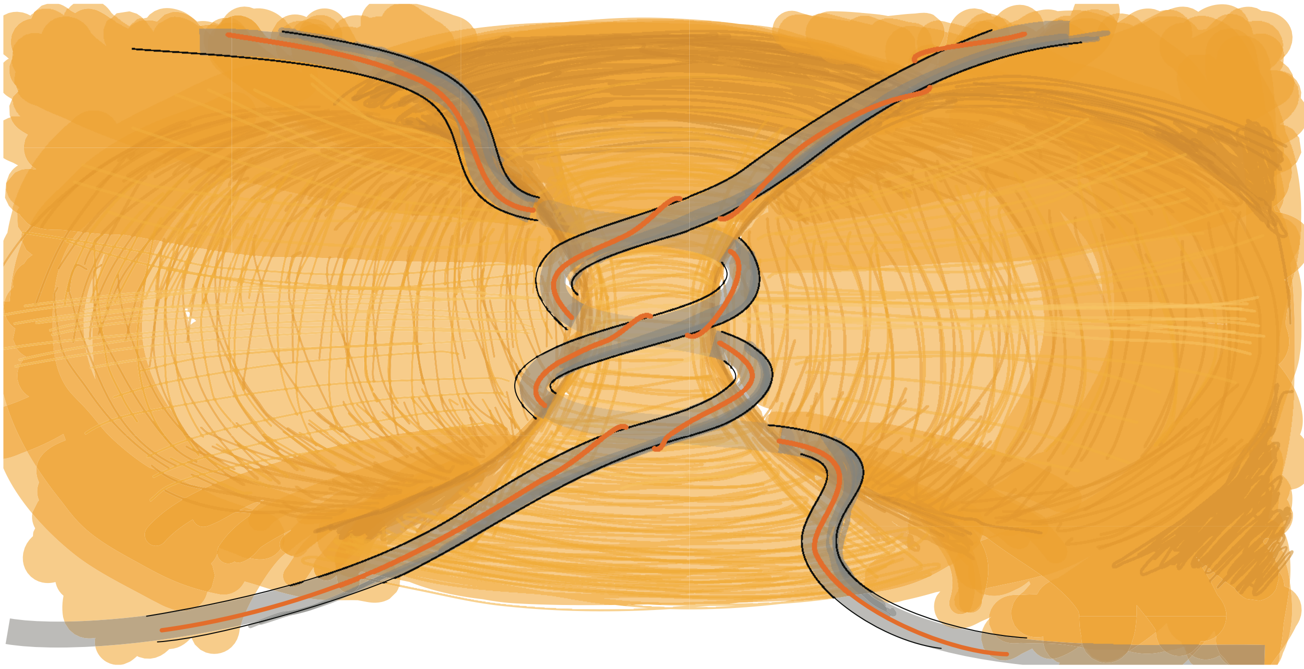
Applications of Instanton Floer Homology



Juanita Pinzón-Caicedo and Daniel Ruberman

1. Introduction

Some thirty-five years ago, Andreas Floer introduced a new tool into geometry: an infinite-dimensional version of Morse theory that is now referred to as Floer homology. The original version of Morse theory computes the homology of a finite-dimensional manifold in terms of critical points of a smooth function and the flow of its gradient vector field. Floer theory does something similar in infinite dimensions, where the critical points and gradient flow equations are now differential geometric objects.

Juanita Pinzón-Caicedo is a professor in the department of mathematics at Notre Dame University. Her email address is `jpinzonc@nd.edu`.

Juanita Pinzón-Caicedo was partially supported by Simons Foundation Collaboration Grant 712377.

Daniel Ruberman is a professor in the department of mathematics at Brandeis University. His email address is `ruberman@brandeis.edu`.

Daniel Ruberman was partially supported by NSF Grant DMS-1811111 and NSF grant DMS-1952755.

DOI: https://doi.org/10.1090/noti2536

The setting for the initial version of Floer homology was in symplectic geometry,[1] but Floer soon understood (following on work of Casson and suggestions of Atiyah and Taubes) that the analytical and geometric tools he had developed were similar to those appearing in gauge theory, and could be used to study 3-dimensional manifolds. The resulting theory is broadly known as Instanton Floer homology. We will not discuss the symplectic version or the closely related Heegaard Floer theory (or the 'monopole Floer homology,' a version that comes from Seiberg-Witten theory), so we will mostly drop the word 'instanton' and refer only to Floer theory.

In this article, we describe a number of applications of this remarkable theory, focusing on two particular areas that have seen much recent activity. The first is the existence of irreducible representations of the fundamental group of 3-manifolds into the simplest non-abelian Lie groups, SU(2) and SO(3). Finding such a representation is an excellent way to show that Dehn surgery on a knot does

[1]*A historical overview by Helmut Hofer of the early days of Floer theory, focusing on its symplectic version, may be found at* `https://www.youtube.com/watch?v=kSNyU71MpgQ`.

not yield a simple manifold such as a sphere or lens space. The second is the use of Floer theory in combination with Yang-Mills gauge theory to study the homology cobordism group of 3-manifolds with the same homology as S^3. The structure of this group is important in low-dimensional topology in the study of the relationship between smooth 4-manifolds and their 3-dimensional boundaries and also in the study of triangulations of high-dimensional manifolds. There are many further applications that we do not have space to mention, including the characterization of the unknot (and other simple knots) via the combinatorially defined Khovanov homology.

2. Morse Theory and Floer Homology

In its most fundamental version, Morse theory gives a recipe for computing the homology groups of a finite-dimensional closed manifold M in terms of some auxiliary data: a real-valued function $f : M \to \mathbb{R}$ and a Riemannian metric g on M. In brief, one assumes that f is a Morse function, meaning that the Hessian matrix of second partial derivatives is non-singular at each critical point. Such critical points are called *non-degenerate*. Each critical point $p \in \mathrm{Crit}(f)$ is assigned an *index* ind_p: the number of negative eigenvalues of the Hessian matrix. The metric turns the differential df into the gradient vector field ∇f, and under some additional transversality assumptions, one can count (with signs) the number of flow lines between critical points with indices differing by one.

There is already a lot to say about the classical theory, so we content ourselves with a brief summary. A downward flow line on a (possibly infinite) interval J is a solution to the ordinary differential equation

$$\dot{\gamma}(t) = -\nabla_{\gamma(t)} f. \tag{1}$$

It is worth keeping in mind that while the set of critical points is determined by f, the gradient flow equation depends on the choice of Riemannian metric.

The counting process already reveals some subtleties. First, a flow line doesn't really 'go from' a critical point x to a critical point y: since $\nabla_x f = 0$, the only flow line that passes through x is the constant solution to (1) given by $\gamma(t) = x$. Rather, we declare a flow line to have (positive and negative) limits y and x if

$$\lim_{t\to\infty} \gamma(t) = y \text{ and } \lim_{t\to-\infty} \gamma(t) = x.$$

In addition, note that for any solution $\gamma(t)$ to (1) and any $c \in \mathbb{R}$, we get a new solution $\gamma(t+c)$ to (1) with the exact same limiting behavior! Let us define an equivalence relation between flow lines γ_1 and γ_2 by saying $\gamma_1 \sim \gamma_2$ if $\gamma_2(t) = \gamma_1(t+c)$ for some $c \in \mathbb{R}$. For any critical points x and y, we therefore define

$$\widetilde{\mathcal{M}}(x,y) = \{\text{flow lines with limits at } x \text{ and } y\}$$
$$\mathcal{M}(x,y) = \widetilde{\mathcal{M}}(x,y)/\sim .$$

In other words, we have an $\mathbb{R}$-action on $\widetilde{\mathcal{M}}(x,y)$, and $\mathcal{M}(x,y)$ is the quotient.

This moduli space has a second interpretation, which leads to a natural topology on $\mathcal{M}(x,y)$. If x is a critical point of index k, then the set of points z such that there is a flow line with limit equal to x as $t \to -\infty$ and passing through z is called the *descending manifold* of x, and is denoted D_x. As the name suggests, D_x is in fact a manifold of dimension k. Similarly, we have the *ascending manifold* A_y consisting of points z with flow lines passing through z and limiting to y as $t \to \infty$.

Using this terminology, we can keep track of flow lines with limits at x and y in the following way. The chain rule, coupled with (1) says

$$\begin{aligned}\frac{d}{dt}(f(\gamma(t))) &= df(\dot{\gamma}(t)) = \langle \nabla(f), \dot{\gamma}\rangle \\ &= -|\nabla(f)|^2 < 0.\end{aligned}$$

This is a familiar fact that we teach in calculus classes: the function f decreases in the direction of the (negative) gradient flow. It follows that if there is a flow line with limits y and x, then $f(x) > f(y)$. Sard's theorem says that there is a regular value $c \in (f(y), f(x))$ for f, which means that the *level set* $M_c = f^{-1}(c)$ is an $(n-1)$-dimensional manifold. Now any flow line from x to y must pass through this level set, exactly once (by the decreasing condition). It follows that

$$\mathcal{M}(x,y) \cong D_x \cap M_c \cap A_y \cong (D_x \cap A_y)/\mathbb{R}, \tag{2}$$

which provides our interpretation of $\mathcal{M}(x,y)$. The structure of $\mathcal{M}(x,y)$ is simplest if we make the assumption that the intersection in (2) is *transverse*: the sum of the tangent spaces to D_x and A_y spans the tangent space to M_c. This transversality property, highlighted by Smale, is called the Morse-Smale condition.

