. If f is smooth then the cohomology class $[df] \in H^1(X; \mathbb{R})$ is the image of the corresponding element of $H^1(X; \mathbb{Z})$ under the map $H^1(X; \mathbb{Z}) \to H^1(X; \mathbb{R})$.

but in fact one can set up Morse theory for an arbitrary closed one-form and this is important as well. Morse theory for circle-valued functions and more generally for closed 1-forms was first considered by Novikov [48], and there have been many subsequent papers on the subject. The style of this chapter is closest to [31].

7.1 Compactness

In some respects, the Morse theory of closed 1-forms is not much different from the Morse theory of real-valued functions. Let X be a closed smooth manifold and let α be a closed 1-form on X. Locally any closed 1-form is d of a real-valued function so it makes sense to define "Morse closed 1-forms". Namely, a "critical point" of α is a zero of α; the critical point p is "nondegenerate" if $\nabla\alpha : T_pX \to T_p^*X$ is invertible; and α is "Morse" if all critical points are nondegenerate. The index of a critical point is defined as before. We choose a metric g on X and let V denote the vector field dual to $-\alpha$ via g. It then makes sense to speak of flow lines of V between critical points. We say the pair (α, g) is "Morse-Smale" if the ascending and descending manifolds of all critical points intersect transversely; if α is Morse, then this condition holds for a generic metric. We let $\mathscr{M}(p,q)$ denote the moduli space of flow lines from p to q as before.

We want to define a chain complex counting gradient flow lines between critical points of index difference one. An important difference with the real-valued case is that compactness does not always hold as before. When $\operatorname{ind}(p) - \operatorname{ind}(q) = 1$ and (α, g) is Morse-Smale, the moduli space $\mathscr{M}(p,q)$, although zero-dimensional, might not be finite. The idea is that there can be a sequence of flow lines which wrap around the manifold more and more times, so that the sequence has no convergent subsequence.

Fortunately, we can still get compactness and finite counts if we classify flow lines according to their some information about their (relative) homology classes. To prepare for this and to clarify the issues with compactness, we will now prove a compactness result. The argument here is pretty standard, cf. [52], and is written in such a way that it generalizes to infinite dimensional settings (although a number of additional issues have to be dealt with to prove compactness in Floer theory).

In the following we regard a flow line as a map $\gamma : \mathbb{R} \to X$; we do not mod out by the $\mathbb{R}$ action. Let p and q be critical points of α

Definition 7.1 A (k-times) **broken flow line** from p to q is a set of flow lines $\widehat{\gamma} = (\widehat{\gamma}_0, \ldots, \widehat{\gamma}_k)$ where k is a nonnegative integer and there exist critical points $r_0, \ldots, r_{k+1}$ with $r_0 = p$ and $r_{k+1} = q$ such that $\widehat{\gamma}_i$ is a flow line from r_i to r_{i+1}.

Definition 7.2 A sequence of flow lines $\gamma_n : \mathbb{R} \to X$ from p to q **converges** to a broken flow line $\widehat{\gamma} = (\widehat{\gamma}_0, \ldots, \widehat{\gamma}_k)$ from p to q if:

- There exist real numbers
$$s_{n,0} < s_{n,1} < \cdots < s_{n,k}$$
such that
$$\gamma_n(s_{n,i} + \cdot) \longrightarrow \widehat{\gamma}_i$$
in C^∞ on compact sets.
- For n sufficiently large, $\gamma_n - \sum_{i=0}^k \widehat{\gamma}_i$ is homologous to zero[37].

Definition 7.3 If $\gamma : \mathbb{R} \to X$ is a map with $\gamma'(s) = V(\gamma(s))$, define the **energy**
$$E(\gamma) := \int_{s=-\infty}^{\infty} |V(\gamma(s))|^2 ds = \int_\gamma (-\alpha) \in [0, \infty]. \tag{22}$$

Lemma 7.4 *(a) $E(\gamma) \geq 0$, with equality if and only if γ is a constant map to a critical point.*

(b) If $E(\gamma) < \infty$ then γ is a flow line between two critical points.

(c) There exists $\delta > 0$ such that any nonconstant flow line γ between two critical points satisfies $E(\gamma) > \delta$.

(Parts (b) and (c) require our assumption that X is compact and α is Morse.)
Proof. (a) is obvious, as the local contribution to the integral (22) is nonnegative, and zero only at critical points.

(b) We need to show that $\gamma(s)$ converges to a critical point as $s \to +\infty$. We can find $\epsilon > 0$ such that the ϵ-balls around the critical points are disjoint.

[37]This already follows from the first condition if the Morse-Smale condition holds. Without the Morse-Smale condition, or in certain infinite dimensional settings, the limiting broken flow line could include a flow line from a critical point to itself, and we want to keep track of this.

It is then enough to show that there exists s_0 such that $\mathrm{dist}(\gamma(s), \mathrm{Crit}(\alpha)) < \epsilon$ for all $s > s_0$. If no such s_0 exists, then we can find a sequence $s_n \to \infty$ with

$$\mathrm{dist}(\gamma(s_n), \mathrm{Crit}(\alpha)) \geq \epsilon \tag{23}$$

for all n. We can pass to a subsequence so that the points $\gamma(s_n)$ converge in X. Then, because the solution to an ODE depends smoothly on the initial condition, the reparametrized maps $\gamma(s_n+\cdot)$ converge in C^∞ on compact sets to a map $\widetilde{\gamma} : X \to \mathbb{R}$ with $\widetilde{\gamma}'(s) = V(\widetilde{\gamma}(s))$. Since $s_n \to \infty$ and $E(\gamma) < \infty$, it follows that $E(\widetilde{\gamma}) = 0$, so $\widetilde{\gamma}$ is a constant map to a critical point, but this contradicts (23) since $\widetilde{\gamma}(0) = \lim_{n\to\infty} \gamma(s_n)$. Likewise, $\gamma(s)$ also converges to a critical point as $s \to -\infty$.

