

ANNALS OF MATHEMATICS

An algorithm for computing some Heegaard Floer homologies

By SUCHARIT SARKAR and JIAJUN WANG



SECOND SERIES, VOL. 171, NO. 2

March, 2010

ANMAAH

An algorithm for computing some Heegaard Floer homologies

By SUCHARIT SARKAR and JIAJUN WANG

Abstract

In this paper, we give an algorithm to compute the hat version of Heegaard Floer homology of a closed oriented three-manifold. This method also allows us to compute the filtration coming from a null-homologous link in a three-manifold.

1. Introduction

Heegaard Floer homology is a collection of invariants for closed oriented three-manifolds, introduced by Peter Ozsváth and Zoltán Szabó [14], [13]. There are four versions, denoted by \widehat{HF} , HF^∞ , HF^+ and HF^- , which are graded abelian groups. The hat version $\widehat{HF}(Y)$ is defined as the homology of a chain complex $\widehat{CF}(Y)$ coming from a Heegaard diagram of the three-manifold Y . The differentials count the number of points in certain moduli spaces of holomorphic disks, which are hard to compute in general.

There is also a relative version of the theory corresponding to pairs (Y, K) , where K is a knot in Y . If K is null-homologous, then a Seifert surface S of K induces a filtration of the chain complex $\widehat{CF}(Y)$, and the chain homotopy type of the filtered chain complex is a knot invariant. The homology groups $\widehat{HFK}(Y, K)$ of successive quotients of filtration levels are called knot Floer homology groups ([12], [18], [16]).

A cobordism between two three-manifolds induces homomorphisms on the Heegaard Floer homology groups of the two three-manifolds. In fact, the homomorphisms on HF^- and HF^+ can be used to construct an invariant of smooth four-manifolds with $b_2^+ > 1$ ([15]), called the Ozsváth-Szabó invariant. Conjecturally, the Ozsváth-Szabó four-manifold invariant is equivalent to the gauge-theoretic Seiberg-Witten invariant.

Heegaard Floer homology turns out to be a fruitful and powerful theory in the study of three-dimensional and four-dimensional topology. It gives an alternate proof of the Donaldson diagonalization theorem and the Thom conjecture for

\mathbb{CP}^2 ([8]). Heegaard Floer homology also detects the Thurston norm of a three-manifold ([17], [6]). Moreover, knot Floer homology detects the genus ([11]) and fiberedness ([1], [7], [2]) of knots and links in the three-sphere. There is an invariant τ coming from the knot filtration, whose absolute value gives a lower-bound of the slice genus for knots in the three-sphere ([9]).

Despite its success, there was no general method to compute the invariants. There were combinatorial descriptions in certain special cases, but the computation for an arbitrary three-manifold was an open problem ([10]). In this paper, we give an algorithm to compute $\widehat{\text{HF}}(Y)$ for a three-manifold Y , and also $\widehat{\text{HFK}}(Y, K)$ for a knot K in any three-manifold. All our computations will be done with coefficients in $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$. We show that one can always find Heegaard diagrams satisfying certain properties (Definition 3.1). Using such Heegaard diagrams, which we call nice, it will be easy to compute $\widehat{\text{HF}}$ and $\widehat{\text{HFK}}$. Our main results are summarized in the following theorems.

THEOREM 1.1. *Given a nice Heegaard diagram of a closed oriented three-manifold Y , $\widehat{\text{HF}}(Y)$ can be computed combinatorially. Similarly, for a knot $K \subset Y$, $\widehat{\text{HFK}}(Y, K)$ can be computed combinatorially in a nice Heegaard diagram.*

THEOREM 1.2. *Every closed oriented three-manifold Y admits a nice Heegaard diagram. For a null-homologous knot K in a closed oriented three-manifold Y , the pair (Y, K) admits a compatible nice Heegaard diagram. In fact, there is an algorithm to convert any pointed Heegaard diagram to a nice Heegaard diagram via isotopies and handleslides.*

It will be interesting to compare our result with the recent work of Ciprian Manolescu, Peter Ozsváth and the first author in [5], where they gave a combinatorial description of knot Floer homology of knots in S^3 , in all versions.

We hope this method can be generalized to compute some of the other versions, notably $\text{HF}^-(Y)$ and $\text{HF}^+(Y)$. It would also be nice to have a proof of the invariance of the combinatorial description without using holomorphic disks.

The paper is organized as follows. In Section 2, we give an overview of certain concepts in Heegaard Floer theory. In Section 3, we give a combinatorial characterization of index one holomorphic disks in nice Heegaard diagrams. In Section 4, we give an algorithm to get such Heegaard diagrams. In Section 5, we give examples to demonstrate our algorithm for three-manifolds and knots in the three-sphere.

2. Preliminaries

In this section, we review the definition of Heegaard Floer homology. See [14], [13], [16], [3] for details.

2.1. *Definition of $\widehat{\text{HF}}$.* The Heegaard Floer homology of a closed oriented three-manifold Y is defined from a pointed Heegaard diagram representing Y .

A Heegaard splitting of Y is a decomposition of Y into two handlebodies glued along their boundaries. We fix a self-indexing Morse function f on Y with k index zero critical points and k index three critical points. (We usually choose $k = 1$.) Then f gives a Heegaard splitting of Y , where the two handlebodies are given by $f^{-1}(-\infty, \frac{3}{2}]$ and $f^{-1}[\frac{3}{2}, \infty)$. If the number of index one critical points or the number of index two critical points of f is $(g + k - 1)$, then $\Sigma = f^{-1}(3/2)$ is a genus g surface. We fix a gradient like flow on Y corresponding to f . We require f to have the property that Y contains a disjoint union of k flow lines, each flowing from an index zero critical point to an index three critical point. We get a collection $\alpha = (\alpha_1, \dots, \alpha_{g+k-1})$ of α circles on Σ which flow down to the index one critical points, and another collection $\beta = (\beta_1, \dots, \beta_{g+k-1})$ of β circles on Σ which flow up to the index two critical points. Note that both $\Sigma \setminus \alpha$ and $\Sigma \setminus \beta$ have k components.

We fix k points w_1, \dots, w_k (called *basepoints*) in the complement of the α circles and the β circles in Σ , such that each component of $\Sigma \setminus \alpha$ contains exactly one w_i and each component of $\Sigma \setminus \beta$ contains exactly one w_j . This is equivalent to the condition that the trajectories of w_i 's under the gradient like flow, hit all the index zero and all the index three critical points. We write $\mathbf{w} = (w_1, \dots, w_k)$. The tuple $(\Sigma, \alpha, \beta, \mathbf{w})$ is called a *pointed Heegaard diagram* for Y .

There are some moves on a Heegaard diagram that do not change the underlying three-manifold. An *isotopy* moves the α curves and β curves in two one-parameter families α_t and β_t in $\Sigma \setminus \mathbf{w}$, moving by isotopy, such that the α curves remain disjoint and the β curves remain disjoint for each t . In a *handleslide* of α , we replace a pair of α curves α_i and α_j with a pair α_i and α'_j , such that the three curves α_i , α_j and α'_j bound a pair of pants in $\Sigma \setminus \mathbf{w}$ disjoint from all the other α curves. A handleslide of β is defined similarly. There is also a move called *stabilization*, but we will not be using it in the present paper. These moves are called *Heegaard moves*.

Heegaard Floer homology is a certain version of Lagrangian Floer homology. The ambient symplectic manifold is the symmetric product $\text{Sym}^{g+k-1}(\Sigma)$. The two half-dimensional totally real subspaces are the tori $T_\alpha = \alpha_1 \times \dots \times \alpha_{g+k-1}$ and $T_\beta = \beta_1 \times \dots \times \beta_{g+k-1}$. The generators for the chain complex $\widehat{\text{CF}}(\Sigma, \alpha, \beta, \mathbf{w})$ are the intersection points between these two tori, and the boundary maps are given by counting certain holomorphic disks. For more details see [14], and see [16] for the issue of *boundary degenerations* when $k > 1$.

