

Notes on the Seiberg-Witten Equations, the Weinstein Conjecture and Embedded Contact Homology

Clifford Henry Taubes*

These notes are supplementary to the author's lectures at the 2007 conference on Current Developments in Mathematics, held on November 16, 17 in Cambridge Massachusetts. They outline various connections between the three subjects that are listed in the title.

My purpose in writing these notes is to give an indication of how the Seiberg-Witten equations on a 3-dimensional manifold can see something of the geometry of a contact 1-form. The new ideas for 3-dimensions led the author to a proof of the 3-dimensional version of the Weinstein conjecture [**W**]:

THEOREM. *The Reeb vector field for any given contact 1-form on any given compact 3-manifold has at least one closed, integral curve.*

The proof appears in [**T1**]; and [**T2**] says more about such integral curves.

Very much related to the story told here is the now decade old story explains how the Seiberg-Witten equations on a 4-dimensional symplectic manifold see something of the symplectic structure [**T3**], [**T4**]. The three and four dimensional stories are connected via a conjectured isomorphism between the Seiberg-Witten Floer cohomology and Michael Hutching's embedded contact homology [**H**], [**HS**], [**HT**]. I say something at the end about embedded contact homology.

A great deal was known about the Weinstein conjecture in dimension 3 prior to [**T1**]; much of this the pioneering work of Helmut Hofer. Hofer proved the conjecture for a huge class of contact structures; in particular contact structures with over-twisted 2-plane fields [**Hof1**]. He also proved the conjecture for any contact structure on S^3 and for any 3-manifold with non-zero π_2 . Other recent work on this conjecture includes [**ACH**], [**Ch**], [**CH**], [**Ga**], and [**Hon**]. Hofer's ICM article [**Hof2**] and his reviews [**Hof3**], [**Hof4**]

*Supported in part by a grant from the National Science Foundation.

say more about Hofer's seminal contributions and give many additional references.

The Weinstein conjecture is still open for the generic contact manifold of dimension greater than 3 (see, e.g., [Gi].) Even so, one of the central themes in the 3-dimensional proof has a direct analog in dimensions greater than 3. This is the notion of *spectral flow* for a Dirac operator. The Seiberg-Witten equations supply a bridge of sorts that connects this essentially linear notion to the very non-linear Weinstein question. There is no known analog of this bridge in dimensions greater than 3.

What follows is an outline for the rest of this article.

- (1) Gauge theory background
 - (a) $\text{Spin}^{\mathbb{C}}$ structures
 - (b) Associated vector bundles
 - (c) Connections
 - (d) The Seiberg-Witten equations
 - (e) General properties
- (2) Contact 3-manifolds
 - (a) The Seiberg-Witten equations and contact 1-forms
 - (b) The proof of the Weinstein conjecture
 - (c) Properties of solutions
 - (d) Implications
- (3) Seiberg-Witten Floer homology
 - (a) A computation
 - (b) A fly in the soup
- (4) Spectral flow
 - (a) An explicit example
 - (b) Spectral flow for the Dirac operator
 - (c) Spectral flow and the Seiberg-Witten equations
 - (d) Spinors with length near 1 everywhere
- (5) Bounds on the energy from the spectral flow
 - (a) Asymptotic spectral flow
 - (b) The Chern-Simons functional and the energy
 - (c) An idealized scenario
 - (d) A realistic scenario
- (6) Some remarks about Hutchings' embedded contact homology
 - (a) Pseudoholomorphic subvarieties
 - (b) The definition of I
 - (c) The definition of δ
 - (d) Embedded contact homology and Seiberg-Witten Floer cohomology

This introduction ends with an acknowledgement of the profound debt owed by the author to Tom Mrowka, Peter Kronheimer and Michael Hutchings for sharing their ideas and thoughts on the subjects in this article.

1. Gauge theory background

This first section introduces the Seiberg-Witten equations on a three manifold. In what follows, M is an oriented three manifold endowed with a Riemannian metric. Those who wish to learn more about the Seiberg-Witten equations on 3-manifolds (and 4-manifolds) are directed to the bible of Kronheimer-Mrowka [KM].

1.1. $\text{Spin}^{\mathbb{C}}$ structures. A Riemannian manifold such as M has its oriented, orthonormal frame bundle; this a principle $SO(3)$ bundle that is denoted in what follows by $Fr \rightarrow M$. A $\text{Spin}_{\mathbb{C}}$ structure on M is an equivalence class of lifts of Fr to a principle $U(2) = \text{Spin}_{\mathbb{C}}(3)$ bundle. Such a lift defines and is defined by a principle $U(2)$ bundle over M whose projection factors in a $U(2)$ equivariant fashion through Fr . The set of $\text{Spin}_{\mathbb{C}}$ structures is in 1-1 correspondence with $H^2(M; \mathbb{Z})$; but there is in general no canonical pairing. Even so, the set of $\text{Spin}_{\mathbb{C}}$ structures is an affine space modeled on $H^2(M; \mathbb{Z})$.

This business about $\text{Spin}_{\mathbb{C}}$ structures can be seen in three dimensions fairly explicitly given that the tangent bundle of an oriented 3-manifold has a global frame. The choice of such a section identifies Fr with $M \times SO(3)$. This trivialization suggests an obvious $\text{Spin}_{\mathbb{C}}$ structure, this the trivial $U(2)$ bundle $F = M \times U(2)$. There are others. Let $S \rightarrow M$ denote a principle $U(1)$ bundle. The principle bundle $S \times_{U(1)} U(2)$ is a different $\text{Spin}_{\mathbb{C}}$ structure if S is non-trivial. Here, $U(1)$ acts on $U(2)$ as its center. This construction identifies the set of $\text{Spin}_{\mathbb{C}}$ structures on M with the set of principle $U(1)$ bundles up to isomorphism. The latter set is a group that is isomorphic to $H^2(M; \mathbb{Z})$.

1.2. Associated vector bundles. Let $F \rightarrow M$ denote the $U(2)$ bundle for a given $\text{Spin}_{\mathbb{C}}$ structure. Two associated, bundles play a central role. The first is the Hermitian \mathbb{C}^2 bundle $F \times_{U(2)} \mathbb{C}^2$ denoted in what follows by \mathbb{S} . The second is the principle $U(1)$ bundle $S_F = F \times_{U(2)} U(1)$; here $U(2)$ acts on $U(1)$ via the determinant representation.