The Morse-Smale condition is remarkably powerful: combined with (2) it implies that $\mathcal{M}(x,y)$ is a smooth manifold of dimension $\mathrm{ind}_x - \mathrm{ind}_y - 1$.

With some additional data, one can consistently assign an orientation to $\mathcal{M}(x,y)$. If $\mathrm{ind}_x = \mathrm{ind}_y + 1$, then $\mathcal{M}(x,y)$ is a finite set of points. Let us write $\#\mathcal{M}(x,y)$ for the algebraic count: each point in $\mathcal{M}(x,y)$ counted as ± 1 depending on its sign.

All of this counting creates the *Morse-Witten chain complex* $(C_*(M,f), \partial)$ where $C_k(M,f)$ is the free abelian group with basis $\{x\}$ corresponding to the critical points of index k. The boundary operator $\partial : C_k(M,f) \to C_{k-1}(M,f)$ is the linear map whose matrix with respect to the critical points x and y of indices k and $k-1$ has (y,x) entry $\#\mathcal{M}(x,y)$. There are several remarkable aspects of the Morse complex: it is indeed a chain complex, or in symbols $\partial^2 = 0$. Moreover, ∂ is the differential of the chain complex associated to a cell decomposition of M, and so its homology is exactly the usual singular homology of M.

It follows from this last statement that the homology of $(C_*(M,f),\partial)$ is independent of the Morse function f and the Riemannian metric used in its definition; it is important for later developments that this can be proven directly by analytical means.

It can certainly happen that a given function has degenerate critical points or that the flow associated to a metric is not Morse-Smale. Fortunately, one can perturb the function and/or the metric to achieve these conditions. The subject of perturbations is a bit technical, and it is best on a first reading to pretend that our fairy godmother has waved a magic wand to make all such issues go away. An interesting byproduct of this discussion is a proof of the Poincaré-Hopf formula. This formula states that the Euler characteristic $\chi(M) = \sum_{k=0}^{\dim M}(-1)^k \dim H_k(M;\mathbb{R})$ may be calculated in terms of the indices of the zeros of ∇f as follows:

$$\chi(M) = \sum_{p\in \mathrm{Crit}(f)} (-1)^{\mathrm{ind}_p}. \tag{3}$$

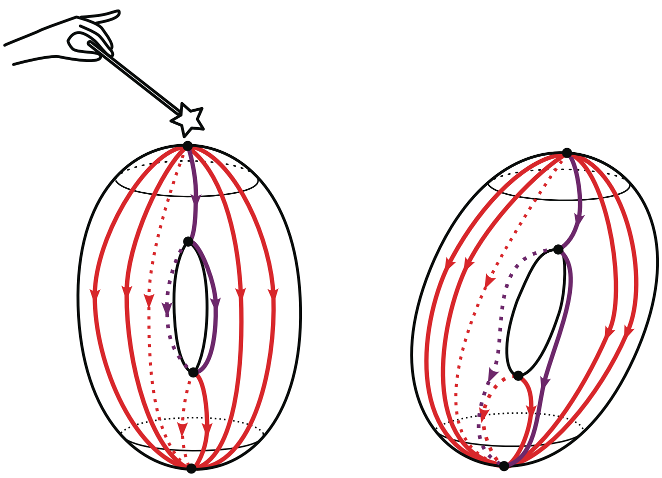


Figure 1. The magic of perturbations.

In a historical parallel with the ordinary Euler characteristic (whose definition preceded the introduction of homology by more than a hundred years!) our path to Floer homology starts with its Euler characteristic, introduced by Casson in a series of MSRI lectures in 1985. Casson worked with an oriented homology sphere Y, and studied the set of representations

$$\alpha : \pi_1(Y) \to \mathrm{SU}(2)$$

up to the equivalence relation of conjugacy. This means that representations α' and α are equivalent if there is $h \in \mathrm{SU}(2)$ with $\alpha'(g) = h^{-1}\alpha(g)h$ for all $g \in \pi_1(Y)$. The set of SU(2) representations of a finitely presented group is a real algebraic variety. Something similar is true for the set of conjugacy classes of representations, referred to as the *character variety* and denoted $\mathcal{R}(Y)$. The biggest caveat is that the trivial representation, while an isolated point in the character variety, is a singular point. The initial workaround is to simply discard this point and look only at non-trivial representations; for homology spheres this is the same as irreducible. But there is a great deal of information in the trivial connection, as we will see in Section 4.

A familiar idea in 3-manifold theory and knot theory is to simply count the number of conjugacy classes of irreducible representations. However, Casson did something deeper by counting them *with signs* defined in terms of a Heegaard splitting of Y so that the count appears as an intersection number. He defined an invariant, which in the non-degenerate case reads

$$\lambda(Y) = \frac{1}{2}\#\Big(\{\text{irreducible } \alpha : \pi_1(Y) \to \mathrm{SU}(2)\}/\text{conj}\Big).$$

The sign of $\lambda(Y)$ changes when the orientation of Y is reversed.

Almost immediately[2] Taubes realized that a more analytic interpretation of representations as 'flat connections' (see below) exhibits them as zeros of a vector field: the gradient of the *Chern-Simons* function. With this interpretation, Taubes [Tau90] was able to write down an expression analogous to (3) and prove that it is equal to Casson's invariant. There are some major challenges stemming from the fact that the vector field whose zeros we want to count is on an infinite-dimensional manifold. In particular, the local index ind_α would seem to be the dimension of an infinite-dimensional space! The issue of perturbations in the infinite-dimensional setting is even trickier, and our fairy godmother will need a better wand. Floer took these ideas even further and defined a homology theory formally similar to Morse theory, whose Euler characteristic is (twice) Casson's invariant. In the next section, we will give some background for the transition from finite-dimensional Morse theory into Floer homology.

2.1. Some basic gauge theory. The analytical setup we need goes under the name of gauge theory. To mathematicians, gauge theory means the study of connections on a vector bundle, up to a natural action of the group of automorphisms of the bundle. This is a highly developed and complicated subject, but we will try to keep it reasonable by focusing on a particularly simple bundle: the trivial bundle $E = Y \times \mathbb{C}^2 \xrightarrow{p} Y$ over a 3-manifold Y where p is projection onto the first factor. Many of the applications discussed below make use of a similar theory associated to a non-trivial $\mathbb{R}^3$ bundle with structure group SO(3); the notation is a bit more complicated in that case.

Recall that the Lie group SU(2) is the set of 2×2 unitary matrices with determinant 1. The unitary condition $AA^* = I$ means that multiplication by A is an isometry of the standard Hermitian metric on $\mathbb{C}^2$. The tangent space to

[2] *Taubes told the authors that this realization came to him while listening to Casson's MSRI lectures.*

SU(2) at the identity matrix is called $\mathfrak{su}(2)$ and is identified with the 2×2 skew-Hermitian matrices with trace 0.

A *connection* on a vector bundle is a way to differentiate sections of the bundle; we will be interested in those that are compatible with the action of the group SU(2) on the $\mathbb{C}^2$ fibers of E. In the case $E = Y \times \mathbb{C}^2$, this notion can be expressed as follows. A section of E is a function $\sigma : Y \to E$ with $p(\sigma(y)) = y$; it may be written in the form $\sigma(y) = (y, s(y))$ for a smooth function $s : Y \to \mathbb{C}^2$. Write $\Gamma(E)$ for the set of sections of E. Now we can certainly differentiate the function s using the usual exterior derivative d to get a 1-form ds, which in turn can be paired with a tangent vector field to get a new section of E.

However, there are many other ways of differentiating sections, all of which can be expressed in the following way. Write $\Omega^k(Y; \mathfrak{su}(2))$ for the $\mathfrak{su}(2)$-valued k-forms. A nice concrete way to think of this is as 2×2 matrices with k-form entries. For any $\alpha \in \Omega^1(Y; \mathfrak{su}(2))$ we get the operator $\nabla_\alpha : \Gamma(E) \to \Gamma(E)$ defined by

$$\nabla_\alpha \sigma(y) = (y, ds + \alpha(s)).$$

In short, $\nabla_\alpha = d + \alpha$. In this way, the space $\mathcal{A}$ of all SU(2) connections is identified with $\Omega^1(Y; \mathfrak{su}(2))$. Note for future reference that this identification depends on a choice of trivialization of the bundle E.

The basic geometric invariant of a connection is its curvature, which is the $\mathfrak{su}(2)$-valued 2-form

$$F_\alpha = d\alpha + \alpha \wedge \alpha.$$

Note that the wedge product here is defined using matrix multiplication and so is not antisymmetric. A connection with curvature $F_\alpha = 0$ is said to be *flat*. Flat connections will be to Floer homology what critical points of a function were to Morse homology, namely, the generators of the chain complex.