THEOREM 2.1 (Ozsváth-Szabó, [14], [16]). *When $k = 1$, the homology of the chain complex $\widehat{\text{CF}}(\Sigma, \alpha, \beta, \mathbf{w})$ is an invariant for the three-manifold Y , written as*

$\widehat{\mathrm{HF}}(Y)$. For a general k , and Y a rational homology three-sphere, we have

$$H_*(\widehat{\mathrm{CF}}(\Sigma, \alpha, \beta, w)) \cong \widehat{\mathrm{HF}}(Y) \otimes H_*(T^{k-1}),$$

where $H_*(T^{k-1})$ is the singular homology of the $(k-1)$ -dimensional torus with coefficients in \mathbb{F}_2 .

When we have a link L in Y , we ensure that L is a union of flow lines from index zero critical points to index three critical points. We also ensure that L contains all index zero and index three critical points and contains no index one or two critical points. We orient L and Σ and define w_i 's as the positive intersection points between L and Σ . We write the other k intersections as $z = (z_1, \dots, z_k)$. Such a Heegaard diagram, denoted by $(\Sigma, \alpha, \beta, w, z)$, is called a *pointed Heegaard diagram* for the pair (Y, L) . (For a link with l components, we usually choose $k = l$.) The knot (link) Floer homology $\widehat{\mathrm{HFK}}(Y, L)$ is defined similarly, where the boundary maps count a more restricted class of holomorphic disks. See [12], [18], [16] for details.

2.2. Cylindrical reformulation of $\widehat{\mathrm{HF}}$. In the present paper, we will use the cylindrical reformulation of the Heegaard Floer homology by Lipshitz. See [3] for details.

Given a pointed Heegaard diagram $(\Sigma, \alpha, \beta, w)$, the generators of the chain complex $\widehat{\mathrm{CF}}$ are given by formal sums of $(g + k - 1)$ distinct points in Σ , $x = x_1 + \dots + x_{g+k-1}$, such that each α circle contains some x_i and each β circle contains some x_j . A connected component of $\Sigma \setminus (\alpha \cup \beta)$ is called a *region*. A formal sum of regions with integer coefficients is called a *2-chain*. Given two generators x and y , we define $\pi_2(x, y)$ to be the collection of all 2-chains ϕ such that $\partial(\partial(\phi)|_\alpha) = y - x$. Such 2-chains are called *domains*. Given a point $p \in \Sigma \setminus (\alpha \cup \beta)$, let $n_p(\phi)$ be the coefficient of the region containing p in ϕ . A domain ϕ is *positive* if $n_p(\phi) \geq 0$ for all points $p \in \Sigma \setminus (\alpha \cup \beta)$. We define $\pi_2^0(x, y) = \{\phi \in \pi_2(x, y) \mid n_{w_i}(\phi) = 0 \ \forall i\}$. A Heegaard diagram is *admissible*, if, for every generator x , any positive domain $\phi \in \pi_2^0(x, x)$ is trivial. If the three-manifold Y has $b_1(Y) > 0$, we require the Heegaard diagram to be admissible.

Fix two generators x, y and a domain $\phi \in \pi_2^0(x, y)$. Let S be a surface with boundary, with $2(g + k - 1)$ marked points $(X_1, \dots, X_{g+k-1}, Y_1, \dots, Y_{g+k-1})$ on ∂S , such that the X points and the Y points alternate. The $2(g + k - 1)$ arcs on ∂S in the complement of the marked points are divided into two groups A and B , each containing $(g + k - 1)$ arcs, such that the A arcs and the B arcs alternate. Let p_1 and p_2 be the projection maps from $\Sigma \times D^2$ onto its first and second factors. Look at maps $u : S \rightarrow \Sigma \times D^2$ such that the image of $p_1 \circ u$ is ϕ (as 2-chains) and the image of $p_2 \circ u$ is $(g + k - 1)D^2$ (in second homology). We also want the X points on ∂S to map injectively by $p_1 \circ u$ to the x_i 's and to map to $-i$ in the unit

disk by $p_2 \circ u$. Similarly, we want the Y points to map injectively to the y_i 's by $p_1 \circ u$ and to i in the unit disk by $p_2 \circ u$. Furthermore we also require the A arcs in ∂S to map to α arcs by $p_1 \circ u$ and under $p_2 \circ u$ to map to the arc e_1 in $\partial(D^2)$ joining $-i$ to i in half-plane $\operatorname{Re}(s) > 0$. Similarly, we require the B arcs to map to β arcs by $p_1 \circ u$ and to map to the arc e_2 in $\partial(D^2)$ in the half-plane $\operatorname{Re}(s) < 0$ by $p_2 \circ u$.

Now fix complex structures on Σ and D^2 and take the product complex structure on $\Sigma \times D^2$. A generic perturbation gives an almost complex structure which achieves transversality for the homology class ϕ . In our case, we can achieve this by a generic perturbation of the α curves and the β curves ([3, Lemma 3.10]).

The holomorphic embeddings u which satisfy the above conditions and whose homology class is ϕ form a moduli space, which we denote by $\mathcal{M}(\phi)$. The *Maslov index* $\mu(\phi)$ of ϕ gives the expected dimension of $\mathcal{M}(\phi)$. It can be computed combinatorially in terms of the Euler measure and the point measures, which are defined as follows. For a generator $\mathbf{x} = \sum x_i$ and a domain ϕ , $\mu_{x_i}(\phi)$ is defined to be the average of the coefficients of the four regions around x_i in ϕ . The *point measure* $\mu_{\mathbf{x}}(\phi)$ is defined as $\sum \mu_{x_i}(\phi)$. If we fix a metric on Σ which makes all the α and β circles geodesic, intersecting each other with right angles, then the *Euler measure* $e(\phi)$ is defined to be $\frac{1}{2\pi}$ of the integral of the curvature on ϕ . The Euler measure is clearly additive, and if D is a $2n$ -gon region, then $e(D) = 1 - \frac{n}{2}$.

PROPOSITION 2.2 (Lipshitz, [3]). *For a domain $\phi \in \pi_2(\mathbf{x}, \mathbf{y})$, the Maslov index is given by*

$$\mu(\phi) = e(\phi) + \mu_{\mathbf{x}}(\phi) + \mu_{\mathbf{y}}(\phi).$$

If ϕ is nontrivial, the moduli space $\mathcal{M}(\phi)$ admits a free \mathbb{R} -action coming from the one-parameter family of holomorphic automorphisms of D^2 which preserve $\pm i$ and the boundary arcs e_1 and e_2 . In particular, if $\mu(\phi) = 1$, the *unparametrized moduli space* $\mathcal{M}(\phi)/\mathbb{R}$ is a zero-dimensional manifold, and then the *count function* $c(\phi)$ is defined to be the number of points in $\mathcal{M}(\phi)/\mathbb{R}$, counted modulo 2. The boundary map in the chain complex $\widehat{\text{CF}}$ is given by

$$\partial \mathbf{x} = \sum_{\mathbf{y}} \sum_{\{\phi \in \pi_2^0(\mathbf{x}, \mathbf{y}) \mid \mu(\phi)=1\}} c(\phi) \mathbf{y}.$$

THEOREM 2.3 (Lipshitz, [3]). *For a three-manifold Y , the homology of the chain complex $(\widehat{\text{CF}}, \partial)$ is isomorphic to $H_*(\widehat{\text{CF}}(\Sigma, \alpha, \beta, \mathbf{w}))$.*

Note that the only noncombinatorial part of the theory is the count function $c(\phi)$.

2.3. Positivity of domains with holomorphic representatives. We will need the following proposition, which asserts that only positive domains can have holomorphic representatives.

