Embedded surfaces and gauge theory in three and four dimensions

P. B. Kronheimer[1]

Department of Mathematics
Harvard University,
Cambridge MA 02138
kronheim@math.harvard.edu

Introduction

Given a 2-dimensional homology class σ in a smooth 4-manifold X, what is the least possible genus for a smoothly embedded, oriented surface Σ in X whose fundamental class is σ? Gauge theory has been a successful tool in answering a collection of basic questions of this sort. In [22, 25, 20, 24], information extracted from Donaldson's polynomial invariants of 4-manifolds [6] gave some strong lower bounds, which were in many cases sharp. To give just one example, if X is a smooth quintic surface in $\mathbb{CP}^3$ and Σ is a smooth

[1]Partially supported by NSF grant number DMS-9531964

algebraic curve obtained as the intersection of X with any other complex surface $H \subset \mathbb{CP}^3$, then Σ is known to achieve the smallest possible genus in its homology class.

The introduction of the Seiberg-Witten monopole equations and the replacement of Donaldson's polynomial invariants by the apparently equivalent monopole invariants [42] lead to much simpler proofs of essentially the same results, though often in rather greater generality. For example, the statement about the quintic surface is now known to hold for complex algebraic surfaces in general, and the theorems can be extended quite cleanly to the case of symplectic manifolds [36].

At the same time, the simpler understanding that the Seiberg-Witten techniques afforded made clear that the case of complex and symplectic 4-manifolds was rather special. Although gauge theory gives lower bounds on the genus of embedded surfaces in general 4-manifolds, these lower bounds should no longer be expected to be sharp, at least in the form in which they are usually phrased. The situation is clarified by the work of Meng and Taubes [30], in which a 3-dimensional version of the monopole invariants is studied. For 3-manifolds with non-zero first Betti number, the monopole invariants are closely related to Milnor torsion, and are a cousin of the familiar Alexander invariant of a knot. These invariants do contain information about embedded surfaces, but they do not lie very deep: the degree of the Alexander polynomial gives a lower bound for the genus of a knot, but not always a very good one. There seems no reason to expect the outcome in four dimensions to be any better.

With a little more work, however, there is more information to be extracted from the monopole equations, at least on 3-manifolds, where one can use the equations to define a 'Floer homology'. There is much here that is not yet worked out in detail, but it is already clear that one obtains sharp lower bounds for the genus of embedded surfaces in general 3-manifolds.

One of our aims in this article is to give an account of this 3-dimensional story, which is closely tied up with a well-established theory of embedded surfaces, foliations and contact structures, due to Bennequin, Eliashberg, Gabai and Thurston, among others. We include a leisurely summary of those parts of the foliation story that are most relevant to the gauge theory side. We give a basic account of gauge theory on 3-manifolds, defining the Seiberg-Witten monopole equations, touching on Floer homology and explaining the connection with foliations and contact structures.

By and large, the flow of ideas is one-way: from the geometric and topological results concerning foliations of 3-manifolds, we learn something about the gauge theory invariants. But it seems likely that gauge theory has something to offer in return. We mention some potential applications in section 6, among them the question of whether one can make a homotopy 3-sphere by surgery on a knot. For such applications, one needs not only the Seiberg-Witten techniques, but the older Yang-Mills invariants of Donaldson and Floer. We shall also look at 4-manifolds, to explore the limitations of the existing tools.

Acknowledgment. Amongst the material presented here, the theorems in which the author had a hand are the result of joint work with Tom Mrowka, and have all appeared (or are appearing) elsewhere. In particular, the results which collect around Theorem 3.6 are contained in the joint paper [21] or are easily deduced from results presented there. The fact that one can use the Seiberg-Witten monopole invariants to define invariants of 3-manifolds (either the integer invariants that we call $SW(Y)$ in this paper or the more problematic Floer homology) has been pursued by several people, as has the fact that one can bound the genus of embedded surfaces in terms of these invariants. See for example [2, 4, 27, 29, 30, 40].

1 Surfaces in 3-manifolds

The Thurston norm

We begin in three dimensions. For a 3-manifold Y, the fact that any class σ in $H_2(Y;\mathbb{Z})$ is represented by a smoothly embedded surface can be seen as follows. Take a smooth map $f_\sigma\colon Y \to S^1$ such that the pull-back of the generator of $H^1(S^1)$ is Poincaré dual to σ. Then for any regular value θ of f_σ, the set $f_\sigma^{-1}(\theta)$ is a suitable surface. Any surface representing σ arises in this way for some map f_σ.

It is not always possible to represent σ as the fundamental class of a *connected* surface. Even when a connected representative exists, it is profitable to consider disconnected representatives also and to try and minimize not the genus but the quantity

$$\chi_-(\Sigma) = \sum_{g(\Sigma_i)>0} (2g(\Sigma_i) - 2),$$

over all oriented embedded surfaces Σ whose fundamental class is σ. One might call this the complexity of Σ. The minimum complexity in this sense, as a function of the homology class represented, was considered by Thurston [39], who made the following observation. Define

$$\chi_{\min}(\sigma) = \min\{\, \chi_-(\Sigma) \mid [\Sigma] = \sigma \,\}.$$

Proposition 1.1 (Thurston). *On any closed, oriented 3-manifold Y, the function $\chi_{\min}$ on $H_2(Y;\mathbb{Z})$ satisfies the triangle inequality and is linear on rays, in that $\chi_{\min}(n\sigma) = n\,\chi_{\min}(\sigma)$ for $n \geq 0$. It is the restriction of a semi-norm on $H_2(Y;\mathbb{R})$.*

The triangle inequality and linearity have straightforward geometrical explanations. Suppose that Σ is an oriented embedded surface realizing the minimal complexity in its homology class $\sigma = [\Sigma]$. Let $\tilde{\Sigma}$ be the surface obtained by taking n disjoint, parallel copies of Σ inside a product neighborhood $\Sigma \times [-1,1] \subset Y$. This surface represents the class $n\sigma$ and has complexity

$$\chi_-(\tilde{\Sigma}) = n\chi_-(\Sigma).$$

The linearity assertion is that $\tilde{\Sigma}$ also has minimal complexity in its class. The reason this is true is that *any* surface $\hat{\Sigma}$ representing $n\sigma$ must be a disjoint union of n surfaces, each representing σ, so the complexity of $\hat{\Sigma}$ cannot be less than $n\,\chi_{\min}(\sigma)$. Indeed, we can realize $\hat{\Sigma}$ as a regular inverse image $\hat{f}^{-1}(\hat{\theta})$ for a suitable $\hat{f}\colon Y \to S^1$ as above, and the divisibility of $[\hat{\Sigma}]$ implies that this $\hat{f}$ lifts through the n-fold covering map $\mu_n\colon S^1 \to S^1$:

$$\hat{f} = \mu_n \circ f.$$

Thus $\hat{\Sigma}$ is the disjoint union of the surfaces $f^{-1}(\theta)$, as θ runs through the n preimages of $\hat{\theta}$.

To see that that $\chi_{\min}$ satisfies the triangle inequality, let classes σ and τ be represented by surfaces Σ and T of minimal complexity. If these are moved into general position, they will intersect along a collection of circles. Each of the circles has a neighborhood $D^2 \times S^1$ which meets $\Sigma \cup T$ in a standard $K \times S^1$, where K is a pair of intersecting arcs in D^2. Replacing $K \times S^1$ with $J \times S^1$, where J is a pair of disjoint arcs in D^2 with the same four endpoints (connected differently), we obtain a new surface in Y representing $\sigma + \tau$. Its complexity is the sum of the complexities of Σ and T, thus

$$\chi_{\min}(\sigma + \tau) \leq \chi_{\min}(\sigma) + \chi_{\min}(\tau). \tag{1}$$

We shall write $|\sigma|$ for the norm (or semi-norm) $\chi_{\min}(\sigma)$ in the 3-dimen/-sional case. The dual norm on $H^2(Y;\mathbb{R})$ can be characterized by

$$|\alpha|_* = \sup_{\Sigma} \frac{\langle \alpha, [\Sigma] \rangle}{(2g(\Sigma) - 2)}, \qquad \alpha \in H^2(Y;\mathbb{R}), \tag{2}$$

where the supremum need only be taken only over connected embedded surfaces of genus 1 or more on which α has non-zero pairing, with the understanding that the norm is infinite if α has non-zero pairing with an embedded torus. The *Thurston polytope* $B(Y)$ is the unit ball for this dual Thurston norm. It is a convex polytope lying in the subspace of $H^2(Y;\mathbb{R})$ on which the norm is finite. Its vertices are lattice points (that is, they are the reduction of integer classes).

Foliations

If working with a closed 3-manifold is difficult, one can get a feel for the problem of determining the Thurston norm, or equivalently the polytope $B(Y)$, by looking at a version of the question for a manifold with boundary, such as a knot complement. Let $K \subset S^3$ be a knot and Y the 3-manifold obtained by removing an open tubular neighborhood of K. The boundary of Y is a torus T, on which there is a simple closed curve λ having the property that it is homologous to zero in Y: this is the longitude of the knot. Being null-homologous, λ is the boundary of an oriented surface $\Sigma \subset Y$. Any such surface is a *spanning surface* for K. The *genus* of the knot K is the least genus of any spanning surface.

Finding a spanning surface for K is never hard. Harder is to provide a spanning surface Σ with a certificate assuring us that it is of least possible genus. It is another observation of Thurston's [39] that a suitable *foliation* of the 3-manifold may supply such a certificate.

To state this result, we return to the closed case and suppose that the 3-manifold Y has a smooth foliation $\mathcal{F}$ by oriented 2-dimensional leaves L. Such a foliation determines a field of oriented 2-planes $T\mathcal{F} \subset TY$, the tangent directions to the leaves. Let $e(\mathcal{F})$ denote the Euler class of $T\mathcal{F}$ in $H^2(Y;\mathbb{Z})$.

Theorem 1.2 (Thurston). *Suppose the foliation $\mathcal{F}$ of Y has no Reeb components, and suppose that Y is not $S^1 \times S^2$. Then the Euler class $e(\mathcal{F})$ belongs to the polytope $B(Y)$. In other words, if Σ is any embedded surface, its complexity satisfies the lower bound*

$$\chi_-(\Sigma) \geq \langle e(\mathcal{F}), [\Sigma] \rangle.$$

A *Reeb component* is a foliation of a solid torus in which the boundary torus is a leaf and all interior leaves are planes.

Corollary 1.3. *Let Σ be an oriented embedded surface in Y which is a union of compact leaves of an oriented foliation $\mathcal{F}$ without Reeb components. Then Σ has minimal complexity in its homology class.*

The Corollary follows from the Theorem because if Σ is a union of correctly-oriented compact leaves, then

$$\begin{aligned}\langle e(\mathcal{F}), [\Sigma]\rangle &= \langle e(T\Sigma), [\Sigma]\rangle \\ &= -\chi_-(\Sigma),\end{aligned}$$

as long as no component of Σ is a sphere. Changing the orientation of the leaves then gives a foliation $\bar{\mathcal{F}}$ for which the inequality of the Theorem is an equality. Note that the Theorem contains the statement that no compact leaf of $\mathcal{F}$ can be a sphere. The proof of Theorem 1.2 itself is given in [39]; see also Theorem 7.1 of [13], where a detailed proof is given of a more general version of the basic lemma which underlies the result. The original version of the lemma was first proved by Thurston [38] and Roussarie [34]:

Lemma 1.4. *Let $\mathcal{F}$ and Y be as in Theorem 1.2, and let Σ be an incompressible surface in Y. Then there is a surface Σ' isotopic to Σ which is transverse to $\mathcal{F}$ except at a finite number of circle and saddle tangencies.*

In this statement, *incompressible* means as usual that no loop in Σ' bounds a disk in $Y \setminus \Sigma'$, unless it already bounds a disk in Σ'. This condition is certainly necessary if Σ is to be of least complexity, since by cutting along a compressing disk one reduces the complexity of a surface. A *circle* tangency is what one sees along the level rim of a volcano, in the foliation of 3-space by level planes. A *saddle* tangency needs no explanation. In each case, the tangency may be given one of two signs, according as the orientation agrees with that of the leaves or not. One can calculate $\chi_-(\Sigma')$ and the evaluation of $e(\mathcal{F})$ on Σ' in terms of the number of tangencies of each type [39], and the Theorem is an elementary consequence.

Corollary 1.3 gives a criterion with which to ascertain that a given surface is of minimal complexity. But it is of little use in itself unless we also have a handle on constructing foliations without Reeb components. In [14], Gabai gave a practical algorithm for finding the genus of a large class of knots, based on exhibiting a spanning surface as a compact leaf of a foliation of the knot complement. Further, in [13], Gabai proved that every surface of minimal complexity can be certified as such by a suitable foliation. We state the result only in the closed case:

Theorem 1.5 (Gabai). *Let Y be a closed, irreducible, oriented 3-manifold. Let Σ be an embedded surface representing a non-trivial homology class σ. Suppose that $\chi_{-}(\Sigma)$ is least possible amongst surfaces representing this class. Then there exists a taut, oriented foliation $\mathcal{F}$ of Y of class C^0, having Σ as an oriented union of compact leaves. The foliation can be taken to be smooth except along the components of Σ which are tori.*

Remarks. The *irreducibility* of Y is the condition that every embedded 2-sphere bounds a ball. It excludes $S^1 \times S^2$ as well as non-trivial connected sums. One definition of *taut* is that every leaf L is met by a closed curve γ in Y which is everywhere transverse to the leaves. This is a stronger condition than the absence of Reeb components. In Lemma 1.4 for example, the taut condition allows one to dispose of circle tangencies, leaving only saddles. Generally, in a taut foliation, no correctly-oriented union of compact leaves can bound: if $W \subset Y$ had oriented boundary which was a union of leaves, then no transverse curve which left W to enter $Y \setminus W$ could return. The possibility that the foliation $\mathcal{F}$ in the Theorem may not be smooth when Σ has tori amongst its components presents a difficulty at some points: we will sometimes legislate against it in the statement of our results. In many cases one can construct a foliation that is smooth, despite the presence of tori.

Gabai's result combines with Thurston's to characterize the polytope $B(Y)$, and hence the Thurston norm, in terms of smooth foliations:

Corollary 1.6. *Let Y be a closed, irreducible oriented 3-manifold in which embedded tori do not form a basis for the homology. Then the unit ball $B(Y) \subset H^2(Y;\mathbb{R})$ for the dual Thurston norm is the convex hull of the classes $e(\mathcal{F})$, as $\mathcal{F}$ runs through smooth, taut foliations. In other words, the Thurston norm is given by*

$$|\sigma| = \max_{\mathcal{F}\ \mathrm{taut}} \langle e(\mathcal{F}), \sigma \rangle.$$

The extra hypothesis in this Corollary ensures that at least one of the foliations which result for Theorem 1.5 is smooth. When this is the case, we can throw out all the non-smooth foliations corresponding which might arise when Σ contains tori, without changing the convex hull.

