JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 15, Number 1, Pages 203–271
S 0894-0347(01)00374-5
Article electronically published on September 24, 2001

# QUIVERS, FLOER COHOMOLOGY, AND BRAID GROUP ACTIONS

MIKHAIL KHOVANOV AND PAUL SEIDEL

## 1. Introduction

1a. **Generalities.** This paper investigates the connection between symplectic geometry and those parts of representation theory which revolve around the notion of categorification. The existence of such a connection, in an abstract sense, follows from simple general ideas. The difficult thing is to make it explicit. On the symplectic side, the tools needed for a systematic study of this question are not yet fully available. Therefore we concentrate on a single example, which is just complicated enough to indicate the depth of the relationship. The results can be understood by themselves, but a glimpse of the big picture certainly helps to explain them, and that is what the present section is for.

Let $\mathcal{Q}$ be a category. An *action of a group $G$ on $\mathcal{Q}$* is a family $(\mathcal{F}_g)_{g\in G}$ of functors from $\mathcal{Q}$ to itself, such that $\mathcal{F}_e \cong \mathrm{id}_{\mathcal{Q}}$ and $\mathcal{F}_{g_1}\mathcal{F}_{g_2} \cong \mathcal{F}_{g_1g_2}$ for all $g_1, g_2 \in G$; here $\cong$ denotes isomorphism of functors.[1] We will not distinguish between two actions $(\mathcal{F}_g)$ and $(\widetilde{\mathcal{F}}_g)$ such that $\mathcal{F}_g \cong \widetilde{\mathcal{F}}_g$ for all $g$. A particularly nice situation is when $\mathcal{Q}$ is triangulated and the $\mathcal{F}_g$ are exact functors. Then the action induces a linear representation of $G$ on the Grothendieck group $K(\mathcal{Q})$. The inverse process, in which one lifts a given linear representation to a group action on a triangulated category, is called categorification (of group representations).

The connection with symplectic geometry is based on an idea of Donaldson. He proposed (in talks circa 1994) to associate to a compact symplectic manifold $(M^{2n}, \o)$ a category $\mathrm{Lag}(M,\omega)$ whose objects are Lagrangian submanifolds $L \subset M$, and whose morphisms are the Floer cohomology groups[2] $HF(L_0, L_1)$. The composition of morphisms would be given by products $HF(L_1, L_2) \times HF(L_0, L_1) \longrightarrow HF(L_0, L_2)$, which are defined, for example, in [39]. Let $\mathrm{Symp}(M,\omega)$ be the group of symplectic automorphisms of $M$. Any $\phi \in \mathrm{Symp}(M,\omega)$ determines a family of isomorphisms $HF(L_0, L_1) \cong HF(\phi L_0, \phi L_1)$ for $L_0, L_1 \in \mathrm{Ob}(\mathrm{Lag}(M,\omega))$ which are compatible with the products. In other words, $\phi$ induces an equivalence $\mathcal{F}_\phi$ from $\mathrm{Lag}(M,\omega)$ to itself. This is just a consequence of the fact that $\mathrm{Lag}(M,\omega)$ is

Received by the editors July 6, 2000 and, in revised form, April 3, 2001.
2000 *Mathematics Subject Classification.* Primary 18G10, 53D40, 20F36.
The first author was supported by NSF grants DMS 96-27351 and DMS 97-29992 and, later on, by the University of California at Davis. The second author was supported by NSF grant DMS-9304580 and by the Institut Universitaire de France.

[1] Strictly speaking, this should be called a *weak* action of $G$ on $\mathcal{Q}$; a full-fledged action comes with preferred isomorphisms $\mathcal{F}_{g_1}\mathcal{F}_{g_2} \cong \mathcal{F}_{g_1g_2}$ that satisfy obvious compatibility relations.

[2] The definition of Floer cohomology in general involves difficult analytic and algebraic questions. To simplify the exposition, we tacitly ignore them here.

an object of symplectic geometry, hence natural under symplectic maps. Assume for simplicity that $H^1(M;\mathbb{R}) = 0$, so that all symplectic vector fields are Hamiltonian. Then a smooth isotopy of symplectic automorphisms $(\phi_t)_{0\leq t\leq 1}$ gives rise to distinguished elements in $HF(\phi_0 L, \phi_1 L)$ for all $L$. This means that the $\mathcal{F}_\phi$ define a canonical action of the symplectic mapping class group $\pi_0(\mathrm{Symp}(M,\omega))$ on $\mathrm{Lag}(M,\omega)$. As a consequence, any symplectic fibre bundle with fibre $(M,\omega)$ and base $B$ gives rise to a $\pi_1(B)$-action on $\mathrm{Lag}(M,\omega)$, through the monodromy map $\pi_1(B) \to \pi_0(\mathrm{Symp}(M,\omega))$. Interesting examples can be obtained from families of smooth complex projective varieties.

While it is thus easy to construct potentially interesting group actions on the categories $\mathrm{Lag}(M,\omega)$, both the group actions and the categories themselves are difficult to understand. A particularly intriguing question is whether, in any given case, one can relate them to objects defined in a purely algebraic way. In homological algebra there is a standard technique for approaching similar comparison problems (an example is Beilinson's work [2] on coherent sheaves on $\mathbb{C}\mathrm{P}^k$). A very crude attempt to adapt this technique to our situation goes like this: pick a finite number of Lagrangian submanifolds $L_0, \dots, L_m \subset M$ which, for some reason, appear to be particularly important. Using the product on Floer cohomology, turn

$$A = \bigoplus_{i,j=0}^{m} HF(L_i, L_j) \tag{1.1}$$

into a ring. Now associate to an arbitrary Lagrangian submanifold $L \subset M$ the $A$-module $\bigoplus_i HF(L, L_i)$. This defines a functor from $\mathrm{Lag}(M,\omega)$ to the category of $A$-modules. As it stands the functor is not particularly useful, since it does not take into account all the available structure.

To begin with, Floer cohomology groups are graded, and hence $A$ is a graded ring (depending on $M$, this may be only a $\mathbb{Z}/N$-grading for some finite $N$). To make more substantial progress one needs to refine the category $\mathrm{Lag}(M,\omega)$. This theory is as yet under construction, and we can only give a vague outline of it. As pointed out by Fukaya [15], working directly with the Floer cochain complexes should enable one to construct an $A_\infty$-category underlying $\mathrm{Lag}(M,\omega)$. Kontsevich [24] suggested that the derived category of this $A_\infty$-category be considered. This "derived Fukaya category" is expected to be triangulated, and to contain $\mathrm{Lag}(M,\omega)$ as a full subcategory; we denote it by $D^b\mathrm{Lag}(M,\omega)$ (which is an abuse of notation). The action of $\pi_0(\mathrm{Symp}(M,\omega))$ on $\mathrm{Lag}(M,\omega)$ should extend to an action on $D^b\mathrm{Lag}(M,\omega)$ by exact functors. Moreover, it seems natural to suppose that the Grothendieck group of $D^b\mathrm{Lag}(M,\omega)$ is related to $H_n(M;\mathbb{Z})$; that would mean that the group actions coming from families of projective varieties could be considered as categorifications of the classical monodromy representations. Returning to the rings (1.1), one expects to get from Fukaya's construction a canonical (up to quasi-isomorphism) $A_\infty$-algebra $\mathcal{A}$ with cohomology $A$. The functor introduced above would lift to an $A_\infty$-functor from the $A_\infty$-category underlying $\mathrm{Lag}(M,\omega)$ to the $A_\infty$-category of $A_\infty$-modules over $\mathcal{A}$. In a second step, this would induce an exact functor

$$D^b\mathrm{Lag}(M,\omega) \longrightarrow D(\mathcal{A}),$$

where $D(\mathcal{A})$ is the derived category of $A_\infty$-modules. A standard argument based on exactness indicates that this functor, when restricted to the triangulated subcategory of $D^b\mathrm{Lag}(M,\omega)$ generated by $L_0, \dots, L_m$, would be full and faithful. This is a much stronger comparison theorem than one could get with the primitive approach which we had mentioned first. We will not say more about this; what the reader should keep in mind is the expected importance of the rings $A$.

It should also be mentioned that there is a special reason to study actions of braid groups and mapping class groups on categories. In 1992–1994 a number of mathematicians independently suggested extending the formalism of TQFTs to manifolds with corners. While three-dimensional extended TQFTs are rather well understood, there has been much less progress in the four-dimensional case. Part of the data which make up a four-dimensional extended TQFT are categories $\mathcal{Q}(S)$ associated to closed surfaces $S$; these would govern the process of cutting and pasting for the vector spaces associated to closed three-manifolds. The mapping class group of $S$ should act on $\mathcal{Q}(S)$ in a natural way, and for interesting TQFTs this action is likely to be highly nontrivial. As a matter of historical interest, we note that the categories $\mathrm{Lag}(M, \omega)$ were invented by Donaldson in an attempt to make gauge theory fit into the extended TQFT formalism; see [16] for more recent developments.

There is a version of this for invariants of two-knots, where the categories are associated to punctured discs and carry actions of the braid groups. The papers [4] and [23] go some way towards constructing such extended TQFTs in a representation-theoretic way.

1b. **The results.** In the algebraic part of the paper, we study the derived categories $D^b(A_m\text{-mod})$ for a particular family $A_m$ $(m \geq 1)$ of graded rings. The definition of $A_m$ starts with the quiver (oriented graph) $\Gamma_m$ shown in Figure 1, whose vertices are labelled $0, 1, \dots, m$. Recall that the path ring of an arbitrary quiver $\Gamma$ is the abelian group freely generated by the set of all paths in $\Gamma$, with multiplication given by composition. Paths of length $l$ in $\Gamma_m$ correspond to $(l+1)$-tuples $(i_1|i_2|\dots|i_{l+1})$ of numbers $i_k \in \{0, 1, \dots, m\}$ such that the difference of any two consecutive ones is $\pm 1$. Therefore, the path ring of $\Gamma_m$ is the abelian group freely generated by such $(l+1)$-tuples for all $l \geq 0$, with the multiplication

$$(i_1|\dots|i_{l+1})(i'_1|\dots|i'_{l'+1}) = \begin{cases} (i_1|\dots|i_l|i'_1|\dots|i'_{l'+1}) & \text{if } i_{l+1} = i'_1, \\ 0 & \text{otherwise.} \end{cases}$$

The paths $(i)$ of length zero are mutually orthogonal idempotents, and their sum is the unit element. We make the path ring of $\Gamma_m$ into a graded ring by setting $\deg(i) = \deg(i|i+1) = 0$ and $\deg(i+1|i) = 1$ for all $i$, and extending this in the obvious way (this is *not* the same as the grading by lengths of paths). Finally, $A_m$ is the quotient of the path ring of $\Gamma_m$ by the relations

$$(i-1|i|i+1) = (i+1|i|i-1) = 0, \quad (i|i+1|i) = (i|i-1|i), \quad (0|1|0) = 0$$

for all $0 < i < m$. These relations are homogeneous with respect to the above grading, so that $A_m$ is a graded ring. As an abelian group $A_m$ is free and of finite rank; a basis is given by the $4m+1$ elements $(0), \dots, (m), (0|1), \dots, (m-1|m), (1|0), \dots, (m|m-1), (1|0|1), \dots, (m|m-1|m)$.

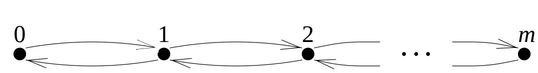



FIGURE 1.

Let $A_m$-mod be the category of finitely generated graded left $A_m$-modules, and let $D^b(A_m\text{-mod})$ be its bounded derived category. We will write down explicitly exact self-equivalences $\mathcal{R}_1, \dots, \mathcal{R}_m$ of $D^b(A_m\text{-mod})$ which satisfy

$$\begin{aligned} \mathcal{R}_i\mathcal{R}_{i+1}\mathcal{R}_i &\cong \mathcal{R}_{i+1}\mathcal{R}_i\mathcal{R}_{i+1} && \text{for } 1 \leq i < m, \text{ and} \\ \mathcal{R}_j\mathcal{R}_k &\cong \mathcal{R}_k\mathcal{R}_j && \text{for } |j-k| \geq 2. \end{aligned}$$

These are the defining relations of the braid group $B_{m+1}$, which means that the functors $\mathcal{R}_i$ generate an action $(\mathcal{R}_\sigma)_{\sigma \in B_{m+1}}$ of $B_{m+1}$ on $D^b(A_m\text{-mod})$. The first step in analyzing this action is to look at the induced linear representation on the Grothendieck group $K(D^b(A_m\text{-mod})) \cong K(A_m\text{-mod})$. The category $A_m$-mod carries a self-equivalence $\{1\}$ which raises the grading of a module by one (this should not be confused with the translation functor $[1]$ in the derived category). The action of $\{1\}$ makes $K(A_m\text{-mod})$ into a module over $\mathbb{Z}[q, q^{-1}]$. In fact

$$K(A_m\text{-mod}) \cong \mathbb{Z}[q, q^{-1}] \otimes \mathbb{Z}^{m+1};$$

a basis is given by the indecomposable projective modules $P_i = A_m(i) \subset A_m$. The functors $\mathcal{R}_i$ commute with $\{1\}$. Therefore the induced representation is a homomorphism $B_{m+1} \longrightarrow GL_{m+1}(\mathbb{Z}[q, q^{-1}])$. A computation shows that this is the well-known Burau representation.

At this point, it is helpful to recall the topological meaning of the Burau representation. Take a closed disc $D$ and a set $\Delta \subset D \setminus \partial D$ of $(m+1)$ marked points on it. Let $\widetilde{D}$ be the infinite cyclic cover of $D \setminus \Delta$ whose restriction to a small positively oriented loop around any point of $\Delta$ is isomorphic to $\mathbb{R} \longrightarrow S^1$. Fix a point $z$ on $\partial D$, and let $\tilde{z} \subset \widetilde{D}$ be its preimage. Let $\mathcal{G} = \mathrm{Diff}(D, \partial D; \Delta)$ be the group of diffeomorphisms $f$ of $D$ which satisfy $f|\partial D = \mathrm{id}$ and $f(\Delta) = \Delta$. Any $f \in \mathcal{G}$ can be lifted in a unique way to a $\mathbb{Z}$-equivariant diffeomorphism of $\widetilde{D}$ which acts trivially on $\tilde{z}$. The Burau representation can be defined as the induced action of $\pi_0(\mathcal{G}) \cong B_{m+1}$ on $H_1(\widetilde{D}, \tilde{z}; \mathbb{Z}) \cong \mathbb{Z}[q, q^{-1}] \otimes \mathbb{Z}^{m+1}$. Our first result will show that the action $(\mathcal{R}_\sigma)$ itself, or rather certain numbers attached to it, admits a similar topological interpretation. For any $\sigma \in B_{m+1}$ and $0 \leq i, j \leq m$ consider the bigraded abelian group

$$\bigoplus_{r_1, r_2 \in \mathbb{Z}} \mathrm{Hom}_{D^b(A_m\text{-mod})}(P_i, \mathcal{R}_\sigma P_j[r_1]\{-r_2\}). \tag{1.2}$$

Fix a collection of curves $b_0, \dots, b_m$ on $(D, \Delta)$ as in Figure 2; this determines a preferred isomorphism $\pi_0(\mathcal{G}) \cong B_{m+1}$. Take some $\sigma \in B_{m+1}$ and a diffeomorphism $f_\sigma \in \mathcal{G}$ representing it. The geometric intersection number $I(b_i, f_\sigma(b_j)) \geq 0$ counts the number of essential intersection points between the curves $b_i$ and $f_\sigma(b_j)$. The exact definition, which is a slight variation of the usual one, will be given in the body of the paper.

**Theorem 1.1.** *The total rank of the bigraded group* (1.2) *is* $2\, I(b_i, f_\sigma(b_j))$.

We will actually prove a slightly stronger result, which describes (1.2) up to isomorphism in terms of *bigraded intersection numbers* invented for that purpose. Theorem 1.1, together with standard properties of geometric intersection numbers, has an interesting implication. Call a group action $(\mathcal{F}_g)_{g \in G}$ on a category $\mathcal{Q}$ faithful if $\mathcal{F}_g \not\cong \mathrm{id}_{\mathcal{Q}}$ for all $g \neq e$.

**Corollary 1.2.** $(\mathcal{R}_\sigma)_{\sigma \in B_{m+1}}$ *is faithful for all* $m \geq 1$.

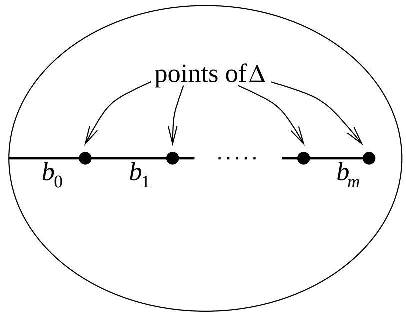



FIGURE 2.

Note that this cannot be proved by considerations on the level of Grothendieck groups, since the Burau representation is not faithful (in the usual sense of the word) for $m \gg 0$ [27]. Corollary 1.2 can be used to derive the faithfulness of certain braid group actions on categories which occur in algebraic geometry [38].

On the symplectic side we will study the Milnor fibres of certain singularities. The $n$-dimensional singularity of type $(A_m)$, for $m, n \geq 1$, is the singular point $x = 0$ of the hypersurface $\{x_0^2 + \cdots + x_{n-1}^2 + x_n^{m+1} = 0\}$ in $\mathbb{C}^{n+1}$. To define the Milnor fibre one perturbs the defining equation, so as to smooth out the singular point, and then intersects the outcome with a ball around the origin. In the present case one can take

$$M = \{x_0^2 + \cdots + x_{n-1}^2 + h(x_n) = 0,\ |x| \leq 1\},$$

where $h(z) = z^{m+1} + w_m z^m + \cdots + w_1 z + w_0 \in \mathbb{C}[z]$ is a polynomial with no multiple zeros, with $w = (w_0, \ldots, w_m) \in \mathbb{C}^{m+1}$ small. Equip $M$ with the restriction $\o$ of the standard symplectic form on $\mathbb{C}^{n+1}$, and its boundary with the restriction $\alpha$ of the standard contact one-form on $S^{2n+1}$. Assume that the disc $D$ considered above is in fact embedded in $\mathbb{C}$ as the subset $\{z \in \mathbb{C},\ |z|^2 + |h(z)| \leq 1\}$, and that the set of marked points is $\Delta = h^{-1}(0)$. Then one can associate to any curve $c$ in $(D, \Delta)$ a Lagrangian submanifold $L_c$ of $(M, \o, \alpha)$. We postpone the precise definition, and just mention that $L_c$ lies over $c$ with respect to the projection to the $x_n$-variable. In particular, from curves $b_0, \ldots, b_m$ as in Figure 2 one obtains Lagrangian submanifolds $L_0, \ldots, L_m$. These are all $n$-spheres except for $L_0$, which is an $n$-ball with boundary in $\partial M$; they intersect each other transversally, and

$$L_i \cap L_j = \begin{cases} \text{one point} & \text{if } |i-j| = 1, \\ \emptyset & \text{if } |i-j| \geq 2. \end{cases} \tag{1.3}$$

The case $m = 3$ and $n = 1$ is shown in Figure 3.

By varying the polynomial $h$ one gets a family of Milnor fibres parametrized by an open subset $W \subset \mathbb{C}^{m+1}$. Parallel transport in this family, for a suitable class of connections, defines a *symplectic monodromy map*

$$\rho_s : \pi_1(W, w) \longrightarrow \pi_0(\mathrm{Symp}(M, \partial M, \o)).$$

The actual construction is slightly more complicated than this rough description may suggest. Similar maps can be defined for general hypersurface singularities. In

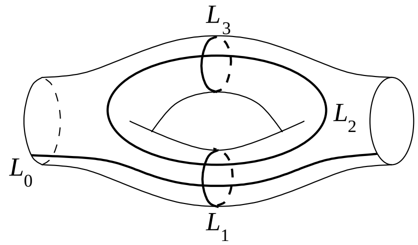



FIGURE 3.

our particular case, there are isomorphisms

$$\pi_1(W, w) \cong \pi_1(\mathrm{Conf}_{m+1}(D \setminus \partial D), \Delta) \cong \pi_0(\mathrm{Diff}(D, \partial D; \Delta)) \cong B_{m+1}; \tag{1.4}$$

the first one comes from a canonical embedding $W \subset \mathrm{Conf}_{m+1}(D \setminus \partial D)$, the second one is well known and again canonical, and the third one is determined by the choice of curves $b_0, \dots, b_m$ as before. Using these isomorphisms, one can consider $\rho_s$ as a homomorphism from the braid group to $\pi_0(\mathrm{Symp}(M, \partial M, \emptyset))$. We mention, although this has no relevance for the present paper, that there is a more direct definition of $\rho_s$ in terms of generalized Dehn twists along $L_1, \dots, L_m$. This is explained in [37, Appendix] for $n = 2$, and the general case is similar.

Take some $\sigma \in B_{m+1}$ and a map $\phi_\sigma \in \mathrm{Symp}(M, \partial M, \emptyset)$ which represents $\rho_s(\sigma)$ in the sense of (1.4). If $n = 1$ we also assume that $\phi_\sigma$ has a particular property called $\theta$-exactness, which will be defined in the body of the paper; this assumption is necessary because in that dimension $H^1(M, \partial M; \mathbb{R}) \neq 0$. The Floer cohomology $HF(L_i, \phi_\sigma(L_j))$ is a finite-dimensional vector space over $\mathbb{Z}/2$ constructed from the intersection points of $L_i$ and $\phi_\sigma(L_j)$, and it is independent of the choice of $\phi_\sigma$.

**Theorem 1.3.** *The dimension of $HF(L_i, \phi_\sigma(L_j))$ is $2\, I(b_i, f_\sigma(b_j))$.*

Here $f_\sigma \in \mathrm{Diff}(D, \partial D; \Delta)$ is as in Theorem 1.1. There is also a faithfulness result analogous to Corollary 1.2:

**Corollary 1.4.** *$\rho_s$ is injective for all $m, n \geq 1$.*

It is instructive to compare this with the outcome of a purely topological consideration. The classical geometric monodromy $\rho_g$ is the composition

$$\begin{tikzcd} & \pi_0(\mathrm{Symp}(M, \partial M, \emptyset)) \arrow[d] \\ B_{m+1} \arrow[ur, "\rho_s"] \arrow[r, "\rho_g"] & \pi_0(\mathrm{Diff}(M, \partial M)) \end{tikzcd}$$

where the vertical arrow is induced by inclusion. For $n = 1$ this arrow is an isomorphism, and the injectivity of $\rho_g$ is an old result of Birman and Hilden [7] (see Proposition 3.11). The situation changes drastically for $n = 2$, where Brieskorn's simultaneous resolution [8] implies that $\rho_g$ factors through the symmetric group $S_{m+1}$. The difference between $\rho_s$ and $\rho_g$ in this case shows that the symplectic structure of the Milnor fibre contains essential information about the singularity, which is lost when one considers it only as a smooth manifold. The situation for $n \geq 3$ is less clear-cut. One can use unknottedness results of Haefliger [19], together with an easy homotopy computation, to show that $\rho_g$ is not injective for all $m \geq 2$

if $n > 2$ is even, and for all $m \geq 3$ if $n \equiv 1 \bmod 4$, $n \neq 1$ (with more care, these bounds could probably be improved).

By combining Theorem 1.1 and Theorem 1.3 one sees that the dimension of $HF(L_i, \phi_\sigma(L_j))$ is equal to the total rank of (1.2). One can refine the comparison by taking into account the graded structure of Floer cohomology, as follows:

**Corollary 1.5.** *For all $\sigma \in B_{m+1}$, $0 \leq i, j \leq m$, and $r \in \mathbb{Z}$,*

$$HF^r(L_i, \phi_\sigma(L_j)) \cong \bigoplus_{r_1 + nr_2 = r} \mathrm{Hom}_{D^b(A_m\text{-mod})}(P_i, \mathcal{R}_\sigma P_j[r_1]\{-r_2\}) \otimes \mathbb{Z}/2. \tag{1.5}$$

Strictly speaking, the grading in Floer theory is canonical only up to an overall shift (in fact, for $n = 1$ there is an even bigger ambiguity), and the isomorphism holds for a particular choice. One can cure this problem by using Kontsevich's idea of graded Lagrangian submanifolds. We will not state the corresponding improved version of Corollary 1.5 explicitly, but it can easily be extracted from our proof. Note that if one looks at the graded groups $HF^*(L_i, \phi_\sigma(L_j))$ in isolation, they appear to depend on $n$ in a complicated way; whereas the connection with the bigraded group (1.2) makes this dependence quite transparent.

One can see this collapsing of the bigrading on a much simpler level, which corresponds to passing to Euler characteristics on both sides in Corollary 1.5. Namely, the representation of the braid group on $H_n(M, \partial L_0; \mathbb{Z}) \cong \mathbb{Z}^{m+1}$ induced by $\rho_s$ is the specialization $q = (-1)^n$ of the Burau representation.

While our proof of Corollary 1.5 is by dimension counting, it is natural to ask whether there are canonical homomorphisms between the two graded groups; these should be compatible with the product structures existing on both sides. Unfortunately, the techniques used here are not really suitable for that; while an inspection of our argument suggests a possible definition of a homomorphism in (1.5), showing that it is independent of the various choices involved would be extremely complicated. And there is no easy way to adapt the proof of Theorem 1.3 to include a description of the product on Floer cohomology. It seems that to overcome these problems, one would have to use a more abstract approach, in line with the general picture of Section 1a. While the details of this are far from clear, an outline can easily be provided, and we will do that now.

It is convenient to modify the algebraic setup slightly. Let $A_{m,n}$ be the graded algebra obtained from $A_m$ by multiplying the grading with $n$ and changing the coefficients to $\mathbb{Z}/2$; regard this as a differential graded algebra $\widehat{A}_{m,n}$ with zero differential, and take the derived category of differential graded modules $D(\widehat{A}_{m,n})$. From the way in which $L_0, \dots, L_m \subset M$ intersect each other, and standard properties of Floer cohomology, one can see that

$$A_{m,n} \cong \bigoplus_{i,j=0}^{m} HF^*(L_i, L_j)$$

(actually, this was the starting point of our work). It seems reasonable to assume that the $A_\infty$-algebra of Floer cochain complexes underlying the right-hand side should be formal (see [38] for much more about this), which would make its derived category of $A_\infty$-modules equivalent to $D(\widehat{A}_{m,n})$. As mentioned in Section 1a, the general features of the construction should imply that one has an exact functor $D^b\mathrm{Lag}(M, \omega) \to D(\widehat{A}_{m,n})$, which would be full and faithful on the triangulated subcategory of $D^b\mathrm{Lag}(M, \omega)$ generated by $L_0, \dots, L_m$. $\rho_s$ defines a braid group action on $\mathrm{Lag}(M,\omega)$, and there should be an induced action on $D^b\mathrm{Lag}(M,\omega)$. On the other hand, one can define a braid group action on $D(\widehat{A}_{m,n})$ in a purely algebraic way, in the same way as we have done for the closely related category $D^b(A_m\text{-mod})$. It is natural to conjecture that the exact functor should intertwine the two actions. This would imply Corollary 1.5, in a stronger form which repairs the deficiencies mentioned above.

Finally, the reader may be wondering whether the objects which we have considered can be used to define invariants of two-knots. The category $D^b(A_m\text{-mod})$ is certainly not rich enough for this. However, it has been conjectured [4] that one can construct invariants of two-knots from a certain family of categories, of which $D^b(A_m\text{-mod})$ is the simplest nontrivial example. The homology groups for classical knots introduced in [23] are part of this program.

*Note.* During the final stage of work on this paper we learned of a preprint [33] by Rouquier and Zimmermann. They construct a braid group action in the derived category of modules over multiplicity one Brauer tree algebras, which are close relatives of $A_m$ (the idempotent $(0)$ of $A_m$ is set to zero; these smaller algebras also appear in [38] and [21]). Our action essentially coincides with the one of Rouquier and Zimmermann.

## 2. The braid group action on $D^b(A_m\text{-mod})$

2a. **The category of $A_m$-modules.** Fix $m \geq 1$. We recall some of the objects introduced in Section 1b:

- The graded ring $A_m$. In the future all left modules, right modules, or bimodules over $A_m$ are understood to be graded.
- The category $A_m$-mod, whose objects are finitely generated $A_m$-modules, and whose morphisms are grading-preserving module homomorphisms. In the future, we call the objects of $A_m$-mod just $A_m$-modules, and write $\mathrm{Hom}(M,N)$ instead of $\mathrm{Hom}_{A_m\text{-mod}}(M,N)$; throughout the paper, all homomorphisms are assumed grading-preserving unless otherwise mentioned.
- The self-equivalence $\{1\}$ of $A_m$-mod which shifts the grading upwards by one.
- The $A_m$-modules $P_i = A_m(i)$. These are indecomposable and projective. In fact, if one considers $A_m$ as a left module over itself, it splits into the direct sum $\bigoplus_{i=0}^m P_i$. Conversely, any indecomposable projective $A_m$-module is isomorphic to $P_i\{k\}$ for some $i$ and $k$.

Similarly, if one considers $A_m$ as a right module over itself, it splits into the direct sum of the indecomposable projective right modules ${}_iP = (i)A_m$.

**Proposition 2.1.** *$A_m$ has finite homological dimension.*

*Proof.* Introduce $A_m$-modules $S_0, S_1, \dots, S_m$ as follows. As a graded abelian group $S_i$ is isomorphic to $\mathbb{Z}$, placed in degree 0. The idempotent $(i)$ acts as the identity on $S_i$, and all other paths in the quiver act as zero. Any $A_m$-module has a finite length composition series with subsequent quotients isomorphic, up to shifts in the grading, to $S_i$ or $S_i/pS_i$ for various $i$ and primes $p$. Moreover, $S_i/pS_i$ has a resolution by modules $S_i$:

$$0 \longrightarrow S_i \longrightarrow S_i \longrightarrow S_i/pS_i \longrightarrow 0.$$

Therefore, the proposition will follow if we construct a finite length projective resolution of $S_i$ for each $i$. Consider the commutative diagram

$$\begin{array}{ccccccc}
P_m\{m-i\} & \to & \cdots & \to & P_{i+1}\{1\} & \to & P_i \\
\uparrow & & & & \uparrow & & \uparrow \\
P_{m-1}\{m-i\} & \to & \cdots & \to & P_i\{1\} & \to & P_{i-1} \\
\uparrow & & & & \uparrow & & \uparrow \\
\cdots & & \cdots & & \cdots & & \cdots \\
\uparrow & & & & \uparrow & & \uparrow \\
P_{m-i}\{m-i\} & \to & \cdots & \to & P_1\{1\} & \to & P_0 \\
\uparrow & & \cdots & & \uparrow & & \\
\cdots & \to & \cdots & \to & P_0\{1\} & & \\
\cdots & \to & \cdots & & & & \\
\uparrow & & & & & & \\
P_0\{m-i\} & & & & & &
\end{array}$$

of projective $A_m$-modules. The maps go from $P_j$ to $P_{j\pm1}$ for various $j$ and are given by right multiplications by $(j|j\pm1)$, which maps $P_j \subset A_m$ to $P_{j\pm1} \subset A_m$. The commutativity of the diagram follows from the relations

$$(j|j-1|j) = (j|j+1|j).$$

Moreover, every row and every column is a complex, since

$$(j|j+1|j+2) = (j|j-1|j-2) = 0,$$

so that, after inserting minus signs in appropriate places, the diagram above becomes a bicomplex. The associated total complex is a finite length projective resolution of the module $S_i$ (use that every column is acyclic except in the uppermost component). $\square$

2b. **The functors $\mathcal{U}_i$.** For $1 \leq i \leq m$ let $P_i \otimes {}_iP$ be the $A_m$-bimodule obtained by tensoring $P_i$ and ${}_iP$ over $\mathbb{Z}$. Let $\mathcal{U}_i : A_m\text{-mod} \to A_m\text{-mod}$ be the functor of tensoring with this bimodule,

$$\mathcal{U}_i(M) = P_i \otimes {}_iP \otimes_{A_m} M. \tag{2.1}$$

This functor is exact (since $P_i \otimes {}_iP$ is right projective) and takes projective modules to projective modules.[3]

**Theorem 2.2.** *There are functor isomorphisms*

$$\mathcal{U}_i\mathcal{U}_{i+1}\mathcal{U}_i \cong \mathcal{U}_i\{1\} \quad \textit{for } 1 \leq i \leq m-1, \tag{2.2}$$

$$\mathcal{U}_i\mathcal{U}_{i-1}\mathcal{U}_i \cong \mathcal{U}_i\{1\} \quad \textit{for } 2 \leq i \leq m, \tag{2.3}$$

$$\mathcal{U}_i^2 \cong \mathcal{U}_i \oplus \mathcal{U}_i\{1\} \quad \textit{for } 1 \leq i \leq m, \tag{2.4}$$

$$\mathcal{U}_i\mathcal{U}_j = 0 \quad \textit{for } |i-j| > 1. \tag{2.5}$$

*Proof.* We will consider only (2.2); the other isomorphisms are proved by similar arguments. The functor on the left-hand side of (2.2) is given by tensoring with the bimodule $P_i \otimes {}_iP \otimes_{A_m} P_{i+1} \otimes {}_{i+1}P \otimes_{A_m} P_i \otimes {}_iP$. The graded abelian groups ${}_iP \otimes_{A_m} P_{i+1}$ and ${}_{i+1}P \otimes_{A_m} P_i$ are free cyclic, one in degree 0, the other in degree 1. Thus $\mathcal{U}_i\mathcal{U}_{i+1}\mathcal{U}_i$ is isomorphic to the composition of tensoring with $P_i \otimes {}_iP$ and

[3] The $i = 0$ case is excluded, for although $\mathcal{U}_0$ can be defined, it does not satisfy the isomorphisms in Theorem 2.2.

shifting the grading by $\{1\}$. This is exactly the functor on the right-hand side of (2.2). $\square$

*Remark* 2.3. The functors $\mathcal{U}_i$ provide a functor realization of the Temperley-Lieb algebra (compare to Section 2e.4 and [4, Section 4.1.2]).

2c. **Categories of complexes.** For an abelian category $\mathcal{Q}$ denote by $P(\mathcal{Q})$ the homotopy category of bounded complexes of projective objects in $\mathcal{Q}$. An object $M$ of $P(\mathcal{Q})$ is a bounded complex of projective objects:

$$M = (M^i, \partial^i), \;\; \partial^i : M^i \to M^{i+1}, \;\; \partial^{i+1}\partial^i = 0.$$

A morphism $f$ from $M$ to $N$ is a collection of maps $f^i : M^i \to N^i$ that intertwine differentials in $M$ and $N$, $f^{i+1}\partial^i_M = \partial^i_N f^i$ for all $i$. Two morphisms $f, g$ are equal (homotopic) if there are maps $h^i \in \mathrm{Hom}_{\mathcal{Q}}(M^i, N^{i-1})$ such that $f^i - g^i = h^{i+1}\partial^i_M + \partial^{i-1}_N h^i$ for all $i$.

As usual, $[k]$ denotes the self-equivalence of $P(\mathcal{Q})$ which shifts a complex $k$ degrees to the left: $M[k]^i = M^{k+i}$, $\partial_{M[k]} = (-1)^k \partial_M$. Given a map of complexes $f : M \to N$, the cone of $f$ is the complex $C(f) = M[1] \oplus N$ with the differential $\partial_{C(f)}(x, y) = (-\partial_M x, f(x) + \partial_N y)$. We refer the reader to [17] for more information about homotopy and derived categories.

If $\mathcal{Q}$ has finite homological dimension, then $P(\mathcal{Q})$ is equivalent to the bounded derived category $D^b(\mathcal{Q})$ of $\mathcal{Q}$. This applies to $\mathcal{Q} = A_m$-mod by Proposition 2.1. We will use $P(A_m\text{-mod})$ instead of $D^b(A_m\text{-mod})$ throughout, and denote it by $\mathcal{C}_m = P(A_m\text{-mod})$.

The shift automorphism $\{k\}$ of $A_m$-mod extends in the obvious way to an automorphism of $\mathcal{C}_m$ which we also denote by $\{k\}$. Note that $\{k\}$ and $[k]$ are quite different. In fact, if we forget about the module structure and differential, objects of $\mathcal{C}_m$ are bigraded abelian groups, so that it is natural to have two shift-like functors. Correspondingly, for any two such objects $M, N$ there is a bigraded morphism group (of finite total rank)

$$\bigoplus_{r_1, r_2 \in \mathbb{Z}} \mathrm{Hom}(M, N[r_1]\{-r_2\}).$$

Let $R = (R^j, \partial^j)$ be a bounded complex of $A_m$-bimodules which are left and right projective, i.e., each $R^j$ is a projective left $A_m$-module and a projective right $A_m$-module. One associates to $R$ the functor $\mathcal{R} : \mathcal{C}_m \to \mathcal{C}_m$ of tensoring with $R$ :

$$\mathcal{R}(M) = R \otimes_{A_m} M.$$

2d. **The functors $\mathcal{R}_i$.** Define homomorphisms $\beta_i, \gamma_i$, $1 \le i \le m$, of $A_m$-bimodules,

$$\begin{aligned} \beta_i &: P_i \otimes {}_iP \longrightarrow A_m, \\ \gamma_i &: A_m \longrightarrow P_i \otimes {}_iP\{-1\}, \end{aligned}$$

by

$$\beta_i((i) \otimes (i)) = (i), \tag{2.6}$$

$$\begin{aligned} \gamma_i(1) = {} & (i-1|i) \otimes (i|i-1) + (i+1|i) \otimes (i|i+1) \\ & +(i) \otimes (i|i-1|i) + (i|i-1|i) \otimes (i) \end{aligned} \tag{2.7}$$

(when $i = m$ we omit the term $(i+1|i) \otimes (i|i+1)$ from the sum for $\gamma_i(1)$).