PROPOSITION 2.4. *Let ϕ be a domain in $\pi_2^0(\mathbf{x}, \mathbf{y})$. If ϕ has a holomorphic representative, then ϕ is a positive domain. In particular, if $c(\phi) \neq 0$, then ϕ is a positive domain.*

Proof. If ϕ has a holomorphic representative, then there exists some holomorphic embedding u of the type described above. Then for any point $p \in \Sigma \setminus (\alpha \cup \beta)$, $n_p(\phi)$ is simply the intersection number of $u(S)$ and $\{p\} \times D^2$. Since both of them are holomorphic objects in the product complex structure, they have positive intersection number and hence $n_p(\phi) \geq 0$. Here we require the complex structure on $\Sigma \times D^2$ to be standard near the basepoints. See [3] for a general discussion. \square

If a domain ϕ has a holomorphic representative, the number of branch points of $p_2 \circ u$ is given by $\mu_{\mathbf{x}} + \mu_{\mathbf{y}} - e(\phi)$ ([3], [18]). Furthermore, in such a situation the Maslov index can also be calculated as $\mu(\phi) = 2e(\phi) + g + k - 1 - \chi(S) = e(\phi) + b + \frac{1}{2}(g + k - 1 - t)$, where b denotes the number of branch points of $p_1 \circ u$, and t denotes the number of *trivial disks*, i.e. the components of S which are mapped to a point by $p_1 \circ u$ (which correspond to coordinates x_i of \mathbf{x} with $\mu_{x_i} = 0$).

3. Holomorphic disks in nice Heegaard diagrams

In this section, we study index one holomorphic disks in nice Heegaard diagrams.

Definition 3.1. Let $\mathcal{H} = (\Sigma, \alpha, \beta, \mathbf{w})$ be a pointed Heegaard diagram for a three-manifold Y . \mathcal{H} is called *nice* if any region that does not contain any basepoint w_i in \mathbf{w} is either a bigon or square.

Let Y be a closed oriented three-manifold. Suppose Y has a nice admissible Heegaard diagram $\mathcal{H} = (\Sigma, \alpha, \beta, \mathbf{w})$. We choose a product complex structure on $\Sigma \times D^2$.

Definition 3.2. A domain $\phi \in \pi_2^0(\mathbf{x}, \mathbf{y})$ with coefficients 0 and 1 is called an *empty embedded $2n$ -gon*, if it is topologically an embedded disk with $2n$ vertices on its boundary, such that at each vertex v , $\mu_v(\phi) = \frac{1}{4}$, and it does not contain any x_i or y_i in its interior.

The following two theorems show that, for a domain $\phi \in \pi_2^0(\mathbf{x}, \mathbf{y})$, the count function $c(\phi) \neq 0$ if and only if ϕ is an empty embedded bigon or an empty embedded square, and in that case $c(\phi) = 1$. Thus $c(\phi)$ can be computed combinatorially in a nice Heegaard diagram.

THEOREM 3.3. *Let $\phi \in \pi_2^0(\mathbf{x}, \mathbf{y})$ be a domain such that $\mu(\phi) = 1$. If ϕ has a holomorphic representative, then ϕ is an empty embedded bigon or an empty embedded square.*

Proof. We know that only positive domains can have holomorphic representatives. We also know that bigons and squares have nonnegative Euler measure. We will use these facts to limit the number of possible cases.

Suppose $\phi = \sum a_i D_i$, where D_i 's are regions containing no basepoints. Since ϕ has a holomorphic representative, we have $a_i \geq 0, \forall i$. Since each D_i is a bigon or a square, we have $e(D_i) \geq 0$ and hence $e(\phi) \geq 0$. So, by Lipshitz' formula $\mu(\phi) = e(\phi) + \mu_x(\phi) + \mu_y(\phi)$, we get $0 \leq \mu_x + \mu_y \leq 1$.

Now let $x = x_1 + \cdots + x_g$ and $y = y_1 + \cdots + y_g$, with $x_i, y_i \in \alpha_i$. We say ϕ hits some α circle if $\partial\phi$ is nonzero on some part of that α circle. Since $\phi \neq n\Sigma$, it has to hit at least one α circle, say α_1 , and hence $\mu_{x_1}, \mu_{y_1} \geq \frac{1}{4}$ as $\partial(\partial\phi|_{\alpha}) = y - x$. Also if ϕ does not hit α_i , then $x_i = y_i$ and they must lie outside the domain ϕ , since otherwise we have $\mu_{x_i} = \mu_{y_i} \geq \frac{1}{2}$ and hence $\mu_x + \mu_y$ becomes too large.

We now note that $e(\phi)$ can only take half-integral values, and thus only the following cases might occur:

- *Case 1.* ϕ hits α_1 and another α circle, say α_2 , ϕ consists of squares, $\mu_{x_1} = \mu_{x_2} = \mu_{y_1} = \mu_{y_2} = \frac{1}{4}$, and there are $(g + k - 3)$ trivial disks.
- *Case 2.* ϕ hits α_1 , $D(\phi)$ consists of squares and exactly one bigon, $\mu_{x_1} = \mu_{y_1} = \frac{1}{4}$, and there are $(g + k - 2)$ trivial disks.
- *Case 3.* ϕ hits α_1 , $D(\phi)$ consists of squares, $\mu_{x_1} + \mu_{y_1} = 1$, and there are $(g + k - 2)$ trivial disks.

Using the reformulation by Lipshitz, in each of these cases, we will try to figure out the surface S which maps to $\Sigma \times D^2$. Recall that a trivial disk is a component of S which maps to a point in Σ after post-composing with the projection $\Sigma \times D^2 \rightarrow \Sigma$.

The first case corresponds to a map from S to Σ with $\chi(S) = (g + k - 2)$, and S has $(g + k - 3)$ trivial disk components. If the rest of S is F , then F is a double branched cover over D^2 with $\chi(F) = 1$ and 1 branch point (for holomorphic maps, the number of branch points is given by $\mu_x + \mu_y - e(\phi)$); i.e., F is a disk with four marked points on its boundary. Call the marked points corners, and call F a square.

In the other two cases, S has $(g + k - 2)$ trivial disk components, so if F denotes the rest of S , then F is just a single cover over D^2 . Thus the number of branch points has to be 0. But in the third case the number of branch points is 1, so the third case cannot occur. In the second case, F is a disk with two marked points on its boundary. Call the marked points corners, and call F a bigon.

Thus in both the first and the second cases, ϕ is the image of F and all the trivial disks map to the x -coordinates (which are also the y -coordinates) which do not lie in ϕ . Note that in both cases, the map from F to ϕ has no branch point, so it is a local diffeomorphism, even at the boundary of F . Furthermore using the

condition that μ_{x_i} (or μ_{y_i}) = $\frac{1}{4}$ whenever it is nonzero, we conclude that there is exactly one preimage for the image of each corner of F .

All we need to show is that the map from F to Σ is an embedding, or in other words, the local diffeomorphism from F to ϕ is actually a diffeomorphism. We will prove this case by case.

Case 1. In this case we have an immersion $f : F \rightarrow \Sigma$, where F is a square (with boundary). Look at the preimage of all the α and β circles in F . Using the fact f is a local diffeomorphism, we see that each of the preimages of α and β arcs are also 1-manifolds, and by an abuse of notation, we will also call them α or β arcs. Using the embedding condition near the four corners, we see that at each corner only one α arc and only one β arc can come in. The different α arcs cannot intersect and the different β arcs cannot intersect, and all intersections between α and β arcs are transverse.

Note that since the preimage of each square region is a square, F (with all the α and β arcs) is also tiled by squares. Thus the α arcs in F cannot form a closed loop, for in that case $F \setminus \{\text{inside of loop}\}$ has negative Euler measure and hence cannot be tiled by squares. Similarly the β arcs cannot form a loop. Also no α arc can enter and leave F through the same β arc on the boundary, for again the outside will have negative Euler measure. Thus the α arcs slice up F into vertical rectangles, and in each rectangle, no β arc can enter and leave through the same α arc. This shows that the α arcs and β arcs make the standard co-ordinate chart on F , as in Figure 1.

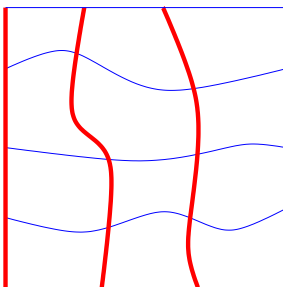


Figure 1. *Preimage of α and β arcs for a square.* We make the convention for all figures in the paper that the thick solid arcs denote α arcs and the thin solid arcs denote β arcs.

