Heegaard Floer homology
Version: August 20, 2024

Peter S. Ozsváth, András I. Stipsicz and Zoltán Szabó

Chapter 1

Topological preliminaries

In this chapter we collect some basic facts, constructions and examples of three-manifolds, and knots and links in them. In Section 1.1 we discuss handlebodies and Morse functions. A short discussion about three-manifolds is given in Section 1.2, and in Section 1.3 we discuss knots and links in S^3, with special attention to their Alexander polynomials. In Section 1.4 we introduce Euler structures on three-manifolds. These structures will play a fundamental role in Heegaard Floer homology.

1.1. Handle decompositions and Morse theory

We consider smooth, connected, oriented manifolds possibly with boundary. Unless otherwise stated, these manifolds will also be compact.

Handle decompositions. Handle decompositions are convenient ways to present and study smooth manifolds. In the following the n-dimensional closed *disk* $\{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$ will be denoted by D^n; its boundary, the $(n-1)$-dimensional *sphere* is $S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$.

Definition 1.1.1. *Suppose that X is a smooth n-dimensional manifold with boundary ∂X. For $0 \leq k \leq n$, an n-**dimensional** k-**handle** h is a copy of $D^k \times D^{n-k}$, attached to the boundary of X along $(\partial D^k) \times D^{n-k}$ by a smooth embedding $\varphi\colon (\partial D^k) \times D^{n-k} \to \partial X$. The manifold $X_h = X \cup_\varphi h$ is the result of the handle attachment, which is a smooth n-manifold with boundary (after rounding corners).*

The disk $D^k \times \{0\}$ is called the *core* of the handle, $\{0\} \times D^{n-k}$ is the *cocore*, φ is the *attaching map*, $(\partial D^k) \times D^{n-k}$ (or its image $\varphi((\partial D^k) \times D^{n-k})$) is the *attaching region*, $(\partial D^k) \times \{0\}$ (or its image) is the *attaching sphere* and $\{0\} \times (\partial D^{n-k})$ is the *belt sphere* of the handle h. The integer k is the *index* of the handle.

Smoothly isotopic attaching maps result in diffeomorphic manifolds. Moreover, the attaching map $\varphi\colon (\partial D^k) \times D^{n-k} \to \partial X$ is determined up to isotopy by an embedding $\varphi_0\colon (\partial D^k) \times 0 \to \partial X$ together with an identification of the normal bundle of $\varphi_0((\partial D^k) \times 0)$ with $(\partial D^k) \times \mathbb{R}^{n-k}$, i.e. a (normal) *framing*. Since the difference of two framings is a map $S^{k-1} \to GL(n-k)$, and $GL(n-k)$ is homotopy equivalent to $O(n-k)$, the set of framings can be parametrized by homotopy classes of maps from S^{k-1} to $O(n-k)$.

By the above reasoning, in the oriented case, the framing of a 0-, 1-, $(n-1)$- and an n-handle is unique. To specify an n-dimensional 2-handle attachment we need to specify a knot in ∂X, together with a framing. The set of framings in this case is parametrized by $\pi_1(SO(n-2))$, which is the trivial group for $n = 3$, is isomorphic to $\mathbb{Z}$ for $n = 4$, and is $\mathbb{Z}/2\mathbb{Z}$ for $n \geq 5$.

Definition 1.1.2. *Let X be a compact n-dimensional manifold with boundary ∂X, and suppose that the boundary is decomposed as a disjoint union $\partial X = \partial_+ X \cup \partial_- X$ of two compact submanifolds. Suppose that X is oriented, and orient $\partial_\pm X$ so that $\partial X = \partial_+ X \cup -\partial_- X$ in the boundary orientation (where $-\partial_- X$ means the manifold $\partial_- X$ with the reversed orientation). A* ***handle decomposition*** *of X relative to $\partial_- X$ is a diffeomorphism of X with a manifold we get from $[0,1] \times \partial_- X$ by attaching handles to $\{1\} \times \partial_- X$. A manifold X equipped with a handle decomposition is called a* ***relative handlebody****, or if $\partial_- X = \emptyset$, simply a* ***handlebody****.*

Let $\phi_0\colon S^{k-1} \to Y^{n-1}$ be an embedding with trivial normal bundle, together with a framing λ. Remove a neighborhood of $\phi_0(S^{k-1})$, identified via the framing with $S^{k-1} \times D^{n-k}$ and glue back in $D^k \times S^{n-k-1}$ using the framing on the boundary. The resulting manifold $Y_\lambda(\phi_0)$ is called the *λ-framed surgery on Y along ϕ_0*. Clearly, attaching an n-dimensional k-handle to X^n changes its boundary ∂X by such a surgery.

For example, if Y_1 and Y_2 are two oriented m-manifolds equipped with points $y_i \in Y_i$, then the connected sum $Y_1 \# Y_2$ of Y_1 and Y_2 is the result of framed surgery along the embedded 0-sphere $\{y_1, y_2\} \subset Y_1 \cup Y_2$.

Morse functions. Morse functions and related constructions will play a crucial role in our subsequent discussions. We recall here the basics of Morse theory, and refer the reader to [**77**] for a thorough treatment.

Suppose that $f\colon X \to \mathbb{R}$ is a smooth function on X. The point $p \in X$ is a *critical point* of f if df vanishes at p. If p is a critical point then $f(p)$ is called a *critical value*; $c \in \mathbb{R}$ is a *regular value* if $f^{-1}(c)$ does not contain any critical points. The critical point p is *non-degenerate* if in some local coordinate chart around p, the Hessian matrix $(\frac{\partial^2 f(p)}{\partial x_i \partial x_j})$ is non-degenerate.

The Hessian has the following coordinate-free description. Fix a Riemannian metric g on X and consider the map $\mathrm{Hess}_f\colon T_pX \to T_pX$ at a critical point p defined by

$$g(\mathrm{Hess}_f(\xi), \eta) = \xi(\widetilde{\eta} f), \tag{1.1}$$

where $\xi, \eta \in T_pX$ and $\widetilde{\eta}$ is a vector field extending η to a neighbourhood of p. (Here, $\tilde{\eta} f$ denotes the smooth function defined as the directional derivative of f in the direction specified by $\widetilde{\eta}$.)

Note that this definition doesn't depend on an extension $\tilde{\eta}$ of η:

Exercise 1.1.3. (a) *Let p be a critical point of a function f, and let $\xi_p, \eta_p \in T_pX$. If $\widetilde{\xi}$ and $\widetilde{\eta}$ are any two vector fields defined near p that extend ξ_p and η_p respectively, show that $\xi_p(\widetilde{\eta} f) = \eta_p(\widetilde{\xi} f)$. Conclude that this value is independent of the extensions $\widetilde{\xi}$ and $\widetilde{\eta}$.*

(b) *Suppose that $X = \mathbb{R}^n$, and consider the Riemannian metric for which the coordinate vector fields $\{\frac{\partial}{\partial x_i}\}_{i=1}^n$ form an orthonormal basis at each tangent space. Using the identification $T_pX \cong \mathbb{R}^n$ induced by this basis, show that the matrix representing the linear transformation $\mathrm{Hess}_f\colon \mathbb{R}^n \to \mathbb{R}^n$ is the matrix $(\frac{\partial^2 f(p)}{\partial x_i \partial x_j})$.*

Definition 1.1.4. *A smooth function $f\colon X \to \mathbb{R}$ is a **Morse function** if each critical point of f is non-degenerate.*

Morse functions exist by the following result:

Proposition 1.1.5. [**77**, Corollary 6.8] *On a compact, smooth n-manifold X, any continuous function can be approximated by a Morse function.* □

Suppose that p is a non-degenerate critical point of f. The *index* $\lambda(p)$ of p is the maximal dimension of a subspace of T_pX on which the Hessian is negative definite. According to the Morse lemma [**77**], an index-k critical point p of a Morse function f admits a coordinate chart on which

$$f = f(p) - x_1^2 - x_2^2 - \cdots - x_k^2 + x_{k+1}^2 + \cdots + x_n^2.$$

It follows that critical points of Morse functions are isolated.

A Morse function f on a smooth, closed, n-manifold can be used to give a handle decomposition of X. The main idea is to use sublevel and level sets: Let X_t denote $f^{-1}((-\infty, t])$ and $Y_t = f^{-1}(t)$. Since X is compact, X_t is empty for sufficiently large negative t, while $X_t = X$ for large enough positive t. Now we can study X by understanding how X_t changes as t passes through a critical value. Suppose that $t_1 < t_2$ are two points that are both regular values. Then the manifold-with-boundary X_{t_2} can be given by attaching a cobordism to X_{t_1}, and this cobordism depends on the critical points as follows:

Theorem 1.1.6. *([**77**]) Let f be a Morse function on the smooth closed n-dimensional manifold X, and assume that $t_1 < t_2 \in \mathbb{R}$ are two regular values. If $f^{-1}([t_1, t_2])$ contains no critical point, then X_{t_1} is diffeomorphic to X_{t_2}. If $f^{-1}([t_1, t_2])$ contains a unique critical point p of index k then X_{t_2} can be constructed from X_{t_1} by attaching a smooth n-dimensional k-handle. More generally, if $f^{-1}([t_1, t_2])$ contains m critical points of index k, all of which have the same value, then X_{t_2} is constructed from X_{t_1} by attaching m disjoint, smooth, n-dimensional k-handles.* □

The proof of Theorem 1.1.6 (as given in [**77**, Chapter (I.3)]) uses the concept of *gradient vector fields* and *flows*. Here we give an idea of the proof. Choose a metric g on X and consider the gradient vector field $\vec{\nabla} f$, specified by

$$g(\vec{\nabla} f, w) = df(w)$$

for all $w \in TX$. The vector field $-\vec{\nabla} f$ induces the *downward gradient flow*, i.e. a one-parameter family of diffeomorphisms $\{\phi_t \colon X \to X\}$ with $\phi_0 = \mathrm{Id}_X$ and $\frac{d\phi_t}{dt} = (\phi_t)_*(-\vec{\nabla} f)$. Similarly $\vec{\nabla} f$ induces the *upward gradient flow*.

Suppose first that $f^{-1}([t_1, t_2])$ contains no critical points. In this case, starting at a point of Y_{t_2}, we can follow the (downward) gradient flow and end up at a point in Y_{t_1}. Since the process can be reversed using the upward gradient flow, we get a diffeomorphism between Y_{t_1} and Y_{t_2}. Indeed, we get that $f^{-1}([t_1, t_2])$ is diffeomorphic to the product $Y_{t_1} \times [0, 1]$ and hence X_{t_1} is diffeomorphic to X_{t_2}.

When $[t_1, t_2]$ contains a unique critical value corresponding to the critical point p of index k, then there is a subset $S_2 \subset Y_{t_2}$ that flows into the critical point, meaning that the trajectory of the gradient flow converges to the critical point p. Similarly there is a subset $S_1 \subset Y_{t_1}$ that flows to p under the upward gradient flow. The local behavior of the Morse function f at the critical point p shows that S_1 and S_2 are smoothly embedded spheres with dimensions $k-1$ and $n-k-1$ respectively, and the points that flow into p under the upward flow form a disk D_1 of dimension k (with boundary S_1) and the points which flow to p under the downward

flow form a disk D_2 of dimension $n-k$ (with boundary S_2). Indeed, these disks are the core and cocore disks of a k-handle and the spheres S_1 and S_2 are the attaching and belt spheres of the handle. Consequently X_{t_2} is constructed from X_{t_1} by attaching a k-handle. The case of more critical points of index k follows similarly. Consequently, a Morse function induces a handle decomposition for any closed, oriented, smooth manifold X.

The set of points of the manifold which flow to a critical point p under the upward flow is called the *ascending manifold* $A(p)$, while the set of points flowing to p under the downward flow is the *descending manifold* $D(p)$ at p.

By Proposition 1.1.5 and Theorem 1.1.6, any smooth, compact n-manifold admits a handle decomposition. In fact, we can arrange for the handles to be added in increasing order of index, according to the next theorem.

Definition 1.1.7. *The Morse function f on a smooth n-manifold is* ***self-indexing*** *if for all critical points p, $f(p) = \lambda(p)$.*

Theorem 1.1.8. *([**79**]) Let X be a smooth, closed, connected n-dimensional manifold. Then X admits a self-indexing Morse function with a unique local minimum and a unique local maximum.* □

Exercise 1.1.9. **(a)** *Embed the torus in $\mathbb{R}^3$ so that the x coordinate is a self-indexing Morse function.*

(b) *Consider the sphere $S^{2n+1} \subset \mathbb{C}^{n+1}$ consisting of $(z_0, \ldots, z_n)$ with*

$$\sum_{i=0}^{n} |z_i|^2 = 1.$$

The quotient of this sphere by the natural S^1-action gives $\mathbb{CP}^n$. Show that the function $\sum_{j=1}^{n} j|z_j|^2$ is a self-indexing Morse function on $\mathbb{CP}^n$.

(c) *Show that a closed, smooth n-manifold admits a CW decomposition.*

1.2. Three-manifolds

We give now some basic examples and constructions of three-manifolds. According to [**83**] any topological three-manifold Y admits a unique (up to diffeomorphism) smooth structure, and any homeomorphism between topological three-manifolds can be isotoped to a diffeomorphism; hence without loss of generality we can assume that Y is a smooth three-manifold. Unless otherwise stated, in the following we will always assume that Y is compact and oriented with possibly non-empty boundary.

When studying three-manifolds, it will be convenient to consider their submanifolds. An m-component *link* L in Y is a collection of m disjoint

smoothly embedded simple closed curves. A 1-component link K is called a *knot*. The links L_1, L_2 are *equivalent* if there is an ambient isotopy taking L_1 to L_2; i.e. a map $F\colon Y \times [0,1] \to Y$ with the properties that $F_t = F|_{Y\times\{t\}}$ is a diffeomorphism for all t, $F_0 = \mathrm{Id}_Y$, and $F_1(L_1) = L_2$. The equivalence class of a link is called the *link type*. An *oriented link*, denoted $\vec{L}$, is a link equipped with orientations on each component. The above notion of isotopy adapts readily to the oriented context. For the oriented knot $\vec{K}$ the same knot with the reversed orientation is called the *reverse* of $\vec{K}$ and is denoted by $r(\vec{K})$. There are oriented knots that are not (oriented) isotopic to their reverses; a family of such examples was found by Trotter [**137**]. For example, three-stranded pretzel knots $P(p,q,r)$ with p,q,r odd, all distinct and greater than 1 in absolute value have this property.

1.2.1. Some basic three-manifolds. A fundamental three-dimensional closed manifold is the three-dimensional sphere $S^3 = \{x \in \mathbb{R}^4 \mid ||x|| = 1\}$.

Other simple examples of closed, oriented three-manifolds are given by products of lower dimensional manifolds: $S^1 \times S^2$, the *three-dimensional torus* $T^3 = S^1 \times S^1 \times S^1$, and $S^1 \times \Sigma_g$, where Σ_g denotes an oriented, closed surface of genus g. More generally we can consider S^1-bundles over an orientable surface Σ_g: such a three-manifold $Y_{g,k}$ is classified by the genus g of the base surface and by the first Chern number k of the complex line bundle associated to the S^1-bundle.

In a different direction, the product $S^1 \times \Sigma_g$ can be generalized by considering surface bundles over S^1. Such a bundle can be constructed from an orientation-preserving diffeomorphism $\phi\colon \Sigma_g \to \Sigma_g$, by taking the product $[0,1] \times \Sigma_g$ and identifying $(1,x)$ with $(0,\phi(x))$, for all $x \in \Sigma_g$. The resulting three-manifold is called the *mapping torus* of ϕ. In turn, every surface bundle over S^1 can be constructed in this way; the isotopy class of ϕ is called the *monodromy* of the fibration. The *mapping class group* in genus g, denoted Γ_g, is the group of orientation-preserving self-diffeomorphisms of Σ_g, modulo isotopy; and multiplication corresponds to composition of maps. Indeed, the monodromy of a surface bundle depends on the chosen identification of the fiber with Σ_g, hence the monodromy is well-defined only up to conjugation in the group Γ_g.

Another family of three-manifolds is given as follows. Fix a pair of relatively prime integers p and q, with $1 \le q < p$. Think of S^3 as the subset of $\mathbb{C}^2$ specified by $\{(z_1,z_2) \in \mathbb{C}^2 \mid |z_1|^2 + |z_2|^2 = 1\}$. The map

$$(z_1, z_2) \mapsto (e^{\frac{2\pi i}{p}} z_1, e^{\frac{2\pi i q}{p}} z_2) \tag{1.2}$$

generates an action of $\mathbb{Z}/p\mathbb{Z}$ on S^3. The quotient of S^3 by this action is the *lens space* $L(p,q)$.

Exercise 1.2.1. *Show that all non-trivial circle bundles over S^2 are lens spaces. Which lens spaces can be realized in this manner?*

Lens spaces and circle bundles over surfaces have the following common generalization. A *Seifert fibered three-manifold* is a closed three-manifold Y together with a decomposition into a disjoint union of circles, called *fibers*, each of which has a closed tubular neighbourhood U with the following structure. Either U decomposes as $U = D^2 \times S^1$, so that the fibers correspond to the circles $x \times S^1$ (these fibers are called *regular fibers*); or U has the following structure. Fix relatively prime, positive integers p and q with $p > 1$, and consider the product $D^2 \times S^1$, equipped with the free $\mathbb{Z}/p\mathbb{Z}$ action $(w, z) \mapsto (\eta w, \eta^q z)$, where $\eta = e^{2\pi i/p}$. The quotient space $V_{p,q} = (D^2 \times S^1)/(\mathbb{Z}/p\mathbb{Z})$ is a manifold (also diffeomorphic to $D^2 \times S^1$), and it has an action by S^1, rotating on the second factor. The local neighborhood U is identified with $V_{p,q}$ so that the fibers in U correspond to the S^1-orbits in $V_{p,q}$. The fiber corresponding to $w = 0$ is called a *singular fiber*, and the pair (p, q) is called its *Seifert invariant* of the singular fiber. Seifert fibered spaces are classified; see [**97**].

Example 1.2.2. *Any circle bundle over a two-manifold can be viewed as a Seifert fibered space with no singular fibers. Seifert fibered spaces over the sphere S^2 with at most two singular fibers are either lens spaces or $S^1 \times S^2$; in fact, all lens spaces can be described as Seifert fibered spaces with one singular fiber.*

Example 1.2.3. *If Y is any three-manifold with a circle action, with the property that each orbit has only finite stabilizers, then Y is a Seifert fibered space. Indeed, by* [**97**, Theorem 2, page 88] *an oriented Seifert fibered three-manifold admits such a cirle action once the space of fibers (a two-dimensional orbifold) is orientable.*

Exercise 1.2.4. **(a)** *For p, q relatively prime, give a diffeomorphism between $V_{p,q}$ and $D^2 \times S^1$.*

(b) *Let $\phi\colon T^2 \to T^2$ be the map $\phi(x, y) = (-x, -y)$, where we view $(x, y) \in (\mathbb{R} \oplus \mathbb{R})/(\mathbb{Z} \oplus \mathbb{Z})$. Show that the mapping torus of ϕ is a Seifert fibered space. How many singular fibers does it have, and what are their Seifert invariants?*

(c) *Let $\phi\colon \Sigma_g \to \Sigma_g$ be an orientation preserving map with $\phi^m = \mathrm{Id}_{\Sigma_g}$ for some values of m and g. Construct a circle action on the mapping torus Y of ϕ, and conclude that Y is a Seifert fibered space. The map ϕ induces a $\mathbb{Z}/m\mathbb{Z}$ action on Σ_g. Give a correspondence between the orbits of this action and the fibers of the Seifert fibration. When does an orbit correspond to a singular fiber?*

Interesting examples of Seifert fibered spaces occur when considering the complex algebraic equation

$$z_1^p + z_2^q + z_3^r = 0 \tag{1.3}$$

in $\mathbb{C}^3$, for positive integers p, q, and r. Equation (1.3) then defines a (complex) codimension-one subset $X_{p,q,r} \subset \mathbb{C}^3$, which is a smooth four-manifold away from the origin. If S_ϵ denotes the (five-dimensional) sphere of radius $\epsilon > 0$ in $\mathbb{C}^3$, then the intersection $X_{p,q,r} \cap S_\epsilon$ is a smooth three-manifold.

Exercise 1.2.5. *Prove that $X_{p,q,r}$ is smooth away from the origin. Show that the diffeomorphism type of $X_{p,q,r} \cap S_\epsilon$ is independent of the choice of $\epsilon > 0$.*

The intersection $X_{p,q,r} \cap S_\epsilon$ is called a *Brieskorn manifold*, and it is denoted $\Sigma(p,q,r)$. Brieskorn manifolds naturally inherit the structure of a Seifert fibered space. When p, q, r are pairwise relatively prime integers, the Brieskorn manifold is called a *Brieskorn sphere*, because in this case $H_1(\Sigma(p,q,r);\mathbb{Z}) = 0 = H_1(S^3;\mathbb{Z})$.

Exercise 1.2.6. **(a)** *Consider the Brieskorn sphere $\Sigma(p,q,r)$, and let the subspaces Γ_i for $i = 1,2,3$ be the intersections of $\Sigma(p,q,r)$ with the hyperplanes $\{z_i = 0\}$. Show that Γ_i are smoothly embedded circles inside $\Sigma(p,q,r)$.*

(b) *Construct a circle action on $\Sigma(p,q,r)$ that is free away from the three circles Γ_1, Γ_2, and Γ_3, each of which is an orbit with finite stabilizers. Conclude that a Brieskorn sphere is a Seifert fibered three-manifold with three singular fibers. What are the Seifert invariants of the three singular fibers?*

(c) *Show that if p, q, r are pairwise relatively prime, then, as stated above, $H_1(\Sigma(p,q,r);\mathbb{Z}) = 0$.*

(d) *Show that if q and r are relatively prime, then $\Sigma(1,q,r)$ is diffeomorphic to S^3.*

For a similar, lower-dimensional example, consider the complex algebraic equation $z_1^p + z_2^q = 0$ in $\mathbb{C}^2$, and intersect the resulting (complex) curve with the unit sphere S^3. When p and q are relatively prime, this gives a knot in S^3 called the *torus knot* $T_{p,q}$; when $\gcd(p,q) = k$, we obtain the k-component *torus link* $\vec{T}_{p,q}$, oriented as the boundary of the complex curve $z_1^p + z_2^q = 0$ in D^4.

Exercise 1.2.7. *Show that $\Sigma(2,3,5)\backslash\Gamma_3$ is diffeomorphic to the complement of the left-handed trefoil knot $T = T_{2,-3}$ in S^3.*

The following terminology will be used throughout:

Definition 1.2.8. *A **homology three-sphere** is a closed oriented connected three-manifold Y for which $H_*(Y;\mathbb{Z}) \cong H_*(S^3;\mathbb{Z})$. Similarly a **rational homology three-sphere** is a closed oriented connected three-manifold Y for which $H_*(Y;\mathbb{Q}) \cong H_*(S^3;\mathbb{Q})$.*

It follows from Poincare duality that a closed, oriented, connected, three-dimensional manifold Y is a homology three-sphere if and only if $H_1(Y,\mathbb{Z}) = 0$, and it is a rational homology three-sphere if and only if its first Betti number $b_1(Y)$ vanishes. Note that all the lens spaces are rational homology spheres, while Brieskorn spheres are examples for homology three-spheres.

1.2.2. Branched covers of three-manifolds. New three-manifolds can be constructed out of old ones, using the following notion:

Definition 1.2.9. *The smooth map $f\colon X \to Y$ between compact n-dimensional manifolds is a **smooth branched covering** if there exists an $(n-2)$-dimensional submanifold $B_f \subset Y$, called the **branching locus** of f, such that the restriction of f to $X \setminus f^{-1}(B_f)$ is an ordinary covering of some finite degree d, and for each $p \in f^{-1}(B_f)$ there is a positive integer m, a coordinate chart $U_p \cong \mathbb{C}\times\mathbb{R}^{n-2}$ around $p \in X$, and a coordinate chart $V_p \cong \mathbb{C}\times\mathbb{R}^{n-2}$ around $f(p) \in Y$, with respect to which f is modeled on the map $\mathbb{C}\times\mathbb{R}^{n-2} \to \mathbb{C}\times\mathbb{R}^{n-2}$ given by $(z,x) \mapsto (z^m,x)$; i.e. we have the following commutative diagram:*

$$\begin{array}{ccc} p \in U_p & \xrightarrow{\phi_p} & \mathbb{C}\times\mathbb{R}^{n-2} \\ \Big\downarrow f & & \Big\downarrow (z,x)\mapsto(z^m,x) \\ f(p) \in V_p & \xrightarrow{\psi_p} & \mathbb{C}\times\mathbb{R}^{n-2} \end{array}$$

*The integer m is called the **branching index** at p.*

Given a link $L \subset Y$ and a homomorphism ϕ from $\pi_1(Y\setminus L)$ to the symmetric group on d letters, there is an associated branched d-fold cover of Y, whose branch locus is L. The homomorphism ϕ specifies the d-sheeted covering space away from the branch locus.

If $Y = S^3$ and $d = 2$, there is a canonical representation of $S^3 \setminus L$ to $\mathbb{Z}/2\mathbb{Z}$ taking each meridian to the non-trivial element. Thus, in this case the branch locus alone specifies the branched cover, which we denote $\Sigma(L)$. More generally, consider an oriented link $\vec{L}$ equipped with the canonical map $\pi_1(S^3\setminus L) \to \mathbb{Z}/d\mathbb{Z}$, taking each oriented meridian to the preferred generator

of $\mathbb{Z}/d\mathbb{Z}$. Thinking of $\mathbb{Z}/d\mathbb{Z}$ as a subgroup of the symmetric group on d letters, we have specified a d-fold branched cover of S^3 along L, called the *cyclic branched d-fold cover* of S^3 along $\vec{L}$.

Exercise 1.2.10. *Verify that the Brieskorn manifold $\Sigma(p,q,r)$ is the p-fold cyclic branched cover of S^3 whose branch locus is the torus link $T_{q,r}$.*

Not every three-manifold can be presented as a cyclic branched cover of S^3. On the other hand, by a theorem of Montesinos [**84**] and Hilden [**49**] every closed, oriented three-manifold can be given as the triple branched cover of S^3 along a knot. Interestingly, there are knots $K \subset S^3$ (called *universal knots*) for which every three-manifold can be presented as some branched cover of S^3 along K. Indeed, the figure-8 knot is such a universal knot [**50**].

1.2.3. Surgeries. A further useful construction of three-manifolds is given by *Dehn filling*. To define it, let M be a three-manifold with torus boundary and fix a homologically non-trivial, embedded curve $\gamma \subset \partial M$. Attaching a three-dimensional 2-handle to M along γ, and a 3-handle to the resulting S^2 (or equivalently, gluing a copy of $D^2 \times S^1$ to M along ∂M so that $(\partial D^2) \times \{p\}$ glues to γ), we obtain a closed three-manifold M_γ called the *Dehn filling* of M along γ.

Let K be a knot in the three-manifold Y. The complement of an open tubular neighborhood of K in Y is a three-manifold M with $\partial M \cong T^2$. A framing ϕ on K can be thought of as a normal vector field to K. The normal push-off of K gives a closed curve γ in ∂M, which is uniquely determined by the framing, up to isotopy. The three-manifold $Y_\gamma(K)$ is the Dehn filling of $M = Y \setminus \nu(K)$ along this curve. Moreover, an orientation on K specifies an identification $f\colon \mathbb{Z} \oplus \mathbb{Z} \xrightarrow{\cong} H_1(\partial M;\mathbb{Z})$ as follows: Let $\lambda \subset \partial M$ be a push-off of K specified by the framing, and let $\mu \subset \partial M$ be the boundary of a normal disk to K, oriented so that oriented intersection number $\mu \cdot \lambda = -1$ in ∂M. (The intersection number is computed using the orientation on $T^2 = \partial M$, which is opposite to the one it inherits as the boundary of a closed tubular neighborhood of K.) When p and q are relatively prime integers, the homology class $p\mu + q\lambda$ can be represented by an embedded, closed, connected curve $\gamma_{p,q}$ on ∂M, and the Dehn filling of M along this curve is called the *p/q Dehn surgery* along the framed knot, and is denoted by $Y_{p/q}(K,\phi)$. This notation is justified by the observation that $Y_{p/q}(K,\phi) = Y_{-p/-q}(K,\phi)$. When $q = \pm 1$ this operation is called an *integral Dehn-surgery*. Another special case is when $q = 0$: since in this case $f(1,0) = \mu$, we have that $Y_{1/0}(K,\phi) = Y_\infty(K,\phi) = Y$.

Exercise 1.2.11. *Let X be a four-manifold with boundary $\partial X = Y$. Show that if we add a four-dimensional 2-handle to X then the new boundary is*

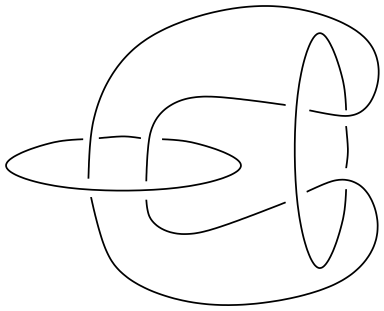


Figure 1.1. The Borromean rings.

given by an ***integral*** *Dehn-surgery on Y, where the knot $K \subset Y$ is given by the attaching sphere of the 2-handle.*

Definition 1.2.12. *When Y is an integral homology sphere, there is a canonical framing for K, called the* Seifert framing*, where λ generates the kernel of the map*

$$i_*\colon H_1(\partial(Y \setminus \nu(K)); \mathbb{Z}) \longrightarrow H_1(Y \setminus \nu(K); \mathbb{Z}) \cong \mathbb{Z}.$$

The name comes from the fact that a curve representing λ can be chosen to lie on a Seifert surface of K.

Since $H_1(Y \setminus \nu(K); \mathbb{Z}) = \mathbb{Z}$ is generated by $i_*(\mu)$, it follows from the Mayer-Vietoris sequence that $H_1(Y_0(K); \mathbb{Z}) \cong \mathbb{Z}$ and $H_1(Y_{p/q}(K); \mathbb{Z}) \cong \mathbb{Z}/p\mathbb{Z}$ for $p \neq 0$.

Dehn surgery naturally generalizes from framed knots to framed links, where we fix a framing for each component of the link. Whereas not all closed, oriented three-manifolds can be realized as Dehn surgery on a knot, they can be realized as Dehn surgeries on links:

Theorem 1.2.13. *(Lickorish, Wallace* [**67, 144**]*) Every connected, closed, oriented three-manifold can be obtained by surgery along a link in S^3, where all surgery coefficients can be assumed to be integers.* □

Exercise 1.2.14. (a) *Let $\mathcal{U}$ denote the unknot in S^3. Show that $S^3_1(\mathcal{U})$ is diffeomorphic to S^3, and $S^3_0(\mathcal{U})$ is diffeomorphic to $S^1 \times S^2$.*

(b) *Given non-zero, relatively prime integers p and q, show that $S^3_{p/q}(\mathcal{U}) \cong -L(p,q)$.*

(c) *Show that 0-surgery on each component of the Borromean rings (given in Figure 1.1) provides the three-dimensional torus T^3.*

(d) *Show that (-1)-surgery along the right-handed trefoil $T_{2,3}$ is diffeomorphic to $\Sigma(2,3,7)$.*

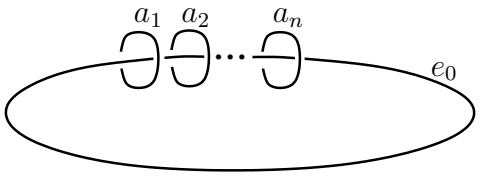


Figure 1.2. Surgery diagram for Seifert fibered three-manifolds. We assume that $e_0 \in \mathbb{Z}$ and $a_1, \dots, a_n \in \mathbb{Q}$.

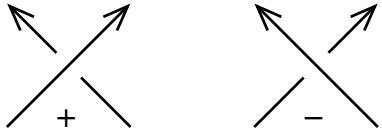


Figure 1.3. Signs of a crossing. The oriented arcs on the left show a positive crossing, while on the right a negative crossing.

(e) *Show that* $S^3_{-1}(T)$ *with* $T = T_{2,-3}$ *(the left-handed trefoil knot) is diffeomorphic to* $\Sigma(2,3,5)$.

(f) *Verify that the three-manifolds* Y *described by surgery diagrams of the form shown in Figure 1.2 are all Seifert fibered three-manifolds.*

(g) *Show that* T^3 *cannot be constructed from* S^3 *by performing Dehn surgery along a knot.*

1.3. Knots and links in S^3

In this section we turn to the study of knots and links in S^3, concentrating on those aspects of this rich theory that are relevant to our subsequent discussions. For a more thorough treatment, the reader is referred to [**14, 68, 119**].

Let L be a link $\mathbb{R}^3$ and consider the orthogonal projection to the (x, y) plane. By isotopying L into general position we can arrange for the projection restricted to L to be an immersion with finitely many double points. The isotopy class of the link can be reconstructed from the immersed curve, equipped with the local information at each double point, specifying which of the two strands at the crossing is higher. This local information is pictorially represented by interrupting the lower strand at the crossing. We will call the immersed curve in the plane, together with this local information, a *link projection* (or *knot projection*, if the link is a knot). An orientation of the link also orients its projection, and signs can be associated to the crossings, following the conventions of Figure 1.3.

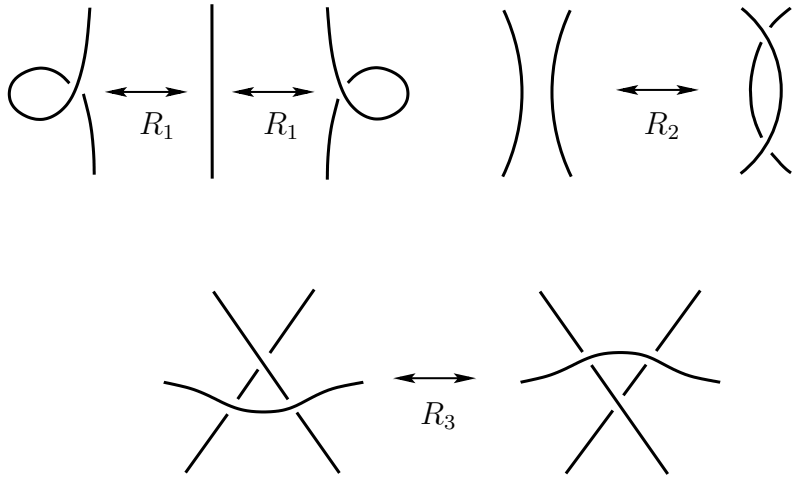


Figure 1.4. The Reidemeister moves R_1, R_2, R_3**.**

The local modifications of a link projection shown in Figure 1.4 are called the *Reidemeister moves* R_1, R_2 and R_3. (The figures indicate changes within a small disk, while the rest of the projection remains unchanged.)

Theorem 1.3.1. [**116**] *The link projections* P_1 *and* P_2 *correspond to isotopic links if and only if these projections can be connected by a finite sequence of Reidemeister moves, their inverses, and orientation-preserving diffeomorphisms of the plane.* □

It follows that any quantity associated to a link projections that is invariant under the three Reidemeister moves is in fact a link invariant.

The Alexander polynomial. The *Alexander polynomial of a knot* $K \subset S^3$ is an integral Laurent polynomial $\Delta_K(t) \in \mathbb{Z}[t, t^{-1}]$, depending only on the knot type of K.

One definition of the Alexander polynomial [**16**] can be given as follows. Let $X = S^3 \setminus K$. Since $H_1(X; \mathbb{Z}) \cong \mathbb{Z}$, there is a naturally associated normal covering $\widetilde{X} \to X$, whose automorphism group is identified with $\mathbb{Z}$. Thus, the homology $H_1(\widetilde{X}; \mathbb{Z}) = M$ inherits an action by the group-ring of $\mathbb{Z}$, which can be thought of as the ring of Laurent polynomials in one variable t: $\mathbb{Z}[\mathbb{Z}] \cong \mathbb{Z}[t, t^{-1}]$. There is a presentation of the $\mathbb{Z}[t, t^{-1}]$-module M as

$$R^m \xrightarrow{A} R^n \to M \to 0,$$

where A is an $m \times n$ matrix over $R = \mathbb{Z}[t, t^{-1}]$. The greatest common divisor of the determinants of the $(n-1) \times (n-1)$ minors of A is, by definition, the Alexander polynomial $\Delta_K(t)$, up to a multiple of $\pm t^k$. To

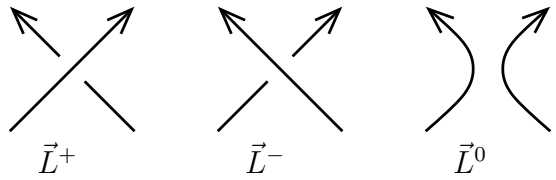


Figure 1.5. The diagrams in the small disk for the skein relation.

pin down this indeterminacy, we require further that $\Delta_K(t) = \Delta_K(t^{-1})$ (to remove the indeterminacy in the t-power) and $\Delta_K(1) = 1$ (to pin down the sign). The fact that Δ_K has a representative with these properties can be seen as a consequence of Poincaré duality [**119**, Corollary 7,p. 207].

For example, for the unknot $\mathcal{U}$ we have $X = S^1 \times D^2$, $\widetilde{X} = \mathbb{R} \times D^2$, hence $M = 0$ and $A = \mathrm{Id}$. This implies that $\Delta_{\mathcal{U}}(t) = 1$.

The Alexander polynomial has a natural generalization for oriented links $\vec{L}$, denoted $\Delta_{\vec{L}}(t) \in \mathbb{Z}[t^{1/2}, t^{-1/2}]$. This generalization satisfies a *skein relation* which we recall presently. Let $\vec{L}^+$, $\vec{L}^-$, and $\vec{L}^0$ be three oriented links that differ in a single crossing; more precisely, they admit oriented link projections P^+, P^-, and P^0 that are identical outside of a small disk, in which they are as shown in Figure 1.5. Then, the Alexander polynomials of these three links are related by the *skein relation*

$$\Delta_{\vec{L}^+}(t) - \Delta_{\vec{L}^-}(t) = (t^{\frac{1}{2}} - t^{-\frac{1}{2}})\Delta_{\vec{L}^0}(t). \tag{1.4}$$

Exercise 1.3.2. *Show that if $\vec{L}$ is a split link (that is, it admits a disconnected projection), then $\Delta_{\vec{L}}(t) = 0$.*

Conway has shown that the skein relation gives an algorithm for computing the Alexander polynomial of an arbitrary link, bearing in mind that $\Delta_{\mathcal{U}}(t) = 1$ for the unknot $\mathcal{U}$.

Recall that any knot $K \subset S^3$ bounds an oriented, connected, compact surface $\Sigma \subset S^3$; such a surface is called a *Seifert surface* for K, see [**68**]. The minimal genus of a Seifert surface for K is the *Seifert genus* of K, denoted by $g_3(K)$.

A Seifert surface allows us to define the Alexander polynomial in more geometric terms: for a Seifert surface Σ of K we define the Seifert form $S(x, y)$ on $H_1(\Sigma; \mathbb{Z})$ by taking the linking number of a geometric representative of x and of the positive normal push-off y^+ of $y \in H_1(\Sigma; \mathbb{Z})$ (where in the push-off we use the orientation of Σ to determine the positive normal direction). By fixing a basis of $H_1(\Sigma; \mathbb{Z})$, S can be described by a matrix, which we also denote by S. Then the symmetrized Alexander polynomial can be

given as

$$\Delta_K(t) = \det(t^{-\frac{1}{2}}S - t^{\frac{1}{2}}S^T).$$

We say that the knot K is *fibered* if there is a locally trivial fibration map $\varphi\colon S^3 \setminus K \to S^1$ so that the closure of a fiber of φ is a Seifert surface of K. For example, the torus knots $T_{p,q}$ are fibered knots.

The main properties of the Alexander polynomial relevant in our future discussions are summarized:

Theorem 1.3.3. *Suppose that $K \subset S^3$ is a knot with Alexander polynomial $\Delta_K(t)$. Then*

(A-1) $\Delta_K(t) = \Delta_K(t^{-1})$, *that is, Δ_K is symmetric, hence can be written as*

$$\Delta_K(t) = a_0 + \sum_{i=1}^{d} a_i(t^i + t^{-i})$$

with $a_d \neq 0$. This d is called the degree of Δ_K.

(A-2) For the degree d of Δ_K we have $d \leq g_3(K)$, where $g_3(K)$ is the Seifert genus of K.

(A-3) $\Delta_K(1) = 1$.

(A-4) If K is a fibered knot, then the leading coefficient a_d of $\Delta_K(t)$ is equal to ± 1.

Properties (A-1)-(A-3) are proved in [**68**, Chapter 6]; Property (A-4) is in [**119**, Chapter 10]

We note that there is a class of knots for which the Alexander polynomial carries precise topological information. A knot is called *alternating* if it admits a diagram in which the crossing alternate as over- and under-crossings as we traverse through the knot. (See [**40, 55**] for a more intrinsic characterization of these knots.)

Theorem 1.3.4. *Suppose that K is an alternating knot. Then the degree d of $\Delta_K(t)$ is equal to the Seifert genus $g_3(K)$ of K. In particular, if K is alternating and $\Delta_K(t) \equiv 1$ then K is the unknot. The leading coefficient a_d of $\Delta_K(t)$ for an alternating knot K is equal to ± 1 if and only if K is fibered.* □

We do not prove the above result here, but refer the reader to the literature. The Alexander polynomial of an alternating knot determines the genus by a theorem of Crowell [**15**] and Murasugi [**89**]. The last statement is another theorem of Murasugi [**90**].

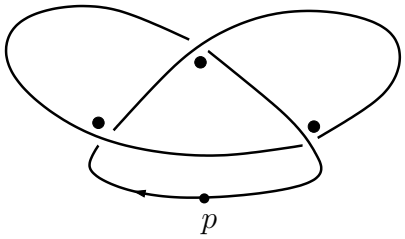


Figure 1.6. A Kauffman state of a decorated projection of the left-handed trefoil knot. The Kauffman state is indicated by a dot placed in the chosen quadrant at each crossing. The arrow indicates an orientation on the knot.

An explicit expression for the Alexander polynomial can be given in terms of a projection P for the oriented link $\vec{L}$, equipped with the following additional choice. Distinguish an edge in the projection P by marking it with a point p. The projection, together with this choice of edge is called a *decorated link projection.*

Let $\mathcal{C}(P)$ denote the set of crossings in the projection and let $Dom(P)$ denote the set of domains in the plane (i.e. the connected components of the complement of P) which do not contain the marking p on their boundary.

Exercise 1.3.5. (a) *Show that for a knot K the cardinality $|\mathcal{C}(P)|$ is equal to $|Dom(P)|$.*

(b) *Show that for a disconnected projection of a split link L we have $|\mathcal{C}(P)| \neq |Dom(P)|$.*

Definition 1.3.6. *A **Kauffman state** is a map that associates to each crossing in $\mathcal{C}(P)$ one of the four quadrants around that crossing, so that the induced map $\sigma\colon \mathcal{C}(P) \to Dom(P)$ is a bijection. The set of Kauffman states in a decorated link projection P will be denoted by $\mathcal{K}(P)$.*

When illustrating Kauffman states, we mark the quadrant associated to the crossing in the diagram, as shown in Figure 1.6.

Exercise 1.3.7. (a) *Find all Kauffman states in the decorated projection of Figure 1.6.*

(b) *Show that if K admits a projection with a single Kauffman state, then K is the unknot.*

(c) *Show that a split link admits a decorated projection with no Kauffman states.*

The complement of the projection of a knot in the plane can be colored with two colors (say black and white) so that each component of the complement

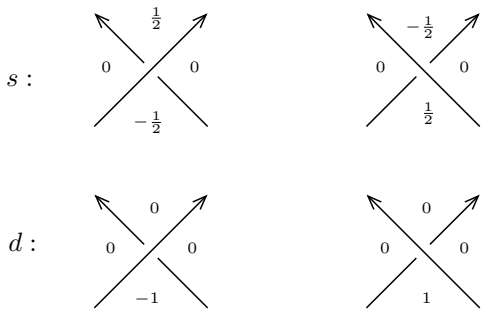


Figure 1.7. Local coefficients s and d at a crossing c_i.

is colored by one of the colors, and components with the same color do not share edge. Such a coloring is called a *chessboard coloring*. Fix a chessboard coloring for the projection P of the knot K, and define a planar graph called the *black graph* Γ_B as follows: the vertex set of the graph is the set of black domains, and two vertices v_1, v_2 are connected by an edge for each crossing of the projection which is in the closure of both components D_1, D_2 corresponding to v_1 and v_2. In the same manner we can define the *white graph* Γ_W of P, which is simply the planar dual of the black graph.

Exercise 1.3.8. (a) *Verify the existence of a chessboard coloring of the diagram D.*

(b) *Construct a bijection between the set of Kauffman states of the decorated projection (P, p) of a knot K and the set of spanning trees of the black graph of a chessboard coloring of P.*

(c) *Conclude that a projection P of a knot has the same number of Kauffman states independent from the choice of the decoration p.*

Two quantities can be associated to a Kauffman state κ of P:

$$s(\kappa) = \sum_{c_i \in \mathcal{C}(P)} s(\kappa(c_i)) \qquad d(\kappa) = \sum_{c_i \in \mathcal{C}(P)} d(\kappa(c_i)), \tag{1.5}$$

where the local coefficients $s(\kappa(c_i)) \in \{0, \pm\frac{1}{2}\}$ and $d(\kappa(c_i)) \in \{0, \pm 1\}$ for $\kappa \in \mathcal{K}(P)$ at a crossing $c_i \in \mathcal{C}(P)$ are shown in Figure 1.7: we take $s(\kappa(c_i))$ and $d(\kappa(c_i))$ to be the value in the quadrant selected by κ at c_i.

The Alexander polynomial for $\vec{L}$ can be expressed in terms of Kauffman states, as follows:

Proposition 1.3.9. *Let $\vec{L}$ be an oriented link, and consider a decorated projection (P, p) for $\vec{L}$. Then, the Alexander polynomial is computed by the*

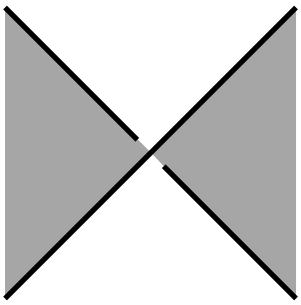


Figure 1.8. Coloring conventions for an alternating diagram.

expression

$$\sum_{\kappa\in\mathcal{K}(P)} (-1)^{d(\kappa)} t^{s(\kappa)} \in \mathbb{Z}[t^{-\frac{1}{2}}, t^{\frac{1}{2}}]. \tag{1.6}$$

Note that in the formula only the mod 2 value of $d(\kappa)$ matters; we keep the present definition for future reference.

Exercise 1.3.10. (a) *Compute the Alexander polynomial of the torus knot $T_{2,2n+1}$ by enumerating all its Kauffman states in an appropriate decorated projection.*

(b) *The determinant* $\det(K)$ *of the knot K is defined as $|\Delta_K(-1)|$. Show that for an alternating diagram the number of Kauffman states is equal to the determinant of the knot.*

Consider the difference $\delta(\kappa) = s(\kappa) - d(\kappa)$ for a Kauffman state κ. As we shall soon see (Theorem 1.3.13 below), for an alternating, δ is independent of the choice of κ. Indeed, we shall identify this constant with another classical knot invariant, the *signature* $\sigma(K)$. The signature in turn can be defined in terms of a Seifert surface Σ of K through the Seifert form associated to a basis of $H_1(\Sigma;\mathbb{Z})$; see for example [**107**, Section 2.3] for a definition. There is a formula, the Gordon-Litherland formula, which gives an explicit description of the signature in terms of a decorated knot projection.

We state without proof a special case of the Gordon-Litherland formula for alternating projections. (The reader interested in a proof is referred to [**37**]; see also [**107**, Corollary 2.7.11].) Let $\mathrm{Pos}(P)$ denote the set of positive crossings in the projection P and $\mathrm{Black}(P)$ denote the set of black regions. Note that any alternating projection can be colored so that at each quadrant, the coloring has the form shown in Figure 1.8.

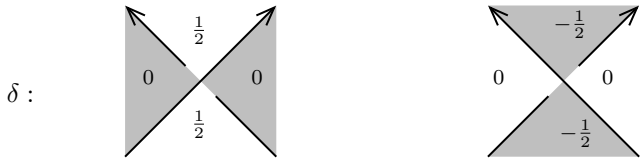


Figure 1.9. Computing δ for alternating diagrams.

Theorem 1.3.11 (Gordon-Litherland [**37**])**.** *Let K be a knot with an alternating projection P, given the checkerboard coloring as indicated in Figure 1.8. Then,*

$$\sigma(K) = \#\mathrm{Pos}(P) + \#\mathrm{Black}(P) - 1.$$

□

Exercise 1.3.12. *Use the Gordon-Litherland formula to compute the signatures of the left- and right-handed torus knots $T_{2,\pm 3}$.*

Theorem 1.3.13. *Suppose that (P, p) is an alternating marked diagram of the knot K. Then there is a constant $C_{P,p}$ for which $\delta(\kappa) = \frac{\sigma(K)}{2}$. for all Kauffman states of (P, p).*

Proof. For an alternating link, the chessboard coloring can be chosen so that at each crossing, it has the form from Figure 1.8.

Consider the following further local quantity b associated to a Kauffman state at a crossing. Let $b_c(\kappa) = 1$ if the Kauffman state lies in one of the two black quadrants and $b_c(\kappa) = 0$ otherwise. Moreover, let $\epsilon_c = \pm 1$ be the local sign at the crossing c, following conventions from Figure 1.3.

Comparing the local contributions to $\delta = s - d$ at each crossing c with with the coloring conventions, as shown in Figure 1.9, we see that $\delta_c = \frac{(1+\epsilon-b)}{2}$.

For any Kauffman state κ, all but one of the black regions in the knot projection contain (exactly) one component in κ; so

$$\sum_c b_c(\kappa) = \#(\mathrm{Black}(P) - 1).$$

It follows that for any Kauffman state κ, the quantity $\delta(\kappa)$ is given by

$$\delta(\kappa) = \frac{1}{2}\#\mathrm{Pos}(P) + \#\mathrm{Black}(P) - 1,$$

which is a constant independent of the choice of Kauffman state κ. The Gordon-Litherland formula (Theorem 1.3.11) identifies this constant with the signature of K. □

Theorem 1.3.13 has the following immediate consequence for alternating knots:

Corollary 1.3.14 (Cromwell [**15**]; Murasugi [**91**])**.** *For an alternating knot K, the non-zero coefficients of the Alexander polynomial alternate in sign.* □

In a smooth four-manifold X there are two natural notions of embedded surfaces S: those which are covered by open neighborhoods $U \subset X$ so that the pair $(U, U \cap S)$ is homeomorphic to the local model $(\mathbb{R}^4, 0 \times 0 \times \mathbb{R}^2)$, the *local flatly embedded surfaces*; and those for which $(U, U \cap S)$ is diffeomorphic to the local model, *smoothly embedded surfaces.*

Definition 1.3.15. *The knot $K \subset S^3$ is said to be* ***smoothly slice*** *resp.* ***topologically slice*** *if there is a smooth resp. locally flat embedding*

$$f\colon (D^2, \partial D^2) \to (D^4, K \subset S^3 = \partial D^4).$$

According to the Fox-Milnor condition, the Alexander polynomial provides an obstruction for sliceness:

Theorem 1.3.16 (Fox-Milnor [**24**])**.** *Suppose that $K \subset S^3$ is a topologically slice knot. Then there is a polynomial $f \in \mathbb{Z}[t]$ such that*

$$\Delta_K(t) = f(t) \cdot f(t^{-1}).$$

Exercise 1.3.17. *Show that the Figure-8 and the trefoil knots (the knots 4_1 and 3_1 in Rolfsen's table* [**119**]*) are not topologically slice.*

Obviously, every smoothly slice knot is topologically slice. The fact that there are topologically slice knots that are not smoothly slice is one of the first manifestations of the subtlety of differential four-manifold topology in knot theory.

1.4. Euler structures on three-manifolds

Spin^c structures will play an important role in our subsequent discussions and constructions. These structures can be defined in any dimension, but the three-dimensional case has a more concrete, geometric formulation, which are the *Euler structures* of Turaev [**138**]. We describe Turaev's construction presently.

Recall that the tangent bundle of a closed, oriented three-manifold Y is trivial; in particular Y admits a nowhere vanishing vector field. Fix a Riemannian metric on TY, and let $\mathrm{Vec}^*(Y)$ denote the set of unit length

vector fields over Y up to homotopy (through unit vector fields). We have the following definition from [**138**].

Definition 1.4.1. *Two unit vector fields v_1 and v_2 on a connected three-manifold Y are said to be **homologous** if the vector fields are homotopic through unit vector fields in the complement of a point $y \in Y$. A homology class $[v]$ of a unit vector field v is called an **Euler structure**[1] on Y. The set of Euler structures on Y is denoted $[\mathrm{Vec}^*(Y)]$.*

Euler structures over Y can be made into an affine space for $H^2(Y;\mathbb{Z})$, with the help of the following notions.

Definition 1.4.2. *Let Y be a connected three-manifold, and let $f_1, f_2 \colon Y \to S^2$ be two maps. We say that f_1 and f_2 are **homologous** if for some $p \in Y$ their restrictions to $Y \setminus \{p\}$ are homotopic.*

Lemma 1.4.3. *Fix a generator $\mu \in H^2(S^2;\mathbb{Z}) \cong \mathbb{Z}$. The map sending $f \colon Y \to S^2$ to $f^*(\mu)$ induces a one-to-one correspondence between $H^2(Y;\mathbb{Z})$ and homology classes of maps from Y to S^2.*

Proof. Consider the diagram

$$\begin{array}{ccc} [Y, S^2] & \xrightarrow{\quad\alpha\quad} & H^2(Y;\mathbb{Z}) \\ \Big\downarrow{\scriptstyle\beta} & & \Big\downarrow{\scriptstyle\gamma} \\ [Y \setminus \{p\}, S^2] & \xrightarrow{\quad\delta\quad} & H^2(Y \setminus \{p\};\mathbb{Z}) \end{array} \tag{1.7}$$

where the horizontal maps are induced by pullback $f \mapsto f^*(\mu)$, and the vertical ones by restriction. In particular, the space of homology classes of maps from Y to S^2 coincides with the image of β. It is elementary to verify that γ is an isomorphism.

For any CW complex X, $H^2(X;\mathbb{Z})$ corresponds to homotopy classes of maps from X to the infinite dimensional complex projective space $\mathbb{CP}^\infty$, a space which can be given the structure of a CW complex obtained by attaching cells of dimension ≥ 4 to S^2. By the Cellular Approximation Theorem (cf. [**47**, Theorem 4.8]), for a three-dimensional CW complex, such as Y, the pull-back map $f \mapsto f^*(\mu)$ induces a surjective map $[Y, S^2] \to H^2(Y;\mathbb{Z})$; i.e. the map α in Euqation (1.7) is surjective.

Moreover, applying cellular approximations to homotopies, it also follows that for a two-dimensional CW complex Z, the pull-back map induces an

[1] Turaev calls these *smooth Euler structures*, to distinguish them from his combinatorial Euler structures, which we will not discuss here.

isomorphism $[Z, S^2] \cong H^2(X;\mathbb{Z})$. Since $Y \setminus \{p\}$ is homotopy equivalent to a two-dimensional CW complex, it follows that δ is an isomorphism.

We conclude that the pull-back map induces a one-to-one correspondence between $H^2(Y;\mathbb{Z})$ and the image of β, as stated.

□

Fix a trivialization of the tangent bundle $\tau\colon TY \to Y \times \mathbb{R}^3$. This induces an identification of $\mathrm{Vec}^*(Y)$ with the set of homotopy classes of maps from Y to S^2, denoted $[Y, S^2]$: compose the vector field $v\colon Y \to TY$ with τ, followed by projection to the $\mathbb{R}^3$ factor and (since v is nowhere zero) by the natural map $\mathbb{R}^3 \setminus \{0\} \to S^2$. By pulling back the generator $\mu \in H^2(S^2;\mathbb{Z})$ with the above composition, we obtain a map

$$f_\tau\colon [\mathrm{Vec}^*(Y)] \to H^2(Y;\mathbb{Z}) \tag{1.8}$$

which is a one-to-one correspondence by Lemma 1.4.3. Thus, given $[v] \in [\mathrm{Vec}^*(Y)]$ and $a \in H^2(Y;\mathbb{Z})$, there is a unique $[v'] \in [\mathrm{Vec}^*(Y)]$ so that $f_\tau([v]) - f_\tau([v']) = a$. This defines an action of $H^2(Y;\mathbb{Z})$ on $[\mathrm{Vec}^*(Y)]$, where $a \in H^2(Y;\mathbb{Z})$ sends $[v]$ to $[v']$. Although the value $f_\tau([v])$ might depend on the choice of τ, the action of $H^2(Y;\mathbb{Z})$ defined above does not, according to the following:

Proposition 1.4.4. *The action defined above gives a transitive, faithful action of $H^2(Y;\mathbb{Z})$ on $[\mathrm{Vec}^*(Y)]$ that is independent of the choice of the trivialization τ.*

Proof. The fact that the action is transitive and faithful is an immediate consequence of Lemma 1.4.3. It remains to verify that the action is independent of τ. To this end, let τ, τ' be two trivializations of the tangent bundle of Y. We want to show that for any two homology classes $[v], [v'] \in [\mathrm{Vec}^*(Y)]$,

$$f_\tau([v]) - f_\tau([v']) = f_{\tau'}([v]) - f_{\tau'}([v']);$$

or equivalently, that

$$f_{\tau'}([v]) - f_\tau([v]) = f_{\tau'}([v']) - f_\tau([v']). \tag{1.9}$$

Observe first that given τ and τ' there is a map $\gamma\colon Y \to SO(3)$ with the property that $\tau'(y) = \tau(y) * \gamma(y)$, where $*$ represents the pointwise action of $SO(3)$ on the trivializations of T_yY. In particular, for any non-zero vector field v, we have that $\tau' \circ v = (\tau \circ v) \star (\gamma(y))$, where now $\star$ is induced from the rotation action $r\colon S^2 \times SO(3) \to S^2$.

Recall that $H^2(SO(3);\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z}$ and let x denote its non-zero element. Under the Künneth decomposition we have

$$H^2(S^2 \times SO(3);\mathbb{Z}) \cong H^2(S^2;\mathbb{Z}) \oplus H^2(SO(3);\mathbb{Z}) \cong \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}.$$

If $\mu \in H^2(S^2;\mathbb{Z})$ is a generator, then a straightforward computation shows that $r^*(\mu) = \mu + x$. It follows that $(\tau' \circ v)^*(\mu) = (\tau \circ v)^*(\mu) + \gamma^*(x)$; i.e.

$$f_{\tau'}([v]) - f_\tau([v]) = \gamma^*(x).$$

Since the right-hand-side is independent of $[v]$, Equation (1.9) follows. □

Given an Euler structure $[v]$, there is another Euler structure $J[v]$, represented by $-v$. An Euler structure gives rise to a two-dimensional cohomology class $c_1([v]) = [v] - [Jv]$.

Exercise 1.4.5. *Show that $c_1([v])$ is the first Chern class of the oriented two-plane bundle $v^\perp$.*

Example 1.4.6. *For $Y = S^3$ we have $\mathrm{Vec}^*(S^3) \cong \mathbb{Z}$, as it can be identified with $[S^3, S^2] \cong \pi_3(S^2) \cong \mathbb{Z}$; whereas $[\mathrm{Vec}^*(S^3)]$ consists of a single point, since $H^2(S^3;\mathbb{Z}) = 0$.*

Having introduced the notion of homology classes of vector fields $[\mathrm{Vec}^*(Y)]$, we turn to the study of the larger set $\mathrm{Vec}^*(Y)$ itself.

Note that a non-vanishing vector field on a three-manifold can be modified in a neighborhood of any point. Such local modifications are parameterized by $\pi_3(S^2) \cong \mathbb{Z}$, the space of homotopy classes of (non-vanishing) vector fields on S^3. Equivalently, this action can be realized as forming connected sum with the vector fields on S^3. This construction provides a $\mathbb{Z}$-action on the set $\mathrm{Vec}^*(Y)$.

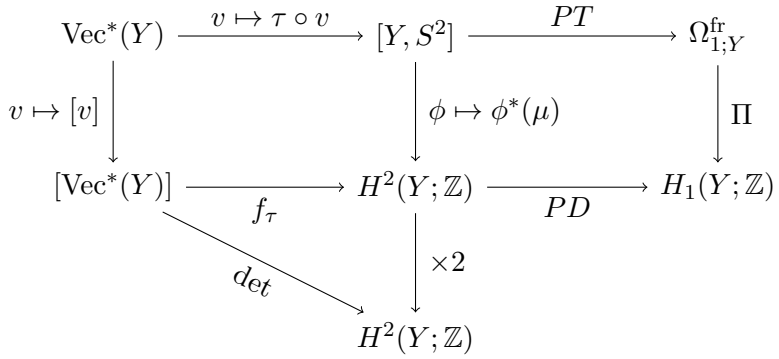
Definition 1.4.7. *The **divisibility** of a class $\xi \in H^2(Y;\mathbb{Z})$ is the integer defined to be 0 if ξ is torsion; otherwise, it is the largest positive integer n so that $\xi \in n \cdot H^2(Y;\mathbb{Z})$. Equivalently, the divisibilty is the non-negative generator of the image of the map $H_2(Y;\mathbb{Z}) \to \mathbb{Z}$ obtained by evaluating against ξ.*

Proposition 1.4.8. *The quotient of $\mathrm{Vec}^*(Y)$ by the above $\mathbb{Z}$ action is the set $[\mathrm{Vec}^*(Y)]$. Moreover, the stabilizer of $v \in \mathrm{Vec}^*(Y)$ is the subgroup $d\mathbb{Z} \subset \mathbb{Z}$, where d is the divisibility of $c_1(v) \in H^2(Y;\mathbb{Z})$.*

Proof. Fix a trivialization τ for TY. As noted earlier, τ induces an identification between $\mathrm{Vec}^*(Y)$ and $[Y, S^2]$. The Pontryagin-Thom construction [**80**] identifies $[Y, S^2]$ with the framed cobordism classes of framed 1-manifolds in Y; this space is denoted by $\Omega^{\mathrm{fr}}_{1;Y}$. This map can be defined

by taking a smooth representative of the homotopy class, and taking the pre-image of a regular value, together with a basis of its tangent plane (the latter providing the framing).

We can organize the information in the following diagram:



The horizontal maps here are all isomorphisms: PT denotes the Pontryagin-Thom isomorphism; PD denotes Poincaré duality; and f_τ is the map from Equation (1.8). The top three spaces have $\mathbb{Z}$ actions: two have been discussed before, and the $\mathbb{Z}$-action on $\Omega^{\mathrm{fr}}_{1;Y}$ is defined by changing the framing. The vertical maps connecting the top two rows are quotients; Π forgets the framing on the one-cycle.

The maps on the top row respect the natural $\mathbb{Z}$-actions on the three sets.

Thus, to complete the argument, it suffices to understand the stabilizer of the $\mathbb{Z}$-action on $\Omega^{\mathrm{fr}}_{1;Y}$.

Suppose that the link $L \subset Y$ with framings f and $f + n$ are framed cobordant. The framed cobordism F_0 in $[0,1] \times Y$ can be closed up to give a surface $F \subset S^1 \times Y$ with self-intersection n. Consider the surface $G = S^1 \times L \subset S^1 \times Y$; it has self-intersection 0. Moreover, since both G and F meet a slice $\{0\} \times Y$ along the same oriented one-manifold L, the homology class $[G] - [F] \in H_2(S^1 \times Y)$ can be represented by a surface that is disjoint from $\{0\} \times Y$; hence, it is homologous to the image of a surface $A \subset Y$, included in some slice $\{p\} \times Y$ in $[0,1] \times Y$.

Under the Künneth decomposition we have

$$H_2(S^1 \times Y) \cong H_2(Y) \oplus \big(H_1(Y) \otimes H_1(S^1)\big) = H_2(Y) \otimes H_1(Y).$$

Homology classes in the first summand are those contained in the image of $i_*\colon H_2(Y) \to H_2(Y \times S^1)$, where $i\colon Y \to S^1 \times Y$ is the inclusion map into some slice $\{p\} \times Y$; the homology class $[F] - [G]$ is contained in this first summand. By construction, $[G]$ is contained in the $H_1(Y) \otimes H_1(S^1)$ summand.

Since homology classes from either summand have self-intersection number equal to 0, it follows that

$$\begin{aligned} n &= \#([F] \cap [F]) \\ &= \#(([F]-[G])+[G]) \cap ([F]-[G])+[G]) \\ &= 2\#(([F]-[G]) \cap [G]) \\ &= 2\langle i_*[A], [S^1 \times L])\rangle_{S^1 \times Y} \\ &= 2\langle [A], \mathrm{PD}[L]\rangle_Y \\ &= \langle [A], c_1([v])\rangle, \end{aligned}$$

where $i_*[A] = [F] - [G]$. In particular, the divisibility d of $c_1([v])$ divides n.

Conversely, given $[A] \in H_2(Y;\mathbb{Z})$, let $A \subset Y$ be an embedded representative, thought of as included in $\{\frac{1}{2}\} \times Y$. The above computation shows that F_0 obtained by smoothing out the double points in $A \cup ([0,1] \times L) \subset [0,1]$ gives a framed cobordism from L with framing f and L with framing $f + 2\langle [A], \mathrm{PD}[L]\rangle_Y$. In particular, framings f and $f+d$ (for the divisibility d of $c_1([v])$) along L are framed cobordant, concluding the proof. □

Remark 1.4.9. *We will henceforth refer to the set of Euler structures* $[\mathrm{Vec}^*(Y)]$ *on a three-manifold as the set of spin^c structures over Y, and accordingly denote the set by* $\mathrm{Spin}^c(Y)$. *Implicit in this notation is an identification with the more classical notion of spin^c structures on Y. The classical construction is recalled in Appendix 32; indeed its indentification with Euler structures is spelled out in Section 32.3.*

Chapter 2

Heegaard diagrams

Heegaard Floer homology is defined using a combinatorial description of a three-manifold, called a *Heegaard diagram*, which is a certain picture drawn on a surface. We define these diagrams in Section 2.1, and in Section 2.2 we describe a complete set of moves that can be used to connect any two Heegaard diagrams for the same three-manifold. In Section 2.3 we explain how to extract basic algebro-topological information about the underlying three-manifold (i.e. its fundamental group and homology groups) from a Heegaard diagram representing it. In Section 2.4 we introduce *Heegaard states*, and study some of their basic properties. Heegaard states play a prominent role in our subsequent discussions: as we shall see in Chapter 9, they generate the Heegaard Floer chain complex.

Heegaard Floer homology will be constructed in terms of a Heegaard diagram equipped with an auxiliary choice of a basepoint. In Section 2.5, we formalize this choice, introduce *pointed Heegaard diagrams*, and refine the notion of Heegaard moves for pointed Heegaard diagrams. In Section 2.6 we associate spin^c structures to Heegaard states. This construction is further refined in Section 2.7 to a grading by nowhere vanishing vector fields. We conclude this chapter in Section 2.8 with some topological motivation for studying Heegaard states.

2.1. Definitions and examples

Before introducing Heegaard diagrams, we start with some preliminaries.

Definition 2.1.1. *Let Σ be a closed surface of genus g. A* ***complete set of attaching circles*** *is a collection of disjoint, embedded simple closed*

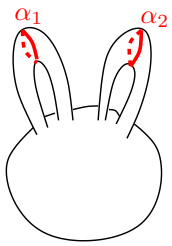


Figure 2.1. A complete set of attaching circles for a surface of genus 2.

curves $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_g\}$ *whose homology classes* $[\alpha_1], \dots, [\alpha_g]$ *are linearly independent in* $H_1(\Sigma; \mathbb{Z}/2\mathbb{Z})$. *See Figure 2.1 for an example.*

The following elementary result provides several alternative ways to think about complete sets of attaching circles.

Proposition 2.1.2. *Let* Σ *be a surface of genus* g, *and let* $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_g\}$ *be a set of* g *disjoint, embedded simple closed curves in* Σ. *The following conditions are equivalent:*

- *(A-1) The homology classes in* $H_1(\Sigma; \mathbb{Z}/2\mathbb{Z})$ *induced by* $\alpha_1, \dots, \alpha_g$ *are linearly independent.*
- *(A-2) The homology classes in* $H_1(\Sigma; \mathbb{Z})$ *induced by* $\alpha_1, \dots, \alpha_g$ *are linearly independent (over* $\mathbb{Z}$*).*
- *(A-3) The complement* $\Sigma \setminus (\alpha_1 \cup \dots \cup \alpha_g)$ *is a connected surface.*
- *(A-4) The complement* $\Sigma \setminus (\alpha_1 \cup \dots \cup \alpha_g)$ *is a planar surface.*

Proof. Consider the surface F obtained by cutting Σ along $\boldsymbol{\alpha}$. The boundary of F contains two copies of each α_i, which we denote α_i' and α_i''. If F is disconnected, then it must contain some component that contains α_i' but not α_i''. That component would then give rise to a non-trivial homological relation between the homology class $[\alpha_i]$ and the $[\alpha_j]$ with $j \neq i$, all viewed as classes in $H_1(\Sigma; \mathbb{Z}/2\mathbb{Z})$. Thus, that (A-1)$\Rightarrow$ Condition (A-3).

Conversely if Condition (A-3) holds, for each i, we can connect α_i' and α_i'' by an arc that closes up in Σ to give a curve β_i which is disjoint from all the α_j with $i \neq j$, meeting α_i transversely in one point. The mod 2 intersection number with the homology classes $[\beta_1], \dots, [\beta_g]$ now gives a homomorphism from $H_1(\Sigma; \mathbb{Z}) \to \mathbb{Z}/2\mathbb{Z}^g$, sending $[\alpha_1], \dots, [\alpha_g]$ to a basis. Thus, Condition (A-3)$\Rightarrow$(A-1).

We have verified the equivalence Condition (A-3)$\Leftrightarrow$(A-1). Verifying the equivalence Condition (A-3)$\Leftrightarrow$(A-2) is similar, only now we use homolgy with coefficients in $\mathbb{Z}$ and signed intersection numbers instead.

To verify Conditions (A-3)$\Leftrightarrow$(A-4), observe that $\chi(\Sigma \setminus \boldsymbol{\alpha}) = \chi(\Sigma)$; so, decomposing $\Sigma \setminus \boldsymbol{\alpha}$ into path-connected components $\cup_{i=1}^{n} F_i$ and letting g_i denote the genus of F_i, we have that

$$2 - 2g = \left(\sum_{i=1}^{n} (2 - 2g_i) \right) - 2g.$$

It follows that $n = 1$ (i.e. Condition (A-3) holds) if and only if each $g_i = 0$ (i.e. Condition (A-4) holds). □

Proposition 2.1.3. *Let Σ be an oriented genus g surface, and suppose that $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are two complete sets of attaching circles for Σ. Then there is an orientation-preserving diffeomorphism $\phi\colon \Sigma \to \Sigma$ that takes $\boldsymbol{\beta}$ to $\boldsymbol{\gamma}$.*

Proof. Cut Σ along $\beta_1, \ldots, \beta_g$ and get a compact planar surface P_β with $2g$ boundary components, together with an orientation reversing involution on ∂P_β. Cutting along $\gamma_1, \ldots, \gamma_g$ gives a similar surface P_γ. By the classification theorem for surfaces, we can find a diffeomorphism from P_β to P_γ that respects these involutions. This map glues together to give the desired diffeomorphism ϕ. □

Definition 2.1.4. *A complete set of attaching circles $\boldsymbol{\alpha}$ in Σ specifies a three-manifold–with–boundary U_α together with an identification $\partial U_\alpha \cong \Sigma$. The manifold U_α is built from $[0,1] \times \Sigma$ by attaching g three-dimensional 2-handles along $\{0\} \times \alpha_1, \ldots, \{0\} \times \alpha_g \subset \{0\} \times \Sigma$ to get a three-manifold with two boundary components, one of which is $\{1\} \times \Sigma$ (naturally identified with Σ), and the other one is a two-sphere. Attaching a three-dimensional 3-handle, we obtain the desired handlebody U_α.*

The three-manifold U_α described above depends only on the choice of the genus g; in fact, it is homeomorphic to a regular neighborhood of a bouquet of g circles in $\mathbb{R}^3$. This three-manifold is called the *genus g handlebody*.

Remark 2.1.5. *It is costumary to call the above three-manifold U_α a handlebody, although earlier (in Definition 1.1.2) a manifold with a handle decomposition was called the same name. We hope no confusion will arise from this.*

The genus g handlebody can be built from D^3 by regarding it as a three-dimensional 0-handle and attaching g (three-dimensional) 1-handles to it.

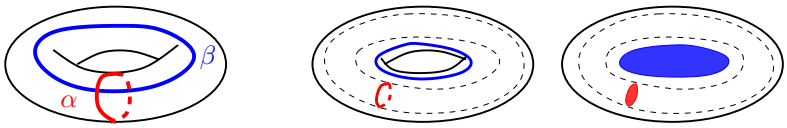


Figure 2.2. Standard genus 1 Heegaard diagram for S^3. At the left is the Heegaard diagram. For the middle, we consider $[-1,1]\times\Sigma$, which is the shown solid torus with the dotted one removed, and equipped with a copy of α on $\{-1\}\times\Sigma$ and a copy of β on $\{1\}\times\Sigma$. At the right, we fill in the copies of α and β by disks. The remaining two 2-spheres are filled with balls to obtain S^3.

This description equips the boundary with a complete set of attaching circles, provided by the belt circles of the 1-handles; see Figure 2.1 for a picture with genus 2. A simpler example is the genus 1 handlebody, which is diffeomorphic to the solid torus $S^1\times D^2$, equipped with a circle of the form $\{p\}\times(\partial D^2)$ with $p\in S^1$.

Definition 2.1.6. *A **Heegaard diagram** is a triple $\mathcal{H}=(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta})$, where Σ is a closed, connected, oriented surface of genus g, and $\boldsymbol{\alpha}=\{\alpha_1,\dots,\alpha_g\}$ and $\boldsymbol{\beta}=\{\beta_1,\dots,\beta_g\}$ are each complete sets of attaching circles in Σ. A Heegaard diagram is called **generic** if each α_i intersects each β_j transversally.*

Note that the g-element sets $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are not ordered; sometimes, however, we will find it convenient to fix orderings on these sets (i.e. thinking of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ as ordered g-tuples of curves).

A Heegaard diagram $\mathcal{H}=(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta})$ can be used to build a closed, oriented three-manifold Y, by gluing the boundary of U_α (which is identified with Σ) to the boundary of $-U_{\boldsymbol{\beta}}$ (which is identified with $-\Sigma$). We say that the Heegaard diagram $\mathcal{H}$ *represents* Y.

Suppose that Y is an oriented, connected three-manifold. A separating surface $\Sigma\subset Y$ that decomposes Y as a union of two handlebodies U_α and U_β provides a *Heegaard splitting* for Y. The three-manifold associated to a Heegaard diagram constructed above is equipped with a natural Heegaard splitting.

Example 2.1.7. *The two-sphere, equipped with two empty sets of curves $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ is the genus zero Heegaard diagram for S^3. A more interesting Heegaard diagram for S^3 is the genus one surface T^2, equipped with a pair of curves α and β that intersect in a single point. This diagram is called the **standard genus one Heegaard diagram for** S^3, shown on the left in Figure 2.2.*

The genus one Heegaard diagram for S^3 has the following generalization:

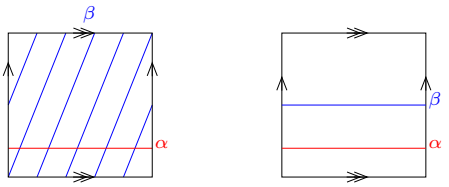


Figure 2.3. More genus 1 Heegaard diagrams. At the left, we have the Heegaard diagram for $L(5,2)$. (For clarity, we have translated the α-circle up a little.) At the right, a generic Heegaard diagram for $S^1 \times S^2$ is shown.

Example 2.1.8. *Fix relatively prime integers (p,q) with $p \geq 1$ and $q \neq 0$. Consider the Heegaard diagram on the genus one surface $\Sigma_1 = (\mathbb{R}\oplus\mathbb{R})/(\mathbb{Z}\oplus\mathbb{Z})$, where α is the image of the horizontal line $\mathbb{R}\oplus 0$ under the quotient map and β is the image of the line $\{(qt,pt) \mid t \in \mathbb{R}\}$. Note that the curves α and β meet transversally in p points in the torus Σ_1. This is a genus one Heegaard diagram for the lens space $L(p,q)$; see the left diagram of Figure 2.3.*

Example 2.1.9. *Consider the genus one surface Σ_1 equipped with the same two homologically non-trivial curves α. This is a (non-generic) genus one Heegaard diagram for $S^1 \times S^2$. We can find a generic Heegaard diagram $(\Sigma_1, \{\alpha\}, \{\beta\})$ for $S^1 \times S^2$, where β is isotopic to, but disjoint from, α; see the right diagram of Figure 2.3.*

Exercise 2.1.10. **(a)** *Show that the diagram from Example 2.1.8 indeed represents the lens space $L(p,q)$.*

(b) *Show that any three-manifold that admits a Heegaard diagram with genus 1 is either a lens space $L(p,q)$ (including S^3), or $S^1 \times S^2$.*

Example 2.1.11. *We construct a family of Heegaard diagrams $\mathcal{H}_n$ parametrized by $n \in \mathbb{Z}$; in Figure 2.4 we illustrated the diagram for $n = 2$. In describing the family, consider the plane compactified with a point at infinity to give S^2, and delete four open disks from S^2 to obtain a planar surface with four boundary components. Think of the genus 2 surface Σ_2 as obtained from this planar surface by identifying the boundary circles in pairs via reflections (preserving the vertical coordinate). Consider the four curves indicated in Figure 2.4. Note that after identifying the two left circles, the arc α_1 becomes a closed curve; and similarly, after the identification of the two right circles, α_2 becomes a closed curve. The curve β_1 is evidently a closed curve. Our family of Heegaard diagrams is obtained from the curve β_2, changing it so that it winds around the right circle various times. For $n \geq 0$ instead of winding around the right circle two times as in the picture, we wind around n times. When $n < 0$, we wind around $-n$ times in the opposite direction. Let W_n be the three-manifold represented by $\mathcal{H}_n$.*

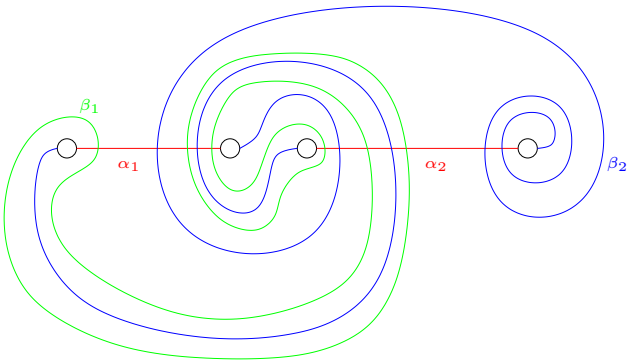


Figure 2.4. Heegaard diagram $\mathcal{H}_2$ from Example 2.1.11. Here the four circles are the feet of the two 1-handles, hence the line segments α_1 and α_2 close up to circles in the genus 2 surface.

Exercise 2.1.12. *Let $\mathcal{H}_n = (\Sigma, \{\alpha_1, \alpha_2\}, \{\beta_1, \beta_2\})$ be the diagram from Example 2.1.11. Consider any orientation-preserving self-diffeomorphism of Σ_2 that maps β_1 and β_2 to α_1 and α_2. Draw the images of α_1 and α_2 under this diffeomorphism.*

Example 2.1.13. *Consider a genus three Heegaard diagram, constructed as follows. Start from S^2 with six distinguished points $\{v_i\}_{i\in\{\pm1,\pm2,\pm3\}}$ chosen so that v_1, v_2, v_3 is an ortho-normal basis for $\mathbb{R}^3$, and $v_{-i} = -v_i$. Draw 12 edges e_{ij}, connecting v_i to v_j for all $j \neq \pm i$. Combinatorially, this gives the octahedron. Next, choose open disks $\{D_i\}_{i\in\{\pm1,\pm2,\pm3\}}$ so that D_i is a neighbourhood of v_i and reflection through the plane orthogonal to v_i sends D_i diffeomorphically to D_{-i}. Consider the genus 3 surface Σ_3 obtained from $S^2 \setminus \cup_{i\in\{\pm1,\pm2,\pm3\}} D_i$ by identifying $\partial\overline{D}_i$ with $\partial\overline{D}_{-i}$. The edges glue up to give three closed curves in Σ, which we label $\boldsymbol{\beta} = \{\beta_1, \beta_2, \beta_3\}$; for $i = 1, \ldots, 3$, $\partial\overline{D}_i$ can be viewed as a closed curve in Σ_3, which we denote α_i. Letting $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \alpha_3\}$, show that $(\Sigma_3, \boldsymbol{\alpha}, \boldsymbol{\beta})$ is a Heegaard diagram for the three-torus T^3. See Figure 2.5.*

We can construct new Heegaard diagrams out of old ones by the following construction:

Definition 2.1.14. *Let $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ and $\mathcal{H}' = (\Sigma', \boldsymbol{\alpha}', \boldsymbol{\beta}')$ be two Heegaard diagrams, and choose points*

$$w \in \Sigma\setminus(\alpha_1\cup\cdots\cup\alpha_g\cup\beta_1\cup\cdots\cup\beta_g) \qquad \text{and} \qquad w' \in \Sigma'\setminus(\alpha'_1\cup\cdots\cup\alpha'_g\cup\beta'_1\cup\cdots\cup\beta'_g).$$

*There is a new Heegaard diagram $\mathcal{H}\#\mathcal{H}'$, the **connected sum of $\mathcal{H}$ and $\mathcal{H}'$**, whose underlying Heegaard surface is the connected sum $\Sigma\#\Sigma'$ of Σ and Σ' at w and w', and whose α-circles are $\boldsymbol{\alpha} \cup \boldsymbol{\alpha}'$ and whose β-circles are $\boldsymbol{\beta} \cup \boldsymbol{\beta}'$.*

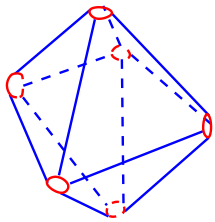


Figure 2.5. Heegaard diagram for T^3. The Heegaard surface is the octahedron with six open disks removed, and oppositely placed boundary circles identified in pairs. The blue arcs glue up to give the three curves in $\boldsymbol{\beta}$. (Each β-circle is glued up from 4 coplanar blue arcs.) The six red circles give three closed curves $\boldsymbol{\alpha}$.

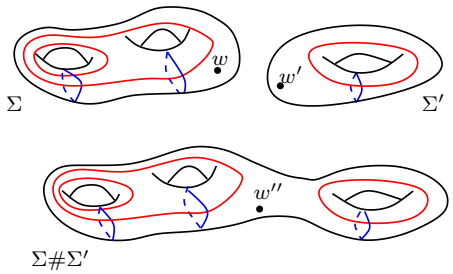


Figure 2.6. The connected sum of $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$ with $\mathcal{H}' = (\Sigma', \boldsymbol{\alpha}', \boldsymbol{\beta}', w')$. We have also indicated a point w'' in the connected sum neck that is disjoint from the $\boldsymbol{\alpha}$- and $\boldsymbol{\beta}$-circles.

It is straightforward to see that if $\mathcal{H}$ represents the three-manifold Y and $\mathcal{H}'$ represents Y', then $\mathcal{H}\#\mathcal{H}'$ represents the connected sum of the underlying three-manifolds, $Y\#Y'$. Also, if $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ is a Heegaard diagram for Y, then $\mathcal{H}_{op} = (\Sigma, \boldsymbol{\beta}, \boldsymbol{\alpha})$ defines $-Y$.

We have the following general existence result:

Theorem 2.1.15. *Let Y be any closed, connected, oriented three-manifold. There exists a Heegaard diagram that represents Y.*

Proof. Consider a self-indexing Morse function f with a unique maximum and minimum on Y, the existence of which is provided by Theorem 1.1.8. According to Theorem 1.1.6, the function f induces a handle decomposition of Y. Let Y_1 be the union of the 0-handle and all the 1-handles in this decomposition.

The Heegaard surface Σ is the oriented boundary of Y_1. Let the belt circles of the 1-handles be the $\boldsymbol{\alpha}$-circles (regarded as circles in $\Sigma = \partial Y_1$), and the attaching circles of the 2-handles be the $\boldsymbol{\beta}$-circles (again, regarded as circles in $\Sigma = \partial Y_1$).

If f has g index-1 critical points, then Y_1 is a genus g handlebody, Σ is a surface of genus g, and we get g $\boldsymbol{\alpha}$-curves in this way. Considering $3-f$ instead of f, we get another self-indexing Morse function on Y, with the same critical points as f. An index-i critical point of f is an index-$(3-i)$ critical point of $3-f$, and the attaching circles of the index-2 critical points of f are the belt circles of the index-1 critical points of $3-f$. The union of the 0-handle and the 1-handles of the handle decomposition induced by $3-f$ give the complement of Y_1 in Y, and since the genus of the common boundary determines the number of 1-handles in the handlebody, we get that the diagram defined above has g $\boldsymbol{\alpha}$- and g $\boldsymbol{\beta}$-circles.

The resulting Heegaard diagram presents Y, concluding the proof. □

Remark 2.1.16. *Each $\boldsymbol{\alpha}$-circle corresponds to an index-1 critical point, and after fixing a metric the circle can be chosen to be those points of Σ which flow under the downward gradient flow $-\nabla f$ to the critical point. Similarly, the $\boldsymbol{\beta}$-circles flow up to the index-2 critical points by the upward gradient flow ∇f; cf. Theorem 1.1.6.*

2.2. Heegaard moves

In this section we describe three basic moves that change the Heegaard diagram while preserving the underlying three-manifold.

Suppose that $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_g\}$ and $\boldsymbol{\alpha}' = \{\alpha'_1, \dots, \alpha'_g\}$ are two sets of attaching circles in Σ. We say that $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'$ are *isotopic* if there is a diffeomorphism $f\colon \Sigma \to \Sigma$ that is isotopic to the identity, so that $\boldsymbol{\alpha}'$ is the image of $\boldsymbol{\alpha}$ under f.

We say that $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'$ *differ by a handle slide* if, after renumbering the curves if necessary, the following conditions hold:

- $\alpha'_i = \alpha_i$ for $i > 1$.
- α'_1 is disjoint from $\alpha_1, \dots, \alpha_g$
- The curves α'_1, α_1, and α_2 bound an embedded subsurface of $\Sigma \setminus (\alpha_3 \cup \dots \cup \alpha_g)$ that is diffeomorphic to the sphere minus three disks (a so-called *pair-of-pants*).

In this case, we say that α'_1 is obtained by sliding α_1 over α_2. See Figure 2.7 for a picture.

A handle slide on $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_g\}$ can be specified by an embedded arc δ connecting some point in α_1 to some point in α_2 in the complement of the other $\boldsymbol{\alpha}$-circles. The pair-of-pants of the handle slide is then the thickening of $\alpha_1 \cup \alpha_2 \cup \delta$.

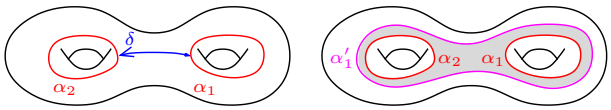


Figure 2.7. Handle slide. The oriented arc δ from α_1 to α_2 on the left specifies a handle slide, as pictured on the right. On the right, the pair-of-pants (which can be thought of as a neighborhood $\delta \cup \alpha_1 \cup \alpha_2$) is shaded.