The bundle \mathbb{S} turns out to be a Clifford module for T^*M . This means that there is a canonical homomorphism $cl : T^*M \rightarrow \text{End}(\mathbb{S})$ that obeys the following:

$$(1.1) \quad \begin{aligned} cl(a) cl(b) &= -\langle a, b \rangle \mathbb{I} - cl(*(a \wedge b)) \\ cl(a)^\dagger &= -cl(a) \end{aligned}$$

Here, \langle, \rangle denotes the metric pairing, and $* : \wedge^2 T^*M \rightarrow T^*M$ is the metric's Hodge dual operator. The latter is defined with the help of the volume 3-form $d \text{vol}_M$ by the rule:

$$(1.2) \quad a \wedge *b = \langle a, b \rangle d \text{vol}_M$$

Note that the multiplication rules in (1.1) are obeyed on \mathbb{R}^3 (where $\mathbb{S} = \mathbb{R}^3 \times \mathbb{C}^2$) if $cl(\cdot)$ assigns the basis covectors $\{e^1, e^2, e^3\}$ to the correspondingly

labeled Pauli matrices:

$$(1.3) \quad \tau^1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad \tau^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \tau^3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

There are two other maps that are associated with cl . The first is denoted by \hat{c} and it defines a homomorphism from $\mathbb{S} \times T^*M$ to \mathbb{S} . It is defined by what it does to a decomposable element, $\eta \otimes a$; it sends the latter to $\text{cl}(a)\eta$. The second constitutes a map from $\mathbb{S} \otimes \mathbb{S}$ to $T^*M \otimes_R \mathbb{C}$. It sends a pair (η, ψ) to what is denoted as $\eta^\dagger \tau \psi$ and it is defined so that $\langle a, \eta^\dagger \tau \psi \rangle = \eta^\dagger \text{cl}(a)\psi$. This map is anti-linear on the first factor and linear on the second. It has the following key properties:

$$(1.4) \quad \begin{aligned} \psi^\dagger \tau \psi &\in iT^*M \\ |\psi^\dagger \tau \psi|^2 &= |\psi|^4 \end{aligned}$$

1.3. Connections. This section constitutes a short primer on the subject of connections on principle bundles. Three bundles in particular are relevant for what follows; these being Fr , F and the principle $U(1)$ bundle $S_F = F \times_{U(2)} U(1)$. In any event, let $\pi : P \rightarrow M$ denote a principle bundle with structure group G , this a compact Lie group. The map π induces an exact sequence

$$(1.5) \quad 0 \rightarrow \ker(\pi_*) \rightarrow TP \rightarrow \pi^*TM \rightarrow 0$$

of vector bundles over P . These bundle maps are equivariant with respect to the action of G on P . Note in this regard that the G action on P can be used to identify the bundle $\ker(\pi_*)$ with the bundle

$$(1.6) \quad \ker(\pi_*) = P \times l_G$$

where l_G denotes the Lie algebra of G . This identification is equivariant with respect to the G action on P if it is understood that G acts on l_G by its adjoint action.

A connection on P is neither more nor less than a G -equivariant splitting of (1.5). Viewed from the perspective of (1.6), a connection on P is a G -equivariant, linear from T^*P to l_G that restricts to $\ker(\pi_*)$ at any given point as the identity map from g to itself. Let \mathbb{A} now denote a connection on P . The kernel of \mathbb{A} is a G -equivariant distribution in TP that is isomorphic to π^*TM . Note that any two connections on P differ by the pull-back via π of a section over M of the vector bundle $(P \times_G l_G) \otimes T^*M$.

The curvature of \mathbb{A} is the G -equivariant, l_G -valued section of $\pi^*(\wedge^2 T^*M)$ given by the restriction of $d\mathbb{A}$ to $\text{kernel}(\mathbb{A}) \subset TP$. Being G -equivariant, it is most profitable to view the curvature as a section, $F_{\mathbb{A}}$, over M of the vector bundle $(\wedge^2 T^*M) \otimes (P \times_G l_G)$. A connection is said to be flat when $F_{\mathbb{A}} = 0$. In this case, the kernel of \mathbb{A} defines an involutive distribution in TP .

A connection on P allows derivatives to be taken of sections over M of any given associated bundle. Such a derivative is called the covariant derivative. To elaborate, suppose that $\rho : G \rightarrow \text{End}(V)$ is a representation

of G on some finite dimensional vector space V . Use ρ to define the bundle $P \times_G V = (P \times V)/G$ where G acts as ρ on V and on P as usual. Let s denote a section of $P \times_G V$. Its covariant derivative as defined by \mathbb{A} is the section $\nabla_{\mathbb{A}} s = \mathfrak{s}^*(\mathbb{A})$ of $(P \times_G V) \otimes T^*M$. The novice reader is invited to unwind the definitions (or work in a local chart where P is trivialized) to prove that $\nabla_{\mathbb{A}}$ is linear with respect to the vector bundle structure, and obeys the Leibnitz rule $\nabla_{\mathbb{A}}(f\mathfrak{s}) = f\nabla_{\mathbb{A}}\mathfrak{s} + \mathfrak{s} \otimes df$ when f is a function on M .

There is a unique connection on the frame bundle Fr determined by the chosen Riemannian metric. This is the Levi-Civita connection; it is characterized as follows: Let \mathfrak{s} denote a section of T^*M . Then $\nabla_{\mathbb{A}}\mathfrak{s}$ is a section of $T^*M \otimes T^*M$ and so anti-symmetrization defines from $\nabla_{\mathbb{A}}\mathfrak{s}$ a section of $\wedge^2 T^*M$. The exterior derivative of \mathfrak{s} is also a section of $\wedge^2 T^*M$. These two sections are equal for all such \mathfrak{s} if and only if \mathbb{A} is the Levi-Civita connection. A connection on Fr does not by itself determine a connection on a $\text{Spin}_{\mathbb{C}}$ lift, F , of Fr . The extra data needed is a connection on the $U(1)$ bundle S_F .

