Khovanov homology is an unknot-detector

P. B. Kronheimer and T. S. Mrowka

Harvard University, Cambridge MA 02138
Massachusetts Institute of Technology, Cambridge MA 02139

Abstract. We prove that a knot is the unknot if and only if its reduced Khovanov cohomology has rank 1. The proof has two steps. We show first that there is a spectral sequence beginning with the reduced Khovanov cohomology and abutting to a knot homology defined using singular instantons. We then show that the latter homology is isomorphic to the instanton Floer homology of the sutured knot complement: an invariant that is already known to detect the unknot.

1 Introduction

1.1 Statement of results

This paper explores a relationship between the Khovanov cohomology of a knot, as defined in [16], and various homology theories defined using Yang-Mills instantons, of which the archetype is Floer's instanton homology of 3-manifolds [8]. A consequence of this relationship is a proof that Khovanov cohomology detects the unknot. (For related results, see [10, 11, 12]).

Theorem 1.1. *A knot in S^3 is the unknot if and only if its reduced Khovanov cohomology is $\mathbb{Z}$.*

In [23], the authors construct a Floer homology for knots and links in 3-manifolds using moduli spaces of connections with singularities in codimension 2. (The locus of the singularity is essentially the link K, or $\mathbb{R} \times K$ in a cylindrical 4-manifold.) Several variations of this construction are already considered in [23], but we will introduce here one more variation, which we call $I^\natural(K)$. Our invariant $I^\natural(K)$ is an invariant for unoriented links $K \subset S^3$ with a marked point $x \in K$ and a preferred normal vector v to K at x. The purpose of the normal vector is in making the invariant functorial for

The work of the first author was supported by the National Science Foundation through NSF grants DMS-0405271 and DMS-0904589. The work of the second author was supported by NSF grant DMS-0805841.

link cobordisms: if $S \subset [0,1] \times S^3$ is a link cobordism from K_1 to K_0, not necessarily orientable, but equipped with a path γ joining the respective basepoints and a section v of the normal bundle to S along γ, then there is an induced map,

$$I^\natural(K_1) \to I^\natural(K_0)$$

that is well-defined up to an overall sign and satisfies a composition law. (We will discuss what is needed to resolve the sign ambiguity in section 4.4.) The definition is set up so that $I^\natural(K) = \mathbb{Z}$ when K is the unknot. We will refer to this homology theory as the reduced singular instanton knot homology of K. (There is also an unreduced version which we call $I^\sharp(K)$ and which can be obtained by applying $I^\natural$ to the union of K with an extra unknotted, unlinked component.) The definitions can be extended by replacing S^3 with an arbitrary closed, oriented 3-manifold Y. The invariants are then functorial for suitable cobordisms of pairs.

Our main result concerning $I^\natural(K)$ is that it is related to reduced Khovanov cohomology by a spectral sequence. The model for this result is a closely-related theorem due to Ozsváth and Szabó [29] concerning the Heegaard Floer homology, with $\mathbb{Z}/2$ coefficients, of a branched double cover of S^3. There is a counterpart to the result of [29] in the context of Seiberg-Witten gauge theory, due to Bloom [3].

Proposition 1.2. *With $\mathbb{Z}$ coefficients, there is a spectral sequence whose E_2 term is the reduced Khovanov cohomology, $Khr(\bar{K})$, of the mirror image knot $\bar{K}$, and which abuts to the reduced singular instanton homology $I^\natural(K)$.*

As an immediate corollary, we have:

Corollary 1.3. *The rank of the reduced Khovanov cohomology $Khr(K)$ is at least as large as the rank of $I^\natural(K)$.* □

To prove Theorem 1.1, it will therefore suffice to show that $I^\natural(K)$ has rank bigger than 1 for non-trivial knots. This will be done by relating $I^\natural(K)$ to a knot homology that was constructed from a different point of view (without singular instantons) by Floer in [9]. Floer's knot homology was revisited by the authors in [24], where it appears as an invariant $KHI(K)$ of knots in S^3. (There is a slight difference between $KHI(K)$ and Floer's original version, in that the latter leads to a group with twice the rank). It is defined using $SU(2)$ gauge theory on a closed 3-manifold obtained from the knot complement. The construction of $KHI(K)$ in [24] was motivated by Juhász's work on sutured manifolds in the setting of Heegaard Floer theory [13, 14]: in the context of sutured manifolds, $KHI(K)$ can be defined as the instanton Floer homology of the sutured 3-manifold obtained from the knot complement by placing two meridional sutures on the torus boundary. It is defined in [24] using complex coefficients for convenience, but one can just as well use $\mathbb{Q}$ or $\mathbb{Z}[1/2]$. The authors establish in [24] that $KHI(K)$ has rank larger than 1 if K is non-trivial. The proof of Theorem 1.1 is therefore completed by the following proposition (whose proof turns out to be a rather straightforward application of the excision property of instanton Floer homology).

Proposition 1.4. *With $\mathbb{Q}$ coefficients, there is an isomorphism between the singular instanton homology $I^\natural(K;\mathbb{Q})$ and the sutured Floer homology of the knot complement, $KHI(K;\mathbb{Q})$.*

Remark. We will see later in this paper that one can define a version of $KHI(K)$ over $\mathbb{Z}$. The above proposition can then be reformulated as an isomorphism over $\mathbb{Z}$ between $I^\natural(K)$ and $KHI(K)$.

Corollary 1.3 and Proposition 1.4 yield other lower bounds on the rank of the Khovanov cohomology. For example, it is shown in [19] that the Alexander polynomial of a knot can be obtained as the graded Euler characteristic for a certain decomposition of $KHI(K)$, so we can deduce:

Corollary 1.5. *The rank of the reduced Khovanov cohomology $Khr(K)$ is bounded below by the sum of the absolute values of the coefficients of the Alexander polynomial of K.* □

For alternating (and more generally, quasi-alternating) knots and links, it is known that the rank of the reduced Khovanov cohomology (over $\mathbb{Q}$ or over the field of 2 elements) is equal to the lower bound which the above corollary provides [25, 27]. Furthermore, that lower bound is simply the absolute value of the determinant in this case. We therefore deduce also:

Corollary 1.6. *When K quasi-alternating, the spectral sequence from $Khr(K)$ to $I^\natural(K)$ has no non-zero differentials after the E_1 page, over $\mathbb{Q}$ or $\mathbb{Z}/2$. In particular, the total rank of $I^\natural(K)$ is equal to the absolute value of the determinant of K.* □

The group $KHI(K)$ closely resembles the "hat" version of Heegaard knot homology, $\widehat{HF}(K)$, defined in [28] and [31]: one can perhaps think of $KHI(K)$ as the "instanton" counterpart of the "Heegaard" group $\widehat{HF}(K)$. The present paper provides a spectral sequence from $Khr(K)$ to $KHI(K)$, but at the time of writing it is not known if there is a similar spectral sequence from $Khr(K)$ to $\widehat{HF}(K)$ for classical knots. This was a question raised by Rasmussen in [32], motivated by observed similarities between reduced Khovanov cohomology and Heegaard Floer homology. There are results in the direction of providing such a spectral sequence in [26], but the problem remains open.

1.2 Outline

Section 2 provides the framework for the definition of the invariant $I^\natural(K)$ by discussing instantons on 4-manifolds X with codimension-2 singularities along an embedded surface Σ. This is material that derives from the authors' earlier work [20], and it was developed further for arbitrary structure groups in [23]. In this paper we work only with the structure group $SU(2)$ or $PSU(2)$, but we extend the previous framework in two ways. First, in the previous development, the locus Σ of the singularity was always taken to be orientable. This condition can be dropped, and we will be considering non-orientable surfaces. Second, the previous expositions always assumed that the bundle which carried the singular connection had an extension across Σ to a bundle on all of X (even though the connection did not extend across the singularity). This condition can also be relaxed. A simple example of such a situation, in dimension 3, arises from a 2-component link K in S^3: the complement of the link has non-trivial second cohomology and there is therefore a $PSU(2)$ bundle on the link complement that does not extend across the link. The second Stiefel-Whitney class of this bundle on $S^3 \backslash K$ is dual to an arc ω running from one component of K to the other.

Section 3 uses the framework from section 2 to define an invariant $I^\omega(Y, K)$ for suitable links K in 3-manifolds Y. The label ω is a choice of representative for the dual of w_2 for a chosen $PSU(2)$ bundle on $Y \backslash K$: it consists of a union of circles in $S^3 \backslash K$ and arcs joining components of K. The invariant $I^\natural(K)$ for a classical knot or link K arises from this more general construction as follows. Given K with a framed basepoint, we form a new link,

$$K^\natural = K \amalg L$$

where L is the oriented boundary of a small disk, normal to K at the basepoint. We take ω to be an arc joining the basepoint of K to a point on L: a radius of the disk in the direction of the vector of the framing. We then define

$$I^\natural(K) = I^\omega(S^3, K^\natural).$$

This construction and related matters are described in more detail in section 4.

Section 5 deals with Floer's excision theorem for Floer homology, as it applies in the context of this paper. In order to work with integer coefficients, some extra work is needed to deal with orientations and $PSU(2)$ gauge transformations that do not lift to $SU(2)$. The excision property is used to prove the relationship between $I^\natural(K)$ and $KHI(K)$ asserted in Proposition 1.4, and also to establish a multiplicative property of $I^\natural(K)$ for split links K.

Sections 6, 7 and 8 are devoted to the proof of Proposition 1.2, concerning the spectral sequence. The first part of the proof is to show that $I^\natural(K)$ can be computed from a "cube of resolutions". This essential idea comes from Ozsváth and Szabó's paper on double-covers [29], and is closely related to Floer's surgery long exact sequence for instanton Floer homology. It can be seen as an extension of a more straightforward property of $I^\natural(K)$, namely that it has a long exact sequence for the unoriented skein relation: that is, if K_2, K_1 and K_0 are knots or links differing at one crossing as shown in Figure 6, then the corresponding groups $I^\natural(K_i)$ form a long exact sequence in which the maps arise from simple cobordisms between the three links. The cube of resolutions provides a spectral sequence abutting to $I^\natural(K)$. Section 8 establishes that the E_2 term of this spectral sequence is the reduced Khovanov cohomology of $\bar{K}$.

Remark. Although it is often called Khovanov *homology*, Khovanov's invariant is a cohomology theory, and we will follow [16] in referring to it as such. The groups we write as $Kh(K)$, or in their bigraded version as $Kh^{i,j}(K)$, are the groups named $\mathcal{H}^{i,j}(K)$ in [16].

Acknowledgments. The authors wish to thank Ciprian Manolescu and Jake Rasmussen for thoughts and conversations that were important in the development of these ideas. Comments from Ivan Smith led to an understanding that the version of Proposition 8.10 in an earlier draft was not stated in a sharp enough form. Finally, this paper owes a great deal to the work of Peter Ozsváth and Zoltán Szabó, particularly their paper [29].

2 Singular instantons and non-orientable surfaces

2.1 Motivation

In [20, 21], the authors considered 4-dimensional connections defined on the complement of an embedded surface and having a singularity along the surface. The basic model is an $SU(2)$ connection in the trivial bundle on $\mathbb{R}^4 \backslash \mathbb{R}^2$ given by the connection matrix

$$i \begin{pmatrix} \alpha & 0 \\ 0 & -\alpha \end{pmatrix} d\theta$$

where α is some parameter in the interval $(0, 1/2)$. Given a closed, embedded surface Σ in an oriented 4-manifold X, one can study anti-self-dual connections on $X\backslash\Sigma$ whose behavior near Σ is modeled on this example. The moduli spaces of such connections were defined and studied in [20]. In [23], a corresponding Floer homology theory was constructed for knots in 3-manifolds; but for the Floer homology theory it was necessary to take $\alpha = 1/4$.

We will now take this up again, but with a slightly more general setup than in the previous paper. We will continue to take $\alpha = 1/4$, and the local model for the singularities of our connections will be the same: only the global topology will be more general. First of all, we will allow our embedded surface Σ to be non-orientable. Second, we will not require that the bundle on the complement of the surface admits any extension, globally, to a bundle on the 4-manifold: specifically, we will consider $PU(2)$ bundles on $X^4\backslash\Sigma^2$ whose second Stiefel-Whitney class is allowed to be non-zero on some torus $\gamma \times S^1$, where γ is a closed curve on Σ with orientable normal and the S^1 factor is the unit normal directions to $\Sigma \subset X$ along this curve. It turns out that the constructions of [20] and [23] carry over with little difficulty, as long as we take $\alpha = 1/4$ from the beginning.

Our first task will be to carefully describe the models for the sort of singular connections we will study. A singular $PU(2)$ connection on $X\backslash\Sigma$ of the sort we are concerned with will naturally give rise to 2-fold covering space $\pi : \Sigma_\Delta \to \Sigma$. To understand why this is so, consider the simplest local model: a flat $PU(2)$ connection A_1 on $B^4\backslash B^2$ whose holonomy around the linking circles has order 2. The eigenspaces of the holonomy decompose the associated $\mathbb{R}^3$ bundle as $\xi \oplus Q$, where ξ is a trivial rank-1 bundle and Q is a 2-plane bundle, the -1 eigenspace. Suppose we wish to extend Q from $B^4\backslash B^2$ to all of B^4. To do this, we construct a new connection A_0 as

$$A_0 = A_1 - \frac{1}{4}\mathbf{i}\, d\theta$$

where θ is an angular coordinate in the 2-planes normal to B^2 and $\mathbf{i}$ is a section of the adjoint bundle which annihilates ξ and has square -1 on Q. Then A_0 is a flat connection in the same bundle, with trivial holonomy, and it determines canonically an extension of the bundle across B^2. In this process, there is a choice of sign: both $d\theta$ and $\mathbf{i}$ depend on a choice (orientations of B^2 and Q respectively). If we change the sign of $\mathbf{i}d\theta$ then we obtain a different extension of the bundle.

In the global setting, when we have a surface $\Sigma \subset X$, we will have two choices of extension of our $PU(2)$ bundle at each point of Σ. Globally, this will determine a double-cover (possibly trivial),

$$\pi : \Sigma_\Delta \to \Sigma.$$

It will be convenient to think of the two different extensions of the $PU(2)$ bundle as being defined simultaneously on a non-Hausdorff space X_Δ. This space comes with a projection $\pi : X_\Delta \to X$ whose fibers are a single point of each point of $X\backslash\Sigma$ and whose restriction to $\pi^{-1}(\Sigma)$ is the double-covering Σ_Δ.

2.2 The topology of singular connections

To set this up with some care, we begin with a closed, oriented, Riemannian 4-manifold X, a smoothly embedded surface $\Sigma \subset X$. We identify a tubular neighborhood ν of Σ with the disk bundle of the normal 2-plane bundle $N_\Sigma \to \Sigma$. This identification gives a tautological section s of the pull-back of N_Σ to ν. The section s is non-zero over $\nu\backslash\sigma$, so on $\nu\backslash\sigma$ we can consider the section

$$s_1 = s/|s|. \tag{1}$$

On the other hand, let us choose any smooth connection in $N_\Sigma \to \Sigma$ and pull it back to the bundle $N_\Sigma \to \nu$. Calling this connection ∇, we can then form the covariant derivative ∇s_1. We can identify the adjoint bundle of the $O(2)$ bundle N_Σ as $i\mathbb{R}_{o(\Sigma)}$, where $\mathbb{R}_{o(\Sigma)}$ is the real orientation line bundle of Σ. So the derivative of s_1 can be written as

$$\nabla s_1 = i\eta s_1$$

for η a 1-form on $\nu\backslash\Sigma$ with values in $\mathbb{R}_{o(\Sigma)}$. This η is a global angular 1-form on the complement of Σ in ν.

Fix a local system Δ on Σ with structure group ± 1, or equivalently a double-cover $\pi : \Sigma_\Delta \to \Sigma$. This determines also a double-cover $\pi : \tilde{\nu}_\Delta \to \nu$. We form a non-Hausdorff space X_Δ as an identification space of $X\backslash\Sigma$ and $\tilde{\nu}_\Delta$, in which each $x \in \tilde{\nu}_\Delta\backslash\Sigma_\Delta$ is identified with its image under π in $X\backslash\Sigma$. We write $\nu_\Delta \subset X_\Delta$ for the (non-Hausdorff) image of the tubular neighborhood $\tilde{\nu}_\Delta$.

The topological data describing the bundles in which our singular connections live will be the following. We will have first a $PU(2)$-bundle $P_\Delta \to X_\Delta$. (This means that we have a bundle on the disjoint union of $\tilde{\nu}_\Delta$ and $X\backslash\Sigma$ together with a bundle isomorphisms between the bundle on $\tilde{\nu}_\Lambda\backslash\Sigma$ and pull-back of the bundle from $X\backslash\Sigma$.) In addition, we will have a reduction of the structure group of $P_\Delta|_{\nu_\Delta}$ to $O(2)$. We will write Q for the associated real 2-plane bundle over ν_Δ.

The bundle Q will be required to have a very particular form. To describe this, we start with 2-plane bundle $\tilde{Q} \to \Sigma_\Delta$ whose orientation bundle is identified with the orientation bundle of Σ_Δ:

$$o(\tilde{Q}) \xrightarrow{\cong} o(\Sigma_\Delta).$$

We can pull $\tilde{Q}$ back to $\tilde{\nu}_\Sigma$. In order to create from this a bundle over the non-Hausdorff quotient ν_Δ, we must give for each $x \in \tilde{\nu}_\Delta \backslash \Sigma_\Delta$ an identification of the fibers,

$$Q_{\tau(x)} \to Q_x \tag{2}$$

where τ is the covering transformation. Let us write

$$\mathrm{Hom}^-(\tilde{Q}_{\tau(x)}, \tilde{Q}_x) \tag{3}$$

for the 2-plane consisting of linear maps that are scalar multiples of an orientation-reversing isometry (i.e. the complex-anti-linear maps if we think of both Q_x and $Q_{\tau(x)}$ as oriented). Like $\tilde{Q}_x$ and $\tilde{Q}_{\tau(x)}$, the 2-plane (3) has its orientation bundle canonically identified with $o(\Sigma)$: our convention for doing this is to fix any vector in $\tilde{Q}_{\tau(x)}$ and use it to map $\mathrm{Hom}^-(\tilde{Q}_{\tau(x)}, \tilde{Q}_x)$ to $\tilde{Q}_x$. We will give an identification (2) by specifying an orientation-preserving bundle isometry of 2-plane bundles on Σ_Δ,

$$\rho : N_{\Sigma_\Delta} \to \mathrm{Hom}^-(\tau^*(\tilde{Q}), \tilde{Q}).$$

This ρ should satisfy

$$\rho(v)\rho(\tau(v)) = 1$$

for a unit vector v in N_{Σ_Δ}. The identification (2) can then be given by $\rho(s_1)$, where s_1 is as in (1). The existence of such a ρ is a constraint on $\tilde{Q}$.

To summarize, we make the following definition.

Definition 2.1. Given a pair (X, Σ), by *singular bundle data* on (X, Σ) we will mean a choice of the following items:

(a) a double-cover $\Sigma_\Delta \to \Sigma$ and an associated non-Hausdorff space X_Δ;

(b) a principal $PU(2)$-bundle P_Δ on X_Δ;

(c) a 2-plane bundle $\tilde{Q} \to \Sigma_\Delta$ whose orientation bundle is identified with $o(\Sigma_\Delta)$;

(d) an orientation-preserving bundle isometry

$$\rho : N_{\Sigma_\Delta} \to \mathrm{Hom}^-(\tau^*(\tilde{Q}), \tilde{Q})$$

(e) an identification, on the non-Hausdorff neighborhood ν_Δ, of the resulting quotient bundle Q_Δ with an $O(2)$ reduction of $P_\Delta|_{\nu_\Delta}$.

◊

Remark. When the conditions in this definition are fulfilled, the double-cover Σ_Δ is in fact determined, up to isomorphism, by the $PU(2)$ bundle P on $X\backslash\Sigma$. It is also the case that not every double-cover of Σ can arise as Σ_Δ. We shall return to these matters in the next subsection.

Given singular bundle data on (X, Σ), we can now write down a model singular connection. We start with a smooth connection a_0 in the bundle $\tilde{Q} \to \Sigma_\Delta$, so chosen that the induced connection in $\mathrm{Hom}^-(\tau^*(\tilde{Q}), \tilde{Q})$ coincides with the connection ∇ in N_{Σ_Δ} under the isomorphism ρ. (Otherwise said, ρ is parallel.) By pull-back, this a_0 also determines a connection in $\tilde{Q}$ on $\tilde{\nu}_\Delta$. On the deleted tubular neighborhood $\tilde{\nu}_\Delta\backslash\Sigma_\Delta$, the bundles $\tau^*(\tilde{Q})$ and $\tilde{Q}$ are being identified by the isometry $\rho(s_1)$; but under this identification, the connection a_0 is not preserved, because s_1 is not parallel. So a_0 does not by itself give rise to a connection over $\nu\backslash\Sigma$ downstairs. The covariant derivative of s_1 is $i\eta s_1$, where η is a 1-form with values in $\mathbb{R}_{o(\Sigma)}$, or equivalently in $\mathbb{R}_{o(\tilde{Q})}$ on $\tilde{\nu}_\Delta$. We can therefore form a new connection

$$\tilde{a}_1 = a_0 + \frac{i}{2}\eta$$

as a connection in $\tilde{Q} \to \tilde{\nu}_\Delta\backslash\Sigma_\Delta$. With respect to this new connection, the isometry $\rho(s_1)$ is covariant-constant, so $\tilde{a}_1$ descends to a connection a_1 on the resulting bundle $Q \to \nu\backslash\Sigma$ downstairs.

Since Q is a reduction of the $PU(2)$ bundle P on $\nu\backslash\Sigma$, the $O(2)$ connection a_1 gives us a $PU(2)$ connection in P there. Let us write A_1 for this $PU(2)$ connection. We may extend A_1 in any way wish to a connection in P over all of $X\backslash\Sigma$. If we pick a point x in Σ_Δ, then a standard neighborhood of x in X_Δ is a B^4 meeting Σ_Δ in a standard B^2. In such a neighborhood, the connection A_1 on $B^4\backslash B^2$ can be written as

$$A_1 = A_0 + \frac{1}{4}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\eta,$$

where A_0 is a smooth $PU(2)$ connection with reduction to $O(2)$. Here are notation identifies the Lie algebra of $PU(2)$ with that of $SU(2)$, and the element i in $\mathrm{Lie}(O(2))$ corresponds to the element

$$\frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

in $\mathfrak{su}(2)$. Note that in this local description, the connection A_0 depends in a significant way on our choice of $x \in \Sigma_\Delta$, not just on the image of x in Σ. Two different points x and x' with the same image in Σ give rise to connections A_0 and A_0' which differ by a term

$$\frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\eta.$$

We will refer to any connection A_1 arising in this way as a *model singular connection.* Such an A_1 depends on a choice of singular bundle data (Definition 2.1), a choice of a_0 making ρ parallel, and a choice of extension of the resulting connection from the tubular neighborhood to all of $X\backslash\Sigma$. The latter two choices are selections from certain affine spaces of connections, so it is the singular bundle data that is important here.

2.3 Topological classification of singular bundle data

When classifying bundles over X_Δ up to isomorphism, it is helpful in the calculations to replace this non-Hausdorff space by a Hausdorff space with the same weak homotopy type. We can construct such a space, X^h_Δ, as an identification space of the disk bundle $\tilde{\nu}_\Delta$ and the complement $X\backslash\mathrm{int}(\nu)$, glued together along $\partial\nu$ using the 2-to-1 map $\partial(\tilde{\nu}_\Delta) \to \partial\nu$. There is a map

$$\pi : X^h_\Delta \to X$$

which is 2-to-1 over points of $\mathrm{int}(\nu)$ and 1-to-1 elsewhere. The inverse image of ν under the map π is a 2-sphere bundle

$$S^2 \hookrightarrow D \to \Sigma.$$

In the case that Δ is trivial, this 2-sphere bundle D is the double of the tubular neighborhood ν, and a choice of trivialization of Δ determines an orientation of D. When Δ is non-trivial, D is not orientable: its orientation bundle is Δ. There is also an involution

$$t : X^h_\Delta \to X^h_\Delta$$

with $\pi \circ t = \pi$ whose restriction to each S^2 in D is an orientation-reversing map.

In $H_4(X^h_\Delta; \mathbb{Q})$ there is a unique class $[X_\Delta]$ which is invariant under the involution t and has $\pi_*[X_\Delta] = [X]$. This class is not integral: in terms of a triangulation of X^h_Δ, the class can be described as the sum of the 4-simplices belonging to $X^h_\Delta \backslash \tilde{\nu}_\Delta$, plus half the sum of the simplices belonging to D, with all simplices obtaining their orientation from the orientation of X. In the case that the double-cover $\Sigma_\Delta \to \Sigma$ is trivial, there is a different fundamental class to consider. In this case, X_Δ is the union of two copies of X, identified on the complement of Σ. A choice of trivialization of Δ picks out one of these two copies: call it $X_+ \subset X_\Delta$. The fundamental class $[X_+]$ is an integral class in $H_4(X_\Delta)$: it can be expressed as

$$[X_+] = ([X_\Delta] + [D]/2)$$

where D is oriented using the trivialization of Δ. A full description of H_4 with $\mathbb{Z}$ coefficients is given by the following, whose proof is omitted.

Lemma 2.2. *The group $H_4(X^h_\Delta; \mathbb{Z})$ has rank $1 + s$, where s is the number of components of Σ on which Δ is trivial. If Δ is trivial and trivialized, then free generators are provided by (a) the fundamental classes of the components of the 2-sphere bundle D, and (b) the integer class $[X_+]$ determined by the trivialization. If Δ is non-trivial, then free generators are provided by (a) the fundamental classes of the orientable components of D, and (b) the integer class $2[X_\Delta]$.* □

The top cohomology of X_Δ contains 2-torsion if Δ is non-trivial:

Lemma 2.3. *If a is the number of components of Σ on which Δ is non-trivial, then the torsion subgroup of $H^4(X^h_\Delta; \mathbb{Z})$ is isomorphic to $(\mathbb{Z}/2)^{a-1}$ if $a \geq 2$ and is zero otherwise. To describe generators, let $x_1, \ldots, x_a$ be points in the different non-orientable components of D which map under π to points in $\mathrm{int}(\nu)$. Let ξ_i be the image in $H^4(X^h_\Delta)$ of a generator of $H^4(X^h_\Delta, X^h_\Delta \backslash x_i) \cong \mathbb{Z}$, oriented so that $\langle \xi_i, [X_\Delta] \rangle = 1$. Then generators for the torsion subgroup are the elements $\xi_i - \xi_{i+1}$, for $i = 1, \ldots, a-1$.*

Proof. This is also straightforward. The element $\xi_i - \xi_{i+1}$ is non-zero because it has non-zero pairing with the $\mathbb{Z}/2$ fundamental class of the i'th component of D. On the other hand, $2\xi_i = 2\xi_{i+1}$, because both of these classes are equal to the pull-back by π of the generator of $H^4(X; \mathbb{Z})$. □

Because we wish to classify $SO(3)$ bundles, we are also interested in H^2 with $\mathbb{Z}/2$ coefficients.

Lemma 2.4. *The group $H^2(X^h_\Delta; \mathbb{Z}/2)$ lies in an exact sequence*

$$0 \to H^2(X;\mathbb{Z}/2) \xrightarrow{\pi^*} H^2(X^h_\Delta;\mathbb{Z}/2) \xrightarrow{e} (\mathbb{Z}/2)^N \to H_1(X)$$

where N is the number of components of Σ. The map e is the restriction map to

$$\bigoplus_i H^2(S^2_i;\mathbb{Z}/2)$$

for a collection of fibers $S^2_i \subset D$, one from each component of D. If $\lambda_i \in H_1(\Sigma)$ denotes the Poincaré dual of $w_1(\Delta|_{\Sigma_i})$, then the last map is given by

$$(\epsilon_1, \ldots, \epsilon_N) \mapsto \sum_i \epsilon_i \lambda_i \in H_1(X;\mathbb{Z}/2).$$

Proof. With $\mathbb{Z}/2$ coefficients understood, we have the following commutative diagram, in which the vertical arrows from the bottom row are given by π^*, the rows are exact sequences coming from Mayer-Vietoris, and the middle column is a short exact sequence:

$$\begin{array}{ccccccccc}
 & & & & (\mathbb{Z}/2)^N & & & & \\
 & & & & \uparrow & & & & \\
\cdots \longrightarrow & H^2(X^h_\Delta) & \longrightarrow & & H^2(X\backslash\Sigma)\oplus H^2(D) & \longrightarrow & H^2(\partial\nu) & \longrightarrow \cdots \\
 & \uparrow & & & \uparrow & & \uparrow & \\
\cdots \longrightarrow & H^2(X) & \longrightarrow & & H^2(X\backslash\Sigma)\oplus H^2(\nu) & \longrightarrow & H^2(\partial\nu) & \longrightarrow \cdots
\end{array}$$

The lemma follows from an examination of the diagram. □

Let us recall from [5] that $SO(3)$ bundles P on a 4-dimensional simplicial complex Z can be classified as follows. First, P has a Stiefel-Whitney class $w_2(P)$, which can take on any value in $H^2(Z;\mathbb{Z}/2)$. Second, the isomorphism classes of bundles with a given $w_2(P) = w$ are acted on transitively by $H^4(Z;\mathbb{Z})$. In the basic case of a class in H^4 represented by the characteristic function of a single oriented 4-simplex, this action can be described as altering the bundle on the interior of the simplex by forming a connect sum with an $SO(3)$ bundle $Q \to S^4$ with $p_1(Q)[S^4] = -4$ (i.e. the $SO(3)$ bundle associated to an $SU(2)$ bundle with $c_2 = 1$). Acting on a bundle P by a class $z \in H^4(Z;\mathbb{Z})$ alters $p_1(P)$ by $-4z$. The action is may not be effective if the cohomology of Z has 2-torsion: according to [5], the kernel of the action is the subgroup

$$\mathcal{T}^4(Z;w) \subset H^4(Z;\mathbb{Z})$$

given by

$$\mathcal{T}^4(Z;w) = \{\, \beta(x) \smile \beta(x) + \beta(x \smile w_2(P)) \mid x \in H^1(Z;\mathbb{Z}/2)\,\},$$

where β is the Bockstein homomorphism $H^i(Z;\mathbb{Z}/2) \to H^{i+1}(Z;\mathbb{Z})$. There are two corollaries to note concerning the class $p_1(P)$ here. First, if $H^4(Z;\mathbb{Z})$ contains classes z with $4z = 0$ which do *not* belong to $\mathcal{T}^4(Z;w)$, then an $SO(3)$ bundle $P \to Z$ with $w_2(P) = w$ is not determined up to isomorphism by its Pontryagin class. Second, $p_1(P)$ is determined by $w = w_2(P)$ to within a coset of the subgroup consisting of multiples of 4; on in other words, the image of $\bar{p}_1(P)$ of $p_1(P)$ in $H^4(Z;\mathbb{Z}/4)$ is determined by $w_2(P)$. According to [5] again, the determination is

$$\bar{p}_1(P) = \mathcal{P}(w_2(P)),$$

where $\mathcal{P}$ is the Pontryagin square, $H^2(Z;\mathbb{Z}/2) \to H^4(Z;\mathbb{Z}/4)$.

Now let us apply this discussion to X^h_Δ and the bundles P_Δ arising from singular bundle data as in Definition 2.1. The conditions of Definition 2.1 imply that $w_2(P_\Delta)$ is non-zero on every 2-sphere fiber in D. A first step in classifying such bundles P_Δ is to classify the possible classes w_2 satisfying this condition. Referring to Lemma 2.4, we obtain:

Proposition 2.5. *Let $\lambda \subset \Sigma$ be a 1-cycle with $\mathbb{Z}/2$ coefficients dual to $w_1(\Delta)$. A necessary and sufficient condition that there should exist a bundle $P_\Delta \to X_\Delta$ with w_2 non-zero on every 2-sphere fiber of D is that λ represent the zero class in $H_1(X;\mathbb{Z}/2)$. When this condition holds the possible values for w_2 lie in a single coset of $H^2(X;\mathbb{Z}/2)$ in $H^2(X_\Delta;\mathbb{Z}/2)$.* □

Let us fix w_2 and consider the action of $H^4(X_\Delta;\mathbb{Z})$ on the isomorphism classes of bundles P_Δ. We orient all the 4-simplices of X^h_Δ using the map π to X. We also choose trivializations of Δ on all the components of Σ on which it is trivial. We can then act by the class in H^4 represented by the characteristic function of a single oriented 4-simplex σ. We have the following cases, according to where σ lies.

(a) If σ is contained in $X \backslash \nu \subset X^h_\Delta$, then we refer to this operation as *adding an instanton.*

(b) If σ is contained in D, then we have the following subcases:

   (i) if the component of D is orientable, so that it is the double of ν, and if σ belongs to the distinguished copy of ν in D picked out by our trivialization of Δ, then we refer this operation as *adding an anti-monopole* on the given component;

(ii) if the component of D is again orientable, but σ lies in the other copy of ν, then we refer to the action of this class as *adding a monopole*;

(iii) if the component of D is not orientable then the characteristic function of any 4-simplex in D is cohomologous to any other, and we refer to this operation as adding a monopole.

We have the following dependencies among these operations, stemming from the fact that the corresponding classes in $H^4(X^h_\Delta;\mathbb{Z})$ are equal:

(a) adding a monopole and an anti-monopole to the same orientable component is the same as adding an instanton;

(b) adding two monopoles to the same non-orientable component is the same as adding an instanton.

Further dependence among these operations arises from the fact that the action of the subgroup $\mathcal{T}^4(X_\Delta; w_2)$ is trivial. The definition of $\mathcal{T}^4$ involves $H^1(X^h_\Delta;\mathbb{Z}/2)$, and the latter group is isomorphic to $H^1(X;\mathbb{Z}/2)$ via π^*. Since $H^4(X;\mathbb{Z})$ has no 2-torsion, the classes $\beta(x)\smile\beta(x)$ are zero. Calculation of the term $\beta(x\smile w_2)$ leads to the following interpretation:

(c) For any class x in $H^1(X;\mathbb{Z}/2)$, let n be the (necessarily even) number of components of Σ on which $w_1(\Delta)\smile(x|_\Sigma)$ is non-zero. Then the effect of adding in n monopoles, one on each of these components, is the same as adding $n/2$ instantons.

Our description so far gives a complete classification of $SO(3)$ (or $PSU(2)$) bundles $P_\Delta\to X_\Delta$ having non-zero w_2 on the 2-sphere fibers of D. Classifying such bundles turns out to be equivalent to classifying the (a priori more elaborate) objects described as singular bundle data in Definition 2.1.

Proposition 2.6. *For fixed Δ, the forgetful map from the set of isomorphism classes of singular bundle data to the set of isomorphism classes of $SO(3)$ bundles P_Δ on X_Δ is a bijection onto the isomorphism classes of bundles P_Δ with $w_2(P_\Delta)$ odd on the 2-sphere fibers in D.*

Proof. Let $P_\Delta\to X_\Delta$ be given. We must show that the restriction of P_Δ to the non-Hausdorff neighborhood ν_Δ (or equivalently, on X^h_Δ, the restriction of P_Δ to the 2-sphere bundle P) admits a reduction to an $O(2)$ bundle Q_Δ of the special sort described in the definition. We must also show that this reduction is unique up to homotopy.

Consider the reduction of P_Δ to the 2-sphere bundle D. The space D is the double of the tubular neighborhood, and as a sphere-bundle it therefore comes with a 2-valued section. Fix a point x in Σ and consider the fiber D_x over x, written as the union of 2 disks D^+ and D^- whose centers are points x^+ and x^- given by this 2-valued section at the point x. The bundle P_Δ on D_x can be described canonically up to homotopy as arising from a clutching function on the equatorial circle:

$$\gamma : (D^+ \cap D^-) \to \mathrm{Hom}(P_\Delta|_{x^-}, P_\Delta|_{x^+}).$$

The target space here is an isometric copy of $SO(3)$. The loop γ belongs to the non-trivial homotopy class by our assumption about w_2. The space of loops $\mathrm{Map}(S^1, SO(3))$ in the non-zero homotopy class contains inside it the simple closed geodesics; and the inclusion of the space of these geodesics is an isomorphism on π_1 and π_2 and surjective map on π_3, as follows from a standard application of the Morse theory for geodesics. Since Σ is 2-dimensional, the classification of bundles P_Δ is therefore the same as the classification of bundles with the additional data of being constructed by clutching functions that are geodesics on each circle fiber $D^+ \cap D^-$. On the other hand, describing P_Δ on D by such clutching functions is equivalent to giving a reduction of P_Δ to an $O(2)$ bundle Q_Δ arising in the way described in Definition 2.1. □

2.4 Instanton and monopole numbers

Consider again for a moment the case that Δ is trivial and trivialized, so that we have a standard copy X_+ of X inside X_Δ. Inside X_+ is a preferred copy, Σ_+, of the surface Σ. The orientation bundle of Q_Δ and the orientation bundle of Σ_+ are canonically identified along Σ_+. In this situation we can define two characteristic numbers,

$$\begin{aligned} k &= -\frac{1}{4}\langle p_1(P_\Delta), [X_+]\rangle \\ l &= -\frac{1}{2}\langle e(Q_\Delta), [\Sigma_+]\rangle. \end{aligned}$$

(Here the euler class $e(Q_\Delta)$ is regarded as taking values in the second cohomology of Σ_+ with coefficients twisted by the orientation bundle.) We call these the *instanton number* and the *monopole number* respectively: in the case that Σ is orientable, these definitions coincide with those from the authors earlier papers [20, 23]. Since the $O(2)$ bundle Q_Δ has degree -1 on the fibers of the 2-sphere bundle D, a short calculation allows us to express l also in terms of the Pontryagin class of the bundle P_Δ:

$$l = \frac{1}{4} p_1(P_\Delta)[D] + \frac{1}{4}\Sigma \cdot \Sigma. \tag{4}$$

(The term $\Sigma \cdot \Sigma$ is the "self-intersection" number of Σ. Recall that this is a well-defined integer, even when Σ is non-orientable, as long as X is oriented.) When Δ is non-trivial, we cannot define either k or l in this way. We can always evaluate the Pontryagin class on $[X_\Delta]$ however.

Recall that for an $SU(2)$ bundle on a closed 4-manifold X, the characteristic number $c_2(P)[X]$ can be computed by the Chern-Weil formula

$$\frac{1}{8\pi^2}\int_X \mathrm{tr}(F_A \wedge F_A)$$

where A is any $SU(2)$ connection and the trace is the usual trace on 2-by-2 complex matrices. For a $PSU(2)$ bundle, the same formula computes $-(1/4)p_1(P)[X]$. (Here we must identify the Lie algebra of $PSU(2)$ with that of $SU(2)$ and define the trace form accordingly.)

Consider now a model singular connection A, as defined in section 2.2, corresponding to singular bundle data $P_\Delta \to X_\Delta$. We wish to interpret the Chern-Weil integral in terms of the Pontryagin class of P_Δ. We have:

Proposition 2.7. *For a model singular connection A corresponding to singular bundle data $P_\Delta \to X_\Delta$, we have*

$$\frac{1}{8\pi^2}\int_{X\setminus\Sigma} \mathrm{tr}(F_A \wedge F_A) = -\frac{1}{4}p_1(P_\Delta)[X_\Delta] + \frac{1}{16}\Sigma\cdot\Sigma,$$

where $[X_\Delta] \in H_4(X_\Delta;\mathbb{Q})$ is again the fundamental class that is invariant under the involution t and has $\pi_*[X_\Delta] = [X]$.