To understand the meaning of flatness, let us introduce the idea of parallel transport with respect to a connection α. For a smooth curve $\gamma : I \to Y$, a section $\sigma(t) = (\gamma(t), s(t))$ is *parallel* if $(\nabla_\alpha)_{\gamma'(t)}(\sigma(t)) = 0$. From the point of view of the connection, σ acts as if it is constant. This equation is a linear ODE, to which the usual sort of existence and uniqueness theorems apply. In particular, given $\sigma(0) \in E_{\gamma(0)}$, there is a unique parallel section $\sigma(t)$ along γ with the given initial condition. The assignment $\sigma(0) \to \sigma(1)$ is the parallel transport of $\sigma(0)$ along γ; it is an invertible linear map $P^\gamma_\alpha : E_{\gamma(0)} \to E_{\gamma(1)}$.

Unlike the situation in Euclidean space, parallel transport may depend on the path γ. This is the phenomenon of holonomy. In particular, if γ is a loop, then P^γ_α might not be the identity. Indeed, the curvature at a point y measures the failure of P^γ_α to be the identity for small loops based at y. Fixing a base point $y \in Y$, the set of holonomies of loops based at a is a closed Lie subgroup of SU(2). If the centralizer of the holonomy subgroup associated to α is $\{\pm I\}$, then we say that α is irreducible; otherwise it is reducible.

What is special about a flat connection is that its holonomy is locally trivial. Globalizing this gives us a link between differential geometry and topology.

Proposition 2.1. *Suppose that α is a flat connection, and that γ and $\tilde{\gamma}$ are (piecewise) smooth curves between points a and b. If γ is homotopic to $\tilde{\gamma}$ relative to their endpoints, then $P^\gamma_\alpha = P^{\tilde{\gamma}}_\alpha$. Applied to loops based at a, this means that holonomy gives a well-defined representation $h_\alpha : \pi_1(Y, a) \to$ SU(2).*

Conversely, any given SU(2) representation of $\pi_1(Y)$ is the holonomy representation of some flat connection. However, the 'holonomy correspondence' $\alpha \to h_\alpha$ is far from perfect: there are many flat connections with the same holonomy. This stems from an all-important property of the curvature equation $F_\alpha = 0$, called its *gauge symmetry*. Let us define the gauge group $\mathcal{G}$ to be the group of SU(2) automorphisms of the bundle E. By looking at what an automorphism does to each fiber, we can identify $\mathcal{G}$ with Map$(Y, \mathrm{SU}(2))$. An element $g \in \mathcal{G}$ acts on the space of connections by the rule

$$g^*\alpha = g^{-1}\alpha g + g^{-1}dg.$$

One says that α and $g^*\alpha$ are *gauge equivalent*.

A direct calculation shows that $F_{g^*\alpha} = g^{-1}F_\alpha g$ so that $F_{g^*\alpha} = 0$ if and only if $F_\alpha = 0$. In other words, the equation $F_\alpha = 0$ is *equivariant* with respect to the action of the gauge group.

In order to fit this discussion into the framework of Morse theory, as outlined above, we need a manifold and a function. The action of the gauge group $\mathcal{G}$ on $\mathcal{A}$ (almost) provides the manifold: we define $\mathcal{B}$ to be the quotient $\mathcal{A}/\mathcal{G}$. This would be an infinite-dimensional manifold (this is a bit technical) except that it fails to have the correct structure at some special points–the orbits of reducible connections, at which $\mathcal{G}$ fails to act freely. For Y a homology sphere, the only flat reducible connection is the trivial connection θ.

In the original literature, there are two ways to deal with the singularity arising from θ. One, adopted by Floer (following Casson), is to simply delete the orbit of θ and work on $\mathcal{B}^* = \mathcal{A}^*/\mathcal{G}$, where $\mathcal{A}^*$ denotes the space of irreducible connections. A second is to replace $\mathcal{G}$ by the *based* gauge group $\mathcal{G}^o$ consisting of gauge transformations that are the identity at a fixed base point in Y. The quotient $\mathcal{B}^o = \mathcal{A}/\mathcal{G}^o$ and $\mathcal{B}^*$ are both manifolds. The trivial connection turns out not to be so trivial, and an important part of the current research highlighted in Section 4 is concerned with topological information coming from θ.

Taking account of the gauge group action, we refine Proposition 2.1 to say that holonomy defines a bijection between gauge equivalence classes of flat connections and conjugacy classes of representations. One can go further, and show that the holonomy map is in fact a homeomorphism with respect to the natural topologies carried by each of those sets. The space $\mathrm{Hom}(\pi_1(Y), \mathrm{SU}(2))$ of SU(2) representations of a group with n generators is a closed subset of the compact space $\mathrm{SU}(2)^n$. Hence it is compact, as is its quotient by conjugation, the character variety $\mathcal{R}(Y)$. It follows that the space of flat connections up to gauge equivalence is therefore compact, a surprising and important fact. From now on, we will identify the space of flat connections mod $\mathcal{G}$ with $\mathcal{R}(Y)$.

The relation between flat connections and representations was well-understood for years before the introduction of Floer homology. The new insight was the realization that the flat connections can be seen as the critical points of a gradient vector field on the manifold $\mathcal{B}^*$. The function in question is the *Chern-Simons* function, introduced by Chern and Simons as an odd-dimensional refinement of the Chern-Weil definition of characteristic classes in terms of connections and curvature. To a connection α, they associate the real number

$$\mathrm{CS}(\alpha) = \frac{1}{8\pi^2}\int_Y \mathrm{tr}(\alpha \wedge d\alpha + \frac{2}{3}\alpha \wedge \alpha \wedge \alpha). \qquad (4)$$

How would one calculate the gradient of CS? Since $\mathcal{A}^*$ is locally an affine space, its tangent space at α is identified with the vector space $\Omega^1(Y; \mathfrak{su}(2))$. We compute the differential (or directional derivative) of CS in the direction of $\beta \in \Omega^1(Y; \mathfrak{su}(2))$ as the derivative of $\mathrm{CS}(\alpha + t\beta)$ with respect to t evaluated at $t = 0$; the result is

$$d\,\mathrm{CS}_\alpha(\beta) = \frac{1}{4\pi^2}\int_Y \mathrm{tr}(F_\alpha \wedge \beta). \qquad (5)$$

The vanishing of this integral for *all* β is equivalent to the condition that $F_\alpha = 0$. In other words, the critical points of CS are exactly the flat connections!

To continue with the analogy with finite-dimensional Morse theory, we ought to have isolated critical points, and this means that we should pass to the quotient of $\mathcal{A}^*$ by the gauge group. Doing so brings out an important twist: CS is *not* invariant under gauge transformations, and so does not induce a real-valued function on $\mathcal{B}^*$. A gauge transformation $g : Y \to \mathrm{SU}(2)$ is a map between oriented 3-manifolds, and so it has a degree $\deg(g)$. One then computes

$$\mathrm{CS}(g^*\alpha) = \mathrm{CS}(\alpha) + \deg(g). \qquad (6)$$

Equation (6) says that CS defines a function from $\mathcal{B}^*$ to the circle $\mathbb{R}/\mathbb{Z}$, rather than to the real numbers.

Following this outline of ordinary Morse theory, we need to understand more about gradient flow lines between critical points of CS. To pass from $d\,\mathrm{CS}$ to the gradient $\nabla\,\mathrm{CS}$ requires a metric on $\mathcal{B}^*$. The choice of a Riemannian metric h on Y determines a natural metric on $\mathcal{B}^*$ (called the L^2 metric) for which the gradient of CS is given by

$$\nabla_A(\mathrm{CS}) = -\frac{1}{4\pi^2} * F_A. \qquad (7)$$

The notation $*$ indicates the Hodge star operator determined by h, which for an oriented n-manifold interchanges k and $n-k$ forms. It is defined by $\eta \wedge *\omega = \langle \eta, \omega \rangle \mathrm{vol}_h$. Hence, the (downward) gradient flow equation for CS is given by

$$\frac{dA_t}{dt} - \frac{1}{4\pi^2} * F_{A_t} = 0. \qquad (8)$$

The gradient flow equation can be written in a different way that is crucial to the interaction between Floer theory and gauge theory on 4-manifolds. On an oriented Riemannian 4-manifold X, the $*$ operator is an involution on 2-forms and so the 2-forms split into the ± 1 eigenspaces of $*$, which are known as self-dual and anti-self-dual 2-forms. The dimension of the space of ordinary *closed* self-dual 2-forms is a topological invariant, written as $b_2^+(X)$. In particular, the curvature form F_B of a connection B on a bundle E over a 4-manifold splits as $F_B = F_B^+ + F_B^-$. The self-dual part is given by $F_B^+ = \frac{1}{2}(F_B + *F_B)$; a connection with $F_B^+ = 0$ is anti-self-dual (ASD), and is called an *instanton*. The ASD equation is preserved by gauge transformations, and one can study the *moduli space* of instantons modulo gauge transformations, which we write as $\mathcal{M}(E, h)$.

This moduli space depends on the metric, and one shows that for a generic choice of h, it is a finite-dimensional oriented manifold. In the nicest situation when $\dim(\mathcal{M}(E,h)) = 0$ and $b_2^+(X) > 1$, then the moduli space is compact and following Donaldson [DK90], one can count the points with sign to get the Donaldson invariant $D_X(E) = \#\mathcal{M}(E,h)$. Donaldson's construction is considerably more general and provides an invariant powerful enough to distinguish infinitely many smooth structures on 4-manifolds.