(c) If not, then we can find a sequence γ_n of nonconstant flow lines between two fixed critical points with $E(\gamma_n) \to 0$. For each n there exists a real number s_n satisfying (23) (or else γ_n would be supported in a neighborhood of a critical point, in which case for homological reasons $E(\gamma_n) = 0$ so γ_n would be constant). We can pass to a subsequence so that $\gamma_n(s_n + \cdot)$ converges, in C^∞ on compact sets, to a flow line γ, which must have energy zero and thus must be a constant map to a critical point; but (23) implies that $\gamma(0)$ has distance at least ϵ from all critical points, a contradiction. □

Proposition 7.5 *Let α be a Morse closed 1-form and g a metric on a closed smooth manifold X. Let p and q be critical points of α, and let $\gamma_n : \mathbb{R} \to X$ be a sequence of flow lines from p to q. Assume (this is crucial) that*

- *There exists a constant C such that $E(\gamma_n) < C$ for all n.*

Then after passing to a subsequence, γ_n converges to a broken flow line $\widehat{\gamma}$.

Proof. Before starting, we pass to a subsequence so that $E(\gamma_n) \to C_0$. As before there exists $\epsilon > 0$ such that the ϵ-balls around the critical points are disjoint.

We can assume that the flow lines γ_n are nonconstant for sufficiently large n, as otherwise the proposition is trivially true. It then makes sense to define

$$s_{n,0} := \inf\{s \in \mathbb{R} \mid \mathrm{dist}(\gamma_n(s), p) \geq \epsilon\}.$$

We can pass to a subsequence so that $\gamma_n(s_{n,0}+\cdot)$ converges in C^∞ on compact sets to a map $\widehat{\gamma}_0$ with $\widehat{\gamma}_0'(s) = V(\widehat{\gamma}_0(s))$. By the C^∞ convergence on compact sets, $E(\widehat{\gamma}_0) \leq 0$, and in particular $\widehat{\gamma}_0$ is a flow line from p to some critical point r_1 by Lemma 7.4.

If $E(\widehat{\gamma}_0) = C_0$, then $\gamma_n \to \widehat{\gamma} = (\widehat{\gamma}_0)$ and we are done.

Suppose $E(\widehat{\gamma}_0) < C_0$. Since $\lim_{s\to\infty} \widehat{\gamma}_0(s) = r_1$, there exists $\psi_0 \in \mathbb{R}$ such that $\mathrm{dist}(\widehat{\gamma}_0(s), r_1) < \epsilon/2$ whenever $s > \psi_0$. For large n, the flow line γ_n cannot be in the same relative homology class as $\widehat{\gamma}_0$, so $\gamma_n((s_{n,0} + \psi_0, \infty)) \not\subset B(r_1, \epsilon)$. We then define

$$s_{n,1} := \inf\{s > s_{n,0} + \psi_0 \mid \mathrm{dist}(\gamma_n(s), r_1) \geq \epsilon\}.$$

We can pass to a subsequence so that $\gamma_n(s_{n,1} + \cdot)$ converges in C^∞ on compact sets to a flow line $\widehat{\gamma}_1$, from r_1 to some critical point r_2, with $E(\widehat{\gamma}_0) + E(\widehat{\gamma}_1) \leq C_0$.

If $E(\widehat{\gamma}_0) + E(\widehat{\gamma}_1) = C_0$, then $\gamma_n \to \widehat{\gamma} = (\widehat{\gamma}_0, \widehat{\gamma}_1)$ and we are done. If not, we continue this process, inductively defining

$$s_{n,j} := \inf\{s > s_{n,j-1} + \psi_{j-1} \mid \mathrm{dist}(\gamma_n(s), r_j) \geq \epsilon\}.$$

By Lemma 7.4, this process must terminate in at most $\lfloor C_0/\delta \rfloor$ steps. □

7.2 Novikov rings

We now need to introduce the Novikov ring, cf. [30], which is basically an algebraic bookkeeping device. It is a simple generalization of the group ring of a group and the ring of Laurent series.

Definition 7.6 Let G be an abelian group and let $N : G \to \mathbb{R}$ be a homomorphism. Define the **Novikov ring** $\mathrm{Nov}(G; N)$ as follows. An element of $\mathrm{Nov}(G; N)$ is a formal (possibly infinite) linear combination[38].

$$a = \sum_{g\in G} a_g g$$

where the a_g's are integers, such that

[38] More precisely, a Novikov ring element is a function $a : G \to \mathbb{Z}$ satisfying the finiteness condition (*). Writing these as formal linear combinations can be confusing because the expression $g_1 + g_2$ has two possible meanings: it could be the function sending $g_1, g_2 \mapsto 1$, which is ususally what we mean, or the function sending $g_1 + g_2 \mapsto 1$ (and all other elements to zero in both cases). To avoid this ambiguity, some people write elements of the Novikov ring as $\sum_{g\in G} a_g e^g$, with e^g regarded as a formal symbol. Then also the multiplication rule has the nice form $e^g e^{g'} = e^{g+g'}$.

(*) For all $R \in \mathbb{R}$, there are only finitely many $g \in G$ with $a_g \neq 0$ and $N(g) < R$.

If $b = \sum_{g \in G} b_g g$ we define $a + b := \sum_{g\in G}(a_g + b_g)g$ and

$$ab := \sum_{g \in G} \left(\sum_{g' \in G} a_{g'} b_{g-g'} \right) g.$$

It is an exercise in logic to check that the finiteness condition (*) implies that the coefficient of g in ab is a sum of only finitely many nonzero terms, and that ab again satisfies the finiteness condition (*).

Note that there is an inclusion of the group ring into the Novikov ring, $\mathbb{Z}[G] \to \mathrm{Nov}(G; N)$, which is an isomorphism if and only if $N \equiv 0$.

Example 7.7 The simplest example is when $G = \mathbb{Z}$ and $N : \mathbb{Z} \to \mathbb{R}$ is the inclusion. Then we can identify $\mathrm{Nov}(G; N)$ with the ring $\mathbb{Z}((t))$ of formal integer Laurent series $\sum_{m=m_0}^{\infty} a_m t^m$ where m_0 and the a_m's are integers. (The identification sends an integer $m \in \mathbb{Z}$ to the symbol t^{-m}.)

7.3 The Novikov complex

Now let α be a Morse closed 1-form on a closed connected smooth manifold X and let g be a metric such that the pair (α, g) is Morse-Smale.