We shall write κ for this Chern-Weil integral:

$$\kappa(A) = \frac{1}{8\pi^2}\int_{X\setminus\Sigma} \mathrm{tr}(F_A \wedge F_A). \tag{5}$$

Proof of the proposition. A formula for κ in terms of characteristic classes was proved in [20] under the additional conditions that Σ was orientable and Δ was trivial. In that case, after choosing a trivialization of Δ, the formula from [20] was expressed in terms of the instanton and monopole numbers, k and l, as

$$\kappa(A) = k + \frac{1}{2}l - \frac{1}{16}\Sigma\cdot\Sigma.$$

(In [20] the formula was written more generally for a singular connection with a holonomy parameter α. The formula above is the special case $\alpha = 1/4$.) Using the expression (4) for the monopole number, this formula becomes

$$\kappa(A) = -\frac{1}{4}p_1(P_\Delta)[X_+] + \frac{1}{8}p_1(P_\Delta)[D] + \frac{1}{16}\Sigma \cdot \Sigma,$$

which coincides with the formula in the proposition, because $[X_\Delta] = [X_+] - (1/2)[D]$. Thus the formula in the proposition coincides with the formula from [20] in this special case. The proof in [20] is essentially a local calculation, so the result in the general case is the same. □

2.5 The determinant-1 gauge group

If $P \to X$ is an $SO(3)$ bundle then there is a bundle $G(P) \to X$ with fiber the group $SU(2)$, associated to P via the adjoint action of $SO(3)$ on $SU(2)$. We refer to the sections of $G(P) \to X$ as *determinant-1 gauge transformation*, and we write $\mathcal{G}(P)$ for the space of all such sections, the *determinant-1 gauge group*. The gauge group $\mathcal{G}(P)$ acts on the bundle P by automorphisms, but the map

$$\mathcal{G}(P) \to \mathrm{Aut}(P)$$

is not an isomorphism: its kernel is the two-element group $\{\pm 1\}$, and its cokernel can be identified with $H^1(X; \mathbb{Z}/2)$.

Suppose we are now given singular bundle data over the non-Hausdorff space X_Δ, represented in particular by an $SO(3)$ bundle $P_\Delta \to X_\Delta$. We can consider the group of determinant-1 gauge transformations, $\mathcal{G}(P_\Delta)$, in this context. Up to this point, we have not been specific, but let us now consider simply *continuous* sections of $G(P_\Delta)$ and denote the correspond group as $\mathcal{G}^{\mathrm{top}}$.

To understand $\mathcal{G}^{\mathrm{top}}$, consider a 4-dimensional ball neighborhood U of a point in Σ_Δ. A section g of the restriction of $G(P_\Delta)$ to U is simply a map $U \to SU(2)$, in an appropriate trivialization. Let U' be the image of U under the involution t on X_Δ, so that $U \cap U' = U \backslash \Sigma_\Delta$. The section g on U determines a section g' of the same bundle on $U' \backslash \Sigma_\Delta$. In order for g to extend to a section of $G(P_\Delta)$ on $U \cup U'$, it is necessary that g' extend across $U' \cap \Sigma_\Delta$. In a trivialization, g' is an $SU(2)$-valued function on $B^4 \backslash B^2$ obtained by applying a discontinuous gauge transformation to the function $g : B^4 \to SU(2)$. It has the local form

$$g'(x) = \mathrm{ad}(u(\theta))g(x)$$

where θ is an angular coordinate in the normal plane to $B^2 \subset B^4$ and u is the one-parameter subgroup of $SU(2)$ that respects the reduction to an $O(2)$ bundle Q_Δ on U. In order for g' to extend continuously over B^2, it is necessary and sufficient that $g(x)$ commutes with the one-parameter subgroup $u(\theta)$ when $x \in B^2$: that is, $g(x)$ for $x \in U \cap \Sigma_\Delta$ should itself lie in the S^1 subgroup that preserves the subbundle Q_Δ as well as its orientation.

To summarize, the bundle of groups $G(P_\Delta) \to X_\Delta$ has a distinguished subbundle over Σ_Δ,

$$H_\Delta \subset G(P_\Delta) \to \Sigma_\Delta$$

whose fiber is the group S^1; and the continuous sections of $G(P_\Delta)$ take values in this subbundle along Σ_Δ. The local model is an $SU(2)$-valued function on B^4 constrained to take values in S^1 on $B^2 \subset B^4$.

The bundle $H_\Delta \to \Sigma_\Delta$ is naturally pulled back from Σ. Indeed, we can describe the situation in slightly different terms, without mentioning Δ. We have an $SO(3)$ bundle $P \to X \backslash \Sigma$ and a reduction of P to an $O(2)$ bundle Q on $\nu \backslash \Sigma$. The bundle $G(P) \to X \backslash \Sigma$ has a distinguished subbundle H over $\nu \backslash \Sigma$, namely the bundle whose fiber is the group $S^1 \subset SU(2)$ which preserves Q and its orientation. This subbundle has structure group ± 1 and is associated to the orientation bundle of Q. This local system with structure group ± 1 on $\nu \backslash \Sigma$ is pulled back from Σ itself, so H extends canonically over Σ. Thus, although the bundle $G(P)$ on $X \backslash \Sigma$ does not extend, its subbundle H does. There is a topological space over $\mathbf{G} \to X$ obtained as an identification space of $G(P)$ over $X \backslash \Sigma$ and H over ν:

$$\mathbf{G} = (H \cup G(P)) / \sim .$$

The fibers of $\mathbf{G}$ over X are copies of S^1 over Σ and copies of $SU(2)$ over $X \backslash \Sigma$. The group $\mathcal{G}^{\text{top}}$ is the space of continuous sections of $\mathbf{G} \to X$.

We now wish to understand the component group, $\pi_0(\mathcal{G}^{\text{top}})$. To begin, note that we have a restriction map

$$\mathcal{G}^{\text{top}} \to \mathcal{H}$$

where $\mathcal{H}$ is the space of sections of $H \to \Sigma$.

Lemma 2.8. *The group of components $\pi_0(\mathcal{H})$ is isomorphic to $H_1(\Sigma; \mathbb{Z}_\Delta)$, where $\mathbb{Z}_\Delta$ is the local system with fiber $\mathbb{Z}$ associated to the double-cover Δ. The map $\pi_0(\mathcal{G}^{\text{top}}) \to \pi_0(\mathcal{H})$ is surjective.*

Proof. Over Σ there is a short exact sequence of sheaves (essentially the real exponential exact sequence, twisted by the orientation bundle of Q):

$$0 \to \mathbb{Z}_Q \to C^0(\mathbb{R}_Q) \to C^0(H) \to 0.$$

From the resulting long exact sequence, one obtains an isomorphism between $\pi_0(\mathcal{H})$ and $H^1(\Sigma;\mathbb{Z}_Q)$, which is isomorphic to $H_1(\Sigma;\mathbb{Z}_\Delta)$ by Poincaré duality. (Recall that the difference between Δ and the orientation bundle of Q is the orientation bundle of Σ.) Geometrically, this isomorphism is realized by taking a section of H in a given homotopy class, perturbing it to be transverse to the constant section -1, and then taking the inverse image of -1. This gives a smooth 1-manifold in Σ whose normal bundle is identified with the orientation bundle of H, and whose tangent bundle is therefore identified with the orientation bundle of Δ. This 1-manifold, C, represents the element of $H_1(\Sigma;\mathbb{Z}_\Delta)$ corresponding to the given element of $\pi_0(\mathcal{H})$.

To prove surjectivity, we consider a class in $H_1(\Sigma;\mathbb{Z}_\Delta)$ represented by a 1-manifold C in Σ whose normal bundle is identified with $\mathbb{R}_Q$, and we seek to extend the corresponding section h of $H \to \Sigma$ to a section g of $\mathbf{G} \to X$. We can take h to be supported in a 2-dimensional tubular neighborhood V_2 of C, and we seek a g that is supported in a 4-dimensional tubular neighborhood V_4 of C. Let $V_4' \subset V_4$ be smaller 4-dimensional tubular neighborhood. The section h determines, by extension, a section h' of $\mathbf{G}$ on V_4', and we need to show that $h'|_{\partial V_4'}$ is homotopic to the section 1. The fiber of $\partial V_4'$ over a point $x \in C$ is a 2-sphere, and on this 2-sphere h' is equal to -1 on an equatorial circle E (the circle fiber of the bundle $\partial\nu \to \Sigma$ over x). To specify a standard homotopy from h' to 1 on this 2-sphere it is sufficient to specify a non-vanishing section of Q over E. To specify a homotopy on the whole of $\partial V_4'$ we therefore seek a non-vanishing section of the 2-plane bundle Q on the circle bundle $T = \partial\nu\_C \to C$. This T is a union of tori or a Klein bottles, and on each component the orientation bundle of Q is identified with the orientation bundle of T. The obstruction to there being a section is therefore a collection of integers k, one for each component. Passing the Δ-double cover, we find our bundle Q extending from the circle bundle to the disk bundle, as the 2-plane bundle $Q_\Delta \to \nu_\Delta|_C$. The integer obstructions k therefore satisfy $2k = 0$. So $k = 0$ and the homotopy exists. □

Next we look at the kernel of the restriction map $\mathcal{G}^{\text{top}} \to \mathcal{H}$, which we denote temporarily by $\mathcal{K}$, in the exact sequence

$$1 \to \mathcal{K} \to \mathcal{G}^{\text{top}} \to \mathcal{H}.$$

Lemma 2.9. *The group of components, $\pi_0(\mathcal{K})$, admits a surjective map*

$$\pi_0(\mathcal{K}) \to H_1(X \backslash \Sigma; \mathbb{Z})$$

whose kernel is either trivial or $\mathbb{Z}/2$. The latter occurs precisely when $w_2(P) = w_2(X\backslash\Sigma)$ in $H^2(X\backslash\Sigma;\mathbb{Z}/2)$.

Proof. This is standard [1]. A representative g for an element of $\pi_0(\mathcal{K})$ is a section of $\mathbf{G}$ that is 1 on Σ. The corresponding element of $H_1(X \setminus \Sigma; \mathbb{Z})$ is represented by $\tilde{g}^{-1}(-1)$, where $\tilde{g} \simeq g$ is a section transverse to -1. The kernel is generated by a gauge transformation that is supported in a 4-ball and represents the non-trivial element of $\pi_4(SU(2))$. This element survives in $\pi_0(\mathcal{K})$ precisely when the condition on w_2 holds. □

2.6 Analysis of singular connections

Having discussed the topology of singular connections, we quickly review some of the analytic constructions of [23] which in turn use the work in [20]. Fix a closed pair (X, Σ) and singular bundle data $\mathbf{P}$ on (X, Σ), and construct a model singular connection A_1 on $P \to X \backslash \Sigma$. (See section 2.2.) We wish to define a Banach space of connections modeled on A_1. In [20] two approaches to this problem were used, side by side. The first approach used spaces of connections modeled on L^p_1, while the second approach used stronger norms. The second approach required us to equip X with a metric with an orbifold singularity along Σ rather than a smooth metric.

It is the second of the two approaches that is most convenient in the present context. Because we are only concerned with the case that the holonomy parameter α is $1/4$, we can somewhat simplify the treatment: in [20], the authors to used metrics g^ν on X with cone angle $2\pi/\nu$ along Σ, with ν a (possibly large) natural number. In the present context we can simply take $\nu = 2$, equipping X with a metric with cone angle π along Σ.

Equipped with such a metric, X can be regarded as an orbifold with point-groups $\mathbb{Z}/2$ at all points of Σ. We will write $\check{X}$ for X when regarded as an orbifold in this way, and we write $\check{g}$ for an orbifold Riemannian metric. The holonomy of A_1 on small loops linking Σ in $X \setminus \Sigma$ is asymptotically of order 2; so in local branched double-covers of neighborhoods of points of Σ, the holonomy is asymptotically trivial. We can therefore take it that P extends to an orbifold bundle $\check{P} \to \check{X}$ and A_1 extends to a smooth orbifold connection $\check{A}_1$ in this orbifold $SO(3)$-bundle. (Note that if we wished to locally extend an $SU(2)$ bundle rather than an $SO(3)$ bundle in this context, we should have required a 4-fold branched cover and we would have been led to use a cone angle of $\pi/2$, which was the approach in [20].)

Once we have the metric $\check{g}$ and our model connection $\check{A}_1$, we can define Sobolev spaces using the covariant derivatives $\nabla^k_{\check{A}_1}$ on the bundles $\Lambda^p(T^*\check{X}) \otimes \mathfrak{g}_{\check{P}}$, where the Levi-Civita connection is used in $T^*\check{X}$. The Sobolev space $\check{L}^2_k(\check{X}; \Lambda^p \otimes \mathfrak{g}_{\check{P}})$ is the completion of space of smooth orbifold sections with respect to the norm

$$\|a\|^2_{\check{L}^2_{k,\check{A}_1}} = \sum_{i=0}^{k} \int_{X\backslash\Sigma} |\nabla_{\check{A}_1} a|^2 d\mathrm{vol}_{\check{g}}.$$

We fix $k \geq 3$ and we consider a space of connections on $P \to X\backslash\Sigma$ defined as

$$\mathcal{C}_k(X, \Sigma, \mathbf{P}) = \{\, A_1 + a \mid a \in \check{L}^2_k(\check{X}) \,\}. \tag{6}$$

As in [20, Section 3], the definition of this space of connections can be reformulated to make clear that it depends only on the singular bundle data $\mathbf{P}$, and does not otherwise depend on A_1. The reader can look there for a full discussion.

This space of connections admits an action by the gauge group

$$\mathcal{G}_{k+1}(X, \Sigma, \mathbf{P})$$

which is the completion in the $\check{L}^2_{k+1}$ topology of the group $\mathcal{G}(\check{P})$ of smooth, determinant-1 gauge transformations of the orbifold bundle. The fact that this is a Banach Lie group acting smoothly on $\mathcal{C}_k(X, \Sigma, \mathbf{P})$ is a consequence of multiplication theorems just as in [20]. Note that the center ± 1 in $SU(2)$ acts trivially on $\mathcal{C}_k$ via constant gauge transformations. Following the usual gauge theory nomenclature we call a connection whose stabilizer is exactly ± 1 *irreducible* and otherwise we call it *reducible*. The homotopy-type of the gauge group $\mathcal{G}_{k+1}(X, \Sigma, \mathbf{P})$ coincides with that of $\mathcal{G}^{\mathrm{top}}$, the group of continuous, determinant-1 gauge transformations considered earlier.

Here is the Fredholm package for the present situation. Let $A \in \mathcal{C}_k(X, \Sigma, \mathbf{P})$ be a singular connection on (X, Σ) equipped with the metric $\check{g}$, and let d^+_A be the linearized anti-self-duality operator acting on $\mathfrak{g}_P$-valued 1-forms, defined using the metric $\check{g}$. Let $\mathcal{D}$ be the operator

$$\mathcal{D} = d^+_A \oplus -d^*_A \tag{7}$$

acting on the spaces

$$\check{L}^2_k(\check{X}; \mathfrak{g}_{\check{P}} \otimes \Lambda^1) \to \check{L}^2_{k-1}(\check{X}; \mathfrak{g}_{\check{P}} \otimes (\Lambda^+ \oplus \Lambda^0)).$$

Then in the orbifold setting $\mathcal{D}$ is a Fredholm operator. (See for example [15] and compare with [20, Proposition 4.17].)

We now wish to define a moduli space of anti-self-dual connections as

$$M(X, \Sigma, \mathbf{P}) = \{\, A \in \mathcal{C}_k \mid F^+_A = 0 \,\} \,/\, \mathcal{G}_{k+1}.$$

Following [20], there is a Kuranishi model for the neighborhood of a connection $[A]$ in $M(X, \Sigma, \mathbf{P})$ described by a Fredholm complex. The Kuranishi theory then tells us, in particular, that if A is irreducible and the operator d_A^+ is surjective, then a neighborhood of $[A]$ in $M(X, \Sigma, \mathbf{P})$ is a smooth manifold, and its dimension is equal to the index of $\mathcal{D}$.

2.7 Examples of moduli spaces

The quotient of $\bar{\mathbb{CP}}^2$ by the action of complex conjugation can be identified with S^4, containing a copy of $\mathbb{RP}^2$ as branch locus. The self-intersection number of this $\mathbb{RP}^2$ in S^4 is $+2$. The pair $(S^4, \mathbb{RP}^2)$ obtains an orbifold metric from the standard Riemannian metric on $\bar{\mathbb{CP}}^2$. We shall describe the corresponding moduli spaces $M(S^4, \mathbb{RP}^2, \mathbf{P})$ for various choices of singular bundle data $\mathbf{P}$.

On $\bar{\mathbb{CP}}^2$ with the standard Riemannian metric, there is a unique anti-self-dual $SO(3)$ connection with $\kappa = 1/4$. This connection $A_{\bar{\mathbb{CP}}^2}$ is reducible and has non-zero w_2: it splits as $\mathbb{R} \oplus L$, where L is an oriented 2-plane bundle with $e(L)[\mathbb{CP}^1] = 1$. (See [7] for example.) We can view L as the tautological line bundle on $\bar{\mathbb{CP}}^2$, and as such we see that the action of complex conjugation on $\bar{\mathbb{CP}}^2$ lifts to an involution on L that is orientation-reversing on the fibers. This involution preserves the connection. Extending the involution to act as -1 on the $\mathbb{R}$ summand, we obtain an involution on the $SO(3)$ bundle, preserving the connection. The quotient by this involution is an anti-self-dual connection A on $S^4 \setminus \mathbb{RP}^2$ for the orbifold metric. It has $\kappa = 1/8$. This orbifold connection corresponds to singular bundle data $\mathbf{P}$ on $(S^4, \mathbb{RP}^2)$ with Δ trivial. The connection is irreducible, and it is regular (because d^+ is surjective when coupled to $A_{\bar{\mathbb{CP}}^2}$ upstairs).

This anti-self-connection is unique, in the following strong sense, amongst solutions with Δ trivial and $\kappa = 1/8$. To explain this, suppose that we have $[A] \in M(S^4, \mathbb{RP}^2, \mathbf{P})$ and $[A'] \in M(S^4, \mathbb{RP}^2, \mathbf{P}')$ are two solutions with $\kappa = 1/8$, and that trivializations of the corresponding local systems Δ and Δ' are given. When lifted to $\bar{\mathbb{CP}}^2$, both solutions must give the same $SO(3)$ connection $A_{\bar{\mathbb{CP}}^2}$ up to gauge transformation, in the bundle $\mathbb{R} \oplus L$. There are two different ways to lift the involution to L on $\bar{\mathbb{CP}}^2$, differing in overall sign, but these two involutions on L are intertwined by multiplication by i on L. It follows that, as $SO(3)$-bundles with connection on $S^4 \setminus \mathbb{RP}^2$, the pairs (P, A) and (P', A') are isomorphic. Such an isomorphism of bundles with connection extends canonically to an isomorphism ψ from $\mathbf{P}$ to $\mathbf{P}'$. A priori, ψ may not preserve the given trivializations of Δ and Δ'. However, the connection A on P has structure group which reduces to $O(2)$, so it has a $\mathbb{Z}/2$ stabilizer in the $SO(3)$ gauge group on $S^4 \setminus \mathbb{RP}^2$. The non-trivial element of this stabilizer is an automorphism of P that extends to an automorphism of P_Δ covering the non-trivial involution t on X_Δ. So ψ can always be chosen to preserve the chosen trivializations.

When Δ is trivialized, singular bundle data $\mathbf{P}$ is classified by the evaluation of $p_1(P_\Delta)$ on $[D]$ and on $[X_+]$ in the notation of subsection 2.4, or equivalently by its instanton and monopole numbers k and l. The uniqueness of $[A]$ means that the corresponding singular bundle data $\mathbf{P}$ must be invariant under the involution t, which in turn means that $p_1(P_\Delta)[D]$ must be zero. Using the formulae for k, l and κ, we see that $\mathbf{P}$ has $k = 0$ and $l = 1/2$; or equivalently, $p_1(P_\Delta) = 0$. We summarize this discussion with a proposition.

Proposition 2.10. *Let $(S^4, \mathbb{RP}^2)$ be as above, so that the branched double-cover is $\bar{\mathbb{CP}}^2$. Fix a trivial and trivialized double-cover Δ of $\mathbb{RP}^2$, and let S^4_Δ be the corresponding space. Then, amongst singular bundle data with $\kappa = 1/8$, there is exactly one $P_\Delta \to S^4_\Delta$ with a non-empty moduli space, namely the one with $p_1(P_\Delta) = 0$ in $H^4(S^4_\Delta) = \mathbb{Z} \oplus \mathbb{Z}$. The corresponding moduli space is a single point, corresponding to an irreducible, regular solution.* □

On the same pair $(S^4, \mathbb{RP}^2)$, there is also a solution in a moduli space corresponding to singular bundle data $\mathbf{P}$ with Δ non-trivial. This solution can be described in a similar manner to the previous one, but starting with *trivial* $SO(3)$ bundle on $\bar{\mathbb{CP}}^2$, acted on by complex conjugation, lifted as an involution on the bundle as an element of order 2 in $SO(3)$. The resulting solution on $(S^4, \mathbb{RP}^2)$ has $\kappa = 0$ and Δ non-trivial. Knowing that $\kappa = 0$ is enough to pin down $P_\Delta \to S^4_\Delta$ up to isomorphism in this case, because $H^4(S^4_\Delta)$ is now $\mathbb{Z}$. This solution $[B]$ is reducible. It is regular, for similar reasons as arise in the previous case. The index of $\mathcal{D}$ in this case is therefore -1.

One can also consider the quotient of $\mathbb{CP}^2$, rather then $\bar{\mathbb{CP}}^2$, with respect to the same involution, which leads to a pair $(S^4, \mathbb{RP}^2_-)$ with $\mathbb{RP}^2_- \cdot \mathbb{RP}^2_0 = -2$. There is an isolated anti-self-dual $PSU(2)$ connection on $\mathbb{CP}^2$, arising from the $U(2)$ connection given by the Levi-Civita derivative in $T\mathbb{CP}^2$. This gives rise to a solution on $(S^4, \mathbb{RP}^2_-)$ with Δ trivial and $\kappa = 3/8$. As solutions with Δ trivialized (rather than just trivial), this solution gives rise to two different solutions, with $(k, l) = (0, 1/2)$ and $(k, l) = (1, -3/2)$, as the reader can verify. The operator $\mathcal{D}$ is again invertible for these two solutions. The flat connection $[B]$ on $\mathbb{CP}^2 \setminus \mathbb{RP}^2$ provides another solution, with $\kappa = 0$ and Δ non-trivial. This solution is reducible and the operator d^+_B now has 2-dimensional cokernel, so that $\mathcal{D}$ has index -3.

2.8 The dimension formula

We next compute the index of the operator $\mathcal{D}$ for a connection A in $\mathcal{C}_k(X, \Sigma, \mathbf{P})$. The index of $\mathcal{D}$ will coincide with the dimension of the moduli space $M(X, \Sigma, \mathbf{P})$ in the neighborhood of any irreducible, regular solution.

Lemma 2.11. *The index of $\mathcal{D}$ is given by*

$$8\kappa(A) - \frac{3}{2}\big(\chi(X) + \sigma(X)\big) + \chi(\Sigma) + \frac{1}{2}(\Sigma \cdot \Sigma)$$

where κ is again the action (5)*.*

Proof. For the case of orientable surfaces with Δ trivial, this formula reduces to the formula proved in [20], where it appears as

$$8k + 4l - \frac{3}{2}\big(\chi(X) + \sigma(X)\big) + \chi(\Sigma). \tag{8}$$

Just as in [20], the general case can be proved by repeatedly applying excision and the homotopy invariance of the index, to reduce the problem to a few model cases. In addition to the model cases from the proof in [20], it is now necessary to treat one model case in which $w_1(\Delta)^2$ is non-zero on Σ. Such an example is provided by the flat connection $[B]$ on $(S^4, \mathbb{RP}^2)$ from the previous subsection. In this case, κ is 0 and the formula in the lemma above predicts that the index of $\mathcal{D}$ should be -3. This is indeed the index of $\mathcal{D}$ in this case, as we have already seen. □

2.9 Orientability of moduli spaces

We continue to consider the moduli space $M(X, \Sigma, \mathbf{P})$ associated to a closed pair (X, Σ) equipped with an orbifold metric and singular bundle data $\mathbf{P}$. The irreducible, regular solutions form a subset of $M(X, \Sigma, \mathbf{P})$ that is a smooth manifold of the dimension given by Lemma 2.11, and our next objective is to show that this manifold is orientable. As usual, the orientability of the moduli space is better expressed as the triviality of the real determinant line of the family of operators $\mathcal{D}$ over the space $\mathcal{B}^*_k(X, \Sigma, \mathbf{P})$ of all irreducible connections modulo the determinant-1 gauge group.

Proposition 2.12. *The real line bundle* $\det \mathcal{D}$ *of the family of operators $\mathcal{D}$ over $\mathcal{B}^*_k(X, \Sigma, \mathbf{P})$ is trivial.*

Proof. The proof follows that of the corresponding result in [21], which in turn is based on [6]. We must show that the determinant line is orientable along all closed loops in $\mathcal{B}^*_k(X,\Sigma,\mathbf{P})$. The fundamental group of $\mathcal{B}^*_k(X,\Sigma,\mathbf{P})$ is isomorphic to the group of components of $\mathcal{G}_{k+1}(X,\Sigma,\mathbf{P})/\{\pm 1\}$. The group $\pi_0(\mathcal{G}_{k+1}(X,\Sigma,\mathbf{P})$ is the same as $\pi_0(\mathcal{G}^{\text{top}})$, for which explicit generators can be extracted from Lemmas 2.8 and 2.9. These generators correspond to loops in $X\backslash\Sigma$, loops in Σ along which Δ is trivial, and a possible additional $\mathbb{Z}/2$. Of these, the only type of generator that is new in the present paper is a loop Σ along which Δ is trivial but Σ is non-orientable. (The orientable case is essentially dealt with in [21].)

So let γ be such a loop in Σ along which Δ is trivialized. There is an element g in $\pi_0(\mathcal{G}_{k+1})$ which maps to the class $[\gamma]$ in $H_1(\Sigma;\mathbb{Z}_\Delta)$ under the map in Lemma 2.8. We wish to describe loop γ^* in $\mathcal{B}^*_k$ that represents a corresponding element in $\pi_1(\mathcal{B}^*_k)$. As in [21, Appendix 1(i)], we construct γ^* by gluing in a monopole at and "dragging it around γ". To do this, we first let $\mathbf{P}'$ be singular bundle data such that $\mathbf{P}$ is obtained from $\mathbf{P}'$ by adding a monopole. We fix a connection A' in $\mathbf{P}'$. We let J be a standard solution on (S^4,S^2) with monopole number 1 and instanton number 0, carried by singular bundle data with Δ trivialized. For each x in γ, we form a connected sum of pairs,

$$(X,\Sigma)\#_x(S^4,S^2)$$

carrying a connection $A(x)=A'\#J$. Even though Σ is non-orientable along Δ, a closed loop in $\mathcal{B}_k(X,\Sigma,\mathbf{P})$ can be constructed this way, because the solution on (S^4,S^2) admits a symmetry which reverses the orientation of S^2 while preserving the trivialization of Δ. This is the required loop γ^*. If we write $\delta_{A(x)}$ and $\delta_{A'}$ for the determinant lines of $\mathcal{D}$ for the two connections, then (much as in [21]) we have the relation

$$\delta_{A(x)}=\delta_{A'}\otimes\det(T_x\Sigma\oplus\mathbb{R}\oplus\mathbb{R}_Q)$$

where the first $\mathbb{R}$ can be interpreted as the scale parameter in the gluing and the factor $\mathbb{R}_Q$ is the real orientation bundle of Q, which arises as the tangent space to the S^1 gluing parameter. Since the product of the orientation bundles of Σ and Q is trivial along γ (being the orientation bundle of Δ), we see that δ is trivial along the loop γ^*. □

The factor $\det(T_x\Sigma\oplus\mathbb{R}\oplus\mathbb{R}_Q)$ above can be interpreted as the orientation line on the moduli space of framed singular instantons on the orbifold (S^4,S^2) (or in other words, the moduli space of finite-action solutions on $(\mathbb{R}^4,\mathbb{R}^2)$ modulo gauge transformations that are asymptotic to 1 at infinity. This is a complex space, and therefore has a preferred orientation, for any choice of instanton and monopole charges. In the situation that arises in the proof of the previous proposition, there is therefore a preferred way to orient $\delta_{A(x)}$, given an orientation of $\delta_{A'}$. To make use of this, we consider the following setup. Let $\mathbf{P}$ be singular bundle data, given on (X, Σ), and let us consider the set of pairs $(\mathbf{P}', g')$, where $\mathbf{P}'$ is another choice of singular bundle data and

$$g' : \mathbf{P}'|_{(X_\Delta \setminus \mathbf{x}')} \to \mathbf{P}|_{(X_\Delta \setminus \mathbf{x}')}$$

is an isomorphism defined on the complement of a finite set $\mathbf{x}'$. We say that $(\mathbf{P}', g')$ is isomorphic to $(\mathbf{P}'', g'')$ if there is an isomorphism of singular bundle data, $h : \mathbf{P}' \to \mathbf{P}''$, such that the composite $g'' \circ (g')^{-1}$ can be lifted to a determinant-1 gauge transformation on its domain of definition, $X_\Delta \setminus (\mathbf{x}' \cup \mathbf{x}'')$.

Definition 2.13. We refer to such a pair $(\mathbf{P}', g')$ as a $\mathbf{P}$-marked bundle. $\diamondsuit$

The classification of the isomorphism classes of $\mathbf{P}$-marked bundles on (X, Σ) can be deduced from the material of section 2.3. Every $\mathbf{P}$-marked bundle can be obtained from $\mathbf{P}$ by "adding instantons and monopoles". Furthermore, as in (a) and (b) on page 14, adding a monopole and an anti-monopole to the same orientable component of Σ is equivalent to adding an instanton, while adding two monopoles to the same non-orientable component is also the same as adding an instanton. There are no other relations: in particular, because of the determinant-1 condition in our definition of the equivalence relation, there is no counterpart here of the relation (c) from page 14. We now have, as in [20]:

Proposition 2.14. *Using the complex orientations of the framed moduli spaces on* (S^4, S^2)*, an orientation for the determinant line of* $\mathcal{D}$ *over* $\mathcal{B}(X, \Sigma, \mathbf{P})$ *determines an orientation of the determinant line also over* $\mathcal{B}(X, \Sigma, \mathbf{P}')$ *for all* $\mathbf{P}$*-marked bundles* $(\mathbf{P}', g')$*. These orientations are compatible with equivalence of* $\mathbf{P}$*-marked bundles, in that if* $h : \mathbf{P}' \to \mathbf{P}''$ *is an equivalence (as above), then the induced map on the corresponding determinant line is orientation-preserving.* $\square$

Remark. The reason for the slightly complex setup in the above proposition is that we do not know in complete generality whether gauge transformations that do not belong to the determinant-1 gauge group give rise to orientation-preserving maps on the moduli spaces $M(X, \Sigma, \mathbf{P})$. We will later obtain some partial results in this direction, in section 5.1; but those results apply, as they stand, only to the case that Σ is orientable.

3 Singular instantons and Floer homology for knots

In this section we review how to adapt the gauge theory for singular connections on pairs (X, Σ) to the three-dimensional case of a link K in a 3-manifold Y, as well as the case of a 4-manifold with cylindrical ends. This is all adapted from [23]: as in section 2, the new ingredient is that we are no longer assuming that our $SO(3)$ bundle P on $Y \backslash K$ extends over K.

3.1 Singular connections in the 3-dimensional case

We fix a closed, oriented, connected three manifold Y containing a knot or link K. The definition of *singular bundle data* from Definition 2.1 adapts in a straightforward way to this 3-dimensional case: such data consists of a double-cover $K_\Delta \to K$, an $SO(3)$ bundle P_Δ on the corresponding non-Hausdorff space Y_Δ, and a reduction of structure group to $O(2)$ in a neighborhood of $K_\Delta \subset Y_\Delta$, taking a standard form locally along K_Δ. Note that K is always orientable, but has not been oriented. A choice of orientation for K will fix an isomorphism between the local system Δ and the orientation bundle of the $O(2)$ reduction Q in the neighborhood of K. For a given Δ, we can also form the Hausdorff space Y^h_Δ, which contains a copy of 2-sphere bundle over K (non-orientable if Δ is non-trivial). The $SO(3)$-bundle P_Δ has w_2 non-zero on the S^2 fibers. Such $w_2's$ form an affine copy of $H^2(Y; \mathbb{Z}/2)$ in $H^2(Y_\Delta; \mathbb{Z}/2)$, and they classify singular bundle data for the given Δ.

Equipping Y with an orbifold structure along K with cone angle $\pi/2$ and a compatible orbifold Riemannian metric $\check{g}$, we can regard singular bundle data $\mathbf{P}$ as determining an orbifold bundle $\check{P} \to \check{Y}$, as in the 4-dimensional case. Thus we construct Sobolev spaces

$$\check{L}^2_k(\check{Y}; \mathfrak{g}_{\check{P}} \otimes \Lambda^q)$$

as before, leading to spaces of $SO(3)$ connections $\mathcal{C}_k(Y, K, \mathbf{P})$ and determinant-1 gauge transformations $\mathcal{G}_{k+1}(Y, K, \mathbf{P})$. The space of connections $\mathcal{C}_k(Y, K, \mathbf{P})$ is an affine space, and on the tangent space

$$T_B\mathcal{C}_k(Y, K, \mathbf{P}) = \check{L}^2_k(\check{Y}; \mathfrak{g}_{\check{P}} \otimes T^*\check{Y})$$

we define an L^2 inner product (independent of B) by

$$\langle b, b' \rangle_{L^2} = \int_Y -\mathrm{tr}(*b \wedge b'), \tag{9}$$

Here tr denotes the Killing form on $\mathfrak{su}(2)$ and the Hodge star is the one defined by the singular metric $\check{g}$. We have the *Chern-Simons functional* on

$\mathcal{C}_k(Y, K, \mathbf{P})$ characterized by

$$(\operatorname{grad} \mathrm{CS})_B = *F_B.$$

Critical points of CS are the flat connections in $\mathcal{C}_k(Y, K, \mathbf{P})$. We denote the set of gauge equivalence classes of critical points by $\mathfrak{C} = \mathfrak{C}(Y, K, \mathbf{P})$.

3.2 The components of the gauge group on a three manifold.

We can analyze the component group $\pi_0(\mathcal{G}_{k+1}(Y, K, \mathbf{P}))$ much as we did in the 4-dimensional case. The group $\mathcal{G}_{k+1}(Y, K, \mathbf{P})$ is homotopy equivalent to the $\mathcal{G}^{\mathrm{top}}$, the continuous gauge transformations of P_Δ respecting the reduction, and we have a short exact sequence,

$$0 \to \mathbb{Z} \to \pi_0(\mathcal{G}_{k+1}(Y, K, \mathbf{P})) \to H_0(K; \mathbb{Z}_\Delta) \to 0.$$

(See Lemma 2.8.) The $\mathbb{Z}$ in the kernel has a generator g_1 represented by a gauge transformation supported in a ball disjoint from K. The group $H_0(K; \mathbb{Z}_\Delta)$ arises as $\pi_0(\mathcal{H})$, where $\mathcal{H}$ is the group of sections of the bundle $H_\Delta \to K$ with fiber S^1. The group $H_0(K; \mathbb{Z}_\Delta)$ is a direct sum of one copy of $\mathbb{Z}$ for each component of K on which Δ is trivial and one copy of $\mathbb{Z}/2$ for each component on which Δ is non-trivial.

The above sequence is not split in general. We can express the component group $\pi_0(\mathcal{G}_{k+1}(Y, K, \mathbf{P}))$ as having generators g_1 (in the kernel) and one generator h_i for each component K_i of K, subject to the relations

$$2h_i = g_1$$

whenever $\Delta|_{K_i}$ is non-trivial.

The Chern-Simons functional is invariant under the identity component of the gauge group. Under the generators g_1 and h_i it behaves as follows:

$$\begin{aligned} \mathrm{CS}(g_1(A)) &= \mathrm{CS}(A) - 4\pi^2 \\ \mathrm{CS}(h_i(A)) &= \mathrm{CS}(A) - 2\pi^2. \end{aligned}$$

3.3 Reducible connections and the non-integral condition

For the construction of Floer homology it is important to understand whether there are reducible connections in $\mathcal{C}(Y, K, \mathbf{P})$, or at least whether the critical points of CS are irreducible.

This is discussed in [23] for the case that P extends to Y (or equivalently, the case that Δ is trivial). See in particular the "non-integral condition" of [23, Definition 3.28]. For this paper, we make the following adaptation of the definition, whose consequences are summarized in the following proposition.

Definition 3.1. Let singular bundle data $\mathbf{P}$ on (Y, K) be given. We say that an embedded closed oriented surface Σ is a *non-integral surface* if either

- Σ is disjoint from K and $w_2(P)$ is non-zero on Σ; or
- Σ is transverse to K and $K \cdot \Sigma$ is odd.

We say that $\mathbf{P}$ satisfies the *non-integral condition* if there is a non-integral surface Σ in Y. $\Diamond$

Proposition 3.2. *If singular bundle data $\mathbf{P}$ on (Y, K) satisfies the non-integral condition, then the Chern-Simons functional on $\mathcal{C}_k(Y, K, \mathbf{P})$ has no reducible critical points. Furthermore, if Δ is non-trivial on any component of K, then a stronger conclusion holds: the configuration space $\mathcal{C}_k(Y, K, \mathbf{P})$ contains no reducible connections at all.*

Proof. For the first part, the point is that there are already no reducible flat connections in the restriction of $\mathbf{P}$ to Σ: since Δ becomes trivial when restricting to the surface, there is nothing new in this statement beyond the familiar case where $\mathbf{P}$ extends across K.

If Δ is non-trivial on a component K' of K, then if T' is a torus which is the boundary of a small tubular neighborhood of K' we have

$$\langle w_2(P), [T'] \rangle \neq 0.$$

Since T' is disjoint from K, we see that T' is a non-integral surface, and so the non-integral condition is automatically satisfied. The asymptotic holonomy group of a connection in $\mathcal{C}(Y, K, \mathbf{P})$ in the neighborhood of K' lies in $O(2)$ and contains both (a) an element of $SO(2) \subset O(2)$ of order 2 (the meridional holonomy), and (b) an element of $O(2) \setminus SO(2)$, namely the holonomy along a longitude. Such a connection therefore cannot be reducible, irrespective of whether it is flat or not. $\square$

3.4 Perturbations

We now introduce standard perturbations of the Chern-Simons functional to achieve suitable transversality properties for the both the set of critical points and the moduli spaces of trajectories for the formal gradient flow. Section 3.2 of [23] explains how to do this, following work of Taubes [33] and Donaldson; and the approach described there needs almost no modification in the present context.

The basic function used in constructing perturbations is obtained as follows. Choose a lift of the bundle $P \to Y \setminus K$ to a $U(2)$ bundle $\tilde{P} \to Y \setminus K$, and fix a connection θ on $\det \tilde{P}$. Each B in $\mathcal{C}_k(Y, K, \mathbf{P})$ then gives rise to a connection $\tilde{B}$ in P inducing the connection θ in $\det \tilde{P}$. Take an immersion $q : S^1 \times D^2$ to $Y \backslash K$, and choose a base point $p \in S^1$. For each $x \in D^2$ the holonomy of a connection $\tilde{B}$ about $S^1 \times x$ starting at (p, x) gives an element $\mathrm{Hol}_x(\tilde{B}) \in U(2)$. Taking a class function $h : U(2) \to \mathbb{R}$, we obtain a gauge-invariant function

$$H_x : \mathcal{C}(Y, K, \mathbf{P}) \to \mathbb{R}$$

as the composite $H_x(B) = h \circ \mathrm{Hol}_x(\tilde{B})$. For analytic purposes it is useful to mollify this function by introducing

$$f_q(B) = \int_{D^2} H_x(B)\mu$$

where μ is choice of volume form on D^2 which we take to be supported in the interior of D^2 and have integral 1.