We call the intersection points between α and β arcs in F vertices (and we are still calling the four original vertices on the boundary of the square F corners). Note that to show f is an embedding, it is enough to show that no two different vertices map to the same point. Assume $p, q \in F$ are distinct vertices with $f(p) = f(q)$. There could be two subcases:

- Both p and q are in $\overset{\circ}{F}$.
- At least one of p and q is in ∂F .

We will reduce the first subcase to the second. Assume both p and q are in the interior of F . Choose a direction on the α arc passing through $f(p) = f(q)$ in Σ , and keep looking at successive points of intersection with β arcs, and locate their inverse images in F . For each point, we will get at least a pair of inverse images, one on the α arc through p , and one on the α arc through q , until one of the points falls on ∂F , and thus we have reduced it to the second subcase.

In the second subcase, without loss of generality, we assume that p lies on a β arc on ∂F . Then choose a direction on the β arc in Σ through $f(p) = f(q)$ and proceed as above, until one of the preimages hits an α arc on ∂F . If that preimage is on the β arc through q , then reverse the direction and proceed again, and this time we can ensure that the preimage which hits α arc on ∂F first is the one that was on the β arc through p . Thus we get two distinct vertices in F mapping to the same point in Σ , one of them being a corner. This is a contradiction to the embedding assumption near the corners.

Case 2. In this case we have an immersion $f : F \rightarrow \Sigma$ with F being a bigon. Again look at the preimage of α and β circles. All intersections will be transverse (call them vertices), and at each of the two corners there can be only one α arc and only one β arc. Again there cannot be any closed loops. We get an induced tiling on F with squares and 1 bigon.

This time the α arcs can (in fact they have to) enter and leave F through the same β arc, but they have to do it in a completely nested fashion; i.e., there is only one bigon piece in $F \setminus \alpha$, the “innermost bigon”. Thus F decomposes into two pieces, the innermost bigon and the rest. In case there are no α arcs in $\overset{\circ}{F}$, the rest might be empty, but otherwise it is a square. From the arguments in the earlier case, the β arcs must cut up the square piece in a standard way, and from the previous argument the β arcs must enter and leave the bigon in a nested fashion, as in Figure 2.

Again to show f is an embedding, it is enough to show that it is an embedding restricted to vertices. Take two distinct vertices p, q mapping to the same point, and follow them along α arcs in some direction, until one of them hits a β arc on ∂F . Then follow them along β arcs, and there exists some direction such that one of them will actually hit a corner, giving the required contradiction.

So in either case, f is an embedding. □

THEOREM 3.4. *If $\phi \in \pi_2^0(x, y)$ is an empty embedded bigon or an empty embedded square, then the product complex structure on $\Sigma \times D^2$ achieves transversality for ϕ under a generic perturbation of the α and the β circles, and $\mu(\phi) = c(\phi) = 1$.*

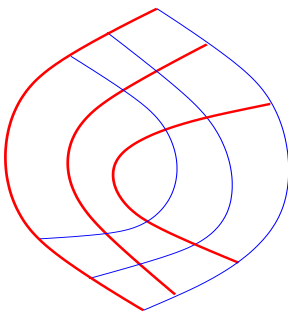


Figure 2. Preimage of α and β arcs for a bigon.

Proof. Let ϕ be an empty embedded $2n$ -gon. Each of the corners of ϕ must be an x -coordinate or a y -coordinate, and at every other x (resp. y) coordinate the point measure μ_{x_i} (resp. μ_{y_i}) is zero. Therefore $\mu_x(\phi) + \mu_y(\phi) = 2n \cdot \frac{1}{4} = \frac{n}{2}$. Also ϕ is topologically a disk, so it has Euler characteristic 1. Since it has $2n$ corners each with an angle of $\frac{\pi}{4}$, the Euler measure $e(\phi) = 1 - \frac{2n}{4} = 1 - \frac{n}{2}$. Thus the Maslov index $\mu(\phi) = 1$.

By [3, Lemma 3.10], we see that ϕ satisfies the boundary injective condition, and hence under a generic perturbation of the α and the β circles, the product complex structure achieves transversality for ϕ .

When ϕ is an empty embedded square, we can choose F to be a disk with four marked points on its boundary, which is mapped to ϕ diffeomorphically. Given a complex structure on Σ , the holomorphic structure on F is determined by the cross-ratio of the four points on its boundary, and there is an one-parameter family of positions of the branch point in D^2 which gives that cross-ratio. Thus there is a holomorphic branched cover $F \rightarrow D^2$ satisfying the boundary conditions, unique up to reparametrization. Hence ϕ has a holomorphic representative, and from the proof of Theorem 3.3 we see that this determines the topological type of F , and hence it is the unique holomorphic representative.

When ϕ is an empty embedded bigon, we can choose F to be a disk with two marked points on its boundary, which is mapped to ϕ diffeomorphically. A complex structure on Σ induces a complex structure on F , and there is a unique holomorphic map from F to the standard D^2 after reparametrization. Thus again ϕ has a holomorphic representative, and similarly it must be the unique one. \square

Proof of Theorem 1.1. Theorems 3.3 and 3.4 make the count function $c(\phi)$ combinatorial in a nice Heegaard diagram. For a domain $\phi \in \pi_2^0(x, y)$ with $\mu(\phi) = 1$, we have $c(\phi) = 1$ if ϕ is an empty embedded bigon or an empty embedded square, and $c(\phi) = 0$ otherwise. \square

4. Algorithm to get nice Heegaard diagrams

In this section, we prove Theorem 1.2. We will demonstrate an algorithm which, starting with an admissible pointed Heegaard diagram, gives an admissible nice Heegaard diagram by doing isotopies and handleslides on the β curves.

For a Heegaard diagram, we call bigon and square regions *good* and all other regions *bad*. We will first do some isotopies to ensure all the regions are disks. We will then define a complexity for the Heegaard diagram which attains its minimum only if all the regions not containing the basepoints are good. We will do an isotopy or a handleslide which will decrease the complexity if the complexity is not the minimal one.

4.1. *The algorithm.* Let $\mathcal{H} = (\Sigma, \alpha, \beta, w)$ be a pointed Heegaard diagram with a single basepoint w . We consider Heegaard diagrams with more basepoints in the last subsection.

Step 1. Killing nondisk regions. We do finger moves on β circles to create new intersections with α circles. After doing this sufficiently many times, every region in \mathcal{H} becomes a disk. We first ensure that every α circle intersects some β circle and every β circle intersects some α circle.

If α_i does not intersect any β circle, we can find an arc c connecting α_i to some β_j avoiding the intersections of α and β circles, as indicated in Figure 3(a). We can select c such that c intersects β_j just at the endpoint. Doing a finger move of β_j along c as in Figure 3(b) will make α_i intersect some β circle.

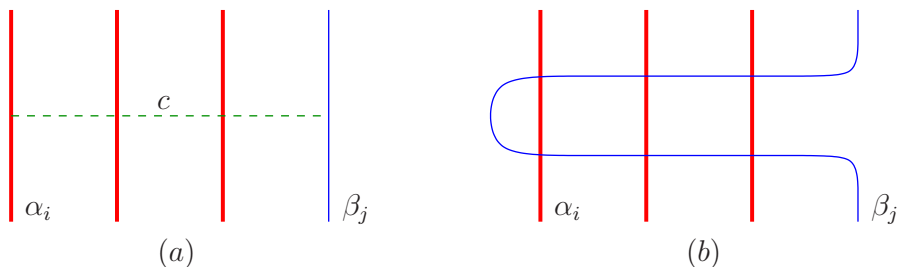


Figure 3. Making each α circle intersect some β circle.

Similarly, if β_i does not intersect any α circle, we find an arc c connecting β_i to some α_j so that $c \cap \alpha$ contains a single point as in Figure 4(a). We then do the operation as depicted in Figure 4(b).