In [13], the inequality of Theorem 1.2 is extended to the case that Σ is not an embedded surface, but simply a surface mapped into Y by an arbitrary map, whose image may therefore have singularities. Combining this strengthened inequality with the existence theorem for foliations, one obtains:

Corollary 1.7 (Gabai, [13]). *If a homology class σ in an irreducible 3-manifold Y is represented as $f_*[\Sigma]$ for some map $f : \Sigma \to Y$, then the homology class is also the fundamental class of an embedded surface of the same complexity.*

Non-trivial examples of plane polygons arising as the Thurston polytopes of various 3-manifolds with $b_1 = 2$ are given in Thurston's original paper. Gabai's theorem and its refinements have many applications in 3-manifold topology [13, 15, 16, 17].

2 Gauge theory on 3-manifolds

The Seiberg-Witten *monopole equations* were introduced as a tool in 4-dimensional topology in [42]. One of their first applications was to questions of embedded surfaces [23]. Here we shall explore their relationship to the Thurston norm in dimension 3.

The monopole equations

A Spin^c structure $\mathfrak{c}$ on an oriented Riemannian 3-manifold Y consists of a rank-2 complex bundle $W = W_{\mathfrak{c}}$ with a hermitian metric (the spinor bundle) and an action ρ of 1-forms on spinors, called Clifford multiplication:

$$\rho : T^*Y \to \mathrm{End}(W).$$

If e^1, e^2, e^3 are an oriented orthonormal frame for the cotangent space at a point, then there should be an orthonormal basis for the fiber of W at that point so that the $\rho(e^i)$ are represented by the Pauli matrices. The standard convention on orientations has $\rho(e^1)\rho(e^2)\rho(e^3) = -1$. Clifford multiplication is extended to forms of higher degree, by $\rho(e^1 \wedge e^2) = \rho(e^1)\rho(e^2)$, for example.

If $\mathfrak{c}$ is a Spin^c structure and $e \in H^2(Y;\mathbb{Z})$ a 2-dimensional cohomology class, there is a new Spin^c structure $\mathfrak{c}+e$. Its spin bundle is $W_{\mathfrak{c}} \otimes L_e$, where L_e is the unique line bundle with first Chern class e. Conversely, if $\mathfrak{c}$ and $\mathfrak{c}'$ are two Spin^c structures, there is a unique difference element $\mathfrak{c}' - \mathfrak{c} \in H^2(Y;\mathbb{Z})$. We write $c_1(\mathfrak{c})$ for the first Chern class of $W_{\mathfrak{c}}$, and we note that, since

$$c_1(\mathfrak{c} + e) = c_1(\mathfrak{c}) + 2e,$$

the class $c_1(\mathfrak{c})$ determines $\mathfrak{c}$ in the absence of 2-torsion in the cohomology. Because every 3-manifold is parallelizable, there is always at least one Spin^c structure with a topologically trivial spin bundle. It follows that $c_1(\mathfrak{c})$ is always divisible by 2 in $H^2(Y;\mathbb{Z})$. Although our definition involves a Riemannian metric, the set of isomorphism classes of Spin^c structures can be viewed as metric-independent.

Given a Spin^c structure, a unitary connection A on W is a *spin* connection if ρ is parallel; that is, the resulting connection on $\mathrm{End}(W)$ should coincide with the Levi-Civita connection on the image of ρ. This leaves only the central part of A undetermined, so if A and A' are two spin connections, then their difference is scalar:

$$A' = A + a\mathbf{1}_W,$$

for an imaginary-valued 1-form a. A spin connection A is therefore determined by its 'trace', the induced connection $\hat{A}$ in the line bundle $\Lambda^2 W$.

The monopole equations are the following equations for a spin connection A and a section Φ of W on a Riemannian 3-manifold Y equipped with Spin^c structure:

$$\begin{aligned} \rho(F_{\hat{A}}) - \{\Phi \otimes \Phi^*\} &= 0 \\ D_A \Phi &= 0. \end{aligned} \tag{3}$$

In the first equation, $F_{\hat{A}}$ is the curvature of the connection in the line bundle (an imaginary-valued 2-form), and the curly brackets denote the trace-free part of the endomorphism. In the second equation, D_A is the Dirac operator for the spin connection A, which is defined as the composite

$$\Gamma(W) \xrightarrow{\nabla_A} \Gamma(T^*Y \otimes W) \xrightarrow{\rho} \Gamma(W).$$

An application of the Weitzenbock formula

When a Riemannian 3-manifold Y is given and a Spin^c structure $\mathfrak{c}$ is specified, we can ask first whether the monopole equations (3) have any solutions at all. There is a constraint which must be satisfied if a solution is to exist, which comes from the Weitzenbock formula for the Dirac operator D_A:

$$D_A^* D_A \Phi = \nabla_A^* \nabla_A \Phi + \frac{s}{4}\Phi + \frac{1}{2}\rho(F_{\hat{A}})\Phi.$$

Here s is the scalar curvature of the Riemannian metric. If (A, Φ) is a solution of the equations, then the left and right-hand sides are zero. For a solution, then, we calculate

$$\begin{aligned}\Delta|\Phi|^2 &= 2\langle \nabla_A^* \nabla_A \Phi, \Phi\rangle - 2\langle \nabla_A \Phi, \nabla_A \Phi\rangle \\ &\le -\frac{s}{2}|\Phi|^2 - \langle \rho(F_{\hat{A}})\Phi, \Phi\rangle.\end{aligned}$$

(Our inner products are real.) The first of the equations (3) can be used to rewrite the last term as $|\Phi|^4/2$ (an elementary calculation with vectors in $\mathbb{C}^2$). So we have

$$2\Delta|\Phi|^2 \le -s|\Phi|^2 - |\Phi|^4, \tag{4}$$

which we may integrate to obtain

$$\int |\Phi|^4 d\mathrm{vol} \le \int (-s)|\Phi|^2 d\mathrm{vol},$$

and hence

$$\int |\Phi|^4 d\mathrm{vol} \le \int s^2 d\mathrm{vol} \tag{5}$$

by Cauchy-Schwartz. The first equation of (3) is used again to rewrite this as

$$\int |F_{\hat{A}}|^2 d\mathrm{vol} \le \int \frac{|s|^2}{4} d\mathrm{vol},$$

which is an inequality between the L^2 norms:

$$\|F_{\hat{A}}\| \le \|s\|/2. \tag{6}$$

The two form $(i/2\pi)\|F_{\hat{A}}\|$ represents the first Chern class of $\Lambda^2 W_{\mathfrak{c}}$, which is $c_1(\mathfrak{c})$. The de Rham representative of this class with smallest L^2 norm is the *harmonic* representative. Let us define an L^2 norm on $H^2(Y;\mathbb{R})$ by defining $\|\alpha\|$ to be the L^2 norm of the harmonic representative. Then we have, from (6)

$$\|c_1(\mathfrak{c})\| \leq \|s\|/4\pi.$$

(Both sides depend, of course, on the Riemannian metric.) We state the conclusion as a lemma:

Lemma 2.1. *A necessary condition for the existence of a solution to the monopole equations on Y for a given Riemannian metric and a given* Spin^c *structure $\mathfrak{c}$ is that the harmonic L^2 norm of the class $c_1(\mathfrak{c})$ satisfy*

$$\|c_1(\mathfrak{c})\| \leq \|s\|/4\pi. \tag{7}$$

□

Note that we could have taken a little more care over the argument, to arrive at the inequality

$$\|c_1(\mathfrak{c})\| \leq \|s_-\|/4\pi, \tag{8}$$

where s_- is defined pointwise as $\max(0, -s)$.

The calculation above, and the inequality (6), were first described by Witten in [42] (though in a 4-dimensional version). As pointed out there, it follows that, for a fixed Riemannian metric, there can be solutions for only finitely many different Spin^c structures on Y. Indeed, there are only finitely many integer classes with norm less than any given constant, and the cohomology class $c_1(\mathfrak{c})$ (as an integer class) determines the Spin^c structure $\mathfrak{c}$ to within the addition of a 2-torsion element.

One can also obtain a pointwise bound on $|\Phi|$ from the inequality (4). At a point where $|\Phi|$ is maximum, the Laplacian is positive, and as long as $|\Phi|$ is not zero at this point one may divide to obtain the estimate

$$|\Phi|^2 \leq -s$$

at the maximum.

Scalar curvature and the Thurston norm

The relationship between the monopole equations and the genus of embedded surfaces arises in its simplest form from the following simple lemma. Like the inequality (7), it compares the norm of a class in H^2 to the norm of the scalar curvature.

Lemma 2.2. *Let $\alpha \in H^2(Y;\mathbb{R})$ be a two-dimensional cohomology class. Then the dual Thurston norm $|\alpha|_*$ satisfies the inequality*

$$|\alpha|_* \leq 4\pi \sup_h \frac{||\alpha||_h}{||s_h||_h}, \tag{9}$$

in which the supremum is taken over all Riemannian metrics h, the norms on the right-hand side are the L^2 norms, and s_h denotes the scalar curvature of h.

Proof. Let Σ be an oriented embedded surface of genus 1 or more on which α is non-zero. Let h_1 be a Riemannian metric on Y such that some neighborhood of Σ is isometric to a product $\Sigma \times [0,1]$, with Σ having constant scalar curvature $-4\pi(2g-2)$ and unit area. Let h_r be a metric which contains a product region $\Sigma \times [0,r]$ and is isometric to h_1 outside that region. Then

$$||s_h||_h = 4\pi r^{1/2}(2g-2) + O(1)$$

as $r \to \infty$, while any 2-form ω representing a class α must satisfy

$$||\omega||_h \geq r^{1/2}\langle \alpha, [\Sigma]\rangle.$$

Thus if g is at least two we have

$$\langle \alpha, [\Sigma]\rangle/(2g-2) \leq \sup_h 4\pi ||\alpha||_h/||s_h||_h,$$

and in the case that Σ is a torus we see that the right-hand side is infinite. This is the desired result, in view of the characterization of the dual Thurston norm at (2). □

Combining this lemma with the previous one, we obtain:

Corollary 2.3. *If $\mathfrak{c}$ is a* Spin^c *structure on Y and the dual Thurston norm of $c_1(\mathfrak{c})$ is bigger than 1, then there exists a Riemannian metric on Y for which the monopole equations have no solution.* □

We shall see later that the inequality in Lemma 2.2 is actually an equality for many 3-manifolds Y (Proposition 3.8).

3 The monopole invariants

Obtaining invariants from the monopole equations

There is a well-understood procedure by which we can extract some metric-independent data from the set of solutions to the monopole equations, to obtain an invariant

$$SW(Y, \mathfrak{c}) \in \mathbb{Z}$$

depending only on Y (a closed, oriented 3-manifold) and the Spin^c structure $\mathfrak{c}$. A careful account of the most general case is given in [4, 27], and there is a model for this construction in [35], where Casson's invariant is given a gauge-theory interpretation. In this subsection, we review this construction, restricting ourselves sometimes to the simpler cases.

The monopole equations on a Riemannian 3-manifold Y are the variational equations for a functional

$$\mathrm{CSD}(A, \Phi) = \frac{1}{2}\int_Y (\hat{A} - \hat{A}_0) \wedge (F_{\hat{A}} + F_{\hat{A}_0}) - \frac{1}{2}\int_Y \langle \Phi, D_A \Phi \rangle \, d\mathrm{vol},$$

the *Chern-Simons-Dirac* functional of A and Φ. (A reference connection A_0 is chosen to define the first term, but a change of reference connection only changes the functional by addition of a constant.) The equations are invariant under the natural symmetry group of the Spin^c bundle $W_{\mathfrak{c}}$, which is the group G of maps $u : Y \to S^1$ acting on $W_{\mathfrak{c}}$ by scalar endomorphisms. This group acts on A and Φ by

$$\begin{aligned} A &\mapsto A - (u^{-1}du)\mathbf{1} \\ \Phi &\mapsto u\Phi. \end{aligned}$$

The functional CSD is invariant only under the identity component of G. The component group of G is the group of homotopy classes of maps $Y \to S^1$, which is isomorphic to $H^1(Y; \mathbb{Z})$: the change in the functional under a general element of G is given by

$$\mathrm{CSD}(A - (u^{-1}du)\mathbf{1}, u\Phi) - \mathrm{CSD}(A, \Phi) = -4\pi^2\big([u] \smile c_1(\mathfrak{c})\big)[Y].$$

So CSD does not descend to a real-valued function on the quotient space $\mathcal{C} = \{(A, \Phi)\}/G$, but does descend to a circle-valued function whose periods are multiples of $4\pi^2$.

The orbit space $\mathcal{C}$ is an infinite-dimensional manifold except at points where the action has non-trivial stablilizer. These are the configurations with $\Phi = 0$ (the *reducible* configurations), whose stabilizer is the circle group of constant maps u. If $c_1(\mathfrak{c})$ is zero or a torsion class, then there are spin connections for which $\hat{A}$ is flat, and these are reducible solutions to the equations.

To define $SW(Y, \mathfrak{c})$, one first chooses a perturbation of the functional CSD so as to make the critical points non-degenerate. To do this, one may add an extra term, setting

$$\mathrm{CSD}_\mu(A, \Phi) = \mathrm{CSD}(A, \Phi) + \int_Y (\hat{A} - \hat{A}_0) \wedge (i\mu),$$

where μ is an exact 2-form. The variational equations are now

$$\begin{aligned} \rho(F_{\hat{A}} + i\mu) - \{\Phi \otimes \Phi^*\} &= 0 \\ D_A \Phi &= 0, \end{aligned} \tag{10}$$

and for a suitable choice of μ (such μ are dense), the Hessian will be non-degenerate at all irreducible solutions in $\mathcal{C}$. A proof is given in [12]. (The condition that μ be exact means, among other things, that the function CSD_μ still descends to a circle-valued function on $\mathcal{C}$ and that its periods are the same as those of CSD.)

Critical points of CSD_μ on $\mathcal{C}$ have infinite index: the Hessian is a self-adjoint operator which, like the Dirac operator, has a discrete spectrum which is infinite in both the positive and negative directions, so one cannot define an index $i(a)$ at a critical point a as the dimension of the sum of the negative eigenspaces. However, there is a way to define a relative index between any two critical points. If a and b are two non-degenerate critical points of CSD_μ and $\gamma(t)$ is a path in $\mathcal{C}$ which joins them, then the Hessians $H_{\gamma t}$ a family of operators for which one can define a spectral flow – the number of eigenvalues which move from negative to positive in the family. One can defines the relative index $i(a, b)$ as the spectral flow. There is a further point to attend to, which is the possible dependence of the spectral flow on the choice of path. The fundamental group of $\mathcal{C}$ is again isomorphic to $H^1(Y; \mathbb{Z})$, and for a *closed* path γ, the spectral flow is given by

$$\mathrm{SF}(\gamma) = \big([u] \smile c_1(\mathfrak{c})\big)[Y],$$

where u is the corresponding element of H^1. As noted above, $c_1(\mathfrak{c})$ is always divisible by 2, so at least the parity of $i(a, b)$ is well-defined. After settling on a convention as to which is to be which, we can divide the critical points into even and odd, using the relative index.