More generally taking a collection of such immersions $\mathbf{q} = (q_1, \ldots, q_l)$ so that they all agree on $p \times D^2$, the holonomy about the l loops determined by $x \in D^2$ gives a map $\mathrm{Hol}_x : \mathcal{C}_k \to U(2)^l$. Now taking an conjugation invariant function $h : U(2)^l \to \mathbb{R}$ we obtain again a gauge invariant function $H_x = h \circ \mathrm{Hol}_x \colon \mathcal{C}_k \to \mathbb{R}$. Mollifying this we obtain

$$f_{\mathbf{q}} : \mathcal{C}_k(Y, K, \mathbf{P}) \to \mathbb{R}$$
$$f_{\mathbf{q}}(B) = \int_{D^2} H_x(B)\mu.$$

These are smooth gauge invariant functions on $\mathcal{C}$ and are called *cylinder functions.*

Our typical perturbation f will be an infinite linear combination of such cylinder functions. In [23] it is explained how to construct an infinite collection $\mathbf{q}^i$ of immersions and a separable Banach space $\mathcal{P}$ of sequences $\pi = \{\pi_i\}$, with norm

$$\|\pi\|_{\mathcal{P}} = \sum_i C_i |\pi_i|$$

such that for each $\pi \in \mathcal{P}$, the sum

$$f_\pi = \sum_i \pi_i f_{\mathbf{q}^i}$$

is convergent and defines a smooth, bounded function f_π on $\mathcal{C}_k$. Furthermore, the formal L^2 gradient of f_π defines a smooth vector field on the Banach space $\mathcal{C}_k$, which we denote by V_π. The analytic properties that we require for V_π (all of which can be achieved by a suitable choice of $\mathcal{P}$) are summarized in [23, Proposition 3.7].

We refer to a function f_π of this sort as a *holonomy perturbation.* Given such a function, we then consider the perturbed Chern-Simons functional $\mathrm{CS}+f_\pi$, The set of gauge equivalence classes of critical points for the perturbed functional is denoted

$$\mathfrak{C}_\pi(Y, K, P, \rho) \subset \mathcal{B}_k(Y, K, \mathbf{P}),$$

or simply as $\mathfrak{C}_\pi$. Regarding the utility of these holonomy perturbations, $\mathcal{P}$ we have [23, Proposition 3.12]:

Proposition 3.3. *There is a residual subset of the Banach space $\mathcal{P}$ such that for all π in this subset, all the irreducible critical points of the perturbed functional $\mathrm{CS} + f_\pi$ in $\mathcal{C}^*(Y, K, P, \rho)$ are non-degenerate in the directions transverse to the gauge orbits.* □

This says nothing yet about the reducible critical points. We will be working eventually with configurations $(Y, K, \mathbf{P})$ satisfying the non-integral condition of Definition 3.1. In case the bundle does not extend there can be no reducible connections at all (by the second part of Proposition 3.2), but in the case where P does extend, there may be reducible critical points if the perturbation is large. However, for small perturbations, there are none:

Lemma 3.4 (Lemma 3.11 of [23]). *Suppose $(Y, K, \mathbf{P})$ satisfies the non-integral condition. Then there exists $\epsilon > 0$ such that for all π with $\|\pi\|_{\mathcal{P}} \le \epsilon$, the critical points of $\mathrm{CS} + f_\pi$ in $\mathcal{C}_k(Y, K, \mathbf{P})$ are all irreducible.* □

This lemma is corollary of the compactness properties of the perturbed critical set: in particular, the fact that the projection

$$\mathfrak{C}_\bullet \to \mathcal{P}$$

from the parametrized critical-point set $\mathfrak{C}_\pi \subset \mathcal{P} \times \mathcal{B}_k(Y, K, \mathbf{P})$ is a proper map. This properness also ensures that the $\mathfrak{C}_\pi$ is finite whenever all the critical points of the perturbed functional are non-degenerate.

3.5 Trajectories for the perturbed gradient flow

Given a holonomy perturbation f_π for the Chern-Simons functional on the space $\mathcal{C}_k(Y, K, \mathbf{P})$, we get a perturbation of the anti-self-duality equations on the cylinder. The "cylinder" here is the product pair,

$$(Z, S) = \mathbb{R} \times (Y, K)$$

viewed as a 4-manifold with an embedded surface. Since Y is equipped with a singular Riemannian metric to become an orbifold $\check{Y}$, so also Z obtains a product orbifold structure: it becomes an orbifold $\check{Z}$ with singularity along S, just as in our discussion from section 2.6 for a closed pair (X, Σ).

To write down the perturbed equations, let the 4-dimensional connection be expressed as

$$A = B + c\,dt,$$

with B a t-dependent (orbifold) connection on (Y, K) and c a t-dependent section of $\mathfrak{g}_{\check{P}}$. We write

$$\hat{V}_\pi(A) = P_+(dt \wedge V_\pi(B))$$

where P_+ the projection onto the self-dual 2-forms, and $V_\pi(B)$ is viewed as a $\mathfrak{g}_{\check{P}}$-valued 1-form on $\mathbb{R} \times \check{Y}$ which evaluates to zero on multiples of d/dt. The 4-dimensional self-duality equations on $\mathbb{R} \times Y$, perturbed by the holonomy perturbation V_π, are the equations

$$F_A^+ + \hat{V}_\pi(A) = 0. \tag{10}$$

These equations are invariant under the 4-dimensional gauge group. Solutions to the downward gradient flow equations for the perturbed Chern-Simons functional correspond to solutions of these equations which are in temporal gauge (i.e. have $c = 0$ in the above decomposition of A). The detailed mapping properties of $\hat{V}_\pi$ and its differential are given in Proposition 3.15 of [23].

Let π be chosen so that all critical points in $\mathfrak{C}_\pi$ are irreducible and non-degenerate. Let B_1 and B_0 be critical points in $\mathcal{C}_k(Y, K, \mathbf{P})$, and let $\beta_1, \beta_0 \in \mathfrak{C}_\pi$ be their gauge-equivalence classes. Let A_o be a connection on $\mathbb{R} \times Y$ which agrees with the pull-back of B_1 and B_0 for large negative and large positive t respectively. The connection A_o determines a path $\gamma : \mathbb{R} \to \mathcal{B}_k(Y, K, \mathbf{P})$, from β_1 to β_0, which is constant outside a compact set. The relative homotopy class

$$z \in \pi_1(\mathcal{B}_k, \beta_1, \beta_0)$$

of the path γ depends on the choice of B_1 and B_0 within their gauge orbit. Given A_o, we can construct a space of connections

$$\mathcal{C}_{k,\gamma}(Z, S, \mathbf{P}; B_1, B_0)$$

as the affine space

$$\{\, A \mid A - A_o \in \check{L}^2_{k,A_o}(\check{Z}; T^*\check{Z} \otimes \mathfrak{g}_{\check{P}}) \,\}.$$

There is a corresponding gauge group, the space of sections of the bundle $G(P) \to Z\backslash S$ defined by

$$\mathcal{G}_{k+1}(Z,S,\mathbf{P}) = \{\, g \mid \ \nabla_{A_o} g, \dots, \nabla^k_{A_o} g \in \check{L}^2(Z\backslash S)\,\}.$$

We have the quotient space

$$\mathcal{B}_{k,z}(Z,S,\mathbf{P};\beta_1,\beta_0) = \mathcal{C}_{k,\gamma}(Z,S,\mathbf{P};B_1,B_0)/\mathcal{G}_{k+1}(Z,S,\mathbf{P})$$

Here z again denotes the homotopy class of γ in $\pi_1(\mathcal{B}_k(Y,K,\mathbf{P});\beta_1,\beta_0)$.

We can now construct the moduli space of solutions to the perturbed anti-self-duality equations (the π-ASD connections) as a subspace of the above space of connections modulo gauge:

$$M_z(\beta_1,\beta_0) = \big\{\, [A] \in \mathcal{B}_{k,z}(Z,S,\mathbf{P};\beta_1,\beta_0) \bigm| F^+_A + \hat{V}_\pi(A) = 0 \,\big\}.$$

This space is homeomorphic to the space of trajectories of the formal downward gradient-flow equations for $\mathrm{CS} + f_\pi$ running from β_1 to β_0 in the relative homotopy class z. Taking the union over all z, we write

$$M(\beta_1,\beta_0) = \bigcup_z M_z(\beta_1,\beta_0).$$

The action of $\mathbb{R}$ by translations on $\mathbb{R}\times Y$ induces an action on $M(\beta_1,\beta_0)$. The action is free except in the case of $M(\beta_1,\beta_1)$ and constant trajectory. The quotient of the space of non-constant solutions by this action of $\mathbb{R}$ is denoted $\breve{M}(\beta_1,\beta_0)$ and typical elements are denoted $[\breve{A}]$.

The linearization of the π-ASD condition at a connection A in $\mathcal{C}_{k,\gamma}(B_1,B_0)$ is the map

$$d^+_A + D\hat{V}_\pi : \check{L}^2_{k,A}(\check{Z};\Lambda^1\otimes\mathfrak{g}_{\check{P}}) \to \check{L}^2_{k-1}(\check{Z};\Lambda^+\otimes\mathfrak{g}_{\check{P}}).$$

When B_1 and B_1 are irreducible and non-degenerate, we have a good Fredholm theory for this linearization together with gauge fixing. For A as above, we write (as in (7), but now with the perturbation)

$$\mathcal{D}_A = (d^+_A + D\hat{V}_\pi) \oplus -d^*_A$$

which we view as an operator

$$\check{L}^2_{k,A}(\check{Z};\Lambda^1\otimes\mathfrak{g}_{\check{P}}) \to \check{L}^2_{k-1}(\check{Z};(\Lambda^+\oplus\Lambda^0)\otimes\mathfrak{g}_{\check{P}})$$

Viewed this was, $\mathcal{D}_A$ is a Fredholm operator. When $\mathcal{D}_A$ is surjective we say that A is a *regular* solution, and in this case $M_z(\beta_1,\beta_0)$ is a smooth manifold near the gauge equivalence class of $[A]$, of dimension equal to the index of $\mathcal{D}_A$. The index of the operator, which can be interpreted as a spectral flow, will be denoted by

$$\mathrm{gr}_z(\beta_1, \beta_0).$$

We will need the moduli spaces of fixed relative grading

$$M(\beta_1, \beta_0)_d = \bigcup_{z \,|\, \mathrm{gr}_z = d} M_z(\beta_1, \beta_0)$$

For the four-dimensional equations we have the following transversality result.

Proposition 3.5 ([23, Proposition 3.18]). *Suppose that π_0 is a perturbation such that all the critical points in $\mathfrak{C}_{\pi_0}$ are non-degenerate and have stabilizer ± 1. Then there exists $\pi \in \mathcal{P}$ such that:*

(a) *$f_\pi = f_{\pi_0}$ in a neighborhood of all the critical points of* $\mathrm{CS} + f_{\pi_0}$*;*

(b) *the set of critical points for these two perturbations are the same, so that $\mathfrak{C}_\pi = \mathfrak{C}_{\pi_0}$;*

(c) *for all critical points β_1 and β_0 in $\mathfrak{C}_\pi$ and all paths z, the moduli spaces $M_z(\beta_1, \beta_0)$ for the perturbation π are regular.* □

From this point on, we will always suppose that our perturbation has been chosen in this way. For reference, we state:

Hypothesis 3.6. We assume that Y is a connected, oriented, closed 3-manifold, that K is a link in Y (possibly empty), and that $\mathbf{P}$ is singular bundle data satisfying the non-integral condition. We suppose that an orbifold metric $\check{g}$ and perturbation $\pi \in \mathcal{P}$ are chosen so that $\mathfrak{C}_\pi$ consists only of non-degenerate, irreducible critical points, and all the moduli spaces $M_z(\beta_1, \beta_0)$ are regular.

3.6 Orientations and Floer homology

The Fredholm operators $\mathcal{D}_A$ form a family over the space $\mathcal{B}_{k,z}(\beta_1, \beta_0)$ whose determinant line $\det(\mathcal{D}_A)$ is orientable. This follows from the corresponding result for the closed pair $(X, \Sigma) = S^1 \times (Y, K)$ (Proposition 2.12) by an application of excision. It follows that the (regular) moduli space $M_z(\beta_1, \beta_0)$ is an orientable manifold. Moreover, given two different paths z and z' between the same critical points, if we choose an orientation for the determinant line over $\mathcal{B}_{k,z}(\beta_1, \beta_0)$, then it canonically orients the determinant line over $\mathcal{B}_{k,z'}(\beta_1, \beta_0)$ also: this follows from the corresponding result for closed manifolds (Proposition 2.14), because any two paths are related by the addition of instantons and monopoles. We may therefore define

$$\Lambda(\beta_1, \beta_0)$$

as the two-element set of orientations of $\det(\mathcal{D}_A)$ over $\mathcal{B}_{k,z}(\beta_1, \beta_0)$, with the understanding that this is independent of z.

If β_1 and β_0 are arbitrary connections in $\mathcal{B}_k(Y, K, \mathbf{P})$ rather than critical points, then we can still define $\Lambda(\beta_1, \beta_0)$ in essentially the same way. The only point to take care of is that, if the Hessian of the perturbed functional is singular at either β_1 or β_0, then the corresponding operator $\mathcal{D}_A$ is not Fredholm on the usual Sobolev spaces. As in [23], we adopt the convention that $\mathcal{D}_A$ is considered as a Fredholm operator acting on the weighted Sobolev spaces

$$e^{-\epsilon t} \check{L}^2_{k,A_o} \tag{11}$$

for a small positive weight ϵ; and we then define $\Lambda(\beta_1, \beta_0)$ for any β_i using this convention.

In particular, we can choose any basepoint θ in $\mathcal{B}_k(Y, K, \mathbf{P})$ and *define*

$$\Lambda(\beta) = \Lambda(\theta, \beta) \tag{12}$$

for any critical point β. Without further input, there is no a priori way to rid this definition of its dependence on θ. In the case considered in [23], when K was oriented and Δ was trivial, we had a preferred choice of θ arising from a reducible connection. When Δ is non-trivial however, the space $\mathcal{B}_k(Y, K, \mathbf{P})$ contains no reducibles, so an arbitrary choice of θ is involved.

Having defined $\Lambda(\beta)$ in this way, we have canonical identifications

$$\Lambda(\beta_1, \beta_0) = \Lambda(\beta_1)\Lambda(\beta_0),$$

where the product on the right is the usual product of 2-element sets (defined, for example, as the set of bijections from $\Lambda(\beta_1)$ to $\Lambda(\beta_2)$). What this implies is that a choice of orientation for a component of the moduli space $M_z(\beta_1, \beta_0)$ (or equivalently, a choice of trivializations of the determinant on $\mathcal{B}_{k,z}(\beta_1, \beta_0)$) determines an identification $\Lambda(\beta_1) \to \Lambda(\beta_0)$. In particular, each one-dimensional connected component

$$[\breve{A}] \subset M(\beta_1, \beta_0)_1,$$

being just a copy of $\mathbb{R}$ canonically oriented by the action of translations, determines an isomorphism $\Lambda(\beta_1) \to \Lambda(\beta_0)$. As in [23, 22], we denote by

$\mathbb{Z}\Lambda(\beta)$ the infinite cyclic group whose two generators are the two elements of $\Lambda(\beta)$, and we denote by

$$\epsilon[\breve{A}] : \mathbb{Z}\Lambda(\beta_1) \to \mathbb{Z}\Lambda(\beta_0)$$

the resulting isomorphism of groups.

We now have everything we need to define Floer homology groups. Let (Y, K) be an unoriented link in a closed, oriented, connected 3-manifold Y, and let $\mathbf{P}$ be singular bundle data satisfying the non-integral condition, Definition 3.1. Let a metric $\check{g}$ and perturbation π be chosen satisfying Hypothesis 3.6. Finally, let a basepoint θ in $\mathcal{B}_k(Y, K, \mathbf{P})$ be chosen. Then we define the chain complex $(C_*(Y, K, \mathbf{P}), \partial)$ of free abelian groups by setting

$$C_*(Y, K, \mathbf{P}) = \bigoplus_{\beta \in \mathfrak{C}_\pi} \mathbb{Z}\Lambda(\beta), \tag{13}$$

and

$$\partial = \sum_{(\beta_1, \beta_0, z)} \sum_{[\breve{A}] \subset M(\beta_1, \beta_0)_1} \epsilon[\breve{A}] \tag{14}$$

where the first sum runs over all triples with $\mathrm{gr}_z(\beta_1, \beta_0) = 1$. That this is a finite sum follows from the compactness theorem, Corollary 3.25 of [23]. (As emphasized in [23], it is the compactness result here depends crucially on that fact that our choice of holonomy for our singular connections satisfies a "monotone" condition: that is, the formula for the dimension of moduli spaces in Lemma 2.11 involves the topology of the bundle $\mathbf{P}$ only though the action κ.

Definition 3.7. For (Y, K) as above, with singular bundle data $\mathbf{P}$ satisfying the non-integral condition, Definition 3.1, and choice of $\check{g}$, π and θ as above, we define the instanton Floer homology group

$$I(Y, K, \mathbf{P})$$

to be the homology of the complex $(C_*(Y, K, \mathbf{P}), \partial)$. ◊

As usual, we have presented the definition of $I(Y, K, \mathbf{P})$ as depending on some auxiliary choices. The standard type of cobordism argument (using the material from the following subsection) shows that $I(Y, K, \mathbf{P})$ is independent of the choice of $\check{g}$, π and θ. There is a slight difference from the usual presentation of (for example) [23] however, which stems from our lack of a canonical choice of basepoint θ. The result of this is that, if $(\check{g}, \pi, \theta)$ and $(\check{g}', \pi', \theta')$ are two choices for the auxiliary data, then the isomorphism between the corresponding homology groups $I(Y, K, \mathbf{P})$ and $I'(Y, K, \mathbf{P})$ is well-defined only up to an overall choice of sign.

3.7 Cobordisms and manifolds with cylindrical ends

Let (W, S) be a cobordism of pairs, from (Y_1, K_1) to (Y_0, K_0). We assume that W is connected and oriented, but S need not be orientable. Let $\mathbf{P}$ be singular bundle data on (W, S), and let $\mathbf{P}_i$ be its restriction to (Y_i, K_i). We shall recall the standard constructions whereby $(W, S, \mathbf{P})$ induces a map on the Floer homology groups,

$$I(W, S, \mathbf{P}) : I(Y_1, K_1, \mathbf{P}_1) \to I(Y_0, K_0, \mathbf{P}_0),$$

which is well-defined up to an overall sign.

To set this up, we assume that auxiliary data is given,

$$\begin{aligned} \mathbf{a}_1 &= (\check{g}_1, \pi_1, \theta_1) \\ \mathbf{a}_0 &= (\check{g}_0, \pi_0, \theta_0), \end{aligned} \tag{15}$$

on $(Y_1, K_1, \mathbf{P}_1)$ and $(Y_0, K_0, \mathbf{P}_0)$ respectively, and that our standing assumptions hold (Hypothesis 3.6). We equip the interior of W with an orbifold metric $\check{g}$ (with orbifold singularity along S) having two cylindrical ends

$$\begin{gathered} (-\infty, 0\,] \times \check{Y}_1 \\ [\,0, \infty) \times \check{Y}_0 \end{gathered}$$

where the metric $\check{g}$ is given by

$$\begin{gathered} dt^2 + \check{g}_1 \\ dt^2 + \check{g}_0 \end{gathered}$$

respectively. To do this, we consider an open collar neighborhood $[\,0, 1) \times \check{Y}_1$ of $\check{Y}_1$, with coordinate r for the first factor, and we set $t = \ln(r)$. Symmetrically, we take an open collar $(-1, 0\,] \times \check{Y}_0$ at the other end, and set $t = -\ln(-r)$ there.

Now we make the following choices (an adaptation, for the case of manifolds with boundary, of the definition of $\mathbf{P}$-marked bundle from Definition 2.13. We choose singular bundle data $\mathbf{P}'$ on (W, S) differing from $\mathbf{P}$ by the addition of instantons and monopoles, so that $\mathbf{P}$ and $\mathbf{P}'$ are identified outside a finite set by a preferred map $g' : \mathbf{P}' \to \mathbf{P}$. We choose gauge representatives B_1' and B_0' for β_1 and β_0, as connections in $\mathbf{P}_1$ and $\mathbf{P}_0$ respectively. And we choose an orbifold connection A_o in $\mathbf{P}'$ that coincides with the pull-back of B_1' and B_0' in the collar neighborhoods of the two ends. We can then construct as usual an affine space of connections

$$\mathcal{C}_k(\mathbf{P}', A_o)$$

consisting of all A with

$$A - A_o \in \check{L}_{k,A_o}(\check{W}, \Lambda^1 \otimes \mathfrak{g}_{\mathbf{P}'})$$

where the Sobolev space is defined using the cylindrical-end metric $\check{g}$. As in the case of closed manifolds (Definition 2.13), we say that choices $(\mathbf{P}', B_1', B_0')$ and $(\mathbf{P}'', B_1'', B_0'')$ are isomorphic if there is a determinant-1 gauge transformation $\mathbf{P}' \to \mathbf{P}''$ pulling back B_i'' to B_i'. (The notion of "determinant 1" has meaning here, because both $\mathbf{P}'$ and $\mathbf{P}''$ are identified with $\mathbf{P}$ outside a finite set.) We denote by z a typical isomorphism-classes of choices:

$$z = [\mathbf{P}', B_1', B_0'].$$

The quotient of $\mathcal{C}_k(\mathbf{P}', A_o)$ by the determinant-1 gauge group $\mathcal{G}_{k+1}(\mathbf{P}', A_o)$ depends only on z, and we write

$$\mathcal{B}_z(W, S, \mathbf{P}; \beta_1, \beta_0) = \mathcal{C}_k(\mathbf{P}', A_o) \big/ \mathcal{G}_{k+1}(\mathbf{P}', A_o)$$

Remark. The discussion here is not quite standard. If S has no closed components, then every isomorphism class z has a representative $[\mathbf{P}, B_0, B_1]$ in which the singular bundle data is $\mathbf{P}$. If S has closed components however, then we must allow $\mathbf{P}' \neq \mathbf{P}$ in order to allow differing monopole charges on the closed components. In the case of a cylinder $(W, S) = I \times (Y, K)$, the set of z's coincides with the previous space of homotopy-classes of paths from β_1 to β_0.

The perturbed version of the ASD equations that we shall use is defined as follow π_i be the chosen holonomy perturbations on the Y_i. We consider perturbation of the 4-dimensional equations on $\check{W}$

$$F_A^+ + \hat{V}(A) = 0 \tag{16}$$

where $\hat{V}$ is a holonomy-perturbation supported on the cylindrical ends. To define this on the cylindrical end $[\,0, \infty) \times Y_0$, we take the given π_0 from the auxiliary data $\mathbf{a}_0$ and an additional term π_0'. We then set

$$\hat{V}(A) = \phi(t)\hat{V}_{\pi_0}(A) + \psi(t)\hat{V}_{\pi_0'}(A), \tag{17}$$

where $\phi(t)$ is a cut-off function equal to 1 on $[\,1, \infty)$ and equal to 0 near $t = 0$, while $\psi(t)$ is a bump-function supported in $[0, 1]$. The perturbation is defined similarly on the other end, using π_1 and an additional π_1'. This sort of perturbation is used in [22, Section 24] and again in [23]. We write

$$M_z(W, S, \mathbf{P}; \beta_1, \beta_0) \subset \mathcal{B}_{k,z}(W, S, \mathbf{P}; \beta_1, \beta_0)$$

for the moduli space of solutions to the perturbed anti-self-duality equations. Note that π_i' contributes a perturbation that is compactly supported in the cobordism. We refer to these additional terms as *secondary perturbations.*

Just as in the cylindrical case, we can combine the linearization of the left hand side of (16) with gauge fixing to obtain a Fredholm operator $\mathcal{D}_A$. We say that the moduli space is regular if $\mathcal{D}_A$ is surjective at all solutions. Regular moduli spaces are smooth manifolds, of dimension equal to the index of $\mathcal{D}_A$, and we write as

$$\mathrm{gr}_z(W, S, \mathbf{P}; \beta_1, \beta_0).$$

We again set

$$M(W, S, \mathbf{P}; \beta_1, \beta_0)_d = \bigcup_{\mathrm{gr}_z = d} M_z(W, S, \mathbf{P}; \beta_1, \beta_0).$$

The arguments use to prove Proposition 24.4.7 of [22] can be used to prove the following genericity result.

Proposition 3.8. *Let π_1, π_0 be perturbations satisfying Hypothesis 3.6. Let $(W, S, \mathbf{P})$ be given, together with a cylindrical-end metric $\breve{g}$ as above. Then there are secondary perturbations π_1', π_0' so that for all β_1, β_0 and z, all the moduli spaces $M_z(W, S, \mathbf{P}; \beta_1, \beta_0)$ of solutions to the perturbed equations* (16) *are regular.* □

3.8 Maps from cobordisms

The moduli spaces $M_z(W, S, \mathbf{P}; \beta_1, \beta_0)$ are orientable, because the determinant line $\det(\mathcal{D}_A)$ over $\mathcal{B}_z(W, S, \mathbf{P}; \beta_1, \beta_0)$ is trivial. Furthermore, when we orient $\det(\mathcal{D}_A)$ for one particular $[A]$ in $\mathcal{B}_z(W, S, \mathbf{P}; \beta_1, \beta_1)$, then this canonically determines an orientation for the moduli spaces $M_{z'}(W, S, \mathbf{P}; \beta_1, \beta_0)$ for all other z'.

To see what is involved in specifying an orientation for the moduli spaces, let us recall first that we have chosen basepoints θ_i in $\mathcal{B}(Y_i, K_i; \mathbf{P}_i)$ for $i = 0, 1$, as part of the auxiliary data $\mathbf{a}_i$. We have defined $\Lambda(\beta_i)$ to be $\Lambda(\theta_i, \beta_i)$. Choose gauge representatives Θ_i for the connections θ_i, an let A_θ be a connection on $(W, S, \mathbf{P})$ which is equal to the pull-back of Θ_i on the two cylindrical ends. The operator $\mathcal{D}_{A_\theta}$ is Fredholm on the weighted Sobolev spaces (11) for small ϵ, and we define

$$\Lambda(W, S, \mathbf{P})$$

to be the two-element set of orientations of $\det(\mathcal{D}_{A_\theta})$. (This set is dependent on θ_1 and θ_0, though our notation hides this.) In this context, we make the a definition:

Definition 3.9. Let $(Y_1, K_1, \mathbf{P}_1)$ and $(Y_0, K_0, \mathbf{P}_0)$ be manifolds with singular bundle data, and let $(W, S, \mathbf{P})$ be a cobordism from the first to the second. Let auxiliary data $\mathbf{a}_1$, $\mathbf{a}_0$ be given on the two ends, as above. We define an *I-orientation* of $(W, S, \mathbf{P})$ to be a choice of element from $\Lambda(W, S, \mathbf{P})$, or equivalently, an orientation of the determinant line $\det(\mathcal{D}_{A_\theta})$. ◊

The definition of $\Lambda(W, S, \mathbf{P})$ is constructed so that the two-element set of orientations of the determinant line over $\mathcal{B}_z(W, S, \mathbf{P}; \beta_1, \beta_0)$ is isomorphic to the product

$$\Lambda(\beta_1)\Lambda(\beta_0)\Lambda(W, S, \mathbf{P}).$$

Thus, once an I-orientation of $(W, S, \mathbf{P})$ is given, a choice of orientation for a component of any moduli space $M_z(W, S, \mathbf{P}; \beta_1, \beta_0)$ determines an isomorphism $\mathbb{Z}\Lambda(\beta_1) \to \mathbb{Z}\Lambda(\beta_0)$. In particular each point $[A]$ in a zero-dimensional moduli space $M_0(W, S, \mathbf{P}; \beta_1, \beta_0)$ determines such an isomorphism,

$$\epsilon([A]) : \mathbb{Z}\Lambda(\beta_1) \to \mathbb{Z}\Lambda(\beta_0).$$

In this way, given a choice of an I-orientation of $(W, S, \mathbf{P})$, we obtain a homomorphism

$$m = \sum_{\beta_1, \beta_0} \sum_{[A] \in M(W,S,\mathbf{P};\beta_1,\beta_0)_0} \epsilon([A]), \tag{18}$$

which by the usual arguments is a chain map

$$m : C_*(Y_1, K_1, \mathbf{P}_1) \to C_*(Y_0, K_0, \mathbf{P}_0).$$

The induced map in homology depends only on $(W, S, \mathbf{P})$, the original auxiliary data $\mathbf{a}_0$ and $\mathbf{a}_1$ on the two ends, and the I-orientation, not on the choice of the secondary perturbations π'_i or on the choice of Riemannian metric $\breve{g}$ on the interior of W. Furthermore, if $(W, S, \mathbf{P})$ is expressed as the union of two cobordisms $(W', S', \mathbf{P}')$ and $(W'', S'', \mathbf{P}'')$, then the chain map m is chain-homotopic to the composite,

$$m \simeq m'' \circ m'.$$

By the standard approach, taking $(W, S, \mathbf{P})$ to be a cylinder, we deduce that $I(Y_0, K_0, \mathbf{P}_0)$ is independent of the choice of auxiliary data $\mathbf{a}_0$. More precisely, if $\mathbf{a}_0$ and $\mathbf{a}_1$ are two choices of auxiliary data for $(Y, K, \mathbf{P})$, then there is a canonical pair of isomorphisms $\{ m_*, -m_* \}$ differing only in sign,

$$\pm m_* : I(Y, K, \mathbf{P})_{\mathbf{a}_1} \to I(Y, K, \mathbf{P})_{\mathbf{a}_0}.$$

Thus $I(Y, K, \mathbf{P})$ is a topological invariant of $(Y, K, \mathbf{P})$. Note that we have no a priori way of choosing I-orientations to resolve the signs in the last formulae: the dependence on the auxiliary data $\mathbf{a}_i$ means that cylindrical cobordisms do not have canonical I-orientations.

3.9 Families of metrics and compactness

The proof from [23] that the chain-homotopy class of the map m above is independent of the choice of π_i' and the metric on the interior of W follows standard lines and exploits a parametrized moduli space, over a family of Riemannian metrics and perturbations. For our later applications, we will need to consider parameterized moduli spaces where the metric is allowed to vary in certain more general, controlled non-compact families.

Let $(Y_1, K_1, \mathbf{P}_1)$ and $(Y_0, K_0, \mathbf{P}_0)$ be given and let $\mathbf{a}_1$ and $\mathbf{a}_0$ be auxiliary data as in (15), so that the transversality conditions of Proposition 3.5 hold. Let $(W, S, \mathbf{P})$ be a cobordism between these, equipped with singular bundle data $\mathbf{P}$ restricting to the $\mathbf{P}_i$ at the ends. By a family of metrics parametrized by a smooth manifold G we mean a smooth orbifold section of $\mathrm{Sym}^2(T^*\check{W}) \times G \to \check{W} \times G$ which restricts to each $W \times \{\check{g}\}$ as a Riemannian metric denoted $\check{g}$. These metrics will always have an orbifold singularity along S. Choose a family of secondary perturbations $\pi_{i,\check{g}}'$, supported in the collars of the two boundary components as before, but now dependent on the parameter $\check{g} \in G$. We have then a corresponding perturbing term $\hat{V}_g$ for the anti-self-duality equations on $\check{W}$. For any pair of critical points $(\beta_1, \beta_0) \in \mathfrak{C}_{\pi_1}(Y_1) \times \mathfrak{C}_{\pi_2}(Y_2)$, we can now form a moduli space

$$M_{z,G}(W, S, \mathbf{P}; \beta_1, \beta_0) \subset \mathcal{B}_z(W, S, \mathbf{P}; \beta_1, \beta_0) \times G$$

of pairs $([A], g)$ and where $[A]$ solves the equation

$$F_A^{+_g} + \hat{V}_g(A) = 0. \tag{19}$$

Here $+_g$ denote the projection onto g-self-dual two forms. A solution $([A], g)$ is called to Equation (19) is called regular if the differential of the map

$$(A, g) \mapsto F_A^{+_g} + \hat{V}_g(A)$$

is surjective. The arguments use to prove Propostion 24.4.10 of [22] can be used to prove the following genericity result.

Proposition 3.10. *Let π_i be perturbations satisfying the conclusions of Proposition 3.5 and let G be a family of metrics as above. There is a family of secondary perturbations $\pi_{i,g}'$, for $i = 0, 1$, parameterized by $g \in G$ so that for all $\beta_i \in \mathfrak{C}_{\pi_i}$ and all paths z, the moduli space*

$$M_{z,G}(W, S, \mathbf{P}; \beta_1, \beta_0)$$

consists of regular solutions. $\square$

In the situation of the proposition, the moduli space $M_{z,G}(W, S, \mathbf{P}; \beta_1, \beta_0)$ is smooth of dimension $\mathrm{gr}_z(\beta_1, \beta_0) + \dim G$. We use $M_G(W, S, \mathbf{P}; \beta_1, \beta_0)_d$ to denote its d-dimensional components. To orient the moduli space, we orient both the determinant line bundle $\det(\mathcal{D})$ on $\mathcal{B}_z(W, S, \mathbf{P}; \beta_1, \beta_0)$ and the parametrizing manifold G, using a fiber-first convention.

Consider now the case that G is a compact, oriented manifold with oriented boundary ∂G. We omit $(W, S, \mathbf{P})$ from our notation for brevity, and denote the moduli space by $M_G(\beta_1, \beta_0)$. Let us suppose (as we may) that the secondary perturbations are chosen so that both $M_G(\beta_1, \beta_0)$ and $M_{\partial G}(\beta_1, \beta_0)$ are regular. The first moduli space will then be a (non-compact) manifold with boundary, for we have

$$\partial M_G(\beta_1, \beta_0)_d = M_{\partial G}(\beta_1, \beta_0)_{d-1}.$$

But this will not be an equality of *oriented* manifolds. Our fiber-first convention for orienting M_G and the standard outward-normal-first convention for the boundary orientations interact here to give

$$\partial M_G(\beta_1, \beta_0)_d = (-1)^{d-\dim G} M_{\partial G}(\beta_1, \beta_0)_{d-1}. \tag{20}$$

When an orientation of G and an element of $\Lambda(W, S, \mathbf{P})$ are chosen, the count of the solutions $[A]$ in $M_G(\beta_1, \beta_0)_0$ defines a group homomorphism

$$m_G : C(Y_1, K_1, \mathbf{P}_1) \to C(Y_0, K_0, \mathbf{P}_0).$$

Similarly $M_{\partial G}(\beta_1, \beta_0)_0$ defines a group homomorphism $m_{\partial G}$. These two are related by the following chain-homotopy formula:

$$m_{\partial G} + (-1)^{\dim G} m_G \circ \partial = \partial \circ m_G.$$

(See [22, Proof of Proposition 25.3.8], though there is a sign error in [22] at this point.) The proof of this formula is to count the endpoints in the 1-dimensional moduli spaces $M_G(\beta, \alpha)_1$ on $(W, S, \mathbf{P})$.

Remarks. Although we will not have use for greater generality here, it is a straightforward matter here to extend this construction by allowing G to parametrize not just a family of metrics on a fixed cobordism (W, S), but a smooth family of cobordisms (a fibration over G) with fixed trivializations of the family at the two ends Y_1 and Y_0. There is an obvious notion of an I-orientation for such a family (whose existence needs to be a hypothesis), and one then has a formula just like the one above for a compact family G with boundary. One can also consider the case that G is a simplex in a simplicial complex Δ. In that case, given a choice of perturbations making the moduli spaces transverse over every simplex, and a coherent choice of I-orientations of the fibers, the above formula can be interpreted as saying that we have a chain map

$$\underline{\mathrm{m}} : C(\Delta) \otimes C(Y_1, K_1, \mathbf{P}_1) \to C(Y_0, K_0, \mathbf{P}_0)$$

where $C(\Delta)$ is the simplicial chain complex (and the usual convention for the signs of the differential on a product complex apply). If the local system defined by the I-orientations of the fibers is non-trivial, then there is a similar chain map, but we must then use the chain complex $C(\Delta; \xi)$ with the appropriate local coefficients ξ.

Next we wish to generalize some of the stretching arguments that are used (for example) in proving the composition law for the chain-maps m induced by cobordisms. For this purpose we introduce the notion of a broken Riemannian metric on a cobordism (W, S) from (Y_1, K_1) to (Y_0, K_0). A *cut* of (W, S) is an orientable codimension-1 submanifold $Y_c \subset \mathrm{int}(W)$ so that the intersection $Y_c \cap S = K_c$ is transverse. A cylindrical-end metric $\check{g}$ on (W, S) is *broken along a cut* Y_c if it is a complete Riemannian metric on $(\mathrm{int}W)\backslash Y_c$ and there is a normal coordinate collar neighborhood $(-\epsilon, \epsilon) \times Y_c$ with normal coordinate $r_c \in (-\epsilon, \epsilon)$ so that

$$g = (dr_c/r_c)^2 + \check{g}_{Y_c}$$

where $\check{g}_{Y_c}$ is a metric on Y_c with orbifold singularity along K_c. Note that Y_c may be disconnected, may have parallel components, and may have components parallel to the boundary. The manifold $(\mathrm{int}W)\backslash Y_c$ equipped with this metric has two extra cylindrical ends (each perhaps with several components), namely the ends

$$\begin{aligned} (-\epsilon, 0) &\times Y_c \\ (0, \epsilon) &\times Y_c. \end{aligned} \tag{21}$$

We will assume that each component of $(Y_c, K_c, \mathbf{P}|_{Y_c})$ satisfies the non-integral condition.

Given perturbations π_i for the (Y_i, K_i) (for $i = 0, 1$) and π_c for (Y_c, K_c) and secondary perturbations π'_i and $\pi'_{c,+}$ and $\pi'_{c,-}$, we can write down perturbed ASD-equations on $\mathrm{int}(W)\backslash Y_c$ as in (16). The perturbation $\hat{V}$ is defined as in (17), using the perturbation π_c together with the secondary perturbations $\pi'_{c,-}$, $\pi'_{c,+}$ on the two ends (21) respectively.

Given a pair of configurations β_i in $\mathcal{B}(Y_i, K_i, \mathbf{P}_i)$, for $i = 1, 0$, and a cut Y_c, a *cut path* from β_1 to β_0 along $(W, S, \mathbf{P})$ is a continuous connection A in a $\mathbf{P}' \to (W, S)$ in singular bundle data $\mathbf{P}'$ equivalent to $\mathbf{P}$, such that A is smooth on $\mathrm{int}(W)\backslash Y_c$ and its restriction to (Y_i, K_i) belongs to the gauge equivalence class β_i. A *cut trajectory* from β_1 to β_0 along $(W, S, \mathbf{P})$ is a cut path from β_1 to β_0 along $(W, S, \mathbf{P})$ whose restriction to $\mathrm{int}(W)\backslash Y_c$ is a solution to the perturbed ASD equation (16).

Given a cut Y_c of (W, S) we can construct a family of Riemannian metrics on $\mathrm{int}(W)$ which degenerates to a broken Riemannian metric. To do this, we start with a Riemannian metric $\check{g}_o$ on W (with orbifold singularity along S) which contains a collar neighborhood of Y_c on which the metric is a product

$$dr^2 + \check{g}_{Y_c}$$

where $r \in [-1, 1]$ denotes the signed distance from Y_c. Let

$$f_s : \mathbb{R} \to \mathbb{R}$$

be a family of functions parametrized by $s \in [0, \infty]$ that smooths out the function which is given by

$$\frac{1 + 1/s^2}{r^2 + 1/s^2}$$

for $r \in [-1, 1]$ and 1 otherwise. Note that the above expression is 1 on the boundary $r = \pm 1$ as well as when $s = 0$. Also note that $\lim_{s \mapsto \infty} f_s(r) = 1/r^2$. For each component of

$$Y_c = \coprod_{i=1}^{N} Y_c^i$$

we introduce a parameter s_i and modify the metric by

$$f_{s_i}(r)dr^2 + \check{g}_{Y_c}.$$

When $s_i = \infty$ the metric is broken along the Y_c^i. In the way we get a family of Riemannian metrics parametrized by $[0, \infty)^N$ which compactifies naturually to a family of broken metrics parameterized by $[0, \infty]^N$. In the neighborhood of $\{\infty\}^N$, each metric is broken along some subset of the components of Y_c. We can further elaborate this construction slightly by allowing the original metric also to vary in a family G_1, while remaining unchanged in the collar neighborhood of Y_c, so that we have a family of metrics parametrized by

$$[0, \infty]^N \times G_1 \tag{22}$$

for some G_1. Given the cut Y_c, we say that a family of singular metrics on (W, S) is a "model family" for the cut Y_c if it is a family of this form.