These homomorphisms induce natural transformations, also denoted $\beta_i, \gamma_i$, between the functor $\mathcal{U}_i$, suitably shifted, and the identity functor:

$$\begin{aligned} \beta_i &: \mathcal{U}_i \longrightarrow \mathrm{Id}, \\ \gamma_i &: \mathrm{Id} \longrightarrow \mathcal{U}_i\{-1\}. \end{aligned}$$

Let $R_i$ be the complex of bimodules

$$R_i = \{0 \longrightarrow P_i \otimes {}_iP \xrightarrow{\beta_i} A_m \longrightarrow 0\}$$

with $A_m$ in degree 0. Denote by $R_i'$ the complex of bimodules

$$R_i' = \{0 \longrightarrow A_m \xrightarrow{\gamma_i} P_i \otimes {}_iP\{-1\} \longrightarrow 0\}$$

with $A_m$ in degree 0. Next, we introduce functors $\mathcal{R}_i$ and $\mathcal{R}_i'$ of tensoring with the complexes $R_i$ and $R_i'$ respectively:

$$\mathcal{R}_i(M) = R_i \otimes_{A_m} M, \qquad \mathcal{R}_i'(M) = R_i' \otimes_{A_m} M, \qquad M \in \mathrm{Ob}(\mathcal{C}_m).$$

The functor $\mathcal{R}_i$ can be viewed as the cone of $\beta_i$, and $\mathcal{R}_i'$ as the cone of $-\gamma_i$, shifted by $[-1]$.

**Proposition 2.4.** *The functors $\mathcal{R}_i$ and $\mathcal{R}_i'$ are mutually inverse equivalences of categories, i.e., there are functor isomorphisms*

$$\mathcal{R}_i\mathcal{R}_i' \cong \mathrm{Id} \cong \mathcal{R}_i'\mathcal{R}_i.$$

*Proof.* The functor $\mathcal{R}_i\mathcal{R}_i'$ is given by tensoring with the complex $R_i \otimes_{A_m} R_i'$ of $A_m$-bimodules. We write down this complex explicitly below. First consider a commutative square of $A_m$-bimodules and $A_m$-bimodule homomorphisms:

$$\begin{array}{ccc} P_i \otimes {}_iP & \xrightarrow{\beta_i} & A_m \\ \Big\downarrow{\scriptstyle\tau} & & \Big\downarrow{\scriptstyle\gamma_i} \\ P_i \otimes Q \otimes {}_iP\{-1\} & \xrightarrow{\delta} & P_i \otimes {}_iP\{-1\} \end{array} \tag{2.8}$$

Here $Q$ is the graded abelian group ${}_iP \otimes_{A_m} P_i$. It is a free group of rank 2 with a basis $u_1 = (i) \otimes (i), u_2 = (i|i-1|i) \otimes (i)$. The maps $\tau$ and $\delta$ are given by

$$\begin{aligned} \tau(x \otimes y) &= x \otimes u_1 \otimes (i|i-1|i)y + x \otimes u_2 \otimes y, \\ \delta(x \otimes u_1 \otimes y) &= x \otimes y, \\ \delta(x \otimes u_2 \otimes y) &= x(i|i-1|i) \otimes y. \end{aligned}$$

Define a map $\xi : A_m \to P_i \otimes Q \otimes {}_iP\{-1\}$ as the composition of $\gamma_i$ and the map $P_i \otimes {}_iP\{-1\} \to P_i \otimes Q \otimes {}_iP\{-1\}$ given by $x \otimes y \to x \otimes u_1 \otimes y$. Notice that $\delta\xi = \gamma_i$.

We make the diagram (2.8) anticommutative by putting a minus sign in front of $\delta$. Denote by $N = (N^j, \partial^j)$ the complex of bimodules, associated to this anticommutative diagram. It has 3 nonzero components

$$\begin{aligned} N^{-1} &= P_i \otimes {}_iP, \\ N^0 &= A_m \oplus (P_i \otimes Q \otimes {}_iP\{-1\}), \\ N^1 &= P_i \otimes {}_iP\{-1\}, \end{aligned}$$

and the differential $\partial^j : N^j \to N^{j+1}$ reads

$$\begin{aligned}\partial^{-1} &= \beta_i + \tau, \\ \partial^0 &= (\gamma_i, -\delta), \\ \partial^j &= 0 \qquad \text{for } j \neq -1, 0.\end{aligned}$$

It is easy to verify that $N$ is isomorphic to the complex $R_i \otimes_{A_m} R'_i$. Thus, we want to show that, in the homotopy category of complexes of $A_m$-modules, tensoring with $N$ is equivalent to the identity functor, i.e., to the functor of tensoring with $A_m$, considered as an $A_m$-bimodule. We will decompose $N$ as a direct sum of a complex isomorphic to $A_m$ and an acyclic complex. Namely, $N$ splits into a direct sum of 3 complexes, $N = T_{-1} \oplus T_0 \oplus T_1$, where

$$\begin{aligned}T_{-1} &= \{0 \to N^{-1} \to \partial^{-1}(N^{-1}) \to 0\}, \\ T_0 &= \{(a, \xi(a)) | a \in A_m \subset N^0\}, \\ T_1 &= \{0 \to P_i \otimes u_1 \otimes {}_iP \to N^1 \to 0\}.\end{aligned}$$

$T_{-1}$ is a subcomplex of $N$ generated by $N^{-1}$. Notice that $\partial^{-1} : N^{-1} \to N^0$ is injective (since $\tau$, a composition of $d^{-1}$ and a projection, is injective) and, therefore, $T_{-1}$ is acyclic. The complex $T_1$ is nonzero only in degrees 0 and 1. Its degree 0 component is $P_i \otimes u_1 \otimes {}_iP$, considered as a sub-bimodule of $N^0$, and the degree 1 component is $N^1$. The differential of $T_1$ induces an isomorphism between degree 0 and degree 1 components, and, thus, $T_1$ is acyclic. Finally, the complex $T_0$ is concentrated in horizontal degree 0, and consists of the set of pairs $(a, \xi(a)) \in N^0$, where $a \in A_m$. Clearly, $T_0$ is isomorphic to the complex which has $A_m$ in horizontal degree 0 and is trivial in all other degrees. Therefore, tensoring with $T_0$ is equivalent to the identity functor, while tensoring with $T_{\pm 1}$ is equivalent to tensoring with the zero complex.

It now follows that $\mathcal{R}_i \mathcal{R}'_i$ is isomorphic to the identity functor. A similar computation gives an isomorphism of $\mathcal{R}'_i \mathcal{R}_i$ and the identity functor. $\square$

We have seen that the functor $\mathcal{R}_i$ is invertible, with the inverse functor isomorphic to $\mathcal{R}'_i$. In view of this and for future convenience we will denote $\mathcal{R}'_i$ by $\mathcal{R}_i^{-1}$.

**Theorem 2.5.** *There are functor isomorphisms*

$$\mathcal{R}_i \mathcal{R}_{i+1} \mathcal{R}_i \cong \mathcal{R}_{i+1} \mathcal{R}_i \mathcal{R}_{i+1}, \tag{2.9}$$

$$\mathcal{R}_i \mathcal{R}_j \cong \mathcal{R}_j \mathcal{R}_i \quad \textit{for } |i-j| > 1. \tag{2.10}$$

*Proof.* The commutativity relation (2.10) follows at once from the obvious isomorphism $R_i \otimes_{A_m} R_j \cong R_j \otimes_{A_m} R_i$ of complexes of bimodules. Indeed, both complexes are isomorphic to the complex

$$0 \longrightarrow (P_i \otimes {}_iP) \oplus (P_j \otimes {}_jP) \xrightarrow{\beta_i + \beta_j} A_m \longrightarrow 0.$$

Therefore, we concentrate on proving (2.9), which is equivalent to

$$\mathcal{R}_{i+1}^{-1} \mathcal{R}_i \mathcal{R}_{i+1} \cong \mathcal{R}_i \mathcal{R}_{i+1} \mathcal{R}_i^{-1}. \tag{2.11}$$

The functor on the left side is given by the tensor product with the complex of bimodules $R'_{i+1} \otimes_{A_m} R_i \otimes_{A_m} R_{i+1}$, isomorphic to the cone of the map of complexes

$$R'_{i+1} \otimes_{A_m} P_i \otimes {}_iP \otimes_{A_m} R_{i+1} \longrightarrow R'_{i+1} \otimes_{A_m} R_{i+1}. \tag{2.12}$$

From the proof of Proposition 2.4 we know that $R'_{i+1} \otimes_{A_m} R_{i+1}$ is isomorphic to the direct sum of an acyclic complex and the complex $\{\cdots \longrightarrow 0 \longrightarrow A_m \longrightarrow$

$0 \longrightarrow \cdots\}$. Factoring out by the acyclic subcomplex we get a complex of bimodules (quasi-isomorphic to (2.12))

$$R'_{i+1} \otimes_{A_m} P_i \otimes {}_iP \otimes_{A_m} R_{i+1} \longrightarrow A_m. \tag{2.13}$$

Notice that $R'_{i+1} \otimes_{A_m} P_i$ is isomorphic to the complex of left modules $\{0 \longrightarrow P_i \longrightarrow P_{i+1} \longrightarrow 0\}$ where $P_i$ sits in the horizontal degree 0; and ${}_iP \otimes_{A_m} R_{i+1}$ is isomorphic to the complex of right modules $\{0 \longrightarrow {}_{i+1}P \longrightarrow {}_iP \longrightarrow 0\}$, where ${}_iP$ sits in degree 0. Therefore, the map (2.13) has the form

$$\begin{tikzcd} 0 \arrow[r] \arrow[d] & P_i \otimes {}_{i+1}P \arrow[r, "\partial^{-1}"] \arrow[d] & \begin{matrix}(P_i \otimes {}_iP)\\ \oplus \\ (P_{i+1} \otimes {}_{i+1}P)\end{matrix} \arrow[r, "\partial^0"] \arrow[d, "e"'] & P_{i+1} \otimes {}_iP \arrow[r] \arrow[d] & 0 \arrow[d] \\ 0 \arrow[r] & 0 \arrow[r] & A_m \arrow[r] & 0 \arrow[r] & 0 \end{tikzcd}$$

for some map $e$ of bimodules. $e$ is completely described by two integers $a_1$ and $a_2$, where

$$e((i) \otimes (i)) = a_1, \quad e((i+1) \otimes (i+1)) = a_2.$$

The condition $e\partial^{-1} = 0$ implies $a_1 + a_2 = 0$. Moreover, if $a_1 \neq \pm 1$, we can quickly come to a contradiction by picking a prime divisor $p$ of $a_1$, reducing all modules and functors mod $p$ and concluding that the functor $\mathcal{R}_{i+1}^{-1}\mathcal{R}_i\mathcal{R}_{i+1}$ (invertible over any base field) decomposes as a nontrivial direct sum in characteristic $p$. Thus, $a_1 = 1$ or $-1$. Multiplying each element of the bimodule $A_m$ by $-1$, if necessary, we can assume that $a_1 = 1$ and $a_2 = -1$.

The right-hand side functor of (2.11) is given by tensoring with the complex of bimodules $R_i \otimes_{A_m} R_{i+1} \otimes_{A_m} R'_i$. Since $R_i \otimes_{A_m} P_{i+1}$ is isomorphic to $\{0 \to P_i \to P_{i+1} \to 0\}$ and ${}_{i+1}P \otimes_{A_m} R'_i$ is isomorphic to $\{0 \to {}_{i+1}P \to {}_iP \to 0\}$, the right-hand side of (2.11) is isomorphic to the functor of tensoring with the cone

$$\begin{tikzcd} 0 \arrow[r] \arrow[d] & P_i \otimes {}_{i+1}P \arrow[r, "\partial^{-1}"] \arrow[d] & \begin{matrix}(P_i \otimes {}_iP)\\ \oplus \\ (P_{i+1} \otimes {}_{i+1}P)\end{matrix} \arrow[r, "\partial^0"] \arrow[d, "f"'] & P_{i+1} \otimes {}_iP \arrow[r] \arrow[d] & 0 \arrow[d] \\ 0 \arrow[r] & 0 \arrow[r] & A_m \arrow[r] & 0 \arrow[r] & 0 \end{tikzcd}$$

for some bimodule map $f$. Using the same argument as for $e$ we can assume that $f((i)\otimes(i)) = 1$ and $f((i+1)\otimes(i+1)) = -1$. Therefore, $e = f$ and the two functors in (2.11) are isomorphic. $\square$

**Definition 2.6.** A weak action of a group $G$ on a category $\mathcal{Q}$ is a choice of functors $\mathcal{F}_g : \mathcal{Q} \to \mathcal{Q}$ for each $g \in G$ such that $\mathcal{F}_1$ is the identity functor and functors $\mathcal{F}_{fg}$ and $\mathcal{F}_f\mathcal{F}_g$ are isomorphic for all $f, g \in G$.

To each element $\sigma$ of the braid group $B_{m+1}$ we associate a complex $R_\sigma$ of graded $A_m$-bimodules as follows. To $\sigma_i$ associate the complex $R_i$ and to $\sigma_i^{-1}$ the complex $R'_i$. In general, fix a decomposition of $\sigma$ as a product of generators $\sigma_i$ of the braid group and their inverses, $\sigma = \tau_1 \cdots \tau_k$. Then define $R_\sigma$ as the tensor product of corresponding complexes:

$$R_\sigma \overset{\text{def}}{=} R_{\tau_1} \otimes_{A_m} R_{\tau_2} \otimes \cdots \otimes_{A_m} R_{\tau_k}.$$

Finally, define $\mathcal{R}_\sigma : \mathcal{C}_m \to \mathcal{C}_m$ as the functor of tensoring with the bimodule $R_\sigma$:

$$\mathcal{R}_\sigma(M) \stackrel{\text{def}}{=} R_\sigma \otimes_{A_m} M, \qquad M \in \mathrm{Ob}(\mathcal{C}_m).$$

Proposition 2.4 and Theorem 2.5 together are equivalent to

**Proposition 2.7.** *The functors $\mathcal{R}_\sigma$ for $\sigma \in B_{m+1}$ define a weak action of the braid group on the category $\mathcal{C}_m$.*

A weak action of $G$ on a category $\mathcal{Q}$ is called *faithful* if for any $g \in G,\ g \neq 1$, the functor $\mathcal{F}_g$ is not isomorphic to the identity functor. We will prove in Section 4 that the above braid group action is faithful.

2e. **Miscellaneous.** In this subsection we link the Burau representation with the braid group action in the derived category $D^b(A_m\text{-mod})$. We also relate $A_m$ and highest weight categories.

2e.1. *The Grothendieck group of* $A_m$-mod *and the Burau representation.* Recall that the Grothendieck group $K(A_m\text{-mod})$ of the category $A_m$-mod is the group formed by virtual graded $A_m$-modules $[M]-[N]$, with a suitable equivalence relation. The functor $\{1\}$ makes $K(A_m\text{-mod})$ into a module over $\mathbb{Z}[q,q^{-1}]$. In fact, it is a free $\mathbb{Z}[q,q^{-1}]$-module of rank $m+1$. As a basis one can take the images of projective modules $P_i$. The functor $\mathcal{U}_i$ is exact and induces a $\mathbb{Z}[q,q^{-1}]$-linear map $[\mathcal{U}_i]$ on $K(A_m\text{-mod})$. The functor $\mathcal{R}_i$ is the cone of $\beta_i : \mathcal{U}_i \to \mathrm{Id}$ and so it induces a linear map $[\mathcal{R}_i] = [\mathrm{Id}] - [\mathcal{U}_i]$ on $K(A_m\text{-mod})$. Since the $\mathcal{R}_i$ satisfy braid group relations, the operators $[\mathcal{R}_i]$ define a representation of the braid group on $K(A_m\text{-mod})$. The generators of the braid group act as follows:

$$\begin{aligned}
\sigma_i[P_i] &= -q[P_i], \\
\sigma_i[P_{i+1}] &= [P_{i+1}] - [P_i], \\
\sigma_i[P_{i-1}] &= [P_{i-1}] - q[P_i], \\
\sigma_i[P_j] &= [P_j] \text{ for } |i-j| > 1.
\end{aligned}$$

The braid group action on the Grothendieck group preserves the $\mathbb{Z}[q,q^{-1}]$-submodule of $K(A_m\text{-mod})$ spanned by $[P_1],[P_2],\ldots,[P_m]$. Denote this submodule by $K'$. Comparing with the Burau representation (see [6]) we obtain

**Proposition 2.8.** *The representation on* $K(A_m\text{-mod})$ *is equivalent to the Burau representation, and its restriction to $K'$ is equivalent to the reduced Burau representation.*

2e.2. *Ring $A_m$ and highest weight categories.* Let $\mathfrak{sl}_{m+1}$ be the Lie algebra of traceless $(m+1)\times(m+1)$-matrices with complex coefficients. Let $\mathfrak{h}$ be the subalgebra of traceless diagonal matrices, and let $\mathfrak{p}$ be the subalgebra of matrices with zeros in all off-diagonal entries of the first column. Let $Z$ be the center of the universal enveloping algebra $U(\mathfrak{sl}_{m+1})$ and denote by $Z_0$ the maximal ideal of $Z$ which is the kernel of the augmentation homomorphism $Z \to \mathbb{C}$.

Denote by $\mathcal{O}'_m$ the category whose objects are finitely-generated $U(\mathfrak{sl}_{m+1})$-modules that are $U(\mathfrak{h})$-diagonalizable, $U(\mathfrak{p})$-locally finite and annihilated by some power of the maximal central ideal $Z_0$. Morphisms are $U(\mathfrak{sl}_{m+1})$-module maps (a $U(\mathfrak{sl}_{m+1})$-module $M$ is said to be $U(\mathfrak{p})$-locally finite if for any $x \in M$ the vector space $U(\mathfrak{p})x$ is finite dimensional).

By a base change we can make the ring $A_m$ into an algebra over the field of complex numbers, $A_m^{\mathbb{C}} \overset{\text{def}}{=} A_m \otimes_{\mathbb{Z}} \mathbb{C}$. Let $A_m^{\mathbb{C}}$-mod denote the category of finitely-generated left $A_m^{\mathbb{C}}$-modules (no grading this time). We learned the following result and its proof from Maxim Vybornov:

**Proposition 2.9.** *The categories $\mathcal{O}'_m$ and $A_m^{\mathbb{C}}$-*mod *are equivalent.*

*Proof.* The category $\mathcal{O}'_m$ consists of modules from a regular block of the Bernstein-Gelfand-Gelfand category $\mathcal{O}$ for $\mathfrak{sl}_{m+1}$ which are locally $U(\mathfrak{p})$-finite, where $\mathfrak{p}$ is the parabolic subalgebra described above. Let $Q_0, \dots, Q_m$ be indecomposable projective modules in $\mathcal{O}'_m$, one for each isomorphism class of indecomposable projectives. Form the projective module $Q = Q_0 \oplus \cdots \oplus Q_m$. The category $\mathcal{O}'_m$ is equivalent to the category of modules over the algebra $E = \mathrm{End}_{\mathcal{O}'_m}(Q)$. We claim that $E$ is isomorphic to $A_m^{\mathbb{C}}$. Indeed, from [3] we know that $E$ has a structure of a $\mathbb{Z}_+$-graded algebra, $E = \bigoplus_{i \geq 0} E_0$, such that $E$ is multiplicatively generated by $E_1$ over $E_0$ with quadratic defining relations. Moreover, $E$ is Koszul and the category of representations of its Koszul dual algebra $E^!$ is equivalent to a certain singular block of the category $\mathcal{O}$ for $\mathfrak{sl}_{m+1}$ (this is just a special case of the Beilinson-Ginzburg-Soergel parabolic-singular duality [3]). The algebra $E^!$, describing this singular block, was written down by Irving (see [22, Section 6.5]). Recall that $A_m$ is the quotient ring of the path ring of the quiver associated to the graph $\Gamma_m$ by certain relations among paths of length 2. The algebra $E^!$ is the quotient algebra of the path algebra of the quiver with the same graph $\Gamma_m$ but different set of quadratic relations:

$$(i|i-1|i) = (i|i+1|i), \qquad (m|m-1|m) = 0.$$

An easy computation establishes that the Koszul dual to $E^!$ is isomorphic to $A_m^{\mathbb{C}}$. Therefore, algebras $E$ and $A_m^{\mathbb{C}}$ are isomorphic. □

*Remark.* The grading which makes $A_m$ into a quadratic and Koszul algebra is the grading by the length of paths. This is different from the grading on $A_m$ defined in Section 1b and used throughout this paper.

An alternative proof of Proposition 2.9, which also engages Koszul duality, goes as follows. The category $\mathcal{O}'_m$ is equivalent to the category of perverse sheaves on the projective space $\mathbb{P}^m$, which are locally constant on each stratum $\mathbb{P}^i \setminus \mathbb{P}^{i-1}$ of the stratification of $\mathbb{P}^m$ by a chain of projective spaces $\mathbb{P}^0 \subset \mathbb{P}^1 \subset \cdots \subset \mathbb{P}^m$ (this is a special case of [3, Theorem 3.5.3]). Since the closure of each strata is smooth, the simple perverse sheaf $\mathcal{L}_i$ is, up to a shift in the derived category, the continuation by 0 of the constant sheaf on $\mathbb{P}^i$. The ext algebra $\bigoplus_{i,j} \mathrm{Ext}(\mathcal{L}_i, \mathcal{L}_j)$ can be easily written down, in particular,

$$\mathrm{Ext}^k(\mathcal{L}_i, \mathcal{L}_j) = H^{k-|i-j|}(\mathbb{P}^i \cap \mathbb{P}^j, \mathbb{C}).$$

This algebra is then seen to be Koszul dual to $A_m^{\mathbb{C}}$, which implies Proposition 2.9.

We thus matched the category of $A_m^{\mathbb{C}}$-modules with the category $\mathcal{O}'_m$. The functors $\mathcal{U}_i$ in the category $A_m^{\mathbb{C}}$-mod have a simple interpretation in the framework of highest weight categories. This interpretation is easy to guess, since the $\mathcal{U}_i$'s are exact functors while translation functors, and, more generally, projective functors (see [5] and references therein), constitute the main examples of exact functors in $\mathcal{O}$.

Precisely, under the equivalence of categories $A_m^{\mathbb{C}}$-mod $\cong \mathcal{O}'_m$, the functor $\mathcal{U}_i$ (formula (2.1)) becomes the functor of translation across the $i$-th wall. $\mathcal{U}_i$ is the composition of tensoring with ${}_iP$ over $A_m$ and then tensoring with $P_i$ over $\mathbb{C}$. Translation across the wall functor, on the other hand, is the composition of translation on and off the wall functors. For our particular parabolic subcategory of a regular block, translation on the wall functor takes it to the category equivalent to the category of $\mathbb{C}$-vector spaces. Translation on and off the $i$-th wall functors are two-sided adjoints, and the adjointness morphisms give rise to natural transformations between $\mathcal{U}_i$ and the identity functor. These transformations should coincide with natural transformations $\beta_i$ and $\gamma_i$ (see formulas (2.6) and (2.7)).

2e.3. *The nature of bimodule maps $\beta_i, \gamma_i$.* The bimodule homomorphism $\beta_i : P_i \otimes {}_iP \to A_m$ is simply a direct summand of the multiplication map $A_m \otimes A_m \to A_m$. The homomorphism $\gamma_i$, which goes the other way, is only slightly more mysterious and can be interpreted as follows. For simplicity of the discussion change the base ring from $\mathbb{Z}$ to a field $\mathbb{F}$ and ignore the grading of $A_m$. One checks that for $i > 0$ the projective module $P_i$ is also injective, and for any $A_m$-module $M$ there is a nondegenerate pairing

$$\mathrm{Hom}(P_i, M) \times \mathrm{Hom}(M, P_i) \to \mathrm{Hom}(P_i, P_i) \to \mathbb{F},$$

which defines an isomorphism $\mathrm{Hom}(P_i, M) \cong \mathrm{Hom}(M, P_i)^*$, functorial in $M$. Armed with the latter, we can rewrite the evaluation morphism

$$M \otimes \mathrm{Hom}(M, P_i) \to P_i$$

as a functorial map $M \to P_i \otimes \mathrm{Hom}(P_i, M)$, which, in turn, determines a bimodule map $A_m \to P_i \otimes {}_iP$. This map is precisely $\gamma_i$.

2e.4. *Temperley-Lieb algebra at roots of unity.* The Temperley-Lieb algebra $TL_n(q)$ has generators $U_1, \ldots, U_n$ and defining relations

$$\begin{aligned} U_i^2 &= (q + q^{-1})U_i, \\ U_i U_{i\pm 1} U_i &= U_i, \\ U_i U_j &= U_j U_i, \qquad |i - j| > 1. \end{aligned}$$

The Temperley-Lieb algebra is a quotient of the Hecke algebra of the symmetric group. Just like the Hecke algebra, the Temperley-Lieb algebra is semisimple for a generic $q \in \mathbb{C}$ and acquires a radical at certain roots of unity. Let $q = e^{\frac{i\pi}{r}}$ and $r > 2$. A theorem of Paul Martin (see [25], Chapter 7, for the exact statement) says that $TL_n(q)$ is isomorphic to the direct sum of matrix algebras and algebras Morita equivalent to $A_m$. Alternative proofs of this result can be found in Westbury [42] and Goodman and Wenzl [18].

## 3. Geometric intersection numbers

This part of the paper collects the necessary results about curves on surfaces. The first section, in which geometric intersection numbers are defined, is a mild generalization of [10, Exposé III]. As an easy application, we give a proof of the theorem of Birman and Hilden which was mentioned in Section 1b. After that we introduce the bigraded version of geometric intersection numbers. Finally, we consider a way of decomposing curves on a disc into certain standard pieces. The basic idea of this is again taken from [10, Exposé IV].

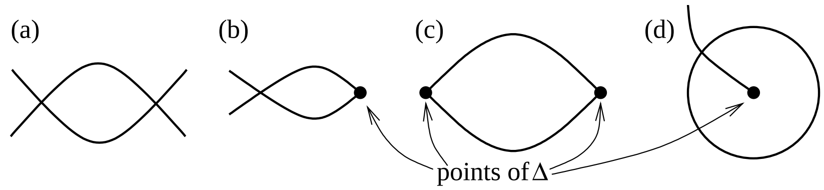



FIGURE 4.

3a. **Geometric intersection numbers.** Let $S$ be a compact oriented surface, possibly with boundary, and $\Delta \subset S \backslash \partial S$ a finite set of marked points. $\mathrm{Diff}(S, \partial S; \Delta)$ is the group of diffeomorphisms $f : S \longrightarrow S$ with $f|\partial S = \mathrm{id}$ and $f(\Delta) = \Delta$. By a *curve in* $(S, \Delta)$ we mean a subset $c \subset S$ which is either a simple closed curve in $S^o = S \setminus (\partial S \cup \Delta)$ and essential (not contractible in $S^o$), or else the image of an embedding $\gamma : [0; 1] \longrightarrow S$ which is transverse to $\partial S$ and such that $\gamma^{-1}(\partial S \cup \Delta) = \{0; 1\}$. Thus, our curves are smooth and unoriented. Two curves $c_0, c_1$ are called isotopic if one can be deformed into the other by an isotopy in $\mathrm{Diff}(S, \partial S; \Delta)$; we write $c_0 \simeq c_1$ for this. Note that endpoints on $\partial S$ may not move during an isotopy.

Let $c_0, c_1$ be two curves in $(S, \Delta)$. We say that they have *minimal intersection* if they intersect transversally, satisfy $c_0 \cap c_1 \cap \partial S = \emptyset$, and also satisfy the following condition: take any two points $z_- \neq z_+$ in $c_0 \cap c_1$ which do not both lie in $\Delta$, and two arcs $\alpha_0 \subset c_0$, $\alpha_1 \subset c_1$ with endpoints $z_-, z_+$, such that $\alpha_0 \cap \alpha_1 = \{z_-, z_+\}$. Let $K$ be the connected component of $S \setminus (c_0 \cup c_1)$ which is bounded by $\alpha_0 \cup \alpha_1$. Then if $K$ is topologically an open disc, it must contain at least one point of $\Delta$.

Among the examples in Figure 4, (a) and (b) do not have minimal intersection while (c) and (d) do.

**Lemma 3.1.** *Assume that $\Delta = \emptyset$. Let $c_0, c_1$ be two curves in $(S, \emptyset)$ which intersect transversally and satisfy $c_0 \cap c_1 \cap \partial S = \emptyset$. They have minimal intersection iff the following property holds: there is no continuous map $v : [0; 1]^2 \longrightarrow S$ with $v(s, 0) \in c_0$, $v(s, 1) \in c_1$, $v(0, t) = z_-$ and $v(1, t) = z_+$ for all $(s, t)$, where $z_- \neq z_+$ are points of $c_0 \cap c_1$.*

The proof is a very slightly modified version of the equivalence $(2^o) \Leftrightarrow (3^o)$ of [10, Proposition III.10]. Note that one of the implications is obvious. There is a similar result for $\Delta \neq \emptyset$ but we will not need it.

Given two curves $c_0, c_1$ in $(S, \Delta)$ with $c_0 \cap c_1 \cap \partial S = \emptyset$, one can always find a $c_1' \simeq c_1$ which has minimal intersection with $c_0$, by successively killing the unnecessary intersection points. We define the geometric intersection number $I(c_0, c_1) \in \frac{1}{2}\mathbb{Z}$ to be $I(c_0, c_1) = 2$ if $c_0, c_1$ are simple closed curves with $c_0 \simeq c_1$, and by the formula

$$I(c_0, c_1) = |(c_0 \cap c_1') \setminus \Delta| + \tfrac{1}{2}|c_0 \cap c_1' \cap \Delta|$$

in all other cases. In words, one counts the intersection points of $c_0$ and $c_1'$, common endpoints in $\Delta$ having weight $1/2$. The exceptional case in the definition can be motivated by Floer cohomology, or by looking at the behaviour of $I$ on double branched covers (see Section 3c). The fact that $I(c_0, c_1)$ is independent of the choice of $c_1'$ follows from the next two lemmas. Once this has been established, it is obvious that it is an invariant of the isotopy classes of $c_0$ and $c_1$.

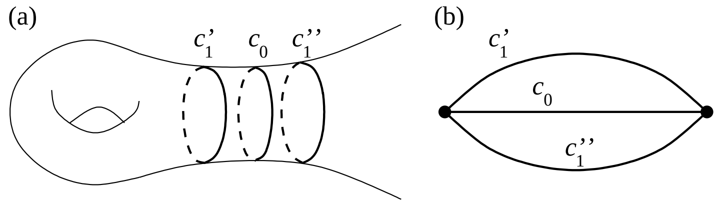



FIGURE 5.

**Lemma 3.2.** *Let $c_0$ be a curve in $(S, \Delta)$. Let $c_1', c_1''$ be two other curves such that both $(c_0, c_1')$ and $(c_0, c_1'')$ have minimal intersection, and with $c_1' \simeq c_1''$. Assume moreover that $c_0 \not\simeq c_1', c_1''$. Then there is an isotopy rel $c_0$ which carries $c_1'$ to $c_1''$ (more formally, there is a smooth path $(f_t)_{0 \leq t \leq 1}$ in $\mathrm{Diff}(S, \partial S; \Delta)$ such that $f_t(c_0) = c_0$ for all $t$, $f_0 = \mathrm{id}$, and $f_1(c_1') = c_1''$).*

For closed curves this is [10, Proposition III.12]; the proof of the general case is similar. The assumption that $c_0 \not\simeq c_1', c_1''$ cannot be removed, but there is a slightly weaker statement which fills the gap:

**Lemma 3.3.** *Let $c_0, c_1$ be two curves in $(S, \Delta)$ which are isotopic and have minimal intersection (by definition, this implies that the $c_i$ are either closed, or else that their endpoints all lie in $\Delta$). Then there is an isotopy rel $c_0$ which carries $c_1$ to one of the curves $c_1', c_1''$ shown (in the two cases) in Figure 5.*

The geometric intersection number can be extended to curves $c_0, c_1$ which intersect on $\partial S$, as follows: take a nonvanishing vector field on $\partial S$ which is positively oriented, and extend it to a smooth vector field $Z$ on $S$ which vanishes on $\Delta$. Let $(f_t)$ be the flow of $Z$. Set $c_0^+ = f_t(c_0)$ for some small $t > 0$, and define $I(c_0, c_1) := I(c_0^+, c_1)$. This depends on the orientation of $S$, and is also no longer symmetric.

3b. **Curves on a disc.** Let $D$ be a closed disc, with some fixed orientation, and let $\Delta \subset D \setminus \partial D$ be a set of $m+1 \geq 2$ marked points. $\mathrm{Conf}_{m+1}(D \setminus \partial D)$ denotes the configuration space of unordered $(m+1)$-tuples of points in $D \setminus \partial D$. As in Section 1b we write $\mathcal{G} = \mathrm{Diff}(D, \partial D; \Delta)$. There is a well-known canonical isomorphism

$$\pi_1(\mathrm{Conf}_{m+1}(D \setminus \partial D), \Delta) \cong \pi_0(\mathcal{G}) \tag{3.1}$$

defined as follows: take the evaluation map $\mathrm{Diff}(D, \partial D) \longrightarrow \mathrm{Conf}_{m+1}(D \setminus \partial D)$, $f \longmapsto f(\Delta)$. This is a Serre fibration, with the fibre at $\Delta$ being exactly $\mathcal{G}$, and (3.1) is the boundary map in the resulting sequence of homotopy groups.

From now on, by a curve we mean a curve in $(D, \Delta)$, unless otherwise specified. A *basic set of curves* is a collection of curves $b_0, \dots, b_m$ which looks like that in Figure 2 (more rigorously, this means that the $b_i$ can be obtained by applying some element of $\mathcal{G}$ to the collection drawn in Figure 2).

**Lemma 3.4.** *Let $b_0, \dots, b_m$ be a basic set of curves. If $f \in \mathcal{G}$ satisfies $f(b_i) \simeq b_i$ for all $i$, then $[f] \in \pi_0(\mathcal{G})$ is the identity class.*

The proof is not difficult (it rests on the fact that one can isotop such an $f$ to a $g$ which satisfies $g(b_i) = b_i$ for all $i$). A basic set of curves determines a preferred isomorphism between $B_{m+1}$, which we think of as an abstract group defined by the standard presentation, and $\pi_0(\mathcal{G})$, resp. $\pi_1(\mathrm{Conf}_{m+1}(D \setminus \partial D), \Delta)$. Concretely, the

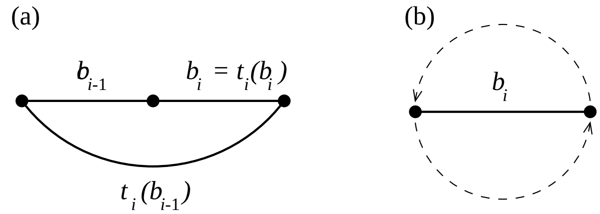



FIGURE 6.

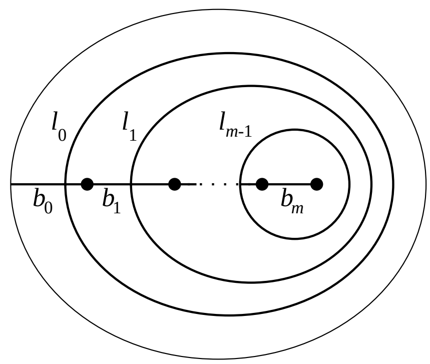



FIGURE 7.

$i$-th generator $\sigma_i$ $(1 \leq i \leq m)$ of $B_{m+1}$ goes to the isotopy class $[t_i] \in \pi_0(\mathcal{G})$ of the half-twist along $b_i$. This half-twist is a diffeomorphism which is trivial outside a small neighbourhood of $b_i$, which reverses $b_i$ itself, and such that the image of $b_{i-1}$ is as shown in Figure 6(a). The corresponding element in the fundamental group of the configuration space is the path which rotates the two endpoints of $b_i$ around each other by $180^{\text{o}}$; see Figure 6(b).

Fix a basic set of curves $b_0, \dots, b_m$. A curve $c$ will be called admissible if it is equal to $f(b_i)$ for some $f \in \mathcal{G}$ and $0 \leq i \leq m$. The endpoints of an admissible curve must lie in $\Delta \cup (b_0 \cap \partial D)$; conversely any curve with such endpoints is admissible. There is an obvious action of $\pi_0(\mathcal{G})$ on the set of isotopy classes of admissible curves. Lemma 3.4 is a faithfulness result for this action: it shows that only the identity element acts trivially.