The invariant $SW(Y, \mathfrak{c})$ can now be defined as, roughly speaking, the euler number of the vector field CSD_μ on $\mathcal{C}$. We restrict our attention to the case that $c_1(\mathfrak{c})$ is not a torsion class, so that there are no reducible solutions to the equations, and after choosing μ so that the solutions are non-degenerate we set

$$SW(\mathfrak{c}) = \#(\text{even}) - \#(\text{odd}). \tag{11}$$

The main technical point here is that the set of critical points in $\mathcal{C}$ is compact, and hence finite under our non-degeneracy assumption. The proof of the compactness property of solutions to the monopole equations is a straightforward application of standard techniques, starting from the C^0 bound obtained by applying the maximum principle to (4). The quantity (11) is independent of the choice of Riemannian metric on Y and the choice of μ. Our failure to fix a convention about which is even and which is odd leaves an overall sign ambiguity in the invariant.

If $c_1(\mathfrak{c})$ is torsion, a similar definition leads to an invariant in the case that $b_1(Y)$ is non-zero: one can perturb by a small, non-exact, closed 2-form μ to remove the reducible solutions. The case of $b_1(Y) = 0$ is a little more subtle, but an invariant can be defined: see [4, 27].

Basic classes

We call a Spin^c structure $\mathfrak{c}$ on a closed 3-manifold Y *basic* if the monopole invariant $SW(Y, \mathfrak{c})$ is non-zero. In this case, we also refer to the first Chern class $c_1(\mathfrak{c})$ as a *basic class* if Y. For our present purposes, these definitions are only interesting when $b_1(Y)$ is non-zero and $c_1(\mathfrak{c})$ is not torsion. The definition of the invariant SW means that, if $\mathfrak{c}$ is basic, then the corresponding perturbed monopole equations (10) have at least one solution, for every Riemannian metric and a dense set of exact 2-forms μ. The compactness properties of the equations imply that the non-emptiness of the solution space is an open condition, so it is also true that the original equations (3) have solutions for every Riemannian metric h. This observation, together with Corollary 2.3, yields a relationship between basic classes and the genus of embedded surfaces:

Proposition 3.1. *If Y is a closed, oriented 3-manifold with $b_1 \neq 0$ and α is a basic class on Y, then the dual Thurston norm of α satisfies $|\alpha|_* \leq 1$. In other words, for any oriented embedded surface Σ in Y representing a class σ, we have*

$$\chi_-(\Sigma) \geq \langle \alpha, \sigma \rangle.$$

□

Monopole invariants and the Alexander invariant

Whether the above proposition is useful depends on what else one knows about the invariant SW and the basic classes. Meng and Taubes [30] showed that when $b_1(Y)$ is non-zero, $SW(Y, \mathfrak{c})$ is completely determined by a classical invariant, the Milnor torsion. While it is interesting that torsion is calculated by the gauge theory route, this result does make Proposition 3.1 look less interesting.

The situation is easiest to describe when the 3-manifold Y is the result of zero surgery on a knot K in S^3: that is, Y is obtained by removing a solid torus neighborhood of K and replacing it while interchanging the longitude and meridian curves on its boundary. In this case, the information contained in the Milnor torsion is the *Alexander polynomial* of K. We shall use the symmetrized Alexander polynomial, whose shape is

$$A_K(t) = a_{-r}t^{-r} + \cdots + a_{-1}t^{-1} + a_0 + a_1 t + \cdots + a_r t^r$$

with $a_i = a_{-i}$. We refer to r as the degree of the polynomial. The betti number of $Y = Y(K)$ is 1, and the second cohomology is $\mathbb{Z}$. There is therefore exactly on Spin^c structure $\mathfrak{c}_k$ with $c_1(\mathfrak{c}_k) = 2k$. Its monopole invariant can be expressed in terms of the Alexander polynomial:

Theorem 3.2 (Meng-Taubes [30]). *On the manifold $Y(K)$ obtained by zero-surgery on K, the monopole invariants are given by*

$$SW(Y, \mathfrak{c}_k) = \sum_{j>0} j a_{j+|k|}. \tag{12}$$

In particular, $SW(z, \mathfrak{c}_k) = 0$ for $k > r - 1$. □

(The symmetry $SW(Y,\mathfrak{c}_k) = SW(Y,\mathfrak{c}_{-k})$ is a general property of the monopole invariants, and follows from a symmetry of the equations.)

With this interpretation of the monopole invariants, we can reformulate Proposition 3.1 as

Statement 3.3. *If Σ is a connected, oriented surface in $Y(K)$ representing the generator of H_2, then*

$$g(\Sigma) \geq r,$$

where r is the degree of the Alexander polynomial.

Indeed, the theorem tells us that the extreme basic classes in $H^2 \cong \mathbb{Z}$ arise from the Spin^c structures $\mathfrak{c}_{\pm(r-1)}$, with $c_1(\mathfrak{c})[\Sigma] = \pm(2r-2)$. (The restriction to connected surfaces has no effect on our statement when $b_1 = 1$.)

The statement above, however, is an elementary consequence of the definition of the Alexander polynomial, in the formulation which expresses $A_K(t)$ in terms of the homology of the infinite cyclic cover of $Y(K)$. Furthermore, the inequality between the genus of Σ and the degree of the Alexander polynomial is not a sharp one in general. Here is a relevant result:

Theorem 3.4 (Gabai [15]). *If Σ is an embedded surface in $Y(K)$ representing the generator of H_2, then $g(\Sigma)$ is no smaller than the genus of the knot.*

This theorem is not self-evident. The proof of the result comes from refinement of the existence theorem for taut foliations quoted earlier: a minimal-genus spanning surface for K can be filled out to a taut foliation $\mathcal{F}$ of the complement of a neighborhood $N(K)$ in S^3, and this can be done so that the boundaries of the leaves are a family of longitudinal circles. The solid torus $N(K)$ has a trivial foliation by meridianal disks, and in the surgered manifold $Y(K)$, this foliation joins with $\mathcal{F}$ to give a taut foliation of the closed manifold.

The genus of a knot K and the degree of A_K are in general different. The untwisted Whitehead double of any non-trivial knot has genus 1 and Alexander polynomial 1. Figure 1 shows a Whitehead double of a trefoil knot; a spanning surface is formed from a ribbon which follows the course of the trefoil, together with a small band with one full twist at the clasp. Under the connected sum operation, the genus of a knot is additive, while the Alexander polynomial is multiplicative, so one can easily obtain knots of large genus whose polynomial has small degree.

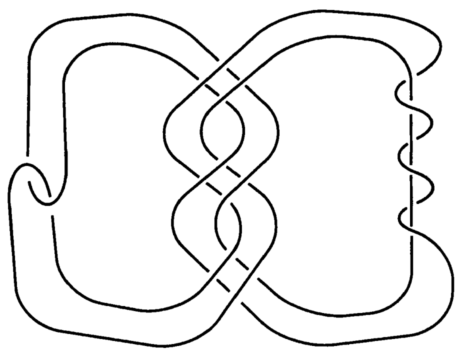

Figure 1: A doubled trefoil.

Monopole classes

The result of the previous subsection is disappointing. The lower bound for the genus of embedded surfaces in terms of basic classes is no better than the lower bound for the genus of a knot which the Alexander polynomial provides: it does not capture the topology. However, the reason for the disappointing result is that too much of the content of the gauge theory has been disposed of in passing from the solution set of the monopole equations (3) to the integer invariant $SW(Y, \mathfrak{c})$ which counts these solutions.

Rather than count the solutions, let us take a step backwards and simply make the following definition:

Definition 3.5. A class $\alpha \in H^2(Y;\mathbb{Z})$ is a *monopole class* if it arises as $c_1(\mathfrak{c})$ for some Spin^c structure $\mathfrak{c}$ on Y for which the equations (3) admit a solution *for every choice of Riemannian metric h on Y.*

This definition ensures that the basic classes are monopole classes. Also, Proposition 3.1 applies to monopole classes, just as it applies to basic classes; so a monopole class α has dual Thurston norm at most 1. (This is simply another step backwards, to Corollary 2.3, from which Proposition 3.1 was deduced.)

However, the monopole classes are in general a *larger* set than the basic classes. The following theorem is our central result. We shall give an outline of the proof in section 5.

Theorem 3.6. *If Y is a closed, irreducible, oriented 3-manifold with a smooth, taut foliation $\mathcal{F}$ by oriented leaves, then $e(\mathcal{F})$ is a monopole class.*

Combining this statement with Corollary 1.6, one obtains

Corollary 3.7. *If Y is a closed, irreducible oriented 3-manifold, then the unit ball $B(Y) \subset H^2(Y;\mathbb{R})$ for the dual Thurston norm is the convex hull of the monopole classes (reduced to real coefficients). In other words, the Thurston norm on H_2 is given by*

$$|\sigma| = \max_{\alpha} \langle \alpha, \sigma \rangle,$$

where the maximum is taken over all monopole classes.

(The extra hypothesis in Corollary 1.6, that there is not a basis for H_2 consisting of tori, was there to ensure that Y had at least one smooth, taut foliation. The hypothesis is unnecessary in Corollary 3.7, for the accidental reason that 0 is always a monopole class, on account of the reducible solution with $\Phi = 0$.)

Thus the monopole classes give us sharp information about the genus of embedded surfaces, while the basic classes (in general) do not. For example, in the 3-manifold $Y(K)$ obtained by zero-surgery on a knot K of genus g, the classes $\pm(2g-2)$ in $H^2 \cong \mathbb{Z}$ are monopole classes. This means that, for the Spin^c structure which we called $\mathfrak{c}_{g-1}$, solutions of the monopole equations always exist, even though the algebraic count of the solutions will be zero if the Alexander polynomial has small degree.

We can also return to our discussion of the scalar curvature, and see that our inequalities there were sharp also. Our previous results stated that the unit ball for the norm

$$4\pi \sup_h \frac{||\alpha||_h}{||s_h||_h}$$

was sandwiched between the convex hull of the monopole classes and the unit ball for the dual Thurston norm (Lemmas 2.1 and 2.2). Knowing that these last two coincide, we can replace the inequality of Lemma 2.2 with an equality:

Proposition 3.8. *If Y is a closed, irreducible oriented 3-manifold, then the dual Thurston norm on $H^2(Y;\mathbb{R})$ is given by*

$$|\alpha|_* = 4\pi \sup_h \frac{||\alpha||_h}{||s_h||_h},$$

where the supremum is taken over all Riemannian metrics on Y. □

Remark. Up until this point, it was by no means clear that the supremum on the right hand side was ever finite. Our results now say that the supremum is reached, in the limit, by stretching the metric along a cylinder $[-R, R] \times \Sigma$, where Σ is a minimum-genus representative for a class σ with $\langle \alpha, \sigma \rangle = |\alpha|_* |\sigma|$.

4 Detecting monopole classes

How can one detect that a given class is a monopole class, without it being a basic class? An existence theorem is needed for solutions to the equations. A simple scenario in which one can see that solutions must exist arises when the 3-manifold in question is embedded in a suitable 4-manifold. We therefore turn to dimension four.

The 4-dimensional equations

The equations we have been discussing are a 3-dimensional version of the monopole equations which were first introduced, by Witten [42], in dimension 4. On an oriented Riemannian 4-manifold X, a Spin^c structure $\mathfrak{c}$ consists of a hermitian vector bundle W of rank 4, together with a Clifford multiplication

$$\rho : T^*X \to \mathrm{End}(W)$$

with the property that, if $e^1, \ldots, e^4$ are an orthonormal coframe at a point in X, then the endomorphisms $\rho(e^i)$ are skew-adjoint and satisfy the Clifford relations

$$\rho(e^i)\rho(e^j) + \rho(e^j)\rho(e^i) = -2\delta_{ij}.$$

Clifford multiplication is extended to forms of higher degree as before. It is a consequence of this definition that the spin bundle W decomposes into two bundles of rank 2, $W^+ \oplus W^-$, whose determinants are equal. The action of 1-forms maps $W^+ \to W^-$, the action of 2-forms preserves the decomposition, and one can characterize W^- as the subspace annihilated by $\rho(\omega)$ for all *self-dual* 2-forms ω (forms satisfying $*\omega = \omega$). A spin connection is defined as before, and given a spin connection A, one has a Dirac operator D_A acting on sections of W. We write D_A^+ for the restriction of D_A to W^+, which is an operator

$$D_A^+ : \Gamma(W^+) \to \Gamma(W^-).$$

The connections on $\Lambda^2 W^+$ and $\Lambda^2 W^-$ induced by a spin connection A are equal, and we write $\hat{A}$ for either. We write $c_1(\mathfrak{c})$ for the first Chern class of W^+, which is the class represented in de Rham cohomology by the form $(i/2\pi)F_{\hat{A}}$.

Once again, the set of Spin^c structures is acted on transitively by $H^2(X;\mathbb{Z})$, and we have the same rule,

$$c_1(\mathfrak{c}+e) = c_1(\mathfrak{c}) + 2e,$$

which shows that $c_1(\mathfrak{c})$ determines $\mathfrak{c}$ to within a finite ambiguity measured by the 2-torsion subgroup of $H^2(X;\mathbb{Z})$.

The 4-dimensional monopole equations are the following pair of equations for a section Φ of W^+ and a spin connection A:

$$\begin{aligned} \rho(F^+_{\hat{A}}) - \{\Phi \otimes \Phi^*\} &= 0 \\ D^+_A \Phi &= 0. \end{aligned} \tag{13}$$

The first equation is to be interpreted as an equality between endomorphisms of W^+. The curly brackets denote the trace-free part on W^+, not on W, and F^+ denotes the projection of the curvature onto the self-dual forms, as usual.

The *moduli space* $M_\mathfrak{c}$ is the space of solutions (A, Φ) modulo the action of the gauge group, $G = \mathrm{Map}(X, S^1)$. We can also perturb the equations, rather as in the 3-dimensional case, by an arbitrary self-dual 2-form η:

$$\begin{aligned} \rho(F^+_{\hat{A}} + i\eta) - \{\Phi \otimes \Phi^*\} &= 0 \\ D^+_A \Phi &= 0. \end{aligned} \tag{14}$$

We write $M_{\mathfrak{c},\eta}$ for the solution space.

In dimension 3, for a generic perturbation, the irreducible solutions are isolated. In dimension 4, the equations have an index. We suppose X is compact and write

$$d(\mathfrak{c}) = \frac{1}{4}\big(c_1(\mathfrak{c})^2[X] - 2\chi(X) - 3\sigma(X)\big),$$

which one can also recognize as the second Chern number, $c_2(W^+)[X]$. The basic facts about the moduli space are these:

Proposition 4.1. *The moduli space $M_{\mathfrak{c},\eta}$ is compact. For an open, dense set of perturbations η, the irreducible part of the moduli space (the locus of solutions with $\Phi \neq 0$) is a smooth manifold of dimension $d(\mathfrak{c})$, cut out transversely by the equations.*

Remark. The proof of compactness runs much as in the 3-dimensional case, beginning with an essentially identical calculation leading to (4).

