

BUILDING HOMOLOGY THEORIES (ALA FLOER)

JOHN B. ETNYRE

ABSTRACT. These notes are an expanded version of evening talks at the 2025 Georgia International Topology Conference, and an abbreviated version of talks at Georgia Tech, which were aimed at graduate students. The hope was to indicate a common framework that has been used since the late 1980s to construct homology theories in low-dimensional topology and symplectic and contact geometry. In addition to this, we also try to indicate how the specific nature of the situation being studied dictates the algebraic nature of the chain groups used to define the homology.

CONTENTS

1. Introduction	1
Acknowledgements	5
2. Morse Homology	5
2.1. Morse functions and gradient flows	6
2.2. Spaces of gradient flow lines	10
2.3. Constructing Morse homology	15
2.4. Invariance of Morse homology	17
2.5. A few details of the main theorems	21
3. Lagrangian Floer homology	30
3.1. A little symplectic geometry	30
3.2. Pseudo-holomorphic curves	33
3.3. Gromov compactness	34
3.4. Defining Lagrangian Floer homology	36
3.5. Applications	43
4. Legendrian contact homology	45
4.1. A little contact geometry	46
4.2. Defining Legendrian contact homology	47
4.3. Applications	52
References	53

1. INTRODUCTION

In the last 40 or so years, there have been lots of “homology theories” in topology and geometry that are built through a similar process. This goes back to work of Floer in [20, 22] that built Symplectic Floer Homology and Lagrangian Floer Homology by looking

at the “moduli space” of solutions to a partial differential equation. Since then, the classical Morse Homology has been reinterpreted in this manner, see [41], and countless other homology theories have been defined. For example, various flavors of Instanton Floer Homology [6, 21, 33] and Monopole Floer Homology [32] are defined in a similar way, and Contact Homology [13], Embedded Contact Homology [29, 31], Legendrian Contact Homology [5, 10, 13], Heegaard Floer Homology [38], and Knot Contact Homology [11] are defined using variants of the Floer approach. (To be clear, the work to define each of these homology theories is considerable. While the approach was laid out by Floer, there are considerable details to be understood and dealt with.)

In these notes, we will discuss a general procedure for constructing the homology theories above. We will illustrate the process in three examples. The simplest case, with which we start, is Morse Homology. We then discuss Lagrangian Floer Homology and Legendrian Contact Homology. The main purpose of this paper is to (1) demonstrate the general process for constructing such homology theories and (2) give a good indication of how one must choose the correct algebraic structure for the chain groups of the homology theory based on the geometry of the problem at hand. There are several great references for the details of Morse Homology [30, 41], and some for the other theories [1, 2], but we hope by not going into a lot of detail on the proofs (though Section 2.5 gives a flavor for these), we can help people see the general approach to building homology theories and how “compactness issues” force the necessary algebraic structures for the homology theory.

Before diving into the examples, we outline a general process for constructing invariants. Any unfamiliar terms in the examples below will be explained later in the text.

Step 0: Start with something you are interested in studying.

Examples:

- For Morse homology, we want to study manifolds.
- For Lagrangian Floer homology, we want to study pairs of Lagrangian submanifolds of symplectic manifolds.
- For Legendrian contact homology, we want to study Legendrian submanifolds of contact manifolds.

Step 1: Choose extra data.

Examples:

- For Morse homology, we choose a pair (f, g) where $f: M \rightarrow \mathbb{R}$ is a Morse function and g is a Riemannian metric on M .
- For Lagrangian Floer homology, we choose an almost-complex structure compatible with the symplectic structure.
- For Legendrian contact homology, we choose a contact form α for the contact structure and a complex structure compatible with $d\alpha$ on the contact hyperplanes.

Step 2: Find an algebraic object, usually a chain complex, that counts something of geometric interest.

Examples:

- For Morse homology, choose a chain complex that is a vector space generated by critical points of f and the differential counts gradient flow lines of f with respect to g .
- For Lagrangian Floer homology, choose a chain complex that is a vector space generated by intersection points between the two Lagrangian submanifolds and the differential counts holomorphic disks “between intersection points”.
- For Legendrian contact homology, choose a chain complex that is an algebra generated by “Reeb cords” and the differential counts holomorphic disks with certain boundary conditions.

Step 3: Show something about the computation in Step 2, usually the homology of the complex, does not depend on the extra data chosen in Step 1.

Before getting into the more complicated examples, we give a couple of very simple examples of the above process.

Example 1.1. We will be interested in the topology of a smooth manifold M . We carry out the above procedure as follows.

Step 1. Choose a “generic” section of the cotangent bundle T^*M .

Step 2. Define a number by the signed count of the zeros of the generic section.

Step 3. Show that this integer is independent of the “generic” section.

We call the number defined above the *Euler characteristic* of M .

Exercise 1.2. Rigorously go through Steps 1. to 3. to show that you have a well-defined integer associated with M . If you know another definition of the Euler characteristic of a smooth manifold, show that it agrees with the definition above.

Example 1.3. Suppose we are interested in the topology of knots in $\mathbb{R}^3$. We carry out the above procedure as follows.

Step 1. Choose a “generic” projection of the knot K to $\mathbb{R}^2$ and remember over and under-crossings in the projection. (That is, choose a knot diagram for K .)

Step 2. Say the diagram D of K is *3-colorable* if the strands of the diagram (that is the arcs the projection is broken into, when small neighborhoods of the crossings are removed from the undercrossing stand) can be colored by three colors so that at each crossing either one or all three colors are used and at least two colors are used.

Step 3. Show that being 3-colorable is independent of the diagram chosen for K .

Exercise 1.4. Rigorously go through Steps 1. to 3. to show that being 3-colorable is an invariant of a knot K in $\mathbb{R}^3$.

Hint: Use Reidemeister moves. We note that the Reidemeister moves simply described a “generic” path of projection of K .

Remark 1.5. We note that there are lots of invariants of knots that can be defined diagrammatically, and so we see the process above to construct invariants is quite robust for knots. (We also note that many of these diagrammatic invariants — including 3-colorability — have other, more intrinsic, definitions that do not come from the above process.)

While this is a “high-level” look at defining invariants, we indicate how Steps 1, 2, and 3 go for “Floer-type homologies”. Above, we indicated what the chain groups would be, so we now need to consider what the boundary maps in the chain complex will be. This will necessarily be fairly abstract, but we will get into the details of the above homology theories in the following sections.

To carry out Step 1 and 2 we let $\mathcal{M}_b^a$ be the space of “objects” counted in the differential (for example, for Morse homology they will be gradient flow lines from a to b , for Lagrangian Floer homology they will be holomorphic disk “from” a “to” b , and for Legendrian contact homology they are a bit more complicated). The first theorem we need is a transversality theorem that takes the following form.

Theorem 1 (Transversality)

For generically chosen data $\mathcal{M}_b^a$ is a manifold of an easily computed dimension.

To see why this is called a transversality theorem, we note that $\mathcal{M}_b^a$ can be defined as the preimage under some function F defined on some function spaces of a submanifold of the target space. So if F is transverse to this submanifold (in some appropriate sense), the preimage will be a manifold. Next, we would like to know that $\mathcal{M}_b^a$ can be oriented.

Theorem 2 (Orientability)

One can make a priori choices so that $\mathcal{M}_b^a$ can be given an orientation for all choices of a and b .

We would also like to know if $\mathcal{M}_b^a$ is compact and if not, how it fails to be compact.

Theorem 3 (Compactness)

If $\{u_i\}$ is a sequence in $\mathcal{M}_b^a$ that has no subsequence that converges to a point in $\mathcal{M}_b^a$, then specify what some subsequence does “converge to”.

We first note that this theorem should really be considered a “lack of compactness” theorem since frequently the spaces are not actually compact, but it is still common to call this a “compactness” theorem. In all relevant situations, we will see that when $\mathcal{M}_b^a$ has dimension 0, then $\mathcal{M}_b^a$ is compact. Thus $\mathcal{M}_b^a$ will be a finite number of points with signs, and we can define the boundary map for the chain complex to be

$$\partial a = \sum |\mathcal{M}_b^a| b,$$

where the sum is over all b such that $\mathcal{M}_b^a$ has dimension 0 and $|\mathcal{M}_b^a|$ is the signed count of points in $\mathcal{M}_b^a$ (we note if you only use the modulo two count of points in $\mathcal{M}_b^a$ then one can skip Theorem 2 above). We will see below that the precise form of the compactness theorem will dictate the exact algebraic structure of the chain complex. The above theorems and comments indicate that ∂a is well-defined.

To show that ∂ is really a boundary map, that is, that $\partial \circ \partial = 0$, we need one last type of theorem.

Theorem 4 (Gluing)

If we have some “element” that might be the limit of a sequence in Theorem 3 above, then it is actually such a limit, and in a nice way.

In this abstract setting, it is not so clear how this will imply that $\partial \circ \partial = 0$, but below, in specific examples, we will see that this is indeed the case.

So the above four theorems are exactly what is needed to prove that we have a chain complex! To carry out Step 3, that is to see that the homology of this complex does not depend on the extra data chosen in Step 1, we need to prove analogs of Theorems 1 through 4 when one considers “generic” 1-parameter families of extra data from Step 1.

While the above discussion is a bit abstract, in the next section, we will see precise versions of the above theorems and prove that they allow us to define chain complexes whose homology is an invariant of manifolds.

We have tried to make these notes largely self-contained and accessible to anyone with a background in differential topology at the level of the standard textbooks [28, 34].

Acknowledgements. We thank Bülent Tosun and an anonymous referee for valuable feedback on early drafts of the paper. The author was partially supported by the National Science Foundation grant DMS-2203312 and the Georgia Institute of Technology Elaine M. Hubbard Distinguished Faculty Award

2. MORSE HOMOLOGY

In this section, we will study a compact manifold without boundary. We do not really need these assumptions on the manifold, but it is easier on a first pass. Once the reader has understood this case, they are encouraged to consider the necessary modifications for manifolds with boundary and non-compact manifolds. But for now

all manifolds in this section are compact and without boundary.

In the first section, we will recall facts about Morse functions and gradient flows, while in the second section, we will discuss the space of gradient flow lines. In the next section, we will define Morse homology, and in the final section, we will briefly indicate what goes into the proofs of the main theorems.

We would also like to point out that the approach to Morse homology given below fits nicely with our discussion of other homology theories, but is not the original way in which Morse homology was defined. There were several previous approaches that were much

more in the spirit of topology and dynamics. We refer the reader to the beautiful exposition of Bott [3] on the history of Morse homology.

2.1. Morse functions and gradient flows. We recall that a function

$$f: M \rightarrow \mathbb{R}$$

is called a *Morse function* if any critical point $p \in M$ of f — that is the points where $df_p = 0$ — is non-degenerate. Here by *non-degenerate* we mean that the *Hessian of f at p* ,

$$Hf_p: T_p M \times T_p M \rightarrow \mathbb{R}: (v, w) \mapsto w \cdot (v \cdot f),$$

is a non-degenerate symmetric 2-form. Here we think of a vector field as a derivation, so $v \cdot f$ is another function that can be thought of as the directional derivative of f at each point.

Remark 2.1. We note that, as defined above, the Hessian is *not* well-defined. This is because for a derivation $v \in T_p M$, we have that $v \cdot f$ is a real number, not a function on M , so we cannot apply w to this. What we mean by the notation above is given $v \in T_p M$, we extend v to a vector field $\tilde{v}$ on M so that $\tilde{v} \cdot f$ is another function on M and then $w \cdot (\tilde{v} \cdot f)$ makes sense. But the value, in general, depends on the extension $\tilde{v}$ of v .

Now at a critical point p of f we will see that Hf_p is well-defined. To this end, let $\tilde{w}$ be an extension of w to a vector field on M and recall that

$$v \cdot f = df_p(v).$$

So we see that

$$(\tilde{v} \cdot (\tilde{w} \cdot f))(p) - (\tilde{w} \cdot (\tilde{v} \cdot f))(p) = ([\tilde{v}, \tilde{w}])_p \cdot f = df_p([\tilde{v}, \tilde{w}]) = 0$$

since $df_p = 0$ at a critical point. Here $[\tilde{v}, \tilde{w}]_p$ is the Lie bracket of the vector fields evaluated at p . Thus, we see that

$$(\tilde{v} \cdot (\tilde{w} \cdot f))(p) = (\tilde{w} \cdot (\tilde{v} \cdot f))(p)$$

and so

$$(\tilde{v} \cdot (\tilde{w} \cdot f))(p) = d(\tilde{w} \cdot f)_p(\tilde{v}) = d(\tilde{w} \cdot f)_p(v)$$

since $d(\tilde{w} \cdot f)_p(\tilde{v})$ only depends on the value of $\tilde{v}$ at p . We can similarly see $(\tilde{v} \cdot (\tilde{w} \cdot f))(p)$ only depends on the value of $\tilde{w}$ at p and thus $Hf_p(v, w)$ is well-defined when p is a critical point. We also notice that Hf_p is symmetric at a critical point.

Exercise 2.2. In local coordinates $(x_1, \dots, x_n)$ near p , show that the symmetric, bilinear form Hf_p can be given by the matrix

$$\left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i,j=1}^n$$

in the basis $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$ for the tangent space given by the coordinate system.

Exercise 2.3. Show that the matrix expression for Hf_p is independent of local coordinates at a critical point. (That is, show that in another local coordinate system $(y_1, \dots, y_n)$ you will get a different matrix, but it will be related by a change of basis for the tangent space given by the change of coordinates.)

Here is another way to define the Hessian.

Exercise 2.4. Suppose that g is a Riemannian metric on M and let ∇ be the Levi-Civita connection on T^*M associated to g (we note that the Levi-Civita connection is usually defined on TM , but as TM and T^*M can be identified using g we will consider the connection on T^*M). We note that one may actually use any connection on T^*M , but this is a natural one. We recall that ∇ allows one to differentiate sections of T^*M . So, given a vector field w , we obtain the 1-form $\nabla_w df$. Show that at a critical point p of f , $\nabla_w df$ is independent of the connection and $Hf_p(v, w) = (\nabla_w df)(v)$.

Exercise 2.5. Note that $df: M \rightarrow T^*M$ is a section of the cotangent bundle. Show that f is a Morse function if and only if df is transverse to the zero section Z of T^*M .

The first thing one would like to know is that any manifold has a Morse function.

Theorem 2.6. A generic smooth function $f: M \rightarrow \mathbb{R}$ is a Morse function.

In this theorem, by generic, we mean that the space of Morse functions $\mathcal{M}(M)$ on M as a subset of the set $C^\infty(M)$ of all C^∞ functions on M is dense. (Or more precisely, $\mathcal{M}(M)$ is residual, meaning it is the intersection of countably many open dense sets in $C^\infty(M)$.)

There are many proofs of the above theorem. Most involve embedding M into some Euclidean space, but see [26, Proposition 6.13] for an alternate and beautiful proof using jet transversality.

Given a Morse function $f: M \rightarrow \mathbb{R}$ on an n -manifold M and a critical point $p \in M$, the Hessian Hf_p can be represented by a symmetric $n \times n$ matrix. Thus, the matrix can be diagonalized. This implies that

$$T_p M = V_0 \oplus V_+ \oplus V_-$$

where V_+ , respectively V_- , is the subspace of $T_p M$ on which Hf_p is positive, respectively negative, definite and V_0 is the subspace where Hf_p is degenerate. Since f is Morse and p is a non-degenerate critical point, we see that $V_0 = \{0\}$, and so

$$n = \dim T_p M = \dim V_+ + \dim V_-$$

We call the dimension of V_- the *index* of p .

Theorem 2.7 (Morse Lemma). Let p be a non-degenerate critical point of

$$f: M \rightarrow \mathbb{R}.$$

If k is the index of p then there is a coordinate chart $\phi: U \rightarrow V$ around p such that

$$f \circ \phi^{-1}(x_1, \dots, x_n) = f(p) - x_1^2 - \dots - x_k^2 + x_{k+1}^2 + \dots + x_n^2.$$

Remark 2.8. Recall from differential topology that if the first derivative of a function $f: M \rightarrow \mathbb{R}$ is non-degenerate at a point p , that is $df_p \neq 0$, then there are local coordinates where f takes the form

$$\mathbb{R}^n \rightarrow \mathbb{R}: (x_1, \dots, x_n) \mapsto x_1.$$

Said another way, if the first derivative of f at p is non-degenerate, then f is locally equivalent to the “simplest” function with non-degenerate first derivative.

One may think of the Morse lemma as a second derivative version of this. So if $df_p = 0$ but Hf_p is non-degenerate (that is the total second derivative is non-degenerate), then f is locally equivalent to the “simplest” function with vanishing first derivative and non-degenerate second derivative.

There are several proofs of the Morse Lemma; see, for example, [1, 26, 36].

Given a Riemannian metric g on a manifold M and a function $f: M \rightarrow \mathbb{R}$, the *gradient* of f is the unique vector field ∇f such that

$$g(\nabla f, v) = df(v)$$

for all vector fields v .

Exercise 2.9. Prove that ∇f exists and is unique.

Hint: Recall that g is non-degenerate and so the map $T_p M \rightarrow T_p^* M: v \mapsto \iota_v g$ is an isomorphism for all p , where $\iota_v g$ is the contraction of v into g . That is, $\iota_v g$ is the 1-form that takes a vector w and gives the real number $g(v, w)$.

The *gradient flow* of f is the flow of $-\nabla f$. Specifically, given a point $x \in M$ the *gradient flow line* of f through x is

$$\gamma: \mathbb{R} \rightarrow M$$

such that

$$\begin{aligned} \gamma'(t) &= -\nabla f_{\gamma(t)} \\ \gamma(0) &= x. \end{aligned}$$

Since we are considering compact manifolds without boundary, the flow line is defined for all time.

Recall that the gradient of a function points in the direction in which f is increasing the fastest. Thus, the gradient flow is going in the direction of decreasing values of f . So if our manifold is the graph of some function h and we think of f as a height function on a manifold, then one can think of the gradient flow as the flow of water poured on the manifold.

Example 2.10. Consider the manifold $\mathbb{R}^n$ and the function

$$f(x_1, \dots, x_n) = -x_1^2 - \dots - x_k^2 + x_{k+1}^2 + \dots + x_n^2.$$

We also choose the Euclidean metric

$$g = \sum_{i=1}^n dx_i \otimes dx_i.$$

One easily computes that

$$df = -\sum_{i=1}^k 2x_i dx_i + \sum_{i=k+1}^n 2x_i dx_i$$

and thus

$$\nabla f = 2 \left(-\sum_{i=1}^k x_i \frac{\partial}{\partial x_i} + \sum_{i=k+1}^n x_i \frac{\partial}{\partial x_i} \right).$$

The gradient flow lines are indicated in Figure 1.

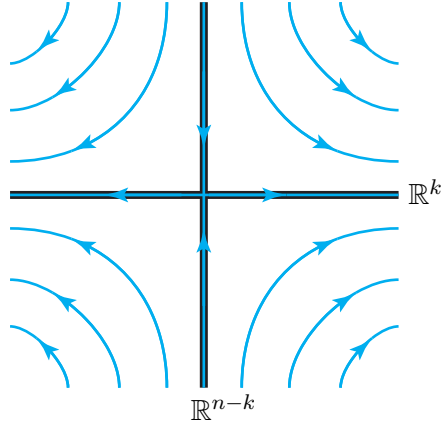


FIGURE 1. The gradient flow of $f(x_1, \dots, x_n) = -x_1^2 - \dots - x_k^2 + x_{k+1}^2 + \dots + x_n^2$.

Example 2.11. Consider the unit sphere S^2 in $\mathbb{R}^3$ with the metric induced from the Euclidean metric on $\mathbb{R}^3$. If the function $f: S^2 \rightarrow \mathbb{R}$ is given by projection to the last coordinate, then we see that there are exactly two critical points $N = (0, 0, 1)$ and $S = (0, 0, -1)$. It is clear that N has index 2 and S has index 0. Moreover, the gradient flow is shown in Figure 2.

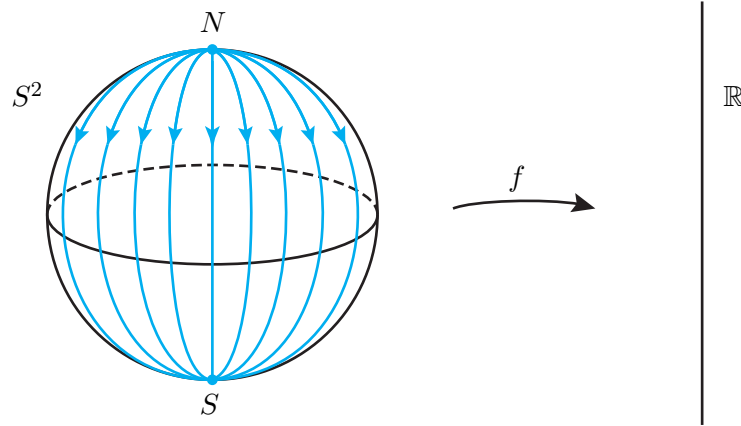


FIGURE 2. The gradient flow of the height function on the sphere.