It is not hard to see that if $\boldsymbol{\alpha}'$ and $\boldsymbol{\alpha}$ differ by a handleslide, then both sets of attaching circles specify the same handlebody. More generally, we have the following:

Exercise 2.2.1. *Show that if $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'$ are disjoint complete sets of attaching circles in Σ, then they are attaching sets for the same handlebody, in the sense that there is diffeomorphism from U_α to $U_{\alpha'}$ that fixes the boundary pointwise.*

Definition 2.2.2. *Let $\mathcal{H} = (\Sigma_g, \{\alpha_1, \dots, \alpha_g\}, \{\beta_1, \dots, \beta_g\})$ be a Heegaard diagram. The* stabilization *$\mathcal{H}'$ of $\mathcal{H}$ is the diagram obtained by taking the connected sum of $\mathcal{H}$ with the standard genus 1 Heegaard diagram for S^3. Dually, we call $\mathcal{H}$ a* destabilization *of $\mathcal{H}'$.*

Clearly, $\mathcal{H}'$ contains a new pair of circles α_{g+1} and β_{g+1}, which are disjoint from all other attaching circles and which meet each other in one point. For an illustration, note that the connected sum in Figure 2.6 is, in fact, a stabilization.

Definition 2.2.3. *A **Heegaard move** on the Heegaard diagram $\mathcal{H}$ gives a new Heegaard diagram $\mathcal{H}'$ so that one of the three conditions holds:*

- *$\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ and $\mathcal{H}' = (\Sigma, \boldsymbol{\alpha}', \boldsymbol{\beta})$, where $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'$ are isotopic; or $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ and $\mathcal{H}' = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}')$, where $\boldsymbol{\beta}$ and $\boldsymbol{\beta}'$ are isotopic. In this case, we say that $\mathcal{H}'$ is obtained from $\mathcal{H}$ by an **isotopy**.*
- *$\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ and $\mathcal{H}' = (\Sigma, \boldsymbol{\alpha}', \boldsymbol{\beta})$ (or $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ and $\mathcal{H}' = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}')$), where $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'$ (or $\boldsymbol{\beta}$ and $\boldsymbol{\beta}'$) are related by a handle slide. In this case, we say that $\mathcal{H}'$ is obtained from $\mathcal{H}$ by a **handle slide**.*
- *$\mathcal{H}'$ is a **stabilization** or a **destabilization** of $\mathcal{H}$.*

Exercise 2.2.4. *Suppose that $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ is a Heegaard diagram, and suppose that α_1 and β_1 meet in exactly one point. Show that all the other circles can be moved by isotopies and handle slides until they are disjoint from both α_1 and β_1.*

Definition 2.2.5. *We say that two diagrams* $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ *and* $(\Sigma', \boldsymbol{\alpha}', \boldsymbol{\beta}')$ *are* ***equivalent****, if there is an orientation-preserving diffeomorphism* $f\colon \Sigma \to \Sigma'$ *between the Heegaard surfaces such that* $f(\boldsymbol{\alpha}) = \boldsymbol{\alpha}'$ *and* $f(\boldsymbol{\beta}) = \boldsymbol{\beta}'$.

For example, in Example 2.1.8, the Heegaard diagrams for $L(p,q)$ and $L(p,q+p)$ are equivalent.

It is easy to see that Heegaard diagrams which are equivalent, or are connected by Heegaard moves, determine diffeomorphic three-manifolds. In fact, Heegaard moves are complete, in the following sense:

Theorem 2.2.6. *Let* Y *be a closed, oriented three-manifold and suppose that*

$$\mathcal{H} = (\Sigma_g, \boldsymbol{\alpha}, \boldsymbol{\beta}), \qquad \text{and} \qquad \mathcal{H}' = (\Sigma_{g'}, \boldsymbol{\alpha}', \boldsymbol{\beta}')$$

are two Heegaard diagrams for Y. *Then, there is a sequence of Heegaard moves that transform the first diagram into one that is equivalent to the second.*

A proof of this theorem will be given in Chapter 31.

Exercise 2.2.7. **(a)** *Use Heegaard moves to show that* W_{-1} *and* W_{-3} *of Example 2.1.11 can both be represented by genus* 1 *Heegaard diagrams.*

(b) *Find* q *and* q' *so that* $W_{-1} = L(5,q)$ *and* $W_{-3} = L(7,q')$.

(c) *Show that* W_{-2} *can be decomposed as the connected sum of two lens spaces. What are these lens spaces?*

2.3. Heegaard diagrams and algebraic topology

Our goal here is to describe topological information about a three-manifold Y explicitly in terms of a Heegaard diagram representing Y.

As a first example, consider the fundamental group $\pi_1(Y,w)$ of a three-manifold Y, specified by a Heegaard diagram $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$, based at a point $w \in \Sigma \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta}) \subset Y$. Orient each $\boldsymbol{\alpha}$- and $\boldsymbol{\beta}$-circle. Consider the group whose generators $\{\gamma_1, \dots, \gamma_g\}$ correspond to the $\boldsymbol{\alpha}$-curves, as follows. The generator γ_i is a closed loop in Σ based at w that crosses α_i exactly once, in a single, transverse intersection point, and is disjoint from all α_j with $j \neq i$. Given each curve β_j, define a word $w(\beta_j)$ in $\{\gamma_1, \dots, \gamma_g\}$ as follows. Fix any point on β_j, and traverse it in the direction specified by the orientation. Terms in $w(\beta_j)$ correspond to the intersection points of β_j with the various $\boldsymbol{\alpha}$-curves; the term corresponding to some $p \in \alpha_i \cap \beta_j$ is either γ_i or γ_i^{-1}, depending on the local intersection number of α_i and β_j at p; and the order of these terms is specified by the orientation of β_j. See Figure 2.8. It is an easy exercise using the Seifert-Van Kampen theorem to

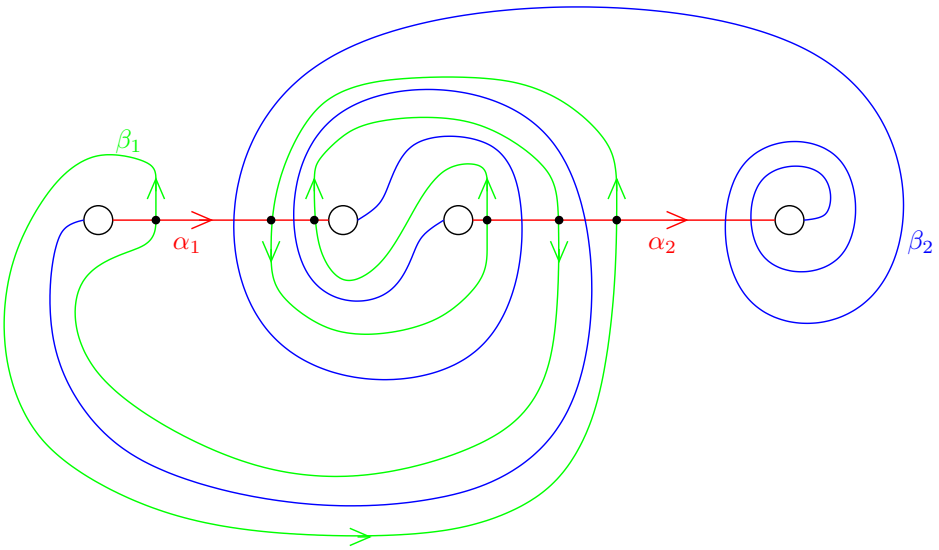


Figure 2.8. Presentation of π_1 in terms of Heegaard diagrams. Orient β_1, α_1, and α_2 as indicated. The six intersection points of β_1 with $\alpha_1 \cup \alpha_2$, traversed in the cyclic order specified by the orientation on β_1, gives the word $w(\beta_1) = \gamma_1\gamma_2\gamma_1^{-1}\gamma_2\gamma_1\gamma_2^{-1}$.

verify that

$$\pi_1(Y) \cong \langle \gamma_1, \dots, \gamma_g \big| w(\beta_1), \dots, w(\beta_g) \rangle.$$

Exercise 2.3.1. *Consider the manifold W_n defined by the Heegaard diagram $\mathcal{H}_n$ of Example 2.1.11 (cf. also Figure 2.8).*

(a) Give a presentation of $\pi_1(W_n)$ with two generators and two relations.

(b) Show that $\pi_1(W_{-1}) \cong \mathbb{Z}/5\mathbb{Z}$, $\pi_1(W_{-2}) \cong (\mathbb{Z}/2\mathbb{Z})*(\mathbb{Z}/3\mathbb{Z})$, and $\pi_1(W_{-3}) \cong \mathbb{Z}/7\mathbb{Z}$.

We now turn to the homology groups of Y, phrased in terms of the Heegaard diagram. The oriented curves $\alpha_1, \dots, \alpha_g$ and $\beta_1, \dots \beta_g$ determine a homomorphism $\phi\colon \mathbb{Z}^g \oplus \mathbb{Z}^g \to H_1(\Sigma;\mathbb{Z})$, by the formula

$$\phi\colon (m_1, \dots, m_g, n_1, \dots, n_g) \mapsto \sum m_i[\alpha_i] + n_i[\beta_i].$$

The cokernel of this homomorphism is $\frac{H_1(\Sigma;\mathbb{Z})}{\mathrm{Span}([\alpha_1],\dots,[\alpha_g],[\beta_1],\dots,[\beta_g])}$, and its kernel can be thought of as the group of homological relations in $H_1(\Sigma;\mathbb{Z})$ between the α- and β-curves.

Proposition 2.3.2. *There is an identification*

$$\frac{H_1(\Sigma;\mathbb{Z})}{\mathrm{Span}([\alpha_1], \dots, [\alpha_g], [\beta_1], \dots, [\beta_g])} \cong H_1(Y;\mathbb{Z}).$$

Similarly, there is an identification between the kernel of ϕ and $H_2(Y;\mathbb{Z})$.

Proof. Consider the long exact sequence of the pair (Y, Σ) (with the understanding that homology is taken with $\mathbb{Z}$ coefficients),

$$\cdots \to H_2(\Sigma) \to H_2(Y) \to H_2(Y, \Sigma) \xrightarrow{\delta} H_1(\Sigma) \to H_1(Y) \to H_1(Y, \Sigma) \to \dots$$

By excision, $H_i(Y, \Sigma; \mathbb{Z}) \cong H_i(U_\alpha, \Sigma; \mathbb{Z}) \oplus H_i(U_\beta, \Sigma; \mathbb{Z})$. Clearly, for a handlebody U with $\partial U = \Sigma$ we have $H_1(U, \Sigma; \mathbb{Z}) \cong H^2(U; \mathbb{Z}) = 0$ and $H_2(U, \Sigma; \mathbb{Z}) \cong H^1(U; \mathbb{Z}) \cong \mathbb{Z}^g$. Therefore the rightmost displayed term in the above sequence is 0. Moreover, the image of $\partial_* \colon H_2(U, \Sigma; \mathbb{Z}) \to H_1(\Sigma; \mathbb{Z})$ is the span of the homology classes of the attaching circles. Thus, the map labeled δ in the above sequence can be identified with the map ϕ, and the H_1 computation follows. The map from $\mathbb{Z} \cong H_2(\Sigma; \mathbb{Z}) \to H_2(Y; \mathbb{Z})$ is trivial, as the Heegaard surface bounds a handlebody. Thus, the H_2 computation follows, as well. □

The above proposition has the following consequence which will be of interest to us:

Corollary 2.3.3. *Let $|H_1(Y; \mathbb{Z})|$ denote the number of elements in $H_1(Y; \mathbb{Z})$ if the latter is a finite group; and let it be 0 otherwise. Then,*

$$\det(\#(\alpha_i \cap \beta_j)_{i,j=1}^g) = |H_1(Y; \mathbb{Z})|.$$

Here, $\#(\alpha_i \cap \beta_j)$ denotes the algebraic intersection number of the oriented curves α_i and β_j in Σ.

Proof. Consider the map $H_1(\Sigma) \to \mathbb{Z}^g$ defined by

$$\gamma \mapsto (\#(\alpha_1 \cap \gamma), \dots, \#(\alpha_g \cap \gamma)).$$

This map induces an isomorphism

$$\frac{H_1(\Sigma)}{\mathrm{Span}([\alpha_1], \dots, [\alpha_g])} \cong \mathbb{Z}^g;$$

furthermore the map identifies the quotient space $\frac{H_1(\Sigma)}{\mathrm{Span}([\alpha_1], \dots, [\alpha_g], [\beta_1], \dots, [\beta_g])}$ with the cokernel of the $g \times g$ matrix

$$M = \#(\alpha_i \cap \beta_j)_{i,j=1}^g.$$

Thus, Proposition 2.3.2 identifies $H_1(Y; \mathbb{Z})$ with the cokernel of M. The corollary follows. □

Exercise 2.3.4. *Compute $H_1(W_n; \mathbb{Z})$ and $H_2(W_n; \mathbb{Z})$.*

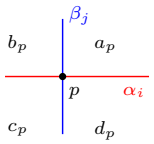


Figure 2.9. Local multiplicities of a domain at an intersection point $p \in \alpha_i \cap \beta_j$.

The curves $\bigcup_{i=1}^{g} \alpha_i \cup \beta_i$ divide Σ into a collection of path-connected components

$$\Sigma \setminus (\alpha_1 \cup \ldots \cup \alpha_g \cup \beta_1 \cup \ldots \cup \beta_g) = \bigcup_{k=1}^{m} \text{int } \mathcal{D}_k^\circ.$$

The closed sets $\mathcal{D}_k = \overline{\mathcal{D}_k^\circ}$ are called the *elementary domains*. A *domain* $\mathcal{D}$ is a formal linear combination of elementary domains:

$$\mathcal{D} = \sum_k m_k \mathcal{D}_k,$$

with $m_k \in \mathbb{Z}$. The *multiplicity* of $\mathcal{D}$ at a point $p \in \Sigma \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta})$ is the coefficient of the elementary domain containing p in the expression for $\mathcal{D}$.

We have the following concrete description of two-dimensional homology classes in terms of domains of $\mathcal{H}$.

Definition 2.3.5. *Let $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ be a generic Heegaard diagram, and fix a domain $\mathcal{D}$. At each intersection point p of α_i and β_j, we have at most four neighbouring regions $\{\mathcal{D}_k\}$ which meet at the corner p. We label the local multiplicities of $\mathcal{D}$ by a_p, b_p, c_p, d_p according to the conventions in Figure 2.9. A domain $\mathcal{D}$ is called **cornerless** if at each point $p \in \alpha_i \cap \beta_j$ (for all $i, j \in \{1, \ldots, g\}$), the four local multiplicities satisfy the relation*

$$a_p + c_p = b_p + d_p.$$

*Fix a point $w \in \Sigma \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta})$. A **periodic domain** is a cornerless domain whose multiplicity at w is 0. Let $\widetilde{\mathcal{P}}$ denote the Abelian group of cornerless domains, and P_w the subgroup of periodic domains.*

Example 2.3.6. *In the standard genus 1 Heegaard diagram for S^3, at the single intersection point all four local multiplicities are the same, so we find that $\widetilde{\mathcal{P}} \cong \mathbb{Z}$ and $P_w = 0$. Consider the genus 1 Heegaard diagram for $S^1 \times S^2$, where the curves α and β are disjoint. In this diagram, all domains are cornerless, so $\widetilde{\mathcal{P}} \cong \mathbb{Z} \oplus \mathbb{Z}$ and $P_w \cong \mathbb{Z}$.*

Lemma 2.3.7. *There is an isomorphism $\widetilde{\mathcal{P}} \cong \mathbb{Z} \oplus H_2(Y; \mathbb{Z})$; and for any $w \in \Sigma \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta})$ we have $P_w \cong H_2(Y; \mathbb{Z})$.*

Proof. A domain can be thought of as a two-chain in Σ, and as such, we can consider the boundary of a domain. Indeed, cornerless domains are exactly those for which the boundary of the corresponding two-chain is in the span of the $[\alpha_i]$ and the $[\beta_j]$. In fact, the boundary $\partial\mathcal{D}$ of a cornerless domain $\mathcal{D} \in \widetilde{\mathcal{P}}$ is in $\ker(\phi)$ (as the domain itself shows that the corresponding 1-chain is a boundary).

Consider this map $\partial\colon \widetilde{\mathcal{P}} \to \ker(\phi)$. By definition each element of $\ker(\phi)$ bounds a domain, which is obviously cornerless, hence ∂ is onto. The kernel of ∂ is generated by the homology class $[\Sigma]$ of the surface Σ, thought of as the domain all of whose local multiplicities are 1. The short exact sequence

$$0 \to \mathbb{Z} \to \widetilde{\mathcal{P}} \to \ker(\phi) \to 0 \tag{2.1}$$

splits since all groups are free, hence $\widetilde{\mathcal{P}} \cong \mathbb{Z} \oplus \ker(\phi)$. Now Proposition 2.3.2 completes the proof. □

For any choice of $w \in \Sigma \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta})$ there is a map

$$n_w\colon \widetilde{\mathcal{P}} \to \mathbb{Z}, \tag{2.2}$$

which is the local multiplicity of $\mathcal{D} \in \widetilde{\mathcal{P}}$ at w. This map induces a splitting of the short exact sequence from Equation (2.1), whose kernel consists of the group of periodic domains P_w.

Definition 2.3.8. *In the proof of Lemma 2.3.7, we have constructed a map from cornerless domains to $H_2(Y;\mathbb{Z})$, which we shall denote*

$$h\colon \widetilde{\mathcal{P}} \to H_2(Y;\mathbb{Z}).$$

According to Lemma 2.3.7, the restriction of h to $P_w \subset \widetilde{\mathcal{P}}$ is an isomorphism.

Every two-dimensional homology class in a manifold can be represented by an oriented surface. For three-manifolds, we can understand this construction concretely from the point of view of Heegaard diagrams, as follows. (This construction will be used explicitly Chapter 8.) We find it convenient to work with domains satisfying the following positivity hypothesis:

Definition 2.3.9. *The domain $\mathcal{D} = \sum_i m_i\mathcal{D}_i$ is **nonnegative** if $m_i \geq 0$ for all i. For nonnegative domains we will write $\mathcal{D} \geq 0$.*

Exercise 2.3.10. *Find the nonnegative, cornerless domains in the Heegaard diagram of T^3 given by Figure 2.5.*

Given a non-negative cornerless domain $\mathcal{D}$, we construct a surface $F_{\mathcal{D}}$ and a map $f\colon F_{\mathcal{D}} \to Y$ that represents $h(\mathcal{D}) \in H_2(Y;\mathbb{Z})$. As a first step, we construct a surface-with-boundary associated to $\mathcal{D}$, which maps into Σ.

Construction 2.3.11. *Let $\mathcal{D} = \sum n_i \mathcal{D}_i$ be a non-negative, cornerless domain. For each elementary domain $\mathcal{D}_i$ consider the surfaces-with-boundary*

$$S^i_1, S^i_2, \dots, S^i_{n_i}$$

where each S^i_j is homeomorphic to $\mathcal{D}_i^c$, together with a map $f^i_j \colon S^i_j \to \mathcal{D}_i$ (the latter viewed as a subset in Σ), which is surjective, maps boundary to boundary and is a diffeomorphism on the interior $\mathcal{D}_i^o$. Glue these pieces together to form a topological space F_0 as follows.

- *If $\mathcal{D}_i$ and $\mathcal{D}_j$ share an $\boldsymbol{\alpha}$-arc a_{ij} in their boundaries, then copies of this interval appear in the surfaces S^i_k and S^j_ℓ. Identify the interval in S^i_k with the corresponding interval in S^j_k for all $1 \leq k \leq \min\{n_i, n_j\}$.*
- *Similarly, if $\mathcal{D}_i$ and $\mathcal{D}_j$ share a $\boldsymbol{\beta}$-arc b_{ij} in their boundaries, then copies of this interval appear in the surfaces S^i_k and S^j_ℓ. Identify the interval in $S^i_{n_i-k+1}$ with the corresponding interval in $S^j_{n_j-k+1}$ for all $1 \leq k \leq \min\{n_i, n_j\}$.*

The map $f_0 \colon F_0 \to \Sigma$ is obtained by patching together the functions f^i_j.

(Compare also [**107**, Chapter 3].)

We claim that F_0 is a two-manifold with boundary. Over points in the interior of the elementary domains, F_0 is homeomorphic to a union of planes, whose number coincides with the multiplicity of the domain. Fix next a point p in $\boldsymbol{\alpha} \cup \boldsymbol{\beta} \setminus (\boldsymbol{\alpha} \cap \boldsymbol{\beta})$, where a domain $\mathcal{D}_i$ with multiplicity n_i meets a domain $\mathcal{D}_j$ with multiplicity n_j, labelled so that $n_i \geq n_j$. Over such a point F_0 consists of a union of n_j planes and $n_i - n_j$ half-planes. Finally, the choices above ensure that even over $p \in (\boldsymbol{\alpha} \cap \boldsymbol{\beta})$, the half planes are glued together along complete edges, so that the preimage of F consists of a union of half-planes and planes. See Figure 2.10 for a picture.

More can be said about the structure of F_0 at its boundary. There are integers $\{k_i\}_{i=1}^g$ and $\{\ell_i\}_{i=1}^g$ associated to the cornerless domain $\mathcal{D}$, defined by

$$\partial \mathcal{D} = \sum_{i=1}^{g} k_i[\alpha_i] + \ell_i[\beta_i].$$

The surface-with-boundary F_0 constructed above has exactly $\sum_{i=1}^g (|k_i| + |\ell_i|)$ boundary components: over α_i, there are k_i disjoint boundary components, and they are each mapped homeomorphically to α_i; similarly, over β_i, there are ℓ_i disjoint boundary components.

Construction 2.3.12. *Given a cornerless domain $\mathcal{D}$, we construct a closed, oriented surface $F_{\mathcal{D}}$, as follows. First, add sufficiently many copies of Σ*

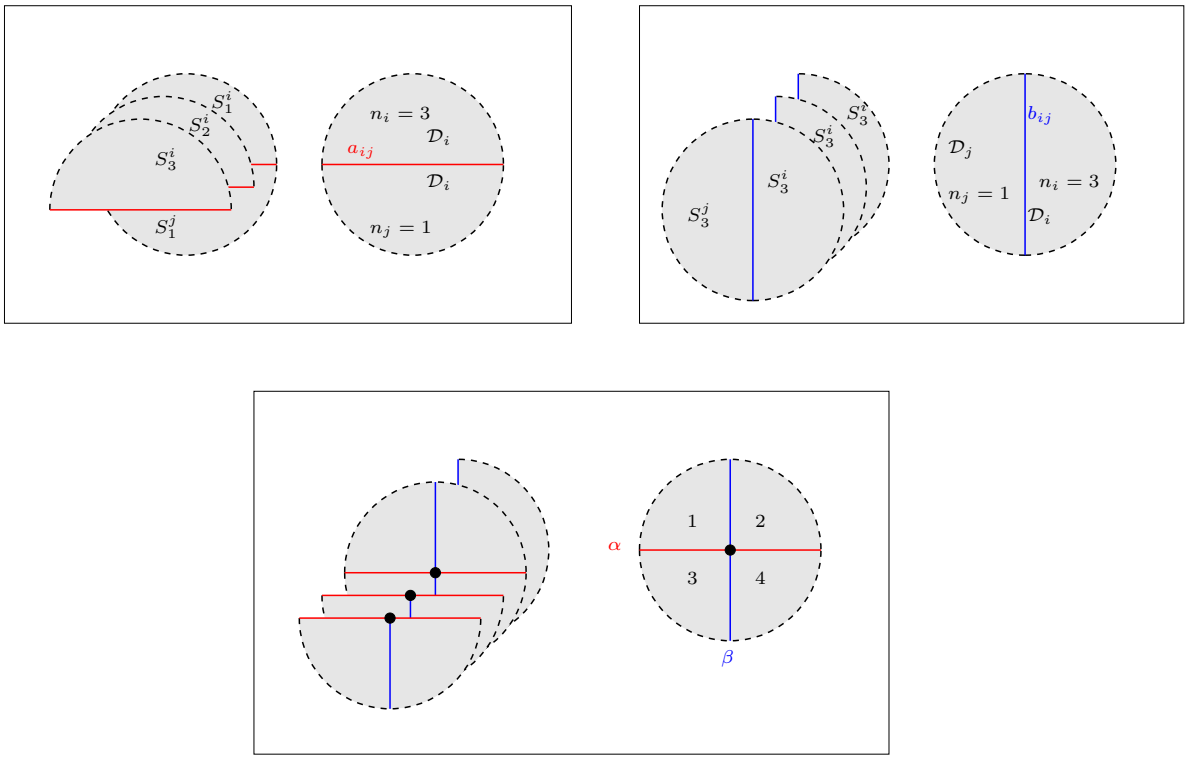


Figure 2.10. Surface associated to a domain. We have drawn heree three examples of portions of domains in the Heeegaar diagrams and the smooth surfaces over theem. The integers on the diagrams represent loca multiplicities.

until the domain becomes non-negative; apply Construction 2.3.11 to get $f_0 \colon F_0 \to \Sigma$. *Next, add copies of the attaching disks along the boundary of* F_0 *to obtain the surface* $F_{\mathcal{D}}$. *Specifically, cap off the* k_i *resp.* ℓ_i *boundary components in* F_0 *over* α_i *resp* β_i *with copies of the* α_i*- resp.* β_i*-attaching disk in the* α*-handlebody resp.* β*-handlebody. This construction gives a closed surface* $F_{\mathcal{D}}$ *containing* F_0*, together with a map* $f \colon F_{\mathcal{D}} \to Y$ *extending the map* $f_0 \colon F_0 \to \Sigma$.

It is straightforward to verify that the homology class represented by f coincides with the one associated to the cornerless domain as in Definition 2.3.8. We added copies of Σ to obtain a positive domain; this clearly does not affect the represented homology class in Y, since the Heegaard surface is null-homologous.

Indeed, every two-dimensional homology class in a three-manifold can represented by an *embedded* surface. (This fact is described in detail in Proposition 17.1.1 below.)

Exercise 2.3.13. *Explain how to modify the above construction of $f\colon F_{\mathcal{D}} \to Y$ to make it an embedding.*

2.4. Heegaard states and domains

Heegaard Floer homology is the homology of a chain complex associated to a Heegaard diagram (and some further analytical choices). The generators of this complex correspond to combinatorial objects, called *Heegaard states*, which we will define presently. For pairs of Heegaard states, there are other combinatorial objects, called *domains connecting Heegaard states*, which we will also define in this section. As we will explain in Chapter 9, these domains play an important role in the definition of the differential of the Heegaard Floer chain complex.

Definition 2.4.1. *Let $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ be a generic Heegaard diagram, and fix an ordering of $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_g\}$ and of $\boldsymbol{\beta} = \{\beta_1, \dots, \beta_g\}$. A **Heegaard state** is a g-tuple of points $\mathbf{x} = \{x_1, \dots, x_g\} \subset \Sigma$, with the property that there is a permutation $\sigma\colon \{1, \dots, g\} \to \{1, \dots, g\}$ so that $x_i \in \alpha_i \cap \beta_{\sigma(i)}$. Let $\mathcal{S}(\mathcal{H})$ denote the set of Heegaard states of $\mathcal{H}$.*

Clearly, a generic Heegaard diagram has only finitely many Heegaard states.

Example 2.4.2. *The genus 0 Heegaard diagram for S^3 has a unique Heegaard state (the empty set). The standard genus 1 Heegaard diagram for S^3 also has a unique Heegaard state. More generally, the genus 1 Heegaard diagram for $L(p, q)$ from Example 2.1.8 has exactly p Heegaard states. The Heegaard diagram for $S^1 \times S^2$ from Example 2.1.9 is not generic. We can perturb α and β so that they become disjoint from one another; in this manner, we obtain a Heegaard diagram with no Heegaard states.*

Exercise 2.4.3. **(a)** *Consider the matrix $M = (m_{i,j})_{i,j=1}^{g}$ where $m_{i,j}$ is the number of points in $\alpha_i \cap \beta_j$. Express the number of Heegaard states in terms of M.*

(b) *Find the number of Heegaard states in the Heegaard diagram for the three-torus T^3 from Example 2.1.13.*

(c) *Find the number of Heegaard states in the diagram $\mathcal{H}_n$ from Example 2.1.11, as a function of n.*

(d) *Suppose that Y is a rational homology sphere (that is, $b_1(Y) = 0$ and hence $|H_1(Y; \mathbb{Z})|$ is finite), and assume that $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ is a Heegaard diagram for Y. Let $\mathcal{S}(\mathcal{H})$ denote the set of Heegaard states of $\mathcal{H}$. Show that $|\mathcal{S}(\mathcal{H})| \geq |H_1(Y; \mathbb{Z})|$.*

Heegaard states have the following interpretation in terms of Morse theory. Let $f\colon Y \to \mathbb{R}$ be a self-indexing Morse function with a unique maximum

and minimum, with g index-1 critical points, and fix a metric h on Y. A *simultaneous trajectory* is a set of g gradient flowlines connecting all the index-1 critical points to all the index-2 critical points of f. When the metric h is generic with respect to f, the associated Heegaard diagram is generic, and the simultaneous trajectories are in natural one-to-one correspondence with Heegaard states.

Having defined Heegaard states, we now consider the domains that connect them.

Definition 2.4.4. *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two Heegaard states. A* ***domain connecting*** $\mathbf{x}$ ***to*** $\mathbf{y}$ *is a domain* $\mathcal{D} = \sum_k m_k \mathcal{D}_k$ *satisfying the linear relation*

$$a_p + c_p = b_p + d_p + \begin{pmatrix} 1 & \text{if } p \in \mathbf{x} \\ 0 & \text{otherwise} \end{pmatrix} + \begin{pmatrix} -1 & \text{if } p \in \mathbf{y} \\ 0 & \text{otherwise} \end{pmatrix}, \tag{2.3}$$

at each $p \in \alpha_i \cap \beta_j$ *(in the notation of Figure 2.9). The set of domains from* $\mathbf{x}$ *to* $\mathbf{y}$ *is denoted* $D(\mathbf{x}, \mathbf{y})$.

It is easy to see that if $\mathcal{D}_1 \in D(\mathbf{x}, \mathbf{y})$ and $\mathcal{D}_2 \in D(\mathbf{y}, \mathbf{z})$, then the formal linear combination $\mathcal{D} = \mathcal{D}_1 + \mathcal{D}_2$ is an element of $D(\mathbf{x}, \mathbf{z})$. Moreover, the group $D(\mathbf{x}, \mathbf{x})$ is independent of the choice of the Heegaard state $\mathbf{x}$; in fact, for any Heegaard state $\mathbf{x}$, $D(\mathbf{x}, \mathbf{x})$ can be identified with the group $\widetilde{\mathcal{P}}$ of cornerless domains, introduced in Definition 2.3.5.

Given two Heegaard states $\mathbf{x}, \mathbf{y}$, there is an obstruction $\epsilon(\mathbf{x}, \mathbf{y}) \in H_1(Y; \mathbb{Z})$ to the existence of a domain connecting $\mathbf{x}$ to $\mathbf{y}$, defined as follows. Suppose the Heegaard diagram is associated to a self-indexing Morse function on Y, and let $\gamma_{\mathbf{x}}$ denote the simultaneous trajectory associated to the Heegaard state $\mathbf{x}$. Clearly, $\gamma_{\mathbf{x}}$ can be thought of as a one-chain in Y, with $\partial \gamma_{\mathbf{x}} = (\sum_{i=1}^{g} [b_i]) - (\sum_{i=1}^{g} [a_i])$, where $\{a_i\}_{i=1}^{g}$ resp. $\{b_i\}_{i=1}^{g}$ are the index-2 resp. index-1 critical points in Y. Thus, if $\mathbf{x}$ and $\mathbf{y}$ are two Heegaard states, $\gamma_{\mathbf{y}} - \gamma_{\mathbf{x}}$ is a one-cycle in Y; let $\epsilon(\mathbf{x}, \mathbf{y}) \in H_1(Y; \mathbb{Z})$ be the homology class it represents.

The element $\epsilon(\mathbf{x}, \mathbf{y})$ has an alternative formulation purely within the Heegaard diagram. Pick paths $\xi_i \subset \alpha_i$ for $i = 1, \ldots, g$ from $\mathbf{x} \cap \alpha_i$ to $\mathbf{y} \cap \alpha_i$; and also pick paths $\eta_i \subset \beta_i$ for $i = 1, \ldots, g$ from $\mathbf{x} \cap \beta_i$ to $\mathbf{y} \cap \beta_i$. The one-chain $\sum_{i=1}^{g} (\xi_i - \eta_i)$ is clearly a cycle, whose homology class in $H_1(\Sigma; \mathbb{Z})$ depends on the choices of paths ξ_i and η_i. The corresponding element $\epsilon(\mathbf{x}, \mathbf{y})$ in the quotient group

$$H_1(\Sigma; \mathbb{Z}) / \mathrm{Span}(\{[\alpha_i], [\beta_i]\}_{i=1}^{g})\rangle \cong H_1(Y; \mathbb{Z})$$

(using the isomorphism from Proposition 2.3.2), is independent of these choices, depending only on the Heegaard states $\mathbf{x}$ and $\mathbf{y}$.

Definition 2.4.5. *Two Heegaard states* $\mathbf{x}, \mathbf{y}$ *in* $\mathcal{H}$ *are* ***equivalent*** *if* $\epsilon(\mathbf{x}, \mathbf{y}) = 0$.

This relation is an equivalence relation, partitioning the Heegaard states of a generic Heegaard diagram into equivalence classes.

Exercise 2.4.6. **(a)** *Show the equivalence of the above two definitions of* $\epsilon(\mathbf{x}, \mathbf{y})$.

(b) *Show that for any three Heegaard states* $\mathbf{x}, \mathbf{y}, \mathbf{z}$,

$$\epsilon(\mathbf{x}, \mathbf{y}) + \epsilon(\mathbf{y}, \mathbf{z}) = \epsilon(\mathbf{x}, \mathbf{z}).$$

(c) *Consider the Heegaard diagram from Example 2.1.8. Show that the* p *different Heegaard states in this example are inequivalent.*

(d) *Determine the equivalence classes of the Heegaard states of the Heegaard diagram of Figure 2.5 (representing* T^3*).*

(e) *Exhibit two distinct Heegaard states in* W_2 *with* $\epsilon = 0$. *Identify all the domains connecting them.*

Proposition 2.4.7. *Suppose that* $\mathbf{x}, \mathbf{y}$ *are two Heegaard states in a generic Heegaard diagram* $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$. *If* $\epsilon(\mathbf{x}, \mathbf{y})$ *is non-zero, then* $D(\mathbf{x}, \mathbf{y})$ *is empty. If* $\epsilon(\mathbf{x}, \mathbf{y}) = 0$, *then* $D(\mathbf{x}, \mathbf{y})$ *is an affine space for* $\mathbb{Z} \oplus H_2(Y; \mathbb{Z})$.

Proof. Clearly, $\epsilon(\mathbf{x}, \mathbf{y}) = 0$ if and only if there are choices of ξ_i and η_i as above so that $\sum_{i=1}^{g}([\xi_i] - [\eta_i])$ is a homologically trivial cycle in Σ. This is equivalent to the existence of a two-chain $\mathcal{D}$ with

$$\partial \mathcal{D} = \sum_{i=1}^{g}([\xi_i] - [\eta_i]).$$

It is straightforward to see that such a two-chain is a domain from $\mathbf{x}$ to $\mathbf{y}$, in the sense of Definition 2.4.4.

If $D(\mathbf{x}, \mathbf{y})$ is non-empty, then the set $D(\mathbf{x}, \mathbf{y})$ is clearly an affine space for the group $\widetilde{\mathcal{P}}$ of cornerless domains: Given $\mathcal{D}_0 \in D(\mathbf{x}, \mathbf{y})$, any other $\mathcal{D} \in D(\mathbf{x}, \mathbf{y})$ can be uniquely written as $\mathcal{D}_0 + P$, where P is a cornerless domain (viewed as an element of $D(\mathbf{x}, \mathbf{x})$). The result then follows from Lemma 2.3.7. □

2.5. Pointed Heegaard diagrams

In the definition of Heegaard Floer homology, we will choose a basepoint on the Heegaard surface. We formalize this choice as follows.

Definition 2.5.1. *A **pointed Heegaard diagram** is a quadruple $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$ where $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ is a Heegaard diagram, and $w \in \Sigma \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta})$. The point w is called the **basepoint**. Two pointed Heegaard diagrams $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$ and $(\Sigma', \boldsymbol{\alpha}', \boldsymbol{\beta}', w')$ are **equivalent**, if there is an orientation-preserving diffeomorphism $f\colon \Sigma \to \Sigma'$ between the Heegaard surfaces such that $f(\boldsymbol{\alpha}) = \boldsymbol{\alpha}'$, $f(\boldsymbol{\beta}) = \boldsymbol{\beta}'$ and $f(w) = w'$.*

We say that a pointed Heegaard diagram $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$ represents Y if the Heegaard diagram $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ represents Y. Heegaard moves have pointed analogues. We say that $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'$ are *pointed isotopic* if there is a diffeomorphism $f = f_1\colon \Sigma \to \Sigma$ that is isotopic to the identity by a one-parameter family of diffeomorphisms $\{f_t\}_{t\in[0,1]}$ ($f_0 = id_\Sigma$), so that $\boldsymbol{\alpha}'$ is the image of $\boldsymbol{\alpha}$ under f and $f_t(\boldsymbol{\alpha})$ is disjoint from w for all t.

Similarly, $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'$ *differ by a pointed handle slide* if they differ by a handle slide so that the pair-of-pants that connects α_1', α_1, and α_2 in the description of the handle slide (cf. Figure 2.7) does not contain the basepoint w.

In Definition 2.1.14 we have defined a connected sum operation on Heegaard diagrams. Note that, although we did not explicitly express it as such, this construction starts from two pointed Heegaard diagrams $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$ and $\mathcal{H}' = (\Sigma', \boldsymbol{\alpha}', \boldsymbol{\beta}', w')$, to form a new Heegaard diagram $\mathcal{H}\#\mathcal{H}'$. Choosing a basepoint w'' in the connected sum region of Σ and Σ' (as shown in Figure 2.6), we obtain a new pointed Heegaard diagram, which we also denote $\mathcal{H}\#\mathcal{H}'$. There were some additional choices made in the definition: to form the connected sum of Σ and Σ', we deleted small neighbourhoods of w and w', identified the boundaries of those neighbourhoods, and chose w'' in the connected sum region. Note that all of these choices give equivalent pointed diagrams.

The connected sum $\mathcal{H}'$ of $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$ with the standard pointed genus one Heegaard diagram for S^3 is called the *pointed stabilization* of $\mathcal{H}$; also, $\mathcal{H}$ is called the *pointed destabilization* of $\mathcal{H}'$.

Pointed isotopies, pointed handle slides, and pointed stabilizations (or destabilizations) are called *pointed Heegaard moves.*

Theorem 2.5.2. *Let Y be a closed, oriented three-manifold and suppose that*

$$\mathcal{H} = (\Sigma_g, \boldsymbol{\alpha}, \boldsymbol{\beta}, w), \qquad \text{and} \qquad \mathcal{H}' = (\Sigma_{g'}, \boldsymbol{\alpha}', \boldsymbol{\beta}', w')$$

are two pointed Heegaard diagrams for Y. Then, there is a sequence of pointed Heegaard moves that transform the first diagram into one that is equivalent to the second.

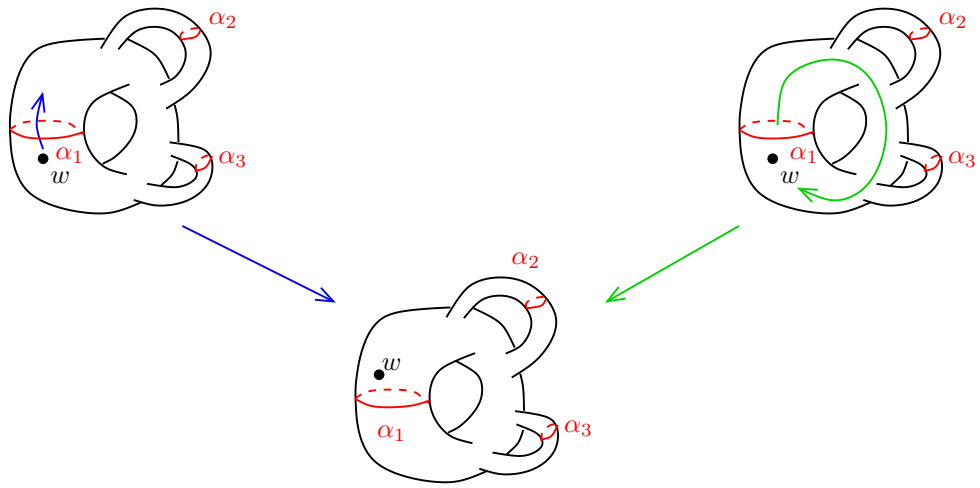


Figure 2.11. Trading isotoping w across α_1 for a sequence of handleslides of α_1 across the other curves.

Proof. This follows quickly from Theorem 2.2.6. The key point is that an isotopy that crosses the basepoint w can be exchanged for a sequence of handle slides and isotopies that are disjoint from w, as follows. Consider the torus obtained by surgering out all the other $\boldsymbol{\alpha}$-circles $\alpha_2, \dots, \alpha_g$ in Σ. This gives a torus containing α_1, the basepoint w, and a collection of $2g$ points $\{p_i, q_i\}_{i=2}^{g}$, where p_i and q_i are the centers of the disks obtained by surgering out α_i. Instead of isotoping w across α_1, we can isotop α_1 around in the torus in the other direction (see Figure 2.11) without crossing w, to obtain an equivalent diagram. Each isotopy of α_1 across p_i or q_i can be lifted to a handleslide in Σ_g, giving the desired sequence of handleslides; see Figure 2.12.

□

2.6. Heegaard states and spin^c structures

A Heegaard state $\mathbf{x} \in \mathcal{S}(\mathcal{H})$ in the pointed diagram $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$ for the three-manifold Y determines an Euler structure, and hence a spin^c structure $\mathfrak{s}_w(\mathbf{x}) \in \mathrm{Spin}^c(Y)$ as follows. By fixing a self-indexing Morse function f and a Riemannian metric h on Y providing the Heegaard diagram $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ (as it is described after Exercise 2.4.3), the Heegaard states of $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ can be identified with the simultaneous trajectories of f, i.e., with the set of g gradient flow lines connecting the g index-1 critical points with the g index-2 critical points. In addition, the basepoint w also determines a

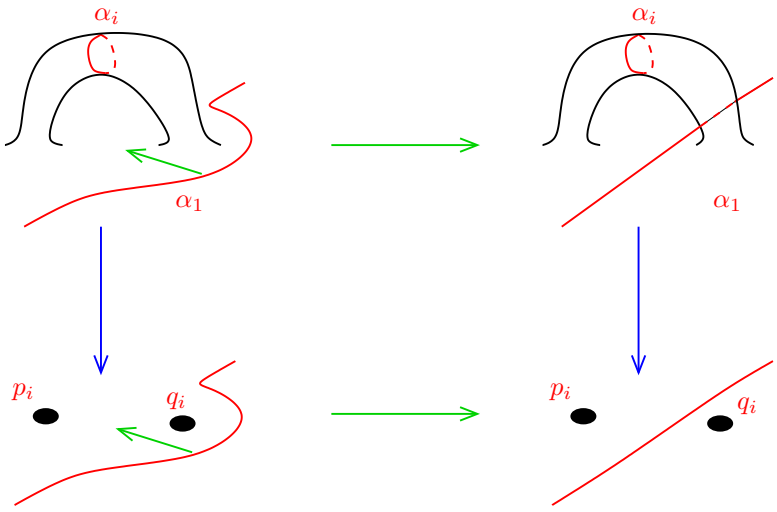


Figure 2.12. Trading isotopies in the torus for handleslides in the Heegaard surface. The vertical maps are surgeries that replace α_i by a pair of points p_i and q_i. Isotoping α_1 across q_i in the torus corresponds to a handlesliding α_1 over α_i in Σ.

gradient flow line, now from the unique index-0 critical point to the unique index-3 critical point of f.

Fix a Heegaard state $\mathbf{x}$ and consider the corresponding simultaneous trajectories, together with the gradient flowline passing through w. (This latter flow connects the minimum and the maximum of f.) Outside a small tubular neighbourhood of these $g+1$ gradient flow lines, the gradient $\nabla_h f$ of f is a nowhere zero vector field that can be extended as a nonvanishing vector field over the $g+1$ balls, since the indices of two critical points in each ball have opposite parity. (Note that the Lefschetz number of the index-i critical point is $(-1)^i$.) In this manner, the Heegaard state, together with the basepoint, determines a nowhere zero vector field on Y, well-defined up to homotopy away from finitely many balls. This object, in turn, is an Euler structure on Y, in the sense of Definition 1.4.1; and as such, it can be viewed as a spin^c structure. (Either by fiat, as in Remark 1.4.9; or by appealing to the identification between these two sets from Proposition 32.3.1.) The obove construction gives a map

$$\mathfrak{s}_w \colon \mathcal{S}(\mathcal{H}) \to \mathrm{Spin}^c(Y). \tag{2.4}$$

In terms of the obstruction class $\epsilon(\mathbf{x},\mathbf{y}) \in H_1(Y;\mathbb{Z})$ we have:

Lemma 2.6.1. *Suppose that $\mathbf{x},\mathbf{y} \in \mathcal{S}(\mathcal{H})$ are two Heegaard states in the pointed Heegaard diagram $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$. Then,*

$$\mathfrak{s}_w(\mathbf{y}) - \mathfrak{s}_w(\mathbf{x}) = PD(\epsilon(\mathbf{x},\mathbf{y})). \tag{2.5}$$

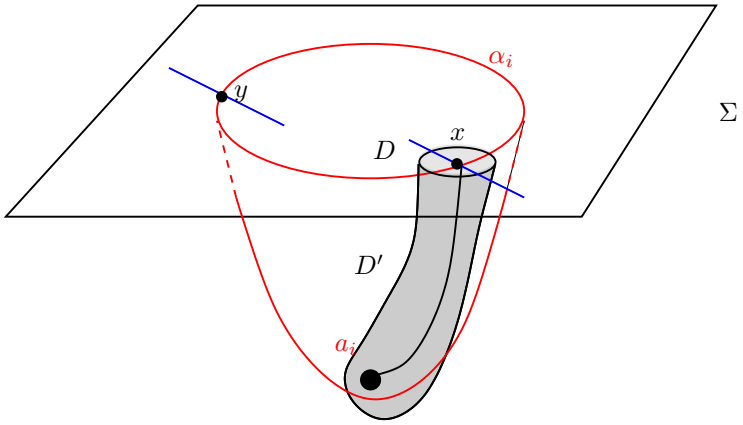


Figure 2.13. Verifying Equation (2.5)**.** We have drawn a disk D in the Heegaard surface around some point $x \in \alpha_i$. Digging into the α-handelbody, we have another disk D' so that $\partial D' = \partial D$, and the two-sphere $D \cup -D'$ bounds a ball containing the index 1 critical point a_i corresponding to α_i.

In particular, $\mathfrak{s}_w(\mathbf{x}) = \mathfrak{s}_w(\mathbf{y})$ *if and only if* $\epsilon(\mathbf{x}, \mathbf{y}) = 0$.

Proof. Let $\gamma_{\mathbf{x}}$ and $\gamma_{\mathbf{y}}$ denote the simultaneous trajectories corresponding to $\mathbf{x}$ and $\mathbf{y}$. Since away from these trajectories the vector fields of the Euler structures corresponding to $\mathbf{x}$ and $\mathbf{y}$ can be chosen to be equal, the difference $s_w(\mathbf{x}) - \mathfrak{s}_w(\mathbf{y})$ (which is a two-dimensional cohomology class on Y) is supported in a neighborhood of $\gamma_{\mathbf{x}} - \gamma_{\mathbf{y}}$; thus, it must be some integral multiple of $PD(\epsilon(\mathbf{x}, \mathbf{y}))$.

To complete the argument, we need to compute this multiple. If $\mathbf{x} = \mathbf{y}$ then the claim obviously follows. In case $\mathbf{x} \neq \mathbf{y}$, there is a gradient flow line in $\gamma_{\mathbf{x}}$ which is not in $\gamma_{\mathbf{y}}$; let D be a small transverse disk to this flowline, intersecting it (and hence $\gamma_{\mathbf{x}} - \gamma_{\mathbf{y}}$) transversally in a unique point. The nowhere zero vector field $v_{\mathbf{y}}$ associated to $\mathbf{y}$ along this disk is ∇f, while the nowhere zero vector field $v_{\mathbf{x}}$ associated to $\mathbf{x}$ is ∇f only near ∂D (assuming ∂D is not in the small neighbourhood of the flowline where ∇f will be changed to get the nowhere zero vector field). Since the two vector fields are equal on ∂D, and (after fixing a trivialization of TY) both give maps to S^2, we can consider their degree difference and it follows that

$$\mathfrak{s}_w(\mathbf{x}) - \mathfrak{s}_w(\mathbf{y}) = (\deg_D(v_{\mathbf{x}}) - \deg_D(v_{\mathbf{y}}))PD(\gamma_{\mathbf{x}} - \gamma_{\mathbf{y}}).$$

Consider now another disk D' with the same boundary as D, and with the property that $D \cup D'$ bounds a 3-ball in Y containing the index-1 critical point of f which is the endpoint of the gradient flowline to which D was normal. On D' the vector field $v_{\mathbf{x}}$ can be chosen to agree with ∇f; hence

(since $v_{\mathbf{x}}$ is nowhere zero) we get that $\deg_D(v_{\mathbf{x}}) = -\deg_{D'}(\nabla f)$. (See Figure 2.13.) Therefore

$$\deg_D(v_{\mathbf{x}}) - \deg_D(v_{\mathbf{y}}) = -\deg_{D'}(\nabla f) - \deg_D(\nabla f),$$

and since ∇f vanishes at the index-1 critical point with Lefschetz number -1, the above difference is equal to 1, and Equation (2.5) is verified. □

A similar formula describes the change of the map $\mathfrak{s}$ when the basepoint w is moved. To this end, suppose that w_1 and w_2 are two base points on the two sides of the $\boldsymbol{\alpha}$-curve α_i; that is, there is a path ϵ interval in Σ joining w_1 and w_2 that intersects α_i transversally once and is disjoint from all the other $\boldsymbol{\alpha}$-curves. As the complement of the $\boldsymbol{\alpha}$-curves is connected, the two endpoints w_1 and w_2 of ϵ can also be connected in the complement of the $\boldsymbol{\alpha}$-curves, hence we get a simple closed curve $\delta_i \subset \Sigma$ that intersects α_i once and is disjoint from all the other $\boldsymbol{\alpha}$-curves. Although this curve δ_i is not unique, its associated homology class $[\delta_i] \in H_1(Y;\mathbb{Z})$ is uniquely determined; we denote the homology class by α_i^*.

Lemma 2.6.2. *For the two basepoints w_1, w_2 described above, and for any fixed Heegaard state $\mathbf{x} \in \mathcal{S}(\mathcal{H})$ of the diagram $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ we have*

$$\mathfrak{s}_{w_1}(\mathbf{x}) - \mathfrak{s}_{w_2}(\mathbf{x}) = PD(\alpha_i^*),$$

where Poincaré duality is taken in the three-manifold Y.

Proof. As above, we think of the Heegaard diagram as specified by a self-indexing Morse function f, and we will consider gradient trajectories for f on Y, equipped with an auxiliary Riemannian metric.

The difference $\mathfrak{s}_{w_1}(\mathbf{x}) - \mathfrak{s}_{w_2}(\mathbf{x})$ of the two spin^c structures is a cohomology class that is supported near $\gamma_{w_1} - \gamma_{w_2}$, where for $i = 1, 2$, γ_{w_i} is the grading flowline through w_i. (Note that γ_i connects the minimum of f to the maximum of f.)

We claim that in Y, the closed curve $\gamma_{w_1} - \gamma_{w_2}$ is homotopic to a curve δ_i in Σ that is crosses α_i exactly once and is disjoint from the α_j for $j \neq i$. To see this, connect w_1 to w_2 by a path ϵ in Σ that crosses only α_i, and another path η in Σ that is disjoint from all the $\boldsymbol{\alpha}$. Thus, $\delta_i = \epsilon - \eta$. The gradient flowlines from η into the maximum provide a homotopy between between δ_i and $\gamma_{w_1} - \gamma_{w_2}$. It follows now that $\gamma_{w_1} - \gamma_{w_2}$ represents the homology class α_i^*; and hence, as in the proof of Lemma 2.6.1, $\mathfrak{s}_{w_1}(\mathbf{x}) - \mathfrak{s}_{w_2}(\mathbf{x})$ is a multiple of $PD[\alpha_i^*]$. In computing the multiple, we argue as before: we choose a small disk D transversal to γ_{w_1} and D' with the same boundary and with the property that $D \cup D'$ bounds a 3-ball containing the index-0 critical point. Similar degree calculations as before imply the result. □

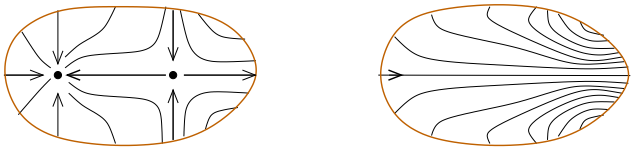


Figure 2.14. Cancellation of index one and zero in dimension two. On the left, we have illustrated the gradient vector field around a gradient flowline from an index one to and index zero critical point. On the right, we have illustrated a new flow in the disk with no zeros, which agrees with the above flow on the boundary of the disk.

2.7. Heegaard states and vector fields

The assignment from Heegaard states to spin^c structures can be further refined to give a grading of Heegaard states by isotopy classes of nowhere vanishing vector fields over Y, denoted $\mathrm{gr}\colon \mathcal{H} \to \mathrm{Vec}^*(Y)$. We explain this construction here, following Gripp and Yang [**139**]. (Note that Gripp and Yang describe the grading set as isotopy classes oriented two-plane fields, by analogy with Seiber-Witten theory [**63**]. These latter objects are in one-to-one correspondence with nowhere vanishing vector fields, by taking orthogonal complements.)

The spin^c structure was constructed by noting that the zeros of the gradient vector field can be cancelled in pairs in a neighborhood of the simultaneous trajectories $\gamma_{\mathbf{x}} \cup \gamma_w$ associated to the Heegaard state $\mathbf{x}$ and the basepoint w, and observing (either by definition or by appealing to Proposition 32.3.1) that the underlying spin^c structure associated to the non-vanishing vector field is independent of the manner in which this cancellation is done. To refine the construction to $\mathrm{Vec}^*(Y)$, we must specify exactly how we cancel the zeros of the gradient vector field in pairs. This must be specified with care, since there is a choice on how to extend a nowhere vanishing vector field from a two-sphere over a three-ball: the space of these extensions is parameterized by $\pi_3(S^2) \cong \mathbb{Z}$.

We describe first the cancellation of index 1 and index 2 critical points, which happens in a neighborhood of $\gamma_{\mathbf{x}}$; and then return to the cancellation of the index 0 and 3 critical points.

As a warm-up, we consider the cancellation of index 0 and 1 critical points in dimension 2. To this end, consider a gradient flow connecting an index 0 and index 1 critical point, as pictured on the left in Figure 2.14. As we know from the Lefschetz numbers, the map $S^1 \to S^1$ obtained by restricting to the boundary disk has degree zero; this is also clear from the picture. Thus, the vector field can be replaced by the non-vanishing vector field shown on

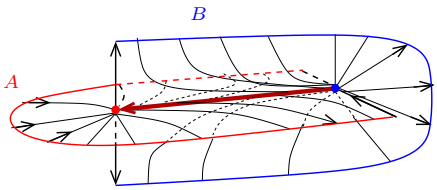


Figure 2.15. Cancel index one and index two critical points (in dimension three).

the right. Of course, there is a unique extension up to homotopy, since the set of extensions is parameterized by $\pi_2(S^1) = 0$.

With the warm-up in hand, we turn to the cancellation of index one and two critical points. To this end, fix a generic gradient flowline γ connecting an index one critical point a and index two critical point b. We construct a neighborhood of γ wherein $\vec{\nabla} f$ is diffeomorphic to the vector field in Figure 2.14 (thought of as lying the xy plane) plus $z\frac{\partial}{\partial z}$. More precisely, we find a two-manifold A containing γ with the following properties:

(A-1) $\vec{\nabla} f$ is tangent to A.

(A-2) In a neighborhood U of A, at each point of $U \setminus A$, $\vec{\nabla} f$ points away from A.

(A-3) The restriction of $\vec{\nabla} f$ to A is diffeomorphic to the index zero and one picture from Figure 2.14

The two-manifold A is a neighborhood of union of points that flow into a under the gradient flow: the *stable manifold* of a. In a Morse chart around a, the subspace A is a disk. In a regular neighborhood U of γ, the unstable manifold extends to a disk whose boundary is an arc in $\partial\overline{U}$ and the union of the two flowlines into b. (See Figure 2.15.) Our manifold A is obtained by extending the unstable manifold past the boundary, so that it is still $\vec{\nabla} f$-invariant. Properties (A-1) and (A-2) follow from the fact that A is $\vec{\nabla} f$-invariant. Property (A-3) follows from the Morse lemma. (In Figure 2.15, we have illustrated both the stable manifold of a and the unstable manifold B of b.)

The above properties state that $\vec{\nabla} f$ near γ is diffeomorphic to the suspension of the flow obtained by the index zero and one cancellation. Thus, we can replace the gradient vector field near γ by the suspension of the nowhere vanishing vector field that appeared in the index zero and one cancellation.

We now turn to the modification in a neighborhood of the gradient flowline γ_w from the index zero critical point, passing through w, and going to the index three critical point. The Morse lemma gives the picture of the gradient

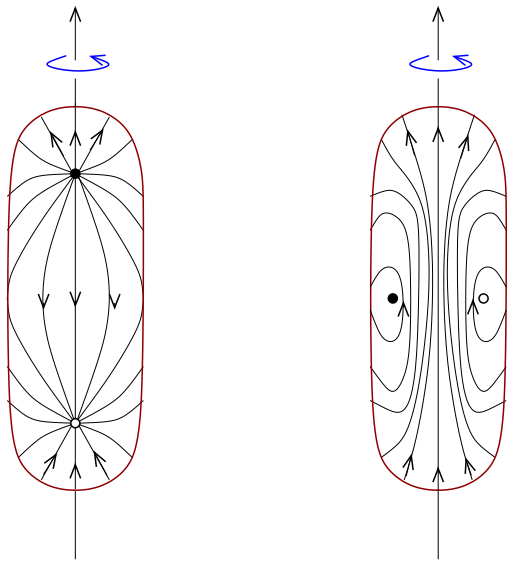


Figure 2.16. Cancel index zero and index three critical points.

flows at the two endpoints. This flow is extended throughout the path to the flow illustrated on the left in Figure 2.16. Specifically, although that is a two-dimensional picture – an index zero and index two critical point in the plane – it can be rotated through the vertical axis to describe a flow in the three-ball.

That flow in turn agrees on the boundary of the disk with a different two-dimensional flow shown on the right of Figure 2.16. Once again, there are two transverse zeros of the vector field in the disk. Rotating the picture in three dimensions through the (indicated) vertical axis gives a new flow which vanishes the circle obtained by rotating the two zeros in the plane. To obtain a nowhere vanishing vector field, we add a vector normal to the pictured surface, pointing out from the plane to the left, vanishing along the axis of rotation, and pointing into the plane to the right – i.e. we can take this to be the vector field generating rotation through the vertical axis.

To state the properties of this assignment, we will refer to the *Hopf vector field* on S^3. This is the vector field generating the one-parameter family of diffeomorphisms of $S^3 = \{(w,z) \in \mathbb{C}^2 \big| |w|^2 + |z|^2 = 1\}$ specified by $t \times (w,z) \mapsto (e^{it}w, e^{it}z)$.

For the statement, recall that if $\mathcal{H}$ is a genus g pointed Heegaard diagram, its stabilization $\mathcal{H}'$ is the genus $g+1$ diagram obtained by forming the connected sum of $\mathcal{H}$ with the standard genus one Heegaard diagram for S^3. There is an induced one-to-one correspondence $\mathcal{S}(\mathcal{H}) \cong \mathcal{S}(\mathcal{H}')$.

Proposition 2.7.1. *The assignment* $\mathrm{gr}\colon \mathcal{S}(\mathcal{H}) \to \mathrm{Vec}(Y)$ *satisfies the following properties:*

*(*gr*-1) The nowhere vanishing vector field* $\mathrm{gr}(\mathbf{x})$ *represents the spin*^{c} *equivalence class* $\mathfrak{s}_w(\mathbf{x})$*.*

*(*gr*-2) Suppose that* $\mathcal{H}'$ *is obtained from* $\mathcal{H}$ *by a single stabilization, and* $\mathrm{gr}\colon \mathcal{S}(\mathcal{H}) \to \mathrm{Vec}^*(Y)$*,* $\mathrm{gr}'\colon \mathcal{S}(\mathcal{H}') \to \mathrm{Vec}^*(Y)$ *denote the corresponding maps. Then, under the one-to-one correspondence* $\mathbf{x} \to \mathbf{x}'$ *between* $\mathcal{S}(\mathcal{H})$ *and* $\mathcal{S}(\mathcal{H}')$*, we have that* $\mathrm{gr}(\mathbf{x}) = \mathrm{gr}'(\mathbf{x}')$*.*

*(*gr*-3) For the single Heegaard state* $\mathbf{x}_0$ *of* S^3 *for the standard genus one Heegaard diagram for* S^3*,* $\mathrm{gr}(\mathbf{x}_0)$ *is the Hopf vector field on* S^3*.*

Proof. Property (gr-1) is immediate from the definitions. Property (gr-2) follows from the fact that the vector field replacing the cancellation of index 1 and 2 zeros of the gradient vector field is isotopic to the gradient vector field after the two critical points are cancelled. In view of Property (gr-2), it suffices to show that if we equip S^3 with its height function, i.e. with a single index 0 and index 3 critical point, and cancel those two zeros in a neighborhood of some gradient flowline through a point w in the Heegaard sphere, the resulting vector field is isotopic to the Hopf vector field. This follows readily from an explicit picture of the Hopf field; see for example [**136**, Figure 2.31]. □

Remark 2.7.1. *In fact, the above proof shows that for the genus zero Heegaard diagram for* S^3*, the grading of the unique (empty) Heegaard state is the Hopf vector field on* S^3*.*

2.8. Bounding the number of Heegaard states

Heegaard states are of fundamental importance in Heegaard Floer homology: they will correspond to the generators of the chain complex. In this section, we give a few remarks about their topological content.

Given a closed, oriented, connected three-manifold Y, let $N(Y)$ denote the minimal number of Heegaard states among all Heegaard diagrams representing Y. For example, $N(S^1 \times S^2) = 0$. It follows from Corollary 2.3.3 that $N(Y) \geq |H_1(Y;\mathbb{Z})|$.

Proposition 2.8.1. *For a closed, connected three-manifold* Y*,* $N(Y) = 1$ *holds if and only if* $Y \cong S^3$*.*

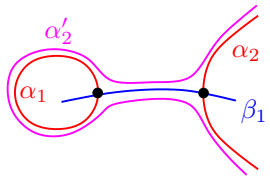


Figure 2.17. If α_1 meets only β_1, and β_1 meets α_2 as shown, we can form a handle slide that eliminates the intersection point with β_1 but does not change the number of Heegaard states.

Proof. One direction of the claim is obvious: see Figure 2.2 for a Heegaard diagram of S^3 with one Heegaard state.

The proof of the converse proceeds by induction on the genus of the Heegaard diagram. The case $g = 1$ is again obvious.

Let $\mathbf{x} = \{x_1, \dots, x_g\}$ be the unique Heegaard state in $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$. We can label the curves so that $x_i \in \alpha_i \cap \beta_i$. Clearly, the hypothesis that $N(Y) = 1$ ensures that α_i and β_i intersect in a single point. We claim that there is some α_i that does not intersect any β_j with $j \neq i$, for otherwise we would be able to construct a Heegaard state $\mathbf{y} \neq \mathbf{x}$ via the following procedure. Suppose that α_1 intersects both β_1 and another $\boldsymbol{\beta}$-curve that we can label β_2. By assumption, α_2 intersects both β_2 and a different $\boldsymbol{\beta}$-curve which is either β_1 or it is a new one, which we can label β_3. Proceeding in this manner, we will eventually find a sequence of intersection points in $\alpha_1 \cap \beta_2, \alpha_2 \cap \beta_3, \dots, \alpha_k \cap \beta_j$ for some $j < k$. There is a Heegaard state $\mathbf{y} = \{y_1, \dots, y_g\}$ with

$$y_i \in \begin{cases} \alpha_i \cap \beta_{i+1} & \text{for all } i \in \{j, \dots, k-1\} \\ \alpha_i \cap \beta_j & \text{for } i = k \\ \alpha_i \cap \beta_i & \text{for } i \notin \{j, \dots, k\}. \end{cases}$$

Clearly, $\mathbf{x} \neq \mathbf{y}$.

The above argument shows that, after renumbering, we can assume that α_1 is disjoint from all β_j with $j > 1$, and it meets β_1 in a single point. We claim that after possibly handle sliding other α_i over α_1, without changing the number of Heegaard states, we can assume that also β_1 meets only α_1. We see this as follows. Suppose that β_1 meets α_1 in one point, and it also meets some other curve α_i. After further renumbering if needed, we can find an arc in β_1 whose interior is disjoint from all α_i and whose endpoints are in α_1 and α_2. Handle sliding α_2 over α_1 along this arc is easily seen to leave the number of Heegaard states unchanged, and to decrease the number of intersection points of β_1 with α_i for $i > 1$. Thus, repeating this step as many times as needed, we obtain a new diagram with one Heegaard state

and which can be destabilized. Clearly, the destabilized diagram also has a unique Heegaard state, and the result follows by induction. □

Remark 2.8.2. *According to Proposition 2.8.1, there is only one three-manifold with $N(Y) = 1$. This has the following generalization, due to Greene and Levine* [**41**]*: for any $m \geq 1$, there are only finitely many rational homology three-spheres Y with $N(Y) = m$.*

It is interesting to consider three-manifolds for which $N(Y) = |H_1(Y; \mathbb{Z})|$, the so-called *strong L-spaces*; see [**41**]. Examples include all lens spaces, and more generally, branched double-covers of alternating knots [**39**]. The fundamental group of such a three-manifold is heavily constrained [**66**].

Chapter 3

Heegaard diagrams and knots

Just as Heegaard diagrams represent closed three-manifolds, Heegaard diagrams with two basepoints represent knots in three-manifolds. We study this construction in the present chapter, formalizing "doubly-pointed Heegaard diagrams" in Section 3.1 and formulating the Heegaard moves that connect them. In Section 3.2 we explain how to construct such diagrams in several particular cases. In Section 3.3, we explain the Heegaard diagrams for describing surgeries on a knot, and Section 3.4 is devoted to present double branched covers from this perspective.

3.1. Doubly-pointed Heegaard diagrams of knots

For our present purposes, a knot K is a smoothly embedded, closed, connected one-dimensional submanifold in a three-manifold Y. A knot endowed with a preferred orientation will be written $\vec{K}$.

Definition 3.1.1. *A **doubly-pointed Heegaard diagram** is a 5-tuple $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w, z)$, where $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ is a Heegaard diagram and $w, z \in \Sigma \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta})$ are two distinct points. Two doubly-pointed Heegaard diagrams are equivalent if there is an orientation preserving diffeomorphism between them. The doubly-pointed Heegaard diagram is **generic** if the $\boldsymbol{\alpha}$- and $\boldsymbol{\beta}$-curves intersect transversally; unless otherwise stated, we will always assume this extra property.*

A doubly-pointed Heegaard diagram specifies a three-manifold Y (forgetting the two basepoints), and the two basepoints specify an oriented knot $\vec{K} \subset Y$.

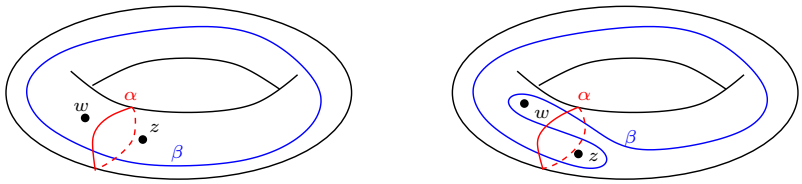


Figure 3.1. Doubly-pointed Heegaard diagrams for the the unknot and the trefoil.

The knot is constructed as follows. First, draw an embedded, oriented arc ξ in $\Sigma \setminus \boldsymbol{\alpha}$ that goes from w to z. (This can be done since $\Sigma \setminus \boldsymbol{\alpha}$ is connected.) Similarly, draw an oriented arc η in $\Sigma \setminus \boldsymbol{\beta}$ that goes from z to w. Note that ξ and η might intersect each other. Thinking of ξ and η as arcs in $\{0\} \times \Sigma \subset [-1,1] \times \Sigma \subset Y$, we can push the interior of ξ down into $(-1,0) \times \Sigma$, and the interior of η up into $(0,1) \times \Sigma$, to obtain two embedded arcs ξ' and η' in $[-1,1] \times \Sigma$, whose union $\vec{K}$ is an embedded, oriented, closed curve in $[-1,1] \times \Sigma$. Although the arcs ξ and η are not unique (even up to isotopy) on Σ, different choices are isotopic in the respective handlebodies.

Definition 3.1.2. *Let $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w, z)$ be a doubly-pointed Heegaard diagram, and let Y be the associated three-manifold equipped with an oriented knot $\vec{K} \subset Y$ constructed above. We say that* $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w, z)$ represents *the pair $(Y, \vec{K})$.*

Notice that the roles of w and z are not symmetric. If $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w, z)$ represents $(Y, \vec{K})$, then $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, z, w)$ represents the same three-manifold with the orientation-reversed knot $(Y, -\vec{K})$.

When the Heegaard diagram arises from a Morse function, the knot can be viewed as the union of the two gradient trajectories (connecting the index 0 and index 3 critical point) through w and z.

Example 3.1.3. *See Figure 3.1 for doubly-pointed Heegaard diagrams representing the unknot and the trefoil. In Figure 3.2, we verify that the second doubly-pointed Heegaard diagram in fact represents the trefoil.*

Exercise 3.1.4. (a) *Construct a doubly-pointed Heegaard diagram for a knot $K \subset S^1 \times S^2$ with the property that the homology class represented by K is a generator of $H_1(S^1 \times S^2; \mathbb{Z}) \cong \mathbb{Z}$.*

(b) *Construct a doubly-pointed Heegaard diagram for a knot $K \subset S^1 \times S^2$ with the property that the homology class of K is twice a generator for $H_1(S^1 \times S^2 : \mathbb{Z}) \cong \mathbb{Z}$.*

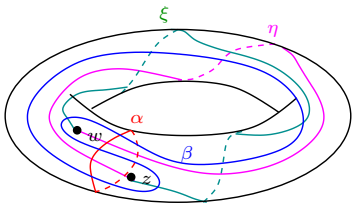


Figure 3.2. Drawing the trefoil on its doubly-pointed diagram. The arcs ξ and η connecting w and z in the complement of α and β are drawn; their union is evidently the (right-handed) trefoil

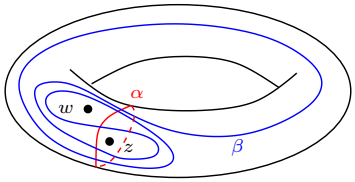


Figure 3.3. Doubly-pointed Heegaard diagram for the Figure-8 knot.

Proposition 3.1.5. *Let Y be a three-manifold and $\vec{K} \subset Y$ be an oriented knot. Then there is a doubly-pointed Heegaard diagram $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w, z)$ that represents $(Y, \vec{K})$.*

Proof. Consider the complement $Y \setminus \nu(K)$ of an open tubular neighbourhood $\nu(K)$ of the knot K and fix a self-indexing Morse function f on this three-manifold with values in $[0, 2.5]$, which admits a unique minimum, and with the property that $f^{-1}(2.5)$ is the torus boundary. This function gives a handle decomposition of $Y \setminus \nu(K)$. The boundary Σ of the union of the single 0-handle and the g 1-handles contains the belt circles $\alpha_1, \ldots, \alpha_g$ of the 1-handles, and the attaching circles $\beta_1, \ldots, \beta_{g-1}$ of the 2-handles. When gluing back the closed tubular neighbourhood of K to $Y \setminus \nu(K)$ to get Y, we attach the last 2-handle along the meridian of the knot. Choosing this attaching circle as β_g, and putting two basepoints w, z to the two sides of β_g gives the desired doubly-pointed Heegaard diagram for $(Y, \vec{K})$. (The orientation of K decides to which side of β_g the basepoint w is positioned.) The resulting Heegaard diagram represents $(Y, \vec{K})$. □

It is natural to consider *doubly-pointed Heegaard moves* for doubly-pointed Heegaard diagrams, consisting of isotopies that do not cross either basepoint (w or z), handle slides where the pairs-of-pants contain neither w nor z, and arbitrary stabilizations and destabilizations. Theorem 2.2.6 has the following refinement in the doubly-pointed case:

Theorem 3.1.6. *Let Y be a closed oriented three-manifold equipped with an oriented knot $\vec{K} \subset Y$. For any two doubly-pointed Heegaard diagrams representing $(Y, \vec{K})$, there is a sequence of doubly-pointed Heegaard moves that transforms the first diagram into one that is equivalent to the second.*

A proof will be given in Chapter 31.

3.2. Constructions of doubly-pointed Heegaard diagrams

We give here a few handy constructions of doubly-pointed Heegaard diagrams for knots.

3.2.1. Heegaard diagrams from knot projections. Fix a knot $K \subset \mathbb{R}^3 \subset S^3$, with a generic decorated projection (P, p) to $\mathbb{R}^2 \subset \mathbb{R}^3$. We give an algorithm for constructing a doubly-pointed Heegaard diagram from this data.

Thinking of K as contained in $\mathbb{R}^3$, the projection is given by $(x, y, z) \mapsto (x, y)$. Singularize the knot to obtain a planar graph X in the (x, y)-plane, which we think of as the locus of points with $z = 0$. Let $U_\beta = \nu(X)$ be a neighborhood of this singular knot. We can think of K as supported inside U_β. The Heegaard surface $\Sigma = -\partial\overline{U_\beta}$ will have genus $g = n + 1$, where n denotes the number of crossings in the knot projection.

The intersection $\Sigma \cap \{z = 0\}$ consists of $n + 2$ circles, corresponding to the connected components of $\mathbb{R}^2 \setminus X$. We will delete one of these circles, and the remaining ones will comprise the $\boldsymbol{\alpha}$-circles. Indeed, two of the components of $\mathbb{R}^2 \setminus X$ are adjacent to the decoration $p \in P$; we discard one of the two corresponding circles.

To construct the $\boldsymbol{\beta}$-circles, we proceed as follows. Add n $\boldsymbol{\beta}$-circles that correspond to crossings in the knot diagram. Specifically, we can find a neighborhood of each crossing in X that meets Σ_{n+1} in a sphere with four disks removed; the crossing governs the choice of the $\boldsymbol{\beta}$-circle, as shown in the top part of Figure 3.4.

Let β_g be the meridian for the knot that is supported in a neighborhood of the fixed point $p \in P$: i.e. the neighborhood of the edge meets Σ in an annulus $S^1 \times [0, 1]$, and we choose β_g to be a curve representing the S^1 factor.

Finally, choose the basepoints w and z to lie in a small neighborhood of β_g, one on each side of β_g; see also Figure 3.4. (The orientation of the knot will specify the sides for w and z.) See Figure 3.5 for a global example.

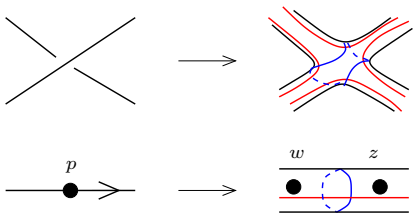


Figure 3.4. From knot projections to Heegaard diagrams, locally. The crossing on the top left is transformed into part of a Heegaard diagram indicated on the top right. The red locus corresponds to $z = 0$ (so they are components of $\boldsymbol{\alpha}$, except if some of the edges contain the point $p \in P$). The blue curve is the corresponding β curve. Similarly, the marked point p on the bottom left is transformed to the portion of the Heegaard diagram in the bottom right.

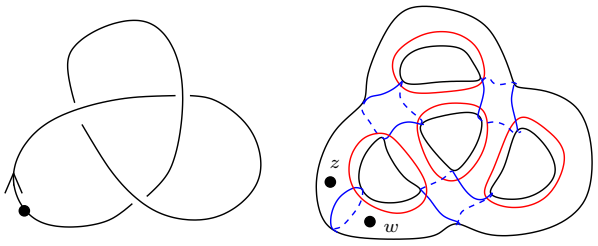


Figure 3.5. From a knot projection of the trefoil to its Heegaard diagram.

Note that we are using an orientation on Σ which is opposite to the one it inherits as the boundary of a regular neighborhood of the graph X.

The resulting diagram is called the *doubly-pointed Heegaard diagram associated to the decorated knot projection* (P, p).

Proposition 3.2.1. *There is a one-to-one correspondence between the Kauffman states of a decorated knot projection and the Heegaard states of the Heegaard diagram associated to the projection.*

Proof. The meridional $\boldsymbol{\beta}$-curve β_g intersects a unique $\boldsymbol{\alpha}$-curve in a single point, hence the component of a Heegaard state on β_g is unique. The other $\boldsymbol{\beta}$-curves, which are associated to crossings c in the projection, intersect at most four $\boldsymbol{\alpha}$-curves, which in turn corresponding to the four regions adjacent to the crossing. Thus, the β_c component of a Heegaard state associates one of the four c-adjacent regions to c. The constraints on a Heegaard state (e.g. that each α_i contains exactly one component of the Heegaard state) correspond to the constraints on a Kauffman state (e.g. that unmarked each region is associated to exactly one crossing). In effect, we have given a

map from Heegaard states to Kauffman states, which is easily seen to be a bijection. □

Exercise 3.2.2. (a) *Show that the doubly-pointed Heegaard diagram associated to a decorated knot projection for $\vec{K} \subset \mathbb{R}^3 \subset S^3$ indeed represents $(S^3, \vec{K})$, in the sense of Definition 3.1.2.*

(b) *Suppose that P_1 and P_2 are two knot diagrams that differ by any of the three Reidemeister moves (away from the basepoint p). Find doubly-pointed Heegaard moves that connect their associated Heegaard diagrams.*

3.2.2. Heegaard diagrams from plat closures of knots. A *plat representation of a knot* is a knot diagram in $\mathbb{R}^2$, equipped with an $\mathbb{R}$-valued function, called the "height function" (which, in practice, we think of as given by the y-coordinate), whose restriction to the knot projection has n global maxima and n global minima, and no further critical points. Thus, such a diagram is represented by n caps on top (representing the n maxima) connected via some braid to n cups on the bottom (representing the minima). Every knot has such a representation.

A knot K is called an *n-bridge knot* if it can be represented by an n-plat, but it cannot be represented by a k-plat for $k < n$. Clearly the unknot is the only 1-bridge knot. The 2-bridge knots are classified by their branched double-covers, which are lens spaces [**124**].

We wish to use the plat description to construct a doubly-pointed Heegaard diagram for K. Consider the sphere S^2 with a linearly arranged collection of points $p_1, q_1, p_2, q_2, \ldots, p_n, q_n$. Let P be the planar surface obtained by cutting out small disks D_{p_i} and D_{q_i} about each of p_i and q_i. Consider next the configuration of disjoint embedded closed curves $\boldsymbol{\gamma} = \gamma_1, \ldots, \gamma_n$ in P, chosen so that for $i = 1, \ldots, n-1$ the curve γ_i encircles both p_i and q_i (and no other p_j or q_j); and an additional circle γ_n encircles only q_n. Next, place two basepoints w_0 and z_0 on either side of γ_n. Our Heegaard surface Σ is obtained by attaching cylinders connecting the boundary component of P corresponding to p_i with the boundary component corresponding to q_i. The braid can be thought of as a diffeomorphism $\phi\colon P \to P$. Let $\boldsymbol{\beta}$ be the image of $\boldsymbol{\gamma}$ under ϕ, and choose basepoints $w = \phi(w_0)$ and $z = \phi(z_0)$. To determine the $\boldsymbol{\alpha}$-circles, draw an arc from p_i to q_i. Let α_i be the closed circle in Σ obtained by closing off the arc inside the cylinder connecting D_{p_i} to D_{q_i}. The data $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w, z)$ is a doubly-pointed Heegaard diagram for $K \subset S^3$. See Figures 3.6 and 3.7; see also [**34**, Figure 12.32].

Note that the above Heegaard diagram can always be destabilized, to obtain one of genus $n-1$.

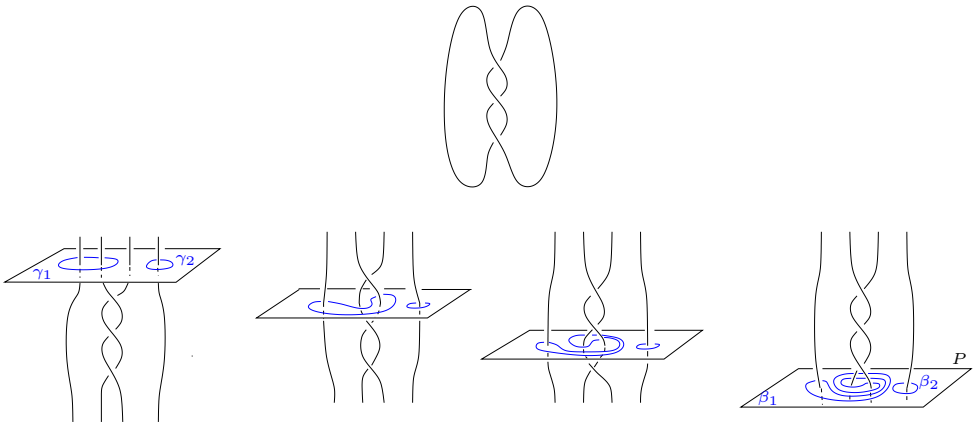


Figure 3.6. Constructing a Heegaard diagram for the trefoil from a plat. The plat is illustrated on the top. Draw the planar diagram P with curves γ_1 and γ_2 on the left; pushing these curves down along the braid (middle two pictures), we end up with a diffeomorphic diagram P with curves β_1 and β_2 on the right.

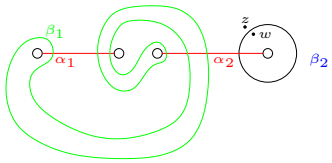


Figure 3.7. Heegaard diagram for the trefoil knot. Completing the diagram from Figure 3.6, we obtain this doubly-pointed Heegaard diagram for the trefoil.

Exercise 3.2.3. (a) *Destabilize the diagram in Figure 3.7 to find a genus* 1 *doubly-pointed Heegaard diagram for the trefoil.*

(b) *Generalize the above construction to give a genus* 1 *doubly-pointed Heegaard diagram for the* $(2, 2n+1)$ *torus knot, for all* $n \in \mathbb{Z}$.

3.2.3. Connected sum. The doubly-pointed Heegaard diagram of the connected sum $K_1 \# K_2$ can be easily determined from the diagrams $\mathcal{H}_i = (\Sigma_i, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, w_i, z_i)$: perform the connected sum of the two diagrams (as in Definition 2.1.14) at the basepoints z_1 and w_2.

3.2.4. Knots on genus 1 diagrams. Let H be a handlebody and a a union of n disjoint arcs in H, whose endpoints lie on ∂H. We say that a is a collection of n *unknotted arcs* if there is an isotopy of a fixing the boundary that takes a to n disjoint arcs in ∂H.

Fix a knot $K \subset Y$. A (g, n) *representation of* K is a genus g Heegaard splitting for Y so that K meets the two handlebodies in n unknotted arcs.

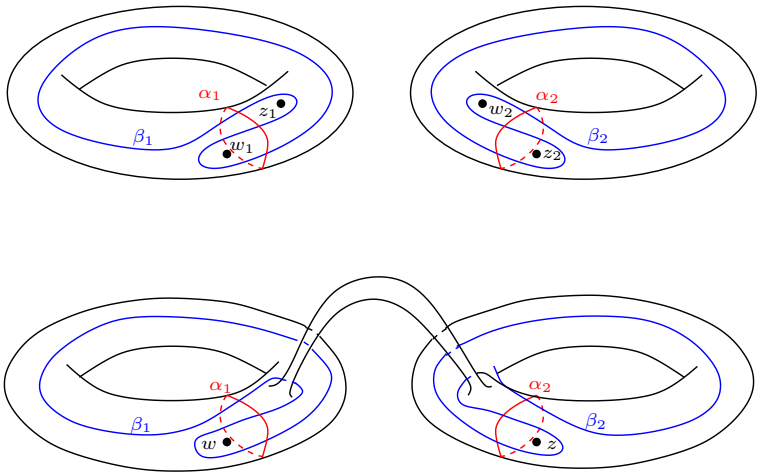


Figure 3.8. Connected sums of knots. Starting from two doubly-pointed Heegaard diagrams on the top (both representing the right-handed trefoil $T_{2,3}$), we form their connected sum on the bottom (representing $T_{2,3}\#T_{2,3}$).

If $K \subset Y$ has a (g,n) representation, then it is called a (g,n) knot. For example, if $K \subset S^3$ has bridge number n, then K is a $(0,n)$ knot.

Exercise 3.2.4. (a) *Show that for $n>1$, a (g,n) knot is also a $(g+1,n-1)$-knot.*

(b) *Show that a $(g,1)$ knot can always be represented by a doubly-pointed Heegaard diagram of genus g.*

An interesting class of knots are the $(1,1)$ ***knots*** in S^3; these are the knots that can be represented on a genus 1, doubly-pointed Heegaard diagram $(\Sigma,\alpha,\beta,w,z)$, where α and β are isotopic (via an isotopy that might cross the basepoints) on the genus one surface $\Sigma = T^2$ to two curves that intersect in exactly one point.

All 2-bridge knots are $(1,1)$-knots. Another family of $(1,1)$-knots are given by torus knots. The knot $K = T_{p,q}$ has a $(1,1)$ representation, as follows. Draw K on the standard torus, thought of as the Heegaard surface for a genus one Heegaard splitting of S^3, divide K into two intervals, and push them into the two (genus one) handlebodies.

For more on $(1,1)$ knots, see [**32, 113, 111**].

Example 3.2.5. *The torus knot $T_{2,2n+1}$ can be given by the (1,1)-diagram shown in Figure 3.9.*

Example 3.2.6. *The torus knot $T_{3,4}$ can be given by Figure 3.10. Indeed,*

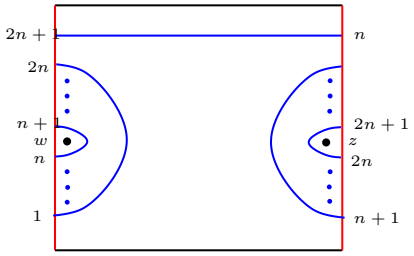


Figure 3.9. A $(1,1)$ diagram for the torus knot $T_{2,2n+1}$ in S^3. The left and right sides of the rectangle are identified with a shift by n, and these two sides form the (red) α-circle. Under this gluing, the blue arcs glue to togeth to form a signle circle, the β-circle.

.

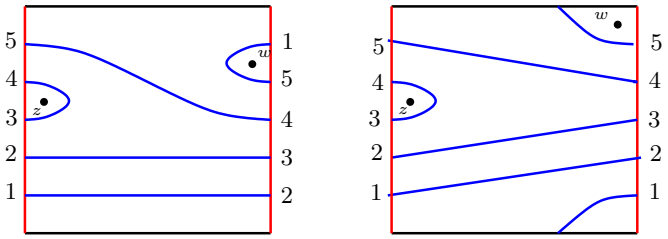


Figure 3.10. A $(1,1)$ diagram for the torus knot $T_{3,4}$ in S^3. We have drawn two pictures: in the left diagram the left and right arc are identified with a shear, while on the right, they are identified directly. As before, the left sides is the (red) α-curve, while the blue arcs glue together to form the β-curve.

.

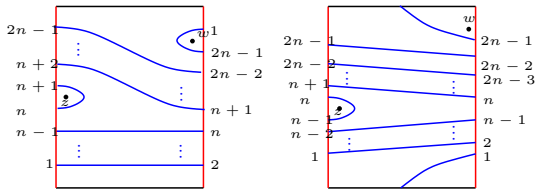


Figure 3.11. A $(1,1)$ diagram for the torus knot $T_{n,n+1}$ with $n \geq 3$ in S^3. Again, on the left the identification is made with a shear, while on the right it is not.