To describe suitable perturbations for the equations over such a model family of singular metrics, we choose again a generic perturbation π_c for Y_c and write its component belonging to Y_c^i as π_c^i. When the coordinate $s_i \in [0, \infty]$ is large, the metric contains a cylindrical region isometric to a product

$$[-T_i, T_i] \times Y_c^i$$

where $T_i \to \infty$ as $s_i \to \infty$. We require that for large s_i, the perturbation on this cylinder has the form

$$\hat{V}_{\pi_c^i} + \psi_-(t)\hat{V}_{(\pi_{c,-}^i)'} + \psi_+(t)\hat{V}_{(\pi_{c,+}^i)'}$$

where the functions ψ_+ and ψ_- are bump-functions supported near $t = -T_i$ and $t = T_i$ respectively. Here $(\pi_{c,\pm}^i)'$ are secondary perturbations which are allowed to vary with the extra parameters G_1.

We can now consider a general family of metrics which degenerates like the model family at the boundary. Thus we consider a manifold with corners, G, parametrizing a family of broken Riemannian metrics $\breve{g}$ on W (with orbifold singularity along S), and we ask that in the neighborhood of each point of every codimension-n facet, there should be a cut Y_c with exactly n components, so that in the neighborhood of this point the family is equal to a neighborhood of

$$\{\infty\}^n \times G_1$$

in some model family for the cut Y_c. (Note that the cut Y_c will vary.) Whenever we talk of a "family of broken metrics", we shall mean that the family has this model structure at the boundary. When considering perturbations, we shall need to have fixed perturbations π_c^i for every component of every cut, and also secondary perturbations which have the form described above, in the neighborhood of each point of the boundary. We suppose that these are chosen so that the parametrized moduli spaces over all strata of the boundary are regular.

Suppose now that we are given a family of broken metrics of this sort, parametrized by a compact manifold-with-corners G. Over the interior $\mathrm{int}(G)$, we have a smooth family of metrics and hence a parametrized moduli space

$$M_{z,\mathrm{int}(G)}(\beta_1, \beta_0)$$

on the cobordism $(W, S, \mathbf{P})$. This moduli space has a natural completion involving cut paths (in the above sense) on (W, S), as well as broken trajectories on the cylinders over (Y_c^i, K_c^i) for each component $Y_c^i \subset Y_c$. We denote the completion by

$$M^+_{z,G}(\beta_1, \beta_0) \supset M_{z,\mathrm{int}(G)}(\beta_1, \beta_0).$$

The completion is a space stratified by manifolds, whose top stratum is $M_{z,\mathrm{int}(G)}(\beta_1, \beta_0)$. The codimension-1 strata are of three sorts.

(a) First, there are the cut paths on $(W, S, \mathbf{P})$ from β_1 to β_0, cut along some *connected* Y_c. These form a codimension-1 stratum of the compactification lying over a codimension-1 face of G (the case $N = 1$ in (22)).

(b) Second there are the strata corresponding to a trajectory sliding off the incoming end of the cobordism, having the form

$$\breve{M}_{z_1}(\beta_1, \alpha_1) \times M_{z-z_1,G}(\alpha_1, \beta_0)$$

where the first factor is a moduli space of trajectories on Y_1 and α_1 is a critical point.

(c) Third there is the symmetrical case of a trajectory sliding off the outgoing end of the cobordism:

$$M_{z-z_0,G}(\beta_1, \alpha_0) \times \breve{M}_{z_0}(\alpha_0, \beta_0).$$

In the neighborhood of a point in any one of these codimension-1 strata, the compactification $M^+_{z,G}(\beta_1, \alpha_0)$ has the structure of a C^0 manifold with boundary.

The completion $M^+_{z,G}(\beta_1, \beta_0)$ is not in general compact, because of bubbling off of instantons and monopoles. However it will be compact when it has dimension less than 4.

An I-orientation for $(W, S, \mathbf{P})$, together with a choice of element from $\Lambda(\beta_1)$ and $\Lambda(\beta_0)$, gives rise not only to an orientation of the moduli space $M_z(W, S\mathbf{P}; \beta_1, \beta_0)$ for a fixed smooth metric $\breve{g}$, but also to an orientation of the moduli spaces of cut paths, for any cut Y_c, by a straightforward generalization of the composition law for I-orientations. Thus, given such an I-orientation and an orientation of G, we obtain oriented moduli spaces over both G and over the codimension-1 strata of ∂G.

Just as in the case of a smooth family of metrics parametrized by a compact manifold with boundary, a family of broken Riemannian metrics parametrized by an oriented manifold-with-corners G, together with an I-orientation of $(W, S, \mathbf{P})$, gives rise to a chain-homotopy formula,

$$m_{\partial G} + (-1)^{\dim G} m_G \circ \partial = \partial \circ m_G.$$

Again, the proof is obtained by considering the endpoints of 1-dimensional moduli spaces over G: the three terms correspond to the three different types of codimension-1 strata. The map m_G is defined as above by counting solutions in zero-dimensional moduli spaces over G; and $m_{\partial G}$ is a sum of similar terms, one for each codimension-1 face of G (with the outward-normal first convention).

The case of most interest to us is when each face of G corresponds to a cut Y_c whose single connected component separates Y_1 from Y_0, so separating W into two cobordisms: W' from Y_1 to Y_c, and W'' from Y_c to Y_0. Let us suppose that G has l codimension-1 faces, $G_1, \ldots, G_l$, all of this form. Each face corresponds to a cut which expresses W as a union,

$$W = W'_j \cup W''_j, \quad j = 1, \ldots, l.$$

In the neighborhood of such a point of G_j, G has the structure

$$(0, \infty] \times G_j.$$

Let us suppose also that G_j has the form of a product,

$$G_j = G'_j \times G''_j,$$

where the two factors parametrize families of metrics on W'_j and W''_j. Let us also equip W'_j and W''_j with I-orientations so that the composite I-orientation is that of W. We can write the chain-homotopy formula now as

$$\Big(\sum_j m_{G_j}\Big) + (-1)^{\dim G} m_G \circ \partial = \partial \circ m_G,$$

where we are still orienting G_j as part of the boundary of G. On the other hand, we can interpret m_{G_j} (which counts cut paths on W) as the composite of the two maps obtained from the cobordisms W'_j and W''_j. When we do so, there is an additional sign, because of our convention that puts the G factors last: we have

$$m_{G_j} = (-1)^{\dim G''_j \dim G'_j} m''_j \circ m'_j$$

where m'_j and m''_j are the maps induced by the two cobordisms with their respective families of metrics. Thus the chain-homotopy formula can be written

$$\Big(\sum_j (-1)^{\dim G''_j \dim G'_j} m''_j \circ m'_j\Big) + (-1)^{\dim G} m_G \circ \partial = \partial \circ m_G. \qquad (23)$$

4 Topological constructions

In this section, we shall consider how to "package" the constructions of section 3, so as to have a functor on a cobordism category whose definition can be phrased in terms of more familiar topological notions.

4.1 Categories of 3-manifolds, bundles and cobordisms

When viewing Floer homology as a "functor" from a category of 3-manifolds and cobordisms to the category of groups, one needs to take care in the definition of a "cobordism". Thus, by a cobordism from a 3-manifold Y_1 to Y_0 we should mean a 4-manifold W whose boundary ∂W comes equipped with a diffeomorphism $\phi : \partial W \to Y_1 \cup Y_0$. If the manifolds Y_i are oriented (as our 3-manifolds always are), then (W, ϕ) is an *oriented cobordism* if ϕ is orientation-reversing over Y_1 and orientation-preserving over Y_0. Two oriented cobordisms (W, ϕ) and (W', ϕ') between the same 3-manifolds are isomorphic if there is an orientation-preserving diffeomorphism $W \to W'$ intertwining ϕ with ϕ'. Closed, oriented 3-manifolds form the objects of a category in which the morphisms are isomorphism classes or oriented cobordisms. We can elaborate this idea a little, to include $SO(3)$-bundles over our 3-manifolds, defining a category B as follows.

An object of the category B consists of the following data:

- a closed, connected, oriented 3-manifold Y;
- an $SO(3)$-bundle $P \to Y$ satisfying the *non-integral condition* (so that $w_2(P)$ has odd evaluation on some oriented surface).

A morphism from (Y_1, P_1) to (Y_0, P_0) is an *equivalence class* of data of the following sort:

- an oriented cobordism W from Y_1 to Y_0;
- a bundle P on W;
- an isomorphism of the bundle $P|_{\partial W}$ with $P_1 \cup P_0$, covering the given diffeomorphism of the underlying manifolds.

Here we say that (W, P) and (W', P') are *equivalent* (and are the same morphism in this category) if there is a diffeomorphism

$$\psi : W \to W'$$

intertwining the given diffeomorphism of ∂W and $\partial W'$ with $-Y_0 \cup Y_1$, together a bundle isomorphism on the complement of a finite set x,

$$\Psi : P|_{W\setminus x} \to P'_{W'\setminus\psi(x)},$$

covering the map ψ and intertwining the given bundle maps at the boundary.

The category B is essentially the category for which Floer constructed his instanton homology groups: we can regard instanton homology as a *projective functor* I from B to groups. Here, the word "projective" means that we regard the morphism $I(W, P)$ corresponding to a cobordism as being defined only up to an overall sign. Alternatively said, the target category is the category P-GROUP of abelian groups in which the morphisms are taken to be unordered pairs $\{h, -h\}$ of group homomorphisms, so that I is a functor

$$I : \text{B} \to \text{P-GROUP}$$

On a 3-manifold Y a bundle P is determined by its Stiefel-Whitney class; but knowing the bundle only up to isomorphism is not sufficient to specify an actual object in B, nor can we make any useful category whose objects are such isomorphism classes of bundles. However, if instead of just specifying the Stiefel-Whitney class we specify a particular geometric representative for this class, then we will again have a sensible category with which to work, as we now explain.

Consider first the following situation. Let ω be a finite, 2-dimensional simplicial complex and let V be a 2-disk bundle over ω (not necessarily orientable). Let $T \subset V$ be the circle bundle and V/T the Thom space. There is a distinguished class τ in $H^2(V, T; \mathbb{Z}/2)$, the Thom class. If (P, ϕ) is an $SO(3)$ bundle on V with a trivialization ϕ on S, then it has a well-defined relative Stiefel-Whitney class in $H^2(V, T; \mathbb{Z}/2)$. We can therefore seek a pair (P, ϕ) whose relative Stiefel-Whitney class is the Thom class. By the usual arguments of obstruction theory applied to the cells of V/T, one sees that such a (P, ϕ) exists, and that its isomorphism class is unique up to the action of adding instantons to the 4-cells of V. Furthermore, if (P', ϕ') is already given over some subcomplex $\omega' \subset \omega$, then it can be extended over all of ω. If ω is a closed surface, perhaps non-orientable, in a 4-manifold X, then we can apply the observation of this paragraph to argue (as in the proof of the theorem of Dold and Whitney [5]) that there exists bundle $P \to X$ with a trivialization outside a tubular neighborhood of ω, whose relative Stiefel-Whitney class is dual to ω.

This leads us to consider a category in which an object is a closed, connected, oriented 3-manifold Y together with an embedded, unoriented

1-manifold $\omega \subset Y$ (thought of as a dual representative for w_2 of some bundle). A morphism in this category, from (Y_0, ω_0) to (Y_1, ω_1) is simply a cobordism of pairs (W, ω), with W an oriented cobordism, and ω unoriented. Morphisms are composed in the obvious way. We call this category W.

It is tempting to suppose that there is a functor from W to B, which should assign to an object (Y, ω) a pair (Y, P) where P is a bundle with $w_2(P)$ dual to ω. As it stands, however, P is unique only up to isomorphism. To remedy this, we analyze the choice involved. Given an object (Y, ω) in W, let us choose a tubular neighborhood V for ω and choose a bundle (P_0, ϕ_0), with a trivialization ϕ_0 outside V, whose relative w_2 is the Thom class. If (P_1, ϕ_1) is another such choice, then the objects (Y, P_1) and (Y, P_0) are connected by a uniquely-determined morphism in B. Indeed, in the cylindrical cobordism $[0, 1] \times Y$, we have a product embedded surface $[0, 1] \times \omega$, and there is an $SO(3)$ bundle P with trivialization outside the tubular neighborhood of $[0, 1] \times \omega$, which extends the given data at the two ends of the cobordism. The resulting bundle is unique up to the addition of instantons, so it gives a well-defined morphism – in fact an isomorphism – in B.

In this way, from an object (Y, ω) in W, we obtain not an object in B, but a commutative diagram in B, consisting of all possible bundles P_i arising this way and the canonical morphisms between them. Applying the projective functor I, we nevertheless obtain from this a well-defined projective functor,

$$I : \text{W} \to \text{P-GROUP}$$

We shall write $I^\omega(Y)$ for the instanton homology of (Y, ω) in this context.

4.2 Introducing links: singular bundles

The discussion from the previous subsection can be carried over to the more general situation in which a we have a knot or link $K \subset Y$, rather than a 3-manifold Y alone. We can define a category BLINK in which objects are pairs (Y, K) carrying singular bundle data $\mathbf{P}$. Thus Y is again a closed, oriented, connected 3-manifold, K is an unoriented link in Y, and we are given singular bundle data $\mathbf{P}$ in the form of a double-cover $K_\Delta \to K$, an $SO(3)$ bundle $P_\Delta \to K_\Delta$ and an $O(2)$ reduction in the neighborhood of K_Δ. The singular bundle data is required to satisfy the non-integral condition from Definition 3.1. A morphism in this category, from $(Y_1, K_1, \mathbf{P}_1)$ to $(Y_0, K_0, \mathbf{P}_0)$ is a cobordism of pairs, (W, S), with W an oriented cobordism from Y_1 to Y_0 and S and unoriented cobordism from K_1 to K_0, equipped with singular bundle data $\mathbf{P}$ and an identification of all this data with the given data at the two ends of the cobordism. Two such cobordisms with singular bundle data are regarded as being the same morphism if they are equivalent, in the same sense that we described for B. Just as in the case of B, singular instanton homology defines a functor,

$$I : \text{BLINK} \to \text{P-GROUP}.$$

This is essentially the content of section 3.8.

In addition to BLINK, we would like to have a version WINK an analogous to W above, in which we replace the bundle data $\mathbf{P}$ with a codimension-2 locus ω representing the dual of w_2. To begin again with the closed case, suppose that we are given a 4-manifold X, an embedded surface Σ and a surface-with-boundary, $(\omega, \partial\omega) \subset (X, \Sigma)$. We require that $\omega \cap \Sigma$ is a collection of circles and points: the circles are $\partial\omega$, along which ω should meet Σ normally; and the points are transverse intersections of ω and Σ in X. From the circles $\partial\omega \subset \Sigma$ we construct a double-cover Σ_Δ by starting with a trivialized double-cover of $\Sigma \backslash \partial\omega$, say

$$(\Sigma \backslash \partial\omega) \times \{1, -1\},$$

and then identifying across the cut with an interchange of the two sheets. Thus $w_1(\Delta)$ is dual to $[\partial\omega]$ in $H^1(\Sigma; \mathbb{Z}/2)$.

Because Δ is trivialized, by construction, on $\Sigma \backslash \partial\omega$, we have distinguished copies of $\Sigma \backslash \partial\omega$ inside the non-Hausdorff space X_Δ. Let

$$\Sigma_- \subset X_\Delta$$

be the closure of the copy $(\Sigma \backslash \partial\omega) \times \{-1\}$ in X_Δ. This is a surface whose boundary is two copies of $\partial\omega$. Let ω_1 denote the two-dimensional complex

$$\omega_1 = \pi^{-1}(\omega) \cup \Sigma_-$$

in X_Δ. Figure 1 shows the inverse image ω_1^h of ω_1 in the Hausdorff space X_Δ^h, in a schematic lower-dimensional picture. A regular neighborhood of ω_1^h in the complex X_Δ^h is a disk bundle over ω_1^h, so there is a well-defined dual class, as in our previous discussion. There is therefore a bundle P_Δ on X_Δ^h, with a trivialization outside a regular neighborhood of ω_1^h, whose w_2 is the Thom class. In this way, the original surface ω determines a bundle $P_\Delta \to X_\Delta$, uniquely up to the addition of instantons and monopoles. This P_Δ in turn determines singular bundle data $\mathbf{P}$ on X_Δ, up to isomorphism and the addition of instantons and monopoles.

We now set up the category WINK in which objects are triples (Y, K, ω), where:

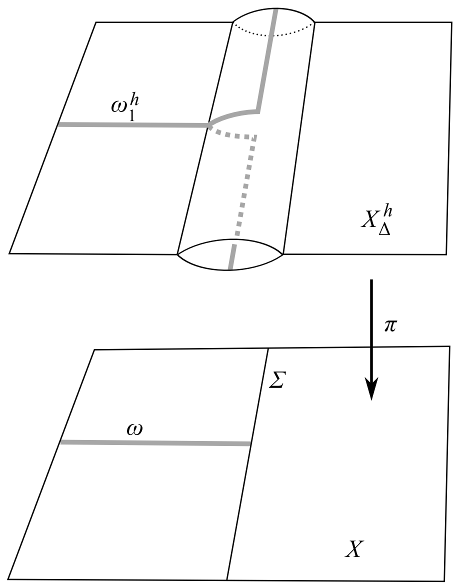

Figure 1: The codimension-2 subcomplex ω_1^h, illustrated in a lower-dimensional picture.

- Y is a closed, oriented, connected 3-manifold;
- K is an unoriented link in Y;
- ω is an embedded 1-manifold with $\omega \cap K = \partial\omega$, meeting K normally at its endpoints.

The triples (Y, K, ω) are required to satisfy the non-integral condition, Definition 3.1. The morphisms from (Y_1, K_1, ω_1) to (Y_0, K_0, ω_0) are isomorphism classes of triples (W, S, ω), where

- (W, S) is a cobordism of pairs, with W an oriented cobordism;
- $\omega \subset W$ is a 2-manifold with corners, whose boundary is the union of ω_1, ω_0, and some arcs in S, along which ω normal to S. The intersection $\omega \cap S$ is also allowed to contain finitely many points where the intersection is transverse.

Just as with W, an object (Y, K, ω) in WINK gives rise to a commutative diagram of objects $(Y, K, \mathbf{P})$ in BLINK. Singular instanton homology therefore defines a functor,

$$I : \text{WINK} \to \text{P-GROUP}$$

as in the previous arguments. We denote this by $I^\omega(Y,K)$. Thus a morphism from (Y_1,K_1,ω_1) to (Y_0,K_0,ω_0) represented by a cobordism (W,S,ω) gives rise to a group homomorphism (well-defined up to an overall sign),

$$I^\omega(W,S) : I^{\omega_1}(Y_1,K_1) \to I^{\omega_0}(Y_0,K_0).$$

Most often, the particular choice of ω is clear from the context, and we will often write simply a generic "ω" in place of the specific ω_1, ω_0, etc.

4.3 Constructions for classical knots

If K is a classical knot or link in S^3, or more generally a link in a homology sphere Y, then the triple (Y,K,ω) satisfies the non-integral condition if and only if some component of K contains an odd number of endpoints of ω. In particular, we cannot apply the functor I to such a triple when ω is empty, so we do not directly obtain an invariant of classical knots and links without some additional decoration.

As described in the introduction however, we can use a simple construction to obtain an invariant of a link with a given *basepoint*. More precisely, we consider a link K in an arbitrary 3-manifold Y, together with a basepoint $x \in K$ and a given normal vector v to K at x. Given this data, we let L be a circle at the boundary of a standard disk centered at x in the tubular neighborhood of K, and we let ω be a radius of the disk: a standard arc joining $x \in K$ to a point in L, with tangent vector v at x. We write $K^\natural$ for the new link

$$K^\natural = K \amalg L$$

and we define

$$I^\natural(Y,K,x,v) = I^\omega(Y,K^\natural).$$

See Figure 2. We shall usually omit x and v (particularly v) from our notation, and simply write $I^\natural(Y,K)$ for this invariant.

The role of the normal vector v here is in making the construction functorial. We can construct a category in which a morphism from (Y_0,K_0,x_0,v_0) to (Y_1,K_1,x_1,v_1) is a quadruple, (W,S,γ,v), where (W,S,γ) is a cobordism of triples (so that γ is 1-manifold with boundary $\{x_0,x_1\}$) and v is a normal vector to S along γ, coinciding with the given v_0, v_1 on the boundary. Only W is required to be oriented here, providing an oriented cobordism between the 3-manifolds as usual. We call this category LINK$*$.

The construction that forms $K^\natural$ from K can be applied (in a self-evident manner) to morphisms (W,S,γ,v) in this category: one replaces S with a

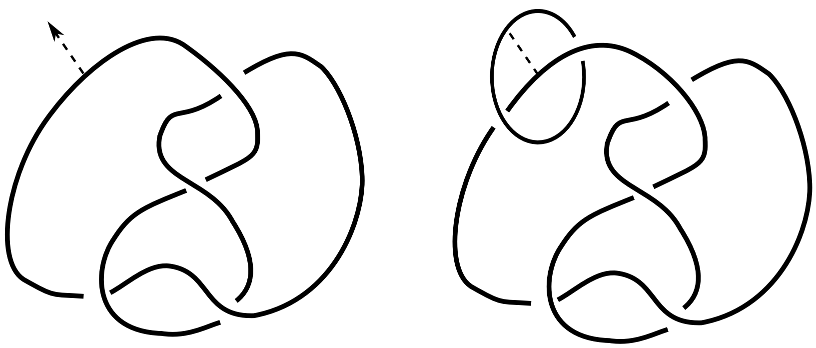

Figure 2: The link K (left) and the link $K^\natural$ (right) obtained from it by adding a meridional circle.

new cobordism

$$S^\natural = S \amalg T$$

where T is a normal circle bundle along γ, sitting in the tubular neighborhood of $S \subset W$; and one takes ω to be the I-bundle over γ with tangent direction V along γ. Thus we obtain a functor

$$\natural : \text{LINK}* \to \text{WINK}.$$

Applying instanton homology gives us a projective functor,

$$I^\natural : \text{LINK}* \to \text{P-GROUP}.$$

Of course, when considering $I^\natural(Y, K)$ as a group only up to isomorphism, we can regard it as an invariant of a link $K \subset Y$, with a marked *component*.

We have the following basic calculation:

Proposition 4.1. *For the unknot $U \subset S^3$, we have $I^\natural(S^3, U) = \mathbb{Z}$.*

Proof. The link $K^\natural$ is a Hopf link and ω is an arc joining the components. The set of critical points for the unperturbed Chern-Simons functional is the set of representations (up to conjugacy) of the fundamental group of $S^3 \setminus (K^\natural \cup \omega)$ in $SU(2)$, subject the constraint that the holonomy on the links of $K^\natural$ is conjugate to $\mathbf{i}$ and the holonomy on the links of ω is -1. There is one such representation up to conjugation, and it represents a non-degenerate critical point. So the complex that computes $I^\natural(S^3, U)$ has just a single generator. $\square$

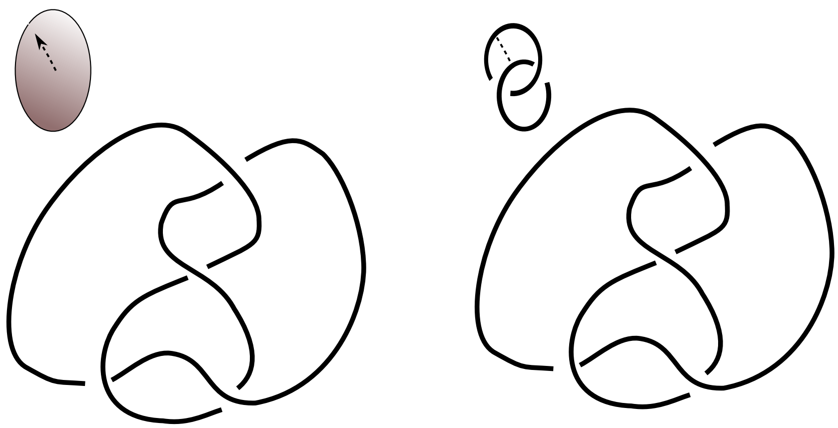

Figure 3: The link K (left) and the link $K^\sharp$ obtained as the union of K with a Hopf link located at a marked basepoint.

To obtain an invariant of a link $K \subset Y$ without need of a basepoint or marked component, we can always replace K by $K \amalg U$, where U is a new unknotted circle contained in a ball disjoint from K. We put the basepoint x on the new component U, and we define

$$K^\sharp = (K \amalg U)^\natural.$$

To say this more directly, $K^\sharp$ is the disjoint union of K and a Hopf link H contained in a ball disjoint from K, as illustrated in Figure 3. It comes with an ω which is a standard arc joining the two components of H. Thus we can define an invariant of links K without basepoint by defining

$$\begin{aligned} I^\sharp(Y, K) &= I^\omega(Y, K^\sharp) \\ &= I^\omega(Y, K \amalg H) \end{aligned}$$

To make the construction $\sharp$ functorial, we need to be given (Y, K) together with:

- a basepoint $y \in Y$ disjoint from K;
- a preferred 2-plane p in $T_y Y$;
- a vector v in p.

We can then form H by taking the first circle U to be a standard small circle in p, and then applying the $\natural$ construction to U, using v to define the framed basepoint on U. Thus $I^\sharp$ is a projective functor from a category whose objects are such marked pairs (Y, K, y, p, v). The appropriate definition of the morphisms in this category can be modeled on the example of LINK$*$ above.

The constructions $I^\natural$ and $I^\sharp$ are the "reduced" and "unreduced" variants of singular instanton Floer homology. The $I^\sharp$ variant can be applied to the empty link, and the following proposition is a restatement of Proposition 4.1.

Proposition 4.2. *For the empty link $\emptyset$ in S^3, we have $I^\sharp(S^3, \emptyset) = \mathbb{Z}$.* □

When dealing with $I^\sharp$ for classical knots and links K, we will regard K as lying in $\mathbb{R}^3$ and take the base-point at infinity in S^3. We will simply write $I^\sharp(K)$ in this context.

4.4 Resolving the ambiguities in overall sign

Thus far, we have been content with having an overall sign ambiguity in the homomorphisms on Floer homology groups which arise from cobordisms. We now turn to consider what is involved in resolving these ambiguities.

We begin with the case of BLINK, in which a typical morphism from $(Y_1, K_1, \mathbf{P}_1)$ to $(Y_0, K_0, \mathbf{P}_0)$ is represented by a cobordism with singular bundle data, $(W, S, \mathbf{P})$. We already have a rather tautological way to deal with the sign issue in this case: we need to enrich our category by including all the data we used to make the sign explicit. Thus, we can define a category $\widetilde{\text{BLINK}}$ in which an object is a tuple

$$(Y_0, K_0, \mathbf{P}_0, \mathbf{a}_0),$$

where $\mathbf{a}_0$ is the auxiliary data consisting of a choice of Riemannian metric $\check{g}_0$, a choice of perturbation π_0, and a choice of basepoint θ_0. (See equation (15) above.) A morphism in $\widetilde{\text{BLINK}}$, from $(Y_1, K_1, \mathbf{P}_1, \mathbf{a}_1)$ to $(Y_0, K_0, \mathbf{P}_0, \mathbf{a}_0)$, consists of the previous data $(W, S, \mathbf{P})$ together with a choice of an I-orientation for $(W, S, \mathbf{P})$ (Definition 3.9). With such a definition, we have a functor (rather than a projective functor)

$$I : \widetilde{\text{BLINK}} \to \text{GROUP}.$$

We can make something a little more concrete out of this for the functor

$$I^\natural : \text{LINK}* \to \text{P-GROUP}.$$

We would like to construct a category $\widetilde{\text{LINK}*}$ and a functor

$$I^\natural : \widetilde{\text{LINK}*} \to \text{GROUP}.$$

To do this, we first define the objects of $\widetilde{\text{LINK}*}$ to be quadruples (Y, K, x, v), where K is now an *oriented* link in Y and x and v are a basepoint and normal vector to K as before. (The orientation is the only additional ingredient here.) Given this data, we can orient the link $K^\natural$ by orienting the new component L so that it has linking number 1 with K in the standard ball around the basepoint x. There is a standard cobordism of oriented pairs, (Z, F), from (Y, K) to $(Y, K^\natural)$: the 4-manifold Z is a product $[0, 1] \times Y$ and the surface F is obtained from the product surface $[0, 1] \times K$ by the addition of a standard embedded 1-handle. Alternatively, (Z, F), is a boundary-connect-sum of pairs, with the first summand being $[0, 1] \times (Y, K)$ and the second summand being (B^4, A), where A is a standard oriented annulus in B^4 bounding the oriented Hopf link in S^3. The arc ω in Y joining K to L in the direction of v is part of the boundary of a disk ω_Z in Z whose boundary consists of $\omega \subset Y$ together with a standard arc lying on F.

Although $(Y, K, \emptyset)$ may not satisfy the non-integral condition, we can nevertheless form the space of connections $\mathcal{B}(Y, K) = \mathcal{B}^\emptyset(Y, K)$ as before: this is a space of $SU(2)$ connections on the link complement, with holonomy asymptotic to the conjugacy class of the element

$$\mathbf{i} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \tag{24}$$

We can now exploit that fact that there is (to within some inessential choices) a preferred basepoint θ in $\mathcal{B}(Y, K)$ arising from a reducible connection, as in section 3.6 of [23]. Thus, the singular connection θ is obtained from the trivial product connection in $SU(2) \times Y$ by adding a standard singular term

$$\beta(r)\frac{\mathbf{i}}{4}\eta$$

where β is a cut-off function on a tubular neighborhood of K and η is (as before) a global angular 1-form, constructed this time using the given orientation of K. If β is a critical point for the perturbed functional in $\mathcal{B}^\omega(Y, K^\natural)$, we let $A_{\theta,\beta}$ be any chosen connection in $\mathcal{B}^{\omega_Z}(Z, F; \theta, \beta)$ (i.e. a connection on the cobordism, asymptotic to θ and β on the two ends). We then *define*

$$\Lambda^\natural(\beta) \tag{25}$$

to be the two-element set of orientations for the determinant line $\det(\mathcal{D}_{A_{\theta,\beta}})$.

To summarize, we have defined $\Lambda^\natural(\beta)$ much as we defined $\Lambda(\beta)$ earlier in (12). The differences are that we are now using the non-trivial cobordism (Z, F, ω_Z) rather than the product, and we are exploiting the presence of a distinguished reducible connection on the other end of this cobordism, defined using the given orientation of K. We can define a chain complex

$$C^{\natural}(Y,K) = \bigoplus_{\beta} \mathbb{Z}\Lambda^{\natural}(\beta),$$

and we can regard $I^{\natural}(Y,K)$ as being defined by the homology of this chain complex.

Consider next a morphism in LINK∗, say (W,S,γ,v) from (Y_1,K_1,x_1,v_1) to (Y_0,K_0,x_0,v_0). We shall choose orientations for K_1 and K_0, as we did in the previous paragraphs. We do not assume that the surface S is an oriented cobordism, but we do require that it looks like one in a neighborhood of the path γ: that is, we assume that if u_i is an oriented tangent vector to K_i at x_i, then there is an oriented tangent vector to S along γ, normal to γ, which restricts to u_1 and u_0 at the two ends.

We have an associated morphism $(W,S^{\natural},\omega)$ between $(Y_1,K_1^{\natural},\omega_1)$ and $(Y_0,K_0^{\natural},\omega_0)$. Let

$$\begin{aligned}\beta_1 &\in \mathcal{B}^{\omega_1}(Y_1,K_1)\\ \beta_0 &\in \mathcal{B}^{\omega_0}(Y_0,K_0)\end{aligned}$$

be critical points, let $[A_{\beta_1,\beta_0}]$ be a connection in $\mathcal{B}^{\omega}(W,S^{\natural};\beta_1,\beta_0)$, and consider the problem of orienting the determinant line

$$\det(\mathcal{D}_{A_{\beta_1,\beta_0}}).$$

Let (Z_1,F_1) and (Z_0,F_0) be the standard cobordisms, described above,

$$(Z_i,F_i) : (Y_i,K_i) \to (Y_i,K_i^{\natural}).$$

There is an evident diffeomorphism, between two different composite cobordisms, from (Y_1,K_1) to $(Y_0,K_0^{\natural})$:

$$(Z_1,F_1,\omega_{Z_1}) \cup_{Y_1} (W,S^{\natural},\omega) = (W,S,\emptyset) \cup_{Y_0} (Z_0,F_0,Y_0).$$

From this we obtain an isomorphism of determinant lines,

$$\det(\mathcal{D}_{A_{\theta_1,\beta_1}}) \otimes \det(\mathcal{D}_{A_{\beta_1,\beta_0}}) = \det(\mathcal{D}_{A_{\theta_1,\theta_0}}) \otimes \det(\mathcal{D}_{A_{\theta_0,\beta_0}}). \tag{26}$$

Here θ_i are the preferred reducible connections in $\mathcal{B}(Y_i,K_i)$ as above, and A_{θ_1,θ_0} is a connection joining them across the cobordism (W,S).

If we wish the cobordism $(W,S^{\natural},\omega)$ to give rise to a chain map

$$C^{\natural}(Y_1,K_1) \to C^{\natural}(Y_0,K_0)$$

with a well-defined overall sign, then we need to specify an isomorphism

$$\det(\mathcal{D}_{A_{\beta_1,\beta_0}}) \to \mathrm{Hom}\big(\mathbb{Z}\Lambda^\natural(\beta_1), (\mathbb{Z}\Lambda^\natural(\beta_0)\big);$$

and by the definition of $\Lambda^\natural$ and the isomorphism (26), this means that we must orient

$$\det(\mathcal{D}_{A_{\theta_1,\theta_0}}).$$

Thus we are led to the following definition:

Definition 4.3. Let (Y_1, K_1) and (Y_0, K_0) be two pairs of oriented links in closed oriented 3-manifolds, and let et (W, S) be a cobordism of pairs, from (Y_1, K_1) to (Y_0, K_0), with W an oriented cobordism, and S an oriented cobordism (and possible non-orientable). Then an $I^\natural$*-orientation* for (W, S) will mean an orientation for the determinant line $\det(\mathcal{D}_{A_{\theta_1,\theta_0}})$, where θ_i are the reducible singular $SU(2)$ connections on Y_i, described above and determined by the given orientations of the K_i.

In the special case that W is a product $[0, 1] \times Y$, an $I^\natural$-orientation for an embedded cobordism S between oriented links will mean an $I^\natural$-orientation for the pair $([0, 1] \times Y, S)$. ◊

We are now in a position to define $\widetilde{\text{LINK}*}$. Its objects are quadruples (Y_i, K_i, x_i, v_i) with K_i an *oriented* link, and its morphisms are quintuples $(W, S, \gamma, v, \lambda)$, where

- (W, S) is a cobordism of pairs, with W and oriented cobordism;
- γ is a path from x_1 to x_0, with the property that S is an oriented cobordism along γ;
- v is a normal vector to S along γ, restricting to v_1, v_0 at the two ends; and
- λ is a choice of $I^\natural$-orientation for (W, S).

With this definition, we have a well-defined functor to groups.

Remarks. Our definition of $I^\natural$-orientation still rests on an analytic index, so some comments are in order. First of all, the definition makes it apparent that there is a natural composition law for $I^\natural$-orientations of composite cobordisms. Second, if S is actually an *oriented* cobordism from K_1 to K_0, then an $I^\natural$-orientation of (W, S) becomes equivalent to a homology-orientation of the cobordism, as discussed in [22] and [23]. Indeed, the case that S is oriented is precisely the case considered in [23], where homology-orientations of the cobordisms are shown to fix the signs of the corresponding chain-maps. In particular, if W is a product $[0,1] \times Y$, then an oriented cobordism S between oriented links in Y has a canonical $I^\natural$-orientation, and these canonical $I^\natural$-orientations are preserved under composition.

When using the unreduced functor $I^\sharp$ for knots in $\mathbb{R}^3$, we have adopted the convention of putting the extra Hopf link "at infinity" in a standard position. With this setup, we have a category

$$\widetilde{\text{LINK}}(\mathbb{R}^3) \tag{27}$$

whose objects are oriented links in $\mathbb{R}^3$ and whose morphisms are $I^\sharp$-oriented cobordisms in $[0,1] \times \mathbb{R}^3$.

4.5 Absolute $\mathbb{Z}/4$ gradings

For a general $(Y, K, \mathbf{P})$ and its corresponding configuration space $\mathcal{B}(Y, K, \mathbf{P})$, the path-dependent relative grading, $\text{gr}_z(\beta_1, \beta_0) \in \mathbb{Z}$, descends to a path-independent relative grading,

$$\bar{\text{gr}}(\beta_1, \beta_0) \in \mathbb{Z}/4.$$

(This is because any two paths differ by the addition of instantons and monopoles, both of which contribute multiples of 4 to the relative grading.) As a consequence, both $I^\natural(K)$ and $I^\sharp(K)$ are homology theories with an affine $\mathbb{Z}/4$ grading. In the case of $I^\natural$ however (and hence also for $I^\sharp$), we can define an *absolute* $\mathbb{Z}/4$ grading in a fairly straightforward manner. Such an absolute $\mathbb{Z}/4$ grading should assign to each $\beta \in \mathfrak{C}_\pi(Y, K^\natural)$ an element

$$\bar{\text{gr}}(\beta) \in \mathbb{Z}/4$$

(depending π and the metric, as well as on β).

Proposition 4.4. *There is a unique way to define an absolute $\mathbb{Z}/4$ grading, $\bar{\text{gr}}(\beta) \in \mathbb{Z}/4$ for $\beta \in \mathfrak{C}_\pi(Y, K^\natural)$ such that the following two conditions hold:*

(a) *the grading is normalized by having $\bar{\text{gr}}(\beta_0) = 0$ for the unique critical point in the case of the unknot in S^3 with $\pi = 0$;*

(b) *if $(W, S, \gamma, v, \lambda)$ is a morphism from (Y_1, K_1, x_1, v_1) to (Y_0, K_0, x_0, v_0) in the category* LINK∗, *and β_1, β_0 are corresponding critical points, then*

$$\text{gr}_z(W, S^\natural, \beta_1, \beta_0) = \bar{\text{gr}}(\beta_1) - \bar{\text{gr}}(\beta_0) + \iota(W, S) \pmod 4$$

where

$$\iota(W,S) = -\chi(S) + b_0(\partial^+ S) - b_0(\partial^- S) \\ - \frac{3}{2}(\chi(W) + \sigma(W)) + \frac{1}{2}(b^1(Y_0) - b^1(Y_1)).$$

Similarly, for $I^\sharp$ there is a canonical $\mathbb{Z}/4$ grading such that the generator for $I^\sharp(\emptyset)$ is in degree 0.