Now suppose we have a path A_t of connections on Y. These can be assembled into a connection B on the trivial bundle over $X = \mathbb{R} \times Y$: at the point (t, y) it is given by $d_4 + A_t$. Here we are distinguishing the exterior derivative d_4 on X from d, the analogous operator on Y. Now the curvature of B is given by

$$F_B = \frac{dA_t}{dt} dt + F_{A_t}$$

and we learn that, up to gauge equivalence, flow lines for the Chern-Simons function correspond to the ASD connections on $\mathbb{R} \times Y$. This formula also gives rise to an alternate expression for CS that works for SO(3) connections and is well-suited to 4-dimensional applications. Any SO(3) bundle over Y extends to a bundle over some oriented 4-manifold X. A connection α will likewise extend to a connection B having the form $d_4 + \alpha$ on a collar neighborhood of the boundary. Then

$$\mathrm{CS}(\alpha) = \frac{1}{8\pi^2}\int_X \mathrm{tr}(F_B \wedge F_B). \tag{9}$$

The connection between dimensions 3 and 4 gave a strong early clue that Floer theory would be intimately tied to the study of Donaldson's invariants and hence the smooth classification of 4-manifolds. For future reference we note that if X is closed, then the integral (9) is an integer and is called the Chern number $c_2(E)$. For SO(3) bundles, this is $-1/4$ times the Pontryagin number $p_1(E)$.

2.2. Floer homology. Treating CS as a Morse function, we want to define a chain complex similar to the Morse complex. We are really reasoning by analogy here; the resulting homology groups are not really computing the homology of a manifold. But surprisingly, those homology groups tell us a lot about the underlying 3-manifold Y. Let us assume that the irreducible flat connections are all non-degenerate and define the instanton chain groups to be the free abelian group generated by the finitely many gauge equivalence classes of irreducible flat connections. The grading presents a challenge: the Hessian at a critical point α is a self-adjoint differential operator that has infinitely many negative (and positive, for that matter) eigenvalues. Hence the definition of index in finite dimensions does not make sense. Floer's resolution of this dilemma is remarkable: the important thing is the *difference* between the indices of two critical points, and this has a sensible definition. Well, almost sensible, as we shall see.

The Morse theory analogy suggests that the boundary operator should be defined by counting points in the space of flow lines (for ∇ CS) between two critical points α and β whose indices differ by one. In contrast to the finite-dimensional case, flow lines do not automatically converge to critical points. To ensure this we assume the 'finite energy' condition

$$\mathcal{E}(B) = \|F_B\|_{Lf^2} = \left(-\frac{1}{2}\int_{\mathbb{R}\times Y} \mathrm{tr}(F_B \wedge *F_B)\right)^{\frac{1}{2}} < \infty.$$

This space is denoted $\widetilde{\mathcal{M}}(\alpha,\beta)$, and is more formally defined by

$$\begin{aligned}\widetilde{\mathcal{M}}(\alpha,\beta) = \Big\{ B = d_4 + A_t \;\Big|\; \frac{dA_t}{dt} - *F_{A_t} = 0,\\ \mathcal{E}(B) < \infty,\ \lim_{t\to-\infty} A_t = \alpha,\ \lim_{t\to\infty} A_t = \beta \Big\} / \mathcal{G}\end{aligned} \tag{10}$$

where $\mathcal{G}$ is the gauge group.

As is the case for moduli spaces on closed manifolds, $\widetilde{\mathcal{M}}(\alpha,\beta)$ is finite-dimensional, and its expected dimension $i(\alpha,\beta)$ can be computed by the Atiyah-Patodi-Singer index theorem. If α and β are non-degenerate, then we may assume after perturbation that $\widetilde{\mathcal{M}}(\alpha,\beta)$ is a smooth manifold of dimension $i(\alpha,\beta)$. As we did in discussing ordinary Morse theory, we note that $\widetilde{\mathcal{M}}(\alpha,\beta)$ has a free action of $\mathbb{R}$ given by reparameterization $t \to t + c$, so that the moduli space of unparameterized flow lines is given by $\mathcal{M}(\alpha,\beta) = \widetilde{\mathcal{M}}(\alpha,\beta)/\mathbb{R}$.

Once again, gauge invariance complicates our analogy: if one changes (say) α by a gauge transformation, then $i(\alpha,\beta)$ changes by a multiple of 8. In other words, we consider the *relative index* $i([\alpha],[\beta])$ as being well-defined mod 8. To lift this to an *absolute* $\mathbb{Z}/8$ grading we make use of the trivial connection θ and set $i([\alpha]) = \dim \mathcal{M}(\alpha,\theta)$.

Floer defined the instanton chain complex $(IC_*(Y),\partial)$ to be the $\mathbb{Z}/8$-graded group generated by the irreducible flat connections. The boundary operator is defined for $[\alpha] \in \mathcal{R}(Y)$ with $i([\alpha]) = k$

$$\partial[\alpha] = \sum_{[\beta] \text{ with } i([\beta])=k-1} \#\mathcal{M}([\alpha],[\beta]) \cdot [\beta]. \tag{11}$$

On a formal level, this is the same definition as the boundary operator for the ordinary Morse complex, and one must establish several important facts.

- There is a perturbation of CS whose critical points are all non-degenerate.
- The 0-dimensional moduli space $\mathcal{M}(\alpha,\beta)$ is compact and oriented.
- $(IC_*(Y),\partial)$ is a complex, i.e., $\partial^2 = 0$.
- The resulting homology groups $I_*(Y) = H(C_*(Y),\partial)$ are independent of the choices of metric and perturbation.

Finding a perturbation of CS is a delicate matter; the idea is to use a gauge invariant function on the space of connections defined by a loop in γ in Y. Roughly speaking, for each connection α, one takes the trace of the holonomy of α around γ, and adding a linear combination of such functions (a *holonomy perturbation*) to CS will make it behave like a Morse function.

The proof of independence from all choices involves a fundamental construction that connects Floer theory with 4-dimensional topology, and underlies many of the applications discussed below. Suppose that Y_0 and Y_1 are oriented homology spheres, and that W is an oriented manifold with $\partial W = -Y_0 \cup Y_1$. We say that W is a cobordism from Y_0 to Y_1, and for technical reasons introduce a noncompact manifold $\widehat{W} = W \cup_{\partial W} \partial W \times [0,\infty)$. With some assumptions on the topology of W, there is an induced map $\Phi_W : I_*(Y_0) \to I_*(Y_1)$ given, roughly speaking, by counting finite energy instantons on $\widehat{W}$ with specified flat limits on Y_0 and Y_1.

The whole construction is functorial in the sense that if $W = W_0 \cup_{Y_1} W_1$, then $\Phi_W = \Phi_{W_1} \circ \Phi_{W_0}$. The composition law is an instance of a *gluing theorem* for instantons.

After considerable work the gluing theorem yields the following relation between instanton Floer homology and the Donaldson invariants of closed 4-manifolds.

Theorem 2.2. *If $X = X_1 \cup_Y X_2$ where $b_2^+(X_i) > 0$ and $D_X \neq 0$, then $I_*(Y) \neq 0$.*

There is a largely parallel theory involving the Seiberg-Witten equations in place of the ASD equations that produces an invariant SW_X, and a Floer-type *monopole Floer homology theory* $HM(Y)$, with a similar non-vanishing result.

3. Instantons and $SU(2)$ Representations

3.1. Surgery and the pillowcase.

A key tool in Floer homology is the relationship between flat connections on a 3-manifold Y and connections on a manifold constructed by surgery on a knot K in Y. A tubular neighborhood $\nu(K)$ is of the form $S^1 \times D^2$, and we choose a non-separating simple closed curve c on $\partial\nu(K)$, called the surgery slope. Gluing $D^2 \times S^1$ to $Y - \nu(K)$ in such a way that ∂D^2 is identified with c yields a manifold $Y_c(K)$. In the special case that Y is S^3, the slope is often denoted by p/q whenever $c = p\mu + q\lambda$, with μ and λ the meridian and longitude of K.

The Seifert-van Kampen theorem says that flat connections on $Y_c(K)$ are given by the SU(2) representations of $\pi_1(Y - \nu(K))$ that extend over $D^2 \times S^1$. The extension condition is precisely that the holonomy around c is trivial, and the dependence of this condition on the slope c can be visualized via the following picture, affectionately known in the trade as the *pillowcase*. The pillowcase itself parameterizes SU(2) representations of $\pi_1(\partial\nu(K))$, which are determined by the holonomies of the meridian and longitude. The red curve in Figure 2 indicates the holonomies for the irreducible representations of $\pi_1(Y - \nu(K))$, where K in this case is the *left-handed* trefoil knot and the surgery slope is -1. The 3-manifold $S^3_{-1}(K)$ is the Poincaré homology sphere $\mathbb{P}$, which can also be described as the quotient S^3/I^* where I^* is the binary icosahedral group.