Repeating the above process, we can make sure that every α circle intersects some β circle and every β circle intersects some α circle.

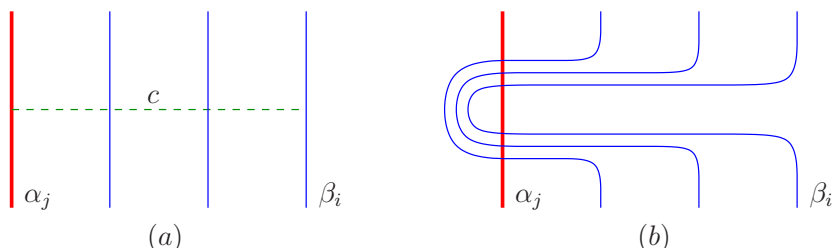


Figure 4. Making each β circle intersect some α circle.

Note that the complement of the α curves is a punctured sphere. Thus every region is a planar surface. A nondisk region D has more than one boundary component. Every boundary component must contain both α and β arcs since every α (resp. β) circle intersects some β (resp. α) circle. Then we make a finger move on the β curve to reduce the number of boundary components of D without generating other nondisk regions. See Figure 5 for this finger move operation. Repeating this process as many times as necessary, we will kill all the nondisk regions.

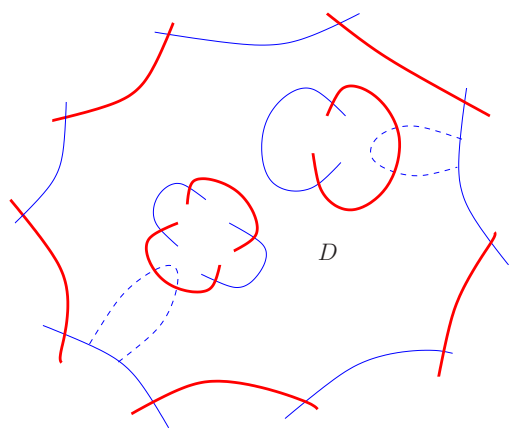


Figure 5. Killing nondisk regions. The dotted arcs indicate our finger moves. After our finger move, the region D becomes a disk region.

Step 2. Making all but one region bigons or squares. We consider Heegaard diagrams with only disk regions. Note that our algorithm will not generate nondisk regions.

Let D_0 be the disk region containing the basepoint w . For any region D , pick an interior point $w' \in D$ and define the *distance* of D , denoted by $d(D)$, to be the smallest number of intersection points between the β curves and an arc connecting w and w' in the complement of the α circles. For a $2n$ -gon disk region D , define the *badness* of D as $b(D) = \max\{n - 2, 0\}$.

For a pointed Heegaard diagram \mathcal{H} with only disk regions, define the *distance* $d(\mathcal{H})$ of \mathcal{H} to be the largest distance of bad regions. Define the *distance d complexity* of \mathcal{H} to be tuple

$$c_d(\mathcal{H}) = \left(\sum_{i=1}^m b(D_i), -b(D_1), -b(D_2), \dots, -b(D_m) \right),$$

where D_1, \dots, D_m are all the distance d bad regions, ordered so that $b(D_1) \geq b(D_2) \geq \dots \geq b(D_m)$. We call the first term the *total badness of distance d* of \mathcal{H} , and denote it by $b_d(\mathcal{H})$. If there are no distance d bad regions, then $c_d(\mathcal{H}) = (0)$. We order the set of distance d complexities lexicographically.

LEMMA 4.1. *For a distance d pointed Heegaard diagram \mathcal{H} with only disk regions, if $c_d(\mathcal{H}) \neq (0)$, we can modify \mathcal{H} by isotopies and handleslides to get a new Heegaard diagram \mathcal{H}' with only disk regions, satisfying $d(\mathcal{H}') \leq d(\mathcal{H})$ and $c_d(\mathcal{H}') < c_d(\mathcal{H})$.*

Proof. We order the bad regions of distance d as in the definition of the distance d complexity. Now we look at D_m . It is a $(2n)$ -gon with $n \geq 3$. Pick an adjacent region D_* with distance $d - 1$ having a common β edge with D_m . Let b_* be (one of) their common β edge(s). We order the α edges of D_m counterclockwise, and denote them by a_1, a_2, \dots, a_n starting at b_* .

We try to make a finger move on b_* into the D_m and out of D_m through a_2 , as indicated in Figure 6 when D_m is an octagon. Our finger will separate D_m into two parts, $D_{m,1}$ and $D_{m,2}$.

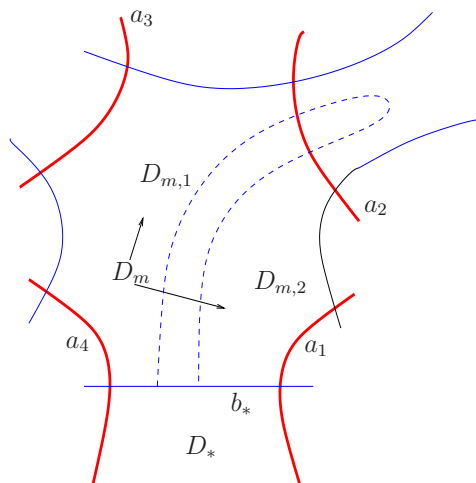


Figure 6. Starting our finger move.

If we reach a square region of distance $\geq d$, we push up our finger outside the region via the opposite edge, as in Figure 7. Note that doing a finger move through regions of distance $\geq d$ does not change the distance of any of the bad regions, since they all have distance $\leq d$.

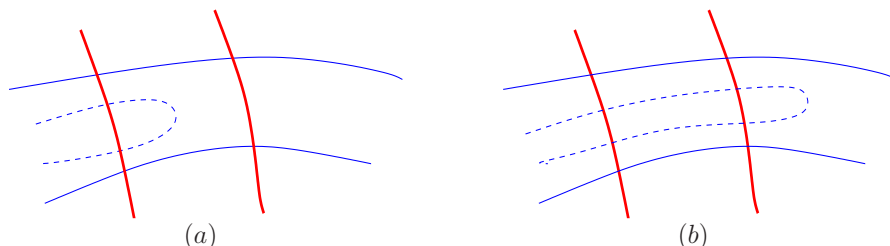


Figure 7. Moving across a square region.

We continue to push up our finger as far as possible, until we reach one of the following:

- (1) a bigon region.
- (2) a region with distance $\leq d - 1$.
- (3) a bad region with distance d other than D_m , i.e., D_i with $i < m$.
- (4) D_m .

We will prove our lemma case by case.

Case 1. A bigon is reached. Before we reach the bigon region, all regions in between are square regions with distance $\geq d$. After our finger moves inside a bigon region, our finger separates the bigon into a square and a new bigon, as in Figure 8.

Denote the new Heegaard diagram by \mathcal{H}' . We have $b(D_{m,1}) = b(D_m) - 1$. Since $D_{m,2}$ is a square and is good, we get $b_d(\mathcal{H}') = b_d(\mathcal{H}) - 1$. Note that we will not increase the distance of any bad region since we do not pass through any region of distance $\leq d - 1$ and all bad regions has distance $\leq d$. Hence $d(\mathcal{H}') \leq d(\mathcal{H})$ and $c_d(\mathcal{H}') < c_d(\mathcal{H})$.

Case 2. A smaller distance region is reached. Let D' be the region with distance $< d$ we reached by our finger. Suppose $d(D') = d'$. Let \mathcal{H}' be the new Heegaard diagram. See Figure 9. Note that D' might be a bigon, which could be covered in both Case 1 and Case 2.

We have $b(D_{m,1}) = b(D_m) - 1$ and $D_{m,2}$ is good. Our finger separates D' into a bigon region D'_1 and the other part D'_2 . When D' is a square or a bad region, D'_2 will be a bad region of distance $d' < d$. We might have increased the distance d' complexity, but we have $d(\mathcal{H}') \leq d(\mathcal{H})$ and $c_d(\mathcal{H}') < c_d(\mathcal{H})$.