Exercise 2.12. Show that a flow line of the gradient in the previous example through the point $x = (\cos \theta, \sin \theta, 0)$ is given by

$$\gamma(t) = \left(\frac{\cos \theta}{\cosh t}, \frac{\sin \theta}{\cosh t}, \tanh t \right).$$

Hint: Show that in stereographic coordinates given by

$$\phi(u, v) = \frac{1}{1 + u^2 + v^2} (2u, 2v, u^2 + v^2 - 1)$$

the metric is given by

$$\frac{4}{(1 + u^2 + v^2)^2} (du \otimes du + dv \otimes dv)$$

or in polar coordinates

$$\frac{4}{(1 + r^2)^2} (dr \otimes dr + r^2 d\theta \otimes d\theta).$$

Show that in these coordinates, f is given by $r \mapsto \frac{r^2-1}{r^2+1}$ and

$$\nabla f = r \frac{\partial}{\partial r}.$$

Thus the flow line through $(1, \theta)$ is $t \mapsto (e^{-t}, \theta)$.

Here are a few simple facts about gradient flow lines that follow from general facts about flows.

Exercise 2.13. Show that if $\gamma: \mathbb{R} \rightarrow M$ is a gradient flow line, then γ_a define by $\gamma_a(t) = \gamma(t + a)$ is also a gradient flow line (through the point $\gamma_a(0) = \gamma(a)$).

Exercise 2.14. If γ and η are gradient flow lines with the same image (this happens if and only if they have one point in common), then there is some a such that $\gamma(t) = \eta(t + a)$.

2.2. Spaces of gradient flow lines. We will now define the Morse homology of a manifold. We begin by discussing the “moduli space of gradient flow lines”. Given a Morse function $f: M \rightarrow \mathbb{R}$ and two critical points p and q of f we define the space

$$\begin{aligned} \mathcal{M}_q^p = \{ \gamma: \mathbb{R} \rightarrow M \text{ such that} \\ 1) \gamma \text{ is a flow line of } -\nabla f, \\ 2) \lim_{t \rightarrow -\infty} \gamma(t) = p, \text{ and} \\ 3) \lim_{t \rightarrow \infty} \gamma(t) = q \} / \mathbb{R} \end{aligned}$$

where $\mathbb{R}$ acts on the gradient flow lines as indicated at the end of the last section. We can think of

$$\mathcal{M}_q^p = \text{the space of unparameterized gradient flow lines from } p \text{ to } q.$$

Example 2.15. We will consider the function and metric in Example 2.11 above. It is clear that there are only two critical points, N and S . The only non-empty moduli space is

$$\mathcal{M}_S^N \cong S^1.$$

To see this, notice that any gradient flow line intersects the equator of the sphere exactly once, and for every point on the equator, there is exactly one (unparameterized) flow line through that point.

Example 2.16. We now consider a Morse function on the torus. One standardly draws the torus as, for example, the result of taking the circle of radius 1 about $(0, 2)$ in the xz -plane and rotating it about the x -axis, see Figure 5. If we restrict the projection map $\mathbb{R}^3 \rightarrow \mathbb{R}: (x, y, z) \mapsto z$ to this torus we will not get a “generic” Morse function as discussed below, so in this example we will slightly skew this torus and let f the restriction of the projection

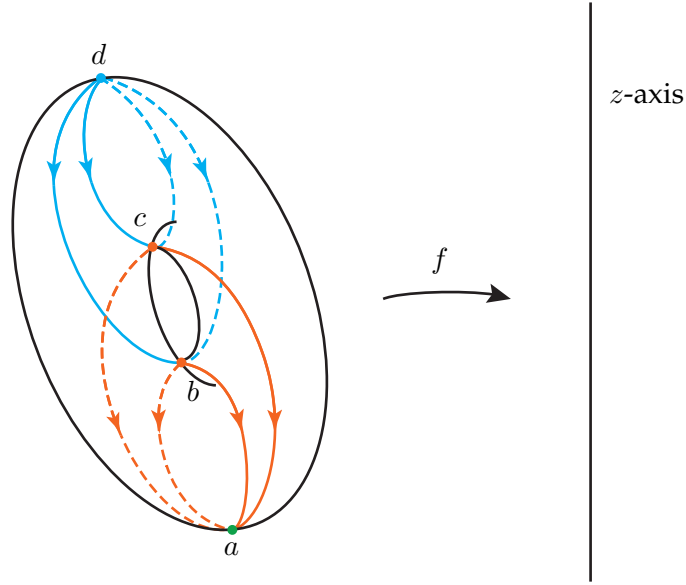


FIGURE 3. The gradient flow of projection the the z -axis for the skew torus in $\mathbb{R}^3$.

map to this torus. See Figure 3. We note there are 4 critical points. A maximum, and hence index 2 critical point, at d , a minimum, and hence index 0 critical point, at a , and two index 1 critical points at b and c . A few gradient flow lines for f are shown in the figure. As it is difficult to draw more flow lines in this figure, we represent the torus as a square with opposite edges identified in Figure 4. There, we can draw more of the flow lines. We note that $\mathcal{M}_b^d, \mathcal{M}_c^d, \mathcal{M}_a^b$, and $\mathcal{M}_a^c$ all consist of 2 points, the first two spaces correspond to the blue lines and the second two correspond to the horizontal and, respectively, vertical orange lines. The spaces $\mathcal{M}_c^b$ and $\mathcal{M}_b^c$ are empty, and $\mathcal{M}_a^d$ is the union of 4 open intervals, which are indicated by the purple curves in the figure.

Morse Theorem 1 (Transversality)

For a generic choice of metric g on M and Morse function $f: M \rightarrow \mathbb{R}$ the space

$$\mathcal{M}_q^p$$

is a manifold of dimension $\text{index } p - \text{index } q - 1$.

The next example shows that the above theorem is not true for any metric and Morse function.

Example 2.17. In Figure 5 we see the symmetric torus in $\mathbb{R}^3$. If f is the restriction of the z -coordinate to the torus then we see the critical points and some of the gradient flow lines in the figure. Note that $\mathcal{M}_b^c$ consists of two points, but if we had chosen a generic Morse function and metric (here the metric is induced from $\mathbb{R}^3$) in the sense of Morse Theorem 1, then $\mathcal{M}_b^c$ would have dimension -1 and hence be empty. Thus it is clear that Morse Theorem 1 is not true for all Morse functions and metrics.

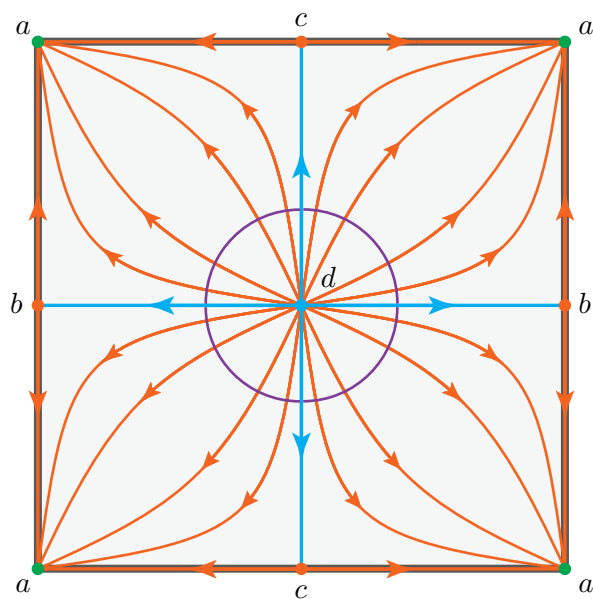


FIGURE 4. The torus represented as the square with opposite edges identified by translation. We see the gradient flow of the function f . The points in the purple curves are in one-to-one correspondence with flow lines from d to a .

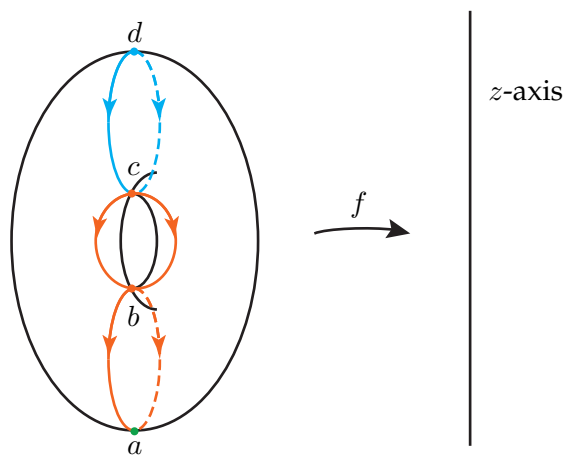


FIGURE 5. The symmetric torus in $\mathbb{R}^3$.

We now move to the second main theorem in the case of Morse homology.

Morse Theorem 2 (Orientability)

If M is oriented and we choose orientations on all the negative eigenspaces of $(T_p M, Hf_p)$ for all critical points p , then the $\mathcal{M}_q^p$ are “consistently” oriented for all p and q .

We will clarify what we mean by “consistently” below when discussing “gluing”, but now turn to the “compactness” theorem (or, as mentioned above, it can really be thought of as the “lack of compactness” theorem).

Morse Theorem 3 (Compactness)

If $\{\gamma_i\}$ is a Cauchy sequence in $\mathcal{M}_q^p$ that does not converge to a point in $\mathcal{M}_q^p$, then $\{\gamma_i\}$ converges to a “broken flow line”.

We say that $\{\gamma_i\}$ converges to a broken flow line if there are critical points

$$p = r_1, r_2, \dots, r_l = q$$

such that $\text{index } r_j > \text{index } r_{j+1}$ and there are elements $\tilde{\gamma}^j \in \mathcal{M}_{r_{j+1}}^{r_j}$ such that for all $\epsilon > 0$ there is some N such that for all $i > N$ we have

$$\text{image}(\gamma_i) \subset \epsilon\text{-neighborhood of } \left(\bigcup_{j=1}^{l-1} \text{image}(\tilde{\gamma}_j) \right)$$

and

$$\left(\bigcup_{j=1}^{l-1} \text{image}(\tilde{\gamma}_j) \right) \subset \epsilon\text{-neighborhood of } \text{image}(\gamma_i).$$

Example 2.18. In Figure 6 we see a sequence of gradient flow lines γ_i shown as orange curves converging to a broken flow line given as the union of $\tilde{\gamma}_1$, and $\tilde{\gamma}_2$. The light green region is a neighborhood of the union of the images of $\tilde{\gamma}_1$, and $\tilde{\gamma}_2$, and we see that for large enough i , the images of the γ_i are contained in the neighborhood. A similar figure will show that for large i the images of $\tilde{\gamma}_1$, and $\tilde{\gamma}_2$ will be in a small neighborhood of the images of the γ_i .

Corollary 2.19. *If $\text{index } p = \text{index } q + 1$, then $\mathcal{M}_q^p$ is a compact 0-manifold.*

Proof. There are no critical points with index strictly between $\text{index } p$ and $\text{index } q$, so any Cauchy sequence in $\mathcal{M}_q^p$ must converge to a point in $\mathcal{M}_q^p$. $\square$

Corollary 2.20. *If $\text{index } p = \text{index } q + 2$, then*

$$\overline{\mathcal{M}_q^p} \subset \left(\mathcal{M}_q^p \cup \left(\bigcup_{\text{index } r = \text{index } p - 1} (\mathcal{M}_r^p \times \mathcal{M}_q^r) \right) \right),$$

where $\overline{\mathcal{M}_q^p}$ is the “completion” formed by adding all broken trajectories.

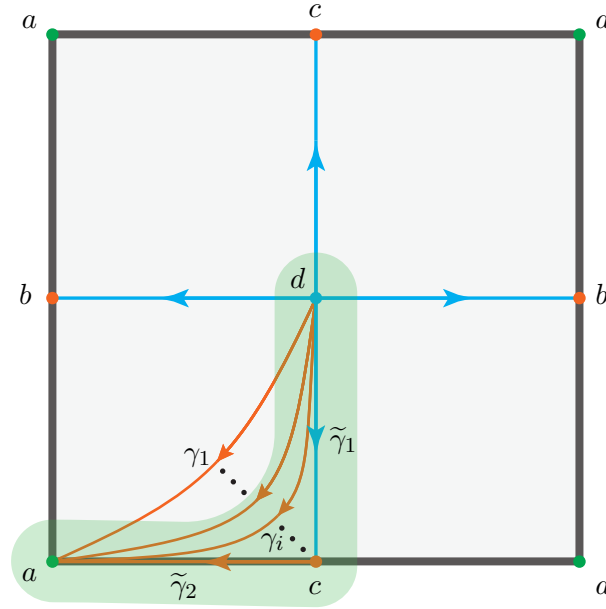


FIGURE 6. A sequence of flow lines converging to a broken flow line.

Exercise 2.21. *Prove this corollary.*

We finally turn to the final main theorem for Morse homology.

Morse Theorem 4 (Gluing)

Let p , r , and q be critical points with

$$\text{index } p = \text{index } r + 1 = \text{index } q + 2.$$

There is an orientation-preserving diffeomorphism

$$\Phi: ((0, \epsilon) \times \mathcal{M}_r^p \times \mathcal{M}_q^r) \rightarrow \mathcal{M}_q^p$$

such that for all $\gamma \in \mathcal{M}_r^p$, $\eta \in \mathcal{M}_q^r$, and sequence $\{s_i\}$ in $(0, \epsilon)$ converging to 0 we have that $\Phi(\gamma, \eta, s_i)$ converges to the broken flow line $\gamma \cup \eta$.

Remark 2.22. Here is where we need the orientations from Morse Theorem 2 to be “consistent”. This means that we can relate the orientations on the moduli spaces as indicated in this theorem.

Example 2.23. We consider the compactification of the moduli spaces for the skew torus in Example 2.16 above. We note that $\mathcal{M}_a^d$ contains an open interval corresponding to the pink curves in Figures 7. The compactification of this interval consists of the broken flow lines $\gamma_1 \cup \delta_4$ and $\gamma_3 \cup \delta_2$. More generally we see that

$$\begin{aligned} \mathcal{M}_b^d &= \{\gamma_1, \gamma_2\}, & \mathcal{M}_c^d &= \{\gamma_3, \gamma_4\}, \\ \mathcal{M}_a^b &= \{\delta_3, \delta_4\}, & \text{and } \mathcal{M}_a^c &= \{\delta_1, \delta_2\}, \end{aligned}$$

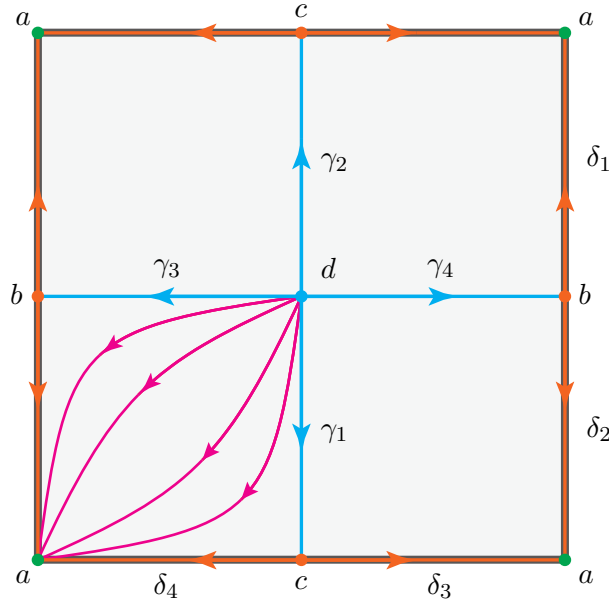


FIGURE 7. The compactification of the gradient flow lines in the lower left corner are the broken flow lines $\gamma_1 \cup \delta_4$ and $\gamma_3 \cup \delta_2$.

and

$$\partial \overline{\mathcal{M}}_a^d = (\mathcal{M}_b^d \times \mathcal{M}_a^b) \cup (\mathcal{M}_c^d \times \mathcal{M}_a^c).$$

So we saw that $\mathcal{M}_a^d$ consists of 4 open intervals and thus its closure consists of 4 closed intervals with boundary being 8 broken flow lines.

This example indicates that the proof of the following result.

Corollary 2.24. *If $k = \text{index } p = \text{index } q + 2$, then $\overline{\mathcal{M}}_q^p$ is a compact 1-manifold with boundary*

$$\bigcup_{r \text{ of index } k-1} (\mathcal{M}_r^p \times \mathcal{M}_q^r).$$

Exercise 2.25. *Prove this corollary.*

2.3. Constructing Morse homology. Let $f: M \rightarrow \mathbb{R}$ be a Morse function and g a metric on M so that all the theorems in the past section hold. We define the chain groups for Morse homology to be

$$C_i = \text{free abelian group generated by critical points of index } i.$$

We then define

$$\partial_i: C_i \rightarrow C_{i-1}$$

on the generators $p \in C_i$ by

$$\partial_i p = \sum_{q \text{ of index } i-1} |\mathcal{M}_q^p| q$$

where $|\mathcal{M}_q^p|$ is the signed count of points in $\mathcal{M}_q^p$.

Notice that Morse Theorem 1 and 3 above show that $\mathcal{M}_q^p$ is a compact 0-manifold, and Morse Theorem 2 shows that $\mathcal{M}_q^p$ is oriented. So $\partial_i q$ is well-defined. We note that if we considered C_i to be a vector space over $\mathbb{Z}_2$ then we could just take the mod 2 count of points in $\mathcal{M}_q^p$, and then we would not need Morse Theorem 2 (or the part of Morse Theorem 4 concerning orientations).

Theorem 2.26. *With the notation above we have*

$$\partial_{i-1} \circ \partial_i = 0.$$

Proof. We note that

$$\begin{aligned} \partial(\partial p) &= \partial \left(\sum_{r \text{ of index } i-1} |\mathcal{M}_r^p| r \right) \\ &= \sum_{q \text{ of index } i-2} |\mathcal{M}_r^p| \left(\sum_{r \text{ of index } i-1} |\mathcal{M}_q^r| q \right) \\ &= \sum_{q \text{ of index } i-2} \left(\sum_{r \text{ of index } i-1} |\mathcal{M}_r^p| |\mathcal{M}_q^r| \right) q. \end{aligned}$$

So the coefficient on q is $\sum_{r \text{ of index } i-1} |\mathcal{M}_r^p| |\mathcal{M}_q^r|$, but note this is exactly

$$\left| \bigcup_{r \in R} (\mathcal{M}_r^p \times \mathcal{M}_q^r) \right| = |\partial \overline{\mathcal{M}_r^p}|.$$

Since $\overline{\mathcal{M}_r^p}$ is a compact 1-manifold, we know the signed count of its boundary components is 0. Thus $\partial \circ \partial = 0$. $\square$

We now give a “picture proof” of this result. This is the version most people would give!

Picture Proof. We begin by noting on the left-hand side of Figure 8 we see how q would show up in $\partial^2 p$. That is, it will show up as a broken flow line. But Morse Theorem 4 says that this will be a boundary component of $\overline{\mathcal{M}_q^p}$ as indicated in the middle of the figure. This component of $\overline{\mathcal{M}_q^p}$ is an oriented 1-dimensional manifold. So this component of $\overline{\mathcal{M}_q^p}$ will look like the figure on the right of Figure 8. So you see if q is in $\partial^2 p$ then it will show up again with the opposite sign. Some people will just draw the diagram on the right of Figure 8 and call that the proof that $\partial^2 = 0$, and once you have thought through it, it is a great proof! $\square$

We can now define the i^{th} Morse homology of M to be

$$H_i(M) = \frac{\ker(\partial: C_i \rightarrow C_{i-1})}{\text{image}(\partial: C_{i+1} \rightarrow C_i)}.$$

Exercise 2.27. *Show that $H_i(M)$ agrees with the i^{th} homology of M as is usually defined in algebraic topology.*

Hint: This might be more challenging than some previous exercises. Maybe read Section 2.5.5 first

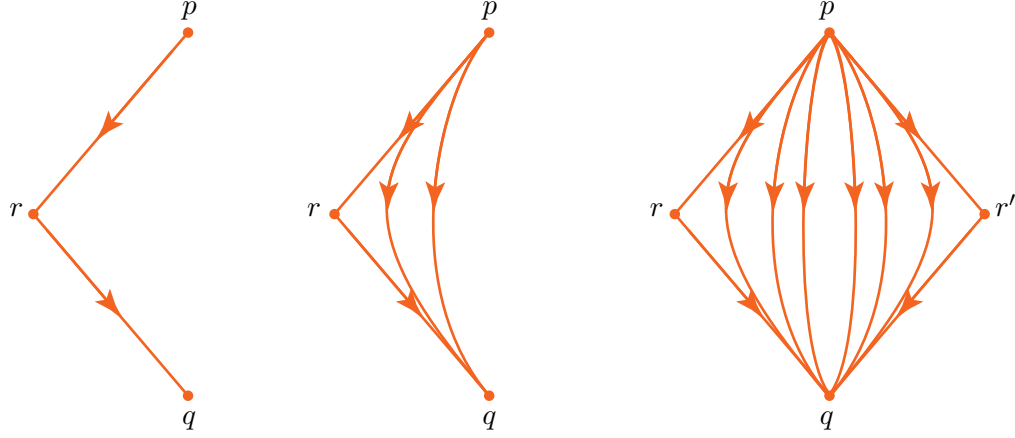


FIGURE 8. The compactification of the gradient flow lines in the lower left corner are the broken flow lines $\gamma_1 \cup \delta_4$ and $\gamma_3 \cup \delta_2$.

and think about cellular homology and how you might get a CW structure on M from a Morse function.