**Lemma 3.5.** *Let $c$ be an admissible curve. Assume that there is a $k \in \{0, \dots, m\}$ such that $I(b_i, c) = I(b_i, b_k)$ for all $i = 0, \dots, m$. Then*

$$c \simeq \begin{cases} b_0 \ \textit{or}\ \tau_0^{-1}(b_0) & \textit{if } k = 0, \\ b_k \ \textit{or}\ \tau_k^{\pm 1}(b_k) & \textit{if } 1 \leq k < m, \\ b_m & \textit{if } k = m. \end{cases}$$

*Here $\tau_0, \dots, \tau_{m-1}$ are the positive Dehn twists along the closed curves $l_0, \dots, l_{m-1}$ shown in Figure 7.*

*Proof.* We will only explain the proof for $k = m$; the other cases are similar. We may assume that $m \geq 2$, since the statement is fairly obvious for $m = 1$. Let $c$ be a curve satisfying the conditions in the lemma. Since $I(b_{m-1}, c) = I(b_{m-1}, b_m) = 1/2$

is not integral, $c$ and $b_{m-1}$ have precisely one endpoint in common. This common endpoint could either be $b_{m-2} \cap b_{m-1}$ or $b_{m-1} \cap b_m$, but the first case is impossible because $I(b_{m-2}, c) = I(b_{m-2}, b_m) = 0$. We have now shown that $c$ and $b_m$ have one endpoint in common. It follows that the other endpoints of these two curves must also be the same, since $I(b_m, c) = I(b_m, b_m) = 1$ is integral. Next, we can assume that $c$ has minimal intersection with all the $b_i$ (this is because the "shortening moves" which kill unnecessary points of $b_i \cap c$ do not create new intersection points with any other $b_j$). Again by using geometric intersection numbers, it follows that only the endpoints of $c$ lie on $b_0 \cup \dots \cup b_m$. This means that we can consider $c$ as a curve on the surface obtained by cutting $D$ open along $b_0 \cup \dots \cup b_m$. The rest is straightforward. $\square$

**Lemma 3.6.** *If $f \in \mathcal{G}$ satisfies $I(b_j, f^2(b_k)) = I(b_j, f(b_k)) = I(b_j, b_k)$ for all $j, k = 0, \dots, m$, then $[f] \in \pi_0(\mathcal{G})$ is the identity class.*

*Proof.* Applying Lemma 3.5 to $c = f(b_j)$ shows that there are numbers $\nu_0 \in \{-1; 0\}$ and $\nu_1, \dots, \nu_{n-1} \in \{-1; 0; 1\}$, such that

$$f(b_j) \simeq \begin{cases} \tau_j^{\nu_j}(b_j) & \text{if } j = 0, \dots, m-1, \\ b_m & \text{if } j = m. \end{cases}$$

Since the $\tau_j$ commute, and $\tau_j(b_k) = b_k$ for all $k \neq j$, the map $g = \tau_0^{\nu_0} \tau_1^{\nu_1} \cdots \tau_{m-1}^{\nu_{m-1}} \in \mathcal{G}$ satisfies $f(b_j) \simeq g(b_j)$ for all $j$. By Lemma 3.4 it follows that $f$ and $g$ lie in the same isotopy class. Applying the same argument to $f^2$ shows that it can be written, up to isotopy, as a product $\tau_0^{\mu_0} \cdots \tau_{m-1}^{\mu_{m-1}}$ with numbers $\mu_j$ satisfying the same properties as $\nu_j$. In $\pi_0(\mathcal{G})$ we have therefore

$$[\tau_0^{2\nu_0} \cdots \tau_{m-1}^{2\nu_{m-1}}] = [f^2] = [\tau_0^{\mu_0} \cdots \tau_{m-1}^{\mu_{m-1}}].$$

Since the $\tau_j$ generate a free abelian subgroup of rank $m$ in $\pi_0(\mathcal{G})$, one has $2\nu_j = \mu_j$ for all $j$. Now $|\mu_j| \leq 1$ which implies that $\nu_j = 0$, hence that $[f] = [\mathrm{id}]$. $\square$

3c. **The double branched cover.** In this section we assume that $D$ is embedded in $\mathbb{C}$. The double cover of $D$ branched along $\Delta$ is, by definition,

$$p_S : S = \{(x_0, x_1) \in \mathbb{C} \times D \mid x_0^2 + h(x_1) = 0\} \longrightarrow D, \quad p_S(x_0, x_1) = x_1,$$

where $h \in \mathbb{C}[z]$ is a polynomial which has simple zeros exactly at the points of $\Delta$.

**Lemma 3.7.** *Let $c$ be a curve in $(D, \Delta)$ such that $c \cap \Delta \neq \emptyset$ (in the future, we will say that $c$ meets $\Delta$ if this is the case). Then $p_S^{-1}(c)$ is a curve in $(S, \emptyset)$.*

*Proof.* Let $z$ be a point of $c \cap \Delta$, and let $\gamma : [0; 1) \longrightarrow c$ be a local parametrization of $c$ near that point. One can write $h(\gamma(t)) = -t\psi(t)$ for some $\psi \in C^\infty([0; 1), \mathbb{C}^*)$. Choose a square root $\sqrt{\psi}$. Then a local smooth parametrization of $p_S^{-1}(c)$ near $\{(0, z)\} = p_S^{-1}(z)$ is given by $(-1; 1) \longrightarrow S$, $t \longmapsto (t\sqrt{\psi(t^2)}, \gamma(t^2))$. The smoothness of $p_S^{-1}(c)$ at all other points is obvious. As for the topology, there are (because of the assumption $c \cap \Delta \neq \emptyset$) only two possibilities: either $c$ is a path joining two points of $\Delta$, and then $p_S^{-1}(c)$ is a simple closed curve which is not homologous to zero; or $c$ connects a point of $\Delta$ with a point of $\partial D$, and then $p_S^{-1}(c)$ is a path connecting two points of $\partial S$. $\square$

**Lemma 3.8.** *Let $c_0, c_1$ be curves on $(D, \Delta)$, both of which meet $\Delta$. Assume that they have minimal intersection and are not isotopic. Then $p_S^{-1}(c_0), p_S^{-1}(c_1)$ also have minimal intersection.*

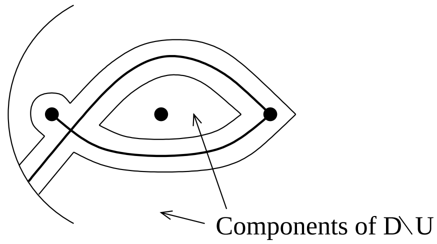



FIGURE 8.

*Proof.* It is clear that $p_S^{-1}(c_0)$ and $p_S^{-1}(c_1)$ intersect transversally, and do not intersect on $\partial S$. Let $U \subset D$ be a thickening of $c_0 \cup c_1$, such that $D \setminus U$ is a surface with corners. More precisely, the thickening (shown in Figure 8) should be done in such a way that a point of $(c_0 \cap c_1) \setminus \Delta$ gives rise to four corners of $D \setminus U$, and a point of $c_0 \cap c_1 \cap \Delta$ to two corners. Take some connected component $K$ of $D \setminus U$ which satisfies $K \cap \partial D = \emptyset$. Then at least one of the following three assertions is true: (1) $K$ is not contractible; (2) it has at least three corners; (3) it has two corners and contains at least one point of $\Delta$. The point here is that $K$ cannot be a 2-gon (a disc with two corners) and disjoint from $\Delta$, since our assumptions exclude all the four situations shown in Figure 4: (a), (b) because of the minimal intersection, (c) because $c_0 \not\simeq c_1$, and (d) because $c_0, c_1$ are not closed curves.

If (1) holds, then the preimage $p_S^{-1}(K)$ may consist of one or two connected components, but none of them can be contractible. If (2) holds, the connected components of $p_S^{-1}(K)$ may be contractible, but each of them has at least three corners. If (3) holds, then $p_S^{-1}(K)$ is connected with at least four corners, since its boundary is a double cover of the boundary of $K$. We have now shown that no connected component of $S \setminus p_S^{-1}(U)$ can be a 2-gon and disjoint from $\partial S$. Essentially by definition, this means that $p_S^{-1}(c_0)$ and $p_S^{-1}(c_1)$ have minimal intersection. □

**Lemma 3.9.** *Let $c_0, c_1$ be two curves on $(D, \Delta)$, both of which meet $\Delta$, such that $c_0 \cap c_1 \cap \partial D = \emptyset$. Assume that $c_0 \not\simeq c_1$. Then $p_S^{-1}(c_0) \not\simeq p_S^{-1}(c_1)$.*

*Proof.* **Case 1:** *$c_0$ or $c_1$ has an endpoint on $\partial D$.* Then $c_0 \cap \partial D \neq c_1 \cap \partial D$, because we have assumed that the intersection of these two sets is zero. It follows that $p_S^{-1}(c_0) \cap \partial S \neq p_S^{-1}(c_1) \cap \partial S$, so these two curves cannot be isotopic. **Case 2:** *$c_0, c_1$ are contained in $D \setminus \partial D$.* This means that both of them are paths joining two points of $\Delta$. Since the statement is one about isotopy classes, we may assume that $c_0, c_1$ have minimal intersection. Then the same holds for $p_S^{-1}(c_0), p_S^{-1}(c_1)$ by Lemma 3.8. Assume that, contrary to what we want to show, $p_S^{-1}(c_0) \simeq p_S^{-1}(c_1)$. Then Lemma 3.3 implies that $p_S^{-1}(c_0) \cap p_S^{-1}(c_1) = \emptyset$, so that $c_0 \cap c_1 = \emptyset$, and in particular $c_0 \cap \Delta \neq c_1 \cap \Delta$. But then $p_S^{-1}(c_0)$ and $p_S^{-1}(c_1)$ are not even homologous, which is a contradiction. □

**Lemma 3.10.** *Let $c_0, c_1$ be curves in $(D, \Delta)$, both of which meet $\Delta$. Then*

$$I(p_S^{-1}(c_0), p_S^{-1}(c_1)) = 2\, I(c_0, c_1).$$

*Proof.* It is enough to prove the statement for $c_0 \cap c_1 \cap \partial D = \emptyset$, since the other situation reduces to this by definition. **Case 1:** $c_0 \simeq c_1$. Because we are assuming that $c_0 \cap c_1 \cap \partial D = \emptyset$, this means that $c_0, c_1$ are paths connecting two points of

$\Delta$; then $I(c_0, c_1) = 1$. On the other hand, $p_S^{-1}(c_0)$ and $p_S^{-1}(c_1)$ are simple closed curves and isotopic, so that $I(p_S^{-1}(c_0), p_S^{-1}(c_1)) = 2$. **Case 2:** $c_0 \not\simeq c_1$. Then $p_S^{-1}(c_0) \not\simeq p_S^{-1}(c_1)$ by Lemma 3.9. Choose a $c_1' \simeq c_1$ which has minimal intersection with $c_0$. By Lemma 3.8 one has $I(p_S^{-1}(c_0), p_S^{-1}(c_1)) = |p_S^{-1}(c_0) \cap p_S^{-1}(c_1')| = 2|(c_0 \cap c_1') \setminus \Delta| + |c_0 \cap c_1' \cap \Delta| = 2I(c_0, c_1)$. □

Any $f \in \mathcal{G}$ can be lifted in a unique way to a homeomorphism $f_S$ of $S$ with $f_S|\partial S = \mathrm{id}$. If $f$ is holomorphic in some neighbourhood of each point of $\Delta$, then $f_S$ is again smooth. This defines a lifting homomorphism $\pi_0(\mathcal{G}) \longrightarrow \pi_0(\mathrm{Diff}(S, \partial S))$.

**Proposition 3.11** (Birman and Hilden)**.** *The lifting homomorphism is injective.*

*Proof.* Choose a set of basic curves $b_0, \dots, b_{m-1}$ in $(D, \Delta)$. Consider a map $f \in \mathcal{G}$ which is holomorphic near each point of $\Delta$, and assume that $[f]$ lies in the kernel of the lifting homomorphism. Then $p_S^{-1}(f(c)) = f_S(p_S^{-1}(c)) \simeq p_S^{-1}(c)$ for any curve $c$ which meets $\Delta$. Therefore, by Lemma 3.10,

$$I(f(b_j), b_k) = \tfrac{1}{2} I(p_S^{-1}(f(b_j)), p_S^{-1}(b_k)) = \tfrac{1}{2} I(p_S^{-1}(b_j), p_S^{-1}(b_k)) = I(b_j, b_k)$$

for $j, k = 0, \dots, m$. The same clearly holds for $f^2$. Lemma 3.6 shows that $[f]$ must be the identity class. □

We have chosen this proof, which is not the most direct one, because it serves as a model for later arguments. One can in fact avoid geometric intersection numbers entirely, by generalizing Lemma 3.9 to the case when $c_0 \cap c_1 \cap \partial D$ is not necessarily empty, and then using Lemma 3.4. Birman and Hilden's original paper [7] contains a stronger result (they identify the image of the lifting homomorphism) and their approach is entirely different.

3d. **Bigraded intersection numbers.** Let $P = PT(D \setminus \Delta)$ be the (real) projectivization of the tangent bundle of $D \setminus \Delta$. We want to introduce a particular covering of $P$ with covering group $\mathbb{Z}^2$. Take an oriented trivialization of $TD$; this allows one to identify $P = \mathbb{RP}^1 \times (D \setminus \Delta)$. For every point $z \in \Delta$ choose a small loop $\lambda_z : S^1 \longrightarrow D \setminus \Delta$ winding positively once around it. The classes $[\mathit{point} \times \lambda_z]$ together with $[\mathbb{RP}^1 \times \mathit{point}]$ form a basis of $H_1(P; \mathbb{Z})$. Let $C \in H^1(P; \mathbb{Z}^2)$ be the cohomology class which satisfies $C([\mathit{point} \times \lambda_z]) = (-2, 1)$ and $C([\mathbb{RP}^1 \times \mathit{point}]) = (1, 0)$. Define $\widetilde{P}$ to be the covering classified by $C$, and denote the $\mathbb{Z}^2$-action on it by $\chi$. This is independent of the choices we have made, up to isomorphism.

Any $f \in \mathcal{G}$ induces a diffeomorphism of $P$ which, since it preserves $C$, can be lifted to an equivariant diffeomorphism of $\widetilde{P}$. There is a unique such lift, denoted by $\tilde{f}$, which acts trivially on the fibre of $\widetilde{P}$ over any point $T_z \partial D \in P$, $z \in \partial D$. For any curve $c$ there is a canonical section $s_c : c \setminus \Delta \longrightarrow P$ given by $s_c(z) = T_z c$. A *bigrading* of $c$ is a lift $\tilde{c}$ of $s_c$ to $\widetilde{P}$. Pairs $(c, \tilde{c})$ consisting of a curve and a bigrading are called *bigraded curves*; we will often write $\tilde{c}$ instead of $(c, \tilde{c})$. $\chi$ defines a $\mathbb{Z}^2$-action on the set of bigraded curves, and the lifts $\tilde{f}$ yield a $\mathcal{G}$-action on the same set. There is an obvious notion of isotopy for bigraded curves, and one obtains induced actions of $\mathbb{Z}^2$ and $\pi_0(\mathcal{G})$ on the set of isotopy classes.

An alternative formulation is as follows. Assume that $D$ is embedded in $\mathbb{C}$, and let $h \in \mathbb{C}[z]$ be as in the previous section. The embedding $D \subset \mathbb{C}$ gives a preferred trivialization of $TD$, which we use to identify $P$ with $\mathbb{C}^*/\mathbb{R} \times (D \setminus \Delta)$. The map

$$\delta_P : P \longrightarrow (\mathbb{C}^*/\mathbb{R}^{>0}) \times (\mathbb{C}^*/\mathbb{R}^{>0}), \quad (\zeta, z) \longmapsto (h(z)^{-2}\zeta^2, -h(z))$$

represents $C$. Hence one can take $\widetilde{P}$ to be the pullback of the universal covering $\mathbb{R}^2 \longrightarrow (\mathbb{C}^*/\mathbb{R}^{>0})^2$, $(\xi_1, \xi_2) \mapsto (\exp(2\pi i \xi_1), \exp(2\pi i \xi_2))$. Then a bigrading of a curve $c$ is just a continuous lift $\tilde{c} : c \setminus \Delta \longrightarrow \mathbb{R}^2$ of the map $c \setminus \Delta \longrightarrow (\mathbb{C}^*/\mathbb{R}^{>0})^2$, $z \mapsto \delta_P(T_z c)$; and the $\mathbb{Z}^2$-action is by adding constant functions to $\tilde{c}$. To put it even more concretely, take an embedding $\gamma : (0;1) \longrightarrow D$ which parametrizes an open subset of $c \setminus \Delta$. Then the map which one has to lift can be written as

$$\delta_P(T_{\gamma(t)} c) = (h(\gamma(t))^{-2} \gamma'(t)^2, -h(\gamma(t))). \tag{3.2}$$

**Lemma 3.12.** *A curve $c$ admits a bigrading iff it is not a simple closed curve.*

*Proof.* If $c$ is not a simple closed curve, $c \setminus \Delta$ is contractible, so that the pullback $s_c^* \widetilde{P} \longrightarrow c \setminus \Delta$ is the trivial covering. Conversely, if $c$ is a simple closed curve and bounds a region of $D$ containing $k > 0$ points of $\Delta$, then $s_c^* C([c]) = \pm(2-2k, k) \neq 0$, which means that $s_c^* \widetilde{P}$ does not admit a section. □

**Lemma 3.13.** *The $\mathbb{Z}^2$-action on the set of isotopy classes of bigraded curves is free. Equivalently, a bigraded curve $\tilde{c}$ is never isotopic to $\chi(r_1, r_2)\tilde{c}$ for any $(r_1, r_2) \neq 0$.*

*Proof.* Let $c$ be a curve which connects two points $z_0, z_1 \in \Delta$. Let $(f_t)_{0 \leq t \leq 1}$ be an isotopy in $\mathcal{G}$ with $f_0 = \mathrm{id}$ and $f_1(c) = c$. Without any real loss of generality we may assume that $f_1 | c = \mathrm{id}$. **Claim:** *the closed path $\kappa_z : [0;1] \longrightarrow D \setminus \Delta$, $\kappa_z(t) = f_t(z)$ is freely nullhomotopic for each $z \in c \setminus \Delta$.* The free homotopy class of $\kappa_z$ is evidently independent of $z$. On the other hand, for $z$ close to $z_0$ it must be a multiple of the free homotopy class of $\lambda_{z_0}$, and the corresponding statement holds for $z$ close to $z_1$; which proves our claim.

As a consequence one can deform the isotopy $(f_t)_{0 \leq t \leq 1}$, rel endpoints, into another isotopy $(g_t)$ such that $g_t(c)$ agrees with $c$ in a neighbourhood of $z_0, z_1$ for all $t$. By considering the preferred lifts $\tilde{g}_t$ one sees immediately that $\tilde{f}_1(\tilde{c}) = \tilde{g}_1(\tilde{c}) = \tilde{c}$ for any bigrading $\tilde{c}$. The proof for the remaining types of curves, which satisfy $c \cap \partial\Delta \neq \emptyset$, is much easier, and we leave it to the reader. □

**Lemma 3.14.** *Let $c$ be a curve which joins two points of $\Delta$, let $t_c \in \mathcal{G}$ be the half twist along it, and let $\tilde{t}_c$ be its preferred lift to $\widetilde{P}$. Then $\tilde{t}_c(\tilde{c}) = \chi(-1,1)\tilde{c}$ for any bigrading $\tilde{c}$ of $c$.*

*Proof.* Since $t_c(c) = c$, one has $\tilde{t}_c(\tilde{c}) = \chi(r_1, r_2)\tilde{c}$ for some $(r_1, r_2) \in \mathbb{Z}^2$. Take an embedded smooth path $\beta : [0;1] \longrightarrow D \setminus \Delta$ from a point $\beta(0) \in \partial D$ to the unique point $\beta(1) \in c$ which is a fixed point of $t_c$; Figure 9 illustrates the situation. Consider the closed path $\pi : [0;2] \longrightarrow P$ given by $\pi(t) = \mathbb{R}\beta'(t) \subset T_{\beta(t)} D$ for $t \leq 1$, and by $\pi(t) = Dt_c(\mathbb{R}\beta'(2-t))$ for $t \geq 1$. It is not difficult to see that (with respect to the basis of $H_1(P;\mathbb{Z})$ which we have used before) $[\pi] = -[\mathbb{R}\mathrm{P}^1 \times point] - [point \times \lambda_z]$, where $z$ is one of the endpoints of $c$. Therefore

$$(r_1, r_2) = -C([\pi]) = C([\mathbb{R}\mathrm{P}^1 \times point]) + C([point \times \lambda_z]) = (-1, 1).$$

□

Let $(c_0, \tilde{c}_0)$ and $(c_1, \tilde{c}_1)$ be two bigraded curves, and let $z \in D \setminus \partial D$ be a point where $c_0$ and $c_1$ intersect transversally; $z$ may lie in $\Delta$ or not. Fix a small circle $l \subset D \setminus \Delta$ around $z$. Let $\alpha : [0;1] \longrightarrow l$ be an embedded arc which moves clockwise around $l$, such that $\alpha^{-1}(c_0) = \{0\}$ and $\alpha^{-1}(c_1) = \{1\}$. If $z \in \Delta$, then $\alpha$ is unique up to a change of parametrization; otherwise there are two possibilities, which are

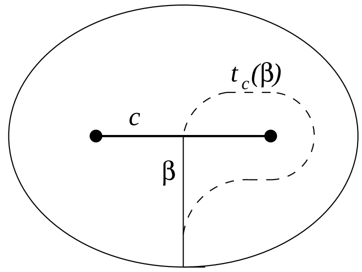



FIGURE 9.

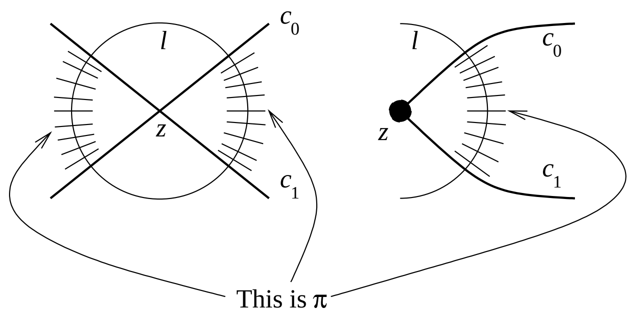



FIGURE 10.

distinguished by their endpoints (see Figure 10). Take a smooth path $\pi : [0;1] \longrightarrow P$ with $\pi(t) \in P_{\alpha(t)}$ for all $t$, going from $\pi(0) = T_{\alpha(0)}c_0$ to $\pi(1) = T_{\alpha(1)}c_1$, such that $\pi(t) \neq T_{\alpha(t)}l$ for all $t$. One can picture $\pi$ as a family of tangent lines along $\alpha$ which are all transverse to $l$ (again see Figure 10). Lift $\pi$ to a path $\tilde{\pi} : [0;1] \longrightarrow \widetilde{P}$ with $\tilde{\pi}(0) = \tilde{c}_0(\alpha(0))$; necessarily $\tilde{c}_1(\alpha(1)) = \chi(\mu_1, \mu_2)\tilde{\pi}(1)$ for some $\mu_1, \mu_2$. We define the local index of $\tilde{c}_0, \tilde{c}_1$ at $z$ to be $\mu^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1; z) = (\mu_1, \mu_2) \in \mathbb{Z}^2$; it is not difficult to see that this is independent of all choices.

Now let $\tilde{c}_0$, $\tilde{c}_1$ be two bigraded curves such that $c_0 \cap c_1 \cap \partial D = \emptyset$. Choose a $c_1' \simeq c_1$ which has minimal intersection with $c_0$. There is a bigrading $\tilde{c}_1'$ of $c_1'$, which is unique by Lemma 3.13, such that $\tilde{c}_1' \simeq \tilde{c}_1$. The bigraded intersection number $I^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1) \in \mathbb{Z}[q_1, q_1^{-1}, q_2, q_2^{-1}]$ is defined by

$$I^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1) = (1 + q_1^{-1}q_2)\Big(\sum_{z \in (c_0 \cap c_1') \setminus \Delta} q_1^{\mu_1(z)} q_2^{\mu_2(z)}\Big) + \Big(\sum_{z \in c_0 \cap c_1' \cap \Delta} q_1^{\mu_1(z)} q_2^{\mu_2(z)}\Big),$$

where $(\mu_1(z), \mu_2(z)) = \mu^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1'; z)$. The proof that this is independent of the choice of $c_1'$, and is an invariant of the isotopy classes of $\tilde{c}_0, \tilde{c}_1$, is basically the same as for ordinary geometric intersection numbers. The only nonobvious case is when $c_0$ and $c_1$ are isotopic, because then there are two essentially different choices for $c_1'$. However, an explicit computation shows that both choices lead to the same result. Bigraded intersection numbers can be extended to curves with $c_0 \cap c_1 \cap \partial D \neq \emptyset$ just like ordinary ones, by taking a flow $(f_t)$ which moves $\partial D$ in the positive sense, lifting the induced flow on $P$ to one $(\tilde{f}_t)$ on $\widetilde{P}$ such that $\tilde{f}_0 = \mathrm{id}$, and setting $I^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1) = I^{\mathrm{bigr}}(\tilde{f}_t(\tilde{c}_0), \tilde{c}_1)$ for small $t > 0$. We list some elementary properties of $I^{\mathrm{bigr}}$:

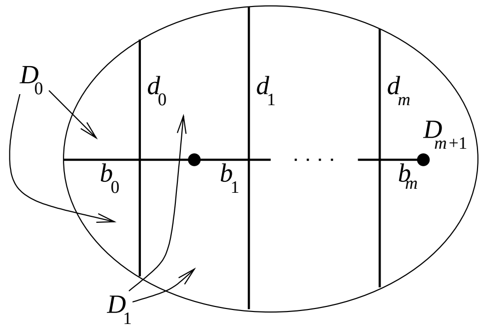



Figure 11.

(B1) $I(c_0, c_1)$ is obtained from $I^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1)$ by setting $q_1 = q_2 = 1$ and dividing by two.
(B2) $I^{\mathrm{bigr}}(\tilde{f}(\tilde{c}_0), \tilde{f}(\tilde{c}_1)) = I^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1)$ for any $f \in \mathcal{G}$.
(B3) $I^{\mathrm{bigr}}(\tilde{c}_0, \chi(r_1, r_2)\tilde{c}_1) = I^{\mathrm{bigr}}(\chi(-r_1, -r_2)\tilde{c}_0, \tilde{c}_1) = q_1^{r_1} q_2^{r_2} I^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1)$.
(B4) If $c_0 \cap c_1 \cap \partial D = \emptyset$ and $I^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1) = \sum_{r_1, r_2} a_{r_1, r_2} q_1^{r_1} q_2^{r_2}$, then $I^{\mathrm{bigr}}(\tilde{c}_1, \tilde{c}_0) = \sum_{r_1, r_2} a_{r_1, r_2} q_1^{-r_1} q_2^{1-r_2}$.

The only item which requires some explanation is the last one. Assume that $c_0, c_1$ have minimal intersection. The local index has the following symmetry property:

$$\mu^{\mathrm{bigr}}(\tilde{c}_1, \tilde{c}_0; z) = \begin{cases} (1,0) - \mu^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1; z) & \text{if } z \notin \Delta, \\ (0,1) - \mu^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1; z) & \text{if } z \in \Delta. \end{cases}$$

One sees immediately that (B4) holds for the contributions coming from points of $c_0 \cap c_1 \cap \Delta$. Now take a point $z \in (c_0 \cap c_1) \setminus \Delta$, and assume for simplicity that $\mu^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1; z) = (0, 0)$. Then the contribution of $z$ to $I^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1)$ is $1 + q_1^{-1} q_2$. Since $\mu^{\mathrm{bigr}}(\tilde{c}_1, \tilde{c}_0; z) = (1, 0)$, the same point contributes $q_1(1 + q_1^{-1} q_2) = q_1 + q_2$ to $I^{\mathrm{bigr}}(\tilde{c}_1, \tilde{c}_0)$, which is in agreement with (B4). We point out that the term $(1 + q_1^{-1} q_2)$ in the definition of bigraded intersection numbers was essential in this computation.

3e. **Normal form.** Fix a set of basic curves $b_0, \dots, b_m$. In addition choose curves $d_0, \dots, d_m$ as in Figure 11 which divide $D$ into regions $D_0, \dots, D_{m+1}$. We say that an admissible curve $c$ is in normal form if it has minimal intersection with all the $d_k$. This can always be achieved by an isotopy. The corresponding uniqueness theorem is a slight generalization of Lemma 3.2:

**Lemma 3.15.** *Let $c_0, c_1$ be two isotopic admissible curves, both of which are in normal form. Then there is an isotopy rel $d_0 \cup d_1 \cup \dots \cup d_m$ which carries $c_0$ to $c_1$.*

Let $c$ be a curve in normal form. Each connected component of $c \cap D_k$ belongs to one of finitely many classes or types. For $1 \leq k \leq m$ there are six such types, denoted by 1, 1′, 2, 2′, 3, 3′ (see Figure 12; the numbers $(r_1, r_2)$ here and in the subsequent figures are for later use, and the reader should ignore them at present). For $k = m+1$ there are two types, which are analogues of types 2 and 3 considered before; we use the same notation for them (Figure 13). For $k = 0$ there is a single type, for which we will not need a proper name (Figure 14). Conversely, it is not difficult to prove that an admissible curve $c$ which intersects all the $d_k$ transversally,

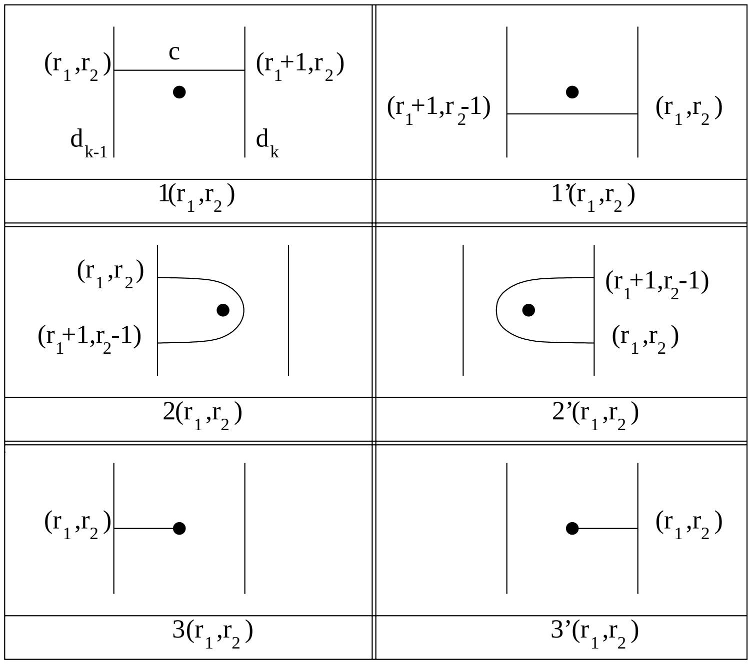



FIGURE 12. Connected components of $d \cap D_k$, for $d$ in normal form and $1 \leq k \leq m$.

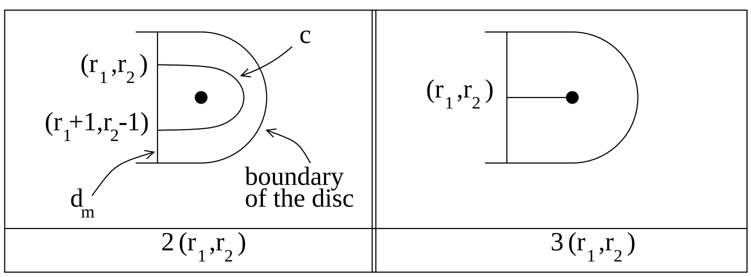



FIGURE 13. The same for $k = m + 1$.

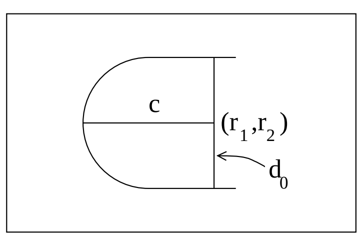



FIGURE 14. The same for $k = 0$.

and such that each connected component of $c \cap D_k$ belongs to one of the types listed in Figures 12–14, is already in normal form.

*Remark* 3.16. For a completely accurate formulation, one should consider the group $\mathcal{G}_k$ of diffeomorphisms $f : D_k \longrightarrow D_k$ which satisfy $f|\partial D \cap D_k = \mathrm{id}$, $f(d_{k-1}) = d_{k-1}$, $f(d_k) = d_k$, and $f(\Delta \cap D_k) = \Delta \cap D_k$. Our list of types classifies each connected component of $c \cap D_k$ up to an isotopy in $\mathcal{G}_k$.

Lemma 3.15 implies that the number of connected components of each type, as well as their relative position and the way in which they join each other, is an invariant of the isotopy class of $c$.

For the rest of this section, $c$ is an admissible curve in normal form. The points of $\mathrm{cr}(c) = c \cap (d_0 \cup d_1 \cup \cdots \cup d_m)$ are called *crossings*, and those in the subset $c \cap d_k$ are called *$k$-crossings* of $c$. The connected components of $c \cap D_k$, $0 \leq k \leq m$, are called *segments* of $c$. A segment is *essential* if its endpoints are both crossings (as opposed to points of $\Delta \cup \partial D$). Thus, the essential segments are precisely those of type $1, 1', 2$ or $2'$. Note that the basic curves $b_0, \ldots, b_m$ have no essential segments.

$c$ can be reconstructed up to isotopy by listing its crossings, plus the types of essential segments bounded by consecutive crossings as one travels along $c$ from one end to the other. Conversely, Lemma 3.15 shows that this combinatorial data is an invariant of the isotopy class of $c$.

We now discuss how the normal form changes under the half-twist $t_k$ along $b_k$. The curve $t_k(c)$, in general, is not in normal form. $t_k(c)$, however, has minimal intersection with each $d_i$ for $i \neq k$. To bring $t_k(c)$ into normal form we only need to simplify its intersections with $d_k$. More precisely, there is a bijection between connected components of the intersection of $c$ and $t_k(c)$ with $D_k \cup D_{k+1}$. Take a connected component $y$ of $c \cap (D_k \cup D_{k+1})$ and deform (inside $D_k \cup D_{k+1}$) the twisted component $t_k(y)$, keeping its intersection with $d_{k-1} \cup d_{k+1} \cup \Delta$ fixed, so that it has minimal intersection with $d_k$. After simplifying each connected component of $c \cap (D_k \cup D_{k+1})$ in this way we end up with normal form for $t_k(c)$. We collect our observations into the following.

**Proposition 3.17.** (a) *The normal form of $t_k(c)$ coincides with $c$ outside of $D_k \cup D_{k+1}$. The curve $t_k(c)$ can be brought into normal form by an isotopy inside $D_k \cup D_{k+1}$.*

(b) *Now assume that $t_k(c)$ is in normal form. There is a natural bijection between $i$-nodes of $c$ and $t_k(c)$ for each $i \neq k$. There is a natural bijection between connected components of intersections of $c$ and $t_k(c)$ with $D_k \cup D_{k+1}$.*

Define a *$k$-string of $c$* as a connected component of $c \cap (D_k \cup D_{k+1})$. Denote by $\mathrm{st}(c, k)$ the set of $k$-strings of $c$.

Define a *$k$-string* as a curve in $D_k \cup D_{k+1}$ which is a connected component of $c \cap (D_k \cup D_{k+1})$ for some admissible curve $c$ (recall that we assume $c$ to be in normal form).

We say that two $k$-strings are isotopic (or belong to the same isotopy class) if there is a deformation of one into the other via diffeomorphisms $f$ of $D' = D_k \cup D_{k+1}$ which satisfy $f(d_{k-1}) = d_{k-1}, f(d_{k+1}) = d_{k+1}$ and $f(\Delta \cap D') = \Delta \cap D'$.