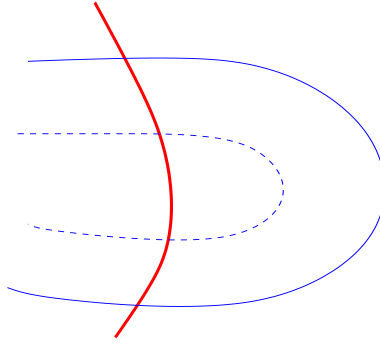


Figure 8. *Case 1. A bigon is reached.*

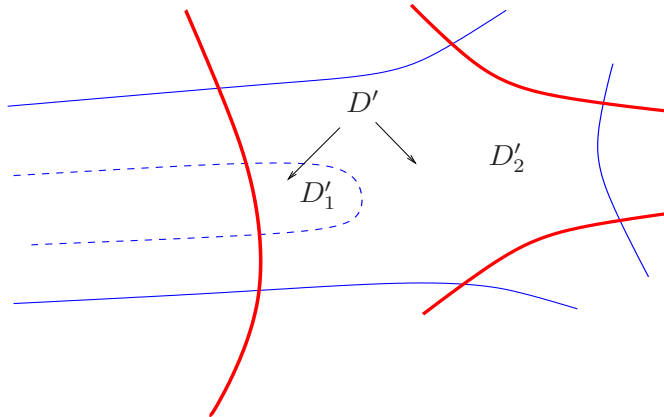


Figure 9. *Case 2. A smaller distance region is reached.*

Case 3. Another distance d bad region is reached. In this case, we reach some distance d bad region D_i with $i < m$. See Figure 10 for an indication. Denote by $D_{i,1}$ and $D_{i,2}$ the two parts of D_i separated by our finger. Then $D_{i,1}$ is good while $D_{i,2}$ is a bad region of distance d . We have $b(D_{i,2}) = b(D_i) + 1$ and $b(D_{m,1}) = b(D_m) - 1$. Thus the total badness of distance d remains the same. But we are decreasing the distance d complexity since we are moving the badness from a later bad region to an earlier bad region. Hence for the new Heegaard diagram \mathcal{H}' , we have $d(\mathcal{H}') = d(\mathcal{H})$ and $c_d(\mathcal{H}') < c_d(\mathcal{H})$.

Case 4. Coming back to D_m . This is the worst case and we need to pay more attention. We divide this case into two subcases, according to which edge the finger is coming back through.

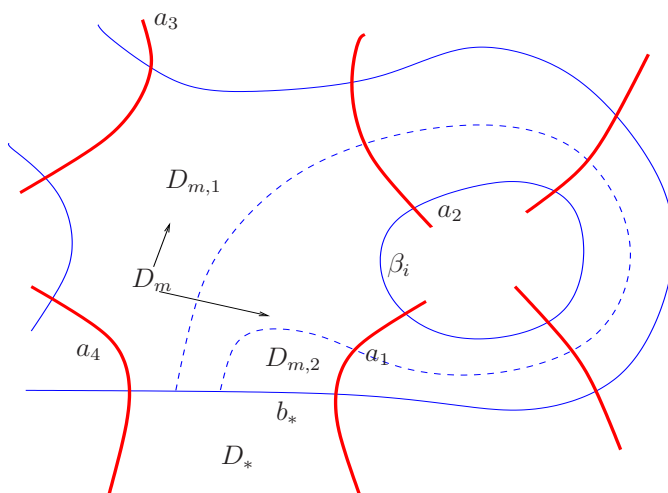


Figure 12. Cases 4.1 and 4.2 Coming back via an adjacent edge-handleslide. The dotted arc denotes the β curve after the handleslide.

$D_{m,2}$ is a bigon region and $b(D_{m,1}) = b(D_m) - 1$. Thus for the new Heegaard diagram \mathcal{H}' after the handleslide, the total badness of distance d is decreased by 1. We have $d(\mathcal{H}') \leq d(\mathcal{H})$ and $c_d(\mathcal{H}') < c_d(\mathcal{H})$.

Subcase 4.2. Coming back via a nonadjacent edge. If we return through a_k with $3 < k \leq n$, then, instead of the finger move through a_2 , we do a finger move through a_3 (starting from b_*). If we reach one of the first three cases, we are decreasing the distance d complexity by similar arguments as before.

Suppose instead that we come back to D_m , say via a_i . We claim that $3 < i < k$. Certainly we can not come back via a_3 . The finger can not come back via a_k since the chain of squares from a_k is connected to a_2 . If $i > k$ or $i < 3$, we could close the cores the two fingers to get two simple closed curves c_1 and c_2 , as indicated in Figure 13. Then c_1 and c_2 intersect transversely at exactly one point and they are in the complement of the β curves. The complement of the β curves is a punctured sphere. Attach disks to get a sphere. Then as homology classes, we get $[c_1] \cdot [c_2] = 1$. But $H_1(S^2) \cong 0$. This is a contradiction. Thus we must have $3 < i < k$. (The argument of this claim was suggested by Dylan Thurston.)

Now, instead of the finger move through a_3 , we do another finger move through a_4 . Continuing the same arguments, we see that we either end up with a finger which does not come back, or we get some finger that starts at a_j and comes back via a_{j+1} . If the finger does not come back, we reduce it to the previous cases and the lemma follows.

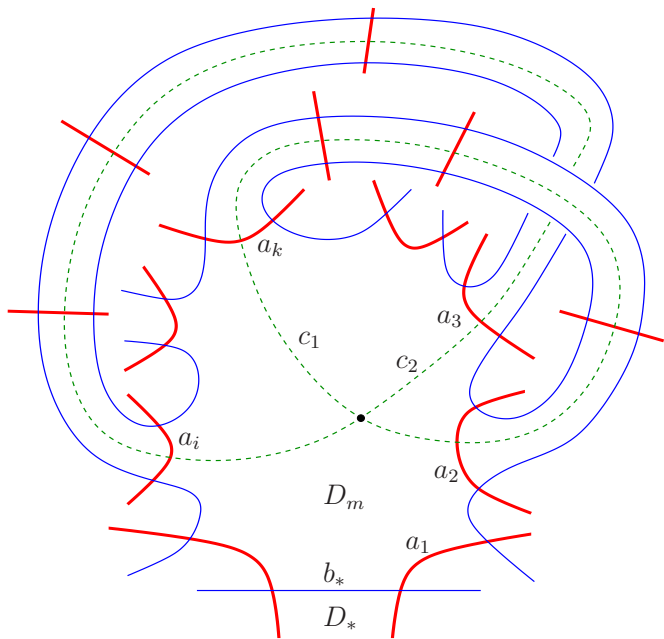


Figure 13. *Case 4.2 There are no crossing fingers. The fingers are not showed here. Instead, the two dotted arcs denote the cores of the two fingers.*

If there is a finger which starts at a_j and comes back at a_{j+1} , we see a full β circle. We do a handleslide similar to the one in Subcase 4.1. We have $b(D_{m,1}) = \max\{n - j - 1, 0\}$ and $b(D_{m,2}) = \max\{j - 2, 0\}$. We also have $b(D_{m,1}) + b(D_{m,2}) \leq n - 3$. Thus for the new Heegaard diagram \mathcal{H}' after the handleslide, the total badness of distance d decreases. We have $d(\mathcal{H}') \leq d(\mathcal{H})$ and $c_d(\mathcal{H}') < c_d(\mathcal{H})$.

Thus we end the proof of our lemma. \square

Repeat this process to make $c_d = (0)$. Repeating the whole process sufficiently many times will eventually kill all the bad regions other than D_0 .

4.2. Admissibility. In this subsection, we show that our algorithm will not change the admissibility, that is, if we start with an admissible Heegaard diagram, then our algorithm ends with an admissible Heegaard diagram. There are two operations involved in our algorithm: isotopies and handleslides, and we will consider them one by one.