Choose a connected abelian covering $\pi : \tilde{X} \to X$ such that $\pi^*\alpha$ is exact. We can always do this; for example, we can take $\tilde{X}$ to be the universal abelian covering of X, which has $H_1(\tilde{X}) = 0$. For a general abelian covering, the group H of covering transformations is the quotient of $H_1(X)$ by the subgroup consisting of homology classes of loops that lift to $\tilde{X}$. That is, we have a short exact sequence

$$0 \longrightarrow H_1(\tilde{X}) \longrightarrow H_1(X) \longrightarrow H \longrightarrow 0.$$

Since $\pi^*\alpha$ is assumed exact, the pairing with $[\alpha]$ from $H_1(X) \to \mathbb{R}$ descends to a map $H \to \mathbb{R}$.

We now define the **Novikov complex** $(C_*^{\mathrm{Nov}}, \partial^{\mathrm{Nov}})$ as follows. (This depends on α, g, and the choice of covering π.) Choose $\tilde{f} : X \to \mathbb{R}$ with

$$d\tilde{f} = \pi^*\alpha.$$

Let C_i^{Nov} be the set of formal linear combinations

$$\sum_{\tilde{p}\in\mathrm{Crit}_i(\tilde{f})} a_{\tilde{p}}\tilde{p}$$

where the $a_{\tilde{p}}$'s are integers, such that

(**) for all $R \in \mathbb{R}$, there are only finitely many $\tilde{p}$ with $\tilde{f}(\tilde{p}) > R$ and $a_{\tilde{p}} \neq 0$.

It is another exercise in logic to check that C_i^{Nov} is a module over the Novikov ring

$$\Lambda := \mathrm{Nov}(H, [-\alpha]),$$

where the module structure is induced by the action of H on $\tilde{X}$ by covering transformations. Moreover, this module is free: one can obtain a basis by choosing a lift of each index i critical point in X to $\tilde{X}$.

We now define the differential $\partial^{\mathrm{Nov}} : C_i^{\mathrm{Nov}} \to C_{i-1}^{\mathrm{Nov}}$ by counting flow lines as usual: if $\tilde{p} \in \mathrm{Crit}_i(\tilde{f})$ then

$$\partial\tilde{p} := \sum_{\tilde{q}\in\mathrm{Crit}_{i-1}(\tilde{f})} \#\mathscr{M}(\tilde{p},\tilde{q})\cdot\tilde{q}.$$

Here $\mathscr{M}$ denotes the moduli space of flow lines of $\tilde{f}$ with respect to the pullback to $\tilde{X}$ of our chosen metric g on X. The signs are determined as in the Morse complex; one chooses orientations of the descending manifolds of the critical points in X, and pulls these back to orientations of the descending manifolds in $\tilde{X}$. It is a third exercise in logic to check that the finiteness condition (**) and the compactness proposition 7.5 imply that ∂ is well defined.

Note that if p and q are critical points in X and $\tilde{p}$ and $\tilde{q}$ are lifts to $\tilde{X}$ then a flow line from $\tilde{p}$ to $\tilde{q}$ projects to a flow line from p to q, although a flow line from p to q might not lift to a flow line from $\tilde{p}$ to $\tilde{q}$; the obstruction to finding such a lift is an element of H. Although there may be infinitely many flow lines from p to q, the point is that by working in a covering such that $\pi^*\alpha$ is exact, we classify flow lines by enough homotopy information to ensure that the coefficients in the differential are finite.

The usual argument shows that $(\partial^{\mathrm{Nov}})^2 = 0$. We denote the homology of the complex $(C_*^{\mathrm{Nov}}, \partial^{\mathrm{Nov}})$ by H_*^{Nov}.

7.4 The Novikov homology

Lemma 7.8 *The Novikov homology H_*^{Nov} depends only on the cohomology class $[\alpha] \in H^1(X;\mathbb{R})$ and the choice of connected abelian cover $\pi : \tilde{X} \to X$.*

The proof of this lemma follows the usual continuation argument. The only subtlety is that one has to restrict to families $\{(\alpha_t, g_t)\}$ in which all of the forms α_t are in the same cohomology class (or at least in the same ray emanating from the origin in $H^1(X;\mathbb{R})$). This is necessary so that one can apply a version of the compactness proposition 7.5 to insure that the continuation maps involve finite counting and so are well-defined.

Note that if $[\alpha_0]$ and $[\alpha_1]$ are in different rays in $H^1(X;\mathbb{R})$, then in general it is difficult to compare the Novikov homologies for α_0 and α_1, since they are modules over different Novikov rings. (We will see one situation where this can be done in the proof of Theorem 7.11 below.)

Remark 7.9 One might guess that $H_*^{\mathrm{Nov}} \simeq H_*(X) \otimes \Lambda$. But in fact that is hardly ever true except in some trivial cases. For example, if $[\alpha] \neq 0$ and $\dim(X) = n$, then H_n^{Nov} always vanishes, see eg. example 7.10(b) below.

Example 7.10 We now consider three examples with $X = S^1$.

(a) Let $\alpha = df$ where $f : S^1 \to S^1$ is the identity. We use the covering $\pi : \mathbb{R} \to S^1$ with covering group $\mathbb{Z}$ so that $\Lambda \simeq \mathbb{Z}((t))$. Since f has no critical points the Novikov homology is trivial.

(b) Now perturb f above so that it has a local maximum p and a local minimum q. By the above lemma, the Novikov homology is still trivial; let us try to understand this explicitly. There are two flow lines from p to q but they are not in the same relative homology class. We can choose lifts $\tilde{p}$ and $\tilde{q}$ of p and q such that

$$\partial^{\mathrm{Nov}} \tilde{p} = \pm(1-t)\tilde{q}. \tag{24}$$

Now $(1-t)$ is invertible in $\mathbb{Z}((t))$:

$$(1-t)^{-1} = 1 + t + t^2 + \cdots .$$

Hence $H_0^{\mathrm{Nov}} = 0$ because

$$\tilde{q} = \pm \partial^{\mathrm{Nov}} (1-t)^{-1} \tilde{p}.$$

Also, $H_1^{\text{Nov}} = 0$ because there are no cycles since $\partial\tilde{p} \neq 0$. (The ring $\mathbb{Z}((t))$ has no zero divisors, although Novikov rings of abelian groups with torsion do.) It is tempting to try to define a 1-cycle as $\sum_{n\in\mathbb{Z}} t^n\tilde{p}$, but this expression is not in C_1^{Nov} because it does not satisfy the finiteness condition (**).