Isotopy classes of $k$-strings can be divided into types as follows: there are five infinite families $I_u$, $II_u$, $II'_u$, $III_u$, $III'_u$ ($u \in \mathbb{Z}$) and five exceptional types $IV$, $IV'$, $V$, $V'$ and $VI$. The exceptional types, and the members $u = 0$ of the families, are drawn in Figure 15. The rule for generating the other members is that the $(u+1)$-st is obtained from the $u$-th by applying $t_k$. As an example, Figure 16 shows type $I_3$. For $k = m$ there is a similar list, which consists of two families and two exceptional

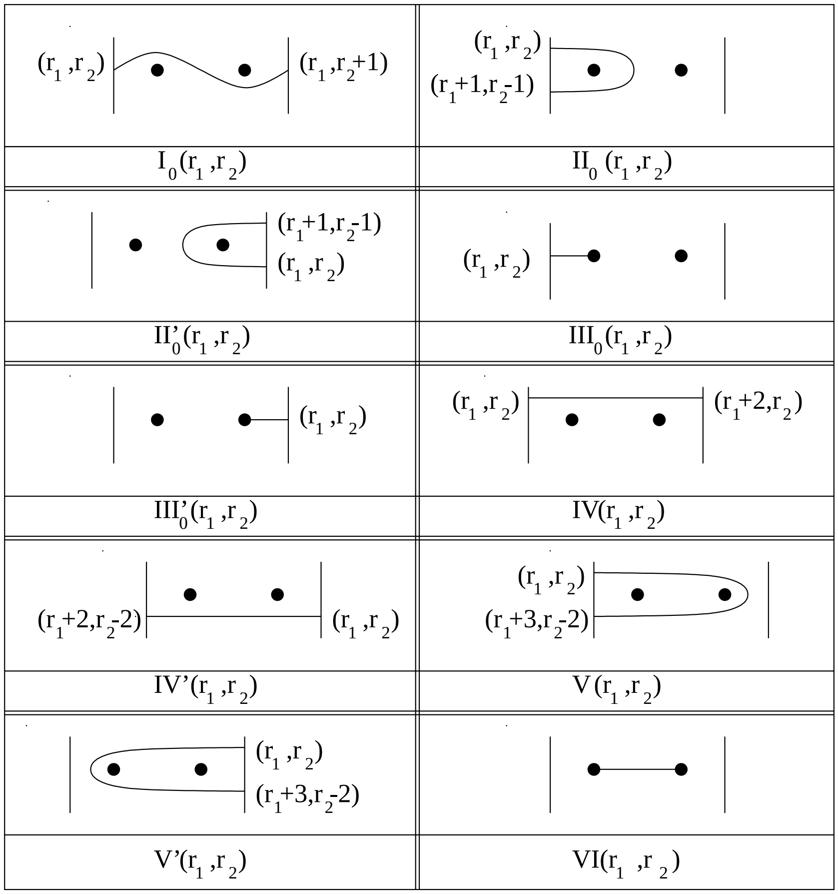



FIGURE 15. Isotopy classes of $k$-strings, for $1 \leq k < m$.

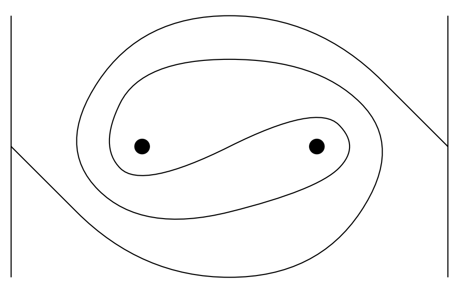


FIGURE 16. Type $I_3$.

types (Figure 17). Finally, for $k = 0$ there are just five exceptional types (Figure 18).

According to our definition, $k$-strings are already assumed to be in "normal form", i.e., to have minimal intersection with $d_k$. Define *crossings* and *essential segments* of $k$-strings in the same way as for admissible curves in normal form. Denote by $\mathrm{cr}(g)$ the set of crossings of a $k$-string $g$.

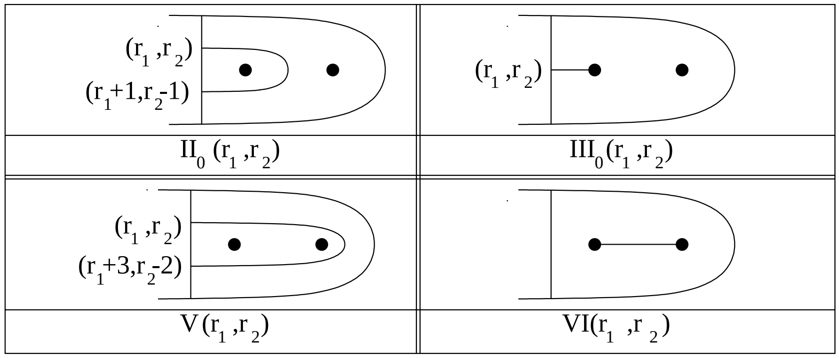



FIGURE 17. Isotopy classes of $m$-strings.

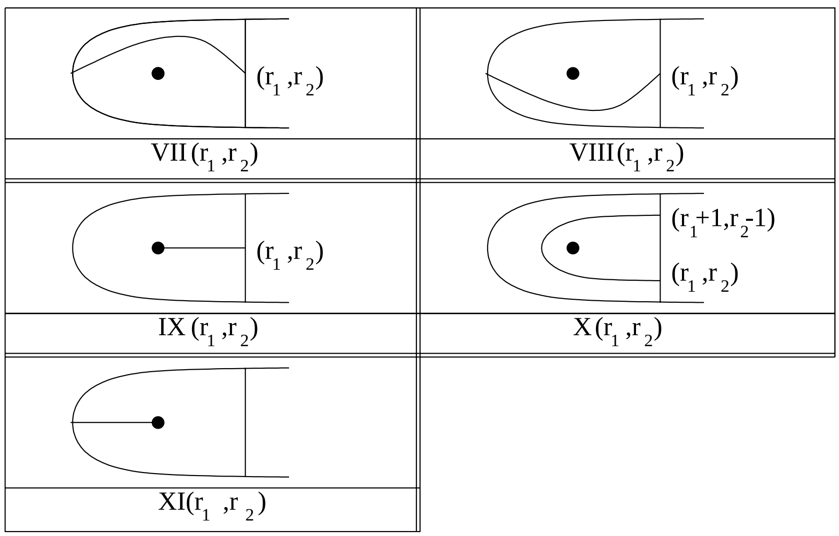



FIGURE 18. Isotopy classes of 0-strings.

**Lemma 3.18.** *If $k > 0$, the geometric intersection number $I(b_k, c)$ can be computed as follows: every $k$-string of $c$ which is of type $I_u$, $II_u$, $II'_u$ or $VI$ contributes 1, those of type $III_u$, $III'_u$ contribute 1/2, and the rest zero. Similarly, for $k = 0$ the types $VIII$ and $X$ contribute 1, the types $IX$ and $XI$ contribute 1/2, and $VII$ zero.*

*Outline of the proof.* We will discuss only the case $k > 0$. The first step is the following result, which is quite easy to prove: there is an isotopy rel $d_0 \cup \cdots \cup d_{k-1} \cup d_{k+1} \cup \cdots \cup d_m$ which brings $c$ into minimal intersection with $b_k$. This yields a lower bound for the geometric intersection number under discussion. For instance, if $c \cap (D_k \cup D_{k+1})$ has a $k$-string of type $I_0$, this gives rise to one intersection point with $b_k$ which cannot be removed by our isotopy. Conversely, one can construct an explicit isotopy whereby this lower bound is attained. $\square$

We will now adapt the discussion to bigraded curves. Choose bigradings $\tilde{b}_k, \tilde{d}_k$ of $b_k, d_k$ for $0 \leq k \leq m$, such that

$$I^{\mathrm{bigr}}(\tilde{d}_k, \tilde{b}_k) = 1 + q_1^{-1} q_2,$$
$$I^{\mathrm{bigr}}(\tilde{b}_k, \tilde{b}_{k+1}) = 1.$$

These conditions determine the bigradings uniquely up to an overall shift $\chi(r_1, r_2)$. Let $\tilde{c}$ be a bigrading of an admissible curve $c$ (as before, $c$ is assumed to be in normal form). Let $a \subset c$ be a connected component of $c \cap D_k$ for some $k$, and let $\tilde{a}$ be the part of $\tilde{c}$ which lies over $a$. Evidently, $\tilde{a}$ is determined by $a$ together with the local index $\mu^{\mathrm{bigr}}(\tilde{b}_{k-1}, \tilde{a}; z)$ or $\mu^{\mathrm{bigr}}(\tilde{d}_k, \tilde{a}; z)$ at any point $z \in (d_{k-1} \cup d_k) \cap a$ (if there is more than one such point, the local indices determine each other). The resulting classification of the pieces $(a, \tilde{a})$ into various types, together with the local indices, is indicated by the pairs of integers in Figures 12–14. For example, if we have a connected component $(a, \tilde{a})$ of type $1(r_1, r_2)$ whose endpoints are $z_0 \in d_{k-1} \cap a$ and $z_1 \in d_k \cap a$, the local indices are $\mu^{\mathrm{bigr}}(\tilde{d}_{k-1}, \tilde{a}; z_0) = (r_1, r_2)$ and $\mu^{\mathrm{bigr}}(\tilde{d}_k, \tilde{a}; z_1) = (r_1 + 1, r_2)$.

In Section 3d we described a canonical lift $\tilde{f}$ of a diffeomorphism $f \in \mathcal{G}$ to a diffeomorphism of $\widetilde{P}$. Denote by $\tilde{t}_k$ the canonical lift of the twist $t_k$ along the curve $b_k$.

**Proposition 3.19.** *Diffeomorphisms $\tilde{t}_i$ for $1 \leq i \leq m$ induce a braid group action on the set of isotopy classes of admissible bigraded curves. In particular, if $\tilde{c}$ is an admissible bigraded curve, the following pairs of bigraded curves are isotopic:*

$$\begin{aligned} \tilde{t}_i \tilde{t}_{i+1} \tilde{t}_i(\tilde{c}) &\simeq \tilde{t}_{i+1} \tilde{t}_i \tilde{t}_{i+1}(\tilde{c}), \\ \tilde{t}_i \tilde{t}_j(\tilde{c}) &\simeq \tilde{t}_j \tilde{t}_i(\tilde{c}), \qquad |i-j| > 1. \end{aligned}$$

A crossing of $c$ will also be called a *crossing* of $\tilde{c}$. We denote by $\mathrm{cr}(\tilde{c})$ the set of crossings of $\tilde{c}$, it is the same set as $\mathrm{cr}(c)$. Crossings of $\tilde{c}$ come with more information than crossings of $c$, namely, each crossing comes with a local index, which is an element of $\mathbb{Z}^2$.

The local index $(r_1, r_2)$ of a crossing $x$ of $\tilde{c}$ will be denoted $(x_1, x_2)$, to emphasize that the index is a function of the crossing. Let $x_0$ denote the index of the vertical curve which contains the crossing: $x \in d_{x_0} \cap c$.

*Essential segments* of $\tilde{c}$ are defined as essential segments of $c$ together with bigradings; the bigrading can be expressed by assigning local indices to the ends of the segment.

Define a *$k$-string* of $\tilde{c}$ as a connected component of $\tilde{c} \cap (D_k \cup D_{k+1})$, together with the bigrading induced from that of $\tilde{c}$. Denote by $\mathrm{st}(\tilde{c}, k)$ the set of $k$-strings of $\tilde{c}$.

Define a *bigraded $k$-string* as a bigraded curve in $D_k \cup D_{k+1}$ which is a connected component of $\tilde{c} \cap (D_k \cup D_{k+1})$ for some bigraded curve $\tilde{c}$.

Figures 15–18 depict the isotopy classes of bigraded $k$-strings. The types $VI$ and $XI$ do not fit immediately into our notational scheme, since they do not intersect $d_{k-1} \cup d_{k+1}$. Instead, given a bigraded $k$-string $\tilde{g}$ with the underlying $k$-string $g$ of type $VI$ or $XI$ ($\tilde{g}$ is then a bigraded admissible curve), we say that $\tilde{g}$ has type $VI(r_1, r_2)$, respectively $XI(r_1, r_2)$, if $\tilde{g} = \chi(r_1, r_2)\tilde{b}_k$.

The next result is the bigraded analogue of Lemma 3.18.

**Lemma 3.20.** *Let $(c, \tilde{c})$ be a bigraded curve. Then $I^{\mathrm{bigr}}(\tilde{b}_k, \tilde{c})$ can be computed by adding up contributions from each bigraded $k$-string of $\tilde{c}$. For $k > 0$ the contributions are listed in the following table:*

| $I_0(0,0)$ | $II_0(0,0)$ | $II_0'(0,0)$ | $III_0(0,0)$ | $III_0'(0,0)$ | $IV$ | $IV'$ | $V$ | $V'$ | $VI(0,0)$ |
|---|---|---|---|---|---|---|---|---|---|
| $q_1 + q_2$ | $q_1 + q_2$ | $q_1 q_2^{-1} + 1$ | $q_2$ | $1$ | $0$ | $0$ | $0$ | $0$ | $1 + q_2$ |

*and the remaining ones can be computed as follows: to determine the contribution of a component of type, say, $I_u(r_1, r_2)$ one takes the contribution of $I_0(0,0)$ and multiplies it by $q_1^{r_1} q_2^{r_2} (q_1^{-1} q_2)^u$. A parallel result holds for $k = 0$, where the relevant contributions are:*

| $VII(0,0)$ | $VIII(0,0)$ | $IX(0,0)$ | $X(0,0)$ | $XI(0,0)$ |
|---|---|---|---|---|
| $0$ | $q_1 q_2^{-1} + 1$ | $1$ | $q_1 q_2^{-1} + 1$ | $1$ |

The proof consists of a series of tedious elementary verifications, which we omit. The only point worth discussing is the dependence of the parameter $u$. By combining Lemma 3.14 with properties (B2) and (B3) of $I^{\mathrm{bigr}}$, one computes that

$$I^{\mathrm{bigr}}(\tilde{b}_k, \tilde{t}_k(\tilde{c})) = I^{\mathrm{bigr}}(\tilde{t}_k^{-1}(\tilde{b}_k), \tilde{c}) = I^{\mathrm{bigr}}(\chi(1,-1)\tilde{b}_k, \tilde{c}) = (q_1^{-1} q_2) I^{\mathrm{bigr}}(\tilde{b}_k, \tilde{c}).$$

The same holds for the local contributions coming from a single $k$-string, which explains the occurrence of the factor $(q_1^{-1} q_2)^u$.

The bigraded intersection number of $\tilde{b}_k$ and a bigraded $k$-string $\tilde{g}$ (denoted by $I^{\mathrm{bigr}}(\tilde{b}_k, \tilde{g})$) is defined as the contribution, described in Lemma 3.20, of $\tilde{g}$ to the bigraded intersection number $I^{\mathrm{bigr}}(\tilde{b}_k, \tilde{c})$. Note that $I^{\mathrm{bigr}}(\tilde{b}_k, \tilde{g})$ depends only on the isotopy class of $\tilde{g}$.

## 4. Admissible curves and complexes of projective modules

In this section we prove Theorem 1.1 and Corollary 1.2 of the introduction. In Section 4a we associate a complex of $A_m$-modules to a bigraded curve. In Section 4b we prove that this construction relates the braid group actions on bigraded curves and in the category of complexes. Section 4c interprets bigraded intersection numbers as dimensions of homomorphism spaces between complexes associated to bigraded curves.

Throughout this section $\tilde{c}$ denotes a bigraded admissible curve in normal form.

4a. **The complex associated to an admissible curve.** We associate to $\tilde{c}$ an object $L(\tilde{c})$ of the category $\mathcal{C}_m$. It is a complex of projective modules, containing one suitably shifted copy of $P_i$ for every $i$-crossing of $\tilde{c}$. Start by defining $L(\tilde{c})$ as a bigraded $A_m$-module:

$$L(\tilde{c}) = \bigoplus_{x \in \mathrm{nd}(\tilde{c})} P(x), \tag{4.1}$$

where $P(x) = P_{x_0}[-x_1]\{x_2\}$ (see Section 3e, after Proposition 3.19, for the definition of $x_0, x_1, x_2$). For every $x, y \in \mathrm{nd}(\tilde{c})$ define

$$\partial_{yx} : P(x) \to P(y)$$

by the following rules:

- If $x$ and $y$ are the endpoints of an essential segment and $y_1 = x_1 + 1$, then
  (1) If $x_0 = y_0$ (then also $x_2 = y_2 + 1$), then
$$\partial_{yx} : P(x) \longrightarrow P(y) \cong P(x)[-1]\{1\}$$
  is the multiplication on the right by $(x_0|x_0 - 1|x_0) \in A_m$.
  (2) If $x_0 = y_0 \pm 1$, then $\partial_{yx}$ is the right multiplication by $(x_0|y_0) \in A_m$.
- Otherwise $\partial_{xy} = 0$.

Now define the differential as
$$\partial = \sum_{x,y} \partial_{xy}.$$

**Lemma 4.1.** *$(L(\tilde{c}), \partial)$ is a complex of graded projective $A_m$-modules with a grading-preserving differential.*

*Proof.* The equation $\partial^2 = 0$ follows from relations $\partial_{zy}\partial_{yx} = 0$ for any triple $z, y, x$ of crossings, which are implied by defining relations in the ring $A_m$: if there is an arrow from $x$ to $y$ and from $y$ to $z$, then $\partial_{zy}\partial_{yx} : P(x) \to P(z)$ is the right multiplication by a certain product of generators $(i|i \pm 1)$ of $A_m$, and this product is equal to zero in all cases (the whole point is that $(i|i + 1)$ can never be followed by $(i + 1|i)$ since that would mean that we're dealing with a closed curve around one point of $\Delta$, and that has been excluded). $\partial$ is grading-preserving since each $\partial_{yx}$ is. The latter property is easily checked on a case-by-case basis, for each type of essential segments. □

Here is a less formal way to describe this complex. Assign the module $P_i$ to each intersection of $c$ with $d_i$. Notice (see Figures 12–14) that $x_1 = y_1 \pm 1$ if $x$ and $y$ are two ends of an essential segment. Consequently, every essential segment can be oriented in a canonical way, from the endpoint $x$ to the endpoint $y$ with $y_1 = x_1 + 1$. If the segment has type 1, the orientation is from left to right; if type $1'$, the orientation is from right to left; if type 2 or $2'$, the orientation is clockwise. In general, the orientation is clockwise around the marked point in the center of the region $D_k$ containing the segment.

There is a natural choice of a homomorphism from $P_{x_0}$ to $P_{y_0}$. Depending on the type of the segment, the homomorphism is the right multiplication by $(i|i - 1|i), (i|i \pm 1)$ or $(i \pm 1|i)$. Namely, if $x_0 = y_0 = i$ (i.e., both crossings belong to the same vertical line $d_i$), the homomorphism is the right multiplication by $(i|i - 1|i)$. If $x_0 = y_0 \pm 1 = i$ (i.e., the two crossings lie on neighbouring vertical curves $d_i$ and $d_{i\pm1}$), the homomorphism is the right multiplication by $(i|i \pm 1)$.

This construction gives us a chain of projective modules and maps between them, as an example in Figures 19–21 indicates.

The composition of any two consecutive homomorphisms is zero. The homomorphisms are not grading-preserving, in general. We can fix that by starting with one of these modules and appropriately shifting the grading of its neighbours in the chain, then the grading of the neighbours of the neighbours, etc.

We get different results by starting with different modules, but, in fact, the difference could only be in the overall shift of the grading.

Fold the resulting diagram made of projective modules and arrows (homomorphisms) between them so that all arrows go from left to right. We call this diagram the *folded diagram* of $\tilde{c}$. By summing up modules in different columns of the folded diagram we get a complex of projective $A_m$-modules. For the example depicted in Figures 19–23 the complex is

$$\cdots \longrightarrow 0 \longrightarrow P_1 \oplus (P_3\{1\}) \xrightarrow{\partial} P_2 \oplus P_2 \xrightarrow{\partial} P_2\{-1\} \longrightarrow 0 \longrightarrow \cdots .$$

This object is defined only up to an overall shift by $[j]\{k\}$. This is where the bigrading of $\tilde{c}$ kicks in. We pick a crossing $x$ of $\tilde{c}$ and make the overall shift in the bigrading so that the projective module $P_{x_0}$, corresponding to $x$, gets the bigrading $[-x_1]\{x_2\}$. The resulting complex, $L(\tilde{c})$, defined earlier more formally, does not depend on the choice of a crossing $x$.

**Lemma 4.2.** *An $(r_1, r_2)$ shift of a bigraded curve translates into the $[-r_1]\{r_2\}$ shift in the category $\mathcal{C}_m$:*

$$L(\chi(r_1, r_2)\tilde{c}) \cong L(\tilde{c})[-r_1]\{r_2\}.$$

*Proof.* The proof immediately follows from formula (4.1) and the description of the differential in $L(\tilde{c})$. ☐

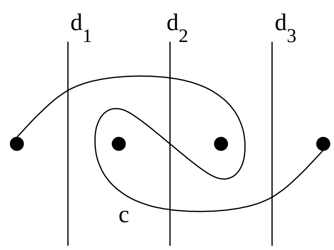



FIGURE 19. A curve $c$.

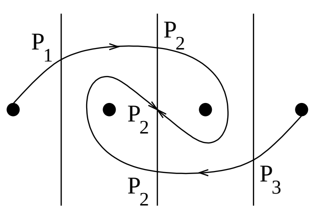



FIGURE 20. Orient essential segments and assign projective modules $P_i$ to intersections of $c$ with $d_i$.

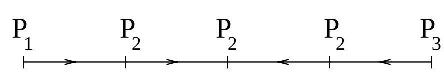



FIGURE 21. Stretch out $c$ and remove two inessential segments at the ends of $c$.

$$P_1 \longrightarrow P_2 \longrightarrow P_2\{-1\} \longleftarrow P_2 \longleftarrow P_3\{1\}$$

FIGURE 22. Replace endpoints of essential segments with the modules associated to these endpoints and shift gradings to make homomorphisms grading-preserving.

$$\begin{tikzcd}[column sep=small, row sep=small]
P_1 \arrow[r] & P_2 \arrow[dr] & \\
 & & P_2\{-1\} \\
P_3\{1\} \arrow[r] & P_2 \arrow[ur] &
\end{tikzcd}$$

FIGURE 23. Fold the diagram to make all arrows go from left to right.

Just like admissible curves, $k$-strings have crossings and essential segments (see Section 3e). In particular, to a bigraded $k$-string $\tilde{g}$ we will associate an object $L(\tilde{g})$ of the category $\mathcal{C}_m$. As a bigraded abelian group,

$$L(\tilde{c}) = \bigoplus_{x \in \mathrm{nd}(g)} P(x),$$

where the sum is over all crossings of $g$, i.e., intersections of $g$ with $d_{k-1} \cup d_k \cup d_{k+1}$ and the differential is read off the essential segments of $g$ in the same way as for admissible curves.

**Example.** The folded diagram of the bigraded $k$-string $I_0(0,0)$ (see Figure 15) is depicted as:

$$\begin{tikzcd}[column sep=small, row sep=small]
P_{k-1} \arrow[dr] & \\
 & P_k \\
P_{k+1}\{1\} \arrow[ur] &
\end{tikzcd}$$

The complex $L(I_0(0,0))$ is $0 \to P_{k-1} \oplus (P_{k+1}\{1\}) \to P_k \to 0$.

*Remark.* If $\tilde{g}$ is a bigraded $k$-string of $\tilde{c}$, then $L(\tilde{g})$ is an abelian subgroup of $L(\tilde{c})$, but not, in general, a subcomplex or a quotient complex.

*Remark.* Lemma 4.2 holds for bigraded $k$-strings as well:

$$L(\chi(r_1, r_2)\tilde{g}) \cong L(\tilde{g})[-r_1]\{r_2\}. \tag{4.2}$$

4b. **Twisting by an elementary braid.** In the previous section we associated an element of $\mathcal{C}_m$ to an admissible bigraded curve. Here we will prove that this map intertwines the braid group action on bigraded curves with the braid group action in the category $\mathcal{C}_m$:

**Theorem 4.3.** *For a braid $\sigma \in B_m$ and an admissible bigraded curve $\tilde{c}$ the objects $\mathcal{R}_\sigma L(\tilde{c})$ and $L(\sigma\tilde{c})$ of the category $\mathcal{C}_m$ are isomorphic.*

It suffices to prove the theorem when $\sigma$ is an elementary braid, $\sigma \in \{\sigma_1, \dots, \sigma_m\}$. In our notations (see Proposition 3.19), the braid $\sigma_k$ acts on bigraded curves as the Dehn twist $\tilde{t}_k$ and on an element of $\mathcal{C}_m$ by tensoring it with the complex of bimodules $R_k$. Thus, we want to prove

**Proposition 4.4.** *For an admissible bigraded curve $\tilde{c}$ and $1 \le k \le m$, the objects $R_k \otimes_{A_m} L(\tilde{c})$ and $L(\tilde{t}_k(\tilde{c}))$ of $\mathcal{C}_m$ are isomorphic.*

*Proof of the proposition.* Denote $L(\tilde{c})$ by $L$. The complex

$$\mathcal{R}_k L = R_k \otimes_{A_m} L$$

is the cone of the map $\mathcal{U}_k(L) \to L$. We start by carefully analyzing $\mathcal{U}_k(L) = P_k \otimes \,_kP \otimes_{A_m} L$. Note that

$$\mathcal{U}_k(P_j) = \begin{cases} P_k \oplus P_k\{1\} & \text{if } j = k, \\ P_k & \text{if } j = k+1, \\ P_k\{1\} & \text{if } j = k-1, \\ 0 & \text{otherwise.} \end{cases}$$

In particular, if a crossing $x$ of $\tilde{c}$ does not belong to any $k$-string of $\tilde{c}$, then $\mathcal{U}_k(P(x)) = 0$.

Pick a bigraded $k$-string $\tilde{g}$ of $\tilde{c}$. We have an inclusion of abelian groups $L(\tilde{g}) \subset L(\tilde{c})$.

**Lemma 4.5.** *The functor $\mathcal{U}_k$, applied to this inclusion of groups, produces an inclusion of complexes*

$$\mathcal{U}_k L(\tilde{g}) \subset \mathcal{U}_k L(\tilde{c}).$$

*Proof of the lemma.* To prove that $\mathcal{U}_k L(\tilde{g})$ is a subcomplex of $\mathcal{U}_k L(\tilde{c})$ we check that

$$\mathcal{U}_k(\partial_{xy}) : \mathcal{U}_k P(x) \longrightarrow \mathcal{U}_k P(y)$$

is zero whenever crossings $x$ and $y$ do not lie in the same $k$-string of $c$. Indeed, for such $x$ and $y$ there are two possibilities:

(1) Either $x$ or $y$ does not lie in a $k$-string of $c$. Then $\mathcal{U}_k P(x) = 0$, respectively $\mathcal{U}_k P(y) = 0$ and $\mathcal{U}_k(\partial_{xy}) = 0$.
(2) $x$ and $y$ belong to different $k$-strings. Then either $\partial_{xy} = 0$ or $x_0 = y_0 = k \pm 1$ and the homomorphism $\partial_{xy} : P(x) \to P(y)$ is the right multiplication by $(x_0|k|x_0)$. Applying $\mathcal{U}_k$ to this homomorphism produces the trivial homomorphism.

The lemma follows. $\square$

From this lemma we deduce

**Proposition 4.6.** *There is a natural direct sum decomposition of complexes*

$$\mathcal{U}_k L \cong \bigoplus_{\tilde{g} \in \mathrm{st}(\tilde{c},k)} \mathcal{U}_k L(\tilde{g}).$$

Earlier we established a bijection between $k$-strings of $\tilde{c}$ and $\tilde{t}_k(\tilde{c})$ (Proposition 3.17)

From the same proposition we have a natural bijection between $j$-crossings of $\tilde{c}$ and $\tilde{t}_k(\tilde{c})$, for each $j$ such that $|j-k| > 1$. If $x$ is such a crossing of $\tilde{c}$ and $x'$ is the associated crossing of $\tilde{t}_k(\tilde{c})$, then $x_1 = x_1'$ and $x_2 = x_2'$. Therefore, the complexes $L$ and $L(\tilde{t}_k(\tilde{c}))$ have a large common piece, namely, the abelian group $\bigoplus_x P(x)$, where the sum is over all crossings $x$ of $\tilde{c}$ such that $|k - x_0| > 1$. Denote this direct sum by $W$.

We want to construct a homotopy equivalence

$$\mathcal{R}_k(L) \cong L(\tilde{t}_k(\tilde{c})).$$

We will obtain it as a composition of homotopy equivalences

$$\mathcal{R}_k(L) = T_0 \cong T_1 \cong \cdots \cong T_{s-1} \cong T_s = L(\tilde{t}_k(\tilde{c})).$$

Every intermediate complex $T_i$ will have a structure similar to that of $\mathcal{R}_k(L)$ and $L(\tilde{t}_k(\tilde{c}))$. As an $A_m$-module, $T_i$ will be isomorphic to the direct sum of $W$ and $A_m$-modules underlying certain complexes $T_i(\tilde{g})$, homotopy equivalent to $\mathcal{R}_k(\tilde{g})$, over all bigraded $k$-strings $\tilde{g}$ in $\tilde{c}$:

$$T_i = W \oplus \left( \bigoplus_{\tilde{g} \in \mathrm{st}(\tilde{c},k)} T_i(\tilde{g}) \right).$$

Elementary homotopies $T_i \cong T_{i+1}$ will be of two different kinds:

(i) Homotopies "localized" inside $T_i(\tilde{g})$, for some bigraded $k$-string $\tilde{g}$. Namely, the homotopy will be the identity on $W$ and on $T_i(\tilde{h})$, for all $k$-strings $\tilde{h}$ of $\tilde{c}$ different from $\tilde{g}$. The complexes $T_i(\tilde{g})$ and $T_{i+1}(\tilde{g})$ will differ as follows: $T_{i+1}(\tilde{g})$ will be the quotient of $T_i(\tilde{g})$ by an acyclic subcomplex, or $T_{i+1}(\tilde{g})$ will be realized as a subcomplex of $T_i(\tilde{g})$ such that the quotient complex is acyclic. The "localized" homotopy $T_i(\tilde{g}) \cong T_{i+1}(\tilde{g})$ will respect the position of these complexes as subcomplexes of $T_i$ and $T_{i+1}$ and will naturally extend to homotopy equivalence $T_i \cong T_{i+1}$.

(ii) Isomorphisms of complexes $T_i \cong T_{i+1}$. These will come from natural isomorphisms of $A_m$-modules that underlie complexes $T_i$ and $T_{i+1}$ and a direct sum decomposition of the module $T_i = (T_i)^+ \oplus (T_i)^-$. The isomorphism will take $(v^+, v^-) \in T_i$ to $(v^+, -v^-)$.

From now on we proceed on a case-by-case basis, working with different types of bigraded $k$-strings (Figures 15-18). We start with some easy cases.

**Case 1.** If $\tilde{g}$ is of type $VI$, then $\tilde{c} \simeq \tilde{c}_k$, up a shift in the bigrading. The half twist $\tilde{t}_k(\tilde{c})$ shifts the bigrading by $(-1, 1)$. On the other hand, the complex $R_k \otimes_{A_m} P_k$ is chain homotopic to $P_k[1]\{1\}$. The topological shift matches the algebraic shift (see Lemma 4.2 and formula (4.2)). This gives an isomorphism $L(\tilde{t}_k(\tilde{c}_k)) \cong \mathcal{R}_k(L(\tilde{c}_k))$.

**Case 2.** If $\tilde{g}$ has one of the types $IV, IV', V$ or $V'$, then the half-twist $\tilde{t}_k$ preserves $\tilde{g}$. At the same time, we have

**Lemma 4.7.** *If $\tilde{g}$ is of type $IV, IV', V$ or $V'$, then the complex $\mathcal{U}_k L(\tilde{g})$ is acyclic.*

*Proof of the lemma.* Explicit computation. For instance, if $\tilde{g}$ has type $IV$, then $L(\tilde{g})$ is the complex

$$0 \longrightarrow P_{k-1}\{r_2\} \longrightarrow P_k\{r_2\} \longrightarrow P_{k+1}\{r_2\} \longrightarrow 0. \tag{4.3}$$

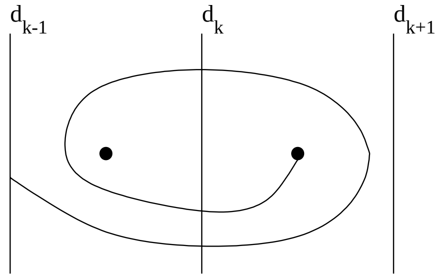
The functor $\mathcal{U}_k$ is the tensor product with the bimodule $P_k \otimes {}_kP$. First we tensor over $A_m$ with the right module ${}_kP$, which gets us a complex of abelian groups. Then we tensor the result with $P_k$ over $\mathbb{Z}$. The tensor product of the complex (4.3) with ${}_kP$ over $A_m$ is an acyclic complex of abelian groups (exercise for the reader). The lemma follows for $\tilde{g}$ of type $IV$. Similar computations establish the lemma for the other three types of $\tilde{g}$. □

Since $\tilde{t}_k(\tilde{g}) \simeq \tilde{g}$, we have a natural isomorphism of complexes $L(\tilde{t}_k(\tilde{g})) \cong L(\tilde{g})$. Moreover, $\mathcal{U}_k(L(\tilde{g}))$, which is a direct summand of the complex $\mathcal{U}_k(L)$, is acyclic, and we see that the complex $\mathcal{R}_k L$ is homotopic to its subcomplex obtained by throwing away $\mathcal{U}_k(L(\tilde{g}))$ for all $\tilde{g}$ of types $IV, IV', V$ and $V'$.

**Case 3.** Types $III$ and $III'$.

Assume $\tilde{g}$ has type $III_u(r_1, r_2)$. The homotopies constructed below commute with shifts in the bigrading, and, without loss of generality, we set $r_1 = r_2 = 0$. We will also set $r_1 = r_2 = 0$ in our treatment of types $II, II'$ and $I$.

Consider first the case $u \geq 0$. A $k$-string of type $III_3$ is depicted as:




The complex $L(\tilde{g})$ is

$$0 \longrightarrow P_k\{u\} \longrightarrow P_k\{u-1\} \longrightarrow \cdots \longrightarrow P_k\{1\} \longrightarrow P_{k-1} \longrightarrow 0.$$

Let us understand the relation between $L(\tilde{g})$ and $L$. The complex $L$ contains $L(\tilde{g})$ as an abelian subgroup. The part of the folded diagram of $\tilde{c}$ that contains the folded diagram of $\tilde{g}$ has the form

$$P_k\{u\} \to P_k\{u-1\} \to \cdots \to P_k\{1\} \to P_{k-1} \leftrightarrow \nabla,$$

where $\nabla$ denotes the complement to $L(\tilde{g})$ in $L$. More precisely, $\nabla$ denotes the direct sum $\bigoplus_x P(x) \subset L$ over all crossings $x$ of $\tilde{c}$ that do not belong to $\tilde{g}$. The notation $P_{k-1} \leftrightarrow \nabla$ means that $\nabla$ relates to $L(\tilde{g})$ through the module $P_{k-1}$. Namely, in the folded diagram of $L$ there might be an arrow $P(x) \to P_{k-1}$ or an arrow $P_{k-1} \to P(x)$, where $x$ is a crossing of $\tilde{c}$ which does not belong to $\tilde{g}$.

We will use this notation throughout the rest of the proof of Proposition 4.4. $\nabla$ will denote either such a complement to $L(\tilde{g})$ in $L$, or the complement to $\mathcal{R}_k(L(\tilde{g}))$ in $\mathcal{R}_kL$, or any similar complement, and the meaning will always be clear from the context. Whenever we modify $\mathcal{R}_kL(\tilde{g})$ to a homotopy equivalent complex $L'$, we will do so relative to $\nabla$, so that our homotopy equivalence extends to a homotopy equivalence between $\mathcal{R}_kL$ and the complex naturally built out of $L'$ and $\nabla$. When we are in such a situation, we say that the homotopy *respects* $\nabla$.

The complex $\mathcal{R}_k(L(\tilde{g}))$, written as a bicomplex, and considered as a part of $\mathcal{R}_kL$, has the form

$$\begin{array}{ccccccccc}
P_k\{u+1\} \oplus P_k\{u\} & \to & \cdots & \to & P_k\{2\} \oplus P_k\{1\} & \to & P_k\{1\} & & \\
\downarrow & & \cdots & & \downarrow & & \downarrow & & \\
P_k\{u\} & \to & \cdots & \to & P_k\{1\} & \to & P_{k-1} & \leftrightarrow & \nabla.
\end{array}$$

Note that, in view of Lemma 4.5, $\nabla$ connects to $\mathcal{R}_k(L(\tilde{g}))$ only through the module $P_{k-1}$ in the bottom right corner of the diagram.

The complex $\mathcal{U}_k(L(\tilde{g}))$ (the top row in the above diagram) is isomorphic to the direct sum of $P_k\{u+1\}$ and acyclic complexes

$$0 \longrightarrow P_k\{i\} \xrightarrow{\text{id}} P_k\{i\} \longrightarrow 0, \qquad 1 \leq i \leq u$$

(we leave this as an exercise for the reader; it follows at once after the differential in $\mathcal{U}_k(L(\tilde{g}))$ is written down), so that $\mathcal{R}_k(L(\tilde{g}))$ is homotopic to

$$P_k\{u+1\} \longrightarrow P_k\{u\} \longrightarrow \cdots \longrightarrow P_k\{1\} \longrightarrow P_{k-1} \longleftrightarrow \nabla,$$

which is exactly $L(\tilde{t}_k(\tilde{g}))$ (note that the homotopy respects $\nabla$).

We next treat the case $III_u$ where $u < 0$. As before, set $r_1 = r_2 = 0$; this does not reduce the generality since our homotopies will commute with bigrading shifts.