The blue curve in Figure 2 shows the representations that extend over D^2 and is therefore determined by c. In this case, we see that there are two such representations, say ρ_1 and ρ_2, corresponding to the intersection of the red and blue curves. The Floer grading of these representations can be computed using a formula described by Fintushel and Stern and we get $\mathrm{grad}(\rho_1) = 1$, $\mathrm{grad}(\rho_2) = 5$. Since the difference of these gradings is not 1, the differentials of the instanton complex are 0, proving that

$$I_*(\mathbb{P}) = \mathbb{Z} \text{ for } * = 1, 5 \text{ and } 0 \text{ otherwise.} \qquad (12)$$

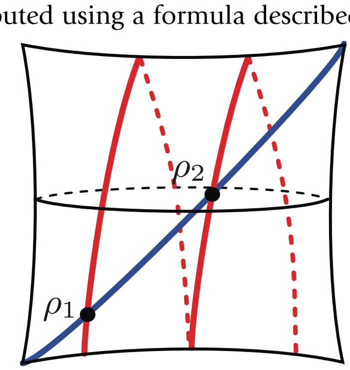


Figure 2. The pillowcase for the left-handed trefoil.

A powerful tool closely related to this picture is Floer's surgery exact triangle [BD95], an exact sequence of Floer homology groups associated to three manifolds realized as surgery along a knot K in a homology sphere Y as follows:

$$\longrightarrow I_{k+1}(Y_0(K)) \longrightarrow I_k(Y) \longrightarrow I_k(Y_1(K)) \longrightarrow$$

A point to note is that $H_1(Y_0(K)) = \mathbb{Z}$ and so $Y_0(K)$ is not a homology sphere. Its instanton homology is defined as above, but now using flat connections on a non-trivial SO(3) bundle. There is an analogous surgery triangle associated to surgery curves that meet in a similar configuration on the boundary torus of the knot as in Figure 3. Casson's surgery formula for his invariant, which was surely an inspiration for this exact sequence, can be derived from it in retrospect.

The grey triangle in Figure 3 is the key to Floer's proof (beautifully exposed in [BD95]) of exactness of the surgery triangle. The curves in the pillowcase corresponding to the surgery curves also meet in a triangle, and Floer's idea was to use holonomy perturbations to deform one side of this triangle to the other two, in order to "simulate the effect of surgery."

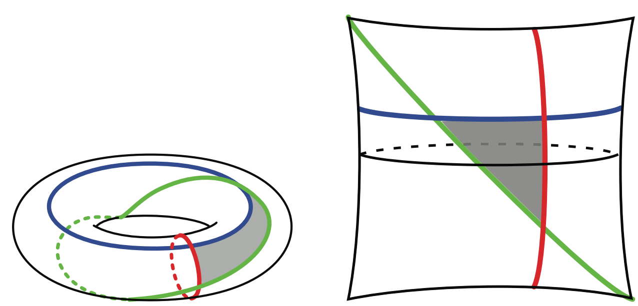


Figure 3. Surgery slopes in the torus and pillowcase.

3.2. Casson's invariant and Floer homology.

From the beginning of topology in the early 20^{th} century, it was understood that a fruitful way to study a 3-manifold was to investigate the representations of its fundamental group into a non-abelian group, often a group of matrices. For instance, one can readily show that the trefoil is knotted by finding a homomorphism from its fundamental group into SO(3) with non-abelian image. Let us restrict our attention to homology spheres, and pose a basic question:

> For a homology sphere Y, is there a non-abelian (and hence irreducible) representation $\pi_1(Y) \to SU(2)$?

Kronheimer and Mrowka [KM04b] proved that this holds when Y is surgery on a non-trivial knot in S^3, establishing the famous Property P conjecture (that such a Y is not simply connected).

The very definitions of the Casson invariant and Floer homology yield the following observation.

Proposition 3.1. *If $\lambda(Y) \neq 0$, or if $I_*(Y)$ is not zero, then there must be irreducible* SU(2) *representations of $\pi_1(Y)$.*

When Y is given by Dehn surgery on a knot K, Casson gave a formula for computing $\lambda(Y)$ in terms of the Alexander polynomial of K. If we normalize that polynomial so that $\Delta_K(t) = \Delta_K(t^{-1})$, then Casson's result says that if $\Delta_K''(1) \neq 0$, then $\lambda(S^3_{1/n}(K)) \neq 0$, and hence there are irreducible SU(2) representations. For example, for $\mathbb{P}$ realized as -1 surgery on the left-handed trefoil knot this formula predicts at least two non-conjugate representations. As we saw in Figure 2, this is in fact the case.

For many homology spheres λ vanishes, and hence is of no use in finding representations. A salutary calculation is $\lambda(Y\#-Y) = \lambda(Y)+\lambda(-Y) = \lambda(Y)-\lambda(Y) = 0$. But if $\pi_1(Y)$ has SU(2) representations, then so does $\pi_1(Y\# - Y)$ even though this is invisible to Casson's invariant. Nevertheless, the existence of SU(2) representations can sometimes be detected by I_* as we will see in the following subsection.

3.3. Existence of SU(2)-representations.

For some manifolds, such as Seifert fibered homology spheres, one can directly find all of the SU(2) representations and compute I_* from its definition. For our purposes, a 3-manifold Y is Seifert fibered if it admits a circle action where some circle orbits have finite (cyclic) isotropy. If Y is a homology sphere, it is determined by the orders of the isotropy groups, and Y is denoted $\Sigma(a_1, \dots, a_n)$. The Poincaré homology sphere $\mathbb{P}$ is the Seifert fibered manifold $\Sigma(2,3,5)$, whose instanton homology is described in Equation (12).

In general, to make use of Proposition 3.1 we need techniques to calculate I_* without first counting all of the representations. The first proof of Property P by Kronheimer and Mrowka [KM04b] showed that $I_*(S^3_0(K))$ is non-zero if K is non-trivial, and then applied the Floer exact triangle. The non-vanishing of $I_*(S^3_0(K))$ is proved using a dizzying array of important results from the last 35 years of low-dimensional topology, in conjunction with the gluing theory presented as Theorem 2.2. Leaving out many details, the argument is summarized in this diagram.

$K \neq$ Unknot $\Longrightarrow$ $S^3_0(K)$ has a taut foliation $\Longrightarrow$ $S^3_0(K)$ has a tight contact structure $\Longrightarrow$

$S^3_0(K) \subset X$, a symplectic manifold $\Longrightarrow$ $\mathrm{SW}_X \neq 0 \Longrightarrow D_X \neq 0$

A second proof appears in [KM04a] showing that Dehn surgery on a knot with slope in the interval $[-2,2]$ (including those which are *rational* homology spheres) has irreducible SU(2) representations. A third argument for the non-vanishing of Floer homology, avoiding the connection between the Seiberg-Witten and Donaldson invariants, was given by Kronheimer and Mrowka in [KM10]. This version uses the Chern-Simons functional for $SO(3)$-connections on a non-trivial bundle to define invariants $I(Y|R)_w$ as subspaces of $I(Y)$ that depend on a surface R and a class $w \in H^2(Y;\mathbb{Z})$. A special case of this construction is *framed instanton homology* $I^\#(Y)$, a certain subspace of $I(Y\#T^3)$. This theory was used by Baldwin-Sivek [BS18] to provide a criterion for the existence of $SU(2)$-representations for integer homology spheres much in the spirit of Proposition 3.1. Namely, they show that if $rk(I^\#(Y)) > 1$, then irreducible SU(2) representations of $\pi_1(Y)$ exist. With this in place, Baldwin-Sivek prove that being the boundary of a Stein 4-manifold (a complex surface with a convexity condition at the boundary) with nontrivial reduced homology guarantees the existence of irreducible representations.

Recent work of Zentner [Zen18] shows that the fundamental group of any homology sphere (well, except S^3) admits an irreducible $\mathrm{SL}(2,\mathbb{C})$ representation. A key step in his proof uses holonomy perturbations directly (rather than via the Floer triangle) to show that the splicing of two non-trivial knot complements (the complements glued along their boundary tori, with meridians glued to longitudes) in fact has an irreducible SU(2) representation. Combining arguments about holonomy perturbations with the Floer triangle, Lidman–Pinzón-Caicedo–Zentner [LPZ21] show the same for any homology sphere that contains an incompressible torus. This fact is now a corollary of a much stronger result of Baldwin-Sivek [BS21] establishing the non-triviality of $I^\#$ for any homology sphere obtained by gluing the complements of two nontrivial knots in homology spheres Y_1, Y_2 that are instanton L-spaces (i.e., satisfy the technical condition $\dim(I^\#(Y_i)) = 1$ for $i = 1, 2$).

The results covered by [KM04a] about the existence of irreducible representations for surgery $S^3_r(K)$ along a nontrivial knot K have been extended to include the following cases:

(i) either $r = p/q$ or $r = -p/q$ [Lin16],
(ii) $|r|$ sufficiently large [SZ17], and
(iii) $S^3_r(K)$ for infinitely many values of r in $[3,5)$, always including $S^3_3(K)$ and $S^3_4(K)$ [BLSY21].

4. Chern-Simons and Homology Cobordism

Our second suite of applications is concerned with the interaction between Floer theory and 4-manifolds. A common theme will be the use of the Floer chain complex and filtrations that are defined on it via the Chern-Simons function. We start by recalling a fundamental question in low-dimensional topology:

> Which unimodular symmetric bilinear forms over the integers arise as the intersection form of a closed (smooth) 4-manifold X? Equivalently, what are the possible homotopy types of simply connected closed smooth 4-manifolds?

In 1952, Rokhlin showed that if X is spin (so all self-intersections of embedded surfaces are even) then the signature of X is divisible by 16. This was the first indication that something special is going on in dimension 4.