.

this figure generalizes to a diagram for the family $T_{n,n+1}$ of torus knots for any $n \in \mathbb{N}$ (with $n \geq 3$), as shown by Figure 3.11.

Exercise 3.2.7. (a) *Find genus 1 doubly-pointed Heegaard diagrams that represent the torus knots $T_{n,kn+1}$ ($k, n \in \mathbb{N}$ and both at least 2) and $T_{3,5}$.*

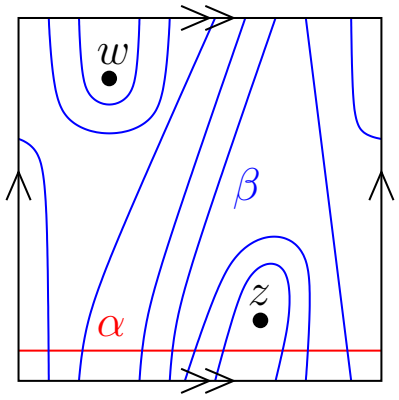


Figure 3.12. A $(1,1)$ diagram for the knot 9_{42} in S^3.

.

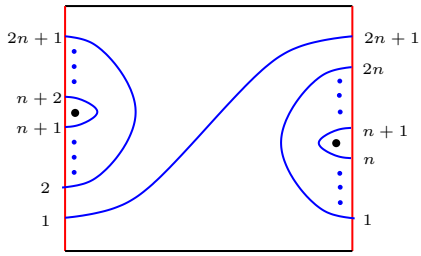


Figure 3.13. A $(1,1)$ diagram for the twist knot Tw_n with $n \geq 1$ in S^3. For $n = 1$ we recover the diagram of the right-handed trefoil knot $T_{2,3}$, and for $n = 2$ we have a diagram for the Figure-8 knot (4_1 in Rolfsen's table).

.

(b) *Show that the knot in Figure 3.12 is neither a 2-bridge knot nor a torus knot.*

(c) *Show that the (1,1)-diagram of Figure 3.13 presents the twist knot Tw_n (where Tw_1 is the right-handed trefoil knot and Tw_2 is the Figure-8 knot).*

Although we typically restrict our attention to knots in S^3, there is an interesting class of $(1,1)$ knots in lens spaces: take the standard genus 1 Heegaard diagram (Σ, α, β) for the lens space $L(p,q)$ (as in Example 2.1.8), and fix a basepoint $w \in \Sigma \setminus \alpha \cup \beta$. There are p different regions where we can put z, giving rise to p different knots $K_i \subset L(p,q)$ for $i = 1, \ldots, p$. These knots are called *simple knots*; see for example [**48**].

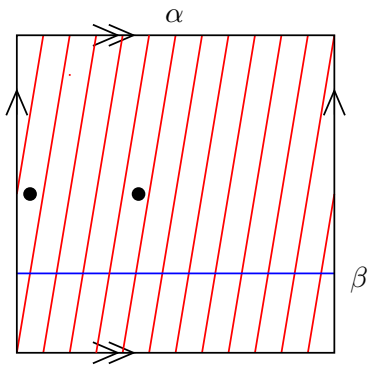


Figure 3.14. A doubly-pointed Heegaard diagram for a simple knot in the lens space $L(11,2)$. The two dots represent the basepoints.

Exercise 3.2.8. **(a)** *Show that for a* $(1,1)$ *knot, the knot group (i.e. the fundamental group of the complement of the knot) has a presentation with two generators and one relation.*

(b) *Consider the simple knot in the lens space pictured in Figure 3.14. Write down an explicit presentation of its knot group (with two generators and one relation).*

(c) *Show that for fixed relatively prime integers* (p,q) *with* $1 \leq q < p$*, the* p *different simple knots represent the* p *different homology classes in* $H_1(L(p,q);\mathbb{Z})$.

3.3. Heegaard diagrams for surgeries

We will now describe how to draw the Heegaard diagram for surgery on a knot K in a three-manifold Y. To this end, let

$$\mathcal{H} = (\Sigma, \{\alpha_1, \ldots, \alpha_g\}, \{\beta_1, \ldots, \beta_{g-1}, \beta_g^*\})$$

be a Heegaard diagram for a three-manifold Y, with a distinguished $\boldsymbol{\beta}$-curve, labeled β_g^*. There is an induced knot $K = K(\beta_g^*) \subset Y$ supported in $U_\beta \subset Y$, with the following two properties:

- K is disjoint from the attaching disks $D_1, \ldots, D_{g-1}$ for $\beta_1, \ldots, \beta_{g-1}$
- K represents the core of the solid torus obtained by removing neighborhoods of $D_1, \ldots, D_{g-1}$.

Given an orientation $\vec{K}$ on $K(\beta_g^*)$, a doubly-pointed Heegaard diagram representing $\vec{K}$ is given by adding w and z on the two sides of β_g^*.

Disregarding β_g^*, the data $(\Sigma, \{\alpha_1, \ldots, \alpha_g\}, \{\beta_1, \ldots, \beta_{g-1}\})$ specifies a handle decomposition of the three-manifold with (torus) boundary $M = Y \setminus$

$\nu(K)$: start from the handlebody determined by $\boldsymbol{\alpha}$, and attach 2-handles along the $g-1$ β-curves. Thus, the various choices for β_g (so that $\{\beta_1, \dots, \beta_g\}$ form a complete set of attaching circles) specify Heegaard diagrams for the various Dehn fillings of M. Indeed, $\Sigma \setminus \nu(\beta_1 \cup \cdots \cup \beta_{g-1})$ is a genus one surface with $2g-2$ boundary components, that is naturally identified with a subset T_0 of the torus boundary of M with $2g-2$ disks removed. This gives a correspondence between filling curves λ for M inside T_0 and choices of β_g. Isotopies in ∂M between curves in T_0 can be followed by handle slides and isotopies of the resulting choices for β_g. In particular, given $(\Sigma, \{\alpha_1, \dots, \alpha_g\}, \{\beta_1, \dots, \beta_g\})$, we can view β_g as specifying a meridian for the knot $K(\beta_g)$. A different choice of attaching circle λ_g that meets β_g transversally in one point corresponds to a framing of $K \subset Y$. If λ_g corresponds to one framing, any other framing can be obtained by twisting λ_g some number of times in the direction specified by β_g, resulting in the curve γ. In conclusion, when replacing β_g^* of $\mathcal{H}$ with γ, we get a Heegaard diagram $\mathcal{H}_\lambda = (\Sigma, \{\alpha_1, \dots, \alpha_g\}, \{\beta_1, \dots, \beta_{g-1}, \gamma\})$ for the surgered three-manifold $Y_\lambda(K(\beta_g^*))$.

Recall that $K \subset S^3$, has a distinguished distinguished framing λ_0, the Seifert framing (Definition 1.2.12), which is characterized by the property that $b_1(S^3_\lambda(K)) = 1$. (All other Dehn fillings Y of $S^3 \setminus \nu(K)$ have $b_1(Y) = 0$.) Therefore we can find λ_g representing this distinguished framing on a Heegaard diagram, by demanding that $[\lambda_g]$ lies in the span of $[\alpha_1], \dots, [\alpha_g]$. (See, for example, Proposition 2.3.2.) Now a diagram for $S^3_n(K)$ can be given by starting with $\mathcal{H}$ so that $K = K(\beta_g^*)$ and choose γ so that it is homologous to $\lambda_0 + n\beta_g^*$.

We have shown how to get Heegaard diagrams for surgeries on a knot, when the knot is specified by a distinguished $\boldsymbol{\beta}$-circle. We show now how to get, in turn, a Heegaard diagram with distinguished β-circle in terms of a doubly-pointed Heegaard diagram

$$\mathcal{H} = (\Sigma, \{\alpha_1, \dots, \alpha_g\}, \{\beta_1, \dots, \beta_g\}, w, z)$$

for $\vec{K}$. Stabilize Σ, if necessary, to obtain a new Heegaard surface Σ', by attaching a 1-handle to Σ with feet near w and z, choose β_{g+1} supported inside the one-handle, and construct α_{g+1} out of two arcs, one of which runs thorough the one-handle, and the other one connects w and z in $\Sigma \setminus \boldsymbol{\alpha}$. This gives the desired Heegaard diagram $\mathcal{H}' = (\Sigma', \{\alpha_1, \dots, \alpha_{g+1}\}, \{\beta_1, \dots, \beta_g, \beta_{g+1}^*\})$ for Y with β_{g+1}^* distinguished. (See Figure 3.15 for an illustration.) Clearly, the knot $K \subset Y$ specified by $\mathcal{H}$ is equivalent to the knot $K(\beta_{g+1}^*)$ specified in the stabilized Heegaard diagram $\mathcal{H}'$.

Example 3.3.1. *By applying the above algorithm to the genus 1 Heegaard diagram of the left-handed trefoil T (which can be derived from the diagram*

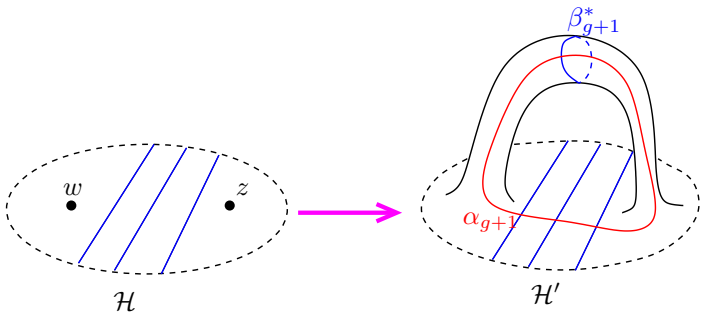


Figure 3.15. Constructing a diagram with distinguished β-curve from a doubly-pointed Heegaard diagram. A portion of the doubly-pointed Heegaard diagram $\mathcal{H}$ is shown on the left; its stabilization $\mathcal{H}'$, with distinguished β-circle β^*_{g+1}, is shown on the right.

.

of $T_{2,3}$ in Figure 3.2) we get the diagram of Figure 2.4, showing that the 3-manifold W_n (defined in Example 2.1.11) is diffeomorphic to $S^3_{n-4}(T)$.

Exercise 3.3.2. **(a)** *Show that W_3 is diffeomorphic to the Poincaré sphere $\Sigma(2,3,5)$, and prove that $\pi_1(W_3)$ has order 120.*

(b) *Consider the simple knot from Figure 3.14. Show that its complement is diffeomorphic to the complement of some knot K in S^3. Draw this knot K.*

3.4. Branched double-covers

The branched double-cover construction can be successfully applied not only in constructing interesting three-manifolds, but also in studying knots and their various properties. It is therefore important to get suitable Heegaard diagrams for these three-manifolds.

The branched double-cover of a genus g surface branched at $2n$ points is a surface of genus $2g+n-1$. Similarly, the branched double-cover of a genus g handlebody branched at n unknotted arcs is a genus $2g+n-1$ handlebody. It follows that the branched double-cover along a (g,n) knot admits a genus $2g+n-1$ Heegaard decomposition. In particular, the double branched-cover of a $(1,1)$-knot always admits a genus 2 Heegaard decomposition.

We can describe branched double-covers along a knot $K \subset Y$ in terms of Heegaard diagrams, as follows. Let $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w, z)$ be a doubly-pointed Heegaard diagram for $K \subset Y$. Let $f\colon \widetilde{\Sigma} \to \Sigma$ be the branched double-cover of Σ at the two points w and z, inherited from the branched double-cover $\Sigma(K) \to Y$. Notice that since the curves α_i and β_j bound disks in Y, eash of these curves will lift to two disjoint curves on $\widetilde{\Sigma}$. Indeed,

$(\widetilde{\Sigma}, f^{-1}(\boldsymbol{\alpha}), f^{-1}(\boldsymbol{\beta}), f^{-1}(w), f^{-1}(z))$ is a doubly-pointed Heegaard diagram for the branched double-cover of Y along K.

Exercise 3.4.1. *Consider the double branched cover of S^3 along the alternating torus knot $T_{2,2n+1}$. Pull back the Heegaard surface and the diagram of $T_{2,2n+1}$ from Figure 3.9 to $\Sigma(T_{2,2n+1})$ as above. Identify the Heegaard states in the genus 2 diagram of the double branched cover.*

3.4.1. Heegaard diagram for $\Sigma(K)$ from a diagram of K. A slightly different diagram can be presented as follows, see [**39**] for details. It has the advantage that the generators can be easily visualized, as being exactly the Kauffman state of the marked diagram of K we start our procedure with.

Consider a marked diagram (P, p) for the knot $K \subset S^3$. Let us also fix a spanning tree T for the black graph, and consider its dual spanning tree T^* for the white graph. (Recall that the two spanning trees have the property that their edges are all disjoint; indeed, there is an edge of T or T^* through every crossing of P.) An edge of $T \cup T^*$ instructs us how to take a resolution at the vertex at hand: pick the one which joins the two domains connected by the edge of T or T^*. It is not hard to see that in this way we will get a single unknot, with all black and all white domains merged by the chosen resolutions.

The Heegaard diagram of $\Sigma(K)$ will closely follow the construction described in Subsection 3.2.1. Indeed, consider the diagram in the plane, and viewing it as the image of an immersion to $\mathbb{R}^2 \subset \mathbb{R}^3$, take its tubular neighborhood, a solid genus g handlebody with g being one more than the crossing number of the knot diagram. The boundary of this handlebody will be our Heegaard surface Σ, and the $\boldsymbol{\alpha}$-curves are chosen exactly as in Subsection 3.2.1: as the intersections of Σ and the plane containing the diagram, with one component deleted.

The definition of the $\boldsymbol{\beta}$-curves is, however, slightly different, and rests on the choice of the spanning tree T. Firs of all take the meridian β_g at the marking p on the knot diagram; recall that the marking is on an edge neighbouring the unbounded domain. The further $\boldsymbol{\beta}$-curves in Subsection 3.2.1 were taken at the crossings by a simple rule (cf. Figure 3.4). Now we will modify the construction of these curves. The plane containing the diagram cuts Σ into two components, an upper (Σ_+) and a lower (Σ_-) one. In the upper component Σ_+ we choose the same curves as in Subsection 3.2.1, which then results two arcs at each crossing. We position them so that each arc has one of its endpoint in an intersection point of Σ with $T \cup T^*$ (where these trees are also viewed as part of the plane containing the diagram of the knot).

In the lower portion Σ_- we look for two further arcs at every crossing to complete the upper two arcs to provide the $\boldsymbol{\beta}$-circles. One of these arcs is easy to define: just project the portion of the edge of $T \cup T^*$ passing through the crossing to Σ_-. Deleting these arcs from Σ_- we get an annulus (corresponding to the fact that the resolution dictated by $T \cup T^*$ has a single component), and when cutting it further along the lower half of β_g (the meridian chosen at the marking), we get a disk as complement. Now there is (up to isotopy) a unique way to complete the arcs to closed curves in a manner that they stay disjoint, providing the $\boldsymbol{\beta}$-curves. For the resulting diagram we have the following (see [**39**] for the proof):

Proposition 3.4.2 ([**39**])**.** *When starting with a marked diagram (P, p) of the knot K, the above procedure provides a Heegaard diagram $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ corresponds to $\Sigma(K)$.* □

Exercise 3.4.3. *Applying the algorithm above, find Heegaard diagrams of the three-manifolds defined as branched double-covers of S^3, branched along*
(a) *the right-handed trefoil knot $T_{2,3}$,*

(b) *the Figure-8 knot and*

(c) *the knot 8_5 of the knot tables (the first alternating knot in Rolfsen's knot table* [**119**] *which is not two-bridge),*

(d) *along 8_{10}.*

Notice that since the diagram of $L \subset S^3$ given in Subsection 3.2.1 and the one derived above for $\Sigma(K)$ are identical in the upper half Σ_+, the intersections between the $\boldsymbol{\alpha}$- and the $\boldsymbol{\beta}$-curves are the same in the two diagrams. In particular, the Heegaard states in this Heegaard diagram of $\Sigma(K)$ can be naturally identified with the Kauffman states of the marked diagram (P, p) of then knot $K \subset S^3$.

Chapter 4

Symplectic geometry

Heegaard Floer homology is built upon constructions from symplectic geometry. In this chapter, we review some of the basic definitions from this subject, and refer the reader to [**76, 12**] for more thorough treatments. In Section 4.1 we recall the basic definitions and constructions of symplectic manifolds, while in Section 4.2 we discuss two types of submanifolds: symplectic and Lagrangian. Almost-complex structures are introduced in Section 4.3, and in Section 4.4 this notion is investigated for four-manifolds. In Section 4.5 we define a characteristic cohomology class associated to Lagrangian submanifolds, which will play an important role in the sequel.

4.1. Symplectic manifolds

We start by recalling the definition of a symplectic manifold:

Definition 4.1.1. *Let M be a $2n$-dimensional smooth manifold. A **symplectic form** on M is a smooth 2-form $\omega \in \Omega^2(M;\mathbb{R})$ that is closed ($d\omega = 0$) and that satisfies the non-degeneracy hypothesis that the n-fold wedge product $\overbrace{\omega \wedge ... \wedge \omega}^{n} \in \Omega^{2n}(M)$ vanishes nowhere. A **symplectic manifold** is a pair (M,ω), where ω is a symplectic form on M.*

Exercise 4.1.2. *Let V be a $2n$-dimensional real vector space equipped with an alternating 2-form $\omega_0 \in \Lambda^2(V^*)$ (i.e. a skew-symmetric bilinear form $\omega_0 \colon V \otimes V \to \mathbb{R}$). Show that the following two conditions are equivalent:*

(1) The n-fold wedge product of ω_0 with itself is non-zero, as an element of $\Lambda^{2n}(V^*) \cong \mathbb{R}$.

(2) For each non-zero vector $v \in V$, there is a vector $w \in V$ so that $\omega_0(v, w) \neq 0$.

In particular, for a symplectic manifold the top exterior power of ω can be viewed as a volume form for M, and hence inducing an orientation on M. We will always think of our symplectic manifolds as oriented with this preferred orientation.

Definition 4.1.3. *A symplectic manifold (M, ω) is called **exact** if ω is exact, that is, can be written as $\omega = d\alpha$ for some 1-form α.*

The *standard symplectic form* ω_{st} on $\mathbb{R}^{2n}$ (with coordinates $x_1, y_1, \ldots, x_n, y_n$) is given by

$$\omega_{st} = \sum_{i=1}^{n} dx_i \wedge dy_i.$$

This form is clearly non-degenerate. It is also closed; in fact, it is exact, as

$$\omega = d\left(\sum_i x_i dy_i\right).$$

Exercise 4.1.4. (a) *In what dimensions m can the sphere S^m be given the structure of a symplectic manifold?*

(b) *Show that a closed symplectic manifold cannot be exact.*

A diffeomorphism $\phi\colon (M, \omega) \to (M', \omega')$ between two symplectic manifolds is a *symplectomorphism* if $\phi^*(\omega') = \omega$. If there is such a symplectomorphism, then we say that (M, ω) and (M', ω') are *symplectomorphic*.

Symplectic manifolds arise in many contexts. Perhaps the first basic example is the cotangent bundle of any smooth manifold, which we describe in Example 4.1.5. We will use the following notation: if $f\colon X \to Y$ is a smooth map between manifolds, let $T_x f\colon T_x X \to T_{f(x)} Y$ denote the induced map of tangent spaces.

Example 4.1.5. *Let L be a real n-dimensional manifold. Its cotangent bundle T^*L is equipped with a tautological one-form λ, called the **Liouville form**, defined as follows. Consider the projection map $\pi\colon T^*L \to L$, and take its differential $T_\eta \pi\colon T_\eta(T^*L) \to T_{\pi(\eta)} L$ at any $\eta \in T^*L$. The map $T_\eta(T^*L) \to \mathbb{R}$, sending $v \in T_\eta(T^*L)$ to the evaluation $\eta(T\pi_\eta(v))$, gives a one-form $\lambda \in \Omega^1(T^*M)$ on $M = T^*L$. The above one-form is natural in the sense that if $f\colon L \to L'$ is a diffeomorphism and $f^*\colon T^*L' \to T^*L$ is its induced map on the contangent bundle, then $(f^*)^*(\lambda) = \lambda'$. The exact 2-form $-d\lambda$ on T^*L is clearly closed, and it is non-degenerate, as the local calculation given below shows. In conclusion, $(T^*L, -d\lambda)$ is a symplectic manifold.*

The Liouville form has the following explicit description in local coordinates. Suppose $L \cong \mathbb{R}^n$, with local coordinates $x_1, \ldots, x_n$. This induces a coordinate system on $T^*\mathbb{R}^n$ by the parameterization $\mathbb{R}^n \times \mathbb{R}^n \to T^*\mathbb{R}^n$ given by

$$(x_1, \ldots, x_n, y_1, \ldots, y_n) \to (x_1, \ldots, x_n, \sum_{i=1}^{n} y_i dx_i).$$

With respect to this trivialization, $\lambda = \sum y_i dx_i$.

A further class of examples for symplectic manifolds is furnished by two-manifolds: any volume form ω on a two-manifold is a symplectic form. These examples are not exact (Definition 4.1.3) if Σ is closed.

Another construction is given as follows:

Example 4.1.6. *If (M_1, ω_1) and (M_2, ω_2) are symplectic manifolds, form their product $M_1 \times M_2$, which has projection maps $\pi_i \colon M_1 \times M_2 \to M_i$ for $i = 1, 2$. For any two positive real numbers α, β, the 2-form $\alpha\pi_1^*(\omega_1) + \beta\pi_2^*(\omega_2)$ is a symplectic form on $M_1 \times M_2$.*

Many more non-exact examples are furnished by algebraic varieties. To give the construction, we will use some standard notation from several complex variables. (See, for example, [**42**, Chapter 0].) Consider $\mathbb{C}^n$ with its coordinates $(z_1, \ldots, z_n)$, where $z_i = x_j + iy_j$. The complex valued one-forms naturally split into a direct sum of complex-linear one-forms and complex anti-linear one-forms; $\Omega^1 \cong \Omega^{1,0} \oplus \Omega^{0,1}$, where $dz_j \in \Omega^{1,0}$ is written as $dz_j = dx_j + idy_j$ and $d\overline{z}_j \in \Omega^{1,0}$ is written as $d\overline{z}_j = dx_j - iy_j$. With respect to this splitting, we can write $d = \partial + \overline{\partial}$, where

$$\partial f = \sum_j \frac{1}{2}\left(\frac{\partial f}{\partial x_i} - i\frac{\partial f}{\partial y_j}\right) dz_j, \qquad \text{and} \qquad \overline{\partial} f = \sum_j \frac{1}{2}\left(\frac{\partial f}{\partial x_i} + i\frac{\partial f}{\partial y_j}\right) d\overline{z}_j$$

More generally, there is a splitting $\Omega^m \cong \bigoplus_{j+k=m} \Omega^{j,k}$, and the exterior derivative has components $\partial \colon \Omega^{j,k} \to \Omega^{j+1,k}$ and $\overline{\partial} \colon \Omega^{j,k} \to \Omega^{j,k+1}$.

Example 4.1.7. *The **Fubini Study form** ω_{FS} on $\mathbb{C}P^n$ is constructed as follows. Consider first $\mathbb{C}^{n+1}$ with coordinates $(z_0, \ldots, z_n)$, and let $|z| = \sqrt{\sum_{i=0}^n |z_i|^2}$. Define the 2-form on $\mathbb{C}^{n+1} \setminus 0$ by*

$$\omega_0 = \frac{i}{2\pi}\left(\frac{\sum_{j=0}^n dz_j \wedge d\overline{z}_j}{|z|^2} - \frac{\left(\sum_{j=0}^n \overline{z}_j dz_j\right) \wedge \left(\sum_{j=0}^n z_j d\overline{z}_j\right)}{|z|^4}\right).$$

This 2-form is invariant under rescaling and hence descends to a 2-form ω_{FS} over $\mathbb{C}P^n$. Since $\omega_0 = \frac{i}{2\pi}\partial\overline{\partial}\log|z|^2$, it follows that ω_{FS} is invariant under the action of $U(n+1)$. It is now a straightforward computation to verify that ω_{FS} is closed and non-degenerate. (See [**42**, Chapter 0.7]*)*

More generally, for any complex submanifold $X \subset \mathbb{C}P^n$, $\omega_{FS}|_X$ is a symplectic form; see Proposition 4.3.20 below.

4.2. Lagrangian and symplectic submanifolds

There are two particularly nice classes of submanifolds of symplectic manifolds:

Definition 4.2.1. *Let (M^{2n},ω) be a symplectic manifold and $W^m \subset M^{2n}$ be a submanifold. We say that $W^m \subset M^{2n}$ is a **symplectic submanifold** if the restriction of ω to W^m is a symplectic form on W^m. We say that $W^m \subset M^{2n}$ is **isotropic** if the restriction of ω to W^m vanishes identically. When $m = n$, an isotropic submanifold is called **Lagrangian**.*

Exercise 4.2.2. *Show that if $W^m \subset M^{2n}$ is isotropic, then $m \leq n$.*

The canonical example of a Lagrangian submanifold is the zero-section L of the cotangent bundle T^*L for any smooth manifold L, equipped with the symplectic structure of Example 4.1.5. In fact, Weinstein's Lagrangian tubular neighborhood theorem states that any Lagrangian submanifold L in a symplectic manifold admits a neighborhood that is symplectomorphic to a neighborhood of the zero section in T^*L; see for example [**76**, Theorem 3.31]. Analogous theorems exist for symplectic submanifolds [**76**, Theorem 3.30].

Remark 4.2.3. *It follows from the above neighbourhood theorem of Weinstein that the normal bundle NL of a Lagrangian submanifold L is isomorphic to T^*L. Indeed, this topological result can be seen directly: the map $f\colon NL \to T^*L$ defined on a normal vector $v \in NL$ by $v \mapsto \omega(v,\cdot)$ is a bundle map between equal dimensional vector bundles, and it is injective (hence an isomorphism) because L is Lagrangian.*

Exercise 4.2.4. *For (Σ_1,ω_1) and (Σ_2,ω_2) two oriented surfaces equip the product $\Sigma_1 \times \Sigma_2$ with the symplectic form $\pi_1^*\omega_1 + \pi_2^*\omega_2$. Find a basis of $H_2(\Sigma_1 \times \Sigma_2;\mathbb{Z})$ represented by embedded surfaces, each of which is either symplectic or Lagrangian.*

Further Lagrangian submanifolds can be constructed as follows:

Exercise 4.2.5. *Fix a smooth 1-form $\eta \in \Omega^1(L)$. The **graph** Γ_η **of** η is the submanifold of T^*L given as the subset $\{(p,\eta_p)\big|p \in L\}$. Show that the restriction of the Liouville form λ to Γ_η is identified with η via the projection $\Gamma_\eta \to L$; i.e. if π is the bundle map $T^*L \to L$, then $\pi^*(\eta) = \lambda|_{\Gamma_\eta}$. In particular, the graph of a closed one-form is a Lagrangian submanifold of $(T^*L,-d\lambda)$.*

Example 4.2.6. *Let (M^{2n}, ω) be a symplectic manifold, and let $\pi_i \colon M \times M \to M$ for $i = 1, 2$ denote the projection to the i^{th} factor. Then the 2-form $\Omega = \pi_1^*(\omega) - \pi_2^*(\omega)$ is a symplectic form (inducing an orientation on $M \times M$ that is $(-1)^n$ times the product orientation), and the diagonal $\Delta = \{(p, p) \mid p \in M\} \subset M \times M$ is a Lagrangian submanifold of $(M \times M, \Omega)$. More generally, if $\phi \colon (M, \omega) \to (M, \omega)$ is a symplectomorphism, then*

$$\Gamma_\phi = \{(p, \phi(p)) \mid p \in M\} \subset M \times M$$

is a Lagrangian submanifold of $(M \times M, \Omega)$.

Example 4.2.7. *The inclusion of $\mathbb{R}^{n+1} \to \mathbb{C}^{n+1}$ induces an embedding of $\mathbb{R}P^n \subset \mathbb{C}P^n$. With respect to the Fubini-Study form on $\mathbb{C}P^n$, the submanifold $\mathbb{R}P^n$ is Lagrangian.*

4.3. Symplectic manifolds and almost-complex structures

It is often convenient to endow symplectic manfiolds with an auxiliary structure, called an *almost-complex structure*, defined below. As motivation, we start with a more restrictive notion, that of a complex structure.

Definition 4.3.1. *A **complex n-manifold** M is a smooth manifold, equipped with distinguished local charts modeled on $\mathbb{C}^n$, whose transition functions are holomorphic. We say that this collection of local charts gives the underlying smooth manifold a **complex structure**.*

For a manifold with a complex structure, the tangent bundle TM is naturally a bundle of complex vector spaces; i.e. for each $m \in TM$, multiplication by i, defined on $T_m M \cong T_z \mathbb{C}^n$, is independent of the choice of coordinate chart around $m \in M$. Intrinsically, this multiplication by i can be thought of as a certain kind of endomorphism of the tangent bundle, formalized in the following definition:

Definition 4.3.2. *Given a complex n-manifold M, a submanifold $N^k \subset M$ is called a **complex submanifold** if it can be covered by open charts U modeled on $\mathbb{C}^n$ in M as in Definition 4.3.1, with the property that $N \cap U$ is modelled on $\mathbb{C}^k \otimes 0 \subset \mathbb{C}^n$.*

The above definition has the following natural generalization to the almost-complex case:

Definition 4.3.3. *Let (X, J) be an almost-complex manifold. A submanifold $Y \subset X$ is called a **J-holomorphic** if TY is preserved by J; i.e. $JT_yY = T_yY$ for each $y \in Y$.*

We explore now the connection between complex notions and symplectic ones.

Definition 4.3.4. *An **almost-complex structure** J on a smooth manifold X is a bundle automorphism $J\colon TX \to TX$ satisfying $J \circ J = -\operatorname{Id}_{TX}$.*

As noted above, a complex structure on X induces an almost-complex structure on X. But not every almost-complex structure arises in this manner: those which are are called *integrable*.

Remark 4.3.5. *To an almost-complex manifold, there is a naturally associated tensor, called the Nijenhuis tensor. A theorem of Newlander and Nirenbeg* [**93**] *states that an almost-complex structure is integrable exactly when its Nijenhuis tensor vanishes. In particular, if the manifold is (real) 2-dimensional, every almost-complex structure is integrable.*

Definition 4.3.6. *Let (M, ω) be a symplectic manifold. An almost-complex structure J on M is said to be **compatible** with ω if at each $p \in M$ and for $v, w \in T_pM$ we have $\omega(v, w) = \omega(Jv, Jw)$, and for each non-zero $v \in T_pM$ the value $\omega(v, Jv)$ is positive.*

The space of almost-complex structures on a manifold M is denoted $\mathcal{J}(M)$; it inherits its topology from the endomorphism space of the tangent bundle, which we endow with the C^∞ topology. The proof of the next result will be given after we analyze similar notions on linear spaces.

Theorem 4.3.7. *A symplectic manifold (M^{2n}, ω) admits a compatible almost-complex structure J. Moreover, for the given (M, ω) the ω-compatible almost-complex structures form a contractible space, hence J is unique up to homotopy.*

In view of Theorem 4.3.7, the tangent bundle of a symplectic manifold can be viewed as a complex vector bundle. The Chern classes of a symplectic manifold $c_i(M, \omega)$ are defined as $c_i(TM, J) \in H^{2i}(M; \mathbb{Z})$, for a compatible almost-complex structure J. Since any two choices are homotopic, the Chern classes are independent of the choice of the compatible J.

- The total Chern class $c(\mathbb{C}P^n) = \sum_{i=0}^{n} c_i(\mathbb{C}P^n)$ of the complex projective space $\mathbb{C}P^n$ is equal to $(1+h)^{n+1}$, where $h \in H^2(\mathbb{C}P^n; \mathbb{Z})$ is the generator which evaluates as 1 on the homology class of $\mathbb{C}P^1 \subset \mathbb{C}P^n$; see [**82**]. In particular, $c_1(\mathbb{C}P^n) = (n+1)h$.
- Fix an integer $d \geq 1$ and consider the degree d hypersurface

$$S_d = \{[z_1 : z_2 : z_3 : z_4] \in \mathbb{C}P^3 \mid \sum_i z_i^d = 0\} \subset \mathbb{C}P^3.$$

The Chern classes of S_d are given by the following formula: if $i\colon S_d \to \mathbb{C}P^3$ denotes the embedding and $x = i^*(h)$ for the above generator $h \in H^2(\mathbb{C}P^3;\mathbb{Z})$, then $c_1(S_d) = (4-d)x$ and $c_2(S_d) = (d^2-4d+6)x^2$; see [**34**]. In particular, the *K3-surface* S_4 has vanishing first Chern class.

- The (complex) n-dimensional torus $\mathbb{T}^n = \mathbb{C}^n/\mathbb{Z}^{2n}$ has trivial tangent bundle, hence for $i > 0$, $c_i(\mathbb{T}^n) = 0$.
- For a closed Riemann surface Σ of genus g, equipped with a symplectic form ν, the Chern class $c_1(\Sigma,\nu)$ equals the Euler class of the tangent bundle, and its integral over the fundamental cycle $[\Sigma]$ of Σ is given by

$$\langle c_1(\Sigma,\nu), [\Sigma]\rangle = 2-2g.$$

The compatibility condition from Definition 4.3.6 fits into the following framework. Consider the following three structures on a manifold M^{2n}: a Riemannian metric g, an almost-complex structure J, and a non-degenerate 2-form ω. There are compatibility conditions one can define for any two of those three pieces. We start our discussion with the linear case, that is, we study these structures first on vector spaces.

Definition 4.3.8. *Suppose that V is an even-dimensional vector space. A **complex structure** on V is a linear endomorphism J so that $J\circ J = -\mathrm{Id}$.*

Remark 4.3.9. *Note that this notion is different from the notion of "complex structure" on manifolds used earlier. To distinguish the above linear notion from the version for complex manifolds, sometimes one refers to the objects in Definition 4.3.8 as "linear complex structures". For example, an almost-complex structure on a manifold M is a (continuous) choice of linear complex structure on its tangent spaces.*

Definition 4.3.10.

- *A non-degenerate anti-symmetric bilinear form ω and a complex structure J on V are **compatible** if $\omega(Ju,Jv) = \omega(u,v)$ for all $u,v \in V$ and the symmetric form $g(u,v) = \omega(u,Jv)$ is positive definite; i.e. if $\omega(v,Jv) > 0$ for all $0 \neq v \in V$.*
- *A symmetric, positive definite form g and a complex structure J on V are said to be **compatible** if $g(u,v) = g(Ju,Jv)$. In this case, we can define an associated non-degenerate 2-form ω by*

$$\omega(u,v) = g(Ju,v). \tag{4.1}$$

- *A non-degenerate 2-form ω and a symmetric, positive definite form g are **compatible** if the endomorphism J characterized by Equation (4.1) is a complex structure.*

*A **compatible triple** is a triple (ω,J,g) where any two are compatible.*

Let $\mathbb{R}^{2n}$ be the vector space, equipped with the basis $\{e_i, f_i\}_{i=1}^n$. The *standard symplectic form* Ω_0 is the anti-symmetric bilinear form determined by

$$\Omega_0(e_i, e_j) = \Omega_0(f_i, f_j) = 0 \qquad \text{and} \qquad \Omega_0(e_i, f_j) = \delta_{i,j}. \tag{4.2}$$

The *standard complex structure* on $\mathbb{R}^{2n}$ in the above basis is determined by

$$J_0(e_i) = f_i \qquad J_0(f_i) = -e_i.$$

Let $\mathbb{J}_0$ be the matrix representing J_0 with respect to this basis. Let $\mathbb{G}_0$ denote the metric on $\mathbb{R}^{2n}$ for which $\{e_i, f_i\}_{i=1}^n$ is an orthonormal basis. It is easy to see that $(J_0, \Omega_0, \mathbb{G}_0)$ is a compatible triple; we call it the *standard compatible triple*.

If V^{2n} is any vector space equipped with a non-degenerate anti-symmetric bilinear form ω, then there is a choice of basis $\{e_i, f_i\}_{i=1}^n$ with respect to which ω satisfies the identities from Equation (4.2); i.e. there is an isomorphism of symplectic vector spaces

$$(V^{2n}, \omega) \cong (\mathbb{R}^{2n}, \Omega_0). \tag{4.3}$$

By pulling back $\mathbb{G}_0$ and J_0 along this identification, the symplectic structure ω can be extended to a compatible triple. In particular, any symplectic structure on a vector space admits a compatible almost-complex structure. (For further discussion about these structures, see [**12**].)

In fact,

Proposition 4.3.11. *A compatible triple (ω, J, g) on a $2n$-dimensional vector space V is uniquely determined by any two of its components.* □

Proof. This is a straightforward consequence of Equation (4.1): two of ω (assumed to be non-degenerate), g (assumed to be non-degenerate), and J determines the third. □

Remark 4.3.12. *The fact that any two components of a compatible triple uniquely determines the third can be formulated as the following identity for Lie groups. Let $(\mathbb{G}_0, \Omega_0, J_0)$ be the standard compatible triple on $\mathbb{R}^{2n}$, equipped with the basis $\{e_i, f_i\}_{i=1}^n$, and (as above) let $\mathbb{J}_0$ be the matrix representing J_0 in this basis. The symmetry groups of $\mathbb{G}_0$, Ω_0, and J_0, which are $O(2n)$, $Sp(2n)$ and $GL_n(\mathbb{C})$ respectively, can be specified as the subgroups $GL_{2n}(\mathbb{R})$:*

$$O(2n) = \{A \in GL_{2n}(\mathbb{R}) \mid AA^T = \mathrm{Id}\},$$
$$Sp(2n) = \{A \in GL_{2n}(\mathbb{R}) \mid A^T \mathbb{J}_0 A = \mathbb{J}_0\},$$
$$GL_n(\mathbb{C}) = \{A \in GL_{2n}(\mathbb{R}) \mid A\mathbb{J}_0 = \mathbb{J}_0 A\}.$$

The claimed identity, then, is that the intersection of any two of the above three subgroups equals their triple intersection (which in turn is the unitary group $U(n)$); i.e.

$$O(2n) \cap Sp(2n) = Sp(2n) \cap GL_n(\mathbb{C}) = GL_n(\mathbb{C}) \cap O(2n) = U(n),$$

cf. [**76**, Lemma 2.17].

Exercise 4.3.13. *Show that a 2-form ω and a symmetric positive definite form g on a vector space V are compatible if and only if the map J determined by $g(Jv, w) = \omega(v, w)$ preserves g, that is, $g(Jv, Jw) = g(v, w)$ for all $v, w \in V$.*

The following linear algebraic result is the crucial step in verifying Theorem 4.3.7.

Proposition 4.3.14. *(*[**76**, Proposition 2.48]*) The space of complex structures on a vector space V compatible with a given non-degenerate antisymmetric form ω is contractible.*

Proof. We can assume that (V, ω) is the standard symplectic vector space $(\mathbb{R}^{2n}, \Omega_0)$.

Consider the following map from the space $\mathfrak{M}(V)$ of symmetric positive definite forms on V to the space $\mathcal{J}_{\Omega_0}$ of Ω_0-compatible complex structures on V: for $g \in \mathfrak{M}(V)$ define $A \in GL_{2n}(\mathbb{R})$ as

$$\Omega_0(v, w) = g(Av, w)$$

for all $w \in V$. Since $\Omega_0(v, w) = -\Omega_0(w, v)$, it follows that for the g-adjoint A^* of A, we have that $A^* = -A$. Thus, $P = A^*A = -A^2$ is g-positive definite. This implies that P can be diagonalized with all positive eigenvalues λ_i. By taking the appropriate conjugate of the diagonal matrix with $\sqrt{\lambda_i}$ in the diagonal, we get the unique g-self-adjoint, g-positive definite symmetric matrix Q satisfying $P = Q^2$. We claim that A and Q commute. Indeed, if V_i is the eigenspace of P corresponding to the eigenvalue λ_i (or alternatively, V_i is the eigenspace of Q with eigenvalue $\sqrt{\lambda_i}$), then for $v \in V_i$ we have $Av \in V_i$ since

$$PAv = -A^3v = APv = A(\lambda_i v) = \lambda_i Av.$$

On V_i, however, Q and A commute, hence they commute on V.

Now we define the map $r\colon \mathfrak{M}(V) \to \mathcal{J}_{\Omega_0}$ by $r(g) = Q^{-1}A$. This map is obviously continuous, and since $r(g)^2 = Q^{-2}A^2 = -\operatorname{Id}$, it maps into the space of complex structures. Furthermore, since $r(g)^* = A^*(Q^*)^{-1} =$

$-AQ^{-1} = -r(g)$, we get that $r(g)^* r(g) = \mathrm{Id}$. This identity then shows that the complex structure $r(g)$ is compatible with Ω_0, since

$$\begin{aligned}\Omega_0(r(g)v, r(g)w) &= g(Ar(g)v, r(g)w) = g(r(g)Av, r(g)w) = \\ &= g(Av, r(g)^* r(g)w) = g(Av, w) = \Omega_0(v, w),\end{aligned}$$

and

$$\Omega_0(v, r(g)v) = g(Av, r(g)v) = g(-r(g)Av, v) = g(-Q^{-1}A^2 v, v) = g(Qv, v) > 0.$$

If Ω_0, g and J form a compatible triple, then $A = J$ and $Q = \mathrm{Id}$, hence $r(g) = J$. Let $t \in [0, 1]$ and define the map $f_t \colon \mathcal{J}_{\Omega_0} \to \mathcal{J}_{\Omega_0}$ by

$$f_t(J) = r((1-t)g_{\mathbb{J}_0} + t g_J)$$

where g_J is defined by the property that Ω_0, J and g_J form a compatible triple. Now $f_1 = \mathrm{Id}$ and $f_0 \equiv \mathbb{J}_0$ shows that we have a contraction of the space of Ω_0-compatible complex structures on V. □

It is sometimes convenient to weaken the compatibility condition between complex structures and anti-symmetric forms, as follows:

Definition 4.3.15. *Suppose that V is a vector space and ω is a non-degenerate anti-symmetric 2-form on V. The complex structure J on V is ω-**tame** if $\omega(v, Jv) > 0$ when $v \in V$ is non-zero.*

A complex structure J compatible with ω is also ω-tame. Unlike ω-compatibility, the ω-tameness condition on J is an open condition. Given ω and J so that J is ω-tame, we can construct an associated positive definite quadratic form

$$g(u, v) = \frac{1}{2}(\omega(u, Jv) + \omega(v, Ju)). \tag{4.4}$$

Proposition 4.3.16. *The space $\mathcal{J}^t_\omega$ of ω-tame complex structures on V is contractible.*

Proof. As before, we will work in the standard symplectic vector space $(\mathbb{R}^{2n}, \Omega_0)$, equipped with its basis $\{e_i, f_i\}_{i=1}^n$. The matrix of representing any endomorphism J of $\mathbb{R}^{2n}$ with respect to this basis can be written as $-\mathbb{J}_0 Z$ for some matrix Z. We say that $Z > 0$ if $\mathbb{G}_0(v, Zv) > 0$ for each $v \neq 0$; obviously, if $Z > 0$, then Z is invertible.

The endomorphism J is a complex structure if and only if $J^2 = \mathbb{J}_0 Z \mathbb{J}_0 Z = -\mathrm{Id}$, which is equivalent to $Z^{-1} = \mathbb{J}_0^{-1} Z \mathbb{J}_0$; and J is Ω_0-tame if, in addition, $Z > 0$. (Furthermore, it is Ω_0-compatible if we also have that Z is symmetric).

Let $\mathbb{I}$ denote the identity matrix. Note that if $Z > 0$, then $\mathrm{Id} + Z > 0$, so we can form $F(Z) = (\mathbb{I} - Z)(\mathbb{I} + Z)^{-1}$. Letting $\|M\| = \max\{\|M(v)\| \mid \|v\| = 1\}$, we have that

$$\|F(Z)\| = \frac{\|\mathbb{I} - Z\|}{\|\mathbb{I} + Z\|} \leq \frac{1 - \|Z\|}{1 - \|Z\|} = 1.$$

The equation $Z^{-1} = \mathbb{J}_0^{-1} Z \mathbb{J}_0$ transforms to $-F(Z) = \mathbb{J}_0^{-1} F(Z) \mathbb{J}_0$:

$$\mathbb{J}_0^{-1} F(Z) \mathbb{J}_0 = \mathbb{J}_0^{-1}(\mathbb{I} - Z)\mathbb{J}_0 \mathbb{J}_0^{-1}(\mathbb{I} + Z)^{-1}\mathbb{J}_0 = \mathbb{J}_0^{-1}(\mathbb{I} - Z)\mathbb{J}_0 (\mathbb{J}_0^{-1}(\mathbb{I} + Z)\mathbb{J}_0)^{-1} =$$

$$(\mathbb{I} - Z^{-1})(\mathbb{I} + Z^{-1})^{-1} = -F(Z).$$

Hence, F provides a homeomorphism between the space of Ω_0-tame complex structures and the space of matrices $\{W \mid \|W\| \leq 1, \mathbb{J}_0^{-1} W \mathbb{J}_0 = -W\}$. Since this latter set is a convex set, the contractibility claim follows at once. □

Exercise 4.3.17. *Show that if Z is symmetric, then $F(Z)$ is also symmetric. Combine this with the proof of Proposition 4.3.16, and give an alternative proof of Proposition 4.3.14.*

For a smooth manifold the notions of Definition 4.3.10 generalize to almost-complex structures, non-degenerate 2-forms and Riemannian metrics by requiring the same compatibility conditions fiberwise. In particular, we can define compatible triples for manifolds. Proposition 4.3.11 has the following consequence for manifolds:

Proposition 4.3.18. *A compatible triple on a smooth manifold M is uniquely determined by two pieces of data; that is, any two compatible of ω, J, or g can be completed to form a compatible triple.* □

We now prove Theorem 4.3.7:

Proof. [Proof of Theorem 4.3.7] An almost-complex structure is a section of the bundle $\mathrm{End}(TM)$; to be compatible with ω, the section is required to have the property that for every $p \in M$, $\sigma(p)$ lies in a contractible subset $\mathrm{End}(T_pM)$ for every p. So compatible almost-complex structures for (M, ω) are sections of a bundle with contractible fibers, implying the claim of the theorem. □

For a non-degenerate 2-form ω on M, we say that an almost-complex structure J is ω-tame if its restriction to each fiber is tame; i.e. $\omega_p(v, J_p v) > 0$ for each $p \in M$ and each non-zero $v \in T_pM$. Proposition 4.3.16, combined with the argument from the proof of Theorem 4.3.7, gives:

Theorem 4.3.19. *A symplectic manifold (M, ω) always admits ω-tame almost-complex structures and the space of ω-tame almost-complex structures is contractible.* □

Given (M, ω) and an ω-tame J, there is an associated Riemannian metric on M defined as in Equation (4.4) by

$$g_p(v, w) = \frac{1}{2}(\omega_p(v, J_p w) + \omega(w, J_p v)), \tag{4.5}$$

for each $p \in M$ and each $v, w \in T_p M$.

A metric g on a complex manifold (M, J) is a *Kähler metric* if it is compatible with J, and the 2-form $\omega(u, v) = g(Ju, v)$ is a symplectic form. (The 2-form ω defined above is always non-degenerate, so the Kähler condition is really $d\omega = 0$.) In this case the symplectic form ω is called a *Kähler form*. Not every complex manifold is Kähler. For example, the quotient of $\mathbb{C}^n \setminus \{0\}$ by the $\mathbb{Z}$-action induced by the map

$$(z_1, \ldots, z_n) \mapsto (2z_1, \ldots, 2z_n)$$

is a complex manifold, which is diffeomorphic to $S^1 \times S^{2n-1}$. For $n \geq 2$ this manifold has vanishing second cohomology, hence it admits no non-degenerate closed 2-form. The Kähler condition has strong topological implications: for example, the odd Betti numbers b_{2i+1} of a Kähler manifold are all even [**42**].

The Fubini-Study 2-form defined in Example 4.1.7 is a Kähler form equipping $\mathbb{C}P^n$ with a Kähler metric.

Proposition 4.3.20. *Any complex submanifold of a Kähler manifold is also Kähler.*

Proof. Let Y be a complex submanifold of a Kähler manifold (X, J, g, ω). Since Y is a complex submanifold, J preserves TY. It follows that the restriction of ω to Y is non-degenerate. The restriction of ω to Y is closed, since ω is closed on X. □

It follows that any complex submanifold of $\mathbb{C}P^n$ is Kähler. These submanifolds are called *smooth projective varieties*. Note that not every Kähler manifold is projective. (See for example [**88**, Chapter 2]; see also [**142**].)

There are also manifolds which are symplectic but not Kähler. In fact, the following *Kodaira-Thurston manifold* K [**134**] admits both a complex and a symplectic structure, but the two structures are not compatible. The four-manifold K is defined as the quotient of $T^3_{x,y,z} \times \mathbb{R}_t$ by a free $\mathbb{Z}$-action,

where $n \in \mathbb{Z}$ acts on (x, y, z, t) as

$$(x, y, z, t) \mapsto (x, y + nx, z, t + n).$$

The symplectic form $dx \wedge dy + dz \wedge dt$ descends to a symplectic form ω_K to K. The manifold has $b_1(K) = 3$, therefore it is not Kähler.

Exercise 4.3.21. *Let K be the Kodaira-Thurston manifold.* **(a)** *Compute the fundamental group of K, and show that $b_1(K) = 3$.*

(b) *Show that for an almost-complex structure J_1 on K compatible with the symplectic structure ω_K we have $c_1(K, J_1) = 0$ (i.e. $c_1(K, \omega_K) = 0$).*

(c) *Exhibit a complex structure J_2 on K with $c_1(K, J_2) = 0$. (Note that J_2 cannot be compatible with ω_K since $b_1(K)$ is odd.)*

In fact, there are many symplectic manifolds that are not Kähler. For example, for any finitely presented group G and integer $n \geq 2$, there is a symplectic $2n$-manifold X^{2n} with $\pi_1(X^{2n}) \cong G$ [**33**], whereas for Kähler manifolds, $b_1(X^{2n})$, which coincides with the rank of the abelianization of G, must be even.

A Riemann surface (that is, a two-dimensional smooth manifold equipped with a complex structure) is always Kähler, since for any compatible metric g the compatible non-degenerate 2-form ω is necessarily closed (as is any 2-form on a two-manifold). Moreover, according to a classical result of Riemann, a Riemann surface is always projective [**92**].

A symplectic manifold with a fixed (compatible) almost-complex structure is called an *almost-Kähler* manifold. In an almost-Kähler manifold (M, ω, J) a submanifold $\Sigma \subset M$ is a *J-holomorphic submanifold* if $T\Sigma$ is invariant under J, that is, $J(T\Sigma) = T\Sigma$.

Proposition 4.3.20 has the following immediate generalization:

Proposition 4.3.22. *In an almost-Kähler manifold, (M, ω, J) all J-holomorphic submanifolds are almost-Käler.*

□

Remark 4.3.23. *It is not true that all symplectic submanifolds of an almost-Kähler manifold (M, ω, J) are J-holomorphic. However, given a symplectic submanifold N of (M, ω), one can find a compatible almost-complex structure J on M for which N is J-holomorphic.*

For a J-holomorphic submanifold $\Sigma \subset (M, J)$ the tangent bundle TM restricted to Σ splits as a complex bundle

$$TM|_\Sigma = T\Sigma \oplus N\Sigma,$$

where $N\Sigma$ denotes the normal bundle of Σ. From the product formula for Chern classes (see for example [**82**, Chapter 14]), it follows that

$$c_1(TM|_\Sigma) = c_1(T\Sigma) + c_1(N\Sigma) \in H^2(\Sigma;\mathbb{Z}). \tag{4.6}$$

For example, if Σ is a J-holomorphic curve in an almost-complex four-manifold (M^4, J), and $[\Sigma]$ denotes the positve generator of $H_2(\Sigma;\mathbb{Z})$, then the evaluation of $c_1(T\Sigma)$ on Σ can be computed from the genus of Σ, while the evaluation of $c_1(N\Sigma)$ is equal to the self-intersection of the homology class represented by Σ, $[\Sigma]\cdot[\Sigma]$, giving the *adjunction formula*

$$\langle c_1(TM), [\Sigma]\rangle = 2 - 2g(\Sigma) + [\Sigma]\cdot[\Sigma]. \tag{4.7}$$

In particular, if $C \subset \mathbb{C}P^2$ is a smooth complex curve of degree $d > 0$, then since $c_1(\mathbb{C}P^2) = 3h$ Equation (4.7) (see [**42**, Chapter 2]) expresses the genus of C as

$$g(C) = \frac{(d-1)(d-2)}{2}.$$

If (M^{2n}, J) is a manifold equipped with an almost-complex structure, one can study J-holomorphic curves; these are maps $u\colon (\Sigma, j) \to (M, J)$ on a Riemann surface (Σ, j) with the property that the derivative of u is complex linear, that is, $J\circ du = du\circ j$. Gromov revolutionalized symplectic geometry by studying the moduli spaces of J-holomorphic curves in a symplectic manifold (M, ω), where J is chosen to be ω-compatible [**43**]. We return to these ideas in Chapter 6.

Let (M, J) be an almost-complex manifold. A submanifold $L \subset M$ is called *totally real* if for every $p \in L$ we have $J_p(T_pL) \cap T_pL = 0$. If (M, ω) is symplectic and $L \subset M$ is Lagrangian, then for any compatible almost-complex structure J the Lagrangian L is totally real.

Let (M^{2n}, J) be an almost-complex manifold, and suppose that $c\colon M \to M$ is an anti-holomorphic involution; i.e. $c^*(J) = -J$. Suppose furthermore that the fixed point set L of c is a smooth n-dimensional manifold. In this case L is totally real. For example, if $L \subset \mathbb{R}P^k$ is a smooth, real algebraic variety whose complexification $M^{2n} \subset \mathbb{C}P^k$ is also smooth, then L is totally real in M^{2n}. In fact, it is Lagrangian for the restriction of the Fubini-Study form to M^{2n}.

4.4. Almost-complex structures on four-manifolds

Suppose that X is a closed, connected, oriented, smooth four-manifold. The existence of an almost-complex structure on X can be easily determined

from the algebraic topology of X, according to the following theorem of Hirzebruch and Hopf.

To formulate the theorem, recall that a four-manifold X as above is equipped with an *intersection form* (see for example [**34**]), which is a non-degenerate bilinear form

$$Q_X \colon H^2(X;\mathbb{Z}) \otimes H^2(X;\mathbb{Z}) \to \mathbb{Z}.$$

In terms of algebraic topology, this form is given by the cup product pairing $\smile \colon H^2(X;\mathbb{Z}) \otimes H^2(X;\mathbb{Z}) \to H^4(X;\mathbb{Z})$, composed with the identification $H^4(X;\mathbb{Z}) \cong \mathbb{Z}$ induced by the orientation on X. Geometrically, the intersection form computes the oriented intersection number of the surfaces that are Poincaré dual to the in-coming two-dimensional cohomology classes. Sometimes we abbreviate $Q_X(\xi \otimes \xi)$ by ξ^2, and $Q_X(\xi \otimes \eta)$ by $\xi \cdot \eta$.

Let $\sigma(X)$ denotes the signature of the intersection form of X, and $\chi(X)$ denotes the Euler characteristic of X.

We state without proof the following celebrated *Hirzebruch signature theorem*, in the case of 4-manifolds, refering the interested reader to [**82**, Theorem 19.4] for a proof (of a more general version, for manifolds of dimension $4k$). For the statement, we use conventions from [**82**]; in particular the first Pontrjagin class $p_1(TX)$ is given by $p_1(TX) = -c_2(TX \otimes \mathbb{C})$.

Theorem 4.4.1 (Hirzebruch)**.** *Let X be a closed, smooth oriented four-manifold, then*

$$\langle p_1(TX), [X] \rangle = 3\sigma(X),$$

where $p_1(TX) \in H^4(X)$ is the first Pontrjagin class of the tangent bundle of X. □

Theorem 4.4.2 (Hirzebruch-Hopf [**51**])**.** *The closed, connected, oriented, smooth four-manifold X admits an almost-complex structure if and only if there is a cohomology class $h \in H^2(X;\mathbb{Z})$ such that*

$$h \equiv w_2(TX) \pmod 2 \qquad \text{and} \qquad h^2 = 3\sigma(X) + 2\chi(X). \tag{4.8}$$

Before the proof, recall the classification of complex 2-plane bundles over a four-manifold.

Lemma 4.4.3. *A complex 2-plane bundle $E \to X$ over a closed four-manifold X is determined by its first and second Chern classes $c_1(E)$ and $c_2(E)$. In addition, for any pair $c_1 \in H^2(X;\mathbb{Z})$ and $c_2 \in H^4(X;\mathbb{Z})$ there is a complex 2-plane bundle E over X with $c_1(E) = c_1$ and $c_2(E) = c_2$.*

Proof. Consider the restriction of $E \to X$ to $X \setminus \{pt.\}$. As $X \setminus \{pt.\}$ is homotopic to a three-complex, the restriction admits a nowhere vanishing section, hence $E|_{X\setminus\{pt.\}}$ splits as $L \oplus \underline{\mathbb{C}}$ with a line bundle L satisfying $c_1(L) = c_1(E)$. The extension of the bundle from $X \setminus \{pt.\}$ to X is determined by the clutching function $S^3 \to U(2)$, i.e. by an element of $\pi_3(U(2)) \cong \mathbb{Z}$. Indeed, this number determines the signed number of zeros of the extension of a nowhere vanishing section from $X \setminus \{pt.\}$ to X. The Poincaré dual of the zero-set is equal to the Euler class $e(E)$ of the bundle [**9**, Theorem 11.17], and the second Chern class $c_2(E)$ of a $U(2)$-bundle is equal to its Euler class (cf. [**82**, Chapter 14]). Consequently $U(2)$-bundles over four-manifolds with equal Chern classes are isomorphic.

As for any c_1 there is a line bundle with this given first Chern class, and the clutching function $S^3 \to U(2)$ can be chosen arbitrarily, we get a complex 2-plane bundle for all $(c_1, c_2) \in H^2(X;\mathbb{Z}) \times H^4(X;\mathbb{Z})$. □

We also have the following:

Lemma 4.4.4. *$SO(4)$-bundles over closed four-manifolds are uniquely characterized by their second Stiefel-Whitney class w_2, their first Pontrjagin class p_1 and their Euler class e.*

Proof. Indeed, over $X \setminus \{pt.\}$ an $\mathbb{R}^4$-bundle F splits as $V \oplus \underline{\mathbb{R}}$, and V is determined by $w_2(F)$. The extension of the bundle to X is determined by the clutching function $S^3 \to SO(4)$, an element of $\pi_3(SO(4)) \cong \mathbb{Z} \oplus \mathbb{Z}$. This pair of integers is determined by the first Pontrjagin and the Euler numbers. (See the proof of [**82**, Lemma 20.10], where this statement is proved for $X = S^4$. The case for general X follows from excision. See also Exercise 4.4.5 below.) □

Exercise 4.4.5. *Consider the standard dimensional representation of $SO(4)$. Its second exterior product Λ^2, splits (as a representation space) as a direct sum of two three-dimensional representations $\Lambda^2 = \Lambda^+ \oplus \Lambda^-$; and correspondingly, there is a Lie group homomorphism $SO(4) \to SO(3)_+ \times SO(3)_-$.*

(a) *Show that the above map is a double cover.*

(b) *Consider the identification*

$$\pi_3(SO(4)) \cong \pi_3(SO(3)_+) \oplus \pi_3(SO(3)_-) = \mathbb{Z} \oplus \mathbb{Z}.$$

Given $\phi \in \pi_3(SO(4))$, denote the components of this isomorphism $\deg_+(\phi)$ and $\deg_-(\phi)$. For a fixed $SO(4)$-bundle P and $\phi \in \pi_3(SO(4))$, let P_ϕ obtained from P by modifying P in a neighborhood of a point, using the clutching function by ϕ.

Verify that

$$p_1(P_\phi) = p_1(P) - 2(\deg_+(\phi)) + \deg_-(\phi))$$
$$e(P_\phi) = e(P) - (\deg_+(\phi) - \deg_-(\phi).$$

(c) *Use the above to conclude that not every triple* $H^2(M;\mathbb{Z}/2\mathbb{Z})\times H^4(M;\mathbb{Z})\times H^4(M;\mathbb{Z})$ *is realized as* (w_2, p_1, e) *for an* $SO(4)$*-bundle.*

Proof. [of Theorem 4.4.2] Let E be a $U(2)$-bundle, and let $E|_{\mathbb{R}}$ denote its underlying $SO(4)$-bundle. Then, we have the following relations between the Chern classes of E and the characteristic classes (i.e. the Stiefel-Whitney, Pontrjagin, and Euler classes) of $E|_{\mathbb{R}}$:

$$w_2(E|_{\mathbb{R}}) \equiv c_1(E) \pmod 2$$
$$p_1(E|_{\mathbb{R}}) = c_1(E)^2 - 2c_2(E)$$
$$e(E|_{\mathbb{R}}) = c_2(E).$$

Theorem 4.4.1 gives $\langle p_1(TX), [X]\rangle = 3\sigma(X)$; and also $\langle e(TX), [X]\rangle = \chi(X)$. Thus, if X admits an almost-complex structure, i.e. $TX \cong E|_{\mathbb{R}}$ for some $U(2)$-bundle E, then Equation 4.8 follows.

Conversely, if we choose $h \in H^2(X;\mathbb{Z})$ satisfying Equation (4.8), the $U(2)$-bundle E with $c_1(E) = h$ and $c_2(E) = e(TX)$, which exists by Lemma 4.4.3, has the property that $E|_{\mathbb{R}} \cong TX$, according to Lemma 4.4.4. This latter isomorphism now endows X with the desired almost-complex structure. □

Exercise 4.4.6. *Show that* $X_n = \#_n \mathbb{C}P^2$ *admits an almost-complex structure if and only if* n *is odd. (By a result of Taubes* [**133**]*, proved using the Seiberg-Witten invariants, the four-manifold* X_n *admits a symplectic structure if and only if* $n = 1$*.)*

For four-manifolds with boundary, we can formulate a relative variant of Theorem 4.4.2, as follows. An almost-complex structure J on the four-manifold X with boundary $\partial X = Y$ naturally induces an oriented two-plane field ξ over Y, which is the intersection of TY, thought of as a subbundle of $TX|_Y$, with $J(TY)$. Indeed, any pair (Y, ξ), where Y is a closed, oriented three-manifold and ξ is an oriented, 2-plane field in TY, can be realized in this way:

Proposition 4.4.7. *Suppose that* (Y, ξ) *is a closed three-manifold equipped with an oriented two-plane field. Then, there is a compact, oriented, almost-complex four-manifold* (X, J) *with* $\partial X = Y$ *such that* J *induces* ξ *on* Y*; in short,* $\partial(X, J) = (Y, \xi)$*.*

Proof. We check the claim for S^3 first. According to the discussion in Section 1.4, the space of oriented two-plane fields is parametrized by $\pi_3(S^2) \cong \mathbb{Z}$. Indeed, the Hopf fibration $S^3 \to S^2$ induces the two-plane field that generates all other oriented two-plane fields on S^3, via connected sums and orientation reversal.

Restricting the complex structure $\mathbb{CP}^2$ to $X_0 = \mathbb{CP}^2 \setminus D^4$, we obtain an almost-complex structure on X_0 which induces the Hopf fibration on $S^3 = \partial X_0$. Boundary connected sums of this example (and its orientation reversed version) verifies the statement for all two-plane fields on S^3.

For (Y, ξ) consider a simply connected four-manifold X with $\partial X = Y$. We argue that the spin^c structure can be extended over X. To see why, note first that there is some spin^c structure over X – any oriented four-manifold admits a spin^c structure, by Theorem 32.2.7 from Appendix 32. Consider the difference $a = \mathbf{t}_\xi - \mathfrak{s}_0|_Y \in H^2(Y;\mathbb{Z})$. Observe that since $H_1(X;\mathbb{Z}) \cong H^3(X,Y;\mathbb{Z}) = 0$, the map $H^2(X;\mathbb{Z}) \to H^2(Y;\mathbb{Z})$ induced by the embedding is onto. Choosing any extension $\tilde{a} \in H^2(X;\mathbb{Z})$ of a to X, the spin^c structure $\mathfrak{s} = \mathfrak{s}_0 + \tilde{a}$ extends $\mathbf{t}_\xi$.

Let L be the line bundle with $c_1(\mathfrak{s})$. Since that $c_1(\mathfrak{s})$ is an integral lift of $w_2(TX)$, it follows that we have an isomorphism of bundles $\mathbb{C} \oplus L \cong TX$. (This is follows as in Lemma 4.4.4, noting now that since X has boundary, it is homotopy equivalent to a 3-complex, so the four-dimensional obstructions p_1 and e vanish.) The bundle isomorphism $\underline{\mathbb{C}} \oplus L \cong TX$ endows X with an almost-complex structure J_0.

The almost-complex structure J_0 on X induces a two-plane field ζ on Y which differs from ξ by the action of $\mathbb{Z}$; i.e. ζ and ξ differ by the connected sum of S^3 equipped with an appropriate two-plane field η. Therefore the boundary connected sum of (X, J) with the almost-complex four-manifold inducing η on its boundary givesthe desired almost-complex four-manifold. □

The identity $c_1^2(X, J) = 3\sigma(X) + 2\chi(X)$ for closed almost-complex four-manifolds can be used to define invariants of two-plane fields on closed three-manifolds as follows.

Given a three-manifold Y, equipped with an oriented two-plane field ξ Proposition 4.4.7 gives a compact, oriented almost-complex four-manifold (X, J) with $\partial X = Y$ and J inducing ξ on Y. Suppose that $c_1(\xi) \in H^2(Y;\mathbb{Z})$ is a torsion element; i.e. there is an integer n so that $nc_1(\xi) = 0$ in $H^2(Y;\mathbb{Z})$ or, equivalently, the image of $c_1(\xi)$ in $H^2(Y;\mathbb{Q})$ vanishes. For such an element, we can still define the square $c_1(\xi)^2$, though that quantity now is *a priori* a rational number. To define it, consider the image of $c_1(\xi)$

in $H^2(X, \partial X; \mathbb{Q})$, and let $c_1(\xi)^2$ denote the evaluation of $c_1(\xi) \cup c_1(\xi) \in H^4(X, \partial X; \mathbb{Q})$ against the fundamental cycle $H_4(X, \partial X; \mathbb{Q}) \cong \mathbb{Q}$ coming from the orientation of X. This construction has the following more concrete formulation. If $n \cdot c_1(\xi) = 0$ in $H^2(Y; \mathbb{Z})$, then $nc_1(X, J) \in H^2(X; \mathbb{Z})$ lifts to an element (which we continue to denote $n \cdot c_1(X, J)$) in $H^2(X, \partial X; \mathbb{Z})$ under the map $H^2(X, \partial X; \mathbb{Z}) \to H^2(X; \mathbb{Z})$; let

$$c_1^2(X, J) = \frac{1}{n^2} Q_{X,\partial X}(nc_1(X, J), nc_1(X, J)) \in \mathbb{Q} \tag{4.9}$$

Now define the invariant $d_3(\xi)$ — sometimes called the *Hopf invariant* or *three-dimensional invariant* of ξ – by the formula

$$d_3(\xi) = \frac{1}{4}\left(c_1^2(X, J) - 3\sigma(X) - 2\chi(X) + 2\right). \tag{4.10}$$

Proposition 4.4.8. *Suppose that (Y, ξ) is a closed three-manifold with an oriented two-plane field, and assume furthermore that $c_1(\xi)$ is torsion. Then the value $d_3(\xi)$ is independent of the chosen almost-complex four-manifold (X, J).*

Proof. Compute $d_3(\xi)$ using (X_1, J_1) and (X_2, J_2). Fix the almost-complex four-manifold (Z, J) with $\partial(Z, J) = -(Y, \xi)$, that is, $\partial Z = -Y$ and the orientation of $-\xi$ is the opposite of ξ (equivalently, if ξ is the oriented orthogonal of the nowhere vanishing vector field v, then $-\xi$ is the oriented orthogonal of $-v$). Then (Z, J) can be glued to both (X_1, J_1) and (X_2, J_2) to get the almost-complex four-manifolds $(X_1 \cup Z, \tilde{J}_1)$ and $(X_2 \cup Z, \tilde{J}_2)$. The quantities σ and χ are additive under this gluing operation; since $c_1(J_i)|_Y$ vanishes, c_1^2 is also additive under this summing operation. By Theorem 4.4.2, for $X_i \cup Z$ we have $c_1^2(X_i \cup Z, \tilde{J}_i) - 3\sigma(X_i \cup Z) - 2\chi(X_i \cup Z) = 0$, hence the claim follows at once. □

If $c_1(\xi) \in H^2(Y; \mathbb{Z})$ is not a torsion class, Equation (4.9) (and hence Equation (4.10)) does not make sense, since $c_1^2(X, J)$ is not well defined. Nonetheless, we can make sense of a difference between Hopf invariants, via the following construction.

Let ξ_1 and ξ_2 be two two-plane fields on Y with $c_1(\xi_1) = c_1(\xi_2) = b$. There is a compact four-manifold X with $\partial X = Y$ admitting two almost-complex structures J_1 and J_2 such that $\partial(X, J_i) = (Y, \xi_i)$. Since J_1 and J_2 induce the same spin^c structure on $\partial X = Y$, $c_1(X, J_1) - c_1(X, J_2)$ is represented by a relative two-dimensional cohomology class. Since $c_1(X, J_1)$ and $c_1(X, J_2)$ are characteristic cohomology classes, both $\frac{1}{2}(c_1(X, J_1) + c_1(X, J_2))$ and $\frac{1}{2}(c_1(X, J_1) - c_1(X, J_2))$ are represented by integral cohomology classes.

Thus, we can define the *relative Hopf invariant* by the expression:

(4.11)
$$d_3(\xi_1,\xi_2) = \left(\frac{1}{2}(c_1(X,J_1) - c_1(X,J_2))\right)\cdot\left(\frac{1}{2}(c_1(X,J_1) + c_1(X,J_2)\right) \in \mathbb{Z}/\mathfrak{d}(b)\mathbb{Z},,$$

where $\mathfrak{d}(b)$ is the divisibility of b; see Definition 1.4.7.

Proposition 4.4.9. *Fix $b \in H^2(Y;\mathbb{Z})$, and suppose that ξ_1 and ξ_2 are two two-plane fields over Y with $c_1(\xi_i) = b$. Then, the relative Hopf invariant $d_3(\xi_1,\xi_2)$, as specified in Equation (4.11), gives a well-defined element of the cyclic group $\mathbb{Z}/\mathfrak{d}(b)\mathbb{Z}$. In the case where b is torsion,*

$$d_3(\xi_1,\xi_2) = d_3(\xi_1) - d_3(\xi_2). \tag{4.12}$$

Proof. As explained above, $\frac{1}{2}(c_1(X,J_1) - c_1(X,J_2))$ can be represented by an element $A \in H^2(X,Y;\mathbb{Z})$; $\frac{1}{2}(c_1(X,J_1) + c_1(X,J_2))$ by an element of $B \in H^2(X;\mathbb{Z})$, so we can pair these to get an element of $H^4(X,\partial X;\mathbb{Z}) \cong \mathbb{Z}$. However, this pairing depends on the choice of the lift A of $A_0 = \frac{1}{2}(c_1(X,J_1) - c_1(X,J_2)) \in H^2(X;\mathbb{Z})$. Any other lift of A_0 is of the form $A' = A + \delta C$ for some $C \in H^1(Y;\mathbb{Z})$, and

$$\begin{aligned}\langle A' \smile B - A \smile B, [X,\partial]\rangle &= \langle \delta(C) \smile B, [X,\partial]\rangle \\ &= \langle C \smile B|_{\partial X}, [Y]\rangle = \langle C \smile b, [Y]\rangle.\end{aligned}$$

Thus, $d_3(\xi_1,\mathbf{x}_2)$, thought of as an element of the cyclic group $\mathbb{Z}/H^1(Y;\mathbb{Z})\cdot b$, is independent of this choice. By Poincaré duality, $H^1(Y;\mathbb{Z})\cdot b = \mathfrak{d}(b)\cdot\mathbb{Z}$.

Equation (4.12), which is the motivation for the above definition of $d_3(\xi_1,\mathbf{x}_2)$, is straightforward to verify:

$$\begin{aligned}&d_3(\xi_1) - d_3(\xi_2)\\ &\quad= \frac{1}{4}(c_1^2(X,J_1) - 3\sigma(X) - 2\chi(X) + 2) - \frac{1}{4}(c_1^2(X,J_2) - 3\sigma(X) - 2\chi(X) + 2)\\ &\quad= \frac{1}{4}(c_1(X,J_1) - c_1(X,J_2)) \cup (c_1(X,J_1) + c_1(X,J_2)).\end{aligned}$$

□

Exercise 4.4.10. *Show that for any two-plane ξ field over Y, $\mathfrak{d}(c_1(\xi))$ is even.*

4.5. The Maslov class

The *Maslov class* is a distinguished one-dimensional cohomology class in the the space of Lagrangian subspaces, which will play a key role in our constructions; see [**3**]. We describe this class in this section and give some of its properties.

Definition 4.5.1. *Let (V^{2n}, ω) be a symplectic vector space. The linear subspace $\Lambda \subset V^{2n}$ is **Lagrangian** if $\dim \Lambda = n$ and $\omega|_\Lambda = 0$. The **Lagrangian Grassmannian** $\mathcal{L}Gr(V, \omega)$ is the space of Lagrangian subspaces of (V^{2n}, ω).*

Since two equal dimensional symplectic vector spaces (V_1, ω_1) and (V_2, ω_2) are symplectic isomorphic, it follows that the associated Lagrangian Grassmannians $\mathcal{L}Gr(V_1, \omega_1)$ and $\mathcal{L}Gr(V_2, \omega_2)$ are diffeomorphic. The Lagrangian Grassmannian $\mathcal{L}Gr(\mathbb{R}^{2n}, \Omega_0)$ of the standard symplectic vector space $(\mathbb{R}^{2n}, \Omega_0)$ will be simply denoted by $\mathcal{L}Gr(n)$.

The Lagrangian Grassmannian $\mathcal{L}Gr(V, \omega)$ is identified with the quotient space $U(n)/O(n)$ as follows. Fix an ω-compatible complex structure J on V and let g be the induced positive definite symmetric form. A Hermitian form on V is specified by $\langle v, w\rangle = g(v, w) + i\omega(v, w)$. Any n-dimensional subspace Λ of V^{2n} can be given a g-orthonormal basis $e_1, \ldots, e_n$. Clearly, Λ is Lagrangian if and only if for all $i, j \in \{1, \ldots, n\}$, $\omega(e_i, e_j) = 0$, i.e. $e_1, \ldots, e_n$ is a unitary orthonormal basis for V. Such bases are parametrized by elements of $U(n)$. Two orthonormal bases specify the same subspace Λ if and only if they can be transformed into one another by an element of $O(n)$. Thus, there is a diffeomorphism $\mathcal{L}Gr(V, \omega) \cong U(n)/O(n)$. In particular $\mathcal{L}Gr(V, \omega)$ is a compact manifold, and since $\dim U(n) = n^2$ and $\dim O(n) = \frac{1}{2}n(n-1)$, the dimension of $\mathcal{L}Gr(V, \omega)$ is $\frac{1}{2}n(n+1)$.

Remark 4.5.2. *The choices made in identifying $\mathcal{L}Gr(V, \omega)$ with the homogeneous space $U(n)/O(n)$ are: (1) an identification of (V, ω) with the standard symplectic space $(\mathbb{R}^{2n}, \omega_{st})$; and (2) a choice of Lagrangian subspace $\Lambda_0 \in \mathcal{L}Gr(V, \omega)$.*

The first choice allows us to choose use $\mathbb{J}_0$ as the compatible almost-complex structure in the identification. This specifies the identification $\mathcal{L}Gr(V, \omega) \cong U(n)/O(n)$ up to left translation by $U(n)$. That indeterminacy in turn is eliminated by requiring Λ_0 to correspond to the identity coset.

Example 4.5.3. *It is easy to see that $\mathcal{L}Gr(1) \cong S^1$. The space $\mathcal{L}Gr(2)$ can be described as follows. Given two-dimensional subspace W^2 in a four-dimensional vector space V^4, there is 2-form $\psi_W \in \Lambda^2(V^*)$, so that W^2 is*

$$\ker \psi = \{u \in V \mid \psi(u, v) = 0 \text{ for all } v \in V\}$$

The 2-form ψ_W determining W is unique up to multiplication by a non-zero element of $\mathbb{R}$, hence ψ_W is well-defined as an element of the projectivized space $\mathbb{P}(\Lambda^2(V^*))$. Indeed, a 2-form $\psi \in \Lambda^2(V^*)$ defines a two-dimensional subspace in this way if $\psi \neq 0$ and $\psi \wedge \psi = 0$. Suppose now that ω is a non-degenerate 2-form on V^4. The two-dimensional subspace $\Lambda \subset V$ is

Lagrangian if $\psi_\Lambda \wedge \omega = 0$. *Therefore* $\mathcal{L}Gr(V)$ *can be described as*

$$\{\psi \in \mathbb{P}(\Lambda^2(V^*)) \mid \psi \wedge \psi = \psi \wedge \omega = 0\}.$$

Choose a basis $\{e_1, \dots, e_4\}$ *for* V, *so that* $\{e_i^* \wedge e_j^*\}$ *is a basis for* $\Lambda^2(V^*)$. *For the non-degenerate form* $\omega = e_1^* \wedge e_2^* + e_3^* \wedge e_4^*$, *we get that* $\psi = \sum_{i<j} y_{ij} \cdot e_i^* \wedge e_j^*$ *corresponds to a Lagrangian subspace in* V *if and only if*

$$y_{12}y_{34} - y_{13}y_{24} + y_{14}y_{23} = 0, \qquad y_{12} + y_{34} = 0. \tag{4.13}$$

Eliminating y_{34} *and introducing new coordinates* $z_1 = y_{12}, z_2 = \frac{1}{2}(y_{13} + y_{24}), z_3 = \frac{1}{2}(y_{13} - y_{24}), z_4 = \frac{1}{2}(y_{14} + y_{23}), z_5 = \frac{1}{2}(y_{14} - y_{23})$, *we identify*

$$\mathcal{L}Gr(2) = \{[z_1 : z_2 : z_3 : z_4 : z_5] \in \mathbb{R}P^4 \mid z_1^2 + z_2^2 + z_4^2 = z_3^2 + z_5^2\};$$

i.e. $\mathcal{L}Gr(2)$ *is identified with* $S^2 \times S^1/(\mathbb{Z}/2\mathbb{Z})$, *where* $\mathbb{Z}/2\mathbb{Z}$ *acts by the antipodal map on both factors. Since this action preserves orientation on* S^1, *but reverses it on* S^2, *the quotient three-manifold is non-orientable.*

The determinant gives a homomorphism $\det\colon U(n) \to S^1$, that sends all of $O(n)$ to $\{\pm 1\}$. Thus $\det^2$ gives a well-defined map $\det^2\colon U(n)/O(n) \to S^1$. When $n = 1$, $\det^2$ induces an isomorphism $\mathcal{L}Gr(1) \cong S^1$. More generally:

Proposition 4.5.4. *There is an isomorphism* $H^1(\mathcal{L}Gr(n); \mathbb{Z}) \cong \mathbb{Z}$, *and the generator is the pull-back of the generator of* $H^1(S^1; \mathbb{Z})$ *by* $\det^2$.

Proof. Consider the homotopy long exact sequence of the fibration $O(n) \to U(n) \to \mathcal{L}Gr(n)$; cf. [**129**, Section 17.11]. We have

$$\pi_1(O(n)) \xrightarrow{f} \pi_1(U(n)) \longrightarrow \pi_1(\mathcal{L}Gr(n)) \xrightarrow{g} \pi_0(O(n)) \cong \mathbb{Z}/2\mathbb{Z}.$$

Since $\pi_1(U(n)) \cong \mathbb{Z}$ and, for $n > 2$, $\pi_1(O(n)) \cong \mathbb{Z}/2\mathbb{Z}$, it follows that f is the zero map when $n > 2$. When $n = 2$, a straightforward computation shows that f factors through $\pi_1(SU(2)) \subset \pi_1(U(2))$, so $f = 0$ again. Consider the closed path $A\colon [0, 1] \to \mathcal{L}Gr(n)$ given by $\theta \mapsto A(\theta)$, where $A(\theta)$ is the $n \times n$ diagonal matrix with $a_{1,1}(\theta) = e^{i\pi\theta}$ and $a_{j,j}(\theta) = 1$ for $j = 2, \dots, n$. Clearly, $g(A)$ is the generator of $\pi_0(O(n))$, while $A * A$ is the image of the generator of $\pi_1(U(n))$. Thus, $\pi_1(\mathcal{L}Gr(n)) \cong \mathbb{Z}$; and so $H_1(\mathcal{L}Gr(n); \mathbb{Z}) \cong H^1(\mathcal{L}Gr(n); \mathbb{Z}) \cong \mathbb{Z}$. Clearly, $\det^2(A)$ is the generator of $\pi_1(S^1)$, so the second statement follows. □

Definition 4.5.5. *The above generator* μ_n *of* $H^1(\mathcal{L}Gr(n); \mathbb{Z}) \cong \mathbb{Z}$ *is called the* ***universal Maslov class***.

4.5.1. The relative Maslov class. In a related direction, given a Lagrangian submanifold $L \subset (M, \omega)$ in the symplectic manifold (M, ω), there is a relative two-dimensional cohomology class $\mu_L \in H^2(M, L; \mathbb{Z})$ called the *relative Maslov class*, defined as follows. Choose an almost-complex structure J compatible with ω on M^{2n}, and consider the corresponding complex line bundle $\det_{\mathbb{C}}(M) = \Lambda^n_{\mathbb{C}} T_{\mathbb{C}} M$. Over $L \subset M$ this line bundle comes with a real line subbundle, $\det_{\mathbb{R}}(L) = \Lambda^n_{\mathbb{R}} TL$. The manifold L is orientable if and only if $\det_{\mathbb{R}}(L)$ is the trivial line bundle; but in any case, $\det_{\mathbb{R}}(L) \otimes_{\mathbb{R}} \det_{\mathbb{R}}(L)$ is a trivial line bundle, equipped with a canonical trivialization. Thus, $\det_{\mathbb{C}}(M) \otimes_{\mathbb{C}} \det_{\mathbb{C}}(M)$ is a complex line bundle with a specified trivialization over L. The Maslov class $\mu_L \in H^2(M, L; \mathbb{Z})$ is the relative first Chern class of $\det_{\mathbb{C}}(M) \otimes_{\mathbb{C}} \det_{\mathbb{C}}(M)$, relative to the specified trivialization $\det_{\mathbb{R}}(L) \otimes_{\mathbb{R}} \det_{\mathbb{R}}(L)$ over L. (See Section 31.1.)

In particular, if F is an oriented surface-with-boundary, and $u\colon (F, \partial F) \to (M, L)$ is a smooth map, the evaluation $\langle \mu_L, [F, \partial F] \rangle$ is computed by taking any generic section s of $u^*(\det_{\mathbb{C}}(M) \otimes_{\mathbb{C}} \det_{\mathbb{C}}(M))$, so that $s(p)$ is a non-zero element in $u^*(\det_{\mathbb{R}}(L) \otimes_{\mathbb{R}} \det_{\mathbb{R}}(L))_p$ for all $p \in \partial F$, and counting the zeros of s with sign.

If L is an orientable manifold, $\mu_L \in H^2(M, L; \mathbb{Z})$ is divisible by 2: an orientation on L trivializes $\det_{\mathbb{R}}(L)$, and hence specifies a relative first Chern class $\det_{\mathbb{C}}(M)$ relative to this trivialization of $\det_{\mathbb{R}}(L)$, which is clearly half the relative Maslov class.

Example 4.5.6. *As an example, consider the equator $L \subset S^2$, with the standard symplectic form on S^2. Let $S^2 \setminus L = D_1 \cup D_2$. The evaluation $\langle \mu_L, [\overline{D}_i, L] \rangle = 2$ for $i = 1, 2$.*

The above example has the following analogue when $n > 1$:

Proposition 4.5.7. *Consider the Lagrangian submanifold*

$$\mathbb{R}P^n = \{[x_0 : \ldots : x_n] \in \mathbb{C}P^n \mid x_i \in \mathbb{R}\}.$$

When $n > 1$, $H^2(\mathbb{C}P^n, \mathbb{R}P^n; \mathbb{Z}) \cong \mathbb{Z}$, and $\mu_{\mathbb{R}P^n}$ is $n+1$ times the generator.

Proof. By the cohomology long exact sequence of the pair $(\mathbb{C}P^n, \mathbb{R}P^n)$, $H^2(\mathbb{C}P^n, \mathbb{R}P^n; \mathbb{Z}) \cong \mathbb{Z}$. We wish to show that $H_2(\mathbb{C}P^n, \mathbb{R}P^n; \mathbb{Z}) \cong \mathbb{Z}$, as well. The long exact sequence shows

$$\mathbb{Z} \cong H^2(\mathbb{C}P^n; \mathbb{Z}) \xrightarrow{f} H^2(\mathbb{R}P^n; \mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z} \to H^3(\mathbb{C}P^n, \mathbb{R}P^n; \mathbb{Z}) \to H^3(\mathbb{C}P^n; \mathbb{Z}) = 0.$$

To see that f is surjective, note that the map $\mathbb{Z}/2\mathbb{Z} \cong H^2(\mathbb{C}P^n; \mathbb{Z}/2\mathbb{Z}) \to H^2(\mathbb{R}P^n; \mathbb{Z}/2\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z}$ is an isomorphism. This can be seen by intersecting the Poincaré dual $\{z_1 + iz_2 = 0\}$ of the generator $H^2(\mathbb{C}P^n; \mathbb{Z}/2\mathbb{Z})$ with

$\mathbb{R}P^n$, to get $\mathbb{R}P^{n-2} \subset \mathbb{R}P^n$, which is Poincaré dual to the generator of $H^2(\mathbb{R}P^n; \mathbb{Z}/2\mathbb{Z})$. It follows that the above map f is surjective as well; so $H^3(\mathbb{C}P^n, \mathbb{R}P^n; \mathbb{Z}) = 0$. By the universal coefficient theorem for cohomology, this implies that $H_2(\mathbb{C}P^n, \mathbb{R}P^n; \mathbb{Z}) \cong \mathbb{Z}$.

A generator of $H_2(\mathbb{C}P^n, \mathbb{R}P^n; \mathbb{Z}) \cong \mathbb{Z}$ can be specified as follows. Consider $\mathbb{C}P^1 = \{[z_0 : z_1 : 0 : \ldots : 0]\} \subset \mathbb{C}P^n$, and take the closure of one of the two components of $\mathbb{C}P^1 \setminus \mathbb{R}P^1$, where $\mathbb{R}P^1 = \mathbb{C}P^1 \cap \mathbb{R}P^n$. In fact, the quadratic hypersurface $Q = \{\sum_{i=0}^n z_i^2\} \subset \mathbb{C}P^n$ is disjoint from $\mathbb{R}P^n$ and intersects $\mathbb{C}P^1$ in the two points $[\pm i : 1 : \ldots : 0]$, one in each disk component of $\mathbb{C}P^1 \setminus \mathbb{R}P^1$. Since both intersections are positive, we get that the two disks $\overline{D}_1$ and $\overline{D}_2$ represent the same relative homology class in $H_2(\mathbb{C}P^n, \mathbb{R}P^n; \mathbb{Z})$. Consider any section of $\det_{\mathbb{C}} \otimes_{\mathbb{C}} \det_{\mathbb{C}}$ over $\overline{D}_1$ and $\overline{D}_2$ extending the trivialization supplied by $\det_{\mathbb{R}}(\mathbb{R}P^n) \otimes_{\mathbb{R}} \det_{\mathbb{R}}(\mathbb{R}P^n)$ over $\mathbb{R}P^1$; these sections fit together to give a section σ of $\det_{\mathbb{C}} \otimes_{\mathbb{C}} \det_{\mathbb{C}}$ over $\mathbb{C}P^1$. Since there are $2n+2$ zeros of σ, $\langle \mu_{\mathbb{R}P^n}, [\overline{D}_1, \partial\overline{D}_1] \rangle + \langle \mu_{\mathbb{R}P^n}, [\overline{D}_2, \partial\overline{D}_2] \rangle = 2n+2$, and from $[\overline{D}_1, \partial\overline{D}_1] = [\overline{D}_2, \partial\overline{D}_2]$ we conclude that $\langle \mu_{\mathbb{R}P^n}, [\overline{D}_1, \partial\overline{D}_1] \rangle = n+1$. □

The universal Maslov class and the relative Maslov class of a Lagrangian are related, as follows. Suppose that $L \subset (M^{2n}, \omega)$ is a Lagrangian submanifold, and suppose $u\colon (D, \partial D) \to (M, L)$ is a smooth map on the disk D. The pullback $u^*(TM)$ over the disk D is a trivial bundle of symplectic vector spaces. We wish to choose a trivialization of this bundle, so that $(u|_{\partial D})^*(TL)$ can be thought of as a loop γ_L of Lagrangians in a single symplectic vector space (V, ω).