(c) Let $\alpha = df$ where $f : S^1 \to \mathbb{R}$ is a real-valued function with two critical points. We could choose the covering $\tilde{X} = X$, but that would be boring because then the Novikov ring $\Lambda = \mathbb{Z}$ and we would be reduced to the usual Morse complex. So let us choose the covering $\tilde{X} = \mathbb{R}$ so that the Novikov ring is the group ring

$$\Lambda = \mathbb{Z}[H_1(X)] \simeq \mathbb{Z}[t, t^{-1}].$$

Then (24) still holds so that $H_1^{\text{Nov}} = 0$ as before, but now $H_0^{\text{Nov}} \neq 0$ because $(1 - t)$ is not invertible in the group ring. All we can say is that H_0^{Nov} is a $\mathbb{Z}[t, t^{-1}]$ module with one generator which is annihilated by $1 - t$.

The Novikov complex does have a topological counterpart. Choose a cell decomposition of X. We can lift the cells to obtain a cell decomposition of $\tilde{X}$. The cell-chain complex $C_*^{\text{cell}}(\tilde{X})$ is then a module over $\mathbb{Z}[H]$, where H acts by covering transformations. We then have

Theorem 7.11 *We have an isomorphism*

$$H_*^{\text{Nov}} \simeq H_*\left(C_*^{\text{cell}}(\tilde{X}) \otimes_{\mathbb{Z}[H]} \Lambda\right).$$

By standard arguments, the homology of the complex on the right hand side is isomorphic to the homology of the complex of "half-infinite singular chains", namely locally finite singular chains in $\tilde{X}$ such that for each real number R, only finitely many simplices hit $\tilde{f}^{-1}((R, \infty))$. Example 7.10(b) now makes sense: $H_0^{\text{Nov}} = 0$ because a point is the boundary of half the line, and $H_1^{\text{Nov}} = 0$ because there are no 1-cycles because a 1-chain can only be infinite in the downward direction.

One can prove Theorem 7.11 along the lines of the proof of Theorem 3.1, and in fact such a proof shows that the isomorphism is canonical. The isomorphism sends a critical point in $\tilde{X}$ to its descending manifold in $\tilde{X}$, viewed as a half-infinite chain. However, in order to introduce some useful ideas in finite-dimensional Morse theory, we will give a different proof here.

Proof of Theorem 7.11. We consider two cases.

Case A: suppose $[\alpha] = 0$ so that $\alpha = df$ where $f : X \to \mathbb{R}$. Then $\Lambda = \mathbb{Z}[H]$, and we need to show that

$$H_*^{\mathrm{Nov}} \simeq H_*(\tilde{X})$$

as $\mathbb{Z}[H]$-modules. This can be proved almost the same way as Theorem 3.1, where one just does everything in $\tilde{X}$. Since $\tilde{f}$ is the pullback of a real-valued function on X, there are no compactness difficulties, even though $\tilde{X}$ need not be compact.

Just for fun, here is a sketch of another proof of Case A. By Lemma 7.8 it is sufficient to prove the theorem for a single Morse-Smale pair (f, g) of our choice. Choose a smooth triangulation of X. One can apparently[39] find a Morse-Smale pair (f, g) such that f has a critical point of index i at the center of each i-simplex and one gradient flow line from the center of a simplex to the center of each face, so that there is an isomorphism of chain complexes (and differentials omitted from the notation) over $\mathbb{Z}[H]$,

$$C_*^{\mathrm{Nov}} = C_*^{\mathrm{cell}}(\tilde{X}).$$

Case B: Now suppose α is an arbitrary Morse closed 1-form. We use a neat trick due by Latour and Sikorav to approximate α by an exact 1-form (!) and reduce to case A. We can find a Morse function $f : X \to \mathbb{R}$ such that the pair (f, g) is Morse-Smale for our given metric g. Now let $\epsilon > 0$ be small and consider the closed 1-form

$$\beta := \alpha + \epsilon^{-1} df.$$

Since β is cohomologous to α, Lemma 7.8 gives

$$H_*^{\mathrm{Nov}}(\alpha) \simeq H_*^{\mathrm{Nov}}(\beta).$$

(Here we are fixing the covering $\tilde{X} \to X$ throughout the discussion.) Now scaling a 1-form does not change the Novikov complex since the flow lines are the same up to reparametrization. Thus we have an isomorphism of chain complexes inducing an isomorphism on homology

$$H_*^{\mathrm{Nov}}(\beta) = H_*^{\mathrm{Nov}}(df + \epsilon\alpha).$$

[39] I don't know if there is a rigorous proof of this in the literature, but it is widely accepted folklore and I think it is doable. One would like the gradient in an i-simplex to be tangent to the i-simplex, but to do this one will generally have to modify the triangulation a bit first due to smoothness issues.

If ϵ is sufficiently small, then the flow lines for $df + \epsilon\alpha$ are just perturbations of the flow lines for df (exercise), so that we have an isomorphism of chain complexes

$$C_*^{\mathrm{Nov}}(df + \epsilon\alpha) = C_*^{\mathrm{Nov}}(df) \otimes_{\mathbb{Z}[H]} \Lambda. \tag{25}$$

By homological algebra, tensoring a chain complex by a ring changes the homology in a manner which depends only on the homology of the original chain complex. So Case A and the above equation imply that

$$H_*^{\mathrm{Nov}}(df + \epsilon\alpha) \simeq H_* \left(C_*^{\mathrm{cell}}(\tilde{X}) \otimes_{\mathbb{Z}[H]} \Lambda \right).$$

Applying the previous two isomorphisms on homology completes the proof. □

7.5 Reidemeister torsion

When $\chi(X) = 0$, the Novikov homology often vanishes, at least after tensoring with a field. This is true, for example, if X is a 3-manifold obtained by zero-surgery on a knot in S^3, and $\alpha = df$ where $f : X \to S^1$ is in a nontrivial homotopy class. In this case we can still extract some interesting topological information out of the Morse theoretic data, such as the Alexander polynomial of the knot K in the above example.

