GRID DIAGRAMS IN HEEGAARD FLOER THEORY

CIPRIAN MANOLESCU

ABSTRACT. We review the use of grid diagrams in the development of Heegaard Floer theory. We describe the construction of the combinatorial link Floer complex, and the resulting algorithm for unknot detection. We also explain how grid diagrams can be used to show that the Heegaard Floer invariants of 3-manifolds and 4-manifolds are algorithmically computable (mod 2).

1. Introduction

Invariants coming from gauge theory and symplectic geometry have played a major role in the study of low-dimensional manifolds. The topological applications of gauge theory started with the work of Donaldson, who used the Yang-Mills equations to get constraints on the intersection forms of smooth 4-manifolds [Don83]. Counting solutions to the anti-self-dual Yang-Mills equations yielded invariants that were able to distinguish between homeomorphic, but not diffeomorphic, 4-manifolds [Don90]. Floer [Flo88] used the same equations to construct an invariant of 3-manifolds, which became known as instanton Floer homology.

In the 1990’s came the advent of the Seiberg-Witten (or monopole) equations [SW94a, SW94b, Wit94] in four dimensions. These can replace the Yang-Mills equations for most applications, and have better compactness properties. The corresponding monopole Floer homology for 3-manifolds was fully developed by Kronheimer and Mrowka in [KM07]; see also [MW01, Man03, Frø96].

Ozsváth and Szabó [OS04d, OS04c, OS06] developed Heegaard Floer theory as a more computable alternative to Seiberg-Witten theory. Instead of gauge theory, they used pseudo-holomorphic curve counts in symplectic manifolds to define invariants of 3-manifolds, knots, links, and 4-manifolds. In particular, their mixed Heegaard Floer invariants of 4-manifolds are conjecturally the same as the Seiberg-Witten invariants, and can be used for the same applications—for example, to detect exotic smooth structures. In dimension 3, the equivalence between Heegaard Floer theory and Seiberg-Witten theory has recently been established, thanks to work of Kutluhan-Lee-Taubes [KLTa, KLTb, KLTc, KLTd, KLTe] and Colin-Ghiggini-Honda [CGHa, CGHc, CGHd, CGHb].

The definitions of the Yang-Mills, Seiberg-Witten, and Heegaard Floer invariants have in common the use of solution counts to nonlinear elliptic PDE’s. (In the Heegaard Floer case, these are the nonlinear Cauchy-Riemann equations, which define pseudo-holomorphic curves.) Consequently, the invariants above are fundamentally different from more traditional topological invariants such as homology and homotopy groups, Reidemeister torsion, etc. The latter are known to be algorithmically computable—their definitions do not involve analysis. However, the traditional invariants are insufficient to detect the subtle information about 3-manifolds and 4-manifolds that comes from gauge theory or symplectic geometry.

The purpose of this article is to survey recent advances in the field that resulted in combinatorial descriptions for most of the Heegaard Floer invariants. These advances started with an idea of Sarkar, who observed that pseudo-holomorphic curves can be counted explicitly if one uses a certain kind of underlying Heegaard diagram. In the case of knots and links in $S^3$, grid diagrams can be successfully used for this purpose. This allowed the Heegaard Floer invariants of knots and links in $S^3$ to be described combinatorially [MOS09, MOST07]. Further, 3-manifolds and 4-manifolds can be represented in terms of links in $S^3$ via surgery diagrams and Kirby diagrams, respectively. One can

The author was partially supported by NSF grant number DMS-1104406.
then show that the Heegaard Floer invariants (modulo 2) of 3- and 4-manifolds are algorithmically computable, by expressing these invariants in terms of those of the corresponding link [OS08b, MO, MOT].

There are several alternate combinatorial approaches to computing some of the Heegaard Floer invariants. These approaches include nice diagrams [SW10, LMW08, OSS11], cubes of resolutions [OS09, BL], and bordered Floer homology [LOT]. In some cases, the resulting algorithms are more efficient than the ones based on grid diagrams. Nevertheless, grid diagrams are the most encompassing method, and they are the focus of our survey.

Acknowledgements. I owe an intellectual debt to my collaborators Robert Lipshitz, Peter Ozsváth, Sucharit Sarkar, Zoltán Szabó, Dylan Thurston, and Jiajun Wang, who all played an important role in the development of the subject. I would also like to thank Tye Lidman for many helpful expository suggestions.

2. Heegaard Floer homology and related invariants

This section contains a quick overview of Heegaard Floer theory.

The theory started with the work of Ozsváth and Szabó, who in [OS04d, OS04c] introduced a set of 3-manifold invariants in the form of modules over the polynomial ring \( \mathbb{Z}[U] \). Roughly, their construction goes as follows. We represent a closed, oriented 3-manifold \( Y \) by a marked Heegaard diagram, consisting of the following data:

- \( \Sigma \), a closed oriented surface of genus \( g \);
- \( \alpha = \{\alpha_1, \ldots, \alpha_g\} \), a collection of disjoint, homologically linearly independent, simple closed curves on \( \Sigma \). By attaching \( g \) disks to \( \Sigma \) along the curves \( \alpha_i \), and then attaching a 3-ball, we obtain a handlebody \( U_\alpha \) with boundary \( \Sigma \);
- \( \beta = \{\beta_1, \ldots, \beta_g\} \) a similar collection of curves on \( \Sigma \), specifying a handlebody \( U_\beta \);
- a basepoint \( z \in \Sigma - \cup \alpha_i - \cup \beta_i \);

such that \( Y = U_\alpha \cup \Sigma U_\beta \). Any 3-manifold can be represented in this way; see Figure 1 for a schematic picture.