A natural trivialization is supplied by *parallel transport*; we make a brief digression to recall this notion. Suppose that M is a smooth manifold. A *connection* on the tangent bundle of M is an operator taking pairs of vector fields X and Y to a new vector field, denoted $\nabla_X Y$, which is bilinear over the two factors, and satisfies the further property that for any smooth function f on M

$$\nabla_{fX} Y = f \nabla_X Y \quad \text{and} \quad \nabla_X (fY) = (Xf)Y + f \nabla_X Y.$$

It is straightforward to construct connections; indeed, the space of connections is an affine space. (For more, see for example [**82**, Appendix C].)

Fix now a connection ∇ on TM. A vector field v along a path $\gamma\colon [0,1] \to M$ is called *parallel* if it satisfies $\nabla_{\gamma'(t)} v(t) = 0$. For a parallel vector field, the value $v(0)$ uniquely determines $v(t)$ for all $t \in [0,1]$. The vector $v(1)$ is called the *parallel transport* of $v(0)$ along γ; and so a path from p to q in M induces, by parallel transport, and identification $T_pM \cong T_qM$. (Hence the terminology: the connection and path specifices a way to connect the various tangent spaces).

When M is symplectic, the connection ∇ is called a *symplectic connection* if $\omega(\nabla_Z X, Y) + \omega(X, \nabla_Z Y) = 0$ for any choice of vector fields X, Y, and Z. The space of symplectic connections is a non-empty affine space. For a symplectic connection, parallel transport along a path γ from p to q gives an isomorphism of symplectic vector spaces $(T_pM, \omega_p) \cong (T_qM, \omega_q)$.

Having completed the digression, we return to our map $u\colon (D, \partial D) \to (M, L)$. Fix a symplectic connection on (M, ω). Parallel transport along radial lines trivializes the vector bundle $u^*(TM)$, identifying its fibers with $(T_{u(0,0)}M, \omega_{u(0,0)})$. In this manner, we can think of $(u|_{\partial D})^*(TL)$ as a loop γ_L of Lagrangian submanifolds in a single symplectic vector space.

Proposition 4.5.8. *The evaluation of the universal Maslov class μ_n on this loop γ_L of Lagrangians coincides with the evaluation of $u^*\mu_L$ on $[D, \partial D]$.*

Proof. Choose a compatible almost-complex structure. Trivialize the pullback $u^*(\det_{\mathbb{C}}(M) \otimes_{\mathbb{C}} \det_{\mathbb{C}}(M))$ (for example, by radial parallel transport from the origin). The trivialization provided by $\det_{\mathbb{R}}(L) \otimes \det_{\mathbb{R}}(L)$ over ∂D can be thought of as a non-vanishing map of the circle ∂D to $\mathbb{C}$. On the one hand, the degree of this map is the evaluation of the universal Maslov class on the loop of Lagrangians coming from the boundary; on the other hand, the degree is also the evaluation of $u^*\mu_L$ on $[D, \partial D]$. □

4.5.2. Geometric representation of the universal Maslov class. The universal Maslov class can be computed on an element in $H_1(\mathcal{L}Gr(n); \mathbb{Z})$ (which is represented by a closed loop in $\mathcal{L}Gr(n)$) as an intersection number of the loop with a stratified subspace of $\mathcal{L}Gr(n)$. In the following we describe the details of this construction, following [**3**]; we will revisit this construction when discussing the Maslov index in Chapter 6.

Let (V, ω) be a symplectic vector space and fix a Lagrangian subspace $\Lambda_0 \in \mathcal{L}Gr(V, \omega)$. The *Maslov cycle relative to* Λ_0, denoted

$$\Sigma(\Lambda_0) = \{L \in \mathcal{L}Gr(V, \omega) \mid L \cap \Lambda_0 \neq \emptyset\} \subset \mathcal{L}Gr(V, \omega).$$

The set $\Sigma = \Sigma(\Lambda_0)$ can be given the structure of a stratified space

$$\Sigma = \cup_{k=1}^{n} \Sigma_k$$

whose strata are given by

$$\Sigma_k(\Lambda_0) = \{L \in \mathcal{L}Gr(V, \omega) \mid \dim(L \cap \Lambda) = k\}.$$

Basic properties of this stratification are formalized in Lemma 4.5.10 below.

We can think of $\Sigma(\Lambda)$ in the following concrete terms. Choose a linear identification of (V, ω) with the standard symplectic vector space $(\mathbb{R}^{2n}, \Omega_0)$ of the

same dimension, cf. Equation (4.3). Choose this identification so that the specified Lagrangian $\Lambda_0 \subset V^{2n}$ maps to $\mathbb{R}^n$ (spanned by e_i, $i = 1, \dots, n$) in $(\mathbb{R}^{2n}, \Omega_0)$. (We think of $(\mathbb{R}^{2n}, \Omega_0)$ as equipped with the standard almost complex structure $\mathbb{J}_0$.) These choices uniquely specify the identification $\mathcal{L}Gr(V, \omega) \cong U(n)/O(n)$, under which the specified Lagrangian Λ_0 corresponds to the identity coset. and the subset $\Sigma_k(\Lambda_0)$ corresponds to the set of matrices

$$\Sigma_k = \{A \in U(n) \mid \dim(A(\mathbb{R}^n) \cap \mathbb{R}^n) = k\}/O(n) \subset U(n)/O(n). \tag{4.14}$$

In Proposition 4.5.11 we will explain how Σ can be used to represent the universal Maslov class.

Example 4.5.9. *Clearly, $\mathcal{L}Gr(1) = U(1)/O(1)$ is a circle, and Σ_1 is a point in it. We describe next $\Sigma \subset \mathcal{L}Gr(2)$, continuing notation from Example 4.5.3, where the basis for V^4 is $e_1, \dots, e_4$, the symplectic form is chosen to be $\omega = e_1^* \wedge e_2^* + e_3^* \wedge e_4^*$, and a Lagrangian subspace Λ is given as the kernel of the 2-form ψ_Λ satisfying $\psi_\Lambda \wedge \psi_\Lambda = \psi_\Lambda \wedge \omega = 0$. Fix now Λ_0 spanned by e_1 and e_3, corresponding to the 2-form $\psi_0 = e_2^* \wedge e_4^*$. Obviously $\Sigma_i = \emptyset$ if $i > 2$ and $\Sigma_2 = \{\Lambda_0\}$.*

The 2-plane $W \subset V^4$ is in Σ_1 if $\psi_0 \wedge \psi_W = 0$ for the corresponding 2-form ψ_W, but ψ_W is distinct from ψ_0 in $\mathbb{P}(\Lambda^2(V^*))$. Writing $\psi_W = \sum_{i,j} y_{ij} e_i^* \wedge e_j^*$ as before, this condition means $y_{13} = 0$. Together with the two constraints from Equation (4.13) we get that Σ_1 is the two-manifold given by $-y_{12}^2 + y_{14}y_{23} = 0$ in the three-dimensional projective space (with coordinates $[y_{12} : y_{24} : y_{14} : y_{23}]$). The surface Σ_1 then can be easily seen to be an annulus, hence $\Sigma = \Sigma_1 \cup \Sigma_2$ is the image of S^2 with one double point.

In describing the structure of Σ_k we will use the following proposition. For a vector space W let $s(W)$ denote the vector space of symmetric bilinear forms on W.

Lemma 4.5.10. *For any $\Lambda \in \mathcal{L}Gr(V, \omega)$, there is a natural identification $T_\Lambda \mathcal{L}Gr(V, \omega) \cong s(\Lambda)$. Moreover, $\Sigma_k(\Lambda)$ is a submanifold of $\mathcal{L}Gr(V, \omega)$, whose normal space at any $\Lambda \in \Sigma_k(\Lambda_0)$ is identified with $s(\Lambda_0 \cap \Lambda)$.*

Proof. Let $V_0 \cong (\mathbb{R}^{2n}, \Omega_0)$ be the standard symplectic vector space with the fixed Lagrangian $\mathbb{R}^n$ defining Σ. Fix a Lagrangian $\Lambda \in \Sigma_k \subset \mathcal{L}Gr(n)$ with $k = \dim(\Lambda \cap \mathbb{R}^n)$, and fix a splitting $V_0 \cong \Lambda \oplus P$, so that P is Lagrangian. Any n-dimensional vector space L transverse to P (i.e. which is sufficiently close to Λ) can be realized as a graph of a linear function $A \colon \Lambda \to P$. Clearly,

$$\omega(x + A(x), y + A(y)) = \omega(x, A(y)) + \omega(A(x), y) = \omega(x, A(y)) - \omega(y, A(x)),$$

so the graph of A is Lagrangian if and only if the quadratic function on Λ defined by $(x, y) \mapsto \omega(x, A(y))$ is symmetric. In this way, a Lagrangian splitting $V_0 \cong \Lambda \oplus P$ determines an open neighborhood $U_\Lambda(P) \subset \mathcal{L}Gr(n)$ of Λ, and a diffeomorphism

$$\phi_{\Lambda,P} \colon U_\Lambda(P) \to s(\Lambda). \tag{4.15}$$

Since $s(\Lambda)$ is a vector space, Equation (4.15) verifies that $T_\Lambda \mathcal{L}Gr(n) \cong s(\Lambda)$. Clearly, the subset $\Sigma_k \cap U_\Lambda(P) \subset U_\Lambda(P)$ corresponds under $\phi_{\Lambda,P}$ to the symmetric bilinear forms on Λ with k-dimensional nullspace. Such a form is determined by its nullspace N, which can be any element in the $k(n-k)$-dimensional Grassmanian of k-planes in Λ, and the induced symmetric bilinear form on $N^\perp$, a space which has dimension $\frac{1}{2}(n-k)(n-k+1)$. It follows that $\Sigma_k \cap U_\Lambda(P)$ is a submanifold with codimension $\frac{1}{2}k(k+1)$ in the $\frac{1}{2}n(n+1)$-dimensional space $\mathcal{L}Gr(n)$. It is easy to see that for any $L \in \Sigma_k$ there is some Lagrangian subspace P that is transverse to both L and Λ, so that for this choice we have $L \in \Sigma_k \cap U_\Lambda(P)$. It follows now that Σ_k is a smooth submanifold of $\mathcal{L}Gr(n)$.

At any $L \in U_\Lambda(P)$, the differential $d_L\, \phi_{\Lambda,P}$ of $\phi_{\Lambda,P}$ induces an isomorphism

$$d_L\, \phi_{\Lambda,P} \colon T_L \mathcal{L}Gr(n) \to s(\Lambda).$$

When $L = \Lambda$, we claim that the induced isomorphism

$$d_\Lambda\, \phi_{\Lambda,P} \colon T_\Lambda \mathcal{L}Gr(n) \to s(\Lambda), \tag{4.16}$$

is independent of the choice of the complementary subspace P. To see this, suppose that Q is another Lagrangian subspaces transverse to Λ, and let $A \colon (-1,1) \to \mathrm{Hom}(\Lambda, P)$ and $B \colon (-1,1) \to \mathrm{Hom}(\Lambda, Q)$ be two one-parameter families of linear maps whose graphs in $\Lambda \oplus P = \Lambda \oplus Q = V_0$ determine the same path $\{L_t\}_{t\in(-1,1)}$ so that $A_0 = B_0 = 0$ (i.e. $L_0 = \Lambda$; equivalently, the quadratic form associated to A_t is $\Phi_{\Lambda,P}(L_t)$, and the one associated to B_t is $\Phi_{\Lambda,Q}(L_t)$). Since both P and Q are transverse to Λ, there is a map $C \colon P \to \Lambda$, so that Q is the graph of C. Explicitly, for each $x \in \Lambda$, there is a unique $y \in \Lambda$ (depending on t, but we suppress this from the notation) so that $x + A_t(x) = y + B_t(y)$; and there is a unique $z \in P$ so that $B_t(y) = z + C(z)$. Thus, $x + A_t(x) = y + z + C(z)$, so $y + C(z) = x$ and $z = A_t(x)$; i.e.

$$x + A_t(x) = (x - C \circ A_t(x)) + B_t(x - C \circ A_t(x)).$$

It follows that for all $v \in \Lambda$ we have

$$\omega(v, A_t(x)) = \omega(v, B_t(x - C \circ A_t(x))).$$

Differentiating with respect to t, and noting that $A_0 = B_0 = 0$, we conclude that $\omega(v, A_0'(x)) = \omega(v, B_0'(x))$; i.e. $d_\Lambda\, \Phi_{\Lambda,P} = d_\Lambda\, \Phi_{\Lambda,Q}$.

Finally, to identify the normal space at Λ to Σ_k, we proceed as follows. Fix a path $\{\Lambda_t \in \Sigma_k \cap U_\Lambda(P)\}_{t\in[-1,1]}$ through $\Lambda = \Lambda_0$, and let $\{A_t \colon \Lambda \to P\}_{t\in[-1,1]}$ be a path of homomorphisms so that the graph of A_t is Λ_t, and $\dim \mathrm{Ker}(A_t) = k$ for all t. We can find a one-parameter family of bases $(e_t^1, \ldots, e_t^k)$ for $\mathrm{Ker}(A_t) \subset \Lambda$. Clearly, for each $i, j \in 1, \ldots, k$, $\omega(e_t^i, A_t e_t^j) = 0$. Differentiating this equation, we find that $\omega(e_0^i, A_0' e_0^j) = 0$. Thus, the tangent space at Λ to Σ_k is identified with the symmetric, bilinear forms on Λ that vanish on $\Lambda \cap \mathbb{R}^n$. Consequently the normal direction in $T_\Lambda \mathcal{L}Gr(n)$ to Σ_k at $\Lambda \in \Sigma_k$ can be given by $s(\Lambda \cap \mathbb{R}^n)$, i.e. symmetric bilinear forms on the intersection $\Lambda \cap \mathbb{R}^n$. This can be summarized in the commutative square

$$\begin{array}{ccc} T_\Lambda \mathcal{L}Gr(n) & \xrightarrow{\pi} & N_\Lambda \Sigma_k \\ {\scriptstyle d_\Lambda \Phi_{\Lambda,P}}\downarrow & & \cong\downarrow \\ s(\Lambda) & \xrightarrow{r} & s(\Lambda \cap \mathbb{R}^n). \end{array}$$

This identification appears to depend on the choice of P. To see that it does not, we use the fact that the isomorphism from Equation (4.16) is independent of P, together with the following easily verified commutative square, where $L \in \Sigma_k \cap U_\Lambda(P)$ and ρ, r are the natural restriction maps:

$$\begin{array}{ccc} N_L \Sigma & \xrightarrow{d_L \phi_{\Lambda,P}} & s(\Lambda) \\ {\scriptstyle d_L \phi_{L,P}}\downarrow & & r\downarrow \\ s(L) & \xrightarrow{\rho} & s(\Lambda \cap L). \end{array}$$

□

The above proposition identifies the fiber of the normal bundle at Λ to $\Sigma_1(\Lambda)$ with the space of quadratic functions on the one-dimensional vector space $\Lambda \cap \mathbb{R}^n$. We can thus fix a co-orientation on Σ_1 by choosing the positive quadratic functions. (Notice that Σ_1 is of codimension 1.)

Using this co-orientation, Σ represents a one-dimensional cohomology class, as follows. Given a closed, oriented loop γ_0 in $\mathcal{L}Gr(n)$, take a generic perturbation γ of γ_0 that meets Σ_1 transversely, and is disjoint from all Σ_k with $k > 1$. Such a representative can be found since $\Sigma_k \subset \mathcal{L}Gr(n)$ with $k > 1$ has codimension ≥ 2. The evaluation of the cohomology class associated to Σ on γ_0 is obtained by summing ± 1 over each intersection point of γ with Σ_1, where the sign is obtained by comparing the orientation of γ with the co-orientation of Σ_1. This quantity is invariant under generic homotopies of γ, and so determines a cohomology class, since the Σ_k with $k > 1$ have codimension ≥ 3.

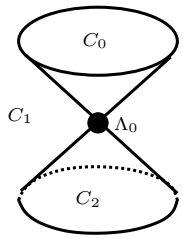


Figure 4.1. Maslov cycle when $n = 4$. When $n = 4$, the Maslov cycle $\Sigma(L_0)$ is locally represented by a cone, as pictured. We have labelled the spaces $C_0(\Lambda_0)$, $C_1(\Lambda_0)$, and $C_2(\Lambda_0)$.

Proposition 4.5.11. *The above defined one-dimensional cohomology class associated to Σ coincides with the universal Maslov class.*

Proof. Equip $\mathbb{C}^n$ with its usual Hermitian form $\langle v, w\rangle = \overline{v}^t \cdot w$, whose imaginary part is the standard symplectic form ω. Consider the path of unitary matrices $[-\frac{\pi}{2}, \frac{\pi}{2}] \to U(n)$ that sends t to the diagonal matrix with one diagonal entry e^{it} and all other diagonal entries $\frac{1+i}{\sqrt{2}}$. This determines a closed loop in $\mathcal{L}Gr(n)$, and its image under $\det^2$ generates $\pi_1(S^1)$. On the other hand, the path crosses Σ_1 at $t = 0$, where it looks like the graph of the map $A_t(e_1) = \tan(t)ie_1$ and $A_t(e_j) = ie_j$ for $j > 1$. In particular, $\langle e_1, A_0' e_1\rangle = 1$. □

The identification $d_\Lambda\, \phi_{\Lambda,P} \colon T_\Lambda \mathcal{L}Gr(n) \to s(\Lambda)$ of Equation (4.16) identifies $T_\Lambda \mathcal{L}Gr(n) \setminus \bigcup_k T_\Lambda \Sigma_k$ with the set of nondegenerate symmetric quadratic forms. This allows us to decompose the above difference as

$$T_\Lambda \mathcal{L}Gr(n) \setminus \bigcup_{k=1}^{n} T_\Lambda \Sigma_k = \bigcup_{k=0}^{n} C_k(\Lambda),$$

where $C_k(\Lambda) = \{v \in T_\Lambda \mathcal{L}Gr(n) \mid d_\Lambda\, \phi_{\Lambda,P}(v) \text{ has index } k\}$. In particular, the subset $C_0(\Lambda)$ corresponds to the positive definite bilinear form s over Λ.

Example 4.5.12. *In the case where V is four-dimensional, if we fix $\Lambda \in \mathcal{L}Gr(V, \omega)$, Lemma 4.5.10 gives a neighborhood of Λ which is identified with $\mathbb{R}^3$, thought of the space of symmetric matrices over $\mathbb{R}^2$. The intersection of $\Sigma(\Lambda)$ with this neighborhood corresponds to the space of symmetric matrices with vanishing determinant, which is a cone in $\mathbb{R}^3$, see Figure 4.1. The cone point corresponds to the singular point in the S^2 from Example 4.5.9.*

Chapter 5

The Morse-Smale complex

Lagrangian Floer homology is based on Morse theory in an infinite-dimensional setting. Before describing that invariant, in this chapter we review the classical case of Morse theory for finite-dimensional manifolds, with an eye towards this infinite-dimensional generalization. For a more thorough account in this spirit see [**125, 5**]; compare also [**146, 77**].

In Section 5.1 we describe the Morse-Smale chain complex of a Morse function on a closed n-dimensional Riemannian manifold, in Section 5.2 some analytic results are quoted and discussed, while in Section 5.3 we show that the homology of the Morse-Smale chain complex is an invariant of the underlying manifold.

5.1. Gradient flowlines and the chain complex

In classical Morse theory, one starts with a closed, finite dimensional manifold M^n, equipped with a Morse function $f\colon M \to \mathbb{R}$. We also equip M with a Riemannian metric g, and consider the downward gradient vector field $-\vec{\nabla}_g f$. For a sufficiently generic choice of g, the classical Morse-Smale complex is then a chain complex whose generators are critical points of the function f, and whose differential counts gradient trajectories in a suitable sense.

Specifically, fix critical points $\mathbf{x}$ and $\mathbf{y}$ of f. A *gradient flowline* (or gradient trajectory) from $\mathbf{x}$ to $\mathbf{y}$ is a path $\gamma\colon \mathbb{R} \to M$ satisfying the following

asymptotic conditions:

$$\lim_{t\mapsto-\infty}\gamma(t)=\mathbf{x},\qquad \lim_{t\mapsto+\infty}\gamma(t)=\mathbf{y} \tag{5.1}$$

and the (downward) gradient flow equation

$$\frac{d\gamma}{dt}=(-\vec{\nabla}_g f)_{\gamma(t)}.$$

Gradient flowlines satisfy an *a priori* energy bound, as follows. Recall that the *energy* of a smooth path $\gamma\colon\mathbb{R}\to M$ is given by

$$E(\gamma)=\int_{-\infty}^{+\infty}\big|\frac{d\gamma(t)}{dt}\big|^2dt,$$

where the integrand is the length squared (with respect to the fixed Riemannian metric g) of the velocity of γ. For a gradient flowline,

$$\big|\frac{d\gamma(t)}{dt}\big|^2=\langle\frac{d\gamma(t)}{dt},\frac{d\gamma(t)}{dt}\rangle=-\frac{d}{dt}(f\circ\gamma(t)),$$

so the energy of any gradient trajectory from $\mathbf{x}$ to $\mathbf{y}$ is given by

$$\int_{-\infty}^{+\infty}\big|\frac{d\gamma(t)}{dt}\big|^2dt=\lim_{t\mapsto-\infty}f\circ\gamma(t)-\lim_{t\mapsto+\infty}f\circ\gamma(t)=f(\mathbf{x})-f(\mathbf{y}). \tag{5.2}$$

Gradient flowlines can be collected into *moduli spaces* $\mathcal{M}(\mathbf{x},\mathbf{y})$, which inherit a natural topology from the C^∞ topology of the space of paths from $\mathbf{x}$ to $\mathbf{y}$. This space admits a natural $\mathbb{R}$-action: if γ is a gradient flowline and s is a real number, then the new path $\tau_s(\gamma)$ given by $t\mapsto\gamma(t+s)$ is also a gradient flowline. The induced action on $\mathcal{M}(\mathbf{x},\mathbf{y})$ is called *time translation.*

Exercise 5.1.1. *Show that if γ is a gradient flowline connecting the critical points $\mathbf{x}$ and $\mathbf{y}$, then $\tau_s(\gamma)$ is a gradient flowline as well. Verify that if for some $s\neq 0$ we have $\tau_s(\gamma)=\gamma$, then γ is a constant path and $\mathbf{x}=\mathbf{y}$.*

Let $\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})$ denote the quotient of $\mathcal{M}(\mathbf{x},\mathbf{y})$ by the time translation action:

$$\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})=\mathcal{M}(\mathbf{x},\mathbf{y})/\mathbb{R}.$$

These moduli spaces satisfy a certain non-degeneracy condition when the Riemannian metric g is sufficiently generic. To formulate this condition, recall that the index $\lambda(p)$ of a critical point p of the function f on M is the index of the Hessian of f at p. The next result follows from standard transversality results (as it is detalied in [**125**], see e.g. Section 2.3.2), by examining intersections of ascending and descending manifolds of critical points (for their definition see [**77**]).

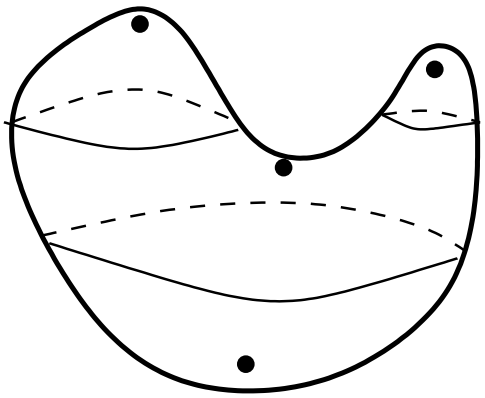


Figure 5.1. The two-sphere with another Morse function

Theorem 5.1.2. *Let M^n be an n-dimensional closed manifold and $f\colon M \to \mathbb{R}$ a given Morse function. For sufficiently generic Riemannian metric g, the ascending and descending manifolds of critical points intersect transversally, hence the space of gradient flows $\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})$ is a smooth manifold whose dimension, when $\mathbf{x} \neq \mathbf{y}$, is given by*

$$\dim \widehat{\mathcal{M}}(\mathbf{x},\mathbf{y}) = \lambda(\mathbf{x}) - \lambda(\mathbf{y}) - 1.$$

Example 5.1.3. *Consider $S^n \subset \mathbb{R}^{n+1}$ with the round metric and the height function $f\colon S^n \to \mathbb{R}$. For the north and south poles (where f takes its maximum and its minimum) $\mathbf{x}$ and $\mathbf{y}$ the gradient flowlines are the great half-circles connecting the two poles, hence the space $\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})$ is diffeomorphic to the equatorial sphere S^{n-1}. For gradient flowlines on S^2 with a different metric and Morse function, see Figure 5.1.*

Theorem 5.1.2 implies that when $\mathbf{x} \neq \mathbf{y}$, and $\lambda(\mathbf{x}) \leq \lambda(\mathbf{y})$, the moduli space $\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})$ is empty for a generic metric.

If M is a smooth manifold, f is a Morse function on M, and g is a Riemannian metric which is generic in the sense of Theorem 5.1.2 for all pairs of critical points (i.e., the ascending and descending manifolds intersect transversally), then we say that the pair (f,g) is *Morse-Smale*.

By choosing a (sufficiently generic) Riemannian metric g, a Morse function f endows M with the structure of a handlebody (cf. Section 1.1) and hence presents M as a CW complex. When the pair (f,g) is Morse-Smale, the resulting CW chain complex, whose homology computes the singular homology of M, has an interpretation in terms of gradient trajectories.

Specifically, the Morse-Smale chain complex, written $\mathrm{CM}_*(M,f,g)$, is the vector space over $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$ generated by the critical points of f, equipped

with the following differential:

$$\partial \mathbf{x} = \sum_{\{\mathbf{y}\in \mathrm{Crit}(f)\,\big|\,\lambda(\mathbf{x})-\lambda(\mathbf{y})=1\}} \#\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})\cdot \mathbf{y}. \tag{5.3}$$

Here, $\#\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})$ is the parity of the number of elements in $\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})$. In order this definition to make sense we need to verify that once $\lambda(\mathbf{x})-\lambda(\mathbf{y})=1$ holds, $\#\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})$ is a finite set — we will return to this issue in the next section. Then, as for a compact manifold M the set of critical points is finite, the above formula of $\partial\colon \mathrm{CM}_*(M,f,g)\to \mathrm{CM}_*(M,f,g)$ makes sense.

Remark 5.1.4. *We work here with coefficients over $\mathbb{Z}/2\mathbb{Z}$ for simplicity. By orienting the moduli spaces of gradient flows, one can lift these constructions to $\mathbb{Z}$; see* [**5, 125**].

Relying on the formulation of singular homology adapted to CW-complexes, one can directly prove that $\mathrm{CM}_*(M,f,g)$ is isomorphic to the cellular chain complex provided by the CW decomposition (with $\mathbb{Z}/2\mathbb{Z}$ coefficients). This implies that $\mathrm{CM}_*(M,f,g)$ is a chain complex. Nonetheless, it is instructive to show that $\partial^2=0$ holds directly via analysis of moduli spaces of gradient flowlines. This is the perspective that generalizes readily to the setting of Floer homology.

The verification rests on a certain compactification of the moduli spaces of gradient flowlines. In fact, this compactness result already plays a role when making sense of Equation (5.3): according to Theorem 5.1.2, when $\lambda(\mathbf{x})-\lambda(\mathbf{y})=1$, the space $\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})$ is zero-dimensional; but the fact that it is compact (and hence consists of finitely many points) needs justification.

5.2. Compactness and gluing

By the homeomorphism $\mathbb{R}\sim(-1,1)$ and the assumptions on the limits at $\pm\infty$, a gradient flowline $\gamma\in\mathcal{M}(\mathbf{x},\mathbf{y})$ extends to a map $\overline{\gamma}\colon[-1,1]\to M$, with $\overline{\gamma}(-1)=\mathbf{x}$ and $\overline{\gamma}(1)=\mathbf{y}$. As such maps, gradient flowlines determine *path homotopy classes*: define $p_1(\mathbf{x},\mathbf{y})$ as the equivalence classes of the set of smooth maps $\overline{\gamma}\colon[-1,1]\to M$ with $\overline{\gamma}(-1)=\mathbf{x}$ and $\overline{\gamma}(1)=\mathbf{y}$, where $\overline{\gamma}$ and $\overline{\gamma}'$ are equivalent if they are homotopic rel endpoints. Obviously, for three critical points $\mathbf{x},\mathbf{y},\mathbf{z}$, if $c_1=[\overline{\gamma}_1]\in p_1(\mathbf{x},\mathbf{y})$ and $c_2=[\overline{\gamma}_2]\in p_1(\mathbf{y},\mathbf{z})$, then (after reparametrizing) we can concatenate them to an element $c_1*c_2\in p_1(\mathbf{x},\mathbf{z})$.

Definition 5.2.1. *A **broken flowline** from $\mathbf{x}$ to $\mathbf{y}$ is a sequence of distinct critical points $\mathbf{x}=\mathbf{x}_1,\dots,\mathbf{x}_{n+1}=\mathbf{y}$ and a collection of gradient flowlines $\alpha_1,\dots,\alpha_n$ where $\alpha_i\in\widehat{\mathcal{M}}(\mathbf{x}_i,\mathbf{x}_{i+1})$. The number $n-1$ is called the **number of breaks** in the broken flowline.*

Definition 5.2.2. *A sequence* $\{\gamma_m\}_{m=1}^{\infty}$ *in* $\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})$ *is said to* ***converge to a broken flowline*** $(\alpha_1,\ldots,\alpha_n)$ *(with* $\alpha_i \in \widehat{\mathcal{M}}(\mathbf{x}_i,\mathbf{x}_{i+1})$ *non-constant flowlines) if, for each* j*, we can find representatives* $\gamma_m^j \in \mathcal{M}(\mathbf{x},\mathbf{y})$ *for* $\gamma_m \in \widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})$ *with the property that* $\{\gamma_m^j\}_{m=1}^{\infty}$ *converges to* α_j *in the* $C^{\infty,\mathrm{loc}}$ *topology. (Recall that* $\gamma_m \in \widehat{\mathcal{M}}(\mathbf{x},\mathbf{y}) = \mathcal{M}(\mathbf{x},\mathbf{y})/\mathbb{R}$ *is itself an equivalence class of gradient flowlines, up to reparametrization.)*

It follows from the definition that the path homotopy classes of γ_m and $\alpha_1 * \ldots * \alpha_n$ are equal. Broken flowlines give a compactification of the space of flowlines in the following sense; see [**125**, Section 2.4.2]:

Theorem 5.2.3. *Any sequence of gradient flowlines from* $\mathbf{x}$ *to* $\mathbf{y}$ *(with* $\mathbf{x} \neq \mathbf{y}$*) has a* $C^{\infty,\mathrm{loc}}$*-convergent subsequence to a broken flowline from* $\mathbf{x}$ *to* $\mathbf{y}$*.*

Proof. [Sketch] The crucial step in the proof of the above theorem is that by Equation (5.2) we have the a priori estimate $\int_{s_1}^{s_2} \left|\frac{d\gamma(t)}{dt}\right|^2 dt \leq f(\mathbf{x}) - f(\mathbf{y})$ for any flowline in $\mathcal{M}(\mathbf{x},\mathbf{y})$, showing that the flowlines (as C^∞ maps from $\mathbb{R}$ to M) are uniformly equicontinuous. A similar estimate shows pointwise convergence, hence the application of the Arzela-Ascoli theorem provides a weakly convergent subsequence. The limit will also satisfy the gradient flow equation, giving a component in the broken flowline compactification (in the sense of Definition 5.2.2). Other components are obtained by a translation. □

This compactification has the following immediate consequence relevant for the Morse-Smale complex; here we provide a sketch of the argument, see [**5**, Corollary 3.2.4] for a detailed proof.

Proposition 5.2.4. *If* g *is a generic metric, then the zero-dimensional moduli spaces* $\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})$ *with* $\lambda(\mathbf{x}) - \lambda(\mathbf{y}) = 1$ *are compact in the* C^∞ *topology.*

Proof. [Sketch] According to Theorem 5.2.3, any sequence of points in $\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})$ has a subsequence which converges to a broken flowline. It is now a consequence of Theorem 5.1.2 that any broken flowline connecting $\mathbf{x}$ to $\mathbf{y}$ with $\lambda(\mathbf{x}) - \lambda(\mathbf{y}) = 1$ is, in fact, unbroken. To see this, note that

$$1 = \sum_{i=1}^{n} \lambda(\mathbf{x}_i) - \lambda(\mathbf{x}_{i+1}),$$

so if $n > 1$, then some $\lambda(\mathbf{x}_i) - \lambda(\mathbf{x}_{i+1}) \leq 0$; but then, by Theorem 5.1.2, there are no flowlines in the corresponding moduli space $\widehat{\mathcal{M}}(\mathbf{x}_i,\mathbf{x}_{i+1})$.

In the case where the limiting broken flowline has no breaks in it, the above notion of convergence coincides with the usual C^∞ convergence of flowlines. □

Proposition 5.2.4 now ensures that the moduli spaces $\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})$ appearing in the definition of the boundary operator from Equation (5.3) (i.e. have $\dim \widehat{\mathcal{M}}(\mathbf{x},\mathbf{y}) = 0$) indeed have finitely many elements, hence ∂ is well-defined.

We turn our attention to the analysis that shows $\partial^2 = 0$. The idea of this proof is to consider two-dimensional moduli spaces $\mathcal{M}(\mathbf{x},\mathbf{z})$ with $\lambda(\mathbf{x}) - \lambda(\mathbf{z}) = 2$, i.e. whose unparameterized versions $\widehat{\mathcal{M}}(\mathbf{x},\mathbf{z})$ are one-dimensional. In order to understand the ends of such a moduli space, we need the following *gluing result*.

Theorem 5.2.5. *Suppose that $\lambda(\mathbf{x}) = \lambda(\mathbf{y}) + 1 = \lambda(\mathbf{z}) + 2$. Then, for a point $(\alpha_1, \alpha_2) \in \widehat{\mathcal{M}}(\mathbf{x},\mathbf{y}) \times \widehat{\mathcal{M}}(\mathbf{y},\mathbf{z})$ there is $\rho_0 \in \mathbb{R}$ and a smooth map*

$$f\colon [\rho_0, \infty) \to \widehat{\mathcal{M}}(\mathbf{x},\mathbf{z})$$

such that $f(\rho)$ converges to the broken flowline (α_1, α_2) as $\rho \to \infty$.

Furthermore, if a sequence $\{\gamma_n\}_{n=1}^\infty \subset \widehat{\mathcal{M}}(\mathbf{x},\mathbf{z})$ with $\lambda(\mathbf{x}) - \lambda(\mathbf{z}) = 2$ converges to the broken flowline (α_1, α_2), then for n large enough, $\gamma_n \in f([\rho_0, \infty))$. □

Remark 5.2.6. *The proof of this result comprises two major steps. First, we construct an approximate flowline for the given $\alpha_1 \in \widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})$ and $\alpha_2 \in \widehat{\mathcal{M}}(\mathbf{y},\mathbf{z})$ using the parameter ρ by following α_1 until it reaches a neighbourhood of $\mathbf{y}$, and then connecting it inside the small neighbourhood to the flowline α_2; see Figure 5.2(b). In the second step, using a contraction principle on Banach spaces, this approximate flowline is modified to an actual flowline, illustrated by Figure 5.2(c).*

From the compactification by broken flowlines and the gluing results (as described in Theorem 5.2.5), the ends of the moduli spaces $\widehat{\mathcal{M}}(\mathbf{x},\mathbf{z})$ with $\lambda(\mathbf{x}) - \lambda(\mathbf{z}) = 2$ now consist of broken flowlines. In fact, a dimension count shows that these ends all consist of moduli spaces where there is exactly one intermediate break, i.e. the ends are contained in the Cartesian product

$$\bigcup_{\{\mathbf{y} \in \mathrm{Crit}(f) \,\big|\, \lambda(\mathbf{x}) - \lambda(\mathbf{y}) = 1\}} \widehat{\mathcal{M}}(\mathbf{x},\mathbf{y}) \times \widehat{\mathcal{M}}(\mathbf{y},\mathbf{z}).$$

Indeed, the above Cartesian product is precisely the set of ends of $\widehat{\mathcal{M}}(\mathbf{x},\mathbf{z})$. That is, given every (one step) broken flowline in $\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y}) \times \widehat{\mathcal{M}}(\mathbf{y},\mathbf{z})$, there is a sequence of unbroken flowlines from $\mathbf{x}$ to $\mathbf{z}$ that converges to it; in fact,

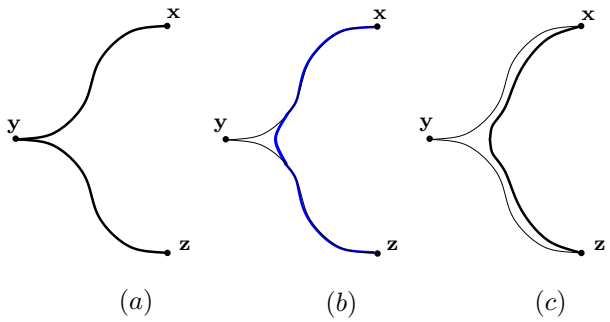


Figure 5.2. The gluing of two gradient trajectories. The two trajectories (from $\mathbf{x}$ to $\mathbf{y}$ and from $\mathbf{y}$ to $\mathbf{z}$) is shown in (a); the figure in (b) shows an approximate flowline, while in (c) the glued up flowline is given.

a neighborhood of the broken flowline is identified with an interval; see [**5**, Theorem 3.2.7].

Armed with this fact, and the elementary fact that any compact one-manifold with boundary has an even number of boundary points, it follows that

$$\sum_{\{\mathbf{y}\in \mathrm{Crit}(f)\,\big|\,\lambda(\mathbf{y})=\lambda(\mathbf{x})-1\}} \#\widehat{\mathcal{M}}(\mathbf{x},\mathbf{y})\cdot\#\widehat{\mathcal{M}}(\mathbf{y},\mathbf{z}) \equiv 0 \pmod 2.$$

Since the left hand side is simply the $\mathbf{z}$-coefficient of $\partial^2\mathbf{x}$, it follows that

Proposition 5.2.7. *For the boundary map ∂ defined in Equation* (5.3) *we have $\partial^2 = 0$; consequently $(\mathrm{CM}_*(M,f,g),\partial)$ is a chain complex.* □

5.3. Topological invariance of the homology

Since the homology of the Morse-Smale complex computes the singular homology of M, it follows immediately that the homology of this complex, which involved the choice of a Morse function and a Riemannian metric, is independent of these auxiliary choices. This topological invariance has an alternative verification, without appealing to the isomorphism with singular homology. As this alternative approach will serve as motivation in Lagrangian Floer theory, we will sketch the main ideas of the invariance proof along these lines presently.

In the following we will consider 1-parameter families of metrics or functions interpolating our given data; using these families we define maps between the chain complexes resulted from different auxiliary choices. The proof of $\partial^2 = 0$ adapts to show that these maps are chain maps. Then a two-parameter family of auxiliary choices provides a chain homotopy, which shows that the above chain maps induce isomorphisms on homology.

First, we consider the independence of the homology of the Morse-Smale chain complex $\mathrm{CM}_*(M, f, g)$ from the choice of the Riemannian metric g, while fixing the Morse function f. (In the following we will drop M from the notation, and denote the chain complex simply by $\mathrm{CM}_*(f, g)$.) Consider two metrics g_0, g_1, with the property that for both the gradient flow of f is Morse-Smale. Metrics form a path-connected space, hence we can find a smooth path of metrics $\{g_t\}_{t\in[0,1]}$ connecting g_0 and g_1. We now consider trajectories $\gamma\colon \mathbb{R} \to M$ satisfying the asymptotic conditions

$$\lim_{t\mapsto-\infty} \gamma(t) = \mathbf{x}, \qquad \lim_{t\mapsto+\infty} \gamma(t) = \mathbf{y} \tag{5.4}$$

for $\mathbf{x}, \mathbf{y} \in \mathrm{Crit}(f)$, and the "time-dependent gradient flow" condition

$$\frac{d\gamma}{dt} = -\vec{\nabla}_{g_{\psi(t)}} f_{\gamma(t)}, \tag{5.5}$$

for some fixed smooth, monotone function $\psi\colon \mathbb{R} \to [0, 1]$ with $\psi(t) = 0$ for $t \leq 0$ and $\psi(t) = 1$ for $t \geq 1$. Once again, we consider moduli spaces of such time-dependent trajectories, which we denote by $\mathcal{M}_{\{g_t\}}(\mathbf{x}, \mathbf{y})$. (We do not record the function ψ in the notation.) Unless $\{g_t\}_{t\in[0,1]}$ is constant, Equation (5.5) is no longer invariant under time translation.

Theorem 5.1.2 has the following adaptation to the time-dependent case:

Theorem 5.3.1. *Let M^n be a closed n-dimensional manifold, and let g_0 and g_1 be two Riemannian metrics on M. For any sufficiently generic path of Riemannian metrics $\{g_t\}$ from g_0 to g_1, and for $\mathbf{x}, \mathbf{y} \in \mathrm{Crit}(f)$ the space of time-dependent gradient flows $\mathcal{M}_{\{g_t\}}(\mathbf{x}, \mathbf{y})$ is a smooth manifold whose dimension, when $\mathbf{x} \neq \mathbf{y}$, is given by*

$$\dim \mathcal{M}_{\{g_t\}}(\mathbf{x}, \mathbf{y}) = \lambda(\mathbf{x}) - \lambda(\mathbf{y}).$$

These time-dependent flows are used to count coefficients in the "continuation map",

$$\Phi_{\{g_t\}}\colon \mathrm{CM}_*(f, g_0) \to \mathrm{CM}_*(f, g_1)$$

defined by

$$\Phi_{\{g_t\}}(\mathbf{x}) = \sum_{\{\mathbf{y}\in\mathrm{Crit}(f)\,|\,\lambda(\mathbf{x})=\lambda(\mathbf{y})\}} \#\mathcal{M}_{\{g_t\}}(\mathbf{x}, \mathbf{y}) \cdot \mathbf{y}. \tag{5.6}$$

A map from $\mathrm{CM}_*(f, g_1)$ to $\mathrm{CM}_*(f, g_0)$ is defined by reversing the path of metrics $\{g_t\}$

$$\Phi_{\{g_{1-t}\}}\colon \mathrm{CM}_*(f, g_1) \to \mathrm{CM}_*(f, g_0).$$

Our aim is to show that $\Phi_{\{g_t\}}$ and $\Phi_{\{g_{1-t}\}}$ are chain maps which are chain homotopy inverses of one another. Concretely, if for $i = 0, 1$, $\partial^{(i)}$ represents

the differential on $\mathrm{CM}_*(f, g_i)$, then the statement that $\Phi_{\{g_t\}}$ and $\Phi_{\{g_{1-t}\}}$ are chain maps amounts to the identities

(5.7)
$$\partial^{(1)} \circ \Phi_{\{g_t\}} + \Phi_{\{g_t\}} \circ \partial^{(0)} = 0 \qquad \text{and} \qquad \partial^{(0)} \circ \Phi_{\{g_{1-t}\}} + \Phi_{\{g_{1-t}\}} \circ \partial^{(1)} = 0.$$

(Note that as we work over the field $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$, so we are free to disregard signs.) The statment that $\Phi_{\{g_t\}}$ and $\Phi_{\{g_{1-t}\}}$ are chain homotopy inverses to one another amounts to finding operators

$$H^{(0)} \colon \mathrm{CM}_*(f, g_0) \to \mathrm{CM}_*(f, g_0) \qquad \text{and} \qquad H^{(1)} \colon \mathrm{CM}_*(f, g_1) \to \mathrm{CM}_*(f, g_1)$$

satisfying the identities

(5.8)
$$\begin{aligned} \partial^{(0)} \circ H^{(0)} + H^{(0)} \circ \partial^{(0)} &= \mathrm{Id} + \Phi_{\{g_{1-t}\}} \circ \Phi_{\{g_t\}} \\ \partial^{(1)} \circ H^{(1)} + H^{(1)} \circ \partial^{(1)} &= \mathrm{Id} + \Phi_{\{g_t\}} \circ \Phi_{\{g_{1-t}\}}. \end{aligned}$$

The verification of all of these identities rest on the analysis of broken flowlines, as we explain below.

There is a broken flowline compactification of the moduli space of time-dependent flowlines $\mathcal{M}_{\{g_t\}}(\mathbf{x}, \mathbf{y})$. In this compactification, we have two sequences of critical points $\mathbf{x} = \mathbf{x}_1, \dots, \mathbf{x}_n$ and $\mathbf{y}_1, \dots, \mathbf{y}_m = \mathbf{y}$, a sequence of gradient flows $\alpha_i \in \widehat{\mathcal{M}}_{g_0}(\mathbf{x}_i, \mathbf{x}_{i+1})$ for $i = 1, \dots, n-1$, another one $\beta_i \in \widehat{\mathcal{M}}_{g_1}(\mathbf{y}_i, \mathbf{y}_{i+1})$ for $i = 1, \dots, m-1$, and a time-dependent flowline $\gamma \in \mathcal{M}_{\{g_t\}}(\mathbf{x}_n, \mathbf{y}_1)$. An analogue of Theorem 5.2.3 states that any sequence of time-dependent gradient flowlines from $\mathbf{x}$ to $\mathbf{y}$ has a $C^{\infty,\mathrm{loc}}$-convergent subsequence to such a broken flowline $(\alpha_1, \dots, \alpha_n, \gamma, \beta_1, \dots, \beta_m)$.

Adapting the argument from the proof of Proposition 5.2.4 to the time-dependent case, we conclude that if $\lambda(\mathbf{x}) = \lambda(\mathbf{y})$, the space $\mathcal{M}_{\{g_t\}}(\mathbf{x}, \mathbf{y})$ is a compact, zero-dimensional manifold. This result provides the necessary finiteness in order to define the coefficients in the definition of $\Phi_{\{g_t\}}$ in Equation (5.6).

The verification that $\Phi_{\{g_t\}}$ is a chain map uses the same logic that was used in the verification that $\partial^2 = 0$. Now, we consider the one-manifolds

$\mathcal{M}_{\{g_t\}}(\mathbf{x},\mathbf{y})$ with $\lambda(\mathbf{y}) = \lambda(\mathbf{x}) - 1$; these moduli spaces have a compactification $\overline{\mathcal{M}}_{\{g_t\}}(\mathbf{x},\mathbf{y})$ with

(5.9)

$$\begin{aligned}\partial\overline{\mathcal{M}}_{\{g_t\}}(\mathbf{x},\mathbf{y}) &= \bigcup_{\{\mathbf{x}'\in\mathrm{Crit}(f)\,\big|\,\lambda(\mathbf{x}')=\lambda(\mathbf{y})\}} \widehat{\mathcal{M}}_{g_0}(\mathbf{x},\mathbf{x}') \times \mathcal{M}_{\{g_t\}}(\mathbf{x}',\mathbf{y}) \quad \cup \\ &\cup \bigcup_{\{\mathbf{y}'\in\mathrm{Crit}(f)\,\big|\,\lambda(\mathbf{y}')=\lambda(\mathbf{x})\}} \mathcal{M}_{\{g_t\}}(\mathbf{x},\mathbf{y}') \times \widehat{\mathcal{M}}_{g_1}(\mathbf{y}',\mathbf{y}).\end{aligned}$$

(This equation is the analogue of Theorem 5.2.5.)

The count of points in the space on the right-hand-side of Equation (5.9) is the coefficient of $\mathbf{y}$ in

$$\Phi_{\{g_t\}}(\partial^{(0)}(\mathbf{x})) + \partial^{(1)}(\Phi_{\{g_t\}}(\mathbf{x})).$$

Since the number of ends of $\{\mathcal{M}_{\{g_t\}}(\mathbf{x},\mathbf{y})\big|\lambda(\mathbf{x})-\lambda(\mathbf{y}) = 1\}$ is zero, counting points in Equation (5.9) gives the first identity from Equation (5.7). Reversing the roles of g_0 and g_1, and using the path of metrics $\{g_{1-t}\}$ gives us the second identity.

The construction of the homotopy operator H involves two-parameter spaces of metrics. Let $\{g_{r,t}\}_{r\in[0,\infty),t\in\mathbb{R}}$ be a smooth two-parameter family of metrics with the following properties:

- $g_{r,t} = g_0$ if $|t| \geq r$
- there is some $R > 1$ so that for all $r \geq R$, and $t \geq 0$, $g_{r,t} = g_{\psi(r-t)}$
- for $r \geq R$ and $t \leq 0$, $g_{r,t} = g_{\psi(r+t)}$.

It follows from the above conditions that $g_{0,t} = g_0$ and $g_{r,t} = g_1$ when $|t| < r - 1$. It is easy to construct a continuous two-parameter space of metrics as above with $R = 1$, which is smooth away from $(r,t) \in [0,1] \times \{0\}$. The desired family is obtained by smoothing out this family in a neighborhood of the interval. In particular, for large enough fixed r the family $\{g_{r,t}\}$ looks like the path of metrics $\{g_t\}$, separated from $\{g_{1-t}\}$ by a long time (depending on r) where it is constant (at g_1). See Figure 5.3 for a schematic illustration.

There is a corresponding moduli space $\mathcal{M}_{\{g_{r,t}\}}(\mathbf{x},\mathbf{y})$ consisting of pairs of a real number $r \in [0,\infty)$ and a path $\gamma\colon \mathbb{R} \to M$ with the usual asymptotic conditions (as given in Equations (5.1) or (5.4)) and the time-dependent gradient flow equation

$$\frac{d\gamma}{dt} = -\vec{\nabla}_{g_{r,t}} f_{\gamma(t)}.$$

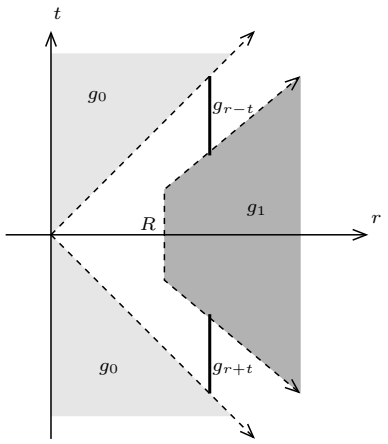


Figure 5.3. Schematics of a two-parameter family of metrics used to construct $H^{(0)}$. The shading indicates regions where the metric is locally constant.

The analogue of Theorem 5.3.1 shows that for generic choice of $\{g_{r,t}\}$, the moduli space $\mathcal{M}_{\{g_{r,t}\}}(\mathbf{x}, \mathbf{y})$ of such pairs is a manifold of dimension $\lambda(\mathbf{x}) - \lambda(\mathbf{y}) + 1$. The additional dimension here comes from the additional parameter r in the moduli space.

There is also a broken flowline compactification of $\mathcal{M}_{\{g_{r,t}\}}(\mathbf{x}, \mathbf{y})$, which has two new kinds of ends, covering the ends $r = 0$ and $r = \infty$ of the parameter space. Explicitly, consider a sequences of points (r_i, γ_i) in $\mathcal{M}_{\{g_{r,t}\}}(\mathbf{x}, \mathbf{y})$. Such a sequence could have a subsequence that converges to another point in $\mathcal{M}_{\{g_{r,t}\}}(\mathbf{x}, \mathbf{y})$. Otherwise, projecting to the r coordinate, and passing to a subsequence, we find the following three kinds of behavior.

(E-1) r_i converge to some value $\rho \in (0, \infty)$. In that case, the γ_i will converge to a broken time-dependent flowline with respect to the path of metrics $\{g_{r=\rho,t}\}_{t\in[-\rho,\rho]}$.

(E-2) r_i converge to 0, in which case $\{\gamma_i\}$ has a subsequence that converge to a (possibly broken) gradient flowline for the g_0 metric.

(E-3) r_i converge to ∞, in which case $\{\gamma_i\}$ converge to a juxtaposition of two (possibly broken) time-dependent flowlines, starting with one for the $\{g_t\}$ metric family, followed by one for the $\{g_{1-t}\}$.

In the cases where $\lambda(\mathbf{y}) = \lambda(\mathbf{x})+1$, the above compactification and transversality show that $\mathcal{M}_{\{g_{r,t}\}}(\mathbf{x}, \mathbf{y})$ is a compact, 0-dimensional space, so we can define

$$H^{(0)} \colon \mathrm{CM}_*(f, g_0) \to CM_{*+1}(f, g_0)$$

by

$$H^{(0)}(\mathbf{x}) = \sum_{\{\mathbf{y}\in \mathrm{Crit}(f)|\lambda(\mathbf{y})=\lambda(\mathbf{x})+1\}} \#\mathcal{M}_{\{g_{r,t}\}}(\mathbf{x},\mathbf{y})\cdot \mathbf{y}.$$

To prove Equation (5.8), we consider moduli spaces $\mathcal{M}_{\{g_{r,t}\}}(\mathbf{x},\mathbf{y})$ where $\lambda(\mathbf{y}) = \lambda(\mathbf{x})$. Transversality ensures that for the ends of these 1-dimensional moduli spaces, the flows appearing in Cases (E-2) and (E-3) are unbroken; while those appearing in Case (E-1) have exactly one break. The gluing result then identifies the ends of $\mathcal{M}_{\{g_{r,t}\}}(\mathbf{x},\mathbf{y})$ with the following union:

$$\left(\bigcup_{\substack{\mathbf{x}'\in \mathrm{Crit}(f)\ s.t.\\ \lambda(\mathbf{x}')=\lambda(\mathbf{x})-1}} \widehat{\mathcal{M}}_{g_0}(\mathbf{x},\mathbf{x}')\times \mathcal{M}_{\{g_{r,t}\}}(\mathbf{x}',\mathbf{y}) \right) \cup \left(\bigcup_{\substack{\mathbf{y}'\in \mathrm{Crit}(f)\ s.t.\\ \lambda(\mathbf{y}')=\lambda(\mathbf{y})+1}} \mathcal{M}_{\{g_{r,t}\}}(\mathbf{x},\mathbf{y}')\times \widehat{\mathcal{M}}_{g_0}(\mathbf{y}',\mathbf{y}) \right)$$

$$\cup \quad \widehat{\mathcal{M}}_{g_0}(\mathbf{x},\mathbf{y})$$

$$\cup \left(\bigcup_{\{\mathbf{z}\in \mathrm{Crit}(f)|\lambda(\mathbf{x})=\lambda(\mathbf{z})\}} \mathcal{M}_{\{g_t\}}(\mathbf{x},\mathbf{z})\times \mathcal{M}_{\{g_{1-t}\}}(\mathbf{z},\mathbf{y}) \right).$$

The three lines correspond to the ends described in Cases (E-1), (E-2), and (E-3) respectively. Counting points in the four unions gives, in order, the four terms appearing in Equation (5.8); and thus the relation holds because the one-dimensonal manifold obtained by compactifying $\mathcal{M}_{g_{r,t}}(\mathbf{x},\mathbf{y})$ with $\lambda(\mathbf{x}) = \lambda(\mathbf{y})$ has an even number of boundary points. The operator $H^{(1)}$ is constructed similarly, reversing the roles of g_0 and g_1 and replacing $\{g_t\}$ by the path $\{g_{1-t}\}$; its homotopy formula follows.

The arguments so far show that the homology $\mathrm{HM}_*(M,f,g) = H_*(\mathrm{CM}_*(M,f,g),\partial)$ is independent of the choice of g. To show that $\mathrm{HM}_*(M,f,g)$ is independent of the choice of f proceeds in a very similar manner. In that case, we fix a one-parameter family $\{f_t\}$ of functions that interpolate between two fixed Morse functions f_0 and f_1. A continuation map is now defined by counting time-dependent gradient trajectories, satisfying

$$\frac{d\gamma}{dt} = -\vec{\nabla}_g f_{\psi(t)} = (-\vec{\nabla}_g f_{\psi(t)})_{\gamma(t)}.$$

Again, we collect these trajectories into moduli spaces $\mathcal{M}_{\{f_t\}}(\mathbf{x},\mathbf{y})$, where $\mathbf{x}\in \mathrm{Crit}(f_0)$ and $\mathbf{y}\in \mathrm{Crit}(f_1)$. For a sufficiently generic choice of g these moduli spaces are manifolds of dimension $\lambda(\mathbf{x}) - \lambda(\mathbf{y})$.

We define a continuation map

$$\Phi_{\{f_t\}}\colon \mathrm{CM}_*(f_0,g) \to \mathrm{CM}_*(f_1,g)$$

by the formula

$$\Phi_{\{f_t\}}(\mathbf{x}) = \sum_{\{\mathbf{y}\in \mathrm{Crit}(f_1)|\lambda(\mathbf{x})=\lambda(\mathbf{y})\}} \#\mathcal{M}_{\{f_t\}}(\mathbf{x},\mathbf{y})\cdot \mathbf{y}.$$

Compactness of moduli spaces ensure that the above counts are finite. Analysis of ends of one-dimensional moduli spaces shows that the map $\Phi_{\{f_t\}}$ is a chain map. Using (generic) two-parameter families of functions, we define a homotopy operator analogous to the one we defined for two-parameter families of metrics. Again, an analysis of one-dimensional moduli spaces with respect to this two-parameter family of functions shows that $\Phi_{\{f_t\}}$ and $\Phi_{\{f_{1-t}\}}$ are homotopy inverses to one another. Thus, $\Phi_{\{f_t\}}$ induces an isomorphism on homology. To fill in the details of this sketch requires more work; the reader is referred to [**10, 6, 5, 125**].

Morse theory can be adapted to more general settings. With little extra work, the compactness hypothesis on M can be weakened, requiring the Morse function to be proper. More interestingly, Morse theory can be done in infinite-dimensional settings, as well. A classical application is to the space of paths in a Riemannian manifold, using an energy functional, whose critical points are geodesics. In this case, although the space is infinite-dimensional, the Hessian nonetheless has finite index, so one can show that the space of paths has the structure of a CW complex; see [**79**] for a beautiful account of this version. Lagrangian Floer homology, however, will live in a setting where the critical points in an infinite dimensional manifold have Hessians with infinitely many positive as well as negative eigenvalues, hence the index no longer makes sense. The new ideas required in this setting will be discussed in the next chapter.

Chapter 6

An overview of Lagrangian Floer homology

Heegaard Floer homology is based on some constructions in symplectic geometry: specifically, Andreas Floer's homology theory for Lagrangian submanifolds of a symplectic manifold. In the construction we will need, the ambient symplectic manifold will always be the symmetric product of a Riemann surface (discussed in Chapter 7) and the Lagrangian submanifolds will always be tori. However, the reader might find it convenient to have these constructions presented in a more general context; and that is the purpose of the present chapter. This chapter is not intended to be a comprehensive introduction to this vast and growing subject: for that, we refer the interested reader to other sources [**28, 120, 126**]. Nevertheless, some of the definitions given in this chapter will be used in our further discussions.

After introducing the main concepts and Arnold's conjecture in Section 6.1, in Section 6.2 we give a short motivation how Lagrangian Floer homology can be used to address these conjectures. In Section 6.3 we motivate the definition of Lagrangian Floer homology, while in Sections 6.4 and 6.5 we start the systematic study of Whitney disks and their Maslov indices. Sections 6.6, 6.7 and 6.8 give some of the analytic details one needs to set up the theory: we briefly discuss transversality, compactness and gluing. The independence on the choice of the almost-complex structures, and invariance under Hamiltonian isotopy for Lagrangian Floer homology is discussed in Section 6.9.

6.1. Hamiltonian diffeomorphisms and their fixed points

Special symplectomorphisms can be generated on symplectic manifolds via the following construction.

Definition 6.1.1. *Fix a symplectic manifold (M,ω). Let $H\colon \mathbb{R}\times M\to\mathbb{R}$ be any smooth function, and consider the (time dependent) vector field X_t satisfying $\iota_{X_t}\omega = dH_t$, where $H_t(x)=H(t,x)$ and $\iota_{X_t}\omega$ denotes the contraction of ω with the vector field X_t. Such a vector field X_t is called a **Hamiltonian vector field**. A diffeomorphism $\phi\colon (M,\omega)\to(M,\omega)$ is called a **Hamiltonian diffeomorphism** if it is the time-one map of a Hamiltonian vector field; i.e. if $\phi(x)=\Phi_1(x)$ where $\{\Phi_t\colon M\to M\}_{t\in[0,1]}$ is the one-parameter family of diffeomorphisms satisfying:*

- $\frac{d\Phi_t}{dt} = (\Phi_t)_*(X_t)$ *and*
- $\Phi_0 = \mathrm{Id}_M$.

Proposition 6.1.2. *A Hamiltonian diffeomorphism is a symplectomorphism.*

Proof. A standard computation shows that

$$\frac{d}{dt}\Phi_t^*(\omega) = \Phi_t^*(\mathcal{L}_{X_t}(\omega)),$$

where $\mathcal{L}_{X_t}(\omega)$ denotes the Lie derivative of ω in the direction X_t. Cartan's formula states that $\mathcal{L}_{X_t}(\omega) = d(\iota_{X_t}\omega) + \iota_{X_t}(d\omega)$. Since ω is closed and $\iota_{X_t}\omega = dH_t$, this Lie derivative vanishes. It follows that $\Phi_t^*(\omega)$ is independent of t, so $\Phi_1^*(\omega)=\Phi_0^*(\omega)=\omega$. □

Using a 1-parameter family $\{\Phi_t\}$ (with $\Phi_0=\mathrm{Id}_M$) of symplectomorphisms (or Hamiltonian diffeomorphisms) we can talk about symplectic (or Hamiltonian) isotopic submanifolds.

Exercise 6.1.3. (a) *Show that if $\eta = dh$ for a function $h\colon L\to\mathbb{R}$ then the graph Γ_η of η in T^*L (as defined in Exercise 4.2.5) is a Legendrian submanifold Hamiltonian isotopic to the 0-section $L\subset(T^*L,-d\lambda)$.*

(b) *Equip $M\times M$ with the symplectic structure described in Example 4.2.6. Show that if the symplectomorphism $\phi\colon M\to M$ is a Hamiltonian diffeomorphism, then the graph $\Gamma_\phi\subset M\times M$ is Hamiltonian isotopic to the diagonal $\Delta=\{(p,p)\mid p\in M\}$.*

Example 6.1.4. *Not every symplectomorphism is Hamiltonian: Hamiltonian symplectomorphisms are isotopic to the identity. Even symplectomorphisms which are isotopic to the identity might not be Hamiltonian; see for example* [**76**, Example 10.6].

Hamiltonian diffeomorphisms are rather special maps. Suppose that $L \subset (M, \omega)$ is a Lagrangian submanifold in a symplectic manifold, and $f\colon M \to M$ is a diffeomorphism smoothly isotopic to the identity. Then basic algebraic topology implies that if L and $f(L)$ intersect transversally, then (as the normal bundle of L is isomorphic to T^*L) we have that for any field $\mathbb{F}$

$$|L \cap f(L)| \geq \big|\chi(L)\big| = \big|\sum_{i=0}^{n} (-1)^i b_i(L; \mathbb{F})\big|,$$

where here $|L \cap f(L)|$ denotes the number of points in the intersection of L with $f(L)$.

By contrast, the Arnold conjecture gives a nice illustration of symplectic rigidity, giving a stronger lower bound on the number of intersection points between L and a Hamiltonian isotopic copy of L.

Conjecture 6.1.5. *(Arnold, cf.* [**76**, Conjecture 1.30]*) Fix a field $\mathbb{F}$. If L is a closed Lagrangian submanifold in the closed symplectic manifold (M^{2n}, ω) and $\phi\colon M^{2n} \to M^{2n}$ is a Hamiltonian diffeomorphism, with the property that L and $\phi(L)$ intersect transversally, then*

$$|L \cap \phi(L)| \geq \dim_{\mathbb{F}} H_*(L; \mathbb{F}) = \sum_{i=0}^{n} b_i(L; \mathbb{F}),$$

where here $|L \cap \phi(L)|$ denotes the number of points in the intersection of L with $\phi(L)$.

As a special case (where the ambient symplectic manifold is $(M \times M, p_1^*(\omega) - p_2^*(\omega))$, and the Lagrangian is the diagonal), we have the following:

Conjecture 6.1.6. *Fix a field $\mathbb{F}$. Let $\phi\colon M \to M$ be a Hamiltonian diffeomorphism of M, all of whose fixed points are non-degenerate. Then the number of fixed points of M is at least $\dim_{\mathbb{F}} H_*(M; \mathbb{F})$.*

Note the importance of having a Hamiltonian isotopy and not only a symplectic isotopy: the rotation of the symplectic torus encountered in Example 6.1.4 has no fixed points, and it is easy to find a Legendrian circle which is displaced disjointly, hence both Conjecture 6.1.5 and Conjecture 6.1.6 fail for this (non-Hamiltonian) symplectomorphism.

When $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$ and $\pi_2(M, L) = 0$, Conjecture 6.1.5 was verified by Floer [**21**]; we will discuss some of the techniques that go into its proof in this chapter.

6.2. Motivation

The main idea Floer invented to prove Conjecture 6.1.5 in the special case mentioned above is now called Floer homology (or, more precisely, Lagrangian Floer homology). In [**22**] these homology groups are introduced for symplectic manifolds (M,ω) with $\pi_2(M)=0$ and $\pi_2(M,L)=0$, see

To this end, Floer defined a chain complex over $\mathbb{F}=\mathbb{Z}/2\mathbb{Z}$, the *Lagrangian Floer complex* $\mathrm{CF}(L_0,L_1)$, associated to pairs of transversally intersecting, compact, oriented Lagrangian submanifolds L_0 and L_1. For simplicity, assume that $\pi_2(M)=\pi_2(M,L_0)=\pi_2(M,L_1)=0$. The chain complex is defined with the help of certain auxiliary choices, and it satisfies the following properties:

(1) Generators for $\mathrm{CF}(L_0,L_1)$ are intersection points of L_0 with L_1.

(2) The homology groups $HF(L_0,L_1)$ of the complex $\mathrm{CF}(L_0,L_1)$ are independent of the auxiliary choices; i.e. they depend only on the symplectic manifold and its two Lagrangian submanifolds. These homology groups $HF(L_0,L_1)$ are called the *Lagrangian Floer homology* of the Lagrangians L_0 and L_1.

(3) The homology groups $HF(L_0,L_1)$ are invariant under Hamiltonian isotopy, in the following sense. If $\phi_0\colon M\to M$ and $\phi_1\colon M\to M$ are two Hamiltonian diffeomorphisms, and $\phi_0(L_0)$ and $\phi_1(L_1)$ intersect transversally, then

$$HF(L_0,L_1)\cong HF(\phi_0(L_0),\phi_1(L_1)).$$

This property allows us to extend $HF(L_0,L_1)$ to cases where L_0 and L_1 do not intersect transversally: in this case, let L_1' be a Hamiltonian translate of L_1 which intersects L_0 transversally, and define $HF(L_0,L_1)$ as $HF(L_0,L_1')$.

(4) The homology groups $HF(L,L)$ are isomorphic to the singular homology of L (with $\mathbb{Z}/2\mathbb{Z}$ coefficients).

(5) The Floer homology groups $HF(L_0,L_1)$ are graded by $\mathbb{Z}/2\mathbb{Z}$, in the following way. If L_0 and L_1 are oriented, each intersection point of L_0 with L_1 has a local intersection number $i(\mathbf{x})$. The $\mathbb{Z}/2\mathbb{Z}$ grading $\mathrm{gr}(\mathbf{x})$ of a generator $\mathbf{x}\in L_0\cap L_1$ is determined by

$$(-1)^{\mathrm{gr}(\mathbf{x})}=i(\mathbf{x}).$$

The existence of a theory satisfying the first four of the above properties immediately implies the Arnold Conjecture. Moreover, in view of the first three properties, Lagrangian Floer homology can be thought of as an obstruction to making L_0 disjoint from L_1 via Hamiltonian isotopies. Note that if L_0

and L_1 are oriented, there is a classical algebro-topological obstruction for making the two Lagrangians disjoint, provided by the oriented intersection number of L_0 with L_1, which we write $\#(L_0 \cap L_1)$ (though, of course, that is invariant under much more general motions of L_0 and L_1). In view of the final property, Lagrangian Floer homology is a homology theory whose Euler characteristic is this intersection number.

The original set-up of Floer can be generalized in various directions. For example, one can relax the hypothesis on the second homotopy groups; or one may wish to work with coefficients over other rings. For further advances in this theory, and generalizations of Floer's result, see [**28, 29, 109**]. For example, Conjecture 6.1.6 was proved for $\mathbb{F} = \mathbb{Q}$ by Fukaya-Ono [**29**].

6.3. The action functional

Lagrangian Floer homology $HF(L_0, L_1)$ is an adaptation of Morse theory for a functional on an infinite-dimensional space, where the critical points of the functional correspond to intersection points between L_0 and L_1, and the gradient flowlines correspond to certain pseudo-holomorphic curves. (See [**63, 18**] for some other versions of Floer homology.) The aim of the present section is to given an informal account of how the Lagrangian Floer complex can be obtained from this infinite-dimensional point of view.

The basic set-up is a symplectic manifold (M^{2n}, ω) equipped with a pair of compact Lagrangian submanifolds L_0 and L_1. Consider the space of paths from L_0 to L_1 in M, i.e.

$$\mathcal{V} = \{\gamma\colon [0,1] \to M \mid \gamma(0) \in L_0, \gamma(1) \in L_1\}. \tag{6.1}$$

For simplicity, think of these paths as being smooth. The space of paths splits into path components, indexed by homotopy classes of paths in $\mathcal{V}$.

We would like to think of the space $\mathcal{V}$ as an infinite-dimensional manifold. A path in the path space is naturally a map $u\colon [0,1] \times [0,1] \to M^{2n}$, where we name coordinates in $[0,1] \times [0,1]$ by (t,s) satisfying the constraints that for all $t \in [0,1]$, $u(t,0) \in L_0$ and $u(t,1) \in L_1$. Let $\gamma(s) = u(0,s) \in \mathcal{V}$. A tangent vector at γ is a tangent vector field $v\colon [0,1] \to TM$ lifting γ, satisfying the conditions that

$$v(0) \in T_{\gamma(0)}L_0 \subset T_{\gamma(0)}M \qquad v(1) \in T_{\gamma(1)}L_1 \subset T_{\gamma(1)}M.$$

To simplify our discussion, we begin with the case where the symplectic manifold (M^{2n}, ω) is exact, so that $\omega = d\alpha$, and the Lagrangians are exact, in the following sense:

Definition 6.3.1. *Let $(M^{2n}, d\alpha)$ be an exact symplectic manifold. A Lagrangian submanifold $L \subset M$ is called **exact** if there is a function $f\colon L \to \mathbb{R}$ so that $\alpha|_L = df$.*

For example, the graph of an exact 1-form on a manifold L is an exact Lagrangian submanifold in the exact symplectic manifold T^*L.

Suppose that L_0 and L_1 are exact Lagrangians, and choose f_i for $i = 0, 1$ with $\alpha|_{L_i} = df_i$. There is a real valued *action functional* on $\mathcal{V}$, defined by the expression

$$\mathcal{A}(\gamma) = f_0(\gamma(0)) - f_1(\gamma(1)) + \int_{[0,1]} \gamma^*(\alpha). \tag{6.2}$$

It is convenient to reformulate the action as follows. Suppose that γ_0 and γ_1 are elements of $\mathcal{V}$ that are in the same path component; i.e. there is a map $u\colon [0,1] \times [0,1] \to M$ so that for all $t \in [0,1]$, $u(t,0) \in L_0$, $u(t,1) \in L_1$; and $u(0,s) = \gamma_0(s)$ and $u(1,s) = \gamma_1(s)$. Then, by Stokes' theorem,

$$\mathcal{A}(\gamma_1) - \mathcal{A}(\gamma_0) = \int_{(t,s)\in[0,1]\times[0,1]} u^*(\omega). \tag{6.3}$$

Floer homology will be constructed by applying Morse theory to this action functional. Thus, the first step is to identify its critical points.

Proposition 6.3.2. *The critical points of $\mathcal{A}$ are the constant paths $\gamma\colon [0,1] \to L_0 \cap L_1$; i.e. they correspond to points in $L_0 \cap L_1$.*

Proof. A path $\gamma\colon [0,1] \to M$ with $\gamma(0) \in L_0$ and $\gamma(1) \in L_1$ is a critical point for $\mathcal{A}$ if for every smooth extension of $\gamma(s)$ to $u\colon [-\epsilon, \epsilon] \times [0,1] \to M$ so that $u(t,0) \in L_0$ and $u(t,1) \in L_1$, and $\gamma(s) = u(0,s)$, the value $t = 0$ is a critical point for the real-valued function on $[-\epsilon, \epsilon]$ given by $t \mapsto \mathcal{A}(u|_{\{t\}\times[0,1]})$. But

$$\begin{aligned}
\frac{d}{dt}\Big|_{t=0} \mathcal{A}(u|_{\{t\}\times[0,1]}) &= \frac{\partial}{\partial t}\Big|_{t=0} \int_{[0,t]\times[0,1]} u^*(\omega) \\
&= \frac{\partial}{\partial t}\Big|_{t=0} \int_{[0,t]\times[0,1]} \omega(\frac{\partial u}{\partial t}, \frac{\partial u}{\partial s})\, dt \wedge ds \\
&= \int_{[0,1]} \omega(\frac{\partial u}{\partial t}, \frac{\partial u}{\partial s}) ds.
\end{aligned}$$

Clearly, if γ is a constant path, we have $\frac{\partial u}{\partial s} \equiv 0$, so $\frac{d}{dt}\Big|_{t=0} \mathcal{A}(u|_{\{t\}\times[0,1]}) = 0$ for all variations u. Conversely, since $v(s) = \frac{\partial u}{\partial t}(0,s)$ can be chosen to be

an arbitrary smooth function with compact support in $(0,1)$, and ω is non-degenerate, we can conclude that at each critical point we have $\frac{\partial u}{\partial s} \equiv 0$; i.e. $\gamma(s) = u(0,s)$ is a constant path. □

Having fixed ω, we choose an ω-compatible almost-complex structure J, which in turn gives a Riemannian metric on M. A Riemannian metric on M naturally gives rise, at least formally, to a Riemannian metric on the space of paths, by the formula

$$\langle v, w \rangle = \int_\gamma \langle v(s), w(s) \rangle ds,$$

where v and w are two vector fields along the path γ.

In computing the gradient of $\mathcal{A}$, we use the computation from the proof of Proposition 6.3.2. For a one-parameter variation of paths, we find that

$$\frac{\partial}{\partial t}\Big|_{t=0} \mathcal{A}(s \mapsto u(t,s)) = \int_{[0,1]} \omega(v, \frac{\partial u}{\partial s}) ds,$$

where $v(s) = \frac{\partial u}{\partial t}(0,s)$. Since $\omega(v, \frac{\partial u}{\partial s}) = -g(v, J\frac{\partial u}{\partial s})$, it follows that the gradient $\vec{\nabla}\mathcal{A}_\gamma \in T_\gamma \mathcal{V}$ of $\mathcal{A}$ at γ is the vector field along γ given by

$$s \mapsto -J_{\gamma(s)} \frac{d\gamma}{ds}.$$

Let $\mathbf{x}$ and $\mathbf{y}$ be two intersection points between L_0 and L_1. An upward gradient flowline for the action functional connecting $\mathbf{x}$ to $\mathbf{y}$ can be formulated as a pseudo-holomorphic strip; i.e. a map $u\colon \mathbb{R} \times [0,1] \to M^{2n}$ satisfying the following partial differential equation, a version of the Cauchy-Riemann equations:

$$\frac{\partial u}{\partial t} + J \frac{\partial u}{\partial s} = 0, \tag{6.4}$$

subject to the boundary conditions

$$u(t,0) \in L_0 \qquad u(t,1) \in L_1$$

(for any $t \in \mathbb{R}$), and asymptotics:

$$\lim_{t \to -\infty} u(t,s) = \mathbf{x} \qquad \lim_{t \to +\infty} u(t,s) = \mathbf{y},$$

in the sense that $u|_{t \times [0,1]}$ converges as $t \to -\infty$ uniformly to the constant function $\mathbf{x}$ (and similarly for $t \to \infty$ with $\mathbf{y}$).

The reader should keep in mind that the endomorphism J on $T_{u(t,s)}M^{2n}$ appearing in Equation (6.4) depends on the value $u(t,s)$; it would be more precise to express Equation (6.4) as

$$\frac{\partial u}{\partial t} + J_{u(t,s)} \frac{\partial u}{\partial s} = 0,$$

calling attention to the non-linear character of the Cauchy-Riemann equations.

The above discussion can be generalized to cases where M and the Lagrangian submanifolds are not necessarily exact. In this case, we define the action functional on the path components of $\mathcal{V}$, using Equation (6.3) as a definition. First, it is well defined only up to an overall additive constant. But more significantly, the value of $\mathcal{A}(\gamma_1) - \mathcal{A}(\gamma_0)$ depends on the choice of the path u in $\mathcal{V}$, at least up to homotopy. However, if we consider elements γ_0 and γ_1 of $\mathcal{V}$ that are sufficiently close, Equation (6.3) makes sense, provided we choose the path u to be sufficiently short; i.e. $\mathcal{A}$ is well-defined locally, so that its derivative

$$d\mathcal{A}_\gamma \colon \Gamma(\gamma^*(TM)) \to \mathbb{R},$$

which is given by $v \mapsto \int_{[0,1]} \omega(v, \frac{\partial u}{\partial s}) ds$ is a closed 1-form on $\mathcal{V}$, which is dual to the vector field $-J\frac{\partial}{\partial s}$. In this case, the Cauchy-Riemann equations are still the suitable analogues of the gradient flow equations.

The term "pseudo-holomorphic strip" is meant to remind the reader that the complex structure J on the target M is not necessarily integrable. In cases where the complex structure is integrable, i.e. the target is a complex manifold, the notion agrees with the standard notion of holomorphic map.

Since we are about to study the Morse-Smale complex, it is important to understand what it means for a critical point (i.e. an intersection point between L_0 and L_1) to be non-degenerate. We claim that if $\gamma \colon [0,1] \to M$ is the constant path at $\mathbf{x}$, thought of as an element of $\mathcal{V}$, then the Hessian of $\mathcal{A}$ at γ is, formally, the operator $v \mapsto -J\frac{dv}{ds}$, where $v \in T_\gamma \mathcal{V}$; and so the null-space of the Hessian is identified with the space of vectors in $T_{\mathbf{x}} L_0 \cap T_{\mathbf{x}} L_1$.

To justify this, recall the description of the Hessian $\mathrm{Hess}_p \colon TX \to TX$ at a critical point p of a function $f \colon X \to \mathbb{R}$ for a finite dimensional manifold X from Equation (1.1). In this spirit, suppose that $u \colon [-\epsilon, \epsilon] \times [-\epsilon, \epsilon] \times [0,1] \to M$ is a two-parameter family of paths, indexed by $(\tau, t) \in [-\epsilon, \epsilon] \times [-\epsilon, \epsilon]$ and with $u(0,0,s) = \gamma(s)$. Denote the two tangent vectors in $T_\gamma \mathcal{V}$ corresponding to $\frac{\partial}{\partial t}\Big|_{t=0} u(0,t,s)$ and $\frac{\partial}{\partial \tau}\Big|_{\tau=0} u(\tau,0,s)$ by $\xi, \eta \in T_\gamma \mathcal{V}$.

To compute the Hessian, we consider the above map u as a map $u \colon [-\epsilon, \epsilon] \times [-\epsilon, \epsilon] \to \mathcal{V}$, with $\frac{\partial u}{\partial \tau} = \xi$ and $\frac{\partial u}{\partial t} = \eta$. Now,

$$\widetilde{\eta}(\mathcal{A} \circ u) = \langle \vec{\nabla} \mathcal{A}, \frac{\partial u}{\partial t} \rangle_{T\mathcal{V}} = \int_{[0,1]} \langle -J\frac{\partial u}{\partial s}, \frac{\partial u}{\partial t} \rangle_{TM}$$

The Hessian, which is characterized by, $\langle \mathrm{Hess}_{\mathcal{A}}(\xi), \eta)\rangle = \xi(\widetilde{\eta}\mathcal{A})$, is computed by

$$\begin{aligned}
\xi(\widetilde{\eta}(\mathcal{A}\circ u)) &= \frac{\partial}{\partial\tau}\big|_{(\tau,t)=(0,0)} \int_{[0,1]} \langle -J\frac{\partial u}{\partial s}, \frac{\partial u}{\partial t}\rangle_{TM} \\
&= \frac{\partial}{\partial\tau}\big|_{(\tau,t)=(0,0)} \int_{[0,1]} \langle J\frac{\partial u}{\partial t}, \frac{\partial u}{\partial s}\rangle ds \\
&= \left[\int_{[0,1]} \langle J\frac{\partial u}{\partial t}, \frac{\partial^2 u}{\partial s\partial\tau}\rangle ds\right]_{(\tau,t)=(0,0)} \\
&= \left[\int_{[0,1]} \langle -J\frac{\partial^2 u}{\partial s\partial\tau}, \frac{\partial u}{\partial t}\rangle ds\right]_{(\tau,t)=(0,0)} \\
&= \int_{[0,1]} \langle -J\frac{\partial \xi}{\partial s}, \eta\rangle ds.
\end{aligned}$$

Thus, the $\mathrm{Hess} = -J\frac{\partial\xi}{\partial s}$.

It follows that the action functional is formally a non-degenerate Morse function if and only if L_0 and L_1 intersect one another transversely. Specifically, at a critical point u_0, which corresponds to an intersection $x \in L_0 \cap L_1$, a tangent vector $\xi \in T_{u_0}\mathcal{V}$ is a map $v\colon [0,1] \to T_xM$ with $v(0) \in T_xL_0$ and $v(1) \in T_xL_1$. That tangent vector lies in the nullspace of the Hessian if the tangent vector v is constant; i.e. v corresponds to a vector in $(T_xL_0) \cap (T_xL_1)$.

Unlike the finite-dimensional case, the gradient does not really define a flow, in the usual sense, on the space of paths from L_0 to L_1. In the finite-dimensional case, the gradient flow equation is an ordinary differential equation, while in the present case, pseudo-holomorphic strips satisfy a partial differential equation.

There is another feature that distinguishes the classical case from the case for Lagrangian intersections: the Hessian at a critical point does not have finite index, that is, has infinitely many negative and positive eigenvalues. For example, if the target manifold is $\mathbb{C}$, $L_0 = \mathbb{R}$ and $L_1 = e^{2\pi i\theta}\cdot\mathbb{R}$, and we consider the constant path (at the origin), the eigenvectors for the Hessian are functions of the form $s \mapsto re^{2\pi is(\theta+n)}$ for $n \in \frac{1}{2}\mathbb{Z}$ and the eigenvalues of the Hessian are $2\pi(\theta + \frac{1}{2}\mathbb{Z})$. In particular, the intersection point $0 \in \mathbb{C}$ is non-degenerate when $L_0 \neq L_1$ (i.e. $\theta \notin \frac{1}{2}\mathbb{Z}$). However, there are infinitely many negative and positive eigenvalues.

Nonetheless, there is a quantity, the *Maslov index*, which plays the role of the *difference* in indices between two critical points. To formulate its properties, it is helpful to formalize the algebraic topology underlying pseudo-holomorphic strips.

6.4. Whitney disks

Having motivated the study of pseudo-holomorphic disks, we turn now to the more systematic study of the objects needed to be defined in the Lagrangian Floer context. In this direction, we find it convenient to formalize the boundary conditions placed on pseudo-holomorphic strips in the previous section, as follows:

Definition 6.4.1. *Fix* $\mathbf{x}, \mathbf{y} \in L_0 \cap L_1$. *A* ***Whitney strip from*** $\mathbf{x}$ ***to*** $\mathbf{y}$ *is a continuous map* $u\colon \mathbb{R} \times [0,1] \to M^{2n}$ *satisfying the boundary conditions* $u(\mathbb{R} \times \{0\}) \subset L_0$, $u(\mathbb{R} \times \{1\}) \subset L_1$, *and the asymptotics*

$$\lim_{t \to -\infty} u(t,s) = \mathbf{x} \qquad \lim_{t \to +\infty} u(t,s) = \mathbf{y}.$$

It is natural to reformulate these conditions in terms of maps of disks, as follows.

Definition 6.4.2. *Let* $\mathbb{D}$ *denote the standard unit disk in* $\mathbb{C}$, *and fix* $\mathbf{x}, \mathbf{y} \in L_0 \cap L_1$. *A* ***Whitney disk from*** $\mathbf{x}$ ***to*** $\mathbf{y}$ *is a continuous map* $u\colon \mathbb{D} \to M^{2n}$ *satisfying the boundary conditions:*

$$u(z) \in L_0 \text{ if } |z| = 1 \text{ and } \mathrm{Re}(z) > 0; \qquad u(z) \in L_1 \text{ if } |z| = 1 \text{ and } \mathrm{Re}(z) < 0.$$
$$u(-i) = \mathbf{x} \qquad u(i) = \mathbf{y}$$

Each pseudo-holomorphic strip gives rise to a Whitney disk in the above sense, using a conformal diffeomorphism $\mathbb{R} \times [0,1] \cong \mathbb{D} \setminus \{\pm i\}$, sending $\mathbb{R} \times 0$ to the semicircle in $\partial\mathbb{D}$ with $\mathrm{Re}(z) > 0$.