1.4. The Seiberg-Witten equations. The Seiberg-Witten equations constitute a system of non-linear, first order differential equations for a pair (A, ψ) where A is a connection on S_F and where ψ is a section of \mathbb{S} . There are many versions of these equations, any two differing by a term lacking derivatives of (A, ψ) . Here is the prototype:

$$(1.7) \quad *F_A - 2\psi^\dagger \tau \psi = 0 \quad \text{and} \quad \hat{c}(\nabla_{\mathbb{A}}\Psi) = 0,$$

where \mathbb{A} is the connection on F that is determined by the Levi-Civita connection and the connection A . In what follows, the operator $\hat{c}(\nabla_{\mathbb{A}}\cdot)$ is denoted by D_A . It is called the Dirac operator by virtue of the fact that when written out on \mathbb{R}^3 , it is Dirac's famous operator.

Note that if (A, ψ) obeys (1.7), then so does $(A - 2u^{-1}du, u\psi)$ when u is a smooth map from M to $U(1)$. Said differently, the space $C^\infty(M; U(1))$ acts on the space of solutions to (1.7). Two solutions on the same $C^\infty(M; U(1))$ orbit are deemed to be *gauge equivalent*, and any given orbit is deemed to be a *gauge equivalence class*. No notational distinction will be made here between a solution to (1.7) and its gauge equivalence class.

Let $\text{Conn}(S_F)$ denote the space of Hermitian connections on S_F . The equations that are depicted in (1.7) are the formal variational equations of a functional on the space $\text{Conn}(S_F) \times C^\infty(M; \mathbb{S})$. The latter functional has the schematic form

$$(1.8) \quad \mathfrak{a} = \frac{1}{2} \mathfrak{cs}(A) + 2 \int_M \psi^\dagger D_A \psi,$$

where \mathfrak{cs} is defined as follows: Fix a fiducial connection A_* on S_F . Then write $A = A_* + \hat{a}$ where \hat{a} is a section of iT^*M . Now set

$$(1.9) \quad \mathfrak{cs}(A) = - \int_M \hat{a} \wedge d\hat{a} - 2 \int_M \hat{a} \wedge F_{A_*}.$$

Note that any two choices for A_* assign the same number to any given connection.

As mentioned above, (1.7) is the prototype for numerous versions of the Seiberg-Witten equations. These other versions are defined with the choice of a suitably benign function on $\text{Conn}(S_F) \times C^\infty(M; \mathbb{S}) \rightarrow \mathbb{R}$ that is invariant under the aforementioned action of $C^\infty(M; U(1))$. Here is an example of such a function: Fix a smooth 1-form μ on M . Use $\epsilon_\mu : \text{Conn}(S_F) \rightarrow \mathbb{R}$ to denote the function

$$(1.10) \quad \epsilon_\mu = -2i \int_M \mu \wedge F_A.$$

The Kronheimer-Mrowka bible [KM] describes a Banach space of “suitably benign” functions. In any event, let ϵ denote such a $C^\infty(M; U(1))$ -invariant function. The formal variational equations for the function $\mathfrak{a} + \epsilon$ give an allowed, perturbed version of the Seiberg-Witten equations. For example, the pair

$$(1.11) \quad *F_A - (2\psi^\dagger \tau \psi - i * du) = 0 \quad \text{and} \quad D_A \psi = 0$$

are the variational equations of $\mathfrak{a} + \epsilon_\mu$.

1.5. General properties. The space of solutions to (1.11) is either empty or infinite dimensional. Indeed, if (A, ψ) solves (1.11), then so does $(A - 2u^{-1} du, u\psi)$ for any map $u \in C^\infty(M; U(1))$. However, the following is true:

LEMMA 1.1. *Fix a Spin^C structure and a 1-form μ . Then the space of gauge equivalence classes of solutions to the corresponding version of (1.11) is compact. Moreover, if μ is suitably generic, then there is but a finite set of such gauge equivalence classes of solutions to (1.11).*

The key to the proof of the lemma’s compactness assertion is the Bochner-Weitzenboch formula for $D_A^2 \psi$. What follows is the formula for a Dirac operator defined by any given metric on M and any given connection A on S_F , and with ψ any given section of \mathbb{S} :

$$(1.12) \quad D_A^2 \psi = \nabla_{\mathbb{A}}^\dagger \nabla_{\mathbb{A}} \psi - \frac{1}{4} cl(*F_A) \psi + \frac{1}{4} R \psi$$

Here, R denotes the scalar curvature of the metric.

To see how this formula is used, suppose that (1.12) holds. Then (1.12) implies that $|\psi|^2$ obeys

$$(1.13) \quad \frac{1}{2} d^\dagger d |\psi|^2 + |\nabla_{\mathbb{A}} \psi|^2 + \frac{1}{2} |\psi|^4 + \frac{1}{4} R |\psi|^2 = 0.$$

An application of the maximum principle to (1.13) implies that $|\psi|$ is zero when M has non-negative scalar curvature. In general, (1.13) implies that $|\psi|^2 \leq \frac{1}{2} \sup_M R$. The analogous Weitzenboch formula when (1.11) holds gives a pointwise bound for ψ that depends only on the scalar curvature and the supremum norm of $d\mu$.

A bound on the sup norm of $|\psi|$ gives apriori bounds for the L^2 norms of $|F_A|$ and $|\nabla_A \psi|$. What with Uhlenbecks theorem in [U], the latter can be parlayed using standard elliptic bootstrapping arguments to give local, uniform bounds on any given C^k norm for A and ψ . Such bounds imply the compactness assertion. The details of this can be found in the Kronheimer-Mrowka bible.

The Kronheimer-Mrowka bible also explains how the Sard-Smale theorem [Sm] can be used to prove that there are only a finite set of equivalence classes of solutions to (1.11) if μ is chosen in a suitably generic fashion.