For the unperturbed equations, a solution with $\Phi = 0$ means a connection $\hat{A}$ in $\Lambda^2 W^+$ with anti-self-dual curvature, and hence an anti-self-dual representative for $c_1(\mathfrak{c})$. An anti-self-dual form has negative square, so there can be no solutions if $c_1(\mathfrak{c})^2[X]$ is positive. If $c_1(\mathfrak{c})^2[X]$ is zero, there can be reducible solutions only if $c_1(\mathfrak{c})$ is a torsion class, so that there is a flat connection. The same is true of the perturbed equations if η is small. Even if $c_1(\mathfrak{c})^2[X]$ is negative, however, there can only be a reducible solution if $2\pi c_1(\mathfrak{c}) - \eta$ is represented by an anti-self-dual form. The real cohomology $H^2(X;\mathbb{R})$ is the direct sum of the self-dual and anti-self-dual harmonic spaces, so the space of η for which such a representative exists is an affine subspace of codimension $b^+(X)$, the dimension of the space of self-dual forms. If $b^+(X)$ is non-zero, there is no reducible solution for generic η, and if $b^+(X)$ is at least 2, there is no solution for all η in a generic path.

Now let $\mathfrak{c}$ be a Spin^c structure with $d(\mathfrak{c}) = 0$, so that $c_1(\mathfrak{c})^2[X] = 2\chi + 3\sigma$. We shall suppose that X either has $b^+ \geq 2$ or has $2\chi + 3\sigma$ non-negative, and if $2\chi + 3\sigma$ is zero we shall also ask that let $c_1(\mathfrak{c})$ is not a torsion class. For a generic η, and also for a generic path of η, the moduli space $M_{\mathfrak{c},\eta}$ then consists of finitely many points which are transverse, irreducible solutions of the equations. The number of solutions, counted with suitable signs, is independent of the choice of perturbation and the choice of metric. We write

$$SW(X, \mathfrak{c}) \in \mathbb{Z}$$

for this number. This is the Seiberg-Witten monopole invariant for X with the Spin^c structure $\mathfrak{c}$ [42]. As before, if $SW(X, \mathfrak{c})$ is well-defined and *non-zero*, we call $c_1(\mathfrak{c})$ a *basic class* of X. Note that if $SW(X, \mathfrak{c})$ is non-zero, then the moduli space $M_{\mathfrak{c}}$ of solutions to (13) is non-empty for *every* choice of Riemannian metric.

The first significant result about basic classes was proved by Witten in [42]. This was the statement that, for a smooth algebraic surface with $b^+ > 1$ (e.g. a hypersurface in $\mathbb{CP}^3$ of degree 4 of more), the first Chern class and its negative (the canonical class) are basic classes. This was soon generalized by Taubes in [36, 37] to symplectic manifolds, in the following form. A symplectic structure ω on a manifold determines an almost-complex structure uniquely up to deformation, and hence has Chern classes $c_i(\omega)$. The *canonical class* K_ω is $-c_1(\omega)$.

Theorem 4.2 (Taubes, [36, 37]). *Let (X, ω) be a compact symplectic 4-manifold. Suppose either that $b^+ > 1$, or that $b^+ = 1$ and $K_\omega \smile [\omega]$ and K_ω^2 are both positive. Then the canonical class is a basic class.*

Remarks. Note that the hypotheses rule out $\mathbb{CP}^2$. (On a Kähler manifold, the sign of $K \smile [\omega]$ is opposite to that of the mean scalar curvature.) The statement for the case $b^+ = 1$ can be sharpened, but not without refining our treatment of the monopole invariants. A full treatment of the monopole invariants in the case $b^+ = 1$ is given in [26].

This version of the theorem is a little careless. As well as determining a canonical class K_ω, a symplectic structure gives rise to a canonical Spin^c structure $\mathfrak{c}_\omega$, with $c_1(\mathfrak{c}_\omega) = K_\omega$. Taubes result asserts that this Spin^c structure is basic, and in fact

$$SW(X, \mathfrak{c}_\omega) = \pm 1 \tag{15}$$

under the hypotheses of the Theorem.

Stretching 4-manifolds

A simple relationship between the 3- and 4-dimensional equations gives the following criterion for a class α on a 3-manifold to be a monopole class.

Proposition 4.3. *Let Y be a closed, oriented 3-manifold embedded in a compact, oriented 4-manifold X. Let α be a basic class on X. Then the restriction $\alpha|_Y$ is a monopole class on Y.*

Proof. Let $\mathfrak{c}$ be the Spin^c structure with $c_1(\mathfrak{c}) = \alpha$ and $SW(X, \mathfrak{c})$ non-zero, so that solutions to (13) exist for all metrics h on X. Let h_Y be a Riemannian on Y, and let h_1 be any metric on X such that a collar neighborhood $[-1, 1] \times Y$ carries a product metric $dt^2 + h_Y$. Let h_R be obtained from h_1 by replacing this short cylinder by a longer cylinder, $[-R, R] \times Y$, for $R > 1$. For each h_R, there exists a solution (A_R, Φ_R) on X.

The idea of the proof is to show that, as R approaches ∞, we can find a subsequence such that the corresponding solutions converge to a translation-invariant solution on some portion of the cylindrical piece. A translation-invariant solution can be interpreted as a solution of the 3-dimensional equations on Y, for the metric h_Y, so showing that a solution exists. Since h_Y is arbitrary, the class $c_1(\mathfrak{c})|_Y$ is a monopole class.

To make this work, we need first to understand the relationship between the equations in dimensions three and four. On a cylinder $[-R, R] \times Y$ with a product metric, the action of $\rho(dt)$ gives an isomorphism between W^+ and W^-. Using this isomorphism, Clifford multiplication by 1-forms orthogonal to dt become endomorphisms of W^+. In this way, Y acquires a Spin^c structure (with spin bundle $W_3 \cong W^+|_Y$) which one can call the restriction of $\mathfrak{c}$. Given a solution of the unperturbed 4-dimensional equations on the cylinder, one can apply a gauge transformation to make the dt component of $\hat{A}$ zero. If A is in such a *temporal gauge*, it can be recovered from the path $A(t)$ in space of spin connections on Y, obtained by restricting A to the slices $\{t\} \times Y$. The spinor Φ on the 4-manifold gives a path $\Phi(t)$ in the space of spinors on the 3-manifold.

In a temporal gauge, the 4-dimensional equations (13), become the following equations for the paths $A(t)$ and $\Phi(t)$,

$$\begin{aligned} \rho(\dot{A}) &= -\rho(F_A) + \{\Phi \otimes \Phi^*\} \\ \dot{\Phi} &= -D_A \Phi, \end{aligned} \tag{16}$$

in which D_A now stands for the 3-dimensional Dirac operator, and the dot is differentiation with respect to t. These equations can be recognized as the downward gradient-flow equations for the functional $\mathrm{CSD}(A, \Phi)$.

Having understood this relationship, we can complete the proof. The solution (A_R, Φ_R) on the cylinder $[-R, R] \times Y$ can now be interpreted as a gradient trajectory for the Chern-Simons-Dirac functional. The compactness properties of the equations can be used to show that the change of CSD along these trajectories is bounded, by a constant independent of R. It follows that, when R is large, there is at least some portion of the cylinder in which the change in the functional is small. Passing to a subsequence, one obtains in the limit a translation-invariant solution to the equations on the cylinder, in a temporal gauge. This is a critical point of CSD, or in other words a solution of the 3-dimensional equations. This outline is filled out in [23]. □

Remark. It is worth commenting that, if $\mathfrak{c}$ restricted to Y is trivial, the solution whose existence is established by this argument may only be the trivial solution. With some additional hypotheses however, one can establish a stronger conclusion.

Using Theorem 4.2, we can draw the following simple corollary.

Corollary 4.4. *Let Y be a closed oriented 3-manifold embedded in a closed, symplectic 4-manifold (X,ω). If $b^+(X)=1$, suppose that the hypotheses of Theorem 4.2 hold. Then $K_\omega|_Y \in H^2(Y)$ is a monopole class on Y.* □

(In fact, solutions on Y exist for the Spin^c structure $\mathfrak{c}_\omega|_Y$.) This observation, with a little ingenuity, is already enough to show that the set of monopole classes can have larger convex hull than the set of basic classes on a 3-manifold. We will need to adapt the corollary, however, before it becomes very useful. It is a puzzling question to characterize the classes $\alpha \in H^2(Y)$ which arise in this way, for a general Y.

Floer homology

There is a well-understood framework in which to place the ideas just discussed, namely the framework of 'Floer homology'. The model in the literature that is closest to what we need is Floer's construction in [11] of an invariant of 3-manifolds, using the gradient-flow of the Chern-Simons functional (of an $SU(2)$ connection).

There appears to be no serious obstacle to adapting [11] to the Chern-Simons-Dirac functional, particularly in the case that $c_1(\mathfrak{c})$ is not torsion, but there is not yet a complete account of such a construction in the literature. Nevertheless, it is clear how to proceed, and we shall content ourselves with some remarks. The starting point of [11] is the basic observation that one can calculate the homology of a compact manifold M by the following recipe. Choose a Morse function f on M whose gradient flow satisfies the additional 'Morse-Smale' condition, that the stable and unstable manifolds of all the critical points meet transversely. This means in particular that the trajectories which run from a critical point a at $t=-\infty$ to a critical point b at $t=+\infty$ form a family of dimension equal to the difference of the indices of a and b. Of these degrees of freedom, one is the freedom to reparametrize the trajectory $\gamma(t)$ as $\gamma(t+c)$. If a and b have index differing by 1, the trajectories are isolated once one forgets the parameterization. Now form the vector space C with a basis e_a indexed by the critical points a, and define a linear transformation ∂ by

$$\partial(e_a)=\sum n_{ab}e_b,$$

where n_{ab} counts the number of trajectories from a to b in the case that $i(b)=i(a)-1$ and is zero otherwise. To avoid questions of orientation, one can take $\mathbb{Z}/2$ as the field of coefficients for C. Then one shows that $\partial^2 = 0$ and the $\ker\partial/\mathrm{im}\partial$ is the homology of M. In particular it is independent of the choice of f and the choice of Riemannian metric used to define the gradient. Without having an alternative definition of the homology however, one can verify this independence directly.

Floer applied this construction in an infinite-dimensional setting, taking the Chern-Simons functional as f. In the monopole setting, one should use the functional CSD on the space $\mathcal{C}$. The situation is simplest in the case that $c_1(\mathfrak{c})$ is not torsion, so that the monopole equations on Y have no solution with $\Phi = 0$. As mentioned above, for a suitable exact 2-form μ, the perturbed functional CSD_μ has non-degenerate critical points, and these are a finite set. Although there is no well-defined difference of indices between a pair of critical points, we can measure the index difference between a and b along a given trajectory γ, as the spectral flow of the Hessian, as before. If a Morse-Smale condition is satisfied, we can then construct C and ∂ as before, defining n_{ab} as the number of trajectories whose spectral flow is 1.

Unfortunately, we cannot expect to achieve the stronger Morse-Smale condition by such a restricted class of perturbations as the addition of an exact 2-form μ. One must seek a larger class of perturbations. At the same time, it is a particular property of the equations involved that the spaces of trajectories have any reasonable compactness properties (as noted in the previous subsection, the gradient trajectories can be interpreted as solutions of the 4-dimensional monopole equations), and one must choose the perturbations of CSD_μ so as not to upset this feature. We make some remarks in the following subsection about how one might define a suitable larger class of perturbations, and for the moment we shall pass over this point. It is precisely here that more work needs to be done to carry through the Floer program for the monopole equations.

After taking care of perturbations and compactness, one should arrive by this construction at a vector space $HF(Y,\mathfrak{c})$ with $\mathbb{Z}/2$ coefficients and an even-odd grading. It should depend only on Y and $\mathfrak{c}$, not on the choice of Riemannian metric or the perturbation chosen for the equations. The construction makes clear that $HF(Y,\mathfrak{c})$ is zero if the original monopole equations, or their perturbation, have no solution. Also, the euler characteristic of $HF(Y,\mathfrak{c})$ (the difference of the odd and even betti numbers) is equal to the integer invariant $SW(Y,\mathfrak{c})$.

The usefulness of Floer homology in the present context is that Proposition 4.3 and Corollary 4.4 can be strengthened, so as to conclude that the

Floer homology is non-zero. (For the 4-dimensional Yang-Mills invariants, this role for the instanton Floer homology of [11] was first noted by Donaldson.) For example, if Y is embedded in a closed symplectic 4-manifold (X, ω) (satisfying the hypotheses of Corollary 4.4 in the case $b^+(X) = 1$), and if $K_\omega|_Y$ is not torsion, then one would conclude that $HF(Y, \mathfrak{c}_\omega|_Y)$ is non-zero, once the definition of this Floer homology was in place.

Perturbing the gradient flow

It may be worth noting that perturbing the Chern-Simons-Dirac functional so as to achieve a Morse-Smale condition for the trajectories of the gradient flow may not be particularly difficult. Let Y be a closed Riemannian 3-manifold with Spin^c structure $\mathfrak{c}$, and let $\mathcal{S}$ be the space of all pairs (A, Φ), where A is a spin connection and Φ is a section of W. The space $\mathcal{C}$ above is the quotient $\mathcal{S}/G$.

Let A_0 be a fixed spin connection, so that we can write the general spin connection as

$$A = A_0 + a\,\mathbf{1},$$

so identifying the space of spin connections with the space of imaginary-valued 1-forms a, as before. We have already considered adding to CSD a function of the form

$$\tau_\mu(a) = i \int a \wedge \mu$$

for an exact 2-form μ. Let us now relax the requirement slightly, and suppose only that μ is *closed.* Take a collection $\mu_1, \ldots, \mu_N$ of closed 2-forms, which included a basis for $H^2(Y)$, and let $\tau_1, \ldots, \tau_N$ be the corresponding functions of a. As noted earlier, the functions τ_i on $\mathcal{S}$ are invariant only under the identity component $G_e \subset G$. The map

$$(\tau_1, \ldots, \tau_N) : \mathcal{S} \to \mathbb{R}^N$$

commutes with the G action, however, when G acts on $\mathbb{R}^N$ through a discrete action of the quotient, $G/G_e \cong H^1(Y; \mathbb{Z})$, by translations. So to obtain a G-invariant function on $\mathcal{S}$, we should take a function

$$F : \mathbb{R}^N \to \mathbb{R}$$

which is invariant under these translations, and define

$$f(a,\Phi) = F\big(\tau_1(a), \dots, \tau_N(a)\big).$$

(Functions of this shape include the 'smoothed-out' holonomy maps used in [11] for non-abelian connections.)

Our function f is not yet sufficiently general, for it does not depend on Φ. To incorporate Φ, we can proceed as follows. Let L be the Greens operator for the ordinary Laplacian on $C^\infty(Y)$. Thus L inverts the Laplacian when restricted to functions with zero mean, and the kernel and cokernel of L are the constant functions. Let $H \subset G_e$ be the subgroup

$$H = \{\, e^{i\xi} \mid \xi : Y \to \mathbb{R}, \text{ with } \textstyle\int \xi = 0 \,\}.$$

The quotient G_e/H is the circle, represented by the constant maps, and G/H is represented by the harmonic maps $Y \to S^1$, a group isomorphic to $S^1 \times H^1(Y;\mathbb{Z})$.