Proof. The uniqueness is clear. For the question of existence, we return first to a closed pair (X, Σ), with $w_2(P) = 0$ so that Δ is trivial. The dimension formula (8) tells us in this case that the index of the linearized operator $\mathcal{D}$ satisfies

$$\operatorname{index} \mathcal{D} = 4l - \frac{3}{2}(\chi(X) + \sigma(X)) + \chi(\Sigma) \pmod 4.$$

The instanton number k is an integer, so the term $8k$ can be omitted; but the monopole number l is potentially a half-integer and $2l$ is congruent to $\chi(\Sigma)$ mod 2. So the formula can be rewritten as

$$\operatorname{index} \mathcal{D} = -\frac{3}{2}(\chi(X) + \sigma(X)) - \chi(\Sigma) \pmod 4.$$

Given this formula, the existence of $\bar{\mathrm{gr}}$ now follows from the additivity of the terms involved. □

5 Two applications of Floer's excision theorem

The proof of Proposition 1.4 rests on Floer's excision theorem (slightly adapted to our situation). The statement of the excision theorem generally involves 3-manifolds Y that may have more than one component, or cobordisms with more than two boundary components. Disconnected 3-manifolds do not create any difficulties when defining instanton homology (see below for a brief review); but they do introduce a new problem when we look at cobordisms and functoriality: this problem stems from the fact that the stabilizer of an irreducible connection on a disconnected 3-manifold is no longer just ± 1, but is $(\pm 1)^n$, where n is the number of components; and not all of these elements of the stabilizer will necessarily extend to locally constant gauge transformations on the cobordism. This results in extra factors of two when gluing. The way to resolve these problems is to carefully enlarge the gauge group, by allowing some automorphisms of the $SO(3)$ bundle that do not lift to determinant-1 gauge transformations. Our first task in this section is to outline how this is done. The issue appears (and is dealt with) already in Floer's original proof of excision (as presented in [4]), but we need a more general framework.

When enlarging the gauge group in this way, the standard approach to orienting moduli spaces breaks down, and an alternative method is needed. We turn to this first.

5.1 Orientations and almost-complex structures

Given an $SO(3)$ bundle P on a closed, oriented, Riemannian 4-manifold X, we have considered the moduli space of ASD connections, $M(X, P)$, by which we mean the quotient of the space of ASD connections by the *determinant-1* gauge group $\mathcal{G}(P)$. This moduli space, when regular, is orientable; and orienting it amounts to trivializing the determinant line $\mathcal{D}_A \to \mathcal{B}(X, P)$. Two approaches to choosing a trivialization are available and described in [6]. The first relies on having a $U(2)$ lift of P (or equivalently, an integral lift v of $w_2(P)$) together with a homology orientation of X: this is the standard approach generally used in defining Donaldson's polynomial invariants. The second approach described in [6] relies on having an almost complex structure J on X: in the presence of J, there is a standard homotopy from the operators $\mathcal{D}_A$ to the complex-linear operator

$$\bar{\partial}_A + \bar{\partial}_A^* : \Omega_X^{0,1} \otimes_{\mathbb{R}} \mathfrak{g}_P \to (\Omega_X^{0,0} \oplus \Omega_X^{0,2}) \otimes_{\mathbb{R}} \mathfrak{g}_P.$$

The complex orientation of the operator at the end of the homotopy provides a preferred orientation for the determinant line.

The second of these approaches has the disadvantage of requiring the existence of J. On the other hand, when J exists, the argument provides a simple and direct proof of the orientability of the determinant line, because the homotopy can be applied to the entire family of operators over $\mathcal{B}(X, P)$. More importantly for us, this second approach to orientations can be used to establish the orientability of the determinant line over the quotient of $\mathcal{C}(X, P)$ by the *full* automorphism group $\mathrm{Aut}(P)$, not just the group $\mathcal{G}(P)$ of determinant-1 gauge transformations.

To set up this approach to orientations, we consider a pair (Y, K) as usual with singular bundle date $\mathbf{P}$. Recall that the center Z of $\mathcal{G}(\mathbf{P})$ is $\{\pm 1\}$ and that we have the isomorphism

$$\mathrm{Aut}(\mathbf{P})/(\mathcal{G}(\mathbf{P})/Z) \cong H^1(Y; \mathbb{Z}/2).$$

Thus $H^1(Y; \mathbb{Z}/2)$ acts on $\mathcal{B} = \mathcal{B}(Y, K, \mathbf{P})$. Fix a subgroup $\phi \subset H^1(Y; \mathbb{Z}/2)$, and consider the quotient

$$\bar{\mathcal{B}} = \mathcal{B}(Y, K, \mathbf{P})/\phi.$$

We shall require that the non-integral condition hold as usual, so that all critical points of the Chern-Simons functional are irreducible; but we also want ϕ to act freely on the set of critical points. This can be achieved (as the reader may verify) by strengthening the non-integral condition as follows.

Definition 5.1. We say that $(Y, K, \mathbf{P})$ satisfies the *ϕ-non-integral* condition if there is a non-integral surface $\Sigma \subset Y$ (in the sense of Definition 3.1) satisfying the additional constraint that $\phi|_\Sigma$ is zero. $\diamondsuit$

When such a condition holds, there is no difficulty in choosing ϕ-invariant holonomy perturbations so that all critical points are non-degenerate and all moduli spaces of trajectories are regular. We will write $\bar{\alpha}$, $\bar{\beta}$ for typical critical points in the quotient $\bar{\mathcal{B}}$, and $M(\bar{\alpha}, \bar{\beta})$ for the moduli spaces of trajectories. To show these moduli spaces are orientable (and to orient eventually orient them), we start on the closed manifold $S^1 \times Y$ which we equip with an S^1-invariant orbifold-complex structure: in a neighborhood of $S^1 \times K$, the model is the $\mathbb{Z}/2$ quotient of a complex disk bundle over the complex manifold $S^1 \times K$. Up to homotopy, such an almost complex structure is determined by giving a non-vanishing vector field on Y that is tangent to K along K. As described above, the family of operators $\mathcal{D}_A$ on the orbifold $S^1 \times \check{Y}$ is homotopic to a family of complex operators. As in the standard determinant-1 story, by excision, we deduce that the moduli spaces $M(\bar{\alpha}, \bar{\beta})$ on $\mathbb{R} \times (Y, K)$ are orientable. We also see that if $\bar{z}$ and $\bar{z}'$ are two homotopy classes of paths from $\bar{\alpha}$ to $\bar{\beta}$, then an orientation for $M_{\bar{z}}(\bar{\alpha}, \bar{\beta})$ determines an orientation for $M_{\bar{z}'}(\bar{\alpha}, \bar{\beta})$. This allows us to define $\bar{\Lambda}(\bar{\alpha}, \bar{\beta})$ as the two-element set that orients all of these moduli spaces simultaneously. For any given $\bar{\beta}$, we can also consider on $\mathbb{R} \times Y$ the 4-dimensional operator that interpolates between $\mathcal{D}_{\bar{\beta}}$ at the $+\infty$ end and its complex version $\bar{\partial}_{\bar{\beta}} + \bar{\partial}^*_{\bar{\beta}}$ at the $-\infty$ end. If we define $\bar{\Lambda}(\bar{\beta})$ to be the set of orientations of the determinant of this operator, then we have isomorphisms

$$\bar{\Lambda}(\bar{\alpha}, \bar{\beta}) = \bar{\Lambda}(\bar{\alpha})\bar{\Lambda}(\bar{\beta}).$$

We are therefore able to define the Floer complex in this situation by the usual recipe: we write it as

$$C(Y, K, \mathbf{P})^\phi = \bigoplus_{\bar{\beta}} \bar{\Lambda}(\bar{\beta})$$

and its homology as $I(Y,K,\mathbf{P})^{\phi}$.

The construction of $I(Y,K,\mathbf{P})^{\phi}$ in this manner depends on the choice of almost-complex structure J on $S^1\times Y$. If J and J' are two S^1-invariant complex structures, then the class

$$(c_1(J)-c_1(J'))/2$$

determines a character $\xi:\phi\to\{\pm1\}$, and hence a local system $\mathbb{Z}_\xi$ with fiber $\mathbb{Z}$ on $\bar{\mathcal{B}}=\mathcal{B}/\phi$. The corresponding Floer homology groups I and I' defined using the orientations arising from J and J' are related by

$$I'(Y,K,\mathbf{P})^{\phi}=I(Y,K,\mathbf{P};\mathbb{Z}_\xi)^{\phi},$$

as can be deduced from the calculations in [6].

In the special case that ϕ is trivial, the group $I(Y,K,\mathbf{P})^{\phi}$ coincides with $I(Y,K,\mathbf{P})$ as previously defined: a choice of isomorphism between the two depends on a choice of trivialization of J-dependent 2-element set $\bar{\Lambda}(\theta)$, where θ is the chosen base-point. Another special case is the following:

Proposition 5.2. *Suppose that ϕ is a group of order 2 in $H^1(Y;\mathbb{Z}/2)$ and that we are in one of the following two cases:*

(a) *the link K is empty and the non-zero element of ϕ has non-trivial pairing with $w_2(P)$; or*

(b) *K is non-empty and its fundamental class has non-zero pairing with ϕ.*

Then we have

$$I(Y,K,\mathbf{P})=I(Y,K,\mathbf{P})^{\phi}\oplus I(Y,K,\mathbf{P})^{\phi}.$$

Proof. In the first case, the complex $C(Y,K,\mathbf{P})$ (or simply $C(Y,P)$) has a relative $\mathbb{Z}/8$ grading, and the action of the non-trivial element of ϕ on the set of critical points shifts the grading by 4, so that $I(Y,P)^{\phi}$ is $\mathbb{Z}/4$ graded. In the second case, there is a relative $\mathbb{Z}/4$ grading on $C(Y,K,\mathbf{P})$, and the action shifts the grading by 2, so that $I(Y,K,\mathbf{P})^{\phi}$ has only a $\mathbb{Z}/2$ grading. In either case, the group $I(Y,K,\mathbf{P})$ is obtained from $I(Y,K,\mathbf{P})^{\phi}$ by simply unwrapping the grading, doubling the period from 4 to 8, or from 2 to 4 respectively. □

As a simple example, we have the following result for the 3-torus:

Lemma 5.3. *In the case that Y is a 3-torus and K is empty, let $P \to T^3$ be an $SO(3)$ bundle and ϕ any two-element subgroup of $H^1(T^3;\mathbb{Z}/2)$. Suppose that $w_2(P)$ pairs non-trivially with the non-zero element of ϕ. Then $I(T^3,P)^\phi = \mathbb{Z}$.*

Proof. This is an instance of the first item in the previous lemma, given that $I(T^3,P) = \mathbb{Z}\oplus\mathbb{Z}$. Alternatively, it can be seen directly, as the representation variety in $\mathcal{B}(T^3,P)/\phi$ consists of a single point. □

We can similarly enlarge the gauge group on 4-dimensional cobordisms, so as to make the groups $I(Y,K,\mathbf{P})^\phi$ functorial. Thus, suppose we have a cobordism $(W,S,\mathbf{P})$ from $(Y_1,K_1,\mathbf{P}_1)$ to $(Y_0,K_0,\mathbf{P}_0)$. Let $\phi \subset H^1(W;\mathbb{Z}/2)$ be chosen subgroup, and let ϕ_i be its image in $H^1(Y_i;\mathbb{Z}/2)$ under the restriction map. Suppose that the singular bundle data $\mathbf{P}_i$ satisfies the ϕ_i-non-integral condition for $i = 0,1$. After choosing metrics and perturbations, the group ϕ acts on the usual moduli spaces on $M(W,S,\mathbf{P};\beta_1,\beta_0)$, and we have quotient moduli spaces $\bar{M}(W,S,\mathbf{P};\bar{\beta}_1,\bar{\beta}_0)$. To set up orientations, we need to choose an almost-complex structure J on (W,S): i.e, an almost-complex structure on W such that S is an almost-complex submanifold. This is always possible when W has boundary, *as long as S is orientable.* (The orientability of S is a restriction on the applicability of this framework.) We denote by J_i the translation-invariant complex structure on $\mathbb{R}\times Y_i$ which arises from restricting J to the ends, and we use J_i in constructing the Floer groups $I(Y_i,K_i,\mathbf{P}_i)^{\phi_i}$ as above. Then J on (W,S) orients the moduli spaces appropriately and we have a well-defined chain map

$$C(Y_1,K_1,\mathbf{P}_1)^{\phi_1} \to C(Y_0,K_0,\mathbf{P}_0)^{\phi_0},$$

and a map on homology,

$$I(W,S,\mathbf{P})^\phi : I(Y_1,K_1,\mathbf{P}_1)^{\phi_1} \to I(Y_0,K_0,\mathbf{P}_0)^{\phi_0}.$$

In this way we have a functor (not just a projective functor) from the category whose morphisms are cobordisms $(W,S,\mathbf{P})$ equipped with a subgroup ϕ of $H^1(W;\mathbb{Z}/2)$ and an almost-complex structure J.

The use of almost-complex structures also provides a canonical mod 2 grading on the Floer groups $I(Y,K,\mathbf{P})^\phi$. In the chain complex, a generator corresponding to a critical points $\bar{\beta}$ is in even or odd degree according to the parity index of the operator whose determinant line is $\bar{\Lambda}(\bar{\beta})$, as defined above.

5.2 Disconnected 3-manifolds

We now extend the above discussion to the case of 3-manifolds that are not necessarily connected: we begin with a possibly disconnected closed, oriented 3-manifold Y, containing a link K with singular bundle data $\mathbf{P}$. The configuration space $\mathcal{B}(Y, K, \mathbf{P})$ is the product of the configuration spaces of the components. This is acted on by $H^1(Y; \mathbb{Z}/2)$, and we suppose that we are given a subgroup ϕ of this finite group with which to form the quotient space $\bar{\mathcal{B}}(Y, K, \mathbf{P})$. We require that the ϕ-non-integral condition hold on each component of Y. We choose a translation-invariant almost-complex structure J on $(\mathbb{R} \times Y, \mathbb{R} \times K)$ in order to determine orientations of moduli spaces, and after choosing metrics and perturbations we arrive at Floer homology groups

$$I(Y, K, \mathbf{P})^{\phi}$$

with no essential changes needed to accommodate the extra generality. We do not exclude the case that Y is empty, in which case its Floer homology is $\mathbb{Z}$.

Since the complex that computes $I(Y, K, \mathbf{P})^{\phi}$ is essentially a product, there is a Künneth-type theorem that describes this homology. To illustrate in the case of two components, if $Y = Y^1 \amalg Y^2$, and $\phi = \phi^1 \times \phi^2$, then there is an isomorphism of chain complexes

$$C(Y^1, K^1, \mathbf{P}^1)^{\phi^1} \otimes C(Y^2, K^2, \mathbf{P}^2)^{\phi^2} \to C(Y, K, \mathbf{P})^{\phi}$$

under which the Floer differential d on the right becomes the differential on the tensor product given by

$$d(a \otimes b) = d^1 a \otimes b + \epsilon^1 a \otimes d^2 b,$$

where ϵ^1 is the sign operator on $C(Y^1, K^1, \mathbf{P}^1)^{\phi^1}$ that is -1 on generators which have odd degree and $+1$ on generators which have even degree. As a result of this isomorphism on chain complexes, there is a short exact sequence relating the homology groups $I = I(Y, K, \mathbf{P})^{\phi}$ to the groups $I_i = I(Y_i, K_i, \mathbf{P}_i)^{\phi^i}$:

$$\mathrm{Tor}(I^1, I^2) \hookrightarrow I \twoheadrightarrow I^1 \otimes I^2. \tag{28}$$

Note that we also have an isomorphism of complexes

$$C(Y^2, K^2, \mathbf{P}^2)^{\phi^2} \otimes C(Y^1, K^1, \mathbf{P}^1)^{\phi^1} \to C(Y, K, \mathbf{P})^{\phi}$$

where the differential on the left is $d^2 \otimes 1 + \epsilon^2 \otimes d^1$. These two isomorphisms are intertwined by the map

$$a \otimes b \mapsto \epsilon^2 b \otimes \epsilon^1 a.$$

The change in sign results from the need to identify the determinant line of the direct sum of two operators with the tensor product of the determinant lines (see [24] for a discussion).

Now consider again a cobordism $(W, S, \mathbf{P})$ between (possibly disconnected) 3-manifolds with singular bundle data, $(Y_1, K_1, \mathbf{P}_1)$ and $(Y_0, K_0, \mathbf{P}_0)$. We do not require that W is connected; but to avoid reducibles, we do suppose that W has no closed components. Let ϕ be a subgroup of $H^1(W; \mathbb{Z}/2)$ and let ϕ_i be its restriction to Y_i. Let J be a complex structure on (W, S) and J_i its restriction to the two ends, $i = 1, 0$. With metrics and perturbations in place, we obtain a chain map and induced map of Floer homology groups,

$$I(W, S, \mathbf{P})^{\phi} : I(Y_1, K_1, \mathbf{P}_1)^{\phi_1} \to I(Y_0, K_0, \mathbf{P}_0)^{\phi_0}.$$

The previous definitions need no modification. There is a difference however when we consider the composition law and functoriality. Suppose that the above cobordism (W, S) is broken into the union of two cobordisms along some intermediate manifold-pair $(Y_{1/2}, K_{1/2})$, so $(W, S) = (W', S') \cup (W'', S'')$, with W' a cobordism from Y_1 to $Y_{1/2}$. By restriction, ϕ gives rise to ϕ' and ϕ'', and J gives rise to J' and J''. The composite map is equal to the map arising from the composite cobordism only if an additional hypothesis holds:

Proposition 5.4. *In the above setting, we have*

$$I(W, S, \mathbf{P})^{\phi} = I(W'', S'', \mathbf{P}'')^{\phi''} \circ I(W', S', \mathbf{P}')^{\phi'}$$

provided that the group $\phi \subset H^1(W; \mathbb{Z}/2)$ contains the image ψ of the connecting map in the Mayer-Vietoris sequence,

$$H^0(Y_{1/2}; \mathbb{Z}/2) \to H^1(W; \mathbb{Z}/2).$$

Remarks. Note that ψ is non-zero only if $Y_{1/2}$ has more than one component. In general, knowing ϕ' and ϕ'' does not determine ϕ without additional information; but the additional hypothesis that ψ is contained in ϕ, is enough to determine ϕ.

Proof of the Proposition. The hypothesis of the proposition ensures that the relevant moduli spaces on the composite cobordism have a fiber-product description. To illustrate the point in a simpler situation (so as to reduce the amount of notation involved), consider a closed Riemannian 4-manifold X decomposed into X' and X'' along a 3-manifold Y with more than one component. Suppose that the metric is cylindrical near Y. Let P be an $SO(3)$ bundle on X, let $M(X, P)$ be the moduli space of anti-self-dual connections, and suppose that all these anti-self-dual connections restrict to irreducible connections on each component of Y. We can then consider the Hilbert manifolds of anti-self-dual connections, $M(X', P')$ and $M(X'', P'')$ on the two manifolds with boundary, in a suitable L^2_k completion of the spaces of connections (as in [22, section 24]). There are restriction maps

$$\begin{array}{ccc} M(X',P') & & M(X'',P'') \\ & \searrow \quad \swarrow & \\ & \mathcal{B}_{k-1/2}(Y, P|_Y) & \end{array}$$

and one can ask whether the fiber product here is the same as $M(X, P)$. The answer is no, in general, because of our use of the determinant-1 gauge group: in the determinant-1 gauge transformation, the stabilizer of an irreducible connection on Y is $\{\pm 1\}^n$, where n is the number of components of Y; but not every element in the stabilizer can be expressed as a ratio $(g')^{-1}(g'')$ of elements in the stabilizer of the connections on X' and X''. Thus $M(X, P)$ may map many-to-one onto the fiber product. The hypothesis on ψ in the proposition ensures that we are using gauge groups for which the corresponding moduli space on X is exactly the fiber product. □

We can summarize the situation by saying that we have a functor to the category of abelian groups from a suitable cobordism category. The morphisms in this category consist of cobordisms $(W, S, \mathbf{P})$ equipped with an almost-complex structure J on (W, S) and a subgroup ϕ of $H^1(W; \mathbb{Z}/2)$. When composing two cobordisms, we must define ϕ on the composite cobordism to be the *largest* subgroup of $H^1(W; \mathbb{Z}/2)$ which restricts to the given subgroups on the two pieces.

There is a variant of the above proposition corresponding to the case that we glue one outgoing component of W to an incoming component. The context for this is that we have (W, S) a cobordism from (Y_1, K_1) to (Y_0, K_0), where

$$Y_1 = Y_{1,1} \cup \cdots \cup Y_{1,r}$$
$$Y_0 = Y_{0,1} \cup \cdots \cup Y_{0,s}$$

with $Y_{1,r} = Y_{1,s}$. We suppose also that $\mathbf{P}$ is given so that its restriction to $Y_{1,r}$ and $Y_{0,s}$ are identified. From $(W, S, \mathbf{P})$ we then form $(W^*, S^*, \mathbf{P}^*)$ by gluing these two components. This is a cobordism between manifolds (Y_1^*, K_1) and (Y_0^*, K_0^*) with $r-1$ and $s-1$ components respectively. Let ϕ^* be a subgroup of $H^1(W^*; \mathbb{Z}/2)$ *which contains the class dual to the submanifold* $Y_{1,r} = Y_{0,s}$ where the gluing has been made. Let ϕ be the subgroup of $H^1(W; \mathbb{Z}/2)$ obtained by pull-back via the map $W \to W^*$. We then have:

Proposition 5.5. *The map*

$$I(W^*, S^*, \mathbf{P}^*)^{\phi^*} : I(Y_1^*, K_1^*, \mathbf{P}_1^*)^{\phi_1^*} \to I(Y_0^*, K_0^*, \mathbf{P}_0^*)^{\phi_0^*}$$

is obtained from the map

$$I(W, S, \mathbf{P})^{\phi} : I(Y_1, K_1, \mathbf{P}_1)^{\phi_1} \to I(Y_0, K_0, \mathbf{P}_0)^{\phi_0}$$

by taking the alternating trace at the chain level over the $\mathbb{Z}/2$*-graded factor*

$$\mathrm{Hom}(C(Y_{1,r}, K_{1,r})^{\phi_{1,r}}, C(Y_{0,s}, K_{0,s})^{\phi_{0,s}}).$$

Proof. There are two issues here. The first is issue of signs, to verify that it is indeed the alternating trace that arises: this issue is dealt with in [24, Lemma 2.4], and the same argument applies here. The second issue is the choice of correct gauge groups, and this is the same point that arises in the previous proposition. □

Remark. As a simple illustration, suppose that Y is connected and (W, S) is a trivial product cobordism from (Y, K) to (Y, K). Suppose that ϕ is trivial and ϕ^* is generated by the class dual to Y in the closed manifold W^* obtained by gluing the two ends. Then the closed pair (W^*, S^*) has an associated moduli space $M(W^*, S^*, \mathbf{P}^*)^{\phi^*}$ and an associated integer invariant obtained by counting points with sign. The proposition asserts that this integer is the euler characteristic of $I(Y, K, \mathbf{P})$. If we had stuck with the determinant-1 gauge group and examined $M(W^*, S^*, \mathbf{P}^*)$ instead, then the integer invariant of this closed manifold would have been twice as large.

5.3 Floer's excision revisited

Floer's excision principle [9, 4] concerns the following situation (see also [24]). Let Y be a closed, oriented 3-manifold *not necessarily connected*, and let T_1, T_2 be a pair of oriented tori in Y, supplied with an identification $h : T_1 \to T_2$. Let Y' be obtained from Y by cutting along T_1 and T_2 and reglueing using the given identification (attaching each boundary component arising from the cut along T_1 to the corresponding boundary component arising from the cut along T_2, respecting the boundary orientations). Again, Y' need not be connected. We suppose that Y contains a link K, disjoint from the tori, and we denote by K' the resulting link in Y'. We also suppose that singular bundle data $\mathbf{P}$ is given on Y, and that the identification $h : T_1 \to T_2$ is lifted to an identification $\mathbf{P}|_{T_1} \to \mathbf{P}|_{T_2}$, so that we may form singular bundle data $\mathbf{P}'$ on Y'. We require as a hypothesis that $w_2(\mathbf{P})$ is non-zero on T_1 (and hence on T_2). From an alternative point of view, we may regard $\mathbf{P}$ as being determined by a 1-manifold ω, in which case we ask that $\omega \cdot T_1$ is odd and that h maps the transverse intersection $\omega \cap T_1$ to $\omega \cap T_2$, so that we may form ω' in Y'.

When K and K' are absent (which was the case in Floer's original setup), the condition that $\omega \cdot T_1$ is odd forces T_1 to be non-separating. In the presence of K, however, it may be that T_1 separates Y, in which case ω must have an arc which joins components of K which lie in different components of $Y \setminus T_1$.

Let $\phi \subset H^1(Y; \mathbb{Z}/2)$ be the subgroup generated by the duals of T_1 and T_2, and let $\phi' \subset H^1(Y'; \mathbb{Z}/2)$ be defined similarly using the tori in Y'. As long as Y and Y' have no components disjoint from the tori, the ϕ-non-integral condition is satisfied on account of the presence of the surfaces T_i. If there are any other components of Y (and therefore of Y'), we impose the non-integral condition as a hypothesis, as usual (though such components are irrelevant in what follows). In order to fix signs, we use almost-complex structures as in section 5.1: we fix an $\mathbb{R}$-invariant complex structure J on $(\mathbb{R} \times Y, \mathbb{R} \times K)$, which we choose in such a way that the manifolds $\{0\} \times T_i$ are almost-complex submanifolds (with their given orientations). By cutting and gluing we also obtain an almost complex structure J' on Y'.

Theorem 5.6. *Under the above hypotheses, there are mutually-inverse isomorphisms*

$$I(Y, K, \mathbf{P})^{\phi} \longleftrightarrow I(Y', K', \mathbf{P}')^{\phi'},$$

or equivalently

$$I^{\omega}(Y, K)^{\phi} \longleftrightarrow I^{\omega'}(Y', K')^{\phi'},$$

arising from standard cobordisms (W, S) and $(\bar{W}, \bar{S})$, from (Y, K) to (Y', K') and from (Y', K') to (Y, K) respectively.

Proof. The standard cobordism W from Y to Y' is the one that appears in Floer's original theorem and is illustrated in [24, Figure 2]. The relevant part of this cobordism is redrawn here in Figure 4. This part is a product $T \times U$, where T is the torus obtained by identifying T_1 and T_2 using h and U is the 2-manifold with corners depicted as the shaded part in the figure. The subset $T \times U \subset W$ meets Y and Y' in 2-sided collar neighborhoods of $T_1 \cup T_1$ and $T_1' \cup T_2'$. The links K and K' are contained in the parts of Y and Y' that are disjoint from these collars, and there is a surface $S = [0, 1] \times K$ lying in the remaining part of W, providing the cobordism between them. Our choices equip W with singular bundle date $\mathbf{P}_W$ and an almost-complex structure J_W which is a product structure on the subset $T \times U$.

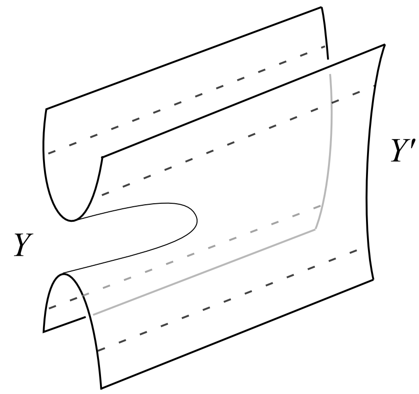

Figure 4: The excision cobordism W from Y to Y'.

Let T_i^+ and T_i^- be positive and negative push-offs of T_i in Y_i (for $i = 1, 2$). The dual classes to these four tori redundantly generate the same group $\phi \subset H^1(Y; \mathbb{Z}/2)$. Let $(T_i')^{\pm}$ be defined similarly in Y'. These eight tori sit over the eight corners of U in the figure. For each of the four dotted edges of U, there is copy of $[0, 1] \times T$ lying above it in W. Let $\phi_W \subset H^1(W; \mathbb{Z}/2)$ be the subgroup generated by the classes dual to these four copies of $[0, 1] \times T$. This subgroup restricts to ϕ and ϕ' on the two ends. Equipped with this data and its complex structure J_W, the cobordism thus gives rise to a map

$$I(W, S, \mathbf{P}_W)^{\phi_W} : I(Y, K, \mathbf{P})^{\phi} \longrightarrow I(Y', K', \mathbf{P}')^{\phi'}.$$

An entirely symmetrical construction gives a map in the opposite direction,

$$I(\bar{W}, \bar{S}, \mathbf{P}_{\bar{W}})^{\phi_{\bar{W}}} : I(Y', K', \mathbf{P}')^{\phi'} \longrightarrow I(Y, K, \mathbf{P})^{\phi}.$$

To show that the maps arising from these cobordisms are mutually inverse, we follow Floer's argument, as described in [4, 24]. The setup is symmetrical, so we need only consider one composite, say the union

$$V = W \cup_{Y'} \bar{W}$$

with its embedded surface $S_V = S \cup \bar{S}$. This composite cobordism, part of which is depicted in Figure 5, contains four copies of $[0, 2] \times T$ (lying above the dotted lines again). We take

$$\phi_V \subset H^1(V; \mathbb{Z}/2)$$

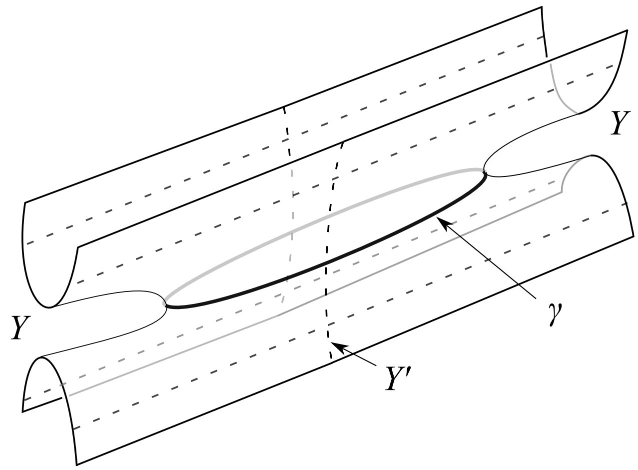

Figure 5: The composite cobordism $V = W \cup_{Y'} \bar{W}$ from Y to Y.

to be the subgroup generated by the dual classes of these four hypersurfaces together with the image of the boundary map in the Mayer-Vietoris sequence: i.e. the classes dual to the components of $Y' \subset V$. The complex structure on V is the product structure on the subset $T \times (U \cup \bar{U})$ shown in the figure. By Proposition 5.4, the composite of the maps obtained from the cobordisms (W, S) and $(\bar{W}, \bar{S})$ is the map

$$I(V, S_V, \mathbf{P}_V)^{\phi_V}.$$

So we must show that this map is the identity.

Floer's proof rests on the fact that the cobordism V can be changed to a product cobordism by cutting along the 3-torus $T \times \gamma$ (where γ is the curve depicted in the figure) and gluing two copies of $T \times D^2$ to the resulting boundary components. If we write V' for the resulting product cobordism, we see that it contains a product surface S_V and acquires from V (by cutting and gluing) a complex structure $J_{V'}$ which respects the product structure. Furthermore, the subgroup ϕ_V becomes a group $\phi_{V'}$ in $H^1(V; \mathbb{Z}/2)$ which is the pull-back of ϕ from $H^1(Y; \mathbb{Z}/2)$. Thus

$$I(V', S_{V'}, \mathbf{P}_{V'})^{\phi_{V'}} = \mathrm{Id}$$

and we are left with the task of showing that V and V' (with their attendant structures) give the same map on $I(Y, S, \mathbf{P})^{\phi}$. This last task is easily accomplished by using the fact that ϕ_V restricts to the 3-torus $T \times \gamma$ to give the non-trivial 2-element subgroup of $H^1(T \times \gamma; \mathbb{Z}/2)$ generated by $T \times \{\text{point}\}$, so that

$$I(T \times \gamma, P_V)^{\phi_V} = \mathbb{Z},$$

by Lemma 5.3. The subgroup ϕ_V also contains the class dual to $T \times \gamma$, so that Proposition 5.5 applies. The relative invariant of $T \times D^2$ is $1 \in \mathbb{Z}$ in the Floer group of $T \times \gamma$, so the result follows, just as in Floer's original argument. □

5.4 Proof of Proposition 1.4

We now apply the excision principle, Theorem 5.6, to prove Proposition 1.4 from the introduction. The proposition can be generalized to deal with knots K in 3-manifolds other than S^3, so we consider a general connected, oriented Y and a knot $K \subset Y$. From (Y, K) we form a new closed 3-manifold T^3_K, depending on a choice of framing for K, as follows. We take a standard coordinate circle $C \subset T^3$ and we glue together the knot complements to form

$$T^3_K = \big(T^3 \backslash N^\circ(C)\big) \cup \big(Y \backslash N^\circ(K)\big),$$

gluing the longitudes of K (for the chosen framing) to the meridians of C and vice versa. In T^3, let R be coordinate 2-torus parallel to C and disjoint from it, and let $\psi \subset H^1(T^3_K; \mathbb{Z}/2)$ be the 2-element subgroup generated by the dual to $[R]$ in T^3_K. Let $\omega_1 \subset T^3$ be another coordinate circle transverse to R and meeting it in one point.

We may now consider the instanton Floer homology group $I^{\omega_1}(T^3_K)$, as well as its companion $I^{\omega_1}(T^3_K)^\psi$ defined using the larger gauge group. From the first half of Proposition 5.2, we have

$$I^{\omega_1}(T^3_K) = I^{\omega_1}(T^3_K)^\psi \oplus I^{\omega_1}(T^3_K)^\psi. \tag{29}$$

Our application of excision is the following result:

Proposition 5.7. *There is an isomorphism $I^\natural(Y, K) \cong I^{\omega_1}(T^3_K)^\psi$, respecting the relative $\mathbb{Z}/4$ grading of the two groups.*

Proof. By definition, we have $I^\natural(Y, K) = I^\omega(Y, K^\natural)$, where $K^\natural$ is the union of K and a meridional circle L and ω is an arc joining the two components. The circle L has a preferred framing, because it is contained in a small ball in Y. Let $N(K)$ be a tubular neighborhood of K that is small enough to be disjoint from L, and let T_1 be its oriented boundary. Let T_2 be the oppositely-oriented boundary of a disjoint tubular neighborhood of L in Y.

Let $h : T_1 \to T_2$ be an orientation-preserving diffeomorphism that maps the longitudes of K to the meridians of L and vice versa, and maps $\omega \cap T_1$ to $\omega \cap T_2$. We can cut along T_1 and T_2 and reglue using h. Theorem 5.6 applies. Note that because T_1 and T_2 are null-homologous, the relevant subgroup ϕ in $H^1(Y;\mathbb{Z}/2)$ is trivial. The pair (Y', K') that results from cutting and gluing has two components. One component is a sphere S^3 containing K' which is the standard Hopf link H. The other component is T^3_K. The tori T'_1 and T'_2 are respectively a standard torus separating the two components of the Hopf link in S^3, and the torus R in T^3_K. The subgroup of $H^1(Y';\mathbb{Z}/2)$ that they generate is the 2-element group ψ, while ω' is the union of an arc joining the two components of H and the coordinate circle ω_1 in T^3. Since $I^{\omega'}(S^3, H) = \mathbb{Z}$, the excision theorem gives

$$\begin{aligned} I^{\omega}(Y, K^{\natural}) &\cong I^{\omega'}(S^3, H) \otimes I^{\omega'}(T^3_K)^{\psi} \\ &= I^{\omega_1}(T^3_K)^{\psi}, \end{aligned}$$

which is what the proposition claims. □

The group $I^{\omega_1}(T^3_K)$ (defined using the determinant-1 gauge group) is exactly the group that Floer associates to a knot K in [9, section 3]. In that paper, this homology group is written $I_*(P, K)$, but to keep the distinctions a little clearer, let us write Floer's group as $I^{\text{Floer}}(Y, K)$. Because of the relation (29), we can recast the result of the previous proposition as

$$I^{\text{Floer}}(Y, K) = I^{\natural}(Y, K) \oplus I^{\natural}(Y, K).$$

On the other hand, in [24], it was explained that, over $\mathbb{Q}$ at least, one can decompose $I^{\text{Floer}}(Y, K)$ into the generalized eigenspaces of degree-4 operators $\mu(\text{point})$, belonging to the eigenvalues 2 and -2. The generalized eigenspace for $+2$ is, by definition, the group $\mathit{KHI}(Y, K)$ of [24]. Since the two generalized eigenspaces are of equal dimension, we at least have

$$I^{\text{Floer}}(Y, K; \mathbb{Q}) = \mathit{KHI}(Y, K; \mathbb{Q}) \oplus \mathit{KHI}(Y, K; \mathbb{Q})$$

as vector spaces. Thus we eventually have

$$I^{\natural}(Y, K) \otimes \mathbb{Q} \cong \mathit{KHI}(Y, K; \mathbb{Q})$$

as claimed in Proposition 1.4. □

5.5 A product formula for split links

For a second application of Floer's excision theorem, consider a pair of connected 3-manifolds Y_1, Y_2, and their connected sum $Y_1 \# Y_2$, as well as their disjoint union $Y = Y_1 \cup Y_2$. Given links $K_i \subset Y_i$ for $i = 1, 2$, chosen so as to be disjoint from the embedded balls that are used in making the connected sum, we obtain a link $K_1 \cup K_2$, in $Y_1 \# Y_2$. In the special case that Y_1 and Y_2 are both S^3, the resulting link is a split link (as long as both K_i are non-empty). Of course, we can also form the union K in $Y = Y_1 \cup Y_2$. Note that any cobordism S from K to K' in the disjoint union Y gives rise also to a "split" cobordism S between the corresponding links in $Y_1 \# Y_2$, as long it is disjoint from the balls.

Let $H_i \subset Y_i$ be a Hopf link contained in a standard ball, and let ω_i be an arc joining its two components, so that

$$I^\sharp(Y_i, K_i) = I^{\omega_i}(Y_i, K_i \cup H_i).$$

Write $\omega = \omega_1 \cup \omega_2$. The group

$$I^\omega(Y, K \cup H_1 \cup H_2)$$

is isomorphic to the tensor product

$$I^\sharp(Y_1, K_1) \otimes I^\sharp(Y_2, K_2)$$

over the rationals, on a account of the Künneth theorem (28). On the other hand, it is also related to the connected sum:

Proposition 5.8. *There is an isomorphism*

$$I^\omega(Y, K \cup H_1 \cup H_2) \cong I^\sharp(Y_1 \# Y_2, K)$$

which respects the $\mathbb{Z}/4$ gradings, and is natural for "split cobordisms".

Proof. In Y, let T_1 be a the boundary of a tubular neighborhood of one of the two components of the Hopf link $H_1 \subset Y_1$, and let $T_2 \subset Y_2$ be defined similarly, but with the opposite orientation. Choose an orientation-preserving diffeomorphism h between these tori, interchanging longitudes with meridians. The manifold (Y', K') obtained by cutting and gluing as in the excision theorem is disconnected. Its first component Y_1' is a 3-sphere containing a standard Hopf link K_1'. Its second component Y_2' is the connected sum $Y_1 \# Y_2$ containing the link

$$K_2' = K_1 \cup K_2 \cup H',$$

where H' is a standard Hopf link. All the tori involved in this application of excision are null-homologous, so no non-trivial subgroups of $H^1(Y;\mathbb{Z}/2)$ are involved. Thus we obtain from Theorem 5.6 an isomorphism

$$\begin{aligned} I^{\omega}(Y, K\cup H_1\cup H_2) &\cong I^{\omega'}(Y_1', K_1')\otimes I^{\omega'}(Y_2', K_2') \\ &= I^{\omega'}(Y_1', K_1')\otimes I^{\omega'}(Y_1\# Y_2, K_1\cup K_2\cup H') \end{aligned} \tag{30}$$

On the right, the first factor is $\mathbb{Z}$. The curve ω' in the second factor is an arc joining the two components of the Hopf link H'. So the right-hand side is simply $I^{\sharp}(Y_1\# Y_2, K_1\cup K_2)$ as desired. □

Corollary 5.9. *If at least one of $I^{\sharp}(Y_i, K_i)$ is torsion-free, we have an isomorphism*

$$I^{\sharp}(Y_1, K_1)\otimes I^{\sharp}(Y_2, K_2)\to I^{\sharp}(Y_1\# Y_2, K_1\cup K_2)$$

arising from an excision cobordism. The isomorphism is natural for the maps induced by "split" cobordisms. □

6 Cubes

6.1 The skein cobordisms

We consider three links K_2, K_1 and K_0 in a closed 3-manifold Y which are related by the unoriented skein moves, as shown in Figure 6. What this means is that there is a standard 3-ball in Y outside which the three links coincide, while inside the ball the three links appear as shown. Two alternative views are given in the figure. In the top row, we draw the picture as it is usually presented for classical links, when a projection in the plane is given: here the links K_1 and K_0 have one fewer crossings that K_2. In the bottom row of the figure, an alternative picture is drawn which brings out the symmetry more clearly: we see that there is a round ball B^3 in Y with the property that all three links meet the boundary sphere in four points. These four points form the vertices of a regular tetrahedron, and the links K_2, K_1 and K_0 are obtained from the three different ways of joining the four vertices in two pairs, by pairs of arcs isotopic to pairs of edges of the tetrahedron.