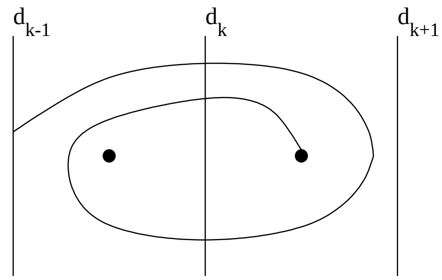



FIGURE 24. $k$-string of type $III_{-3}$.

The complex $L(\tilde{g})$ (considered as a part of $L$) has the form (compare to Figure 24)

$$\nabla \longleftrightarrow P_{k-1} \longrightarrow P_k \longrightarrow P_k\{-1\} \longrightarrow \cdots \longrightarrow P_k\{1+u\}.$$

The complex $\mathcal{R}_k(L(\tilde{g}))$, in relation to $\mathcal{R}_k(L)$, has the form:

$$\begin{array}{ccccccccc}
P_k\{1\} & \to & \begin{pmatrix} P_k \\ \oplus \\ P_k\{1\} \end{pmatrix} & \to & \begin{pmatrix} P_k\{-1\} \\ \oplus \\ P_k \end{pmatrix} & \to & \cdots & \to & \begin{pmatrix} P_k\{1+u\} \\ \oplus \\ P_k\{2+u\} \end{pmatrix} \\
\downarrow & & \downarrow & & \downarrow & & & & \downarrow \\
P_{k-1} & \to & P_k & \to & P_k\{-1\} & \to & \cdots & \to & P_k\{1+u\} \\
\updownarrow & & & & & & & & \\
\nabla & & & & & & & &
\end{array}$$

The complex $\mathcal{U}_k(L(\tilde{g}))$ (top row in the diagram above) is isomorphic to the direct sum of $P_k\{1+u\}$ and acyclic complexes

$$0 \longrightarrow P_k\{i\} \xrightarrow{\mathrm{id}} P_k\{i\} \longrightarrow 0, \qquad \text{for } 1 \geq i \geq 2+u.$$

Therefore, $\mathcal{R}_k(L(\tilde{g}))$ is homotopic to the folding of

$$\begin{array}{ccccccccccccc}
 & & & & & & & & & & & & P_k\{1+u\} \\
 & & & & & & & & & & & & \mathrm{id}\downarrow \\
\nabla & \leftrightarrow & P_{k-1} & \to & P_k & \to & P_k\{-1\} & \to & \cdots & \to & & & P_k\{1+u\}
\end{array}$$

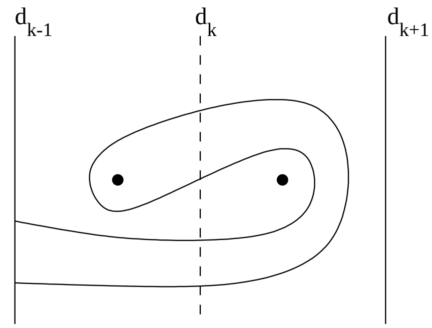
(an explicit computation shows that the vertical arrow is the identity map). This complex has an acyclic subcomplex

$$0 \to P_k\{1+u\} \xrightarrow{\mathrm{id}} P_k\{1+u\} \to 0,$$

and the quotient is isomorphic to $L(\tilde{t}_k(\tilde{g}))$. This sequence of two homotopy equivalences respects $\nabla$. Our consideration of case $III$ is now complete. Type $III'$ can be worked out entirely parallel to $III$, and we omit the computation.

**Case 4.** Types $II$ and $II'$. Let $\tilde{g}$ have type $II_u(r_1, r_2)$. Without loss of generality we can specialize to $r_1 = r_2 = 0$.

We start with the case $u > 0$. The following depicts an example of a $k$-string of that type:




The complex $L(\tilde{g})$ has the form (the first line of arrows continues into the third, and the second into the fourth):

$$\begin{array}{ccccccccc}
P_k\{u\} & \longrightarrow & P_k\{u-1\} & \longrightarrow & \cdots & & & & \\
& \searrow & & & & & & & \\
& & P_k\{u-1\} & \longrightarrow & \cdots & & & & \\
\cdots & \longrightarrow & P_k\{1\} & \longrightarrow & P_{k-1} & \longleftrightarrow & \nabla & & \\
\cdots & \longrightarrow & P_k\{1\} & \longrightarrow & P_k & \longrightarrow & P_{k-1}\{-1\} & \longleftrightarrow & \nabla
\end{array}$$

We next write down the complex $\mathcal{U}_k(L(\tilde{g}))$:

$$\begin{array}{ccccccccc}
P_k\{u\} \oplus P_k\{u+1\} & \to & P_k\{u-1\} \oplus P_k\{u\} & \to & \cdots & & & & \\
& \searrow & & & & & & & \\
& & P_k\{u-1\} \oplus P_k\{u\} & \to & \cdots & & & & \\
\cdots & \to & P_k\{1\} \oplus P_k\{2\} & \to & P_k\{1\} & \to & 0 & & \\
\cdots & \to & P_k\{1\} \oplus P_k\{2\} & \to & P_k \oplus P_k\{1\} & \to & P_k & \to & 0
\end{array}$$

It is easily checked that this complex is isomorphic to the direct sum of acyclic complexes $0 \to P_k\{u\} \xrightarrow{\mathrm{id}} P_k\{u\} \to 0$ (two complexes for each $i$, $1 \le i \le u$, and one for $i = 0$) and the complex

$$\begin{array}{ccccc}
P_k\{u\} \oplus P_k\{u+1\} & \xrightarrow{\psi_1} & P_k\{u\} & \longrightarrow & 0 \\
& \searrow\psi_1 & & & \\
& & P_k\{u\} & \longrightarrow & 0,
\end{array} \tag{4.4}$$

where $\psi_1$ is the projection onto the first summand, taken with the minus sign.

Denote by $L'$ the subcomplex of $\mathcal{R}_k(L(\tilde{g}))$ which, as an abelian group, is the direct sum of $L(\tilde{g})$ and (4.4). Note that $L'$ is homotopy equivalent to $\mathcal{R}_k(L(\tilde{g}))$, and that this homotopy equivalence respects $\nabla$. The complex $L'$ consists of the "central part"

$$\begin{array}{ccccccccc}
& P_k\{u\} & \xleftarrow{\psi_1} & & P_k\{u\} \oplus P_k\{u+1\} & & \xrightarrow{\psi_1} & P_k\{u\} & \\
& \downarrow & & & \psi_2\downarrow & & & \downarrow & \\
\leftarrow & P_k\{u-1\} & \leftarrow & & P_k\{u\} & & \to & P_k\{u-1\} & \to
\end{array} \tag{4.5}$$

and two chains of $P_k\{i\}$'s going to the left and right of the center. $\psi_2$ is given by

$$\psi_2(a,b) = a + b(k|k-1|k), \qquad a \in P_k\{u\}, \quad b \in P_k\{u+1\}. \tag{4.6}$$

The differential in this complex is injective on $P_k\{u\}$ situated in the center of the top row of (4.5). Thus, $P_k\{u\}$ generates an acyclic subcomplex $0 \to P_k\{u\} \xrightarrow{\mathrm{id}} P_k\{u\} \to 0$ of $L'$. The quotient of $L'$ by this acyclic subcomplex is isomorphic to (the diagram below only depicts the central part of the complex, don't forget to add two tails of $P_k$'s)

$$\cdots \longleftarrow P_k\{u-1\} \xleftarrow{-(k|k-1|k)} P_k\{u\} \xrightarrow{(k|k-1|k)} P_k\{u-1\} \longrightarrow \cdots .$$

Denote this quotient by $L''$. The $A_m$-modules underlying complexes $L''$ and $L(\tilde{t}_k(\tilde{g}))$ are canonically isomorphic. The two differentials differ only in one place: in $L''$ there is a minus sign in the left arrow. We can get rid of this minus sign with the isomorphism $L'' \cong L(\tilde{t}_k(\tilde{g}))$ that takes $a \in P(x)$ to $\pm a$ depending on whether $P(x)$ is to the left or right of the "minus" arrow.

We next treat the case $u = 0$. The complex $L(\tilde{g})$ is

$$\nabla \leftrightarrow P_{k-1} \longrightarrow P_{k-1}\{-1\} \leftrightarrow \nabla,$$

and the complex $\mathcal{R}_k(L(\tilde{g}))$, for $\tilde{g}$ of type $II_0$, is the folding of

$$\begin{array}{ccccccc} & & P_k\{1\} & \xrightarrow{0} & P_k & & \\ & & \downarrow & & \downarrow & & \\ \nabla & \leftrightarrow & P_{k-1} & \longrightarrow & P_{k-1}\{-1\} & \leftrightarrow & \nabla. \end{array} \tag{4.7}$$

We want to find a homotopy equivalence with the complex $L(\tilde{t}_k(\tilde{g}))$, which is

$$\begin{array}{ccccccc} & & P_k\{1\} & \xrightarrow{-(k|k-1|k)} & P_k & & \\ & & \downarrow & & \downarrow & & \\ \nabla & \leftrightarrow & P_{k-1} & & P_{k-1}\{-1\} & \leftrightarrow & \nabla. \end{array} \tag{4.8}$$

In fact, these two complexes are isomorphic, via the isomorphism which takes an element

$$\begin{pmatrix} x & y \\ z & w \end{pmatrix}, \qquad x \in P_k\{1\}, y \in P_k, z \in P_{k-1}, w \in P_{k-1}\{-1\},$$

of (4.8) to

$$\begin{pmatrix} x & -y - z(k-1|k) \\ z & -w \end{pmatrix}$$

of the complex (4.7).

Since our map takes $w \in P_{k-1}\{-1\}$ to $-w$, when extending it to the remaining part of the complex we will send $v$ to $-v$ for any $v$ that belongs to $\nabla$ on the right side of the diagram (4.8). This is to insure that our map commutes with differentials, i.e., it is a map of complexes.

Case $u<0$. Let $\tilde{g}$ be of type $II_u$ for $u<0$. The complex $L(\tilde{g})$ is the folding of

$$\begin{array}{ccccccccccc}
 & & \nabla & \longleftrightarrow & P_{k-1}\{-1\} & \longrightarrow & P_k\{-1\} & \longrightarrow & \cdots & & \\
\nabla & \longleftrightarrow & P_{k-1} & \longrightarrow & P_k & \longrightarrow & P_k\{-1\} & \longrightarrow & \cdots & & \\
 & & & & \cdots & \longrightarrow & P_k\{u+1\} & \longrightarrow & P_k\{u\} & & \\
 & & & & & & & \nearrow & & & \\
 & & & & \cdots & \longrightarrow & P_k\{u+1\} & & & &
\end{array}$$

The situation is very similar to the case $u>0$. The complex $\mathcal{U}_k(L(\tilde{g}))$ is isomorphic to the direct sum of acyclic complexes and of the complex which is the folding of

$$P_k\{u+1\} \xrightarrow{\psi_1} P_k\{u\} \oplus P_k\{u+1\} \xleftarrow{\psi_1} P_k\{u+1\}, \tag{4.9}$$

where $\psi_1 = (0, -\mathrm{id})$.

The complex $L'$, which is the direct sum of $L(\tilde{g})$ and (4.9) inside $\mathcal{R}_k(L(\tilde{g}))$, consists of the central part, which is the folding of

$$\begin{array}{ccccccccc}
 & P_k\{u+1\} & \xrightarrow{\psi_1} & P_k\{u\} \oplus P_k\{u+1\} & \xleftarrow{\psi_1} & P_k\{u+1\} & \\
 & {\scriptstyle \mathrm{id}}\downarrow & & {\scriptstyle \psi_2}\downarrow & & {\scriptstyle \mathrm{id}}\downarrow & \\
\rightarrow & P_k\{u+1\} & \rightarrow & P_k\{u\} & \leftarrow & P_k\{u+1\} & \leftarrow
\end{array} \tag{4.10}$$

and of two long tails of $P_k\{i\}$'s. The map $\psi_2$ is given by the formula (4.6).

The module $P_k\{u\}$ in the center of the top row of (4.10) generates an acyclic complex inside $L'$. The quotient of $L'$ by this acyclic subcomplex is isomorphic to $L(\tilde{t}_k(\tilde{g}))$ (things work here in the same way as in the case $u<0$, treated earlier).

**Case 5.** Type $I$. This case is very similar to $II$. Write down the complex $L(\tilde{g})$, compute $\mathcal{U}_k(L(\tilde{g}))$, throw away long tails of acyclic subcomplexes in $\mathcal{U}_k(L(\tilde{g}))$; we will be left with (4.4) (case $u>0$) or (4.9) (case $u \le 0$). Form the direct sum $L'$ of $L(\tilde{g})$ and what's left of $\mathcal{U}_k(L(\tilde{g}))$. The complex $L'$ is a subcomplex of $\mathcal{R}_k(L(\tilde{g}))$. Quotient out $L'$ by an acyclic subcomplex in the same way as in case $II$. The result is isomorphic to $L(\tilde{t}_k(\tilde{g}))$.

This completes our case-by-case analysis. For each $k$-string $\tilde{g}$ of $\tilde{c}$ we constructed a homotopy equivalence $\mathcal{R}_k(L(\tilde{g})) \cong L(\tilde{t}_k(\tilde{g}))$ which extends naturally to $\nabla$. Putting all these equivalences together we obtain a homotopy equivalence $\mathcal{R}_k(L) \cong L(\tilde{t}_k(\tilde{c}))$. The proof of Proposition 4.4 is complete. □

Proposition 4.4 implies (taking $\tilde{c} = \tilde{b}_k$)

**Corollary 4.8.** *For a braid $\sigma \in B_m$ objects $\mathcal{R}_\sigma P_k$ and $L(\sigma\tilde{b}_k)$ of the category $\mathcal{C}_m$ are isomorphic.*

4c. **Bigraded intersection numbers and dimensions of homomorphism spaces.** The main result of this section is:

**Proposition 4.9.** *For $\sigma, \tau \in B_{m+1}, s_1, s_2 \in \mathbb{Z}$ and $0 \le k, j \le m$ the abelian group*

$$\mathrm{Hom}_{\mathcal{C}_m}(\mathcal{R}_\tau P_k, \mathcal{R}_\sigma P_j[s_1]\{-s_2\})$$

*is free. The Poincaré polynomial*

$$\sum_{s_1, s_2} \mathrm{rk}(\mathrm{Hom}_{\mathcal{C}_m}(\mathcal{R}_\tau P_k, \mathcal{R}_\sigma P_j[s_1]\{-s_2\})) q_1^{s_1} q_2^{s_2}$$

*is equal to the bigraded intersection number $I^{\mathrm{bigr}}(f_\tau(\tilde{b}_k), f_\sigma(\tilde{b}_j))$.*

*Proof.* The braid group acts in the category $\mathcal{C}_m$. In particular,

$$\mathrm{Hom}_{\mathcal{C}_m}(\mathcal{R}_\tau P, \mathcal{R}_\tau Q) \cong \mathrm{Hom}_{\mathcal{C}_m}(P, Q).$$

The bigraded intersection number is also preserved by the braid group action:

$$I^{\mathrm{bigr}}(f_\tau(\tilde{c}_1), f_\tau(\tilde{c}_2)) = I^{\mathrm{bigr}}(\tilde{c}_1, \tilde{c}_2).$$

It suffices, therefore, to prove the proposition in the case $\tau = 1$. Moreover, Corollary 4.8 tells us that we can substitute $L(\sigma\tilde{b}_j)$ for $\mathcal{R}_\sigma P_j$ in this proposition.

**Lemma 4.10.** *For a bigraded admissible curve $\tilde{c}$ there is a natural isomorphism*

$$\mathrm{Hom}_{\mathcal{C}_m}(P_k, L(\tilde{c})[s_1]\{-s_2\}) \cong \bigoplus_{\tilde{g} \in \mathrm{st}(\tilde{c},k)} \mathrm{Hom}_{\mathcal{C}_m}(P_k, L(\tilde{g})[s_1]\{-s_2\})$$

*for all $s_1, s_2 \in \mathbb{Z}$ and $0 \le k \le m$.*

In other words, homomorphisms from $P_k$ to the complex associated to an admissible curve come from homomorphisms of $P_k$ to complexes associated to $k$-strings of the curve.

*Proof of the lemma.* The complex $\bigoplus_{\tilde{g} \in \mathrm{st}(\tilde{c},k)} L(\tilde{g})$ is obtained from $L(\tilde{c})$ by throwing away all modules $P_j$ for $|j-k| > 1$, and by changing differentials

$$\partial_{xy} : P(x) \to P(y)$$

to zero whenever $x$ and $y$ belong to boundaries of two different $k$-strings of $\tilde{c}$. The lemma follows since

$$\mathrm{Hom}_{\mathcal{C}_m}(P_k, P_j[s_1]\{s_2\}) = 0$$

for all $s_1, s_2 \in \mathbb{Z}$ and all $j$ such that $|j-k| > 1$; and since maps

$$\mathrm{Hom}_{\mathcal{C}_m}(P_k, P(x)) \longrightarrow \mathrm{Hom}_{\mathcal{C}_m}(P_k, P(y))$$

induced by $\partial_{xy}$ as above are zero maps. □

We know that $I^{\mathrm{bigr}}(\tilde{b}_k, \tilde{c})$ is obtained by summing up contributions from all $k$-strings of $\tilde{c}$ (see Lemma 3.20). Similarly, Lemma 4.10 tells us that the space of homomorphisms from $P_k$ to a (possibly shifted) $L(\tilde{c})$ is the direct sum of contributions corresponding to $k$-strings of $\tilde{c}$. Therefore, it suffices to prove

**Lemma 4.11.** *For any bigraded $k$-string $\tilde{g}$ the abelian group*

$$\mathrm{Hom}_{\mathcal{C}_m}(P_k, L(\tilde{g})[s_1]\{-s_2\})$$

*is free. The Poincaré polynomial*

$$p(\tilde{g}) \stackrel{\mathrm{def}}{=} \sum_{s_1,s_2} \mathrm{rk}(\mathrm{Hom}_{\mathcal{C}_m}(P_k, L(\tilde{g})[s_1]\{-s_2\}))q_1^{s_1}q_2^{s_2}$$

*is equal to the bigraded intersection number $I^{\mathrm{bigr}}(\tilde{b}_k, \tilde{g})$.*

*Proof of the lemma.* Let us examine how shifting a bigraded $k$-string by $(r_1, r_2)$ changes its Poincaré polynomial. If $\tilde{g}' = \chi(r_1, r_2)\tilde{g}$, then $L(\tilde{g}') = L(\tilde{g})[-r_1]\{r_2\}$ and $p(\tilde{g}') = q_1^{r_1}q_2^{r_2}p(\tilde{g})$. This shift matches the corresponding shift in the bigraded intersection number:

$$I^{\mathrm{bigr}}(\tilde{b}_k, \tilde{g}') = q_1^{r_1}q_2^{r_2} I^{\mathrm{bigr}}(\tilde{b}_k, \tilde{g}).$$

Therefore, it suffices to prove the lemma in the case when the bigraded $k$-string $\tilde{g}$ has its parameters $(r_1, r_2)$ (see Figures 15–18) set to 0.

We start by treating the case $0 < k < m$. Figure 15 depicts ten possible types of bigraded $k$-strings, and we have already reduced to the $r_1 = r_2 = 0$ case. Notice that types $I, II, II', III$ and $III'$ have an extra integral parameter $u$. Denote by $Y$ one of the above five types. The $k$-string $Y_u$ is given by twisting $Y_0$ by $(\tilde{t}_k)^u$. Since

$$\begin{aligned} I^{\mathrm{bigr}}(\tilde{b}_k, Y_u) &= I^{\mathrm{bigr}}(\tilde{b}_k, (\tilde{t}_k)^u Y_0) = I^{\mathrm{bigr}}((\tilde{t}_k)^{-u}\tilde{b}_k, Y_0) = I^{\mathrm{bigr}}(\chi(u,-u)\tilde{b}_k, Y_0) \\ &= (q_1^{-1}q_2)^u I^{\mathrm{bigr}}(\tilde{b}_k, Y_0) \end{aligned}$$

(the third equality uses Lemma 3.14) and, at the same time,

$$\begin{aligned} p(Y_u) &= \sum_{s_1,s_2} \mathrm{rk}(\mathrm{Hom}_{\mathcal{C}_m}(P_k, L((\tilde{t}_k)^u Y_0)[s_1]\{-s_2\}))q_1^{s_1}q_2^{s_2} \\ &= \sum_{s_1,s_2} \mathrm{rk}(\mathrm{Hom}_{\mathcal{C}_m}(P_k, (\mathcal{R}_k)^u L(Y_0)[s_1]\{-s_2\}))q_1^{s_1}q_2^{s_2} \\ &= \sum_{s_1,s_2} \mathrm{rk}(\mathrm{Hom}_{\mathcal{C}_m}((\mathcal{R}_k)^{-u} P_k, L(Y_0)[s_1]\{-s_2\}))q_1^{s_1}q_2^{s_2} \\ &= \sum_{s_1,s_2} \mathrm{rk}(\mathrm{Hom}_{\mathcal{C}_m}(P_k[-u]\{-u\}, L(Y_0)[s_1]\{-s_2\}))q_1^{s_1}q_2^{s_2} \\ &= \sum_{s_1,s_2} \mathrm{rk}(\mathrm{Hom}_{\mathcal{C}_m}(P_k, L(Y_0)[s_1+u]\{u-s_2\}))q_1^{s_1}q_2^{s_2} \\ &= (q_1^{-1}q_2)^u p(Y_0) \end{aligned}$$

(the second equality follows from Theorem 4.3), it suffices to treat the case $u = 0$.

**Lemma 4.12.** *For $k > 0$ the Poincaré polynomials of bigraded $k$-strings with parameters $r_1 = r_2 = 0$ and (for types $I, II, II', III, III'$) with $u = 0$ are given by the following table:*

| $I_0(0,0)$ | $II_0(0,0)$ | $II'_0(0,0)$ | $III_0(0,0)$ | $III'_0(0,0)$ | $IV$ | $IV'$ | $V$ | $V'$ | $VI(0,0)$ |
|---|---|---|---|---|---|---|---|---|---|
| $q_1+q_2$ | $q_1+q_2$ | $q_1q_2^{-1}+1$ | $q_2$ | $1$ | $0$ | $0$ | $0$ | $0$ | $1+q_2$ |

*The Poincaré polynomials of* 0*-strings with $r_1 = r_2 = 0$ are:*

| $VII(0,0)$ | $VIII(0,0)$ | $IX(0,0)$ | $X(0,0)$ | $XI(0,0)$ |
|---|---|---|---|---|
| $0$ | $q_1q_2^{-1}+1$ | $1$ | $q_1q_2^{-1}+1$ | $1$ |

*Proof of the lemma.* Direct computation. For $0 < k < m$ and a bigraded $k$-string $\tilde{g}$ as above write down the complex $L(\tilde{g})$ according to the prescription in Section 4a. Then compute the hom complex $\mathrm{Hom}(P_k, L(\tilde{g}))$ and the Poincaré polynomial $P(\tilde{g})$. To simplify the computation we can look at all homomorphisms from $P_k$ to $L(\tilde{g})$, not just the ones that preserve the grading, and later sort them out into graded pieces.

Example: $\tilde{g}$ has type $I_0(0,0)$. The complex $L(\tilde{g})$ is (see also the end of Section 4a)

$$0 \longrightarrow P_{k-1} \oplus P_{k+1}\{1\} \longrightarrow P_k \longrightarrow 0.$$

The complex $\mathrm{Hom}(P_k, L(\tilde{g}))$ is quasi-isomorphic to

$$0 \longrightarrow \mathbb{Z}\{1\} \xrightarrow{0} \mathbb{Z} \longrightarrow 0,$$

and the Poincaré polynomial $p(\tilde{g}) = q_1 + q_2$. The remaining nine cases are equally easy to compute; the details are left to the reader. We note that for $\tilde{g}$ of type $IV, IV', V$ and $V'$ the complex $\mathrm{Hom}(P_k, L(\tilde{g}))$ is acyclic and the Poincaré polynomial is 0.

Cases $k = 0$ and $k = m$ can be treated in the same fashion. For $k = m$, the four types of $m$-strings (see Figure 17) naturally correspond to the four out of ten types of $k$-strings with $0 < k < m$, and the computation for the $0 < k < m$ case extends word-for-word to $k = m$.

If $k = 0$, there are five cases to consider. A 0-string has either one or two crossings, and the Poincaré polynomial is easy to compute. The answer is written in the second table of the lemma. □

The tables in Lemmas 3.20 and 4.12 are identical. Lemma 4.11 and Proposition 4.9 follow. □

Proposition 4.9, specialized to $q_1 = q_2 = 1$, implies Theorem 1.1 of the introduction. Proposition 4.9, specialized to $\tau = 1$, and Lemma 3.6 imply that objects $\mathcal{R}_\sigma P_j$ and $P_j$ are isomorphic if and only if $\sigma$ is the trivial braid (a similar argument was used to prove Proposition 3.11). Therefore, the braid group acts faithfully in the category $\mathcal{C}_m$ (which proves Corollary 1.2).

## 5. Floer cohomology

The definition of Floer cohomology used here is essentially the original one [11]; far more general versions have been developed in the meantime, but we do not need them. In fact, we will even impose additional restrictions whenever this simplifies the exposition. The only exception is that we allow Lagrangian submanifolds with boundary. After the expository part, a result is proved which computes Floer cohomology in certain very special situations. This is based on a simple $\mathbb{Z}/2$-symmetry argument, and will be the main technical tool later on. To conclude, we discuss briefly the question of grading.

5a. **Some basic symplectic geometry.** A symplectic manifold with contact type boundary is a compact manifold with boundary $M$, with a symplectic form $\o \in \Omega^2(M)$ and a contact one-form $\alpha \in \Omega^1(\partial M)$, subject to two conditions. One is that $\o|\partial M = d\alpha$, and the other is that the Reeb vector field $R$ on $\partial M$ should satisfy $\o(N, R) > 0$ for any normal vector field $N$ pointing outwards. A Lagrangian submanifold with Legendrian boundary in $(M, \o, \alpha)$ is a submanifold $L \subset M$ which intersects $\partial M$ transversally, such that $\o|L = 0$ and $\alpha|\partial L = 0$ (note that $\partial L = \emptyset$ is allowed). For brevity, we call these objects simply *Lagrangian submanifolds*.

**Lemma 5.1.** *Let $L \subset M$ be a Lagrangian submanifold. There is a unique $Y \in C^\infty(TL|\partial L)$ which satisfies $i_Y \o|T_x(\partial M) = \alpha$ for any $x \in \partial L$.*

The proof is straightforward, and we omit it. By definition, a Lagrangian isotopy is a smooth family $(L_t)_{0 \leq t \leq 1}$ of Lagrangian submanifolds. Infinitesimally, such an isotopy is described at each time $t$ by a section $X_t$ of the normal bundle to $L_t$, or dually by a one-form $\beta_t$ on $L_t$. $\beta_t$ is always closed. Moreover, if $Y_t \in C^\infty(TL_t|\partial L_t)$ is the vector field from Lemma 5.1 for $L = L_t$, then

$$\beta_t|\partial L_t = d(i_{Y_t}\beta_t). \tag{5.1}$$

It follows that the pair $(\beta_t, i_{Y_t}\beta_t)$ defines a class in $H^1(L_t, \partial L_t; \mathbb{R})$. The Lagrangian isotopy is called *exact* if this class is zero for all $t$. This means that there is a $K_t \in C^\infty(L_t, \mathbb{R})$ with $K_t|\partial L_t = i_{Y_t}\beta_t$ and $dK_t = \beta_t$.

A special class of Lagrangian isotopies are those where $\partial L_t$ remains constant for all $t$, which we call *isotopies rel $\partial M$*.

Another important class consists of those isotopies where $\partial L_t$ moves along the Reeb flow with positive speed. By this we mean that there is a function $b \in C^\infty([0;1], \mathbb{R}^{>0})$ such that $b(t)R$, when projected to the normal bundle of $\partial L_t$, is equal to $X_t|\partial L_t$. Using (5.1) one can see that this is the case iff $i_{Y_t}\beta_t \in C^\infty(\partial L_t, \mathbb{R})$ is a positive constant function for each $t$. Another equivalent condition: if one takes an embedding $\partial L_0 \times [0;1] \longrightarrow \partial M$ such that the image of $\partial L_0 \times \{t\}$ is $\partial L_t$, then the pullback of $\alpha$ under it is $\psi(t)\,dt$ for some positive function $\psi$. We call isotopies of this kind *positive*. There are always many positive isotopies, even exact ones, starting at a given Lagrangian submanifold $L$. For instance, one can carry $L$ along the flow of some Hamiltonian function $H \in C^\infty(M, \mathbb{R})$ with $H(x) = 1$ and $dH_x = ø(\cdot, R_x)$ at all points $x \in \partial M$.

Let $(M, ø, \alpha)$ be a symplectic manifold with contact type boundary. One can always extend $\alpha$ to a one-form $\theta$ on some neighbourhood $U \subset M$ of $\partial M$, such that $d\theta = ø|U$ holds. The vector field $Z$ dual to $\theta$, $i_Z ø = \theta$, is Liouville and points outwards along $\partial M$. The embedding $\kappa : (-r_0; 0] \times \partial M \longrightarrow M$, for some $r_0 > 0$, defined by the flow of $Z$ satisfies $\kappa^* ø = d(e^r\alpha)$ and hence $\kappa^*\theta = \kappa^*(i_Z ø) = e^r\alpha$, where $r$ is the variable in $(-r_0; 0]$. Fix some $\theta$ and the corresponding $\kappa$. A Lagrangian submanifold $L \subset M$ is called *$\kappa$-compatible* if $\kappa^{-1}(L) \cap ([-r_1; 0] \times \partial M) = [-r_1; 0] \times \partial L$ for some $0 < r_1 < r_0$. Equivalently, $\theta|L \in \Omega^1(L)$ should vanish near $\partial L$.

**Lemma 5.2.** (a) *Any Lagrangian submanifold can be deformed, by an exact isotopy rel $\partial M$, to a $\kappa$-compatible one.*

(b) *Let $(L_t)_{0 \le t \le 1}$ be a Lagrangian isotopy, and assume that $L_0, L_1$ are compatible with $\kappa$. Then there is another isotopy $(L'_t)$ with the same endpoints, and with $\partial L'_t = \partial L_t$ for all $t$, such that all $L'_t$ are compatible with $\kappa$. If $(L_t)$ is exact, then $(L'_t)$ may be chosen to have the same property.*

*Proof.* (a) Let $L \subset M$ be a Lagrangian submanifold. Choose an $r_1$, $0 < r_1 < r_0$, such that $\partial L$ is a deformation retract of $K = L \cap \kappa([-r_1; 0] \times \partial M)$. $\theta|K$ is a closed one-form which vanishes on $\partial L \subset K$; hence it can be written as the differential of a function on $K$ which vanishes on $\partial L$. Extend this to a function $h \in C^\infty(M, \mathbb{R})$ with $h|\partial M = 0$, and consider the family of one-forms $\theta_t = \theta - t\,dh$, $0 \le t \le 1$. They satisfy $\theta_t|\partial M = \alpha$ and $d\theta_t = \omega$, and hence define a family of embeddings $\kappa_t : [-r_2; 0] \times \partial M \longrightarrow M$ for some $r_2 > 0$. Note that $\theta_0 = \theta$ while $\theta_1|L$ vanishes on a neighbourhood of $\partial L$, namely on $K$.

Let $U$ be a neighbourhood of $\partial M$ which is contained in the image of $\kappa_t$ for all $t$; we can assume that $\partial M \subset U$ is a deformation retract. $X_t = (\partial\kappa_t/\partial t) \circ \kappa_t^{-1}|U$ is a symplectic vector field on $U$ which vanishes on $\partial M$. The corresponding one-form $ø(\cdot, X_t)$ is zero on $\partial M$, and hence it can be written as the boundary of a function $H_t \in C^\infty(U, \mathbb{R})$ with $H_t|\partial M = 0$. After possibly making $U$ smaller one finds that for each $t$, $X_t$ can be extended to a Hamiltonian vector field $\widetilde{X}_t$ on the whole of $M$. Let $\phi_t$ be the flow of $(\widetilde{X}_t)$ seen as a time-dependent vector field. By definition $\phi_t \circ \kappa_0 = \kappa_t$ on some neighbourhood of the boundary, and, therefore,

$$\phi_1^*\theta = (\kappa_1\kappa_0^{-1})^*\theta = (\kappa_0^{-1})^*(e^r\alpha) = \theta_1 \quad \text{near } \partial M.$$

This shows that $\phi_1^*(\theta|\phi_1(L))$ is equal to $\theta_1|L$ near $\partial L$, so that $\phi_1(L)$ is $\kappa$-compatible. One sees easily that the isotopy $L_t = \phi_t(L)$ is exact, which finishes the argument. The proof of (b) is just a parametrized version of (a). $\square$

Assume now that the relative symplectic class $[ø, \alpha] \in H^2(M, \partial M; \mathbb{R})$ is zero. Then there is a one-form $\theta$ on $M$ with $\theta|\partial M = \alpha$ and $d\theta = ø$. Fix such a $\theta$; the corresponding embedding $\kappa$ is defined on all of $\mathbb{R}^{\leq 0} \times \partial M$. A Lagrangian submanifold $L$ is called *$\theta$-exact* if $[\theta|L] \in H^1(L, \partial L; \mathbb{R})$ is zero. The next lemma describes the relationship between the absolute notion of a $\theta$-exact submanifold and the relative notion of exact isotopy. The lemma after that concerns the existence of $\theta$-exact submanifolds in a given Lagrangian isotopy class. We omit the proofs.

**Lemma 5.3.** *Let $(L_t)_{0 \leq t \leq 1}$ be a Lagrangian isotopy such that $L_0$ is $\theta$-exact. Then the isotopy is exact $\Leftrightarrow$ all the $L_t$ are $\theta$-exact.*

**Lemma 5.4.** (a) *Let $L \subset M$ be a Lagrangian submanifold. Assume that $H^1(M, \partial M; \mathbb{R})$ surjects onto $H^1(L, \partial L; \mathbb{R})$. Then there is a Lagrangian isotopy $(L_t)$ rel $\partial M$ with $L_0 = L$ and such that $L_1$ is $\theta$-exact.*

(b) *Let $L_0, L_1 \subset M$ be two $\theta$-exact Lagrangian submanifolds which are isotopic rel $\partial M$. Assume that $H^1(M, \partial M; \mathbb{R})$ surjects onto $H^1(L_j, \partial L_j; \mathbb{R})$. Then there is an exact isotopy rel $\partial M$ which connects $L_0$ and $L_1$.*

There is a similar notion for symplectic automorphisms. Let $\mathrm{Symp}(M, \partial M, ø)$ be the group of symplectic automorphisms $\phi : M \longrightarrow M$ which are the identity in some neighbourhood (depending on $\phi$) of $\partial M$. By definition, a Hamiltonian isotopy in this group is one generated by a time-dependent Hamiltonian function which vanishes near $\partial M$. Call $\phi \in \mathrm{Symp}(M, \partial M, ø)$ *$\theta$-exact* if $[\phi^*\theta - \theta] \in H^1(M, \partial M; \mathbb{R})$ is zero. The analogue of Lemma 5.3 is this: let $(\phi_t)$ be an isotopy in $\mathrm{Symp}(M, \partial M, ø)$ with $\phi_0$ $\theta$-exact. Then $(\phi_t)$ is a Hamiltonian isotopy iff all the $\phi_t$ are $\theta$-exact. As for Lemma 5.4, the analogous statement is actually simpler: the subgroup of $\theta$-exact automorphisms is always a deformation retract of $\mathrm{Symp}(M, \partial M, ø)$. Finally, note that $\theta$-exact automorphisms preserve the class of $\theta$-exact Lagrangian submanifolds.

5b. **The definition.** Let $(M, ø, \alpha)$ be a symplectic manifold with contact type boundary, such that $[ø, \alpha] \in H^2(M, \partial M; \mathbb{R})$ is zero. Fix a $\theta \in \Omega^1(M)$ with $\theta|\partial M = \alpha$, $d\theta = ø$. Floer cohomology associates a finite-dimensional vector space over $\mathbb{Z}/2$, denoted by $HF(L_0, L_1)$, to any pair $(L_0, L_1)$ of $\theta$-exact Lagrangian submanifolds whose boundaries are either the same or else disjoint. We now give a brief account of its definition, concentrating on the issues which arise in connection with the boundary $\partial M$. For expositions emphasizing other aspects see e.g. [28], [34], [35].

We begin with a digression concerning almost complex structures. Let $\xi \longrightarrow \partial M$ be the symplectic vector bundle $\xi = (\ker \alpha, d\alpha)$. Given a compatible almost complex structure $\bar{j}$ on $\xi$, one can define an almost complex structure $\overline{J}$ on $\mathbb{R} \times \partial M$, compatible with $d(e^r\alpha)$, by $\overline{J}(y_1, X + y_2R) = (-y_2, \bar{j}X + y_1R)$ for $X \in \xi$ and $y_1, y_2 \in \mathbb{R}$. The almost complex structures $\overline{J}$ appeared first in [20, p. 529]. The next lemma describes their main property; it is taken from [9, Lemma 3.9.2] with some modifications.