We begin with an algebraic digression on how to define the "determinant of a chain complex", otherwise known as "Reidemeister torsion". (A good reference on this topic is [68].) Let (C_*, ∂) be a bounded[40] complex over a field F, and let H_* denote its homology. Also let Z_* and B_* denote the spaces of cycles and boundaries respectively. The short exact sequence

$$0 \longrightarrow Z_i \longrightarrow C_i \longrightarrow B_{i-1} \longrightarrow 0$$

induces an isomorphism on top exterior powers,

$$\det(C_i) \xrightarrow{\simeq} \det(Z_i) \otimes \det(B_{i-1}).$$

The short exact sequence

$$0 \longrightarrow B_i \longrightarrow Z_i \longrightarrow H_i \longrightarrow 0$$

[40] "Bounded" means that $\sum_i \dim(C_i) < \infty$.

induces an isomorphism

$$\det(Z_i) \xrightarrow{\simeq} \det(B_i) \otimes \det(H_i).$$

Putting this isomorphism into the previous one and taking the alternating product over i, we obtain an isomorphism

$$\bigotimes_i \det(C_i)^{(-1)^i} \xrightarrow{\simeq} \bigotimes_i \det(H_i)^{(-1)^i}. \tag{26}$$

Now suppose that (C_*, ∂) is acyclic, i.e. $H_* \equiv 0$, and suppose further that we have a chosen (unordered) basis for each C_i. Then the right hand side of (26) is canonically isomorphic to F, and the chosen bases give an element of the left hand side of (26) up to sign, and hence an element of $F/\pm 1$. This element is called the **Reidemeister torsion**

$$T(C_*) \in F/\pm 1.$$

If C_* is not acyclic, we define $T(C_*) := 0$.

For example, the torsion of a 2-term acyclic complex with chosen bases is given by

$$T\left(0 \to C_i \xrightarrow{\partial} C_{i-1} \to 0\right) = \pm \det(\partial)^{(-1)^i}.$$

In general the torsion is an alternating product of determinants of square submatrices of ∂. Namely:

Proposition 7.12 *Let (C_*, ∂) be a bounded acyclic complex over F with chosen bases b_i of C_i. Then we can find a decomposition of the chains $C_* = D_* \oplus E_*$ such that:*

(a) D_i and E_i are spanned by subbases of b_i.

(b) The map $\widehat{\partial}_i := \pi_{E_{i-1}} \circ \partial|_{D_i} : D_i \to E_{i-1}$ is an isomorphism.

For any such decomposition we have

$$T(C_*) = \pm \prod_i \det\left(\widehat{\partial}_i\right)^{(-1)^i}$$

where the determinants are computed with respect to the subbases of b_*.

Now consider a Morse closed 1-form α and a metric g such that the pair (α, g) is Morse-Smale. For simplicity, let us assume that the automorphism group H of our covering $\tilde{X} \to X$ has no torsion. Then the Novikov ring Λ has no zero divisors, so its quotient ring $Q(\Lambda)$ is a field. We define the **Morse-theoretic torsion**

$$T^{\text{Morse}} := T\left(C_*^{\text{Nov}} \otimes_\Lambda Q(\Lambda)\right) \in Q(\Lambda)/\pm H.$$

To explain this, the complex C_*^{Nov} has a preferred set of bases obtained by lifting each critical point in X to $\tilde{X}$. Choosing different lifts will multiply the torsion of the chain complex by some element of H, which is why T^{Morse} is well-defined[41] only in $Q(\Lambda)/\pm H$.

The Morse-theoretic torsion has a topological counterpart which we can try to compare it to. Namely, let $C_*^{\text{cell}}(\tilde{X})$ be the chain complex over $\mathbb{Z}[H]$ obtained by lifting the simplices of a triangulation of X. This has a preferred set of bases consisting of a lift of each simplex from X to $\tilde{X}$, and so we can define the **topological Reidemeister torsion**

$$T^{\text{top}} := T\left(C_*^{\text{cell}}(\tilde{X}) \otimes_{\mathbb{Z}[H]} Q(\mathbb{Z}[H])\right) \in Q(\mathbb{Z}[H])/\pm H.$$

This is known to be a topological invariant depending only on X and the choice of covering. For example, if $X = S^1$ and $\tilde{X} = \mathbb{R}$ then $T^{\text{top}} = (1-t)^{-1}$, as we can easily see by choosing a triangulation of S^1 with one 0-simplex and one 1-simplex. If X is the three-manifold obtained by zero-surgery on a knot $K \subset S^3$, so that $H_1(X) \simeq \mathbb{Z}$, and if $\tilde{X}$ is the infinite cyclic cover of X with $H \simeq \mathbb{Z}$, then it is a result of Milnor that

$$T^{\text{top}} = \frac{\Delta_K(t)}{(1-t)^2},$$

where $\Delta_K(t) \in \mathbb{Z}[t]$ is the Alexander polynomial of K.

The inclusion $\mathbb{Z}[H] \to \Lambda$ induces a map $\imath : Q(\mathbb{Z}[H]) \to Q(\Lambda)$, and we could ask: is

$$\imath\left(T^{\text{top}}\right) = T^{\text{Morse}}?$$

The answer is no; T^{Morse} is not even a topological invariant, as we can see by $X = S^1$ in Examples 7.10(a) and (b). In the first example, $T^{\text{Morse}} = 1$ because there are no critical points, and in the second example $T^{\text{Morse}} = (1-t)^{-1}$.

It is then natural to ask: what is the error $T^{\text{Morse}}/T^{\text{top}}$?

[41] One can get a well-defined element of $Q(\Lambda)/\pm 1$ by choosing an "Euler structure" on X, and one can apparently remove the sign ambiguity by choosing a "homology orientation" of X.

7.6 Periodic orbits and the zeta function

In the Morse theory of circle-valued functions and closed 1-forms, there is a new dynamical feature which does not exist in the real-valued case. Namely, we can consider periodic orbits of the flow V.

A **periodic orbit** of V is a nonconstant map $\gamma : S^1 \to X$ such that $\gamma'(s) = \lambda(V(s))$ for some constant $\lambda > 0$. Here we are not requiring γ to be an embedding. Any periodic orbit factors through an embedding via a p-fold covering map $S^1 \to S^1$; the positive integer p is called the **period** of γ. We declare two periodic orbits to be equivalent if they differ by reparametrization.

For counting purposes, we attach a sign to a generic periodic orbit as follows. For $x \in \gamma(S^1)$, let U be a hypersurface intersecting γ transversely at x, and let $\phi : U \to U$ be the return map (defined near x) which follows the flow p times around $\gamma(S^1)$. The linearized return map induces a map

$$d\phi_x : T_xX/T_x\gamma(S^1) \longrightarrow T_xX/T_x\gamma(S^1)$$

which does not depend on U, and whose eigenvalues do not depend on x. We say that γ is **nondegenerate** if $d\phi_x$ does not have 1 as an eigenvalue, and if so we define the **Lefschetz sign**

$$(-1)^{\epsilon(\gamma)} := \operatorname{sign}\det(1 - d\phi_x) \in \{\pm 1\}.$$

It is not hard to see that if a periodic orbit is nondegenerate then it is isolated.