2. Contact 3-manifolds

A 1-form, a , on a 3-manifold is deemed a contact form when $a \wedge da$ is nowhere zero. It is customary to orient M so that $a \wedge da$ is a positive multiple of the volume 3-form. The 2-form da has everywhere 1-dimensional kernel. In particular, there is a unique vector field, v , with $da(v, \cdot) = 0$ and $a(v) = 1$. This is called the *Reeb* vector field. The purpose of this section is to indicate how the Seiberg-Witten equations detect closed integral curves of Reeb vector fields.

2.1. The Seiberg-Witten equations and contact 1-forms. Fix now a contact 1-form a on M . Here it is assumed that $a \wedge da > 0$. One can now choose a Riemannian metric on M such that $*da = 2a$ and such that $|a| = 1$. Fix such a metric. Let $F \rightarrow M$ denote the principle $\text{Spin}_{\mathbb{C}}$ bundle for a chosen $\text{Spin}_{\mathbb{C}}$ structure on M . The 1-form a also supplies a 1-parameter family of allowed perturbations to the Seiberg-Witten equations. This family is indexed by $r \in [1, \infty)$. The basic version has the form

$$(2.1) \quad *F_A - 2r(\psi^\dagger \tau \psi - ia) = 0 \quad \text{and} \quad D_A \psi = 0.$$

The generic version adds $-i * d\mu$ to the right hand side of the equation for $*F_A$ where μ is a fixed, r -independent 1-form. The very large r versions of (2.1) and its $-i * d\mu$ perturbations are sensitive to the contact structure a and the integral curves of v . This is made precise in the upcoming Proposition 2.1. This proposition provides the link between the Seiberg-Witten equations and integral curves of v .

Proposition 2.1 refers to a splitting of \mathbb{S} into eigen-subbundles for the endomorphism $\text{cl}(a)$. This splitting is written as

$$(2.2) \quad \mathbb{S} = E \oplus EK^{-1},$$

where the convention has $\text{cl}(a)$ acting as i on the left most subbundle and as $-i$ on the right most. The complex line bundle K^{-1} is isomorphic as an $SO(2)$ bundle to the 2-plane bundle $\text{kernel}(a) \subset TM$ with the latter oriented by the 2-form da . Note that the assignment of E 's first Chern class, $c_1(E)$, in $H^2(M; \mathbb{Z})$ to the given $\text{Spin}_{\mathbb{C}}$ structure defines another 1-1 correspondence between the set of $\text{Spin}_{\mathbb{C}}$ structures on M with $H^2(M; \mathbb{Z})$.

The proposition also refers to the $\mu = ra$ version of the functional ϵ_μ . This is given its own name,

$$(2.3) \quad E = ir \int_M a \wedge F_A$$

Here now is the key result:

PROPOSITION 2.1. *Fix $\text{Spin}_\mathbb{C}$ structure and write \mathbb{S} as in (2.2). Suppose that $\{r_n\}_{n=1,2,\dots} \subset [1, \infty)$ is an increasing, unbounded sequence, and that $\{c_n = (A_n, \psi_n)\}_{n=1,2,\dots} \subset \text{Conn}(S_F) \times C^\infty(M; \mathbb{S})$ is a corresponding sequence such that (A_n, ψ_n) solves the $r = r_n$ version of (2.1) (or an $-i^*d\mu$ perturbation) for each n . Suppose, in addition, that $\{\sup_M |1 - |\psi_n||\}_{n=1,2,\dots}$ is bounded away from zero and $\{E(c_n)\}_{n=1,2,\dots}$ is bounded from above.*

- (A) *There exists a non-empty, finite set Θ whose elements have the form (γ, m) where γ is a closed integral curve of v and where m is a positive integer. In this regard, distinct pairs from Θ have distinct integral curve components. In addition, the formal sum $\sum_{(\gamma, m) \in \Theta} m\gamma$ defines an integral homology chain that represents the Poincaré dual to the first Chern class of the line bundle E that appears in (2.2).*
- (B) *The set Θ arises from a subsequence of $\{c_n\}_{n=1,2,\dots}$ as follows: Renummer the subsequence consecutively from one. Write the section ψ_n with respect to the splitting in (2.2) as (α_n, β_n) . Then $\lim_{n \rightarrow \infty} \alpha_n^{-1}(0)$ converges in the Hausdorff sense to $\cup_{(\gamma, m) \in \Theta} \gamma$.*

2.2. The proof of the Weinstein conjecture. This last proposition raises the following questions:

- *When does a sequence $\{r_n, c_n = (A_n, \psi_n)\}_{n=1,2,\dots} \in [1, \infty) \times (\text{Conn}(S_F) \times C^\infty(M; \mathbb{S}))$ exist such that $\{r_n\}$ is unbounded and (A_n, ψ_n) obeys the $r = r_n$ version of (2.1)?*
 - *When does a sequence of this sort exist with*
- $$(2.4) \quad \left\{ \sup_M |1 - |\psi_n|| \right\}_{n=1,2,\dots} \text{ bounded away from zero?}$$
- *When does a sequence of this sort exist with both*
- $$\left\{ \sup_M |1 - |\psi_n|| \right\}_{n=1,2,\dots} \text{ bounded away from zero and}$$
- $$\{E(c_n)\}_{n=1,2,\dots} \text{ bounded?}$$

As it turns out, there is, in all cases, a $\text{Spin}_\mathbb{C}$ structure where the answer is *always* to all three questions. This implies Weinstein's conjecture.

Section 3 addresses the first of these questions. The second question is addressed at the very end of Section 4. Section 5 takes up the third question. The remainder of this section describes various relevant properties of solutions to (2.1) that are used to prove Proposition 2.1 and to answer the questions in (2.4).