Remark 6.4.3. *The domain of a Whitney disk is a disk with two distinguished points on its boundary. Later, we will consider analogues with more distinguished boundary points, leading us to* ***Whitney polygons*** *(see Section 6.10); in those cases, we will refer to the above "Whitney disks" as* ***Whitney bigons****.*

Definition 6.4.4. *Let* $\mathbb{D}$ *denote the standard unit disk in* $\mathbb{C}$, *and fix* $\mathbf{x}, \mathbf{y} \in L_0 \cap L_1$. *Fix two Whitney disks* u_0 *and* u_1 *from* $\mathbf{x}$ *to* $\mathbf{y}$. *A* ***homotopy*** *from* u_0 *to* u_1 *is a continuous, one-parameter family of Whitney disks from* $\mathbf{x}$ *to* $\mathbf{y}$*; more precisely, it is a map* $u\colon \mathbb{D} \times [0,1] \to M$ *satisfying the conditions*

(with $z \in \mathbb{D}$ and $t \in [0,1]$)

$$u(z,t) \in L_0 \text{ if } |z| = 1 \text{ and } \mathrm{Re}(z) > 0 \quad u(z,t) \in L_1 \text{ if } |z| = 1 \text{ and } \mathrm{Re}(z) < 0$$
$$u(-i,t) = \mathbf{x} \quad u(i,t) = \mathbf{y}$$
$$u(z,0) = u_0(z) \quad u(z,1) = u_1(z).$$

The set of homotopy classes of Whitney disks from $\mathbf{x}$ to $\mathbf{y}$ is denoted $W(\mathbf{x},\mathbf{y})$. Whitney disks can be composed, providing a map $*\colon W(\mathbf{x}_1,\mathbf{x}_2) \times W(\mathbf{x}_2,\mathbf{x}_3) \to W(\mathbf{x}_1,\mathbf{x}_3)$ as follows:

Definition 6.4.5. *Given $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3 \in L_0 \cap L_1$, and Whitney disks u_1 from $\mathbf{x}_1$ to $\mathbf{x}_2$ and u_2 from $\mathbf{x}_2$ to $\mathbf{x}_3$, there is a natural **juxtaposition** $u_1 * u_2$, which is a Whitney disk from $\mathbf{x}_1$ to $\mathbf{x}_3$. To define this, use the quotient map $q\colon \mathbb{D} \to \mathbb{D} \vee \mathbb{D}$, where first we take the map from $\mathbb{D}$ to the quotient space gotten by collapsing the real interval to a point, and then using a homeomorphism between this quotient space with*

$$\mathbb{D} \vee \mathbb{D} = \frac{\mathbb{D}_1 \coprod \mathbb{D}_2}{(i \in \mathbb{D}_1) \sim (-i \in \mathbb{D}_2)}.$$

The map $u_1 * u_2$ is the composite

$$\mathbb{D} \xrightarrow{q} \mathbb{D} \vee \mathbb{D} \xrightarrow{u_1 \vee u_2} M.$$

Remark 6.4.6. *The set $W(\mathbf{x},\mathbf{y})$ can be identified with the set of homotopy classes of maps connecting $\mathbf{x}$ and $\mathbf{y}$, where $\mathbf{x}$ and $\mathbf{y}$ are regarded as elements of $\mathcal{V}$ (as constant maps). In this approach, the juxtaposition defined above corresponds to juxtaposition of paths. The set $W(\mathbf{x},\mathbf{y})$ can be identified (but not canonically) with the group $W(\mathbf{x},\mathbf{x}) \cong \pi_1(\mathcal{V},\mathbf{x})$. To understand this group better, observe that evaluation at the two endpoints induces a fibration $\mathcal{V} \to L_0 \times L_1$. The homotopy fiber of this map is identified with the loop space $\Omega(M)$ (based at any point in M). It follows that there is a short exact sequence*

$$\pi_2(M) \to \pi_1(\mathcal{V}) \to \pi_1(L_0) \oplus \pi_1(L_1) \to 0.$$

For each $\mathbf{x} \in L_0 \cap L_1$, there is an action of $\pi_2(M, L_0; \mathbf{x})$ on the set $W(\mathbf{x},\mathbf{y})$. This action can be defined by simply viewing an element ψ of $\pi_2(M, L_0; \mathbf{x})$ as an element of $W(\mathbf{x},\mathbf{x})$, and then letting the action of ψ on $\phi \in W(\mathbf{x},\mathbf{y})$ be the juxtaposition of these two disks. We denote this action by $\psi \star_0 \phi$. It is easy to check that $(\psi \star_0 \phi_1) * \phi_2 = \psi \star_0 (\phi_1 * \phi_2)$.

Suppose that u is a Whitney disk from $\mathbf{x}$ to $\mathbf{y}$, and suppose that $v\colon \mathbb{D} \to M$ is a continuous map with the property that $v(\partial\mathbb{D}) \subset L_0$, and $v(i) = u(1)$. Then, u and v can be combined to give a continuous map from $\mathbb{D}$ to M, by composing $\mathbb{D} \to \mathbb{D}\vee\mathbb{D}$ with the map $u \vee v\colon \mathbb{D}\vee\mathbb{D} \to M$. It is straightforward to check that $[u \vee v] = [v'] \star [u]$, where $v' \in \pi_2(M, L_0, \mathbf{x})$ is the homotopy

class obtained by acting on $v \in \pi_2(M, L_0, u(1))$ by some path from $\mathbf{x}$ to $u(1)$.

Similarly, $\psi \in \pi_2(M, L_1; \mathbf{x})$ also acts on $\phi \in W(\mathbf{x}, \mathbf{y})$; that action is now denoted $\phi \star_1 \psi \in W(\mathbf{x}, \mathbf{y})$.

Finally, $S \in \pi_2(M; \mathbf{x})$ acts on $\phi \in W(\mathbf{x}, \mathbf{y})$ via the natural quotient map $\pi_2(M; \mathbf{x}) \to \pi_2(M, L_0; \mathbf{x})$. We denote the result by $\phi \star S \in W(\mathbf{x}, \mathbf{y})$. This has a similar geometric interpretation: if u is a Whitney disk from $\mathbf{x}$ to $\mathbf{y}$, and $f\colon S^2 \to M$ is a map with $u(0) = f(p)$ for some $p \in S^2$, then we can join these two maps by composing $\mathbb{D} \to \mathbb{D} \vee S^2$ with the map $u \vee f\colon \mathbb{D} \vee S^2 \to M$. Then, $[u \vee f] = [u] \star [f']$, where $f' \in \pi_2(M; \mathbf{x})$ is obtained by acting on $v \in \pi_2(M; u(0))$ by some path. (This path in turn can can be taken to be the image under u of any path connecting 0 with $-i$.)

6.5. The Maslov index

Thinking of an intersection point $\mathbf{x} \in L_0 \cap L_1$ as a critical point of the action functional, the corresponding Hessian has infinitely many negative and positive eigenvalues. Nonetheless, if $\mathbf{x}$ and $\mathbf{y}$ are two intersection points, there is an associated integer, the Maslov index, that behaves like the index difference $\lambda(\mathbf{x}) - \lambda(\mathbf{y})$. The Maslov index will depend on a little more data than just the endpoints $\mathbf{x}$ and $\mathbf{y}$: it will be a homotopy invariant of Whitney disks. We give a construction presently; and refer the interested reader to [**117**].

Suppose that $u\colon \mathbb{R} \times [0,1] \to M^{2n}$ is a Whitney strip. The bundle $u^*(TM)$ is a bundle of symplectic vector spaces over $\mathbb{R} \times [0,1]$, with Lagrangian subbundles $(u|_{\mathbb{R}\times\{0\}})^*(TL_0)$ and $(u|_{\mathbb{R}\times\{1\}})^*(TL_1)$ over $\mathbb{R}\times\{0\}$ and $\mathbb{R}\times\{1\}$, respectively.

In particular, for fixed $t \in \mathbb{R}$, $V_t = u^*(TM)_{(t,0)}$ is a symplectic vector space with Lagrangian subspace $\Lambda_0^t = (u|_{\mathbb{R}\times\{0\}})^*(TL_0)_{(t,0)}$. Parallel transport across $\{t\} \times [0,1]$ (cf. Section 4.5.1) identifies the symplectic vector spaces $u^*(TM)_{(t,1)}$ with $u^*(TM)_{(t,0)}$; thus, the Lagrangian $(u|_{\mathbb{R}\times\{1\}})^*(TL_1)_{(t,1)}$ can be transported to give a new Lagrangian Λ_1^t inside V_t. Under the identification $\mathcal{L}Gr(V_t) \cong U(n)/O(n)$ that identifies Λ_0^t with the identity coset (cf. Remark 4.5.2), the Lagrangian subspace Λ_1^t corresponds to an element $\Lambda_t \in U(n)/O(n)$. (See also the end of Section 4.5 for a discussion on parallel transport.)

Recall that $U(n)/O(n)$ contains a stratified, co-oriented subset $\Sigma = \Sigma(\mathbb{R}^n) \subset U(n)/O(n)$, the Maslov cycle, as specified in Equation (4.14).

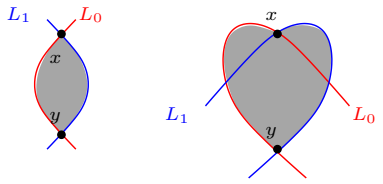


Figure 6.1. The Maslov index in the plane. In the two pictures, we have Lagrangians L_0 and L_1 which are contained in the plane. The shadings represent two homotopy classes of disks from x to y. Comparing the tangent spaces to the red and the blue curves, it is easy to see that the first disk has Maslov index 1 and the second has Maslov index 2.

Definition 6.5.1. *The intersection number of the curve $\{\Lambda_t\}_{t\in\mathbb{R}}$ with the Maslov cycle Σ is the **Maslov index** of u. It is denoted $\mu(u)$.*

The intersection number is well-defined since $\lim_{t\mapsto\pm\infty}\Lambda_t \notin \Sigma$. The following operations change the path by a homotopy fixing the asymptotics: moving u by a homotopy of Whitney disks, and changing the symplectic connection used for parallel transport. It follows that the Maslov index is independent of these choices.

See Figure 6.1 for some examples, in cases where $M^{2n} = \mathbb{C}$, and L_0, L_1 are the red and blue curves.

We establish a few further properties of the Maslov index in Proposition 6.5.7 below, using a slightly different construction for it, which we give after setting up a few preliminaries.

For the following definition, recall that for each $\Lambda_0 \in \mathcal{L}Gr(V,\omega)$, there is a subset $C_0(\Lambda) \subset T_{\Lambda_0}\mathcal{L}Gr$ corresponding to the positive bilinear forms over the vector space Λ_0 (cf. Proposition 4.5.11). For any path $\Lambda\colon [0,1] \to \mathcal{L}Gr(V,\omega)$ with $\Lambda(0) = \Lambda_0$ and

$$\Lambda'(0) = \frac{d\Lambda}{dt}(0) \in C_0(\Lambda_0),$$

there is an $\epsilon > 0$ so that $\Lambda(t) \notin \Sigma(\Lambda)$ for all $0 < t < \epsilon$.

Definition 6.5.2. *Let (V,ω) be a symplectic vector space. Suppose that $\Lambda_0, \Lambda_1 \subset V$ are two Lagrangian subspaces that meet transversely. A smooth path $\{\Lambda_t\}_{t\in[0,1]}$ from Λ_0 to Λ_1 is called **preferred** if it satisfies the following properties:*

- *(P-1)* $\Lambda'(0) \in C_0(\Lambda_0)$.
- *(P-2)* Λ_t *meets* $\Sigma(\Lambda_0)$ *transversely at all* $t \in (0,1]$*; i.e.* Λ_t *is disjoint from* $\Sigma_k(\Lambda_0)$ *for* $k > 1$ *and* Λ_t *meets* $\Sigma_1(\Lambda_0)$ *transversely.*
- *(P-3) The algebraic intersection number of* $\{\Lambda_t\}_{t\in(0,1]}$ *with* $\Sigma(\Lambda_0)$ *is zero.*

Lemma 6.5.3. *Let (V,ω) be a symplectic vector space. For any two Lagrangian subspaces $\Lambda_0, \Lambda_1 \subset V$ that meet transversely the following statements hold:*

(1) (Existence) Λ_0 and Λ_1 can be connected by a preferred path.

(2) (Homotopy uniqueness) Any two preferred paths are homotopic relative to the their endpoints, through paths satisfying Property (P-1).

Proof. A *pre-preferred path* $\{\Lambda_t\}$ is a smooth path from Λ_0 to Λ_1 satisfying Properties (P-1) and (P-2).

Given Λ_0 and Λ_1, we can find a pre-preferred path $\{\Lambda_t\}$ since the Lagrangian Grassmannian is connected, $C_0(\Lambda_0)$ is connected, and Λ_0 is in the closure of $C_0(\Lambda_0)$. Since $\pi_1(\mathcal{L}Gr(V,\omega)) = H_1(\mathcal{L}Gr(V,\omega)) \cong \mathbb{Z}$, and $[\Sigma]$ generates $H^1(\mathcal{L}Gr(V,\omega);\mathbb{Z})$ (Proposition 4.5.11), it follows that there is a closed loop γ in $\mathcal{L}Gr(V,\omega)$ for which $\#(\gamma \cap \Sigma(\Lambda_0)) = 1$. Thus, by modifying $\{\Lambda_t\}$ if necessary by adding as many copies of γ as needed, we can arrange for Property (P-3) to hold, as well. This establishes existence.

To see homotopy uniqueness, we argue as follows. Fix two preferred paths $\Lambda, \widetilde{\Lambda}\colon [0,1] \to \mathcal{L}Gr(V,\omega)$ from Λ_0 to Λ_1. Since the space $C_0(\Lambda_0)$ is contractible, we can find an $\epsilon > 0$ and a homotopy between $\Lambda|_{[0,\epsilon]}$ and $\widetilde{\Lambda}|_{[0,\epsilon]}$ through paths starting at Λ_0, with non-zero derivative at 0, and whose restriction to $(0,\epsilon]$ lies in $C_0(\Lambda_0)$. Extending this homotopy to $[\epsilon,1]$ is equivalent to giving a null homotopy of the closed loop δ obtained by juxtaposing $\Lambda|_{[\epsilon,1]}$, the reverse of $\widetilde{\Lambda}_{[\epsilon,1]}$, and the path in $C_0(\Lambda)$ obtained by following the homotopy between $\Lambda|_{[0,\epsilon]}$ and $\widetilde{\Lambda}|_{[0,\epsilon]}$ at $t = \epsilon$. But this null homotopy can be constructed, since the intersection number of

$$\#\Sigma(\Lambda_0)\cap\delta = \#\Sigma(\Lambda_0)\cap\Lambda|_{[\epsilon,1]} - \#\Sigma(\Lambda_0)\cap\widetilde{\Lambda}|_{[\epsilon,1]} = \#\Sigma(\Lambda_0)\cap\Lambda - \#\Sigma(\Lambda_0)\cap\widetilde{\Lambda} = 0,$$

and $\pi_1(\mathcal{L}Gr(V,\omega)) = H_1(\mathcal{L}Gr(V,\omega)) \cong \mathbb{Z}$, and $[\Sigma]$ generates $H^1(\mathcal{L}Gr(V,\omega);\mathbb{Z})$. See Figure 6.2.

□

Remark 6.5.4. *The intersection number of a path of Lagrangian subspaces with Σ_1 has the following interpretation. Under the coordinate maps $\phi_{\Lambda,P}$ from the proof of Lemma 4.5.10, the path of Lagrangians corresponds to a path $\{A_t\}_{t\in[0,1]}$ of symmetric matrices, and Σ corresponds to those symmetric matrices with non-trivial kernel. If $\{A_t\}_{t\in[0,1]}$ is a generic path of symmetric matrices, then there are n distinct eigenvalues $\lambda_1(t) < \cdots < \lambda_n(t)$, all of which are non-zero at $t = 0, 1$. The intersection number of A_t with Σ_1 counts the number of negative eigenvalues of A_1 minus the number of*

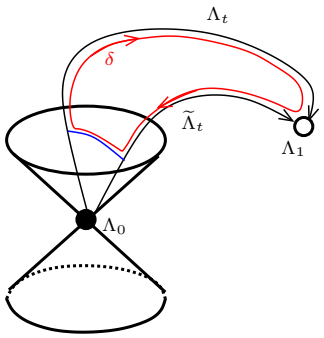


Figure 6.2. Homotoping preferred paths of Lagrangians.

negative eigenvalues of A_0*. This quantity is called the* ***spectral flow*** *of* $\{A_t\}_{t\in[0,1]}$*; see* [**118**].

We can now give our second definition of the Maslov index, given a symplectic manifold M and a pair of Lagrangians L_0 and L_1 meeting transversely.

Definition 6.5.5. *Choose first for each* $\mathbf{x} \in L_0 \cap L_1$ *a preferred path (in the sense of Definition 6.5.2)* $\{\Lambda_{\mathbf{x}}(t)\}_{t\in[0,1]}$ *from* $T_{\mathbf{x}}L_0$ *to* $T_{\mathbf{x}}L_1$ *in* $\mathcal{L}Gr(T_{\mathbf{x}}M)$*.*

Pre-compose $u\colon \mathbb{D} \to M$ *with the quotient map* $[0,1]\times[0,1] \to \mathbb{D}$ *which sends* $[0,1]\times\{0\}$ *to* $-i$ *and* $[0,1]\times\{1\}$ *to* i*, to obtain a continuous map* $\overline{u}\colon [0,1]\times[0,1] \to M$*. Thus,* $\overline{u}^*(TM)$ *is a bundle of symplectic vector spaces over* $[0,1]\times[0,1]$*, equipped with identifications*

$$\overline{u}^*(TM)|_{[0,1]\times\{0\}} \cong [0,1]\times T_{\mathbf{x}}M$$
$$\overline{u}^*(TM)|_{[0,1]\times\{1\}} \cong [0,1]\times T_{\mathbf{y}}M.$$

Consider the Lagrangian subbundle $\mathfrak{L}$ *of* $\overline{u}^*(TM)$ *over the boundary of* $[0,1]\times[0,1]$*, specified by*

$$\mathfrak{L}_{(s,0)} = \Lambda_{\mathbf{x}}(t) \subset T_{\mathbf{x}}M$$
$$\mathfrak{L}_{(s,1)} = \Lambda_{\mathbf{y}}(t) \subset T_{\mathbf{y}}M$$
$$\mathfrak{L}_{(0,t)} = T_{u(s,0)}L_0 \subset T_{u(s,0)}TM$$
$$\mathfrak{L}_{(1,t)} = T_{u(s,1)}L_1 \subset T_{u(s,1)}TM.$$

Using a symplectic trivialization of the bundle $\overline{u}^*(TM)$*, and thinking of the boundary of* $[0,1]\times[0,1]$ *as an oriented loop (using the boundary orientation), the Lagrangian subbundle gives rise to a loop* ℓ_u *of Lagrangian subsapces in a fixed symplectic vector space. The (second definition of the)* ***Maslov index*** *of* u *is obtained by applying the universal Maslov class to this loop.*

Proposition 6.5.6. *The Maslov index of u, as defined in Definition 6.5.1, coincides with evaluation of the universal Maslov class μ_n on the closed loop ℓ_u, as in Definition 6.5.5 above.*

Proof. Use parallel transport to identify the symplectic vector spaces $\overline{u}^*(TM)_{(s,t)}$ with $\overline{u}_{0,t}(TM)$; and hence, their corresponding Lagrangian Grassmannians

$$\mathcal{L}Gr(\overline{u}^*_{(s,t)}(TM)) \cong \mathcal{L}Gr(\overline{u}^*_{(0,t)}(TM)).$$

We compose these identifications with the identifcation $\mathcal{L}Gr(\overline{u}^*_{(0,t)}(TM) \cong U(n)/O(n)$ under which $T_{u(0,t)}L_0$ corresponds to the identity coset. Under these identifications,

- $\mathfrak{L}|_{(0,t)}$ corresponds to the constant path,
- $\mathfrak{L}|_{(s,0)}$ corresponds to the path $\Lambda_{\mathbf{x}}(s)$,
- $\mathfrak{L}|_{(s,1)}$ corresponds to the path $\Lambda_{\mathbf{y}}(s)$, and
- $\mathfrak{L}_{(1,t)}$ corresponds to the path Λ_1^t.

It follows that the algebraic intersection number of the Malsov cycle (represented by $\Sigma(\mathbb{R}^n)$) with three of the four boundary components is zero, and hence the intersection of the Maslov cycle with the loop of Lagrangians on the boundary coincides with the intersection number of Λ_1^t with the Maslov cycle, which in turn is the definition of the Maslov index from Definition 6.5.1. □

Proposition 6.5.7. *Homotopic Whitney disks have the same Maslov index, so we can think of μ as a function on $W(\mathbf{x}, \mathbf{y})$. The Maslov index is additive under juxtapositions, in the following ways:*

(M-1) If $\phi \in W(\mathbf{x}, \mathbf{y})$ and $\psi \in W(\mathbf{y}, \mathbf{z})$, then $\mu(\phi * \psi) = \mu(\phi) + \mu(\psi)$.

(M-2) If $\phi \in W(\mathbf{x}, \mathbf{y})$, and $\eta_i \in \pi_2(M, L_i; \mathbf{x})$ for $i = 0, 1$, then

$$\mu(\eta_0 \star_0 \phi \star_1 \eta_1) = \langle \mu_{L_0}, [\eta_0] \rangle + \mu(\phi) + \langle \mu_{L_1}, [\eta_1] \rangle.$$

(M-3) If $S \in \pi_2(M; \mathbf{x})$, then $\mu(\phi \vee S) = \mu(\phi) + 2\langle c_1(TM), [S] \rangle$.

Proof. A homotopy of Whitney disks moves the curve Λ_t by homotopies whose endpoints do not intersect Σ. By elementary algebraic topology, the intersection number of Λ_t with Σ is invariant under such isotopies.

Composition of Whitney disks corresponds to concatenation of paths, and the intersection number is additive under such concatenations, so Property (M-1) follows.

For the other two properties, we will use the interpretation of the Maslov index from Proposition 6.5.6. Specifically, in the proof of that proposition, the map u is extended to a map $\overline{u}\colon [0,1]\times[0,1]\to M$, with a specified Lagrangian subbundle $\mathcal{L}$ of $\overline{u}^*(TM)$. Analogous to the proof of Proposition 4.5.8, the evaluation of the universal Maslov class on the loop of Lagrangians can be viewed as a relative first Chern class of the bundle $\overline{u}^*(\det_{\mathbb{C}}(TM)\otimes_{\mathbb{C}}\det_{\mathbb{C}}(TM))$ on $[0,1]\times[0,1]$ relative to the trivialization $\det_{\mathbb{R}}(\mathfrak{L})\otimes\det_{\mathbb{R}}(\mathfrak{L})$ on its boundary (which in turn depends on the restriction of u to its boundary). Property (M-3) follows immediately.

Property (M-2) also follows from this interpretation of the Maslov index, now combined with the interpretation of the Maslov index for boundary degenerations from Proposition 4.5.8. □

Suppose next that L_0 and L_1 are oriented Lagrangian submanifolds that meet transversally. Given an intersection point $\mathbf{x}\in L_0\cap L_1$, let $\epsilon(\mathbf{x})\in\{\pm1\}$ be the signed intersection number of L_0 and L_1 at $\mathbf{x}$; i.e. the discrepancy between the orientation of $T_{\mathbf{x}}M$ at $\mathbf{x}$ specified by the orientation of M, and the orientation of $T_{\mathbf{x}}M\cong T_{\mathbf{x}}L_0\oplus T_{\mathbf{x}}L_1$ specifieid by the orientations of L_0 and L_1.

Proposition 6.5.8. *Suppose that L_0 and L_1 are oriented Lagrangian submanifolds in M. For any $\mathbf{x},\mathbf{y}\in L_0\cap L_1$ and $\phi\in W(\mathbf{x},\mathbf{y})$,*

$$(-1)^{\mu(\phi)}=\epsilon(\mathbf{x})\cdot\epsilon(\mathbf{y}).$$

Proof. Orient $\mathbb{R}^n\subset\mathbb{C}^n$. To any oriented Lagrangian subspace Λ that is transverse to $\mathbb{R}^n$, we can assign a sign $e(\Lambda)\in\{\pm1\}$ which is $+1$ exactly when the orientation $\mathbb{C}^n$ induced by its splitting $\mathbb{C}^n\cong\mathbb{R}^n\oplus\Lambda$ coincides with the orientation it inherits from its symplectic structure. The proposition is equivalent to the following claim: if $\{\Lambda_t\}_{t\in[0,1]}$ is a generic one-parameter family of oriented Lagrangian subspaces with Λ_0 and Λ_1 transverse to $\mathbb{R}^n$, then $e(\Lambda_0)-e(\Lambda_1)\equiv\#(\Lambda_t\cap\Sigma_1)\pmod 2$. This is obvious from the interpretation of the intersection number as spectral flow. □

6.6. Transversality

In finite dimensional Morse theory, for a sufficiently generic choice of metrics, the difference in Morse indices $\lambda(\mathbf{x})-\lambda(\mathbf{y})$ at the critical points $\mathbf{x}$ and $\mathbf{y}$ computes the dimension of the moduli space of gradient flows; cf. Theorem 5.1.2.

The present section concerns the analogue of this result for pseudo-holomorphic strips. To make the corresponding transversality work, we must relax slighly the notion of pseudo-holomorphic strip from the previous section, to include one-parameter families of almost-complex structures, as follows.

In the previous sections we chose ω-compatible almost-complex structures, which allowed simpler arguments in the motivational part of this chapter. The ω-tame condition is, however, turns out to be more convenient, as (in contrast wit compatibility) is an open condition. For this reason, from now on, the almost-complex structures will always be tame with respect to ω.

Definition 6.6.1. *Let $\{J^s\}_{s\in[0,1]}$ be a one-parameter family of ω-tame almost-complex structures on (M^{2n},ω). A $\{J^s\}$-**pseudo-holomorphic strip** $u\colon \mathbb{R}\times[0,1]\to M^{2n}$ is a Whitney strip in the sense of Definition 6.4.1, which satisfies the partial differential equation*

$$\frac{\partial u}{\partial t}+J^s\frac{\partial u}{\partial s}=0; \tag{6.5}$$

i.e. at each $(t,s)\in\mathbb{R}\times[0,1]$,

$$\frac{\partial u}{\partial t}+J^s_{u(t,s)}\frac{\partial u}{\partial s}=0,$$

*where $J^s_p\colon T_pM\to T_pM$ is the endomorphism determined by $\{J^s\}$. For fixed $\phi\in W(\mathbf{x},\mathbf{y})$, let the **moduli space** $\mathcal{M}_{\{J^s\}}(\phi)$ denote the set of pseudo-holomorphic representatives of ϕ.*

Since for any $\tau\in\mathbb{R}$, the map $\mathbb{R}\times[0,1]\to\mathbb{R}\times[0,1]$ given by $(t,s)\mapsto(t+\tau,s)$ is holomorphic, it follows that if $u\colon\mathbb{R}\times[0,1]\to M$ is pseudo-holomorphic, then so is the map $(t,s)\mapsto u(t+\tau,s)$. There is a resulting $\mathbb{R}$-action on $\mathcal{M}_{\{J^s\}}(\phi)$; the quotient by this action will be denoted by $\widehat{\mathcal{M}}_{\{J^s\}}(\phi)$.

With these notions in place, we have the following analogue of Theorem 5.1.2, showing the relevance of the Maslov index in the study of pseudo-holomorphic disks.

Theorem 6.6.2. *Let (M^{2n},ω) be a symplectic manifold, equipped with Lagrangians L_0 and L_1. If $\{J^s\}$ is a suitably generic one-parameter family of ω-tame almost-complex structures, then for any non-constant homotopy class $\phi\in W(\mathbf{x},\mathbf{y})$ with $\mu(\phi)\leq 2$, the space $\widehat{\mathcal{M}}_{\{J^s\}}(\phi)$ is a smooth manifold with dimension given by*

$$\dim\widehat{\mathcal{M}}_{\{J^s\}}(\phi)=\mu(\phi)-1.$$

In particular, if ϕ is a non-constant homotopy class with $\mu(\phi)\leq 0$, then $\widehat{\mathcal{M}}_{\{J^s\}}(\phi)$ is empty. Furthermore, if ϕ is the homotopy class represented by a constant Whitney disk, then $\widehat{\mathcal{M}}_{\{J^s\}}(\phi)$ consists of a single point (i.e. it is the constant flowline).

The moduli spaces $\mathcal{M}_{\{J^s\}}(\phi)$ are thought of as the solution set of a nonlinear elliptic partial differential equation whose linearization has index computed by the Maslov index, according to a theorem of Floer [**21**]. The proof of smoothness involves an infinite-dimensional application of the Sard-Smale theorem. See Theorem 35.2.1; compare also [**23**].

6.7. Compactness

By analogy with Morse theory, we define $\mathrm{CF}(L_0, L_1)$ to be the $\mathbb{F}$-vector space generated by intersection points between L_0 and L_1. To turn $\mathrm{CF}(L_0, L_1)$ into a chain complex, we attempt to define the boundary operator as

$$\partial \mathbf{x} = \sum_{\mathbf{y} \in L_0 \cap L_1} \sum_{\{\phi \in W(\mathbf{x},\mathbf{y}) \,\big|\, \mu(\phi)=1\}} \#\widehat{\mathcal{M}}(\phi) \cdot \mathbf{y}. \tag{6.6}$$

(For simplicity, from now on we will omit the indication of the choice of the one-parameter family $\{J^s\}$ from the notation of the moduli space.) As in the case of the Morse-Smale complex, to make sense of the above definition, it remains to address compactness issues, which we do here; and to verify that it is indeed a chain complex, we will need another ingredient, gluing (addressed in the next section).

To make sense of Equation (6.6), we must verify that the coefficient of $\mathbf{y}$ is finite. This problem can be divided into two pieces. The first is showing that for fixed $\phi \in W(\mathbf{x}, \mathbf{y})$, the number of points in $\widehat{\mathcal{M}}(\phi)$ is finite. The second is that for fixed $\mathbf{x}$ and $\mathbf{y}$, there are only finitely many $\phi \in W(\mathbf{x}, \mathbf{y})$ with non-zero contribution.

While the second problem is handled in various ways, as the setting requires, the key point to addressing the first problem is a suitable compactification of the moduli space of pseudo-holomorphic strips analogous to the compactification of gradient flowlines described in Section 5.2.

Definition 6.7.1. *The* ***energy*** *of a strip u is the function $E(u)$ defined as*

$$\begin{aligned} E(u) &= \int_{\mathbb{R}\times[0,1]} \frac{1}{2}\left(\left\| \frac{\partial u}{\partial t} \right\|^2 + \left\| \frac{\partial u}{\partial s} \right\|^2 \right) dtds \\ &= \int_{\mathbb{R}\times[0,1]} \frac{1}{2}\left(g_s\left(\frac{\partial u}{\partial t}, \frac{\partial u}{\partial t} \right) + g_s\left(\frac{\partial u}{\partial s}, \frac{\partial u}{\partial s} \right) \right) dtds, \end{aligned}$$

where g_s denotes the metric on TM associated to ω and the ω-tame (but not necessarily compatible) almost-complex structure J^s, as in Equation (4.5)*.*

The symplectic form ω provides a cohomology class $[\omega] \in H^2(M; \mathbb{R})$, and since L_0 and L_1 are Lagrangian submanifolds, ω also gives rise to a relative class in $H^2(M, L_0 \cup L_1; \mathbb{R})$. Therefore ω can be evaluated on a Whitney

disk from $\mathbf{x}$ to $\mathbf{y}$, and indeed the resulting value will depend only on the homotopy type of the Whitney disk. This value provides an *a priori* bound on the energy of a holomorphic strip, similarly to the bound Equation (5.2) for a gradient flowline. (As it is customary, we will confuse disks and strips by implicitely using the conformal identification $\mathbb{R} \times [0,1] \cong \mathbb{D} \setminus \{\pm i\}$.)

Lemma 6.7.2. *If u is a $\{J^s\}$-holomorphic disk with respect to a one-parameter family of almost-complex structures $\{J^s\}_{s\in[0,1]}$ which are ω-tame, then the energy of u is identified with the evaluation $\int_{\mathbb{D}} u^*(\omega)$. Indeed, this latter quantity can be interpreted as the evaluation of the relative cohomology class of ω evaluated on the relative homology class determined by the disk.*

Proof. Note that for the complex structure on the strip, $j\frac{\partial}{\partial t} = \frac{\partial}{\partial s}$; so for a pseudo-holomorphic map $u\colon \mathbb{R}\times[0,1] \to M$, it follows that $J^s \frac{\partial u}{\partial t} = \frac{\partial u}{\partial s}$. Thus,

$$\begin{aligned}\int_{\mathbb{R}\times[0,1]} \left(\left\| \frac{\partial u}{\partial s} \right\|^2 + \left\| \frac{\partial u}{\partial t} \right\|^2 \right) dtds &= \int_{\mathbb{R}\times[0,1]} \left(\omega(\frac{\partial u}{\partial s}, J^s \frac{\partial u}{\partial s}) + \omega(\frac{\partial u}{\partial t}, J^s \frac{\partial u}{\partial t}) \right) dtds \\ &= \int_{\mathbb{R}\times[0,1]} \left(-\omega(\frac{\partial u}{\partial s}, \frac{\partial u}{\partial t}) + \omega(\frac{\partial u}{\partial t}, \frac{\partial u}{\partial s}) \right) dtds \\ &= 2 \int_{\mathbb{R}\times[0,1]} u^*(\omega).\end{aligned}$$

□

The notion of broken flowlines from Chapter 5 generalizes in the present setting as follows.

Definition 6.7.3. *A **broken holomorphic strip** from $\mathbf{x}$ to $\mathbf{y}$ is a sequence $\mathbf{x} = \mathbf{x}_0, \dots, \mathbf{x}_{n+1} = \mathbf{y}$ of intersection points between L_0 and L_1, and a sequence $u_0, \dots, u_n$ of non-constant $\{J^s\}$-holomorphic strips (modulo translation). A broken holomorphic strip represents a fixed homotopy class $\phi \in W(\mathbf{x}, \mathbf{y})$ if $\phi = [u_0] * \cdots * [u_n]$.*

Under suitable hypotheses on (M, ω, L_0, L_1), the spaces of broken pseudo-holomorphic strips compactify the moduli spaces of pseudo-holomorphic strips. In case M is non-compact (for example, (M,ω) is exact), we need some control on holomorphic strips near the ends of M to achieve such compactness.

Definition 6.7.4. *The symplectic manifold (M,ω) is **convex at infinity** if there is a pair (J, f) of ω-tame almost-complex structure J and smooth,*

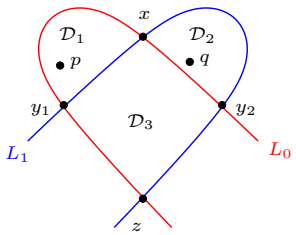


Figure 6.3. A configuration of the curve L_0 and L_1 in the plane $\mathbb{C}$.

proper function $f\colon M \to [0,\infty)$ *and a compact set* $K \subset M$ *such that for all* $p \in M \setminus K$ *the 2-form*

$$\omega_f = -d(d^{\mathbb{C}} f)$$

(where $d^{\mathbb{C}} f(v) = df(Jv)$*) satisfies* $\omega_f(v, Jv) \geq 0$ *for all* $v \in T_p M$*, cf.* [**75**, Section 9.2].

By the maximum modulus principle, if C denotes the supremum of f on $L_0 \cup L_1$, any pseudo-holomorphic curve with boundary on L_0 and L_1 is contained in the compact set $f^{-1}[0, C]$.

The compactness result (motivated by Theorem 5.2.3) in the Lagrangian Floer theoretic setting is due to Gromov [**43**]; see also [**25**]. In its simplest form (under the rather restricted topological conditions on π_2) it is formulated as follows:

Theorem 6.7.5. *Assume that* $\pi_2(M, L_i) = 0$*,* $\pi_2(M) = 0$ *and* L_0, L_1 *are compact Lagrangian submanifolds. Assume furthermore that either* M *is compact, or* (M, ω) *is convex at infinity. Then any sequence of* $\{J^s\}$*-holomorphic strips from* $\mathbf{x}$ *to* $\mathbf{y}$ *(with* $\mathbf{x} \neq \mathbf{y}$*), which has a fixed energy bound, has a* $C^{\infty,\mathrm{loc}}$*-convergent subsequence to a broken holomorphic strip from* $\mathbf{x}$ *to* $\mathbf{y}$*.* □

We do not prove this theorem; see [**95, 96**].

We give an extended example of the broken flowline compactification stated in Theorem 6.7.5. Consider the complex line $\mathbb{C}$, equipped with the two curves L_0 and L_1 pictured in Figure 6.3, with intersection points $x, y_1, y_2, z \in L_0 \cap L_1$.

The region $\mathcal{D}_1$ bounded by the L_0-arc and the L_1-arc containing x and y_1 on its boundary is simply-connected; so by the Riemann mapping theorem, there is, up to translation, a unique holomorphic strip from x to y_1. One way to pin down this indeterminacy is as follows: there is a unique holomorphic strip $u\colon \mathbb{R} \times [0,1] \to \mathbb{C}$ from x to y_1 with the property that there is an s so that $u(0, s) = p$.

Similarly, there are unique holomorphic strips up to reparametrization from x to y_2, y_1 to z, and y_2 to z. Using the notation from Figure 6.3, these holomorphic strips represent the homotopy classes $\mathcal{D}_2$, $\mathcal{D}_2+\mathcal{D}_3$ and $\mathcal{D}_1+\mathcal{D}_3$, respectively.

There is a two-dimensional moduli space of holomorphic Whitney disks from x to z (representing the homotopy class $\phi = \mathcal{D}_1+\mathcal{D}_2+\mathcal{D}_3$), which becomes one-dimensional after dividing out by translation. The image of $\mathbb{R}\times(0,1)$ is contained in the heart-shaped simply-connected region shown in the figure. In fact, the interior must consist of the interior of the heart, with an interval removed, one of whose ends is at x; and the other one is a critical value for the restriction of the map u to $(\mathbb{R}\times\{0\})\cup(\mathbb{R}\times\{1\})$. That interval can be either on the portion of the L_1-arc in the interior of the heart connecting x to y_1, or on the L_0-arc in the interior of the heart connecting x to y_2. This gives a parameterization of $\widehat{\mathcal{M}}(\phi)$ by an interval with two ends, corresponding to y_1 and y_2. As the parameter r goes to y_1 or y_2, the sequence of holomorphic strips limits to a broken flowline in $\widehat{\mathcal{M}}(\mathcal{D}_1)\times\widehat{\mathcal{M}}(\mathcal{D}_2+\mathcal{D}_3)$ or in $\widehat{\mathcal{M}}(\mathcal{D}_2)\times\widehat{\mathcal{M}}(\mathcal{D}_1+\mathcal{D}_3)$.

To make this convergence precise, for $\phi = \mathcal{D}_1+\mathcal{D}_2+\mathcal{D}_3$ let $u_i\in\mathcal{M}(\phi)$ $(i=1,\dots)$ be a sequence of holomorphic strips, normalized so that $u_i(0,s_i)=p$ for a suitable sequence of $s_i\in[0,1]$, with $r(u_i)\to y_1$. By standard complex analysis, u_i converges to some u. Observe that $\lim_{t\to-\infty}u(t,s)=x$ and $\lim_{t\to+\infty}u(t,s)=y_1$.

Note that q is not contained in the image of u. We can retranslate our u_i to obtain a sequence of holomorphic strips v_i with $v_i(t,s)=u_i(t+\tau_i,s)$ (for some sequence $\tau_i\in\mathbb{R}$), so that $v_i(0,s_i')=q$. The sequence v_i converges to a different strip v that obviously contains q in its interior; and in fact $v\in\widehat{\mathcal{M}}(\mathcal{D}_2+\mathcal{D}_3)$. In this case, $[u_i]$ converges to the broken flowline $[u]*[v]\in\widehat{\mathcal{M}}(\mathcal{D}_1)\times\widehat{\mathcal{M}}(\mathcal{D}_2+\mathcal{D}_3)$.

Similarly, if $r(u_i)\to y_2$, then $[u_i]\in\mathcal{M}(\phi)$ converges to the broken flowline $[v']*[u']\in\widehat{\mathcal{M}}(\mathcal{D}_2)\times\widehat{\mathcal{M}}(\mathcal{D}_1+\mathcal{D}_3)$. See Figure 6.4.

Exercise 6.7.6. *Show that there is no holomorphic strip from y_i to x in the above example.*

Theorem 6.7.5 is a special case of *Gromov's compactness theorem*, which compactifies the moduli spaces of pseudo-holomorphic curves in an almost-complex manifold.

To state the hypothesis of this theorem, we generalize the energy of a strip to the case of a map $u\colon\Sigma\to M$, where Σ is a two-manifold equipped with a Riemannian metric. In that case, there is a 1-form $du\in\Omega^1(\Sigma,u^*(TM))$, whose norm at $p\in\Sigma$ is given by $\|du\|^2=|du(e_1)|^2+|du(e_2)|^2$, where e_1,e_2

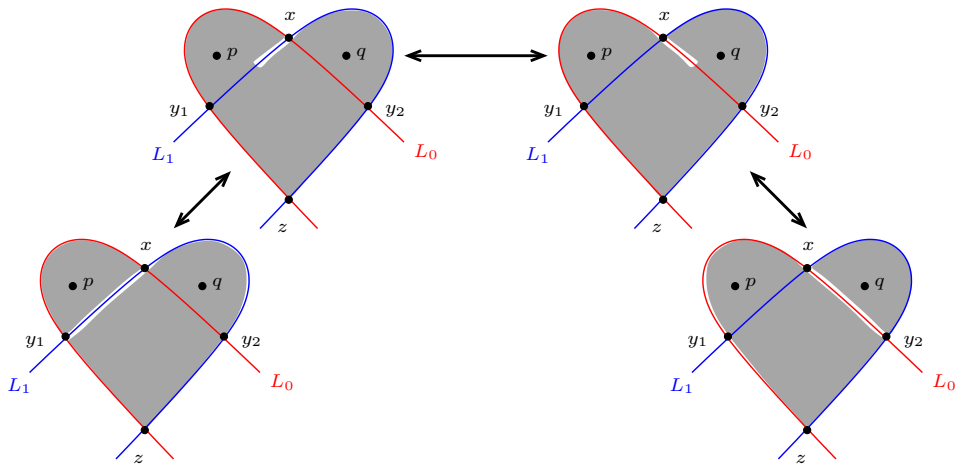


Figure 6.4. The two ends of the one-dimensional moduli space $\widehat{\mathcal{M}}(x,z)$ are pictured on the far left and on the far right.

is an orthonormal frame at $T_p\Sigma$. The *energy of* u is defined by

$$E(u) = \int_\Sigma \|du\|^2 d\sigma, \tag{6.7}$$

where $d\sigma$ is the volume form on (Σ, g). If u satisfies the condition

$$du_p(jv) = Jdu_p(v)$$

at each $p \in \Sigma$ and for every $v \in T_p\Sigma$, then u is called a *pseudo-holomorphic curve*. By construction, if u is a smooth map with zero energy, then the map is constant.

We have the following generalization of Lemma 6.7.2:

Lemma 6.7.7. *Suppose that* $u\colon (\Sigma, \partial\Sigma) \to (M, L_0 \cup L_1)$ *is a pseudo-holomorphic curve, then*

$$E(u) = \int_\Sigma u^*(\omega).$$

In particular, if u *is a pseudo-holomorphic curve that evaluates trivially against* ω*, then* u *is constant.*

Proof. This follows exactly as in Lemma 6.7.2. □

We will often apply Lemma 6.7.7 in cases where $\partial\Sigma = 0$.

Suppose now that M is compact or, more generally, convex at infinity. Gromov's compactness theorem states that a sequence of J-holomorphic maps with uniform energy bound has a subsequence which converges in a suitable sense to a map from a nodal holomorphic curve; i.e. a Riemann

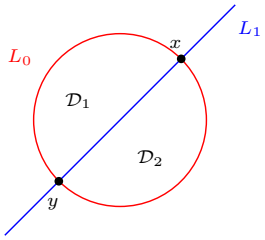


Figure 6.5. Another configuration of the curves L_0 and L_1 in the plane $\mathbb{C}$.

surface with double-points. From this perspective, a broken flowline can be thought of as a chain of holomorphic disks, and each point where two consecutive disks meet is a node.

Even in cases where the domain curve is a disk, one gets convergence to a more general object than a broken holomorphic strip, provided that one relaxes the topological assumptions in Theorem 6.7.5. Without the π_2 hypotheses, the limiting object may contain, in addition to holomorphic strips, J-holomorphic spheres that bubble off from some interior point in the sequence of strips, and J-holomorphic disks with boundary contained entirely inside L_0 or L_1, which we think of as a disk bubbling off the L_0- or the L_1-side. A precise statement needs to allow for multiple gluings in such a manner that the source is, in a suitable sense, a stable map of a disk with nodes. In particular, if spheres do indeed bubble off, then at least one of those spheres needs to be non-constant; similarly for the disks with boundary in L_0 or L_1; compare [**75, 25**, Chapter 5.3].

Rather than stating Gromov's compactness theorem in generality, we give a few illustrative one-dimensional examples. (See 35.5 for a similar case.)

To see disks bubbling off from the side, consider the configuration of curves in $\mathbb{C}$ where L_0 is an embedded circle, and L_1 is a curve that meets L_0 transversely in two points x and y, as shown in Figure 6.5. It is easy to see that in this case $\pi_2(M, L_0) \cong \mathbb{Z}$.

Let $\phi \in W(x, x)$ be the homotopy class generating $\pi_2(M, L_0)$, that is, the disk encircled by L_0. The moduli space of disks representing ϕ is parametrized naturally by a parameter r in the interval (x, y) on L_1. (Here, the curve $u = u_r$ if r is a critical value for $u|_{\mathbb{R}\times 1}$.) As $r \mapsto y$, the sequence of curves $u_r \in \widehat{\mathcal{M}}(\phi)$ converges to a broken flowline in $\widehat{\mathcal{M}}(\mathcal{D}_1) \times \widehat{\mathcal{M}}(\mathcal{D}_2)$. As $r \mapsto x$, on the other hand, the sequence of curves u_r converges to a configuration of curves $u_0 * v$, where u_0 corresponds to the constant holomorphic curve (representing the domain $0 \in W(\phi)$), and v is a holomorphic map of the disk, mapping boundary to all of L_0; see Figure 6.6.

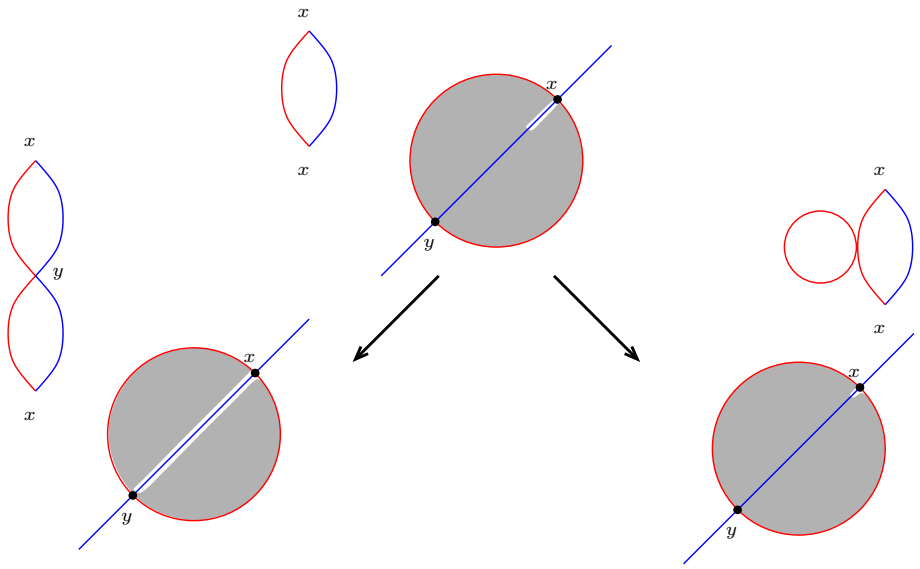


Figure 6.6. The one-dimensional moduli space shown at the top has two ends: a broken flowline (on the left) and a boundary disk end (on the right).

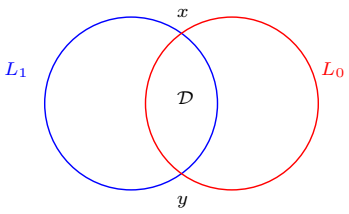


Figure 6.7. A configuration of the curve L_0 and L_1 in S^2.

Next, we investigate how a moduli space can develop a sphere bubble. To this end, consider the pair of circles L_0 and L_1 embedded in $S^2 = \mathbb{CP}^1$, so that they intersect in two points, as shown in Figure 6.7. Fix an intersection point x, choose $\phi \in W(x,x)$, covering all of S^2 with degree one, and consider the four-dimensional moduli space $\mathcal{M}(\phi)$. Then the space $\widehat{\mathcal{M}}(\phi)$ is three-dimensional, parametrized by the three critical values of $u|_{\mathbb{R}\times\{0,1\}}$ on $L_0 \cup L_1$. For concreteness, consider the portion of the moduli space where two of the critical points r_1 and r_2 lie on $L_0 \setminus \{x\}$, and the third q lies on L_1, as indicated in Figure 6.8. For a sequence of curves u_t, the parameter $q \in L_1$ can approach x in two possible ways. In one direction, as the cut gets long, the holomorphic curve degenerates into a broken flowline. In the other direction, as the cut gets short, the moduli space degenerates to a boundary disk bubble. If all three cuts r_1, r_2 and q degenerate to x

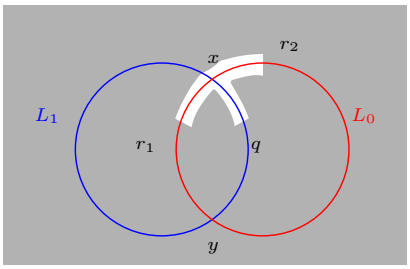


Figure 6.8. A three-dimensional moduli space.

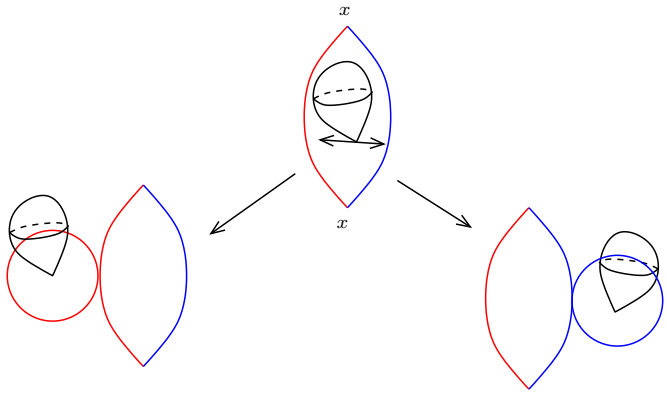


Figure 6.9. Sphere bubbles in a three-dimensional moduli space. Note that all the non-spherical components are constant.

at once, the moduli space degenerates to a constant flowline with a sphere bubbling off. Note that there is a one-dimensional space of these. If first we let q approaches x (so that the cut length goes to zero), and then we degenerate r_1 and r_2, the configuration consists of a (constant) boundary disk on the L_0 side with a sphere degeneration on it. If we let r_1 and r_2 degenerate to x first, and then q, we obtain an L_1-boundary disk with a sphere bubble on it. We can interpolate between these two degenerations to get a one-parameter family of (constant) flowlines from x to x with a sphere bubble off; see Figure 6.9.

For our final example, consider the same configuration of circles as in Figure 6.7, but now consider the homotopy class ϕ in $W(x, y)$ whose local multiplicities are 2 at the component $\mathcal{D}$ in $S^2 \setminus (L_0 \cup L_1)$, and whose local multiplicities are 1 at the other three regions. Then $\mathcal{M}(\phi)$ is of dimension five, the space $\widehat{\mathcal{M}}(\phi)$ is a four-dimensional moduli space, and we will consider some portion of it. Given any two distinct points $z_1, z_2 \in \mathcal{D}$, there is a unique holomorphic representative of ϕ, as a map whose two branch points are at z_1 and z_2. Denote this map by u_{z_1,z_2}. Suppose that z_1 and

z_2 approach some $z \in \mathcal{D}$. If that point z is in the interior of $\mathcal{D}$, the limiting curve consists of a flowline from x to y that meets a sphere bubble at the point z. If the point z is on $L_0 \setminus \{x, y\}$, the configuration consists of three components: a flow from x to y, a (constant) boundary bubble, and sphere that meets the boundary bubble. If $z \in \{x, y\}$, then there is a configuration consisting of a broken flowline one of whose components is constant; and the constant flowline meets a sphere bubble.

Observe now that the hypotheses from Theorem 6.7.5 ensure that these pathologies – sphere bubbles and boundary bubbles – do not occur. For example, $\pi_2(M) = 0$ implies that spheres bubbling off would have to be null-homologous, and hence the integral of ω over such spheres (which can be thought of as evaluating a cohomology class on a homology class) would have to vanish. But this integral can alternatively be thought of as the energy of the sphere, as in Equation (6.7). Clearly, a map with zero energy, though, is constant. Similarly interpreting the integal of ω over $\mathbb{D}$ (with boundary inside L_0 or L_1) as the evaluation of a relative cohomology class ω (recall that L_i were Lagrangian) over a relative homology class, we conclude that when $\pi_2(M, L_i) = 0$, there can be no pseudo-holomorphic disks bubbling off the side.

The significance of this compactness for the boundary operator appearing in Equation (6.6) is the following:

Corollary 6.7.8. *Let (M^{2n}, ω) be a symplectic manifold, equipped with compact Lagrangians L_0 and L_1, satisfying $\pi_2(M^{2n}, L_0) = \pi_2(M^{2n}, L_1) = \pi_2(M^{2n}) = 0$. Suppose furthermore that either M is compact, or convex at infinity. If $\{J^s\}$ is a suitably generic one-parameter family of ω-tame almost-complex structures, then for all non-constant homotopy classes $\phi \in W(\mathbf{x}, \mathbf{y})$ with $\mu(\phi) = 1$, the moduli space $\mathcal{M}_{\{J^s\}}(\phi)$ is a compact, zero-dimensional manifold.*

Proof. Choose $\{J^s\}$ so that for all non-constant homotopy classes ψ with $\mu(\psi) \leq 0$, the corresponding moduli spaces $\widehat{\mathcal{M}}_{\{J^s\}}(\psi)$ are empty, and for all homotopy classes with $\mu(\phi) = 1$, the corresponding moduli space is smooth. This can be done by Theorem 6.6.2. By Theorem 6.7.5, any sequence in $\mathcal{M}_{\{J^s\}}(\phi)$ converges to a broken flowline, so if that broken flowline has at least one break, then one of the component homotopy classes has index ≤ 0 (by the additivity of the Maslov index under juxtapositions) and it is non-constant, contradicting our choice of $\{J^s\}$. □

Having established (at least under the topological constraints of Theorem 6.7.5 and Corollary 6.7.8) that the counts in the Floer differential for each homotopy class ϕ are finite, we must show that only finitely many homotopy classes contribute. We start with the special case where L_0 and L_1 are exact Lagrangians in an exact symplectic manifold.

In the exact case, for each constant path $\mathbf{x}$, the action functional (as defined in Equation (6.2)) is given by $f_0(\mathbf{x}) - f_1(\mathbf{x})$. Thus, for pseudo-holomorphic $u \in W(\mathbf{x}, \mathbf{y})$, the energy of u, which by Lemma 6.7.2 is identified with $\int_{\mathbb{D}} u^*(\omega)$, is computed by

$$\int_{\mathbb{D}} u^*(\omega) = E(u) = -f_0(\mathbf{x}) + f_1(\mathbf{x}) + f_0(\mathbf{y}) - f_1(\mathbf{y}),$$

as in Equation (6.3). This gives the desired energy bound independent of $\phi \in W(\mathbf{x}, \mathbf{y})$, and hence gives compactness for $\bigcup_{\{\phi \in W(x,y) | \mu(\phi)=1\}} \widehat{\mathcal{M}}(\phi)$.

In conclusion, if (M, ω) is an exact symplectic manifold with compact exact Lagrangian submanifolds L_0, L_1 and (M, ω) is convex at infinity, then in the boundary operator defined in Equation (6.6) for every pair $\mathbf{x}, \mathbf{y}$ there are finitely many homotopy classes $\phi \in W(\mathbf{x}, \mathbf{y})$ with nonempty moduli spaces, and if $\mu(\phi) = 1$, all these spaces consist of finitely many points. Therefore the formula of (6.6) makes sense and defines an endomorphism of $\mathrm{CF}(L_0, L_1)$.

When (M, ω) is not exact, we can deal with the possibly infinitely many homotopy classes connecting $\mathbf{x}$ and $\mathbf{y}$ in another way by introducing a new coefficient ring.

Definition 6.7.9. *The **Novikov field** $N_{\mathbb{Z}/2\mathbb{Z}}$ over $\mathbb{Z}/2\mathbb{Z}$ is defined as the collection of those formal sums $x_A = \sum_{a \in A} x_a T^a$ where $x_a \in \mathbb{Z}/2\mathbb{Z}$, T is a formal variable and the set $A \subset \mathbb{R}$ is any discrete subset which is boundary below. Given two such subsets A and B, their Minkowski sum $A + B = \{x \big| x = a + b$ for some $a \in A$ and $b \in B\}$ (where elements are counted with multiplicity in $\mathbb{Z}/2\mathbb{Z}$) has the same property, so we can define $x_A \cdot x_B = x_{A+B}$.*

Define $\mathrm{CF}(L_0, L_1; N_{\mathbb{Z}/2\mathbb{Z}})$ as the vector space over $N_{\mathbb{Z}/2\mathbb{Z}}$ generated by the (finitely many) intersection points of L_0 and L_1. For a homotopy class $\phi \in W(\mathbf{x}, \mathbf{y})$ define $a(\phi)$ to be the integral of ω on (any representative of) ϕ. Equip the module with the endomorphism

$$\partial \mathbf{x} = \sum_{\mathbf{y} \in L_0 \cap L_1} \sum_{\{\phi \in W(\mathbf{x}, \mathbf{y}) \big| \mu(\phi)=1\}} \#\widehat{\mathcal{M}}_{\{J^s\}}(\phi) t^{a(\phi)} \cdot \mathbf{y}. \tag{6.8}$$

Since the quantity $a(\phi)$ is equal to the energy of a pseudo-holomorphic representative of ϕ, the compactness result of Theorem 6.7.5 ensures that there

are only finitely many homotopy classes with pseudo-holomorphic representatives. Moreover, $a(\phi)$ can be thought of as the evaluation of the relative cohomology class $[\omega] \in H^2(M, L_0 \cup L_1; \mathbb{R})$ on the relative homology class in $H_2(M, L_0 \cup L_1)$ associated to ϕ. If M is compact, then $H_2(M, L_0 \cup L_1)$ is finitely generated, and the coefficient of $\mathbf{y}$ on the right hand side of Equation (6.8) is a well-defined element of $N_{\mathbb{Z}/2\mathbb{Z}}$. When M is not compact but convex at infinity, we can work instead with $H_2(f^{-1}[0, C], L_0 \cup L_1)$, with $C = \sup_{L_0 \cup L_1} f$.

6.8. Gluing

As before, assume that $\pi_2(M) = 0$ and $\pi_2(M, L_i) = 0$, and either M is compact, or (M, ω) is convex at infinity. Having made sense of the Floer differential, we turn to the verification that $\partial^2 = 0$. The gluing theorem now implies the following refined version of the compactification theorem for moduli spaces; compare [**5**, Chapter 9] and [**21**].

Theorem 6.8.1. *For generic choices of $\{J^s\}$, and for each $\phi \in W(\mathbf{x}, \mathbf{y})$ with $\mu(\phi) = 2$, $\mathcal{M}_{\{J^s\}}(\phi)$ has a compactification to a one-manifold with boundary, whose boundary is identified with*

$$\bigcup_{\left\{\phi_1, \phi_2 \middle| \begin{smallmatrix} \phi_1 * \phi_2 = \phi \\ \mu(\phi_1) = \mu(\phi_2) = 1 \end{smallmatrix}\right\}} \widehat{\mathcal{M}}_{\{J^s\}}(\phi_1) \times \widehat{\mathcal{M}}_{\{J^s\}}(\phi_2). \tag{6.9}$$

Remark 6.8.2. *From Theorem 6.7.5 it follows that the broken flowlines compactify the moduli space. We must further show two things. First, that each pair $(u_1, u_2) \in \widehat{\mathcal{M}}_{\{J^s\}}(\phi_1) \times \widehat{\mathcal{M}}_{\{J^s\}}(\phi_2)$ arises as the limit of a sequence of flowlines from $\widehat{\mathcal{M}}_{\{J^s\}}(\phi)$. In this step, as it was explained in the Morse-Smale case in Remark 5.2.6, we rely on a gluing result: first for the given pair (u_1, u_2) and a parameter ρ we construct a Whitney disk representing $\phi = \phi_1 * \phi_2$ by considering u_1 on $(-\infty, \rho) \times [0, 1] \subset \mathbb{R} \times [0, 1]$ and u_2 on $(-\rho, \infty) \times [0, 1] \subset \mathbb{R} \times [0, 1]$ and interpolating between the two maps near the intermediate intersection point. In the second step this approximate flowline from $\mathbf{x}$ to $\mathbf{y}$ is modified to an actual flowline $u_\rho \in \widehat{\mathcal{M}}_{\{J^s\}}(\phi)$; it can be shown that as $\rho \to \infty$ the sequence $\{u_\rho\}$ has a subsequence converging to the given broken flowline (u_1, u_2). Finally, we must show that any point in the moduli space which is sufficiently close to a broken flowline is actually contained in this gluing neighborhood.*

Using the above theorem, we can now show that $\partial^2 = 0$. By the definition of the boundary map, the mod 2 number of points in the space given in Equation (6.9) gives the coefficient of $\mathbf{y}$ in $\partial^2 \mathbf{x}$.

In the exact case, our assumptions on the topology of (M, L_0, L_1), and the above gluing result implies that the points of the space described in Equation (6.9) corresponds to the ends of the 1-dimensional manifold $\widehat{\mathcal{M}}_{\{J^s\}}(\phi)$. Since a compact 1-manifold has an even number of boundary points, it follows that the coefficient of $\mathbf{y}$ in $\partial^2\mathbf{x}$ is zero mod 2. Applying the same reasoning for all further pairs of intersection points, we conclude that $\partial^2 = 0$, as desired.

The same argument applies for $\mathrm{CF}(L_0, L_1; N_{\mathbb{Z}/2\mathbb{Z}})$. In this adaptation, to see that the terms in ∂^2 cancel, it is important to note that $a(\phi * \psi) = a(\phi) + a(\psi)$.

Thus, (under the by now usual topological constraints) the pair $(\mathrm{CF}(L_0, L_1), \partial)$ is a chain complex, and we define $HF(L_0, L_1)$ as the homology of this chain complex. Note that as it is defined, $HF(L_0, L_1)$ depends on the symplectic manifold (M, ω), the two Lagrangians L_0, L_1 and the chosen ω-tame 1-parameter family of almost-complex structures $\{J^s\}$. We show that it is independent of the choice $\{J^s\}$ in the following section, and it is invariant under Hamiltonian isotopy of the Lagrangians.

Remark 6.8.3. *Notice that the chain complex* $\mathrm{CF}(L_0, L_1)$ *depends only on the path* $\{J^s\}$ *of almost-complex structures; the symplectic form* ω *plays only an indirect role in the definition of the boundary map* ∂ *by specifying which almost-complex structures are* ω*-tame.*

Recall that the compactness results of this section were stated under rather restrictive conditions. Indeed, these conditions will not be satisfied in the applications of the Lagrangian Floer homology package we will encounter later. In the specific context of Heegaard Floer homology the arising complications related to compactness issues will be handled by other means, as it will be detailed in Chapter 9.

6.9. The invariance proof

Consider for definiteness the case of $HF(L_0, L_1; N_{\mathbb{Z}/2\mathbb{Z}})$:

Theorem 6.9.1. *The homology* $HF(L_0, L_1; N_{\mathbb{Z}/2\mathbb{Z}})$ *is an invariant of the Lagrangian subspaces of* (M, ω)*.*

We must show that it is independent of the choice of the path of almost-complex structures $\{J^s\}$ used in its definition. This proof is analogous to the proof sketched in Section 5.3. Specifically, suppose that $\{J_0^s\}$ and $\{J_1^s\}$ are two one-parameter families of almost-complex structures that are suitably generic for the complexes $\mathrm{CF}_{\{J_0^s\}}(L_0, L_1; N_{\mathbb{Z}/2\mathbb{Z}})$ and $\mathrm{CF}_{\{J_1^s\}}(L_0, L_1; N_{\mathbb{Z}/2\mathbb{Z}})$ to be well-defined. Connect $\{J_0^s\}$ and $\{J_1^s\}$ by a one-parameter family

of paths of almost-complex structures, $\{J_\tau^s\}$ ($\tau \in [0,1]$). Given $\mathbf{x}, \mathbf{y} \in L_0 \cap L_1$ and $\phi \in W(\mathbf{x}, \mathbf{y})$, we now have a parameterized moduli space, $\mathcal{M}_{\{J_\tau^s\}_{s,\tau\in[0,1]}}(\phi)$ of Whitney strips u from $\mathbf{x}$ to $\mathbf{y}$ representing ϕ and satisfying the conditions

$$\frac{\partial u}{\partial t} + J^s_{\psi(t)} \frac{\partial u}{\partial s} = 0,$$

where (as in the discussion after Equation (5.5)) $\psi\colon \mathbb{R} \to [0,1]$ is some fixed smooth, monotone function $\psi(t) = 0$ for $t \leq 0$ and $\psi(t) = 1$ for $t \geq 1$. Note the dependence of the almost-complex structure J on the (t,s) parameter: whereas the J-term considered earlier depended on the t-parameter only through the image of $u(t,s)$, now we are considering the almost-complex structure $\{J^s_{\psi(t)}\}$ at $u(t,s)$. The moduli spaces associated to a one-parameter family $\{J^s\}_{s\in[0,1]}$ considered earlier have an $\mathbb{R}$ action; this symmetry has been broken in the construction of moduli spaces associated to two-parameter families of almost-complex structures.

Theorem 6.6.2 has the following analogue (compare Theorem 5.3.1, [**23**]):

Theorem 6.9.1. *Let (M^{2n}, ω) be a closed symplectic manifold, equipped with compact Lagrangians L_0 and L_1, and fix two paths of ω-tame almost-complex structures $\{J_0^s\}_{s\in[0,1]}$ and $\{J_1^s\}_{s\in[0,1]}$. Then for all sufficiently generic two-parameter family of ω-tame almost-complex structures $\{J_\tau^s\}_{s,\tau\in[0,1]}$ connecting those two paths, and for all homotopy classes $\phi \in W(\mathbf{x}, \mathbf{y})$ with $\mu(\phi) \leq 1$, the space $\mathcal{M}_{\{J_\tau^s\}_{s,\tau\in[0,1]}}(\phi)$ is a smooth manifold of dimension given by*

$$\dim \mathcal{M}_{\{J_\tau^s\}_{s,\tau\in[0,1]}}(\phi) = \mu(\phi).$$

In particular, if ϕ is a non-constant homotopy class with $\mu(\phi) < 0$, then $\mathcal{M}_{\{J_\tau^s\}_{s,\tau\in[0,1]}}(\phi)$ is empty. □

These moduli spaces $\mathcal{M}_{\{J_\tau^s\}_{s,\tau\in[0,1]}}(\phi)$ can be assembled to construct a map

$$\Phi = \Phi_{\{J_t^s\}_{s,t\in[0,1]}}\colon \mathrm{CF}_{\{J_0^s\}_{s\in[0,1]}}(L_0, L_1; N_{\mathbb{Z}/2\mathbb{Z}}) \to \mathrm{CF}_{\{J_1^s\}_{s\in[0,1]}}(L_0, L_1; N_{\mathbb{Z}/2\mathbb{Z}}),$$

defined by

$$\Phi(\mathbf{x}) = \sum_{\mathbf{y}\in L_0\cap L_1} \sum_{\{\phi\in W(\mathbf{x},\mathbf{y})|\mu(\phi=0\}} \#\mathcal{M}_{\{J_\tau^s\}_{s,\tau\in[0,1]}}(\phi) T^{a(\phi)} \cdot \mathbf{y}, \tag{6.10}$$

where $\#\mathcal{M}_{\{J_\tau^s\}_{s,\tau\in[0,1]}}(\phi)$ denotes the mod 2 count of points in the moduli space $\mathcal{M}_{\{J_\tau^s\}_{s,\tau\in[0,1]}}(\phi)$. Gromov's compactness theorem, and gluing, can be adapted in the present case to prove the following

Proposition 6.9.2. *Assume that $\pi_2(M, L_i) = 0$, $\pi_2(M) = 0$ and L_0, L_1 are compact Lagrangian submanifolds. Assume furthermore that either M is compact, or (M, ω) is convex at infinity. Suppose that $\{J_t^s\}_{s,t\in[0,1]}$ is a generic two-parameter family of ω-tame almost-complex structures. Then*

for all $\mathbf{x}, \mathbf{y} \in L_0 \cap L_1$, *and all homotopy classes of Whitney disks* $\phi \in W(\mathbf{x}, \mathbf{y})$ *with* $\mu(\phi) = 0$, *the moduli space* $\mathcal{M}_{\{J_\tau^s\}}(\phi)$ *is a compact, zero-dimensional manifold. Moreover, if* $\mu(\phi) = 1$, *the compactified moduli spaces are compact one-dimensional manifolds, whose boundaries are identified with the following:*

$$\Big(\bigcup_{\left\{\phi_1 * \phi_2 = \phi \,\middle|\, \begin{smallmatrix} \mu(\phi_1) = 1 \\ \mu(\phi_2) = 0 \end{smallmatrix}\right\}} \widehat{\mathcal{M}}_{\{J_0^s\}_{s\in[0,1]}}(\phi_1) \times \mathcal{M}_{\{J_\tau^s\}_{s,\tau\in[0,1]}}(\phi_2) \Big)$$

$$\cup \Big(\bigcup_{\left\{\phi_1 * \phi_2 = \phi \,\middle|\, \begin{smallmatrix} \mu(\phi_1) = 0 \\ \mu(\phi_2) = 1 \end{smallmatrix}\right\}} \mathcal{M}_{\{J_\tau^s\}_{s,\tau\in[0,1]}}(\phi_1) \times \widehat{\mathcal{M}}_{\{J_1^s\}_{s\in[0,1]}}(\phi_2) \Big). \tag{6.11}$$

The topological hypotheses in the beginning rule out the possibility of Gromov convergence to a more general type of object (with spheres and/or boundary bubbles).

The above proposition implies that Φ is a chain map. Specifically, the $\mathbf{y}$ component of $(\partial \circ \Phi + \Phi \circ \partial)(\mathbf{x})$ counts points in the space displayed in Equation (6.11). Proposition 6.9.2 exhibits this space as the boundary of a compact one-manifold, hence the number of points in this space is even; i.e.

$$\partial \circ \Phi + \Phi \circ \partial = 0.$$

In order to show that $\Phi_{\{J_\tau^s\}_{s,\tau\in[0,1]}}$ induces an isomorphism on homology, we show that $\Phi_{\{J_{1-\tau}^s\}_{s,\tau\in[0,1]}}$ (defined analogously) is its homotopy inverse; i.e. we construct an operator

$$H \colon \mathrm{CF}_{\{J_0^s\}_{s\in[0,1]}}(L_0, L_1) \to \mathrm{CF}_{\{J_0^s\}_{s\in[0,1]}}(L_0, L_1)$$

satisfying the equation

$$\partial \circ H + H \circ \partial = \mathrm{Id} + \Phi_{\{J_{1-\tau}^s\}_{s,\tau\in[0,1]}} \circ \Phi_{\{J_\tau^s\}_{s,\tau\in[0,1]}}.$$

The operator H is defined by counting holomorphic disks using a three-parameter family of almost-complex structures $\{J_{r,t}^s\}$; compare Equation (5.8). We start with $\{J_\tau^s\}$ as before, and choose $\{J_{r,t}^s\}_{\{s\in[0,1], r\in[0,\infty), t\in\mathbb{R}\}}$ so that the following properties hold:

- $J_{r,t}^s = J_0^s$ if $|t| \geq r$
- there is some $R > 1$ so that for all $r \geq R$, and $t \geq 0$, $J_{r,t}^s = J_{\psi(r-t)}^s$
- $r \geq R$ and $t \leq 0$, $J_{r,t}^s = J_{\psi(r+t)}^s$.

Given $\phi \in W(\mathbf{x}, \mathbf{y})$, we can consider the moduli space $\mathcal{M}_H(\phi)$ consisting of pairs $r \in (0, \infty)$, and Whitney strips u representing ϕ, satisfying the equation

$$\frac{\partial u}{\partial t} + J_{r,t}^s \frac{\partial u}{\partial s} = 0.$$

The usual transversality arguments show that $\mathcal{M}_H(\phi)$ is a manifold of dimension $\mu(\phi) - 1$. In particular, when the dimension is zero, the moduli space is compact.

When the moduli space is one-dimensional, it can have three kinds of boundary. One kind of boundary occurs when a $\{J^s_{r,t}\}$-holomorphic disk breaks off, for some $r \in (0, \infty)$. Another kind of boundary occurs when $r \to 0$; such a boundary point is a constant flowline for the family of almost-complex structures $\{J^s_0\}_{s\in[0,1]}$. The third kind of boundary occurs as $r \to \infty$, where the boundary point consists of a $\{J^s_t\}_{s,t\in[0,1]}$-pseudo-holomorphic disk, juxtaposed with a $\{J^s_{1-t}\}_{s,t\in[0,1]}$-pseudo-holomorphic one. Counting boundary points, we obtain the relation

$$\partial \circ H + H \circ \partial = \mathrm{Id} + \Phi_{\{J^s_{1-t}\}_{s,t\in[0,1]}} \circ \Phi_{\{J^s_t\}_{s,t\in[0,1]}}.$$

A simple adaptation of the above construction shows that $\Phi_{\{J^s_t\}_{s,t\in[0,1]}} \circ \Phi_{\{J^s_{1-t}\}_{s,t\in[0,1]}}$ is also homotopic to the respective identity map. Thus, the chain homotopy type of $\mathrm{CF}(L_0, L_1; N_{\mathbb{Z}/2\mathbb{Z}})$ is independent of the choice of the path of ω-tame almost-complex structures $\{J^s\}_{s\in[0,1]}$ that goes into its definition.

6.9.1. Hamiltonian isotopy invariance. To see that the (chain homotopy type of the) complex is invariant under Hamiltonian isotopies, we argue similarly. Suppose that $\mathcal{H}\colon M \times [0,1] \to \mathbb{R}$ is a bounded Hamiltonian function, and $\{\Psi_t\}_{t\in[0,1]}\colon M \to M$ is the corresponding one-parameter family of Hamiltionian diffeomorphisms with $\Psi_0(x) = x$; i.e. for any other vector field Y and $x \in M$,

$$\omega(\frac{d\Psi_t}{dt}(x), Y) = Y_x \mathcal{H}_t.$$

We wish to construct a homotopy equivalence

$$\Phi_{\{\Psi_t\}_{t\in[0,1]}}\colon \mathrm{CF}_{\{J^s\}_{s\in[0,1]}}(L_0, L_1) \to \mathrm{CF}_{\{J^s\}_{s\in[0,1]}}(L_0, \Psi_1(L_1)). \tag{6.12}$$

To this end we consider pseudo-holomorphic disks with fixed one-parameter family of almost-complex structures $\{J^s\}_{s\in[0,1]}$ and with moving boundary conditions; i.e. maps $u\colon \mathbb{R}\times[0,1] \to M^{2n}$ satisfying the boundary conditions $u(t,0) \in L_0$, $u(t,1) \in \Psi_t(L_1)$ for all $t \in \mathbb{R}$, and

$$\lim_{t\mapsto-\infty} u(t,s) = \mathbf{x} \qquad \lim_{t\mapsto+\infty} u(t,s) = \mathbf{y},$$

for $\mathbf{x} \in L_0 \cap L_1$ and $\mathbf{y} \in L_1 \cap \Psi_1(L_1)$, and the usual Cauchy-Riemann equations (Equation (6.5)). Continuous maps satisfying the above boundary conditions can be assembled into homotopy classes, which we now denote $W'(\mathbf{x}, \mathbf{y})$; and the space of pseudo-holomorphic representatives of a given homotopy class $\phi \in W'(\mathbf{x}, \mathbf{y})$ is denoted $\mathcal{M}_{\{\Psi_t\}_{t\in[0,1]}}(\phi)$.

The by now familiar transversality theorem (analogous to Theorem 6.9.1) shows that the moduli spaces $\mathcal{M}_{\{\Psi_t\}_{t\in[0,1]}}(\phi)$ with $\mu(\phi) \leq 1$ are smooth, with dimension computed by $\mu(\phi)$.

We define the map promised in Equation (6.12), as follows. Given $\mathbf{x} \in L_0 \cap L_1$, let

$$\Phi_{\{\Psi_t\}_{t\in[0,1]}}(\mathbf{x}) = \sum_{\mathbf{y}\in L_0\cap\Psi_1(L_1)} \sum_{\{\phi\in W'(\mathbf{x},\mathbf{y})\,\big|\,\mu(\phi)=0\}} \#\mathcal{M}_{\{\Psi_t\}_{t\in[0,1]}}(\phi)\cdot t^{a(\phi)}\mathbf{y}.$$

These maps are called *continuation maps*. To see that the sums are finite, we apply a suitable variant of Gromov's compactness theorem, now with energy bounds supplied by the following:

Lemma 6.9.3. *For each $\phi \in W'(\mathbf{x},\mathbf{y})$, there is a constant $C(\phi)$ with the property that all pseudo-holomorphic representatives of ϕ have $E(u) \leq C(\phi)$.*

Proof. We show that if $m = \sup_{M\times[0,1]} \mathcal{H}_t(x)$, then for any two pseudo-holomorphic representatives u_0 and u_1 of the same moving pseudo-holomorphic homotopy class $\phi \in W'(\mathbf{x},\mathbf{y})$ we have $|E(u_1) - E(u_0)| \leq 2m$.

To this end, fix a homotopy $\{u_\tau\}_{\tau\in[0,1]}$ from u_0 to u_1, i.e. a map $u\colon \mathbb{R}\times[0,1]\times[0,1]\times M$ with the property that for all $t\in\mathbb{R}$ and $\tau\in[0,1]$,

$$u(t,0,\tau)\in L_0 \qquad \text{and} \qquad u(t,1,\tau)\in\Psi_t(L_1)$$

and $u(t,s,0) = u_0(t,s)$, $u(t,s,1) = u_1(t,s)$ for $i = 0,1$. Then,

$$\begin{aligned} 0 &= \int_{\mathbb{R}\times[0,1]\times[0,1]} du^*(\omega) = \int_{\partial(\mathbb{R}\times[0,1]\times[0,1])} u^*(\omega) \\ &= E(u_1) - E(u_0) + \int_{\mathbb{R}\times\{0\}\times[0,1]} u^*(\omega) + \int_{\mathbb{R}\times\{1\}\times[0,1]} u^*(\omega). \end{aligned}$$

Since $u|_{\mathbb{R}\times\{0\}\times[0,1]} \in L_0$, the restriction of ω to this region vanishes. Now, $u(t,1,\tau)\in L_{\psi(t)}$, so

$$\frac{du}{d\tau} \in T_{u(t,1,\tau)}\Psi_t(L_1) \qquad \text{and} \qquad \frac{du}{dt} - \frac{d\Psi_t}{dt}(u(t,1,\tau)) \in T_{u(t,1,\tau)}\Psi_t(L_1).$$

Moreover, $\frac{d\Psi_t}{dt}$ is supported in $t\in[0,1]$, so

$$\begin{aligned} \int_{\mathbb{R}\times\{1\}\times[0,1]} u^*(\omega) &= \int_{\mathbb{R}\times\{1\}\times[0,1]} \omega(\frac{\partial u}{\partial t},\frac{\partial u}{\partial\tau})dt\wedge d\tau = \int_{[0,1]\times[0,1]} \omega(\frac{d\Psi_t}{dt},\frac{\partial u}{\partial\tau})dt\wedge d\tau \\ &= \int_{[0,1]}\left(\int_{[0,1]} \frac{\partial}{\partial\tau}(\mathcal{H}_t(u(t,\tau)))d\tau\right)dt \\ &= \int_{[0,1]} (\mathcal{H}_t(u(t,1)) - \mathcal{H}_t(u(t,0)))dt; \end{aligned}$$

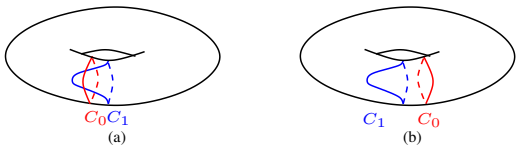


Figure 6.10. Curves in T^2.

so $|E(u_1) - E(u_0)| \leq 2m$, as claimed. □

A suitable adaptation of Gromov's compactness theorem, now with the energy bound supplied by Lemma 6.9.3 shows that the map from Equation (6.10) is well defined. The usual gluing results show that it is a chain map. The homotopy inverse is supplied by $\Phi_{\{\Psi_{-t}\}}$; a homotopy can be constructed as before, now using moving boundary conditions, to show that $\Phi_{\{\Psi_t\}}$ and $\Phi_{\{\Psi_{-t}\}}$ are homotopy inverses, as claimed.

Remark 6.9.4. *The following example shows that Floer homology is not invariant under general symplectomorphisms. Consider the two-torus T^2 and two homologically essential circles C_0, C_1 on it with two intersection points, see Figure 6.10(b). It is not hard to see that $HF(C_0, C_1)$ is $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$. By applying a symplectomorphism φ on C_1 coming from a sufficiently large rotation, we can make C_0 and $\varphi(C_1)$ disjoint, so that the Floer homology is 0.*

6.10. The Maslov index for polygons

In Chapter 12, we will need a slight extension of Whitney disks to polygons and their Maslov indices. We sketch these notions here. As usual, fix a symplectic manifold M.

Definition 6.10.1. *Suppose that $\mathbb{L} = (L_1, \ldots, L_m)$ is an ordered m-tuple of Lagrangian submanifolds of (M, ω). We say that $\mathbb{L}$ is a **transversal chain** if L_i intersects L_j transversally for any $i \neq j \in \{1, \ldots, m\}$, and the triple intersections $L_i \cap L_j \cap L_k$ is empty for any three distinct indices $i, j, k \in \{1, \ldots, m\}$.*

Consider the disk $\mathbb{D} \subset \mathbb{C}$ with m distinct marked points $v_1, \ldots, v_m$ on its boundary in this order, and denote the punctured disk $\mathbb{D} \setminus \{v_1, \ldots, v_m\}$ by $\dot{\mathbb{D}}$. The arc on $\partial\mathbb{D}$ between the marked points v_i and v_{i+1} is denoted by $\overline{v_i v_{i+1}}$. (Here, and in the following we will regard indices mod m, hence for example $m+1$ is considered to be equal to 1.) Fix also disjoint closed neighbourhoods $V_i \subset \mathbb{D}$ of the marked points v_i together with conformal isomorphisms $V_i \setminus \{v_i\} \to (-\infty, 0] \times [0, 1]$. We complete $\dot{\mathbb{D}}$ by replacing each

$V_i \setminus \{v_i\}$ with the compactification $\widehat{V}_i = [-\infty, 0] \times [0,1]$, to construct the compact surface $\widehat{\mathbb{D}}$.

Suppose that $\vec{\mathbf{x}} = (\mathbf{x}_{i,i+1})$ with $\mathbf{x}_{i,i+1} \in L_i \cap L_{i+1}$ is a chain of intersection points for the transversal chain $\mathbb{L}$ of Lagrangian submanifolds in (M, ω). As in Definitions 6.4.1 and 6.4.2, we get the following

Definition 6.10.2. *A **Whitney** m**-gon** through $\vec{\mathbf{x}}$ is a smooth map $u\colon \dot{\mathbb{D}} \to M$ which extends to $\mathbb{D}$ continuously and satisfies the conditions*

$$\lim_{z \to v_i} u(z) = x_{i,i+1} \qquad u(\overline{v_i v_{i+1}}) \subset L_i.$$

The set of homotopy classes of such maps will be denoted by $W^P(\vec{\mathbf{x}})$.

The Maslov index, following the construction in Proposition 6.5.6, can be readily extended to the case of holomorphic polygons, as follows. For each $\mathbf{x} \in L_i \cap L_{i+1}$, choose a preferred path $\gamma_{\mathbf{x}} = \{\Lambda_{\mathbf{x}}(t)\}_{t \in [0,1]}$ in $\mathcal{L}Gr(T_{\mathbf{x}}M)$ from $(TL_i)_{\mathbf{x}}$ to $(TL_{i+1})_{\mathbf{x}}$.

Given a Whitney m-gon u, we can define a closed loop in the bundle of Lagrangian subspaces of $u^*_{\partial}(TM)$ as follows. Consider the the pull-backs of paths $u^*(TL_1 \cup \ldots TL_m)$, and the close it up by attaching $\gamma_{\mathbf{x}_{i,i+1}}$ at the corners. Using a trivialization of $u^*(TM)$ over $\mathbb{D}$, we can view this as a closed path of the Lagrangian Grassmannian in a fixed symplectic vector space. Evaluating the universal Maslov cycle on this closed path gives the Maslov index of u.

Chapter 7

Symmetric products

This chapter is devoted to the discussion of relevant properties of the symmetric products of two-dimensional manifolds.

In Section 7.1 we describe symmetric products of two-manifolds, and in Section 7.2 the algebraic topology of these symmetric products is discussed. In Sections 7.3 and 7.4 we identify the cohomology class corresponding to the diagonal and the first Chern class of the symmetric product of a Riemann surface. In Section 7.5 the second homology classes representable by spheres will be identified, and finally in Section 7.6 we present the description of a symplectic structure on the symmetric product of a two-dimensional manifold. The material presented here draws heavily from the references [**7, 19, 71, 110**].

7.1. Symmetric products of a topological space

Let X be a topological space, and fix an integer $m > 0$. We can form the *m-fold symmetric product of X*, a new topological space consisting of unordered m-tuples of points in X. More precisely, we form the m-fold Cartesian product $\prod_{i=1}^{m} X = \times^m(X)$ of X, and then divide it out by the natural action of the symmetric group $\mathfrak{S}_m$ on m letters, which acts by permuting factors in the Cartesian product. In this manner, we form a new topological space $\mathrm{Sym}^m(X)$. By construction, there is a quotient map

$$\pi\colon \times^m(X) \longrightarrow \mathrm{Sym}^m(X). \tag{7.1}$$

For example, the second symmetric product of the circle S^1 is homeomorphic to the Möbius strip, where the boundary is modeled on the pairs of points which coincide. Of course, the Möbius strip is homotopy equivalent to its

central circle which, in the present model, is identified with pairs of antipodal points in S^1.

Exercise 7.1.1. *Show that for any positive integer m, the m-fold symmetric product of the circle S^1 is homotopy equivalent to the circle.*

If $f\colon X \to Y$ is a continuous map, then there is a naturally induced continuous map $\mathrm{Sym}^m(f)\colon \mathrm{Sym}^m(X) \to \mathrm{Sym}^m(Y)$, defined by

$$\mathrm{Sym}^m(f)\{x_1, \ldots, x_m\} = \{f(x_1), \ldots, f(x_m)\}.$$

In particular, if f is a homeomorphism, then $\mathrm{Sym}^m(f)$ is a homeomorphism.

Let $\Delta \subset \mathrm{Sym}^m(X)$ be the *diagonal* in the symmetric product, consisting of those points where at least two of the coordinates coincide. Thus, the diagonal consists of orbits in $\times^m(X)$ where the symmetric group action is not free. Away from this set, the symmetric group acts freely, and hence if X is an n-dimensional manifold, $\mathrm{Sym}^m(X) \setminus \Delta$ is an $(m \cdot n)$-dimensional manifold.

Consider for simplicity $\mathrm{Sym}^2(X^n)$ when X^n is an n-dimensional manifold. In this case, the diagonal Δ is identified with X itself, and a neighborhood of the diagonal in $\mathrm{Sym}^2(X^n)$ is a bundle over X, whose fibers are modeled on the cone on $\mathbb{RP}^{n-1}$.

Exercise 7.1.2. *Show that the second symmetric product of an n-manifold is not a manifold if $n > 2$.*

We consider now symmetric products of two-manifolds. The fundamental theorem of algebra gives an identification of $\mathrm{Sym}^m(\mathbb{C})$ with $\mathbb{C}^m$: given an unordered m-tuple of complex numbers $\{x_1, \ldots, x_m\}$, there is a unique monic polynomial whose roots are $\{x_1, \ldots, x_m\}$, which in turn is identified with an element of $\mathbb{C}^m$ by taking its various coefficients.