Definition 7.13 The pair (α, g) is **admissible** if it is Morse-Smale and if all periodic orbits are nondegenerate.

One can show that for a fixed cohomology class $[\alpha]$, a generic pair (α, g) is admissible. If (α, g) is admissible, we count the periodic orbits using the **zeta function**[42]

$$\zeta := \exp \sum_{\gamma \in \mathcal{O}} \frac{(-1)^{\epsilon(\gamma)}}{p(\gamma)} [\gamma] \in \Lambda.$$

Here $\mathcal{O}$ denotes the set of periodic orbits modulo reparametrization, and if γ is a periodic orbit then $[\gamma]$ denotes the image of its homology class under the projection $H_1(X) \to H$. Also exp denotes the formal power series operation $\exp(t) := \sum_{n=0}^{\infty} t^n/n!$.

[42] As we are defining it, the zeta function is not a function, just an element of Λ. When say $\Lambda \simeq \mathbb{Z}((t))$, if one is lucky the power series might converge when one substitutes some complex numbers for t, thus giving an actual function.

Lemma 7.14 *ζ is a well-defined element of the Novikov ring Λ.*

Proof. We first show that

$$\sum_{\gamma\in\mathscr{O}} \frac{(-1)^{\epsilon(\gamma)}}{p(\gamma)}[\gamma] \in \Lambda\otimes\mathbb{Q}.$$

It is enough to show that for any constant C there are only finitely many periodic orbits γ with energy $E(\gamma) < C$. If there are infinitely many, then a compactness argument as in Proposition 7.5 shows that there is a subsequence converging to either (i) a non-isolated periodic orbit, or (ii) a flow line from a critical point to itself. Both cases violate admissibility: in the former case, the limiting periodic orbit is not isolated and hence degenerate, and in the latter case the broken flow line must include a flow line in a moduli space of negative expected dimension, violating the Morse-Smale condition.

It is easy to see that exp sends the Novikov ring to itself so we have $\zeta \in \Lambda\otimes\mathbb{Q}$.

To see that ζ is actually in Λ, we note that there is a product formula

$$\zeta = \prod_{\gamma\in\mathscr{E}} \left(1-(-1)^{i_-(\gamma)}[\gamma]\right)^{(-1)^{-i_0(\gamma)}}. \tag{27}$$

Here $\mathscr{E}$ denotes the set of *embedded* periodic orbits; $i_-(\gamma)$ is the number of real eigenvalues of the linearized return map in the interval $(-\infty,-1)$, and $i_0(\gamma)$ is the number of eigenvalues in $(-1,1)$. One can verify the product formula (27) by taking the formal logarithm of both sides. Clearly the right side of equation (27) has integer coefficients. $\square$

Example 7.15 Let $X = S^1$. In Example 7.10(a),

$$\zeta = \exp\sum_{n=1}^{\infty}\frac{t^n}{n} = (1-t)^{-1}.$$

In Example 7.10(b), there are no periodic orbits so $\zeta = 1$.

Now define

$$I := T^{\mathrm{Morse}}\cdot\zeta \in \frac{\Lambda}{\pm H}.$$

We then have:

Theorem 7.16 *If (α, g) is admissible then:*

(a) I is a topological invariant depending only on X, the cohomology class $[\alpha]$, and the choice of cover.

(b) Moreover $I = \imath(T^{\mathrm{top}})$.

For a proof, see [31]; for an earlier version and a connection with Seiberg-Witten invariants of 3-manifolds see [34]; for more on the connection with Seiberg-Witten theory see [43]. Of course part (a) implies part (b), but one can prove (a) first which leads to an easy proof of (b) similar to our proof of Theorem 7.11 above. There are many other papers on Reidemeister torsion in circle-valued Morse theory; for example, an algebraic refinement of (b) above is given in [50]. Part (a) can be generalized to define a notion of Reidemeister torsion in Floer theory, see [42], where one does not necessarily have an interpretation of the invariant in terms of classical topology.

Exercises for §7.

1. Do the three "exercises in logic" in §7.2 and §7.3.
2. Verify equation (25).
3. Prove Proposition 7.12.
4. Fill in the details in the proof of Lemma 7.14.
5. Let $f : X^n \to S^1$ be a circle-valued function with no critical points. Assume that the fiber is a connected manifold Σ. Choose a generic metric on X and let $\phi : \Sigma \to \Sigma$ be the diffeomorphism defined by following the flow V from Σ back to itself. There is a natural covering $\tilde{X} \simeq \mathbb{R} \times \Sigma$ with $H \simeq \mathbb{Z}$ and $\Lambda \simeq \mathbb{Z}((t))$. Formally, $\tilde{X}$ is the fiber product of X and $\mathbb{R}$ over S^1.
   (a) Check that
   $$\zeta = \exp \sum_{k=1}^{\infty} \# \operatorname{Fix}(\phi^k) \frac{t^k}{k}.$$
   (This is analogous to the zeta function introduced in number theory by Weil [69], which is an ancestor of dynamical zeta functions such as the one considered here.)
   (b) Use the Lefschetz fixed point theorem to deduce that
   $$\zeta = \prod_{i=0}^{n-1} \det(1 - tH_i(\phi))^{(-1)^{i+1}}.$$

8 What we did in the rest of the course, with references

8.1 Pseudoholomorphic curves in symplectic manifolds

(A reference for much of the following is Gromov's seminal paper [28], together with the expository articles in [5] and the second edition of the essential text [44].)

ω-tame and ω-compatible almost complex structures, and contractibility of the space of these. Pseudoholomorphic curves.

Energy and symplectic area; calibration argument for ω-compatible almost complex structures.

Trivial examples of pseudoholomorphic curves: nullhomologous curves and curves in products.

Transversality of somewhere injective curves for a generic almost complex structure. Special cases where transversality is automatic. Dimension of the moduli space.

Introduction to Gromov compactness.

Gromov's nonsqueezing theorem; Gromov-Witten invariants in a special case, monotonicity lemma for minimal surfaces.