Next, we consider the tori

\[ T_\alpha = \alpha_1 \times \cdots \times \alpha_g, \quad T_\beta = \beta_1 \times \cdots \times \beta_g \]

inside the symmetric product \( \text{Sym}^g(\Sigma) = (\Sigma \times \cdots \times \Sigma)/S_g \), which can be viewed as a symplectic manifold, cf. [Per08]. After small perturbations of the alpha or beta curves, we can assume that \( T_\alpha \) and \( T_\beta \) intersect transversely.

To define the Heegaard Floer complex, we consider the intersection points

\[ x = \{x_1, \ldots, x_g\} \in T_\alpha \cap T_\beta \]
with \(x_i \in \alpha_i \cap \beta_j \subset \Sigma\). Given two such intersection points \(x, y\), a Whitney disk from \(x\) to \(y\) is defined to be a map \(u\) from the unit disk \(D^2\) to \(\text{Sym}^2(\Sigma)\), such that \(u\) maps the lower half of the boundary \(\partial D^2\) to \(T_\alpha\), the upper half to \(T_\beta\), and \(u(-1) = x, u(1) = y\). The space of relative homotopy classes of Whitney disks from \(x\) to \(y\) is denoted \(\pi_2(x, y)\). There is a natural map

\[
T_\alpha \cap T_\beta \rightarrow \text{Spin}^c(Y), \quad x \rightarrow s_z(x)
\]
sending each intersection point to a \(\text{Spin}^c\) structure on \(Y\), such that \(\pi_2(x, y)\) is empty when \(s_z(x) \neq s_z(y)\).

For every \(\phi \in \pi_2(x, y)\), we can consider the moduli space of pseudo-holomorphic disks in the class \(\phi\). These disks are solutions to the nonlinear Cauchy-Riemann equations, which depend on the choice of a family \(J = \{J_t\}_{t \in [0, 1]}\) of almost complex structures on the symmetric product. The moduli space has an expected dimension \(\mu(\phi)\) (the Maslov index), and it comes with a natural action of \(\mathbb{R}\) by automorphisms of the domain. If \(\mu(\phi) = 1\), then for generic \(J\), after dividing by the \(\mathbb{R}\) action the moduli space becomes just a finite set of points, which can be counted with certain signs. We let \(c(\phi, J) \in \mathbb{Z}\) denote the resulting count. Moreover, we denote by \(n_2(\phi)\) the intersection number between \(\phi\) and the divisor \(\{s\} \times \text{Sym}^{g-1}(\Sigma)\). If \(\phi\) admits any pseudo-holomorphic representatives, the principle of positivity of intersections for holomorphic maps implies that \(n_2(\phi) \geq 0\).

Fix a \(\text{Spin}^c\) structure \(s\) on \(Y\). Under a certain assumption on the Heegaard diagram (admissibility), one defines the Heegaard Floer complex \(CF^-(Y, s)\) to be the \(\mathbb{Z}[U]\)-module freely generated by intersection points \(x \in T_\alpha \cap T_\beta\) with \(s_z(x) = s\). The differential on \(CF^-(Y, s)\) is given by the formula:

\[
\partial x = \sum_{\{y \in T_\alpha \cap T_\beta | s_z(y) = s\}} \sum_{\{\phi \in \pi_2(x, y) | \mu(\phi) = 1\}} c(\phi, J) \cdot U^{n_2(\phi)} y.
\]

It can be shown that \(\partial^2 = 0\). Starting from the complex \(CF^- = CF^-(Y, s)\) one defines the various versions of Heegaard Floer homology to be:

\[
HF^- = H_*(CF^-), \quad HF^+ = H_*(U^{-1}CF^-/CF^-),
\]

\[
HF^\infty = H_*(U^{-1}CF^-), \quad \widehat{HF} = H_*(CF^-/(U = 0)).
\]

Although the complex \(CF^-\) depends on the choices of Heegaard diagram and almost complex structure, the homology groups do not:

**Theorem 2.1** ([OS04d]). The Heegaard Floer homology modules \(HF^-(Y, s), HF^+(Y, s), HF^\infty(Y, s), \widehat{HF}(Y, s)\) are invariants of the 3-manifold \(Y\) equipped with the \(\text{Spin}^c\) structure \(s\).

If \(\circ\) is any of the four flavors of Heegaard Floer homology \((-\), \(+\), \(\infty\), or \(\wedge\)), we will denote by \(HF^\circ(Y)\) the direct sum of \(HF^\circ(Y, s)\) over all \(\text{Spin}^c\) structures \(s\).

Heegaard Floer homology has found numerous applications to 3-dimensional topology. Among them we mention: detection of the Thurston norm [OS04a], detection of whether the 3-manifold fibers over the circle [Ni09], and a characterization of which Seifert fibrations admit tight contact structures [LS09].

The reader may wonder about the differences between the different variants of Heegaard Floer homology. The full power of the theory comes from the versions \(HF^-\) and \(HF^+\), which contain (roughly) equivalent information. The version \(\widehat{HF}\) is weaker: it suffices for many 3-dimensional applications, but it does not give any non-trivial information about closed 4-manifolds. (As explained below, the mixed invariants of 4-manifolds are constructed by combining the plus and minus theories.) The version \(HF^\infty\) is the least useful, being determined by classical topological information, at least when we work modulo 2 and we restrict to torsion \(\text{Spin}^c\) structures; see [Lid].