Exercise 7.1.3. *Prove that $\mathrm{Sym}^m(S^2)$ is homeomorphic to $\mathbb{CP}^m$.*

The *discriminant* of the monic complex polynomial $p(z) = z^m + a_1 z^{m-1} + \cdots + a_m$ with roots $\{x_1, \ldots, x_m\}$ is defined as the product

$$D(p) = \prod_{i<j} (x_i - x_j)^2.$$

Since this expression is a symmetric polynomial of the roots, and the coefficients of $p(z)$ are elementary symmetric polynomials of the roots, $D(p)$ is a polynomial of the coefficients a_i of $p(z)$. In fact, $D(p)$ is equal to $(-1)^{\frac{m(m-1)}{2}} R(p, p')$, where $R(p, p')$ is the *resultant* of p and its derivative p', and is defined by the following $(2m-1) \times (2m-1)$ determinant:

$$\begin{vmatrix}
1 & a_1 & a_2 & \dots & a_{m-1} & a_m & 0\dots & \dots & 0 \\
0 & 1 & a_1 & a_2 & \dots & a_{m-1} & a_m & 0\dots & 0 \\
\vdots & & & & & & & & \vdots \\
0 & \dots & 0 & 1 & a_1 & a_2 & \dots & a_{m-1} & a_m \\
m & (m-1)a_1 & (m-2)a_2 & \dots & a_{m-1} & 0 & \dots & \dots & 0 \\
0 & m & (m-1)a_1 & (m-2)a_2 & \dots & a_{m-1} & 0 & \dots & 0 \\
\vdots & & & & & & & & \vdots \\
0 & 0 & \dots & 0 & m & (m-1)a_1 & (m-2)a_2 & \dots & a_{m-1}
\end{vmatrix}$$

(See for example [**140**, Chapter 5.8].) It follows from the definition that $D(p) = 0$ if and only if p has repeated roots. In the space of monic polynomials we define the *discriminant locus* as the vanishing set of the above determinant, corresponding to the set of polynomials with repeated roots. Under the identification $\mathrm{Sym}^m(\mathbb{C}) \cong \mathbb{C}^m$, the diagonal $\Delta \subset \mathrm{Sym}^m(\mathbb{C})$ is mapped to the discriminant locus.

The fact that $\mathrm{Sym}^m(\mathbb{C})$ is a manifold has the following generalization:

Proposition 7.1.4. *If Σ is an orientable two-manifold, then $\mathrm{Sym}^m(\Sigma)$ admits a smooth structure. Furthermore, a complex structure j on Σ induces a complex structure on $\mathrm{Sym}^m(\Sigma)$, uniquely characterized by the property that if $\times^m(\Sigma)$ is equipped with its product complex structure induced from j, then the quotient map $\pi\colon \times^m(\Sigma) \to \mathrm{Sym}^m(\Sigma)$ is holomorphic.*

The proof will use a lemma which, in turn, will rely on the following Riemann Extension Theorem from several complex variables, a simple consequence of the Cauchy integral formula; see for example [**42**]. We state the result here for the reader's convenience:

Theorem 7.1.5. *(Riemann Extension Theorem) Let $\Omega \subset \mathbb{C}^m$ be an open set. Suppose that $f\colon \Omega \to \mathbb{C}$ is a bounded holomorphic function, and let $g\colon \Omega \setminus f^{-1}(0) \to \mathbb{C}$ be a bounded, holomorphic function. Then g extends to a holomorphic function on all of Ω.* □

Lemma 7.1.6. *Let D_1 and D_2 be two bounded open sets in $\mathbb{C}$ and $\phi\colon D_1 \to D_2$ a biholomorphism (that is, a bijective holomorphic map) between them. Then the induced map*

$$\mathrm{Sym}^m(\phi)\colon \mathrm{Sym}^m(D_1) \to \mathrm{Sym}^m(D_2)$$

is a biholomorphism between subsets of $\mathrm{Sym}^m(\mathbb{C}) \cong \mathbb{C}^m$.

Proof. Denote the quotient map by $\pi\colon \mathbb{C}^m = \times^m(\mathbb{C}) \to \mathrm{Sym}^m(\mathbb{C})$, and the identification $\mathrm{Sym}^m(\mathbb{C}) \cong \mathbb{C}^m$ coming from the fundamental theorem of algebra by $\theta\colon \mathrm{Sym}^m(\mathbb{C}) \to \mathbb{C}^m$.

We wish to prove that $\theta \circ \mathrm{Sym}^m(\phi) \circ \theta^{-1}$, defined on the open set $\theta(\mathrm{Sym}^m(D_1)) \subset \mathbb{C}^m$, is a holomorphic map. To this end, consider the diagram of continuous maps

$$\begin{array}{ccc} \times^m(D_1) & \xrightarrow{\times^m(\phi)} & \times^m(D_2) \\ \pi \downarrow & & \pi \downarrow \\ \mathrm{Sym}^m(D_1) & \xrightarrow{\mathrm{Sym}^m(\phi)} & \mathrm{Sym}^m(D_2) \\ \theta \downarrow & & \theta \downarrow \\ \mathbb{C}^m & & \mathbb{C}^m \end{array}$$

The diagonal $\Delta \subset \mathrm{Sym}^m(\mathbb{C})$ is mapped by θ to the discriminant locus $\Delta_0 \subset \mathbb{C}^m$, which is the set of points $a_1, \ldots, a_m$ where the polynomial $p(z) = z^m + a_1 z^{m-1} + \cdots + a_m$ has a repeated root. The inverse function theorem guarantees that $(\theta \circ \pi)|_{\mathbb{C}^m \setminus \pi^{-1}(\Delta)} \colon \mathbb{C}^m \setminus \pi^{-1}(\Delta) \to \mathbb{C}^m \setminus \Delta_0$ is a holomorphic covering map. Thus, since $\times^m(\phi)$ is holomorphic, the restriction of $\theta \circ \mathrm{Sym}^m(\phi) \circ \theta^{-1}$ to $(\mathbb{C}^m \setminus \Delta_0) \cap \theta(\mathrm{Sym}^m(D_1))$ can be written as a composition of (locally defined) holomorphic maps. Since $\theta \circ \mathrm{Sym}^m(\phi) \circ \theta^{-1}$ is a continuous map, and $\Delta_0 \subset \mathbb{C}^m$ is an algebraic variety (the zero set of the discriminant in $a_1, \ldots, a_m$), it follows from the Riemann Extension Theorem that $\theta \circ \mathrm{Sym}^m(\phi) \circ \theta^{-1}$ is holomorphic. Clearly, its (holomorphic) inverse is supplied by $\theta \circ \mathrm{Sym}^m(\phi^{-1}) \circ \theta^{-1}$. □

Proof. [of Proposition 7.1.4] We start by constructing the desired complex structure on $\mathrm{Sym}^m(\Sigma)$.

The complex structure on Σ provides an atlas $\{\phi_i \colon U_i \to \Sigma\}_i$, where $U_i \subset \mathbb{C}$ are open subsets, whose transition maps

$$\phi_j^{-1} \circ \phi_i \colon \phi_i^{-1}(\phi_j(U_j)) \to U_j$$

are holomorphic. Clearly, $\mathrm{Sym}^m(\Sigma)$ is covered by charts indexed by partitions $(d_1, \ldots, d_k)$ of m, and k-tuples of charts $\{\phi_{n_i} \colon U_{n_i} \to \Sigma\}_{i=1}^k$ whose images are disjoint. The parametrization of the corresponding chart $\mathrm{Sym}^{d_1}(U_{n_1}) \times \cdots \times \mathrm{Sym}^{d_k}(U_{n_k})$ is given by $\mathrm{Sym}^{d_1}(\phi_{n_1}) \times \cdots \times \mathrm{Sym}^{d_k}(\phi_{n_k})$. The fact that the transition maps for this system of charts are holomorphic is a direct consequence of Lemma 7.1.6.

To see that the complex structure is uniquely characterized by the property that π is holomorphic, suppose that $\mathrm{Sym}^m(\Sigma)$ is equipped with two different complex structures J_1 and J_2 for which the quotient map $\pi \colon \times^m(\Sigma) \to \mathrm{Sym}^m(\Sigma)$ is holomorphic. Then, the identity map, thought of as a map from $\mathrm{Sym}^m(\Sigma)$ equipped with the complex structure J_1 to the same space equipped with the complex structure J_2 is continuous and holomorphic away

from Δ. Using the Riemann Extension Theorem, the identity extends as a holomorphic function across Δ. It follows that $J_1 = J_2$. □

Remark 7.1.7. *The proof of Lemma 7.1.6 shows that the symmetric product of a holomorphic map between Riemann surfaces is a holomorphic map. By contrast, the symmetric product of a diffeomorphism need not be smooth. For example, consider* $\mathrm{Sym}^2(\mathbb{C})$*, let* $\phi\colon \mathbb{C} \to \mathbb{C}$ *be the map* $\phi(x+iy) = x+2iy$*, and consider the smooth path* $\gamma\colon \mathbb{R} \to \mathrm{Sym}^2(\mathbb{C})$ *given by* $t \mapsto \{\sqrt{t}, -\sqrt{t}\}$*. Clearly,* $\theta \circ \gamma$ *is the smooth path in* $\mathbb{C}^2$ *given by* $t \mapsto (0, -t)$*. By contrast,* $\theta \circ \mathrm{Sym}^2(\phi) \circ \gamma$ *is the path in* $\mathbb{C}^2$ *that sends* $t \geq 0$ *to* $(0, -t)$ *and* $t \leq 0$ *to* $(0, -4t)$*.*

Following [**19**], we enhance Proposition 7.1.4 to cases where the complex structure on Σ varies. The key point is the Ahlfors-Bers generalization of the Riemann mapping theorem [**1**], which we state without proof. Recall that a metric g on an oriented two-manifold induces an almost-complex structure j_g (defined by rotation of tangent vectors by 90° counterclockwise); indeed, any almost-complex structure on a two-manifold arises in this way. Moreover, this almost-complex structure can be integrated to give a complex structure on Σ; see Remark 4.3.5.

Theorem 7.1.8. *(Ahlfors-Bers) Suppose that* $\{g_t\}_{t\in[0,1]}$ *is a smooth family of smooth metrics on the disk* D*, and* j_{g_t} *denotes the almost-complex structure on* D *corresponding to* g_t*. Let* $\psi^{g_t}\colon (D, j) \to (D, j_{g_t})$ *denote unique conformal identification sending* 0 *to* 0*, where* j *is the standard complex structure on* D *induced from* $\mathbb{C}$*. Then, the map* $(t, z) \mapsto (t, \psi^{g_t}(z))$ *is smooth. Moreover, if* $\{g_t\}_{t\in D}$ *is a holomorphically varying family of metrics, then the map* $(t, z) \mapsto (t, \psi^{g_t}(z))$ *is holomorphic.* □

Proposition 7.1.9. *For different choices of complex structures, the complex manifolds* $\mathrm{Sym}^m(\Sigma)$ *are diffeomorphic.*

Proof. Note first that any two complex structures j_0 and j_1 over Σ can be thought of as associated to two smooth metrics g_0 and g_1 over the same smooth manifold Σ. Any two metrics, in turn, can be connected by a smooth, one parameter family of metrics $\{g_t\}_{t\in[0,1]}$ over Σ. We will now give $[0,1] \times \mathrm{Sym}^m(\Sigma)$ the structure of a smooth manifold so that the projection map $p\colon [0,1] \times \mathrm{Sym}^m(\Sigma) \to [0,1]$ is a smooth submersion, and $p^{-1}(t)$ is identified with $\mathrm{Sym}^m(\Sigma)$, equipped with the smooth structure it inherits from the metric g_t through the almost-complex structure induced by g_t.

To this end, let $U_i \subset \Sigma$ be a simply-connected coordinate chart on Σ. Consider the map

$$\psi_i\colon [0,1] \times D \to [0,1] \times \Sigma$$

defined by $\psi_i(t,z) = (t, \psi^t(z))$, where ψ^t is the Ahfors-Bers unformization of U_i with the metric $g_t|_{U_i}$. These specify coordinate charts on $[0,1] \times \mathrm{Sym}^m(\Sigma)$ mapping $(t, \mathbf{w}_1, \ldots, \mathbf{w}_k) \in ([0,1], \mathrm{Sym}^{d_1}(D) \times \cdots \times \mathrm{Sym}^{d_k}(D))$ to the point $(t, \mathrm{Sym}^{d_1}(\psi^t_{n_1})(\mathbf{w}_1) \times \cdots \times \mathrm{Sym}^{d_k}(\psi^t_{n_k})(\mathbf{w}_k))$. The transition maps are smooth in t by Theorem 7.1.8, and they are smooth in the other variables by Lemma 7.1.6.

We have argued that the transition functions are smooth in the parameter t and the parameters in the symmetric product separately. In fact, the transition functions are smooth in all the variables at once. This is true because of the following extension of Lemma 7.1.6: if D_1 and D_2 are two bounded open subsets in $\mathbb{C}$, and $\Phi\colon [0,1] \times D_1 \to D_2$ is smooth map whose restriction $\Phi_t = \Phi|_{\{t\}\times D_1}\colon D_1 \to D_2$ is a biholomorphism, then the map

$$\mathrm{Sym}^m(\Phi)\colon [0,1] \times \mathrm{Sym}^m(D_1) \to [0,1] \times \mathrm{Sym}^m(D_2),$$

whose restriction to $\{t\}\times\mathrm{Sym}^m(D_1)$ coincides with the map $\mathrm{Sym}^m(\Phi_t)$ from Lemma 7.1.6 is smooth. This in turn follows from a parametrized version ot Theorem 7.1.5: if $\Omega \subset \mathbb{C}^m$ is an open set, $f\colon [0,1]\times\Omega \to \mathbb{C}$ is a bounded holomorphic function, and $G\colon [0,1]\times(\Omega\setminus f^{-1}(0)) \to \mathbb{C}$ is a bounded function whose restriction $G_t = G|_{\{t\}\times(\Omega\setminus f^{-1}(0))}$ is bounded and holomorphic, then the function $G'\colon [0,1]\times\Omega \to \mathbb{C}$ whose restriction to $\{t\}\times\Omega$ coincides with the holomorphic extension of G_t from Theorem 7.1.5 is smooth. This is true because the Riemann extension is defined using the Cauchy integral formula (see [**42**]); and the integral depends smoothly on the integrand.

Thus, the above charts on $[0,1]\times\mathrm{Sym}^m(\Sigma)$ make it into a smooth manifold with boundary, endowed with a submersion, the projection to $[0,1]$, whose fiber over $t \in [0,1]$ is diffeomorphic to with $\mathrm{Sym}^m(\Sigma, j_t)$. Integrating gradient trajectories gives a diffeomorphism between $\mathrm{Sym}^m(\Sigma)$ with the smooth structures it inherits from j_0 and j_1 respectively. □

Remark 7.1.10. *There are other approaches to constructing a holomorphic structure on* $\mathrm{Sym}^m(\Sigma)$. *For example, one can show that* $\mathrm{Sym}^m(\Sigma)$, *equipped with the sheaf of* $\mathfrak{S}_m$*-invariant holomorphic functions on* $\times^m(\Sigma)$, *gives* $\mathrm{Sym}^m(\Sigma)$ *the structure of an analytic space; see* [**57**]. *In a related vein, if we start with a projective algebraic curve, geometric invariant theory gives* $\mathrm{Sym}^m(\Sigma)$ *as a projective algebraic variety; i.e. in addition to a complex structure, this equips* $\mathrm{Sym}^m(\Sigma)$ *also with a compatible symplectic structure. We outline the construction of another symplectic structure on* $\mathrm{Sym}^m(\Sigma)$ *in Section 7.6.*

7.2. The algebraic topology of $\mathrm{Sym}^m(\Sigma)$

We review some facts about the algebraic topology of the m-fold symmetric product of a two-dimensional manifold, mostly following [**71**]; see also [**7**].

As a warm-up, we start with the fundamental group.

Lemma 7.2.1. *For all $m > 1$, there is an isomorphism*

$$\pi_1(\mathrm{Sym}^m(\Sigma)) \cong H_1(\Sigma;\mathbb{Z}).$$

Proof. Fix a reference point $w \in \Sigma$, and let $\mathbf{w} = \{w,\dots,w\} \in \mathrm{Sym}^m(\Sigma)$. Consider the map $(j_1)_*\colon \pi_1(\Sigma,w) \longrightarrow \pi_1(\mathrm{Sym}^m(\Sigma),\mathbf{w})$ induced by the inclusion map

$$j_1\colon \Sigma \times \{w\} \times \ldots \times \{w\} \hookrightarrow \mathrm{Sym}^m(\Sigma).$$

To see that $(j_1)_*$ is surjective, observe that a generic representative of a given homotopy class in $\pi_1(\mathrm{Sym}^m(\Sigma),\mathbf{w})$ meets the diagonal only at the basepoint $\mathbf{w}$. Thus, every homotopy class has a representative which can be written as a product of m curves $\{\gamma_1,\dots,\gamma_m\}$. Such a curve, in turn, can easily be seen to be homotopic to the curve

$$\{\gamma_1 * \ldots * \gamma_m, \overbrace{w,\dots,w}^{m-1}\}$$

in $\mathrm{Sym}^m(\Sigma)$, verifying surjectivity.

Next, we claim that $\pi_1(\mathrm{Sym}^m(\Sigma),\mathbf{w})$ is Abelian, when $m \geq 2$. This follows from the fact that if γ_1 and γ_2 are two closed curves in Σ based at w, then the composite of $(j_1)_*(\gamma_1)$ and $(j_1)_*(\gamma_2)$ is homotopic to the curve induced from

$$t \mapsto \{\gamma_1(t), \gamma_2(t), \overbrace{w,\dots,w}^{m-2}\} \in \mathrm{Sym}^m(\Sigma),$$

where γ_1 and γ_2 play symmetric roles.

From the above, it follows that, $\pi_1(\mathrm{Sym}^m(\Sigma);\mathbf{w}) \cong H_1(\mathrm{Sym}^m(\Sigma);\mathbb{Z})$, and j_1 induces a surjection $H_1(\Sigma;\mathbb{Z}) \to H_1(\mathrm{Sym}^m(\Sigma);\mathbb{Z})$. It remains to show that j_1 induces an injection on $H_1(\Sigma;\mathbb{Z})$. To this end, we construct a map $\theta\colon H^1(\Sigma;\mathbb{Z}) \to H^1(\mathrm{Sym}^m(\Sigma);\mathbb{Z})$ with the property that for any $\alpha \in H_1(\Sigma;\mathbb{Z})$ and $\xi \in H^1(\Sigma;\mathbb{Z})$, there is an identification of Kronecker pairings

$$\langle \theta(\xi), j_1(\alpha)\rangle = \langle \xi, \alpha\rangle. \tag{7.2}$$

This map θ is constructed as follows. Any $\xi \in H^1(\Sigma;\mathbb{Z})$ is the pull-back of the generator of $H^1(S^1;\mathbb{Z}) \cong \mathbb{Z}$ by some map $f\colon \Sigma \to S^1$. Given such a map f, we can form another map $F\colon \mathrm{Sym}^m(\Sigma) \to S^1$ by

$$F(\{x_1,\dots,x_m\}) = \prod_{i=1}^{m} f(x_i).$$

Define $\theta(\xi)$ to be the pullback of the generator of $H^1(S^1;\mathbb{Z}) \cong \mathbb{Z}$ by F. Equation (7.2) follows readily from this definition; and the stated injectivity follows. □

Next we describe the cohomology ring of $\mathrm{Sym}^m(\Sigma)$. In order to do so, we need a few preparatory definitions and constructions.

There is a map

$$\mu\colon H_1(\Sigma;\mathbb{Z}) \longrightarrow H^1(\mathrm{Sym}^m(\Sigma);\mathbb{Z})$$

defined as the composition of maps

$$\begin{aligned} H_1(\Sigma;\mathbb{Z}) \cong H^1(\Sigma;\mathbb{Z}) &\cong \mathrm{Hom}(H_1(\Sigma;\mathbb{Z});\mathbb{Z}) \cong \mathrm{Hom}(H_1(\mathrm{Sym}^m(\Sigma);\mathbb{Z}),\mathbb{Z}) \\ &\cong H^1(\mathrm{Sym}^m(\Sigma);\mathbb{Z}), \end{aligned}$$

where the first map is provided by Poincaré duality on Σ, and the third is provided by Lemma 7.2.1.

A further two-dimensional cohomology class U is defined as follows. A point $w \in \Sigma$ gives rise to a codimension two submanifold $\{w\} \times \mathrm{Sym}^{m-1}(\Sigma)$ of $\mathrm{Sym}^m(\Sigma)$, consisting of m-tuples of points $\{x_1, \ldots, x_m\}$, where at least one of the $x_i = w$. The Poincaré dual to the homology class represented by this closed submanifold is denoted $U \in H^2(\mathrm{Sym}^m(\Sigma);\mathbb{Z})$. (One can think of U as $\mu(\{w\})$, where $\{w\}$ is the 0-dimensional homology class generating $H_0(\Sigma;\mathbb{Z})$ and $\mu\colon H_0(\Sigma;\mathbb{Z}) \to H^2(\mathrm{Sym}^m(\Sigma);\mathbb{Z})$.)

Let $\{A_i, B_i\}_{i=1}^g$ be a standard set of oriented, simple closed curves in Σ; i.e. for all distinct $i, j \in \{1, \ldots, g\}$, $A_i \cap A_j = \emptyset$, $A_i \cap B_j = \emptyset$, and $B_i \cap B_j = \emptyset$; while $A_i \cap B_i$ consists of a single intersection point, which is transverse and positive.

The next theorem (from [**71**]) states that the cohomology ring of $\mathrm{Sym}^m(\Sigma)$ is generated by the image of μ and U. More precisely:

Theorem 7.2.2. *(MacDonald) The cohomology ring $H^*(\mathrm{Sym}^m(\Sigma);\mathbb{Z})$ is generated by the μ-classes and by U. The relations in this ring are of the form*

$$0 = U^r \cdot \prod_{i\in I}(U - \mu(A_i)\cdot\mu(B_i)) \prod_{j\in J}\mu(A_j) \prod_{k\in K}\mu(B_k) \tag{7.3}$$

where $I, J, K \subset \{1, \ldots, g\}$ are arbitrary pairwise disjoint subsets, and r is a non-negative integer with $r + 2|I| + |J| + |K| \geq m + 1$.

We will prove the above statement first with rational coefficients, and return to the integral case afterward. Before turning to the proof, we make a few preliminary remarks.

Think of $\mathrm{Sym}^m(\Sigma)$ as the quotient of $\times^m(\Sigma)$ by the action of the permutation group $\mathfrak{S}_m$ on m letters, with quotient map $\pi\colon \times^m(\Sigma) \to \mathrm{Sym}^m(\Sigma)$. Clearly, the pull-back map

$$\pi^*\colon H^*(\mathrm{Sym}^m(\Sigma)) \to H^*(\times^m(\Sigma))$$

maps into the subring $H^*(\times^m(\Sigma))^{\mathfrak{S}_m}$ of cohomology classes that are invariant under the action of $\mathfrak{S}_m$. The following general result (for a proof, see for example [**11**, Theorem 2.4]) ensures that π^* induces an isomorphism over $\mathbb{Q}$:

Theorem 7.2.3. *Let X be a simplicial complex, and let G be a finite group that acts simplicially on X. Then, the pull-back map $p\colon X \to X/G$ induces an isomorphism between $H^*(X/G;\mathbb{Q})$ and the G-invariant cohomology classes in X, $H^*(X;\mathbb{Q})^G$.* □

Specializing to the Cartesian product of Σ with itself, we have the following:

Proposition 7.2.4. *Let $\mathfrak{S}_m$ act on $\otimes_{i=1}^m H^*(\Sigma;\mathbb{Q})$ by permuting the tensor factors (with a sign when permuting odd generators). The map π^* induces an isomorphism over $\mathbb{Q}$:*

$$\pi^*\colon H^*(\mathrm{Sym}^m(\Sigma);\mathbb{Q}) \to H^*(\times^m(\Sigma);\mathbb{Q})^{\mathfrak{S}_m}.$$

Proof. The symmetric group $\mathfrak{S}_m$ acts on the space $\times^m(\Sigma)$, inducing an action of $\mathfrak{S}_m$ on the cohomology ring $H^*(\times^m(\Sigma);\mathbb{Q})$. Under the Künneth formula $H^*(\times^m(\Sigma);\mathbb{Q}) \cong \otimes^m H^*(\Sigma;\mathbb{Q})$ that group action corresponds to the group action stated in the proposition. Thus, the result follows from Theorem 7.2.3. □

Lemma 7.2.5. *The vector space $H^*(\mathrm{Sym}^m(\Sigma);\mathbb{Q})$ is spanned by the elements*

$$w(r,I,J,K) = U^r \cdot \prod_{i\in I}(U - \mu(A_i)\cdot\mu(B_i)) \prod_{j\in J}\mu(A_j) \prod_{k\in K}\mu(B_k)$$

where $I,J,K \subset \{1,\dots,g\}$ are disjoint subsets, r is a non-negative integer, and $r + 2|I| + |J| + |K| \le m$; moreover, those elements $w(r,I,J,K)$ with $r + 2|I| + |J| + |K| > m$ are zero.

Proof. $H^*(\times^m(\Sigma);\mathbb{Q})$ is generated by cohomology classes of the form $a_1 \otimes \cdots \otimes a_m$, where $a_i \in H^{n_i}(\Sigma;\mathbb{Q})$, for some sequence $n_1,\dots,n_m$. Given any class ξ in $H^*(\times^m(\Sigma);\mathbb{Q})$, we can construct a $\mathfrak{S}_m$-invariant class by adding up all the classes in the $\mathfrak{S}_m$-orbit through ξ; i.e. forming

$$\sum_{\sigma\in\mathfrak{S}_m} \sigma^*(\xi).$$

Indeed, all the $\mathfrak{S}_m$-invariant cohomology classes in $H^*(\times^m(\Sigma);\mathbb{Q})$ arise in this way.

Now, $w(r,I,J,K)$ is the $\mathfrak{S}_m$-orbit of the class

(7.4)
$$\overbrace{\nu\otimes\cdots\otimes\nu}^{r}\otimes(\otimes_{i\in I}\mathrm{PD}(A_i)\otimes\mathrm{PD}(B_i))\otimes(\otimes_{j\in J}\mathrm{PD}(A_j))\otimes(\otimes_{k\in K}\mathrm{PD}(B_k)),$$

where $\nu\in H^2(\Sigma;\mathbb{Z})$ is a generator, and $\mathrm{PD}([\alpha])\in H^1(\Sigma;\mathbb{Z})$ is the Poincaré dual to $[\alpha]\in H_1(\Sigma;\mathbb{Z})$. Since these latter classes generate $H^*(\times^m(\Sigma);\mathbb{Q})$, it follows from $H^*(\times^m(\Sigma);\mathbb{Q})^{\mathfrak{S}_m}\cong H^*(\mathrm{Sym}^m(\Sigma);\mathbb{Q})$ that the classes $w(r,I,J,K)$ generate $H^*(\mathrm{Sym}^m(\Sigma);\mathbb{Q})$. The quantity $r+2|I|+|J|+|K|$ is the number of tensor factors in $\otimes^m H^*(\Sigma;\mathbb{Q})$ in which the corresponding class specified by Equation (7.4) has non-zero grading. Thus, when $r+2|I|+|J|+|K|>m$, the class is obviously zero. □

Proposition 7.2.6. *Theorem 7.2.2 holds with $\mathbb{Q}$ coefficients; i.e. $H^*(\mathrm{Sym}^m(\Sigma);\mathbb{Q})$ is generated by the μ-classes and by U, and the relations are given by Equation* (7.3).

According to Lemma 7.2.5, the relations from Equation (7.3) hold in $H^*(\mathrm{Sym}^m(\Sigma);\mathbb{Q})$.

To prove that these are precisely the relations, we proceed as follows: we construct the (graded) model ring R_* which is freely generated by the μ-classes and U, divide it by the ideal $\mathcal{I}$ generated by the stated relations, verify that it satisfies an analogue of Poincaré duality (see Lemma 7.2.7 below), and with the help of that property, prove that the natural map Φ from R_* to $H^*(\mathrm{Sym}^m(\Sigma);\mathbb{Q})$ induces an isomorphism from $R/\mathcal{I}$.

In more detail, let $R=R_*$ be the exterior algebra over $\mathbb{Q}$ on generators $\{a_i,b_i\}_{i=1}^g$ tensored with the polynomial algebra in a formal variable u. There is a grading gr specified by the following requirements: the grading is additive under multiplication, the elements with $\mathrm{gr}=0$ are $\mathbb{Q}$-multiples of the multiplicative unit 1, R_1 is the vector space spanned by $\{a_i,b_i\}_{i=1}^g$, and u lies in R_2. For each triple of disjoint subsets $I,J,K\subset\{1,\dots,g\}$ and each integer $r\geq 0$, let

$$w_0(r,I,J,K)=u^r\cdot\left(\prod_{i\in I}(u-a_i\cdot b_i)\right)\left(\prod_{j\in J}a_j\right)\left(\prod_{k\in K}b_k\right). \tag{7.5}$$

It is clear that the elements $w_0(r,I,J,K)$ form a basis for R, thought of as a vector space over $\mathbb{Q}$, and

$$\mathrm{gr}(w_0(r,I,J,K))=2r+2|I|+|J|+|K|. \tag{7.6}$$

We equip R with a filtration, meaning a sequence of vector space inclusions

$$R = R^0 \supset R^1 \supset \cdots \supset R^k \supset R^{k+1} \supset \dots$$

where R^i are vector subspaces with $\cap_k R^k = 0$. The filtration is specified by the function F on the basis

$$F(w_0(r, I, J, K)) = r + 2|I| + |J| + |K|;$$

in particular, $F(1) = 0$, $F(a_i) = F(b_i) = F(u) = 1$, $F(u - a_i b_i) = 2$. (The reader should be warned that the equations specifying gr and F are very similar, but not the same.) An element lies in R^t if it lies in the vector space span of those elements $w_0(r_k, I_k, J_k, K_k)$ with $F(w_0(r_k, I_k, J_k, K_k) \geq t$.

We collect some properties of this filtration in the following lemma, to assist in the proof of Proposition 7.2.6.

Lemma 7.2.7. *The graded ring R_* has the following properties.*

(R-1) For any integer $t \geq 0$, the vector subspace $R^t \subset R$ is an ideal; and hence the quotient R/R^t inherits the grading gr *on R_*.*

(R-2) The portion of R/R^{m+1} in grading $2m$ is one-dimensional, generated by u^m.

(R-3) Multiplication in R/R^{m+1} induces a bilinear pairing

$$(R_d/R_d^{m+1}) \otimes (R_{2m-d}/R_{2m-d}^{m+1}) \to R_{2m}/R_{2m}^{m+1} \cong \mathbb{Q},$$

which is a perfect pairing; i.e. if $a \in R_d/R_d^{m+1}$ is a non-zero algebra element, then there is a $b \in R_{2m-d}/R_{2m-d}^{m+1}$ so that $a \cdot b = u^m$.

Proof. It is easy to verify the following relations:

$$\begin{array}{ll} (u - a_i b_i)^2 = 2u(u - a_i b_i) - u^2, & a_i \cdot b_i = u - (u - a_i b_i), \\ (u - a_i b_i) \cdot a_i = u a_i & (u - a_i b_i) \cdot b_i = u b_i. \end{array} \tag{7.7}$$

It follows quickly that for any integer $t \geq 0$, $R^t \subset R$ is an ideal. Since each basis vectors $w_0(r, I, J, K)$ is homogeneous with respect to the grading gr, it follows that gr descends to R/R^t, completing Property (R-1).

To see Condition (R-2), observe that the conditions that $w_0(r, I, J, K)$ lies in grading $2m$ and its projection to R/R^{m+1} is non-zero (equivalently, $F(w_0(r, I, J, K)) \leq m$) can be expressed as

$$\begin{aligned} 2m &= 2r + 2|I| + |J| + |K| \\ m &\geq r + 2|I| + |J| + |K|; \end{aligned}$$

so $I = J = K = \emptyset$ and $r = m$.

We prove Property (R-3) by showing that if (r_1, I_1, J_1, K_1) and (r_2, I_2, J_2, K_2) are chosen so that

$$(2r_1 + 2|I_1| + |J_1| + |K_1|) + (2r_2 + 2|I_2| + |J_2| + |K_2|) = 2m,$$

then

(7.8)

$$[w_0(r_1, I_1, J_1, K_1)]\cdot[w_0(r_2, I_2, J_2, K_2)] = \begin{cases} [u]^m & \text{if } I_1 = I_2,\ J_1 = K_2,\ K_1 = J_2 \\ 0 & \text{otherwise.} \end{cases}$$

To see this, first observe that

$$w_0(r_1, I_1, J_1, K_1) \cdot w_0(r_2, I_1, K_1, J_1) = u^{r_1+r_2} + R^{r_1+r_2+1} \tag{7.9}$$

Then the first line in Equation (7.8) follows quickly from Equation (7.7). Conversely, suppose that the expansion of $w_0(r_1, I_1, J_1, K_1)\cdot w_0(r_2, I_2, J_2, K_2)$ in the basis $\{w_0(r, I, J, K)\}_{r,I,J,K}$ contains a non-zero multiple of $w_0(r_3, I_3, J_3, K_3)$. Then, from Equation (7.7), it follows easily that:

$$J_3 \supseteq (I_1 \cap J_2) \cup (I_2 \cap J_1) \cup (J_2 \setminus K_1) \cup (J_1 \setminus K_2);$$
$$K_3 \supseteq (I_1 \cap K_2) \cup (I_2 \cap K_1) \cup (K_2 \setminus J_1) \cup (K_1 \setminus J_2).$$

In particular, if $J_3 = K_3 = \emptyset$, then $J_1 = K_2$, $J_2 = K_1$, and $I_1 \cup I_2$ is disjoint from $J_1 \cup J_2 = K_1 \cup K_2$. In that case, we also have (using Equation (7.7)) that

$$I_3 = I_1 \cup I_2 \setminus I_1 \cap I_2;$$

so if $I_3 = \emptyset$, then $I_1 = I_2$. The second line of Equation (7.8) follows immediately.

Equation (7.8) shows that the basis for R_d/R^{m+1} resp. R_{2m-d}/R^{m+1}_{2m-d} specified by

$$\{[w_0(r_1, I, J, K)]\}_{2r_1+2|I|+|J|+|K|=d}$$

and

$$\{[w_0(r_2, I, J, K)]\}_{2r_1+2|I|+|J|+|K|=2m-d}$$

are dual to each other under the bilinear form induced by multiplication. Thus, that form induces a non-degenerate pairing, verifying Property (R-3).

□

Proof. [of Proposition 7.2.6] Consider the graded ring homomorphism $\Phi\colon R \to H^*(\mathrm{Sym}^m(\Sigma); \mathbb{Q})$, specified by sending $a_i \to \mu(A_i)$, $b_i \to \mu(B_i)$, and u to U.

Lemma 7.2.5 ensures that Φ is surjective; and indeed, since the elements $\Phi(a_i) = A_i$, $\Phi(b_i) = B_i$ and $\Phi(u) = U$ satisfy Equation (7.3), Φ descends to a (surjective) ring homomorphism

$$\phi\colon R/\mathcal{I} \to H^*(\mathrm{Sym}^m(\Sigma); \mathbb{Q}),$$

where $\mathcal{I} = R^{m+1}$.

To verify injectivity, note first that

$$\phi(u^m) \in H^{2m}(\mathrm{Sym}^m(\Sigma); \mathbb{Q}) \cong \mathbb{Q}$$

is non-zero. This follows from the fact that $\phi(u^m)$ is the m-fold tensor product of the two-dimensional generator of $H^2(\Sigma; \mathbb{Q})$, as in Proposition 7.2.4.

More generally, if a is a non-zero element of $R/\mathcal{I}$, consider the element $b \in R/\mathcal{I}$ with $a \cdot b = [u]^m$, provided by Property (R-3). Since

$$\phi(a) \smile \phi(b) = \phi(a \cdot b) = \phi(u^m) \neq 0,$$

it follows that $\phi(a) \neq 0$, completing the proof that ϕ is injective. □

Before turning to the computation of the integral cohomology of the symmetric product, we recall some standard constructions from Riemann surface theory.

There is an algebro-geometric interpretation of $\mathrm{Sym}^m(\Sigma)$ as the space of effective, degree m divisors. Specifically, endow Σ with a complex structure, and consider the set of holomorphic line bundles $\mathcal{E}$ over Σ, equipped with a non-trivial section ϕ. This space can be naturally partitioned into components, indexed by non-negative integers m specifying the topological type of $\mathcal{E}$, giving its Euler number. The integer m measures the number of zeros of ϕ (counted with multiplicity). Declare $(\mathcal{E}_1, \phi_1)$ to be equivalent to $(\mathcal{E}_2, \phi_2)$ if there is a holomorphic isomorphism $u\colon \mathcal{E}_1 \longrightarrow \mathcal{E}_2$ of bundles, with $u^*(\phi_2) = \phi_1$. The set of equivalance classes of pairs $(\mathcal{E}, \phi)$ where the Euler number of $\mathcal{E}$ is m coincides with the symmetric product $\mathrm{Sym}^m(\Sigma)$.

The *degree m Jacobian* $\mathfrak{J}^m(\Sigma)$ over Σ is the space of holomorphic structures on the topological complex line bundle over Σ with Euler number equal to m. This space can be identified with a torus with real dimension $2g$. There is a forgetful map, called the *Abel-Jacobi map* which sends a point in $\mathrm{Sym}^m(\Sigma)$, thought of as a pair $(\mathcal{E}, \phi)$, to the isomorphism class of the line bundle $\mathcal{E}$.

Example 7.2.8. *If Σ_2 is a surface of genus 2, the Abel-Jacobi map gives a map from $\mathrm{Sym}^2(\Sigma_2)$ to T^4. In fact, standard tools in Riemann surface theory* [**42**] *show that $\mathrm{Sym}^2(\Sigma_2)$ is the blow-up of T^4 at a point, as follows. The Riemann-Roch formula shows that if $\mathcal{E}$ is a degree 2 line bundle, then $h^0(\mathcal{E}) - h^1(\mathcal{E}) = 1$, i.e. $h^0(\mathcal{E}) > 0$, so the Abel-Jacobi map is surjective.*

By Serre duality, the only degree 2 line bundle over Σ_2 that admits more than one holomorphic section is the canonical bundle $K = T^*_{\mathbb{C}}$; a bundle with $h^0(K) = 2$. Thus, the fiber over $K \in T^4$ is identified with $\mathbb{CP}^1$, the projectivization of the space of holomorphic 1-forms over Σ. It follows then that $\mathrm{Sym}^2(\Sigma_2)$ is diffeomorphic to $T^4 \# \overline{\mathbb{CP}}^2$. Note also that in this case $\pi_2(\mathrm{Sym}^2(\Sigma))$ is a countable direct sum of copies of $\mathbb{Z}$; compare with Proposition 7.5.4.

Lemma 7.2.9. *If $m > 2g-2$, the Abel-Jacobi map realizes $\mathrm{Sym}^m(\Sigma)$ as a fiber bundle over T^{2g}, whose fiber is identified with $\mathbb{CP}^{m-g}$.*

Proof. Serre duality [**42**] ensures that $h^1(\mathcal{E}) = h^0(K \otimes \mathcal{E}^*)$, where K is the canonical bundle (having degree $2g-2$). Thus $\deg(K \otimes \mathcal{E}^*) = 2g-2-m$. If $m > 2g-2$, then $h^1(\mathcal{E}) = 0$, so by Serre duality, $h^0(\mathcal{E}) = m+1-g$. □

The above is a significant step in the proof of the following result from [**71**]:

Proposition 7.2.10. *The cohomology group $H^k(\mathrm{Sym}^m(\Sigma); \mathbb{Z})$ is torsion free for all $k \in \mathbb{N}$.*

Proof. For $m > 2g-2$, the statement follows from the Leray-Hirsch theorem (see for example [**47**]), and Lemma 7.2.9.

To decrease m, fix $w \in \Sigma$ and consider the inclusion $j\colon \mathrm{Sym}^{m-1}(\Sigma) \to \mathrm{Sym}^m(\Sigma)$ defined by $\mathbf{x} \mapsto \mathbf{x} \cup \{w\}$. Since $\mathrm{Sym}^m(\Sigma) \setminus j(\mathrm{Sym}^{m-1}(\Sigma)) = \mathrm{Sym}^m(\Sigma \setminus \{w\})$ is an affine variety of (real) dimension $2m$, it follows that $\mathrm{Sym}^m(\Sigma)$ is obtained from $\mathrm{Sym}^{m-1}(\Sigma)$ by attaching cells of dimension $\geq m$; see [**79**]. Thus,

$$j^r\colon H^r(\mathrm{Sym}^m(\Sigma); \mathbb{Z}) \to H^r(\mathrm{Sym}^{m-1}(\Sigma); \mathbb{Z})$$

is an isomorphism for $0 \leq r < m-1$, and it is injective for $r = m-1$, with cokernel a free abelian group. It follows from descending induction on m that $H^r(\mathrm{Sym}^m(\Sigma); \mathbb{Z})$ is a free Abelian group for all $0 \leq r \leq m$. The claim of the proposition then follows from Poincaré duality. □

Proof. [Proof of Theorem 7.2.2] We claim that the class U^m generates $H^{2m}(\mathrm{Sym}^m(\Sigma); \mathbb{Z})$. Indeed, the class U is Poincaré dual to the submanifold $\{w_1\} \times \mathrm{Sym}^{m-1}(\Sigma)$; choosing distinct points $w_1, \ldots, w_m \in \Sigma$, the class U^m is Poincaré dual to the intersection

$$\big(\{w_1\} \times \mathrm{Sym}^{m-1}(\Sigma)\big) \cap \cdots \cap \big(\{w_m\} \times \mathrm{Sym}^{m-1}(\Sigma)\big),$$

which consists of the single point $\{w_1, \ldots, w_m\}$, transversely cut out in $\mathrm{Sym}^m(\Sigma)$.

Since U^m generates $H^{2m}(\mathrm{Sym}^m(\Sigma);\mathbb{Z})$, the proof of Proposition 7.2.6 (using algebras over $\mathbb{Z}$ in place of the algebra R over $\mathbb{Q}$, and noting that $w_0(r,I,J,K)$ in fact gives a $\mathbb{Z}$-basis) shows that the ring $H^*(\mathrm{Sym}^m(\Sigma);\mathbb{Z})/\mathrm{Tors}$ is generated by the μ-classes and U, subject to the relations from Equation (7.3). Proposition 7.2.10 now completes the proof. □

Remark 7.2.11. *A classical theorem of Dold and Thom* [**47**, p. 475] *states that the infinite symmetric product of a topological space X (which can be thought of as a limit of the finite symmetric products of X) has the weak homotopy type of a product of Eilenberg-MacLane spaces; in fact,*

$$\mathrm{Sym}^\infty(X) \sim \prod_{i=0}^{\infty} K(H_i(X), i).$$

It follows that $\mathrm{Sym}^\infty(\Sigma)$ *has the weak homotopy type of the product of the torus of dimension $b_1(\Sigma)$ with $\mathbb{CP}^\infty$; so its cohomology ring is an exterior algebra on $2g$ 1-dimensional generators and (a polynomial algebra on) a 2-dimensional generator. This is the ring R appearing in the proof of Lemma 7.2.5.*

7.3. The cohomology class of the diagonal

In this section, we identify the (integral) cohomology class represented by the diagonal $\Delta \subset \mathrm{Sym}^m(\Sigma)$. A little work is needed to make sense of this object, since Δ is not a submanifold. We pause therefore for some generalities concerning cohomology classes represented by analytic subsets.

Let X be a compact, complex manifold, and let $Y \subset X$ be a complex codimension k analytic subset; i.e. a subset that is locally described as the zero set of some collection of holomorphic functions. If Y were smooth, we could apply the Thom isomorphism to its normal bundle to get $H^{2k}(X, X \setminus Y;\mathbb{Z}) \cong H^0(Y;\mathbb{Z})$, and take the image of the generator under the inclusion $H^{2k}(X, X \setminus Y;\mathbb{Z}) \to H^{2k}(X;\mathbb{Z})$ to construct the Poincaré dual of Y.

Even when Y is not smooth, one can generalize the above construction to obtain a $2k$-dimensional cohomology class dual to Y. We sketch the construction and refer to [**143**, Chapter 11] for more details. The key point is the following:

Theorem 7.3.1. *If $Y \subset X$ is a complex codimension k subvariety, we can filter Y by closed analytic subsets $\emptyset = Y_k \subset Y_{k-1} \subset \cdots \subset Y_0 = Y$, so that $Y_j \setminus Y_{j+1}$ is a closed complex submanifold in $X \setminus Y_{j+1}$ of (complex) dimension n_j with $n_j > n_{j+1}$.* □

Use the exact sequence

$$H^{2k}(X\backslash Y_{j+1}, X\backslash Y_j) \to H^{2k}(X\backslash Y_{j+1}) \to H^{2k}(X\backslash Y_j) \to H^{2k+1}(X\backslash Y_{j+1}, X\backslash Y_j)$$

and the Thom isomorphism theorem, which gives $H^r(X \setminus Y_{j+1}, X \setminus Y_j; \mathbb{Z}) = H^{r-2(n-n_j)}(Y_j; \mathbb{Z}) = 0$ for $r = 2k, 2k+1$ to prove by induction on j that the restriction map induces an isomorphism $H^{2k}(X \setminus Y_j; \mathbb{Z}) \cong H^{2k}(X \setminus Y_{j+1}; \mathbb{Z})$. In particular, $H^{2k}(X; \mathbb{Z}) \to H^{2k}(X \setminus Y_1; \mathbb{Z})$ is an isomorphism.

The Thom class gives an element in $H^{2k}(X \setminus Y_1, X \setminus Y_0; \mathbb{Z})$, whose image in $H^{2k}(X \setminus Y_1; \mathbb{Z}) \cong H^{2k}(X; \mathbb{Z})$ is what we call the *cohomology class associated to the subvariety* Y, and will denote by $[\![Y]\!]$.

Let $Y \subset X$ have complex codimension 1, and let S be an oriented two-manifold, with fundamental class $[S] \in H_2(S; \mathbb{Z})$. A smooth map $f\colon S \to X$ is said to be *transverse* to Y if $f(S) \cap Y_1 = \emptyset$, and $f(S)$ intersects the submanifold $Y \setminus Y_1$ transversely.

Lemma 7.3.2. *If $f\colon S \to X$ is transverse to Y, the evaluation $\langle [\![Y]\!], f_*[S]\rangle$ is the algebraic intersection number of f with Y. More generally, if $f(S) \cap Y_1 = \emptyset$, and $f^{-1}(Y)$ consists of finitely many points $x_1, \ldots, x_m$, then around each $x_i \in S$ we can consider the map $D_i \to D$ sending a small neighborhood D_i around x_i, mapped via f to X, followed by the projection to a normal disk to $Y \setminus Y_1$ at $f(x_i)$. The sum of the local degrees of these maps at each x_i is the evaluation $\langle [\![Y]\!], f_*[S]\rangle$.*

Proof. This follows from standard intersection theory; see [**44**]. □

We will apply the above constructions to the diagonal $\Delta \subset \mathrm{Sym}^m(\Sigma)$, to obtain a two-dimensional cohomology class $[\Delta] \in H^2(\mathrm{Sym}^m(\Sigma); \mathbb{Z})$. In that case, the filtration of Theorem 7.3.1 has the following shape:

$$\Delta_i = \{\{x_1, \ldots, x_m\} \in \mathrm{Sym}^m(\Sigma) \Big| |\{x_1, \ldots, x_m\}| \leq m - i - 1\},$$

in particular $\Delta_0 \setminus \Delta_1$ consists of points in Δ where exactly two of the points collide.

Exercise 7.3.3. **(a)** *Show that $\Delta_{m-1} = \emptyset$ and $\Delta_{m-2} = \Sigma$.*

(b) *Identify a component of $\Delta_i \setminus \Delta_{i+1}$ with $\mathrm{Sym}^{m-i-1}(\Sigma)$.*

(c) *Determine the number of components of $\Delta_i \setminus \Delta_{i+1}$.*

Remark 7.3.4. *Given an embedded, oriented curve $\gamma \subset \Sigma$, the cohomology class $\mu(\gamma)$ can be interpreted also as the Poincaré dual of a stratified space, as follows. There is an associated subspace $\gamma \times \mathrm{Sym}^{m-1}(\Sigma)$, consisting of m-tuples $\{x_1, \ldots, x_m\} \subset \Sigma$ so that at least one $x_i \in \gamma$. Although $B =$*

$\gamma \times \mathrm{Sym}^{m-1}(\Sigma)$ *is not a manifold, it is a filtered space, with filtration given by*

$$B_i = \{\{a, \dots, a, x_{i+2}, \dots, x_m\} \mid a \in \gamma\}$$

for $i = 0, \dots, m$. *In particular,* $B_m = \emptyset$ *and* $B_{m-1} = \{\{a, \dots, a\} \mid a \in \gamma\}$. $B_0 \setminus B_1$ *is the image of an immersion of a manifold of dimension* $2m-1$, *with trivialized normal bundle. The subset* B_1 *is of codimension 2 in it, so one can make sense of the algebraic intersection number of this cycle* B *with any one-dimensional homology class. Now,* $\mu(\gamma)$ *is the cocycle whose value on a one-dimensional cycle* c *in* $\mathrm{Sym}^m(\Sigma)$ *is given by the intersection number of* c *with* $\gamma \times \mathrm{Sym}^{m-1}(\Sigma)$.

Exercise 7.3.5. (a) *Identify the double points in* $B_0 \setminus B_1$.

(b) *Given* $\gamma \in H_1(\Sigma; \mathbb{Z})$, *its Poincaré dual is represented by some map* $u\colon \Sigma \to S^1$, *so that the preimage of a regular value in* S^1 *is* γ. *Find the associated map* $\mathrm{Sym}^m(\Sigma) \to S^1$ *that represents* $\mu(\gamma)$.

With the notation of Theorem 7.2.2 we have

Proposition 7.3.6. *The two-dimensional cohomology class* $[\![\Delta]\!] \in H^2(\mathrm{Sym}^m(\Sigma); \mathbb{Z})$ *associated to* Δ *can be written as*

$$[\![\Delta]\!] = 2(m-1)U + 2\sum_{i=1}^{g}(U - \mu(A_i)\mu(B_i)).$$

Proof. Consider the embedded surfaces in $\mathrm{Sym}^m(\Sigma)$ of the form $\Sigma \times \{w_1, \dots, w_{m-1}\}$, $A_i \times B_i \times \{w_1, \dots, w_{m-2}\}$, $A_i \times B_j \times \{w_1, \dots, w_{m-2}\}$ $A_i \times A_j \times \{w_1, \dots, w_{m-2}\}$ $B_i \times B_j \times \{w_1, \dots, w_{m-2}\}$ where $i \neq j$, and $\{w_1, \dots, w_{m-2}, w_{m-1}\}$ are distinct fixed points. The verification that these are indeed embedded surfaces is a straightforward local computation near the intersection points with the diagonal; the computations of the intersection numbers will be given in detail below. Before doing so, notice that it is straightforward to verify that the Poincaré duals to those classes are, respectively, U^{m-1}, $(U-\mu(A_i)\mu(B_i))U^{m-2}$, $\mu(A_i)\mu(B_j)U^{m-2}$, $\mu(A_i)\mu(A_j)U^{m-2}$, and $\mu(B_i)\mu(B_j)U^{m-2}$; and in particular these homology classes generate $H_2(\mathrm{Sym}^m(\Sigma); \mathbb{Z})$. Therefore, to determine $[\![\Delta]\!]$, it will suffice to compute its intersection number with the above homology generators.

To compute these intersection numbers, note first that Δ is disjoint from $A_i \times B_j$, $A_i \times A_j$, $B_i \times B_j$ for $i \neq j$. Next, $\Sigma \times \{w_1, \dots, w_{m-1}\}$ meets Δ in $m-1$ points, where the point on Σ agrees with one of $w_1, \dots, w_{m-1}$; i.e. it is contained in $\Delta \setminus \Delta_1$. We claim that the intersection is not transverse, and its local multiplicity is 2. This can be seen by performing a computation in a local model around the point $\mathbf{w} = \{w_1, w_1, w_2, \dots, w_{m-1}\}$. Indeed, since

all w_i $(i = 1, \dots, m-1)$ are chosen to be distinct, we can do this local computation in the case when $m = 2$; the general case follows from this special case at once.

Fix a neighborhood D of $w_1 \in \Sigma$ with an identification of D with $\mathbb{C}$ (in such a way that w_1 maps to 0). In the chart $\mathrm{Sym}^2(D) \subset \mathrm{Sym}^2(\Sigma)$ (when $\mathrm{Sym}^2(D)$ is identified with $\mathrm{Sym}^2(\mathbb{C})$ and this latter with the space of monic, degree 2 polynomials) the surface $\Sigma \times \{w_1\}$ corresponds to the set $\{t(t-z) = t^2 - zt \mid z \in \mathbb{C}\}$, while Δ corresponds to those polynomials $t^2 + bt + c$ with $b^2 - 4c = 0$, i.e. to $\{t^2 + bt + \frac{b^2}{4} \mid b \in \mathbb{C}\}$. These two curves intersect at $(0, 0)$, and the multiplicity of the intersection is 2, which can be easily checked by displacing one of the curves with an additive constant.

We claim that $A_j \times B_j$ intersects Δ in 1 point. This is obvious: A_j and B_j intersect in a single point. We claim that the intersection is non-transverse, and that the local multiplicity is -2. This is done by computing in a neighborhood where $A_j \times B_j$ is parametrized by $(x, y) \in \mathbb{R}^2$. Explicitly, we have that $A_j \times B_j$ corresponds to $\{x, iy\}$ which in turn corresponds to $t^2 - (x+iy)t + ixy$ (and again our subset is a submanifold near its intersection with the diagonal), a polynomial whose discriminant is given by $(x - iy)^2$; i.e. the intersection is locally modeled on the zero of the map $z \mapsto \overline{z}^2$; which is a single zero with local multiplicity -2.

The form of $[\![\Delta]\!] \in H^2(\mathrm{Sym}^m(\Sigma); \mathbb{Z})$ now follows from the above evaluations and Lemma 7.3.2. □

7.4. The first Chern class

We compute the first Chern class of $\mathrm{Sym}^m(\Sigma)$ by relating it to $[\![\Delta]\!]$. Before doing this, we recall some standard terminology from algebraic geometry. A holomorphic map $\phi\colon X \to Y$ between compact, complex manifolds of dimension m is called a *branched cover* if the set $R \subset X$ of critical points of ϕ is a proper subset of X. In this case, the set of critical values $\Delta \subset Y$ of ϕ forms an analytic subset called the *branch locus* of ϕ, and R is called the *ramification locus* of ϕ. In general both R and Δ are subvarieties, hence we can apply the filtration result of Theorem 7.3.1 in studying them.

Example 7.4.1. *The map $\pi\colon \times^m(\Sigma) \to \mathrm{Sym}^m(\Sigma)$ is an $m!$-fold branched cover, with the diagonal in $\times^m(\Sigma)$ (the critical points of π) as ramification locus and $\Delta \subset \mathrm{Sym}^m(\Sigma)$ as branch locus.*

Lemma 7.4.2. *Suppose that S is a compact, oriented two-manifold and $f\colon S \to \mathrm{Sym}^m(\Sigma)$ is transverse to Δ. Then, there is a diagram*

$$\begin{array}{ccc} \widetilde{S} & \xrightarrow{\widetilde{f}} & \times^m(\Sigma) \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle \pi} \\ S & \xrightarrow{f} & \mathrm{Sym}^m(\Sigma) \end{array} \tag{7.10}$$

where $\widetilde{S}$ is a compact, oriented two-manifold equipped with a smooth action by the symmetric group $\mathfrak{S}_m$, $\widetilde{f}$ is $\mathfrak{S}_m$-equivariant, and p is a branched covering map with degree $m!$. Moreover,

$$\langle c_1(T\mathrm{Sym}^m(\Sigma)), [S]\rangle = \frac{1}{m!}\langle c_1(T(\times^m(\Sigma))), [\widetilde{S}]\rangle + \frac{\#(\Delta \cap S)}{2}.$$

Proof. Construct $\widetilde{S}$ by pull-back. We compute the first Chern class as follows. Let K_X denote the canonical line bundle over $X = \times^m(\Sigma)$ or $\mathrm{Sym}^m(\Sigma)$, i.e. the m-fold wedge product of the complex cotangent bundle $\Omega^{n,0}$; in particular $c_1(K_X) = -c_1(TX)$. Pulling back forms induces a map of complex line bundles, which we denote here $D\pi^*$:

$$D\pi^*\colon \pi^*(K_{\mathrm{Sym}^m(\Sigma)}) \to K_{\times^m(\Sigma)},$$

which is an isomorphism of line bundles away from the diagonal Δ. Thus, if we think of $D\pi^*$ as a section of the line bundle

$$\mathrm{Hom}(\pi^*(K_{\mathrm{Sym}^m(\Sigma)}), K_{\times^m(\Sigma)}) \cong \left(\pi^*(K_{\mathrm{Sym}^m(\Sigma)})\right)^* \otimes K_{\times^m(\Sigma)},$$

then $D\pi^*$ is non-vanishing away from Δ. Consider a point $\mathbf{x} \in \times^m(\Sigma)$ on the top stratum of the ramification locus R. This point has a coordinate neighborhood U, parametrized by $(z_1, \ldots, z_m)$ so that $\mathbf{x}$ corresponds to $z_1 = \cdots = z_m = 0$; moreover, $\pi(\mathbf{x})$ has a coordinate system $(w_1, \ldots, w_m)$ so that π is modelled on the map

$$q\colon (z_1, z_2, \ldots, z_m) \mapsto (z_1^2, z_2, \ldots, z_m) = (w_1, \ldots, w_m);$$

i.e. $w_1 = 0$ corresponds to the branch locus Δ (around $\pi(\mathbf{x})$), and $z_1 = 0$ corresponds to the ramification locus R (around $\mathbf{x}$) of π. Clearly,

$$Dq^*(dw_1 \wedge \cdots \wedge dw_m) = 2z_1 dz_1 \wedge \cdots \wedge dz_m;$$

i.e. the section $D\pi^*$ vanishes with multiplicity 1 along the branch locus Δ.

Pulling back the section $D\pi^*$ using $\widetilde{f}^*$ we obtain a section of the bundle $\left(\pi^*(K_{\mathrm{Sym}^m(\Sigma)})\right)^* \otimes K_{\times^m(\Sigma)}$, restricted to $\widetilde{S}$, with transverse zeros along $\widetilde{f}^{-1}(R)$. It follows at once that

$$-\langle c_1(\pi^*(K_{\mathrm{Sym}^m(\Sigma)})), [\widetilde{S}]\rangle + \langle c_1(K_{\times^m(\Sigma)}), [\widetilde{S}]\rangle = \#(R \cap \widetilde{S}).$$

Since p has degree $m!$, we have

$$\langle c_1(\pi^*(K_{\mathrm{Sym}^m(\Sigma)})), [\widetilde{S}]\rangle = -m! \cdot \langle T\mathrm{Sym}^m(\Sigma), [S]\rangle$$

(notice the sign caused by switching from cotangent to tangent bundle), and

$$\#(R \cap \widetilde{S}) = \frac{m!}{2}\#(\Delta \cap S),$$

since R intersects $\widetilde{f}$ with half the multiplicity that Δ intersects $\pi \circ \widetilde{f}$. □

Remark 7.4.3. *The relationship between the Chern class of a ramified cover and the ramification divisor described above can be found for example in* [**46**].

Proposition 7.4.4. *The first Chern class of* $T\mathrm{Sym}^m(\Sigma)$ *is given by*

$$(1+m-2g)U + \sum_{i=1}^{g}(U - \mu(A_i)\mu(B_i)) = (1+m-g)U - \sum_{i=1}^{g}\mu(A_i)\mu(B_i).$$

Proof. Consider the embedded surfaces generating $H_2(\mathrm{Sym}^m(\Sigma);\mathbb{Z})$ as in the proof of Proposition 7.3.6. As shown in the proof of that proposition, some of these surfaces do not meet the diagonal transversely; nonetheless, one can construct the $m!$-sheeted branched covers of these as in Diagram (7.10). For example, when $m = 2$, the surface $S = T^2$ embedded as $A_1 \times B_1$ in $\mathrm{Sym}^m(\Sigma)$ is double covered by the nodal surface $\widetilde{S} = T^2 \vee T^2$. Thus, Lemma 7.4.2 holds for these homology classes.

The first Chern class of $\times^m(\Sigma)$ is $2-2g$ times the cohomology class

$$\sum_{i=1}^{m} \overbrace{1 \otimes \cdots \otimes 1}^{i-1} \otimes \nu \otimes \overbrace{1 \otimes \cdots \otimes 1}^{m-i},$$

where ν is a volume form on Σ. Evidently, the lifts of the embedded surfaces $A_i \times B_i \times \{w_1, \ldots, w_{m-2}\}$, $A_i \times B_j \times \{w_1, \ldots, w_{m-2}\}$, $A_i \times A_j \times \{w_1, \ldots, w_{m-2}\}$, $B_i \times B_j \times \{w_1, \ldots, w_{m-2}\}$ where $i \neq j$ and $\{w_1, \ldots, w_{m-2}\}$ are distinct, are all collections of Lagrangian tori: in particular, $c_1(T(\times^m(\Sigma)))$ vanishes on them. The lift of $\Sigma \times \{w_1, \ldots, w_{m-1}\}$ is a union of $m!$ copies of Σ, and the first Chern class of $\times^m(\Sigma)$ evaluates $2-2g$ on each of those.

The form of the first Chern class now follows. □

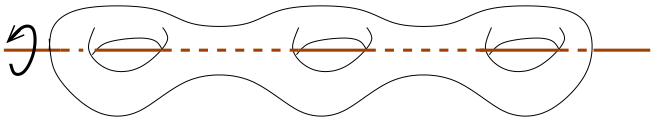


Figure 7.1. Rotate through the axis by 180°.

7.5. Spherical homology classes

When understanding Gromov compactifications of spaces of curves in a complex manifold, it is important to understand how the first Chern class of the manifold evaluates on spherical 2-dimensional homology classes. Having computed $c_1(T\mathrm{Sym}^m(\Sigma))$, it remains to understand the homology classes in $\mathrm{Sym}^m(\Sigma)$ that are represented by spheres.

To this end, it will be helpful to note that any compact, connected, oriented two-manifold Σ admits a branched double-cover map $f\colon \Sigma \to S^2$; equivalently, any such two-manifold Σ is equipped with an involution $\tau\colon \Sigma \to \Sigma$ acting with finitely many fixed points, so that the orbit space Σ/τ is identified with S^2. This follows from a symmetric drawing of Σ, as shown in Figure **??**. There is then a naturally induced continuous map $F\colon S^2 \to \mathrm{Sym}^2(\Sigma)$, which associates to each $z \in S^2$, the two points in $f^{-1}(\{z\})$, thought of as a point in $\mathrm{Sym}^2(\Sigma)$. Here, if p is a branch point, we think of $f^{-1}(\{z\}) \subset \Sigma$ as a single point with multiplicity 2; i.e. it still gives a well-defined point in $\mathrm{Sym}^2(\Sigma)$ (indeed, in its diagonal).

Remark 7.5.1. *If Σ is thought of as equipped with a complex structure, then such a covering does not necessarily exist holomorphically: when it does, the involution τ is called a* ***hyperelliptic involution****, and Σ, as a complex curve is called a* ***hyperelliptic curve****. In explicit terms, consider the algebraic curve specified by*

$$y^2 = (x - a_1)\cdots(x - a_{2g+1}),$$

where $(x, y) \in \mathbb{C}^2$ and $a_1, \ldots, a_{2g+1}$ are distinct in $\mathbb{C}$. It is easy to check that this is a compact two-manifold of genus g punctured in one point, and the hyperelliptic involution is given by $(x, y) \mapsto (x, -y)$. The compact version is realized by projectivizing the curve.

We formalize the above construction of the two-sphere in $\mathrm{Sym}^2(\Sigma)$, as follows:

Proposition 7.5.2. *Let F, S, and Σ be surfaces, $f\colon F \to \Sigma$ be a map and $\phi\colon F \to S$ be a branched m-fold covering. Then there is an induced map $\Phi\colon S \to \mathrm{Sym}^m(\Sigma)$ defined as follows: given $s \in S$, if $\{x_1, \ldots, x_m\} = \phi^{-1}(s)$ (thought of as a multi-set, i.e. a set with possibly repeated entries), then*

$\Phi(s) = \{f(x_1), \ldots, f(x_m)\}$. *Moreover, if* F, S, *and* Σ *have complex structures, and* f *and* ϕ *are holomorphic, then the induced map* Φ *is holomorphic.*

Proof. Continuity is straightforward. For the holomorphic version, Φ is clearly holomorphic away from the branch points; as it is continuous, its extension across the finitely many branch points is holomorphic, as well. □

Remark 7.5.3. *In Lemma 7.4.2 we associated a branched* $m!$*-fold cover* $\widetilde{S}$ *of the sphere* S, *associated to the map* f *from* S *to* $\mathrm{Sym}^m(\Sigma)$. *This space* $\widetilde{S}$ *is equipped with an action by* $\mathfrak{S}_m$. *In Proposition 7.5.2, we constructed a sphere in* $\mathrm{Sym}^m(\Sigma)$ *from a branced* m*-fold cover* F. *These two covers are related by* $F = \widetilde{S}/\mathfrak{S}_{m-1}$*; i.e. we have the following commutative diagram:*

$$\begin{array}{ccc} \widetilde{S} & \xrightarrow{\widetilde{f}} & \times^m(\Sigma) \\ \downarrow & & \downarrow{\scriptstyle p_1} \\ F = \widetilde{S}/\mathfrak{S}_{m-1} & \longrightarrow & \Sigma \\ \downarrow{\scriptstyle p} & & \\ S & & \end{array}$$

where $p_1 \colon \times^m(\Sigma) \to \Sigma$ *is projection onto the first factor, and* $\mathfrak{S}_{m-1} \subset \mathfrak{S}_m$ *is the subgroup that fixes the first letter.*

The sphere in $\mathrm{Sym}^2(\Sigma)$ associated to a hyperelliptic involution on Σ is constructed using Proposition 7.5.2 by choosing $F = \Sigma$, f the identity map, and ϕ the branched double-cover $\Sigma \to S^2$. To obtain an induced sphere in $\mathrm{Sym}^m(\Sigma)$ for arbitrary $m \geq 2$, we fix $w_1, \ldots, w_{m-2} \in \Sigma$ and use the induced inclusion $\mathrm{Sym}^2(\Sigma) \to \mathrm{Sym}^m(\Sigma)$ given by

$$\{x, y\} \to \{x, y, w_1, \ldots, w_{m-2}\}.$$

To see that the sphere associated to a hyperelliptic involution is homologically non-trivial, we argue as follows. Fix $w \in \Sigma$ and consider the corresponding submanifold $\{w\} \times \mathrm{Sym}^{m-1}(\Sigma) \subset \mathrm{Sym}^m(\Sigma)$. Given any map $\varphi \colon S^2 \to \Sigma$, let $n_w(\varphi)$ denote the algebraic intersection number of $\varphi(S^2)$ with $\{w\} \times \mathrm{Sym}^{m-1}(\Sigma)$. In the notation of Theorem 7.2.2, this quantity is the evaluation of $\varphi^*(U)$ on the fundamental cycle of S^2. In particular, if $n_w(\varphi) \neq 0$, then $\varphi(S^2)$ is homologically non-trivial. For the map Φ associated to the involution f above, we note that $n_w(\Phi) = 1$. This is most easily seen if w is not a branch point of f. There is a unique point $z \in \Sigma$

so that $w \in p^{-1}(z)$. When w is not a branch point, it is straightforward to check that Φ is transverse to the submanifold $\{w\} \times \Sigma$, and that the local intersection number is $+1$.

We turn to the more systematic study of homology classes represented by spheres in $\mathrm{Sym}^m(\Sigma)$. Recall that for any topological space X and $x_0 \in X$, there is a natural homomorphism, the *Hurewicz homomorphism*

$$h_n \colon \pi_n(X, x_0) \to H_n(X; \mathbb{Z}),$$

which carries a sphere $f \colon S^n \longrightarrow X$ to f_* of the fundamental cycle $[S^n] \in H_n(S^n; \mathbb{Z}) \cong \mathbb{Z}$, , cf. [**47**]. A homology class in the image of this homomorphism is called *spherical*. The spherical classes of $H_2(\mathrm{Sym}^m(\Sigma); \mathbb{Z})$ are identified in the following:

Proposition 7.5.4. *For $m > 1$, the image of the Hurewicz homomorphism h_2 (i.e., the image of $\pi_2(\mathrm{Sym}^m(\Sigma), \mathbf{x}_0)$ in $H_2(\mathrm{Sym}^m(\Sigma); \mathbb{Z})$) consists of integral multiples of the Poincaré duals of the cohomology class*

$$(7.11)\quad (U - \sum_{i=1}^{g} (U - \mu(A_i)\mu(B_i)))U^{m-2} = ((1-g)U + \sum_{i=1}^{g} \mu(A_i)\mu(B_i))U^{m-2}.$$

Moreover, the map n_w induces an isomorphism from the image of $\pi_2(\mathrm{Sym}^m(\Sigma), \mathbf{x}_0)$ in $H_2(\mathrm{Sym}^m(\Sigma); \mathbb{Z})$ to $\mathbb{Z}$, taking the above generator to $1 \in \mathbb{Z}$.

Proof. By Theorem 7.2.2, any cohomology class in $H^{2m-2}(\mathrm{Sym}^m(\Sigma); \mathbb{Z})$ is a sum of elements of the form U^{m-1}, $(U - \mu(A_i)\mu(B_i))U^{m-2}$, $\mu(A_i)\mu(A_j)U^{m-2}$, $\mu(B_i)\mu(B_j)U^{m-2}$, and $\mu(A_i)\mu(B_j)U^{m-2}$ with $i \neq j$. Clearly, any spherical class $[S]$ has the property that

$$\langle \mu(A_i)\mu(A_j), [S] \rangle = \langle \mu(B_i)\mu(B_j), [S] \rangle = \langle \mu(A_i)\mu(B_j), [S] \rangle = 0,$$

since the pull-backs of $\mu(A_i)$ and $\mu(B_j)$ to S^2 vanish. The above equations for $i \neq j$ imply that $\mathrm{PD}[S]$ is in the span of U^{m-1} and $(U - \mu(A_i)\mu(B_i))U^{m-2}$ for $i = 1, \ldots, g$. Observe that if $i \neq j$,

$$\begin{aligned} &\mu(A_i) \cdot \mu(B_i) \cdot (U - \mu(A_j) \cdot \mu(B_j))U^{m-2} \\ &\qquad = (U - (U - \mu(A_i) \cdot \mu(B_i)) \cdot (U - \mu(A_j) \cdot \mu(B_j))U^{m-2} \ = 0 \end{aligned}$$

while

$$\begin{aligned} &\mu(A_i) \cdot \mu(B_i) \cdot (U - \mu(A_i) \cdot \mu(B_i))U^{m-2} \\ &\qquad = \mu(A_i) \cdot \mu(B_i)U^{m-1} = (U - (U - \mu(A_i) \cdot \mu(B_i)))U^{m-1} = U^m. \end{aligned}$$

It now follows that the Poincaré dual of a spherical class is a multiple of the class from Equation (7.11).

It remains to show that there is a map $\Phi \colon S^2 \to \mathrm{Sym}^m(\Sigma)$ with $n_w(\Phi) = 1$. When $m = 2$, such a map Φ was constructed in the discussion preceding

the proposition. Clearly, the map $\phi\colon \mathrm{Sym}^m(\Sigma) \to \mathrm{Sym}^{m+1}(\Sigma)$ induced by $\mathbf{x} \to \{w\} \cup \mathbf{x}$ has the property that $n_w(\phi \circ \Phi) = n_w(\Phi)$. Thus, we can promote the sphere in $\mathrm{Sym}^2(\Sigma)$ with $n_w = 1$ to spheres with $n_w = 1$ in arbitrarily large symmetric products. □

For our applications, we will need the following:

Proposition 7.5.5. *For any $m > 1$, and sphere $S \subset \mathrm{Sym}^m(\Sigma)$ we have*

$$\langle c_1(T\mathrm{Sym}^m(\Sigma)), [S]\rangle = (1 + m - g)n_w([S]). \tag{7.12}$$

Proof. By Proposition 7.5.4, it suffices to verify the formula for the choice

$$\mathrm{PD}[S] = ((1-g)U + \sum_{i=1}^{g} \mu(A_i)\mu(B_i))U^{m-2},$$

as in Equation (7.11). The result now follows from Proposition 7.4.4, and the ring structure on $H^*(\mathrm{Sym}^m(\Sigma); \mathbb{Z})$ given in Theorem 7.2.2. □

Remark 7.5.6. *The case of primary interest to us will be when $m = g$, reducing Equation* (7.12) *to $\langle c_1(T\mathrm{Sym}^m(\Sigma)), [S]\rangle = n_w([S])$.*

Proposition 7.5.7. *Given any $m > 1$, if S is any sphere in $\mathrm{Sym}^m(\Sigma)$, then*

$$\langle [\![\Delta]\!], [S]\rangle = (2m + 2g - 2) \cdot n_w([S]).$$

Proof. This is a straightforward computation from Propositions 7.3.6 and 7.5.4. □

We conclude this section with a consistency check for our computations thus far. First, we formalize the mechanism that constructed the spheres in the symmetric product. This gives some insight into how to construct maps into the symmetric product. Versions of this result will be used later.

Suppose that Σ_g is a hyperelliptic surface. In this case the hyperelliptic involution induces a holomorphic map from S^2 to $\mathrm{Sym}^2(\Sigma)$, whose image S has $n_w(S) = 1$. By an elementary Euler characteristic count (the Riemann-Hurwitz formula), the branched double-cover has $2g + 2$ double points, so we expect S to intersect Δ in $2g + 2$ points. This is consistent with Proposition 7.5.7 above.

Exercise 7.5.8. *Suppose that Σ_g is a hyperelliptic complex curve.*

(a) *Compute the self-intersection number of the holomorphic sphere S in $\mathrm{Sym}^2(\Sigma_g)$ induced by the hyperelliptic involution.*

(b) *Use the adjunction formula*

$$\langle c_1(T\mathrm{Sym}^2(\Sigma)), [S]\rangle = 2 + [S]\cdot[S]$$

(which holds since S is holomorphic) to give an alternate verification of Proposition 7.5.5 when $m = 2$.

In the above exercise, we chose a particular complex structure on $\mathrm{Sym}^m(\Sigma)$ (associated to a hyperelliptic complex structure on Σ). Proposition 7.1.9 shows that the answer is independent of this choice.

7.6. Symplectic structures on the symmetric product

Let Σ be equipped with a complex structure j, and let $\times^m(j)$ denote the naturally induced $\mathfrak{S}_m$-invariant complex structure on the Cartesian product $\times^m(\Sigma)$. Proposition 7.1.4 supplies a complex structure $\mathrm{Sym}^m(j)$ on $\mathrm{Sym}^m(\Sigma)$, with the property that the projection map $\pi\colon \times^m(\Sigma) \to \mathrm{Sym}^m(\Sigma)$ is holomorphic. In the definition of Heegaard Floer homology, we will need a symplectic structure on $\mathrm{Sym}^m(\Sigma)$ that is compatible with $\mathrm{Sym}^m(j)$.

Fix a symplectic form ν on Σ compatible with j. There is a naturally induced product symplectic form $\nu^\times$ over $\times^m(\Sigma)$, defined by $\nu^\times = \sum_{i=1}^m p_i^*(\nu)$, where $p_i\colon \times^m(\Sigma) \to \Sigma$ is projection onto the i^{th} factor. Over $\mathrm{Sym}^m(\Sigma)\setminus\Delta$, the $\mathfrak{S}_m$-invariant form $\nu^\times$ determines a symplectic form, which we write $\nu^\times/\mathfrak{S}_m$. Obviously, $\nu^\times$ is not the pull-back of a symplectic form over $\mathrm{Sym}^m(\Sigma)$, since the pull-back of any two-form by π is degenerate at the critical points of π, i.e. at the preimage of the diagonal $\pi^{-1}(\Delta) \subset \times^m(\Sigma)$. On the other hand, by Theorem 7.2.3, since $\nu^\times$ is invariant under the action of $\mathfrak{S}_m$, its underlying cohomology class $[\nu^\times]$ is the pull-back of some cohomology class over $\mathrm{Sym}^m(\Sigma)$. Our aim is to find a symplectic structure on $\mathrm{Sym}^m(\Sigma)$ that represents $[\nu^\times]$ and coincides with $\nu^\times$ on the complement of an open neighbourhood U of the diagonal Δ. (Indeed, the form will depend on the chosen neighbourhood U.)

The desired symplectic form on $\mathrm{Sym}^m(\Sigma)$ is provided by the following theorem of Perutz [**110**], based on complex analytic constructions of Varouchas [**141**]. Perutz's theorem states that the cohomology class $[\nu^\times]$ is the pull-back of a cohomology class which represents a symplectic form over $\mathrm{Sym}^m(\Sigma)$, and it gives control of the symplectic form away from Δ.

Theorem 7.6.1. *(Perutz [**110**]) Let (j, ν) be a Kähler structure on the oriented two-manifold Σ, and let $\nu^\times$ be the induced product Kähler form over $\times^m(\Sigma)$. Given any open set U containing the diagonal $\Delta \subset \mathrm{Sym}^m(\Sigma)$,*

there is a Kähler form ω on $\mathrm{Sym}^m(\Sigma)$, equipped with its induced complex structure $\mathrm{Sym}^m(j)$ from Proposition 7.1.4, so that $\pi^*(\omega) - \nu^\times = d\eta$, where $\eta \in \Omega^1(\times^m(\Sigma))$ is a form with support inside $\pi^{-1}(U)$.

We will use this result in the context of Heegaard Floer theory with the following roles: suppose that $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$ is a Heegaard diagram, and let

$$\mathbb{T}_\alpha = \alpha_1 \times \ldots \times \alpha_g \quad \text{and} \quad \mathbb{T}_\beta = \beta_1 \times \ldots \times \beta_g;$$

be two tori in $\mathrm{Sym}^g(\Sigma)$. For a complex structure j and area form ν on Σ there is a branched cover map $\pi\colon \times^g(\Sigma) \to \mathrm{Sym}^g(\Sigma)$ and a symplectic two-form

$$\nu^\times = \sum_{i=1}^{g} p_i^*(\nu)$$

on $\times^g(\Sigma)$, where $p_i\colon \times^g(\Sigma) \to \Sigma$ is projection to the i^{th} factor. The preimage of $\mathbb{T}_\alpha$ under π is the union of $g!$ tori

$$\bigcup_{\sigma \in \mathfrak{S}_g} \alpha_{\sigma(1)} \times \cdots \times \alpha_{\sigma(g)}$$

(and similarly for the preimage of $\mathbb{T}_\beta$); these preimages $\pi^{-1}(\mathbb{T}_\alpha)$ and $\pi^{-1}(\mathbb{T}_\beta)$ are Lagrangian with respect to $\nu^\times$ in $\times^g(\Sigma)$.

Since $\mathbb{T}_\alpha$ and $\mathbb{T}_\beta$ are disjoint from the diagonal Δ, Theorem 7.6.1 has the following immediate corollary:

Corollary 7.6.2. *(Perutz [**110**]) For a Heegaard diagram $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$ and a complex structure j over Σ there is a Kähler form ω on the complex manifold $(\mathrm{Sym}^g(\Sigma), \mathrm{Sym}^g(j))$, for which $\mathbb{T}_\alpha$ and $\mathbb{T}_\beta$ are Lagrangian. Moreover, for any positive area form ν over Σ, we can choose this Kähler form ω so that the cohomology classes $[\pi^*(\omega)]$ and $[\nu^\times]$ coincide, when thought of as elements in the relative cohomology group $H^2(\times^g(\Sigma), \pi^{-1}(\mathbb{T}_\alpha) \cup \pi^{-1}(\mathbb{T}_\beta); \mathbb{R})$.*

Proof. As the tori $\mathbb{T}_\alpha$ and $\mathbb{T}_\beta$ are disjoint from Δ, we can choose the open neighbourhood U of Δ so that it is disjoint from $\mathbb{T}_\alpha \cup \mathbb{T}_\beta$. For a given Kähler form ν on Σ the symplectic form ω provided by Theorem 7.6.1 will have the desired properties. □

The proof of Theorem 7.6.1 will rest on smoothing techniques from complex analysis. We will not need these methods in the rest of the present work; but we include the proof for completeness, nonetheless.

We will need some definitions. Recall first that a smooth function $\phi\colon \Omega \to \mathbb{R}$ on an open domain $\Omega \subset \mathbb{R}^2$ is *harmonic*, if

$$\Delta\phi = \frac{\partial^2 \phi}{\partial x_1^2} + \frac{\partial^2 \phi}{\partial x_2^2} = 0.$$

A continuous function $\psi\colon \Omega \to \mathbb{R}$ is called *subharmonic* if for any $x \in U$ there is $r > 0$ with $B_r(x) \subset \Omega$ such that for any harmonic function h with $\psi \leq h$ on $\partial B_r(x)$ we have that $\psi \leq h$ on $B_r(x)$. Alternatively, ψ is subharmonic if it satisfies

$$\psi(x) \leq \frac{1}{2\pi} \int_0^{2\pi} \psi(x + re^{it}) dt.$$

It is called *strictly subharmonic* if the above inequality is strict.

The above definitions extend to multivariable functions as follows:

Definition 7.6.3. *Suppose that $\Omega \subset \mathbb{C}^n$ is an open domain. A smooth function $\phi\colon \Omega \to \mathbb{R}$ is **pluriharmonic** if the restriction $\phi|_{\Omega \cap \{a+bz | z \in \mathbb{C}\}}$ $(a, b \in \mathbb{C}^n)$ to any complex line is harmonic. Similarly, a continuous function $\psi\colon \Omega \to \mathbb{R}$ is **plurisubharmonic** resp. **strictly plurisubharmonic** if the restrictions $\psi|_{\Omega \cap \{a+bz | z \in \mathbb{C}\}}$ $(a, b \in \mathbb{C}^n)$ are subharmonic resp. strictly plurisubharmonic.*

Notice that harmonic and pluriharmonic functions are smooth by definition, while subharmonic and plurisubharmonic functions might be only continuous. (Indeed, these notions are also defined for broader classes of functions, but we will not need that generality in the present context.)

Example 7.6.4. *For any real number $c > 0$, the function on $\mathbb{C}^n \to \mathbb{R}$ defined by $x \mapsto c\|x\|^2$ is strictly plurisubharmonic.*

Exercise 7.6.5. *Show that $\psi\colon \Omega \to \mathbb{R}$ is ψ is strictly plurisubharmonic if and only if for every $\mathbf{x} \in \Omega$, there is a neighborhood U with $x \in U \subset \Omega$ and a real number $c > 0$ so that the function $y \mapsto \psi(y) - c\|y\|^2$ is plurisubharmonic.*

Since for a holomorphic function $F\colon \Omega_1 \to \Omega_2$ and a pluriharmonic (plurisubharmonic) function $f\colon \Omega_2 \to \mathbb{R}$ we have that $f \circ F$ is also pluriharmonic (plurisubharmonic, resp.), the above notions extend to functions on complex manifolds.

Suppose that (X, J) is a complex manifold. For a smooth function $f\colon X \to \mathbb{R}$ define the 1-form $d^{\mathbb{C}} f$ as $df \circ J$. (We are using here a Let g_f denote the symmetric bilinear form defined as

$$g_f(v, w) = -dd^{\mathbb{C}} f(v, Jw).$$

(Note that since X is complex $-dd^{\mathbb{C}}$ coincides with the two-form $\frac{i}{2}\partial\overline{\partial}$ which appeared in Example 4.1.7.) Using this notion, it can be shown that the smooth function $f\colon X \to \mathbb{R}$ is pluriharmonic if $dd^{\mathbb{C}} f = 0$, is plurisubharmonic if g_f is positive semi-definite and f is strictly plurisubharmonic if g_f is positive definite; in this latter case $-dd^{\mathbb{C}} f$ is a symplectic (in fact, Kähler) form and g_f is a compatible Riemannian (and indeed Kähler) metric.

Remark 7.6.1. *Our definition of $d^{\mathbb{C}}$ follows conventions from symplectic geometry; see for example* [**13**]. *The definition of $d^{\mathbb{C}}$ from* [**42**], *which we denote $d^{\mathbb{C}}_{GH}$, is related to the above definition of $d^{\mathbb{C}}$ by $d^{\mathbb{C}}_{GH} = -\frac{1}{4\pi} d^{\mathbb{C}}$.*

Definition 7.6.6. *For a complex manifold (X, J) the set $\{(U_i, \varphi_i)\}_{i\in I}$ is a **Kähler cocycle** if $\{U_i\}_{i\in I}$ is an open cover of X, and $\varphi_i \colon U_i \to \mathbb{R}$ are continuous functions such that*

- *φ_i is strictly plurisubharmonic on U_i for all $i \in I$, and*
- *the difference $\varphi_i - \varphi_j$ is pluriharmonic on $U_i \cap U_j$ for all $i, j \in I$.*

The Kähler cocyle is called *smooth* if all φ_i are smooth functions. In this case the second property ensures that the cocycle $-dd^{\mathbb{C}}\varphi_i$ patches together to a global 2-form on X, which (by the first property) is a J-compatible symplectic, and therefore a Kähler form.

Conversely, if ω is a Kähler form on the complex manifold (X, J), then by taking an open cover $\{U_i\}$ of X with all U_i biholomorphic to a ball, the solutions of the equations $\omega|_{U_i} = -dd^{\mathbb{C}}\varphi_i$ provide a smooth Kähler cocycle. Indeed, ω is a closed form (since it is symplectic) and $d^{\mathbb{C}}\omega = 0$ also holds (as $d^{\mathbb{C}} = J^{-1} \circ d \circ J$ and ω is J-invariant). According to the $dd^{\mathbb{C}}$-lemma [**57**, Lemma 3.A.22], since $\omega|_{U_i}$ is also exact (by our topological assumption on U_i) we can solve the equation $\omega|_{U_i} = -dd^{\mathbb{C}}\varphi_i$ for φ_i. Note that in the above definition φ_i might be only continuous, hence the Kähler cocycle is a more general notion than a Kähler form. This feature will be important in our later discussion.

Suppose that $\pi \colon X \to X'$ is a branched covering of complex manifolds of degree d and ω is a Kähler form on X. Represent ω by a smooth Kähler cocycle $\{(U_i, \varphi_i)\}_{i\in I}$ (so that components of the possibly disconnected U_i are all biholomorphic to balls) that has the additional property that X' admits a locally finite cover $\{U'_i\}_{i\in I}$ such that $\pi^{-1}(U'_i) \subset U_i$ for all $i \in I$. The map π can be used to push the smooth Kähler cocycle forward: define $\varphi'_i = \pi_*\varphi_i$ at the point $x' \in U'_i$ by

$$\varphi'_i(x') = \frac{1}{d} \sum_{x\in\pi^{-1}(x')} \varphi_i(x),$$

counting the branch points counted with multiplicity. The resulting functions φ'_i are obviously continuous, but not necessarily smooth.

Example 7.6.7. *For an example when the push-forward of a smooth function is not smooth, consider the double branched cover map $\pi \colon \mathbb{D} \to \mathbb{D}$ given by the formula $z \mapsto z^2$ on the unit disk $\mathbb{D} \subset \mathbb{C}$. Let $f \colon \mathbb{D} \to \mathbb{R}$ be given by $f(z) = (\mathrm{Re}(z))^2$. The map $g = \pi_* f$, defined by $g(z') = \frac{1}{2}\sum_{z\in\pi^{-1}(z')} f(z)$, is obviously continuous, but not differentiable at the origin. Indeed, for*

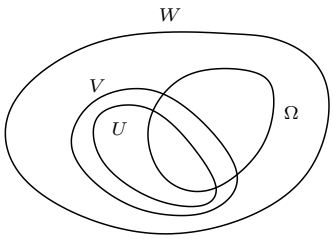


Figure 7.2. The open sets U, V, W, Ω of Lemma 7.6.8.

$z' = x + iy \in \mathbb{D}$ with vanishing imaginary part y and positive real part x we have that $\pi^{-1}(z') = \{\sqrt{x}, -\sqrt{x}\}$, hence $g(z') = x$. On the other hand, if $z' = x + iy$ has $y = 0$ and $x < 0$, then $\pi^{-1}(z') = \{i\sqrt{|x|}, -i\sqrt{|x|}\}$, and so $g(z') = 0$, hence g is not differentiable even along the line $y = 0$.

The key result of Perutz [**110**] for smoothing continuous Kähler cocycles, and ultimately verifying Theorem 7.6.1, rests on the complex analytic theorem of Varouchas [**141**], which we give below. (For the proof of this theorem see Subsection 7.6.1.)