Of course, if you finish this exercise, we know that the Morse homology does not depend on the choice of Morse function or the Riemannian metric, but we now turn to a more direct proof of this fact.

2.4. Invariance of Morse homology. Suppose that (f_0, g_0) and (f_1, g_1) are two pairs of Morse functions and Riemannian metrics so that all the theorems in the previous section hold. Each (f_i, g_i) gives us a chain complex (C^i, ∂^i) . We want to build a chain map

$$\Phi: (C^0, \partial^0) \rightarrow (C^1, \partial^1)$$

that will induce an isomorphism on homology. To this end, let

$$\begin{aligned} \Gamma = \{ (f_t, g_t) : t \in \mathbb{R} \text{ such that} \\ f_t = f_0 \text{ and } g_t = g_0 \text{ for } t \leq 0 \\ f_t = f_1 \text{ and } g_t = g_1 \text{ for } t \geq 1 \}. \end{aligned}$$

We note that the space of functions and the space of metrics are contractible, so such a Γ exists. Now let V_t be the gradient $\nabla_{g_t} f_t$ of f_t with respect to the metric g_t and set V to be the time-dependent vector field such that at time t the vector field V is $-V_t$:

$$V = \{-V_t\}.$$

A flow line of V is a map $\gamma: \mathbb{R} \rightarrow M$ such that $\gamma'(t) = (V_t)_{\gamma(t)}$. Now given a critical point p of f_0 and a critical point q of f_1 we set

$$\begin{aligned} \widetilde{\mathcal{M}}_q^p = \{ \gamma: \mathbb{R} \rightarrow M \text{ such that} \\ 1) \gamma \text{ is a flow line of } V, \\ 2) \lim_{t \rightarrow -\infty} \gamma(t) = p, \text{ and} \\ 3) \lim_{t \rightarrow \infty} \gamma(t) = q \}. \end{aligned}$$

We note that since V is a time-dependent vector field, there is no $\mathbb{R}$ action on $\widetilde{\mathcal{M}}_q^p$ as there was above. We now have analogs of Morse Theorems 1 through 4.

Morse Theorem 1' (Transversality)

For a generic choice of Γ the space

$$\widetilde{\mathcal{M}}_q^p$$

is a manifold of dimension index p – index q .

We note that there is no -1 as in Morse Theorem 1 since there is no $\mathbb{R}$ action on flow lines of the time-dependent vector field V .

Morse Theorem 2' (Orientability)

If M is oriented and we choose orientations on all the negative eigenspaces of the critical points of f_0 and f_1 , then the $\widetilde{\mathcal{M}}_q^p$ are “consistently” oriented for all p and q .

The meaning of “consistently” is similar to its meaning above. In particular, it means the orientations in the gluing theorem below are consistent. We now move to the compactness theorem.

Morse Theorem 3' (Compactness)

If $\{\gamma_i\}$ is a Cauchy sequence in $\widetilde{\mathcal{M}}_q^p$ that does not converge to a point in $\widetilde{\mathcal{M}}_q^p$, then $\{\gamma_i\}$ converges to “broken flow line”.

The notion of convergence is similar to the one above, but slightly different. We say that $\{\gamma_i\}$ converges to a broken flow line if there are critical points $p = p_1, p_2, \dots, p_k$ of f_0 and critical points $p_{k+1}, \dots, p_l = q$ of f_1 such that the indices are strictly decreasing except for possibly p_k and p_{k+1} which can have the same index, and elements $\tilde{\gamma}_j \in \mathcal{M}(f_0)_{p_{j+1}}^{p_j}$ for $j < k$, $\tilde{\gamma}_j \in \mathcal{M}(f_1)_{p_{j+1}}^{p_j}$ for $j > k$, and $\tilde{\gamma}_k \in \widetilde{\mathcal{M}}_{p_{k+1}}^{p_k}$ such that the γ_i converge to $\{\tilde{\gamma}_j\}_{j=1}^l$ as in Morse Theorem 3.

We now turn to the “gluing” theorem, but state it a bit vaguely and leave it to the reader to see how to adopt the statement of Morse Theorem 4 to the current situation.

Morse Theorem 4' (Gluing)

Can “undo” breaking with index difference 1.

We now have the immediate corollaries whose proofs are very similar to the analogous ones in the previous section.

Corollary 2.28. *If p is a critical point of f_0 , q is a critical point of f_1 such that index $p = \text{index } q$, then $\widetilde{\mathcal{M}}_q^p$ is a compact 0-manifold.*

Corollary 2.29. *If p is a critical point of f_0 , q is a critical point of f_1 such that $k = \text{index } p = \text{index } q + 1$, then $\widetilde{\mathcal{M}}_q^p$ is a compact 1-manifold with boundary*

$$\left(\bigcup_{r \in \text{Crit}_{k-1}(f_0)} (\mathcal{M}_r^p \times \widetilde{\mathcal{M}}_q^r) \right) \cup \left(\bigcup_{r \in \text{Crit}_k(f_1)} (\widetilde{\mathcal{M}}_r^p \times \mathcal{M}_q^r) \right)$$

where $\text{Crit}_k(f)$ is the set of critical points of f with index k .

We can now define $\Phi: C_n^0 \rightarrow C_n^1$ on a generator $p \in C_n^0$ by

$$\Phi_\Gamma(p) = \sum_{q \in \text{Crit}_n(f_1)} |\widetilde{\mathcal{M}}_q^p| q.$$

Where, again, $|\widetilde{\mathcal{M}}_q^p|$ means the signed count of points in $\widetilde{\mathcal{M}}_q^p$. Clearly, Morse Theorems 1' through 3' show Φ_Γ is a well-defined homomorphism. We now see that this is a chain map.

Theorem 2.30. *With the notation above*

$$\Phi_\Gamma \circ \partial_0 + \partial_1 \circ \Phi_\Gamma = 0.$$

From basic homological algebra, we have the immediate corollary.

Corollary 2.31. *The map Φ_Γ induces a homomorphism*

$$H(C^0, \partial^0) \rightarrow H(C^1, \partial^1),$$

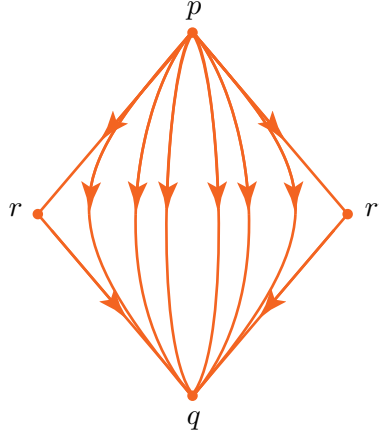
where $H(C^i, \partial^i)$ is the homology of (C^i, ∂^i) .

Proof of Theorem 2.30. We will denote by $\text{Crit}_k(f_i)$ the critical points of f_i with index k . Now take a critical point p of f_0 with index n . Then we have

$$\begin{aligned} (\Phi_\Gamma \circ \partial_0)(p) + (\partial_1 \circ \Phi_\Gamma)(p) &= \Phi_\Gamma \left(\sum_{r \in \text{Crit}_{n-1}(f_0)} |\mathcal{M}_r^p| r \right) + \partial_1 \left(\sum_{r \in \text{Crit}_n(f_1)} |\widetilde{\mathcal{M}}_r^p| r \right) \\ &= \sum_{r \in \text{Crit}_{n-1}(f_0)} \left(|\mathcal{M}_r^p| \sum_{q \in \text{Crit}_{n-1}(f_1)} |\widetilde{\mathcal{M}}_q^r| q \right) + \sum_{r \in \text{Crit}_n(f_1)} \left(|\widetilde{\mathcal{M}}_r^p| \sum_{q \in \text{Crit}_{n-1}(f_1)} |\mathcal{M}_q^r| q \right) \\ &= \sum_{q \in \text{Crit}_{n-1}(f_0)} \left(\sum_{r \in \text{Crit}_{n-1}(f_0)} |\mathcal{M}_r^p| |\widetilde{\mathcal{M}}_q^r| + \sum_{r \in \text{Crit}_n(f_1)} |\widetilde{\mathcal{M}}_r^p| |\mathcal{M}_q^r| \right) q \\ &= \sum_{q \in \text{Crit}_{n-1}(f_0)} |\partial \widetilde{\mathcal{M}}_q^p| q = 0, \end{aligned}$$

since $\widetilde{\mathcal{M}}_q^p$ is a compact 1-manifold with boundary. $\square$

In Figure 9 we see the “picture proof” of this theorem. Here $p \in \text{Crit}_n(f_0)$, $q \in \text{Crit}_{n-1}(f_1)$ and r and r' are either in $\text{Crit}_n(f_1)$ or $\text{Crit}_{n-1}(f_0)$. So we see the terms showing up in $(\Phi_\Gamma \circ \partial_0 + \partial_1 \circ \Phi_\Gamma)(p)$ are simply the boundary of $\widetilde{\mathcal{M}}_q^p$.

FIGURE 9. A component of $\widehat{\mathcal{M}}_q^p$.

We would now like to see that Φ_Γ induces an isomorphism. To achieve this we consider two paths Γ_0 and Γ_1 from (f_0, g_0) to (f_1, g_1) and let D be a path from Γ_0 to Γ_1 . That is,

$$D = \{(f_{s,t}, g_{s,t}) \text{ for } (s, t) \in [0, 1] \times \mathbb{R} \text{ and } \Gamma_0 = (f_{0,t}, g_{0,t}), \Gamma_1 = (f_{1,t}, g_{1,t})\}.$$

Now if $p \in \text{Crit}(f_0)$ and $q \in \text{Crit}(f_1)$ then we set

$$\widehat{\mathcal{M}}_q^p = \{(\gamma, s) : \gamma \in \widetilde{\mathcal{M}}_q^p(s)\},$$

where $\widetilde{\mathcal{M}}_q^p(s)$ is $\widetilde{\mathcal{M}}_q^p$ for the path $\Gamma_s = (f_{s,t}, g_{s,t})$. We have the “same” Morse Theorems 1 through 4 for $\widehat{\mathcal{M}}_q^p$ except that

$$\dim \widehat{\mathcal{M}}_q^p = \text{index } p - \text{index } q + 1$$

and the “obvious” changes to the compactness and gluing theorems.

Exercise 2.32. Figure out what the “obvious” changes are.

Now we can define $K: C_n^0 \rightarrow C_{n+1}^1$ by

$$K(p) = \sum_{q \in \text{Crit}_{n+1}(f_1)} |\widehat{\mathcal{M}}_q^p| q.$$

This gives us a chain homotopy equivalence between Φ_{Γ_0} and Φ_{Γ_1} .

Theorem 2.33. With the notation above

$$\partial_1 \circ K + K \circ \partial_0 + \Phi_{\Gamma_0} + \Phi_{\Gamma_1} = 0.$$

Exercise 2.34. Using the previous exercise, prove this theorem.

Hint: Just as in the previous theorems, notice that the count of points in $\partial \widehat{\mathcal{M}}_q^p$ is exactly the coefficient of q in the application of the left-hand side of the formula in the theorem to p .

Of course, standard homological algebra tells us that the previous theorem shows Φ_{Γ_0} and Φ_{Γ_1} induce the same map on homology (or the way it is stated, one map is minus the other).

Corollary 2.35. *With the notation above the maps Φ_{Γ_0} and $-\Phi_{\Gamma_1}$ induce the same maps on homology.*

To complete our proof that the Morse homology is independent of the choices of Morse function and Riemannian metric, we have the following exercises.

Exercise 2.36. *Let Γ be the constant path from (f, g) to (f, g) . Show that Φ_Γ induces the identity map on homology.*

Hint: Of course, one must perturb Γ so that it is “generic” and Theorems 1' through 4' hold, but this perturbation can be chosen to be arbitrarily small.

Exercise 2.37. *If Γ_1 is a path from (f_0, g_0) to (f_1, g_1) and Γ_2 is a path from (f_1, g_1) to (f_2, g_2) , then we can define the concatenated path Γ that agrees with Γ_1 for $t \leq 1$ and agrees with Γ_2 , shifted by 2, for $t \geq 1$. Show that $\Phi_{\Gamma_2} \circ \Phi_{\Gamma_1} = \Phi_\Gamma$ on homology.*

Theorem 2.38. *The Morse homology of M is independent of the Morse function f and Riemannian metric g .*

Exercise 2.39. *Use the above exercises and theorems to prove this.*

2.5. A few details of the main theorems. In this section, we will give an indication of how to prove Morse Theorems 1 through 4 (and hence 1' through 4' too). We will not give complete proofs as they are quite intricate and involve a lot of analysis that can obscure the main ideas on a first pass at the subject. One may find more details in [24, 30, 41]; in particular, the presentation here was partially inspired by the second listed reference. We also note that one can define Morse homology with chain groups being $\mathbb{Z}/2\mathbb{Z}$ vector spaces generated by critical points, and in this case, one does not need Morse Theorem 2 about the orientability of the spaces $\mathcal{M}_q^p$. In many situations, it is common for people to work over $\mathbb{Z}/2\mathbb{Z}$ and ignore this theorem. We will do that here and only discuss the other theorems.

We will not describe the ideas of the proof of the main theorems for the other homology theories, but they mostly follow the same general idea (though details are quite different), except for the proof of the compactness result, which are quite different.

2.5.1. Morse Theorem 1. Given a Morse function $f: M \rightarrow \mathbb{R}$ and a Riemannian metric g on M we define the *path space* between two points p and q to be

$$\mathcal{P}_q^p = \{\gamma: \mathbb{R} \rightarrow M : \lim_{t \rightarrow -\infty} \gamma(t) = p \text{ and } \lim_{t \rightarrow \infty} \gamma(t) = q\}.$$

We will need to know what the tangent space of $\mathcal{P}_q^p$ is. (At this point, we are ignoring issues of smoothness. That is, we are not specifying if γ is smooth or C^k or something else. We will discuss this later.) To this end, recall that one may think of a tangent vector at a point as the equivalence class of paths through that point that have the same first derivative at that point. So if $\gamma \in \mathcal{P}_q^p$, then we can consider a path $s \rightarrow \gamma_s$ for $s \in (-\epsilon, \epsilon)$ with $\gamma_0 = \gamma$. Notice that $\gamma_s(t_0)$ for a fixed t_0 is a path in M such that $\gamma_0(t_0)$ is $\gamma(t_0)$. Thus

$$v(t_0) = \left. \frac{\partial \gamma_s(t_0)}{\partial s} \right|_{s=0}$$

is a vector in $T_{\gamma(t_0)}M$. See Figure 10. So $v(t)$ for $t \in \mathbb{R}$ is a vector field along γ , by this we mean that for each $\gamma(t)$ the vector $v(t) \in T_{\gamma(t)}M$. Summing up, given a path in $\mathcal{P}_q^p$ through

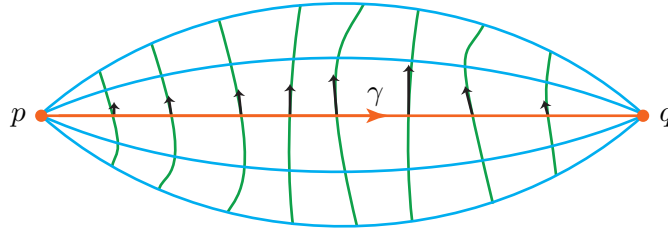


FIGURE 10. The blue is a path γ_s in $\mathcal{P}_q^p$ through γ (shown in orange). The green curves are the paths $s \mapsto \gamma_s(t_0)$ for some fixed $t_0 \in \mathbb{R}$ and the black vectors give the vector field v along γ .

γ its “derivative” at $\gamma(t)$ is a vector field $v(t)$ along γ . We also know that since $\gamma_s \in \mathcal{P}_q^p$ we must have that $v(t) \rightarrow 0$ as $t \rightarrow \pm\infty$.

We can easily see the converse too. Given any vector field $v(t)$ along γ that decays to 0 as t approaches $\pm\infty$, extend v to a vector field V on all of M . Let $\Gamma: M \times \mathbb{R} \rightarrow M$ be the flow of V . We can now define

$$\gamma_s(t) = \Gamma(\gamma(t), s).$$

This gives a path in $\mathcal{P}_q^p$ through γ and by the properties of the flow of a vector field, we see that the vector field associated to γ_s along γ is simply v . Thus, vectors tangent to $\mathcal{P}_q^p$ at γ are in one-to-one correspondence with vector fields along γ that decay to 0 as t approaches $\pm\infty$. So we see

$$\begin{aligned} T_\gamma(\mathcal{P}_q^p) &= \{\text{vector fields along } \gamma \text{ that decay to 0 as } t \text{ approaches } \pm\infty\} \\ &= \Gamma_d(\gamma^*TM) \end{aligned}$$

where γ^*TM is the pull-back of TM to $\mathbb{R}$ by γ and Γ_d denotes sections of this bundle that decay as indicated.

We now suppose that p and q are critical points of f and consider the map

$$S: \mathcal{P}_q^p \rightarrow T\mathcal{P}_q^p: \gamma \mapsto \gamma'(t) + \nabla f_{\gamma(t)},$$

where ∇f is the gradient of f with respect to g . Clearly $\gamma'(t) + \nabla f_{\gamma(t)}$ is an element of $\Gamma_d(\gamma^*TM)$ as desired and also that the space of *parameterized* flow lines, which we denote by $\underline{\mathcal{M}}_q^p$, is

$$\underline{\mathcal{M}}_q^p = S^{-1}(Z)$$

where Z is the zero section of $T\mathcal{P}_q^p$. We note that $\mathbb{R}$ acts on $\underline{\mathcal{M}}_q^p$ freely by reparameterization and thus if $\underline{\mathcal{M}}_q^p$ is a manifold then so is $\mathcal{M}_q^p$. So we have managed to describe the space we are interested in, that is $\mathcal{M}_q^p$, as the preimage of a submanifold Z under some function S (after modding out by reparameterization). Thus, from standard differential topology, if S were “transverse” to Z then we would expect $\underline{\mathcal{M}}_q^p$ to be a manifold. Of course, the problem here is that $\mathcal{P}_q^p$ and $T\mathcal{P}_q^p$ are infinite-dimensional, and a standard differential topology course does not deal with these. So we will take a brief digression to address this and then return to a sketch of the proof that $\mathcal{M}_q^p$ is a manifold.

2.5.2. Recollections on transversality and infinite dimensional manifolds. We begin by recalling the “standard way” to create functions transverse to submanifolds in finite dimensions. Given a smooth function $f: X \rightarrow Y$ between manifolds and a submanifold Z of Y we would like to “make” f transverse to Z , by which we mean we would like to find arbitrarily small perturbations of f so that it is transverse to Z (if it is not already transverse). The idea is to try to choose a perturbation space P and extend f to a map $F: X \times P \rightarrow Y$. Since P adds dimensions to the domain, F has a better chance of being transverse to Z (and of course, for the right space P , it will be). We then use the following standard theorem from differential topology.

Theorem 2.40. *With the notation in the previous paragraph, assume that F is transverse to Z and let $\pi: X \times P \rightarrow P$ be the projection map. If p is a regular value of $\pi|_{F^{-1}(Z)}: F^{-1}(Z) \rightarrow P$, then $f_p(\cdot) = F(\cdot, p): X \rightarrow Y$ is transverse to Z .*

Exercise 2.41. *Prove this theorem for X, Y , and P of arbitrary dimension.*

In finite dimensions we have Sard’s Theorem, which tells us when there is a dense set of regular values p .

Theorem 2.42 (Sard’s Theorem). *If $h: M \rightarrow N$ is a C^k -map then the set of regular values of h is dense in N if $k > \max\{0, m - n\}$.*

Thus returning to our original problem stated above, in the finite dimensional setting, if F is smooth enough, then there will be $f_p: X \rightarrow Y$ transverse to Z for a dense set of p and thus f may be “approximated” by a map transverse to Z (once we find F !).