The second view in Figure 6 makes clear the cyclic symmetry of the three links. Note that there *is* a preferred cyclic ordering determined by the pictures: if the picture of K_i is rotated by a right-handed one-third turn about any of the four vertices of the tetrahedron, then the result is the

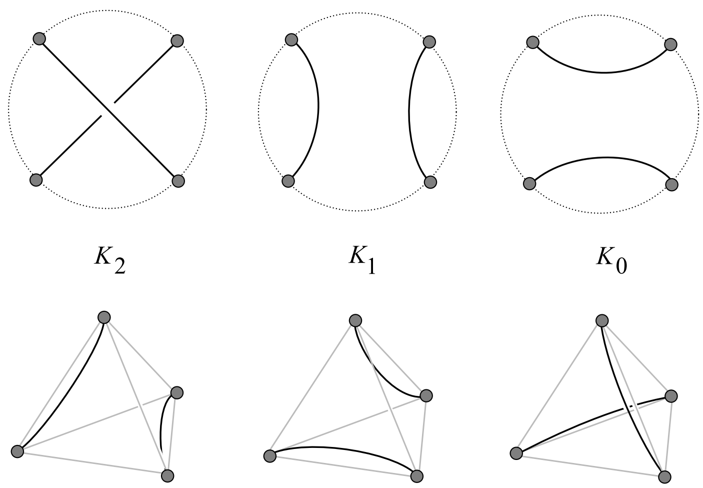

Figure 6: Knots K_2, K_1 and K_0 differing by the unoriented skein moves, in two different views.

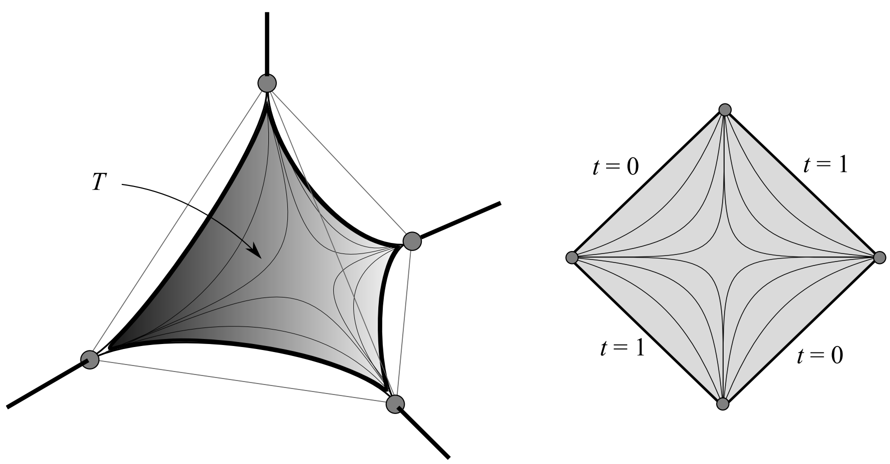

Figure 7: The twisted rectangle T that gives rise to the cobordism $S_{2,1}$ from K_2 to K_1.

picture of K_{i-1}. We may consider links K_i for all integers i by repeating these three cyclically.

We next describe a standard cobordism surface $S_{i,i-1}$ from K_i to K_{i-1} inside the cylindrical 4-manifold $[0,1] \times Y$, for each i. Because of the cyclic symmetry, it is sufficient to describe the surface $S_{2,1}$. This cobordism will be a product surface outside $[0,1] \times B^3$. Inside $[0,1] \times B^3$, the first coordinate $t \in [0,1]$ will have a single index-1 critical point on $S_{2,1}$. The intrinsic topology of $S_{2,1}$ is therefore described by the addition of a single 1-handle. To describe the geometry of the embedding, begin with the 1/4-twisted rectangular surface $T \subset B^3$ shown in Figure 7, and let T^o be the complement of the 4 vertices of T. Let t be a Morse function on T^o with $t = 0$ on the two arcs of K_2 and $t = 1$ on the two arcs of K_1, and with a single critical point on the center of T with critical value 1/2. The graph of this Morse function places T^o into $[0,1] \times B^3$. The cobordism $S_{2,1}$ is the union of this graph with the product part outside $[0,1] \times B^3$.

We put an orbifold Riemannian metric $\check{g}$ on $[0,1] \times Y$ (with cone-angle π along the embedded surface $S_{2,1}$ as usual). We choose the metric so that it is a cylindrical product metric of the form

$$dt^2 + \check{g}_Y$$

on the subset $[0,1] \times (Y \setminus B^3)$. We also require that the metric be cylindrical in collar neighborhoods of $\{0\} \times Y$ and $\{1\} \times Y$.

Suppose now that instead of a single ball we are given N disjoint balls $B_1, \ldots, B_N$ in Y. Generalizing the above notation, we may consider a collection of links $K_v \subset Y$ for $v \in \{0,1,2\}^N$: all these links coincide outside the union of the balls, while $K_v \cap B_i$ consists of a pair of arcs as in Figure 6, according as the i-th coordinate v_i is 0, 1 or 2. We extend this family of links to a family parametrized by $v \in \mathbb{Z}^N$, making the family periodic with period 3 in each coordinate v_i.

We give $\mathbb{Z}^N$ the product partial order. We also define norms

$$|v|_\infty = \sup_i |v_i|$$
$$|v|_1 = \sum_i |v_i|.$$

Then for a pair $v, u \in \mathbb{Z}^N$ with $v \geq u$ and $|v-u|_\infty = 1$, we define a cobordism S_{vu} from K_v to K_u in $[0,1] \times Y$ by repeating the construction of S_{v_i,u_i} for each ball B_i for which $v_i = u_i + 1$. For the sake of uniform notation, we also write S_{vv} for the product cobordism. These cobordisms satisfy

$$S_{wu} = S_{vu} \circ S_{wv}$$

whenever $w \geq v \geq u$ with $|w - u|_\infty \leq 1$.

It is notationally convenient to triangulate $\mathbb{R}^N$ as a simplicial complex R^N with vertex set $\mathbb{Z}^N$ by declaring the n-simplices to be all ordered $(n+1)$-tuples of vertices $(v^0, \dots, v^n)$ with

$$v^0 > v^1 > \cdots > v^n$$

and $|v^0 - v^n|_\infty \leq 1$. In this simplicial decomposition, each unit cube in $\mathbb{Z}^N$ is decomposed into $N!$ simplices of dimension N. The non-trivial cobordisms S_{vu} correspond to the 1-simplices (v, u) of R^N. We also talk of *singular n-simplices* for this triangulation, by which we mean $(n+1)$-tuples $(v^0, \dots, v^n)$ with

$$v^0 \geq v^1 \geq \cdots \geq v^n$$

and $|v^0 - v^n|_\infty \leq 1$. We can regard these singular simplices as the generators of the singular simplicial chain complex, which computes the homology of $\mathbb{R}^N$.

We will be applying instanton homology, to associate a chain complex $C^\omega(Y, K_v)$ to each link K_v. To do so, we need to have an ω so that (Y, K_v, ω) satisfies the non-integral condition. We choose an ω which is disjoint from all the balls $B_1, \dots, B_N$. Such a choice for one K_v allows us to use the same ω for all other K_u. When considering the cobordisms S_{vu}, we extend ω as a product. We impose as a hypothesis the condition *that (Y, K_v, ω) satisfies the non-integral condition, for all v.*

We put an orbifold metric $\check{g}_v$ on (Y, K_v) for every v, arranging that these are all isometric outside the union of the N balls. As in the case $N = 1$ above, we then put an orbifold metric $\check{g}_{vu}$ on the cobordism of pairs, $([0,1] \times Y, S_{vu})$ for every 1-simplex (v, u) of R^N. We choose these again so that they are product metrics in the neighborhood of $\{0\} \times Y$ and $\{1\} \times Y$, and also on the subset $[0,1] \times (Y \setminus B)$, where B is the union of the balls. Inside $[0,1] \times B_i$ we can take standard metric for the cobordism, which depends only on i, not otherwise on u or v.

In order to have chain-maps with a well-defined sign, we need to choose I-orientations for all the cobordisms that arise. From the definition, this entails first choosing singular bundle data

$$\mathbf{P}_v \to (Y, K_v)$$

for each v, corresponding to the chosen ω. It also entails choosing auxiliary data $\mathbf{a}_v$ (a metric, perturbation and basepoint in $\mathcal{B}^\omega(Y, K_v)$), for all v. The metric is something we have already discussed, but for the perturbation and basepoint we make arbitrary choices. We fix $(\mathbf{P}_v, \mathbf{a}_v)$ once and for all, and make no further reference to them. For each cobordism S_{vu}, we have a well-defined notion of an I-orientation, in the sense of Definition 3.9. We wish to choose I-orientations for all the cobordisms, so that they behave coherently with respect to compositions. The following lemma tells us that this is possible.

Lemma 6.1. *It is possible to choose I-orientations μ_{vu} for each cobordism S_{vu}, so that whenever (w, v, u) is a singular 2-simplex of R^N, the corresponding I-orientations are consistent with the composition, so that*

$$\mu_{wu} = \mu_{vu} \circ \mu_{wv}. \tag{31}$$

Proof. The proof only depends on the fact that the composition law for I-orientations is associative. Begin by choosing an arbitrary I-orientation μ'_{vu} for each singular 1-simplex (v, u). For any singular 2-simplex (w, v, u), define $\eta(w, v, u) \in \mathbb{Z}/2$ according to whether or not the desired composition rule (31) holds: that is,

$$\mu'_{wu} = (-1)^{\eta(w,v,u)} \mu'_{vu} \circ \mu'_{wv}.$$

We seek new orientations

$$\mu_{vu} = (-1)^{\theta(v,u)} \mu'_{vu}$$

so that the (31) holds for all 2-simplices. The θ that we seek can be viewed as a 1-cochain on R^N with values in $\mathbb{Z}/2$, and the desired relation (31) amounts to the condition that the coboundary of θ is η:

$$\delta\theta = \eta.$$

Because the second cohomology of $\mathbb{R}^N$ is zero, we can find such a θ if and only if η is coclosed.

To verify that η is indeed coclosed, consider a singular 3-simplex (w, v, u, z). The value of $\delta\eta$ on this 3-simplex is the sum of the values of η on its four faces. Choose any I-orientation ν for S_{wz}. There are four paths from w to z along oriented 1-simplices, $\gamma_0, \ldots, \gamma_3$. Viewing these as singular 1-chains, they are:

$$\begin{aligned}\gamma_0 &= (w,z)\\ \gamma_1 &= (w,u)+(u,z)\\ \gamma_2 &= (w,v)+(v,u)+(u,z)\\ \gamma_3 &= (w,v)+(v,z).\end{aligned}$$

For each path γ_a, $a = 0,\ldots,3$, define $\phi_a \in \mathbb{Z}/2$ by declaring that $\phi_a = 0$ if and only if the composite of the chosen I-orientations μ' along the 1-simplices of γ_a is equal to ν. Thus, for example,

$$\mu'_{wv} \circ \mu'_{vu} \circ \mu'_{uz} = (-1)^{\phi_2}\nu.$$

In the given cyclic ordering of the 1-chains γ_a, the differences $\gamma_{a+1} - \gamma_a$ is the boundary of a face of the 3-simplex, for each $a \in \mathbb{Z}/4$. Thus we see,

$$\begin{aligned}\phi_1 - \phi_0 &= \eta(w,u,z)\\ \phi_2 - \phi_1 &= \eta(w,v,u)\\ \phi_3 - \phi_2 &= \eta(v,u,z)\\ \phi_0 - \phi_3 &= \eta(w,v,z).\end{aligned}$$

The sum of the value of η on the 4 faces of the singular 3-simplex is therefore zero, as required. □

To define maps in the Floer homology we make the pair $([0,1]\times Y, S_{vu})$ into a pair of cylindrical manifolds, by adding the half cylinders $(-\infty,0\,]\times(Y,K_v)$ and $[\,1,\infty)\times(Y,K_u)$. The four manifold is simply $\mathbb{R}\times Y$, and we call the $\mathbb{R}$ coordinate t. The metric is locally a product metric everywhere except on the subset

$$[0,1]\times(B_{i_1}\cup\cdots\cup B_{i_d})$$

where the union is over all i with $v_i \neq u_i$. With a slight abuse of notation, we continue to denote by S_{vu} the non-compact embedded surface with cylindrical ends,

$$S_{vu} \subset \mathbb{R}\times Y.$$

This orbifold metric on $\mathbb{R}\times Y$ is one of a natural family of metrics, which we now describe. Let us write

$$\begin{aligned}I &= \{\, i_1,\ldots,i_d \,\}\\ &= \operatorname{supp}(v-u)\\ &\subset \{\, 1,\ldots,N \,\},\end{aligned}$$

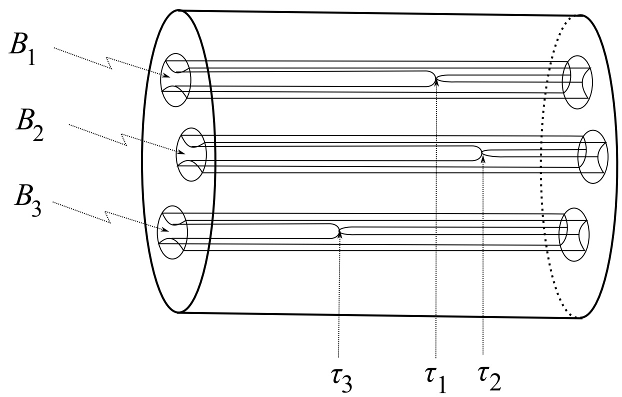

Figure 8: The family of metrics G_{vu} parametrized by $\tau \in \mathbb{R}^I$ ($\mathbb{R}^3$ in this example).

and let us denote our original orbifold metric by $\check{g}_{vu}(0)$. For each

$$\tau = (\tau_{i_1}, \ldots, \tau_{i_d}) \in \mathbb{R}^I \cong \mathbb{R}^d,$$

we construct a Riemannian manifold

$$(\mathbb{R} \times Y, \check{g}_{vu}(\tau))$$

by starting with $(\mathbb{R} \times Y, \check{g}_{vu}(0))$, cutting out the subsets $\mathbb{R} \times B_{i_m}$ for $m = 1, \ldots, d$, and gluing them back via the isometry of the boundaries $\mathbb{R} \times S^3$ given by translating the t coordinate by τ_{i_m}. After an adjustment of our parametrization, we can assume that the t coordinate on $\mathbb{R} \times \check{Y}$ has exactly d critical points when restricted to the singular locus S_{vu}, and that these occur in $\{\tau_{i_m}\} \times B_{i_m}$ for $m = 1, \ldots, d$. (See Figure 8.) We write G_{vu} for this family of Riemannian metrics.

Suppose now that (w, v, u) be a singular 2-simplex. Let I and J be the support of $w - v$ and $v - u$ respectively, so that $I \cup J$ is the support of $w - u$ and $I \cap J$ is empty. There is a natural identification

$$G_{wu} \to G_{wv} \times G_{vu} \tag{32}$$

arising from $\mathbb{R}^{I \cup J} \to \mathbb{R}^I \times \mathbb{R}^J$.

Lemma 6.2. *Orientations can be chosen for G_{vu} for all singular 1-simplices (v, u) of R^N such that for all singular 2-simplices (w, v, u), the natural identification (32) is orientation-preserving.*

Proof. The proof is essentially the same as the proof of the previous lemma. □

There is an action of $\mathbb{R}$ on the space of metrics G_{vu} by translation (adding a common constant to each coordinate of $\tau \in \mathbb{R}^d$). We can therefore normalize τ by requiring that

$$\sum \tau_{i_a} = 0.$$

We write $\breve{G}_{vu} \subset G_{vu}$ for this normalized family (which we can also regard as the quotient of G_{vu} by the action of translations). As coordinates on $\breve{G}_{vu}$ we can take the differences

$$\tau_{i_{a+1}} - \tau_{i_a}, \qquad a = 1, \ldots, d.$$

There is a natural compactification of $\breve{G}_{vu}$, which we can think of informally as resulting from allowing some of the differences to become infinite: it parametrizes a family of broken Riemannian metrics of the sort considered in section 3.9. This compactification, which we call $\breve{G}^+_{vu}$, is constructed as follows. We consider all simplices $\sigma = (v^0, \ldots, v^n)$ with $v^0 = v$ and $v^n = u$ (including the 1-simplex (v, u) itself amongst these). For each such simplex σ, we write

$$\breve{G}_\sigma = \breve{G}_{v^0 v^1} \times \cdots \times \breve{G}_{v^{n-1} v^n}.$$

The compactification $\breve{G}^+_{vu}$ is then the union

$$\breve{G}^+_{vu} = \bigcup_\sigma \breve{G}_\sigma. \tag{33}$$

The definition of the topology on this union follows the usual approach for broken trajectories: see for example [22]. The space $\breve{G}^+_{vu}$ is a polytope: if σ is an n-simplex, then the corresponding subset $\breve{G}_\sigma \subset \breve{G}^+_{vu}$ is the interior of a face of codimension $n - 1$. In particular, the codimension-1 faces of the compactification are the parts

$$\breve{G}_{vs} \times \breve{G}_{su}$$

for all s with $v > s > u$. Thus each face parametrizes broken Riemannian metrics broken along a single cut (Y, K_s), for some s. (A family of Riemannian metrics with much the same structure occurs in the same context in [3], where it is observed that the polytope is a permutohedron.)

In section 3.9, when we considered general families of broken metrics, we chose to orient the boundary faces of the family using the boundary orientation. In our present situation, we need to compare the naturally-arising orientations to the boundary orientation:

Lemma 6.3. *Suppose that orientations have been chosen for G_{vu} for all 1-simplices (v,u) so as to satisfy the conditions of Lemma 6.2. Orient $\breve{G}_{vu}$ by making the identification*

$$G_{vu} = \mathbb{R} \times \breve{G}_{vu}$$

where the $\mathbb{R}$ coordinate is the center of mass of the coordinates τ_{i_a} on G_{vu}. Then for any 2-simplex (w,v,u) the product orientation on $\breve{G}_{wv} \times \breve{G}_{vu}$ differs from the boundary orientation of

$$\breve{G}_{wv} \times \breve{G}_{vu} \subset \partial \breve{G}_{wu}$$

by the sign $(-1)^{\dim \breve{G}_{wv}}$.

Proof. From the identification $G_{wu} = G_{wv} \times G_{vu}$ we obtain an orientation-preserving identification

$$\mathbb{R}_0 \times \breve{G}_{wu} \cong (\mathbb{R}_1 \times \breve{G}_{wv}) \times (\mathbb{R}_2 \times \breve{G}_{vu}),$$

where the $\mathbb{R}_0$, $\mathbb{R}_1$, $\mathbb{R}_2$ factors correspond to centers of mass of the appropriate τ_i. Thus, the $\mathbb{R}_0$ coordinate on the left is a positive weighted sum of the $\mathbb{R}_1$ and $\mathbb{R}_2$ coordinates on the right. This becomes an orientation-preserving identification

$$\breve{G}_{wu} \cong \breve{G}_{wv} \times \mathbb{R}_3 \times \breve{G}_{vu},$$

where the coordinate $\mathbb{R}_3$ is related to the previous $\mathbb{R}_1$ and $\mathbb{R}_2$ coordinates by $t_3 = t_2 - t_1$. The boundary component $\breve{G}_{wv} \times \breve{G}_{vu}$ in $\partial \breve{G}^+_{wu}$ arises by letting the $\mathbb{R}_3$ coordinate go to $+\infty$, so

$$\breve{G}^+_{wu} \supset \breve{G}_{wv} \times (-\infty, +\infty] \times \breve{G}_{vu}.$$

The orientation of the boundary is determined by the outward-normal-first convention, which involves switching the order of the first two factors on the right. This introduces the sign $(-1)^{\dim \breve{G}_{wv}}$. $\square$

6.2 Maps from the cobordisms

We continue to consider the collection of links K_v in Y indexed by $v \in \mathbb{Z}^N$. Recall that we have singular bundle data $\mathbf{P}_v$ over each (Y, K_v) satisfying the non-integral condition, and fixed auxiliary data $\mathbf{a}_v$, so that we realize explicit chain complexes for each Floer homology group: we write

$$C_v = C_*(Y, K_v, \mathbf{P}_v)$$

for this chain complex. Its homology is $I^{\omega}(Y,K_v)$. We write $\mathfrak{C}_v$ for the critical points, so that

$$C_v = \bigoplus_{\beta\in\mathfrak{C}_v} \mathbb{Z}\Lambda(\beta).$$

Now suppose that (v,u) is a singular 1-simplex, and let $\beta\in\mathfrak{C}_v$, $\alpha\in\mathfrak{C}_u$. The cobordism of pairs, with cylindrical ends attached, namely the pair

$$(\mathbb{R}\times Y, S_{vu}),$$

carries the family of Riemannian metrics G_{vu} (trivial in the case that $v=u$). We choose generic secondary perturbations as in section 3.9, and we write

$$M_{vu}(\beta,\alpha)\to G_{vu}$$

for the corresponding parametrized moduli space.

There is an action of $\mathbb{R}$ on M_{vu} by translations, covering the action of $\mathbb{R}$ on G_{vu} given by

$$(\tau_{i_1},\dots,\tau_{i_d})\mapsto(\tau_{i_1}-t,\dots,\tau_{i_d}-t).$$

(The choice of sign here is so as to match a related convention in [22].) In the special case that $v=u$, the cobordism is a cylinder, and $t\in\mathbb{R}$ acts on M_{vv} by pulling back by the translation $(\tau,y)\mapsto(\tau+t,y)$ of $\mathbb{R}\times Y$. When $v=u$, we exclude the translation-invariant part of $M_{vv}(\alpha,\alpha)$, so $\breve{M}_{vv}$ is the quotient by $\mathbb{R}$ only of the non-constant solutions. When $v\neq u$, the action of $\mathbb{R}$ on G_{vu} is free and we can form the quotient $\breve{G}_{vu}$ considered earlier, so that we have

$$\breve{M}_{vu}(\beta,\alpha)\to\breve{G}_{vu}.$$

We can also choose to normalize by the condition

$$\sum_{a=1}^{d}\tau_{i_a}=0$$

and so regard $\breve{G}_{vu}$ as a subset of G_{vu}.

We have not specified the instanton and monopole numbers here, so each $\breve{M}_{vu}(\beta,\alpha)$ is a union of pieces of different dimensions. We write

$$\breve{M}_{vu}(\beta,\alpha)_d\subset\breve{M}_{vu}(\beta,\alpha)$$

for the union of the d-dimensional components, if any.

As in section 3.9, we may consider a natural completion of the space $\breve{M}_{vu}(\beta,\alpha)$ over the polytope of broken Riemannian metrics $\breve{G}^{+}_{vu}$ (defined at (33) above): we call this completion $\breve{M}^{+}_{vu}(\beta,\alpha)$. To describe it explicitly in the present set-up, we consider all singular simplices $\sigma = (v^0, \dots, v^n)$ with $v^0 = v$ and $v^n = u$ (including the 1-simplex (v, u) itself amongst these). For each such simplex σ and each sequence

$$\boldsymbol{\beta} = (\beta^0, \dots \beta^n)$$

with $\beta^0 = \beta$ and $\beta^n = \alpha$, we consider the product

$$\breve{M}_{\sigma}(\boldsymbol{\beta}) = \breve{M}_{v^0 v^1}(\beta^0, \beta^1) \times \cdots \times \breve{M}_{v^{n-1} v^n}(\beta^{n-1}, \beta^n). \tag{34}$$

As a set, the completion is the union

$$\breve{M}^{+}_{vu}(\beta,\alpha) = \bigcup_{\sigma} \bigcup_{\boldsymbol{\beta}} \breve{M}_{\sigma}(\boldsymbol{\beta}).$$

For each singular simplex σ, there is a map

$$\breve{M}_{\sigma}(\boldsymbol{\beta}) \to \breve{G}_{\sigma'}$$

where σ' is obtained from σ by removing repetitions amongst the vertices. The union of these maps is a map

$$\breve{M}^{+}_{vu}(\beta,\alpha) \to \breve{G}^{+}_{vu}.$$

The space $\breve{M}^{+}_{vu}(\beta,\alpha)$ is given a topology by the same procedure as in the spaces of broken trajectories (see [22] again).

Proposition 6.4. *For a fixed singular* 1*-simplex* (v, u)*, and any* β*,* α *for which the* 1*-dimensional part* $\breve{M}_{vu}(\beta,\alpha)_1$ *is non-empty, the completion* $\breve{M}^{+}_{vu}(\beta,\alpha)_1$ *is a compact 1-manifold with boundary. Its boundary consists of all zero-dimensional products of the form* (34) *with* $n = 2$*,*

$$\breve{M}_{v,v^1}(\beta,\beta^1) \times \breve{M}_{v^1,u}(\beta^1,\alpha),$$

corresponding to singular 2*-simplices* (v, v^1, u)*.*

Proof. This follows from the general discussion in section 3.9. There it is explained that the compactified moduli space over the family has three types of boundary points, described as (a)–(c) on page 46. The three case described there corrrespond to the cases:

(a) the case $v > v^1 > u$ (i.e. a face of $\breve{G}^+_{vu}$ arsing from a non-degenerate 2-simplex);

(b) the case $v^1 = v$, corresponding to a singular 2-simplex;

(c) the case $v^1 = u$, which is also a singular 2-simplex, and which may coincide with the previous case if $v = u$, i.e. if the original 1-simplex is singular.

□

We now consider orientations in the context of the proposition above. For this purpose, let us fix I-orientations μ_{vu} for all cobordism S_{vu} satisfying the conclusion of Lemma 6.1, and let us fix also orientations for all G_{vu} satisfying the conclusion of Lemma 6.2. If we are then given elements of $\Lambda(\alpha$ and $\Lambda(\beta)$, we may orient $M_{vu}(\beta, \alpha)$ using the fiber-first convention as in section 3.9. Having oriented $M_{vu}(\beta, \alpha)$ we then orient $\breve{M}_{vu}(\beta, \alpha)$ as a quotient: giving $\mathbb{R}$ its standard orientation, and putting it first, we write

$$M_{vu}(\beta, \alpha) = \mathbb{R} \times \breve{M}_{vu}(\beta, \alpha) \tag{35}$$

as oriented manifolds. Note that there is another way to orient $\breve{M}_{vu}$ that is different from this one: we could orient $\breve{G}_{vu}$ (as we have done) as the quotient of G_{vu}, and then orient $\breve{M}_{vu}$ as a parametrized moduli space over $\breve{G}_{vu}$. The difference between these two orientations is a sign

$$(-1)^{\dim G_{vu} - 1} = (-1)^{|v-u|_1 - 1}. \tag{36}$$

We shall always use the first orientation (35) for $\breve{M}_{vu}$.

Having so oriented our moduli space $\breve{M}_{vu}(\beta, \alpha)$, we obtain a group homomorphism

$$\breve{m}_{vu} : C_v \to C_u$$

by counting points in zero-dimensional moduli spaces, as in section 3.9.

Lemma 6.5. *For every singular* 1*-simplex* (w, u) *we have*

$$\sum_v (-1)^{|v-u|_1(|w-v|_1 - 1) + 1} \breve{m}_{vu} \circ \breve{m}_{wv} = 0.$$

where the sum is over all v *with* $w \geq v \geq u$.

Proof. There is a degenerate case of this lemma, when $w = u$. In this case, $\breve{m}_{ww}$ is the Floer differential d_w on C_w, and the lemma states that $-d_w^2 = 0$. For the non-degenerate case, where $w > u$, the sum involves two special terms, namely the terms where $v = w$ and $v = u$. Extracting these terms separately, we can recast the formula as

$$\Big(\sum_{v \neq w,u} (-1)^{|v-u|_1(|w-v|_1-1)+1} \breve{m}_{vu} \circ \breve{m}_{wv} \Big) \\ + (-1)^{|w-u|_1+1} \breve{m}_{wu} \circ d_w - d_u \circ \breve{m}_{wu} = 0,$$

or equivalently

$$\Big(\sum_{v \neq w,u} (-1)^{(\dim \breve{G}_{vu}+1)\dim \breve{G}_{wv}+1} \breve{m}_{vu} \circ \breve{m}_{wv} \Big) \\ + (-1)^{\dim \breve{G}_{wu}} \breve{m}_{wu} \circ d_w - d_u \circ \breve{m}_{wu} = 0. \quad (37)$$

This formula is simply a special case of the general chain-homotopy formula (23), but to verify this we need to compare the signs here to those in (23). To make this comparison, let us first rewrite the last formula in terms of the homomorphism $\bar{m}_{vu}$, defined in the same way as $\breve{m}_{vu}$ but using the orientation convention of section 3.9, so that

$$\breve{m}_{vu} = (-1)^{\dim \breve{G}_{vu}} \bar{m}_{vu}$$

as in (36). After multiplying throughout by $(-1)^{\dim \breve{G}_{wu}}$, the formula becomes

$$\Big(\sum_{v \neq w,u} (-1)^{(\dim \breve{G}_{vu}+1)\dim \breve{G}_{wv}} \bar{m}_{vu} \circ \bar{m}_{wv} \Big) \\ + (-1)^{\dim \breve{G}_{wu}} \bar{m}_{wu} \circ d_w - d_u \circ \bar{m}_{wu} = 0.$$

In this form, the formula resembles the formula (23), with the only difference being the extra $+1$ in the first factor of the first exponent. This extra term is accounted for by Lemma 6.3 and is present because the product orientation on $\breve{G}_{wv} \times \breve{G}_{vu}$ is not equal to its boundary orientation as a face of $\breve{G}_{wu}$. $\square$

In order to define away some of the signs, we observe that the formula in the lemma above can be written as

$$(-1)^{\sum w_i + \mathsf{s}(w,u)} \sum_v (-1)^{\mathsf{s}(v,u)+\mathsf{s}(w,v)} \breve{m}_{vu} \circ \breve{m}_{wv} = 0,$$

where s is given by the formula

$$\mathsf{s}(v,u) = \frac{1}{2}|v-u|(|v-u|-1) + \sum v_i. \tag{38}$$

With our choices of homology orientations μ_{vu} etc. still understood, we make the following definition:

Definition 6.6. In the above setting, we define homomorphisms

$$f_{vu} : C_v \to C_u$$

by the formula

$$f_{vu} = (-1)^{\mathsf{s}(v,u)} \breve{m}_{vu}$$

for all singular 1-simplices (v,u). ◊

Note that in the case $v = u$ we have

$$f_{vv} = (-1)^{\sum_i v_i} d_v.$$

With these built-in sign adjustments, the previous lemma takes the following form.

Proposition 6.7. *For any singular* 1*-simplex* (w,u)*, we have*

$$\sum_v f_{vu} f_{wv} = 0,$$

where the sum is over all v *with* $w \geq v \geq u$*.* □

For each singular 1-simplex (v,u), we now introduce

$$\mathbf{C}[vu] = \bigoplus_{v \geq v' \geq u} C_{v'}.$$

(Here v' runs over the unit cube with extreme vertices v and u.) We similarly define

$$\mathbf{F}[vu] = \bigoplus_{v \geq v' \geq u' \geq u} f_{v'u'},$$

so we have

$$\mathbf{F}[vu] : \mathbf{C}[vu] \to \mathbf{C}[vu].$$

The previous proposition then becomes the statement

$$\mathbf{F}[vu]^2 = 0,$$

so that $(\mathbf{C}[vu], \mathbf{F}[vu])$ is a complex. The basic case here is to take $v = (1, \ldots, 1)$ and $u = (0, \ldots, 0)$, in which case $\mathbf{C}[vu]$ becomes

$$\mathbf{C} := \bigoplus_{v' \in \{0,1\}^N} C_{v'}, \tag{39}$$

with one summand for each vertex of the unit N-cube. The corresponding differential $\mathbf{F}$ has a summand $f_{v'v'} : C_{v'} \to C_{v'}$ for each vertex, together with summands $f_{v'u'}$ for each $v' > u'$. The general $\mathbf{C}[vu]$ is also a sum of terms indexed by v' running over the vertices of a cube of side-length 1, though the dimension of the cube is $|v - u|_1$ in general, which may be less than N.

The remainder of this subsection and the following one are devoted to proving the following theorem, which states that the homology of the cube $\mathbf{C}[vu]$ coincides with the homology of a single C_w for appropriate w:

Theorem 6.8. *Let (v, u) be any 1-simplex, and let $w = 2v - u$. Then there is a chain map*

$$(C_w, f_{ww}) \to (\mathbf{C}[vu], \mathbf{F}[vu])$$

inducing an isomorphism in homology. Thus in the case that $v = (1, \ldots, 1)$ and $u = (0, \ldots, 0)$, the homology of $(\mathbf{C}, \mathbf{F})$, where $\mathbf{C}$ is as in (39), is isomorphic to the homology of (C_w, f_w), where $w = (2, \ldots, 2)$.

To amplify the statement of the theorem a little, we can point out first that in the case $v = u$, the result is a tautology, for both of the chain complexes then reduce to (C_v, f_{vv}). Next, we can look at the case $|v - u|_1 = 1$. In this case, K_v and K_u differ only inside one of the N balls, and we may as well take $N = 1$. In the notation of Figure 6, we can identify K_v and K_u with K_1 and K_0, in which case K_w is K_2. The complex $\mathbf{C}[vu]$ is the sum of the chain complexes for the two links K_1 and K_0,

$$C_1 \oplus C_0$$

and

$$\mathbf{F}[vu] = \begin{pmatrix} -d_1 & 0 \\ f_{10} & d_0 \end{pmatrix}$$

where f_{10} is minus the chain map induced by S_{10}. Thus $(\mathbf{C}[vu], \mathbf{F}[vu])$ is, up to sign, the mapping cone of the chain map induced by the cobordism. The theorem then expresses the fact that the homologies of C_2, C_1 and C_0 are related by a long exact sequence, in which one of the maps arises from the cobordism S_{10}. The proof of the theorem will also show that the remaining maps in the long exact sequence can be taken to be the ones arising from S_{21}

and S_{02}. This is the long exact sequence of the unoriented skein relations, mentioned in the introduction. For larger values of $|v-u|_1$, the differential $\mathbf{F}[vu]$ still has a lower triangular form, reflecting the fact that there is a filtration of the cube (by the sum of the coordinates) that is preserved by the differential.

Corollary 6.9. *In the situation of the theorem above, there is a spectral sequence whose E_1 term is*

$$\bigoplus_{v'\in\{0,1\}^N} I^\omega_*(Y, K_{v'})$$

and which abuts to the instanton Floer homology $I^\omega_*(Y, K_w)$, for $w = (2, \ldots, 2)$. □

The signs of the maps in this spectral sequence are determined by choices of I-orientations for the cobordisms S_{vu} and orientations of the families of metrics G_{vu}, subject to the compatibility conditions imposed by Lemmas 6.1 and 6.2. These compatibility conditions still leave some freedom. For both lemmas, the compatibility conditions mean that the orientations for all 1-simplices (v, u) are determined by the orientations of S_{vu} and G_{vu} for the *edges*, i.e. the 1-simplices (v, u) for which $|v - u|_1 = 1$. These are the simplices that will contribute to the differential d_1 in the spectral sequence above. For these, the space $\breve{G}_{vu}$ is a point, so an orientation is determined by a sign $(-1)^{\delta(v,u)}$ for each v, u. The condition of Lemma 6.2 is equivalent to requiring the following: for every 2-dimensional face of the cube, with diagonally opposite vertices $w > u$ and intermediate vertices v and v', we need

$$\delta(w, v)\delta(v, u) = 1 + \delta(w, v')\delta(v', u) \pmod 2. \tag{40}$$

We can achieve this condition by an explicit choice of sign, such as

$$\delta(v, u) = \sum_{i=0}^{i_0-1} v_i, \tag{41}$$

where i_0 is the unique index at which v_i and u_i differ.

Corollary 6.10. *Let I-orientations be chosen for the cobordisms $S_{v'u'}$ so that the conditions of Lemma 6.1 hold. For each edge (v', u') of the cube, let $I^\omega(S_{v'u'})$ be the map $I^\omega(Y, K_{v'}) \to I^\omega(Y, K_{u'})$ induced by $S_{v'u'}$ with this I-orientation. Then the spectral sequence in the previous corollary can be set up so that the differential d_1 is the sum of the maps*

$$(-1)^{\tilde{\delta}(v',u')} I^\omega(S_{v'u'})$$

over all edges of the cube, where

$$\tilde{\delta}(v', u') = \sum_{i=i_0}^{N} v'_i$$

and i_0 is the index at which v' and u' differ.

Proof. The difference between δ and the $\tilde{\delta}$ that appears here is $\sum_i v'_i \bmod 2$, which is $\mathsf{s}(v', u') \bmod 2$ in the case of an edge (v', u'). □

Theorem 6.8 will be proved by an inductive argument. The special case described at the end of the theorem is equivalent to the general case, so we may as well take

$$\begin{aligned} w &= (2, \dots, 2) \\ v &= (1, \dots, 1) \\ u &= (0, \dots, 0). \end{aligned}$$

Thus v and u span an N-cube. We again write $(\mathbf{C}, \mathbf{F})$ for $(\mathbf{C}[uv], \mathbf{F}[uv])$ in this context. For each $i \in \mathbb{Z}$, let us set

$$\mathbf{C}_i = \bigoplus_{v' \in \{1,0\}^{N-1}} C_{v',i}.$$

Each $\mathbf{C}_i$ can be described as $\mathbf{C}[v'u']$ for some v', u' with $|v' - u'|_1 = N - 1$. We have

$$\mathbf{C} = \mathbf{C}_1 \oplus \mathbf{C}_0,$$

generalizing the case $N = 1$ considered above, and we can similarly decompose $\mathbf{F}$ in block form as

$$\mathbf{F} = \begin{pmatrix} \mathbf{F}_{11} & 0 \\ \mathbf{F}_{10} & \mathbf{F}_{00} \end{pmatrix},$$

where

$$\mathbf{F}_{ij} = \bigoplus_{v', u' \in \{1,0\}^{N-1}} f_{(v'i)(u'j)}.$$

To prove the theorem, we will establish:

Proposition 6.11. *There is a chain map*

$$(\mathbf{C}_2, \mathbf{F}_{22}) \to (\mathbf{C}, \mathbf{F})$$

inducing isomorphisms in homology.

This proposition expresses an isomorphism between the homologies of two cubes, one of dimension $N-1$, the other of dimension N. Theorem 6.8 for the given w, v and u is an immediate consequence of N applications of this proposition. Just as in the case $N=1$, the proposition will be proved while establishing that there is a long exact sequence in homology arising from the anti-chain maps,

$$\cdots \longrightarrow (\mathbf{C}_3, \mathbf{F}_{33}) \xrightarrow{\mathbf{F}_{32}} (\mathbf{C}_2, \mathbf{F}_{22}) \xrightarrow{\mathbf{F}_{21}} (\mathbf{C}_1, \mathbf{F}_{11}) \xrightarrow{\mathbf{F}_{10}} (\mathbf{C}_0, \mathbf{F}_{00}) \longrightarrow \cdots$$

(In the above sequence, the chain groups and anti-chain maps are periodic mod 3, up to sign.)