For any fixed spinor $\psi \in \Gamma(W)$, let σ_ψ be the complex function on $\mathcal{S}$ defined by the hermitian inner product

$$\sigma_\psi(a,\Phi) = \int \langle \psi^a, \Phi\rangle,$$

where ψ^a denotes the expression

$$\psi^a = e^{-Ld^*a}\psi.$$

The definition of ψ^a is such that it transforms as Φ does under the action of H: if $u = e^{i\xi}$, where ξ has mean zero, then

$$\begin{aligned}\psi^{a-u^{-1}du} &= \psi^{a-id\xi}\\ &= e^{-Ld^*a}e^{iLd^*d\xi}\psi\\ &= u\psi^a.\end{aligned}$$

Thus $\sigma_\psi : \mathcal{S} \to \mathbb{C}$ is invariant under H. Under the circle G_e/H, however, it transforms with weight 1.

Now choose a collection of spinors ψ_i ($i = 1,\dots K$), and let σ_i be the corresponding complex functions on $\mathcal{S}$. We now have a collection of functions

$$(\tau_1, \dots, \tau_N, \sigma_1, \dots, \sigma_K) : \mathcal{S} \to \mathbb{R}^N \times \mathbb{C}^K.$$

These functions are equivariant for G when G/H is made to act suitably on $\mathbb{R}^N \times \mathbb{C}^K$. The quotient of $\mathbb{R}^N \times \mathbb{C}^K$ by this action is a bundle with fiber $\mathbb{C}^K/S^1$ over the base $T^b \times \mathbb{R}^{N-b}$, where b is the Betti number of Y.

Now choose a smooth function

$$F : \mathbb{R}^N \times \mathbb{C}^K \to \mathbb{R}$$

which is invariant under the action of G/H, and define

$$f(a, \phi) = F(\tau_1, \dots \tau_N, \sigma_1, \dots, \sigma_K).$$

Now consider perturbing CSD by the addition of such a function f. With terms of this sort we can C^1-approximate, for example, any smooth function on a compact submanifold of $\mathcal{C}$ (lying in the locus where Φ is non-zero), and quite formally the class of functions is large enough to give the necessary transversality. If the partial derivatives of F are bounded, it seems that the compactness theorems for spaces of trajectories hold up too. Thus, in the crucial calculation (4), one finds amongst other things an additional cubic term in Φ, which is local in the t coordinate, but non-local on Y, involving an expression of the shape

$$\Phi^* \rho\big(d \circ L\langle \psi^a, \Phi \rangle\big)\Phi.$$

But such a term does not break the argument.

5 Monopoles and contact structures

We now turn to the proof of Theorem 3.6. In view of Corollary 4.4, one might hope to prove this by showing that if Y had a taut foliation $\mathcal{F}$, then one could always embed Y in a closed symplectic manifold (X, ω) in such a way that $c_1(\omega)$ restricted to Y was $\pm e(\mathcal{F})$. Perhaps this can be done; such a result would be very interesting, and presumably very hard. A slight modification of this tactic leads to a proof, however. Applying a theorem of Eliashberg and Thurston [8], we shall embed Y in an *open* symplectic manifold whose ends have a cone-like geometry. We shall then extend the 4-dimensional gauge-theory techniques to this setting.

Using the theorem of Eliashberg and Thurston

The following material can be found in [8]. To begin, the following proposition sheds a geometric light on the meaning of taut. (The converse to the proposition is true also, but is rather deeper.)

Proposition 5.1. *If $\mathcal{F}$ is a taut foliation of Y by oriented leaves, then there is a closed 2-form Ω on Y whose restriction to the leaves is positive.*

Proof. Let γ be a closed curve, transverse to the leaves and meeting every leaf. (The existence of γ was our chosen definition of taut.) A small tubular neighborhood $N(\gamma)$ meets the leaves of $\mathcal{F}$ in a foliation by disks, to give a product structure. Using the product structure, pull back to $N(\gamma)$ a 2-form ψ supported in the interior of the disk and non-negative there. The result is a closed form $\Omega(\gamma)$ which is non-negative on the leaves of $\mathcal{F}$. By pushing γ along the leaves, one obtains transverse curves running through any point of Y. By taking a suitable finite collection of such curves γ_i and adding up the corresponding forms $\Omega(\gamma_i)$, one obtains a suitable Ω. □

Now let α be a non-vanishing 1-form on Y whose kernel at each point is the tangent plane to $\mathcal{F}$ and whose orientation is such that $\alpha \wedge \Omega$ is positive. The integrability of the tangents to the foliation means that $\alpha \wedge d\alpha$ is zero. On the cylinder $[-1,1] \times Y$, consider the closed 2-form

$$\omega = d(t \wedge \alpha) + \Omega.$$

In ω^2, the only term to survive is the terms $dt \wedge \alpha \wedge \Omega$, which is positive. So ω is a symplectic form.

We have succeeded in embedding Y in a symplectic 4-manifold with boundary, $[-1,1]\times Y$, and it is not hard to see that the first Chern class $c_1(\omega)$ restricts to $e(\mathcal{F})$ on the 3-manifold. But this elementary step is insufficient for our needs.

A *contact structure* on a 3-manifold is a field of 2-planes ξ which strictly fails to be integrable at every point of Y. This means that, if ξ is defined locally as the kernel of a 1-form β, then $\beta \wedge d\beta$ is nowhere zero. If the 3-manifold and the 2-plane field are oriented (as will always be the case for us), then a suitable form β exists globally. Note however that, as a non-vanishing 3-form, the product $\beta \wedge d\beta$ itself determines an orientation of Y. We shall say that the contact structure ξ is compatible with the orientation of Y if the form $\beta \wedge d\beta$ is positive, for some and hence for all choices of β with kernel

ξ. The theorem of Eliashberg and Thurston which we need states that the tangent planes to a foliation can be deformed to give a contact structure, compatible with either orientations:

Theorem 5.2 (Eliashberg-Thurston [8]). *Let $\mathcal{F}$ be a smooth, oriented foliation of an oriented 3-manifold, other than the foliation of $S^1 \times S^2$ by spheres. Then the 2-plane field $T\mathcal{F}$ can be C^0 approximated by contact structures ξ compatible with either the given orientation of Y or its opposite.* □

(An example of this phenomenon arises when Y is a circle bundle over a surface arising as a compact left quotient of $SL(2,\mathbb{R})$. A left-invariant 2-plane field is determined by a 2-plane π in the Lie algebra $\mathfrak{sl}(2,\mathbb{R})$. If π is tangent to the null cone of the Killing form, then the corresponding 2-plane field is a foliation. If the Killing form is either definite or hyperbolic on π, then the 2-plane field is a foliation, compatible with one or other orientation of the 3-manifold.)

Now let us return to the manifold $X = [-1,1] \times Y$ with the symplectic form ω constructed above. The oriented boundary of X is $\bar{Y} \cup Y$ (the bar denotes the opposite orientation), and $\mathcal{F}$ foliates both components. Using the Theorem, one obtains foliations ξ_- and ξ_+ on $\bar{Y}$ and Y, compatible with their respective orientations, and at a small angle from the tangents to $\mathcal{F}$. If the angle is made small enough, we can arrange that these contact structures are *compatible* with ω, in the weak sense that ω is positive on the 2-planes at the boundary:

$$\omega|_\xi > 0 \quad \text{at } \partial X. \tag{17}$$

Indeed, ω is positive on the tangent planes to $\mathcal{F}$, which $\xi_\pm$ approximate. (To clarify the signs involved, this sort of compatibility holds for the standard, Kähler, symplectic form on a pseudo-convex domain in $\mathbb{C}^2$, such as a ball, when the boundary is given the contact structure defined by the complex tangent directions.)

To summarize, starting from a 3-manifold Y with a foliation $\mathcal{F}$, we have constructed a symplectic 4-manifold (X,ω) in which Y is embedded, with $K_\omega|_Y$ equal to the euler class $e(\mathcal{F})$. The 4-manifold has a contact structure $\xi = \xi_- \cup \xi_+$ on its boundary (compatible with the boundary orientation, in a 'convex' direction), and the symplectic form is compatible with ξ, in the sense described by (17).

Four-manifolds with contact boundary

Although we have not embedded Y in a closed manifold, the convex contact structure on the boundary of X is all we shall need, because we can extend the monopole invariants $SW(X, \mathfrak{c})$ for closed manifolds so as to define similar invariants for 4-manifolds with contact boundary.

We give an account of the construction from [21]. Let X be a compact, connected, oriented 4-manifold with non-empty boundary ∂X, and let ξ be an oriented contact structure on ∂X, compatible with the boundary orientation. In the presence of a metric, any oriented 2-plane field such as ξ determines a Spin^c structure on ∂X, and hence a 4-dimensional Spin^c structure on a collar of the boundary. One can think of this in various ways. For example, define the spin bundle W on ∂X can be defined as the sum $\mathbb{C} \oplus \xi^c$, where the second summand means that the oriented 2-planes of ξ are being regarded as complex lines; then define Clifford multiplication at a point y by picking a basis e_1, e_2, e_3 of tangent vectors at y, with e_1 the positive normal to ξ and declaring that these act on $(\mathbb{C} \oplus \xi^c)_y$ by the Pauli matrices

$$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix},$$

using the basis vector e_2 to trivialize ξ. Alternatively, one can think of this as a special case of the way in which an almost-complex structure determines a Spin^c structure in even dimensions. Note that the spin bundle comes with a canonical section $\Phi_0 = (1, 0)$.

Now let $\mathfrak{c}$ be any extension of $\mathfrak{c}_\xi$ to the interior of X. Given such an extension, we shall define a monopole invariant

$$SW(X, \xi, \mathfrak{c}) \in \mathbb{Z},$$

which is a diffeomorphism invariant of the triple (no condition on $b^+(X)$ is needed).

The invariant $SW(X, \xi, \mathfrak{c})$ is defined as follows. First we enlarge the manifold X by adding expanding cones to the boundary components. In more detail, if Y is a 3-manifold with a positive contact structure defined as the kernel of a 1-form β, then there is a symplectic form on the cone $(0, \infty) \times Y$, given as

$$\omega = d(f(t)\beta)$$

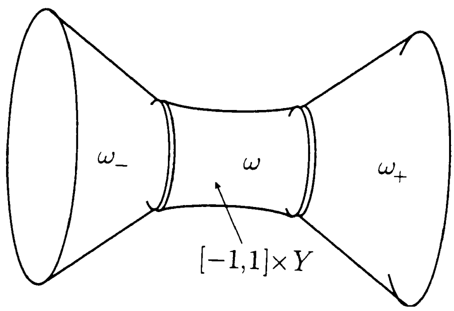

Figure 2: The geometry of Z.

for any monotone increasing function f of $t \in (0,\infty)$. To reproduce the way in which $\mathbb{R}^4$ with its standard symplectic structure arises from S^3, we prefer to set

$$\omega = (1/2)d(t^2\beta).$$

We apply this standard construction to the components of ∂X, and attach conical pieces $[1,\infty) \times \partial X$ to the boundary, to obtain an open 4-manifold Z (diffeomorphic to the interior of X):

$$Z \;=\; X \;\cup\; [1,\infty) \times \partial X.$$

On Z we choose a Riemannian metric h compatible with ω on the conical pieces. This means that there are local orthonormal coframes in which ω can be expressed as $e^1 \wedge e^2 + e^3 \wedge e^4$. Figure 2 shows an illustration of Z in the case that X is the manifold $[-1,1] \times Y$ from the previous subsection. In this case, the contact structures ξ_- and ξ_+ give symplectic forms $\omega_\pm$ on the two conical ends.

The symplectic structure ω on the conical pieces determines a canonical Spin^c structure $\mathfrak{c}_\omega$ there, essentially the same as the Spin^c structure determined by ξ on the boundary of X. The choice of $\mathfrak{c}$ on X gives an extension of $\mathfrak{c}_{omega}$ to all of Z. The spin bundle $W^+ = W^+_{\mathfrak{c}}$ has a canonical section Φ_0 of unit length on the conical pieces of Z, and there is a unique spin connection there, with the property that $D^+_{A_0}\Phi_0 = 0$. (Such a connection is determined by any non-vanishing spinor on a 4-manifold.) We extend Φ_0 and A_0 arbitrarily over the remained of Z.

Motivated by the constructions of [36, 37], we now consider a modified version of the Seiberg-Witten monopole equations on the Riemannian manifold Z. The equations are

$$\begin{aligned} \rho(F^+_{\hat{A}}) - \{\Phi \otimes \Phi^*\} &= \rho(F^+_{\hat{A}_0}) - \{\Phi_0 \otimes \Phi_0^*\} \\ D^+_A\Phi &= 0. \end{aligned} \tag{18}$$

We can also consider, as before, perturbing these equations by the addition of a self-dual 2-form η, which should decay on the ends of Z:

$$\begin{aligned} \rho(F^+_{\hat{A}} + i\eta) - \{\Phi \otimes \Phi^*\} &= \rho(F^+_{\hat{A}_0}) - \{\Phi_0 \otimes \Phi_0^*\} \\ D^+_A\Phi &= 0. \end{aligned}$$

The unperturbed equations are set up so that (A_0, Φ_0) satisfies the equations on the conical ends, though not necessarily in the interior, since Φ_0 may not satisfy the Dirac equation there. We seek solutions (A, Φ) in general which are asymptotic to (A_0, Φ_0) at infinity. It turns out that, even if only mild decay is required of $A - A_0$ and $\Phi - \Phi_0$, any solution of the unperturbed equations will approach the canonical pair exponentially fast after adjustment by a suitable gauge transformation.

We again write $M_{\mathfrak{c}}$ (or $M_{\mathfrak{c},\eta}$ in the perturbed case) for the set of such solutions, considered up to the equivalence relation defined by the gauge transformations. Note that there can be no 'reducible' solutions with $\Phi = 0$, because Φ is required to approach the unit-length spinor Φ_0 at infinity. The main facts about the moduli space are these:

Proposition 5.3 ([21]). *The space $M_{\mathfrak{c},\eta}$ is compact, and for generic η it is a smooth manifold cut out transversely by the equations. In this case, the dimension of the moduli space is given by*

$$d(\mathfrak{c}) = c_2(W^+_{\mathfrak{c}}, \Phi_0)[Z, Z \setminus X],$$

(which is the relative euler class of the bundle $W^+_{\mathfrak{c}}$ on Z, relative to the non-vanishing section Φ_0 on the ends).

(The formula for the dimension coincides with the alternative formula $c_2(W^+[X]$ which we gave in the closed case.)

In the case $d(\mathfrak{c}) = 0$, we can again count the number of solutions, either with suitable signs, or just modulo 2, to obtain a definition of the invariant $SW(X, \xi, \mathfrak{c})$. It is again independent of the choices made, such as the 1-form β, the metric h, and the perturbation η.