One can show that any unimodular form is the intersection form of a smooth 4-manifold with boundary a homology sphere Y; the problem above then comes down to understanding which homology 3-spheres are the boundary of a homology 4-ball. Thus we are led to study oriented homology spheres, modulo the relation of homology cobordism: Y_0 and Y_1 are homology cobordant if there is a manifold W with the homology of $I \times S^3$ and with $\partial W = -Y_0 \cup Y_1$. The set of equivalence classes form the homology cobordism group $\Theta^3_{\mathbb{Z}}$. Rokhlin's theorem gives rise to a surjective homomorphism $\mu : \Theta^3_{\mathbb{Z}} \to \mathbb{Z}/2$ given by $\mu(Y) = \frac{1}{8}\operatorname{sign}(W) \pmod 2$ where W is a spin manifold with $\partial W = Y$; for instance $\mathbb{P}$ bounds a manifold with intersection form the definite form E_8 with signature -8, so $\mu(\mathbb{P}) = 1$ and $\mathbb{P} \neq 0 \in \Theta^3_{\mathbb{Z}}$. Using a version of Seiberg-Witten Floer homology, Manolescu showed that the map μ does not split. Combined with deep work of Galewski-Stern and Matsumoto, this proves that there are non-triangulable topological manifolds in dimensions at least 5 that are not homeomorphic to a simplicial complex! Casson previously showed this in dimension 4, using the fact that $\mu(Y) \equiv \lambda(Y) \pmod 2$.

Thirty years after Rokhlin, Donaldson brought gauge theory into topology by showing that any definite form that is realized is diagonalizable. Donaldson's proof is based on an analysis of a particular 5-dimensional moduli space $\mathcal{M}$ over a smooth 4-manifold X, including its reducible points and its compactness properties. Donaldson's theorem immediately implies something stronger about $\mathbb{P}$: it is of infinite order in $\Theta^3_{\mathbb{Z}}$, since the multiple direct sum of E_8 with itself is not diagonalizable.

A key aspect of Donaldson's argument is to determine precisely how this moduli space $\mathcal{M}$ over X fails to be compact: for any point p in the manifold, there is a sequence of ASD connections whose curvature 2-forms concentrate at p. This phenomenon, called 'bubbling', was understood first in fundamental work of Uhlenbeck and Taubes. In particular, such concentration at points requires a minimum 'quantum' of energy given by the Chern number of the simplest non-trivial bundle P over S^4, so that $c_2(P) \neq 0$. Donaldson refined this and showed how to compactify $\mathcal{M}$ by adjoining a copy of X corresponding to the bubbling points. Another crucial aspect of the moduli space for definite manifolds is that there are reducible connections, so that $\mathcal{M}$ has singularities each modeled on a cone on $\mathbb{CP}^2$. Cutting out a neighborhood of these cones gives a compact 5-manifold that can't exist for homological reasons unless the original intersection form was diagonalizable. A few years later, Fintushel and Stern gave a version of this argument for certain intersection forms that used SO(3) connections to avoid Donaldson's compactification theorem. Using SO(3) connections allowed them to work with a bundle E with $p_1(E) < 4$, which is not enough energy to allow bubbling.

The proof that $\Theta^3_{\mathbb{Z}}$ is infinite can be greatly strengthened by combining Donaldson's analysis with an understanding of the the Chern-Simons functional. We will describe this, and then sketch how one gets much stronger results using ideas from Floer theory.

4.1. Definite manifolds with boundary.

Furuta and Fintushel–Stern generalized Donaldson's argument about definite closed manifolds to the setting of certain manifolds with boundary, proving the remarkable result that $\Theta^3_{\mathbb{Z}}$ is infinitely generated. (Recent work of Dai, Hom, Stoffregen, and Truong, using Heegaard Floer theory, sharpens this to provide a $\mathbb{Z}^\infty$ summand of $\Theta^3_{\mathbb{Z}}$.) We will briefly explain the ideas and some of their descendants, and then discuss how these interact with Floer theory in some very recent work of Daemi and Nozaki-Sato-Taniguchi.

The setting is a bundle E over a manifold X with cylindrical ends $[0,\infty) \times Y_i$ $i = 1, \ldots, m$, where one has specified limiting flat connections α_i in each end. We will refer to the data $(E, \alpha = \{\alpha_i\})$ as an *adapted bundle*. Crucially, a non-empty boundary introduces a new source of potential non-compactness called breaking. Instead of curvature accumulating near a point, one can have on a given end $[0,\infty) \times Y$, a sequence $\{A_n\}_n$ of ASD connections whose restrictions to that end satisfy

(i) the curvature F_{A_n} concentrates on a tubular portion $(L_n, \infty) \times Y$ where $L_n \to \infty$ as $n \to \infty$, and
(ii) $\{A_n|_{\{t\}\times Y}\}$ is close to a flat connection β for t near L_n and as $t \to \infty$ approaches the fixed connection α.

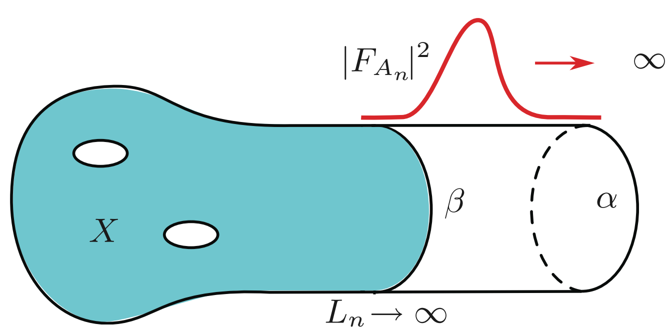


Figure 4. Breaking on an end of X.

To have breaking again requires a quantum of energy that may be much less than the amount required to concentrate at a point.

A key observation is that the Chern-Simons function gives a lower bound for the energy that could allow for breaking. In the setting of SO(3) connections where the CS function is well-defined mod 4, its minimum value on the set of flat connections defines an invariant $\tau(Y) \in [0,4)$. If one works on an adapted bundle with trivial connections α_i at each end and p_1 less than the minimum of all of the $\tau(Y_i)$, then one gets a compact moduli space.

Finding such an adapted bundle is not easy–here is a situation where one can do so. Suppose that Σ is a Seifert fibered homology sphere $\Sigma(a_1, a_2, \dots, a_m)$. This means that Σ supports a circle action where the cyclic subgroups $\mathbb{Z}/a_i$ have fixed points. The quotient Σ/S^1 is topologically a 2-sphere, and the mapping cone C of the projection $\Sigma \to S^2$ is an orbifold: a space with some singularities modeled on a cone over a lens space. Fintushel–Stern showed how to do gauge theory on such orbifolds, and applied their results to study homology cobordism. An important part of their work was an index calculation that gives the dimension of a moduli space over such an orbifold in terms of an invariant $R(a_1, a_2, \dots, a_m)$.

Furuta [Fur90] used this method to great advantage, showing that $\Theta^3_{\mathbb{Z}}$ contains a $\mathbb{Z}^\infty$ subgroup. Combining Furuta's results with work of Fintushel–Stern, and Hedden–Kirk [HK11] yields the following.

Theorem 4.1. *Let $\Sigma = \Sigma(a_1, a_2, \dots, a_m)$ be a Seifert fibered homology sphere. If $R(a_1, a_2, \dots, a_m) > 0$ and*

$$\frac{|H_1(\Sigma;\mathbb{Z})|}{a_1 \cdot a_2 \cdot \dots \cdot a_m} < \min\Big\{\frac{1}{|H_1(\Sigma;\mathbb{Z})|}, \frac{1}{a_1}, \dots, \frac{1}{a_m}, \tau(\pm Y_1), \dots, \tau(\pm Y_N)\Big\},$$

then the class of Σ in $\Theta^3_{\mathbb{Z}/2}$ does not contain any linear combination of elements of $\{Y_i\}_{i=1}^N$.

The conditions in the theorem are used as follows: A relation in $\mathbb{Z}/2$ homology cobordism is the same as a 4-manifold X with $H_1(X;\mathbb{Z}/2) = H_2(X;\mathbb{Z}/2) = 0$ and with boundary Σ plus some number of copies of the Y_i. The orbifold obtained by gluing C and X along Σ then admits an adapted bundle (E, θ) with trivial connections over each end corresponding to a Y_i. The associated moduli space $\mathcal{M} = \mathcal{M}(E,\theta)$ can be shown to be compact using the given bounds on the $\tau(Y_i)$. As in Donaldson's work, this moduli space will have a number of singular points corresponding to reducible connections; removing these singular points gives rise to a compact manifold with boundary a union of complex projective spaces $\mathbb{CP}^n$ that cannot exist by a homological argument.

Theorem 4.1 obstructs not just the existence of $\mathbb{Z}/2$-homology cobordisms, but rather the more general class of 4-manifolds X with $H_1(X;\mathbb{Z}/2) = 0$ and $b_2^+(X) = 0$. This is a really useful extension; a typical application is to knot concordance. Two knots K_0 and K_1 are concordant if they cobound a cylinder in $I \times S^3$, and the set of concordance classes form the much-studied concordance group $\mathcal{C}$. The double branched cover of $I \times S^3$ branched along a concordance is a $\mathbb{Z}/2$-homology cobordism between $\mathbb{Z}/2$ homology spheres. This passage to the double cover gives a homomorphism from the concordance group to $\Theta^3_{\mathbb{Z}/2}$. Hedden and Pinzón-Caicedo [HPC21] used this result on $\mathbb{Z}/2$ homology cobordism to provide a criterion for a satellite operation to have infinite rank as a function on smooth concordance.