7 Proof of Proposition 6.11

7.1 The algebraic setup

The proof is based on an algebraic lemma which appears (in a mod 2 version) as Lemma 4.2 in [29]. We omit the proof of the lemma:

Lemma 7.1 ([29, Lemma 4.2]). *Suppose that for each $i \in \mathbb{Z}$ we have a complex (C_i, d_i) and anti-chain maps*

$$f_i : C_i \to C_{i-1}.$$

Suppose that the composite chain map $f_{i-1} \circ f_i$ is chain-homotopic to 0 via a chain-homotopy j_i, in that

$$d_{i-1} j_i + j_i d_i + f_{i-1} f_i = 0$$

for all i. Suppose furthermore that for all i, the map

$$j_{i-1} f_i + f_{i-2} j_i : C_i \to C_{i-3}$$

(which is a chain map under the hypotheses so far) induces an isomorphism in homology. Then the induced maps in homology,

$$(f_i)_* : H_*(C_i, d_i) \to H_*(C_{i-1}, d_{i-1})$$

form an exact sequence; and for each i the anti-chain map

$$\Phi : s \mapsto (f_i s, j_i s)$$
$$\Phi : C_i \to \mathrm{Cone}(f_i)$$

induces isomorphisms in homology. Here Cone(f_i) *denotes* $C_i \oplus C_{i-1}$ *equipped with the differential*

$$\begin{pmatrix} d_i & 0 \\ f_i & d_{i-1} \end{pmatrix}.$$

□

Given the lemma, our first task is to construct (in the case $i = 2$, for example) a map

$$\mathbf{J}_{20} : \mathbf{C}_2 \to \mathbf{C}_0$$

satisfying

$$\mathbf{F}_{00}\mathbf{J}_{20} + \mathbf{J}_{20}\mathbf{F}_{22} + \mathbf{F}_{10}\mathbf{F}_{21} = 0. \tag{42}$$

After constructing these maps, we will then need to show that the maps such as

$$\mathbf{F}_{10}\mathbf{J}_{31} + \mathbf{J}_{20}\mathbf{F}_{32} : \mathbf{C}_3 \to \mathbf{C}_0$$

(which is a chain map) induce isomorphisms in homology. This second step will be achieved by constructing a chain-homotopy

$$\mathbf{K}_{30} : \mathbf{C}_3 \to \mathbf{C}_0$$

with

$$\mathbf{F}_{00}\mathbf{K}_{30} + \mathbf{K}_{30}\mathbf{F}_{33} + \mathbf{F}_{10}\mathbf{J}_{31} + \mathbf{J}_{20}\mathbf{F}_{32} + \tilde{\mathbf{Id}} = 0, \tag{43}$$

where $\tilde{\mathbf{Id}}$ is a chain-map that is chain-homotopic to ± 1. The construction of $\mathbf{J}$ and $\mathbf{K}$ and the verification of the chain-homotopy formulae (42) and (43) occupy the remaining two subsections of this section of the paper.

7.2 Construction of J

We will construct $\mathbf{J}_{i,i-2}$ for all i so that (in the case $i = 2$, for example) the relation (42) holds. Recalling the definition of $\mathbf{C}_i$, we see that we should write

$$\mathbf{J}_{20} = \sum j_{(v'2)(u'0)} \tag{44}$$

where the sum is over all $v' \geq u'$ in $\{0,1\}^{N-1}$. The desired relation then expands as the condition that, for all w', u' in $\{0,1\}^{N-1}$,

$$\sum_{v'} \Big(f_{(v'0)(u'0)} j_{(w'2)(v'0)} + j_{(v'2)(u'0)} f_{(w'2)(v'2)} + f_{(v'1)(u'0)} f_{(w'2)(v'1)} \Big) = 0. \tag{45}$$

Our task now is to define j_{vu} for $v = v'2$ and $u = u'0$ in $\mathbb{Z}^N$, where (v', u') some singular 1-simplex in the triangulation of R^{N-1} of $\mathbb{R}^{N-1}$.

We have previously considered I-orientations for S_{vu} and orientations of families of metrics G_{vu} in the case that (v,u) is a singular 1-simplex. We now extend our constructions to the case of an arbitrary pair (v,u) with $v \geq u$. (So we now allow $|v-u|_\infty$ to be larger than 1. At present we are most interested in the case $|v-u|_\infty = 2$; and in the next section, 3 will be relevant.) We still have natural cobordisms S_{vu} when $|v-u|_1 > 1$, obtained by concatenating the cobordisms we used previously. So for example, when $N = 1$, the cobordism S_{20} is the composite of the cobordisms S_{21} and S_{10}.

For the I-orientations, we can begin by choosing I-orientations as before for 1-simplices (v,u), so that the conditions of Lemma 6.1 hold. Then we simply extend to all pairs $v \geq u$ so that the consistency condition

$$\mu_{wu} = \mu_{vu} \circ \mu_{wv}$$

holds for all $w \geq v \geq u$.

For arbitrary $v \geq u$, the cobordism $(\mathbb{R} \times Y, S_{vu})$ also carries a family of metrics G'_{vu} of dimension $n = |v-u|_1$. We define G'_{vu} first in the case $N = 1$. In this case, we can regard S_{vu} as the composite of n cobordisms, where $n = |v-u|_1$ and each cobordism is a surface on which the t coordinate as a single critical point. Unlike the previous setup, these critical points cannot be re-ordered, as they all lie in the same copy of $\mathbb{R} \times B_1$ (where B_1 is a 3-ball), rather than in distinct copies $\mathbb{R} \times B_i$. As an appropriate parameter space in this case, we define G'_{vu} to be

$$G'_{vu} = \{\, (\tau_1, \dots \tau_n) \in \mathbb{R}^n \mid \tau_{m+1} \geq 1 + \tau_m,\ \forall m < n \,\}, \tag{46}$$

and construct the metrics so that τ_m is the t coordinate of the critical point in the m'th cobordism (much as we did in the construction of G_{vu} earlier, on page 82). For larger N, we construct G'_{vu} as a product of $G'_{v_i u_i}$ over all $i = 1, \dots, N$. In the case that $|v-u|_\infty = 1$, the space G'_{vu} coincides with $G_{vu} = \mathbb{R}^n$ as defined before; while if $|v_N - u_N| = 2$ and $|v_j - u_j| \leq 1$ for $j \leq N-1$, then G'_{vu} is a half-space. Whenever $w \geq v \geq u$, we have self-evident maps

$$G'_{wu} \to G'_{wv} \times G'_{vu}.$$

These maps are either surjective, or have image equal to the intersection of the codomain with a product of half-spaces. Following Lemma 6.2, we can choose orientations for all the G'_{vu} so these maps are always orientation-preserving.

Next we examine the topology of the cobordism S_{vu} in the case that $N = 1$ and $|v-u|_\infty = 2$. For the following description, we revert to considering S_{vu} as a compact surface in a product $I \times Y$, rather than a surface with cylindrical ends in $\mathbb{R} \times Y$.

Lemma 7.2. *In the situation depicted in Figures 6 and 7, the composite cobordism $S_{2,0} = S_{1,0} \circ S_{2,1}$ from K_2 to K_0 in $I \times Y$ has the form*

$$(I \times Y, V_{2,0}) \# (S^4, \mathbb{RP}^2)$$

where the $\mathbb{RP}^2$ is standardly embedded in S^4 with self-intersection $+2$, as described in section 2.7.

Remark. The cobordism $V_{2,0}$ that appears in the above lemma is diffeomorphic to $S_{3,2}$, viewed as a cobordism from K_2 to K_3 by reversing the orientation of $I \times Y$.

Proof of the Lemma. Arrange the composite cobordism $S_{2,0}$ so that the t coordinate runs from 0 to 1 across $S_{2,1}$ and from 1 to 2 across $S_{1,0}$. The projection of $S_{2,1}$ to Y meets the ball B^3 in the twisted rectangle $T = T_{2,1}$ depicted in Figure 7, while the projection of $S_{1,0}$ similarly meets B^3 in a twisted rectangle $T_{1,0}$. The intersection $T_{2,1} \cap T_{1,0}$ in B^3 is a closed arc δ, joining two points of K_1. The preimages of δ in $S_{2,1}$ and $S_{1,0}$ are two arcs in $S_{2,0}$ whose union is a simple closed curve

$$\gamma \subset S_{2,0}.$$

On γ, the t coordinate takes values in $[1/2, 3/2]$. A regular neighborhood of $[0,2] \times \delta$ in $[0,2] \times Y$ is a 4-ball meeting $S_{2,0}$ in a Möbius band: the band is the neighborhood of γ in $S_{2,0}$.

This Möbius band in the 4-ball can be seen as arising from pushing into the ball an unknotted Möbius band in the 3-sphere. The Möbius band M in the 3-sphere is the union of three pieces:

(a) a neighborhood of $\{0\} \times \delta$ in $\{0\} \times T_{2,1}$;

(b) a neighborhood of $\{2\} \times \delta$ in $\{2\} \times T_{1,0}$;

(c) the pair of rectangles $[0,2] \times \epsilon$, where ϵ is a pair of arcs, one in each component of $K_1 \cap B^3$.

This Möbius band M possesses a left-handed half-twist. The half-twist is the result of two quarter-turns, one in each of the first two pieces of M in the list above. The signs of the quarter-turns can be seen in Figure 9: a neighborhood of δ in $T_{2,1}$ has a *right*-hand quarter turn for the standard orientation of B^3, but this 3-ball occurs in the boundary of the 4-ball with its opposite orientation; and a neighborhood of δ in $T_{1,0}$ has a *left*-hand quarter turn for the standard orientation of B^3, and this 3-ball occurs with

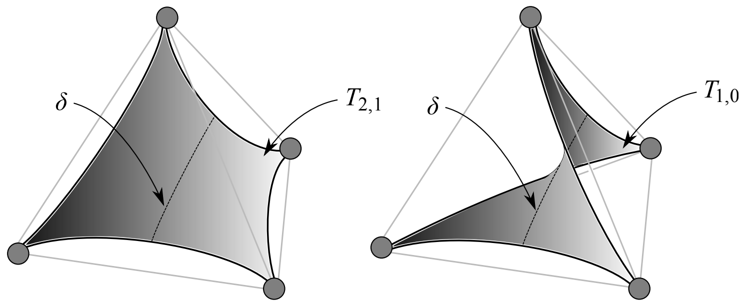

Figure 9: The arc δ as the intersection of $T_{2,1}$ and $T_{1,0}$.

its positive orientation in the boundary of the 4-ball. Thus the Möbius band in the boundary of the 4-ball has a left-hand half-twist resulting from two left-handed quarter-turns.

The $\mathbb{RP}^2$ obtained from Möbius band with a left-handed half-twist is the standard $\mathbb{RP}^2$ with self-intersection $+2$. $\square$

The fact that the composite cobordism S_{20} from (Y, K_2) to (Y, K_0) splits off a summand $(S^4, \mathbb{RP}^2)$ (as stated in the lemma above) implies, by standard stretching arguments, that this composite cobordism induces the zero map in homology:

$$(f_{10} \circ f_{21})_* = 0 : I^\omega(Y, K_2) \to I^\omega(Y, K_0).$$

This is essentially the same point as the vanishing theorem for the Donaldson invariants of connected sums. It is important here that the summand $(S^4, \mathbb{RP}^2)$ carries no reducible solutions, which might live in moduli spaces for which the index of $\mathcal{D}$ is negative: see the examples of moduli spaces in Proposition 2.10. Although it is zero at the level of homology, at the *chain level*, the map induced by the composite may be non-zero. The one-parameter family of metrics involved in the stretching provides a chain homotopy, showing that the map is chain-homotopic to zero. This is what will be used to construct the chain-homotopy $\mathbf{J}_{20}$ in Definition 7.3 below. But first, we must make the family of metrics explicit and extend our notation to the case of more than one ball in Y.

Staying for a moment with the case of one ball, we have already set up a family of metrics G'_{20}, which in this case is a 2-dimensional half-space:

$$G'_{20} = \{ (\tau_1, \tau_2) \mid \tau_2 \geq 1 + \tau_1 \}.$$

We extend this family of metrics to a family

$$G_{20} = G'_{20} \cup G''_{20} \tag{47}$$

as follows. The boundary of G'_{20} consists of the family of metrics with $\tau_2 = \tau_1 + 1$, all of which are isometric to each other, by translation of the coordinates. Fixing any one of these, say at $\tau_1 = 0$, we construct a family of metrics parametrized by the negative half-line $\mathbb{R}^-$, by stretching along the sphere S^3 which splits off the summand $(S^4, \mathbb{RP}^2)$ in Lemma 7.2. (This family of metrics can be completed to a family parametrized by $\mathbb{R}^- \cup \{-\infty\}$, where the added point is a broken metric, cut along this S^3.) Putting back the translation parameter, we obtain our family of metrics G''_{20} parametrized by $\mathbb{R}^- \times \mathbb{R}$. The space G_{20} is the union of these two half-spaces, along their common boundary.

Suppose now that N is arbitrary, and that $v_N - u_N = 2$ and $v_j - u_j = 0$ or 1 for $j < N$. The space of metrics G'_{vu} is a half-space: it is a product

$$G'_{vu} = \mathbb{R}^{m-1} \times G'_{20}$$

where m is the number of coordinates in which v and u differ, and G'_{20} is a 2-dimensional half-space as above. The coordinates on $\mathbb{R}^{m-1}$ are the locations τ of the critical points in the balls B_i corresponding to coordinates $i < N$ where v and u differ. We extend this family of metrics to a family

$$\begin{aligned} G_{vu} &= \mathbb{R}^{m-1} \times (G'_{20} \cup G''_{20}) \\ &= G'_{vu} \cup G''_{vu} \end{aligned}$$

where G''_{20} is as before. We again use the notation $\breve{G}_{vu}$ for the quotient by translations:

$$\breve{G}_{vu} = G_{vu}/\mathbb{R}.$$

Let us consider the natural compactification $\breve{G}^+_{wu}$ of $\breve{G}_{wu}$, where $w = w'2$, $u = u'0$ and $|w' - u'| \leq 1$. This is a family of broken Riemannian metrics whose codimension-1 faces are as follows.

(a) First, there are the families of broken metrics which are cut along (Y, K_v) where $w > v > u$. This face is parametrized by $\breve{G}_{wv} \times \breve{G}_{vu}$. These faces we can classify further into the cases

 (i) the case $v_N = 0$, in which case the first factor $\breve{G}_{wv}$ has the form $\breve{G}'_{wv} \cup \breve{G}''_{wv}$, where $\breve{G}''_{wv}$ involves stretching across the S^3;

 (ii) the similar case $v_N = 2$, where the second factor has the form $\breve{G}'_{vu} \cup \breve{G}''_{vu}$;

(iii) the case $v_N = 1$, in which case $\breve{G}_{wv}$ and $\breve{G}_{vu}$ are both the simpler families described in the previous subsection leading to the construction of the maps f_{vu} etc.

(b) Second, there is the family of broken metrics which are cut along the S^3.

Now let $\beta \in \mathfrak{C}_v$ and $\alpha \in \mathfrak{C}_u$ be critical points, corresponding to generators of the complexes C_v and C_u respectively. The family of metrics G_{vu} gives rise to a parametrized moduli space

$$M_{vu}(\beta, \alpha) \to G_{vu}.$$

Dividing out by the translations, we also obtain

$$\breve{M}_{vu}(\beta, \alpha) \to \breve{G}_{vu}.$$

We have already oriented the subset $G'_{vu} \subset G_{vu}$, so we have chosen orientation for G_{vu}. As before, we orient $M_{vu}(\beta, \alpha)$ using our chosen I-orientations and a fiber-first convention, and we orient $\breve{M}_{vu}(\beta, \alpha)$ as the quotient of $M_{vu}(\beta, \alpha)$ with the $\mathbb{R}$ factor first. The zero-dimensional part

$$\breve{M}_{vu}(\beta, \alpha)_0 \subset \breve{M}_{vu}(\beta, \alpha)$$

(if any) is a finite set of oriented points as usual, and we define j_{vu} by counting these points, with an overall correction factor for the sign:

Definition 7.3. Given $v \geq u$ in $\mathbb{Z}^N$ with $v_N - u_N = 2$ and $v_j - u_j \leq 1$ for $j < N$, we define

$$j_{vu} : C_v \to C_u$$

by declaring the matrix entry from β to α to be the signed count of the points in the zero-dimensional moduli space $\breve{M}_{vu}(\beta, \alpha)_0$ (if any), adjusted by the overall sign $(-1)^{\mathsf{s}(v,u)}$, where $\mathsf{s}(v, u)$ is again defined by the formula (38). ◊

Having defined j_{vu} in this way, we can now construct $\mathbf{J}_{20} : \mathbf{C}_2 \to \mathbf{C}_0$ in terms of j_{vu} by the formula (44). We must now prove the chain-homotopy formula (42), or equivalently the formula (45), which we can equivalently write as

$$\sum_{\{v | v_N = 0\}} f_{vu} j_{wv} + \sum_{\{v | v_N = 2\}} j_{vu} f_{wv} + \sum_{\{v | v_N = 1\}} f_{vu} f_{wv} = 0. \qquad (48)$$

As usual, the proof that this expression is zero is to interpret the matrix entry of this map, from γ to α, as the number of boundary points of an oriented 1-manifold, in this case the manifold $\breve{M}_{wu}(\gamma,\alpha)$. This is in essence an example of the chain-homotopy formula (23), resulting from counting ends of one-dimensional moduli-spaces $\breve{M}^{+}_{wu}(\gamma,\alpha)_1$ over $\breve{G}^{+}_{wv}$. The three types of terms in the above formulae capture the three types of boundary faces (a) above, together with the terms of the form "$\partial \circ m_G \pm m_G \circ \partial$" in (23). This is just the same set-up as the proof that $\mathbf{F}_{10} \circ \mathbf{F}_{10} = 0$ in the previous section, and our signs are once again arranged so that all terms contribute with positive sign.

The only remaining issue for the proof of (48) is the question of why there is no additional term in this formula to account for a contribution from the face (b) of $\breve{G}^{+}_{wu}$. This face does not fall into the general analysis, because the cut (S^3, S^1) does not satisfy the non-integral condition. (We have $w = 0$ on this cut.) Analyzing the contribution from this type of boundary component follows the standard approach to a connected sum – in this case, a connected sum with the pair $(S^4, \mathbb{RP}^2)$ along a standard (S^3, S^1). There is no contribution from this type of boundary component, however, by the usual dimension-counting argument for connected sums, because all solutions on $(S^4, \mathbb{RP}^2)$ are irreducible and the unique critical point for (S^3, S^1) is reducible.

7.3 Construction of K

We turn to the construction of $\mathbf{K}_{30}$ and the proof the formula (43). We start with a look at the topology of the composite cobordism

$$S_{30} = S_{10} \circ S_{21} \circ S_{32}$$

from K_3 to K_0 in the case $N = 1$. Our discussion is very closely modeled on the exposition of [17, section 5.2].

Arrange the t coordinate on the composite cobordism S_{30} so that t runs from $3 - i$ to $3 - j$ across S_{ij}. In the previous subsection we exhibited a Möbius band (called M there) inside S_{20}. Let us now call this Möbius band M_{20}. It is the intersection of S_{20} with a 4-ball arising as the regular neighborhood of $[1, 3] \times \delta$. Just as we renamed M, let us now write δ_{20} for the arc δ. There is a similar Möbius band M_{31} in S_{31}, arising as the intersection of S_{31} with the regular neighborhood of $[0, 2] \times \delta_{31}$.

The arcs δ_{20} and δ_{31} in B^3 both lie on the surface T_{21}, where they meet at a single point at the center of the tetrahedron. The two Möbius bands M_{31}

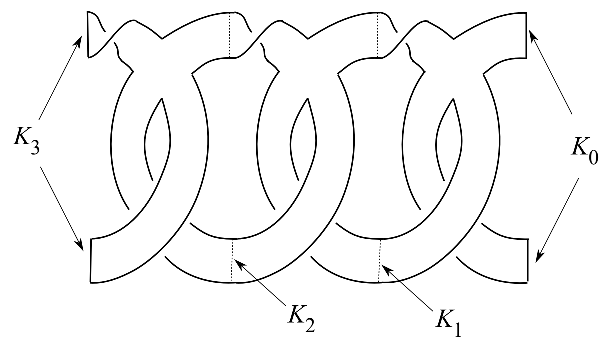

Figure 10: Two intersecting Möbius bands, M_{31} and M_{20} inside S_{30}.

and M_{20} meet T_{21} in regular neighborhoods of these arcs; so the intersection $M_{31} \cap M_{20}$ is a neighborhood in T_{21} of this point. (See Figure 10.) The union

$$M_{30} := M_{31} \cup M_{20}$$

has the topology of a twice-punctured $\mathbb{RP}^2$ and it sits in a 4-ball B_{30} obtained as regular neighborhood of the union of the previous two balls, B_{31} and B_{20}. The 3-sphere $\mathbb{S}_{30}$ which forms the boundary of B_{30} meets M_{30} in two unknotted, unlinked circles. The following lemma helps to clarify the topology of the cobordism S_{30}.

Lemma 7.4. *If we remove (B_{30}, M_{30}) from the pair $([0,3] \times Y, S_{30})$ and replace it with (B_{30}, Δ), where Δ is a union of two standard disks in the 4-ball, then the resulting cobordism $\bar{S}$ from K_3 to $K_0 = K_3$ is the trivial cylindrical cobordism in $[0,3] \times Y$.*

Proof. This is clear. □

Altogether, we can identify five separating 3-manifolds in $[0,3] \times Y$, namely the three 3-spheres $\mathbb{S}_{30}$, $\mathbb{S}_{31}$ and $\mathbb{S}_{20}$ obtained as the boundaries of the three balls, and the two copies of Y,

$$Y_2 = \{1\} \times Y$$
$$Y_1 = \{2\} \times Y$$

which contain the links K_2 and K_1. Just as in [17, section 5.2], each of these five 3-manifolds intersects two of the others transversely, in an arrangement indicated schematically in Figure 11, and each non-empty intersection is a 2-sphere.

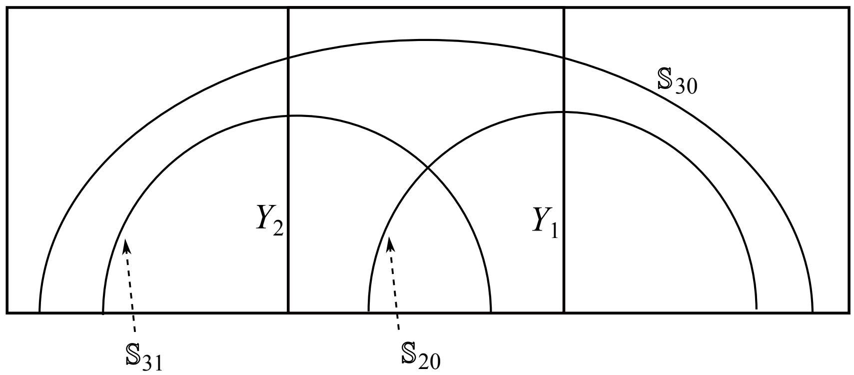

Figure 11: The five 3-manifolds, Y_2, Y_1, $\mathbb{S}_{30}$, $\mathbb{S}_{31}$ and $\mathbb{S}_{20}$ in the composite cobordism $([0,3]\times Y, S_{30})$.

We can form a family of Riemannian metrics $\breve{G}_{30}$ on this cobordism whose compactification $\breve{G}^+_{30}$ is a 2-dimensional manifold with corners – in fact, a pentagon – parametrizing a family of broken Riemannian metrics. The five edges of this pentagon correspond to broken metrics for which the cut is a single one of the five separating 3-manifolds,

$$\mathbb{S} \in \{\mathbb{S}_{30}, \mathbb{S}_{31}, \mathbb{S}_{20}, Y_2, Y_1\}.$$

We denote the corresponding face by

$$Q(\mathbb{S}) \subset \breve{G}^+_{30}.$$

The five corners of the pentagon correspond to broken metrics where the cut has two connected components, $\mathbb{S}\cup\mathbb{S}'$, where

$$\{\mathbb{S}, \mathbb{S}'\} \subset \{\mathbb{S}_{30}, \mathbb{S}_{31}, \mathbb{S}_{20}, Y_2, Y_1\},$$

is a pair of 3-manifolds that do *not* intersect. (There are exactly five such pairs.) In the neighborhood of each edge and each corner, the family of metrics has the model form described in section 3.9. As special cases, we have

$$Q(Y_1) = \breve{G}_{31} \times \breve{G}_{10}$$
$$Q(Y_2) = \breve{G}_{32} \times \breve{G}_{20}.$$

We can, if we wish, regard $\breve{G}_{30}$ as the quotient by translations of a larger family G_{30} of dimension 3. We can regard the previously-defined family

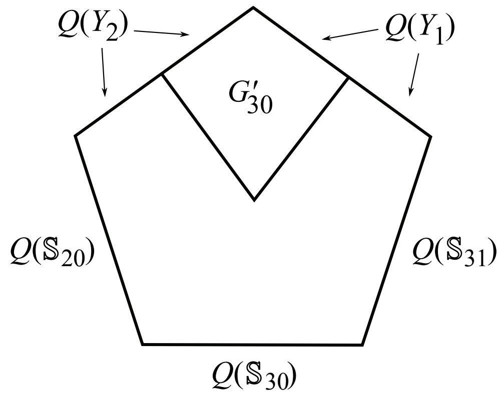

Figure 12: The family of metrics $\breve{G}_{30}$ containing the image of family G'_{30}.

$G'_{30} \cong \mathbb{R}^+ \times \mathbb{R}^+$ as a subset of G_{30} in such a way that its image in $\breve{G}_{30}$ is the indicated quadrilateral in Figure 12.

Turning now to the case of arbitrary N, we proceed as we did in the previous subsection. That is, we suppose that we have w and u in $\mathbb{Z}^N$, with $w_N - u_N = 3$ and $w_j - u_j = 0$ or 1 for $j < N$. The space of metrics G'_{wu} is the product of $\mathbb{R}^{m-1}$ with a 2-dimensional quadrant, where m is again the number of coordinates in which w and u differ:

$$G'_{wu} = \mathbb{R}^{m-1} \times G'_{30}.$$

If we write $w = w'3$ and $u = u'0$ with $w', u' \in \mathbb{Z}^{N-1}$, then we can identify the $\mathbb{R}^{m-1}$ here with $G_{w'u'}$. We extend G'_{wu} to a larger family

$$\begin{aligned} G_{wu} &= \mathbb{R}^{m-1} \times G_{30} \\ &= G_{w'u'} \times G_{30} \end{aligned}$$

where $G_{30} \supset G'_{30}$ is the interior of the pentagon just described, and we set $\breve{G}_{wu} = G_{wu}/\mathbb{R}$. By suitably normalizing the coordinates, we can choose to identify

$$\breve{G}_{wu} = G_{w'u'} \times \breve{G}_{30}.$$

We can complete $\breve{G}_{wu}$ to a family of broken Riemannian metrics $\breve{G}^+_{wu}$ whose codimension-1 faces are as follows.

(a) First, the faces of the form $\breve{G}_{wv} \times \breve{G}_{vu}$ with $w > v > u$, parametrizing metrics broken at (Y, K_v). These we further subdivide as:

(i) the cases with $v_N = w_N$;

(ii) the cases with $v_N = u_N$;

(iii) the cases with $v_N = w_N - 1$ (these correspond to the edge $Q(Y_2)$ of the pentagon, in the case $N = 1$);

(iv) the cases with $v_N = w_N - 2$ (these correspond to the edge $Q(Y_1)$ of the pentagon, in the case $N = 1$).

(b) Second, the faces of the form $G_{w'u'} \times Q(\mathbb{S})$ for $\mathbb{S} = \mathbb{S}_{31}, \mathbb{S}_{20}$ or $\mathbb{S}_{30}$.

Our chosen orientation of $\breve{G}'_{wu}$ determines an orientation for the larger space $\breve{G}_{wu}$. Over the compactification $\breve{G}^+_{wu}$ we have moduli spaces $\breve{M}^+_{wu}(\alpha, \beta)$ as usual. Mimicking Definition 7.3, we define the components of $\mathbf{K}$ as follows:

Definition 7.5. Given $v \geq u$ in $\mathbb{Z}^N$ with $v_N - u_N = 3$ and $v_j - u_j \leq 1$ for $j < N$, we define

$$k_{vu} : C_v \to C_u$$

by declaring the matrix entry from β to α to be the signed count of the points in the zero-dimensional moduli space $\breve{M}_{vu}(\beta, \alpha)_0$ (if any), adjusted by the overall sign $(-1)^{\mathsf{s}(v,u)}$, where $\mathsf{s}(v, u)$ is again defined by the formula (38). ♢

The map $\mathbf{K}_{30} : \mathbf{C}_3 \to \mathbf{C}_0$ is defined in terms of these k_{vu} by

$$\mathbf{K}_{30} = \sum k_{(v'3)(u'0)}. \tag{49}$$

The last stage of the argument is now to prove the formula (43):

Proposition 7.6. *The anti-chain-map*

$$\mathbf{F}_{00}\mathbf{K}_{30} + \mathbf{K}_{30}\mathbf{F}_{33} + \mathbf{F}_{10}\mathbf{J}_{31} + \mathbf{J}_{20}\mathbf{F}_{32}$$

from $\mathbf{C}_3$ to $\mathbf{C}_0$ is chain-homotopic to ± 1.

Remark. Our definitions mean that $\mathbf{C}_0$ and $\mathbf{C}_3$ are the same group, but the differential $\mathbf{F}_{33}$ is $-\mathbf{F}_{00}$, because of the sign $(-1)^{\mathsf{s}(v,u)}$ in (38).

Proof. Let $w \geq u$ be given, with $w = (w', 3)$ and $u = (u', 0)$, with $w', u' \in \{0, 1\}^{N-1}$. We must prove a formula of the shape:

$$\begin{aligned} \sum_{\{v | v_N = 0\}} f_{vu} k_{wv} + \sum_{\{v | v_N = 3\}} k_{vu} f_{wv} + \sum_{\{v | v_N = 1\}} f_{vu} j_{wv} \\ + \sum_{\{v | v_N = 2\}} j_{vu} f_{wv} + \pm n_{wu} = 0 \end{aligned} \tag{50}$$

where n_{wu} are the components of a map $\mathbf{N}$ chain-homotopic to ± 1 from $\mathbf{C}_3$ to $\mathbf{C}_0$. As usual, the proof goes by equating the matrix-entry of the left-hand side, from γ to α, with the number of ends of an oriented 1-manifold, in this case the 1-manifold

$$\breve{M}^+_{wu}(\gamma, \alpha)_1.$$

As such, the above formula has again the same architecture as the general chain-homotopy formula (23). In the latter formula, the terms $\partial \circ m_G$ and $m_G \circ \partial$ correspond to special cases of the first two terms of (50), of the special form

$$f_{uu} k_{wu} \qquad \text{or} \qquad k_{wu} f_{ww}.$$

With the exception of these terms and the term $\pm n_{wu}$, the terms in (50) in the four summations are the contributions from the first four types of faces of $\breve{G}^+_{wu}$. Specifically, the case (a)(i) gives rise to the terms $k_{vu} f_{wv}$ with $w > v > u$ in 50; the case (a)(ii) gives rise similarly to the terms $f_{vu} k_{wv}$; the cases (a)(iii) and (a)(iv) provide the terms $j_{vu} f_{wv}$ and $f_{vu} j_{wv}$.

The terms from faces (b) of type $G_{w'u'} \times Q(\mathbb{S}_{31})$ and $G_{w'u'} \times Q(\mathbb{S}_{20})$ are all zero, for the same reason as in the previous subsection: for these families of broken metrics we have pulled off a connect-summand $(S^4, \mathbb{RP}^2)$.

What remains is the contribution corresponding to the face of the form $G_{w'u'} \times Q(\mathbb{S}_{30})$. We will complete the proof of the lemma by showing that these contributions are the matrix entries of a map n_{wu} which is chain-homotopic to ± 1. That is, we define $n_{wu}(\gamma, \alpha)$ by counting with sign the ends of $M_{wu}(\gamma, \alpha)_1$ which lie over this face; we define n_{wu} to be the map with matrix entries $n_{wu}(\gamma, \alpha)$, and we define $\mathbf{N}$ to be the map $\mathbf{C}_3$ to $\mathbf{C}_0$ whose components are the n_{wu}. With this understood, we then have

$$\mathbf{F}_{00}\mathbf{K}_{30} + \mathbf{K}_{30}\mathbf{F}_{33} + \mathbf{F}_{10}\mathbf{J}_{31} + \mathbf{J}_{20}\mathbf{F}_{32} \pm \mathbf{N} = 0.$$

From this it follows formally that $\mathbf{N}$ is an anti-chain map; and to complete the proof of the proposition, we must show:

the map $\mathbf{N}$ *is chain-homotopic to the identity.* (51)

The face $G_{w'u'} \times Q(\mathbb{S}_{30})$ parametrizes metrics on a broken Riemannian manifold with two components, obtained by cutting along $\mathbb{S}_{30}$. Recall that $\mathbb{S}_{30}$ is a 3-sphere meeting the embedded surface S_{wv} in a 2-component unlink. One component of the broken manifold is the pair (B_{30}, M_{30}), equipped with a cylindrical end, were B_{30} is the standard 4-ball described above: it contains the embedded surface M_{30} obtained by plumbing two Möbius bands. The second component has three cylindrical ends: we denote it by

(W', S'), and it is obtained by removing B_{30} from $(\mathbb{R} \times Y, S_{wu})$ and attaching a cylindrical end. The manifold-pair (B_{30}, M_{30}) carries the 1-parameter family of metrics $Q(\mathbb{S}_{30})$, obtained by stretching along $\mathbb{S}_{20}$ or $\mathbb{S}_{31}$ as $T \to -\infty$ or $+\infty$ respectively, while the cobordism W' with cylindrical ends carries a family of metrics $G_{w'u'}$. The dimension of $G_{w'u'}$ is equal to $|w' - u'|_1$.

As in Lemma 7.4, we consider now the cobordism $(\bar{W}, \bar{S})$ obtained from (W', S') by attaching to the $\mathbb{S}_{30}$ end a pair (B_{30}, Δ), where Δ is a pair of standard disks in the 4-ball B_{30} having boundary the unlink. The manifold $\bar{W}$ is topologically a cylinder on Y, and in the case $N = 1$ the embedded surface $\bar{S}$ is also a trivial, cylindrical cobordism from K_3 $(= K_0)$ to K_0, as Lemma 7.4 states. For larger N, we can identify $\bar{S}$ with the cobordism standard $S_{\bar{w}u}$ from $K_{\bar{w}}$ to K_u, where $\bar{w} = w'0$ and $u = u'0$. (So $K_{\bar{w}}$ is the same link as $K_w = K_{w'3}$.)

Lemma 7.7. *The cobordism $(I \times Y, S_{\bar{w}u})$ from $K_{\bar{w}}$ to K_u, equipped with the family of metrics $G_{\bar{w}u}$, gives rise to the identity map from $C_{\bar{w}}$ to C_u in the case $\bar{w} = u$ and the zero map otherwise.*

Proof. The family of metrics $G_{\bar{w}u}$ includes the redundant $\mathbb{R}$ factor, so the induced map counts only translation-invariant instantons. These exist only when $\bar{w} = u$, in which case they provide the identity map. □

In light of the lemma, we can prove the assertion (51), if we can show that $\mathbf{N}$ is chain-homotopic (up to an overall sign) to the map obtained from the cobordism $\bar{S}$ with the family of metrics $G_{\bar{w}u}$. We will do this by introducing a third map, $\mathbf{N}'$, whose components n'_{wu} count solutions on the pair (W', S') with its three cylindrical ends.

To define $\mathbf{N}'$ in more detail, recall again that the third end of this pair is a cylinder on the pair $(\mathbb{S}_{30}, \partial\Delta)$, which is a 2-component unlink. For the pair $(\mathbb{S}_{30}, \partial\Delta)$, the critical points comprise a closed interval

$$\mathfrak{C}(\mathbb{S}_{30}, \partial\Delta) = [0, \pi].$$

To see this, note that the fundamental group of the link complement is free on two generators and we are looking at homomorphisms from this free group to $SU(2)$ which send each generator to a point in the conjugacy class of our preferred element (24). This conjugacy class is a 2-sphere, and (up to conjugacy) the homomorphism is determined by the great-circle distance between the images of the two generators. In this closed interval, the interior points represent irreducible representations, while the two endpoints are reducible. To be more precise, in order to identify the critical points with $[0,\pi]$ in this way, we need to choose a relative orientation of the two components of the unlink $\partial\Delta$, because without any orientations the two generators of the free group are well-defined only up to sign. Changing our choice of orientation will change our identification by flipping the interval $[0,\pi]$.

For a generic perturbation of the equations, any solution on (W', S') lying in a zero-dimensional moduli space is asymptotic to a critical point in the interior of the interval on this end [18, Lemma 3.2]. We define $n'_{wu}(\gamma,\alpha)$ by counting these solutions over the family of metrics $G_{w'u'}$, and we define $\mathbf{N}'$ as usual in terms of its components $n'_{wu}(\gamma,\alpha)$.

Each critical point on $(\mathbb{S}_{30}, \partial\Delta)$ extends uniquely to a flat connection on the pair (B_{30}, Δ). So we can regard $\mathbf{N}'$ also as obtained by counting solutions on the broken manifold with two pieces: (W', S') and (B_{30}, Δ), with their cylindrical ends. Since this broken manifold is obtained in turn from $(I \times Y, S_{\bar{w}u})$ by stretching across $\mathbb{S}_{30}$, we see by an argument similar to the previous ones that $\mathbf{N}'$ is chain-homotopic to the map arising from the cobordism $(I \times Y, S_{\bar{w}u})$ with its family of metrics $G_{\bar{w}u}$: i.e. to the identity map, by the lemma. (See also [18, section 3.3].) All that remains now is to prove:

$$\mathbf{N} = \mathbf{N}', \textit{ up to an overall sign.} \tag{52}$$

The components $n'_{wu}(\gamma,\alpha)$ of $\mathbf{N}'$ count the points of the moduli spaces $M(W', S'; \gamma, \alpha)_0$ on the three-ended manifold W', and as stated above, this moduli space comes with a map to the space of critical points on the $\mathbb{S}_{30}$ end:

$$\begin{aligned} r : M(W', S'; \gamma, \alpha)_0 &\to \mathfrak{C}(\mathbb{S}_{30}, \partial\Delta) \\ &= [0, \pi]. \end{aligned}$$

The components $n_{wu}(\gamma,\alpha)$ of $\mathbf{N}$ on the other hand count the points of a fiber product of the map r with a map

$$s : M_{Q(\mathbb{S}_{30})}(B_{30}, M_{30})_1 \to \mathfrak{C}(\mathbb{S}_{30}, \partial\Delta)$$

where the left-hand side is the 1-dimensional part of the moduli space on the pair (B_{30}, M_{30}) equipped with a cylindrical end and carrying the 1-parameter family of metrics $Q(\mathbb{S}_{30})$. To show that $\mathbf{N} = \mathbf{N}'$ up to sign, it suffices to show that the map s is a proper map of degree ± 1 onto the interior of the interval $[0,\pi]$. (The actual sign here depends on a choice of orientation for the moduli space $M_{Q(\mathbb{S}_{30})}(B_{30}, M_{30})$.)

The two ends of the family of metrics $Q(\mathbb{S}_{30})$ on (B_{30}, M_{30}) correspond to two different connected-sum decompositions of (B_{30}, M_{30}), both of which have the form

$$(B_{30}, M_{30}) = (B_{30}, A)\#(S^4, \mathbb{RP}^2)$$

where A is a standard annulus in the 4-ball, with $\partial A = \partial\Delta$, and $\mathbb{RP}^2$ is (as before) a standard $\mathbb{RP}^2$ with positive self-intersection. Since these are two different decompositions, we really have two different annuli A involved here; so we should write the first summand as (B_{30}, A_+) or (B_{30}, A_-) to distinguish the two cases. The two annuli can be distinguished as follows: either annulus determines a preferred isotopy-class of diffeomorphisms between its two boundary components (the two components of the unlink $\partial\Delta$); but the annuli A_+ and A_- determine isotopy classes of diffeomorphisms with the opposite orientation.