Adjunction formula and intersection positivity for pseudoholomorphic curves in symplectic 4-manifolds.

Foliation of $S^2 \times S^2$ by pseudoholomorphic spheres. Gromov's theorem on the recognition of $\mathbb{R}^4$. Gromov's theorem on the symplectomorphism group of $S^2 \times S^2$ and introduction to Abreu's generalization of this [2].

8.2 Floer homology

Introduction to the Arnold conjecture. Introduction to Floer theory of Hamiltonian symplectomorphisms, regarded as homology of the symplectic action functional. Rough description of Floer homology of more general symplectomorphisms (see e.g. some of Seidel's papers) and definition of the flux homomorphism (see [45]).

Index of Cauchy-Riemann operators on punctured Riemann surfaces: Conley-Zehnder index and index formula for Cauchy-Riemann operators on the cylinder via spectral flow (see various papers by Salamon and coauthors such as [55]), relative first Chern class [32], additivity of the index under gluing, axiomatic determination of the index formula for Cauchy-Riemann operators on a punctured Riemann surface (see Schwarz's thesis [57] and the second edition of [44]).

Proof of the Arnold conjecture for monotone symplectic manifolds: definition of Floer homology of Hamiltonian symplectomorphisms (gluing analysis omitted), isomorphism of this Floer homology with Morse homology. (For an excellent introduction to this and much more than we did in the course, see [54]. For transversality details see [20].)

Floer homology with Novikov rings and the Piunikhin-Salamon-Schwarz isomorphism [51]. Introduction to quantum cohomology and its relation to the more general quantum product on Floer theory of symplectomorphisms [14].

Remarks on the classification of surface diffeomorphisms [12]. Floer homology and the mapping class group [63]. Floer homology of finite order symplectomorphisms (not just on surfaces). Computation of the Floer homology of a Dehn twist on a surface [59, 24, 35]. Introduction to Seidel's work on generalized Dehn twists (see Seidel's thesis [61] and more recent papers such as [62]).

Introduction to Floer theory for Lagrangian intersections [17] and the Fukaya category. Floer theory for (noncontractible, nonisotopic) Lagrangians in a surface; combinatorial formula for the differental, proof that the number of generators of the Floer homology equals the geometric intersection number (see [25]). Remarks on Massey products and A_∞ category structure, see e.g. [23].

Introduction to TQFT [3, 58]. Introduction to Seiberg-Witten Floer homology; see [40] and the recent series of papers by Ozsváth and Szabó [49].

Introduction to "introduction to symplectic field theory" [15].

8.3 What we would have also liked to do in the course

Coherent orientations [19].

Gluing analysis.

Khovanov's categorification of the Jones polynomial [36].

The literature on this subject is very large. The following list is nowhere near comprehensive but is merely intended to provide some useful starting points.

References

[1] R. Abraham and J. Robbin, *Transversal mappings and flows*, W. A. Benjamin, 1967.

[2] M. Abreu, *Topology of symplectomorphism groups of $S^2 \times S^2$*, Invent. Math. **131** (1998), 1–23.

[3] M. Atiyah, *The geometry and physics of knots*, Cambridge University Press.

[4] M.F. Atiyah and R. Bott, *The Yang-Mills equations over Riemann surfaces*, Phil. Trans. Roy. Soc. London A **308** (1982), 523–615.

[5] M. Audin and J. Lafontaine (ed.), *Holomorphic curves in symplectic geometry*, Progress in Mathematics 117, Birkhäuser, 1994.

[6] Austin and P. Braam, *Morse-Bott theory and equivariant cohomology*, the Floer memorial volume, 123–183, Progr. Math. 133, Birkhäuser, 1995.

[7] M. Betz and R. Cohen, *Graph moduli spaces and cohomology operations*, Turkish J. Math **18** (1994), 23–41.

[8] R. Bott, *The stable homotopy of the classical groups*, Ann. Math. **70** (1959), 313–337.

[9] R. Bott, *Morse theory indomitable*, I.H.E.S. Publ. Math. **68** (1989), 99–114.

[10] R. Bott and L. Tu, *Differential forms in algebraic topology*, Springer-Verlag, 1982.

[11] F. Bourgeois, *A Morse-Bott approach to contact homology*, preprint.

[12] A. Casson and S. Bleiler, *Automorphisms of surfaces after Nielsen and Thurston*, London Math Soc Student Texts 9, Cambridge University Press, 1988.

[13] R. Cohen, J. Jones, G. Segal, *Floer's infinite-dimensional Morse theory and homotopy theory*, The Floer memorial volume, 297–325, Progr. Math. 133, Birkhäuser.

[14] S. Donaldson, *Floer homology and algebraic geometry*, Vector bundles in algeberaic geometry, 119–138, London Math. Soc Lec. Note Ser. 208, Cambridge University Press, 1995.

[15] Y. Eliashberg, A. Givental, H. Hofer, *Introduction to symplectic field theory*, GAFA, 2000.

[16] A. Floer, *Symplectic fixed points and holomorphic spheres*, Comm. Math. Phys. **120** (1989), 575–611.

[17] A. Floer, *Morse theory for Lagrangian intersections*, J. Diff. Geom. **28** (1988) no. 3, 513–547.

[18] A. Floer, *An instanton-invariant for 3-manifolds*, Comm. Math. Phys. **118** (1998), 215–240.

[19] A. Floer and H. Hofer, *Coherient orientations for periodic orbit problems in symplectic geometry*, Math. Z. **212** (1993), 13–38.

[20] A. Floer, H. Hofer, and D. Salamon, *Transversality in elliptic Morse theory for the symplectic action*, Duke Math. J. **80** (1995), 251–292.

[21] K. Fukaya, *Floer homology of connected sum of homology 3-spheres*, Topology **35** (1996), no. 1, 89–136.

[22] K. Fukaya, *The symplectic s-cobordism conjecture: a summary*, Geometry and Physics (Aarhus, 1995), 209–219, Lecture Notes in Pure and Appl. Math. 184, Dekker, 1997.

[23] K. Fukaya and P. Seidel, *Floer homology, A^∞ categories and topological field theories*, Geometry and Physics (Aarhus, 1995), 9–32, Lecture Notes in Pure and Appl. Math. 184, Dekker, 1997.

[24] R. Gautschi, *Floer homology of algebraically finite mapping classes*, math.SG/0204032.