The infinite-dimensional setting is a bit more complicated. For this, we begin with a few definitions. First, recall that a linear space V with a norm $\|\cdot\|: V \rightarrow \mathbb{R}$ is a *Banach space* if any Cauchy sequence in V (that is sequence v_i such that for any $\epsilon > 0$ there is some N such that $\|v_i - v_j\| < \epsilon$ for all $i, j > N$) converges to a point in V . Given two Banach spaces V and W (we will suppress the norms, though they are essential) we say a bounded linear map $L: V \rightarrow W$ is a *Fredholm operator* if

- L has a closed image,
- the dimension of $\ker F$ is finite, and
- the dimension of $\text{coker} F$ is finite,

where the $\text{coker} F$ is simply $W/F(V)$ and is called the *co-kernel*. We define the index of a Fredholm operator to be

$$\text{index } F = \dim \ker(F) - \dim \text{coker}(F).$$

Finally, a *Banach manifold* M is a topological manifold with coordinate charts modeled on a Banach space $(V, \|\cdot\|)$ and having smooth transition maps. We can now generalize Sard’s theorem.

Theorem 2.43 (Sard-Smale). *Let X and Y be separable Banach manifolds and $h: X \rightarrow Y$ be a C^k -map such that*

$$dh_x: T_x X \rightarrow T_{f(x)} Y$$

is Fredholm of index l for all $x \in X$. If $k \geq \max\{1, l + 1\}$, then the regular values of h are dense in Y . (And for such a y , $h^{-1}(y)$ is a manifold of dimension l .)

With this theorem and an improved version of Theorem 2.40 we have the following result that will be useful to prove $\mathcal{M}_q^p$ is a manifold. We first set up some notation. Let X and P be Banach manifolds, and $\pi: E \rightarrow X \times P$ be a Banach vector bundle (the reader should think through what this might mean). Suppose $\phi: (X \times P) \rightarrow E$ is a section and Z is the zero section of E . We want to discuss when ϕ is transverse to Z . Note that $\ker d\pi_e: T_e E \rightarrow T_{\phi(e)}(X \times P)$ is naturally identified with the fiber $E_{\phi(e)}$ of E above $\phi(e)$, since the tangent space to a linear space is naturally the linear space. So $T_e E = \ker d\pi_e \oplus H_e$ where H_e is a “horizontal” subspace of $T_e E$; that is, it is transverse to $\ker d\pi_e$ and its intersection with $\ker d\pi_e$ is $\{0\}$. In general H_e is determined by a “connection” on the bundle E but if $e \in Z$ then we can take $H_e = T_e Z$, so there is no need for a connection. We can now define $\nabla_{\phi(e)} \phi$ to be the composition of $d\phi_{\phi(e)}$ with the projection to $E_{\phi(e)}$ using the splitting $T_e E = E_{\phi(e)} \oplus T_e Z$.

Theorem 2.44. *With the notation above, suppose $\phi: (X \times P) \rightarrow E$ is a section of the bundle and for all $(x, p) \in \phi^{-1}(Z)$, we have*

- (1) $\nabla_{(x,p)} \phi: T_{(x,p)}(X \times P) \rightarrow T_{\phi(x,p)} \pi^{-1}((x, p))$ is surjective, and
- (2) if ϕ_p is the map $\phi_p: X \rightarrow E: x \mapsto \phi(x, p)$ then $\nabla_{(x,p)} \phi_p: T_x X \rightarrow T_{\phi(x,p)} \pi^{-1}((x, p))$ is Fredholm of index l .

Then, for a generic $p \in P$ we have $\phi_p^{-1}(Z)$ is a manifold of dimension l .

We can finally sketch the proof of Morse Theorem 1.

Sketch of the proof of Morse Theorem 1. Recall we have the space $\mathcal{P}_q^p$ of paths between critical points p and q of a Morse function f , and its tangent space at γ is $\Gamma_d(\gamma^* TM)$. Trying to apply the theorems above, we choose our space of perturbations to be the set P of all Riemannian metrics on M . Now let E be the bundle over $\mathcal{P}_q^p \times P$ with fiber at (γ, g) the space $\Gamma_d(\gamma^* TM)$. We can now define the section

$$\phi: (\mathcal{P}_q^p \times P) \rightarrow E: (\gamma, g) \mapsto \gamma'(t) + V(\gamma(t))$$

where $V = \nabla_g f$ is the gradient of f with respect to g .

Exercise 2.45. *Show that $\phi(\gamma, g)$ is the zero vector field along γ if and only if γ is a gradient flow line of $\nabla_g f$ from p to q .*

Now, if we can find the appropriate function spaces for $\mathcal{P}_q^p$, P , and E , such that ϕ is Fredholm and transverse to the zero section in E , then there would be a dense set of g for which the Moduli space $\mathcal{M}_q^p$ is a manifold. Moreover, if we can compute that the index of ϕ is $\text{index } p - \text{index } q$ then it would also have the desired dimensions.

For example, for P one could consider the space of C^k -smooth Riemannian metrics on M , and for the other spaces one could choose appropriate Sobolev spaces where some number of derivatives have bounded L^2 -integrals. To carry out the details would be a long diversion for these lecture notes, so we refer the reader to [30], but we do briefly indicate why ϕ will be transverse to the zero section in the right functional analytic setup.

Given a tangent vector $(\tilde{\gamma}, \tilde{g})$ in $T_{(\gamma, g)}(\mathcal{P}_q^p \times P)$ we can represent it by a path (γ_t, g_t) such that $(\gamma_0, g_0) = (\gamma, g)$ and $(\dot{\gamma}_t|_{t=0}, \dot{g}_t|_{t=0}) = (\tilde{\gamma}, \tilde{g})$ where the dot represents the derivative with respect to the t parameter. Now

$$\phi(\gamma_t, g_t)(s) = \gamma'_t(s) + V_{\gamma_t(s)}^t$$

where V^t is the gradient of f with respect to g_t , and the prime denotes the derivative with respect to s . Now

$$\begin{aligned} d\phi_{(\gamma,g)}(\tilde{\gamma}, \tilde{g}) &= \frac{d}{dt} \phi(\gamma_t, g_t)|_{t=0} \\ &= \frac{d}{dt} \left(\frac{d}{ds} \gamma_t(s) \right) \Big|_{t=0} + \frac{d}{dt} V_{\gamma_t(s)}^0 \Big|_{t=0} + \frac{d}{dt} V_{\gamma_0(s)}^t \Big|_{t=0}. \end{aligned}$$

If we choose our path so that $\gamma_t = \gamma$ for all t , then we see that

$$d\phi_{(\gamma,g)}(\tilde{\gamma}, \tilde{g}) = \frac{d}{dt} V_{\gamma(s)}^t \Big|_{t=0}.$$

We let $\dot{V}$ be the projection of $\frac{d}{dt} V_{\gamma(s)}^t \Big|_{t=0}$ to the fiber of the bundle E (that is to $\Gamma_d(\gamma^*TM)$). To work out $\dot{V}$, we recall that V^t is the gradient of f with respect to g_t . If we choose local coordinates $x_1, \dots, x_n$, then we can write

$$g_t = \sum_{i,j=1}^n (g_t)_{ij} dx^i \otimes dx^j$$

for some real valued functions $(g_t)_{ij}$. Thinking of $((g_t)_{ij})$ as a matrix we let $((g_t)^{ij})$ be its inverse and since

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

we see that

$$V^t = \sum_{i,j=1}^n (g_t)^{ij} \Big|_{t=0} \frac{\partial f}{\partial x_i} \frac{d}{dx_j}.$$

Thus

$$\dot{V} = \sum_{i,j=1}^n \frac{d}{dt} (g_t)^{ij} \Big|_{t=0} \frac{\partial f}{\partial x_i} \frac{d}{dx_j}.$$

We now suppose that v is a vector field along γ in $\Gamma_d(\gamma^*TM)$. In local coordinates we can write it $\sum_{i=1}^n a_i \frac{d}{dx_i}$. Let A be the matrix

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & \\ a_n & 0 & \cdots & 0 \end{pmatrix}$$

with v in the first column and first row and 0's elsewhere. Now if g is given by the matrix $G = (g_{ij})$ then define the matrix $G_t^{-1} = G^{-1} + tA$ and G_t as its inverse. For t small, the inverse will exist and define a Riemannian metric g_t . We can further choose coordinates near some point on γ so that $\nabla f = (1 \ 0 \ \cdots \ 0)^T$. Thus, we see that near this point on γ the vector $\dot{V} = v$. Using a cutoff function, we can find a global family of metrics g_t on M so that near the desired point we still have $\dot{V} = v$. That is, locally any element of $\Gamma_d(\gamma^*TM)$ is in the image of $\nabla_{(\gamma,g)} \phi$.

In the appropriate functional analytic setting, this implies that $\nabla_{(\gamma,g)} \phi$ is onto $E_{\gamma,g} = \Gamma_d(\gamma^*TM)$. We sketch some details here. To this end, we need to choose a Banach structure

on the bundle E . Given a fixed metric on M and we can define the L^2 -inner product on elements $f, g \in \Gamma_d(\gamma^*TM)$ to be

$$\langle f, g \rangle = \int_{\mathbb{R}} \langle f(s), g(s) \rangle ds,$$

where the inner product in the integrand is coming from the metric on M . This gives a norm $\|f\|^2 = \langle f, f \rangle$ and hence a metric $\Gamma_d(\gamma^*TM)$ thought of as smooth sections. We can now complete this space with respect to this metric to obtain $L_d^2(\gamma^*TM)$, the space of L^2 -integrable sections of γ^*TM .

We would now like to show that the orthogonal complement (with respect to the L^2 -inner product) of the image of $d\phi_{(\gamma, g)}$ in $L_d^2(\gamma^*TM)$ is the trivial vector space $\{0\}$. So let ω be an element of $L_d^2(\gamma^*TM)$ orthogonal to the image of $d\phi_{(\gamma, g)}$. Recall that $d\phi_{(\gamma, g)}(0, \tilde{g}) = \dot{V}$, where $\dot{V}$ is defined above. We will show that anything orthogonal to $\dot{V}$ for all $\dot{V}$ is actually zero (notice that this shows that a vector even being orthogonal to a subset of $\text{image}(d\phi_{(\gamma, g)})$ is enough to show it is zero). So we assume that, for any $\dot{V}$, we have

$$\int_{\gamma} \langle \dot{V}, \omega \rangle ds = 0.$$

Now, for any $x \in \mathbb{R}$, we can choose a neighborhood N of $\gamma(x)$ so that $N \cap \text{image}(\gamma)$ is the image of an interval I . Choose a function $h: \mathbb{R} \rightarrow \mathbb{R}$ so that $h = 1$ near x , $h = 0$ outside I , and $h \geq 0$. Now set $v = h\omega$. From above, we can choose $\tilde{g}$ so that $\dot{V} = v$ where h is non-zero. We now compute

$$0 = \int_{\mathbb{R}} \langle \dot{V}, \omega \rangle ds = \int_{\mathbb{R}} h \langle \omega, \omega \rangle ds.$$

This implies that $h\|\omega\|^2 = 0$ and so $\|\omega\| = 0$ near x . But since x was arbitrary $\omega = 0$, as we wanted to prove. Note, there are details here involving the Banach structure we use on the domain and the continuity (or lack thereof) of ω , but we leave it to the reader and the references above to sort out these details. The main point was to give an indication of the flavor of the argument necessary to prove some sort of “transversality” result.

In a similar fashion, but with more work, one may also show that $\nabla_{(x,p)}\phi$ for a fixed g satisfies the conditions in Theorem 2.44, which gives us Morse Theorem 1. $\square$

2.5.3. Morse Theorem 3. Given a Morse function $f: M \rightarrow \mathbb{R}$ we choose a Riemannian metric g so that Morse Theorem 1 holds. We want to consider a sequence of flow lines from the critical point p to the critical point q . Let γ_i be such a sequence and assume it is a Cauchy sequence, so it should converge. To analyze this situation, we note that one can choose a small ball B_1 about p so that B_1 contains no other critical points and any gradient flow line from p transversely intersects ∂B_1 in one point.

Exercise 2.46. Use the Morse Lemma, Theorem 2.7, to prove that such a ball exists.

Let x_i be the unique point at which γ_i intersects ∂B_1 . One may easily check that $\{x_i\}$ is a Cauchy sequence in the compact set ∂B and so it converges to a point x_∞ . See the upper left diagram in Figure 11. Now let $\tilde{\gamma}_1$ be the gradient flow line through x_∞ . If $\tilde{\gamma}_1$ flows to

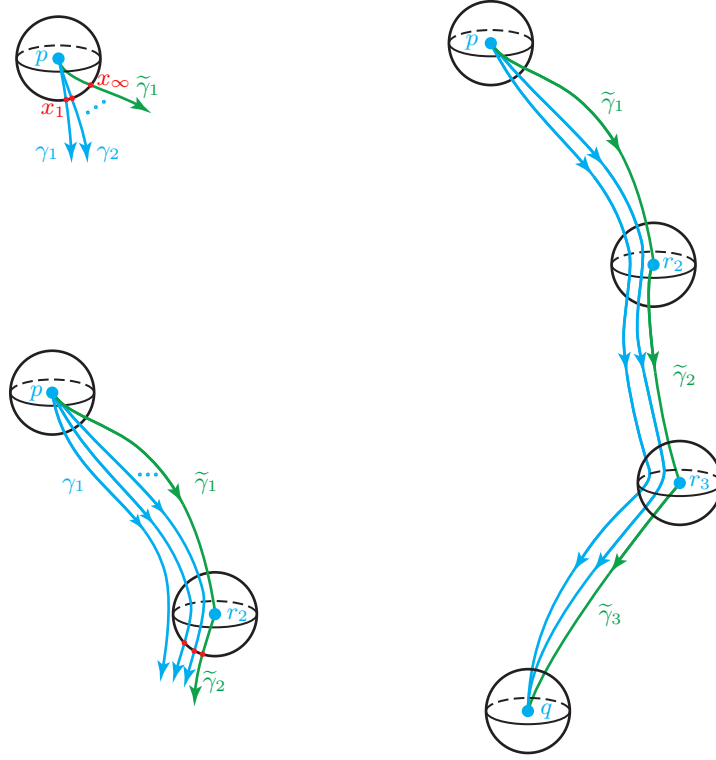


FIGURE 11. Converging to a broken flow line. In the top left, we see the sequence of gradient flow lines in $\mathcal{M}_q^p$ near p . They intersect ∂B_1 in points x_i that converge to x_∞ . The flow line $\tilde{\gamma}_1$ is the flow line through x_∞ . On the bottom left we see the critical point r_2 to which $\tilde{\gamma}_1$ flows. We also see the ball B_2 around r_2 used to find the next flow line in the broken flow line. On the right we see the complete broken flow line to which the γ_i converge.

q then the γ_i converge to an element of $\mathcal{M}_q^p$. We now need to see that the γ_i converge to a broken flow line.

So suppose $\lim_{t \rightarrow \infty} \tilde{\gamma}_1(t)$ is the critical point $r_2 \neq q$ (recall, we are taking $r_1 = p$). Since $\mathcal{M}_{r_2}^p$ is non-empty, we know its dimension is greater than or equal to 0 and so $\text{index } r_2 < \text{index } p$. We can parameterize the γ_i so that $\gamma_i(0) = x_i$ and by definition $\tilde{\gamma}_1(0) = x_\infty$. We can now choose a small neighborhood B_2 of r_2 as above (that is, so that any flow line asymptotic to r_2 is transverse to ∂B_2 , and other flow lines that intersect B_2 either intersect ∂B_2 transversely twice or are tangent to ∂B_2). We now know there is some $t_1 > 0$ such that $\tilde{\gamma}_1(t_1)$ is the unique intersection point of $\tilde{\gamma}_1$ with ∂B_2 . Since the gradient flow is smooth and the initial conditions of the γ_i converge to that of $\tilde{\gamma}_1$, we know that there is some N_1 such that for $i \geq N_1$, γ_i also intersects ∂B_2 transversely. Since the γ_i do not limit to r_2 we know they must intersect ∂B_2 at another point. Denote this point by x_i^2 . The sequence $\{x_i^2\}$ converges to a point x_∞^2 . We now let $\tilde{\gamma}_2$ be the flow line through x_∞^2 .

We can continue with this process until we find a $\tilde{\gamma}_n$ that flows to q .

Exercise 2.47. *Prove this last statement. Specifically, why must one of the trajectories in the broken flow line eventually end at q ?*

To see that the sequence γ_i converges to the broken flow line $\cup_{i=1}^n \tilde{\gamma}_i$ we note that one may take the balls B_i used to find the $\tilde{\gamma}_i$ to be arbitrarily small, so points of the γ_i in these balls are very close to the images of the $\tilde{\gamma}_i$. There are finite intervals of time for which the γ_i and $\tilde{\gamma}_i$ are outside the B_i , and thus by the smoothness of the flow, we know for large enough j , the images of the γ_j will be very close to the images of the $\tilde{\gamma}_i$.

Exercise 2.48. *Carefully write out this convergence argument.*

Remark 2.49. We note that the geometry of gradient flows that determines the “compactness” result is simply the fact that flow lines of vector fields depend smoothly on their initial conditions.

2.5.4. Morse Theorem 4. We now turn to the “gluing” theorem. The basic idea here is to build an “approximate flow line” and then prove there is a unique actual flow line associated with it. The second step is accomplished by a theorem of Floer [19, Proposition 24] who generalized prior results of Picard and Newton. More specifically, Newton gave an iterative scheme to find zeros of a real-valued function, and if one has an “approximate zero” and one has appropriate control over the derivative of the function, then one can guarantee there is actually a zero near the approximate zero. Picard then gave an iterative scheme to find a solution to an initial value problem with similar control over where the solution is relative to an approximate solution. Floer finds a version of these ideas that can be applied to partial differential equations. We will not state Floer’s result here as it is a bit technical, but we give an indication of why one would hope to be able to “guess” an almost solution to an equation and then hope that there is an actual solution nearby. To this end, suppose we have a function $f: \mathbb{R} \rightarrow \mathbb{R}$ and want to find a zero of f . Write f as

$$f(x) = f(0) + L(x) + N(x)$$

where L is a linear map (given by multiplication by $f'(0)$) and $N(x)$ is a non-linear map. Suppose we know that $f(x_0)$ is very small relative to $f'(0)$ (that is, the norm of L), then if N were not there, we know there would be a zero of f very close to x_0 . Now, if $N(x)$ is very small in comparison to $L(x)$ for all x near x_0 , then one can also conclude that there is a zero of f near x_0 , and we can estimate how close the zero is to x_0 in terms of a bound on $L^{-1}N(x)$.

Exercise 2.50. *Work out the details of the above discussion.*

The above gives a very simple indication of the type of information and bounds Floer uses to prove his result.

We are now left to see how to construct an approximate flow line to a give broken flow line. Recall that we saw above that the space of parameterized flow lines $\mathcal{M}_q^p$ is the preimage of the zero section of

$$\phi: (\mathcal{P}_q^p \times P) \rightarrow E: (\gamma, g) \mapsto \gamma'(t) + V(\gamma(t)).$$

So we want to find an approximate zero of ϕ and then use a theorem of Floer to prove that there is an actual zero of ϕ nearby and actually parameterize the nearby zeros.

Suppose $\tilde{\gamma}_1 \in \mathcal{M}_r^p$ and $\tilde{\gamma}_2 \in \mathcal{M}_q^r$ form a broken flow line. We can define a new path from p to q as follows. For a given positive number ρ choose a parameterization of the flow lines so that $\tilde{\gamma}_1(-1)$ and $\tilde{\gamma}_2(1)$ are the points where the flow lines intersect the ball B_2 of radius ρ about r discussed in the previous section. Now the path $\tilde{\gamma}_1 \#_\rho \tilde{\gamma}_2$ will be $\tilde{\gamma}_1$ for $t \leq -1$, $\tilde{\gamma}_2$ for $t \geq 1$ and for $t \in [-1, 1]$ will be a (well-chosen) path in B_2 . See Figure 12.