2.3. Properties of solutions to (2.1). When (A, ψ) obey (2.1), the Bochner-Weitzenboch identity

$$(2.5) \quad \nabla_{\mathbb{A}}^{\dagger} \nabla_{\mathbb{A}} \psi - \frac{1}{4} \text{cl}(*F_A)\psi + \frac{1}{4} R\psi = 0$$

leads to the modification of (1.13) that replaces $|\psi|^4$ with $r|\psi|^2(|\psi|^2 - i\psi^{\dagger} \text{cl}(a)\psi)$. An application of the maximum principle to the resulting equation implies

LEMMA 2.2. *There exists $\kappa > 1$ with the follow significance: If $r \geq 1$ and (A, ψ) obeys (2.1), then $|\psi| \leq 1 + \kappa/r$.*

A somewhat more sophisticated application of the maximum principle to (2.5) when (A, ψ) obeys (2.1) leads to:

LEMMA 2.3. *There exists $\kappa > 1$ such that if $r \geq 1$ and $(A, \psi = (\alpha, \beta))$ obeys (2.1), then*

- $|\alpha| \leq 1 + \kappa/r$
- $|\beta| \leq \kappa r^{\frac{1}{2}}$
- $|\nabla_A \alpha| \leq \kappa r^{\frac{1}{2}}$
- $|\nabla_A \beta| \leq \kappa$

These separate bounds for α and β are obtained by considering separately the E and EK^{-1} summands in (2.5). The bounds for the covariant derivatives of α and β are obtained by differentiating these summands of (2.5).

The Dirac equation $D_A \psi = 0$ has an E and EK^{-1} summand given the decomposition $\mathbb{S} = E \oplus EK^{-1}$. The E summand equates the covariant derivative of α along the Reeb vector field with transversal covariant derivatives of β and linear functions of β . What with Lemma 2.3, it implies that

$$(2.6) \quad |\langle v, \nabla_A \alpha \rangle| \leq \kappa.$$

2.4. Implications. To see what to make of this, fix point in M and local coordinates $(t, p = (x, y))$ centered at the chosen point such that t, x and y take values in $(-\delta, \delta)$ and such that $v = \frac{\partial}{\partial t}$ and such that $\{\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}\}$ are orthonormal at the origin. Here, $\delta \in (0, \frac{1}{2})$ is some fixed, small number. Fix $\delta' < \delta$ and for each $t \in (-\delta, \delta)$, define M_t to be the integral over the constant t disk of radius δ' of $r(1 - |\alpha|^2)^2$. It follows from (2.6) and Lemma 2.3 that this integral is bounded by an r and (A, ψ) independent multiple of ϵ .

Here is another consequence of Lemma 2.3: Fix $\epsilon > 0$ and there an r and (A, α) independent constant κ_{ϵ} such that $\kappa_{\epsilon} \epsilon$ bounds the maximal number of disjoint disks of radius $r^{-\frac{1}{2}}$ in any given constant t slice of the coordinate chart with $|\alpha| \leq 1 - \epsilon$ at the disk's center point.

And, here is one more consequence of (2.6) and Lemma 2.3: Let $(0, p)$ denote a point in the coordinate chart where $|\alpha| \leq 1 - \epsilon$. Then $|\alpha| \leq \frac{1}{2}\epsilon$ at

any point (t, p) if $t \leq \kappa^{-1}\varepsilon$ where κ is again some r and (A, ψ) independent constant.

Further ramifications can be seen by rescaling the $t = 0$ square by writing the coordinates (x, y) as $(r^{-\frac{1}{2}}z_1, r^{-\frac{1}{2}}z_2)$. Given a bound on E , and given that r is sufficiently large, the pull-back via this map of the pair (A, α) to the (z_1, z_2) plane is very close to solving the *vortex equations*. These are the equations for a pair consisting of a connection, A , on the trivial bundle over \mathbb{C} and a section, α_0 of this bundle that demand

$$(2.7) \quad \begin{aligned} F_A &= \frac{1}{2}(1 - |\alpha_0|^2)dz \wedge d\bar{z} \\ \bar{\partial}_A \alpha_0 &= 0 \\ \int_C (1 - |\alpha_0|^2) &< \infty \end{aligned}$$

Here $z = z_1 + iz_2$, and the d-bar operator is defined using the connection A . Two important properties to (2.7) enter the story: First, α_0 can be written with respect to a trivialization of the bundle as $e^{-u}h(z)$ where h is a polynomial in z . In particular, α_0 vanishes at its zeros with positive degree. Second,

$$(2.8) \quad 0 < (1 - |\alpha_0|^2) < \kappa \sum_{z' \in \alpha_0^{-1}(0)} \exp^{-\frac{|z-z'|}{2}}$$

with κ a fixed constant. The reader can learn about vortices from Section 4a of [T5].

The preceding points suggest

LEMMA 2.4. *Fix $\varepsilon > 0$ and there is a constant $\kappa = \kappa(\varepsilon)$ with the following significance: Suppose that $r \geq \kappa$ and that (A, ψ) is a solution to (2.1) with $E < \varepsilon^{-1}$. Suppose in addition that $|\alpha| < 1 - \varepsilon$ at the point $(t = 0, x = 0, y = 0)$. Then the set of points where $\alpha = 0$ in any constant t square with side length δ has at most κ elements. In addition, there is at least one such point in each constant t square and it has distance less than $\kappa r^{-1/2}$ from the origin.*

Proposition 2.1 is little more than a corollary to Lemma 2.4.

3. Seiberg-Witten Floer homology

The Seiberg-Witten Floer homology is used to answer the first question in (2.4). As noted at the outset, Kronheimer and Mrowka wrote the definitive bible on this subject [KM]. To summarize a detailed story, the Seiberg-Witten Floer homology is canonically associated to a pair consisting of a compact, oriented 3-manifold and a $\text{Spin}_{\mathbb{C}}$ structure on the manifold. This homology is an equivalence class of vector spaces with equivalence given by vector space isomorphism. What follows describes a salient property.

THEOREM 3.1. *If the Seiberg-Witten Floer homology is non-trivial, then there are solutions to the equations for any given choice for μ in (1.11) and thus for any given choice for r in (2.1).*

There is, in general, no canonical choice of identification of the Seiberg-Witten Floer homology with a fixed vector spaces. However, given a suitably generic 1-form μ , a representative vector space and a basis can be constructed using the solutions to μ 's version of (1.11). Any given basis element for the representative vector space is a formal, appropriately weighted sum of gauge equivalence classes of solutions to the given version of (1.11).