Lemma 7.6.8 (*Lemme principal*, [**141**])**.** *Suppose that $U, V, W, \Omega \subset \mathbb{C}^n$ are open sets such that U and V are bounded, and $U \subset \overline{U} \subset V \subset \overline{V} \subset W \subset \mathbb{C}^n$ and $\Omega \subset W$. Suppose furthermore that $\phi\colon W \to \mathbb{R}$ is a continuous, strictly plurisubharmonic function which is smooth on Ω. Then there is a function $\psi\colon W \to \mathbb{R}$ satisfying the following properties:*

- *ψ continuous, strictly plurisubharmonic,*
- *$\psi|_{W\setminus\overline{V}} = \phi|_{W\setminus\overline{V}}$, and*
- *ψ is smooth on $\Omega \cup U$.*

For a schematic picture of the sets encountered above, see Figure 7.2.

The proof of Theorem 7.6.1 will follow from:

Proposition 7.6.9 (Perutz, [**110**]; cf. also [**141**])**.** *Suppose that $\{(U_i, \varphi_i)\}_{i\in I}$ is a continuous Kähler cocycle on the complex manifold X, $X_1, X_2 \subset X$ are two open subsets with $X = X_1 \cup X_2$ and the functions $\varphi_i|_{U_i\cap X_1}$ are smooth. Then there is a continuous function $\chi\colon X \to \mathbb{R}$ with $Supp(\chi) \subset X_2$ and a refinement $V_j \subset U_{i(j)}$ $(j \in J)$ such that the family*

$$\{(V_j, \varphi_{i(j)}|_{V_j} + \chi|_{V_j})\}_{j\in J}$$

is a smooth Kähler cocycle on X.

Proof. Refine the open cover $\{U_i\}$ so that each element of the cover is either in X_1 or in X_2; let $\{V_{i,1}\}_{i\in\mathbb{N}}$ denote the ones in X_1 and $\{V_{j,2}\}_{j\in\mathbb{N}}$ in

X_2. (We can assume that the original cover is countable; if it is finite, repeat the last open set infinitely many times.) After taking a further refinement if necessary, we can arrange that the cover is locally finite; that is, each point in X is contained in finitely many distinct open sets of the cover. The Kähler cocycle corresponding to this refinement (obtained by restricting the functions) is $\{(V_{i,1}, \varphi_{i,1})\}_{i\in\mathbb{N}} \cup \{(V_{j,2}, \varphi_{j,2})\}_{j\in\mathbb{N}}$.

Find open sets $V''_{j,2}, V'_{j,2}$ satisfying $V''_{j,2} \subset \overline{V}''_{j,2} \subset V'_{j,2} \subset \overline{V}'_{j,2} \subset V_{j,2}$, such that the union $\cup_j V''_{j,2}$ covers $X_2 \setminus X_1$.

Our next aim is to construct a sequence of continuous Kähler cocycles $\{(V_{i,k}, \psi^n_{i,k})\}_{i\in\mathbb{N},k=1,2}$ for each $n \in \mathbb{N}$ such that $\psi^n_{i,k}$ is smooth on $V_{i,k} \cap (X_1 \bigcup \cup_{j=1}^{n-1} V''_{j,2})$ for every $i \in \mathbb{N}$ and $k = 1, 2$. We define $\psi^1_{i,k} = \varphi_{i,k}$ and find the sequence $\{\psi^n_{i,k}\}$ by defining appropriate continuous functions $\chi_n \colon X \to \mathbb{R}$ supported in $V'_{n-1,2}$; then the functions for the n^{th} cocycle will be inductively defined by

$$\psi^n_{i,k} = \psi^{n-1}_{i,k} + \chi_n. \tag{7.13}$$

Define χ_n as follows. If $\psi^{n-1}_{n-1,2}$ is smooth, let $\psi^n_{n-1,2} = \psi^{n-1}_{n-1,2}$ (and $\chi_n = 0$). If $\psi^{n-1}_{n-1,2}$ is not smooth, apply Lemma 7.6.8 for the following setting:

$$(U, V, W, \Omega) = (V''_{n-1,2}, V'_{n-1,2}, V_{n-1,2}, V_{n-1,2} \cap \left(X_1 \cup \bigcup_{j=1}^{n-2} V''_{j,2} \right))$$

with the strictly plurisubharmonic function $\phi = \psi^{n-1}_{n-1,2}$. Lemma 7.6.8 then provides a new function $\psi^n_{n-1,2}$, and consider the difference $\chi_n = \psi^n_{n-1,2} - \psi^{n-1}_{n-1,2}$ (extended as zero on the complement of $V'_{n-1,2}$). Define the maps for the n^{th} cocycle as $\psi^n_{i,k} = \psi^{n-1}_{i,k} + \chi_n$. Note that by induction we have $\psi^n_{i,k} = \psi^1_{i,k} + \sum_{j=2}^n \chi_j$; in particular,

$$\psi^n_{i,k} - \psi^n_{n-1,2} = \psi^1_{i,k} - \psi^1_{n-1,2}.$$

For x not in $V'_{n-1,2}$ we have that $\psi^n_{i,k}(x) = \psi^{n-1}_{i,k}(x)$, which was already smooth on $V_i \cap (X_1 \cup \bigcup_{j=1}^{n-1} V''_{j,2})$, while for $x \in V'_{n-1,2}$ we have

$$\psi^n_{i,k} = (\psi^n_{i,k} - \psi^n_{n-1,2}) + \psi^n_{n-1,2} = (\psi^1_{i,k} - \psi^1_{n-1,2}) + \psi^n_{n-1,2}$$

where the first summand is pluriharmonic (hence smooth) and the second one is smooth and strictly plurisubharmonic. It also shows that the function $\psi^n_{i,k}$ is strictly plurisubharmonic.

Consider now the function $\chi \colon X \to \mathbb{R}$ defined as

$$\chi(x) = \sum_{n\in\mathbb{N}} \chi_n(x).$$

Since the sum is locally finite (as the cover was chosen to be locally finite), the definition makes sense.

The functions $\zeta_{i,k} = \psi^1_{i,k} + \chi = \varphi_{i,k} + \chi$ now define a Kähler cocyle, which is smooth: indeed, on $X \setminus \cup_j V'_{j,2} \subset X_1$ the cocycle $\varphi_{i,k}$ was already smooth and it has not been changed, while on $V''_{j,2} \subset X_1 \cup \bigcup_{i=1}^{n-1} V''_{i,2}$ the smoothness follows from the construction. □

Proof. [Proof of Theorem 7.6.1] For $x \in \times^m(\Sigma)$ consider a neighbourhood U_x biholomorphic to the disk D^{2m} and let ϕ_x be the solution of the equation $-dd^{\mathbb{C}}\phi_x = \nu^\times|_{U_x}$. For $x' \in \mathrm{Sym}^m(\Sigma)$ let us choose the open neighbourhood $V_{x'}$ so that $\pi^{-1}(V_{x'}) \subset \cup_{x\in\pi^{-1}(x')} U_x$. Consider a finite subcover $\{V_{x'_1}, \ldots, V_{x'_k}\}$ of the cover $\{V_{x'} \mid x' \in \mathrm{Sym}^m(\Sigma)\}$ of $\mathrm{Sym}^m(\Sigma)$. Next consider a finite subcover $\{W_1, \ldots, W_K\}$ of the cover $\{W_x = U_x \cap \pi^{-1}(V_{x'_i}) \mid \pi(x) = x'_i\}$ of $\times^m(\Sigma)$. Restricting ϕ_x to W_x we have a smooth Kähler cocycle over $\times^m(\Sigma)$. We will perform a final adjustment on these functions. Indeed, for $y \in W_x$ consider the value

$$\varphi_x(y) = \frac{1}{m!} \sum_{\sigma\in\mathfrak{S}_m} \phi_{\sigma(x)}(\sigma(y)),$$

where in the sum we count the values with multiplicity. Since the resulting functions are averages of smooth strictly plurisubharmonic functions (and the differences are pluriharmonic on the intersections), and by the $\mathfrak{S}_m$-equivariance of $\nu^\times$ all solve the $dd^{\mathbb{C}}$ equations, we have that the resulting $\{(W_j, \varphi_j)\}$ is a smooth Kähler cocycle on $\times^m(\Sigma)$ which is, in addition, $\mathfrak{S}_m$-equivariant.

Let $\phi'_i \colon V_{x'_i} \to \mathbb{R}$ be the push-forward of the cocycle $\{(W_j, \varphi_j)\}$, that is, $\phi'_i(x') = \frac{1}{m!}\sum_{x\in\pi^{-1}(x')} \varphi_j(x)$, where the points in $\pi^{-1}(x')$ are counted with their branching multiplicity. Since φ_j is continuous, so is ϕ'_i, but (although φ_j is C^∞) the push-forward ϕ'_i might not be differentiable everywhere. Note that by the $\mathfrak{S}_m$-equivariance of the smooth Kähler cocycle $\{(W_j, \psi_j)\}$, we have that the pull-back of the push-forward $\{(V_{x'_i}, \phi'_i)\}$ is equal to $\{(W_j, \varphi_j)\}$.

Nevertheless, $\{(V_{x'_i}, \phi'_i)\}$ is a continuous Kähler cocycle: away from the branch locus the functions ϕ'_i are clearly plurisubharmonic (and the differences are pluriharmonic), and since by [**38**] these properties extend through the diagonal, the properties follow for the functions. We also need that the functions ϕ'_i are strictly plurisubharmonic. For a given point $x' \in V_{x'_i}$ take a neighbourhood A such that on every component of $\pi^{-1}(W_j)$ the function ϕ'_i can be written as the sum of a plurisubharmonic function and $c\|y\|^2$ for

some fixed positive c. The push-forward of the plurisubharmonic part provides a plurisubharmonic function on A, while $c\|y\|^2$ pushed forward gives a similar term, concluding the argument.

Next we replace the continuous Kähler cocycle $\{(V_{x_i'}, \phi_i')\}$ with a smooth one, as follows. Fix an open neighborhood U of the diagonal $\Delta \subset \mathrm{Sym}^m(\Sigma)$. Let $U' \subset U$ be an open set with the property that its closure $\overline{U'}$ is in U and apply Proposition 7.6.9 with the choice $X_1 = \mathrm{Sym}^m(\Sigma) - \overline{U'}$ and $X_2 = U$ for the contiuous Kähler cocycle $\{(V_{x_i'}, \phi_i')\}$. Proposition 7.6.9 provides a continuous function $\chi\colon \mathrm{Sym}^m(\Sigma) \to \mathbb{R}$ such that when adding this to the continuous Kähler cocycle we get a smooth Kähler cocycle with functions $\phi_i' + \chi$ and corresponding symplectic form $\omega = -dd^{\mathbb{C}}(\phi_i' + \chi)$.

The definition of ω (and the property that $\phi_i' + \chi$ are strictly plurisubharmonic functions) shows that it is a Kähler form. The pull-back $\pi^*(\omega)$ can be represented by the pull-back of the smooth Kähler cocycle given by the functions $\{\phi_i' + \chi\}$. By our choice, the pull-back of $\{\phi_i'\}$ is the smooth Kähler cocycle defining $\nu^{\times}$, hence the pull-back $\tilde{\chi}$ of χ is a smooth function with the property that

$$\pi^*(\omega) - \nu^{\times} = -dd^{\mathbb{C}}\tilde{\chi}.$$

Since χ is supported in the chosen open neighbourhood U of the diagonal, by taking $\eta = -d^{\mathbb{C}}\tilde{\chi}$ the claim of the theorem is verified. □

7.6.1. The proof of Varouchas' lemma. For the sake of completeness, we include the proof of Lemma 7.6.8 of Varouchas here; so fix $U, V, W, \Omega \subset \mathbb{C}^n$ and $\phi\colon W \to \mathbb{R}$ as in the statement. The lemma is proved by mollifying ϕ to smooth it out near V, and interpolating between this mollified version with ϕ while preserving strict plurisubharmonicity. We start with finding the convolving functions to smooth out ϕ near V, and turn to the interpolation in Equation (7.15).

To find the the right convolution, we make some preliminary choices and definitions:

- For $\rho > 0$ fix a non-negative smooth function $\alpha_\rho\colon \mathbb{C}^n \to \mathbb{R}_{\geq 0}$ supported in the ball $B(0,\rho)$ of radius ρ and center 0, with $\int_{\mathbb{C}^n} \alpha_\rho = 1$. For example, fix a non-negative smooth function $\alpha\colon \mathbb{C}^n \to \mathbb{R}_{\geq 0}$ with $\mathrm{Supp}(\alpha) \subset B(0,1)$ and $\int_{\mathbb{C}^n} \alpha = 1$, and define $\alpha_\rho(z) = \rho^{-1}\alpha(\frac{z}{\rho})$ for $\rho > 0$.
- Choose a smooth function $\eta\colon W \to [-1,1]$ such that $\eta = 1$ on U and $\eta = -1$ on $W \setminus V$.

- Let $\xi\colon \mathbb{R}^2 \to \mathbb{R}$ be a smooth even function, supported in $(-1,1)\times(-1,1)$ and satisfying

$$\int_{\mathbb{R}^2} \xi(t_1,t_2)dt_1dt_2 = 1.$$

Consider the domain $W_\rho = \{x \in W \mid \operatorname{dist}(x, \mathbb{C}^n \setminus W) > \rho\}$. The convolution

$$(\phi * \alpha_\rho)(x) = \int_{\mathbb{C}^n} \phi(y)\alpha_\rho(x-y)dy$$

is defined on W_ρ; since the convolution of a smooth and a continuous function is smooth, we get that $\phi * \alpha_\rho$ is smooth on W_ρ.

Since ϕ is strictly plurisubharmonic, there is an open set $W' \subset W$ and a positive constant $c > 0$ such that on W' we have $\phi = \phi_1 + \phi_2$, where ϕ_1 is plurisubharmonic and ϕ_2 is the strictly plurisubharmonic function $\phi_2(x) = c \cdot \|x\|^2$.

Since ϕ_2 is strictly plurisubharmonic, and strict subharmonicity is an open condition, we can fix some $t > 0$ so that $\phi_2 + t\eta$ is strictly subharmonic.

Since V is bounded with $\overline{V} \subset W$, for any sufficiently small $\rho > 0$, we have that $\overline{V} \subset W_\rho$. For such a choice of ρ, the convolution $\phi * \alpha_\rho$ defines a function on $\overline{V}$. From the expression

$$(\phi * \alpha_\rho)(x) - \phi(x) = \int_{\mathbb{C}^n} (\phi(y) - \phi(x))\alpha_\rho(x-y)dy$$

and the fact that ϕ is continuous (hence uniformly contiuous) on the compact set $\overline{V}$, it follows that $\phi * \alpha_\rho$ uniformly converges to ϕ on $\overline{V}$ as $\rho \to 0$. Hence, we can pick $\rho > 0$ so that

$$|\phi * \alpha_\rho - \phi| < \frac{t}{2} \qquad \text{on} \qquad \overline{V}. \tag{7.14}$$

Lemma 7.6.10. *For sufficiently small $t > 0$, the function $\phi * \alpha_\rho + t\eta$ is strictly plurisubharmonic on a neighbourhood of $\overline{V}$.*

Proof. Recall that $\phi * \alpha_\rho = \phi_1 * \alpha_\rho + \phi_2 * \alpha_\rho$. Since convolution of a plurisubharmonic function with α_ρ is also plurisubharmonic, it follows that $\phi_1 * \alpha_\rho$ is plurisubharmonic. Thus the lemma follows if $\phi_2 * \alpha_\rho + t\eta$ is strictly plurisubharmonic. For any $0 < t < \frac{c}{\sup|dd^{\mathbb{C}}\eta|}$,

$$dd^{\mathbb{C}}(\phi_2 * \alpha_\rho + t\eta) = (dd^{\mathbb{C}}\phi_2) * \alpha_\rho + tdd^{\mathbb{C}}\eta = c + tdd^{\mathbb{C}}\eta > 0.$$

□

For $\delta \leq \frac{t}{4}$, let $M_\delta \colon \mathbb{R}^2 \to \mathbb{R}$ be the function defined by

$$M_\delta(x,y) = \int_{\mathbb{R}^2} \max\{x - \delta t_1, y - \delta t_2\} \cdot \xi(t_1,t_2) dt_1 dt_2.$$

Lemma 7.6.11. *If $|x-y| \geq 2\delta$ then $M_\delta(x,y) = \max\{x,y\}$.*

Proof. Recall that if $\xi(t_1,t_2) \neq 0$, then $t_1, t_2 \in (-1,1)$. Suppose that $x - y \geq 2\delta$. Then, $x - t_1\delta > y - t_2\delta$, hence the maximum of these two terms is $x - t_1\delta$. Since ξ is even by assumption, it follows that $\int t_1 \xi(t_1,t_2) dt_1 dt_2 = 0$, and so $M_\delta(x,y) = x$. A similar argument applies when $y - x \geq 2\delta$. □

Consider the function

$$\psi_\delta(z) = M_\delta(\phi(z), (\phi * \alpha_\rho + t\eta)(z)). \tag{7.15}$$

Lemma 7.6.12. *The function ψ_δ is strictly plurisubharmonic and it is equal to ϕ on a neighborhood of ∂V.*

Proof. If f, g are two strictly plurisubharmonic functions, then so are $f - \delta t_1$ and $g - \delta t_2$ and also $\max(f - \delta t_1, g - \delta t_2)$. Then the convolution is also strictly plurisubharmonic.

Since $\eta = -1$ on $W \setminus V$, near ∂V the function $t\eta$ is arbitrary close to $-t$. Since on $\overline{V}$ we have $\phi * \alpha_\rho - \phi < \frac{t}{2}$, it follows that on a sufficiently small neighbourhood of ∂V we have

$$\phi * \alpha_\rho + t\eta < \phi - \frac{t}{2} \leq \phi - 2\delta.$$

By Lemma 7.6.11 the claim on ψ_δ follows. □

Define the function ψ by

$$\psi(x) = \begin{cases} \psi_\delta(x) & x \in V \\ \phi(x) & x \in W \setminus V. \end{cases}$$

By the behaviour of ψ_δ near ∂V, the resulting function $\psi \colon W \to \mathbb{R}$ is continuous on W. With this definition now we can turn to the proof of the *Lemme principal*, Lemma 7.6.8.

Proof. [Proof of Lemma 7.6.8] The function ψ defined above clearly equals ϕ on $W \setminus V$ and by Lemma 7.6.12 it is strictly plurisubharmonic. We need to show that it is smooth on $U \cup \Omega$. Once again, as on the complement $W \setminus V$ the function ψ is equal to ϕ (and ϕ was assumed to be smooth on Ω), we need to concentrate on the smoothness of ψ on $(U \cup \Omega) \cap V$.

Combining the fact that $\eta \equiv 1$ on U with Equation (7.14), we have that $\phi * \alpha_\rho + t\eta > \phi + \frac{t}{2} \geq \phi + 2\delta$. By Lemma 7.6.11, the restriction of $\psi_\delta = \psi$ to U agrees with the smooth function $\phi * \alpha_\rho + t\eta$, establishing smoothness on U.

Finally, the smoothness of ψ on $\Omega \cap V$ follows from the fact that ϕ is smooth there, $\phi * \alpha_\rho + t\eta$ is smooth on V, and the function M_δ is smooth on $\mathbb{R}^2$. (Note that we chose ρ so small that $\overline{V} \subset W_\rho$.) Thus, over $\Omega \cap V$, the function ψ is a composition of smooth functions, and therefore is smooth.

□

Chapter 8

The Maslov Index Formula

The Maslov index plays a crucial role in the definition of the boundary map in Lagrangian (and so in particular in Heegaard Floer) homology. In addition, the concept of the Maslov index will be central in finding a grading on Heegaard Floer homology in Chapter 10. The aim of this chapter is to give an explicit formula for the Maslov index of a Whitney disk in the symmetric product of a Heegaard surface, expressed explicitly in terms of combinatorial data associated to the Heegaard diagram. This formula was established by Robert Lipshitz, building on work of Jacob Rasmussen. We start by identifying the intersection points of the Lagrangian tori $\mathbb{T}_\alpha$ and $\mathbb{T}_\beta$ in $\mathrm{Sym}^g(\Sigma)$ with the Heegaard states of the diagram in Section 8.1. Next, we state a the formula for the Maslov index in Section 8.2. We illustrate the use of this formula by some examples in Section 8.3. By constructing the necessary geometric objects in Section 8.4, the proof of the formula will be presented in Section 8.5. In Section 8.6 the relation of the Maslov index of a periodic domain, and the evaluation of (the first Chern class of) a spin^c structure is explained — the resulting formula will be used in Chapter 10. Additivity (and some further basic properties) of the Maslov index from this point of view will be given in Section 8.7, and the chapter concludes with a discussion of the corresponding formula for polygons in Section 8.8. For an alternative treatment of the Maslov index in Heegaard Floer homology see [**64**].

8.1. Heegaard states and intersection points; Whitney disks and shadows

Before turning to the Maslov index formula, we study generators for the Lagrangian Floer homology in the symmetric product of a Heegaard diagram. Specifically, fix a Heegaard diagram $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$ representing a given three-manifold Y. Perutz's construction (Corollary 7.6.2) endows the symmetric product $\mathrm{Sym}^g(\Sigma)$ with a symplectic structure for which the tori

$$\mathbb{T}_\alpha = \alpha_1 \times ... \times \alpha_g \quad \text{and} \quad \mathbb{T}_\beta = \beta_1 \times ... \times \beta_g$$

are Lagrangian submanifolds. Lagrangian Floer homology is a chain complex generated by intersection points of $\mathbb{T}_\alpha \cap \mathbb{T}_\beta$; but these are objects we have met already in Section 2.4. Specifically:

Proposition 8.1.1. *There is a one-to-one correspondence between the Heegaard states from Definition 2.4.1 and the intersection points of $\mathbb{T}_\alpha$ and $\mathbb{T}_\beta$.*

Proof. Label the elements of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ by $\{1, ..., g\}$. An intersection point between $\mathbb{T}_\alpha$ and $\mathbb{T}_\beta$ is a permutation σ in the symmetric group on g letters and an ordered g-tuple of intersection points $x_i \in \alpha_i \cap \beta_{\sigma(i)}$; and this is precisely a Heegaard state. □

For two intersection points $\mathbf{x}, \mathbf{y} \in \mathbb{T}_\alpha \cap \mathbb{T}_\beta$ consider $\phi \in W(\mathbf{x}, \mathbf{y})$. Given $z \in \Sigma \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta})$, let $n_z(\phi)$ be the algebraic intersection number of ϕ with $\{z\} \times \mathrm{Sym}^{g-1}(\Sigma)$. If z and z' are in the same component of $\Sigma \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta})$, then $\{z\} \times \mathrm{Sym}^{g-1}(\Sigma)$ and $\{z'\} \times \mathrm{Sym}^{g-1}(\Sigma)$ are homotopic in the complement of $\mathbb{T}_\alpha \cup \mathbb{T}_\beta$; this implies that $n_z(\phi) = n_{z'}(\phi)$.

Cutting Σ along $\boldsymbol{\alpha} \cup \boldsymbol{\beta}$, we obtain the decomposition of Σ into elementary domains $\{\mathcal{D}_k\}_{k=1}^m$, so that $\Sigma \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta}) = \bigcup_{k=1}^m \mathrm{int}\, \mathcal{D}_k$. Thus, given $\mathbf{x}, \mathbf{y} \in \mathbb{T}_\alpha \cap \mathbb{T}_\beta$, and $\phi \in W(\mathbf{x}, \mathbf{y})$ there is a well-defined formal sum of the $\{\mathcal{D}_k\}_{k=1}^m$ defined by

$$\mathcal{S}(\phi) = \sum_{k=1}^{m} n_{z_k}(\phi) \cdot \mathcal{D}_k, \tag{8.1}$$

where z_k are any points with $z_k \in \mathrm{int}\, \mathcal{D}_k$.

Proposition 8.1.2. *Given $\mathbf{x}, \mathbf{y} \in \mathbb{T}_\alpha \cap \mathbb{T}_\beta$ and $\phi \in W(\mathbf{x}, \mathbf{y})$, the linear combination $\mathcal{S}(\phi)$ is a domain from the Heegaard state $\mathbf{x}$ to the Heegaard state $\mathbf{y}$, in the sense of Definition 2.4.4.*

Proof. Given a Whitney strip from $u\colon \mathbb{R} \times [0,1] \to \mathrm{Sym}^g(\Sigma)$ from $\mathbf{x}$ to $\mathbf{y}$ (as in Definition 6.4.1), the restriction of u to $\mathbb{R} \times \{0\}$ is a path a in

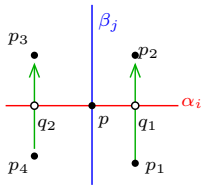


Figure 8.1. Multiplicities and intervals near an intersection point $p \in \alpha_i \cup \beta_j$ of the $\boldsymbol{\alpha}$- and $\boldsymbol{\beta}$-curves.

$\mathbb{T}_\alpha$ from $\mathbf{x}$ to $\mathbf{y}$. Since $\mathbb{T}_\alpha$ is disjoint from the diagonal, the path a can alternatively be interpreted as a g-tuple of paths $\{a_i \subset \alpha_i\}_{i=1}^g$, where a_i connects $x_i = \mathbf{x} \cap \alpha_i$ to $y_i = \mathbf{y} \cap \alpha_i$.

Consider an intersection point $p \in \alpha_i \cap \beta_j$. Draw four nearby basepoints $\{p_1, p_2, p_3, p_4\}$ as in Figure 8.1. Draw an oriented arc from p_1 to p_2 that crosses α_i in a single transverse intersection point q_1. The intersection number of a with the the $g-1$-dimensional torus

$$\{q_i\} \times \alpha_1 \times \cdots \times \widehat{\alpha_i} \times \cdots \times \alpha_g \subset \mathbb{T}_\alpha$$

is computed by $n_{p_2}(\phi) - n_{p_1}(\phi)$; equivalently, the signed number of times a_i passes through q_i (denoted by $\#(a_i \cap q_i)$) is equal to $n_{p_2}(\phi) - n_{p_1}(\phi)$. Similarly, $\#(a_i \cap q_2) = n_{p_3}(\phi) - n_{p_4}(\phi)$, implying

$$n_{p_1}(\phi) - n_{p_2}(\phi) + n_{p_3}(\phi) - n_{p_4}(\phi) = \#(a_i \cap \{q_1\}) - \#(a_i \cap \{q_2\}).$$

The latter quantity, in turn, is the sum of $+1$ if $p = y_i$ and -1 if $p = x_i$. Therefore the linear relation characterizing domains from $\mathbf{x}$ to $\mathbf{y}$ (Equation (2.3)) follows, concluding the proof. □

Definition 8.1.3. *Given $\mathbf{x}, \mathbf{y} \in \mathbb{T}_\alpha \cap \mathbb{T}_\beta$ and $\phi \in W(\mathbf{x}, \mathbf{y})$, the **shadow** of ϕ, denoted $\mathcal{S}(\phi) \in D(\mathbf{x}, \mathbf{y})$ is the domain given by Equation* (8.1).

8.2. Statement of the Maslov index formula

In the following we will express the Maslov index $\mu(\phi)$ of $\phi \in W(\mathbf{x}, \mathbf{y})$ in terms of quantities associated to the shadow $\mathcal{S}(\phi) \in D(\mathbf{x}, \mathbf{y})$ of ϕ. We start with the definition of two functions on $D(\mathbf{x}, \mathbf{y})$, which will play a central role in the formula expressing the Maslov index of ϕ.

Definition 8.2.1. *Suppose that $\phi \in D(\mathbf{x}, \mathbf{y})$ is a domain from the Heegaard state $\mathbf{x}$ to the Heegaard state $\mathbf{y}$. For each $t \in \alpha_i \cap \beta_j$, let $\overline{n}_t(\mathcal{D})$ be the average of the multiplicities of $\mathcal{D}$ at the four (not necessarily distinct) elementary domains having corner at t. Let*

$$\overline{n}_{\mathbf{x}}(\mathcal{D}) = \sum_{x_i \in \mathbf{x}} \overline{n}_{x_i}(\mathcal{D}) \qquad \text{and} \qquad \overline{n}_{\mathbf{y}}(\mathcal{D}) = \sum_{y_i \in \mathbf{y}} \overline{n}_{y_i}(\mathcal{D}).$$

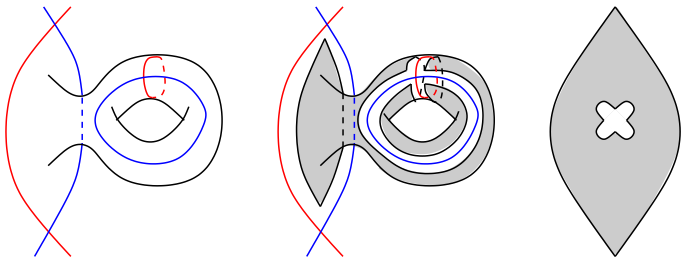


Figure 8.2. Consider the elementary domain $\mathcal{D}$ on the left, which is a connected sum of a bigon with a torus. The region $\mathcal{D}^c$ is shaded in the middle and redrawn on the right. It has $b = 6$ and $\chi = 0$, thus $e(\mathcal{D}) = -\frac{3}{2}$

*The **point measure** $\overline{n}(\mathcal{D})$ of $\mathcal{D} \in D(\mathbf{x}, \mathbf{y})$ is defined by*

$$\overline{n}(\mathcal{D}) = \overline{n}_{\mathbf{x}}(\mathcal{D}) + \overline{n}_{\mathbf{y}}(\mathcal{D}).$$

Example 8.2.2. *For an elementary domain $\mathcal{D}_i$ from $\mathbf{x}$ to $\mathbf{y}$, which is an embedded $2n$-gon, we have $\overline{n}(\mathcal{D}_i) = 2n \cdot \frac{1}{4} = \frac{n}{2}$.*

The second function on domains is the following:

Definition 8.2.3. *Recall that each elementary domain $\mathcal{D}_i$ is the closure of some path component $\mathcal{D}_i^\circ$ in $\Sigma \setminus \boldsymbol{\alpha} \cup \boldsymbol{\beta}$. Let $\mathcal{D}_i^c \subset \mathcal{D}_i^\circ$ denote the subset formed as the complement of a sufficiently small tubular neighborhood of the α-curves and β-curves in $\mathcal{D}_i^\circ$. The boundary $\partial\mathcal{D}_i^c$ is a union of intervals, each of which are parallel to some α- or β-curve. Let $b(\mathcal{D}_i^c)$ be the number of these intervals, and let $\chi(\mathcal{D}_i^c)$ denote the Euler characteristic of the subspace $\mathcal{D}_i^c \subset \Sigma$. The **Euler measure** of the elementary domain $\mathcal{D}_i$, denoted $e(\mathcal{D}_i)$, is defined by*

$$e(\mathcal{D}_i) = \chi(\mathcal{D}_i^c) - \frac{1}{4} b(\mathcal{D}_i^c).$$

*For any $\mathcal{D} = \sum_i n_i \cdot \mathcal{D}_i \in D(\mathbf{x}, \mathbf{y})$ define the **Euler measure** of $\mathcal{D}$ by*

$$e(\mathcal{D}) = \sum_i n_i \cdot e(\mathcal{D}_i).$$

Example 8.2.4. *For an elementary domain $\mathcal{D}_i$ which is an embedded $2n$-gon we have $e(\mathcal{D}_i) = 1 - \frac{n}{2}$. See Figure 8.2 for a more complicated domain.*

Remark 8.2.5. *The Euler measure of an elementary domain has the following geometric interpretation. Equip Σ with a Riemannian metric with the property that all of the curves α_i and β_j are geodesics, and at each intersection point of α_i with β_j, the geodesics meet at right angles. Then, the Euler measure $e(\mathcal{D}_i)$ of $\mathcal{D}_i$ is equal to the integral of the curvature of the metric over $\mathcal{D}_i$, divided by 2π.*

The main result of this chapter is the following:

Theorem 8.2.6. *(Lipshitz [**69**]) Fix a pointed Heegaard diagram $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$ and a pair of Heegaard states $\mathbf{x}, \mathbf{y} \in \mathcal{S}(\mathcal{H})$. For any $\phi \in W(\mathbf{x}, \mathbf{y})$, we have that*

$$\mu(\phi) = \overline{n}(\mathcal{S}(\phi)) + e(\mathcal{S}(\phi)). \tag{8.2}$$

Theorem 8.2.6 is an immediate consequence of the following two theorems:

Theorem 8.2.7. *(Rasmussen [**112**]) Let $\mathbf{x}, \mathbf{y}$ be any two Heegaard states, and $\phi \in W(\mathbf{x}, \mathbf{y})$. Then*

$$\mu(\phi) = 2e(\mathcal{S}(\phi)) + \Delta(\phi), \tag{8.3}$$

where $\Delta(\phi)$ denotes the algebraic intersection number of ϕ with the diagonal $\Delta \subset \mathrm{Sym}^g(\Sigma)$.

Theorem 8.2.8. *(Lipshitz [**69**]) Let $\mathbf{x}, \mathbf{y}$ be any two Heegaard states, and $\phi \in W(\mathbf{x}, \mathbf{y})$, then the intersection number of ϕ with the diagonal Δ is computed as*

$$\Delta(\phi) = \overline{n}(\mathcal{S}(\phi)) - e(\mathcal{S}(\phi))). \tag{8.4}$$

To prove the above results, it helps to have a concrete understanding of Whitney disks in the symmetric product, as appropriately decorated surfaces, equipped with a branched cover map to the standard disk and a map into Σ. We formalize this more explicitly as follows.

Definition 8.2.9. *A **decorated surface** F is an oriented surface with boundary, equipped with a decomposition of its boundary ∂F as a union of connected arcs $\{A_i\}_{i=1}^m \cup \{B_i\}_{i=1}^m$. Each A_i can be thought of as an oriented arc (with its boundary orientation) from some point $p_i \in \partial F$ to some point $q_i \in \partial F$, and B_i from q_i to p_{i+1}. The points $\{p_i\}_{i=1}^m$ are called **initial corners** and $\{q_i\}_{i=1}^m$ are called **terminal corners**.*

A decorated surface can be thought of as a surface with corners. Each point $p \in F$ has a neighborhood U_p which is locally modeled either on the unit disk $\mathbb{D}$ in $\mathbb{C}$, or on the upper half disk

$$\mathbb{D}^+ = \{z = x + iy \in \mathbb{C} \big| |z| \leq 1, y \geq 0\},$$

or on the intersection of the disk with a quadrant

$$\mathbb{D}^{++} = \{(z = x + iy \in \mathbb{C} \big| |z| \leq 1, x, y \geq 0\},$$

where p maps to the origin in all three models. Points of the first kind are called *interior points*; points in the interior of A_i and B_i have local models in $\mathbb{D}^+$ and are called *boundary points*; and points of the third kind are of the form $p = p_i$ or q_i, and are called *corner points*.

Definition 8.2.10. *Fix a genus g Heegaard diagram $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$, and let F be a decorated surface, with $m = g$ (i.e., the genus of Σ). Given states $\mathbf{x}, \mathbf{y}$, a smooth map $f\colon F \to \Sigma$ is called* ***compatible with the decoration*** *if the following conditions hold:*

- $f|_{A_i} \subset \alpha_i$ *for some suitable numbering of* $\boldsymbol{\alpha}$,
- $f|_{B_i} \subset \beta_i$ *for some suitable numbering of* $\boldsymbol{\beta}$,
- $f(\{p_i\}_{i=1}^m) = \mathbf{x}$,
- $f(\{q_i\}_{i=1}^m) = \mathbf{y}$.

Think of $\mathbb{D} \subset \mathbb{C}$ as a decorated surface, where A is the right arc and B is the left arc of $\partial\mathbb{D}$. A branched cover $\pi\colon F \to \mathbb{D}$ is *compatible with the decoration* if $\pi(A_i) \subset A$ and $\pi(B_i) \subset B$. The corner points $\{p_j\}$ and $\{q_j\}$ map to $\pm i \in \mathbb{D}$. Note that branch points occur in the interior of $\mathbb{D}$.

In Proposition 7.5.2, we explained how a branched cover of the sphere mapping into Σ gives rise to a sphere in the symmetric product. This construction has a generalization to disks, as follows.

Definition 8.2.11. *Suppose that F is a decorated surface, and $f\colon F \to \Sigma$ is a smooth map compatible with the decoration. Suppose moreover that $\pi\colon F \to \mathbb{D}$ is a branched m-fold cover which is compatible with the decoration. The pair f and π give rise to an* ***associated Whitney disk****, which is the continuous map $u\colon \mathbb{D} \to \mathrm{Sym}^m(\Sigma)$, defined by $u(z) = f(\pi^{-1}(z))$: i.e. given $z \in \mathbb{D}$, think of the preimage $\pi^{-1}(z)$ as a set of m points (counted possibly with multiplicity) in F, and apply f to each element of $\pi^{-1}(\{z\})$ to obtain a set of m points in Σ (once again, counted with multiplicity), thought of as an element of* $\mathrm{Sym}^m(\Sigma)$.

We are typically interested in the case where m coincides with the genus g of the surface Σ. We will find it convenient to drop this requirement in the next few sections; see for example Definition 8.4.1.

We give some examples to illustrate Theorems 8.2.6 and 8.2.8 in Section 8.3. Theorem 8.2.8 is proved in Section 8.4, after constructing suitable decorated surfaces, and Theorem 8.2.7 will be proved in Section 8.5.

8.3. Examples

We analyze some domains to illustrate Equations (8.2), (8.3), and (8.4).

Polygons. Suppose that $\mathbf{x}$ and $\mathbf{y}$ are two states in some Heegaard diagram so that $x_i = y_i$ for $i = n+1, \dots, g$, and x_i, y_i for $i = 1, \dots, n$ are the corners of an elementary domain, which is a $2n$-gon. That elementary domain can be viewed as a domain $\mathcal{D} \in D(\mathbf{x}, \mathbf{y})$. (See the left diagram of Figure 8.3.)

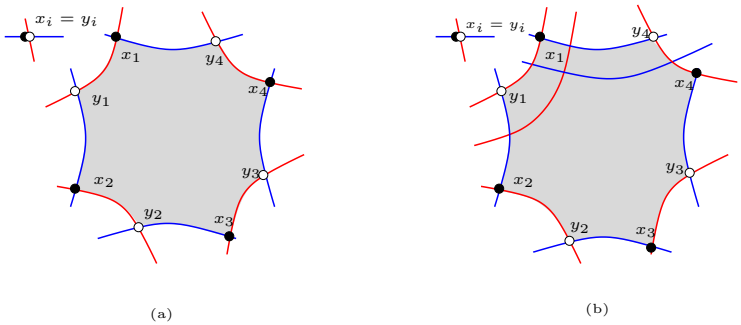


Figure 8.3. The $2n$-gon with $n = 4$. The cross next to the octagon represents all other components of the Heegaard states $\mathbf{x}$ (solid dots) and $\mathbf{y}$ (hollow dots).

In this case,

$$e(\mathcal{D}) = 1 - \frac{n}{2} \qquad \text{and} \qquad \overline{n}(\mathcal{D}) = \frac{n}{2}, \tag{8.5}$$

so $\mu(\mathcal{D}) = 1$.

For these polygons, we take a closer look at Equation (8.4) in cases where $n = 1$ and 2; starting with the case $n = 1$, i.e. where the elementary domain is a bigon. Consider the decorated surface F consisting of g disjoint bigons, and let $f\colon F \to \Sigma$ be the map which sends one of the components homeomorphically to $\mathcal{D}$, mapping the remaining $g-1$ bigons to the $g-1$ points $x_i = y_i$ for $i = 2, \dots, g$. At the same time, we view F also as a trivial (unbranched) covering $\pi\colon F \to \mathbb{D}$ of the standard disk. Consider the associated disk $u\colon \mathbb{D} \to \Sigma$, as in Definition 8.2.11. By construction, $\mathcal{S}(u) = \mathcal{D}$. Moreover, u clearly misses the diagonal Δ, which is consisent with the the fact that, in this case, $\overline{n}(\mathcal{D}) = e(\mathcal{D}) = 1/2$ (cf. Equation (8.4)).

Next, consider $n = 2$. In this case, we consider the surface F consisting of a disjoint union of one rectangle R, which maps homeomorphically to $\mathcal{D}$, and $g-2$ bigons, each of which is mapped by constant maps to the remaining $x_i = y_i$ for $i = 3, \dots, g$. The rectangle R can be thought of as a branched double-cover of a disk, with a single branch point at the center.

Explicitly, let $R = [-a, a] \times [-b, b]$, and consider the involution $j\colon R \to R$ given by $j(x, y) = (-x, -y)$. Clearly, R/j is a disk, and the origin $(0, 0)$ corresponds to the branch point. Consider the map $\pi\colon F \to \mathbb{D}$ which is this branched covering on the first component and the trivial cover on the remaining $g-2$ components. Once again, (f, π) gives rise to a Whitney disk $u \in W(\mathbf{x}, \mathbf{y})$ with $\mathcal{S}(u) = \mathcal{D}$. In this case, u meets the diagonal transversely at the branch point of the covering, as one would expect from the fact that $\overline{n}(\mathcal{D}) = 1$ and $e(\mathcal{D}) = 0$.

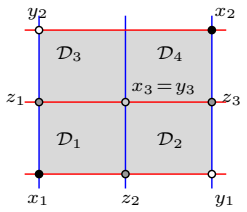


Figure 8.4. Rectangle with a point inside.

Exercise 8.3.1. *Construct a surface F as above for a $2n$-gon with $n > 2$.*

Equation (8.5) has a straightforward generalization to polygons that are not necessarily elementary domains. Let $\mathcal{D} \in D(\mathbf{x}, \mathbf{y})$ be an embedded polygon in Σ, which decomposes as a sum of elementary domains, each counted with multiplicity one. Suppose that its edges alternate between α- and β-arcs, and its corners alternate between the components of $\mathbf{x}$ and $\mathbf{y}$. Suppose also that each corner is the corner of exactly one of the component elementary domains, i.e., at each corner c, $\overline{n}_c(\mathcal{D}) = 1/4$. Suppose finally that for all $i = 2n+1, \dots, g$, the corresponding $x_i = y_i$ is not contained in the polygon. In this case, Equation (8.5) still holds. (For an example, see the right diagram of Figure 8.3.)

Consider next the case of a polygon $\mathcal{D}$ that now contains exactly one component $x_{n+1} = y_{n+1}$ in its interior. For simplicity, we consider first the case where $n = 2$, as shown in Figure 8.4. In this case, the shadow $\mathcal{D}$ can be written as a sum of four elementary domains $\mathcal{D} = \sum_{i=1}^{4} \mathcal{D}_i$; $e(\mathcal{D}) = 0$, and $\overline{n}(\mathcal{D}) = 3$, so $\mu(\mathcal{D}) = 3$.

The computation $\mu(\mathcal{D}) = 3$ can be motivated as follows. One can find intermediate Heegaard states $\mathbf{x}' = \{z_1, z_2, x_2\}$ and $\mathbf{x}'' = \{y_2, z_2, z_3\}$ and domains $\mathcal{D}_1 \in D(\mathbf{x}, \mathbf{x}')$, $\mathcal{D}_{34} \in D(\mathbf{x}', \mathbf{x}'')$, and $\mathcal{D}_2 \in D(\mathbf{x}'', \mathbf{y})$ so that $\mathcal{D}_1$, $\mathcal{D}_{34}$ and $\mathcal{D}_2$ are represented by the domains $\mathcal{D}_1$, $\mathcal{D}_3 + \mathcal{D}_4$ and $\mathcal{D}_2$ as indicated in Figure 8.4. By the above computations, $\mu(\mathcal{D}_1) = \mu(\mathcal{D}_{34}) = \mu(\mathcal{D}_2) = 1$, so by the additivity of the Maslov index under juxtapositions (Proposition 6.5.7), it follows that $\mu(\mathcal{D}) = 3$, as indicated by the theorem.

It is also a consequence of Theorem 8.2.7 that if $\phi \in W(\mathbf{x}, \mathbf{y})$ has $\mathcal{S}(\phi) = \mathcal{D}$ of the above example, then $\Delta(\phi) = 3$. It is interesting to see this explicitly in two ways. Consider first the surface F_1 which is a copy of a rectangle R and a bigon D. Consider the map $f_1 \colon F_1 \to \mathcal{D}$ which maps R homeomorphically to $\mathcal{D}$, and sends D to the constant map at $x_3 = y_3$. Let $p \colon F_1 \to \mathbb{D}$ be the branched triple cover which is a standard branched double-cover on the R component, and a homeomorphism from D to $\mathbb{D}$. Combining f and p, we obtain a map $\phi \in W(\mathbf{x}, \mathbf{y})$. Clearly, $\Delta(\phi) = 3$: one of the intersection points

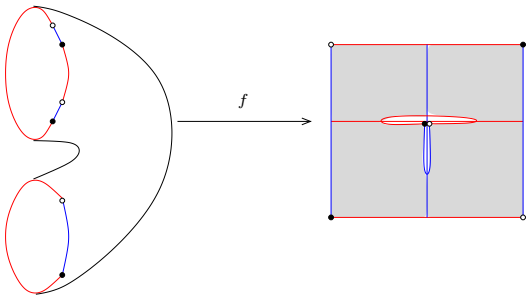


Figure 8.5. Decorated surface for the rectangle.

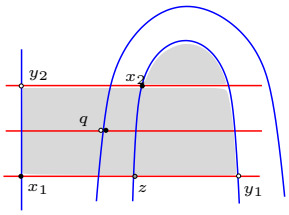


Figure 8.6. A domain with a point (x_2) with local multiplicity 3/4.

of ϕ and Δ corresponds to the branched point in R, and a multiplicity two intersection point occurs over the point where f maps to $x_3 = y_3$.

Another way to see this latter computation is as follows. Introduce three slits around $x_3 = y_3$ in the rectangle underlying $\mathcal{D}$, as shown in Figure 8.5. Topologically, the result is an annulus; and there is an abstract annulus F_2 with two α- and β-arcs on one boundary component, and one α- and one β-arc on the other, equipped with a map $f_2 \colon F_2 \to \Sigma$ which, when restricted to the interior of F_2, induces a homeomorphism to $\mathcal{D}$ with the three introduced slits. If an annulus F_2 is obtained as a branched triple cover of a disk $\mathbb{D}$, then an elementary Euler characteristic argument (the Riemann-Hurwitz formula; cf. Lemma 8.4.10 below) shows that the number of branched points in p equals 3. Those three branched points give an alternative way to see that $\Delta(\phi) = 3$.

Note that in the above example we verified that $\Delta(\phi) = 3$ using two different topological surfaces F_1 and F_2.

Points with local multiplicity 3/4**.** Consider now the domain given by taking the shaded elementary domains of Figure 8.6 with multiplicity 1 and all others with multiplicity zero. Assume that $\mathbf{x}$ and $\mathbf{y}$ have no components inside the domain (so disregard q in Figure 8.6 for the moment); then the Euler measure is $\frac{1}{2}$ and the point measure is $1\frac{1}{2}$, giving $\mu = 2$. This is compatible with the fact that the domain is the composition of the rectangle

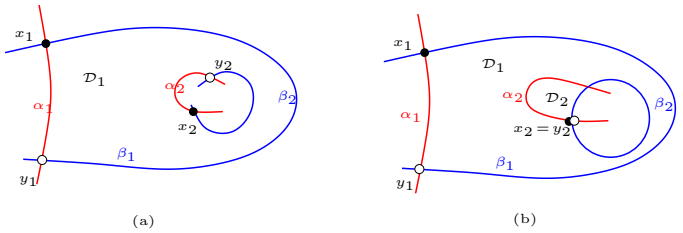


Figure 8.7. The elementary domain $\mathcal{D}_1$ has Maslov index 0, while $\mu(\mathcal{D}_1 \cup \mathcal{D}_2) = 1$.

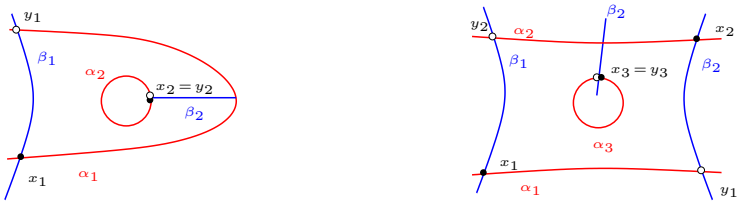


Figure 8.8. An elementary domain with nontrivial π_1.

from x_1x_2 to zy_2 and the bigon from zy_2 to y_1y_2, both of Maslov index one. If $q = x_3 = y_3$ is a common component of $\mathbf{x}$ and $\mathbf{y}$ inside the domain, this increases the point measure, giving $\mu = 4$.

Some non-simply connected examples. The elementary domain $\mathcal{D}_1$ of Figure 8.7(a) has Euler measure -1, and its point measure is 1, hence $\mu = 0$. Adding the small bigon $\mathcal{D}_2$ to it, and assuming $x_2 = y_2$ (as shown on Figure 8.7(b)) we get that the Euler measure is now $-\frac{1}{2}$, the point measure is $1\frac{1}{2}$, hence the Maslov index is equal to 1.

Consider the examples of Figure 8.8. On the left diagram the domain (connecting x_1x_2 and y_1y_2) has Euler measure $-\frac{1}{2}$ and point measure $\frac{3}{2}$, hence its Maslov index is equal to 1. On the right diagram $\overline{n} = 2$, $e = -1$ and hence $\mu = 1$.

Consider the domains in Figure 8.9. On the left we have a bigon, connecting x_1 and y_1; its Maslov index is equal to 1. On the right we depicted the result of a stabilization, performed in the interior of the bigon. This domain connects x_1x_2 with y_1y_2 with $x_2 = y_2$. The point measure of the domain is easy to determine: we need to add 2 (coming from x_2 and y_2) to the point measure of the bigon. The Euler measure is $-\frac{3}{2}$ (see Figure 8.2 above), and so the Maslov index is equal to 1 again.

Exercise 8.3.2. *Consider the four domains pictured in Figure 8.10.*

(a) *Compute the Euler measures and point measures in all four examples. Deduce the Maslov indices.*

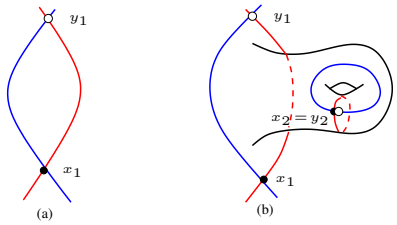


Figure 8.9. The stabilization of a bigon.

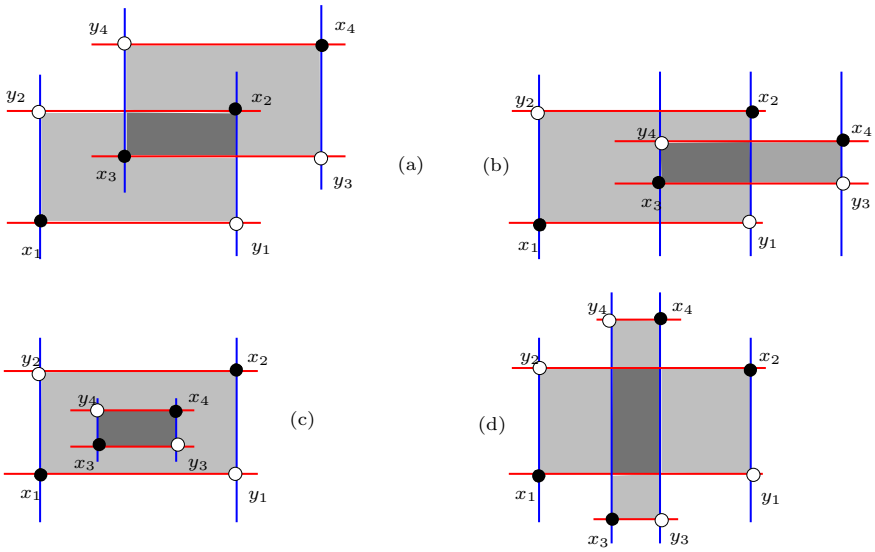


Figure 8.10. Domains by composing two rectangles. The light gray domains come with multiplicity one, while the dark gray ones with multiplicity two.

(b) *In all four examples, decompose the resulting domains as juxtapositions of Maslov index one domains.*

8.4. Constructing decorated surfaces

The aim of the present section is to prove Theorem 8.2.8. The material here is based on [**69, 70**].

With no extra effort, we prove a slightly more general version, working in $\mathrm{Sym}^m(\Sigma)$, when m is not necessarily equal to $g = g(\Sigma)$. To formulate the generalization, we introduce the following definitions:

Definition 8.4.1. *Fix an integer $m \geq 0$. A **generalized Heegaard diagram $\mathcal{H}$ with multiplicity** m is a triple consisting of a surface Σ of genus g, and two sets $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_m\}$ and $\boldsymbol{\beta} = \{\beta_1, \dots, \beta_m\}$ of pairwise*

disjoint, smoothy embedded closed curves in Σ, for some positive integer m. For a generalized Heegaard diagram $\mathcal{H}$, a ***generalized Heegaard state*** *is an m-tuple of points $\mathbf{x} = \{x_1, \dots, x_m\}$ with the property that there is a permutation $\sigma\colon \{1, \dots, m\} \to \{1, \dots, m\}$ so that $x_i \in \alpha_i \cap \beta_{\sigma(i)}$.*

Remark 8.4.2. *While generalized Heeegaard diagrams do not necessarily correspond to closed three-manifolds, the Maslov index formula is still useful; compare* [**58**] *and* [**106**]*. Generalized Heegaard diagrams will also play a prominent role in Chapter 15.*

For a generalized Heegaard diagram, there are two induced tori

$$\mathbb{T}_\alpha = \alpha_1 \times \cdots \times \alpha_m \subset \mathrm{Sym}^m(\Sigma) \qquad \text{and} \qquad \mathbb{T}_\beta = \beta_1 \times \cdots \times \beta_m \subset \mathrm{Sym}^m(\Sigma).$$

Clearly, by the simple extension of Proposition 8.1.1 the generalized Heegaard states correspond to intersection points between $\mathbb{T}_\alpha$ and $\mathbb{T}_\beta$. The notions of homotopy classes of Whitney disks from $\mathbf{x}$ to $\mathbf{y}$, and domains from $\mathbf{x}$ to $\mathbf{y}$ have straightforward adaptation to the generalized settings.

The proof of the following generalization of Theorem 8.2.8 will occupy the rest of this section:

Theorem 8.4.3. *Let $\mathcal{H}$ be a generalized Heegaard diagram, $\mathbf{x}, \mathbf{y}$ be any two (generalized) Heegaard states, and let $\phi \in W(\mathbf{x}, \mathbf{y})$. Then, the intersection number $\Delta(\phi)$ of ϕ with the diagonal $\Delta \subset \mathrm{Sym}^m(\Sigma)$ is computed by*

$$\Delta(\phi) = \overline{n}(\mathcal{S}(\phi)) - e(\mathcal{S}(\phi)). \tag{8.6}$$

We will reduce the proof of the Maslov index formula to particular nice homotopy classes, using the following two results:

Lemma 8.4.4. *Suppose that $\phi_1, \phi_2 \in W(\mathbf{x}, \mathbf{y})$ are two Whitney disks with $\mathcal{S}(\phi_1) = \mathcal{S}(\phi_2)$. Then $\Delta(\phi_1) = \Delta(\phi_2)$.*

Proof. For $i = 1, 2$, choose Whitney disks u_i representing $\phi_i \in W(\mathbf{x}, \mathbf{y})$. Then, $\mathcal{S}(\phi_i)$ gives rise to a pair of one-chains $A_i \subset \mathbb{T}_\alpha$ and $B_i \subset \mathbb{T}_\beta$ with $\partial A_i = \mathbf{y} - \mathbf{x}$ and $\partial B_i = \mathbf{x} - \mathbf{y}$. The condition that $\mathcal{S}(u_1) = \mathcal{S}(u_2)$ ensures that $A_1 - A_2$ is null-homologous in $\mathbb{T}_\alpha$ and $B_1 - B_2$ is null-homologous in $\mathbb{T}_\beta$. Equivalently, since the torus $\mathbb{T}_\alpha$ has Abelian fundamental group, A_1 and A_2 are homotopic as paths in $\mathbb{T}_\alpha$ from $\mathbf{y}$ to $\mathbf{x}$, and B_1 and B_2 are homotopic as paths in $\mathbb{T}_\beta$ from $\mathbf{x}$ to $\mathbf{y}$. Thus, we can construct a sphere S' mapping into $\mathrm{Sym}^m(\Sigma)$ with the following properties. On a disk near the north pole, it is given by u_1; on a disk near the south pole it is given by u_2 (pre-composed with an orientation-reversing identification); and in the remaining annular region, it maps into $\mathbb{T}_\alpha \cup \mathbb{T}_\beta$ (as given by the homotopies between A_1 and A_2 and between B_1 and B_2). Since Δ is disjoint from

$\mathbb{T}_\alpha \cup \mathbb{T}_\beta$, $\Delta(S') = \Delta(u_1) - \Delta(u_2)$. Moreover, since $n_w(u_1) = n_w(u_2)$, and $\{w\} \times \mathrm{Sym}^{m-1}(\Sigma)$ is disjoint from $\mathbb{T}_\alpha \cup \mathbb{T}_\beta$, we have that $n_w(S') = 0$. By Proposition 7.5.4, S' is null-homologous, and so its intersection number with Δ is zero. □

We assume in the next lemma that $m > 1$. The case $m = 1$ will be handled separately.

Lemma 8.4.5. *In a generalized Heegaard diagram with $m > 1$, for any $\phi \in W(\mathbf{x}, \mathbf{y})$ there is another $\phi' \in W(\mathbf{x}, \mathbf{y})$ with $\mathcal{S}(\phi') = \mathcal{S}(\phi) + \Sigma$, where we think of Σ here as the sum of all of its elementary domains. Moreover,*

$$\Delta(\phi') = \Delta(\phi) + 2m + 2g - 2, \tag{8.7}$$

$$e(\phi') = e(\phi) + 2 - 2g \tag{8.8}$$

$$\overline{n}(\phi') = \overline{n}(\phi) + 2m. \tag{8.9}$$

Proof. Let S in $\mathrm{Sym}^m(\Sigma)$ be a sphere with $n_w(S) = 1$; it is not hard to see that $\mathcal{S}(S)$ is the sum of all elementary domains (which we, in a somewhat sloppy manner, write as $\mathcal{S}(S) = \Sigma$). At this point we use $m > 1$.

Form ϕ' by adding a copy of S to ϕ; i.e. if $u\colon \mathbb{D} \to \mathrm{Sym}^m(\Sigma)$ represents ϕ, choose a path from $u(0)$ to S, and let ϕ' be the homotopy class of the map $u'\colon \mathbb{D} \to \mathrm{Sym}^m(\Sigma)$ obtained by joining u to S along the path.

Equation (8.7) follows from the fact that the algebraic intersection number of S with Δ is $2m + 2g - 2$; see Proposition 7.5.7.

By the Gauss-Bonnet theorem, $e(\Sigma) = 2 - 2g$, so Equation (8.8) follows. Equation (8.9) follows immediately from the fact that $\mathbf{x}$ and $\mathbf{y}$ have m components. □

Lemmas 8.4.4 and 8.4.5 reduce the verification of Equations (8.6) to domains with all local multiplicites at least 1. We will focus on that case, representing those homotopy classes by decorated surfaces and maps which have particularly nice behaviour. We formulate this behaviour in the following definitions.

Definition 8.4.6. *Let F be a decorated surface. For generalized Heegaard stated $\mathbf{x}$ and $\mathbf{y}$, a map $f\colon F \to \Sigma$ which is compatible with the decoration induces a two-chain, the **shadow of** f, defined by $\mathcal{S}(f) = \sum_i n_i \cdot \mathcal{D}_i$, where n_i is the local degree of f at a generic point $p \in \mathcal{D}_i$. If the shadow of f is $\mathcal{D} \in D(\mathbf{x}, \mathbf{y})$, we say that f **represents** $\mathcal{D}$.*

This is related to the previous notion of shadows for Whitney disks (Definition 8.1.3): if $f\colon F \to \Sigma$ is as in Definition 8.4.6, and u is the associated

Whitney disk (in the sense of Definition 8.2.11), then $\mathcal{S}(u)$ (in the sense of Definition 8.1.3) coincides with $\mathcal{S}(f)$ (in the sense of Definition 8.4.6).

Definition 8.4.7. *Let F be a decorated surface and suppose $f\colon F \to \Sigma$ is compatible with the decoration. We say that f is **locally branched** if about each $x \in F$, there is an open neighborhood $x \in U \subset F$ and coordinate chart $\Phi\colon X \to U$ with $X \in \{\mathbb{D}, \mathbb{D}^+, \mathbb{D}^{++}\}$, with $\Phi(0) = x$, an open neighborhood $y = f(x) \in V \subset \Sigma$ and a coordinate chart $\Psi\colon \mathbb{D} \to V$ with $\Psi(0) = y$, and a positive integer k so that the following diagram commutes:*

$$\begin{array}{ccc} X & \xrightarrow{\Phi} & U \\ {\scriptstyle z\mapsto z^k}\downarrow & & \downarrow{\scriptstyle f} \\ \mathbb{D} & \xrightarrow{\Psi} & V \end{array}$$

*The number $k = k_x$ is called the branching index at $x \in F$, the points where $k \neq 1$ are called **ramification points** and their images are called **branch points**. The map f is called **simple** if all the branching indices are ≤ 2, and the images of the corner points of f are not branch points; in particular, for a simple map all the corner points have branching index 1.*

*Finally, f is called **sliced** if for each corner point $c \in F$, $f^{-1}(f(c))$ is contained on the boundary of F.*

Note that the above definition allows ramification points on ∂F. For a simple locally branched map $f\colon F \to \Sigma$ let b_f denote the total number of ramification points of f in the interior of F plus half the number of ramification points on ∂F.

Informally, the following lemma ensures that we can represent a non-negative $\mathcal{D} \in D(\mathbf{x}, \mathbf{y})$ by a surface $f\colon F_0 \to \Sigma$, all of whose ramification points lie above the corner points of ϕ.

Lemma 8.4.8. *Let $\mathcal{H}$ be a generalized Heegaard diagram. Given any $\mathbf{x}, \mathbf{y} \in \mathcal{S}(\mathcal{H})$ and $\mathcal{D} \in D(\mathbf{x}, \mathbf{y})$ so that all the local multiplicities of $\mathcal{D}$ are positive, there is a surface-with-corners F_0 and a map $f\colon F_0 \to \Sigma$ with the following properties:*

- *(r-1) f represents $\mathcal{D}$*
- *(r-2) $f(\partial F_0) \subset \boldsymbol{\alpha} \cup \boldsymbol{\beta}$*
- *(r-3) f is a locally branched,*
- *(r-4) The ramification points of f project under f in a one-to-one manner to the points $p \in (\mathbf{x} \cup \mathbf{y}) \setminus (\mathbf{x} \cap \mathbf{y})$.*

(r-5) The preimage under f of $\alpha_i \subset \Sigma$ consists of a connected arc A_i that connects two ramification points, and a disjoint union of circles, each of which maps homeomorphically to α_i.

(r-6) Similarly, the preimage under f of $\beta_i \subset \Sigma$ consists of a connected arc B_i that connects two ramification points, and a disjoint union of circles, each of which maps homeomorphically to β_i.

Proof. Suppose that $\mathcal{D} = \sum_i n_i \mathcal{D}_i$, where $\mathcal{D}_i$ are the elementary domains of the diagram $\mathcal{H}$. We build a surface F_0 and a map $f\colon F_0 \to \Sigma$ that represents $\mathcal{D}$, as follows. To each elementary domain $\mathcal{D}_i$ consider

$$S_1^i, S_2^i, \ldots, S_{n_i}^i$$

where each S_j^i is homeomorphic to $\mathcal{D}_i^c$, together with a map $f_j^i\colon S_j^i \to \mathcal{D}_i$. Glue these pieces together to form a topological space F_0 as follows.

- If $\mathcal{D}_i$ and $\mathcal{D}_j$ share an $\boldsymbol{\alpha}$-arc in their boundaries, for all $1 \leq k \leq \min\{n_i, n_j\}$, glue S_k^i to S_k^j along the corresponding arcs.
- Similarly, if $\mathcal{D}_i$ and $\mathcal{D}_j$ share a $\boldsymbol{\beta}$-arc on their boundaries, then for all $1 \leq k \leq \min\{n_i, n_j\}$, glue $S_{n_i-k+1}^i$ to $S_{n_j-k+1}^i$ along the corresponding arcs.

The map $f_0\colon F_0 \to \Sigma$ is obtained by patching together the functions f_j^i. (This algorithm was introduced in Construction 2.3.11, except there we assumed $\mathcal{D}$ was cornerless.)

Properties (r-1), (r-2) and (r-3) are obvious.

Consider points $p \in \alpha_i \cap \beta_j$, which are of three basic types: either $p \notin (\mathbf{x} \cup \mathbf{y})$, $p \in (\mathbf{x} \cap \mathbf{y})$, or $p \in (\mathbf{x} \cup \mathbf{y}) \setminus (\mathbf{x} \cap \mathbf{y})$. Consider first $p \notin \mathbf{x} \cup \mathbf{y}$. The preimage under f of p is modeled on some number of sheets which project (under f) homemorphically to a neighborhood of p, and some number of half-planes, with the property that $f^{-1}(p) \subset \partial F$. Specifically, if a, b, c, d denote the local multiplicities of $\mathcal{D}$ in the four quadrants around p, then the number of complete sheets is given by $\min(a, b, c, d)$. The number of half-sheets contained in a given quadrant is the local multiplicity of that quadrant minus $\min(a, b, c, d)$. None of the preimages of p is a ramification point. The case where $p \in \mathbf{x} \cap \mathbf{y}$ works the same.

It remains to consider $p \in (\mathbf{x} \cup \mathbf{y}) \setminus (\mathbf{x} \cap \mathbf{y})$. Label local multiplicites around p as in Figure 8.12; we will drop the subscripts p from these local multiplicities. There are two distinct cases, acording to whether $p \in \mathbf{x} \setminus \mathbf{y}$ or $\mathbf{y} \setminus \mathbf{x}$, or, equivalently $a + c - b - d$ is -1 or $+1$. We describe in detail the case where $a + c - b - d = -1$, so that $p \in \mathbf{x} \setminus \mathbf{y}$. After rotating, we can assume that $a = \min(a, b, c, d)$. Assume further that $b >$

a. Let $\{A_1, \dots, A_a\}$, $\{B_1, \dots, B_b\}$, $\{C_1, \dots, C_c\}$, $\{D_1, \dots, D_d\}$ denote the sheets in F_0 that project to the quadrants with multiplicity a, b, c, and d respectively. In our construction of F_0, for all $i \in \{b-a, \dots, b\}$, B_i is glued (along an edge containing the preimage of p) to C_i, which in turn is glued to D_{d-c+i}, which in turn is glued to A_{d-c+i}, which in turn is glued to $B_{b-a+d-c+i} = B_{i+1}$. Thus, the preimage of p in the union of $\{A_1, \dots, A_a\} \cup \{B_{b-a}, \dots, B_b\} \cup \{C_{b-a}, \dots, C_b\} \cup \{D_1, \dots D_{d-c+b}\}$ consists of a single branch point; while the quadrants $C_{b+1}, \dots, C_c$ are glued to $D_{a+2}, \dots, D_d$ to form half-planes; and the quadrants $B_1, \dots B_{b-a-1}$ are glued to $C_1, \dots, C_{b-a-1}$ to form further half-planes. In particular, none of the preimages of p in these boundary components is a ramification point, and hence Property (r-4) follows. When $b = a$, then $d > a$, which works similarly. The cases where $a + b - b - d = 1$ work similarly, as well.

Follow some fixed β_j along an orientation. Let m denote the minimum of all the local multiplicities to the left of α_i, and n denote the minimum of all the local multiplicities to the right of α_i. We claim that F_0 contains $|m - n|$ disjoint boundary components that map homeomorphically to β_j. For example, suppose β_j is oriented bottom to top in Figure 8.12, and that $n > m$; then $a > b$ and $c > d$. Then, the quadrants A_i are glued to D_i for $i = 1, \dots, n - m$, and these are not glued to any of the B_k or C_k. The remaining copies of A_i and D_i, which do not contain the ramification points constructed above or the boundary points are glued together with B_{i-n+m} and C_{i-n+m} to form planes. Property (r-6) follows, and Property (r-5) follows similarly. □

Note that F_0 is not a decorated surface, in the sense of Definition 8.2.9. For example, $f|_{\partial F}^{-1}(\alpha_i)$ need not be connected; also, for components $p \in \alpha_i \cap \beta_j$ with $p \in \mathbf{x} \cap \mathbf{y}$, the component $(f|_{\partial F})^{-1}(\alpha_i)$ and $(f|_{\partial F})^{-1}(\beta_j)$ are disjoint. We find it convenient to modify the above construction, to get a sliced surface compatible with a given domain, following [**70**]:

Lemma 8.4.9. *(Lemma 4.1 [**70**]) Given any $\mathbf{x}, \mathbf{y}$ and $\mathcal{D} \in D(\mathbf{x}, \mathbf{y})$ so that all the local multiplicities of $\mathcal{D}$ are positive, there is a connected decorated surface F and a compatible map $f \colon F \to \Sigma$ with the following properties (some of which are defined in Definition 8.4.7)*

- *f represents $\mathcal{D}$,*
- *the map f is locally branched,*
- *f is simple,*
- *f is sliced.*

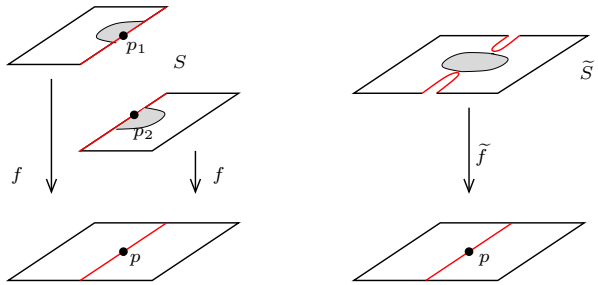


Figure 8.11. Gluing intervals to modify f. Forming the boundary connected sum of S along two f-preimages p_1 and p_2 of p, we obtain a new surface-with-boundary $\widetilde{S}$ and a map $\widetilde{f}$ with two branch points on the boundary.

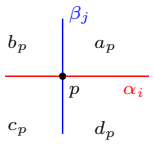


Figure 8.12. Local multiplicities of a domain at an intersection point $p \in \alpha_i \cap \beta_j$.

Proof. For the proof of the lemma, we modify the surface (and the map) $f\colon F_0 \to \Sigma$ constructed in Lemma 8.4.8. The modification is done by a repeated boundary connected sum operation, described presently.

Suppose that S is an oriented surface-with-boundary, fix two points p_1 and p_2 on the boundary, open intervals I_1 and I_2 also in the boundary, containing p_1 and p_2, and an orientation-reversing identification of I_1 and I_2. With this data given, we can identify I_1 with I_2 to obtain a new oriented surface $\widetilde{S}$. Suppose moreover that S is also equipped with a map $f\colon S \to \Sigma$, and f maps I_1 and I_2 diffeomorphically to the same interval in Σ. Then, the map f extends smoothly to a map $\widetilde{f}\colon \widetilde{S} \to \Sigma$. This construction introduces two new branch points in $\widetilde{f}$, which are the endpoints of $\overline{I}_1$ (which in turn are identified with the endpoints of $\overline{I}_2$). See 8.11.

With these preliminaries in place, we explain how to modify $f\colon F_0 \to \Sigma$ in a neighborhood of the preimage of $\mathbf{x} \cup \mathbf{y}$. There are two cases, according to whether or not $p \in \mathbf{x} \cap \mathbf{y}$; i.e. whether or not the sums of local multiplicities in opposing domains are equal.

Consider first corners when $p \in \mathbf{x} \cap \mathbf{y}$; i.e. where the local multiplicities a_p, b_p, c_p, and d_p around p satisfy $a_p + c_p = b_p + d_p$. (We use here conventions from Figure 2.9, which we have included also as Figure 8.12 for the reader's convenience.) The preimage of a neighborhood of p consists of some number of half-planes and a positive number of planes. For exactly

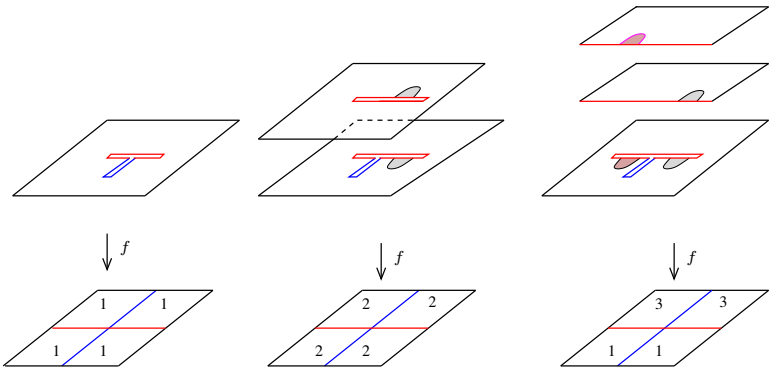


Figure 8.13. Constructing the surface over points in $(\mathbf{x} \cap \mathbf{y})$. Introduce slits and identify pairs of intervals in the boundary as shown by the second and third pictures. After such an identification, the two half-disks shown are glued along their boundary to form a complete disk, via boundary connected sum.

one of these planar components, we introduce slits near the preimage of p in both the α and the β-directions, as shown in the first picture in Figure 8.13; we call this the *distinguished plane*. To ensure that f is sliced, we introduce a short α- or β-slice locally on each additional planar sheet, as shown by the second picture of Figure 8.13. Glue (using the boundary connected sum as explained above) each (sliced) plane or half-plane to the distinguished plane along sufficiently short arcs, as instructed by the second and third pictures of Figure 8.13. This is how f is modified above a neighborhood of $p \in \mathbf{x} \cap \mathbf{y}$.

Consider next the case $p \in (\mathbf{x} \cup \mathbf{y}) \setminus (\mathbf{x} \cap \mathbf{y})$, i.e. following notation from Definition 2.4.4 $a_p + c_p = b_p + d_p \pm 1$. According to Lemma 8.4.8, over p there is a single ramification point, call it p_0; let k denote its ramification index. It is straightforward to see that k is odd and ≥ 3. Insert small cuts along the $\boldsymbol{\alpha}$-curves if $p \in \mathbf{x}$ or along the $\boldsymbol{\beta}$-curves if $p \in \mathbf{y}$ at such a point. This can be done so that the new map has no ramification points over p; the single order k ramification point is traded for $(k-1)/2$ order 2 branchpoints and a single corner point. See Figure 8.14 for an example. Finally, when the projection of p is a component of $\mathbf{x}$, use the boundary connected sum operation to attach the remaining α- and β- half-planes to a neighborhood of the corner point. (The choice of connected sum when p is a component of $\mathbf{x}$, rather than $\mathbf{y}$, is arbitrary; it is important to connect each component of the boundaries of the α- and β-half-planes to the boundary component containing the corner points exactly once, to ensure that $f^{-1}(\alpha_i) \cap \partial F$ is connected.)

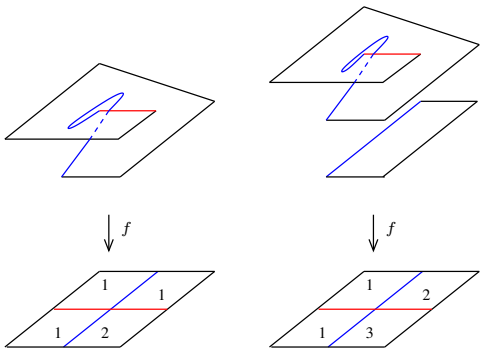


Figure 8.14. Constructing the surface over points in $(\mathbf{x} \cup \mathbf{y}) \setminus (\mathbf{x} \cap \mathbf{y})$. Insert slits to the constructed surface F_0 as indicated here. We have drawn here a ramification points with order $k = 5$.

The above procedure provides a surface F, together with a map $f\colon F \to \Sigma$, which is a simple branched cover which is also sliced. Moreover, the boundary connected sums of the half-plane to the the distinguished plane (in the case where $p \in (\mathbf{x} \cap \mathbf{y})$) or to the neighborhood of the corner (in the case where $p \in \mathbf{y} \setminus (\mathbf{x} \cap \mathbf{y})$) ensure that $f^{-1}(\alpha_i) \cap \partial F$ is connected; similarly, $f^{-1}(\beta_j) \cap \partial F$ is connected. Thus, F now satisfies the hypotheses of Definition 8.2.9.

We argue that F is connected, as follows. Choose any $p \in \mathbf{x} \cap \mathbf{y}$. Note that each component of F meets $f^{-1}(p)$; and each element in $f^{-1}(p)$ is contained in the same boundary component of F. □

Lemma 8.4.10. *Given $m > 1$, any connected, decorated surface F with non-empty boundary can be presented as a simple m-fold branched cover $\pi\colon F \to \mathbb{D}$, where the latter is thought of as a decorated surface whose boundary is decomposed as an A-arc and a B-arc. Moreover, the Euler characteristic $\chi(F)$, the degree m of the branched cover, and the number b_π of branched points of π are related by*

$$\chi(F) = m - b_\pi. \tag{8.10}$$

Proof. Suppose that F is of genus zero and has a single path component. Such a decorated surface is a polygon with $2m$ sides. This can be realized explicitly as a simple m-fold branched cover of $\mathbb{D}$.

Given decorated surfaces F_1 and F_2, and compatible branched covers $\pi_1\colon F_1 \to \mathbb{D}$ and $\pi_2\colon F_2 \to \mathbb{D}$ of degree m_1 and m_2 respectively, we can find a new branched cover $\pi_3\colon F_1 \# F_2 \to \mathbb{D}$, by forming the connected sum, and introducing two branch points in the connected sum region to construct an $(m_1 + m_2)$-fold cover π_3.

Similarly, if $\pi\colon F \to \mathbb{D}$ has multiplicity $m > 1$, then we can find a new branched cover $\pi\colon F\#T^2 \to \mathbb{D}$, where $F\#T^2$ is thought of as F with a one-handle attached, again by introducing two branched points in the one-handle.

Since the general connected, decorated surface F can be realized as the connected sum of a collection of polygons, with some additional genus, the above method can be used to give the desired branched cover in the general case.

Equation (8.10), which can be thought of as a special case of the *Riemann-Hurwitz formula*, can be seen as follows. Choose a CW-complex structure on $\mathbb{D}$ so that the branched points are contained in the 0-skeleton. Thus, those cells have $m-1$ preimages, but all other cells have m preimages. Using this CW-complex structure to compute the Euler characteristic $\chi(F)$, it follows readily that if there are b_π branch points, then

$$\chi(F) = m(\chi(\mathbb{D}) - b_\pi) + (m-1)b_\pi = m - b_\pi.$$

□

Given $\mathcal{D} \in D(\mathbf{x}, \mathbf{y})$ with only positive local multiplicities, let F be the decorated surface and $f\colon F \to \Sigma$ be the map constructed in Lemma 8.4.9. Lemma 8.4.10 gives an m-fold branched cover $\pi\colon F \to \mathbb{D}$.

Consider the map $G\colon F \to \Sigma \times \mathbb{D}$ obtained by $G(z) = (f(z), \pi(z))$. By a generic choice of π the map G can be made to be an embedding away from a finite collection of transverse double points. Let s_+ and s_- denote the number of positive and negative double points of G. It is interesting to note that if both f and π are holomorphic, then $s_- = 0$.

The verification of Equation (8.4) will proceed by comparing the Euler characteristic calculation for F from Equation (8.10), obtained from the projection map π, with another one, obtained from the projection map f, which we give as follows.

Lemma 8.4.11. *Let F be a decorated surface, and let $f\colon F \to \Sigma$ be a map that is compatible with the decoration, and $\pi\colon F \to \mathbb{D}$ be a simple m-fold branched m-fold cover of the disk. Suppose moreover that $G = f \times \pi$ is an embedding away from a finite collection of transverse double points; and let s_+ and s_- as defined above. Then, if $u\colon \mathbb{D} \to \mathrm{Sym}^m(\Sigma)$ is the map associated to (f, π), then*

$$\Delta(u) = b_\pi + 2(s_+ - s_-). \tag{8.11}$$

Proof. Clearly, $u^{-1}(\Delta)$ consists of two kinds of points: the branch points of π, and the double-points of G.

Each branch point of π gives rise to a single (positive) intersection point of $u(\mathbb{D})$ with the diagonal Δ. Each double-point of G also gives rise to an intersection point with Δ: a self-intersection of the image of G projects to a point in $\mathbb{D}$ with $m-1$ preimages, and each self-intersection point counts with multiplicity two. A local calculation (such as the one from the proof of Proposition 7.3.6) shows that such a self-intersection of the image of G gives a double intersection of $u(\mathbb{D})$ with the diagonal. For a positive double point of $\mathrm{Im}(G)$ the intersection multiplicity between u and Δ is 2; while for a negative double point, it is -2. The proof of the lemma is complete. □

Lemma 8.4.12. *Let F be a decorated surface and $f\colon F \to \Sigma$ be a simple, locally branched map. Then,*

$$\chi(F) = e(\mathcal{S}(f)) + \frac{1}{2}m - b_f. \tag{8.12}$$

Proof. Equip F with a metric for which the boundary arcs A_i and B_i are geodesics, each meeting at 90° angles at the corner points. The Euler measure $e(F)$ of F is equal to the integral of the curvature of this metric (divided by 2π). Since all the corners are 90°, the Gauss-Bonnet formula shows that

$$\chi(F) = e(F) + \frac{1}{2}m. \tag{8.13}$$

Endow Σ with a metric for which all the α_i and β_j are geodesics, which meet at 90° angles. Pulling back the curvature form from Σ and integrating over F computes $e(\mathcal{S}(f))$. In fact, the Euler measure of F can be computed by integrating the pull-back of the curvature form on Σ, and correcting at the branch points, giving

$$e(F) = e(\mathcal{S}(f)) - b_f. \tag{8.14}$$

Combining Equation (8.13) with Equation (8.14) we get the claimed equality. □

A final piece of information is provided by the next proposition, whose proof is given in a sequence of lemmas:

Proposition 8.4.13. *If F is a decorated surface, $f\colon F \to \Sigma$ is a locally branched simple map with b_f ramification points (counted with $1/2$ if they occur on the boundary), and $\pi\colon F \to \mathbb{D}$ is a branched m-fold cover, then*

$$b_f = \overline{n}(\mathcal{S}(f)) - \frac{1}{2}m - 2(s_+ - s_-), \tag{8.15}$$

where s_+ and s_- count double-points of $G = f \times \pi$ (with signs).

The proposition is proved after establishing several lemmas, which in turn use the following notation. Fix a conformal equivalence $\mathbb{D}\setminus\{\pm i\}\simeq[0,1]\times\mathbb{R}$ (identifying the ends $\pm i$ with the ends $[0,1]\times\{\pm\infty\}$). Composing $\pi\times f\colon F\to\mathbb{D}\times\Sigma$ with this conformal equivalence on the first factor, we have an induced map

$$G^\circ\colon F\setminus\{p_i,q_i\}_{i=1}^m\to\Sigma\times[0,1]\times\mathbb{R},$$

obtained by composing with this identification. Using coordinates (z,s,t) on $\Sigma\times[0,1]\times\mathbb{R}$, the branch points of f can be identified with those points where the vector field $\frac{\partial}{\partial t}$ is tangent to the image of G°.

For a real number r, consider the map $\tau_r\colon\Sigma\times[0,1]\times\mathbb{R}\to\Sigma\times[0,1]\times\mathbb{R}$ defined by translation in the $\mathbb{R}$-direction, that is, $\tau_r(p,s,t)=(p,s,t+r)$. Equation (8.15) will be proved by comparing the number of intersection points of G° with $\tau_r(G^\circ)$ when r goes from being very small to very large. (Our notation here does not distinguish the map G° from the image of $F\setminus\{p_i,q_i\}_{i=1}^m$ under the map G°.)

To formulate an invariant intersection number, recall that $F\setminus\{p_i,q_i\}$ is oriented; so is the ambient space $\mathbb{D}\times[0,1]\times\Sigma$. Thus, at each transverse intersection point of G° with $\tau_r(G^\circ)$, there is an local transverse intersection number (with values in $\{\pm1\}$) of G° and $\tau_r(G^\circ)$. At a point $p\in G^\circ\cap\tau_r(G^\circ)$, the local intersection number $\epsilon(p)\in\{\pm1\}$ is obtained by comparing the orientation on the vector space $T_p(\mathbb{D}\times[0,1]\times\Sigma)$ induced from the orientation on $\mathbb{D}\times[0,1]\times\Sigma$ with the orientation on $T_pG^\circ\oplus T_p(\tau_r(G^\circ))\cong T_p(\mathbb{D}\times[0,1]\times\Sigma)$ induced from the orientations on G° and $\tau_r(G^\circ)$. The sum of these local intersection points is denoted $\#(G^\circ\cap\tau_r(G^\circ)$.

The boundary correction is formulated by considering intersection points

$$\partial(G^\circ)\cap\tau_r(\partial G^\circ)\subset(\{\boldsymbol{\alpha}\}\times0\times\mathbb{R})\cup(\boldsymbol{\beta}\times\{1\}\times\mathbb{R}).$$

One could orient the two-manifold $(\{\boldsymbol{\alpha}\}\times0\times\mathbb{R})\cup(\boldsymbol{\beta}\times\{1\}\times\mathbb{R})$ and consider the algebraic intersection number in that space; but this is not what we do. Rather we count each intersection point $p\in\partial(G^\circ)\cap\tau_r(\partial G^\circ)$ with ±1 computed by thinking of p as an intersection point in $G^\circ\cap\tau_r(G^\circ)$. We denote this boundary intersection number

$$\#_b((\partial G^\circ)\cap\tau_r(\partial G^\circ).$$

We now have a *corrected intersection number*

$$\widetilde{\#}(G^\circ\cap\tau_r(G^\circ))=\#(G^\circ\cap\tau_r(G^\circ))+\frac{1}{2}\#_b((\partial G^\circ)\cap\tau_r(\partial G^\circ). \tag{8.16}$$

Lemma 8.4.14. *For generic r, G° and $\tau_r(G^\circ)$ meet transversely. Moreover, if r_1 and r_2 are generic, then the corrected intersection numbers coincide; i.e.*

$$\widetilde{\#}(G^\circ \cap \tau_{r_1}(G^\circ)) = \widetilde{\#}(G^\circ \cap \tau_{r_2}(G^\circ)). \tag{8.17}$$

Proof. Consider the moduli spaces

$$M_1 = \{p, q \in \partial(G^\circ), s \in [r_1, r_2] \big| q = \tau_s(p)\},$$
$$M_2 = \{p, q \in (G^\circ)^{\text{int}}, s \in [r_1, r_2] \big| q = \tau_s(p)\}.$$

For generic choices, the spaces M_1 and M_2 are oriented one-dimensional manifolds.