As proved by Ozsváth and Szabó in [OS06], the four variants of Heegaard Floer homology are functorial under \(\text{Spin}^c\)-decorated cobordisms. Precisely, given a 4-dimensional manifold \(W\) with
\( \partial W = (-Y_0) \cup Y_1 \), and a Spin\(^c\) structure \( s \) on \( W \) with restrictions \( s_0 \) to \( Y_0 \) and \( s_1 \) to \( Y_1 \), there are induced maps:
\[
F^0_{W,s} : HF^0(Y_0,s_0) \to HF^0(Y_1,s_1),
\]
where \( \circ \) can stand for any of the four flavors of Heegaard Floer homology. The maps \( F^0_{W,s} \) are defined by counting \emph{pseudo-holomorphic triangles} in the symmetric product, with boundaries on three different tori.

Suppose we have a smooth, closed, oriented 4-manifold \( X \) with \( b^+_2(X) > 1 \). An \emph{admissible cut} for \( X \) is a smoothly embedded 3-manifold \( N \subset X \) which divides \( X \) into two pieces \( X_1 \) and \( X_2 \) with \( b^+_2(X_i) > 0 \) for \( i = 1, 2 \), and such that \( \delta H^1(N; \mathbb{Z}) = 0 \subset H^2(X; \mathbb{Z}) \).

An admissible cut can be found for any \( X \). Given an admissible cut, one can delete a four-ball from the interior of each piece \( X_i \) to obtain two cobordisms \( W_1 \) (from \( S^3 \) to \( N \)) and \( W_2 \) (from \( N \) to \( S^3 \)). Let \( s \) be a Spin\(^c\) structure on \( X \). We consider the following diagram:

\[
\begin{array}{c}
HF^-(S^3) \xrightarrow{F^+_{W_1,s|W_1}} HF^-(N,s|N) \\
\downarrow \hspace{5cm} \downarrow \\
HF_{\text{red}}(N,s|N) \xrightarrow{F^+_{W_2,s|W_2}} HF^+(S^3),
\end{array}
\]

where \( HF_{\text{red}} \) is the image of a natural map \( HF^+ \to HF^- \). (This map exists for any 3-manifold.) The conditions in the definition of an admissible cut ensure that the maps \( F^+_{W_1,s|W_1} \) and \( F^+_{W_2,s|W_2} \) factor through \( HF_{\text{red}} \). The composition of the two lifts (indicated by dashed arrows) is called a mixed map. The image of 1 in \( HF^-(S^3) \cong \mathbb{Z} \) under this map defines the \emph{Ozsváth-Szabó mixed invariant} of the pair \((X,s)\). This is conjecturally equivalent to the well-known Seiberg-Witten invariant [Wit94], and is known to share many of its properties. In particular, it can be used to distinguish homeomorphic 4-manifolds that are not diffeomorphic.

3. Link Floer homology

Ozsváth-Szabó [OS04b] and, independently, Rasmussen [Ras03] used Heegaard Floer theory to define invariants for knots in 3-manifolds: these are the various versions of \emph{knot Floer homology}.

Recall that a marked Heegaard diagram \((\Sigma, \alpha, \beta, z)\) represents a 3-manifold \( Y \). If one specifies another basepoint \( w \) in the complement of the alpha and beta curves, this gives rise to a knot \( K \subset Y \). Indeed, one can join \( w \) to \( z \) by a path in \( \Sigma \setminus \cup \alpha_i \), and then push this path into the interior of the handlebody \( U_\alpha \). Similarly, one can join \( w \) to \( z \) in the complement of the beta curves, and push the path into \( U_\beta \). The union of these two paths is the knot \( K \). For simplicity, we will assume that \( K \) is null-homologous. (Of course, this happens automatically if \( Y = S^3 \).)

In the definition of the Heegaard Floer complex \( CF^- \) we kept track of intersections with \( z \) through the exponent \( n_z(\phi) \) of the variable \( U \). Now that we have two basepoints, we have two quantities \( n_z(\phi) \) and \( n_w(\phi) \). One thing we can do is to count only disks in classes with \( n_z(\phi) = 0 \), and keep track of \( n_w(\phi) \) in the exponent of \( U \). The result is a complex of \( \mathbb{Z}[U] \)-modules denoted \( CFK^{-}(Y,K) \), with homology \( HF^{-}(Y,K) \). If we set the variable \( U \) to zero, we get a complex \( \widehat{CFK}(Y,K) \), with homology \( \widehat{HFK}(Y,K) \). These are two of the variants of knot Floer homology. There exist many other variants, some of which involve classes \( \phi \) with \( n_z(\phi) \neq 0 \) and \( n_w(\phi) \neq 0 \); an example of this, denoted \( \mathfrak{A}^\pm(K) \), will be mentioned in section 5.
Let us focus on the case when $Y = S^3$. The groups $\hat{HF}(S^3, K)$ naturally split as direct sums:

$$\hat{HF}(S^3, K) = \bigoplus_{m, s \in \mathbb{Z}} \hat{HF}_m(S^3, K, s).$$

Here, $m$ and $s$ are certain quantities called the Maslov and Alexander gradings, respectively. We can encode some of the information in $\hat{HF}$ into a polynomial

$$\mathcal{P}_K(t, q) = \sum_{m, s \in \mathbb{Z}} t^m q^s \cdot \text{rank } \hat{HF}_m(S^3, K, s).$$

The specialization $\mathcal{P}_K(-1, q)$ is the classical Alexander polynomial of $K$. However, the applications of knot Floer homology go well beyond those of the Alexander polynomial. In particular, the genus of the knot, which is defined as

$$g(K) = \min \{ g \mid \exists \text{ embedded, oriented surface } \Sigma \subset S^3, \partial \Sigma = K \},$$

can be read from $\hat{HF}$:

**Theorem 3.1** ([OS04a]). For any knot $K \subset S^3$, we have

$$g(K) = \max \{ s \geq 0 \mid \exists m, \hat{HF}_m(S^3, K, s) \neq 0 \}.$$

Since the only knot of genus zero is the unknot, we have:

**Corollary 3.2.** $K$ is the unknot if and only if $\mathcal{P}_K(q, t) = 1$.

By a result of Ghiggini [Ghi08], the polynomial $\mathcal{P}$ has enough information to also detect the right-handed trefoil, the left-handed trefoil, and the figure-eight knot. Ni [Ni07] extended the work in [Ghi08] to show that $S^3 \setminus K$ fibers over the circle if and only if $\oplus_m \hat{HF}_m(S^3, K, g(K)) \cong \mathbb{Z}$. Other applications of knot Floer homology include the construction of a concordance invariant called $\tau$ [OS03], and a complete characterization of which lens spaces can be obtained by surgery on knots [Gre].

If instead of a knot $K \subset S^3$ we have a link $L$ (a disjoint union of knots), we can define invariants $\hat{HFL}(S^3, L), \hat{HFL}^-(S^3, L)$, which are versions of link Floer homology. When $L$ is a knot, $\hat{HFL}$ and $\hat{HFL}^-$ reduce to $\hat{HF}$ and $\hat{HF}^-$, respectively. In general, the definition of link Floer homology involves choosing a new kind of Heegaard diagram for $S^3$, in which the number of alpha (or beta) curves exceeds the genus of the Heegaard surface. The details can be found in [OS08a]. If the diagram has $g + k - 1$ alpha curves, it should also have $g + k - 1$ beta curves, $k$ basepoints of type $z$, and $k$ basepoints of type $w$. In the simplest version, $k$ is the same as the number $\ell$ of components of the link, and joining the $w$ basepoints to the $z$ basepoints in pairs (by a total of $2\ell$ paths) produces the link $L$. Instead of $\mathbb{Z}[U]$, the link Floer complexes are defined over a polynomial ring $\mathbb{Z}[U_1, \ldots, U_\ell]$, with one variable for each component. (In fact, we expect the complexes to be defined over $\mathbb{Z}[U_1, \ldots, U_\ell]$. However, at the moment some orientation issues are not yet settled, and the theory is only defined with mod 2 coefficients.)

More generally, we could have $k \geq \ell$, and break the link into more segments. We can then define a link Floer complex over $\mathbb{Z}/2[U_1, \ldots, U_k]$, with one variable for each basepoint; see [MOS09, MO]. The homology of this complex is still $\hat{HFL}^-$, and all the variables corresponding to basepoints on the same link component act the same way. If we set one $U_i$ variable from each link component to zero in the complex, the resulting homology is $\hat{HFL}$. If we set all the $U_i$ variables to zero in the complex, the homology becomes

$$\hat{HFL}(S^3, L) \otimes V^{k-\ell},$$

where $V$ is a two-dimensional vector space over $\mathbb{Z}/2$. 
Definition 4.1. Let $L \subset S^3$ be a link. A grid diagram for $L$ consists of an $n$-by-$n$ grid in the plane with $O$ and $X$ markings inside, such that:

1. Each row and each column contains exactly one $X$ and one $O$;
2. As we trace the vertical and horizontal segments between $O$'s and $X$'s (with the vertical segments passing over the horizontal segments), we see a planar diagram for the link $L$.

An example is shown on the left hand side of Figure 2. It is not hard to see that every link admits a grid diagram. In fact, as a way of representing links, grid diagrams are equivalent to arc presentations, which originated in the work of Brunn [Bru98]. The minimal number $n$ such that $L$ admits a grid diagram of size $n$ is called the arc index of $L$.

Grid diagrams can be viewed as particular examples of Heegaard diagrams with multiple basepoints, of the kind discussed at the end of the previous section. Indeed, if we identify the opposite sides of a grid diagram $G$ to get a torus, we can let this torus be the Heegaard surface $\Sigma$, the horizontal circles be the $\alpha$ curves, the vertical circles be the $\beta$ curves, the $O$ markings be the $w$ basepoints, and the $X$ markings be the $z$ basepoints. A point $x = \{x_1, \ldots, x_n\}$ in the intersection $T_\alpha \cap T_\beta$ consists an $n$-tuple of points on the grid (one on each vertical and horizontal circle). There are $n!$ such intersection points, and they are precisely the generators of the link Floer complex. We denote the set of these generators by $S(G)$.

Definition 4.2. Let $G$ be a grid diagram, and $x, y \in S(G)$. We define a rectangle from $x$ to $y$ to be a rectangle on the grid torus with the lower left and upper right corner being points of $x$, the lower right and upper right corners being points of $y$, and such that all the other components of $x$ and $y$ coincide. (In particular, for such a rectangle to exist we need $x$ to differ from $y$ in exactly two rows.) A rectangle is called empty if it contains no components of $x$ or $y$ in its interior. The set of empty rectangles from $x$ to $y$ is denoted $\text{Rect}^\circ(x, y)$.

Of course, the space $\text{Rect}^\circ(x, y)$ has at most two elements. An example where it has exactly two is shown on the right hand side of Figure 2.