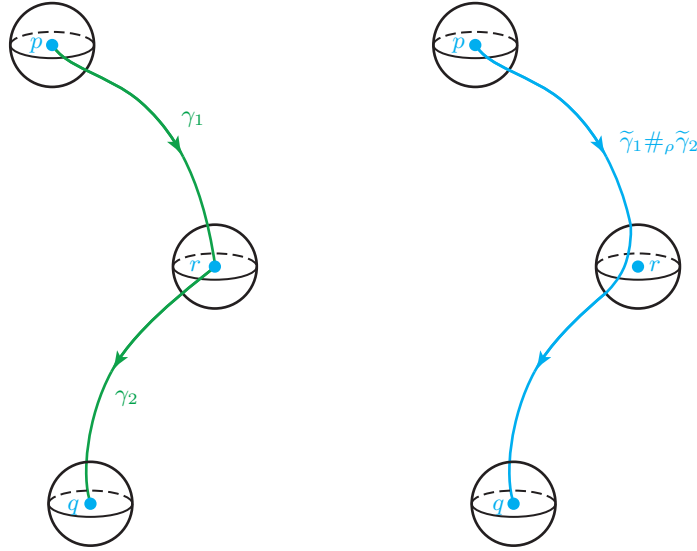


FIGURE 12. On the left we see a broken flow line, and on the right we see an approximate flow line near the broken flow line.

One can see that $\tilde{\gamma}_1 \#_\rho \tilde{\gamma}_2$ is “almost” a geodesic, in the sense that for most t values it is a geodesic, and if one chooses the path correctly for $t \in [-1, 1]$ then it is not far from being a geodesic on all of $\mathbb{R}$ (notice that this path can be chosen to have very small derivative and the gradient is also very small here so their difference is very small). With some work, one may use an infinite-dimensional version of the above mentioned result to show that near $\tilde{\gamma}_1 \#_\rho \tilde{\gamma}_2$ there is a 1-parameter family of flow lines in $\mathcal{M}_q^p$ that converge to the broken flow line $\tilde{\gamma}_1 \cup \tilde{\gamma}_2$.

2.5.5. Another approach. Given a critical point p of a Morse function $f: M \rightarrow \mathbb{R}$ and a metric g on M , we can define the *unstable manifold* of p to be

$$U_p = \{x \in M \text{ such that } \lim_{t \rightarrow -\infty} \gamma_x(t) = p\},$$

and the *stable manifold* of p to be

$$S_p = \{x \in M \text{ such that } \lim_{t \rightarrow \infty} \gamma_x(t) = p\}.$$

These are also sometimes referred to as the *descending manifold* and *ascending manifold*, respectively. One may prove, [2], that U_p and S_p are both open disks of dimension k and $n - k$, respectively, where $k = \text{index } p$.

We note that

$$\mathcal{M}_q^p = U_p \cap S_q.$$

One must interpret this equation correctly. Specifically, $U_p \cap S_q$ is the image of the elements in $\mathcal{M}_q^p$. So we can define Morse homology by counting the flow lines in $U_p \cap S_q$. One can interpret Morse Theorem 1 as the statement that all the stable and unstable manifolds of the critical points of f are transverse to each other. Some approaches to Morse homology do not define the moduli spaces $\mathcal{M}_q^p$, but instead focus on U_p and S_p . While this is a fine approach, the one presented above gives a better introduction to other homology theories. The main reason we mention this approach is the following fact.

Exercise 2.51. *Show that one may give M the structure of a CW complex using the unstable manifolds.*

With this observation, one can easily identify Morse homology with cellular homology (which is, of course, known to be isomorphic to singular homology).

3. LAGRANGIAN FLOER HOMOLOGY

One can think of Lagrangian Floer homology as an infinite-dimensional version of Morse homology applied to the “action functional on the space of paths between two Lagrangian submanifolds”, but it is not possible to carry out the details. The reason for this is that the “gradient flow equation” for the action functional turns out to be an elliptic partial differential equation (as opposed to a parabolic equation which is the infinite dimensional analog of the finite dimensional gradient flow equation discussed above). Elliptic equations do not typically have solutions for “initial value problems”. More specifically, given a path between the two Lagrangian submanifolds, one cannot expect to find a solution to these equations, and thus the definitions of the moduli spaces above do not work in this case. Luckily, we can reformulate the above moduli spaces in terms of this elliptic equation. More specifically, Floer reformulated this idea in terms of pseudo-holomorphic curves. We will forgo this intuitive idea behind the theory and go straight to the definition of Lagrangian Floer theory that can be made rigorous.

We begin by recalling a few basic facts from symplectic geometry to set the stage for Lagrangian Floer theory. The objects counted in the boundary map defining Lagrangian Floer theory are pseudo-holomorphic curves, so we discuss these in the next section. As we saw above, compactness issues are key to defining a homology theory, so in Section 3.3, we discuss Gromov compactness, which details the behavior of limits of pseudo-holomorphic curves (under nice conditions). With this background in hand, we then turn to defining Lagrangian Floer homology in Section 3.4, and then end with a section discussing applications of Lagrangian Floer homology.

3.1. A little symplectic geometry. There are many good introductions to symplectic geometry, for example [4, 35]. We refer the reader to those sources for a more comprehensive introduction to the subject, but discuss a few basic definitions and results here to set the stage for Lagrangian Floer homology.

A *symplectic manifold* is a pair (X, ω) , where X is a manifold and ω is a 2-form that is closed, that is $d\omega = 0$, and non-degenerate, that is if at any point $x \in X$ have have that $\omega(v, w) = 0$ for all $w \in T_x X$ implies that $v = 0$. We call ω a symplectic form, or symplectic structure, on X . Some of the first things one observes about a symplectic manifold is that it

dimension must be even, say $2n$ dimensional, and ω being non-degenerate is equivalently to ω^n being a volume form on X (that is that ω^n is a non-zero $2n$ -form at all points of X), where ω^n means the n -fold wedge product of ω with itself.

Example 3.1. The most basic example of a symplectic manifold is $(\mathbb{R}^{2n}, \omega_{std})$, where

$$\omega_{std} = \sum_{i=1}^n dx_i \wedge dy_i,$$

and $(x_1, y_1, \dots, x_n, y_n)$ are Euclidean coordinates on $\mathbb{R}^{2n}$. One may easily check that $\omega_{std} = -d\lambda_{std}$ where

$$\lambda_{std} = \sum_{i=1}^n y_i dx_i,$$

and so ω_{std} is clearly closed. It is also a simple exercise to see that ω^n is a volume form. So we see that $(\mathbb{R}^{2n}, \omega_{std})$ is indeed a symplectic manifold.

In general, if a symplectic form ω is the differential of a 1-form, then we call it an *exact symplectic form*. So the above example is an example of an exact symplectic form.

Example 3.2. Another key example of a symplectic manifold is (T^*M, ω_{can}) where T^*M is the cotangent bundle of any smooth manifold M and ω_{can} is defined as follows. In local coordinates $(q_1, \dots, q_n)$ on an open set U in M we can take local coordinates for $T^*U = U \times \mathbb{R}^n$ to be $(q_1, \dots, q_n, p_1, \dots, p_n)$ where the last n -coordinates are coordinates for $\mathbb{R}^n$ in the basis $\{dq_1, \dots, dq_n\}$. We can now define

$$\lambda_{can} = \sum_{i=1}^n p_i dq_i$$

on U .

Exercise 3.3. Show that this defines a 1-form on all of T^*M . Specifically, cover M in coordinate charts and show that the λ_{can} defined above in each of these coordinate charts defines the same 1-form whenever the coordinate charts overlap.

Thus we have λ_{can} defined on all of T^*M so we define $\omega_{can} = -d\lambda_{can}$. This gives a closed 2-form that is easy to verify is also non-degenerate.

Symplectic geometry originally grew out of an elegant reformulation of classical mechanics and has many beautiful and deep applications. For the sake of brevity, we refer the reader to the references mentioned at the beginning of this section for more details.

A fundamental theorem in symplectic geometry, called the Darboux theorem, says that for any symplectic manifold (X, ω) and any point $x \in X$ there is a neighborhood U of x in X , a neighborhood V of the origin in $\mathbb{R}^{2n}$, and a diffeomorphism $\phi: U \rightarrow V$ such that $\phi^*\omega_{std} = \omega$. That is, any symplectic manifold, locally, looks like the example above! We note that a diffeomorphism ϕ from one symplectic manifold, (X, ω) , to another, (X', ω') , is called a *symplectomorphism*.

A common means to build symplectomorphisms is via Hamiltonian flows. Specifically, given a function $H: X \rightarrow \mathbb{R}$ on a symplectic manifold (X, ω) there is a unique vector field

v_H on X such that

$$\iota_{v_H}\omega = dH.$$

Exercise 3.4. *Using the non-degeneracy of ω , show that such a v_H exists and is unique.*

The function H is called a *Hamiltonian* and the vector field v_H is called a *Hamiltonian vector field*. Let $\phi_H: M \rightarrow M$ be the time 1 flow of v_H . One may easily check that ϕ_H is a symplectomorphism. We call ϕ_H a *Hamiltonian symplectomorphism*. We note that it is not important that we choose the time 1 flow of v_H because by scaling H we scale v_H and the time 1 flow of these rescaled vector fields is simply the time t flow of v_H for some $t \neq 1$. We note that on most symplectic manifolds, there are symplectomorphisms that are not Hamiltonian symplectomorphisms, but we do not pursue this matter here.

The last ingredient we need for our discussion of Lagrangian Floer homology is Lagrangian submanifolds. A *Lagrangian submanifold* L of a symplectic manifold (X, ω) is a submanifold of X having half the dimension of X and for which ω vanishes on the tangent space of L .

Example 3.5. Any oriented surface Σ can be made a symplectic manifold simply by choosing an area form ω on Σ . We note that any 1-dimensional manifold in Σ will be Lagrangian. So we have lots of examples of Lagrangian submanifolds.

Exercise 3.6. *Show that the product of symplectic manifolds is symplectic in a natural way, and the product of Lagrangian submanifolds in each symplectic manifold is a Lagrangian submanifold of the product symplectic manifold.*

From the previous exercise, we can find many Lagrangian tori in symplectic manifolds built as products of orientable surfaces.

We end this section with an example showing the key nature of Lagrangian submanifolds in the study of symplectic manifolds.

Example 3.7. Let (X, ω) be a symplectic manifold and $\phi: X \rightarrow X$ be a symplectomorphism. We now consider the symplectic manifold (W, ω_W) where $W = X \times X$ and $\omega_W = \omega \oplus -\omega$. To be a bit more precise about the definition of ω_W we let $\pi_i: W \rightarrow X$ be the projection of W to the i^{th} factor of W and then the more proper definition of ω_W is

$$\omega_W = \pi_1^*\omega - \pi_2^*\omega.$$

Now, if we let Γ_ϕ be the graph of ϕ , that is $\Gamma_\phi = \{(x, \phi(x)) : x \in X\}$, then Γ_ϕ is a Lagrangian submanifold of (W, ω_W) . Indeed, consider

$$\Phi: X \rightarrow W: x \mapsto (x, \phi(x)).$$

This map clearly parameterizes Γ_ϕ and the restriction of ω_W to the tangent space of Γ_ϕ is simply $\Phi^*\omega_W$. We now compute

$$\Phi^*\omega_W = \Phi^*\pi_1^*\omega - \Phi^*\pi_2^*\omega = \omega - \phi^*\omega = 0,$$

since ϕ is a symplectomorphism.

Exercise 3.8. *Considering the previous example, show that symplectomorphisms of (X, ω) correspond to a certain subset of Lagrangian submanifolds of (W, ω_W) .*

3.2. Pseudo-holomorphic curves. Pseudo-holomorphic curves have been a cornerstone of the study of symplectic manifolds since their introduction by Gromov in [27], and, in particular, are used in the definition of Lagrangian Floer homology. In this section, we will discuss a few basic facts about pseudo-holomorphic curves.

An *almost-complex structure* on a manifold M is a bundle map

$$J: TM \rightarrow TM$$

covering the identity on M such that $J^2 = -\text{id}_{TM}$. Notice that J gives each tangent space $T_x M$ the structure of a complex vector space.

Example 3.9. Show that a complex manifold has a canonical almost-complex structure.

We note that not all almost-complex structures on a manifold come from a complex structure on the manifold. In particular, there are many manifolds that do not admit a complex structure, but do admit almost-complex structures. However, it is true that on an oriented surface Σ complex structures and almost-complex structures are equivalent.

Now, given a closed Riemannian surface Σ then it has a natural complex structure (namely, one simply rotates a tangent vector anti-clockwise by $\pi/2$). We will denote this almost-complex structure j . Now given a manifold M with an almost-complex structure J , we say a map

$$u: \Sigma \rightarrow M$$

is *pseudo-holomorphic* if it intertwines the almost-complex structures, that is

$$du \circ j = J \circ du$$

where $du: T\Sigma \rightarrow TM$ is the derivative of u .

Example 3.10. If $M = \mathbb{C}$ with the almost-complex structure coming from the complex structure on $\mathbb{C}$, then show that a function $u: \Sigma \rightarrow \mathbb{C}$ is pseudo-holomorphic if and only if it is holomorphic in the sense of complex analysis.

Below, we will need to have our complex structure related to a symplectic structure. To discuss this relation, we consider the symplectic manifold (X, ω) . An almost-complex structure on X is said to be *tamed* by ω if

$$\omega(v, Jv) > 0,$$

for any non-zero tangent vector v . We also say that J is *compatible* with ω if J is tamed by ω and ω is J -invariant, which means that $\omega(Jv, Jw) = \omega(v, w)$. The main result we will need about compatible almost-complex structure is the following.

Lemma 3.11. *The space of almost-complex structures compatible with a fixed symplectic structure is non-empty and convex.*

The proof of this lemma is straightforward, but a bit involved. It can be found in the references mentioned at the beginning of this section.

3.3. Gromov compactness. As we saw with Morse homology, understanding the compactness of the geometric objects counted in the homology boundary map is a key to making a homology theory. So, before we start to define Lagrangian Floer homology, we will discuss what can happen when considering sequences of pseudo-holomorphic curves.

To this end, we first define “singular” Riemannian surfaces. We start with a surface Σ and a collection of properly embedded curves $\gamma_1, \dots, \gamma_n$. See Figure 13. Let $\hat{\Sigma}$ be the result

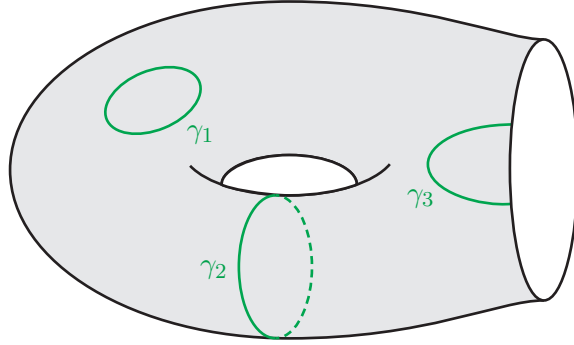


FIGURE 13. The properly embedded curves γ_i on Σ .

of collapsing each γ_i to a point, which we denote by z_i . See Figure 14. We note that if j is an almost-complex structure on Σ , then it induces one on $\hat{\Sigma} - \{z_1, \dots, z_n\}$, which we still denote by j . We call $(\hat{\Sigma}, j)$ a *cusped curve*. We note there is a natural quotient map

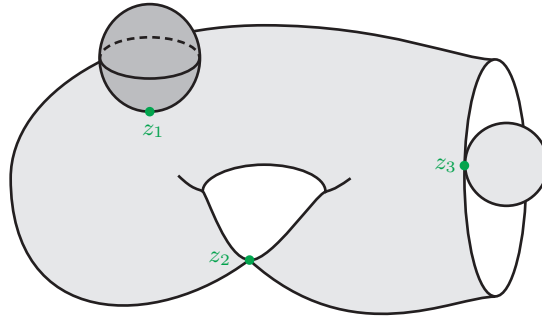


FIGURE 14. The cusped surface $\hat{\Sigma}$ and the images z_i of the γ_i .

$\pi: \Sigma \rightarrow \hat{\Sigma}$ and we define the *singular set* of $\hat{\Sigma}$ to be $s(\hat{\Sigma}) = \{z_1, \dots, z_n\}$.

Now, given an oriented Riemannian surface (Σ, j) we get a metric on Σ by setting $h(v, w) = \Omega(v, jw)$ where Ω is any area form on Σ giving the orientation on Σ . Given an almost-complex structure J compatible with a symplectic form ω on a manifold X can now define the *action* of a map $u: \Sigma \rightarrow X$ to be

$$\mathcal{A}(u) = \int_{\Sigma} \|du\|^2 \Omega,$$

where $\|du\|$ at the point x is defined to be

$$\|du\|_x = \max\{\|du(v)\|_g : v \in T_x \Sigma, \|v\|_h = 1\}$$

and $\|v\|_h$ is the length of v measured with respect to the metric h and similarly for $\|w\|_g$ where $g(v, w) = \omega(v, Jw)$. So pointwise $\|du\|_x$ measures the *maximal distortion of vectors in $T_x\Sigma$ by du_x* and $\mathcal{A}(u)$ is just the integral of this over Σ . We note that one can similarly define the action for a map from a cusped curve to X .

Exercise 3.12. *Show that*

$$\mathcal{A}(u) \leq \int_{\Sigma} u^* \omega,$$

and that one has equality if and only if u is pseudo-holomorphic.

Given a sequence of pseudo-holomorphic curves $u_k: (\Sigma, j_k) \rightarrow (X, J)$ and a cusped curve $\pi: \Sigma \rightarrow \widehat{\Sigma}$ with a almost-complex structure j , then we say the sequence $\{u_k\}$ converges to $u_{\infty}: \widehat{\Sigma} \rightarrow X$ if

- (1) there are maps $\phi_k: \Sigma \rightarrow \widehat{\Sigma}$ and curves $\gamma_1^k, \dots, \gamma_n^k$ such that

$$\phi_k|_{\Sigma - \cup \gamma_i}: (\Sigma - \cup \gamma_i) \rightarrow (\widehat{\Sigma} - s(\widehat{\Sigma}))$$

is a diffeomorphism,

- (2) the sequence of almost-complex structures $\{(\phi_k|_{\Sigma - \cup \gamma_i}^* j_k)\}$ C^∞ -converges to j on compact subsets as $k \rightarrow \infty$,
(3) the sequence $\{u_k \circ \phi_k^{-1}|_{\widehat{\Sigma} - s(\widehat{\Sigma})}\}$ C^∞ -converges to $u_{\infty}|_{\widehat{\Sigma} - s(\widehat{\Sigma})}$ on compact subsets as $k \rightarrow \infty$, and
(4) $\lim_{k \rightarrow \infty} \mathcal{A}(u_k) = \mathcal{A}(u_{\infty})$.

Theorem 3.13 (Gromov compactness, [27]). *Let (X, ω) be a symplectic manifold with an almost-complex structure J . If $u_k: (\Sigma, j_k) \rightarrow (X, J)$ is a sequence of pseudo-holomorphic maps with bounded action (here if, $\partial\Sigma \neq \emptyset$ then $u_k(\partial\Sigma) \subset L$ for some fixed Lagrangian submanifold L of X), then some subsequence of $\{u_k\}$ converges to a cusp curve $u_{\infty}: (\widehat{\Sigma}, j) \rightarrow (X, J)$. (We note that $s(\widehat{\Sigma})$ could be empty, in which case the u_k converge to a pseudo-holomorphic curve.)*

Remark 3.14. We note that Gromov's original theorem did not require a symplectic form, but merely a Riemannian metric, so that the action of the pseudo-holomorphic curves can be measured. We state this version as it is all that we will need below.

To get some sense of this type of convergence, we look at an example.

Example 3.15. Let $\Sigma = S^2$ be the 1-point compactification of $\mathbb{C}$ (with any fixed area form and almost-complex structure) and consider the maps

$$u_k: S^2 \rightarrow S^2 \times S^2: z \mapsto (z, 1/kz).$$

We can see the graphs of these functions in Figure 15. We notice that as $k \rightarrow \infty$ these maps converge to $u_{\infty}^1: S^2 \rightarrow S^2 \times S^2: z \mapsto (z, 0)$ way from $z = 0$. The problem is that the gradient of the u_k is blowing up at $z = 0$. So to see what the u_k converge to in a neighborhood of 0 we need to reparameterize the u_k . This is similar to when we studied geodesics, and we were only interested in the image of the curves and not their parameterization. If we reparameterize the u_k by the holomorphic map $\phi: S^2 \rightarrow S^2: z \mapsto 1/kz$ we see

$$u_k \circ \phi: S^2 \rightarrow S^2 \times S^2: z \mapsto (1/kz, z).$$

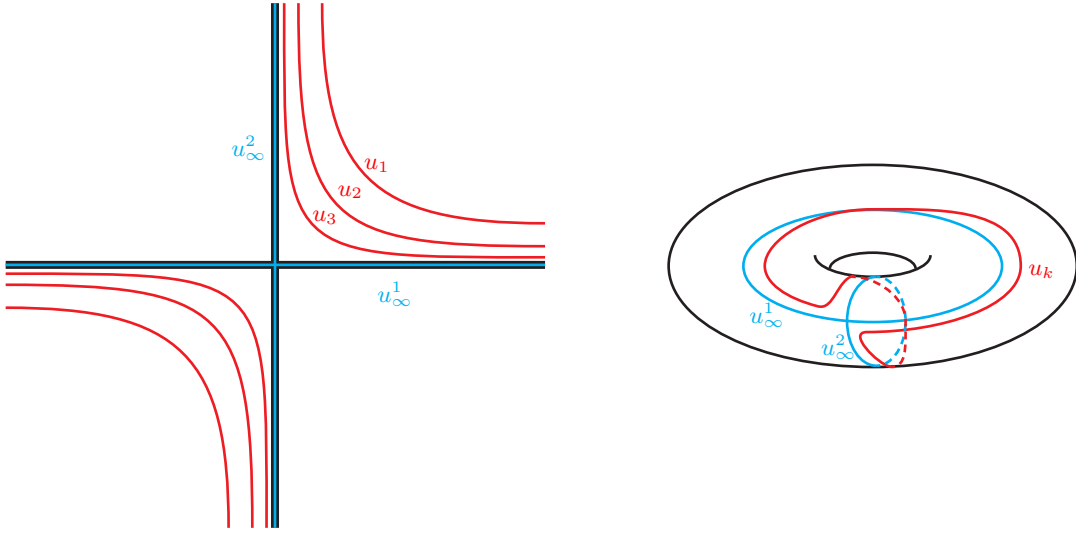


FIGURE 15. On the left are the images of the u_k , shown in red, and the u_∞^l , shown in blue. On the right we see the same picture where the spheres have not been punctured (though they are still, necessarily, half-dimensional, so the figure can be drawn).