As it turns out, the Seiberg-Witten Floer homology can be given a canonical \mathbb{Z} grading if the $\text{Spin}_{\mathbb{C}}$ structure is such that \mathbb{S} has torsion first Chern class in $H^2(M; \mathbb{Z})$. This is to say that the representative vector spaces can be given such a grading; and the isomorphisms that define the equivalence classes preserve the grading. In the case when \mathbb{S} has non-torsion first Chern class, the Seiberg-Witten-Floer homology has a relative $\mathbb{Z}/p\mathbb{Z}$ grading where p here is the divisibility of the first Chern class in $H^2(M; \mathbb{Z})$.

Here is the bottom line: Seiberg-Witten Floer homology constitutes a clever algebraic count of the gauge equivalence classes of solutions to any given version of (1.11); and this count gives *the same answer* when computed using respective versions of (1.11) as defined by distinct choices for the metric on M and the 1-form μ in (1.11), or for the parameter r in (2.1).

3.1. A computation. Theorem 3.1 would be of little use without the following Kronheimer-Mrowka coup:

THEOREM 3.2. *Suppose that \mathbb{S} has torsion first Chern class. Then the Seiberg-Witten Floer homology is non-zero in an unbounded set of negative degrees.*

Theorems 3.1 and 3.2 answer the top question in (2.4). To see why, note first that they jointly assert the following:

(3.1) *If the first Chern class of \mathbb{S} is a torsion class, then there exists at least one $C^\infty(M; U(1))$ equivalence class of solution to (2.1) for each $r \in [1, \infty)$.*

When \mathbb{S} splits as in (2.2), then its first Chern class is twice the first Chern class of E minus that of K ; thus $2c_1(E) - c_1(K)$. As it turns out, the first Chern class of K is always divisible by 2. (This is because orientable 3-manifolds have trivial tangent bundles.) As a consequence, there exists a line bundle $E_K \rightarrow M$ with $2c_1(E_K) - c_1(K) = 0$. Take the $\text{Spin}_{\mathbb{C}}$ structure whose version of \mathbb{S} splits as depicted in (2.2) with $E = E_K$. The latter $\text{Spin}_{\mathbb{C}}$ structure has a version of \mathbb{S} with zero first Chern class, which is a torsion class.

Given Lemma 1.1, one might expect that the Seiberg-Witten Floer homology is by necessity, finitely generated. It is, in fact, finitely generated in each degree, so it is finitely generated when the first Chern class of

\mathbb{S} is not torsion. As indicated by Theorem 3.2, it is definitely not finitely generated when \mathbb{S} has torsion first Chern class. More is said about this in what follows.

3.2. A fly in the soup. The reader may have noticed that there is no need for either Theorem 3.2 or Floer homology to write down a solution to (2.1) in the case when $E = E_K$. In this case the bundle S_F has a flat connection, θ , and so (2.1) has the solution:

$$(3.2) \quad A = \theta - ira \quad \text{and} \quad \psi = 0.$$

In fact, S_F has a flat connection if and only if \mathbb{S} has torsion first Chern class, so any such version of (2.1) has at least one solution that is given by (3.2). Solutions with ψ identically zero are said to be *reducible* and solutions with $\psi \neq 0$ somewhere are said to be irreducible. This distinction is relevant with regards to the action of $C^\infty(M; U(1))$. This group acts with trivial stabilizer on an irreducible solution, but it acts on any reducible solution with stabilizer $U(1)$, the latter in its guise as the subset of constant maps.

On the face of it, a reducible solution is of little interest with regards to Proposition 2.1 by virtue of the fact that

$$(3.3) \quad E = 2r \text{Vol}(M)$$

in the case when A is given by (3.2). Here, $\text{Vol}(M)$ is the volume of M .

Ah, but Theorem 3.1 asserts that the Seiberg-Witten Floer homology is non-zero in an unbounded set of negative degrees. Surely some of these classes are labeled by gauge equivalence classes of irreducible solutions?! Not necessarily: As is explained momentarily, all but at most a finite set of Seiberg-Witten Floer homology classes are labeled by equivalence classes of reducible solutions.

There are two facts to note with regards to this labeling of the reducible equivalence classes. Here is the first: If \mathbb{S} has torsion first Chern class, then any given version of (1.11) has reducible solutions; these given by $(A = \theta - i\mu, \psi = 0)$ where θ is a flat connection on S_F . Thus, reducible solutions are unavoidable. They are always present. Here is the second fact: As r varies in (2.1) or, more generally, as μ or the Riemannian metric varies in (1.11), any given reducible solution can give birth to, or bury one or more distinct, irreducible equivalence classes of solutions. Such a birth or death can happen at any value of r (or choice for μ and the metric) where the operator D_A has a non-trivial kernel. In particular, such births and deaths need not occur in pairs. As a consequence, any non-trivial and *invariant* count of equivalence classes of solutions to the Seiberg-Witten equations (such as Seiberg-Witten Floer homology) must count the reducible equivalence classes in a sophisticated way to account for the fact that they can create or destroy an irreducible equivalence class.

This lone birth/death phenomena is unique to the reducible solutions. By contrast, a *pair* of distinct equivalence classes of irreducible solutions can ‘collide’ and cancel, or appear denovo as r or μ is varied. The latter

phenomena is a precise analog of what happens to the critical points of a varying family of functions on a compact manifold. But a lone equivalence class of irreducible solution can neither appear nor disappear on its own as r or μ varies except via bifurcation from a reducible equivalence class. This special feature of the reducible equivalence classes is a manifestation of the fact that they have non-trivial stabilizer with respect to the action of $C^\infty(M; U(1))$.

Morse theory (see, e.g., [Mil], [Bo1]) supplies an accurate finite dimension model for the treatment in Seiberg-Witten Floer homology of the irreducible equivalence classes. The appropriate analog when reducible solutions are present is what is called equivariant Morse theory [Bo2]. Andreas Floer [F], [A] had the genius to see Morse analogs for functions much like those used here.

To summarize: An infinite set of Seiberg-Witten Floer homology generators must be assigned to the reducible equivalence classes of solutions to (1.11) so as to obtain a count of equivalence classes that is independent of the metric and the 1-form μ .

4. Spectral flow

The notion of spectral flow can be used to answer the second question in (2.4) and it is the key to answering the third question in (2.4). Indeed, spectral flow is central to the whole story. This section constitutes a digression to say some things about this notion of spectral flow.