The reason why grid diagrams are useful in Heegaard Floer theory is that they make pseudoholomorphic disks of Maslov index 1 easy to count:

Proposition 4.3 ([MOS09]). Let $G$ be a grid diagram, and let $x, y \in S(G)$. Then, there is a 1-to-1 correspondence:

$$\left\{ \phi \in \pi_2(x, y) \mid \mu(\phi) = 1, \ c(\phi, J) \equiv 1(\mod \ 2) \right\} \longleftrightarrow \text{Rect}^\circ(x, y).$$
Sketch of proof. In any Heegaard diagram, if we have a relative homotopy class \( \phi \in \pi_2(x, y) \), we can associate to it a two-chain \( D(\phi) \) on the Heegaard surface \( \Sigma \), as follows. Let \( n \) be the number of alpha (or beta) curves. Together, the alpha and the beta curves split \( \Sigma \) into several connected regions \( R_1, \ldots, R_m \). For each \( i \), let us pick a point \( p_i \) in the interior of \( R_i \), and define the multiplicity of \( \phi \) at \( R_i \) to be the intersection number \( n_{p_i}(\phi) \) between \( \phi \) and \( \{p_i\} \times \text{Sym}^{n-1}(\phi) \). We set

\[
D(\phi) = \sum_i n_{p_i}(\phi)R_i.
\]

This is called the domain of \( \phi \). If \( \phi \) admits any pseudo-holomorphic representatives, then the multiplicities \( n_{p_i}(\phi) \) must be nonnegative.

Lipshitz [Lip06] showed that the Maslov index of \( \phi \) can be expressed in terms of the domain:

\[
\mu(\phi) = e(D(\phi)) + N(D(\phi)),
\]

where \( e \) and \( N \) are certain quantities called the Euler measure and total vertex multiplicity, respectively. The Euler measure is additive on regions, that is, we can define \( e(R_i) \) such that \( e(D(\phi)) = \sum_i n_{p_i}(\phi)e(R_i) \). If we take the sum \( \sum_i e(R_i) \) we get the Euler characteristic of the Heegaard surface \( \Sigma \). As for the total vertex multiplicity, it is the sum of 2\( n \) vertex multiplicities \( N_\phi(D(\phi)) \), one for each point \( q \in x \) or \( q \in y \). The quantity \( N_\phi(D(\phi)) \) is the average of the multiplicities of \( \phi \) in the four quadrants around \( q \).

In the case of a grid \( G \), the regions \( R_i \) are the \( n^2 \) unit squares of \( G \). Each square has Euler measure zero. If we have \( \phi \in \pi_2(x, y) \) with \( \mu(\phi) = 1 \) and \( e(\phi, J) \neq 0 \), then the coefficients of \( R_i \) in \( D(\phi) \) are nonnegative. This implies that \( 1 = \mu(\phi) = N(D(\phi)) \) is a sum of vertex multiplicities \( N_\phi(D(\phi)) \). Each \( N_\phi(D(\phi)) \) is either zero or at least \( 1/4 \). A short analysis shows that \( D(\phi) \) must be an empty rectangle.

Conversely, given an empty rectangle, there is a corresponding class \( \phi \) with \( \mu(\phi) = 1 \). An application of the Riemann mapping theorem shows that \( \phi \) has an odd number of pseudo-holomorphic representatives for generic \( J \). \( \square \)

In view of Proposition 4.3, the link Floer complex associated to a grid can be defined in a purely combinatorial way. Precisely, we define \( C^-(G) \) to be freely generated by \( S(G) \) over the ring \( \mathbb{Z}/2[U_1, \ldots, U_n] \), with differential:

\[
\partial x = \sum_{y \in S(G)} \sum_{r \in \text{Rect}^\circ(x, y) \setminus X_i(r) = 0, \forall i} U_1^{O_1(r)} \ldots U_n^{O_n(r)} \cdot y.
\]

Here, \( O_i(r) \) encodes whether or not the \( i \)-th marking of type \( O \) is in the interior of the rectangle \( r \): if it is, we set \( O_i(r) \) to be 1; otherwise it is 0. The quantity \( X_i(r) \in \{0, 1\} \) is defined similarly, in terms of the \( i \)-th marking of type \( X \).

The homology of \( C^-(G) \) is the link Floer homology \( HF^-(S^3, L) \).

Remark 4.4. Although the complex \( C^-(G) \) is defined with mod 2 coefficients, one can add signs in the differential to get a complex over \( \mathbb{Z}[U_1, \ldots, U_n] \), whose homology is still a link invariant. See [MOST07, Gal08].

If in \( C^-(G) \) we set one variable \( U_i \) from each link component to zero, we get a complex \( \tilde{C}(G) \) with homology \( \tilde{HF}(S^3, L) \). Perhaps the simplest complex is

\[
\tilde{C}(G) = C^-(G)/(U_1 = U_2 = \ldots = U_n = 0),
\]

for which we only need to count the empty rectangles with no markings of any type in their interior. The homology of \( \tilde{C}(G) \) is \( \tilde{HF}(S^3, L) \otimes V^{n-\ell} \); compare (1).

In particular, when \( L = K \) is a knot, the homology of \( \tilde{C}(G) \) is \( \tilde{HF}(S^3, K) \otimes V^{n-1} \). There exist simple combinatorial formulas for the Maslov and Alexander gradings of the generators in \( S(G) \), and from them one gets a bi-grading on \( H_*(\tilde{C}(G)) \); see [MOS09], [MOST07]. From here one can
recover \( \widehat{HFK}(S^3, K) \) as a bi-graded group, taking into account that each \( V \) factor is spanned by one generator in bi-degree \((0,0)\) and another in bi-degree \((-1,-1)\). This method of calculating \( \widehat{HFK} \) was implemented on the computer by Baldwin and Gillam [BG]; see also Droz [Dro08] for a more efficient program, using a variation of this method due to Beliakova [Bel10].