And these maps converge to $u_\infty^2 : S^2 \rightarrow S^2 \times S^2 : z \mapsto (0, z)$. So the sequence of pseudo-holomorphic curves $u_k(S^2)$ converges to the cusped curve $u_\infty : \widehat{S}^2 \rightarrow S^2 \times S^2$ where $\widehat{S}^2$ is a wedge of two S^2 's and u_∞ is u_∞^1 on one of the spheres and u_∞^2 on the other.

We note in this simple example that we were able to find one global reparameterization to “see” both the limiting spheres, but in general, one might need to reparameterize each of the u_k separately, and this is the role of the ϕ_k in the definition of convergence.

3.4. Defining Lagrangian Floer homology. We are now ready to define Lagrangian Floer homology. We start with a symplectic manifold (X, ω) and two Lagrangian submanifolds L_0 and L_1 of X . Lagrangian Floer homology will be an invariant of the pair L_0 and L_1 up to Hamiltonian isotopy of either of the Lagrangian submanifolds.

We first note that one may slightly isotop the L_i so that they are transverse to each other. The chain complex for Lagrangian Floer homology is

$$FC(L_0, L_1) = \text{the free } R\text{-module generated by } L_0 \cap L_1,$$

where R is some ring of field like $\mathbb{Z}$ or $\mathbb{Z}/2\mathbb{Z}$.

To define the boundary map, we will consider “pseudo-holomorphic strips”. That is we will consider the surface $\mathbb{R} \times [0, 1]$ with coordinate (s, t) and the complex structure that sends $\frac{\partial}{\partial t}$ to $\frac{\partial}{\partial s}$ and $\frac{\partial}{\partial s}$ to $-\frac{\partial}{\partial t}$. Thus, if we choose an almost-complex structure J on X that is compatible with ω (this is like choosing a Riemannian metric when studying Morse homology), then a map $u : (\mathbb{R} \times [0, 1]) \rightarrow X$ is pseudo-holomorphic if and only if

$$\frac{\partial u}{\partial s} + J \frac{\partial u}{\partial t} = 0$$

where of course $\frac{\partial u}{\partial s}$ is the image of the vector $\frac{\partial}{\partial s}$ under du and similarly for $\frac{\partial u}{\partial t}$.

Exercise 3.16. *Prove the statement characterizing when u is pseudo-holomorphic.*

We can also consider a family of almost complex structures J_t compatible with ω and consider pseudo-holomorphic strips $u: (\mathbb{R} \times [0, 1]) \rightarrow X$ that satisfy

$$\frac{\partial u}{\partial s} + J_t \frac{\partial u}{\partial t} = 0.$$

This extra flexibility in the definition of a pseudo-holomorphic strip can be helpful in proving various technical results.

Now given two point $x, y \in L_0 \cap L_1$ we consider the space

$$\begin{aligned} \mathcal{M}_y^x = \{ & u: (\mathbb{R} \times [0, 1]) \rightarrow X \text{ such that} \\ & 1) u(s, i) \in L_i \text{ for } i = 0, 1, \\ & 2) \lim_{s \rightarrow -\infty} u(s, t) = x, \\ & 3) \lim_{s \rightarrow \infty} u(s, t) = y, \text{ and} \\ & 4) u \text{ is pseudo-holomorphic} \} / \mathbb{R}. \end{aligned}$$

To understand the $\mathbb{R}$ action above we note that $T_{s_0}: (\mathbb{R} \times [0, 1]) \rightarrow (\mathbb{R} \times [0, 1]): (s, t) \mapsto (s + s_0, t)$ is a holomorphic map so given any pseudo-holomorphic $u: (\mathbb{R} \times [0, 1]) \rightarrow X$ the map $u \circ T_{s_0}$ is also pseudo-holomorphic and hence $\mathbb{R}$ acts on the set of such pseudo-holomorphic maps.

It will be important for us to keep track of the homotopy class of an element $u \in \mathcal{M}_y^x$. To this end we define the set

$$\begin{aligned} \pi_2(x, y) = \{ & \text{homotopy classes of maps } u: (\mathbb{R} \times [0, 1]) \rightarrow X \text{ such that} \\ & 1) u(s, i) \in L_i \text{ for } i = 0, 1, \\ & 2) \lim_{s \rightarrow -\infty} u(s, t) = x, \text{ and} \\ & 3) \lim_{s \rightarrow \infty} u(s, t) = y \} \end{aligned}$$

For each $h \in \pi_2(x, y)$ we can now consider the spaces

$$\mathcal{M}_y^x(h) = \{u \in \mathcal{M}_y^x \text{ such that } u \in h\}$$

and notice that

$$\mathcal{M}_y^x = \bigcup_{h \in \pi_2(x, y)} \mathcal{M}_y^x(h).$$

Before we can state the main theorems to construct our boundary map, we need to consider an “index” for x and y . In Morse theory, we could assign an index to each critical point, and we could then determine the dimension of the moduli space of flowlines in terms of these indices. Unfortunately, we do not have a well-defined index for each $x \in L_0 \cap L_1$, but we do have a “relative index” for a pair of points $x, y \in L_0 \cap L_1$, and this will be sufficient to define the boundary map for our chain complex. This relative index is provided by the Maslov index or Maslov class.

3.4.1. *The Maslov class.* We begin by considering $\mathbb{R}^{2n} = \mathbb{C}^n$ with its standard symplectic structure ω_{std} defined in Example 3.1. We can now let $\mathcal{L}(\mathbb{C}^n)$ be the space of linear Lagrangian subspaces in $\mathbb{C}^n$. It is not too hard to show

$$\pi_1(\mathcal{L}(\mathbb{C}^n)) \cong \mathbb{Z}$$

and there is a “preferred” generator, see [40].

Exercise 3.17. *Prove this for when $n = 1$.*

Hint: Notice that $\mathcal{L}(\mathbb{C})$ is simply $\mathbb{R}P^2$ and a natural generator for its fundamental group is the loop that starts with the x -axis and rotates through an angle of π .

Now consider two linear Lagrangian subspaces Λ_0 and Λ_1 in $\mathbb{C}^n$ that intersect transversely.

Exercise 3.18. *Show that there is a complex linear map $I: \mathbb{C}^n \rightarrow \mathbb{C}^n$ such that $I(\Lambda_1) = \Lambda_0$.*

Using I from the exercise, we can define a map a path in $\mathcal{L}(\mathbb{C}^n)$ from Λ_1 to Λ_0 given by

$$\gamma_{\Lambda_0, \Lambda_1}(t) = e^{tI} \Lambda_1 = ((\cos t) \text{id}_{\mathbb{C}^n} + (\sin t)I) \Lambda_1,$$

as t goes from 0 to $\pi/2$. See Figure 16.

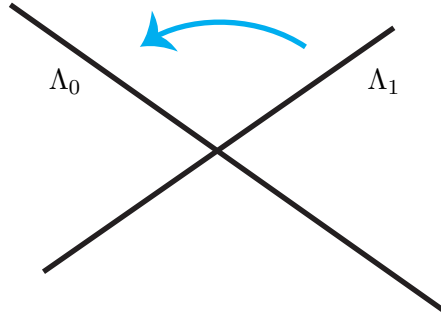


FIGURE 16. Two Lagrangian subspaces of $\mathbb{C}$. The map I is indicated by the arrow. The path $\gamma_{\Lambda_0, \Lambda_1}$ rotates Λ_1 anticlockwise to Λ_0 .

Now suppose that $u: \mathbb{R} \times [0, 1] \rightarrow X$ is a representative of an element in $\pi_2(x, y)$. Consider

$$u^*TX \cong (\mathbb{R} \times [0, 1]) \times \mathbb{C}^n.$$

Thus, if we have a linear Lagrangian subspace in $T_x X$ for some x in the image of u , then we can use the trivialization above to think of it as a Lagrangian subspace of $\mathbb{C}^n$, that is, as an element of $\mathcal{L}(\mathbb{C}^n)$. Thus restricting u to $\mathbb{R} \times \{i\}$, for $i = 0, 1$, will give paths $\gamma_i(t) = T_{u(t, i)} L_i$ in $\mathcal{L}(\mathbb{C}^n)$. We can think of extending the paths to $[-\infty, \infty]$ by continuity. So $\gamma_i(\infty)$ will be $T_y L_i$ and $\gamma_i(-\infty)$ will be $T_x L_i$. We can finally define the path

$$\gamma_u = \gamma_0 * \gamma_{\gamma_1(\infty), \gamma_0(\infty)} * \overline{\gamma_1} * \overline{\gamma_{\gamma_0(-\infty), \gamma_1(-\infty)}}$$

by concatenating the paths, where $*$ denotes concatenation and $\overline{\gamma}$ is the path γ run backwards. See Figure 17. It is clear that γ_u is a loop in $\mathcal{L}(\mathbb{C}^n)$ and so we define the *Maslov index* of u to be the integer associated to the homotopy class of this loop

$$\mu(u) \in \mathbb{Z} \cong \pi_1(\mathcal{L}(\mathbb{C}^n))$$

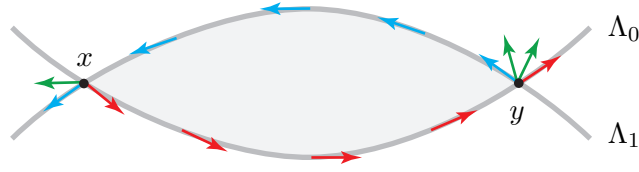


FIGURE 17. The gray is the image of a pseudo-holomorphic curve in $\mathbb{C}$ with boundary on Λ_0 and Λ_1 . We drew red and blue arrows whose span gives the tangent space to the Λ_i . The green arrows give the interpolation between the end of the red and blue arrows given by the paths $\gamma_{\gamma_1(\infty), \gamma_0(\infty)}$ and $\overline{\gamma_{\gamma_0(-\infty), \gamma_1(-\infty)}}$. (Note that the first path rotates anticlockwise a bit more than $\pi/2$ while the second path rotates clockwise a bit less than $\pi/2$.) This gives a loop in $\mathcal{L}(\mathbb{C})$, which we know to be $\mathbb{R}P^2$ from the exercise above. The Maslov index of this disk is 1.

We are now ready to state the first main theorem.

Floer Theorem 1 (Transversality)

Suppose L_0 and L_1 are Lagrangian submanifolds of (X, ω) that intersect transversely. For a generic choice of almost complex structure J on compatible with ω the space

$$\mathcal{M}_y^x(h)$$

is a manifold of dimension $\mu(u)$ for some u the class $h \in \pi_2(x, y)$, where x and y are intersection points of L_0 and L_1 .

One can also orient these moduli spaces. We note in the statement we need the Lagrangian submanifolds to be spin. Spin is a mild condition on a manifold that is not relevant for our main discussion, so we do not discuss it further.

Floer Theorem 2 (Orientability)

Suppose L_0 and L_1 are spin Lagrangian submanifolds of (X, ω) that intersect transversely. There is a natural way to orient the moduli spaces

$$\mathcal{M}_y^x(h).$$

3.4.2. Compactness and gluing. We have discussed compactness for pseudo-holomorphic curves in almost complex manifolds in Section 3.3, but we now consider the specific pseudo-holomorphic curves we will need in the definition of Floer homology.

Floer Theorem 3 (Compactness)

A sequence of pseudo-holomorphic curves (u_n, j_n) in $\mathcal{M}_y^x(h)$ that has $\mathcal{A}(u_n)$ bounded will have a subsequence that converges to either a pseudo-holomorphic curve in $\mathcal{M}_y^x(h)$ or a cusped pseudo-holomorphic curve as indicated in Figure 18.

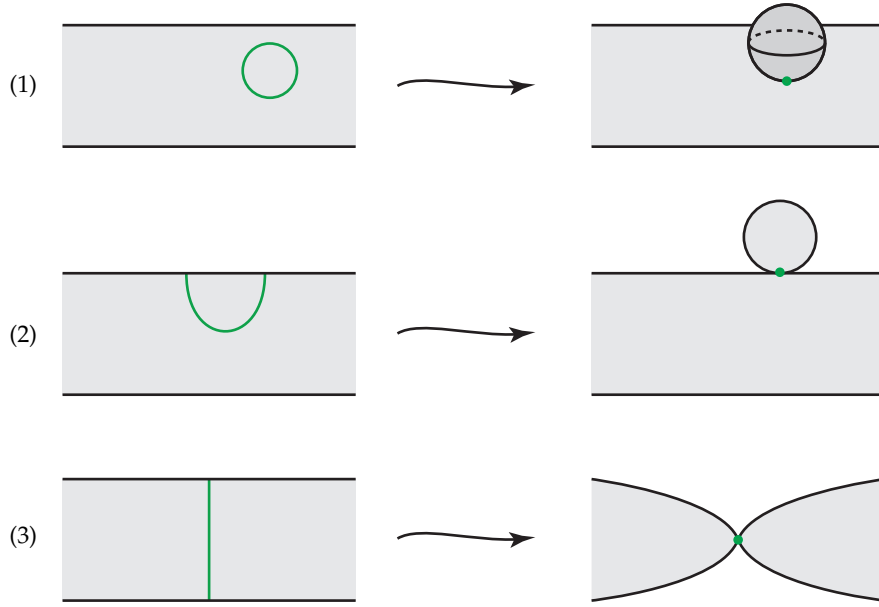


FIGURE 18. Cusped curves to which a pseudo-holomorphic strip can converge. In the first example, a sphere “bubbles off” in the limit. In the second example, a disk “bubbles off” in the limit. In the last case, the pseudo-holomorphic strip limits to a “broken strip”.

Finally, we have a gluing theorem. We state a vague version of the theorem as the full theorem can be a bit technical.

Floer Theorem 4 (Gluing)

Any cusped pseudo-holomorphic strip (that is, anything that can appear in the limit in Floer Theorem 3) can be “glued” to find a family of pseudo-holomorphic strips that converge to it.

We are now ready to return to the definition of Floer homology. Recall from the beginning of this section that given two Lagrangian submanifolds L_0 and L_1 of a symplectic manifold (X, ω) (after isotoping them so they are transverse to each other), we defined the chain complex for Lagrangian Floer homology to be

$$FC(L_0, L_1) = \text{the free } R \text{ module generated by } L_0 \cap L_1.$$

We now define the chain map

$$\partial: FC(L_0, L_1) \rightarrow FC(L_0, L_1)$$

on an element $p_+ \in L_0 \cap L_1$ by

$$\partial p_+ = \sum_{p_- \in L_0 \cap L_1} \left(\sum_{h \in \pi_2(p_+, p_-), \mu(h)=1} |\mathcal{M}_{p_-}^{p_+}(h)/\mathbb{R}| p_- \right).$$

Recall that $\mathcal{M}_{p_-}^{p_+}(h)$ has an $\mathbb{R}$ action so $\mathcal{M}_{p_-}^{p_+}(h)/\mathbb{R}$ is just its quotient by this action and $|\mathcal{M}_{p_-}^{p_+}(h)/\mathbb{R}|$ is the number of points in the 0-dimensional manifold $\mathcal{M}_{p_-}^{p_+}(h)/\mathbb{R}$.

Remark 3.19. For this sum to be finite, we need

$$(*) \quad \{\mathcal{A}(\phi) : [\phi] \in \pi_2(p_+, p_1)\} < \infty$$

This will happen if $|\pi_2(p_+, p_-)|$ is finite.

We now see when ∂ is a differential.

Lemma 3.20. *The map ∂ above is well-defined if $(*)$ holds. Moreover, if the limits (1) and (2) in Figure 18 do not occur, then $\partial^2 = 0$.*

Proof. To see that the map is well-defined, we need to see that $|\mathcal{M}_{p_-}^{p_+}(h)/\mathbb{R}|$ is compact (it is already a 0-manifold, so it will then be finite). Since we are assuming $(*)$ we can use the Gromov Compactness Theorem, Theorem 3.13, to see that a sequence of pseudo-holomorphic strips in $\mathcal{M}_{p_-}^{p_+}(h)$ that does not converge to a pseudo-holomorphic strip must converge to a cusped strip as in Figure 18. But we are assuming that the limits (1) and (2) do not occur, so the only limit we can have is (3). That is, our sequence u_n converges to a broken strip. That is it converges to the union of two strips, say u_∞^1 and u_∞^2 .

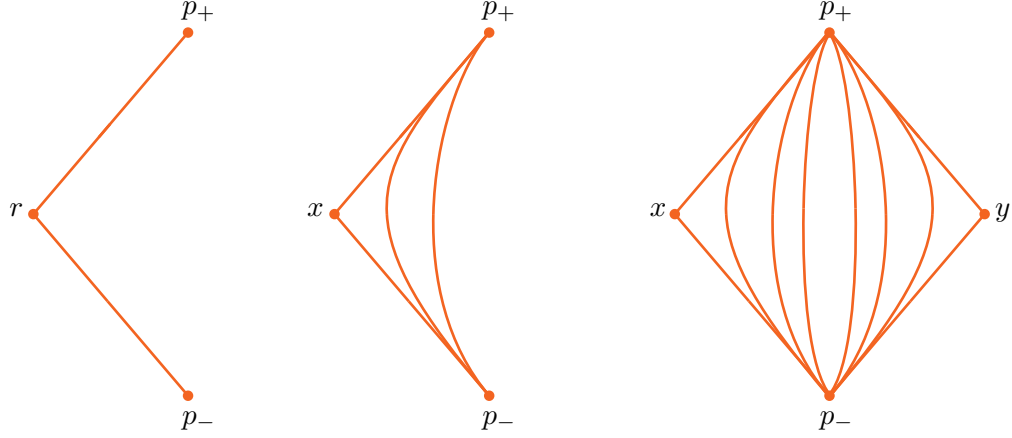
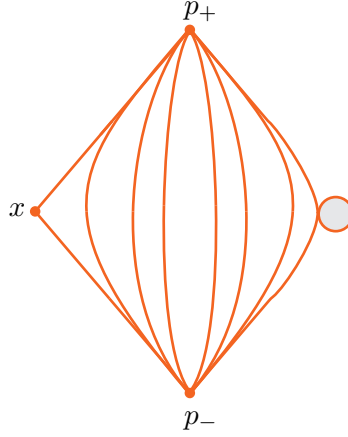
Exercise 3.21. *Show that*

$$\mu(u_n) = \mu(u_\infty^1) + \mu(u_\infty^2).$$

Since we know that $\mu(u_n) = 1$, we see that one of the u_∞^i must be in a 0-dimensional moduli space. But we know any pseudo-holomorphic strip is in a moduli space with an $\mathbb{R}$ -action, so the space cannot have dimension 0. Thus, our sequence cannot limit to (3) and must limit to another pseudo-holomorphic strip. In particular, $\mathcal{M}_{p_-}^{p_+}(h)$ is compact.

The proof that $\partial^2 = 0$ is analogous to the proof given in Theorem 2.26 for $\partial^2 = 0$ in Morse homology. Specifically, see Figure 19. There we see that if p_- is in $\partial^2 p_+$ (that is p_- is in ∂x where x is in ∂p_+), then the gluing theorem says there are nearby pseudo-holomorphic strips from p_+ to p_- that converge to the strip from p_+ to x and from x to p_- . Now, since limits (1) and (2) are not allowed, the only compactification of this moduli space is by another broken strip through some double point y . $\square$

Remark 3.22. Notice that if the limit (2) in Figure 18 could occur, then the last diagram in Figure 19 could look like Figure 20. In that case, p_- might only show up in $\partial^2 p_+$ an odd number of times, and hence $\partial^2 p_+ \neq 0$.