Considering the gluing problem for this connected sum, we see that the parametrized moduli space has two ends, one for each end of the parameter space $Q(\mathbb{S}_{30})$, and that each end is obtained by gluing the standard irreducible solution on $(S^4, \mathbb{RP}^2)$ to a flat, reducible connection on $(B_{30}, A_\pm)$. The limiting value of the map s on the two ends is equal to critical point in $\mathfrak{C}(\mathbb{S}_{30}, \partial\Delta)$ arising as the restriction to $(\mathbb{S}_{30}, \partial\Delta)$ of the unique flat solution on $(B_{30}, A_\pm)$. In each case, this value is one of the two ends of the interval $[0, \pi]$; and if A_+ gives rise to the endpoint $0 \in [0, \pi]$, then A_- will give rise to the endpoint π, because the two different annuli provide identifications of the two boundary components that differ in orientation, as explained above. □

7.4 The absolute $\mathbb{Z}/4$ grading

We return briefly to Theorem 6.8, which expresses the existence of a quasi-isomorphism between two complexes. The complex (C_w, f_{ww}) is just the complex that computes $I^\omega(K_w)$, to within an immaterial change of sign in f_{ww}; so this complex carries a relative $\mathbb{Z}/4$ grading. In the spirit of Proposition 4.4, we can fix absolute $\mathbb{Z}/4$ gradings on all the complexes C_v, for $v \in \mathbb{Z}^N$, in such a way that the maps $f_{vv'}$ when $|v - v'|_1 = 1$ have degree

$$-\chi(S_{vv'}) - b_0(K_v) + b_0(K_{v'}) = 1 - b_0(K_v) + b_0(K_{v'}).$$

For general $v \geq u$, let us also write

$$\begin{aligned}\iota(v, u) &:= -\chi(S_{vu}) - b_0(K_v) + b_0(K_u) \\ &= |v - u|_1 - b_0(K_v) + b_0(K_u)\end{aligned}$$

Let us denote by $\langle n\rangle$ a shift of grading by n mod 4, so that if A has a generator in degree i then $A\langle n\rangle$ has a generator in degree $i - n$. Then the cobordism S_{vu} equipped with just a fixed metric (not a family) induces a chain map of degree 0,

$$C_v \to C_u\langle \iota(v,u)\rangle.$$

Having fixed an absolute $\mathbb{Z}/4$ grading for (C_w, f_{ww}) in this way, we can ask how we may grade the other complex $(\mathbf{C}[vu], \mathbf{F}[vu])$ in Theorem 6.8 so that the quasi-isomorphism respects the $\mathbb{Z}/4$ grading. Let us then refine the definition of $\mathbf{C}[vu]$ by specifying a grading mod 4:

$$\mathbf{C}[vu] = \bigoplus_{v \geq v' \geq u} C_{v'}\langle j(v')\rangle,$$

where

$$\begin{aligned} j(v') &= -\iota(v', w) - |v' - u|_1 \\ &= -\iota(v', u) - |v' - u|_1 - n + b_0(K_u) - b_0(K_w) \end{aligned}$$

where $n = |v - u|_1$ (the dimension of the cube). With this definition, it is easily verified that the differential $\mathbf{F}[vu]$ has degree -1. So $(\mathbf{C}[vu], \mathbf{F}[vu])$ is another $\mathbb{Z}/4$-graded complex. We then have the following refinement of the theorem:

Proposition 7.8. *If C_w and $\mathbf{C}[vu]$ are given absolute $\mathbb{Z}/4$ gradings as above, then the quasi-isomorphism of Theorem 6.8 becomes a quasi-isomorphism*

$$C_w \to \mathbf{C}[vu]$$

of $\mathbb{Z}/4$ graded complexes.

Proof. The quasi-isomorphism is exhibited in the proof of Theorem 6.8 as the composite of n maps, each of which (as is easy to check) has a well-defined $\mathbb{Z}/4$ degree. The composite map has a component

$$C_w \to C_v\langle j(v)\rangle \subset \mathbf{C}[vu]$$

which is the map induced by the cobordism S_{wv}. The Euler number of this surface is $-n$, so the map $C_w \to C_v$ induced by S_{wv} has degree

$$\iota(w, v) = n - b_0(K_w) + b_0(K_v)$$

with respect to the original $\mathbb{Z}/4$ gradings. This last quantity coincides with $j(v)$; so the map has degree 0 as a map

$$C_w \to C_v\langle j(v)\rangle \subset \mathbf{C}[vu].$$

□

8 Unlinks and the E_2 term

8.1 Statement of the result

We now turn to classical knots and links K, and invariants $I^{\natural}(K)$ and $I^{\sharp}(K)$ introduced in section 4.3 above. We will focus on the *unreduced* version, $I^{\sharp}(K)$, and return to the reduced version later. Recall that we have defined

$$I^{\sharp}(K) = I^{\omega}(S^3, K \amalg H)$$

where K is regarded as a link in $\mathbb{R}^3$ and H is a standard Hopf link near infinity, with ω an arc joining the components of H. From this definition, it is apparent that the results of section 6 apply equally well to the invariant $I^{\sharp}(K)$ as they do to $I^{\omega}(Y, K)$ in general. Thus for example, if K_2, K_1 and K_0 are links in S^3 which differ only inside a single ball, as in Figure 6, then there is a skein exact sequence

$$\cdots \to I^{\sharp}(K_2) \to I^{\sharp}(K_1) \to I^{\sharp}(K_0) \to \cdots$$

in which the maps are induced by the elementary cobordisms S_{21} etcetera. More generally, we can consider again a collection of links K_v indexed by $v \in \{0,1,2\}^N$ which differ by the same unoriented skein relations in a collection of N disjoint balls in $\mathbb{R}^3$. From Corollary 6.9 we obtain:

Corollary 8.1. *For links K_v as above, there is a spectral sequence whose E_1 term is*

$$\bigoplus_{v \in \{0,1\}^N} I^{\sharp}(K_v)$$

and which abuts to the instanton Floer homology $I^{\sharp}(K_v)$, for $v = (2, \ldots, 2)$. The differential d_1 is the sum of the maps induced by the cobordisms S_{vu} with $v > u$ and $|v - u| = 1$, equipped with I-orientations satisfying the conditions of Lemma 6.1 and corrected by the signs $(-1)^{\tilde{\delta}(v,u)}$ as given in Corollary 6.10. □

Let K be a link in $\mathbb{R}^3 \subset S^3$ with a planar projection giving a diagram D in $\mathbb{R}^2$. Let N be the number of crossings in the diagram. As in [16], we can consider the 2^N possible smoothings of D, indexed by the points v of the cube $\{0,1\}^N$, with the conventions of [16, 30], for example. This labeling of the smoothings is consistent with the convention illustrated in Figure 6. This gives 2^N different unlinks K_v. For each $v \geq u$ in $\{0,1\}^N$, we have our standard cobordism S_{vu} from K_v to K_u.

We can consistently orient all the links K_v, for $v \in \{0,1\}^N$, and all the cobordisms S_{vu}, so that $\partial S_{vu} = K_u - K_v$. To do this, start with a checkerboard coloring of the regions of the diagram D, and simply orient each K_v so that, away from the crossings and their smoothings, the orientation of K_v agrees with the boundary orientation of the black regions of the checkerboard coloring. We can then give each S_{vu} the I-orientation is obtains as an oriented surface. The resulting I-orientations respect composition: they satisfy the conditions of Lemma 6.1.

We therefore apply Corollary 8.1 to this situation. We learn that there is a spectral sequence abutting to $I^\sharp(K)$ whose E_1 term is

$$E_1 = \bigoplus_{v \in \{0,1\}^N} I^\sharp(K_v).$$

In this sum, each K_v is an unlink. The differential d_1 is

$$d_1 = \sum_{v \geq u} (-1)^{\tilde{\delta}(v,u)} I^\sharp(S_{vu}), \tag{53}$$

where each cobordism S_{vu} is obtained from a "pair of pants" that either joins two components into one, or splits one component into two. We can consider the spectral sequence of Corollary 8.1 in this setting, about which we have the following result.

Theorem 8.2. *In the above situation, the page (E_1, d_1) of the spectral sequence furnished by Corollary 8.1 is isomorphic (as an abelian group with differential) to the complex that computes the Khovanov cohomology of $\bar{K}$ (the mirror image of K) from the given diagram. Therefore, the E_2 term of the spectral sequence is isomorphic to the Khovanov cohomology of $\bar{K}$. The spectral sequence abuts to the instanton homology $I^\sharp(K)$.*

Remarks. The relation expressed by this theorem, between $I^\sharp(K)$ and $Kh(\bar{K})$, pays no attention to the bigrading that is carried by $Kh(\bar{K})$. It is natural to ask, for example, whether at least the filtration of $I^\sharp(K)$ that arises from the spectral sequence is a topological invariant of K. More generally, one can ask whether the intermediate pages of the spectral sequence, as filtered groups, are invariants of K (see [2] for the similar question concerning the spectral sequence of Ozsváth and Szabó). A related question is whether the intermediate pages are functorial for knot cobordisms. Although the bigrading is absent, there is at least a $\mathbb{Z}/4$ grading throughout: carrying Proposition 7.8 over to the present situation, we see that there is a spectral sequence of $\mathbb{Z}/4$ graded groups abutting to $I^\sharp(K)$ whose E_1 term is

$$E_1 = \bigoplus_{v\in\{0,1\}^N} I^\sharp(K_v)\langle k_v\rangle, \tag{54}$$

with

$$k_v = -b_0(K_v) + 2b_0(K_0) - b_0(K) - N. \tag{55}$$

In deriving this formula from the formula for $j(v')$, we have used the fact that the cobordisms between the different smoothings are all orientable, which implies that $|v - v'|_1 = b_0(v) - b_0(v')$ mod 2. The formula for k_v can also be written (mod 4) as

$$k_v = -b_0(K_v) + b_0(K) - N_- + N_+,$$

where N_- and N_+ are the number of positive and negative crossings in the diagram. From this version of the formula, it is straightforward to compare our $\mathbb{Z}/4$ grading to the bigradings in [16]. The result is that the $\mathbb{Z}/4$-graded E_2 page of our spectral sequence is isomorphic to $Kh(\bar{K})$ with the $\mathbb{Z}/4$-grading defined by

$$q - h - b_0(K)$$

where q and h are the q-grading and homological grading respectively. That is, the part of the E_2 term in $\mathbb{Z}/4$-grading α is

$$\bigoplus_{j-i-b_0(K)=\alpha} Kh^{i,j}(\bar{K}).$$

Understanding the E_1 page and the differential d_1 means computing $I^\sharp(K_v)$ for an unlink K_v and computing the maps given by pairs of pants. We take up these calculations in the remaining parts of this section.

8.2 Unlinks

We write U_n for the unlink in $\mathbb{R}^3$ with n components, so that U_0 is the empty link and U_1 is the unknot. We take specific models for these. For example, we may take U_n to be the union of standard circles in the (x, y) plane, each of diameter 1/2, and centered on the first n integer lattice points along the x axis; and we can then orient the components of U_n as the boundaries of the standard disks that they bound.

Given any subset $\mathbf{i} = \{i_1, \ldots, i_m\} \subset \{1, \ldots, n\}$, there is a corresponding m-component sublink $U_\mathbf{i}$ of U_n. We will identify $U_\mathbf{i}$ with U_m in a standard way, via a self-evident isotopy (preserving the ordering of the components).

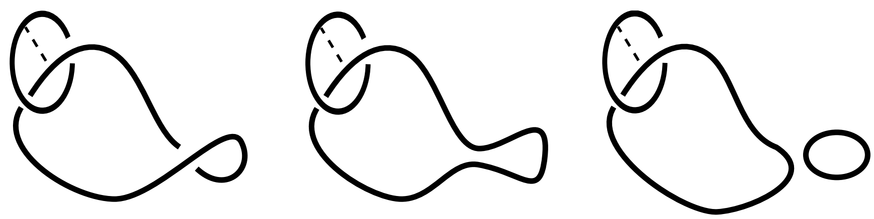

Figure 13: A skein sequence relating U_0 (twice) and U_1.

We have already seen that $I^\sharp(U_0)$ is $\mathbb{Z}$ (Proposition 4.2). There are two possible identifications of $I^\sharp(U_0)$ with $\mathbb{Z}$, differing in sign. We fix one of them, once and for all by specifying a generator

$$\mathbf{u}_0 \in I^\sharp(U_0),$$

so that

$$I^\sharp(U_0) = \mathbb{Z}.$$

This $\mathbb{Z}$ occurs in grading 0 mod 4, by convention.

Lemma 8.3. *For the unknot U_1, the instanton homology $I^\sharp(U_1)$ is free of rank* 2*,*

$$I^\sharp(U_1) \cong \mathbb{Z} \oplus \mathbb{Z},$$

with generators in degrees 0 *and* -2 *mod* 4*.*

Proof. Draw a diagram of the Hopf link H with an extra crossing, so that by smoothing that crossing in two different ways one obtains the links H (again) and $H \amalg U_1$ (see Figure 13). The skein sequence for this situation gives a long exact sequence

$$\cdots \to I^\sharp(U_0) \xrightarrow{a} I^\sharp(U_1) \xrightarrow{b} I^\sharp(U_0) \xrightarrow{c} I^\sharp(U_0) \to \cdots,$$

in which the maps a and b have degree -2 and 0 respectively, while c has degree 1. From our calculation of $I^\sharp(U_0)$ it follows that $c = 0$ and that $I^\sharp(U_1)$ is free of rank 2 with generators in degrees 0 and -2. The generator of degree -2 is the image of a, while the generator of degree 0 is mapped by b to a generator of $I^\sharp(U_0) = I^\omega(S^3, H)$. □

We wish to have explicit generators of the rank-2 group $I^\sharp(U_1)$ defined without reference to the auxiliary Hopf link H. To this end, let D be the standard disk that U_1 bounds in the (x,y) plane. Let D^+ be the oriented cobordism from the empty link U_0 to U_1 obtained by pushing the disk D a little into the 4-dimensional cylinder $[-0,1]\times\mathbb{R}^3$. Similarly, let D^- be the cobordism from U_1 to U_0 obtained from D with its opposite orientation. These oriented cobordisms give preferred maps

$$I^\sharp(D^+): I^\sharp(U_0)\to I^\sharp(U_1)$$
$$I^\sharp(D^-): I^\sharp(U_1)\to I^\sharp(U_0)$$

of degrees 0 and 2 respectively.

Lemma 8.4. *There are preferred generators $\mathbf{v}_+$ and $\mathbf{v}_-$ for the rank-2 group $I^\sharp(U_1)$, in degrees 0 and 2 mod 4 respectively, characterized by the conditions*

$$I^\sharp(D^+)(\mathbf{u}_0)=\mathbf{v}_+$$

and

$$I^\sharp(D^-)(\mathbf{v}_-)=\mathbf{u}_0$$

respectively, where $\mathbf{u}_0$ is the chosen generator for $I^\sharp(U_0)=\mathbb{Z}$.

Proof. The proof of the previous lemma shows that a generator of the degree-0 part of $I^\sharp(U_1)$ is the image of the map b. So to show that $\mathbf{v}_+$ as defined in the present lemma is a generator it suffices to show that the composite map $b\circ I^\sharp(D^+)$ is the identity map on the rank-1 group $I^\sharp(U_0)=I^\omega(S^3,H)$. This in turn follows from the fact that the composite cobordism from the Hopf link H to itself is a product.

Similarly, to show that there is a generator $\mathbf{v}_-$ of the degree-2 part with the property described, it suffices to show that the composite map $I^\sharp(D^-)\circ a$ is the identity map $I^\sharp(U_0)$. The composite cobordism is again a product, so the result follows. □

Corollary 8.5. *Write $V=\langle\mathbf{v}_+,\mathbf{v}_-\rangle\cong\mathbb{Z}^2$ for the group $I^\sharp(U_1)$. We then have isomorphisms of $\mathbb{Z}/4$-graded abelian groups,*

$$\Phi_n: V^{\otimes n}\to I^\sharp(U_n),$$

for all n, with the following properties. First, if D_n^+ denotes the cobordism from U_0 to U_n obtained from standard disks as in the previous lemma, then

$$I^\sharp(D_n^+)(\mathbf{u}_0)=\Phi_n(\mathbf{v}_+\otimes\cdots\otimes\mathbf{v}_+).$$

Second, the isomorphism is natural for split cobordisms from U_n to itself. Here, a "split" cobordism means a cobordism from U_n to U_n in $[0,1] \times \mathbb{R}^3$ which is the disjoint union of n cobordisms from U_1 to U_1, each contained in a standard ball $[0,1] \times B^3$.

Proof. This follows from the general product formula for split links, Corollary 5.9. □

Remark. We will see later, in Proposition 8.10, the extent to which this isomorphism is canonical.

8.3 An operator of degree 2

Let K be a link, let p be a marked point on K, and let an orientation be chosen for K at p. We can then form a cobordism S from K to K by taking the cylinder $[-1,1] \times K$ and forming a connect sum with a standard torus at the point $(0,p)$. This cobordism then determines a map

$$\sigma : I^\sharp(K) \to I^\sharp(K) \tag{56}$$

of degree 2 mod 4.

Lemma 8.6. *For any link K, and any base-point $p \in K$, the map σ is nilpotent.*

Proof. The map σ behaves naturally with respect to cobordisms of links with base-points. So if K_2, K_1 and K_0 are three links related by the unoriented skein relation, then σ commutes with the maps in the long exact sequence relating the groups $I^\sharp(K_i)$. From this it follows that if σ is nilpotent on $I^\sharp(K_i)$ for two of the three links K_i then it is nilpotent also on the third. By repeated use of the skein relation, we see that it is enough to check the case that $K = U_1$. Finally, for U_1, we can use the exact sequence from the proof of Lemma 8.3 to reduce to the case of $I^\sharp(U_0)$, or more precisely to the case of $I^\omega(S^3, H)$, with a marked point p on one of the two components of the Hopf link H. This last case is trivial, however, because the group has a generator only in one of the degrees mod 4, which forces σ to be zero on $I^\sharp(U_0)$. □

8.4 Pairs of pants

Let Π be a pair-of-pants cobordism from U_1 to U_2. We wish to calculate the corresponding map on instanton homology. Via the isomorphisms of Corollary 8.5, this map becomes a map

$$\Delta : V \to V \otimes V.$$

The degree of this map is -2.

There is also a pair-of-pants cobordism Π, from U_2 to U_1, which induces a map

$$\nabla : V \otimes V \to V$$

of degree 0.

Lemma 8.7. *In terms of the generators* $\mathbf{v}_+$ *and* $\mathbf{v}_-$ *for* V*, the map* Δ *is given by*

$$\begin{aligned} \mathbf{v}_- &\mapsto \mathbf{v}_- \otimes \mathbf{v}_- \\ \mathbf{v}_+ &\mapsto (\mathbf{v}_- \otimes \mathbf{v}_+ + \mathbf{v}_+ \otimes \mathbf{v}_-) \end{aligned}$$

and ∇ *is given by*

$$\begin{aligned} \mathbf{v}_+ \otimes \mathbf{v}_+ &\mapsto \mathbf{v}_+ \\ \mathbf{v}_+ \otimes \mathbf{v}_- &\mapsto \mathbf{v}_- \\ \mathbf{v}_- \otimes \mathbf{v}_+ &\mapsto \mathbf{v}_- \\ \mathbf{v}_- \otimes \mathbf{v}_- &\mapsto 0. \end{aligned}$$

Proof. We begin with $\Delta(\mathbf{v}_+)$. Because of the $\mathbb{Z}/4$ grading, we know that

$$\Delta(\mathbf{v}_+) = \lambda_1(\mathbf{v}_- \otimes \mathbf{v}_+) + \lambda_2(\mathbf{v}_+ \otimes \mathbf{v}_-),$$

for some integers λ_1 and λ_2. Consider the composite cobordism $\Pi_0 = D^+ \cup \Pi$ from U_0 to U_2, formed from the pair of pants Π by attaching a disk to the incoming boundary component. The composite cobordism determines a map

$$\Delta_0 : \mathbb{Z} \to V^{\otimes 2}$$

with

$$\Delta_0(\mathbf{u}_0) = \lambda_1(\mathbf{v}_- \otimes \mathbf{v}_+) + \lambda_2(\mathbf{v}_+ \otimes \mathbf{v}_-).$$

Next, form a cobordism $\Pi_1 = \Pi_0 \cup D^-$ from U_0 to U_1 by attaching a disk to the first of the two outgoing boundary components of Π_0. As a cobordism from U_1 to U_0, the disk maps $\mathbf{v}_-$ to 1; so by the naturality expressed in Corollary 5.9, it maps $\mathbf{v}_- \otimes \mathbf{v}_+$ to $\mathbf{v}_+$ when viewed as a cobordism from U_2 to U_1. The composite cobordism Π_1 therefore defines a map

$$\Delta_1 : \mathbb{Z} \to V$$

with

$$\Delta_1(\mathbf{u}_0) = \lambda_1 \mathbf{v}_+.$$

But Π_1 is simply a disk, with the same orientation as the original Π; so $\Delta_1(\mathbf{u}_0) = \mathbf{v}_+$ by definition of $\mathbf{v}_+$, and we conclude that $\lambda_1 = 1$. We also have $\lambda_2 = 1$ by the same argument, so we have computed $\Delta(\mathbf{v}_+)$.

As a preliminary step towards computing $\Delta(\mathbf{v}_-)$, we compute the effect of the degree-2 operator σ on $I^\sharp(U_1)$. The cobordism which defines σ on U_1 can be seen as a composite cobordism $U_1 \to U_1$ obtained from two pairs of pants: first Π from U_1 to U_2, then $\amalg$ from U_2 to U_1. These are oriented surfaces, so the cobordisms are canonically I-oriented, in a manner that is compatible with composition. By an argument dual to the one in the previous paragraph, we see that

$$\nabla(\mathbf{v}_+ \otimes \mathbf{v}_-) = \nabla(\mathbf{v}_- \otimes \mathbf{v}_+) = \mathbf{v}_-.$$

Looking at the composite $\nabla \circ \Delta$, we see that

$$\sigma(\mathbf{v}_+) = 2\mathbf{v}_-.$$

We also know that σ is nilpotent (Lemma 8.6), so we must have

$$\sigma(\mathbf{v}_-) = 0.$$

We now appeal to the naturality of σ with respect to cobordism of links with base-points. This tells us that

$$\Delta \circ \sigma = (1 \otimes \sigma) \circ \Delta.$$

In particular,

$$\Delta(\sigma(\mathbf{v}_+)) = (1 \otimes \sigma)(\Delta(\mathbf{v}_+)).$$

We have already calculated $\Delta(\mathbf{v}_+)$ and $\sigma(\mathbf{v}_+)$, so we have,

$$\begin{aligned} \Delta(2\mathbf{v}_-) &= (1 \otimes \sigma)(\mathbf{v}_- \otimes \mathbf{v}_+ + \mathbf{v}_+ \otimes \mathbf{v}_-) \\ &= 2(\mathbf{v}_- \otimes \mathbf{v}_-). \end{aligned}$$

It follows that $\Delta(\mathbf{v}_-) = \mathbf{v}_- \otimes \mathbf{v}_-$ as claimed. A dual argument determines the remaining terms of ∇ in a similar manner. □

8.5 Isotopies of the unlink

Lemma 8.8. *Let $S \subset [0,1] \times \mathbb{R}^3$ be a closed, oriented surface, regarded as a cobordism from the empty link U_0 to itself. Then the induced map $I^\sharp(S) : I^\sharp(U_0) \to I^\sharp(U_0)$ is multiplication by 2^k if S consists of k tori; and $I^\sharp(S)$ is zero otherwise.*

Proof. The first point is that the map $I^\sharp(S)$ in this situation depends only on S as an abstract surface, not on its embedding in $(0,1)\times\mathbb{R}^3$. This can be deduced from results of [18], which show that the invariants of a closed pair (X,Σ) defined using singular instantons depend on Σ only through its homotopy class. To apply the results of [18] to the present situation, we proceed as follows. Since $I^\sharp(U_0)$ is $\mathbb{Z}$, the map $I^\sharp(S)$ is determined by its trace; and twice the trace can be interpreted as the invariant of a closed pair (X,Σ) obtained by gluing the incoming to the outgoing ends of the cobordism. Thus X is $S^1\times S^3$ and Σ is the union of $S^1\times H$ and the surface S. The results of [18] can be applied directly to any homotopy of S that remains in a ball disjoint from $S^1\times H$, and this is all that we need.

Because of this observation, it now suffices to verify the statement in the case that S is a standard connected, oriented surface of arbitrary genus.If we decompose a genus-1 surface S as an incoming disk D^+, a genus-1 cobordism from U_1 to itself, and an outgoing disk D^-, we find that $I^\sharp(S)$ in this case is given by

$$I^\sharp(S)(\mathbf{u}_0) = I^\sharp(D^-)\circ\sigma\circ I^\sharp(D^+)(\mathbf{u}_0),$$

which is $2\mathbf{u}_0$ by our previous results. For the case of genus g, we look at

$$I^\sharp(S)(\mathbf{u}_0) = I^\sharp(D^-)\circ\sigma^g\circ I^\sharp(D^+)(\mathbf{u}_0),$$

which is zero for all g other than $g=1$. □

Lemma 8.9. *Let S be an oriented concordance from the standard unlink U_n to itself, consisting of n oriented annuli in $[0,1]\times\mathbb{R}^3$. Let τ be the permutation of $\{1,\ldots,n\}$ corresponding to the permutation of the components of U_n arising from S. Then the standard isomorphism Φ_n of Corollary 8.5 intertwines the map*

$$I^\sharp(S) : I^\sharp(U_n)\to I^\sharp(U_n)$$

with the permutation map

$$\tau_* : V\otimes\cdots\otimes V\to V\otimes\cdots\otimes V.$$

In particular, if the permutation τ is the identity, then $I^\sharp(S)$ is the identity.

Proof. We start with the case that the permutation τ is the identity. Let $\sigma_i : I^\sharp(U_n)\to I^\sharp(U_n)$ be the map σ applied with a chosen basepoint on the i'th component of the link. The isomorphism Φ_n intertwines σ_i with

$$1\otimes\cdots\otimes 1\otimes\sigma\otimes 1\otimes\cdots\otimes 1$$

with σ in the i'th spot, by Corollary 8.5. Furthermore, $I^\sharp(S)$ commutes with σ_i, because the corresponding cobordisms commute up to diffeomorphism relative to the boundary. Since $\sigma(\mathbf{v}_+) = 2\mathbf{v}_-$, we therefore see that, to show $I^\sharp(S)$ is the identity, we need only show that

$$I^\sharp(S)(\mathbf{v}_+ \otimes \cdots \otimes \mathbf{v}_+) = \mathbf{v}_+ \otimes \cdots \otimes \mathbf{v}_+.$$

Let λ_m be the coefficient of $\mathbf{v}_+^{\otimes m} \otimes \mathbf{v}_-^{\otimes(n-m)}$ in $I^\sharp(S)(\mathbf{v}_+^{\otimes n})$. (There is no loss of generality in putting the $\mathbf{v}_+$ factors first here: it is only a notational convenience.) We must show that $\lambda_m = 0$ for $m < n$ and $\lambda_n = 1$. From our calculation of σ etc., we see that we can get hold of λ_m by the formula

$$I^\sharp(D_n^-) \circ \sigma_1 \circ \cdots \circ \sigma_m I^\sharp(D_n^+)(\mathbf{u}_0) = 2^m \lambda_m \mathbf{u}_0.$$

On the other hand, the composite map on the left is equal to $I^\sharp(S)$, where S is a closed surface consisting of m tori and $n - m$ spheres, viewed as a cobordism from U_0 to U_0. From the results of the previous lemma, we see that the left-hand side is 0 if $m \neq n$ and is 2^n if $m = n$. This completes the proof in the case that the permutation τ is 1.

For the case that S provides a non-trivial permutation τ of the components, the map $I^\sharp(S)$ intertwines σ_i with $\sigma_{\tau(i)}$. It is again sufficient to show that $I^\sharp(S)$ sends $\mathbf{v}_+^{\otimes n}$ to itself, and essentially the same argument applies. □

As a special case of a cobordism from U_n to itself, we can consider the trace of an isotopy $f_t : U_n \to \mathbb{R}^3$ $(t \in [0,1])$ which begins and ends with the standard inclusion. As an application of the lemma, we therefore have:

Proposition 8.10. *Let $\mathcal{U}_n$ be any oriented link in link-type of U_n, and let its components be enumerated. Then there is canonical isomorphism*

$$\Psi_n : V \otimes \cdots \otimes V \to I^\sharp(\mathcal{U}_n)$$

which can be described as $I^\sharp(S) \circ \Phi_n$, where Φ_n is the standard isomorphism of Corollary 8.5 and S is any cobordism from U_n to $\mathcal{U}_n$ arising from an isotopy from U_n to $\mathcal{U}_n$, respecting the orientations and the enumeration of the components.

If the enumeration of the components of $\mathcal{U}_n$ is changed by a permutation τ, then the isomorphism Ψ_n is changed simply by composition with the corresponding permutation of the factors in the tensor product. □

The following proposition encapsulates the calculations of this section so far.

Proposition 8.11. *Let V be a $\mathbb{Z}/4$-graded free abelian group with generators v_- and v_+ in degrees -2 and 0. Then for each oriented n-component unlink $\mathcal{U}_n$ with enumerated components $K_1, \ldots, K_n$, there is a canonical isomorphism*

$$\Phi_n(\mathcal{U}_n) : V \otimes \cdots \otimes V \to I^\sharp(\mathcal{U}_n)$$

with the following properties.

(a) *Given an orientation-preserving isotopy from $\mathcal{U}_n$ to $\mathcal{U}'_n$, respecting the enumeration of the components, the map $I^\sharp(S)$ arising from the corresponding cobordism S intertwines $\Phi_n(\mathcal{U}_n)$ with $\Phi_n(\mathcal{U}'_n)$.*

(b) *If the components of $\mathcal{U}_n$ are enumerated differently, then $\Phi_n(\mathcal{U}_n)$ changes by composition with the permutation of the factors in the tensor product.*

(c) *If Π_n is the oriented cobordism from $\mathcal{U}_n$ to some $\mathcal{U}_{n+1}$ which attaches a pair of pants Π to the last component, inside a ball disjoint from the other components, then the isomorphisms $\Phi_n(\mathcal{U}_n)$ and $\Phi_n(\mathcal{U}_{n+1})$ intertwine the corresponding map $I^\sharp(\Pi_n) : I^\sharp(\mathcal{U}_n) \to I^\sharp(\mathcal{U}_{n+1})$ with the map*
$$1 \otimes \cdots \otimes 1 \otimes \Delta.$$

(d) *Similarly, for an oriented cobordism $\mathcal{U}_{n+1} \to \mathcal{U}_n$ obtained using a pair of pants $\amalg$ on the last two components, in a ball disjoint from the other components, we obtain the map*

$$1 \otimes \cdots \otimes 1 \otimes \nabla.$$

□

8.6 Khovanov cohomology

We now have all that we need to conclude the proof of Theorem 8.2. If we write $n(v)$ for the number of components of K_v and enumerate those components, then we have a canonical identification

$$E_1 = \bigoplus_{v \in \{0,1\}^N} V^{\otimes n(v)}.$$

The differential d_1 is given by the sum (53) The cobordism S_{vu} is a the union of some product cylinders and a single pair of pants, either Π or $\amalg$.

Proposition 8.11 therefore tells us that, after pre- and post-composing by permutations of the components, the map f_{vu} is given either by

$$(-1)^{\tilde{\delta}(v,u)}(1 \otimes \cdots \otimes 1 \otimes \Delta)$$

or

$$(-1)^{\tilde{\delta}(v,u)}(1 \otimes \cdots \otimes 1 \otimes \nabla).$$

The complex (E_1, d_1) that one arrives at in this way is exactly the complex that computes the Khovanov cohomology of $\bar{K}$. The fact that the mirror $\bar{K}$ of the link K appears in this statement is accounted for by the fact that, in Khovanov's definition, the differential is the sum of contributions from the edges oriented so that $|v|$ *increases* along the edges, whereas in our setup the differential d_1 *decreases* $|v|$. It follows that the E_2 page of the spectral sequence is isomorphic (as an abelian group) to the Khovanov cohomology $Kh(\bar{K})$.

8.7 The reduced homology theories

Recall that for a link K with a marked point x and normal vector v at x, we have defined

$$I^{\natural}(K) = I^{\omega}(S^3, K \cup L),$$

where L is a meridional circle centered at x and ω is an arc in the direction of v. There is a skein exact sequence (illustrated for the case of the unknot in Figure 13),

$$\cdots \to I^{\natural}(K) \to I^{\sharp}(K) \to I^{\natural}(K) \to I^{\natural}(K) \to \cdots .$$

Corollary 8.1 has a straightforward adaptation to this reduced theory, which can again be deduced from the more general result, Corollary 6.9.

The maps in the long exact sequence above have already been described for the unknot U_1. Thus, the map

$$I^{\sharp}(U_1) \to I^{\natural}(U_1)$$

is the same as the map $I^{\sharp}(U_1) \to I^{\sharp}(U_0) = \mathbb{Z}$ given as the quotient map

$$V \mapsto V/\langle v_- \rangle \cong \mathbb{Z}.$$

For the unlink U_n we similarly have

$$I^{\natural}(U_n) = V \otimes \cdots \otimes V \otimes V/\langle v_- \rangle,$$

as a quotient of $V^{\otimes n}$. (The marked point is on the last component here.) The maps ∇ and Δ give rise to maps

$$\nabla_r : V \otimes V/\langle v_- \rangle \to V/\langle v_- \rangle$$

and

$$\Delta_r : V/\langle v_- \rangle \to V \otimes V/\langle v_- \rangle;$$

and these are precisely the maps induced by the pair-of-pants cobordisms ∐ and Π.

In the spectral sequence abutting to $I^\natural(K)$, we can therefore identify E_1 and d_1. The E_1 term is obtained from the unreduced version by replacing (at each vertex of the cube) the factor V corresponding to the marked component by a factor $V/\langle v_- \rangle$. And the differential d_1 is obtained from the unreduced case by replacing ∇ or Δ by ∇_r or Δ_r whenever the marked component is involved. The resulting complex is precisely the complex that computes the reduced Khovanov cohomology of the mirror of K. We therefore have:

Theorem 8.12. *For a knot or link K in S^3, there is a spectral sequence of abelian groups whose E_2 term is the reduced Khovanov cohomology of $\bar{K}$ and which abuts to $I^\natural(K)$.* □

References

[1] S. Akbulut, T. Mrowka, and Y. Ruan. Torsion classes and a universal constraint on Donaldson invariants for odd manifolds. *Trans. Amer. Math. Soc.*, 347(1):63–76, 1995.

[2] J. A. Baldwin. On the spectral sequence from Khovanov homology to Heegaard floer homology. Prepint, 2008.

[3] J. Bloom. A link surgery spectral sequence in monopole Floer homology. Prepint, 2009.

[4] P. J. Braam and S. K. Donaldson. Floer's work on instanton homology, knots and surgery. In *The Floer memorial volume*, volume 133 of *Progr. Math.*, pages 195–256. Birkhäuser, Basel, 1995.

[5] A. Dold and H. Whitney. Classification of oriented sphere bundles over a 4-complex. *Ann. of Math. (2)*, 69:667–677, 1959.

[6] S. K. Donaldson. The orientation of Yang-Mills moduli spaces and 4-manifold topology. *J. Differential Geom.*, 26(3):397–428, 1987.

[7] S. K. Donaldson and P. B. Kronheimer. *The geometry of four-manifolds.* Oxford University Press, New York, 1990.

[8] A. Floer. An instanton-invariant for 3-manifolds. *Comm. Math. Phys.*, 118(2):215–240, 1988.

[9] A. Floer. Instanton homology, surgery, and knots. In *Geometry of low-dimensional manifolds, 1 (Durham, 1989)*, volume 150 of *London Math. Soc. Lecture Note Ser.*, pages 97–114. Cambridge Univ. Press, Cambridge, 1990.

[10] J. E. Grigsby and S. Wehrli. On the colored Jones polynomial, sutured floer homology, and knot floer homology. *Adv. in Math.*, 223:2114–2165, 2010.

[11] M. Hedden. Khovanov homology of the 2-cable detects the unknot. *Math. Res. Lett.*, 16(6):991–994, 2009.

[12] M. Hedden and L. Watson. Does Khovanov homology detect the unknot? Preprint, 2008.

[13] A. Juhász. Holomorphic discs and sutured manifolds. *Algebr. Geom. Topol.*, 6:1429–1457 (electronic), 2006.

[14] A. Juhász. Floer homology and surface decompositions. *Geom. Topol.*, 12(1):299–350, 2008.

[15] T. Kawasaki. The index of elliptic operators over V-manifolds. *Nagoya Math. J.*, 84:135–157, 1981.

[16] M. Khovanov. A categorification of the Jones polynomial. *Duke Math. J.*, 101(3):359–426, 2000.

[17] P. Kronheimer, T. Mrowka, P. Ozsváth, and Z. Szabó. Monopoles and lens space surgeries. *Ann. of Math. (2)*, 165(2):457–546, 2007.

[18] P. B. Kronheimer. An obstruction to removing intersection points in immersed surfaces. *Topology*, 36(4):931–962, 1997.

[19] P. B. Kronheimer and T. S. Mrowka. Instanton Floer homology and the Alexander polynomial. To appear in Algebraic and Geometric Topology.

[20] P. B. Kronheimer and T. S. Mrowka. Gauge theory for embedded surfaces. I. *Topology*, 32(4):773–826, 1993.

[21] P. B. Kronheimer and T. S. Mrowka. Gauge theory for embedded surfaces. II. *Topology*, 34(1):37–97, 1995.

[22] P. B. Kronheimer and T. S. Mrowka. *Monopoles and three-manifolds*. New Mathematical Monographs. Cambridge University Press, Cambridge, 2007.

[23] P. B. Kronheimer and T. S. Mrowka. Knot homology groups from instantons. Preprint, 2008.

[24] P. B. Kronheimer and T. S. Mrowka. Knots, sutures and excision. Preprint, 2008.

[25] E. S. Lee. An endomorphism of the Khovanov invariant. *Adv. Math.*, 197(2):554–586, 2005.

[26] C. Manolescu. An unoriented skein exact triangle for knot Floer homology. *Math. Res. Lett.*, 14(5):839–852, 2007.

[27] C. Manolescu and P. Ozsváth. On the Khovanov and knot Floer homologies of quasi-alternating links. In *Proceedings of Gökova Geometry-Topology Conference 2007*, pages 60–81. Gökova Geometry/Topology Conference (GGT), Gökova, 2008.

[28] P. Ozsváth and Z. Szabó. Holomorphic disks and knot invariants. *Adv. Math.*, 186(1):58–116, 2004.

[29] P. Ozsváth and Z. Szabó. On the Heegaard Floer homology of branched double-covers. *Adv. Math.*, 194(1):1–33, 2005.

[30] J. Rasmussen. Khovanov homology and the slice genus. To appear in *Inventiones Math.*

[31] J. Rasmussen. *Floer homology and knot complements*. PhD thesis, Harvard University, 2003.

[32] J. Rasmussen. Knot polynomials and knot homologies. In *Geometry and topology of manifolds*, volume 47 of *Fields Inst. Commun.*, pages 261–280. Amer. Math. Soc., Providence, RI, 2005.

[33] C. H. Taubes. Casson's invariant and gauge theory. *J. Differential Geom.*, 31(2):547–599, 1990.