In view of Theorem 3.1, we see that grid diagrams yield an algorithm for detecting the genus of a knot. In particular, if we are given a knot diagram and want to see if it represents the unknot, we can turn it into a grid diagram (after some suitable isotopies), then set up the complex \( \tilde{C}(G) \), and check if \( H_*(\tilde{C}(G)) \cong V^{n-1} \).

Among the other applications of the grid diagram method we mention one to contact geometry: the construction of new invariants for Legendrian and transverse knots in \( S^3 \) [OST08]. This led to numerous examples of transverse knots with the same self-linking number that are not transversely isotopic [NOT08].

Another application of knot Floer homology via grid diagrams is Sarkar’s combinatorial proof of the Milnor conjecture [Sar]. The Milnor conjecture states that the slice genus of the torus knot \( T(p, q) \) is \((p-1)(q-1)/2\); a corollary is that the minimum number of crossing changes needed to turn \( T(p, q) \) into the unknot is also \((p-1)(q-1)/2\). The conjecture was first proved by Kronheimer and Mrowka using gauge theory [KM93]; for other proofs, see [OS03], [Ras10].

Slight variations of grid diagrams can be used to compute the knot Floer homology of knots inside lens spaces [BGH08], and of a knot inside its cyclic branched covers [Lev08].

Finally, we mention that there exists a purely combinatorial proof that \( H_*(C^-(G)) \) and \( H_*(\tilde{C}(G)) \) are link invariants [MOST07].

5. THREE-MANIFOLDS AND FOUR-MANIFOLDS

For a general Heegaard diagram, counting pseudo-holomorphic disks in the symmetric product is very difficult. Why is it easy for a grid diagram? If we look at the proof of Proposition 4.3, a key point we find is that the regions \( R_i \) have zero Euler measure. In fact, what is important is that they have \textit{nonnegative} Euler measure: since the total vertex multiplicity is always nonnegative, the fact that \( e(D(\phi)) + N(D(\phi)) = 1 \) imposes tight constraints on the possibilities for \( D(\phi) \).

In general, if a Heegaard surface \( \Sigma \) can be partitioned into regions of nonnegative Euler measure, its Euler characteristic (which is the sum of all the Euler measures) must be nonnegative; that is, \( \Sigma \) must be a sphere or a torus. Our grid diagrams were set on a torus. There is also a variant on the sphere, that produces another combinatorial link Floer complex, and in the end yields the same homology.

Instead of a knot in \( S^3 \), we could take a 3-manifold \( Y \) and try to compute its Heegaard Floer homology using this method. The problem is that a typical 3-manifold does not admit a Heegaard diagram of genus 0 or 1; only \( S^3 \), \( S^1 \times S^2 \) and lens spaces do. However, Sarkar and Wang [SW10] proved that one can find a Heegaard diagram for \( Y \), called a \textit{nice diagram}, in which all regions except one have nonnegative Euler measure. (This is related to the fact that on a surface of higher genus we can move all negative curvature to a neighborhood of a point.) If we put the basepoint \( z \) in the bad region (the one with negative Euler measure), then we can understand pseudo-holomorphic curve counts for all classes \( \phi \) with \( n_\pm(\phi) = 0 \). These are the only classes that appear when defining the complex \( \tilde{CF}(Y) \). Thus, we get an algorithm for computing \( \widehat{HF} \) of any 3-manifold. We refer to [SW10] for more details, and to [OSS11, OSS] for related work.

Similarly, one can compute the cobordism maps \( \tilde{F}_{W} \) for any simply connected \( W \) [LMW08]. These suffice to detect exotic smooth structures on some 4-manifolds with boundary, but not on any closed 4-manifolds.

This line of thought runs into major difficulties if one wants to understand combinatorially the plus and minus versions of \( HF \), or the mixed invariants of 4-manifolds. Instead, what is helpful is
to reduce everything to the case of links in $S^3$, and then appeal to grid diagrams. This program was developed in [MO, MOT], and is summarized below.

Let us recall a theorem of Lickorish and Wallace [Lic62, Wal60], which says that any closed 3-manifold $Y$ is integral surgery on a link in $S^3$:

$$Y = (S^3 \setminus \nu(L)) \cup_\phi (\nu(L)).$$

Here, $\nu(L)$ is a tubular neighborhood, and $\phi$ is a self-diffeomorphism of $\partial \nu(L)$. The diffeomorphism can be specified in terms of a framing of the link, which in turn is determined by choosing one integer for each link component. For example, the Poincaré sphere is surgery on the right-handed trefoil with $+1$ framing. In general, we denote by $S^3_\Lambda(L)$ the result of surgery on $L$ with framing $\Lambda$.

Four-manifolds can also be expressed in terms of links, using Kirby diagrams [Kir78]. By Morse theory, a closed 4-manifold can be broken into a 0-handle, some 1-handles (represented in a Kirby diagram by circles marked with a dot), some 2-handles (represented by framed knots), some 3-handles, and a 4-handle. The positions of the 1-handles and 2-handles determine the manifold. See Figure 3 for a few examples.

The next step in the program is to express the Heegaard Floer homology of surgery on a link in terms of data associated to the link. The first result in this direction was obtained by Ozsváth and Szabó [OS08b], who dealt with surgery on knots:

**Theorem 5.1** ([OS08b]). There is an (infinitely generated) version of the knot Floer complex, $\mathfrak{A}^+(K)$, such that

$$HF^+(S^3_\Lambda(K)) = H_*(\text{Cone}(\mathfrak{A}^+(K) \xrightarrow{\Phi_n^K} \mathfrak{A}^+(\emptyset)))$$

where in $\mathfrak{A}^+(K), \mathfrak{A}^+(\emptyset)$ we count pseudo-holomorphic bigons and in $\Phi_n^K$ we count pseudo-holomorphic triangles.