4.2. The non-trivial trivial connection. The results above use the generators of the instanton chain complex and some facts about ASD moduli spaces, but largely ignore the differentials. The remaining two homology cobordism invariants r_Y and Γ_Y featured in the rest of our story use similar compactness arguments, but incorporate the differentials in their definition. In fact they go considerably further by also considering the trivial connection.

Many of the ideas originate in work of Frøyshov [Frø02], who used the interaction with the trivial connection to define a surjective homomorphism $h : \Theta^3_{\mathbb{Z}} \to \mathbb{Z}$. This map is defined in terms of two homomorphisms at the chain level. The first of these is $u : C_i \to C_{i-4}$ and is defined as the count of elements in a codimension-4 submanifold $\mathcal{N}(\alpha,\beta)$ of $\mathcal{M}(\alpha,\beta)$. To define the second one, recall first that the Floer chain differential $\partial : C_i \to C_{i-1}$ is defined by counting instantons over the tube $\mathbb{R} \times Y$ that limit asymptotically to flat irreducible connections. Even though the trivial connection θ is reducible, it is still possible to apply the same idea to define a homomorphism

$$D_1 : C_1 \to \mathbb{Z}$$

that counts instantons over the tube $\mathbb{R} \times Y$ that limit to θ in the outgoing end. Frøyshov's homomorphism can be interpreted in terms of the set $\mathcal{L}_{4k-3}$ consisting of elements $a \in C_{4k-3}(Y)$ that are cycles, and such that

- $D_1 u^{k-1} a \neq 0$
- $D_1 u^j a = 0$ for any $0 \leq j < k-1$

In words, $a \in \mathcal{L}_{4k-3}$ means that it can be joined to θ by a 'broken flow line' consisting of exactly $(k-1)$ u-flows and one linear combination of honest flow lines. A result of Frøyshov states that $h(Y) > k$ if and only if the set $\mathcal{L}_{4k-3}$ is non-empty. We now describe two homology cobordism invariants that generalize h, obtained using the homomorphism D_1. The underlying idea is that in addition to gradings, Morse homology groups admit a filtration induced by the Morse function itself. One can then extract useful invariants from the first filtration level at which some cycle representing a homology class appears. In order to do that here, we must revisit certain technicalities of CS. Recall that the Chern-Simons function is $\mathbb{R}$-valued over $\mathcal{A}$, but only $\mathbb{R}/\mathbb{Z}$-valued after considering the action of the gauge group since $\mathrm{CS}(g^*\alpha) = \mathrm{CS}(\alpha) + \deg(g)$ as in Equation (6). To produce a function that is both $\mathbb{R}$-valued and takes into consideration the action of the gauge group (at least partially), let $\mathcal{G}_0$ be the subgroup of $\mathcal{G}$ consisting of the gauge transformations that have degree 0 (not to be confused with $\mathcal{G}^0$ from Section 2.1). Then, setting $\mathcal{B}_\infty = \mathcal{A}/\mathcal{G}_0$ results in an $\mathbb{R}$-valued Chern-Simons function $\widetilde{\mathrm{CS}} : \mathcal{B}_\infty \to \mathbb{R}$.

First we describe an invariant defined by Nozaki-Sato-Taniguchi [NST19] that comes in the shape of a function $r(Y) : [-\infty, 0] \to [0, \infty]$ associated to a $\mathbb{Z}$-homology sphere Y. Both the argument and the output of the function depend on a filtration of instanton homology defined by restricting the image of $\widetilde{\mathrm{CS}}$ to intervals $[s,r] \subset \mathbb{R}$ that contain 0. The filtered instanton groups are obtained by restricting the image of the latter function to intervals of $\mathbb{R}$ as in [FS92]. More precisely, for r, s real numbers with $s \le 0 \le r$ neither of which is a critical value of $\widetilde{\mathrm{CS}}$, the chain group $C_k^{[s,r]}$ is the $\mathbb{Z}$-module generated by the set of flat irreducible elements a of $\mathcal{B}_\infty$ such that $\widetilde{\mathrm{CS}}(a) \in (s,r)$ and $\dim \mathcal{M}(a, \theta_0) = k$, where θ_0 denotes the lift of $[\theta]$ to $\mathcal{B}_\infty$ satisfying $\widetilde{\mathrm{CS}}(\theta_0) = 0$. If s is a critical value, we consider values of $\widetilde{\mathrm{CS}}$ that are an ϵ smaller than s, with $\epsilon > 0$ small enough so that $s - \epsilon$ is bigger than the next critical value of $\widetilde{\mathrm{CS}}$. The differential $\partial^{[s,r]}$ is simply the restriction of ∂ to the filtered chain groups $C_k^{[s,r]}$.

Definition 4.2 ([NST19]). For $s \in [-\infty, 0]$, the value $r_s(Y)$ is defined as

$$\inf\{r \in [0,\infty] \mid \exists\, a \in C_*^{[s,r]} \text{ such that } \partial^{[s,r]} a = 0, \text{ but } D_1(a) \neq 0\}.$$

That is, $r_s(Y)$ is the smallest positive real number for which the count of flow lines from some cycle to the trivial connection is non-trivial.

A related invariant was defined by Daemi [Dae20], associating a function $\Gamma_Y : \mathbb{Z} \to \mathbb{R}$ to a $\mathbb{Z}$-homology sphere Y. The output of the function again depends on values of $\widetilde{\mathrm{CS}}$ at irreducible flat connections, and the input is the absolute $\mathbb{Z}$-grading of the critical points of $\widetilde{\mathrm{CS}}$.

Definition 4.3 ([Dae20]). For a positive integer k, the value of $\Gamma_Y(k)$ is defined as

$$\inf\{\mathrm{val}(a) \mid a \in \mathcal{L}_{4k-3}\}.$$

The real number $\mathrm{val}(a)$ is the minimum value that $\widetilde{\mathrm{CS}}$ takes on the set of flat connections that appear in an expression for a as a linear combination.

A first comparison of the two invariants is given by the equality

$$r_{-\infty}(Y) = \Gamma_{-Y}(1).$$

This relation follows almost immediately from the definitions of the invariants after noticing that $\partial^{[-\infty,r]} a = \partial a$ whenever $\widetilde{\mathrm{CS}}(a) < r$. The change in orientation is due to differing conventions used by each set of authors. And, to reiterate, if $r_s(Y) < \infty$ or if $\Gamma_Y(k) \neq 0, \infty$, the function takes values in the image under $\widetilde{\mathrm{CS}}$ of flat connections.

To illustrate the behavior of the invariants, we present computations for the Poincaré homology sphere with both of its orientations. The diagrams in Table 1 represent the portion of the $\mathbb{Z}$-graded instanton complex for $\pm\mathbb{P}$ with filtrations in the interval $[0,1]$. The downward pointing arrows represent the maps u and D_1, and the horizontal and dashed arrows represent the values of the $\widetilde{\mathrm{CS}}$ function at the flat connections from Figure 2. The absence of an arrow or a label means that the maps are zero. The values of $\Gamma_{\pm\mathbb{P}}$ and $r_s(\mp\mathbb{P})$ can then be computed directly from Definitions 4.2 and 4.3.

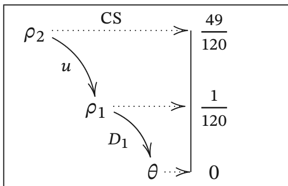
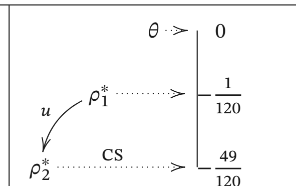
	
$\mathrm{grad}(\rho_1) = 1$ $\mathrm{grad}(\rho_2) = 5$	$\mathrm{grad}(\rho_1^*) = 4$ $\mathrm{grad}(\rho_2^*) = 0$
$D_1(\rho_2) = 0$ $D_1 u \rho_2 = D_1 \rho_1 \neq 0$ $D_1^{[s,r]} \equiv 0, r < \frac{1}{120}$	$D_1 \equiv 0$
$\Gamma_{\mathbb{P}}(k) = \begin{cases} \infty & k > 2 \\ \frac{49}{120} & k = 2 \\ \frac{1}{120} & k = 1 \end{cases}$	$\Gamma_{-\mathbb{P}}(k) = \infty$ for $k > 0$
$r_s(-\mathbb{P}) = \frac{1}{120}$	$r_s(\mathbb{P}) = \infty$

Table 1. Computations of the invariants Γ and r_s for the Poincaré homology sphere with both orientations.

Here are some sample applications of Γ and r_s to both the existence of irreducible representations, and to homology cobordism.

Theorem 4.4 ([NST19]). *Linear Independence: If Y_n is a sequence of integer homology spheres such that $r_0(Y_{n+1}) < r_0(Y_n) < \infty$ and $r_0(-Y_{n+1}) = r_0(-Y_n) = \infty$, then the Y_n are linearly independent in $\Theta^3_{\mathbb{Z}}$.*

Special family: for any $k > 0$, the homology cobordism class of $2\mathbb{P}\# - \Sigma(2,3,6k+5)$ does not bound any definite smooth 4-manifold. Moreover, it has no Seifert fibered representative nor a representative that is surgery on a knot in S^3.