The moduli space M_1 is a compact one-manifold, with boundary identified with $(\partial G^\circ) \cap \tau_{r_1}(\partial G^\circ)) \cup (\partial G^\circ) \cap \tau_{r_2}(\partial G^\circ))$. The orientation on M_1 inherited from G°, $\mathbb{D}$, and Σ has the property that these boundary points are counted as

$$\#_b(\partial G^\circ) \cap \tau_{r_1}(\partial G^\circ)) - \#_b(\partial G^\circ) \cap \tau_{r_2}(\partial G^\circ)).$$

However, this orientation convention does not give a consistent orientation on the one manifold M_1, in the sense that there is a finite set S of special points $p \in M_1$ where the orientation on M_1 and the two arcs of M_1 either both point into or out of p. Thus, counting points in the oriented boundary of M_1, we find that

$$\#_b(\partial G^\circ) \cap \tau_{r_1}(\partial G^\circ)) - \#_b(\partial G^\circ) \cap \tau_{r_2}(\partial G^\circ)) = 2\#S. \tag{8.18}$$

The moduli space M_1 has two kinds of boundary points: those where $s = r_1$, those where $s = r_2$, and those where (p, q, s) arises as the limit of two different paths $(x_1(t), y_1(t), s(t))$ and $(x_2(t), y_2(t), s(t))$. It also has finitely many ends E corresponding to paths $(x(t), y(t), s(t))$ where $x(t)$ limits to ∂G°. Counting ends with sign, we find

$$\#(G^\circ \cap \tau_{r_1}) - \#(G^\circ \cap \tau_{r_2}) = -\#E. \tag{8.19}$$

We claim that the signed count of ends E coincides with the signed count of special points S appearing in the moduli space M_2 described above. This can be seen explicitly by a doubling argument: double G° across its boundary, while doubling $\mathbb{D}$. The ends E correspond to points where intersection points in G° collide with intersection points on their virtual pairs $(G^\circ)'$; the local intersection number of the given intersection point coincides with the local intersection number of its virtual pair. The ends S correspond to pairs of intersection points on the boundary with the same local intersection number colliding. See Figure 8.15 for a picture. Thus, Equations (8.18) and (8.19) give the desired equality. □

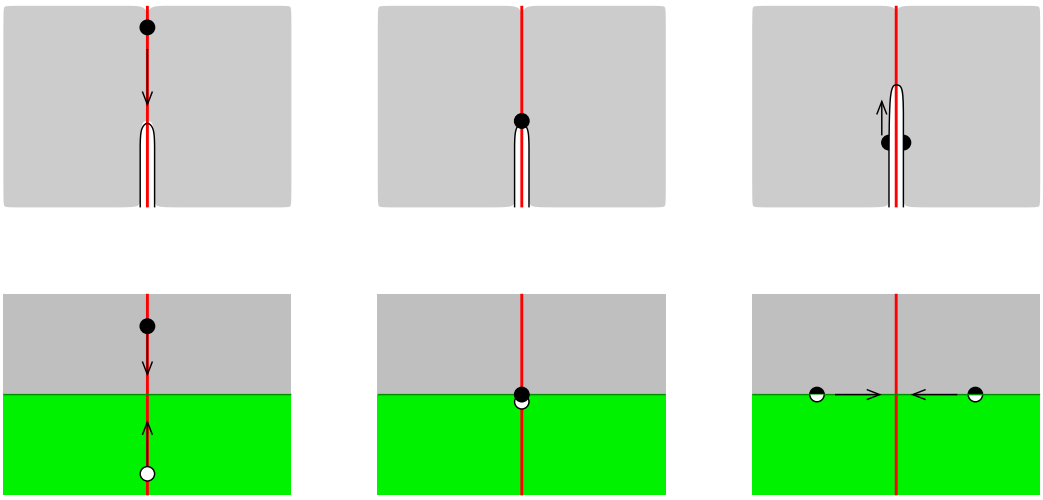


Figure 8.15. Trading an intersection point in the interior for a pair of intersection points on the boundary (top row): at the left is a one-parameter family of (black) double-points that wander out to the boundary (along the ends E), pictured in the middle. These can be continued as pairs of double points on the boundary (right). When doubling the above picture, we obtain the three pictures below. The green region represents the double of the original surface F, containing a one-parameter family of virtual (white) double-points.

The proof of Proposition 8.4.13 will follow from comparing the large r and the small r behaviour of the intersection number.

Lemma 8.4.15. *For all $\epsilon > 0$ sufficiently small, G° and $\tau_\epsilon(G^\circ)$ meet transversely, and*

$$b_f = \#(G^\circ \cap \tau_\epsilon(G^\circ)) - 2(s_+ - s_-). \tag{8.20}$$

Proof. In the special case where G° is an embedding, the tangecies of the vector field $\frac{\partial}{\partial t}$ on G° can be identified with the intersection points $G^\circ \cap \tau_r(G^\circ)$ for sufficiently small r. In general, G° is only an immersion, with s_+ positive and s_- negative double points. Near each double point, the translate $\tau_r(G^\circ)$ intersects G° in 2 points (and the signs of these intersections are determined by the sign of the double point), so for sufficiently small r, the self-intersection points of G° contribute to the intersection number, but not to the number of tangencies to $\frac{\partial}{\partial t}$. The formula then follows from a local count near each self-intersection. □

Lemma 8.4.16. *If R is sufficiently large, then*

$$\widetilde{\#}(G^\circ \cap \tau_R(G^\circ)) = \overline{n}(\mathcal{S}(\phi)) - \frac{1}{2}m. \tag{8.21}$$

Proof. Each corner point $c \in F$ has a neighborhood U_c which, under $\pi\colon F \to \mathbb{D}$, is identified with $(-\infty, -T)\times[0,1]$ (if the corner corresponds to a component of $\mathbf{x}$), or $(T, \infty)\times[0,1]$ (if the corner corresponds to a component of $\mathbf{y}$). Under the map f, the neighborhood is identified with a quadrant near an intersection point of $\alpha_i \cap \beta_j$. Since f is a local homeomorphism on these neighborhoods, it is clear for $(p,q,s) \in M_1 \cup M_2$, for each corner point c and at most one of p or q is in U_c. Moreover, if R is sufficiently large, then exactly one of p or q is in U_c.

Assume for simplicity that the map

$$G^\circ\colon F \setminus \{p_i, q_i\}_{i=1}^m \to \Sigma \times [0,1] \times \mathbb{R},$$

has the property that the restriction of G° to U_c is constant (rather than merely being asymptotic to constant). This can be arranged by a homotopy, and hence it can be done without affecting $\widetilde{\#}(G^\circ \cap \tau_R(G^\circ))$.

With these preliminaries, Equation (8.21) follows quickly from the structure of F above the corner points. Choose an $x \in \mathbf{x} \setminus \mathbf{y}$. The preimage of x under f consists of the initial corner over x, and a union of preimages in edges. The number of such edges is $2\overline{n}_x(\mathcal{D}) - \frac{1}{2}$. Each preimage in the edge contributes an intersection point in $\partial(G^\circ) \cap \tau_R(\partial G^\circ)$; and each preimage contributes $1/2$ to the count in $\widetilde{\#}((G^\circ) \cap \tau_R(G^\circ))$. A similar accounting occurs at $y \in \mathbf{y} \setminus \mathbf{x}$ and indeed for each point $\mathbf{x} \cap \mathbf{y}$. Summing contributions over all the components of $\mathbf{x}$ and $\mathbf{y}$, we obtain Equation (8.21). □

Proof. [Proof of Proposition 8.4.13] The proof is an immediate consequence of Lemmas 8.4.15, 8.4.14, and 8.4.16. □

Proof. [Proof of Theorem 8.4.3] Consider the cases where $m > 1$. By Lemmas 8.4.4 and 8.4.5, it suffices to prove the theorem for positive domains. Let (f, F) be the decorated surface constructed in Lemma 8.4.9.

Applying Equations (8.11), (8.10), (8.12), and (8.15) in this order, we find

$$\begin{aligned}\Delta(u) &= b_\pi + 2(s_+ - s_-)\\ &= m - \chi(F) + 2(s_+ - s_+)\\ &= \frac{1}{2}m - e(\mathcal{S}(\mathcal{D})) + b_f + 2(s_+ - s_+)\\ &= \overline{n}(\mathcal{S}(u)) - e(\mathcal{S}(u)),\end{aligned}$$

establishing Equation (8.4) when $m > 1$.

When $m = 1$, the diagonal is empty, so Equation (8.4) reduces to showing that for any Whitney disk u from x to y in the generalized Heegaard

diagram $(\Sigma, \{\alpha_1\}, \{\beta_1\})$, we have that $\overline{n}(\mathcal{S}(u)) = e(\mathcal{S}(u))$. We establish this by reducing to the case $m = 2$, as follows. Consider the generalized Heegaard diagram $(\Sigma, \{\alpha_1, \alpha_2\}, \{\beta_1, \beta_2\})$ where α_2 and β_2 are two curves chosen to intersect in some point $p \in \alpha_2 \cap \beta_2$. We can construct then a Whitney disk u' from $\{x, p\}$ to $\{y, p\}$ obtained by adding a constant map at p. We claim that

$$\Delta(u') = -2n_p(u), \quad e(u') = e(u), \quad \overline{n}(u') = \overline{n}(u).$$

Combining the above observations with Equation (8.4) for u' now gives the desired $\overline{n}(\mathcal{S}(u)) = e(\mathcal{S}(u))$. □

8.5. Computing the Maslov index

In this section we prove Theorem 8.2.7 and deduce Theorem 8.2.6. More precisely, we prove its obvious generalizations:

Theorem 8.5.1. *(Rasmussen* [**112**]*) Let* $\mathbf{x}, \mathbf{y}$ *be any two generalized Heegaard states in a generalized Heegaard diagram* $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$. *Given* $\phi \in W(\mathbf{x}, \mathbf{y})$, *the Maslov index can be computed by*

$$\mu(\phi) = 2e(\mathcal{S}(\phi)) + \Delta(\phi). \tag{8.22}$$

For the proof, we will consider the following general situation. Suppose L_0 and L_1 are two Lagrangian submanifolds of a symplectic manifold (M, ω), and suppose that we have an almost-complex structure J that is compatible with ω, with the property that at each intersection point $\mathbf{x} \in L_0 \cap L_1$ we have $JT_{\mathbf{x}}L_0 = T_{\mathbf{x}}L_1$. Under these hypotheses, the determinant line $\det_{\mathbb{R}}(T_{\mathbf{x}}L_0)$ thought of as a one-dimensional real subspace of the complex line $\det_{\mathbb{C}}(T_{\mathbf{x}}M)$, is i^n times the real line $\det_{\mathbb{R}}(T_{\mathbf{x}}L_1)$ (when $\dim(M) = 2n$); and therefore the squares of the determinant lines $(\det_{\mathbb{R}}(T_{\mathbf{x}}L_0) \otimes \det_{\mathbb{R}}(T_{\mathbf{x}}L_0))$ and $(\det_{\mathbb{R}}(T_{\mathbf{x}}L_1) \otimes \det_{\mathbb{R}}(T_{\mathbf{x}}L_1))$ give the same real line in $\det_{\mathbb{C}}(T_{\mathbf{x}}M) \otimes \det_{\mathbb{C}}(T_{\mathbf{x}}(M))$.

Thus, the bundles $\det_{\mathbb{R}}(L_0) \otimes \det_{\mathbb{R}}(L_0)$ and $\det_{\mathbb{R}}(L_1) \otimes \det_{\mathbb{R}}(L_1)$, thought of as subbundles of $\det_{\mathbb{C}}(TM) \otimes \det_{\mathbb{C}}(TM)$, are identified over $L_0 \cap L_1$. When L_0 and L_1 are orientable, this identification gives a simultaneous trivialization of $\det_{\mathbb{C}}(TM) \otimes \det_{\mathbb{C}}(TM) = \det_{\mathbb{C}}(TM)^{\otimes 2}$ over $L_0 \cup L_1$, and hence a well-defined relative two-dimensional cohomology class $c_1(\det_{\mathbb{C}}(TM)^{\otimes 2}, L_0 \cup L_1) \in H^2(M, L_0 \cup L_1; \mathbb{Z})$, which is the first Chern class of the complex line bundle $\det_{\mathbb{C}}(TM)^{\otimes 2}$, relative to its compatible trivialization on $L_0 \cup L_1$ induced by orientation $\det(L_0) \otimes \det(L_0)$ over L_0 and $(-1)^n \det(L_1) \otimes \det(L_1)$ over L_1.

Concretely, given an oriented surface-with-boundary F and a map

$$f\colon (F, \partial F) \to (M, L_0 \cup L_1),$$

the relative Chern number $\langle c_1(\det_{\mathbb{C}}(TM)^{\otimes 2}, L_0 \cup L_1), [F]\rangle$ measures the number of transverse zeros of any section σ of $f^* \det_{\mathbb{C}}(TM)^{\otimes 2}$ with the property that $\sigma|_{\partial F}$ agrees with the fixed trivialization of $L_0 \cup L_1$.

Proposition 8.5.2. *Suppose that L_0 and L_1 are orientable Lagrangian submanifolds in a symplectic manifold (M, ω). Suppose that there is an almost-complex structure J compatible with ω for which $JT_{\mathbf{x}}L_0 = T_{\mathbf{x}}L_1$ for all $\mathbf{x} \in L_0 \cap L_1$. Then for any Whitney disk $\phi \in W(\mathbf{x}, \mathbf{y})$, the Maslov index of ϕ agrees with the evaluation of the above relative first Chern class of $\det_{\mathbb{C}}(M) \otimes_{\mathbb{C}} \det_{\mathbb{C}}(M)$ on the relative homology class determined by ϕ.*

Proof. We use the formulation of the Maslov index from Definition 6.5.5. Fix preferred paths $\Lambda_{\mathbf{x}}$ from $T_{\mathbf{x}}L_0$ to $T_{\mathbf{x}}L_1$ and $\Lambda_{\mathbf{y}}$ from $T_{\mathbf{y}}L_0$ to $T_{\mathbf{y}}L_1$ (in the sense of Definition 6.5.2). Fix a representative u of the Whitney disk ϕ. There are two paths of Lagrangian subspaces induced along the edges of u: $\mathfrak{L}_{s,0}$ is given by $u^*(T_{u(s,0)}L_0)$ and $\mathcal{L}_{s,1}$ is given by $u^*(T_{u(s,1)}L_1)$. Trivializing the symplectic bundle $u^*(TM)$, we can view the four paths of Lagrangian subspaces Λ_x, $\mathfrak{L}_{s,0}$, $\Lambda_{\mathbf{y}}$, and $\mathfrak{L}_{s,1}$ as four paths $\lambda_{\mathbf{x}}$, ℓ_0, $\lambda_{\mathbf{y}}$, and ℓ_1 in $U(n)/O(n)$. According to the formulation from Definition 6.5.5, the Maslov index is computed as the degree of the map from S^1 to S^1 obtained by applying $\det^2\colon U(n)/O(n) \to S^1$ to the closed loop $\lambda_{\mathbf{x}} * \ell_1 * \lambda_{\mathbf{y}}^{-1} * \ell_0^{-1}$. (The inverses appearing in this expression reflect the boundary orientation of the square.) Note that the constituent paths $\det^2(\lambda_{\mathbf{x}})$, $\det^2(\ell_1)$, $\det^2(\lambda_{\mathbf{y}}^{-1})$, $\det^2(\ell_0^{-1})$ are not closed paths: $\det^2(\lambda_{\mathbf{x}})$ is a path from some point $\xi \in S^1$ to $(-1)^n \cdot \xi$, $\det^2(\lambda_{\mathbf{y}})$ is a path from $\eta \in S^1$ to $(-1)^n \cdot \eta$.

In order to work with closed paths, we postcompose all paths with the squaring map $s\colon S^1 \to S^1$, $s(z) = z^2$. We claim that if λ is any preferred path from Λ to $J \cdot \Lambda$, then $s \circ \det^2 \lambda$ has degree n. We can see this concretely: the for the Lagrangian $\mathbb{R}^n \subset \mathbb{C}^n$, the path $\lambda\colon [0,1] \to \mathcal{L}Gr(n)$ given by $t \to e^{it}\mathbb{R}^n$ is a preferred path from $\mathbb{R}^n$ to $J\mathbb{R}^n$, and $s \circ \det^2 \circ \lambda$ has degree n.

Applying $\deg(s \circ \det^2)$ to the closed loop, $\lambda_{\mathbf{x}} * \ell_1 * \lambda_{\mathbf{y}}^{-1} * \ell_0^{-1}$, we find that $\det^2(\lambda_{\mathbf{x}} * \ell_1 * \lambda_{\mathbf{y}}^{-1} * \ell_0^{-1})$ agrees with the degree of the closed loop $\det^2(\ell_1) * (-1)^n \det^2(\ell_0^{-1})$, which in turn is the relative first Chern class. (Compare Proposition 4.5.8; the multiple of $(-1)^n$ is explained before the statement of the present proposition.)

□

Lemma 8.5.3. *Suppose that* $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$ *is a pointed Heegaard diagram, and suppose* Σ *is equipped with a complex structure* $\mathfrak{j}$ *with the property that at each* $x \in \alpha_i \cap \beta_j$, $\mathfrak{j}T_x\alpha_i = T_x\beta_j$. *If* $f\colon (F, \partial F) \to (\Sigma, \boldsymbol{\alpha} \cup \boldsymbol{\beta})$ *is a smooth map, then the evaluation of* $c_1(\det(T\Sigma)^{\otimes 2}, \boldsymbol{\alpha} \cup \boldsymbol{\beta})$ *on* $[F, \partial F]$ *is computed by twice the Euler measure of* f.

Proof. This is an immediate application of the Gauss-Bonnet theorem, using a metric for which the α_i and β_j are geodesics, each intersecting at right angles. The factor of two corresponds to the fact that we are now squaring the tangent bundle. □

Proof. [Of Theorem 8.2.7] Observe that $\mathfrak{j}$ can be chosen so that $\mathfrak{j}T_x\alpha_i = T_x\beta_j$ at each $x \in \alpha_i \cap \beta_j$. Choosing J over $\mathrm{Sym}^m(\Sigma)$ as in Corollary 7.6.2, it follows that for each $\mathbf{x} \in \mathbb{T}_\alpha \cap \mathbb{T}_\beta$, $JT_{\mathbf{x}}\mathbb{T}_\alpha = T_{\mathbf{x}}\mathbb{T}_\beta$. Consider a Whitney disk u from $\mathbf{x}$ to $\mathbf{y}$, thought of as a map $u\colon (\mathbb{D}, \partial\mathbb{D}) \to (\mathrm{Sym}^m(\Sigma), \mathbb{T}_\alpha \cup \mathbb{T}_\beta)$. Proposition 8.5.2 gives the formula

$$\mu(u) = \langle c_1(\det_{\mathbb{C}}(T\mathrm{Sym}^m(\Sigma)))^{\otimes 2}, \mathbb{T}_\alpha \cup \mathbb{T}_\beta), u_*([\mathbb{D}, \partial\mathbb{D}])\rangle.$$

After a small perturbation, we can assume that u is transverse to the branch locus $\Delta \subset \mathrm{Sym}^m(\Sigma)$. The argument from Lemma 7.4.2 gives a diagram

$$\begin{array}{ccc} (\widetilde{\mathbb{D}}, \partial\widetilde{\mathbb{D}}) & \xrightarrow{\widetilde{u}} & (\times^m(\Sigma), \widetilde{\mathbb{T}}_\alpha \cup \widetilde{\mathbb{T}}_\beta) \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle \pi} \\ (\mathbb{D}, \partial\mathbb{D}) & \xrightarrow{u} & (\mathrm{Sym}^m(\Sigma), \mathbb{T}_\alpha \cup \mathbb{T}_\beta), \end{array}$$

where

$$\widetilde{\mathbb{T}}_\alpha = \cup_{\sigma \in \mathfrak{S}_m} \alpha_{\sigma(1)} \times \cdots \times \alpha_{\sigma(m)}.$$

As in the proof of Lemma 7.4.2, there is an identification of $\pi^*(T\mathrm{Sym}^m(\Sigma))$ with $T(\times^m(\Sigma))$ away from the diagonal, and hence a corresponding identification of $\pi^*(\det_{\mathbb{C}}(\mathrm{Sym}^m(\Sigma)) \otimes \det_{\mathbb{C}}(\mathrm{Sym}^m(\Sigma)))$ with $\det_{\mathbb{C}}(\times^m(\Sigma)) \otimes \det_{\mathbb{C}}(\times^m(\Sigma))$ away from the diagonal.

Thus, the computation from Lemma 7.4.2 adapted to this relative setting gives

$$\mu(u) = \Delta(u) + \frac{1}{m!}\langle c_1(\det_{\mathbb{C}}(\times^m(\Sigma))^{\otimes 2}, \mathbb{T}_\alpha \cup \mathbb{T}_\beta), \widetilde{u}_*[\widetilde{\mathbb{D}}, \partial\widetilde{\mathbb{D}}]\rangle. \tag{8.23}$$

Note that in the statement of Lemma 7.4.2, the intersection number with Δ appears with a coefficient of $1/2$ whereas here it appears with a coefficient of 1. The reason for this discrepancy is that there we are computing the first Chern number of the tangent bundle, and here we are computing the first Chern number of its square.

We would like to express the first Chern class appearing in Equation (8.23) in $\times^m(\Sigma)$ in terms of Σ. To this end, observe that

$$\det_{\mathbb{C}}(\times^m(\Sigma))^{\otimes 2} = \pi_1^*\det_{\mathbb{C}}(\Sigma)^{\otimes 2} \otimes \cdots \otimes \pi_m^*\det_{\mathbb{C}}(\Sigma)^{\otimes 2};$$

and also the trivialization of $\det_{\mathbb{C}}(\times^m(\Sigma))^{\otimes 2}$ along $\widetilde{\mathbb{T}}_\alpha \cup \widetilde{\mathbb{T}}_\beta$ is induced from the tensor products of the trivializations of $\pi_i^*\det_{\mathbb{C}}(\Sigma)^{\otimes 2}$ along $\boldsymbol{\alpha} \cup \boldsymbol{\beta}$; so

$$c_1(\det_{\mathbb{C}}(\times^m(\Sigma))^{\otimes 2}, \widetilde{\mathbb{T}}_\alpha \cup \widetilde{\mathbb{T}}_\beta) = \sum_{i=1}^m \pi_i^*(c_1(\det_{\mathbb{C}}(\Sigma))^{\otimes 2}, \boldsymbol{\alpha} \cup \boldsymbol{\beta}). \tag{8.24}$$

Observe also that $\widetilde{u}$, $\widetilde{\Sigma}$, and $\times^m(\Sigma)$ all admit actions by $\mathfrak{S}_m$; and the quotient $\widetilde{u}/\mathfrak{S}_{m-1} = F$ fits into the following commutative diagram:

$$\begin{array}{ccc} (\widetilde{\mathbb{D}}, \partial\widetilde{\mathbb{D}}) & \xrightarrow{\widetilde{f}} & (\times^m(\Sigma), \widetilde{\mathbb{T}}_\alpha \cup \widetilde{\mathbb{T}}_\beta) \\ \downarrow & & \downarrow{\scriptstyle \pi_1} \\ (F, \partial F) & \xrightarrow{f} & (\Sigma, \boldsymbol{\alpha} \cup \boldsymbol{\beta}) \\ \downarrow{\scriptstyle p} & & \\ (\mathbb{D}, \partial\mathbb{D}) & & \end{array} \tag{8.25}$$

It follows from Equation (8.24) and Diagram (8.25) that

$$\begin{aligned} &\langle c_1(\det_{\mathbb{C}}^{\otimes 2}(\times^m(\Sigma)), \widetilde{\mathbb{T}}_\alpha \cup \widetilde{\mathbb{T}}_\beta), \widetilde{u}_*[\widetilde{\mathbb{D}}, \partial\widetilde{\mathbb{D}}]\rangle \\ &\qquad = \langle \sum_{i=1}^m \pi_i^*(c_1(\det_{\mathbb{C}}(\Sigma)^{\otimes 2}, \boldsymbol{\alpha} \cup \boldsymbol{\beta}), \widetilde{u}_*[\widetilde{\mathbb{D}}, \partial\widetilde{\mathbb{D}}]\rangle \\ &\qquad = m\pi_1^*(c_1(\det_{\mathbb{C}}(\Sigma)^{\otimes 2}, \boldsymbol{\alpha} \cup \boldsymbol{\beta}), \widetilde{u}_*[\widetilde{\mathbb{D}}, \partial\widetilde{\mathbb{D}}]\rangle \\ &\qquad = m! \cdot \pi_1^*(c_1(\det_{\mathbb{C}}(\Sigma)^{\otimes 2}, \boldsymbol{\alpha} \cup \boldsymbol{\beta}), \widetilde{u}_*[\widetilde{\mathbb{D}}, \partial\widetilde{\mathbb{D}}]\rangle, \end{aligned} \tag{8.26}$$

where the last step uses the fact that the quotient map $\widetilde{S} \to F$ has degree $(m-1)!$.

Combining Equations (8.23), (8.26) and Lemma 8.5.3, we find

$$\begin{aligned} \mu(\phi) &= \Delta(u) + \langle \pi_1^*(c_1(\det_{\mathbb{C}}(\Sigma))^{\otimes 2}, \boldsymbol{\alpha} \cup \boldsymbol{\beta}), f_*[F, \partial F]\rangle \\ &= \Delta(u) + 2e(f). \end{aligned}$$

□

Putting together the above results, we have the following:

Theorem 8.5.4. *Fix a generalized Heegaard diagram $\mathcal{H}$ in the sense of Definition 8.4.1 and a pair of generalized Heegaard states $\mathbf{x}, \mathbf{y} \in \mathcal{S}(\mathcal{H})$. For any $\phi \in W(\mathbf{x}, \mathbf{y})$, we have that*

$$\mu(\phi) = \overline{n}(\mathcal{S}(\phi)) + e(\mathcal{S}(\phi)). \tag{8.27}$$

Proof. This is an immediate consequence of Theorems 8.4.3 and 8.5.1. □

Theorem 8.2.6 is a special case of Theorem 8.5.4.

8.6. The Maslov index for cornerless domains

Given two Heegaard states $\mathbf{x}$ and $\mathbf{y}$, we shall define their relative grading to be the Maslov index of any disk that connects $\mathbf{x}$ and $\mathbf{y}$. (See for example Proposition 9.5.6.) This relative grading is only well defined modulo the Maslov index of cornerless domains P (thought of as $D(\mathbf{x},\mathbf{x})$) which, in turn, can be computed in terms of the Heegaard diagram by

$$\mu(P) = e(P) + 2\overline{n}_{\mathbf{x}}(P), \tag{8.28}$$

according to Theorem 8.2.6. (Noting that in this application, $\mathbf{x} = \mathbf{y}$.)

The aim of this section is to give the following explicit three-dimensional interpretation of the right-hand-side of Equation (8.28) (when $n_w(P) = 0$), in terms of the first Chern class of spin^c structures:

Proposition 8.6.1. *Let $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$ be a pointed Heegaard diagram for a closed, oriented three-manifold Y, let P be a cornerless domain, and let $h(P) \in H_2(Y;\mathbb{Z})$ be its associated two-dimensional homology class, as in Definition 2.3.8. Fix $\mathbf{x} \in \mathbb{T}_\alpha \cap \mathbb{T}_\beta$, and let $\mathfrak{s}_w(\mathbf{x})$ denote its associated spin^c structure, as in Section 2.6. Then,*

$$\langle c_1(\mathfrak{s}_w(\mathbf{x})), h(P)\rangle = e(P) + 2\overline{n}_{\mathbf{x}}(P) - 2n_w(P), \tag{8.29}$$

where $e(P)$ is Euler measure of P, as in Definition 8.2.3, and $\overline{n}_{\mathbf{x}}(P)$ is the point measure, as in Definition 8.2.1.

Remark 8.6.1. *The right-hand-sides of Equation* (8.28) *and* (8.29) *differ by $2n_w(P)$. The need for this correction is illustrated by the following example. When $P = [\Sigma]$, it is straightforward to see that $e(P) + 2n_{\mathbf{x}}(P) = 2$; while $\langle c_1(\mathfrak{s}_w(\mathbf{x})), h(P)\rangle = 0$, as the Heegaard surface Σ is null-homologous in Y.*

We prove the above proposition after a discussion and a lemma.

As in Section 1.4, we think of spin^c structures over a closed, oriented three-manifold Y as equivalence classes of nowhere vanishing vector fields (as described in Section 1.4). The first Chern class of a spin^c structure represented by the vector field v is the first Chern class of the complex line bundle $v^\perp$.

Equation (8.29) will be proved by comparing the tangent bundle of $h(P)$ with the restriction of the vector field representing $\mathfrak{s}_w(\mathbf{x})$, and showing that the difference localizes to the union of gradient trajectories through the components of $\mathbf{x}$ and the one through w.

We formulate this comparising using the notion of relative Chern numbers. Suppose that if X is a manifold and $\mathcal{L}$ is a complex line bundle over X, and τ is a given trivialization of $\mathcal{L}$ over $Y \subset X$. Then, the trivialization allows us to lift the first Chern class $c_1(\mathcal{L}) \in H^2(X;\mathbb{Z})$ to a relative class $c_1(\mathcal{L},\tau) \in H^2(X,Y;\mathbb{Z})$. (See Section 31.1.) When F is a compact, oriented surface F and $Y = \partial F$, we can evaluate this relative first Chern class $c_1(\mathcal{L},\tau) \in H^2(F,\partial F;\mathbb{Z})$ on the (relative) fundamental class $[F,\partial F] \in H_2(F,\partial F;\mathbb{Z})$, to give the *relative first Chern number of $\mathcal{L}$ relative to the trivialization τ*, denoted $\langle c_1(\mathcal{L},\tau),[F,\partial F]\rangle$. Explicitly, this number is the signed count of zeros of any transverse section σ of $\mathcal{L}$ which agrees with τ over ∂F. For example, equip the tangent bundle TF with its canonical trivialization τ over ∂F, specified by the positively oriented unit tangent vector to the boundary. It is easy to see that

$$\langle c_1(TF,\tau),[F,\partial F]\rangle = \chi(F). \tag{8.30}$$

We formulate now another useful property of the relative first Chern number, which is evident from the definition. Suppose that F_1 and F_2 are two closed surfaces with $\partial F_1 = Z = -\partial F_2$. For $i = 1,2$, equip F_i with a line bundle $\mathcal{L}_i$ and a trivialization τ_i over ∂F_i. We can glue these bundles using the specified trivializations to obtain a line bundle $\mathcal{L}$ over $F = F_1 \cup_Z F_2$, whose first Chern number is related to the relative first Chern numbers of the pieces by:

$$\langle c_1(\mathcal{L}),[F]\rangle = \langle c_1(\mathcal{L}_1,\tau_1),[F_1,\partial F_1]\rangle + \langle c_1(\mathcal{L}_2,\tau_2),[F_2,\partial F_2]\rangle.$$

With these remarks in place, we return to the set-up of Proposition 8.6.1, fixing a pointed Heegaard diagram $\mathcal{H}$ for Y, and a self-indexing Morse function $f\colon Y \to \mathbb{R}$ that represents the diagram. Let $\gamma_{\mathbf{x}}$ denote the g-tuple of gradient trajectories connecting the index one and two critical points at each $x \in \mathbf{x}$; and let γ_w denote the additional trajectory connecting the index zero and three critical points of f through $w \in \Sigma$. Following Section 2.6, the spin^c structure $\mathfrak{s}_w(\mathbf{x})$ associated to a Heegaard state $\mathbf{x}$ is represented by any non-vanishing vector field on Y that agrees with the gradient vector field $\vec{\nabla} f$ away from a regular neighborhood of $\gamma_{\mathbf{x}} \cup \gamma_w$.

Consider an immersion $\Phi\colon F \to Y$ of a surface F into Y which is transverse to $\vec{\nabla} f$, and which is transverse to $\gamma_{\mathbf{x}}$ and γ_w. Away from $\Phi^{-1}(\gamma_{\mathbf{x}} \cup \gamma_w)$, there is an identification of $\Phi^*(v^\perp)$ with the tangent bundle of F. Thus, the comparison of the first Chern numbers of TF and $\Phi^{-1}(v^\perp)$ localizes to contributions at the trajectory curves, which we describe next.

Suppose that γ_x is the connecting trajectory between some index one and index two critical points associated to $x \in \mathbf{x}$, and let B_x be a three-ball neighborhood of γ_x that meets the Heegaard surface Σ in a disk D_x. We specify a trivialization τ of $v^\perp|_{\partial D_x}$ as follows. Observe that over ∂D_x, v

and $\vec{\nabla} f$ agree. Moreover, since $\vec{\nabla} f$ is transverse to the Heegaard surface, there is an identification $\vec{\nabla} f^{\perp}|_{\Sigma} = T\Sigma$; in particular, $\vec{\nabla} f^{\perp}|_{D_x} = TD_x$. Moreover, the subbundle $T\partial D_x \subset TD_x|_{\partial D}$ has a canonical trivialization.

Let τ be a trivialization of $v^{\perp}|_{\partial D_x}$ induced from

$$T\partial D_x \subset TD_x|_{\partial D_x} \cong v^{\perp}|_{\partial D_x}.$$

Similarly, let B_w be the three-ball neighborhood of the trajectory γ_w, and let $D_w = B_w \cap \Sigma$. There is an analogous trivialization τ of $v^{\perp}|_{\partial D_w}$.

Lemma 8.6.2. *For D_x, v, and τ as above, the evaluation of the relative first Chern class of $v^{\perp}$ (relative to the trivialization τ) on the disk D_x is given by*

$$\langle c_1(v^{\perp}, \tau), [D_x, \partial D_x] \rangle = 3. \tag{8.31}$$

Similarly, for D_w as above,

$$\langle c_1(v^{\perp}, \tau), [D_w, \partial D_w] \rangle = -1. \tag{8.32}$$

Proof. The ball B_x is divided into two pieces B_α and B_β by the disk D_x, where

$$B_\alpha = B_x \cap U_\alpha \qquad \text{and} \qquad B_\beta = B_x \cap U_\beta;$$

B_α contains the index 1 critical point on the trajectory γ_x and B_β the index 2 critical point. Correspondingly, the sphere $S = \partial B_x$ is divided into two hemispheres D_α and D_β, so that the sphere $D_\alpha \cup D_x$ and $D_x \cup D_\beta$ bound B_α and B_β respectively.

Fix a trivialization $TY|_{B_x}$. With respect to this trivialization, the vector field v (after renormalized) can be viewed as a map from B_x to the two-sphere; in particular the degree of the restriction to the boundary of B_x vanishes: $\deg v|_{D_\alpha \cup D_x} = 0$. Similarly, the restriction of $\vec{\nabla} f/|\vec{\nabla} f|$ to $D_\alpha \cup D_x$ is a map between two two-spheres. Since $\vec{\nabla} f|_{B_\alpha}$ contains a single transverse zero, which has index 1, the induced map $\vec{\nabla} f/|\vec{\nabla} f| \colon D_\alpha \cup D_x \to S^2$ is orientation reversing.

The first of the above equations gives $\langle c_1(v^{\perp}), [D_\alpha \cup D_x] \rangle = 0$; while the second gives $\langle c_1(\vec{\nabla} f^{\perp}), [D_\alpha \cup D_x] \rangle = -2$ (explicitly, $\vec{\nabla} f^{\perp} \cong r^*(TS^2)$, where $r \colon S^2 \to S^2$ is an orientation-reversing map). Using the common trivialization τ of $v^{\perp}|_{\partial D_x} = \vec{\nabla} f^{\perp}|_{\partial D_x}$, we decompose these two equations:

$$\langle c_1(v^{\perp}, \tau), [D_x, \partial D_x] \rangle + \langle c_1(v^{\perp}, \tau), [D_\alpha, \partial D_\alpha] \rangle = 0$$
$$\langle c_1(\vec{\nabla} f^{\perp}, \tau), [D_x, \partial D_x] \rangle + \langle c_1(\vec{\nabla} f^{\perp}, \tau), [D_\alpha, \partial D_\alpha] \rangle = -2.$$

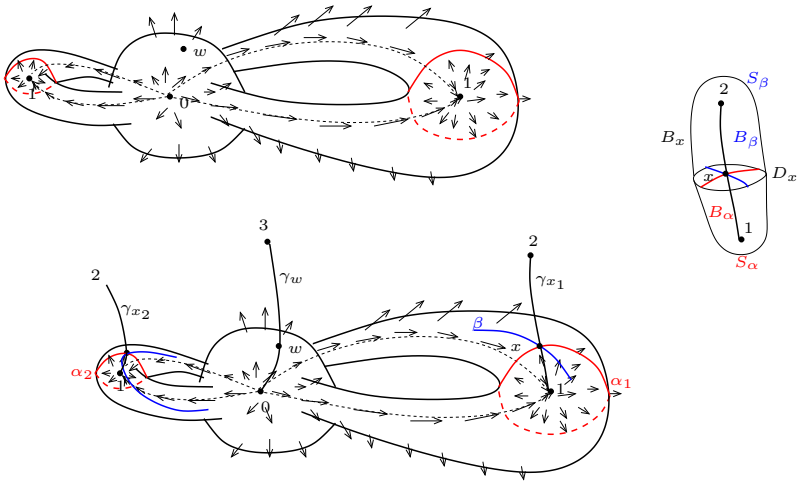


Figure 8.16. Modifying a vector field near a trajectory. At the top, we have indicated the gradient vector field on the handlebody U_α. Critical points are labelled by their indices. At the bottom, we have indicated two trajectories (though U_β is not drawn explicitly). At the right, we have drawn a neighborhood of a trjectory γ_x (dropping subscripts).

Since $\langle c_1(v^\perp, \tau), [D_\alpha, \partial D_\alpha]\rangle = \langle c_1(\vec{\nabla} f^\perp, \tau), [D_\alpha, \partial D_\alpha]\rangle$ (since $v = \vec{\nabla} f$ over D_α) and $\langle c_1(\vec{\nabla} f^\perp, \tau), [D_x, \partial D_x]\rangle = 1$ (using the identification $\vec{\nabla} f^\perp|_\Sigma = T\Sigma$ and then using Equation (8.30)), Equation (8.31) follows.

Equation (8.32) follows similarly. The key difference is that now if we consider neighborhood B_w of γ_w, and let $B_\alpha = B_w \cap U_\alpha$, then the restriction $\vec{\nabla} f|_{B_\alpha}$ contains a single transverse zero which has index 0, so the induced map on the boundary is orientation preserving. □

Proof. [Proof of Proposition 8.6.1] Equation (8.29) will be seen as a consequence of a formula computing $c_1(v^\perp)$ for appropriate surfaces immersed in Y (Equation (8.34) below), suitably reinterpreted in terms of the Heegaard surface.

As a warm-up, suppose that $\Phi\colon F \to Y$ is an immersed surface that misses the critical points of f, and is everywhere transverse to $\vec{\nabla} f$. Then,

$$\langle c_1(\vec{\nabla} f^\perp), [F]\rangle = \chi(F). \tag{8.33}$$

This is an immediate consequence of the transversality condition, which gives an identification of $\Phi^*(\vec{\nabla} f^\perp)$ with TF.

Fix a Heegaard state $\mathbf{x}$, let $\gamma_{\mathbf{x}}$ denote the corresponding g-tuple of gradient flowlines that connect the index 1 and 2 critical points; and let γ_w denote

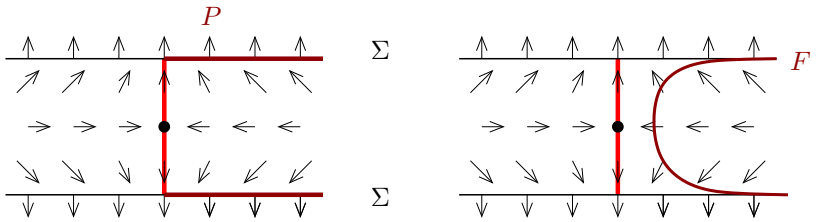


Figure 8.17. Making an attaching disk transverse to the gradient flow. At the left, we have indicated a cross-section of the one-handle, along with the gradient vector field. The actual one-handle is obtained by rotating the picture around a horizontal axis through the black dot in the middle (which represents the index one critical point). The red arc represents a cross section of the attaching disk. A periodic domain P is indicated; the surface F is obtained by attaching a copy of the attaching disk. At the right, we have shown how to perturb this F, making it transverse to the gradient flow.

the gradient flowline that connects the index 0 and 3 critical point, crossing the Heegaard surface at w. Suppose that $\Phi\colon F \to Y$ is as above; and that moreover Φ intersects $\gamma_{\mathbf{x}}$ and γ_w transversely. Then,

$$\langle c_1(v), [F]\rangle = \chi(F) + 2\#(\gamma_{\mathbf{x}} \cap F) - 2\#(\gamma_w \cap F). \tag{8.34}$$

This follows readily from Equation (8.33), with local corrections at the intersection points of $\gamma_{\mathbf{x}}$ and γ_w with F supplied by Lemma 8.6.2.

Given a non-negative, cornerless domain P, a suitable representative $\Phi\colon F_P \to Y$ for $h(P)$ was built in Construction 2.3.12. The underlying surface F_P is constructed from two pieces: a part F_0 that maps with non-zero differential into the Heegaard surface, via the map Φ_0, given in Construction 2.3.11; and another part, which consists of a disjoint union of copies of the α- and β-attaching disks, which we denote D. Thus, $F_P = F_0 \cup D$.

We shall perturb Φ slightly so that it is everywhere transverse to $\vec{\nabla} f$. To this end, observe that near each boundary component of F_0, there is a collar neighborhood $[0,1) \times S^1$. If the given boundary component maps into α_i, that boundary component bounds an attaching disk for U_α. The inclusion of this attaching disk is the extension of Φ across the given boundary component. The gradient flow $\vec{\nabla} f$ vanishes at the center of this disk. To avoid this, we change Φ to Φ' by a homotopy, removing a small annulus $[0,\epsilon)\times S^1$ from each collar neighborhood to obtain F_0' (diffeomorphic to F_0), and then the attaching disk along the new boundary component $\{\epsilon\} \times S^1$. The construction as described gives only a continuous map into Y, which is creased along the circle $\{\epsilon\} \times S^1$. It is easy to smooth out the map locally as in Figure 8.17, to obtain a smooth map to Y which is transverse to $\vec{\nabla} f$.

This construction gives a local immersion

$$\Phi' \colon F \to Y,$$

which misses the critical points of f, and is transverse to $\vec{\nabla} f$, so Equation (8.34) can be used to compute $\langle c_1(v), h(P) \rangle$.

Next, we express the right-hand-side of Equation (8.34) in this case in terms of the Heegaard surface. To this end, observe that the disks $D \subset F$ are disjoint from γ_w; so it follows readily that

$$\#(\gamma_w \cap F) = n_w(P).$$

Around each $x \in \mathbf{x}$, label the four local multiplicities a_x, b_x, c_x, and d_x, so that $a_x + c_x = b_x + d_x$ (compare Figure 2.9). The total number of disks attached to F_0 to obtain F is given by

$$\sum_{x \in \mathbf{x}} \left(|a_x - b_x| + |a_x - d_x| \right).$$

Moreover, since $\chi(F_0) = e(P)$, we see that

$$\chi(F) = e(P) + \sum_{x \in \mathbf{x}} \left(|a_x - b_x| + |a_x - d_x| \right).$$

To compute $\gamma_x \cap F$, observe that the attached disks are also disjoint from γ_x. Thus, the intersection number of γ_x with F can be computed by looking at the local multiplicity of F_0 at x, which in turn is $\min(a_x, b_x, c_x, d_x)$.

Thus, Equation (8.34) reads:

$$\langle c_1(v), [F] \rangle = e(P) + \sum_{x \in \mathbf{x}} \left(|a_x - b_x| + |a_x - d_x| + 2 \min(a_x, b_x, c_x, d_x) \right) - 2 n_w(P).$$

It is elementary to verify that if a, b, c, d are positive numbers, with $a + c = b + d$, then

$$|a - b| + |a - d| + 2 \min(a, b, c, d) = \frac{a + b + c + d}{2}.$$

In particular, Equation (8.29) follows for those domains P that have no negative local multiplicities.

Given an arbitrary domain P, we have

$$e(P + [\Sigma]) = e(P) + 2 - 2g \qquad \overline{n}(P + [\Sigma]) = \overline{n}(P) + g \qquad n_w(P + [\Sigma]) = n_w(P) + 1;$$

i.e. the right-hand-side of Equation (8.29) is invariant under adding copies of Σ. Also, $h(P + [\Sigma]) = h(P)$, so the left-hand-side is similarly invariant. Thus, Equation (8.29) for non-negative domains implies the formula in general. □

Remark 8.6.3. *We have used above the immersion of F into Y. It is not difficult to explicitly describe a perturbation of this map which is an embedding. For this perturbation, the portion F_0 maps into a tubular neighborhood of the Heegaard surface.*

8.7. Combinatorial aspects of the Maslov index

The formula of Theorem 8.2.6 expresses the Maslov index of $\phi \in W(\mathbf{x}, \mathbf{y})$ with the combinatorially defined quantities associated to the shadow of ϕ, specifically, with the sum $\overline{n}(\mathcal{S}(\phi)) + e(\mathcal{S}(\phi))$. Thus, basic properties of the Maslov index imply identities for $\overline{n}(\mathcal{S}(\phi)) + e(\mathcal{S}(\phi))$: integrality, an expression for its parity (compare Proposition 6.5.8), and additivity under juxtaposition (compare Proposition 6.5.7 (M-1)).

Our aim here is to give alternative, combinatorial proofs of these properties, without appealing to the identification with the Maslov index.

8.7.1. Integrality. The terms $\overline{n}(\mathcal{S}(\phi))$ and $e(\mathcal{S}(\phi))$ for $\phi \in W(\mathbf{x}, \mathbf{y})$ are not necessarily integers but their sum is, since that in turn is the Maslov index $\mu(\phi)$ from Chapter 6. We give here a more direct proof that $\overline{n}(\mathcal{S}(\phi)) + e(\mathcal{S}(\phi))$ is an integer.

Proposition 8.7.1. *For a Whitney disk $\phi \in W(\mathbf{x}, \mathbf{y})$ the sum $\overline{n}(\mathcal{S}(\phi)) + e(\mathcal{S}(\phi))$ is an integer.*

Proof. Suppose that F is a two-manifold-with-corners. Then the Gauss-Bonnet formula shows that

$$\chi(F) = e(F) + \sum_c \frac{2 - \ell_c}{4},$$

where $2\pi\ell_c$ measures the angle at c. Consider the surface constructed in Lemma 8.4.8. Note that $e(F_0) = e(\mathcal{S}(\phi))$. Moreover, all the corners of F_0 project to the points of $\mathbf{x} \cup \mathbf{y} \setminus (\mathbf{x} \cap \mathbf{y})$. Fix $p \in \mathbf{x} \cup \mathbf{y}$ with $p \in \alpha_i \cap \beta_j$. Let $N(\alpha_i)$ denote the number of components in ∂F that project to α_i.

When $p \in \mathbf{x} \cup \mathbf{y} \setminus \mathbf{x} \cap \mathbf{y}$, so that $f^{-1}(p)$ contains some corner c, then

$$\overline{n}_p(\phi) = \frac{\ell_c}{4} + \frac{N(\alpha_i) + N(\beta_j)}{2}.$$

When $p \in \mathbf{x} \cap \mathbf{y}$,

$$\overline{n}_p(\phi) = \nu_p + \frac{N(\alpha_i) + N(\beta_j)}{2},$$

where ν_p denotes the number of planes above p.

Summing the second equation over $p \in \mathbf{x}$ and then over all $p \in \mathbf{y}$, and letting $C(F_0)$ denote the set of corners of F_0, we find that

$$e(\phi) + \overline{n}(\phi) = \chi(F_0) + \sum_{c \in C(F_0)} \left(\frac{\ell_c - 1}{2} \right) + \sum_{i=1}^{m} (N(\alpha_i) + N(\beta_i)) + \sum_{p \in \mathbf{x} \cap \mathbf{y}} 2\nu_p. \tag{8.35}$$

Since at each corner point c we have $\ell_c \equiv 1 \pmod 2$, we conclude that the right-hand-side of the above equation is an integer. □

8.7.2. Mod 2 reduction.

Definition 8.7.2. *For a given generalized Heegaard diagram $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta})$, the following auxiliary choices are called **sign data**: orderings of the circles $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_m\}$ and $\boldsymbol{\beta} = \{\beta_1, \dots, \beta_m\}$ and orientations of all of the α_i and β_j.*

For a given sign data, we can associate to each Heegaard state $\mathbf{x}$ a permutation $\sigma_{\mathbf{x}} \in \mathfrak{S}_m$, uniquely specified by the property that $\mathbf{x} = \{x_1, \dots, x_m\}$ has $x_i \in \alpha_i \cap \beta_{\sigma(i)}$.

Given an oriented two-manifold Σ and two oriented curves $\alpha, \beta \subset \Sigma$ that meet transversely, we can associate to each $x \in \alpha \cap \beta$ a local intersection number $\epsilon_x(\alpha, \beta)$.

Sign data can be used to construct a map $s \colon \mathcal{S}(\mathcal{H}) \to \pm 1$, via the formula

$$s(\mathbf{x}) = \varepsilon(\sigma) \prod_{i=1}^{g} \epsilon_{x_i}(\alpha_i, \beta_{\sigma(i)}), \tag{8.36}$$

where $\varepsilon(\sigma)$ is $+1$ or -1 depending on whether the permutation σ is even or odd.

Lemma 8.7.3. *If $\mathbf{x}, \mathbf{y} \in \mathcal{S}(\mathcal{H})$, then $s(\mathbf{x}) \cdot s(\mathbf{y})$ is independent of the choice of sign data.*

Proof. The proof is straightforward. □

Proposition 8.7.4. *For a Whitney disk $\phi \in W(\mathbf{x}, \mathbf{y})$ we have that*

$$(-1)^{\overline{n}(\mathcal{S}(\phi)) + e(\mathcal{S}(\phi))} = s(\mathbf{x}) \cdot s(\mathbf{y}). \tag{8.37}$$

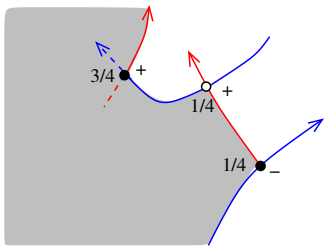


Figure 8.18. Relating the local intersection number ϵ, the type (x or y), and the multiplicity ℓ of a corner of F_0. The solid dots represent components of $\mathbf{x}$, the empty dot represents a component of $\mathbf{y}$. F_0 is the shaded region on the left, and the curves are oriented as ∂F_0. The sign $\epsilon \in \{\pm 1\}$ and the local multiplicity ℓ is indicated near each corner point. Compare with Equation (8.38).

Proof. This will be proved by taking a second more careful look at Equation (8.35).

In view of Lemma 8.7.3, we can assume that $\sigma_{\mathbf{x}}$ is the identity permutation. By Lemma 8.4.5, we can assume without loss of generality that ϕ is non-negative, we can use the surface F_0. The corner points in F_0 can be grouped into equivalence classes: two corner points are equivalent if they are on the same connected component of ∂F_0. Let $\mathcal{Z}$ denote the set of equivalence classes of corners of F_0. In view of Lemma 8.7.3, we can orient the α_i and β_j with the boundary orientation they inherit from F_0, as part of the boundary.

It is easy to see that

$$\ell_c \equiv \begin{cases} 1 & \text{if } \epsilon_{f(c)}(\alpha_i, \beta_j) = +1 \text{ and } f(c) \in \mathbf{x} \\ & \text{or } \epsilon_{f(c)}(\alpha_i, \beta_j) = -1 \text{ and } f(c) \in \mathbf{y} \\ 3 & \text{if } \epsilon_{f(c)}(\alpha_i, \beta_j) = +1 \text{ and } f(c) \in \mathbf{x} \\ & \text{or } \epsilon_{f(c)}(\alpha_i, \beta_j) = -1 \text{ and } f(c) \in \mathbf{y} \end{cases} \pmod 4. \tag{8.38}$$

(See Figure 8.18.) Equation (8.38) implies that for any $Z \in \mathcal{Z}$,

$$\begin{aligned} \sum_{c \in Z} (\ell_c - 1) &\equiv 2|Z^-| + |Z \cap f^{-1}(\mathbf{y})| \pmod 4 \\ &= 2|Z^-| - \frac{|Z|}{2} \pmod 4 \end{aligned} \tag{8.39}$$

where $|Z|$ denotes the number of corner points in the equivalence class Z, and $|Z^-|$ denotes the number of corner points that project to a negative intersection point of α_i with β_j.

Combining Equations (8.35) and (8.39), it follows that
(8.40)

$$e(\phi)+\overline{n}(\phi) \equiv \chi(F_0)+|\mathbf{x}^-|+|\mathbf{y}^-|+\left(\sum_{Z\in\mathcal{Z}}\frac{|Z|}{2}\right)+\sum_{i=1}^{m}(N(\alpha_i)+N(\alpha_j)) \pmod 2,$$

where $|\mathbf{x}^-|$ denotes the number of $x \in \mathbf{x}$ with $\epsilon_x = -1$. Boundary components of F_0 are of three types: those corresponding to the cycles $Z \in \mathcal{Z}$, those counted in $N(\alpha_i)$ (for various choices of α_i), and those counted in $N(\beta_i)$; and hence, for the number of components $|\pi_0(F_0)|$ we have

$$|\pi_0(F_0)| = |\mathcal{Z}| + \sum_{i=1}^{m}(N(\alpha_i) + N(\alpha_j)).$$

For any oriented surface,

$$\chi(F_0) \equiv |\pi_0(F_0)| \pmod 2. \tag{8.41}$$

The permutation $\sigma_{\mathbf{x}} \cdot \sigma_{\mathbf{y}}$ splits as a product of disjoint cycles, which in turn are in one-to-one correspondence with equivalence classes $\mathcal{Z}$ of corners in F_0. Moreover, the cycle corresponding to Z can be written as a product of $\frac{|Z|}{2} - 1$ transpositions, so

$$(-1)^{\sum_{Z\in\mathcal{Z}}|\frac{|Z|}{2}} = \varepsilon(\sigma)(-1)^{|\mathcal{Z}|}. \tag{8.42}$$

Equation (8.37) now follows from Equations (8.40), (8.41), and (8.42). □

Exercise 8.7.5. *Let $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$ be a generalized Heegaard diagram with multiplicity m, and fix sign data as in Definition 8.7.2. Use this data to orient the submanifolds $\mathbb{T}_\alpha$ and $\mathbb{T}_\beta$ in $\mathrm{Sym}^m(\Sigma)$. Under the natural correspondence between $\mathcal{S}(\mathcal{H})$ and intersection points between $\mathbb{T}_\alpha$ and $\mathbb{T}_\beta$ (from Proposition 8.1.1), show that, up to an overall product with ± 1, the function $s(\mathbf{x})$ as constructed above coincides with the local intersection number of $\mathbb{T}_\alpha$ with $\mathbb{T}_\beta$ at $\mathbf{x}$.*

Remark 8.7.6. *Using Exercise 8.7.5 Equation* (8.37) *can now be seen as a special case of the relationship between the parity of a Maslov index and the products of local intersection numbers at the two corners, expressed in Proposition 6.5.8.*

8.7.3. Additivity. Whitney disks can be juxtaposed, and the Maslov index is additive under this operation. On the other hand, the shadow of the juxtaposition of two Whitney disks is the sum of the shadows of the disks.

A function

$$f\colon \bigcup_{\mathbf{x},\mathbf{y}} D(\mathbf{x},\mathbf{y}) \to \mathbb{Z}$$

is called *additive* if for any three states $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{S}(\mathcal{H})$ and two domains $\mathcal{D}_1 \in D(\mathbf{x}, \mathbf{y})$, $\mathcal{D}_2 \in D(\mathbf{y}, \mathbf{z})$,

$$f(\mathcal{D}_1 + \mathcal{D}_2) = f(\mathcal{D}_1) + f(\mathcal{D}_2).$$

In the present section we give an elementary proof following [**122**] (independent of its identification with the Maslov index) of the fact that the function $\overline{n} + e$ is additive in the above sense.

Since the Euler measure is clearly an additive function, it suffices to show that the point measure $\overline{n}$ is also additive.

Proposition 8.7.7. *Suppose that $\mathcal{D}_1 \in D(\mathbf{x}, \mathbf{y})$ and $\mathcal{D}_2 \in D(\mathbf{y}, \mathbf{z})$. Then*

$$\overline{n}(\mathcal{D}_1 + \mathcal{D}_2) = \overline{n}(\mathcal{D}_1) + \overline{n}(\mathcal{D}_2).$$

The above identity will be proved after establishing a lemma. Suppose that $\mathcal{D} \in D(\mathbf{x}, \mathbf{y})$ is a domain connecting the Heegaard states $\mathbf{x}$ and $\mathbf{y}$. Thought of as a two-chain, $\mathcal{D}$ has its boundary $\partial\mathcal{D}$, which we can think of as a one-chain supported in $\boldsymbol{\alpha} \cup \boldsymbol{\beta}$. Define $\partial_\alpha = (\partial\mathcal{D}) \cap \boldsymbol{\alpha}$ and $\partial_\beta = (\partial\mathcal{D}) \cap \boldsymbol{\beta}$.

Consider an oriented interval I on the β-curves that connects two points of $\boldsymbol{\alpha} \cap \boldsymbol{\beta}$. We would like to define the intersection multiplicity of $\partial_\alpha\mathcal{D}$ with I, even in cases when the two curves meet at the boundary. To this end, note that any component of $\partial_\alpha\mathcal{D}$ can be displaced slightly in four different ways, depending on which domains the two endpoints of the immersed interval is moved into. We define the intersection number of $\partial_\alpha\mathcal{D}$ with I, denoted $\#(\partial_\alpha\mathcal{D} \cap I)$, to be the average of the intersection numbers for the four displacements of each connected component of $\partial_\alpha\mathcal{D}$. In a similar manner, we can define the intersection number $\#(A \cap \partial_\alpha\mathcal{D})$, where A is any oriented one-chain supported in $\boldsymbol{\beta}$. We can extend this to one-chains supported in $\boldsymbol{\alpha} \cup \boldsymbol{\beta}$: if A is supported in $\boldsymbol{\alpha}$, then $\#(A \cap \partial_\alpha\mathcal{D})$ is defined to vanish.

In particular, given any two-chain $\mathcal{D}' = \sum_i m_i\mathcal{D}_i$ (where $m_i \in \mathbb{Z}$, and $\mathcal{D}_i$ are elementary domains) and a domain $\mathcal{D} \in D(\mathbf{x}, \mathbf{y})$, the above procedure gives a well-defined intersection number $\#(\partial\mathcal{D}' \cap \partial_\alpha\mathcal{D})$. Using $\boldsymbol{\beta}$ in place of $\boldsymbol{\alpha}$, we can define analogously $\#(\partial\mathcal{D}' \cap \partial_\beta\mathcal{D})$.

Lemma 8.7.8. *Suppose that $\mathcal{D}'$ is a two-chain in Σ, that is, a formal linear combination of elementary domains $\mathcal{D}' = \sum_i m_i\mathcal{D}_i$. Suppose furthermore that $\mathcal{D} \in D(\mathbf{x}, \mathbf{y})$ is a domain connecting the Heegaard states $\mathbf{x}$ and $\mathbf{y}$. Then*

$$\overline{n}_{\mathbf{x}}(\mathcal{D}') - \overline{n}_{\mathbf{y}}(\mathcal{D}') = \#(\partial_\alpha\mathcal{D} \cap \partial\mathcal{D}') = \#(\partial\mathcal{D}' \cap \partial_\beta\mathcal{D}).$$

Proof. Write $\partial_\alpha \mathcal{D} = \sum_{i=1}^{g} A_i$, where $A_i \subset \alpha_i$ is an oriented one-chain with $\partial A_i = y_i - x_i$. It is easy to see that

$$\overline{n}_{x_i}(\mathcal{D}') - \overline{n}_{y_i}(\mathcal{D}') = \#(A_i \cap \partial \mathcal{D}').$$

Adding these up for $i = 1, \dots, g$, gives

$$\overline{n}_{\mathbf{x}}(\mathcal{D}') - \overline{n}_{\mathbf{y}}(\mathcal{D}') = \#(\partial_\alpha \mathcal{D} \cap \partial \mathcal{D}').$$

The other equality follows analogously. □

With that preparation in place, we can now return to the proof of Proposition 8.7.7.

Proof. [Proof of Proposition 8.7.7] Recall that the point measure of a domain is by definition the sum of the point measures at the two endpoints of the domain, hence the identity we would like to prove can be written as

$$\overline{n}_{\mathbf{x}}(\mathcal{D}_1 + \mathcal{D}_2) + \overline{n}_{\mathbf{z}}(\mathcal{D}_1 + \mathcal{D}_2) = \overline{n}_{\mathbf{x}}(\mathcal{D}_1) + \overline{n}_{\mathbf{y}}(\mathcal{D}_1) + \overline{n}_{\mathbf{y}}(\mathcal{D}_2) + \overline{n}_{\mathbf{z}}(\mathcal{D}_2).$$

Since $\overline{n}_{\mathbf{x}}(\mathcal{D}_1 + \mathcal{D}_2) = \overline{n}_{\mathbf{x}}(\mathcal{D}_1) + \overline{n}_{\mathbf{x}}(\mathcal{D}_2)$ (and similarly for $\overline{n}_{\mathbf{z}}$), it suffices to show

$$\overline{n}_{\mathbf{x}}(\mathcal{D}_2) - \overline{n}_{\mathbf{y}}(\mathcal{D}_2) = \overline{n}_{\mathbf{y}}(\mathcal{D}_1) - \overline{n}_{\mathbf{z}}(\mathcal{D}_1).$$

Applying Lemma 8.7.8 twice, and noting that the α-curves in $\partial \mathcal{D}_2$ are disjoint from the displaced copies of $\partial_\alpha \mathcal{D}_1$, we have that

$$\begin{aligned} \overline{n}_{\mathbf{x}}(\mathcal{D}_2) - \overline{n}_{\mathbf{y}}(\mathcal{D}_2) &= \#(\partial_\alpha \mathcal{D}_1 \cap \partial \mathcal{D}_2) = \#(\partial_\alpha \mathcal{D}_1 \cap \partial_\beta \mathcal{D}_2) \\ &= \#(\partial \mathcal{D}_1 \cap \partial_\beta \mathcal{D}_2) = \overline{n}_{\mathbf{y}}(\mathcal{D}_1) - \overline{n}_{\mathbf{z}}(\mathcal{D}_1), \end{aligned}$$

as needed. □

8.8. Formula for the Maslov index for polygons

In Section 6.10, we explained how to extend the Maslov index for Whitney disks to Whitney m-gons. Sucharit Sarkar [**122**] gave an extension of Theorem 8.2.6 to the case of Whitney m-gons, which we state here without proof.

Let Σ be an oriented, closed, smooth two-dimensional manifold of genus g, and $\boldsymbol{\alpha}^i = \{\alpha_1^i, \dots, \alpha_g^i\}$ $(i = 1, \dots, n)$ be a collection of n complete sets of attaching circles, as in Definition 2.1.1; assume moreover that all the distinct curves α_j^i and α_ℓ^k intersect transversely. We call this data a *Heegaard n-tuple*. (Cf. also Definition 12.4.1 in Chapter 12.) The corresponding n tori in $\mathrm{Sym}^g(\Sigma)$, $\mathbb{T}_{\alpha^i} = \times_{j=1}^{g} \alpha_j^i$ for $i = 1, \dots, n$ intersect transversely. Chosing a symplectic structure $\mathrm{Sym}^g(\Sigma)$ as discussed in Section 7.6, the tori $\mathbb{T}_{\alpha^i}$ can be made to be all Lagrangian, hence we get a transversal chain, in the sense of Definition 6.10.1. As we discussed in Section 6.10, by fixing intersection

points $x_{i,i+1} \in \mathbb{T}_{\alpha^i} \cap \mathbb{T}_{\alpha^{i+1}}$ we can define Whitney n-gons, their homotopy classes and their Maslov indices.

The connected components of $\Sigma \setminus \cup_{i,j} \alpha^i_j$ are the elementary domains of the Heegaard n-tuple $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}^1, \dots, \boldsymbol{\alpha}^n)$. As in the disk case, a Whitney n-gon, that is, a map $u\colon \mathbb{D} \setminus \{v_1, \dots, v_n\} \to \mathrm{Sym}^g(\Sigma)$ with $u(\overline{v_i v_{i+1}}) \subset \mathbb{T}_{\alpha_i}$, defines a formal linear combination (the shadow $\mathcal{S}([u])$ of u) of elementary domains: if $\mathcal{D}_i$ is an elementary domain in $\mathcal{H}$, then the coefficient of it in $\mathcal{S}([u])$ is equal to the intersection number of $\{z_i\} \times \mathrm{Sym}^{g-1}(\Sigma) \subset \mathrm{Sym}^g(\Sigma)$ with the image of u for any $z_i \in \mathrm{int}\, \mathcal{D}_i$.

Orient all the circles in the $\boldsymbol{\alpha}^i$'s and recall that Σ is an oriented surface. Suppose now that a_i and a_j are one-chains supported by $\boldsymbol{\alpha}^i$ and $\boldsymbol{\alpha}^j$, respectively. Their intersection number $a_i \cdot a_j$ is defined as before: displace a_j in all four possible directions so that its endpoints are disjoint from a_i (and endpoints of a_i are disjoint from the translates). Then the intersection number of any of the displaced copies can be defined using the orientation of Σ (and the order of the two chains); define $a_i \cdot a_j$ as the average of these four integers. It is easy to see that $a_j \cdot a_i = -a_i \cdot a_j$.

Suppose now that u is a Whitney n-gon and its shadow $\mathcal{S}([u])$ is the two-chain associated to it (or more precisely, to its homotopy class $[u]$) as above. The definition of the Euler measure and the point measure of $\mathcal{S}([u])$ naturally extends from disks (given in Definitions 8.2.1 and 8.2.3) to n-gons.

Then, the Maslov index $\mu([u])$ of a Whitney n-gon can be determined through its shadow $\mathcal{S}([u])$ as follows:

Theorem 8.8.1 (Sarkar [**122**]). *Suppose that $(\Sigma, \boldsymbol{\alpha}^1, \dots, \boldsymbol{\alpha}^n)$ is a Heegaard n-tuple, and $\vec{\mathbf{x}} = (\mathbf{x}_{1,2}, \dots, \mathbf{x}_{n,1})$ is an intersection point of the transverse chain $(\mathbb{T}_{\alpha^1}, \dots, \mathbb{T}_{\alpha^n})$. Suppose furthermore that u is a Whitney n-gon (as defined in Section 6.10). Then the Maslov index $\mu([u])$ is computed by*

$$\mu([u]) = e(\mathcal{S}([u])) + \overline{n}_{\mathbf{x}_{1,2}} + \overline{n}_{\mathbf{x}_{n,1}} + \sum_{1<i<j\leq n} \partial_j(\mathcal{S}([u])) \cdot \partial_i(\mathcal{S}([u])) - \frac{g(n-2)}{2},$$

where $\partial_j(\mathcal{S}([u]))$ is the intersection of the boundary of the two-chain $\mathcal{S}([u])$ with the curves in $\boldsymbol{\alpha}^j$.

Notice that for a bigon (i.e, when $n = 2$), the sum contains no terms, and also the last additive term vanishes, hence this formula generalizes the one given in Theorem 8.2.6. For a proof of Theorem 8.8.1 we refer the interested reader to [**122**]; see also [**64**].

Bibliography

[1] L. Ahlfors and L. Bers. Riemann's mapping theorem for variable metrics. *Ann. of Math. (2)*, 72:385–404, 1960.

[2] E. Arbarello, M. Cornalba, and P. Griffiths. *Geometry of algebraic curves. Volume II*, volume 268 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2011. With a contribution by Joseph Daniel Harris.

[3] V. I. Arnol'd. On a characteristic class entering into conditions of quantization. *Funkcional. Anal. i Prilozhen.*, 1:1–14, 1967.

[4] M. Atiyah and I. Macdonald. *Introduction to commutative algebra*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969.

[5] M. Audin and M. Damian. *Morse theory and Floer homology*. Universitext. Springer, London; EDP Sciences, Les Ulis, 2014. Translated from the 2010 French original by Reinie Erné.

[6] Augustin Banyaga and Christopher Saunders. Floer homology for almost Hamiltonian isotopies. *C. R. Math. Acad. Sci. Paris*, 342(6):417–420, 2006.

[7] A. Bertram and M. Thaddeus. On the quantum cohomology of a symmetric product of an algebraic curve. *Duke Math. J.*, 108(2):329–362, 2001.

[8] J. E. Borzellino and V. Brunsden. Elementary orbifold differential topology. *Topology Appl.*, 159(17):3583–3589, 2012.

[9] R. Bott and L. Tu. *Differential forms in algebraic topology*, volume 82 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.

[10] P. Braam and S. Donaldson. Floer's work on instanton homology, knots, and surgery. In H. Hofer, C. Taubes, A. Weinstein, and E. Zehnder, editors, *The Floer Memorial Volume*, number 133 in Progress in Mathematics, pages 195–256. Birkhäuser, 1995.

[11] G. E. Bredon. *Introduction to compact transformation groups*. Academic Press, New York-London, 1972. Pure and Applied Mathematics, Vol. 46.

[12] Ana Cannas da Silva. *Lectures on symplectic geometry*, volume 1764 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001.

[13] Kai Cieliebak and Yakov Eliashberg. *From Stein to Weinstein and back*, volume 59 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012. Symplectic geometry of affine complex manifolds.

[14] P. Cromwell. *Knots and links*. Cambridge University Press, Cambridge, 2004.

[15] R. Crowell. Genus of alternating link types. *Ann. of Math. (2)*, 69:258–275, 1959.

[16] R. Crowell and R. Fox. *Introduction to knot theory*. Springer-Verlag, New York, 1977. Reprint of the 1963 original, Graduate Texts in Mathematics, No. 57.

[17] S. Donaldson. An application of gauge theory to four-dimensional topology. *J. Differential Geom.*, 18(2):279–315, 1983.

[18] 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.

[19] Simon Donaldson and Ivan Smith. Lefschetz pencils and the canonical class for symplectic four-manifolds. *Topology*, 42(4):743–785, 2003.

[20] N. Elkies. A characterization of the Z^n lattice. *Math. Res. Lett.*, 2(3):321–326, 1995.

[21] A. Floer. Morse theory for Lagrangian intersections. *J. Differential Geom.*, 28(3):513–547, 1988.

[22] A. Floer. Symplectic fixed points and holomorphic spheres. *Comm. Math. Phys.*, 120(4):575–611, 1989.

[23] A. Floer, H. Hofer, and D. Salamon. Transversality in elliptic Morse theory for the symplectic action. *Duke Math. J*, 80(1):251–29, 1995.

[24] R. Fox and J. Milnor. Singularities of 2-spheres in 4-space and cobordism of knots. *Osaka J. Math.*, 3:257–267, 1966.

[25] Urs Frauenfelder. Gromov convergence of pseudoholomorphic disks. *J. Fixed Point Theory Appl.*, 3(2):215–271, 2008.

[26] M. Freedman. A surgery sequence in dimension four; the relations with knot concordance. *Invent. Math.*, 68(2):195–226, 1982.

[27] K. Frøyshov. The Seiberg-Witten equations and four-manifolds with boundary. *Math. Res. Lett*, 3:373–390, 1996.

[28] K. Fukaya, Y-G. Oh, K. Ono, and H. Ohta. *Lagrangian intersection Floer theory—anomaly and obstruction*. Kyoto University, 2000.

[29] Kenji Fukaya and Kaoru Ono. Arnold conjecture and Gromov-Witten invariant. *Topology*, 38(5):933–1048, 1999.

[30] D. Gabai. Foliations and the topology of 3-manifolds. *J. Differential Geom.*, 18(3):445–503, 1983.

[31] P. Ghiggini. Knot Floer homology detects genus-one fibred knots. *Amer. J. Math.*, 130(5):1151–1169, 2008.

[32] H. Goda, H. Matsuda, and T. Morifuji. Knot Floer homology of $(1,1)$-knots. *Geom. Dedicata*, 112:197–214, 2005.

[33] R. Gompf. A new construction of symplectic manifolds. *Ann. of Math. (2)*, 142(3):527–595, 1995.

[34] R. Gompf and A. Stipsicz. *4-manifolds and Kirby calculus*, volume 20 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1999.

[35] Robert E. Gompf. Spin^c-structures and homotopy equivalences. *Geom. Topol.*, 1:41–50, 1997.

[36] C. Gordon. Dehn surgery on knots. In *Proceedings of the International Congress of Mathematicians Vol. I (Kyoto, 1990)*, pages 631–642. Springer-Verlag, 1991.

[37] C. Gordon and R. Litherland. On the signature of a link. *Invent. Math.*, 47(1):53–69, 1978.

[38] H. Grauert and R. Remmert. Plurisubharmonische Funktionen in komplexen Räumen. *Math. Z.*, 65:175–194, 1956.

[39] J. Greene. A spanning tree model for the Heegaard Floer homology of a branched double-cover. *J. Topol.*, 6(2):525–567, 2013.

[40] J. Greene. Alternating links and definite surfaces. *Duke Math. J.*, 166(11):2133–2151, 2017. With an appendix by András Juhász and Marc Lackenby.

[41] J. Greene and A. Levine. Strong Heegaard diagrams and strong L-spaces. *Algebr. Geom. Topol.*, 16(6):3167–3208, 2016.

[42] Phillip Griffiths and Joseph Harris. *Principles of algebraic geometry*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1994. Reprint of the 1978 original.

[43] M. Gromov. Pseudoholomorphic curves in symplectic manifolds. *Invent. Math.*, 82(2):307–347, 1985.

[44] V. Guillemin and A. Pollack. *Differential topology*. AMS Chelsea Publishing, Providence, RI, 2010. Reprint of the 1974 original.

[45] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. Oxford University Press, Oxford, sixth edition, 2008. Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles.

[46] R. Hartshorne. *Algebraic Geometry*. Number 52 in Graduate Texts in Mathematics. Springer-Verlag, 1977.

[47] A. Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002.

[48] M. Hedden. On Floer homology and the Berge conjecture on knots admitting lens space surgeries. *Trans. Amer. Math. Soc.*, 363(2):949–968, 2011.

[49] H. Hilden. Every closed orientable 3-manifold is a 3-fold branched covering space of S^3. *Bull. Amer. Math. Soc.*, 80:1243–1244, 1974.

[50] H. Hilden, M. Lozano, and J. Montesinos. On knots that are universal. *Topology*, 24(4):499–504, 1985.

[51] Friedrich Hirzebruch and Heinz Hopf. Felder von Flächenelementen in 4-dimensionalen Mannigfaltigkeiten. *Math. Ann.*, 136:156–172, 1958.

[52] J. Hom. The knot Floer complex and the smooth concordance group. arXiv:1111.6635.

[53] Jennifer Hom. A note on cabling and L-space surgeries. *Algebr. Geom. Topol.*, 11(1):219–223, 2011.

[54] Jennifer Hom. Satellite knots and L-space surgeries. *Bull. Lond. Math. Soc.*, 48(5):771–778, 2016.

[55] J. Howie. A characterisation of alternating knot exteriors. *Geom. Topol.*, 21(4):2353–2371, 2017.

[56] Yang Huang and Vinicius G. B. Ramos. An absolute grading on Heegaard Floer homology by homotopy classes of oriented 2-plane fields. *J. Symplectic Geom.*, 15(1):51–90, 2017.

[57] D. Huybrechts. *Complex geometry*. Universitext. Springer-Verlag, Berlin, 2005. An introduction.

[58] A. Juhász. Holomorphic discs and sutured manifolds. *Algebr. Geom. Topol.*, 6:1429–1457, 2006.

[59] A. Juhász. The sutured Floer homology polytope. *Geom. Topol.*, 14(3):1303–1354, 2010.

[60] P. Kronheimer and T. Mrowka. Gauge theory for embedded surfaces. I. *Topology*, 32(4):773–826, 1993.

[61] P. Kronheimer and T. Mrowka. The genus of embedded surfaces in the projective plane. *Math. Res. Lett.*, 1(6):797–808, 1994.

[62] P. Kronheimer and T. Mrowka. Scalar curvature and the Thurston norm. *Math. Res. Lett.*, 4(6):931–937, 1997.

[63] P. Kronheimer and T. Mrowka. *Monopoles and three-manifolds*, volume 10 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2007.

[64] R. Krutowski. Combinatorial proof of Maslov index formula in Heegaard Floer theory. 2206.05221.

[65] H. Lawson and M.-L. Michelsohn. *Spin Geometry*. Number 38 in Princeton Mathematics Series. Princeton University Press, 1989.

[66] A. Levine and S. Lewallen. Strong L-spaces and left-orderability. *Math. Res. Lett.*, 19(6):1237–1244, 2012.

[67] R. Lickorish. A representation of orientable combinatorial 3-manifolds. *Ann. of Math. (2)*, 76:531–540, 1962.

[68] R. Lickorish. *An introduction to knot theory*, volume 175 of *Graduate Texts in Mathematics*. Springer-Verlag, 1997.

[69] R. Lipshitz. A cylindrical reformulation of Heegaard Floer homology. *Geom. Topol.*, 10:955–1097 (electronic), 2006.

[70] R. Lipshitz. Correction to the article: A cylindrical reformulation of Heegaard Floer homology. *Geom. Topol.*, 18:17–30 (electronic), 2014.

[71] I. MacDonald. Symmetric products of an algebraic curve. *Topology*, 1:319–343, 1962.

[72] C. Manolescu, P. Ozsváth, and S. Sarkar. A combinatorial description of knot Floer homology. *Ann. of Math. (2)*, 169(2):633–660, 2009.

[73] C. Manolescu, P. Ozsváth, Z. Szabó, and D. Thurston. On combinatorial link Floer homology. *Geom. Topol.*, 11:2339–2412, 2007.

[74] W. S. Massey. On the Stiefel-Whitney classes of a manifold. II. *Proc. Amer. Math. Soc.*, 13:938–942, 1962.

[75] D. McDuff and D. Salamon. *J-holomorphic curves and quantum cohomology*. Number 6 in University Lecture Series. American Mathematical Society, 1994.

[76] D. McDuff and D. Salamon. *Introduction to symplectic topology*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, second edition, 1998.

[77] J. Milnor. *Morse theory*. Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51. Princeton University Press, Princeton, N.J., 1963.

[78] J. Milnor. Spin structures on manifolds. *Enseign. Math. (2)*, 9:198–203, 1963.

[79] J. Milnor. *Lectures on the h-cobordism theorem*. Princeton University Press, 1965. Notes by L. Siebenmann and J. Sondow.

[80] J. Milnor. *Topology from the differentiable viewpoint*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997. Based on notes by David W. Weaver, Revised reprint of the 1965 original.

[81] J. Milnor and D. Husemoller. *Symmetric bilinear forms*. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 73. Springer-Verlag, New York-Heidelberg, 1973.

[82] J. Milnor and J. Stasheff. *Characteristic classes*. Annals of Mathematics Studies, No. 76. Princeton University Press, Princeton, N. J.; University of Tokyo Press, Tokyo, 1974.

[83] E. Moise. *Geometric topology in dimensions* 2 *and* 3. Springer-Verlag, New York-Heidelberg, 1977. Graduate Texts in Mathematics, Vol. 47.

[84] J. Montesinos. A representation of closed orientable 3-manifolds as 3-fold branched coverings of S^3. *Bull. Amer. Math. Soc.*, 80:845–846, 1974.

[85] J. Montesinos. Three-manifolds as 3-fold branched covers of S^3. *Quart. J. Math. Oxford Ser. (2)*, 27(105):85–94, 1976.

[86] J. Morgan, Z. Szabó, and C. Taubes. A product formula for Seiberg-Witten invariants and the generalized Thom conjecture. *J. Differential Geometry*, 44:706–788, 1996.

[87] D. Mumford. *Tata lectures on theta. I.* Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2007. With the collaboration of C. Musili, M. Nori, E. Previato and M. Stillman, Reprint of the 1983 edition.

[88] David Mumford. *Abelian varieties*, volume 5 of *Tata Institute of Fundamental Research Studies in Mathematics*. Tata Institute of Fundamental Research, Bombay; by Hindustan Book Agency, New Delhi, 2008. With appendices by C. P. Ramanujam and Yuri Manin, Corrected reprint of the second (1974) edition.

[89] K. Murasugi. On the genus of the alternating knot. I, II. *J. Math. Soc. Japan*, 10:94–105, 235–248, 1958.

[90] K. Murasugi. On a certain subgroup of the group of an alternating link. *Amer. J. Math.*, 85:544–550, 1963.

[91] K. Murasugi. On the Alexander polynomial of alternating algebraic knots. *J. Austral. Math. Soc. Ser. A*, 39(3):317–333, 1985.

[92] T. Napier and M. Ramachandran. *An introduction to Riemann surfaces*. Cornerstones. Birkhäuser/Springer, New York, 2011.

[93] A. Newlander and L. Nirenberg. Complex analytic coordinates in almost complex manifolds. *Ann. of Math. (2)*, 65:391–404, 1957.

[94] Y. Ni. Knot Floer homology detects fibred knots. *Invent. Math.*, 170(3):577–608, 2007.

[95] Yong-Geun Oh. Floer cohomology of Lagrangian intersections and pseudo-holomorphic disks. I. *Comm. Pure Appl. Math.*, 46(7):949–993, 1993.

[96] Yong-Geun Oh. Floer cohomology of Lagrangian intersections and pseudo-holomorphic disks. II. $(\mathbf{C}\mathrm{P}^n, \mathbf{R}\mathrm{P}^n)$. *Comm. Pure Appl. Math.*, 46(7):995–1012, 1993.

[97] P. Orlik. *Seifert manifolds*. Lecture Notes in Mathematics, Vol. 291. Springer-Verlag, Berlin-New York, 1972.

[98] B. Owens. Unknotting information from Heegaard Floer homology. *Adv. Math.*, 217(5):2353–2376, 2008.

[99] P. Ozsváth and Z. Szabó. The symplectic Thom conjecture. *Ann. of Math.*, 151(1):93–124, 2000.

[100] P. Ozsváth and Z. Szabó. Absolutely graded Floer homologies and intersection forms for four-manifolds with boundary. *Advances in Mathematics*, 173(2):179–261, 2003.

[101] P. Ozsváth and Z. Szabó. Heegaard Floer homology and alternating knots. *Geom. Topol.*, 7:225–254 (electronic), 2003.

[102] P. Ozsváth and Z. Szabó. Holomorphic disks and genus bounds. *Geom. Topol.*, 8:311–334, 2004.

[103] P. Ozsváth and Z. Szabó. Holomorphic disks and knot invariants. *Adv. Math.*, 186(1):58–116, 2004.

[104] P. Ozsváth and Z. Szabó. Holomorphic disks and topological invariants for closed three-manifolds. *Ann. of Math. (2)*, 159(3):1027–1158, 2004.

[105] P. Ozsváth and Z. Szabó. On knot Floer homology and lens space surgeries. *Topology*, 44(6):1281–1300, 2005.

[106] P. Ozsváth and Z. Szabó. Holomorphic disks, link invariants and the multi-variable Alexander polynomial. *Algebr. Geom. Topol.*, 8(2):615–692, 2008.

[107] P. S. Ozsváth, A. I. Stipsicz, and Z. Szabó. *Grid homology for knots and links*, volume 208 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2015.

[108] Peter Ozsváth, András I. Stipsicz, and Zoltán Szabó. Knot lattice homology in L-spaces. *J. Knot Theory Ramifications*, 25(1):1650003, 24, 2016.

[109] John Pardon. An algebraic approach to virtual fundamental cycles on moduli spaces of pseudo-holomorphic curves. *Geom. Topol.*, 20(2):779–1034, 2016.

[110] T. Perutz. Hamiltonian handleslides for Heegaard Floer homology. In *Proceedings of Gökova Geometry-Topology Conference 2007*, pages 15–35. Gökova Geometry/Topology Conference (GGT), Gökova, 2008.

[111] Béla András Rácz. *Geometry of (1,1)-Knots and Knot Floer Homology*. ProQuest LLC, Ann Arbor, MI, 2015. Thesis (Ph.D.)–Princeton University.

[112] J. Rasmussen. *Floer homology and knot complements*. PhD thesis, Harvard University, 2003.

[113] J. Rasmussen. Knot polynomials and knot homologies. In *Geometry and topology of manifolds*, volume 47 of *Fields Inst. Commun.*, pages 261–280. Amer. Math. Soc., Providence, RI, 2005.

[114] J. Rasmussen. Khovanov homology and the slice genus. *Invent. Math.*, 182(2):419–447, 2010.

[115] J. A. Rasmussen. Floer homologies of surgeries on two-bridge knots. *Algebr. Geom. Topol.*, 2:757–789, 2002.

[116] K. Reidemeister. *Knotentheorie*. Springer-Verlag, Berlin, 1974. Reprint.

[117] Joel Robbin and Dietmar Salamon. The Maslov index for paths. *Topology*, 32(4):827–844, 1993.

[118] Joel Robbin and Dietmar Salamon. The spectral flow and the Maslov index. *Bull. London Math. Soc.*, 27(1):1–33, 1995.

[119] D. Rolfsen. *Knots and links*, volume 7 of *Mathematics Lecture Series*. Publish or Perish Inc., Houston, TX, 1990. Corrected reprint of the 1976 original.

[120] Dietmar Salamon. Lectures on Floer homology. In *Symplectic geometry and topology (Park City, UT, 1997)*, volume 7 of *IAS/Park City Math. Ser.*, pages 143–229. Amer. Math. Soc., Providence, RI, 1999.

[121] S. Sarkar. Grid diagrams and the Ozsváth-Szabó tau-invariant. *Math. Res. Lett.*, 18(6):1239–1257, 2011.

[122] Sucharit Sarkar. Maslov index formulas for Whitney n-gons. *J. Symplectic Geom.*, 9(2):251–270, 2011.

[123] M. Scharlemann. Heegaard splittings of 3-manifolds. In *Low dimensional topology*, volume 3 of *New Stud. Adv. Math.*, pages 25–39. Int. Press, Somerville, MA, 2003.

[124] Horst Schubert. Knoten mit zwei Brücken. *Math. Z.*, 65:133–170, 1956.

[125] M. Schwarz. *Morse homology*, volume 111 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 1993.

[126] P. Seidel. *Fukaya categories and Picard-Lefschetz theory*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2008.

[127] J.-P. Serre. *A course in arithmetic*. Graduate Texts in Mathematics, No. 7. Springer-Verlag, New York-Heidelberg, 1973. Translated from the French.

[128] J. Singer. Three-dimensional manifolds and their Heegaard diagrams. *Trans. Amer. Math. Soc.*, 35(1):88–111, 1933.

[129] Norman Steenrod. *The Topology of Fibre Bundles*. Princeton Mathematical Series, vol. 14. Princeton University Press, Princeton, N. J., 1951.

[130] Z. Szabó and P. S. Ozsváth. Algebras with matchings and knot Floer homology. arXiv:1912.01657. 165pp. Submitted.

[131] Z. Szabó and P. S. Ozsváth. Kauffman states, bordered algebras, and a bigraded knot invariant. *Adv. Math.*, 328:1088–1198, 2018.

[132] Z. Szabó and P. S. Ozsváth. Bordered knot algebras with matchings. *Quantum Topol.*, 10(3):481–592, 2019.

[133] Clifford Henry Taubes. The Seiberg-Witten invariants and symplectic forms. *Math. Res. Lett.*, 1(6):809–822, 1994.

[134] W. Thurston. Some simple examples of symplectic manifolds. *Proc. Amer. Math. Soc.*, 55(2):467–468, 1976.

[135] W. Thurston. A norm for the homology of 3-manifolds. *Mem. Amer. Math. Soc.*, 59(339):i–vi and 99–130, 1986.

[136] William P. Thurston. *Three-dimensional geometry and topology. Vol. 1*, volume 35 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1997.

[137] H. F. Trotter. Non-invertible-knots exist. *Topology*, 2:278–280, 1963.

[138] V. Turaev. *Torsions of* 3*-dimensional manifolds*, volume 208 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 2002.

[139] Y. Huang V. Gripp Barros Ramos. An absolute grading on Heegaard Floer homology by homotopy classes of oriented 2-plane fields. arXiv:1109.2168.

[140] B. L. van der Waerden. *Algebra. Vol. I*. Springer-Verlag, New York, german edition, 1991. Based in part on lectures by E. Artin and E. Noether.

[141] J. Varouchas. Stabilité de la classe des variétés kählériennes par certains morphismes propres. *Invent. Math.*, 77(1):117–127, 1984.

[142] C. Voisin. On the homotopy types of Kähler manifolds and the birational Kodaira problem. *J. Differential Geom.*, 72(1):43–71, 2006.

[143] C. Voisin. *Hodge theory and complex algebraic geometry. I*, volume 76 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, english edition, 2007. Translated from the French by Leila Schneps.

[144] A. H. Wallace. Modifications and cobounding manifolds. *Canadian J. Math.*, 12:503–528, 1960.

[145] H. Whitney. On singularities of mappings of euclidean spaces. I. Mappings of the plane into the plane. *Ann. of Math. (2)*, 62:374–410, 1955.

[146] E. Witten. Monopoles and four-manifolds. *Math. Res. Lett.*, 1(6):769–796, 1994.

Index