The complex $\mathfrak{A}^+(\emptyset)$ is a direct sum of infinitely many copies of $CF^+(S^3)$. The inclusion of one of these copies into the mapping cone complex

$$\text{Cone}(\mathfrak{A}^+(K) \xrightarrow{\Phi_n^K} \mathfrak{A}^+(\emptyset))$$

induces on homology the map $F_{W,\mathfrak{s}}^- : HF^+(S^3) \to HF^+(S^3_\Lambda(K))$ corresponding to the surgery cobordism (2-handle attachment along $K$), equipped with a Spin$^c$ structure $\mathfrak{s}$.

The proof of Theorem 5.1 is based on an important property of $HF^+$ called the surgery exact triangle. The version $HF^-$ does not have a similar exact triangle, but a slight variant of it, $HF^-$ does. The version $HF^-$ is obtained from $HF^-$ by completion with respect to the $U$ variable. For torsion Spin$^c$ structures $\mathfrak{s}$, one can recover $HF^-(Y,\mathfrak{s})$ from $HF^-(Y,\mathfrak{s})$, so in that case the two versions contain equivalent information.

There is an analogue of Theorem 5.1 with $HF^-$ instead of $HF^+$, and with a knot Floer complex denoted $\mathfrak{A}^-$ instead of $\mathfrak{A}^+$.

There is also an extension of Theorem 5.1 to surgeries on links rather than single knots. Phrased in terms of $HF^-$, it reads:
Theorem 5.2 ([MO]). If \( L = K_1 \cup K_2 \subset S^3 \) is a link with framing \( \Lambda \), then \( \text{HF}^-(S^3_\Lambda(L)) \) is isomorphic to the homology of a complex \( \mathcal{C}^-(L, \Lambda) \) of the form
\[
\begin{array}{c}
\mathfrak{A}^-(L) \quad \mathfrak{A}^-(K_1) \\
\downarrow \quad \downarrow \\
\mathfrak{A}^-(K_2) \quad \mathfrak{A}^-(\emptyset)
\end{array}
\]
where the edge maps count holomorphic triangles, and the diagonal map counts holomorphic quadrilaterals.

This can be generalized to links with any number of components. The higher diagonals involve counting higher holomorphic polygons. Further, the inclusion of the subcomplex corresponding to \( L' \subseteq L \) corresponds to the cobordism maps given by surgery on \( L - L' \).

Remark 5.3. For technical reasons, at the moment Theorem 5.2 is only established with mod 2 coefficients.

If \( W \) is a cobordism between (connected) 3-manifolds that consists of 2-handles only, then we can express one boundary piece of \( W \) as surgery on a link \( L' \subset S^3 \), and \( W \) as a handle attachment along a link \( L - L' \). Thus, Theorem 5.2 gives a description of the maps on \( \text{HF}^- \) associated to any such cobordism \( W \). In fact, 2-handles are the main source of complexity in 4-manifolds. Once we understand them, it is not hard to incorporate the maps induced by 1-handles and 3-handles into the picture. The result is a description of the Ozsváth-Szabó mixed invariant of a 4-manifold in terms of link Floer complexes. For this one needs to represent \( X \) by a slight variant of a Kirby diagram, called a cut link presentation; we refer to [MO] for more details.

Theorem 5.4 ([MOT]). Given any 3-manifold \( Y \) with a Spin\(^c\) structure \( s \), the Heegaard Floer homologies \( \text{HF}^+(Y, s) \) and \( \text{HF}^-(Y, s) \) (with mod 2 coefficients) are algorithmically computable. So are the mixed invariants \( \Psi_{X, s} (\text{mod} \; 2) \) for closed 4-manifolds \( X \) with \( b^+_2(X) > 1 \) and \( s \in \text{Spin}^c(X) \).

Sketch of proof. We can represent the 3-manifold or the 4-manifold in terms of a link, as above (by a surgery diagram or a cut link presentation). The idea is then to take a grid diagram \( G \) for the link, and apply Theorem 5.2. We know that index 1 holomorphic disks (bigons) on the symmetric product of the grid correspond to empty rectangles. However, to apply Theorem 5.2 we also need to be able to count higher pseudo-holomorphic polygons. In [MOT], it is shown that isolated pseudo-holomorphic triangles on the symmetric product are in 1-to-1 correspondence with domains on the grid of certain shapes, as shown in Figure 4.

No such easy description is available for counts of pseudo-holomorphic \( m \)-gons on \( \text{Sym}^n(G) \) with \( m \geq 4 \). The trouble is that, unlike for \( m = 2 \) or 3, the counts for \( m \geq 4 \) depend on the choice of a generic family \( J \) of almost complex structures on \( \text{Sym}^n(G) \). Still, the counts are required to satisfy certain constraints, coming from positivity of intersections and Gromov compactness. We define a formal complex structure on \( G \) to be any count of domains on the grid that satisfies these constraints.

A formal complex structure is a purely combinatorial object. Each such structure \( c \) gives rise to a complex \( \mathcal{C}^-(G, \Lambda, c) \), similar to (2), but where instead of pseudo-holomorphic polygon counts we use the domain counts prescribed by \( c \). In particular, a family of almost complex structures \( J \) on the symmetric product produces a formal complex structure, whose corresponding complex is exactly (2). There is a definition of homotopy between formal complex structures, and if two such structures are homotopic, they give rise to quasi-isomorphic complexes \( \mathcal{C}^-(G, \Lambda, c) \).