In discussing homology cobordisms in section 3, we did not mention much about their fundamental groups. A striking observation of Akbulut (based on work of Taubes) is that for any homology cobordism W from $\mathbb{P}$ there is an

Daemi	NST
$\Gamma_Y : \mathbb{Z} \to \mathbb{R}$	$r(Y) : [-\infty, 0] \to [0, \infty]$
Monotonicity	
$\Gamma_Y(k) \leq \Gamma_Y(k+1)$	$s_1 \leq s_2 \leq 0$, then $r_{s_1}(Y) \geq r_{s_2}(Y)$
Behaviour under cobordisms $W : Y_0 \to Y_1$, $b_2^+(W) = 0$, $H^1(W;\mathbb{R}) = 0$	
$\Gamma_{Y_0}(k) \geq \Gamma_{Y_1}(k)$ for every $k \in \mathbb{Z}_{>0}$	$r_s(Y_0) \leq r_s(Y_1)$ for every $s \in [-\infty, 0]$
Relationship with minimum CS	
$\Gamma_Y(1) \geq \tau(Y)$	$r_0(Y) \geq \tau(Y)$
Summands of $\Theta^3_{\mathbb{Z}}$	
$\mathbf{\Gamma} : \Theta^3_{\mathbb{Z}} \to \oplus_{k\in\mathbb{Z}}\mathbb{R}$ $Y \to (\Gamma_Y(k))_{k\geq 0}$	$\mathbf{r} : \Theta^3_{\mathbb{Z}} \to \oplus_{s\in\mathbb{R}_{\leq 0}}\mathbb{R}$ $Y \to (r_s(Y))_{s\in\mathbb{R}_{\leq 0}}$

Table 2. Properties of the invariants Γ and r_s.

irreducible SU(2) representation of $\pi_1(W)$. The Γ-invariant extends results like this much further.

Theorem 4.5 ([Dae20]). *For a negative definite cobordism $W : Y_0 \to Y_1$, there exists a constant $\eta(W) \geq 0$ such that*

- $\Gamma_{Y_1}(k) \leq \Gamma_{Y_0}(k) - \eta(W)$ *for* $k > 0$.
- $\eta(W) = 0$ *if and only if there exists a representation* $\rho : \pi_1(W) \to SU(2)$ *that restricts to irreducible representations for both* Y_0 *and* Y_1.

5. Closing Remarks

There is much more to the story of Instanton Floer homology! There are important variations of the theory, such as a relative version for knots in a homology sphere [KM10] and even a version for trivalent graphs introduced by Kronheimer-Mrowka. This latter suggests a surprising approach to the 4-color theorem using methods of Floer theory.

Counting flow lines to the trivial connection as discussed in Section 4.2 can be viewed as part of a construction of Floer theory as an SO(3)-equivariant Morse theory. A construction of such an equivariant theory was given by Miller Eismeier and was recently proved to be an invariant by Daemi-Miller Eismeier; a version for knot Floer homology was given by Daemi-Scaduto. The resulting theories have the same involved algebraic structures found in the parallel Heegaard Floer and Monopole Floer theories, and computational evidence suggests that they are closely related. Even so, these theories have different features: instanton homology has the filtration coming from the $\widetilde{\mathrm{CS}}$ invariant, while the Heegaard Floer and monopole theories are more directly connected to symplectic and contact geometry.

Finally, we mention an intriguing question stemming from an old proposal of Cohen-Jones-Segal: is there a space or other homotopy-theoretic object whose homology groups yield Floer homology (ideally displaying the full algebraic structure mentioned above)? Such spaces have recently been constructed in the monopole setting and for the more combinatorial Khovanov homology, and one would like to see this same enrichment for the instanton theory.

ACKNOWLEDGMENT. The authors wish to thank John Baldwin, Aliakbar Daemi, Paul Kirk, Tye Lidman, Nikolai Saveliev, Steven Sivek, and Cliff Taubes for their insightful comments and valuable suggestions.

References

[BLSY21] John A. Baldwin, Zhenkun Li, Steven Sivek, and Fan Ye, *Small Dehn surgery and SU(2)*, 2021. https://arxiv.org/abs/2110.02874.

[BS18] John A. Baldwin and Steven Sivek, *Stein fillings and* SU(2) *representations*, Geom. Topol. **22** (2018), no. 7, 4307–4380, DOI 10.2140/gt.2018.22.4307. MR3890778

[BS21] John A. Baldwin and Steven Sivek, *Instanton L-spaces and splicing*, 2021. https://arxiv.org/abs/2103.08087.

[BD95] P. J. Braam and S. K. Donaldson, *Floer's work on instanton homology, knots and surgery*, The Floer memorial volume, Progr. Math., vol. 133, Birkhäuser, Basel, 1995, pp. 195–256. MR1362829

[Dae20] Aliakbar Daemi, *Chern-Simons functional and the homology cobordism group*, Duke Math. J. **169** (2020), no. 15, 2827–2886, DOI 10.1215/00127094-2020-0017. MR4158669

[DK90] S. K. Donaldson and P. B. Kronheimer, *The geometry of four-manifolds*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 1990. Oxford Science Publications. MR1079726

[FS92] Ronald Fintushel and Ronald J. Stern, *Integer graded instanton homology groups for homology three-spheres*, Topology **31** (1992), no. 3, 589–604, DOI 10.1016/0040-9383(92)90053-K. MR1174261

[Frø02] Kim A. Frøyshov, *Equivariant aspects of Yang-Mills Floer theory*, Topology **41** (2002), no. 3, 525–552, DOI 10.1016/S0040-9383(01)00018-0. MR1910040

[Fur90] Mikio Furuta, *Homology cobordism group of homology 3-spheres*, Invent. Math. **100** (1990), no. 2, 339–355, DOI 10.1007/BF01231190. MR1047138

[HK11] Matthew Hedden and Paul Kirk, *Chern-Simons invariants,* SO(3) *instantons, and* $\mathbb{Z}/2$ *homology cobordism*, Chern-Simons gauge theory: 20 years after, AMS/IP Stud. Adv. Math., vol. 50, Amer. Math. Soc., Providence, RI, 2011, pp. 83–114. MR2809448

[HPC21] Matthew Hedden and Juanita Pinzón-Caicedo, *Satellites of infinite rank in the smooth concordance group*, Invent. Math. **225** (2021), no. 1, 131–157, DOI 10.1007/s00222-020-01026-w. MR4270665

[KM04a] P. B. Kronheimer and T. S. Mrowka, *Dehn surgery, the fundamental group and SU*(2), Math. Res. Lett. **11** (2004), no. 5-6, 741–754, DOI 10.4310/MRL.2004.v11.n6.a3. MR2106239

[KM04b] P. B. Kronheimer and T. S. Mrowka, *Witten's conjecture and property P*, Geom. Topol. **8** (2004), 295–310, DOI 10.2140/gt.2004.8.295. MR2023280

[KM10] Peter Kronheimer and Tomasz Mrowka, *Knots, sutures, and excision*, J. Differential Geom. **84** (2010), no. 2, 301–364. MR2652464

[LPZ21] Tye Lidman, Juanita Pinzón Caicedo, and Raphael Zentner, *Toroidal homology spheres and SU(2)-representations*, 2021. `https://arxiv.org/abs/2101.02621`.

[Lin16] Jianfeng Lin, SU(2)*-cyclic surgeries on knots*, Int. Math. Res. Not. IMRN **19** (2016), 6018–6033, DOI 10.1093/imrn/rnv326. MR3567265

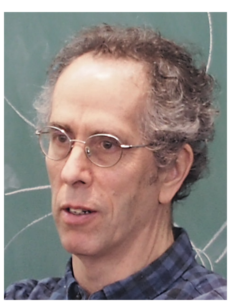
[NST19] Yuta Nozaki, Kouki Sato, and Masaki Taniguchi, *Filtered instanton Floer homology and the homology cobordism group*, 2019. `https://arxiv.org/abs/1905.04001`, J. Eur. Math. Soc., to appear.

[SZ17] Steven Sivek and Raphael Zentner, *SU(2)-cyclic surgeries and the pillowcase*, 2017. `https://arxiv.org/abs/arXiv:1710.01957`.

[Tau90] Clifford Henry Taubes, *Casson's invariant and gauge theory*, J. Differential Geom. **31** (1990), no. 2, 547–599. MR1037415

[Zen18] Raphael Zentner, *Integer homology 3-spheres admit irreducible representations in* $\mathrm{SL}(2,\mathbb{C})$, Duke Math. J. **167** (2018), no. 9, 1643–1712, DOI 10.1215/00127094-2018-0004. MR3813594

Juanita Pinzón-Caicedo



Daniel Ruberman

Credits

Opening image and Figures 1–4 are courtesy of Juanita Pinzón-Caicedo and Daniel Ruberman.

Photo of Juanita Pinzón-Caicedo is courtesy of Mariana Pinzón-Caicedo.

Photo of Daniel Ruberman is courtesy of Stefan Friedl.