Symplectic filling

In the definition of $SW(X, \xi, \mathfrak{c})$ as just described, no symplectic form on X is involved and none is needed in the construction. When one has a symplectic structure on X compatible with ξ at the boundary, then there is a non-vanishing theorem for the invariant. The symplectic form ω determines a canonical Spin^c structure $\mathfrak{c}_\omega$ on X, and the compatibility condition with ξ means in particular that $\mathfrak{c}_\omega$ and $\mathfrak{c}_\xi$ are the same at the boundary, so and invariant $SW(X, \xi, \mathfrak{c}_\omega)$ is defined. The theorem is then:

Theorem 5.4 ([21]). *If X is a compact 4-manifold with contact structure ξ on the boundary, and ω is a symplectic form on X compatible with ξ, then*

$$SW(X, \xi, \mathfrak{c}_\omega = \pm 1.$$

□

Once the analytic framework of the previous subsection is in place, the proof of this result is very much as the same as the proof of (15) from [37]. There is one point to note, however. The proof begins by constructing a symplectic form ω_Z on Z, the manifold with conical ends. Although Z is a union of pieces each of which carries a symplectic form (as illustrated in Figure 2 for the case of $[-1, 1] \times Y$), these forms do not necessarily agree at the joins, and even their cohomology classes may be different. Nevertheless, one can patch the forms together, in that one can find a symplectic form ω_Z which is *asymptotic* to the conical form on the ends, and agrees with the given form ω on X except in a small neighborhood of ∂X. See [21], for example.

The stretching argument used in Theorem 4.3 and Corollary 4.4 works just as well in the present setting of 4-manifolds with contact boundary. From the above theorem, we can therefore deduce:

Corollary 5.5. *Let (X, ω) be a symplectic 4-manifold with a compatible contact structure ξ on the boundary. If Y is an oriented 3-manifold embedded in X, then $K_\omega|_Y$ is a monopole class.* □

Using the construction of Eliashberg and Thurston, with $X = [-1,1]\times Y$, we deduce Theorem 3.6:

Corollary 5.6. *If Y is a closed, irreducible, oriented 3-manifold with a smooth, taut foliation $\mathcal{F}$ by oriented leaves, then $e(\mathcal{F})$ is a monopole class.* □

(Again, the hypothesis of irreducibility excludes $S^1 \times S^2$.) Slightly more generally, one can apply Corollary 5.5 to the case that Y is parallel to ∂X or a component of ∂X. In Eliashberg's terminology, [8], a contact 3-manifold Y is *symplectically fillable* if Y arises as the correctly oriented boundary of a 4-manifold X carrying a compatible symplectic form ω. If Y arises as a union of components of such a boundary, then it is symplectically *semi*-fillable.

Corollary 5.7. *If a contact 3-manifold (Y, ξ) is symplectically semifillable, then $e(\xi)$ is a monopole class.* □

One useful feature of this corollary is one can form a connected sum of semifillable contact structures, and so obtain results about monopole classes on reducible 3-manifolds also.

All these statements can be sharpened a little, because one knows which Spin^c structure is involved, not just the Chern class. More significantly perhaps, one should draw the stronger conclusion that the Floer homology is non-zero in these cases. For example, if (Y, ξ) is symplectically semifillable and e_ξ is not torsion, one would conclude that $HF(Y, \mathfrak{c}_\xi)$ is non-zero. (A more refined statement could be made in the case of a torsion class.) In particular then, one should say:

5.8. *If Y carries a taut foliation, then the Seiberg-Witten Floer homology of Y is non-zero for the corresponding* Spin^c *structure.*

The obstruction to proving such statements is no larger than the problem of verifying a suitable construction of Floer homology.

Invariants of contact structures

There is a way to rephrase part of the construction just described. Given an oriented 3-manifold Y and a contact structure ξ compatible with the orientation, one can form a symplectic cone $[1,\infty)\times Y$ with symplectic form ω_+ as before, and attach to it a cylinder $(-\infty, 1]\times Y$, as shown in Figure 3.

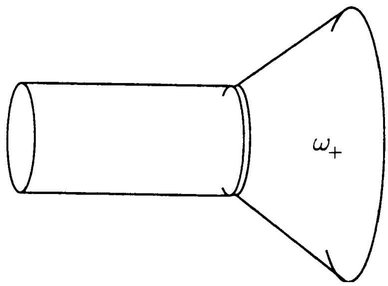

Figure 3: Defining an invariant of (Y, ξ).

Using the Spin^c structure given by ξ, one can then write down a version of the monopole equations on this 4-manifold which resemble the deformed equations (18) on the conical piece and resemble the usual equations (13) on cylindrical end, where they can be interpreted as the gradient flow equations for a trajectory of CSD.

For each critical point a of the functional CSD on Y (or of the perturbed functional CSD_μ considered before), one can look at the moduli space of solutions (A, Φ) which are asymptotic to (A_0, Φ_0) on the cone, and which descend from the critical point a on the cylinder. After perturbation, these moduli spaces are smooth manifolds. For each α, let $n(a, \xi)$ be the number of solutions belonging to zero-dimensional moduli spaces, counted with signs as usual, and consider the expression

$$\sum_a n_a e_a$$

as an element of the chain group C which defines Floer homology. If we suppose that $e(\xi)$ is not torsion, to eliminate the reducible solutions, then the above sum is closed under ∂ (by arguments that are familiar in other applications of Floer homology), and the resulting homology class

$$[\sum_a n_a e_a] \in HF(Y, \mathfrak{c}_\xi) \tag{19}$$

is an invariant of the contact 3-manifold.

In this way, one can define an invariant of contact structures, once Floer homology is in hand. These seems the most natural setting in which to place the constructions of [28], where collections of contact structures are exhibited which are homotopic as 2-plane fields but not isotopic as contact structures. (The manifolds Y in [28] are homology spheres, so one needs to tackle Floer homology for the case that $c_1(\mathfrak{c})$ is zero.)

The non-vanishing result, Corollary 5.7, should be rephrased so as to say that the invariant (19) of (Y, ξ) is a non-zero element of HF if the contact structure is semi-fillable.

6 Potential applications

When the Seiberg-Witten equations were introduced as an alternative gauge theory tool to replace the self-dual Yang-Mills equations exploited by Donaldson, an important link with topology was temporarily lost. The 3-dimensional companions of Donaldson's Yang-Mills invariants are, as we have mentioned, the instanton Floer homology $I(Y)$ introduced in [11], and the Casson invariant $\lambda(Y)$ [1]. These play the roles of $HF(Y, \mathfrak{c})$ and $SW(Y, \mathfrak{c})$ from the monopole theory. Floer defined $I(Y)$ for homology 3-spheres Y by studying the gradient-flow of the $SU(2)$ Chern-Simons functional, whose critical points correspond to flat $SU(2)$ connections, or representations of $\pi_1(Y)$ in $SU(2)$. There is also a version $I_w(Y)$ for the case that Y is a homology $S^1 \times S^2$, where the chain group is built from flat $SO(3)$ connections with non-zero Stiefel-Whitney class [3]. Defining $I(Y)$ in other situations presents technical difficulties related to reducible connections.

In the definition of $I(Y)$, there is an immediate connection with the fundamental group. Thus, for example, Floer's instanton homology vanishes if Y is a homotopy sphere (Floer's definition did not use the trivial representation of π_1). No such statement can be made very easily for the monopole Floer homology (our tentatively defined $HF(Y, \mathfrak{c})$).

On the other hand, in the Seiberg-Witten version, we have a handle on anon-vanishing result: irreducible 3-manifolds with $b_1 \neq 0$ admit taut foliations by Gabai's results, and we have argued that a foliation forces $HF(Y, \xi)$ to be non-zero. If one could establish even a weak relationship between the monopole $HF(Y, \mathfrak{c})$ and instanton Floer homology, then there would be a useful payoff.

Surgery and property 'P'

To elaborate on the last remark, the application we have in mind is in line with Casson's application of the invariant $\lambda(Y)$ to the question of 'property P'. This is the question of whether one can make a simply connected 3-manifold by non-trivial surgery on a non-trivial knot K. (If one cannot, then K is said to have property P.) Work of Gordon an Luecke [18] showed that one cannot make S^3 this way. So if one did manufacture a homotopy-sphere by surgery on a knot, it would be a fake 3-sphere (a counterexample to the Poincaré conjecture).

The basic example to consider is $+1$ surgery, in which a neighborhood $N(K)$ is removed and sewn back in so that the meridian on $N(K)$ is attached to a curve in the class of the meridian plus longitude. Let us call this manifold Y_1, or $Y_1(K)$. We also have the manifold Y_0 obtained by zero-surgery, which is a homology $S^1 \times S^2$. To prove property P, we would like to know that Y_1 has non-trivial fundamental group, and to this end we could seek to show that $\pi_1(Y_1)$ had non-trivial representations in $SU(2)$, or that $I(Y_1)$ is non-zero.

There is a powerful tool at hand, in the following theorem of Floer. (This is a special case of his 'exact triangle' in instanton homology [3]).

Theorem 6.1 (Floer, [3]). *In the above situation the instanton Floer homology $I(Y_1)$ of the homology sphere Y_1 is isomorphic to the instanton Floer homology $I_w(Y_0)$ of the homology $S^1 \times S^2$ obtained by zero-surgery.* □

In this form, the theorem is already hard to prove. But the special case we are interested is actually very *easy* to prove:

Corollary 6.2. *If the instanton Floer homology $I_w(Y_0)$ is non-trivial, then Y_1 is not a homotopy 3-sphere.* □

(In fact, if Y_1 is a homotopy 3-sphere, then one can easily use the holonomy perturbations introduced by Floer to deform the equations $F_A = 0$ on Y_0 so that they have no solutions for $SO(3)$ connections A on Y with non-zero w_2.)

Suppose now that we could prove:

Conjecture 6.3. *If the monopole Floer homology $HF(Y, \mathfrak{c})$ of a homology $S^1 \times S^2$ is non-trivial for some Spin^c structure with $c_1(\mathfrak{c})$ not torsion, then the instanton Floer homology $I_w(Y)$ is non-trivial also.*

Then we would be home, at least in the case that the genus of K is 2 or more. Indeed, we have learned from the foliation theory that in this case, Y_0

admits a smooth, taut foliation having a genus 2 surface has a compact leaf. The euler class of this foliation is non-trivial, and modulo the verification of the definitions, the monopole Floer homology of Y_0 for the corresponding Spin^c structure has been shown to be non-zero. The conjecture would imply that instanton Floer homology is non-trivial also, and it would follow that Y_1 was not a homotopy sphere, by Floer's result. (In the case of genus 1, there are two additional difficulties. The first is that Gabai's foliation may not be smooth. The second is that the relevant Spin^c structure on Y_0 has $c_1 = 0$, so one must treat reducibles with respect. The second point is moot, perhaps, because one would need to consider all Spin^c structures together, most likely, to prove the conjecture above.)

Note that, without using any connection between the instanton and monopole Floer homologies, one could try and reprove the theorem of Gordon and Luecke by establishing an exact triangle for the monopole Floer homology.

The Pidstrigatch-Tyurin program

Conjecture 6.3 does not stand unsupported. Indeed, there is evidence for a closer relationship. In introducing the monopole invariants of 4-manifolds, Witten [42] conjectured a very specific relationship between the monopole invariants and the older instanton invariants, for a large class of 4-manifolds. (This conjecture is extended in [32].) A mathematical approach to proving Witten's conjecture was proposed by Pidstrigatch and Tyurin [33], and although their program does raise some technical challenges at the time of writing, it seems most likely that a proof will be along the lines suggested. There is work in this direction in [9]. The proposed method is to study a larger moduli space of solutions to the $PU(2)$ monopole equations. These equations have the same shape as the equations (13) but involve a non-abelian connection A. The equations contain both the instanton equations and the usual monopole equations.

If one pursues this line in three dimensions instead of four, one can arrive at an elegant proof that the 3-manifold invariants we have called $SW(Y, \mathfrak{c})$ are related to the 'odd' Casson invariant. For example, if Y is obtained by zero surgery on a knot K, there is a Casson-type invariant $\lambda(Y)$ which is the euler characteristic of the instanton Floer homology $I_w(Y)$, and without encountering any of the technical difficulties of the 4-dimensional case, one can prove that

$$\lambda(Y) = \sum_{s} SW(Y, \mathfrak{c}).$$

Of course, one can verify this relationship externally, because both sides can be reduced to the Alexander invariant. (In the case of the left-hand side, this is due to Casson.) But the 'internal' proof using the $PU(2)$ equations seems to show that the relationship extends beyond the euler characteristics. It may well be that one can relate the Floer homologies using this approach.

7 Surfaces in 4-manifolds

Where the theory succeeds

We turn now to the question of representing a 2-dimensional homology class σ in a smooth, oriented 4-manifold X by a smoothly embedded, oriented surface Σ of minimal complexity. There are many statements here that can be made to look very much like their 3-manifold counterparts, particularly when Σ has trivial normal bundle. (This condition implies that the self-intersection number $\sigma \cdot \sigma$ is zero, but the converse is false, because we still allow that Σ may be disconnected.) For example, if $\sigma \cdot \sigma$ is zero, then $\mathrm{PD}[\sigma] \in H^2(X)$ is the pull-back of the generator of $H^2(S^2)$ by some map $f : X \to S^2$, and one can find a representative surface Σ as the inverse image of a regular value of f, so establishing that representative surface exists. Every embedded surface with trivial normal bundle arises this way, just as a surface in a 3-manifold Y arises from a map to S^1.

We have already introduced the 4-dimensional monopole equations and the basic classes, at least in the case that $b^+(X) > 1$. For surfaces with trivial normal bundle, the basic classes provide a lower bound for the genus, just as in dimension three (compare Proposition 3.1):

Proposition 7.1. *Let X be a smooth, oriented, closed 4-manifold with $b^+(X) > 1$, let α be a basic class and Σ an embedded surface representing a class σ. Suppose that the normal bundle of Σ is trivial. Then*

$$\chi_-(\Sigma) \geq \langle \alpha, \sigma \rangle.$$

Proof. Being a basic class means for α that there is a Spin^c structure $\mathfrak{c}$ with $c_1(\mathfrak{c}) = \alpha$ for which the invariant $SW(X, \mathfrak{c})$ is non-zero. All we need, however, is that the moduli space $M_{\mathfrak{c}}$ of solutions to the equations (13) is non-empty, for every choice of Riemannian metric on X. Just as in the 3-dimensional case (see (5)), if (A, Φ) is a solution for a given metric, then the Weitzenbock formula leads to

$$\int_X |\Phi|^4 d\text{vol} \leq \int_X s^2 d\text{vol}.$$

In dimension 4, the first of the two Seiberg-Witten equations leads to the relationship $|F^+_{\hat{A}}|^2 = (1/8)|\Phi|^4$, so there is a relationship between the L^2 norms,

$$\|F^+_{\hat{A}}\|^2 \leq \|s\|^2/8. \tag{20}$$

Now let $\|\alpha\|$ stand for the L^2 norm of the harmonic representative, as before. We have

$$\|\alpha\| \leq (1/2\pi)\|F_{\hat{A}}\|$$

On the other hand, alpha is the orthogonal sum of its self-dual and anti-self dual parts, and

$$\|\alpha^+\|^2 - \|\alpha^-\|^2 = (\alpha \smile \alpha)[X],$$

so from the inequality (20) we obtain the bound

$$\|\alpha\|_h^2 \leq \|s_h\|_h^2/(4\pi)^2 + (1/2)\alpha^2[X],$$

in which we have once again adjusted our notation to indicate the dependence on a Riemannian metric h. We can write

$$\|\alpha\|_h \leq \|s_h\|_h/(4\pi) + C, \tag{21}$$

for a constant C which is independent of h, to reach an inequality that we can compare with (7).