**Lemma 5.5.** *Let $\Sigma$ be an open subset of $\mathbb{R} \times [0; 1]$, and let $(\bar{j}_{s,t})$ be a smooth family of compatible almost complex structures on $\xi$ parametrized by $(s, t) \in \Sigma$. Let $(\overline{J}_{s,t})$ be the corresponding family of almost complex structures on $\mathbb{R} \times \partial M$. Let $\Lambda_0, \Lambda_1 \subset \partial M$ be two Legendrian submanifolds, and let $u = (u_1, u_2) : \Sigma \longrightarrow \mathbb{R} \times \partial M$ be a smooth map which satisfies $\partial u/\partial s + \overline{J}_{s,t}(u)\partial u/\partial t = 0$, such that $u(s, t) \in \mathbb{R} \times \Lambda_t$ for all $(s, t) \in (\mathbb{R} \times \{0; 1\}) \cap \Sigma$. Then $u_1$ has no local maxima.*

*Proof.* Let $pr_1, pr_2$ be the projections from $\mathbb{R} \times \partial M$ to the two factors. For any $(s,t) \in \Sigma$ one has $Dpr_1 = \alpha \circ Dpr_2 \circ \overline{J}_{s,t}$. Hence

$$\partial u_1/\partial s = \alpha(Dpr_2 \circ \overline{J}_{s,t} \circ \partial u/\partial s) = \alpha(\partial u_2/\partial t),$$

and similarly $\partial u_1/\partial t = -\alpha(\partial u_2/\partial s)$. Therefore $\Delta u_1 = -d\alpha(\partial u_2/\partial s, \partial u_2/\partial t) \leq 0$. By assumption $\partial u_2/\partial s \in T\Lambda_t$ at each point $(s,t) \in (\mathbb{R} \times \{0;1\}) \cap \Sigma$, and hence $\partial u_1/\partial t = -\alpha(\partial u_2/\partial s) = 0$. This shows that $u_1$ is subharmonic and satisfies von Neumann boundary conditions on $(\mathbb{R} \times \{0;1\}) \cap \Sigma$. The maximum principle implies that such a function cannot have local maxima. □

Let $\kappa : \mathbb{R}^{\leq 0} \times \partial M \longrightarrow M$ be the embedding associated to $\theta$, as in the previous section. We define $\mathcal{J}$ to be the space of all smooth families $\mathbf{J} = (J_t)_{0 \leq t \leq 1}$ of $\o$-compatible almost complex structures on $M$ which have the following property: there is an $r_0 > 0$ and a family $(\bar{j}_t)$ of compatible almost complex structures on $\xi$ such that for each $t$, $\kappa^* J_t | [-r_0; 0] \times \partial M$ agrees with $\overline{J}_t$.

Now take two $\theta$-exact Lagrangian submanifolds $L_0, L_1 \subset M$ with $\partial L_0 \cap \partial L_1 = \emptyset$. The first step in defining $HF(L_0, L_1)$ is to deform them in a suitable way.

**Lemma 5.6.** *There are Lagrangian submanifolds $L_0', L_1'$ such that $L_j$ and $L_j'$ are joined by an exact isotopy rel $\partial M$, for $j = 0, 1$, and with the following properties: each $L_j'$ is $\kappa$-compatible, and the intersection $L_0' \cap L_1'$ is transverse.*

This follows from Lemma 5.2(a) and a well-known transversality argument. Note that $L_0', L_1'$ are again $\theta$-exact, because of Lemma 5.3. For $\mathbf{J} = (J_t) \in \mathcal{J}$, let $\mathcal{M}(\mathbf{J})$ be the set of smooth maps $u : \mathbb{R} \times [0;1] \longrightarrow M$ satisfying Floer's equation

$$\partial u/\partial s + J_t(u)\partial u/\partial t = 0, \quad u(\mathbb{R} \times \{0\}) \subset L_0', \;\; u(\mathbb{R} \times \{1\}) \subset L_1', \tag{5.2}$$

and whose energy $E(u) = \int u^* \o$ is finite. We remind the reader that (5.2) can be seen as the negative gradient flow equation of the action functional $a$ on the path space $\mathcal{P}(L_0', L_1') = \{\gamma \in C^\infty([0;1], M) \mid \gamma(j) \in L_j',\ j = 0, 1\}$. To define $a$ one needs to choose functions $H_j \in C^\infty(L_j', \mathbb{R})$ with $dH_j = \theta | L_j'$ (such functions exist by the assumption of $\theta$-exactness); then $a(\gamma) = -\int_{[0;1]} \gamma^*\theta + H_1(\gamma(1)) - H_0(\gamma(0))$.

**Lemma 5.7.** *For any $\mathbf{J} \in \mathcal{J}$, there is a compact subset of $M \setminus \partial M$ which contains the image of all $u \in \mathcal{M}(\mathbf{J})$.*

This is mainly a consequence of Lemma 5.5 and of the definition of $\mathcal{J}$. The other elements which enter the proof are: the convergence of each $u \in \mathcal{M}(\mathbf{J})$ towards limits $x_\pm = \lim_{s \to \pm\infty} u(s, \cdot)$ in $L_0' \cap L_1'$; the $\kappa$-compatibility of $L_0', L_1'$; and the fact that $\partial L_0' \cap \partial L_1' = \partial L_0 \cap \partial L_1 = \emptyset$. We leave the details to the reader. The importance of Lemma 5.7 is that it allows one to ignore the boundary of $M$ in all arguments concerning solutions of (5.2).

One can associate to any $u \in \mathcal{M}(\mathbf{J})$ a Fredholm operator $D_{u,\mathbf{J}} : \mathcal{W}_u^1 \longrightarrow \mathcal{W}_u^0$ from $\mathcal{W}_u^1 = \{X \in W^{1,p}(u^*TM) \mid X(\cdot, 0) \in u^*TL_0',\ X(\cdot, 1) \in u^*TL_1'\}$ to $\mathcal{W}_u^0 = L^p(u^*TM)$ ($p > 2$ is some fixed real number) which describes the linearization of (5.2) near $u$. Let $\mathcal{M}_k(x_-, x_+; \mathbf{J}) \subset \mathcal{M}(\mathbf{J})$ be the subset of those $u$ with limits $x_-, x_+ \in L_0' \cap L_1'$ and such that $D_{u,\mathbf{J}}$ has index $k \in \mathbb{Z}$. There is a natural action of $\mathbb{R}$ on each of these sets, by translation in the $s$-variable. Call $u \in \mathcal{M}(\mathbf{J})$ regular if $D_{u,\mathbf{J}}$ is onto. If this is true for all $u \in \mathcal{M}(\mathbf{J})$, then $\mathbf{J}$ itself is called regular.

**Proposition 5.8** (Floer-Hofer-Salamon [14], Oh [29])**.** *The regular $\mathbf{J}$ form a $C^\infty$-dense subset $\mathcal{J}^{\mathrm{reg}}(L_0', L_1') \subset \mathcal{J}$.*

**Proposition 5.9** (Floer). *For $\mathbf{J} \in \mathcal{J}^{\mathrm{reg}}(L_0', L_1')$, the quotients $\mathcal{M}_1(x_-, x_+; \mathbf{J})/\mathbb{R}$ are finite sets. Moreover, if we define $\nu(x_-, x_+; \mathbf{J}) \in \mathbb{Z}/2$ to be the number of points mod 2 in $\mathcal{M}_1(x_-, x_+; \mathbf{J})/\mathbb{R}$, then $\sum_x \nu(x_-, x; \mathbf{J})\nu(x, x_+; \mathbf{J}) = 0$ for each $x_-, x_+$.*

Let $CF(L_0', L_1')$ be the $\mathbb{Z}/2$-vector space freely generated by the points of $L_0' \cap L_1'$, and let $d_{\mathbf{J}}$ be the endomorphism of this space which sends the basis element $\langle x_+ \rangle$ to $\sum_x \nu(x, x_+; \mathbf{J})\langle x \rangle$. Proposition 5.9 implies that $d_{\mathbf{J}}^2 = 0$, and one sets $HF(L_0', L_1'; \mathbf{J}) = \ker d_{\mathbf{J}}/\mathrm{im}\, d_{\mathbf{J}}$. Finally, $HF(L_0, L_1)$ is defined as $HF(L_0', L_1'; \mathbf{J})$ for any choice of $L_0', L_1'$ as in Lemma 5.6 and any $\mathbf{J} \in \mathcal{J}^{\mathrm{reg}}(L_0', L_1')$. One can show that this is independent of the choices up to canonical isomorphism. The proof uses the continuation maps introduced in [35, Section 6]. We will not explain the details, but it seems appropriate to mention two simple facts which are used in setting up the argument: first, an exact Lagrangian isotopy rel $\partial M$ which consists only of $\kappa$-compatible submanifolds can be embedded into a Hamiltonian isotopy in $\mathrm{Symp}(M, \partial M, \emptyset)$. Second, there is an analogue of Lemma 5.7 for solutions of the continuation equation, which is again derived from Lemma 5.5. As well as proving its well-definedness, the continuation argument also establishes the basic isotopy invariance property of Floer cohomology:

**Proposition 5.10.** *$HF(L_0, L_1)$ is invariant (up to isomorphism) under exact isotopies of $L_0$ or $L_1$ rel $\partial M$.*

We will now extend the definition of Floer cohomology to pairs of Lagrangian submanifolds whose boundaries coincide. The basic idea is the same as in the case of geometric intersection numbers. Let $(L_0, L_1)$ be two $\theta$-exact Lagrangian submanifolds with $\partial L_0 = \partial L_1$, and let $L_0^+$ be a submanifold obtained from $L_0$ by a sufficiently small Lagrangian isotopy which is exact and positive. Then $\partial L_0^+ \cap \partial L_1 = \emptyset$, and one sets $HF(L_0, L_1) = HF(L_0^+, L_1)$. The proof that this is independent of the choice of $L_0^+$ requires a generalization of Proposition 5.10:

**Lemma 5.11.** *Let $(L_{0,s})_{0 \leq s \leq 1}$ be an isotopy of $\theta$-exact Lagrangian submanifolds, and let $L_1$ be a $\theta$-exact Lagrangian submanifold such that $\partial L_{0,s} \cap \partial L_1 = \emptyset$ for all $s$. Then $HF(L_{0,s}, L_1)$ is independent of $s$ up to isomorphism.*

*Sketch of the proof.* It is sufficient to show that the dimension of $HF(L_{0,s}, L_1)$ is locally constant in $s$ near some fixed $\hat{s} \in [0; 1]$. Let $L_1'$ be a Lagrangian submanifold which is $\kappa$-compatible and isotopic to $L_1$ by an exact isotopy rel $\partial M$. Let $(L_{0,s}')_{0 \leq s \leq 1}$ be an isotopy of Lagrangian submanifolds, such that each $L_{0,s}'$ is $\kappa$-compatible and isotopic to $L_{0,s}$ by an exact isotopy rel $\partial M$. We may assume that $L_{0,\hat{s}}'$ intersects $L_1'$ transversally. Take a $\mathbf{J} \in \mathcal{J}^{\mathrm{reg}}(L_{0,\hat{s}}', L_1')$. For all $s$ sufficiently close to $\hat{s}$, the intersection $L_{0,s}' \cap L_1'$ remains transverse, and $\mathbf{J}$ lies still in $\mathcal{J}^{\mathrm{reg}}(L_{0,s}', L_1')$. Hence by definition of Floer cohomology

$$HF(L_{0,s}, L_1) = HF(L_{0,s}', L_1'; \mathbf{J})$$

for all $s$ close to $\hat{s}$. By considering the corresponding parametrized moduli spaces one sees that the complex $(CF(L_{0,s}', L_1'), d_{\mathbf{J}})$ is independent of $s$ (near $\hat{s}$) up to isomorphism; compare [11, Lemma 3.4]. □

The same argument can be used to extend Proposition 5.10 to pairs $(L_0, L_1)$ with $\partial L_0 = \partial L_1$. It is difficult to see whether the isomorphisms obtained in this way are canonical or not. This is an inherent weakness of our approach; there are alternatives which do not appear to have this problem [30]. In any case, the question is irrelevant for the purposes of the present paper.

A theorem of Floer [13] says that $HF(L,L) \cong H^*(L;\mathbb{Z}/2)$ for any $\theta$-exact Lagrangian submanifold $L$ (Floer stated the theorem for $\partial L = \emptyset$ only, but his proof extends to the general case). There is no general way of computing $HF(L_0, L_1)$ when $L_0 \neq L_1$, but there are a number of ad hoc arguments which work in special situations. One such argument will be the subject of the next two sections.

5c. **Equivariant transversality.** Let $(M, \o, \alpha)$ and $\theta$ be as before. In addition we now assume that $M$ carries an involution $\iota$ which preserves $\o$, $\theta$, and $\alpha$. The fixed point set $M^\iota$ is again a symplectic manifold with contact type boundary. Moreover, if $L \subset M$ is a Lagrangian submanifold with $\iota(L) = L$, its fixed part $L^\iota = L \cap M^\iota$ is a Lagrangian submanifold of $M^\iota$. We denote by $\mathcal{J}^\iota \subset \mathcal{J}$ the subspace of those $\mathbf{J} = (J_t)_{0 \leq t \leq 1}$ such that $\iota^* J_t = J_t$ for all $t$.

Let $L_0', L_1'$ be a pair of $\theta$-exact and $\kappa$-compatible Lagrangian submanifolds of $M$, which intersect transversally and satisfy $\partial L_0' \cap \partial L_1' = \emptyset$. Assume that $\iota(L_j') = L_j'$ for $j = 0, 1$. If one wants to use this symmetry property to compute Floer cohomology, the following question arises: can one find a $\mathbf{J} \in \mathcal{J}^\iota$ which is regular? The answer is no in general, because of a simple phenomenon which we will now describe. Take a $\mathbf{J} \in \mathcal{J}^\iota$, and let $u \in \mathcal{M}(\mathbf{J})$ be a map whose image lies in $M^\iota$. Then the spaces $\mathcal{W}_u^1$ and $\mathcal{W}_u^0$ carry natural $\mathbb{Z}/2$-actions, and $D_{u,\mathbf{J}}$ is an equivariant operator. Assume that the invariant part

$$D_{u,\mathbf{J}}^{\mathbb{Z}/2} : (\mathcal{W}_u^1)^{\mathbb{Z}/2} \longrightarrow (\mathcal{W}_u^0)^{\mathbb{Z}/2}$$

is surjective, and that $\operatorname{index} D_{u,\mathbf{J}}^{\mathbb{Z}/2} > \operatorname{index} D_{u,\mathbf{J}}$. Then $D_{u,\mathbf{J}}$ can obviously not be onto. This situation is stable under perturbation, which means that for any $\mathbf{J}' \in \mathcal{J}^\iota$ sufficiently close to $\mathbf{J}$ one can find a $u' \in \mathcal{M}(\mathbf{J})$ with the same property as $u$. Hence $\mathcal{J}^\iota \cap \mathcal{J}^{\mathrm{reg}}(L_0', L_1')$ is not dense in $\mathcal{J}^\iota$. This is a general problem in equivariant transversality theory, and it is not difficult to construct concrete examples. By taking the reasoning a little further, one can find cases where $\mathcal{J}^\iota \cap \mathcal{J}^{\mathrm{reg}}(L_0', L_1') = \emptyset$. The aim of this section is to show that the maps $u \in \mathcal{M}(\mathbf{J})$ with $\mathrm{im}(u) \subset M^\iota$ are indeed the only obstruction; in other words, equivariant transversality can be achieved everywhere except at these solutions.

**Lemma 5.12.** *Let $u \in \mathcal{M}(\mathbf{J})$, for some $\mathbf{J} \in \mathcal{J}^\iota$, be a map which is not constant and such that $\mathrm{im}(u) \not\subset M^\iota$. Let $x_-, x_+ \in L_0' \cap L_1'$ be the limits of $u$. Then the subset $R(u) \subset \mathbb{R} \times (0;1)$ of those $(s,t)$ which satisfy*

$$du(s,t) \neq 0, \quad u(s,t) \notin u(\mathbb{R} \setminus \{s\}, t) \cup \iota(u(\mathbb{R}, t)) \cup \{x_\pm\} \cup \{\iota(x_\pm)\}$$

*is open and dense.*

*Proof.* We will use the methods developed in [14]. To begin with, take a more general situation: let $M$ be an arbitrary manifold and $(J_t)_{0<t<1}$ a smooth family of almost complex structures on it. We consider smooth maps $v : \Sigma \longrightarrow M$, where $\Sigma$ is some connected open subset of $\mathbb{R} \times (0;1)$, satisfying

$$\partial v/\partial s + J_t(v) \partial v/\partial t = 0. \tag{5.3}$$

Some basic properties of this equation are:

(J1) If $v$ is a nonconstant solution of (5.3), then $C(v) = \{(s,t) \in \Sigma \mid dv(s,t) = 0\}$ is a discrete subset of $\Sigma$.

(J2) Let $v_1, v_2$ be two solutions of (5.3) defined on the same set $\Sigma$, such that $v_1 \not\equiv v_2$. Then $A(v_1, v_2) = \{(s,t) \in \Sigma \mid v_1(s,t) = v_2(s,t)\}$ is a discrete subset.

(J3) Let $v$ be a nonconstant solution of (5.3). For each $x \in M$, $v^{-1}(x) \subset \Sigma$ is a discrete subset.

(J4) Let $v_1 : \Sigma_1 \longrightarrow M$ and $v_2 : \Sigma_2 \longrightarrow M$ be two solutions of (5.3) such that $v_2$ is an embedding. Assume that for any $(s,t) \in \Sigma_1$ there is an $(s',t') \in \Sigma_2$ with $t = t'$ and $v_1(s,t) = v_2(s',t')$. Then there is a $\sigma \in \mathbb{R}$ such that for all $(s,t) \in \Sigma_1$ we have $(s - \sigma, t) \in \Sigma_2$ and $v_1(s,t) = v_2(s - \sigma, t)$.

(J5) Let $v_1, v_2 : \mathbb{R} \times (0;1) \longrightarrow M$ be two nonconstant solutions of (5.3). Assume that $v_2$ is not a translate of $v_1$ in the $s$-direction. Then for any $\rho > 0$ the subset $S_\rho(v_1, v_2) = \{(s,t) \in \mathbb{R} \times (0;1) \mid v_1(s,t) \notin v_2([-\rho;\rho] \times \{t\})\}$ is open and dense in $\mathbb{R} \times (0;1)$.

(J1) is [14, Corollary 2.3(ii)]. (J2) is a form of the unique continuation theorem [14, Proposition 3.1], and (J3) follows from it by letting $v_2 \equiv x$ be a constant map. (J4) is a variant of [14, Lemma 4.2], and is proved as follows. $\psi = v_2^{-1} v_1 : \Sigma_1 \longrightarrow \Sigma_2$ is a holomorphic map of the form $\psi(s,t) = (\psi_1(s,t), t)$. By looking at the derivative of $\psi$ one sees that it must be a translation in the $s$-direction.

The proof of the final property (J5) is similar to that of [14, Theorem 4.3]; however, the differences are sufficiently important to deserve a detailed discussion. It is clear that $S_\rho(v_1, v_2)$ is open for any $\rho > 0$. Now assume that $(S,T) \in \mathbb{R} \times (0;1)$ is a point which has a neighbourhood disjoint from $S_\rho(v_1, v_2)$. After possibly moving $(S,T)$ slightly, one can assume that $v_1(S,T) \notin v_2(C(v_2))$. This can be achieved because $C(v_2)$ is a countable set by (J1), and hence $v_1^{-1}(v_2(C(v_2)))$ is countable by (J3). Again because of (J3), there are only finitely many $S'_1, \dots, S'_k \in [-\rho;\rho]$ such that $v_1(S,T) = v_2(S'_j, T)$. By the choice of $(S,T)$, we have $dv_2(S'_j, T) \neq 0$ for all these $S'_j$. Choose an $\epsilon > 0$ such that the restriction of $v_2$ to the open disc $D_\epsilon(S'_j, T)$ of radius $\epsilon$ around $(S'_j, T)$ is an embedding for any $j$.

**Claim A.** *There is an $\epsilon' > 0$ with the following property: for each $(s,t) \in D_{\epsilon'}(S,T)$ there is an $(s',t') \in \bigcup_{j=1}^k D_{\epsilon/2}(S'_j, T)$ with $t' = t$ and $v_1(s,t) = v_2(s',t)$.*

To prove this, take a sequence of points $(s_i, t_i)_{i \in \mathbb{N}}$ with limit $(S,T)$. For sufficiently large $i$ we have $(s_i, t_i) \notin S_\rho(v_1, v_2)$. Hence there are $s'_i \in [-\rho;\rho]$, $s'_i \neq s_i$, such that $v_2(s'_i, t_i) = v_1(s_i, t_i)$. If the claim is false, one can choose the $(s_i, t_i)$ and $s'_i$ in such a way that $(s'_i, t_i) \notin \bigcup_j D_{\epsilon/2}(S'_j, T)$. Then, after passing to a subsequence, one obtains a limit $s' \in [-\rho;\rho]$ with $v_2(s', T) = v_1(S,T)$ and $s' \neq S'_1, \dots, S'_k$, which is a contradiction.

**Claim B.** *There is a nonempty connected open subset $\Sigma_1 \subset D_{\epsilon'}(S,T)$ and a $j \in \{1, \dots, k\}$ with the following property: for any $(s,t) \in \Sigma_1$ there is an $(s',t') \in D_\epsilon(S'_j, T)$ with $t' = t$ and $v_1(s,t) = v_2(s',t)$.*

Let $U_j \subset D_{\epsilon'}(S,T)$ be the closed subset of those $(s,t)$ such that $v_1(s,t) = v_2(s',t)$ for some $(s',t) \in \overline{D_{\epsilon/2}(S'_j, T)}$. Claim A says that $U_1 \cup \dots \cup U_k = D_{\epsilon'}(S,T)$; hence at least one of the $U_j$ must have nonempty interior. Define $\Sigma_1$ to be some small open disc inside that $U_j$; this proves Claim B.

Now consider the restrictions $v_1|\Sigma_1$ and $v_2|\Sigma_2$, where $\Sigma_2 = D_\epsilon(S'_j, T)$. By definition $v_2|\Sigma_2$ is an embedding; applying (J4) shows that $v_1|\Sigma_1$ is a translate of $v_2|\Sigma_2$ in the $s$-direction. Because of unique continuation (J2) it follows that $v_1$ must be a translate of $v_2$, which completes the proof of (J5).

After these preliminary considerations, we can now turn to the actual proof of Lemma 5.12. Set $v = \iota \circ u$. $v$ cannot be a translate of $u$ in the $s$-direction. To see this, assume that on the contrary $v(s,t) = u(s-\rho,t)$ for some $\rho \in \mathbb{R}$. The case $\rho = 0$ is excluded by the assumptions, since it would imply that $\mathrm{im}(u) \subset M^\iota$. If $\rho \neq 0$, then $u(s,t) = u(s-2\rho,t)$ which, because of the finiteness of the energy, means that $u$ is constant; this is again excluded by the assumptions.

Clearly $R(u) = R_1(u) \cap R_2(u) \cap R_3(u)$, where

$$\begin{aligned} R_1(u) &= \{(s,t) \in \mathbb{R} \times (0;1) \mid du(s,t) \neq 0,\ u(s,t) \notin u(\mathbb{R} \setminus \{s\}, t),\ u(s,t) \neq x_\pm\},\\ R_2(u) &= \{(s,t) \in \mathbb{R} \times (0;1) \mid u(s,t) \neq \iota(x_\pm)\},\\ R_3(u) &= \{(s,t) \in \mathbb{R} \times (0;1) \mid u(s,t) \notin \iota(u(\mathbb{R},t))\}. \end{aligned}$$

$R_1(u)$ is the set of regular points of $u$ as defined in [14], and it is open and dense by a slight variation of [14, Theorem 4.3]; see also [29]. $R_2(u)$ is open and dense by (J3) above. $R_3(u)$ is the intersection of countably many sets $S_\rho(u,v)$, each of which is open and dense by (J5). Baire's theorem now shows that $R(u)$ is dense. As for its openness, it can be proved by an elementary argument, as in [14, Theorem 4.3]. □

**Proposition 5.13.** *There is a* $\mathbf{J} \in \mathcal{J}^\iota$ *such that every* $u \in \mathcal{M}(\mathbf{J})$ *whose image is not contained in* $M^\iota$ *is regular. In fact, the subspace of such* $\mathbf{J}$ *is* $C^\infty$*-dense in* $\mathcal{J}^\iota$*.*

*Sketch of the proof.* It is convenient to recall first the proof of the basic transversality result in Floer theory, Proposition 5.8. Let $\mathcal{T}_\mathbf{J}$ be the tangent space of $\mathcal{J}$ at some point $\mathbf{J}$. This space consists of smooth families $\mathbf{Y} = (Y_t)_{0 \leq t \leq 1}$ of sections $Y_t \in C^\infty(\mathrm{End}(TM))$ satisfying some additional properties, which we will not write down here. For $\mathbf{J} \in \mathcal{J}$ and $u \in \mathcal{M}(\mathbf{J})$ one defines an operator

$$\widetilde{D}_{u,\mathbf{J}} : \mathcal{W}^1_u \times \mathcal{T}_\mathbf{J} \longrightarrow \mathcal{W}^0_u, \quad \widetilde{D}_{u,\mathbf{J}}(X, \mathbf{Y}) = D_{u,\mathbf{J}}(X) + Y_t(u)\partial u/\partial t.$$

The main point in the proof of Proposition 5.8 is to show that $\widetilde{D}_{u,\mathbf{J}}$ is always onto. Assume that the contrary holds: then, because the image is closed, there is a nonzero $Z \in L^q(u^*TM)$, with $1/p + 1/q = 1$, which satisfies a linear $\bar\partial$-type equation, hence is smooth on $\mathbb{R} \times (0;1)$, and such that

$$\int_{\mathbb{R} \times [0;1]} ø(Y_t(u)\partial u/\partial t, J_t(u)Z)\, ds\, dt = 0 \quad \text{for all } \mathbf{Y} \in \mathcal{T}_\mathbf{J}. \tag{5.4}$$

Let $x_\pm$ be the limits of $u$. Excluding the case of constant maps $u$, for which $D_{u,\mathbf{J}}$ is already onto, one can find an $(s,t) \in \mathbb{R} \times (0;1)$ with $du(s,t) \neq 0$, $u(s,t) \neq x_\pm$, and $u(s,t) \notin u(\mathbb{R} \setminus \{s\}, t)$. In order to avoid problems with the behaviour of $\mathbf{Y}$ near the boundary, it is useful to require that $u(s,t)$ lies outside a fixed small neighbourhood of $\partial M$. Then, assuming that $Z(s,t) \neq 0$, one can construct a $\mathbf{Y} \in \mathcal{T}_\mathbf{J}$ which is concentrated on a small neighbourhood of $u(s,t)$ and which contradicts (5.4). Since the set of points $(s,t)$ with these properties is open, it follows that $Z$ vanishes on some open subset. By the unique continuation for solutions of linear $\bar\partial$-equations, $Z \equiv 0$, which is a contradiction.

The proof of Proposition 5.13 differs from this in that one considers only equivariant $\mathbf{J}$. This means that the domain of $\widetilde{D}_{u,\mathbf{J}}$ must be restricted to those $\mathbf{Y}$ which are equivariant. One considers a supposed $Z$ as before, and takes a point $(s,t) \in R(u)$ as in Lemma 5.12. The same local construction as before yields a nonequivariant $\mathbf{Y} \in \mathcal{T}_\mathbf{J}$ which is concentrated near $u(s,t)$. One then makes this equivariant by taking $Y_t + \iota^* Y_t$; the properties of $R(u)$ ensure that this averaged element of the tangent space still leads to a contradiction with (5.4). The details are straightforward, and we leave them to the reader. $\square$

5d. **A symmetry argument.** The notations $(M, \phi, \alpha)$, $\theta$, $\kappa$ and $\iota$ are as in the previous section. We begin with the most straightforward application of equivariant transversality.

**Lemma 5.14.** *Let $L_0, L_1 \subset M$ be two $\theta$-exact Lagrangian submanifolds with $\partial L_0 \cap \partial L_1 = \emptyset$ and $\iota(L_j) = L_j$. Assume that the following properties hold:*

(S1) *the intersection $L_0 \cap L_1$ is transverse, and each intersection point lies in $M^\iota$;*
(S2) *there is no continuous map $v : [0;1]^2 \longrightarrow M^\iota$ with the following properties: $v(0,t) = x_-$, $v(1,t) = x_+$ for all $t$, where $x_- \neq x_+$ are points of $L_0 \cap L_1$, and $v(s,0) \in L_0$, $v(s,1) \in L_1$ for all $s$.*

*Then* $\dim_{\mathbb{Z}/2} HF(L_0, L_1) = |L_0 \cap L_1|$.

*Proof.* An inspection of the proof of Lemma 5.2(a) shows that one can find $\kappa$-compatible Lagrangian submanifolds $L_j'$, $j = 0, 1$, which are isotopic to $L_j$ by an exact *and $\iota$-equivariant* Lagrangian isotopy rel $\partial M$. Moreover, this can be done in such a way that $L_j'$ is equal to $L_j$ outside a small neighbourhood of $\partial M$, and such that $L_0' \cap L_1' = L_0 \cap L_1$. Then the conditions (S1) and (S2) continue to hold with $(L_0, L_1)$ replaced by $(L_0', L_1')$. Choose a $\mathbf{J} \in \mathcal{J}^\iota$ which has the property stated in Proposition 5.13. We will show that $\mathbf{J} \in \mathcal{J}^{\mathrm{reg}}(L_0', L_1')$. Assume that on the contrary there is a $u \in \mathcal{M}(\mathbf{J})$ which is not regular. Note that $u$ cannot be constant, since the constant maps are regular for any choice of $\mathbf{J}$. By Proposition 5.13 $\mathrm{im}(u) \subset M^\iota$. Then condition (S2) says that the two endpoints $\lim_{s \to \pm\infty} u(s, \cdot)$ must necessarily agree. Because of the gradient flow interpretation of Floer's equation (recall that the action functional is not multivalued in the situation which we are considering), $u$ must be constant, which is a contradiction. This shows that $\mathbf{J}$ is indeed regular.

Composition with $\iota$ defines an involution on $\mathcal{M}(\mathbf{J})$. It is not difficult to see that the indices of $D_{u,\mathbf{J}}$ and $D_{\iota \circ u, \mathbf{J}}$ are the same; indeed, there are obvious isomorphisms between the kernels, resp. cokernels, of these two operators. Moreover, $u$ and $\iota \circ u$ have the same endpoints because $L_0' \cap L_1' \subset M^\iota$. Hence the involution preserves each subset $\mathcal{M}_k(x_-, x_+; \mathbf{J}) \subset \mathcal{M}(\mathbf{J})$. Consider the induced involution $\bar\iota$ on the quotients $\mathcal{M}_1(x_-, x_+; \mathbf{J})/\mathbb{R}$. A fixed point of $\bar\iota$ would be a map $u \in \mathcal{M}(\mathbf{J})$ such that $u(s,t) = \iota(u(s - \sigma, t))$ for some $\sigma \in \mathbb{R}$. By the same argument as in the proof of Lemma 5.12, the only maps with this property are the constant ones, which do not lie in $\mathcal{M}_1(x_-, x_+)$ since the associated operators have index zero. Hence $\bar\iota$ is free for all $x_-, x_+$.

By definition, $HF(L_0, L_1)$ is the homology of $(CF(L_0', L_1'), d_{\mathbf{J}})$. Since each set $\mathcal{M}_1(x_-, x_+; \mathbf{J})/\mathbb{R}$ admits a free involution, $\nu(x_-, x_+; \mathbf{J})$ and hence $d_{\mathbf{J}}$ are zero, so that $\dim HF(L_0, L_1) = \dim CF(L_0', L_1') = |L_0' \cap L_1'| = |L_0 \cap L_1|$. $\square$

**Proposition 5.15.** *Let $L_0, L_1 \subset M$ be two $\theta$-exact Lagrangian submanifolds with $\partial L_0 \cap \partial L_1 = \emptyset$ and $\iota(L_j) = L_j$. Assume that the following properties hold:*

(S1′) *the intersection $N = L_0 \cap L_1$ is clean, and there is an $\iota$-invariant Morse function $h$ on $N$ whose critical points are precisely the points of $N^\iota = N \cap M^\iota$;*
(S2′) *same as (S2) above.*

*Then* $\dim_{\mathbb{Z}/2} HF(L_0, L_1) = |N^\iota| = \dim_{\mathbb{Z}/2} H^*(N; \mathbb{Z}/2)$.

*Proof.* Clean intersection means that $N$ is a smooth manifold and $TN = (TL_0|N) \cap (TL_1|N)$. We will now describe a local model for clean intersections, due to Weinstein [41, Theorem 4.3]. There are a neighbourhood $U \subset L_0$ of $N$, a function $f \in C^\infty(U, \mathbb{R})$ with $df^{-1}(0) = N$ and which is nondegenerate in the sense of Bott, a neighbourhood $V \subset T^*U$ of the zero-section, and a symplectic embedding

$$\psi : V \longrightarrow M \setminus \partial M$$

with $\psi|U = \mathrm{id}$ and $\psi^{-1}(L_1) = \Gamma_{df} \cap V$. Here $T^*U$ carries the standard symplectic structure, $U$ is considered to be embedded in $V$ as the zero-section, and $\Gamma_{df} \subset T^*U$ is the graph of $df$. Weinstein's construction can easily be adapted to take into account the presence of a finite symmetry group. In our situation this means that one can take $U, f$ to be invariant under $\iota|L_0$, $V$ to be invariant under the induced action of $\mathbb{Z}/2$ on $T^*U$, and $\psi$ to be an equivariant embedding.

Let $h$ be a function as in (S1$'$). One can find an $\iota$-invariant function $g \in C^\infty_c(U, \mathbb{R})$ such that $g|N = h$. An elementary argument shows that for all sufficiently small $t > 0$, $f + t\,g$ is a Morse function whose critical point set is $N \cap dg^{-1}(0) = dh^{-1}(0) = N^\iota$. Define a Lagrangian submanifold $L_1' \subset M$ by setting $L_1' \cap (M \setminus \operatorname{im} \psi) = L_1$ and $\psi^{-1}(L_1') = \Gamma_{df + t\,dg} \cap V$ for some small $t > 0$. $L_1'$ is obviously $\iota$-invariant, and can be deformed back to $L_1$ by an exact Lagrangian isotopy rel $\partial M$. Moreover, $L_0 \cap L_1'$ consists of the critical points of $f + tg$, hence is equal to $N^\iota$.

The $\iota$-invariant parts of $L_0, L_1$ are Lagrangian submanifolds of $M^\iota$ which intersect cleanly in $N^\iota$. By assumption (S1$'$) $N^\iota$ is a finite set, so that $L_0^\iota$ and $L_1^\iota$ actually intersect transversally. Since $(L_1')^\iota$ is a $C^1$-small perturbation of $(L_1)^\iota$, there is a homeomorphism of $M^\iota$ which sends $L_0^\iota$ to itself and carries $L_1^\iota$ to $(L_1')^\iota$. As a consequence of this, (S2$'$) continues to hold with $(L_0, L_1)$ replaced by $(L_0, L_1')$. We have now proved that the pair $(L_0, L_1')$ satisfies all the assumptions of Lemma 5.14. Using that lemma and Proposition 5.10 one obtains

$$\dim HF(L_0, L_1) = \dim HF(L_0, L_1') = |L_0 \cap L_1'| = |N^\iota|.$$

It remains to explain why $|N^\iota|$ is equal to the dimension of $H^*(N; \mathbb{Z}/2)$. This is in fact a consequence of the finite-dimensional (Morse theory) analogue of Lemma 5.14. Consider the gradient flow of $h$ with respect to some $\mathbb{Z}/2$-invariant metric on $N$. Since $N^\iota$ is a finite set, there are no $\mathbb{Z}/2$-invariant gradient flow lines. Therefore one can perturb the metric equivariantly so that the moduli spaces of gradient flow lines become regular. Then $\iota$ induces a free $\mathbb{Z}/2$-action on these moduli spaces, which implies that the differential in the Morse cohomology complex (with coefficients in $\mathbb{Z}/2$) for $h$ is zero. Hence $\dim H^*(N; \mathbb{Z}/2) = |dh^{-1}(0)| = |N^\iota|$. □

5e. **The grading on Floer cohomology.** The basic idea in this section is due to Kontsevich [24]; for a detailed exposition see [36]. Let $(M, \o, \alpha, \theta)$ be as in Section 5b, with $\dim M = 2n$; in addition assume that $2c_1(M, \o)$ is zero.