We conjecture that any two formal complex structures on a grid diagram are homotopic. A weaker form of this conjecture, sufficient for our purposes, is proved in [MOT]. Instead of an ordinary grid diagram \( G \), we use its sparse double \( G_{\#} \). This is obtained from \( G \) by introducing \( n \) additional rows, columns, and \( O \) markings, interspersed between the previous rows and columns,
Figure 4. Snail domains. Darker shading corresponds to higher local multiplicities. The domains in each row (top or bottom) are part of an infinite sequence, corresponding to increasing complexities. The larger circles represent certain fixed points on the grid, called destabilization points. Each domain corresponds to a pseudo-holomorphic triangle in the symmetric product of the grid.

Figure 5. The sparse double of the grid diagram from Figure 2.

as shown in Figure 5. The sparse double is not a grid diagram in the usual sense, because the new rows and columns have no $X$ markings. Nevertheless, it can still be viewed as a type of Heegaard diagram for the link, and pseudo-holomorphic bigons and triangles correspond to empty rectangles and snail domains, just as before. One result of [MOT] is that on the sparse double, any two (sparse) formal complex structures are homotopic.

With this in mind, the desired algorithm for computing $\text{HF}^-$ is as follows: Choose any formal complex structure on $G_\#$, and then calculate the homology of $\mathcal{C}^-(G, \Lambda, c)$. This homology is independent of $c$, so it agrees with the homology of the complex (2). By Theorem 5.2, this gives exactly $\text{HF}^-$ of surgery on the framed link. Similar algorithms can be constructed for computing $\text{HF}^+$ and $\Phi_{X,s}$. □
6. Open problems

6.1. Develop more efficient algorithms. A weakness of the grid diagram approach is that the size of the combinatorial knot Floer complex increases super-exponentially (like \(n!\)) with respect to the size of the grid. Nevertheless, in practice, computer programs [BG, Dro08] can calculate knot Floer homology (for knots and links in \(S^3\)) from diagrams of grid number up to 13. The algorithms become much less effective for 3-manifolds, and especially for 4-manifolds: this is because, for example, representing the \(K3\) surface requires a grid of size at least 88.

A related open problem is to decide whether the unknotting problem can be solved in polynomial time.

6.2. Combinatorial proofs. To completely set the theory in elementary terms, it remains to give purely combinatorial proofs that the Heegaard Floer invariants are indeed invariants of the underlying object. For link Floer homology, this was achieved in [MOST07]. For \(\widehat{HF}\) of 3-manifolds (defined from a class of diagrams called convenient, rather than from surgery formulas), a combinatorial proof of invariance appeared in [OSS]. The cases of the other versions of \(HF\), and of the mixed invariants of 4-manifolds, remain open.

Also missing are combinatorial proofs for most of the topological properties of Heegaard Floer theory. For example, it is not known how to prove combinatorially that knot Floer homology detects the genus of a knot.

6.3. Loose ends. In terms of showing algorithmic computability, there are a few aspects of the theory that are not taken care of by Theorem 5.4:

- **Signs.** Extend the combinatorial descriptions to the invariants defined over \(\mathbb{Z}\) (rather than over \(\mathbb{F} = \mathbb{Z}/2\mathbb{Z}\)).
- **Cobordism maps.** One can understand combinatorially the maps on Heegaard Floer homology induced by 2-handle cobordisms, and the mixed map for closed 4-manifolds, but not yet the maps induced by a general cobordism between 3-manifolds (which may include 1- and 3-handles).
- **The uncompleted \(HF^-\) and \(HF^\infty\).** Theorem 5.4 is about \(HF^-\) rather than \(HF^\infty\). For torsion \(\text{Spin}^c\) structures, knowledge of \(HF^-\) determines \(HF^-\). For nontorsion \(\text{Spin}^c\) structures, understanding \(HF^-\) is basically equivalent to understanding \(HF^-\) and \(HF^\infty\); the latter group has not yet been computed.

6.4. Other open problems in Heegaard Floer theory. As mentioned in the introduction, in dimension 3, the Heegaard-Floer and Seiberg-Witten Floer homologies are known to be isomorphic. In dimension 4, it is still open to prove that the mixed Ozsváth-Szabó invariant of 4-manifolds is the same as the Seiberg-Witten invariant.

Another important question is to understand the relationship of Heegaard Floer theory to Yang-Mills theory, and to the fundamental group of a 3-manifold.

References


Lenhard Ng, Peter Ozsváth, and Dylan Thurston, *Transverse knots distinguished by knot Floer homology*, J. Symplectic Geom. 6 (2008), no. 4, 461–490.


Lenhard Ng, Peter Ozsváth, and Dylan Thurston, *Transverse knots distinguished by knot Floer homology*, J. Symplectic Geom. 6 (2008), no. 4, 461–490.


Lenhard Ng, Peter Ozsváth, and Dylan Thurston, *Transverse knots distinguished by knot Floer homology*, J. Symplectic Geom. 6 (2008), no. 4, 461–490.


Lenhard Ng, Peter Ozsváth, and Dylan Thurston, *Transverse knots distinguished by knot Floer homology*, J. Symplectic Geom. 6 (2008), no. 4, 461–490.


Lenhard Ng, Peter Ozsváth, and Dylan Thurston, *Transverse knots distinguished by knot Floer homology*, J. Symplectic Geom. 6 (2008), no. 4, 461–490.


A cube of resolutions for knot Floer homology, J. Topol. 2 (2009), no. 4, 865–910.

Combintorial Heegaard Floer homology and nice Heegaard diagrams, Preprint (2009), arXiv:0912.0830.


Department of Mathematics, UCLA, 520 Portola Plaza, Los Angeles, CA 90095, USA

E-mail address: cm@math.ucla.edu