FIGURE 19. The proof that $\partial^2 = 0$.FIGURE 20. How $\partial^2 = 0$ can go wrong if one has undesirable limits. The limit on the right is supposed to represent a disk bubbling off.

Given Lagrangian submanifolds L_0 and L_1 satisfying $(*)$ and for which we do not have limits (1) and (2) in Figure 18, we can define the *Floer homology* of L_0 and L_1 to be

$$FH(L_0, L_1) = \frac{\ker(\partial: FC(L_0, L_1) \rightarrow FC(L_0, L_1))}{\text{image}(\partial: FC(L_0, L_1) \rightarrow FC(L_0, L_1))}$$

Remark 3.23. We note that one can define a relative grading on $FH(L_0, L_1)$ using the Maslov class, but this is a little involved (it might be only well-defined modulo some integer), so we do not address this here.

Recall, earlier we discussed that given a function $H: X \rightarrow \mathbb{R}$, there is a Hamiltonian vector field v_H defined by $\iota_{v_H}\omega = dH$. Any time t flow of v_H is called a *Hamiltonian diffeomorphism*.

Theorem 3.24. *Given a symplectic manifold (X, ω) and Lagrangian submanifolds L_0 and L_1 such that the limits (1) and (2) in Floer Theorem 3 do not occur. If $(*)$ holds, then $FH(L_0, L_1)$ is*

independent of the almost complex structures J_t chosen and

$$FH(\phi(L_0), \psi(L_1)) = FH(L_0, L_1)$$

for any Hamiltonian diffeomorphisms ϕ and ψ .

The proof of this theorem is similar to the proof that Morse homology is independent of the chosen metric and Morse function.

3.5. Applications. In this section, we will consider a simple situation when Floer homology is well-defined and a surprising construction that allows one to define an invariant of 3-manifolds using Lagrangian Floer homology.

3.5.1. Exact Lagrangians submanifolds. Let (X, ω) be an exact symplectic manifold, meaning that there is a 1-form λ such that $d\lambda = \omega$.

Suppose that Σ is a closed pseudo-holomorphic surface in X . Then $\omega(T_x \Sigma) > 0$ for all $x \in \Sigma$ (since, for any $v \in T_x \Sigma$ we know v and Jv span $T_x \Sigma$ and $\omega(v, Jv) > 0$ since J is tamed by ω). Thus, we see that

$$\mathcal{A}(\Sigma) = \int_{\Sigma} \omega > 0,$$

but

$$\int_{\Sigma} \omega = \int_{\Sigma} d\lambda = \int_{\partial \Sigma} \lambda = 0.$$

This contradiction indicates that there are no closed pseudo-holomorphic surfaces in X and thus limit (1) in Floer Theorem 3 cannot happen.

Now suppose that L is an *exact Lagrangian submanifold* of X . This means that there is some function $f: L \rightarrow \mathbb{R}$ such that $\lambda|_L = df$. Given a holomorphic surface Σ with boundary on L , we note as above that

$$\int_{\Sigma} \omega > 0,$$

but

$$\int_{\Sigma} \omega = \int_{\partial \Sigma} \lambda = \int_{\partial \Sigma} df = \int_{\partial(\partial \Sigma)} f = 0.$$

This contradiction shows that limit (2) in Floer Theorem 3 cannot happen.

Finally, if L_0 and L_1 are two exact Lagrangian submanifolds of X , with corresponding functions f_0 and f_1 , and $u: (\mathbb{R} \times [0, 1]) \rightarrow X$ is a pseudo-holomorphic strip with boundary on L_0 and L_1 from the intersection points p_+ to p_- , then

$$\begin{aligned} \mathcal{A}(u) &= \int_{\mathbb{R} \times [0, 1]} u^* \omega = \int_{\partial(\mathbb{R} \times [0, 1])} u^* \lambda = \int_{-\infty}^{\infty} u^* f_1 - \int_{-\infty}^{\infty} u^* f_0 \\ &= f_1(p_+) - f_1(p_-) - f_0(p_+) + f_0(p_-) \end{aligned}$$

Thus, $\mathcal{A}$ is bounded on $\pi_2(p_+, p_-)$ and so (*) is true.

From the above, we see the following result holds.

Theorem 3.25. *Floer homology is well-defined for exact Lagrangian submanifolds in an exact symplectic manifold.*

Now consider a Morse function $f: M \rightarrow \mathbb{R}$ on a manifold M . Let $\pi: T^*M \rightarrow M$ be the projection map from the cotangent bundle of M to M . If we set $F = f \circ \pi: T^*M \rightarrow \mathbb{R}$ we can let v_F be the Hamiltonian vector field associated to F on T^*M using the canonical symplectic structure on T^*M from Examples 3.2.

Now let L be the image of the zero section Z under the time 1 flow of v_F . It is easy to see that Z and L are exact Lagrangian submanifolds of the exact symplectic manifold T^*M . Thus, Floer homology is well defined. Floer [23] showed that rigid flow lines of the gradient of f are in correspondence with pseudo-holomorphic strips in the Floer complex of L and Z . From this, one can show that the Floer homology of Z and L is isomorphic to the Morse homology of M :

$$FH(Z, L) \cong H(M; \mathbb{Z}).$$

This leads to the following fundamental result.

Theorem 3.26. *If ϕ is a Hamiltonian diffeomorphism of T^*M then*

$$|(\phi(Z) \cap Z)| \geq \sum_{i=0}^n \dim(H_i(M; \mathbb{Z}))$$

where n is the dimension of M .

Proof. We know that the dimension of a chain complex is greater than or equal to the dimension of its homology, and the dimension of $FC(Z, L)$ is $|(\phi(Z) \cap Z)|$, so

$$|(\phi(Z) \cap Z)| \geq \dim FH(Z, L) \cong \dim H(M; \mathbb{Z}).$$

This establishes the theorem. $\square$

We now have the following famous corollary of Gromov [27] (though Gromov's proof was very different).

Theorem 3.27. *There is no closed exact Lagrangian submanifold in $(\mathbb{R}^{2n}, \omega_{std})$. (See Examples 3.1 for the symplectic structure on $\mathbb{R}^{2n}$.)*

Proof. Suppose there was a closed exact Lagrangian submanifold L in $(\mathbb{R}^{2n}, \omega_{std})$. Let $f: \mathbb{R}^{2n} \rightarrow \mathbb{R}$ be the projection onto the x_1 -axis. Then the flow of the Hamiltonian vector field v_f is translation in the y_1 -direction. Thus there is a Hamiltonian diffeomorphism $\phi: \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ such that $\phi(L) \cap L = \emptyset$. This implies that $FH(L, \phi(L)) = 0$ but from the previous theorem, and the fact that any Lagrangian submanifold has a neighborhood symplectomorphic to a neighborhood of the zero section in its cotangent bundle, we know that $FH(L, \phi(L))$ must be isomorphic to $H(L; \mathbb{Z})$ which is not trivial. Thus, no such L exists. $\square$

3.5.2. Heegaard Floer homology. We now discuss an application of Lagrangian Floer homology to 3-manifold topology. We only give a hint at this application and refer the reader to [38] for all the details.

Given a closed oriented 3-manifold M , there is always a surface Σ of some genus g that splits M into two genus g handlebodies. By this we mean that $M \setminus \Sigma = V_0 \cup V_1$ and there are g disjoint curves $\alpha_1, \dots, \alpha_g$ on Σ such that each α_i bounds a disk A_i in V_0 such that $V_0 \setminus \bigcup_{i=1}^g A_i$ is a ball, and similarly there are disjoint curves $\beta_1, \dots, \beta_g$ on Σ such that each β_i

bounds a disk B_i in V_1 such that $V_1 \setminus \bigcup_{i=1}^g B_i$ is a ball. This is called a *Heegaard splitting* of M . We note that the triple $(\Sigma, \{\alpha_i\}, \{\beta_i\})$ determines M , so if we can associate an invariant to this tripple then it might determine an invariant of M . It turns out that different triples can give the same M , so to really get an invariant of M we need to make sure that all triples for M give the same result.

Notice that choosing an area form on Σ , we see that Σ is a symplectic manifold and that any curve on Σ is Lagrangian. Thus the g -fold product of Σ with itself is a symplectic manifold and the products $\alpha_1 \times \cdots \times \alpha_g$ and $\beta_1 \times \cdots \times \beta_g$ are Lagrangian tori in this symplectic manifold. So we could try to apply Lagrangian Floer theory to these two tori, but this is not what we will do. We notice is that different labelings of the α_i and β_i determine the same manifold M , and we would like our construction not to depend on the ordering of these curves. To this end, we consider a quotient of the g -fold product of Σ by the symmetric group S_g (acting by permutting the coordinates). This is called the *symmetric product* and denoted $\text{Sym}^g(\Sigma)$. We can now look at the image of the above two tori in $\text{Sym}^g(\Sigma)$ and denote them by $\mathbb{T}_\alpha$ and $\mathbb{T}_\beta$. With a little work, one can show that $\text{Sym}^g(\Sigma)$ is also a symplectic manifold and the tori $\mathbb{T}_\alpha$ and $\mathbb{T}_\beta$ are Lagrangian. Now our symplectic manifold and Lagrangian submanifolds do not depend on the ordering of the α_i and β_i . With a great deal of effort one can show that the Lagrangian Floer homology of $\mathbb{T}_\alpha$ and $\mathbb{T}_\beta$ is well-defined, so now we could define an invariant of the triple to be the Lagrangian Floer homology

$$FH(\mathbb{T}_\alpha, \mathbb{T}_\beta).$$

This turns out to be an invariant of M , but not a particularly interesting one. To get an interesting invariant of M , we need to add one more thing. If one fixes a point $z \in \Sigma$ that is disjoint from the α_i and β_i then we can consider the image of $\{z\} \times \text{Sym}^{g-1}(\Sigma)$ in $\text{Sym}^g(\Sigma)$, denoted by F_z . This is a symplectic submanifold, and we can modify the differential in the Floer complex to take account of this submanifold in various ways. For example, we could ignore pseudo-holomorphic strips that intersect F_z , or we could add a formal variable to the coefficient ring of the chain complexes and then include a power of this variable that counts the intersections of the pseudo-holomorphic strips with F_z . One may check that doing either of these things will still produce a chain complex, and the homology of these complexes produces a powerful invariant of M . (Of course, to see it is an invariant of M , one needs to check that different triples and different choices of z lead to the same homologies.) One may find details of all of these facts above in [38] and some of the first applications in [37]. A nice, and more details, survey of the construction can be found in [39].

4. LEGENDRIAN CONTACT HOMOLOGY

We end this survey by discussing Legendrian contact homology. This is a version of Floer homology for Legendrian submanifolds of contact manifolds. It is a powerful invariant of Legendrian submanifolds and was first developed by Chekanov [5] and Eliashberg [13] in dimension 3, with the latter also discussing the higher-dimensional situation. They were able to show that there were Legendrian knots in $\mathbb{R}^3$ that had the same classical invariants but were not Legendrian isotopic because they had different Legendrian contact

homology. For more details on applications of Legendrian contact homology, please see [17]. The theory was developed for certain contact manifolds in higher dimensions in [7, 10] and can be used to produce powerful (complete!) invariants of smooth knots in $\mathbb{R}^3$, see [11].

We begin this section by discussing a few facts about contact geometry and Legendrian submanifolds. In the next section, we define Legendrian contact homology. A key feature, and the reason why we are discussing this here, is that the compactness issues that show up in this context do not allow us to define a homology theory as a vector space, but it will need to be an algebra. So we will see that we can follow the same program to create a homology theory as we did above, but the compactness issues will dictate the algebraic structure necessary to define our chain groups. We then end with a section discussing applications of Legendrian contact homology.

4.1. A little contact geometry. There are many good introductions to contact geometry, for example [14, 18, 25]. We refer the reader to those sources for a more comprehensive introduction to the subject, but discuss a few basic definitions and results here to set the stage for Lagrangian contact homology.

A contact manifold is a pair (M, ξ) , where M is a manifold of dimension $2n + 1$, and ξ is a sub-bundle of the tangent bundle TM , that has fibers of dimension $2n$ and is totally non-integrable. That is, ξ is a hyperplane field on M , and being totally non-integrable means that if it is tangent to a submanifold of M along an open set of the submanifold, then the submanifold has dimension less than or equal to n . One may use the Frobenius theorem to show that the total non-integrability is equivalent to ξ being (locally) defined by a 1-form α (that is $\xi = \ker \alpha$) and

$$\alpha \wedge (d\alpha)^n$$

is a never zero $2n + 1$ -form.

Example 4.1. Let $(X, d\lambda)$ be an exact symplectic manifold. Then the manifold $M = X \times \mathbb{R}$ has a contact structure ξ_λ given by the kernel of $dt - \lambda$, where t is the coordinate on $\mathbb{R}$.

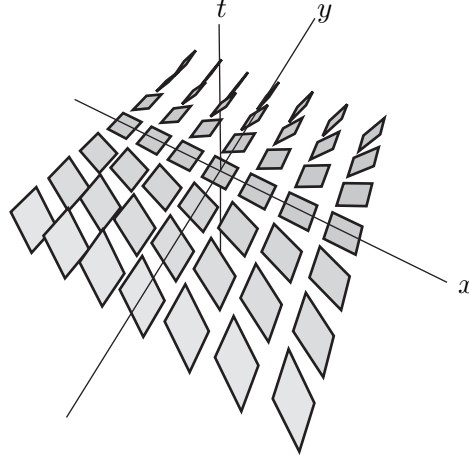
A specific example of this construction is $M = T^*W \times \mathbb{R}$ where we use the Liouville form λ_{std} on the cotangent bundle T^*W .

An even more specific example is $\mathbb{R}^{2n+1} = T^*\mathbb{R}^n \times \mathbb{R}$ with the contact form $\alpha = dt - \sum_{i=1}^n y_i dx_i$, where the x_i are coordinates on $\mathbb{R}^n$ and the y_i are coordinates on the fibers of $T^*\mathbb{R}^n$. See Figure 21 for the case when $n = 1$.

A submanifold L of a $(2n + 1)$ -dimensional contact manifold (M, ξ) is called *Legendrian* if it has dimension n and the tangent space of L is contained in ξ : $T_x L \subset \xi_x$ for all $x \in L$. Legendrian submanifolds are essential objects in contact geometry, see for example, [15, 16].

Suppose that L is a Legendrian submanifold of $(X \times \mathbb{R}, dt - \lambda)$ and let $\pi: (X \times \mathbb{R}) \rightarrow X$ be the projection map.

Exercise 4.2. Show that $\pi(L)$ is an immersed exact Lagrangian submanifold of the symplectic manifold $(X, -d\lambda)$.

FIGURE 21. The standard contact structure on $\mathbb{R}^3$.

The idea of Legendrian contact homology is to try to “do” Lagrangian Floer homology with $\pi(L)$. We will see below that this does not work, but with some modification, we can build a different homology theory that does work.

4.2. Defining Legendrian contact homology. As mentioned at the end of the last section, we would like to try to define some form of Lagrangian Floer homology for $\pi(L)$ where L is a compact Legendrian submanifold of $(X \times \mathbb{R}, dt - \lambda)$ and $\pi: (X \times \mathbb{R}) \rightarrow X$ is the projection map. To that end, observe that, possibly after a small perturbation, we can assume that $\pi(L)$ is embedded except for a finite number of transverse double points. We could then try to define a complex as a vector space generated by the double points of $\pi(L)$ and then, after choosing a complex structure J compatible with $-d\lambda$, define a differential by counting pseudo-holomorphic strips between double points. We begin by seeing that this does not work!

We will consider the situation where $X = T^*\mathbb{R}$ (though the same considerations occur in the more general situation). Consider a pseudo-holomorphic map $u: (\mathbb{R} \times [0, 1]) \rightarrow \mathbb{R}^2$ such that the boundary of $\mathbb{R} \times [0, 1]$ maps to $\pi(L)$ and as one approaches infinity, the map limits to double points. Notice that this map is a pseudo-holomorphic map from $\mathbb{R} \times [0, 1]$ to $\mathbb{R}^2$, so we can use our knowledge of single variable complex analysis to study this map. We first note that if there is a non-regular point of u on the interior of $\mathbb{R} \times [0, 1]$, then generically, we can assume it is of the form $z \mapsto z^n$ and, if $n > 2$, by a small perturbation, the critical point will break into several critical points which have local model $z \mapsto z^2$. So, assuming our critical point has a local model of the form $z \mapsto z^2$, we note that there is a 2-parameter family of maps near u with the same image that come by moving the critical point (since $\mathbb{R}^2$ is 2-dimensional). We also note that we still have the $\mathbb{R}$ actions on such maps, so we really have a 2-dimensional family, even after modding out by this $\mathbb{R}$ action. Since we are only interested in strips in 1-dimensional families (see the section on Lagrangian Floer homology), we can assume the maps u we are interested in do not have interior critical points.

Now consider a critical point on the boundary. Generically, we can assume it is also of the form $z \mapsto z^2$, but now the critical points must map to $\pi(L)$, and thus, there will be a 1-parameter family of maps near u . So again, if we are interested in rigid maps, we will only consider maps u with no critical points on the boundary or interior. That is, u is an immersion. However, when trying to prove $\partial^2 = 0$, we will need to consider 1-parameter families and so will consider u with a critical point on its boundary.

Now suppose x , y , and z are double points of $\pi(L)$ and $u: (\mathbb{R} \times [0, 1]) \rightarrow \mathbb{R}^2$ is a rigid pseudo-holomorphic strip from x to y with boundary in $\pi(L)$ and $u': (\mathbb{R} \times [0, 1]) \rightarrow \mathbb{R}^2$ is a similar strip from y to z . So if we defined our homology theory like in the case of Floer homology, then z would be counted in $\partial^2 x$. We can see such a situation in Figure 22. Figure 23 shows the family of disks formed when u and u' are glued. We notice that this

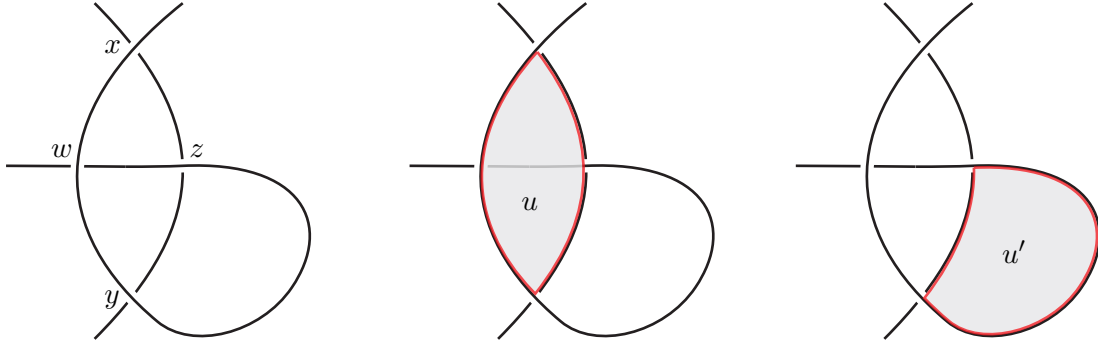


FIGURE 22. Part of the projection $\pi(L)$ is shown on the left. The middle diagram shows the image of u while the right-hand diagram shows the image of u' .