Let $\mathcal{L} \to M$ be the natural fibre bundle whose fibres $\mathcal{L}_x$ are the Lagrangian Grassmannians (the manifolds of all linear Lagrangian subspaces) of $T_xM$. Recall that $\pi_1(\mathcal{L}_x) \cong \mathbb{Z}$ and that $H^1(\mathcal{L}_x; \mathbb{Z})$ has a canonical generator, the Maslov class $C_x$. Because of our assumption on the first Chern class, there is a cohomology class on $\mathcal{L}$ which restricts to $C_x$ for any $x$. Correspondingly there is an infinite cyclic covering $\widetilde{\mathcal{L}} \longrightarrow \mathcal{L}$ such that each restriction $\widetilde{\mathcal{L}}_x \longrightarrow \mathcal{L}_x$ is the universal covering. Fix one such covering (there may be many nonisomorphic ones in general, so that there is some nontrivial choice to be made here) and denote the $\mathbb{Z}$-action on it by $\chi_{\widetilde{\mathcal{L}}}$. For any Lagrangian submanifold $L$ in $(M, \o, \alpha)$ there is a canonical section

$s_L : L \longrightarrow \mathcal{L}$ given by $s_L(x) = T_xL$. A *grading of* $L$ is a lift $\widetilde{L}$ of $s_L$ to $\widetilde{\mathcal{L}}$. Clearly, a grading exists iff $s_L^*\widetilde{\mathcal{L}}$ is a trivial covering of $L$, so that the obstruction lies in $H^1(L)$. Pairs $(L, \widetilde{L})$ are called *graded Lagrangian submanifolds*; we will usually write $\widetilde{L}$ instead of $(L, \widetilde{L})$. $\chi_{\widetilde{\mathcal{L}}}$ defines a $\mathbb{Z}$-action on the set of graded Lagrangian submanifolds. There is also a notion of isotopy, whose definition is obvious.

An equivalent formulation of the theory goes as follows. Let $J$ be an ø-compatible almost complex structure on $M$. By assumption the bicanonical bundle $K^{\otimes 2}$ is trivial; any nowhere zero section $\Theta$ of it determines a map

$$\delta_{\mathcal{L}} : \mathcal{L} \longrightarrow \mathbb{C}^*/\mathbb{R}^{>0},$$

defined by $\delta_{\mathcal{L}}(\mathbb{R}e_1 \oplus \cdots \oplus \mathbb{R}e_n) = \Theta((e_1 \wedge \cdots \wedge e_n)^{\otimes 2})$ for any family of $n$ orthonormal vectors $e_1, \dots, e_n \in T_xM$ spanning a Lagrangian subspace. The pullback of the universal cover $\mathbb{R} \longrightarrow \mathbb{C}^*/\mathbb{R}^{>0}$, $\xi \mapsto \exp(2\pi i\xi)$ by $\delta_{\mathcal{L}}$ is a covering $\widetilde{\mathcal{L}}$ of the kind considered above, and one can show that all such coverings (up to isomorphism) can be obtained in this way. From this point of view, a grading of a Lagrangian submanifold $L$ is just a map $\tilde{L} : L \longrightarrow \mathbb{R}$ which is a lift of $L \longrightarrow \mathbb{C}^*/\mathbb{R}^{>0}$, $x \mapsto \delta_{\mathcal{L}}(T_xL)$; and the $\mathbb{Z}$-action on the set of gradings is by adding constant functions.

**Example 5.16.** Suppose that $M$ is the intersection of a complex hypersurface $g^{-1}(0) \subset \mathbb{C}^{n+1}$ with some ball in $\mathbb{C}^{n+1}$, and that ø is the restriction of the standard symplectic form. In this case one can take $J$ to be the standard complex structure, and $\Theta$ to be the square of the complex $n$-form $\det_{\mathbb{C}}(\overline{dg_x} \wedge \dots)|TM$. Then, if $e_1, \dots, e_n$ is an orthonormal basis of $T_xL$, one has

$$\delta_{\mathcal{L}}(T_xL) = \det_{\mathbb{C}}(\overline{dg_x} \wedge e_1 \wedge \cdots \wedge e_n)^2. \tag{5.5}$$

Let $(L_0, \widetilde{L}_0)$ and $(L_1, \widetilde{L}_1)$ be a pair of graded Lagrangian submanifolds. To any point $x \in L_0 \cap L_1$ one can associate an absolute Maslov index $\mu^{\mathrm{abs}}(\widetilde{L}_0, \widetilde{L}_1; x) \in \mathbb{Z}$. To define it, take a path $\tilde{\lambda}_0 : [0; 1] \longrightarrow \widetilde{\mathcal{L}}_x$ with endpoints $\tilde{\lambda}_0(j) = \widetilde{L}_j(x)$, $j = 0, 1$. The projection of $\tilde{\lambda}$ to $\mathcal{L}$ is a path $\lambda_0 : [0; 1] \longrightarrow \mathcal{L}_x$ with $\lambda_0(j) = T_xL_j$. Let $\lambda_1 : [0; 1] \longrightarrow \mathcal{L}_x$ be the constant path $\lambda_1(t) = T_xL_1$, and set

$$\mu^{\mathrm{abs}}(\widetilde{L}_0, \widetilde{L}_1; x) = \mu^{\mathrm{paths}}(\lambda_0, \lambda_1) + \tfrac{1}{2}(n - \dim(T_xL_0 \cap T_xL_1));$$

here $\mu^{\mathrm{paths}}$ is the "Maslov index for paths" defined in [32]. The absolute Maslov index is well defined, essentially because $\lambda_0$ is unique up to homotopy rel endpoints. It depends on the gradings through the formula

$$\mu^{\mathrm{abs}}(\chi_{\widetilde{\mathcal{L}}}(r_0)(\widetilde{L}_0), \chi_{\widetilde{\mathcal{L}}}(r_1)(\widetilde{L}_1); x) = \mu^{\mathrm{abs}}(\widetilde{L}_0, \widetilde{L}_1; x) + r_1 - r_0. \tag{5.6}$$

For later use, we record a situation in which the index is particularly easy to compute. For any $t \in [0; 1]$ one can define a quadratic form $B_t$ on $\lambda_0(t) \cap \lambda_1(t)$ by setting $B_t(v) = (d/d\tau)_{\tau=t}ø(v, w(\tau))$, where $w$ is a smooth path in $T_xM$ (defined for $\tau$ near $t$) with $w(t) = v$, $w(\tau) \in \lambda_0(\tau)$ for all $\tau$. In particular, taking $t = 1$ yields a quadratic form $B_1$ defined on $T_xL_1$.

**Lemma 5.17.** *Assume that $\lambda_0(t) \cap \lambda_1(t) = \Lambda$ is the same for all $0 \leq t < 1$, and that the quadratic form $B_1$ is nonpositive with nullspace precisely equal to $\Lambda$. Then $\mu^{\mathrm{abs}}(\widetilde{L}_0, \widetilde{L}_1; x) = 0$.*

We omit the proof, which is an immediate consequence of the basic properties of the Maslov index for paths.

Now let $L_0, L_1$ be two $\theta$-exact Lagrangian submanifolds with $\partial L_0 \cap \partial L_1 = \emptyset$, and $\widetilde{L}_0, \widetilde{L}_1$ gradings of them. To define $HF(L_0, L_1)$ one has to deform the given $L_j$ to $\kappa$-compatible Lagrangian submanifolds $L_j'$ as in Lemma 5.6. The $L_j'$ inherit preferred gradings $\widetilde{L}_j'$ from those of $L_j$ (this is done by lifting the isotopy from $L_j$ to $L_j'$ to an isotopy of graded Lagrangian submanifolds). Introduce a $\mathbb{Z}$-grading on $CF(L_0', L_1')$ by setting

$$CF^r(L_0', L_1') = \bigoplus_{\mu^{\mathrm{abs}}(\widetilde{L}_0', \widetilde{L}_1'; x) = r} \mathbb{Z}/2\langle x \rangle.$$

Then $d_{\mathbf{J}}$, for any $\mathbf{J} \in \mathcal{J}^{\mathrm{reg}}(L_0', L_1')$, has degree one, so that one gets a $\mathbb{Z}$-grading on $HF(L_0', L_1'; \mathbf{J})$. One can show that the continuation isomorphisms between these groups for various choices of $L_0', L_1'$ and $\mathbf{J}$ have degree zero. Hence there is a well-defined graded Floer cohomology group $HF^*(\widetilde{L}_0, \widetilde{L}_1)$, which satisfies

$$HF^*(\chi_{\widetilde{\mathcal{L}}}(r_0)\widetilde{L}_0, \chi_{\widetilde{\mathcal{L}}}(r_1)\widetilde{L}_1) = HF^{*+r_0-r_1}(\widetilde{L}_0, \widetilde{L}_1).$$

In view of the definition of Floer cohomology, the whole discussion carries over to pairs of Lagrangian submanifolds with $\partial L_0 = \partial L_1$.

We will now explain the relation between this formalism and the traditional *relative grading* on Floer cohomology, which we have used in the statement of Corollary 1.5. The relative grading is defined under the assumption that a certain class $m_{L_0,L_1} \in H^1(\mathcal{P}(L_0, L_1); \mathbb{Z})$ vanishes; then $HF(L_0, L_1)$ splits as a direct sum of groups corresponding to the different connected components of $\mathcal{P}(L_0, L_1)$, and for each of these summands there is a $\mathbb{Z}$-grading which is unique up to a constant shift; see e.g. [40], [12]. If the manifolds $L_0, L_1$ admit gradings, for some covering $\widetilde{\mathcal{L}}$, then $m_{L_0,L_1}$ must necessarily be zero, and the absolute grading $HF^*(\widetilde{L}_0, \widetilde{L}_1)$ with respect to any $\widetilde{L}_0, \widetilde{L}_1$ is also one of the possible choices of relative grading. The converse statement is false: there are more choices of relative gradings than those which occur as $HF^*(\widetilde{L}_0, \widetilde{L}_1)$.

**Proposition 5.18.** *Let $L_0, L_1 \subset M$ be two Lagrangian submanifolds as in Proposition 5.15, and assume that they have gradings $\widetilde{L}_0, \widetilde{L}_1$ with respect to some covering $\widetilde{\mathcal{L}}$. Then there is a graded isomorphism*

$$HF^*(\widetilde{L}_0, \widetilde{L}_1) \cong \bigoplus_{k=1}^{d} H^{*-\mu_k}(N_k; \mathbb{Z}/2).$$

*Here $N_1, \dots, N_d$ are the connected components of $N = L_0 \cap L_1$, and $\mu_k$ is the index $\mu^{\mathrm{abs}}(\widetilde{L}_0, \widetilde{L}_1; x)$ at some point $x \in N_k$ (this is independent of the choice of $x$ because the intersection $L_0 \cap L_1$ is clean).*

The proof is essentially the same as that of Proposition 5.15. There is just one additional computation, which describes the change in the Maslov index when $L_0, L_1$ are perturbed to make the intersection transverse. Since this is a well-known issue, which has been addressed e.g. in [31], we will not discuss it here.

## 6. The symplectic geometry of the $(A_m)$-singularities

We will first define the general notion of symplectic monodromy map for deformations of isolated hypersurface singularities. After that we consider the Milnor fibres of the singularities of type $(A_m)$ in more detail. Concretely, following a suggestion of Donaldson, we will set up a map

$$\begin{pmatrix}\text{curves on the disc with}\\ (m+1)\text{ marked points}\end{pmatrix} \longmapsto \begin{pmatrix}\text{Lagrangian submani-}\\ \text{folds of the Milnor fibre}\end{pmatrix}.$$

Here "curve" is taken in the sense of Section 3a. At least for those curves which meet at least one marked point, the Floer cohomology of the associated Lagrangian submanifolds is well defined and recovers the geometric intersection number of the curves. This quickly leads to a proof of our main symplectic results, Theorem 1.3 and Corollary 1.4. We will also carry out a refined version of the same argument, which involves Floer cohomology as a graded group. A comparison between this and the results of Section 4c yields Corollary 1.5.

6a. **More basic symplectic geometry.** Throughout this section, $(M, \o, \alpha)$ is a fixed symplectic manifold with contact type boundary, such that the relative symplectic class $[\o, \alpha]$ is zero.

Let $W$ be a connected manifold, and let $\widehat{W} \subset W$ be a nonempty submanifold. A *symplectic fibration* over $(W, \widehat{W})$ consists of the following data:

(F1) A proper smooth fibration $\pi : E \longrightarrow W$ whose fibres $E_w = \pi^{-1}(w)$ are manifolds with boundary. If $W$ also has a boundary, $E$ is a manifold with corners; in that case we adopt the convention that $\partial E$ refers just to the boundary in the fibre direction, which is the union of the boundaries $\partial E_w$ of the fibres.

(F2) The structure of a symplectic manifold with contact type boundary on each fibre. More precisely, we want differential forms $\o_w \in \Omega^2(E_w)$ and $\alpha_w \in \Omega^1(\partial E_w)$, depending smoothly on $w \in W$ and satisfying the obvious conditions. We also require that $[\o_w, \alpha_w] \in H^2(E_w, \partial E_w; \mathbb{R})$ is zero for all $w$.

(F3) A trivialization of the family of contact manifolds $(\partial E_w, \alpha_w)_{w \in W}$. This is a fibrewise diffeomorphism $\tau : \partial M \times W \longrightarrow \partial E$, such that $\tau_w : (\partial M, \alpha) \longrightarrow (\partial E_w, \alpha_w)$ is a contact diffeomorphism for every $w \in W$.

(F4) A trivialization of $E|\widehat{W}$, compatible with all the given structure. This is a fibrewise diffeomorphism $\eta : M \times \widehat{W} \longrightarrow E|\widehat{W}$ such that $\eta_w^* \o_w = \o$ and $(\eta_w|\partial M)^* \alpha_w = \alpha$ for each $w \in \widehat{W}$, and which agrees with $\tau$ on $\partial M \times \widehat{W}$.

Two symplectic fibrations over $(W, \widehat{W})$ are *cobordant* if they can be joined by a symplectic fibration over $(W \times [0;1], \widehat{W} \times [0;1])$. A symplectic fibration is called *strict* if $\tau_w^* \alpha_w = \alpha$ for all $w$ (for a general symplectic fibration, this holds only as an equality of contact structures).

What we have defined is actually a family of symplectic manifolds with contact type boundary, equipped with certain additional data. Still, we prefer to use the name symplectic fibration for the sake of brevity; for the same reason we usually write $E$ rather than $(E, \pi, (\o_w), (\alpha_w), \tau, \eta)$.

**Example 6.1.** Let $\phi$ be a map in $\mathrm{Symp}(M, \partial M, \o)$. Consider the projection $\pi : M \times [0;1] \longrightarrow [0;1]$. Define $\eta_0 : M \longrightarrow \pi^{-1}(0)$ to be the identity, and $\eta_1 : M \longrightarrow \pi^{-1}(1)$ to be $\phi^{-1}$. This, with the remaining data chosen in the obvious way, defines a strict symplectic fibration over $([0;1], \{0;1\})$, which we denote by $E_\phi$.

**Proposition 6.2.** *One can associate to any symplectic fibration $E$ over $(W, \widehat{W})$ a map $\rho_s^E : \pi_1(W, \widehat{W}) \longrightarrow \pi_0(\mathrm{Symp}(M, \partial M, ø))$, such that*

(A1) *$\rho_s^E$ is natural with respect to the pullback of $E$ by smooth maps on the base;*
(A2) *cobordant symplectic fibrations have the same maps $\rho_s^E$;*
(A3) *if $E = E_\phi$ is as in Example 6.1 and $[\beta] \in \pi_1(W, \widehat{W})$ is the class of the identity map $\beta = \mathrm{id}_{[0;1]}$, then $\rho_s^E([\beta]) = [\phi]$.*

*Conversely, these properties determine the assignment $E \mapsto \rho_s^E$ uniquely.*

The nontrivial issue here is that different fibres of a symplectic fibration may not be symplectically isomorphic; not even their volumes need to be the same. We will therefore define $\rho_s^E$ first for strict symplectic fibrations, where the problem does not occur, and then extend the definition using cobordisms. An alternative approach is outlined in Remark 6.5 below.

Take a symplectic fibration $E$ over $(W, \widehat{W})$, and let $\psi \in C^\infty(\partial M \times W, \mathbb{R})$ be the function defined by $\tau_w^* \alpha_w = e^{\psi_w} \alpha$; by assumption, it vanishes on $\partial M \times \widehat{W}$. Choose $\theta \in \Omega^1(M)$ with $\theta | \partial M = \alpha$, $d\theta = ø$, and let $\kappa : \mathbb{R}^{\leq 0} \times \partial M \to M$ be the embedding determined by the corresponding Liouville vector field. Similarly choose a smooth family $\theta_w \in \Omega^1(E_w)$, $w \in W$, with $\theta_w | \partial E_w = \alpha_w$ and $d\theta_w = ø_w$ for all $w$, and such that $\eta_w^* \theta_w = \theta$ whenever $w \in \widehat{W}$. This determines an embedding of $\mathbb{R}^{\leq 0} \times \partial E$ into $E$ which fibers over $W$. By combining this with $\tau$, one obtains an embedding $\gamma : \mathbb{R}^{\leq 0} \times \partial M \times W \to E$ which satisfies $\gamma_w^* ø = d(e^{r+\psi_w}\alpha)$, and such that $\eta_w^{-1} \circ \gamma_w = \kappa$ for $w \in \widehat{W}$.

**Lemma 6.3.** *Any strict symplectic fibration $E$ over $(W, \widehat{W})$ has the structure of a fibre bundle with structure group $\mathrm{Symp}(M, \partial M, ø)$, compatible with $\eta$.*

*Proof.* More precisely, we will show that for any point of $W$ there is a neighbourhood $U$ and a local trivialization $\eta_U : M \times U \to E|U$ which agrees with $\eta$ on $M \times (U \cap \widehat{W})$, such that

$$\eta_U(\kappa(r, x), w) = \gamma(r, x, w)$$

for all $(r, x, w)$ in a neighbourhood of $\{0\} \times \partial M \times U \subset \mathbb{R}^{\leq 0} \times \partial M \times U$. The transition maps between two local trivializations with this property obviously take values in $\mathrm{Symp}(M, \partial M, ø)$, providing the desired structure.

The first step is to find a two-form $\Omega \in \Omega^2(E)$ with the following properties:

($\Omega$1) $\Omega | E_w = ø_w$,
($\Omega$2) $d\Omega = 0$,
($\Omega$3) $\gamma^* \Omega \in \Omega^2(\mathbb{R}^{\leq 0} \times \partial M \times W)$ agrees with the pullback of $d(e^r \alpha) \in \Omega^2(\mathbb{R}^{\leq 0} \times \partial M)$ on some neighbourhood of $\{0\} \times \partial M \times W$,
($\Omega$4) $\eta^* \Omega \in \Omega^2(M \times \widehat{W})$ is the pullback of $\omega$.

This is not difficult. One takes a one-form $\Theta$ on $E$ with $\Theta | E_w = \theta_w$, such that $\gamma^* \Theta$ is equal to the pullback of $e^r \alpha$ near $\{0\} \times \partial M \times W$, and with $\eta^* \Theta$ the pullback of $\theta$ to $M \times \widehat{W}$. Such a one-form exists because the fibration is strict. Then $\Omega = d\Theta$ satisfies the conditions above.

Define, for any $x \in E$, a subspace $H_x = \{X \in T_x E \mid Ø(X, Y) = 0$ for all $Y \in T_x E$ with $D\pi(Y) = 0\}$. The subbundle $H \subset TE$ formed by the $H_x$ is complementary to $\ker D\pi \subset TE$ by ($\Omega$1) and is tangent to $\partial E$ by ($\Omega$3); hence it defines a connection on $\pi : E \longrightarrow W$. The parallel transport maps of this connection are symplectic by ($\Omega 2$) (the reader can find a more extensive discussion of symplectic connections in [26, Chapter 6]). Moreover, with respect to the map $\gamma$, they are trivial near $\partial E$ by ($\Omega 3$). Using those maps, one easily obtains the desired local trivializations. $\square$

As a special case of this construction, one sees that any strict symplectic fibration over $([0;1],\{0;1\})$ is isomorphic to one of those in Example 6.1.

**Lemma 6.4.** *Any symplectic fibration is cobordant to a strict one.*

*Proof.* Take a symplectic fibration $E$, and define $\psi$ and $\gamma$ as before. Let $c : W \longrightarrow \mathbb{R}^{\geq 0}$ be given by $c(w) = \max(\{\psi_w(x) \mid x \in \partial M\} \cup \{0\})$. Choose some $\epsilon > 0$ and a function $\xi \in C^\infty(\mathbb{R}^{\leq 0} \times W, \mathbb{R})$, such that the restriction $\xi_w$ has the following properties for each $w \in W$: $\xi_w(r) = 0$ for $r \leq -c(w) - 2\epsilon$, $\xi_w(r) = 1$ for $r \geq -\epsilon$, and $0 \leq \xi_w'(r) < c(w)^{-1}$ for $r \in \mathbb{R}^{\leq 0}$. For $(w,s) \in W \times [0;1]$ define $\tilde{\o}_{w,s} \in \Omega^2(E_w)$ by setting $\tilde{\o}_{w,s} = \o_w$ outside $\gamma_w([-c(w)-2\epsilon; 0] \times \partial M)$ and

$$\gamma_w^* \tilde{\o}_{w,s} = d(e^{r+\psi_w - s\xi_w(r)\psi_w}\alpha). \tag{6.1}$$

$\tilde{\o}_{w,s}$ is a symplectic form on $E_w$. In fact, the top exterior power of the two-form on the r.h.s of (6.1) is $(1 - s\xi_w'(r)\psi_w)\, dr \wedge \alpha \wedge m(d\alpha)^{m-1}$, and this is everywhere nonzero by assumption on $\xi_w'$. Define $\tilde{\alpha}_{w,s} \in \Omega^1(\partial E_w)$ by $\tau_w^* \tilde{\alpha}_{w,s} = \exp((1-s)\psi_w)\alpha$. In the same way one can construct one-forms $\tilde{\theta}_{w,s}$ on $E_w$ with $d\tilde{\theta}_{w,s} = \tilde{\o}_{w,s}$ and $\tilde{\theta}_{w,s} | \partial E_w = \tilde{\alpha}_{w,s}$, which shows that the relative cohomology class $[\tilde{\o}_{w,s}, \tilde{\alpha}_{w,s}]$ is zero. $(E_w, \tilde{\o}_{w,s}, \tilde{\alpha}_{w,s})$ is a symplectic manifold with contact type boundary for all $(w,s) \in W \times [0;1]$. For $s = 0$ these are the original fibres $(E_w, \o_w, \alpha_w)$; the same holds for arbitrary $s$ when $w \in \widehat{W}$, since then $\psi_w = 0$. Note also that for $s = 1$, $\tau_w^* \tilde{\alpha}_{w,1} = \alpha$. Finally, by putting these forms onto $\widetilde{E} = E \times [0;1] \to W \times [0;1]$ and taking $\tilde{\tau}, \tilde{\eta}$ to be induced from $\tau, \eta$ in the obvious way, one obtains a cobordism between $E$ and the strict symplectic fibration $\widetilde{E}_1 = \widetilde{E} | W \times \{1\}$. $\square$

*Proof of Proposition 6.2.* For a strict symplectic fibration $E$ one defines $\rho_s^E$ in terms of the fibre bundle structure from Lemma 6.3. To show that this is independent of all choices, it is sufficient to observe that any two fibre bundle structures on $E$ obtained from different choices of $\Omega$ can be joined by a fibre bundle structure on $E \times [0;1]$. As mentioned earlier, the definition is then extended to general symplectic fibrations by using Lemma 6.4. There is again a uniqueness part of the proof, which consists in showing that two strict symplectic fibrations which are cobordant are also cobordant by a strict cobordism; this is a relative version of Lemma 6.4. The properties (A1)–(A3) follow immediately from the definition; and the fact that they characterize $\rho_s^E$ is obvious. $\square$

Our construction implies that $\rho_s^E$ is multiplicative in the following sense: if $\beta_0, \beta_1$ are two paths $([0;1],\{0;1\}) \longrightarrow (W, \widehat{W})$ with $\beta_0(1) = \beta_1(0)$, then $\rho_s^E([\beta_1 \circ \beta_0]) = \rho_s^E([\beta_1])\rho_s^E([\beta_0])$. In the special case where $\widehat{W} = \{\hat{w}\}$ is a single point, one obtains a group homomorphism $\pi_1(W, \hat{w}) \longrightarrow \pi_0(\mathrm{Symp}(M, \partial M, \o))$ (because of the way in which we write the composition of paths, our fundamental groups are the opposites of the usual ones). As a final observation, note that $\rho_s^E$ depends only on the homotopy class (rel $\widehat{W}$) of the contact trivialization $\tau$; the reason is that changing $\tau$ within such a class gives a fibration which is cobordant to the original one.

*Remark* 6.5. An alternative way of proving Proposition 6.2 is to attach an infinite cone to $M$, forming a noncompact symplectic manifold

$$(\overline{M}, \bar{\omega}) = (M, \omega) \cup_{\partial M} (\mathbb{R}^{\geq 0} \times \partial M, d(e^r\alpha)).$$

Because $[\omega, \alpha] = 0$, the inclusion $\mathrm{Symp}(M, \partial M, \omega) \hookrightarrow \mathrm{Symp}^c(\overline{M}, \bar{\omega})$ is a weak homotopy equivalence. Given a symplectic fibration $E$, one similarly attaches an infinite cone to each fibre. The resulting family of noncompact symplectic manifolds admits a natural structure of a fibre bundle with the group $\mathrm{Symp}^c(\overline{M}, \bar{\omega})$ as structure group, and this defines a map

$$\pi_1(W, \widehat{W}) \longrightarrow \pi_0(\mathrm{Symp}^c(\overline{M}, \bar{\omega})) \cong \pi_0(\mathrm{Symp}(M, \partial M, \omega)).$$

**Lemma 6.6.** *Let $E$ be a symplectic fibration over $(W, \widehat{W}) = ([0;1], \{0;1\})$. Let $F \subset E \setminus \partial E$ be a compact submanifold which intersects each fibre $E_w$ transversally, such that $F_w = F \cap E_w \subset E_w$ is a closed Lagrangian submanifold for all $w$. Define Lagrangian submanifolds $L_0, L_1 \subset M$ by $L_w = \eta_w^{-1}(F_w)$. Let $\phi \in \mathrm{Symp}(M, \partial M, \omega)$ be a map which represents $\rho_s([\beta])$, where $\beta = \mathrm{id}_{[0;1]}$. Then $L_1$ is Lagrangian isotopic to $\phi(L_0)$.*

*Proof.* The result is obvious when $E$ is strict; as remarked after Lemma 6.3, we may then assume that $E = E_\phi$ is as in Example 6.1, and the submanifolds $F_w \subset E_w = M$ will form a Lagrangian isotopy between $L_0$ and $\phi^{-1}(L_1)$.

Now consider the general case; for this we adopt the notation from the proof of Lemma 6.4. The vector fields dual to the one-forms $\theta_w$ define a semiflow $(\lambda_s)_{s \leq 0}$ on $E$, whose restriction to each fibre is Liouville (rescales the symplectic form exponentially). Define

$$\widetilde{F} = \lambda_s(F) \times [0;1] \subset \widetilde{E},$$

where $s \in \mathbb{R}$ is a number smaller than $-c(w) - 2\epsilon$ for all $w \in W$. The restriction $\widetilde{F}_{w,s} = \widetilde{F} \cap \widetilde{E}_{w,s}$ is a closed Lagrangian submanifold in $(\widetilde{E}_{w,s}, \tilde{\omega}_{w,s}, \tilde{\alpha}_{w,s})$ for all $(w, s) \in W \times [0;1]$. By restricting to the subset $W \times \{1\}$, on which $\widetilde{E}$ becomes strict, and using the observation made above concerning the strict case together with the cobordism invariance of $\rho_s^E$, one finds that $\phi(\eta_0^{-1}(\lambda_s(F_0)))$ is Lagrangian isotopic to $\eta_1^{-1}(\lambda_s(F_1))$. Moreover, the flow $\lambda_s$ provides symplectic isotopies between $\eta_w^{-1}(\lambda_s(F_w))$ and $L_w$ for $w = 0, 1$. ☐

**Lemma 6.7.** *Let $E$ be a symplectic fibration over $(W, \widehat{W}) = ([0;1], \{0;1\})$. Let $F \subset E$ be a compact submanifold which intersects each fibre $E_w$ transversally, and such that $F_w = F \cap E_w \subset E_w$ is a Lagrangian submanifold with nonempty boundary for all $w$. Assume that the path in the space of all Legendrian submanifolds of $(\partial M, \alpha)$ defined by $w \mapsto \tau_w^{-1}(\partial F_w)$ is closed and contractible. Then, for $L_0, L_1$ and $\phi$ as in the previous lemma, $L_1$ is isotopic to $\phi(L_0)$ by a Lagrangian isotopy rel $\partial M$.*

This is the analogue of Lemma 6.6 for Lagrangian submanifolds with nonempty boundary. The proof uses Lemma 5.2 and the same ideas as before; we omit it.

6b. **Singularities.** To begin with, some notation: $B^{2k}(r)$, $\overline{B}^{2k}(r)$, $S^{2k-1}(r)$ denote the open ball, the closed ball, and the sphere of radius $r > 0$ around the origin in $\mathbb{C}^k$; $\theta_{\mathbb{C}^k} = i/4 \sum_j (x_j d\bar{x}_j - \bar{x}_j dx_j) \in \Omega^1(\mathbb{C}^k)$ and $\omega_{\mathbb{C}^k} = d\theta_{\mathbb{C}^k} \in \Omega^2(\mathbb{C}^k)$ are the standard forms.

Let $g \in \mathbb{C}[x_0, \dots, x_n]$ be a polynomial with $g(0) = 0$ and $dg(0) = 0$, such that $x = 0$ is an isolated critical point. A theorem of Milnor says that $H_0 = g^{-1}(0)$ intersects $S^{2n+1}(\epsilon)$ transversally for all sufficiently small $\epsilon > 0$. Fix an $\epsilon$ in that range. Assume that we are given a family $(g_w)_{w=(w_0,\dots,w_m)\in\mathbb{C}^{m+1}}$ of polynomials of the form $g_w(x) = g(x) + \sum_{j=0}^m w_j \tilde{g}_j(x)$, where $\tilde{g}_0 \equiv 1$. The intersection of $H_w = g_w^{-1}(0)$ with $S^{2n+1}(\epsilon)$ remains transverse for all sufficiently small $w$. Fix a $\delta > 0$ such that this is the case for all $|w| < \delta$.

**Lemma 6.8.** *The smooth family of contact manifolds* $(H_w \cap S^{2n+1}(\epsilon), \theta_w | H_w \cap S^{2n+1}(\epsilon))$, $w \in B^{2m+2}(\delta)$, *admits a trivialization, which is unique up to homotopy.*

This is an immediate consequence of Gray's theorem, because $B^{2m+2}(\delta)$ is contractible. Let $W \subset B^{2m+2}(\delta)$ be the subset of those $w$ such that $H_w \cap \overline{B}^{2n+2}(\epsilon)$ contains no singular point of $H_w$. It is open and connected, because its complement is a complex hypersurface. For each $w \in W$ equip $E_w = H_w \cap \overline{B}^{2n+2}(\epsilon)$ with the forms $\o_w = \o_{\mathbb{C}^{n+1}} | E_w$ and $\theta_w = \theta_{\mathbb{C}^{n+1}} | E_w$, and set $\alpha_w = \theta_w | \partial E_w$. Then each $(E_w, \o_w, \alpha_w)$ is a symplectic manifold with contact boundary, and the relative symplectic class $[\o_w, \alpha_w]$ is zero because of the existence of $\theta_w$. Consider the smooth fibration

$$\pi : E = \{(x, w) \in \mathbb{C}^{n+1} \times W \mid |x| \leq \epsilon,\ g_w(x) = 0\} \longrightarrow W$$

whose fibres are the manifolds $E_w$. Choose some base point $\hat{w} \in W$ and set $(M, \o, \alpha) = (E_{\hat{w}}, \o_{\hat{w}}, \alpha_{\hat{w}})$. By restriction, any trivialization as in Lemma 6.8 defines a contact trivialization $\tau : \partial M \times W \longrightarrow \partial E$. Take $\eta : M \longrightarrow E_{\hat{w}}$ to be the identity map. What we have defined is a symplectic fibration $E$ over $(W, \{\hat{w}\})$. By Proposition 6.2 one can associate to it a homomorphism

$$\rho_s = \rho_s^E : \pi_1(W, \hat{w}) \longrightarrow \pi_0(\mathrm{Symp}(M, \partial M, \o)).$$

In the usual terminology, $E$ is the Milnor fibration of the deformation $(H_w)$ of the singularity $0 \in H_0$, and the manifolds $E_w$ are the Milnor fibres. We call $\rho_s$ the symplectic monodromy map associated to the deformation.

We need to discuss briefly how $\rho_s$ depends on the various parameters involved. $\tau$ is unique up to homotopy, so that $\rho_s$ is independent of it. Changing $\delta$ affects the space $W$, but the fundamental group remains the same for all sufficiently small $\delta$, and we will assume from now on that $\delta$ has been chosen in that range. The dependence on the base point $\hat{w}$ is a slightly more complicated issue. Different fibres of the Milnor fibration are not necessarily symplectically isomorphic, but they become isomorphic after attaching an infinite cone, as in Remark 6.5. This means that, given a path in $W$ from $w_0$ to $w_1$, one can identify

$$\pi_0(\mathrm{Symp}(E_{w_0}, \partial E_{w_0}, \o_{w_0})) \cong \pi_0(\mathrm{Symp}(E_{w_1}, \partial E_{w_1}, \o_{w_1})),$$

and this fits into a commutative diagram with the corresponding isomorphism $\pi_1(W, w_0) \cong \pi_1(W, w_1)$ and with the symplectic monodromy maps at these base points. For a similar reason, making $\epsilon$ smaller changes $M$ but does not affect $\pi_0(\mathrm{Symp}(M, \partial M, \o))$ or $\rho_s$. Finally (but we will not need this) one can choose $(g_w)$ to be a miniversal deformation of $g$, and then the resulting symplectic monodromy map is really an invariant of the singularity.

*Remark* 6.9. Apart from the present paper, symplectic monodromy has been studied only for the one-parameter deformations $g_w(x) = g(x) + w$. In that case

$\pi_1(W, \hat{w}) = \mathbb{Z}$, so that it is sufficient to consider the class

$$[\phi] = \rho_s(1) \in \pi_0(\mathrm{Symp}(M, \partial M, \o)).$$

When $g$ is the ordinary singularity $g(x) = x_0^2 + \cdots + x_n^2$, $M$ is isomorphic to a neighbourhood of the zero-section in $T^*S^n$, and $\phi$ can be chosen to be the generalized Dehn twist along the zero-section in the sense of [37], [36]; those papers show that $[\phi]$ has infinite order in all dimensions $n \geq 1$. We should say that the relevance of symplectic geometry in this example was first pointed out by Arnol'd [1].

More generally, in [36] it is proved that $[\phi]$ has infinite order for all weighted homogeneous singularities such that the sum of the weights is $\neq 1$ (the definition of symplectic monodromy in [36] differs slightly from that here, but it can be shown that the outcome is the same). The general question, *does $[\phi]$ have infinite order for all nontrivial isolated hypersurface singularities?*, is open.

6c. **From curves to Lagrangian submanifolds.** We continue to use the notation introduced in the previous section. The $n$-dimensional singularity of type $(A_m)$, for $m, n \geq 1$, is given by $g(x_0, \dots, x_n) = x_0^2 + x_1^2 + \cdots + x_{n-1}^2 + x_n^{m+1}$. We consider the $(m+1)$-parameter deformation $g_w(x) = x_0^2 + \cdots + x_{n-1}^2 + h_w(x_n)$, where $h_w(z) = z^{m+1} + w_m z^m + \cdots + w_1 z + w_0 \in \mathbb{C}[z]$ (this is slightly larger than the usual miniversal deformation, but that makes no difference as far as the monodromy map is concerned). Since $g$ is weighted homogeneous, one can take $\epsilon = 1$. In contrast, we will make use of the right to choose $\delta$ small. For $w = (w_0, \dots, w_m) \in B^{2m+2}(\delta)$ set $D_w = \{z \in \mathbb{C} \mid |h_w(z)| + |z|^2 \leq 1\}$, and $\Delta_w = h_w^{-1}(0) \subset \mathbb{C}$.

**Lemma 6.10.** *Provided that $\delta$ has been chosen sufficiently small, the following properties hold for any $w \in B^{2m+2}(\delta)$:*

(P1) *$D_w \subset \mathbb{C}$ is an compact convex subset with smooth boundary, which contains $B^2(\frac{1}{2})$ and is contained in $B^2(2)$.*

(P2) *Let $\beta : [0;1] \longrightarrow \partial D_w$ be an embedded path which moves in positive sense with respect to the obvious orientation. Then*

$$\tfrac{d}{dt} \arg h_w(\beta(t)) > 0, \quad \tfrac{d}{dt} \arg \beta(t) > 0$$

*for all $t \in [0;1]$, where $\arg = \mathrm{im}\,\log : \mathbb{C}^* \longrightarrow \mathbb{R}/2\pi\mathbb{Z}$ is the argument function.*

(P3) *$\Delta_w$ is contained in $B^2(\frac{1}{3})$.*

*Proof.* For $w = 0$, $h_0(z) = z^{m+1}$ and $\partial D_0 = \{|z|^{m+1} + |z|^2 = 1\}$ is a circle centered at the origin, so the conditions are certainly satisfied. Perturbing $w$ slightly will change $h_w$ and $\partial D_w$ only by a small amount in any $C^k$-topology, and this implies the first two parts. The proof of (P3) is even simpler. ☐

**Lemma 6.11.** *$W \subset B^{2m+2}(\delta)$ is the subset of those $w$ such that $h_w$ has no multiple zeros.*

This follows immediately from property (P3) above. It implies in particular that for $w \in W$, $\Delta_w \subset D_w \setminus \partial D_w$ is a set of $(m+1)$ points.