Now let Σ be an embedded surface representing σ, as in the statement. We may assume that Σ is connected, because the general case follows by linearity. Since σ^2 is zero, the normal bundle of Σ is trivial, and we can find embedded in X a product region $[0,1] \times S^1 \times \Sigma$, as a collar on the boundary of a tubular neighborhood of Σ. Let h_1 be a metric on X whose restriction to this region is a product metric, with the metric on the Σ factor being of unit area and constant non-negative scalar curvature. (We can suppose Σ is not a sphere.) Then we can do just what we did in dimension three (see Lemma 9). Let h_r be a metric on X which contains a product region $[0,r] \times S^1 \times \Sigma$ and coincides with h_1 outside this region. Exactly as before, we can calculate for $h = h_r$,

$$\|s_h\|_h = 4\pi r^{1/2}(2g-2) + O(1)$$

as $r \to \infty$, while

$$\|\alpha\|_h \geq r^{1/2}\langle \alpha, [\Sigma]\rangle.$$

These two estimates are inconsistent with the inequality (21), unless $2g-2 \geq \langle \alpha, \sigma\rangle$. □

One can pass from the proposition above to a statement about surfaces with non-negative normal bundle (that is, surfaces Σ such that each component has non-negative self-intersection number):

Proposition 7.2. *Let X be a smooth, oriented, closed 4-manifold with $b^+(X) > 1$, let α be a basic class and Σ an embedded surface representing a class σ. Suppose that the normal bundle of Σ is non-negative. Then*

$$\chi_-(\Sigma) \geq \sigma \cdot \sigma + \langle \alpha, \sigma\rangle.$$

Proof. The statement can be deduced from the case of trivial normal bundle by using the 'blow-up formula' for the 4-dimensional monopole invariants. It is enough once more to consider the case that Σ is connected. Let $k = \sigma \cdot \sigma$ and et $X_k = X \# k\bar{\mathbb{CP}}^2$ be the connected sum of X and k copies of $\mathbb{CP}^2$ with reversed orientation. Put

$$\alpha_k = \alpha + e_1 + \cdots + e_k, \tag{22}$$

where e_i is the generator of H^2 in the ith copy of $\bar{\mathbb{CP}}^2$. The blow-up formula ([19], Proposition 2) says that α_k is a basic class on X_k if α is a basic class on X, because the monopole invariants of corresponding Spin^c structures are equal. Let $\tilde{\Sigma}$ be the embedded surface in X_k formed by an internal connected sum of Σ with the spheres representing the generators of homology in the $\bar{\mathbb{CP}}^2$ summands. (We orient these spheres so that the class e_i evaluates as $+1$ on the ith sphere, and we form the sum respecting these orientations.)

The surface $\tilde{\Sigma}$ has trivial normal bundle and the same genus as Σ. So the previous proposition gives

$$\begin{aligned}\chi_-(\Sigma) &= \chi_-(\tilde{\Sigma}) \\ &\geq \langle \alpha_k, [\tilde{\Sigma}]\rangle \\ &= k + \langle \alpha, \sigma\rangle \\ &= \sigma\cdot\sigma + \langle \alpha, \sigma\rangle,\end{aligned}$$

which is the desired result. □

Remarks. In some applications of this proposition, a proof of the blow-up formula is not needed. One may know that α is a basic class by an application of Taubes' result, Theorem 15, when α arises as K_ω for some symplectic form, in which case one knows that α_k is a basic class also, because there is a symplectic structure on X_k with this class as its canonical class. Without using the blow-up formula, one can therefore deduce

$$\chi_-(\Sigma) \geq \sigma\cdot\sigma + \langle K_\omega, \sigma\rangle. \tag{23}$$

This inequality is an equality for symplectic submanifolds, or smooth algebraic curves in a complex surface, where it is usually referred to as the adjunction formula. The inequality in the proposition is often called the 'adjunction inequality'.

Unlike the 3-dimensional version of this statement, which led only to such basic facts as the lower bound for the genus of a knot in terms of the degree of its Alexander polynomial, the 4-dimensional version provides information which we can reach in no other way at present. Even with the 4-torus, where the Spin^c structure $\mathfrak{c}$ with $c_1 = 0$ has a monopole invariant $SW(T^4, \mathfrak{c}) = 1$, we learn that embedded surfaces in T^4 satisfy

$$\chi_-(\Sigma) \geq |\Sigma\cdot\Sigma|,$$

which is a significant result. For example, it leads to a proof of Milnor's conjecture [31] on the unknotting number of torus knots [22]. (The absolute value appears on the right-hand side because T^4 looks the same with either orientation.) By contrast, the 3-dimensional result, applied to the 3-torus, is entirely contentless.

The result for T^4 above is sharp: every class σ can be represented by a surface of complexity $|\sigma\cdot\sigma|$. More generally, if X carries a symplectic form ω, then Proposition 7.2 is sharp for at least a significant range of classes σ, by Donaldson's theorem [7] on the existence of symplectic submanifolds. That is, since non-degeneracy of a closed 2-form is an open condition, any rational cohomology class $\Omega' \in H^2(X;\mathbb{Q})$ sufficiently close to the class $\Omega = [\omega]$ in $H^2(X;\mathbb{R})$ is represented by a symplectic form ω', and the theorem of [7] says that for some large k (depending on Ω') an integer class $\mathrm{PD}(k\Omega')$ is represented by a symplectic submanifold Σ, for which the inequality (23) is inevitably an equality.

Where the theory hesitates

One should not, however, expect too much of Proposition 7.2. Symplectic 4-manifolds are, perhaps, close cousins of the 3-manifolds Y which fiber over the circle, such as manifolds obtained by zero-surgery on a fibered knot in S^3. For these knots, the degree of the Alexander polynomial and the genus of the knot are equal, and the lower bounds coming from the basic classes are sharp. For general 3-manifolds, the basic classes (as opposed to the monopole classes) give us information which is rather less than sharp, as explained in section 3, and the 4-dimensional situation is probably no better. In dimension three, we recovered much better information by looking at monopole classes detected using the stretching arguments from one dimension higher, or what is essentially the same, the non-vanishing of Floer homology. In dimension four, we have no analogous tool: if we defined a 'monopole class' to be a class $c_1(\mathfrak{c})$ for a Spin^c structure $\mathfrak{c}$ with the property that the equations had solutions for all metrics, as we did in dimension three, then we would have (at present) no tools for detecting monopole classes, other than the observation that basic classes are monopole classes. We have no Floer homology in dimension four.

The relationship between dimensions three and four is clarified somewhat by focusing on 4-manifolds of the form $X = S^1 \times Y$, where Y is a closed oriented 3-manifold. If $\mathfrak{c}$ is a Spin^c structure on X which is pulled back from Y, then it is quite easy to see that the monopole invariants are related:

$$SW(X,\mathfrak{c}) = SW(Y,\mathfrak{c}).$$

Note that, as long as $c_1(\mathfrak{c})$ is non-zero, the invariant $SW(X,\mathfrak{c})$ is well-defined according to our exposition above, because even if $b^+(X) = 1$ (which occurs when $b_1(Y) = 1$), the class $c_1(\mathfrak{c})$ has square zero.

Thus if Y is the manifold $Y(K)$ obtained by zero-surgery on a knot K, then the invariants $SW(X,\mathfrak{c})$ are determined by the Alexander polynomial of K, through a formula like (12). If K has Alexander polynomial 1, like the unknot, then the monopole invariants of X are all zero. A similar example with $b^+ > 1$ is the 4-manifold $S^1 \times Y_2$, where Y_2 is the connected sum of two copies of the same $Y(K)$. For such 4-manifolds, the inequality of Proposition 7.2 tells us nothing.

For the special manifolds $S^1 \times Y$, this set-back is temporary. There is actually a simple device which gives excellent information (sharp in many cases) concerning the minimum genus problem, despite the failure of the basic classes. It is a variant of the stretching argument.

Proposition 7.3. *Let Y be an oriented 3-manifold embedded in a closed, oriented 4-manifold Z with $b^+(Z) > 1$, and suppose that the image of $H_1(Y)$ in every component of $H_1(Z \setminus Y)$ is zero. Let α be a basic class on Z, and let α' be the class on $S^1 \times Y$ obtained by restricting α to Y and pulling back to the product. Then if Σ is an oriented embedded surface in the 4-manifold $S^1 \times Y$, we have*

$$\chi_-(\Sigma) \geq \sigma \cdot \sigma + \langle \alpha', \sigma \rangle,$$

where σ is the class represented.

Remark. Thus α' behaves like a basic class on $X = S^1 \times Y$, even though the monopole invariants of this 4-manifold may all be zero.

Proof. Let Σ be a surface in $S^1 \times Y$. We may take it that Σ is connected. Cut open $S^1 \times Y$ to form the manifold with boundary $[0,1] \times Y$, and make the cut transverse to Σ, so that we obtain a surface with boundary Σ_1. Let $\Sigma_n \subset [0,n] \times Y$ be the surface in $[0,n] \times Y$ obtained by concatenating n copies of Σ_1. Regard Σ_n as a surface with boundary in Z, by embedding $[0,n] \times Y$ as a collar neighborhood N of Y. The boundary of Σ_n, as a subset of ∂N in Z, is independent of n, and by the hypothesis on H_1 it is the boundary of a surface S in $Z \setminus N$.

Consider now the closed surface $S_n \subset Z$ formed as the union of $\Sigma_n \subset N$ and $S \subset Z \setminus N$. We apply the basic class inequality, Proposition 7.2, to the surface S_n, and we find that all three terms are linear in n:

$$\begin{aligned} \chi_-(S_n) &= n\chi_-(\Sigma) + C_1 \\ S_n \cdot S_n &= n(\sigma \cdot \sigma) + C_2 \\ \langle \alpha, S_n \rangle &= n\langle \alpha', \sigma \rangle + C_3, \end{aligned}$$

for constants C_i independent of n. Taking n sufficiently large, we deduce the required inequality. □

As in section 5, one can apply the same device to the case that Z is not closed but has a boundary carrying a contact structure ξ compatible with the orientation. If Z carries a symplectic structure compatible with ξ, then K_ω is a basic class, and we have a revision of the above proposition:

Proposition 7.4. *Let Z be a compact 4-manifold with boundary carrying a contact structure ξ on the boundary, compatible with the boundary orientation, and having a compatible symplectic form ω. Let Y be a 3-manifold embedded in Z and suppose that the image of $H_1(Y)$ in each component of $Z \setminus Y$ is zeero. Let $K' \in H^2(S^1 \times Y)$ be obtained from the canonical class K_ω as before. Then for any embedded surface Σ in $S^1 \times Y$, one has*

$$\chi_-(\Sigma) \geq \sigma \cdot \sigma + \langle K', \sigma \rangle,$$

where σ is the class represented. □

Using the theorem of Eliashberg and Thurston, we deduce:

Corollary 7.5. *Let Y be a closed, oriented 3-manifold carrying a taut foliation $\mathcal{F}$ by oriented leaves. Let e be the euler class of $T\mathcal{F}$ pulled back to $S^1 \times Y$. Then for any embedded surface Σ in $S^1 \times Y$ representing a class σ, one has*

$$\chi_-(\Sigma) \geq \sigma \cdot \sigma + \langle e, \sigma \rangle,$$

where σ is the class represented.

Proof. Using the result of Eliashberg and Thurston, Theorem 5.2, we construct a 4-manifold $Z = S^1 \times Y$ with contact structure ξ on the boundary and a symplectic structure ω, for which K_ω is $e(\mathcal{F})$. The only thing missing is the condition on $H_1(Y)$. This can be put right by the Legendrian surgery of [41]. That is, we choose Legendrian curves γ_i in the contact 3-manifold ∂Z so as to represent a basis for the homology of each component. We can then form a 4-manifold Z^* by adding 2-handles to these curves. According to [41], if the framing of the 2-handles is correctly chosen, the manifold Z^* has contact boundary and carries a compatible form ω^* extending ω. In the larger manifold Z^*, the condition on H_1 is satisfied. □

Corollary 7.6. *Let Σ be an embedded surface in $S^1 \times Y$ representing a class σ. Suppose that Y does not have a basis for H_2 represented by tori. Then the complexity of Σ satisfies the lower bound*

$$\chi_-(\Sigma) \geq |\sigma \cdot \sigma| + |\pi_*(\sigma)|,$$

in which the last term denotes the Thurston norm of the image of σ under the map $\pi : S^1 \times Y \to Y$.

Proof. Under the given hypotheses, the classes $e(\mathcal{F})$ as $\mathcal{F}$ runs through (smooth) taut foliations have the Thurston polytope $B(Y)$ as their convex hull, by Gabai's result, Theorem 1.5. (Recall also that, even without the hypothesis on tori, enough smooth foliations may well exist.) The absolute value sign is appropriate because the 4-manifold has orientation-reversing diffeomorphisms. □

The last corollary usually gives sharp information. To represent a class σ by a surface Σ whose complexity is as given on the right-hand side, one can proceed as follows. Write

$$\sigma = i_* \pi_*(\sigma) + \tau,$$

where i_* is the inclusion of $\{1\} \times Y$ and τ has the form $[S^1] \times [\gamma]$ for a closed curve γ in Y. Let S be a surface representing $\pi_*(\sigma)$ in Y, whose complexity is as small as possible (so given by the Thurston norm). Try to arrange that S meets γ transversely and without excess intersection in Y, so that the geometric and algebraic intersection numbers coincide. If this can be done, then we form a singular surface

$$\Sigma' = i(S) \cup (S^1 \times \gamma)$$

representing σ, and by smoothing the double points which occur at $i(S \cap \gamma)$ we arrive at a smooth surface Σ in the 4-manifold having the right complexity. The condition on excess intersection easily fulfilled if S is connected or if S is a union of parallel surfaces. Thus we can always apply this construction in the case that $b_1(Y)$ is 1.

This corollary seems to draw more from the 3-dimensional world than it does from gauge theory. Our proof has used the existence theorem for taut foliations, as well as the contact perturbations of foliations obtained from [8]. And yet the gauge theory remains an essential part of the story.

An interesting corollary of the lower bound above is to the question of whether the 4-manifold $S^1 \times Y$ can be symplectic. As we said in the previous subsection, if $S^1 \times Y$ is symplectic, then the lower bounds coming from basic classes of $S^1 \times Y$ are sharp, at least for some homology classes σ of positive square. If Y is obtained by zero-surgery on a knot K, then this lower bound is

$$\chi_-(\Sigma) \geq \sigma \cdot \sigma + (2r - 2)\beta \cdot \pi_*(\sigma),$$

where r is the degree of the symmetrized Alexander polynomial and β is the generator of $H_1(Y)$. On the other hand, the corollary above gives the potentially stronger lower bound

$$\chi_-(\Sigma) \geq \sigma \cdot \sigma + (2g - 2)\beta \cdot \pi_*(\sigma),$$

where g is the genus of the knot, at least if g is at least 2. (If g is 1, we need to know that the taut foliation can be made smooth.)