4.1. An explicit example. To introduce the subject of spectral flow, digress for a moment to consider a relatively simple family of Dirac operators, this a family acting on $C^\infty(S^1; \mathbb{C})$, parameterized by $\sigma \in \mathbb{R}$, and given by:

$$(4.1) \quad \mathfrak{D}_\sigma = i \frac{d}{dt} + \sigma.$$

Here, S^1 is viewed as $\mathbb{R}/2\pi\mathbb{Z}$ and t is the $\mathbb{R}/2\pi\mathbb{Z}$ valued coordinate. The family $\sigma \rightarrow \mathfrak{D}_\sigma$ should be viewed in what follows as a continuously varying family of unbounded, self-adjoint operators on $L^2(S^1; \mathbb{C})$. The spectrum of \mathfrak{D}_σ consists of the numbers

$$(4.2) \quad \{n + \sigma\}_{n \in \mathbb{Z}}.$$

Each eigenvalue is simple when viewed over \mathbb{C} , but multiplicity 2 when viewed over \mathbb{R} . Note that 0 is not an eigenvalue of \mathfrak{D}_σ when σ is not an integer.

Now suppose that $r \geq 0$, and that neither σ nor $\sigma + r$ are integers. The number of values of $\sigma' \in [\sigma, \sigma + r]$ where $\mathfrak{D}_{\sigma'}$ has eigenvalue 0 is equal to the greatest integer less than r . This integer is deemed to be the spectral flow from \mathfrak{D}_σ to $\mathfrak{D}_{\sigma+r}$. It is denoted in here by $f(\mathfrak{D}_\sigma, \mathfrak{D}_{\sigma+r})$. Note in

particular that

$$(4.3) \quad f(\mathfrak{D}_\sigma, \mathfrak{D}_{\sigma+r}) = r + \mathcal{O}(1)$$

as $r \rightarrow \infty$.

4.2. Spectral flow for the Dirac operator. This preceding formula has a generalization that applies to Dirac operators on any odd dimensional, compact manifold. The generalization is described in [T6]. What follows describes the story in the case of a 3-dimensional manifold. Let $F \rightarrow M$ denote a given $\text{Spin}_\mathbb{C}$ structure and let A denote a given connection on S_F . The associated Dirac operator, D_A is a self-adjoint, unbounded operator on $L^2(M; \mathbb{S})$ with purely point spectrum. Each eigenvalue has finite multiplicity and the eigenvalues do not accumulate in \mathbb{R} .

Suppose now that D_A has trivial kernel. Let A' denote a second connection on S_F and suppose that $D_{A'}$ also has trivial kernel. In this instance, one can define the spectral flow from D_A to $D_{A'}$ as follows: Choose a differentiable path $\{A(t)\}_{t \in [0,1]} \in \text{Conn}(S_F)$ that starts at A and ends at A' . If the path is chosen in suitably generic fashion, then the following will be true: First, there is but a finite set of points in $[0, 1]$ where $D_{A(\cdot)}$ has non-trivial kernel. Second, if t is such a point, then the zero eigenvalue of $D_{A(t)}$ will have multiplicity 1 and any $D_{A(t')}$ for t' near t will have exactly one simple eigenvalue very close to zero. Let $\lambda(t')$ denote this simple eigenvalue that is nearly zero. This assignment $t' \rightarrow \lambda(t')$ will be differentiable for t' near t , and its derivative at t will be non-zero. This understood, $t \in [0, 1]$ contributes $+1$ to the spectral flow when the derivative is positive and -1 when the derivative is negative. The analysis in [Ka] can be used to prove that the associated sum of $+1$'s and -1 's is independent of the chosen path between A and A' . This weighted sum is deemed to be the spectral flow from D_A to $D_{A'}$. It is denoted by $f(D_A, D_{A'})$.

Now fix an i -valued 1-form \hat{a} on M . Here is a question: What is the leading order in r behavior of the spectral flow for the family $r \rightarrow D_{A+r\hat{a}}$ as $r \in [0, \infty)$ gets large? Note that this question makes sense even in the case that $D_{A+r\hat{a}}$ has non-zero kernel. Indeed, as explained in [Ka], the spectrum varies with r in a real analytic fashion and so the subset in $[0, \infty)$ where $D_{A+r\hat{a}}$ has a zero eigenvalue is discrete with no accumulation points. This understood, define the spectral flow to be

$$(4.4) \quad f(D_A, D_{A+r\hat{a}}) = \lim_{\varepsilon \rightarrow 0} f(D_A, D_{A+r+\varepsilon}).$$

Here is the answer:

PROPOSITION 4.1. *There is a constant, c , that depends on A and \hat{a} and is such that the spectral flow function $r \rightarrow f(D_A, D_{A+r\hat{a}})$ differs from*

$$-\frac{1}{32\pi^2} r^2 \int_M \hat{a} \wedge d\hat{a},$$

by no more than $c(r^{2-\delta} + 1)$ where $\delta > 0$ is independent of A and \hat{a} .

This proposition is proved in Section 5 of [T1].

For a relevant example, take S_F to be the trivial bundle. Let A denote a given connection on S_F . Take $\hat{a} = -ia$ where a is now a contact 1-form. Take the metric to be such that $*da = 2a$ and $|a| = 1$. In this case, the spectral flow (over \mathbb{C}) for the family $r \rightarrow D_{A+ra}$ is

$$(4.5) \quad \frac{1}{16\pi^2}r^2 + \mathcal{O}(r^{2-\delta}).$$

It is also the case that this spectral flow is one sided in the sense that the zero crossings of eigenvalues when r is large all count with positive weight to the spectral flow.