We will now study the symplectic geometry of the Milnor fibres $E_w = H_w \cap \overline{B}^{2n+2}(1) = \{|x| \leq 1,\ x_0^2 + \cdots + x_{n-1}^2 + h_w(x_n) = 0\}$ equipped with the forms $\o_w, \theta_w, \alpha_w$. Consider the projection $H_w \longrightarrow \mathbb{C}$, $x \longmapsto x_n$, whose fibres are the affine quadrics $Q_{w,z} = \{x_0^2 + \cdots + x_{n-1}^2 = -h_w(z),\ x_n = z\}$. $Q_{w,z}$ is smooth for all $z \notin \Delta_w$, and actually symplectically isomorphic to $T^*S^{n-1}$. If one chooses a symplectic isomorphism which respects the obvious $O(n)$-actions, the zero-section $S^{n-1} \subset T^*S^{n-1}$ corresponds to $\Sigma_{w,z} = \sqrt{-h_w(z)} S^{n-1} \times \{z\} \subset Q_{w,z}$, which is the subset of those $x = (x_0, \dots, x_{n-1}, x_n)$ such that $x_n = z$, $|x_0|^2 + \cdots + |x_{n-1}|^2 = |h_w(z)|$, and $x_i \in \sqrt{-h_w(z)}\mathbb{R}$ for $i = 0, \dots, n-1$. For $z \in \Delta_w$, $Q_{w,z}$ is homeomorphic to $T^*S^{n-1}$ with the zero-section collapsed to a point, that point being precisely $\Sigma_{w,z}$. Moreover, by definition of $D_w$,

$$\Sigma_{w,z} \text{ is } \begin{cases} \text{contained in } E_w \setminus \partial E_w & \text{if } z \in D_w \setminus \partial D_w, \\ \text{contained in } \partial E_w & \text{if } z \in \partial D_w, \\ \text{disjoint from } E_w & \text{otherwise.} \end{cases}$$

For $w \in W$ and $c$ a curve in $(D_w, \Delta_w)$ as defined in Section 3a, set

$$L_{w,c} = \bigsqcup_{z \in c} \Sigma_{w,z} \subset E_w.$$

**Lemma 6.12.** (a) *If $c$ meets $\Delta_w$, then $L_{w,c}$ is a $\theta_w$-exact Lagrangian submanifold of $(E_w, \omega_w, \alpha_w)$.*

(b) *Let $c_0, c_1$ be two curves in $(D_w, \Delta_w)$ which meet $\Delta_w$ and such that $c_0 \simeq c_1$. Then $L_{w,c_0}, L_{w,c_1}$ are isotopic by an exact Lagrangian isotopy rel $\partial E_w$.*

*Proof.* (a) Assume first that $c$ can be parametrized by a smooth embedding $\gamma : [0;1] \longrightarrow D_w$ with $\gamma^{-1}(\Delta_w) = \{0\}$ and $\gamma^{-1}(\partial D_w) = \{1\}$. Because $\gamma(0)$ is a simple zero of $h_w$, one can write $h_w(\gamma(t)) = t\,k(t)$ for some $k \in C^\infty([0;1], \mathbb{C}^*)$. Choose a smooth square root $\sqrt{-k}$, and define a map from the unit ball $\overline{B}^n \subset \mathbb{R}^n$ to $E_w$ by

$$y \longmapsto (y\sqrt{-k(|y|)^2}, \gamma(|y|^2)). \tag{6.2}$$

This is a smooth embedding with image $L_{w,c}$. Since $\gamma$ meets $\partial D_w$ transversally, the embedding intersects $\partial E_w$ transversally. Moreover, the pullback of $\theta_w$ under (6.2) is of the form $\psi(|y|^2)d|y|^2$ for some $\psi \in C^\infty([0;1], \mathbb{R})$. This implies that $\omega_w|L_{w,c} = 0$ and $\theta_w|\partial L_{w,c} = 0$, which means that $L_{w,c}$ is indeed a Lagrangian submanifold.

In the other case, when $c$ joins two points of $\Delta_w$, $L_{w,c}$ is a Lagrangian submanifold diffeomorphic to the $n$-sphere; this is proved in the same way as before, only that now one splits $c$ into two parts, and covers $L_{w,c}$ by two smooth charts.

The $\theta_w$-exactness of $L_{w,c}$ is obvious for all $n \geq 2$, since then $H^1(L_{w,c}, \partial L_{w,c}; \mathbb{R}) = 0$. For $n = 1$ what one has to prove is that

$$\int_{L_{w,c}} \theta_w = 0$$

for some orientation of $L_{w,c}$. Now the involution $\iota(x_0, x_1) = (-x_0, x_1)$ preserves $\theta_w$ and reverses the orientation of $L_{w,c}$, as one can see by looking at the tangent space at any point $(0, z)$, $z \in c \cap \Delta_w$; this implies the desired result. Part (b) of the lemma follows from (a) together with Lemma 5.3. $\square$

$L_{w,c}$ is Lagrangian even if $c$ does not meet $\Delta_w$, but then it is no longer $\theta_w$-exact; which is why we do not use these submanifolds.

**Lemma 6.13.** *Let $c$ be a curve in $(D_w, \Delta_w)$ which joins a point of $\Delta_w$ with a point of $\partial D_w$. Take a vector field on $\partial D_w$ which is nonvanishing and positively oriented. Extend it to a vector field $Z$ on $D_w$ which vanishes near $\Delta_w$, and let $(f_t)$ be the flow of $Z$. Then the Lagrangian isotopy $t \mapsto L_{w,f_t(c)}$, $0 \leq t \leq 1$, is exact and also positive in the sense of Section 5a.*

*Proof.* The exactness follows again from Lemma 5.3. Define a smooth path $\beta : [0;1] \longrightarrow \partial D_w$ by requiring that $\beta(t)$ be the unique point of $f_t(c) \cap \partial D_w$. Choose some square root of $-h_w \circ \beta$, and consider

$$S^{n-1} \times [0;1] \longrightarrow \partial E_w, \quad (y,t) \longmapsto (\sqrt{-h_w(\beta(t))}y, \beta(t)). \tag{6.3}$$

This maps $S^{n-1} \times \{t\}$ to $\partial L_{w,f_t(c)}$. A computation shows that the pullback of $\alpha_w$ under (6.3) is $\psi(t)\,dt$, where $\psi(t) = |h_w(\beta(t))|\frac{d}{dt}(\arg h_w(\beta(t))) + \frac{1}{2}|\beta(t)|^2\frac{d}{dt}(\arg \beta(t))$ is a positive function by (P2) in Lemma 6.10. This is one of the definitions of a positive isotopy. ☐

The intersections of the submanifolds $L_{w,c}$ are governed by those of the underlying curves. In fact $L_{w,c_0} \cap L_{w,c_1} = \coprod_{z \in c_0 \cap c_1} \Sigma_{w,z}$ consists of an $(n-1)$-sphere for any point of $(c_0 \cap c_1) \setminus \Delta_w$ together with an isolated point for each point of $c_0 \cap c_1 \cap \Delta_w$. The next lemma translates this elementary fact into a statement about Floer cohomology.

**Lemma 6.14.** *Let $c_0, c_1$ be two curves in $(D_w, \Delta_w)$, each of which meets $\Delta_w$. Then the dimension of $HF(L_{w,c_0}, L_{w,c_1})$ is $2\,I(c_0, c_1)$.*

*Proof.* We consider first the case when $c_0 \cap c_1 \cap \partial D_w = \emptyset$ and $c_0 \not\simeq c_1$. $I(c_0, c_1)$ depends only on the isotopy classes of $c_0, c_1$; the same is true of $HF(L_{w,c_0}, L_{w,c_1})$, by Lemma 6.12(b) and Proposition 5.10. Hence we may assume that $c_0, c_1$ have minimal intersection.

A straightforward computation of the tangent spaces shows that $L_{w,c_0}, L_{w,c_1}$ have clean intersection. Moreover, since $c_0 \cap c_1 \cap \partial D_w = \emptyset$, $L_{w,c_0} \cap L_{w,c_1} \cap \partial E_w = \emptyset$. Consider the involution $\iota(x_0, \dots, x_n) = (-x_0, \dots, -x_{n-2}, x_{n-1}, x_n)$ on $E_w$, which preserves $\o_w$ and $\theta_w$. Each connected component $\Sigma_{w,z}$ of $L_{w,c_0} \cap L_{w,c_1}$ carries an $\iota$-invariant Morse function $h_z$ whose critical point set is $\{(0, \dots, \pm\sqrt{-h_w(z)}, z)\} = \Sigma_{w,z} \cap M^\iota$. Namely, for $z \in \Delta_w$ one takes $h_z$ to be constant, and otherwise $h_z(x) = (-h_w(z))^{-1/2} x_{n-1}$ for some choice of square root. Hence $L_{w,c_0}, L_{w,c_1}$ satisfy condition (S1$'$) in Proposition 5.15.

The fixed point set $M^\iota$ is precisely the double branched cover of $D_w$ from Section 3c, and $L_{w,c_j} \cap M^\iota \subset M^\iota$ is the preimage of $c_j$ under the covering projection. Lemma 3.8 shows that $L_{w,c_0} \cap M^\iota$ and $L_{w,c_1} \cap M^\iota$ have minimal intersection as curves in $M^\iota$. Hence, using Lemma 3.1, it follows that $L_{w,c_0}, L_{w,c_1}$ also satisfy condition (S2$'$) in Proposition 5.15. Applying that proposition shows that $\dim HF(L_{w,c_0}, L_{w,c_1}) = \dim H^*(L_{w,c_0} \cap L_{w,c_1}; \mathbb{Z}/2) = 2|(c_0 \cap c_1) \setminus \Delta_w| + |c_0 \cap c_1 \cap \Delta_w| = 2\,I(c_0, c_1)$.

Next, consider the case when $c_0 \cap c_1 \cap \partial D_w = \emptyset$ and $c_0 \simeq c_1$. This means that $c_0$ must be a path joining two points of $\Delta_w$; hence $I(c_0, c_1) = I(c_0, c_0) = 1$. On the other hand, since $L_{w,c_0}$ and $L_{w,c_1}$ are isotopic by an exact Lagrangian isotopy, one has $HF(L_{w,c_0}, L_{w,c_1}) \cong HF(L_{w,c_0}, L_{w,c_0}) \cong H^*(L_{w,c_0}; \mathbb{Z}/2) \cong (\mathbb{Z}/2)^2$ by Floer's theorem [13].

The final case is $c_0 \cap c_1 \cap \partial D_w \neq \emptyset$. Then $c_0, c_1$ each have one point on the boundary $\partial D_w$, which is the same for both, so that $\partial L_{w,c_0} = \partial L_{w,c_1}$. To compute $HF(L_{w,c_0}, L_{w,c_1})$ one has to perturb $L_{w,c_0}$ by an exact and positive Lagrangian isotopy. One can take the isotopy to be $L_{w,f_t(c_0)}$ for some flow $(f_t)$ on $D_w$ as in Lemma 6.13. By definition, $I(c_0, c_1) = I(f_t(c_0), c_1)$ for small $t > 0$. This reduces the Floer cohomology computation to the case which we have already proved. ☐

We are now ready to bring in the symplectic monodromy map. Fix a base point $\hat{w} \in W$. Write $(M, \o, \theta, \alpha) = (E_{\hat{w}}, \o_{\hat{w}}, \theta_{\hat{w}}, \alpha_{\hat{w}})$, $(D, \Delta) = (D_{\hat{w}}, \Delta_{\hat{w}})$, $h = h_{\hat{w}}$, and $L_c = L_{\hat{w},c}$ for any curve $c$ in $(D, \Delta)$.

One can embed $W$ as an open subset into $\mathrm{Conf}_{m+1}(D \setminus \partial D)$ by sending $w \mapsto \Delta_w$. The image of this embedding contains all configurations of points which lie close to the origin, and it is mapped into itself by the maps which shrink configurations by some factor $0 < r < 1$. From this one can easily see that the embedding is a weak homotopy equivalence. By combining this with the observation made in Section 3b one obtains canonical isomorphisms

$$\pi_1(W, \hat{w}) \cong \pi_1(\mathrm{Conf}_{m+1}(D), \Delta) \cong \pi_0(\mathcal{G}), \tag{6.4}$$

where $\mathcal{G} = \mathrm{Diff}(D, \partial D; \Delta)$.

**Lemma 6.15.** *Let $f \in \mathcal{G}$ and $\phi \in \mathrm{Symp}(M, \partial M, \o)$ be related in the following way: $[f] \in \pi_0(\mathcal{G})$ corresponds to a class in $\pi_1(W, \hat{w})$ whose image under $\rho_s$ is $[\phi] \in \pi_0(\mathrm{Symp}(M, \partial M, \o))$. Then, for any curve c in $(D, \Delta)$ which meets $\Delta$, $L_{f(c)}$ and $\phi(L_c)$ are isotopic by a Lagrangian isotopy rel $\partial M$.*

*Proof.* Consider first the case when $c \cap \partial D = \emptyset$. Since the result depends only on the isotopy class of $c$, we may assume that $c \subset B^2(1/2)$. Similarly, the result depends only on the isotopy class of $f$, and we may assume that $f$ is the identity outside $B^2(1/2) \subset D$. Take a smooth path $\beta : ([0;1], \{0;1\}) \longrightarrow (W, \hat{w})$ whose class in $\pi_1(W, \hat{w})$ corresponds to $[f] \in \pi_0(\mathcal{G})$ under the canonical isomorphism. This means that there is an isotopy $(f_t)_{0 \leq t \leq 1}$ in $\mathrm{Diff}(D, \partial D)$ with $f_0 = \mathrm{id}$, $f_1 = f$, and $f_t(\Delta) = \Delta_{\beta(t)}$ for all $t$. We may assume that each $f_t$ is equal to the identity outside $B^2(1/2)$. Then $f_t(c)$ is a curve in $(D_{\beta(t)}, \Delta_{\beta(t)})$ for all $t \in [0;1]$. Moreover, the Lagrangian submanifolds $L_{\beta(t), f_t(c)}$ depend smoothly on $t$; by this we mean that they form a submanifold $F$ of the pullback $\beta^* E \longrightarrow [0;1]$, as in Lemma 6.6. By using that lemma and the naturality of the maps $\rho_s^E$ under pullbacks, it follows that $L_{f(c)}$ is Lagrangian isotopic to $\phi(L_c)$.

The case when $c \cap \partial D \neq \emptyset$ is not very different. We may assume that the intersection $c \cap (D \setminus B^2(1/2))$ consists of a straight radial piece $\mathbb{R}^{>0}\zeta \cap (D \setminus B^2(1/2))$, for some $\zeta \in S^1$. We take a smooth path $\beta$ and an isotopy $(f_t)$ as in the previous case. Then $f_t(c) \cap D_{\beta(t)}$ is a curve in $(D_{\beta(t)}, \Delta_{\beta(t)})$ for all $0 \leq t \leq 1$. The rest is as before except that in order to apply Lemma 6.7, one has to show that the path of Legendrian submanifolds of $(\partial M, \alpha)$ given by

$$t \mapsto \tau_{\beta(t)}^{-1}(\partial L_{\beta(t), f_t(c) \cap D_{\beta(t)}})$$

is closed and contractible (in the space of Legendrian submanifolds). The closedness is obvious because $L_{f(c)}$ and $L_c$ have the same boundary. In order to prove contractibility, observe that for any $w \in B^{2m+2}(\delta)$, one can define a Legendrian submanifold $\Lambda_w$ in $(H_w \cap S^{2n+1}(1), \alpha_w | H_w \cap S^{2n+1}(1))$ by

$$\Lambda_w = \Big( \coprod_{1/2 \leq s \leq 2} \Sigma_{w, s\zeta} \Big) \cap S^{2n+1}(1).$$

By using a relative version of Lemma 6.8, one sees that the trivialization $\tau$ can be chosen in such a way that $\tau_w^{-1}(\Lambda_w) \subset \partial M$ is the same for all $w \in W$. Now our path of Legendrian submanifolds consists of $\Lambda_{\beta(t)}$, so that its contractibility is obvious for that specific choice of $\tau$. Since any two $\tau$ are homotopic, it follows that contractibility holds for an arbitrary choice. $\square$

**Lemma 6.16.** *Assume that the symplectic automorphism $\phi$ in Lemma 6.15 has been chosen $\theta$-exact. Then the isotopy between $L_{f(c)}$ and $\phi(L_c)$ in Lemma 6.15 can be made exact.*

*Proof.* For $n > 1$ there is nothing to prove: first of all, $H^1(M, \partial M; \mathbb{R})$ is zero, so that any choice of $\phi$ is $\theta$-exact; and secondly, $H^1(L_c, \partial L_c; \mathbb{R}) = 0$, so that any Lagrangian isotopy between such submanifolds is exact.

Therefore we now assume that $n = 1$. Since $\phi$ is $\theta$-exact, both $\phi(L_c)$ and $L_{f(c)}$ are $\theta$-exact Lagrangian submanifolds. With one exception, it is true that $L_c$ represents a nontrivial class in $H_1(M, \partial M)$, so that the restriction map $H^1(M, \partial M; \mathbb{R}) \longrightarrow H^1(L_c, \partial L_c; \mathbb{R})$ is onto; then one can apply Lemma 5.4(b) to prove the desired result.

The one exception is when $(m, n) = (1, 1)$ and $c$ is the unique (up to isotopy) curve joining the two points of $\Delta$. In this particular case the Milnor fibre $(M, \emptyset)$ is isomorphic to a cylinder $([-s_0; s_0] \times S^1, ds \wedge dt)$, and $L_c$ is a simple closed curve which divides $M$ into two pieces of equal area; to verify this last fact, it is sufficient to note that the involution $\iota$ exchanges the two pieces. Clearly $\phi(L_c)$ also divides $M$ into two pieces of equal area; and this is sufficient to prove that $\phi(L_c)$ is isotopic to $L_c$ by an exact isotopy. On the other hand, $f(c) \simeq c$ for any $f \in \mathcal{G}$, so that $L_c$ is exact isotopic to $L_{f(c)}$ by Lemma 6.12(b). □

Now choose a set of basic curves $b_0, \dots, b_m$ in $(D, \Delta)$, and let $L_0 = L_{b_0}, \dots, L_m = L_{b_m}$ be the corresponding Lagrangian submanifolds of $(M, \emptyset, \alpha)$; from the discussion above, it follows that these do indeed satisfy (1.3). As explained in Section 3b, the choice of $b_0, \dots, b_m$ allows one to identify each of the canonically isomorphic groups (6.4) with $B_{m+1}$. In particular, the symplectic monodromy can now be thought of as a map $\rho_s : B_{m+1} \longrightarrow \pi_0(\mathrm{Symp}(M, \partial M, \emptyset))$.

*Proof of Theorem 1.3.* From the various isomorphisms which we have introduced, it follows that $\phi = \phi_\sigma$ and $f = f_\sigma$ are related as in Lemmas 6.15 and 6.16. Using those two results together with Lemma 6.14, one finds that $\dim HF(L_i, \phi(L_j)) = \dim HF(L_{b_i}, L_{f(b_j)}) = 2\, I(b_i, f(b_j))$. □

*Proof of Corollary 1.4.* Let $\sigma \in B_{m+1}$ be an element whose image under the monodromy map is the trivial class in $\pi_0(\mathrm{Symp}(M, \partial M, \emptyset))$. Then one can take $\phi_\sigma = \mathrm{id}_M$ in Theorem 1.3, so that

$$I(b_i, f_\sigma(b_j)) = \tfrac{1}{2} \dim HF(L_i, \phi_\sigma(L_j)) = \tfrac{1}{2} \dim HF(L_i, L_j) = I(b_i, b_j)$$

for all $0 \leq i, j \leq m$. The same holds for $f_\sigma^2$, since $\rho_s(\sigma^2)$ is again trivial. In view of Lemma 3.6, it follows that $\sigma$ must be the trivial element. □

6d. **Bigraded curves and graded Lagrangian submanifolds.** We use the notation $(D, \Delta)$, $h$, $L_c$, and $(M, \emptyset, \alpha, \theta)$ as in the previous section. As explained in Section 3d, the polynomial $h$ can be used to define a map $\delta_P$ which classifies the covering $\widetilde{P}$ of $P$, and one can express the notion of bigraded curves in terms of this map. On the other hand, $M$ is the intersection of a complex hypersurface $H_{\hat{w}}$ with the unit ball $B^{2n+2}(1)$. As explained in Section 5e, the polynomial $g_{\hat{w}}$ which defines $H_{\hat{w}}$ determines a map $\delta_{\mathcal{L}} : \mathcal{L} \longrightarrow \mathbb{C}^*/\mathbb{R}^{>0}$; and one can use this map to define a covering $\widetilde{\mathcal{L}}$ and the corresponding notion of graded Lagrangian submanifold of $M$.

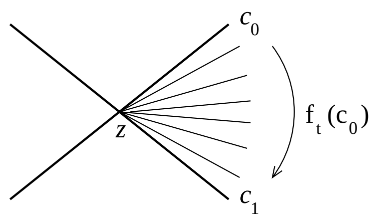



FIGURE 25.

**Lemma 6.17.** *Let $c$ be a curve in $(D, \Delta)$ which meets $\Delta$. Any bigrading $\tilde{c}$ of $c$ determines in a canonical way a grading $\widetilde{L}_c$ of $L_c$. Moreover, if $\widetilde{L}_c$ is the grading associated to $\tilde{c}$, the grading associated to $\chi(r_1, r_2)\tilde{c}$ is $\chi_{\widetilde{\mathcal{L}}}(r_1 + nr_2)\widetilde{L}_c$.*

*Proof.* Let $\gamma : (0;1) \longrightarrow D \setminus \Delta$ be an embedding which parametrizes an open subset of $c$, and let $x \in M$ be a point of $\Sigma_{\hat{w},\gamma(t)} \times \{\gamma(t)\} \subset L_c$. A straightforward computation, starting from (5.5), shows that

$$\delta_{\mathcal{L}}(T_x L_c) = (-h(\gamma(t)))^{n-2}\gamma'(t)^2 \in \mathbb{C}^*/\mathbb{R}^{>0}. \tag{6.5}$$

Now assume that we have a bigrading of $c$, which is a map $\tilde{c} = (\tilde{c}_1, \tilde{c}_2) : c \setminus \Delta \longrightarrow \mathbb{R}^2$ with specific properties. Define

$$\widetilde{L}_c : L \setminus \bigcup_{z \in c \cap \Delta} \Sigma_{\hat{w},z} \times \{z\} \longrightarrow \mathbb{R}$$

by $\widetilde{L}_c(x) = \tilde{c}_1(\gamma(t)) + n\tilde{c}_2(\gamma(t))$ for a point $x$ as before. A comparison of (6.5) with (5.5) and (3.2) shows that $\widetilde{L}_c$ is a grading of $L$, defined everywhere except at finitely many points. It remains to prove that $\widetilde{L}_c$ can be extended continuously over those missing points. If $n \geq 2$, this is true for general reasons. Namely, whenever one has an infinite cyclic covering of an $n$-dimensional manifold (which in our case is $s_L^*\widetilde{\mathcal{L}}$) and a continuous section of it defined outside a finite subset, the values at the missing points can always be filled in continuously. In the remaining case $n = 1$, any one of the missing points divides $L_c$ locally into two components, and one has to check that the limits of $\widetilde{L}_c$ from both directions (which certainly exist) are the same. But this follows from the fact that $\widetilde{L}_c$ is invariant under the involution $\iota$.

The last sentence of the lemma is obvious from the construction. □

**Lemma 6.18.** *Let $c_0, c_1$ be two curves in $(D, \Delta)$ which meet $\Delta$ and which intersect transversally. Let $\tilde{c}_0, \tilde{c}_1$ be two bigradings, and let $(r_1, r_2) = \mu^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1; z)$ be the local index at a point $z \in c_0 \cap c_1$. Let $\widetilde{L}_{c_0}, \widetilde{L}_{c_1}$ be the corresponding gradings of $L_{c_0}, L_{c_1}$. Then the absolute Maslov index at any point $x \in \Sigma_{\hat{w},z} \subset L_{c_0} \cap L_{c_1}$ is*

$$\mu^{\mathrm{abs}}(\widetilde{L}_{c_0}, \widetilde{L}_{c_1}; x) = r_1 + nr_2.$$

*Proof.* In view of the last part of Lemma 6.17 it is sufficient to prove this for $(r_1, r_2) = (0, 0)$. Take an isotopy $(f_t)_{0 \leq t \leq 1}$ in $\mathrm{Diff}(D, \partial D; \Delta)$ with $f_t(z) = z$ for all $t$, $f_0 = \mathrm{id}$, and $f_1(c_0) = c_1$ in some neighbourhood of $z$. We also assume that $Df_t(T_z c_0)$ rotates clockwise with positive speed during the isotopy, and that the total angle by which it moves is less than $\pi$ (see Figure 25). Let $\widetilde{L}_{f_t(c_0)}$ be the family of graded Lagrangian submanifolds corresponding to the bigraded curves $\tilde{f}_t(\tilde{c}_0)$.

By definition of the local index, $\tilde{f}_1(\tilde{c}_0)$ is equal to $\tilde{c}_1$ near $z$. Hence $\widetilde{L}_{f_t(c_0)}$ agrees with $\widetilde{L}_{c_1}$ near $\Sigma_{\hat{w},z}$. It follows that for any $x \in \Sigma_{\hat{w},z}$ one can use the paths $\lambda_0, \lambda_1 : [0;1] \longrightarrow \mathcal{L}_x$, $\lambda_0(t) = T_x L_{f_t(c_0)}$, $\lambda_1(t) = T_x L_{c_1}$, to compute the absolute Maslov index. Now, it is not difficult to see that these two paths satisfy the conditions of Lemma 5.17, so that the absolute Maslov index is in fact zero. □

**Lemma 6.19.** *Let $(c_0, \tilde{c}_0)$ and $(c_1, \tilde{c}_1)$ be two bigraded curves in $(D, \Delta)$, such that both $c_0$ and $c_1$ intersect $\Delta$. Let $\widetilde{L}_{c_0}, \widetilde{L}_{c_1}$ be the corresponding graded Lagrangian submanifolds of $(M, \o, \alpha)$. Then the Poincaré polynomial of the graded Floer cohomology is obtained by setting $q_1 = q$, $q_2 = q^n$ in the bigraded intersection number:*

$$\sum_{r \in \mathbb{Z}} q^r \dim HF^r(\widetilde{L}_{c_0}, \widetilde{L}_{c_1}) = I^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1)_{q_1 = q,\, q_2 = q^n}.$$

*Proof.* This is essentially the same argument as for Lemma 6.14. Consider the case when $c_0 \cap c_1 \cap \partial D = \emptyset$ and $c_0 \not\simeq c_1$. We may assume that $c_0, c_1$ have minimal intersection. Using Proposition 5.18, and Lemma 6.18 to compute the relevant absolute Maslov indices, one finds that the following holds: a point $z \in c_0 \cap c_1$ with local index $(\mu_1(z), \mu_2(z)) \in \mathbb{Z}^2$ contributes

$$\begin{cases} q^{\mu_1(z) + n\mu_2(z)}(1 + q^{n-1}) & \text{if } z \notin \Delta, \\ q^{\mu_1(z) + n\mu_2(z)} & \text{if } z \in \Delta \end{cases}$$

to the Poincaré polynomial of the Floer cohomology. The factor $(1 + q^{n-1})$ in the first case comes from the ordinary cohomology of the $(n-1)$-sphere $\Sigma_{\hat{w},z}$. It is now straightforward to compare this with the definition of the bigraded intersection number $I^{\mathrm{bigr}}(\tilde{c}_0, \tilde{c}_1)$.

As in Lemma 6.14, the case $c_0 \simeq c_1$ can be checked by a simple computation, and that case $c_0 \cap c_1 \cap \partial D \neq \emptyset$ follows from the previously considered one. □

*Proof of Corollary 1.5.* One can forget about graded Lagrangian submanifolds and consider Lemma 6.19 as a statement about the relative grading on $HF^*(L_{c_0}, L_{c_1})$. Using this, and arguing as in the proof of Theorem 1.3, one sees that (for some choice of relative grading on the left-hand side)

$$\sum_{r \in \mathbb{Z}} q^r \dim HF^r(L_i, \phi(L_j)) = I^{\mathrm{bigr}}(b_i, f(b_j))_{q_1 = q,\, q_2 = q^n}.$$

The proof is completed by comparing this with Proposition 4.9. □

## Acknowledgements

The first author would like to thank Maxim Vybornov who introduced him to the algebras $A_m$ by pointing out that they describe perverse sheaves on projective spaces. The construction of the Lagrangian submanifolds $L_c$ was suggested to the second author by Simon Donaldson. Most of this work was done at the Institute for Advanced Study, during the year 97/98 and two shorter visits later on. The authors are very grateful to the Institute, and to Robert MacPherson in particular, for making this collaboration possible.

## References


[1] V. I. Arnol'd, *Some remarks on symplectic monodromy of Milnor fibrations*, The Floer Memorial Volume (H. Hofer, C. Taubes, A. Weinstein, and E. Zehnder, eds.), Progress in Mathematics, vol. 133, Birkhäuser, 1995, pp. 99–104. MR **96m:**32043

[2] A. A. Beilinson, *Coherent sheaves on* $\mathbb{P}^n$ *and problems of linear algebra*, Funct. Anal. Appl. **12** (1978), 214–216. MR **80c:**14010b

[3] A. Beilinson, V. Ginzburg and W. Soergel, *Koszul duality patterns in representation theory*, Journal of the AMS **9**, n.2, (1996), 473–527. MR **96k:**17010

[4] J. Bernstein, I. Frenkel, and M. Khovanov, *A categorification of the Temperley-Lieb algebra and Schur quotients of* $U(sl_2)$ *via projective and Zuckerman functors*, Selecta Math., New ser. 5, (1999), 199–241; math.QA/0002087. MR **2000i:**17009

[5] J. Bernstein and S. Gelfand, *Tensor products of finite and infinite dimensional representations of semisimple Lie algebras*, Compositio Math. **41** (1980), 245–285. MR **82c:**17003

[6] J. Birman, *Braids, links and mapping class groups*, Annals of Math. Studies 82, Princeton University Press, 1974. MR **51:**11477; errata MR **83m:**57006

[7] J. S. Birman and H. M. Hilden, *On isotopies of homeomorphisms of Riemann surfaces*, Annals of Math. **97** (1973), 424–439. MR **48:**4305

[8] E. Brieskorn, *Über die Auflösung gewisser Singularitäten von holomorphen Abbildungen*, Math. Ann. **166** (1966), 76–102. MR **34:**6789

[9] Ya. Eliashberg, H. Hofer, and D. Salamon, *Lagrangian intersections in contact geometry*, Geom. Funct. Anal. **5** (1995), 244–269. MR **96c:**58034

[10] A. Fathi, F. Laudenbach, and V. Poénaru, *Travaux de Thurston sur les surfaces*, Astérisque, vol. 66–67, Soc. Math. France, 1979.

[11] A. Floer, *Morse theory for Lagrangian intersections*, J. Differential Geom. **28** (1988), 513–547. MR **90f:**58058

[12] ———, *A relative Morse index for the symplectic action*, Comm. Pure Appl. Math. **41** (1988), 393–407. MR **89f:**58055

[13] ———, *Witten's complex and infinite dimensional Morse theory*, J. Differential Geom. **30** (1989), 207–221. MR **90d:**58029

[14] A. Floer, H. Hofer, and D. Salamon, *Transversality in elliptic Morse theory for the symplectic action*, Duke Math. J. **80** (1995), 251–292. MR **96h:**58024

[15] K. Fukaya, *Morse homotopy,* $A_\infty$*-categories, and Floer homologies*, Proceedings of GARC workshop on Geometry and Topology (H. J. Kim, ed.), Seoul National University, 1993. MR **95e:**57053

[16] ———, *Floer homology for three-manifolds with boundary I*, Preprint, 1997.

[17] S. Gelfand, Yu. Manin, *Methods of homological algebra*, Springer-Verlag, Berlin, 1996. MR **97j:**18001

[18] F. M. Goodman and H. Wenzl, *The Temperley-Lieb algebra at roots of unity,* Pacific J. Math. **161** (1993), no.2, 307–334. MR **95c:**16020

[19] A. Haefliger, *Plongements différentiables des variétés dans variétés*, Comm. Math. Helv. **36** (1962), 47–82. MR **26:**3069

[20] H. Hofer, *Pseudoholomorphic curves in symplectizations with applications to the Weinstein conjecture in dimension three*, Invent. Math. **114** (1993), 515–563. MR **94j:**58064

[21] S. Huerfano and M. Khovanov, *A category for the adjoint representation*, math.QA/0002087.

[22] R. S. Irving, *Projective modules in the category* $\mathcal{O}$, preprint, 1982.

[23] M. Khovanov, *A categorification of the Jones polynomial*, Duke Math. J. **101** (2000), no.3, 359–426; math.QA/9908171. CMP 2000:08

[24] M. Kontsevich, *Homological algebra of mirror symmetry*, Proceedings of the International Congress of Mathematicians (Zürich, 1994), Birkhäuser, 1995, pp. 120–139. MR **97f:**32040

[25] P. Martin, *Potts models and related problems in statistical mechanics,* Series on Advances in Statistical Mechanics, vol.5, 1991, World Scientific. MR **92m:**82030

[26] D. McDuff and D. Salamon, *Introduction to Symplectic Topology*, Oxford University Press, 1995. MR **97b:**58062

[27] J. Moody, *The Burau representation of the braid group* $B_n$ *is unfaithful for large* $n$, Bull. Amer. Math. Soc. **25** (1991), 379–384. MR **92b:**20041

[28] Y.-G. Oh, *Floer cohomology of Lagrangian intersections and pseudo-holomorphic discs I*, Comm. Pure Appl. Math. **46** (1993), 949–994. MR **95d:**58029a

[29] ———, *On the structure of pseudo-holomorphic discs with totally real boundary conditions*, J. Geom. Anal. **7** (1997), 305–327. MR **99g:**58017
[30] ———, *Floer theory for non-compact Lagrangian submanifolds*, Preprint, 2000.
[31] M. Poźniak, *Floer homology, Novikov rings and clean intersections*, Northern California Symplectic Geometry Seminar, 119–181, AMS Transl. Ser. 2, 196, AMS, Providence, RI, 1999. MR **2001a:**53124
[32] J. Robbin and D. Salamon, *The Maslov index for paths*, Topology **32** (1993), 827–844. MR **94i:**58071
[33] R. Rouquier and A. Zimmermann, *Picard groups for derived module categories*, preprint, 1998, http://www.math.jussieu.fr/~rouquier/.
[34] D. Salamon, *Morse theory, the Conley index, and Floer homology*, Bull. London Math. Soc. **22** (1990), 113–140. MR **92a:**58028
[35] D. Salamon and E. Zehnder, *Morse theory for periodic solutions of Hamiltonian systems and the Maslov index*, Comm. Pure Appl. Math. **45** (1992), 1303–1360. MR **93g:**58028
[36] P. Seidel, *Graded Lagrangian submanifolds*, Bull. Soc. Math. France **128** (2000), 103-149; math.SG/9903049. MR **2001c:**53114
[37] ———, *Lagrangian two-spheres can be symplectically knotted*, J. Differential Geom. **52** (1999), 147–173; math.DG/9803083. MR **2001g:**53139
[38] P. Seidel and R. Thomas, *Braid group actions on derived categories of coherent sheaves*, Duke Math. J. **108** (2001), no. 1, 37–108; math.AG/0001043. CMP 2001:12
[39] V. De Silva, *Products in the symplectic Floer homology of Lagrangian intersections*, Ph.D. thesis, Oxford University, 1998.
[40] C. Viterbo, *Intersection des sous-variétés Lagrangiennes, fonctionelles d'action et indice des systèmes Hamiltoniens*, Bull. Soc. Math. France **115** (1987), 361–390. MR **89b:**58081
[41] A. Weinstein, *Lagrangian submanifolds and hamiltonian systems*, Annals of Math. **98** (1973), 377–410. MR **48:**9761
[42] B.W. Westbury, *The representation theory of the Temperley-Lieb algebras,* Math. Z. **219** (1995), no.4, 539–565. MR **96h:**20029

Department of Mathematics, University of California at Davis, Davis, California 95616-8633

*E-mail address*: mikhail@math.ucdavis.edu

Department of Mathematics, Ecole Polytechnique, F-91128 Palaiseau, France – and – School of Mathematics, Institute for Advanced Study, Einstein Drive, Princeton, New Jersey 08540

*E-mail address*: seidel@math.polytechnique.fr

*E-mail address*: seidel@math.ias.edu