[25] R. Gautschi, J. Robbin, and D. Salamon, *Heegard splittings and Morse-Smale flows*, preprint, 2001, `www.math.ethz.ch/~salamon`; see also the paper with de Silva cited therein.

[26] R. Gompf and A. Stipsicz, *4-manifolds and Kirby calculus*, Graduate Studies in Math. 20, AMS, 1999.

[27] P. Griffiths and J. Harris, *Principles of algebraic geometry*, Wiley, 1978.

[28] M. Gromov, *Pseudoholomorphic curves in symplectic manifolds*, Invent. Math. **82** (1985), 307–347.

[29] R. Harvey and H. B. Lawson, Jr., *Finite volume flows and Morse theory*, Ann. of Math. **153** (2001), 1–25.

[30] H. Hofer and D. Salamon, *Floer homology and Novikov rings*, the Floer memorial volume, 483–524, Progr. Math. 133, Birkhäuser, 1995.

[31] M. Hutchings, *Reidemeister torsion in generalized Morse theory*, Forum Math. **14** (2002), 209–244.

[32] M. Hutchings, *An index inequality for embedded pseudoholomorphic curves in symplectizations*, J. Eur. Math. Soc **4** (2002), 313–361.

[33] M. Hutchings, *Floer homology of families of equivalent objects*, in preparation.

[34] M. Hutchings and Y-J. Lee, *Circle-valued Morse theory, Reidemeister torsion, and Seiberg-Witten invariants of 3-manifolds*, Topology **38** (1999), 861–888.

[35] M. Hutchings and M. Sullivan, *The periodic Floer homology of a Dehn twist*, preprint.

[36] M. Khovanov, *A categorification of the Jones polynomial*, Duke Math. J. **101** (2000), 359–426.

[37] R. Kirby, *A calculus for framed links in S^3*, Invent. Math. **45** (1978), 35–56.

[38] R. Kirby, *The topology of 4-manifolds*, Lecture Notes in Mathematics 1374, Springer-Verlag, 1989.

[39] F. Kirwan, *Cohomology of quiatients in symplectic and algebraic geometry*, Mathematical Notes 31, Princeton Univ. Press, 1986.

[40] P. Kronheimer, *Embedded surfaces and gauge theory in three and four dimensions*, Surveys in differential geometry, Vol. III, 243–298, Int. Press, 1998.

[41] F. Laudenbach, *On the Thom-Smale complex*, Asterisque **205** (1992), 219–233.

[42] Y-J. Lee, *Reidemeister torsion in symplectic Floer theory and counting pseudo-holomorphic tori*, math.DG/0111313.

[43] T. Mark, *Torsion, TQFT, and Seiberg-Witten invariants of 3-manifolds*, math.DG/9912147.

[44] D. McDuff and D. Salamon, *J-holomorphic curves and quantum cohomology*, AMS, 1994.

[45] D. McDuff and D. Salamon, *Introduction to symplectic topology*, Oxford Univ. Press.

[46] J. Milnor, *Morse theory*, Annals of Mathematics Studies no. 51, Princeton University Press, 1963.

[47] J. Milnor, *Lectures on the h-cobordism theorem*, Princeton University Press, 1965.

[48] S. P. Novikov, *Multivalued functions and functionals: an analogue of the Morse theory*, Soviet Math. Dokl. **24** (1981), no. 2, 222–226.

[49] P. Ozsváth and Z. Szabó, series of recent papers at `http://math.princeton.edu/~szabo`.

[50] A. Pajitnov, *Closed orbits of gradient flows and logarithms of non-abelian Witt vectors*, K-theory **21** (2000), 301–324.

[51] S. Piunikhin, D. Salamon, and M. Schwarz, *Symplectic Floer-Donaldson theory and quantum cohomology*, Contact and symplectic geometry (Cambridge, 1994), 171–200, Publ. Newton Inst. 8, Cambridge University Press, 1996.

[52] M. Pozniak, *Floer homology, Novikov rings, and clean intersections*, in AMS Translations series 2 volume 196, 1999.

[53] J. Robbin and D. Salamon, *The spectral flow and the Masolv index*, Bull. London Math. Soc. **27** (1995), 1–33.

[54] D. Salamon, *Lectures on Floer homology*, Symplectic geometry and topology, IAS/Park City Math Ser. 7, AMS, 1999.

[55] D. Salamon and E. Zehnder, *Morse theory for periodic solutions of Hamiltonian systems and the Maslov index*, Comm. Pure Appl. Math. **45** (1992), 1303–1360.

[56] M. Schwarz, *Morse homology*, Progress in Mathematics 111, Birkh auser, 1993.

[57] M. Schwarz, *Cohomology operations from S^1 cobordisms in Floer theory*, PhD thesis, ETH Zürich 1995.

[58] G. Segal, *Topological field theory*, `www.cgtp.duke.edu/ITP99/segal/`.

[59] P. Seidel, *The symplectic Floer homology of a Dehn twist*, Math. Res. Lett. **3** (1996), 829–834.

[60] P. Seidel, π_1 *of symplectic automorphism groups and invertibles in quantum cohomology rings*, GAFA **7** (1997), 1046–1095.

[61] P. Seidel, *Floer homology and the symplectic isotopy problem*, PhD thesis, Oxford University, 1997.

[62] P. Seidel, *A long exact sequence for symplectic Floer cohomology*, math.SG/0105186.

[63] P. Seidel, *Symplectic Floer homology and the mapping class group*, Pacific J. Math **206** (2002), 219–229.

[64] S. Smale, *Morse inequalities for dynamical systems*, Bull. AMS **66** (1960), 43-49.

[65] C. H. Taubes, *The geometry of the Seiberg-Witten invariants*, Proceedings of the ICM 1998, Vol. II, 493–504.

[66] M. Thaddeus, *A perfect Morse function on the moduli space of flat connections*, Topology **39** (2000), 773–787.

[67] R. Thom, *Sur une partition en cellules associée à une fonction sur une variété*, C. R. Acad. Sci. Paris **228** (1949), 973-975.

[68] V. Turaev, *Introduction to combinatorial torsions*, Birkhauser, 2001.

[69] A. Weil, *Numbers of solutions of equations in finite fields*, Bull. AMS **55** (1949), 497–508.

[70] E. Witten, *Supersymmetry and morse theory*, J. Diff. Geom. **17** (1982), no. 4, 661-692.