The isotopy is the operation in Figure 14. Let \mathcal{H} and \mathcal{H}' be the Heegaard diagrams before and after the isotopy. Suppose \mathcal{H} is admissible. For a periodic

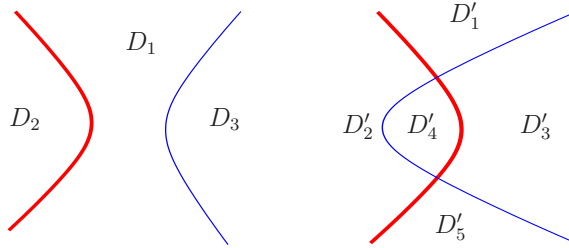


Figure 14. Isotopy of the β curve.

domain in \mathcal{H}'

$$\phi' = c_1 D'_1 + c_2 D'_2 + c_3 D'_3 + c_4 D'_4 + c_5 D'_5 + \cdots,$$

we have $c_2 - c_1 = c_4 - c_3 = c_2 - c_5$ and $c_1 - c_3 = c_2 - c_4 = c_5 - c_3$. Hence $c_1 = c_5$ and $c_4 = c_2 + c_3 - c_1$. Note that the regions are all the same except those in Figure 14. Therefore,

$$\phi = c_1 D_1 + c_2 D_2 + c_3 D_3 + \cdots$$

is a periodic domain for \mathcal{H} . Since \mathcal{H} is admissible, ϕ has both positive and negative coefficients, and so does ϕ' . Hence \mathcal{H}' is admissible.

Our handleslide operation is indicated in Figure 15. Suppose \mathcal{H} is admissible. For a periodic domain in \mathcal{H}'

$$\phi' = c_* D'_* + c_1 D'_{m,1} + c_2 D'_{m,2} + c_{1,1} S'_{1,1} + c_{1,2} S'_{1,2} + \cdots + c_{k,1} S'_{k,1} + c_{k,2} S'_{k,2} + \cdots$$

we get $c_1 - c_* = c_{1,1} - c_{1,2} = \cdots = c_{k,1} - c_{k,2} = c_2 - c_*$. Suppose $c_1 - c_* = c_0$, then $c_{i,1} = c_{i,2} + c_0$ and $c_1 = c_2$. Now

$$\phi = c_* D_* + c_1 D_m + c_{1,1} S_1 + \cdots + c_{k,1} S_k + \cdots$$

is a periodic domain for \mathcal{H} . Since \mathcal{H} is admissible, ϕ has both positive and negative coefficients. Hence ϕ' has both positive and negative coefficients, so \mathcal{H}' is admissible.

Remark. In fact, it can be shown that nice Heegaard diagrams are always (weakly) admissible ([4, Cor. 3.2]).

We have similar conclusions for Heegaard diagrams with multiple basepoints. Our algorithm could be modified to get nice Heegaard diagrams in that case. Note that every region is connected to exactly one region containing some w point in the complement of the α curves, so we can define the distance and hence the complexity in the same way, and thus our algorithm works as before.

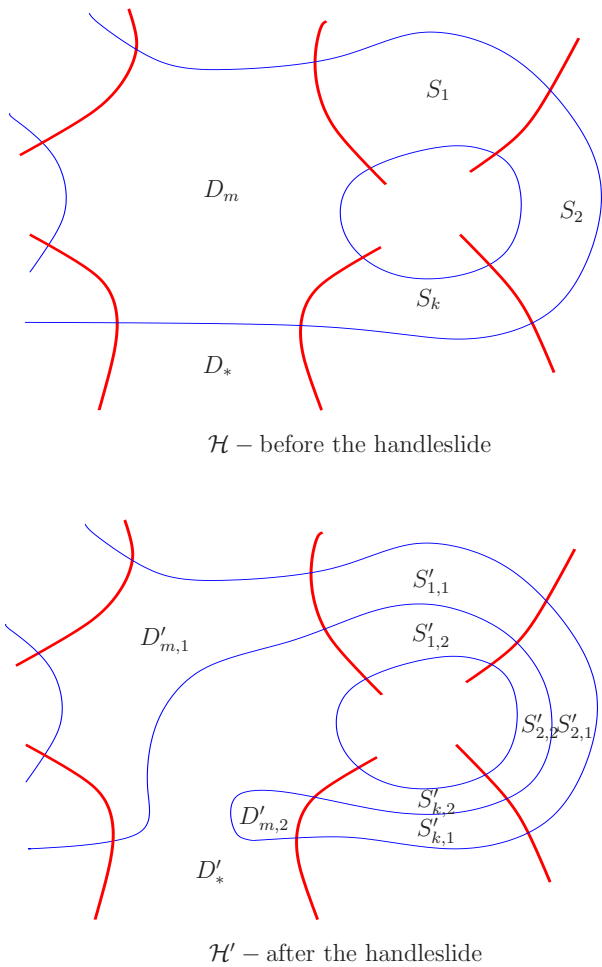


Figure 15. Handleslide of the β curve.

Proof of Theorem 1.2. Starting with an admissible one-pointed Heegaard diagram, our algorithm described in Section 4.1 gives an admissible Heegaard diagram with only one bad region, the one containing the basepoint w . The algorithm can be modified for multiple basepoints as described above. \square

5. Examples

In this section, we give two examples to demonstrate our algorithm. One is on knot Floer homology and the other is on the Heegaard-Floer homology of three-manifolds.

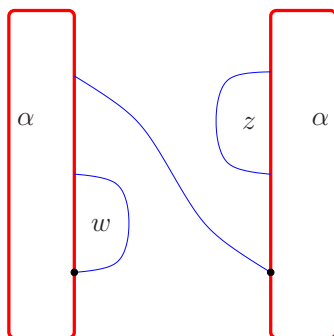


Figure 16. A Heegaard diagram for the trefoil knot. We make the convention that every two thick circles with the same α labels are identified so that the two dark points on them are identified.

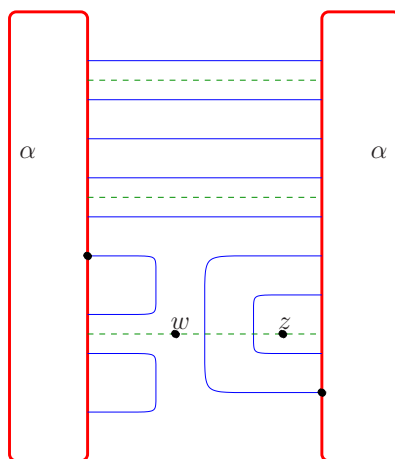


Figure 17. A nice Heegaard diagram for the trefoil knot. We use the same convention as in Figure 16. The trefoil is given by the dotted curve.

5.1. *The trefoil.* We start with the Heegaard diagram of the trefoil knot in Figure 16, where the two circles labeled by α are identified to get a genus one Heegaard diagram.

After isotopy using the algorithm in Section 4, we end up with the Heegaard diagram as in Figure 17. So we have nine generators. It is routine to find all boundary holomorphic disks and determine the Alexander and Maslov gradings of each generator.

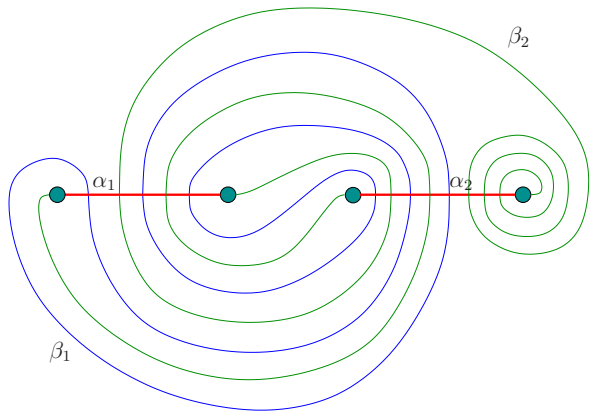


Figure 18. A Heegaard diagram for the Poincaré homology sphere. The two darkly shaded circles on the left are the feet of one handle, and the two darkly shaded circles on the right are the feet of the other handle.