Corollary 7.7. *Let K be a knot of genus g and let r be the degree of its symmetrized Alexander polynomial. Let Y be obtained by zero-surgery on K. If g is 1, suppose in addition that Y carries a smooth taut foliation. Then a necessary condition for the existence of a symplectic structure on $S^1 \times Y$ is that g and r are equal.* □

Where the theory fails

An interesting construction of 4-manifolds was described recently by Fintushel and Stern in [10]. The building brick of this construction is a 4-manifold with boundary, of the form $S^1 \times Y$, where Y is a knot complement, $S^3 \setminus N(K)$. The boundary of $S^1 \times Y$ is a 3-torus. Let X_0 be a $K3$ surface, obtained for example by the Kummer construction which resolves the sixteen double points in $T^4/\pm 1$. Let $T \subset X_0$ be a standard 2-torus. In the Kummer model, T might be the image of a standard 2-torus in T^4. Let X_K be a closed 4-manifold obtained by removing a neighborhood $N(T)$ from X_0 and replacing it with $S^1 \times Y$, attaching the 3-torus boundaries. This should be done in such a way as to leave the homology of X_K the same as that of X_0. (Note that $S^1 \times Y$ and $N(T)$ look the same at the level of homology.)

In [10] it is shown that one can read off the Alexander polynomial of K from the monopole invariants of X_K. In general, X_K has the same homotopy type as X_0, and if K has trivial Alexander polynomial we have no invariants to distinguish the two manifolds. The question raised in [10] is whether X_K can be diffeomorphic to $X_{K'}$ if K and K' are different knots. In particular, if K is a knot with trivial Alexander polynomial, is X_K diffeomorphic to the $K3$ surface X_0?

It is tempting to guess that the *genus* of K is visible in X_K. For example, let $S \subset X_0$ be a sphere meeting T orthogonally in one point and having normal bundle of degree 2. (One can see such a sphere in the Kummer construction.) The cohomology class σ carried by S in X_0 corresponds to a class σ' in X_K. What is the smallest possible genus for a representative of σ'? The simplest surface that is easily visible is the surface Σ obtained from S by removing the disk in which S meets $N(T)$ and replacing the disk with a spanning surface of the knot in $S^1 \times Y$. This representative has

$$\chi_-(\Sigma) = 2g - 2,$$

where g is the genus of the knot. The basic classes give us the lower bound

$$\chi_-(\Sigma') \geq 2r - 2,$$

for any other surface Σ' representing this class, but it seems most likely that Σ is already best possible.

Note that if one carries out this construction using simply the 4-torus for X_0 rather than $K3$, then one has 4-manifolds X_K with an S^1 factor, and the results of the previous subsection can be applied. In particular, the genus of the knot *is* reflected in the minimum genus of embedded surfaces in the 4-manifold, even if it is not reflected in the monopole invariants. This lends some support to the conjecture that $2g - 2$ cannot be bettered in the $K3$ examples, but we seem to have no tools.

Unanswered questions

There are many more questions than answers concerning the minimum genus problem in dimension 4. We know very little about classes of negative square in algebraic surfaces, for example. There are manifolds such as $\mathbb{CP}^2 \# \mathbb{CP}^2$, where the gauge theory invariants have told us very little.

To illustrate the lack of present understanding in another direction, consider the following question.

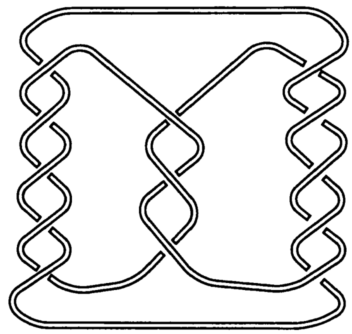

Figure 4: The pretzel knot $(5, -3, 5)$.

Question 7.8. Let $\tau : \tilde{X} \to X$ be a finite covering of a closed 4-manifold X. If Σ is a surface representing the smallest possible complexity for a class $\sigma \in H_2(X;\mathbb{Z})$, does its inverse image $\tilde{\Sigma} = \tau^{-1}(\Sigma)$ have least complexity in its homology class? Is there any large class of 4-manifolds (not simply connected) for which the answer is yes?

This question was posed by Thurston in dimension 3, and answered by Gabai in [13]: since one can lift a taut foliation to a cover, Theorem 1.5 and Corollary 1.3 give an affirmative answer in this case, at least if the manifold is irreducible. There is no comparable tool in dimension 4, which is why the question arises. In a similar spirit, one can ask about branched covers. For example:

Question 7.9. Let $\tau : \tilde{X} \to X$ be a branched double covering of a closed 4-manifold X, branched along a surface $B \subset X$. Let $\tilde{B}$ be its inverse image. If B has least complexity in its homology class, is the same true of $\tilde{B}$? Is there a class of 4-manifolds X for which the answer is yes?

One can also ask whether χ_{min} is linear on the ray generated by a class σ in the case that the normal bundle of a representative is trivial:

Question 7.10. Let Σ be a representative of least complexity for the class $\sigma \in H_2(X;\mathbb{Z})$. Suppose that the normal bundle of Σ is trivial, and let Σ_n be the surface consisting of n parallel copies of Σ in the tubular neighborhood. Is Σ_n a representative of least complexity for the class $n\sigma$? Is there a class of 4-manifolds X for which the answer is yes?

It is easy to invent other questions whose answers seem out of reach. The list of 4-manifolds with $b_2 > 0$ for which the function $\chi_{\min}(\sigma)$ is known is very short. There is a slightly longer list if one only asks about surfaces with non-negative normal bundle.

The following question is of a slightly different nature, but arises naturally from the discussion above.

Question 7.11. Let Y be obtained by zero-surgery on a knot K in S^3. If the 4-manifold $S^1 \times Y$ carries a symplectic structure, is the knot K necessarily fibered?

That K be a fibered knot is a sufficient condition for $S^1 \times Y$ to be symplectic. This is an observation of Thurston's. On the other hand, we have seen that a *necessary* condition is that the genus of K coincide with the degree of its symmetrized Alexander polynomial, at least if the genus is 2 or more. Another necessary condition is that the Alexander polynomial have leading coefficient ± 1 (see [10]). The pretzel knot shown in Figure 4 satisfies both of these necessary conditions (though with genus 1), but is not a fibered knot [5].

One might speculate that there is a role for codimension-two foliations of 4-manifolds, somewhat akin to the role of codimension-one foliations of 3-manifolds, though perhaps without the powerful existence theorem. Let $\mathcal{F}$ be an oriented foliation of a closed, oriented 4-manifold X by 2-dimensional leaves. Let us say that $\mathcal{F}$ is *taut* if there is closed 2-form ω which is positive on the leaves. Let $\kappa(\mathcal{F})$ be the 'canonical class' of the almost complex structure on X which $\mathcal{F}$ determines. That is, κ is minus the sum of the euler classes of the tangential and normal 2-plane fields. If Σ is a compact leaf of $\mathcal{F}$, its genus is determined by an adjunction formula,

$$2g - 2 = \sigma \cdot \sigma + \langle \kappa(\mathcal{F}), \sigma \rangle,$$

where σ is the class carried in homology.

Question 7.12. Is it true that a compact leaf of a taut foliation $\mathcal{F}$ of this sort is always genus-mimizing? More generally, is it possible that embedded surfaces in such a foliated 4-manifold satisfy an adjunction inequality, which bounds there genus from below by a formula such as

$$\chi_{-}(\Sigma) \geq \sigma \cdot \sigma + \langle \kappa(\mathcal{F}), \sigma \rangle,$$

at least in the case that $\langle \omega, \sigma \rangle$ is positive?

(Some extra hypothesis of the sort indicated is needed to deal with the case that $\kappa(\mathcal{F})$ is negative, as the example of $S^2 \times S^2$ shows.) Note that $S^1 \times Y$ has codimension-two foliations arising from the foliations of Y, so the above question includes Corollary 7.5 as a special case. Another special case arises when the foliation is a fibration. The quesion then has an affirmative answer, because the fibration carries a symplectic form.

In the version given above, the question may not be very useful, because of a lack of examples of foliations. The situation looks more interesting if one allows the foliation to have singularities of the sort that arise in holomorphic foliations (taking due account of orientations). What we have in mind is the sort of foliation $\mathcal{F}$ that one obtains in a complex surface when $T\mathcal{F}$ is defined by the vanishing of a holomorphic 1-form, or more generally a 1-form with values in some holomorphic line bundle. One could extend the previous question to singular foliations based on this model.

In this form, the question encompasses the question of mimimum genus in the Fintushel-Stern fake $K3$ surfaces (the spaces we called X_K earlier). The 4-manifolds X_K carry singular codimension-two foliations $\mathcal{F}$ formed by combining a holomorphic foliation of $K3$ with Gabai's foliation of the knot comlement $S^3 \setminus K$ (extended trivially to the product with the circle). The canonical class of this foliation $\mathcal{F}$ is given by $2g\,\mathrm{P.D.}[T]$, where $[T]$ is the class of the torus on which the modification was performed and g is the genus of the knot. Thus an affirmative answer to the last question would imply that the genus of K is visible in X_K, as we speculated earlier. Note that the manifolds X_K provide interesting examples of taut, singular foliations on a manifold admitting no symplectic structure. (It is pointed out in [10] that X_K is not symplectic if the Alexander polynomial of K is not monic.)

References

[1] S. Akbulut and J. D. McCarthy, *Casson's invariant for oriented homology 3-spheres*, Mathematical Notes, vol. 36, Princeton University Press, Princeton, NJ, 1990.

[2] D. Auckly, *The Thurston norm and three-dimensional Seiberg-Witten theory*, Osaka J. Math. **33** (1996), no. 3, 737–750.

[3] 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.

[4] W. Chen, *Dehn surgery formula for Seiberg-Witten invariants of homology 3-spheres*, preprint.

[5] R. H. Crowell and H. F. Trotter, *A class of pretzel knots*, Duke Math. J. **30** (1963), 373–377.

[6] S. K. Donaldson, *Polynomial invariants for smooth four-manifolds*, Topology **29** (1990), no. 3, 257–315.

[7] ———, *Symplectic submanifolds and almost-complex geometry*, J. Differential Geom. **44** (1996), no. 4, 666–705.

[8] Y. M. Eliashberg and W. P. Thurston, *Confoliations*, to appear in Mem. Amer. Math. Soc.

[9] P. M. N. Feehan and T. G. Leness, *Non-abelian monopoles, I: regularity, transversality and compactness*, preprint.

[10] R. Fintushel and R. Stern, *Knots, links and four-manifolds*, preprint.

[11] A. Floer, *An instanton-invariant for* 3*-manifolds*, Comm. Math. Phys. **118** (1988), no. 2, 215–240.

[12] K. A. Frøyshov, *The Seiberg-Witten equations and four-manifolds with boundary*, Math. Res. Lett. **3** (1996), no. 3, 373–390.

[13] D. Gabai, *Foliations and the topology of 3-manifolds*, J. Differential Geom **18** (1983), no. 3, 445–503.

[14] ———, *Foliations and genera of links*, Topology **23** (1984), no. 4, 381–394.

[15] ———, *Foliations and the topology of* 3*-manifolds. II*, J. Differential Geom. **26** (1987), no. 3, 461–478.

[16] ———, *Foliations and the topology of* 3*-manifolds. III*, J. Differential Geom. **26** (1987), no. 3, 479–536.

[17] ———, *Foliations and 3-manifolds*, Proceedings of the International Congress of Mathematicians (Kyoto 1990), vol. I, Math. Soc. Japan, 1991, pp. 601–609.

[18] C. McA. Gordon and J. Luecke, *Knots are determined by their complements*, J. Amer. Math. Soc. **2** (1989), no. 2, 371–415.

[19] D. Kotschick, J. W. Morgan, and C. H. Taubes, *Four-manifolds without symplectic structures but with nontrivial Seiberg-Witten invariants*, Math. Res. Lett. **2** (1995), no. 2, 119–124.

[20] P. B. Kronheimer, *An obstruction to removing intersection points in immersed surfaces*, Topology **36** (1997), no. 4, 931–962.

[21] P. B. Kronheimer and T. S. Mrowka, *Monopoles and contact structures*, Inventiones Mathematicae, to appear.

[22] ———, *Gauge theory for embedded surfaces, I*, Topology **32** (1993), no. 4, 773–826.

[23] ———, *The genus of embedded surfaces in the projective plane*, Math. Res. Lett. **1** (1994), no. 6, 797–808.

[24] ———, *Embedded surfaces and the structure of Donaldson's polynomial invariants*, J. Differential Geom. **41** (1995), no. 3, 573–734.

[25] ———, *Gauge theory for embedded surfaces, II*, Topology **34** (1995), no. 1, 37–97.

[26] T. J. Li and A. Liu, *General wall crossing formula*, Math. Res. Lett. **2** (1995), no. 6, 797–810.

[27] Y. Lim, *Seiberg-Witten invariants for 3-manifolds and product formulae*, preprint.

[28] P. Lisca and G. Matić, *Tight contact structures and the Seiberg-Witten invariants*, Inventiones Mathematicae, to appear.

[29] M. Marcoli and B.-L. Wang, *Equivariant Seiberg-Witten Floer homology*, preprint.

[30] G. Meng and C. H. Taubes, SW = *Milnor torsion*, Math. Res. Lett. **3** (1996), no. 5, 661–674.

[31] J. Milnor, *Singular points of complex hypersurfaces*, Princeton University Press, Princeton, N.J., 1968, Annals of Mathematics Studies, No. 61.

[32] G. Moore and E. Witten, *Integration over the u-plane in Donaldson theory*, preprint.

[33] V. Y. Pidstrigach and A. N. Tyurin, *Localisation of Donaldson invariants along Seiberg-Witten basic classes*, preprint.

[34] R. Roussarie, *Plongements dans les variétés feuilletées et classification de feuilletages sans holonomie*, Inst. Hautes Études Sci. Publ. Math. (1974), no. 43, 101–141.

[35] C. H. Taubes, *Casson's invariant and gauge theory*, J. Differential Geom. **31** (1990), no. 2, 547–599.

[36] ———, *The Seiberg-Witten invariants and symplectic forms*, Math. Res. Lett. **1** (1994), no. 6, 809–822.

[37] ———, *More constraints on symplectic forms from Seiberg-Witten invariants*, Math. Res. Lett. **2** (1995), no. 1, 9–13.

[38] W. P. Thurston, *Foliations of three-manifolds which are circle bundles*, Ph.D. thesis, Berkeley, 1972.

[39] ———, *A norm for the homology of 3-manifolds*, Mem. Amer. Math. Soc. **59** (1986), no. 339, i–vi and 99–130.

[40] R. G. Wang, *On Seiberg-Witten Floer invariants and the generalized Thom problem*, preprint.

[41] A. Weinstein, *Contact surgery and symplectic handlebodies*, Hokkaido Math. J. **20** (1991), no. 2, 241–251.

[42] E. Witten, *Monopoles and four-manifolds*, Math. Res. Lett. **1** (1994), no. 6, 769–796.