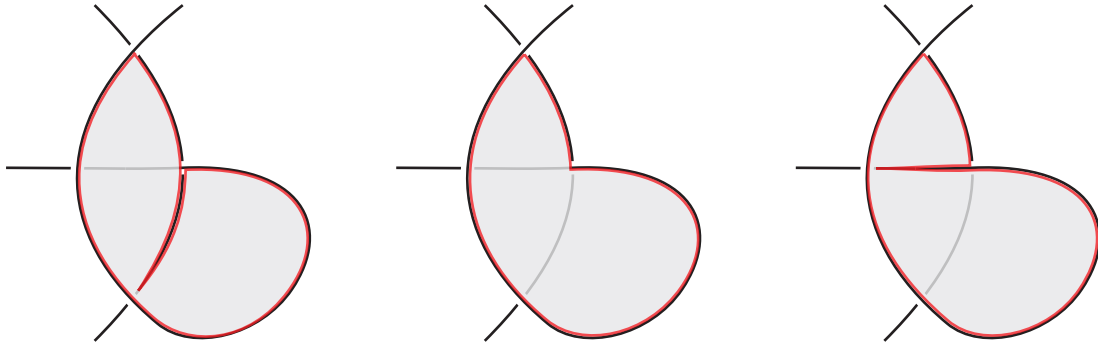


FIGURE 23. On the left, we see the image of a pseudo-holomorphic strip obtained by gluing u and u' . Notice the branched point on the boundary near the crossing y . In the middle, we see the middle of the family of pseudo-holomorphic strips where the “branched point” is “at infinity” or “at z ”. On the right, we see a pseudo-holomorphic strip with the branched point approaching w .

family will now break into pseudo-holomorphic disks as shown in Figure 24. So we see

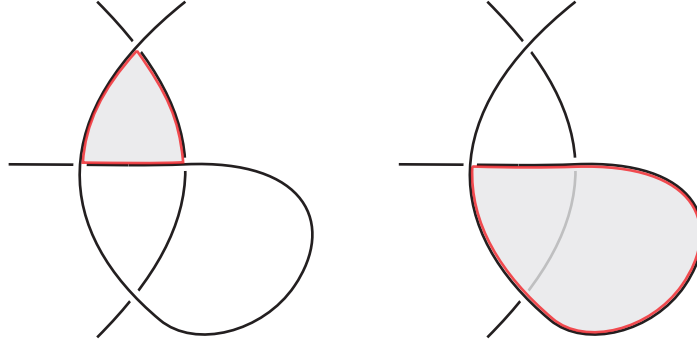


FIGURE 24. The two pseudo-holomorphic disks in the boundary of the family in Figure 23.

that the boundary of the compactified moduli spaces has a broken pseudo-holomorphic strip made out of two rigid pseudo-holomorphic strips and also a broken disk made out of a pseudo-holomorphic disk going through three double points and another pseudo-holomorphic disk going through only one double point.

Let us formalize these sorts of pseudo-holomorphic disks. Let $p_0, \dots, p_k$ be $k + 1$ points arranged anti-clockwise round the boundary of the unit disk D^2 in $\mathbb{R}^2$, and let D_{k+1} be $D^2 - \{p_0, \dots, p_k\}$. We will consider maps

$$u: D_{k+1} \rightarrow \mathbb{R}^2$$

with $u(\partial D_{k+1}) \subset \pi(L)$. We will be interested in such maps where the limits as we approach one of the p_i is a double point of $\pi(L)$. Moreover, we call a puncture p_i of u *positive* if as we transverse ∂D_{k+1} in the anti-clockwise direction we move from a lower sheet of $\pi(L)$ to an upper sheet of $\pi(L)$ and we will call the puncture *negative* if we move from an upper sheet to a lower sheet.

In this language, we see that if we “glue” two rigid pseudo-holomorphic disks from D_2 to $\mathbb{R}^2$, we can obtain a 1-parameter family of disks whose other boundary component consists of pseudo-holomorphic disks with domain D_3 and D_1 . So it is clear that the formalism in the definition of Floer homology for a pair of Lagrangian submanifolds from the last section will not work here! Moreover, it is not hard to see that if we also consider disks with 3 punctures, then in 1-parameter families we will also see disks with more punctures on their boundaries. We do note that if we glue a rigid pseudo-holomorphic disk with n punctures, one being positive and the other negative, to a disk with m punctures of the same type, then we get a 1-parameter family of pseudo-holomorphic disks with punctures that will have boundary two disks with punctures.

So if we try to define a differential of a double point in $\pi(L)$ we need to consider all such disks. So the differential of a double point x will need to count disks with a positive puncture at x and negative punctures at $x_1, \dots, x_k$. Then the chain complex cannot be a vector space generated by the double points, but it will need to be an algebra generated by the double points so that the boundary consists of words in the double points. **So we see that the compactness issues in the problem force us to choose an algebraic structure for the chain complex when trying to define a homology theory.**

Based on the above observations, we can now give the definition of Legendrian contact homology for Legendrian knots in $X \times \mathbb{R}$ where X is an exact symplectic manifold. To this end, if L is such a Legendrian submanifold, we let the chain complex $C(L)$ be the free algebra over R generated by the double points of $\pi(L)$, where R is the ring $\mathbb{Z}/2\mathbb{Z}$ or $\mathbb{Z}$ (there are more sophisticated rings as well).

We will now define a grading on $C(L)$. To simplify this, we assume that the tangent space of X can be globally trivialized: $TX \cong X \times \mathbb{C}^n$. This assumption is not necessary, but without it, the definition is more complicated. Now, given a double point $x \in \pi(L)$; there are two points in L that map to x . Call the one with the larger $\mathbb{R}$ -coordinate x^+ and the other x^- . Choose a path $\gamma_x: [0, 1] \rightarrow L$ from x^+ to x^- . Notice that $T_{\pi(\gamma_x(t))}\pi(L)$ is a Lagrangian subspace of $T_{\pi(\gamma_x(t))}X = \mathbb{C}^n$. Because the tangent bundle of X is trivial, we obtain a path of Lagrangians in $\mathbb{C}^n$. We can make this path a loop as we did in Section 3.4.1 above. The *Conley-Zehnder index* $\nu(x)$ of x is simply the Maslov index of this loop of Lagrangian subspaces. Notice that this number depends on the path γ we chose. To deal with this, we notice that we can get a loop of Lagrangian subspaces in $\mathbb{C}^n$ from any loop l in L . By taking the Maslov index of the loop of Lagrangian subspaces associated to $\pi(l)$, we obtain a map $\mu: H_1(L) \rightarrow \mathbb{Z}$. Let $c(L)$ be the smallest positive integer in the image of this map. It is not hard to see that $\nu(x)$ is well-defined modulo $c(L)$. Finally, we define the *grading* of a double point x to be $|x| = \nu(x) - 1$, and the grading of a word in double points is the sum of the gradings of the elements in the word. Our grading is well-defined modulo $c(L)$.

To define the differential of a point p , we consider pseudo-holomorphic maps $u: D_{k+1} \rightarrow X$ with $u(\partial D_{k+1}) \subset \pi(L)$, with the puncture p_0 being positive and mapping to x_0 and the other punctures being negative where p_i maps to x_i . Denote the moduli space of such maps, modulo pseudo-holomorphic reparameterization, by $\mathcal{M}_{x_1, \dots, x_k}^{x_0}$.

Legendrian Floer Theorem 1 (Transversality)

Suppose L is a Legendrian submanifold of $(X \times \mathbb{R}, dt - \lambda)$ such that $\pi(L)$ has transverse double points. For a generic choice of almost complex structure J compatible with $d\lambda$ or a generic perturbation of L , the space

$$\mathcal{M}_{x_1, \dots, x_k}^{x_0}$$

is a manifold of dimension $|x_0| - \sum_{i=1}^k |x_i| - 1$.

(We note that the dimension in this theorem is only well-defined modulo $c(L)$, since the grading is only defined modulo $c(L)$. This just means that the moduli space can have components of dimension $|x_0| - \sum_{i=1}^k |x_i| - 1 + nc(L)$ for any integer n .)

Legendrian Floer Theorem 2 (Orientability)

Suppose L is a Legendrian submanifold of $(X \times \mathbb{R}, dt - \lambda)$ such that $\pi(L)$ has transverse double points. If L is spin, then $\mathcal{M}_{x_1, \dots, x_k}^{x_0}$ is orientable.

For our compactness result, we need to see to what a sequence of pseudo-holomorphic disks in $\mathcal{M}_{x_1, \dots, x_k}^{x_0}$ can converge. To this end, we define a broken pseudo-holomorphic disk. Given an element $u \in \mathcal{M}_{x_1, \dots, x_k}^{x_0}$ and $v \in \mathcal{M}_{x_{k+1}, \dots, x_l}^{x_i}$ for some i between 1 and k and some $l \geq k$, we call the union of u and v a *broken pseudo-holomorphic disk*. We can define a sequence of disks u_n in $\mathcal{M}_{x_1, \dots, x_{i-1}, x_{k+1}, \dots, x_l, x_{i+1}, \dots, x_k}^{x_0}$ converging to the broken disk defined by u and v just as we did in Section 3.3. See Figure 25.

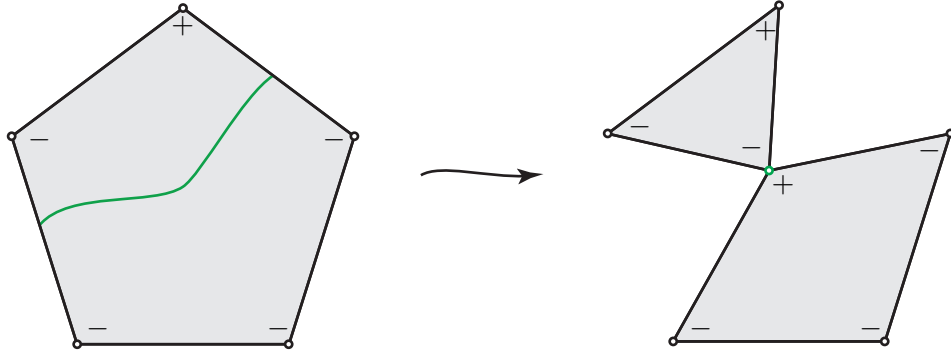


FIGURE 25. A pseudo-holomorphic disk converging to a broken pseudo-holomorphic disk.

We now have our compactness result.

Legendrian Floer Theorem 3 (Compactness)

Suppose L is a Legendrian submanifold of $(X \times \mathbb{R}, dt - \lambda)$ such that $\pi(L)$ has transverse double points. A sequence of pseudo-holomorphic disks in $\mathcal{M}_{x_1, \dots, x_k}^{x_0}$ has a subsequence that converges to a pseudo-holomorphic disk in $\mathcal{M}_{x_1, \dots, x_k}^{x_0}$ or a broken pseudo-holomorphic disk.

Remark 4.3. We note that, for simplicity, this theorem is not quite accurate in some cases, as we defined $\mathcal{M}_{x_1, \dots, x_k}^{x_0}$ to be pseudo-holomorphic disks modded out by pseudo-holomorphic reparameterization. When k is large, there are no such reparameterizations and the theorem is correct as stated, but for small k , the statement should be made before one considers reparameterization.

Just as we did while discussing Floer homology, we roughly state the gluing theorem as follows.

Legendrian Floer Theorem 4 (Gluing)

Any broken pseudo-holomorphic disk (that is, anything that can appear in the limit in Legendrian Floer Theorem 3) can be “glued” to find a family of pseudo-holomorphic disks that converge to it.

We are now ready to define our differential. We will define it on generators and then extend to the algebra using linearity and a signed Leibniz rule

$$\partial(w_1 w_2) = (\partial w_1) w_2 + (-1)^{|w_1|} w_1 \partial(w_2).$$

Let x be a double point of $\pi(L)$, that is, a generator of $C(L)$. We define the *differential of x* to be

$$\partial x = \sum_{\substack{x_1, \dots, x_k \\ |x| - \sum_{i=1}^k |x_i| = 1}} |\mathcal{M}_{x_1, \dots, x_k}^x| x_1 \cdots x_k,$$

where $|\mathcal{M}_{x_1, \dots, x_k}^x|$ is the signed count of the 0 dimensional components of $\mathcal{M}_{x_1, \dots, x_k}^x$, and k is any non-negative integer.

Arguing as in previous sections, we can easily check the following result.

Theorem 4.4. *The differential $\partial: C(L) \rightarrow C(L)$ lowers the grading by 1 and satisfies $\partial^2 = 0$.*

We can now define the *Legendrian contact homology* to be

$$LCH(L) = \frac{\ker(\partial: C(L) \rightarrow C(L))}{\text{image}(\partial: C(L) \rightarrow C(L))}.$$

Moreover, we can also prove the invariance of the homology under Legendrian isotopy and the choice of complex structure J , by using alternate versions of Legendrian Floer Theorems 1 through 4 in a similar way we proved that Morse homology is independent of the chosen metric and Morse function.

Theorem 4.5. *If L and L' are Legendrian isotopic, then $LCH(L) \cong LCH(L')$.*

4.3. Applications. It was an open question whether two Legendrian knots in the standard contact structure on $\mathbb{R}^3$ that were smoothly isotopic and had the same “classical invariants” (these are invariants associated with the tangent information of a Legendrian knot) had to be Legendrian isotopic. In the late 1990s, Chekanov [5] and Eliashberg [13] defined Legendrian contact homology and used it to show that such Legendrian knots do indeed exist. It turns out that it is very difficult to work with $LCH(L)$ since it is a non-abelian object. So it was important to be able to extract computable information from this difficult algebraic object. This was done using “augmentations” and “linearizations”. We refer the reader to [17] for a recent overview of this whole theory, along with several computations. In a series of papers [7, 8, 9, 10], Ekholm, Sullivan, and the author developed this theory in higher dimensions and also used it to distinguish many Legendrian submanifolds.

There is another interesting and surprising application of Legendrian contact homology. Specifically, one can use it to study and distinguish smooth knots in $\mathbb{R}^3$! This actually works for submanifolds in any dimensions, but we will restrict our discussion to knots in $\mathbb{R}^3$. We first construct our contact manifold. If we fix a metric on $\mathbb{R}^3$ and look at the unit cotangent bundle $U^*\mathbb{R}^3 \subset T^*\mathbb{R}^3$, we can restrict the Liouville form λ_{std} on $T^*\mathbb{R}^3$ to obtain a contact form on $U^*\mathbb{R}^3$. Now, if K is a knot in $\mathbb{R}^3$, then we can consider $U^*(K)$, the unit conormal bundles

$$U^*(K) = \{\alpha \in U_x^*\mathbb{R}^3 : x \in K \text{ and } \alpha \text{ vanishes on } T_x K\}.$$

One may easily check that $U^*(K)$ is a Legendrian torus in $U^*\mathbb{R}^3$; and, moreover, if K and K' are smoothly isotopic then $U^*(K)$ and $U^*(K')$ are Legendrian isotopic. Thus, the Legendrian contact homology of $U^*(K)$ is an invariant of the smooth isotopy class of K ! It is known that this is almost a complete invariant (and might be a complete invariant) [12]. For more on this story, see [11].

REFERENCES

- [1] Michèle Audin and Mihai Damian. *Morse theory and Floer homology*. Universitext. Springer, London; EDP Sciences, Les Ulis, 2014. Translated from the 2010 French original by Reinie Ern .
- [2] Augustin Banyaga and David Hurtubise. *Lectures on Morse homology*, volume 29 of *Kluwer Texts in the Mathematical Sciences*. Kluwer Academic Publishers Group, Dordrecht, 2004.
- [3] Raoul Bott. Morse theory indomitable. *Inst. Hautes  tudes Sci. Publ. Math.*, (68):99–114, 1988.
- [4] Ana Cannas da Silva. *Lectures on symplectic geometry*, volume 1764 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001.
- [5] Yuri Chekanov. Differential algebra of Legendrian links. *Invent. Math.*, 150(3):441–483, 2002.
- [6] S. K. Donaldson. *Floer homology groups in Yang-Mills theory*, volume 147 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2002. With the assistance of M. Furuta and D. Kotschick.
- [7] Tobias Ekholm, John Etnyre, and Michael Sullivan. The contact homology of Legendrian submanifolds in $\mathbb{R}^{2n+1}$. *J. Differential Geom.*, 71(2):177–305, 2005.
- [8] Tobias Ekholm, John Etnyre, and Michael Sullivan. Non-isotopic Legendrian submanifolds in $\mathbb{R}^{2n+1}$. *J. Differential Geom.*, 71(1):85–128, 2005.
- [9] Tobias Ekholm, John Etnyre, and Michael Sullivan. Orientations in Legendrian contact homology and exact Lagrangian immersions. *Internat. J. Math.*, 16(5):453–532, 2005.
- [10] Tobias Ekholm, John Etnyre, and Michael Sullivan. Legendrian contact homology in $P \times \mathbb{R}$. *Trans. Amer. Math. Soc.*, 359(7):3301–3335, 2007.
- [11] Tobias Ekholm, John B. Etnyre, Lenhard Ng, and Michael G. Sullivan. Knot contact homology. *Geom. Topol.*, 17(2):975–1112, 2013.
- [12] Tobias Ekholm, Lenhard Ng, and Vivek Shende. A complete knot invariant from contact homology. *Invent. Math.*, 211(3):1149–1200, 2018.
- [13] Yakov Eliashberg. Invariants in contact topology. In *Proceedings of the International Congress of Mathematicians, Vol. II (Berlin, 1998)*, number Extra Vol. II, pages 327–338 (electronic), 1998.
- [14] John B. Etnyre. Introductory lectures on contact geometry. In *Topology and geometry of manifolds (Athens, GA, 2001)*, volume 71 of *Proc. Sympos. Pure Math.*, pages 81–107. Amer. Math. Soc., Providence, RI, 2003.
- [15] John B. Etnyre. Legendrian and transversal knots. In *Handbook of knot theory*, pages 105–185. Elsevier B. V., Amsterdam, 2005.
- [16] John B. Etnyre. Submanifolds in contact geometry. arXiv 2510.05804, 2025.
- [17] John B. Etnyre and Lenhard L. Ng. Legendrian contact homology in $\mathbb{R}^3$. In *Surveys in differential geometry 2020. Surveys in 3-manifold topology and geometry*, volume 25 of *Surv. Differ. Geom.*, pages 103–161. Int. Press, Boston, MA, [2022]  2022.
- [18] John B. Etnyre and B lent Tosun. *Low dimensional contact geometry*. To appear.
- [19] A. Floer. Monopoles on asymptotically flat manifolds. In *The Floer memorial volume*, volume 133 of *Progr. Math.*, pages 3–41. Birkh user, Basel, 1995.
- [20] Andreas Floer. Morse theory for fixed points of symplectic diffeomorphisms. *Bull. Amer. Math. Soc. (N.S.)*, 16(2):279–281, 1987.
- [21] Andreas Floer. An instanton-invariant for 3-manifolds. *Comm. Math. Phys.*, 118(2):215–240, 1988.
- [22] Andreas Floer. Morse theory for Lagrangian intersections. *J. Differential Geom.*, 28(3):513–547, 1988.
- [23] Andreas Floer. Witten’s complex and infinite-dimensional Morse theory. *J. Differential Geom.*, 30(1):207–221, 1989.
- [24] Urs Frauenfelder and Robert Nicholls. The moduli space of gradient flow lines and morse homology, 2020.

- [25] Hansjörg Geiges. *An introduction to contact topology*, volume 109 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2008.
- [26] M. Golubitsky and V. Guillemin. *Stable mappings and their singularities*, volume Vol. 14 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Heidelberg, 1973.
- [27] M. Gromov. Pseudoholomorphic curves in symplectic manifolds. *Invent. Math.*, 82(2):307–347, 1985.
- [28] Victor Guillemin and Alan Pollack. *Differential topology*. AMS Chelsea Publishing, Providence, RI, 2010. Reprint of the 1974 original.
- [29] Michael Hutchings. An index inequality for embedded pseudoholomorphic curves in symplectizations. *J. Eur. Math. Soc. (JEMS)*, 4(4):313–361, 2002.
- [30] Michael Hutchings. Lecture notes on morse homology (with an eye towards floer theory and pseudoholomorphic curves). Lecture notes, <https://math.berkeley.edu/~hutching/teach/276-2010/mfp.ps>, 2002.
- [31] Michael Hutchings. Lecture notes on embedded contact homology. In *Contact and symplectic topology*, volume 26 of *Bolyai Soc. Math. Stud.*, pages 389–484. János Bolyai Math. Soc., Budapest, 2014.
- [32] Peter Kronheimer and Tomasz Mrowka. *Monopoles and three-manifolds*, volume 10 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2007.
- [33] Peter Kronheimer and Tomasz Mrowka. Knots, sutures, and excision. *J. Differential Geom.*, 84(2):301–364, 2010.
- [34] John M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.
- [35] Dusa McDuff and Dietmar Salamon. *Introduction to symplectic topology*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, second edition, 1998.
- [36] J. Milnor. *Morse theory*, volume No. 51 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1963. Based on lecture notes by M. Spivak and R. Wells.
- [37] Peter Ozsváth and Zoltán Szabó. Holomorphic disks and three-manifold invariants: properties and applications. *Ann. of Math. (2)*, 159(3):1159–1245, 2004.
- [38] Peter Ozsváth and Zoltán Szabó. Holomorphic disks and topological invariants for closed three-manifolds. *Ann. of Math. (2)*, 159(3):1027–1158, 2004.
- [39] Peter Ozsváth and Zoltán Szabó. Lectures on Heegaard Floer homology. In *Floer homology, gauge theory, and low-dimensional topology*, volume 5 of *Clay Math. Proc.*, pages 29–70. Amer. Math. Soc., Providence, RI, 2006.
- [40] Joel Robbin and Dietmar Salamon. The Maslov index for paths. *Topology*, 32(4):827–844, 1993.
- [41] Matthias Schwarz. *Morse homology*, volume 111 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 1993.

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA, GEORGIA
 Email address: etnyre@math.gatech.edu