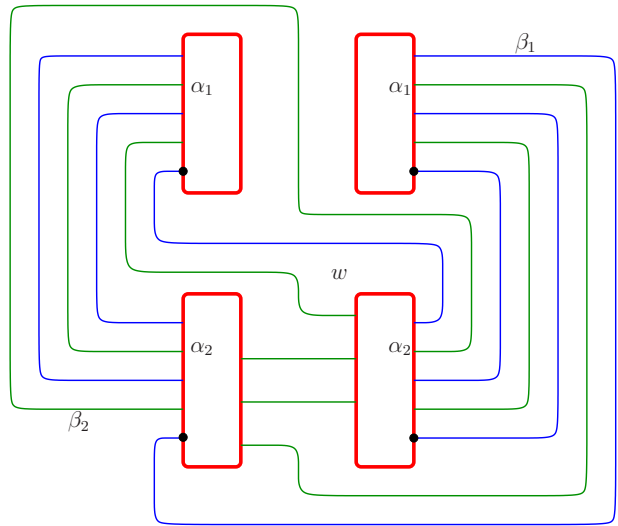


Figure 19. A Heegaard diagram for the Poincaré homology sphere. We use the same convention as in Figure 16.

5.2. *The Poincaré homology sphere $\Sigma(2, 3, 5)$.* We start with the Heegaard diagram of $\Sigma(2, 3, 5)$ in Figure 18, viewed as the +1 surgery on the right-handed trefoil knot. By cutting the Heegaard surface along the α circles, we get a planar presentation of the Heegaard diagram in Figure 19.

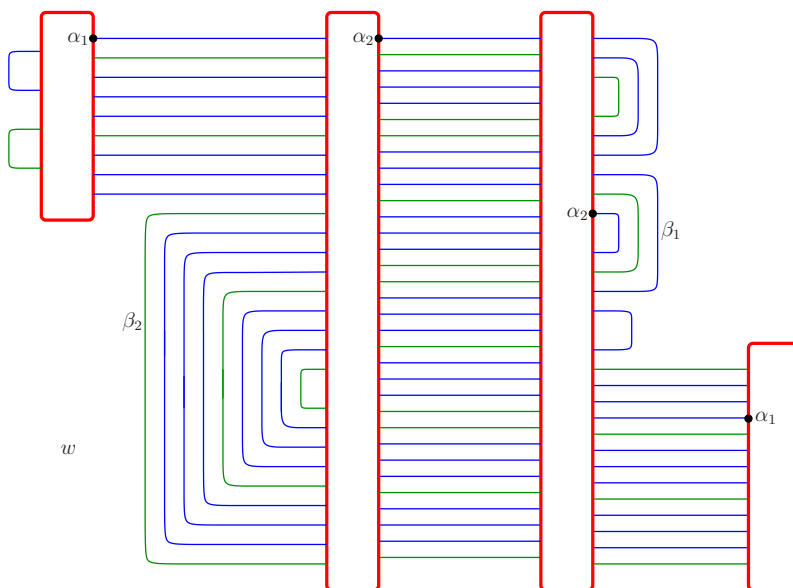


Figure 20. A nice Heegaard diagram for the Poincaré homology sphere. We use the same convention as in Figure 16.

It is easy to see that there are 21 generators for the chain complex. However, the authors do not know how to compute the differentials.

After applying our algorithm, we get a nice Heegaard diagram as in Figure 20. There are 335 generators and 505 differentials for this diagram. We leave the actual computation using this diagram to the patient reader.

Acknowledgements. The first author wishes to thank Zoltán Szabó for introducing him to the subject of Heegaard Floer homology and having many helpful discussions at various points.

This work was done when the second author was an exchange graduate student in Columbia University. He is grateful to the Columbia math department for its hospitality. He also would like to thank Rob Kirby and Peter Ozsváth for their continuous guidance, support and encouragement.

Both authors thank Matthew Hedden, Robert Lipshitz, Ciprian Manolescu, Peter Ozsváth, Jacob Rasmussen and Dylan Thurston for making comments and having helpful discussions during the development of this work. We also thank the referees for their comments and suggestions.

References

- [1] P. GHIGGINI, Knot Floer homology detects genus-one fibred knots, *Amer. J. Math.* **130** (2008), 1151–1169. MR 2450204 Zbl 1149.57019

- [2] A. JUHÁSZ, Floer homology and surface decompositions, *Geom. Topol.* **12** (2008), 299–350. MR 2009a:57021 Zbl 05256859
- [3] R. LIPSHITZ, A cylindrical reformulation of Heegaard Floer homology, *Geom. Topol.* **10** (2006), 955–1097. MR 2007h:57040 Zbl 1130.57035
- [4] R. LIPSHITZ, C. MANOLESCU, and J. WANG, Combinatorial cobordism maps in hat Heegaard Floer theory, *Duke Math. J.* **145** (2008), 207–247. MR 2009f:57050 Zbl 1153.57027
- [5] C. MANOLESCU, P. S. OZSVÁTH, and S. SARKAR, A combinatorial description of knot Floer homology, *Ann. of Math.* **169** (2009), 633–660. MR 2009k:57047 Zbl pre05578753
- [6] Y. NI, Link Floer homology detects the Thurston norm, 2006. arXiv math.GT/0604360
- [7] ———, Knot Floer homology detects fibred knots, *Invent. Math.* **170** (2007), 577–608. MR 2008j:57053 Zbl 1138.57031
- [8] P. OZSVÁTH and Z. SZABÓ, Absolutely graded Floer homologies and intersection forms for four-manifolds with boundary, *Adv. Math.* **173** (2003), 179–261. MR 2003m:57066 Zbl 1025.57016
- [9] ———, Knot Floer homology and the four-ball genus, *Geom. Topol.* **7** (2003), 615–639. MR 2004i:57036 Zbl 1037.57027
- [10] ———, Heegaard diagrams and holomorphic disks, in *Different Faces of Geometry*, *Int. Math. Ser.* **3**, Kluwer/Plenum, New York, 2004, pp. 301–348. MR 2005g:57057 Zbl 1091.57010
- [11] ———, Holomorphic disks and genus bounds, *Geom. Topol.* **8** (2004), 311–334. MR 2004m:57024 Zbl 1056.57020
- [12] ———, Holomorphic disks and knot invariants, *Adv. Math.* **186** (2004), 58–116. MR 2005e:57044 Zbl 1062.57019
- [13] ———, Holomorphic disks and three-manifold invariants: properties and applications, *Ann. of Math.* **159** (2004), 1159–1245. MR 2006b:57017 Zbl 1081.57013
- [14] ———, Holomorphic disks and topological invariants for closed three-manifolds, *Ann. of Math.* **159** (2004), 1027–1158. MR 2006b:57016 Zbl 1073.57009
- [15] ———, Holomorphic triangles and invariants for smooth four-manifolds, *Adv. Math.* **202** (2006), 326–400. MR 2007i:57029 Zbl 1099.53058
- [16] ———, Holomorphic disks, link invariants and the multi-variable Alexander polynomial, *Algebr. Geom. Topol.* **8** (2008), 615–692. MR 2443092 Zbl 1144.57011
- [17] ———, Link Floer homology and the Thurston norm, *J. Amer. Math. Soc.* **21** (2008), 671–709. MR 2009a:57050
- [18] J. RASMUSSEN, Floer homology and knot complements, Ph.D. thesis, Harvard University, 2003.

(Received August 4, 2006)

(Revised May 1, 2008)

E-mail address: sucharit@math.princeton.edu

PRINCETON UNIVERSITY, DEPARTMENT OF MATHEMATICS, FINE HALL - WASHINGTON ROAD,
PRINCETON NJ 08544, UNITED STATES,
<http://www.princeton.edu/~sucharit/>

E-mail address: wjiajun@caltech.edu

CALIFORNIA INSTITUTE OF TECHNOLOGY, MATHEMATICS 253-37,
1200 EAST CALIFORNIA BOULEVARD, PASADENA CA 91125, UNITED STATES,
<http://www.its.caltech.edu/~wjiajun/>