4.3. Spectral flow and the Seiberg-Witten equations. A relative \mathbb{Z} grading of the Seiberg-Witten Floer homology is also defined via spectral flow. What follows describes how this comes about. The linearization of (2.1) at any given $(A, \psi) \in \text{Conn}(S_F) \times C^\infty(M; \mathbb{S})$ and $r \geq 1$ defines a self adjoint, elliptic operator on $L^2(M; iT^*M \oplus \mathbb{S} \oplus i\mathbb{R})$ with purely point spectrum having finite multiplicity and no accumulation points. This operator is denoted by $\mathfrak{L} = \mathfrak{L}_{r,(A,\psi)}$. It sends any given triple (b, η, ϕ) to the section whose respective iT^*M, \mathbb{S} and $i\mathbb{R}$ components are

$$(4.6) \quad \begin{aligned} &* db - d\phi - 2^{-1/2}r^{1/2}(\psi^\dagger\tau\eta + \eta^\dagger\tau\psi), \\ &D_A\eta + 2^{1/2}r^{1/2}(\text{cl}(b)\psi + \phi\psi), \\ &* d * b - 2^{-1/2}r^{1/2}(\eta^\dagger\psi - \psi^\dagger\eta). \end{aligned}$$

Suppose that \mathfrak{L}_0 is defined by (4.6) using r_0 and (A_0, ψ_0) and \mathfrak{L}_1 is defined using $r_1, (A_1, \psi_1)$. Then the spectral flow $f(\mathfrak{L}_0, \mathfrak{L}_1)$ is defined just as it was for the case of the Dirac operator; one chooses a suitably generic path $\{t \rightarrow (r_t, (A_t, \psi_t))\}_{t \in [0,1]}$ that interpolates between the two data sets, and takes an algebraic count of the number of points in $[0, 1]$ where the $(r_t, (A_t, \psi_t))$ version of (4.6) has a zero eigenvalue. Strictly speaking, such a definition makes sense only if neither \mathfrak{L}_0 nor \mathfrak{L}_1 has eigenvalue 0. If this is not the case, then something along the lines of (4.4) is used to make f well defined.

Now fix $r_0 = 1$ and suppose that $(A_0, \psi_0) \in \text{Conn}(S_F) \times C^\infty(M; \mathbb{S})$ has been chosen so that the r_0 and (A_0, ψ_0) version of (4.6) has trivial kernel. (Such a pair always exists). Let \mathfrak{L}_0 denote this operator. Suppose next that $r \geq 1$ and (A, ψ) solve (2.1) and are such that their version of (4.6) also has trivial kernel. Let \mathfrak{L} denote the latter. If $\psi \neq 0$, then the gauge equivalence class of (A, ψ) can appear as part of a generator of the Seiberg-Witten Floer homology. In this regard, the corresponding homology class has degree equal to $-f(\mathfrak{L}_0, \mathfrak{L})$.

Meanwhile, an upper bound for the \mathbb{Z} gradings of the countable set of homology classes that are associated to a reducible solution $(A = \theta - \frac{i}{2}ra, \psi = 0)$ is given by minus the spectral flow from \mathfrak{L}_0 to the r and

($A = \theta - ira, \psi = 0$) version of (4.6). The latter operator is

$$(4.7) \quad \begin{aligned} & * db - d\phi, \\ & D_{A-ira}\eta, \\ & * d * b. \end{aligned}$$

Note in particular, that r appears here only in the Dirac operator. In particular it follows from (4.5) that the reducibles contribute only very negative degree Seiberg-Witten Floer homology classes when r is large. This is a very important point.

Section 3.2 alludes to the birth and death of irreducible solutions to (2.1) at some reducible solution ($A = \theta - ira, \psi = 0$). As noted in Section 3.2, this birth/death phenomena can occur only when D_{A-ira} has a zero eigenvalue. This is a consequence of the fact that (4.7) depicts the linearization of the equations in (2.1) at the given reducible solution. In particular, this birth/death business can be seen purely in the context of perturbation theory.

Note, however, that the one sided nature of the spectral flow for D_{A-ira} as r increases implies that the newly born or newly killed irreducible solutions define Seiberg-Witten Floer homology classes with ever more negative degree as r increases. Of course, such must be the case if the classes that are associated to the equivalence classes of reducible solutions have degree upper bound given by minus what is written in (4.5).

To summarize: Theorem 3.2 and (4.5) imply the following:

PROPOSITION 4.2. *Given $k \in \mathbb{Z}$, there exists r_k with the following significance: Suppose that $r \geq r_k$.*

- *All Seiberg-Witten Floer homology classes with degree k or greater are represented by formal sums of irreducible solutions to (2.1).*
- *Given $N \in \{1, 2, \dots\}$, there exists $k \in \mathbb{Z}$ such that when $r \geq r_k$, there are at least N non-zero Seiberg-Witten Floer homology classes with degree k or greater.*

This last proposition keeps hope alive for a desirable answer to (2.4)'s second question.

4.4. Spinors with length near 1 everywhere. This subsection addresses the second of the three questions posed in (2.4). There are two cases to consider here: The first occurs when the bundle E in (2.2) has non-zero first Chern class. Suppose that this is the case. Then any section of E must vanish at points on M . In particular, the component α of a section $\psi = (\alpha, \beta)$ of \mathbb{S} must be zero at points in M . If ψ comes from a pair (A, ψ) that solves some large r version of (2.1), then the second bullet of Lemma 2.3 implies that $1 - |\psi| \geq \kappa r^{-1/2}$ at points where $\alpha = 0$. This understood, then the third question in (2.4) has answer: Always, if E has non-zero first Chern class.

Now assume that E is isomorphic to the product line bundle $M \times \mathbb{C}$. Fix such an isomorphism and let ψ_0 denote the section of \mathbb{S} given by $(1, 0)$. Take a fixed connection A_0 on S_F . For $r \geq 1$, use \mathfrak{L}_{r0} to denote the r and $(A = A_0, \psi = (1, 0))$ version of (4.6). A quick glance at the analog of (1.12) for \mathfrak{L}_{r0}^2 finds that \mathfrak{L}_{r0}^2 is very positive when r is large: Its smallest eigenvalue is no less than $\frac{1}{8}r$ when r is large. As a consequence, the smallest of the absolute values of the eigenvalues of \mathfrak{L}_{0r} is no less than $\frac{1}{4}r^{1/2}$ when r is large. This implies that there is no spectral flow for the path $r \rightarrow \{\mathfrak{L}_{r0}\}$ as long as r is sufficiently large.

Suppose that r is large and that (A, ψ) is a solution to (2.1) and suppose that $|\psi|$ is everywhere very close to 1. This is to say that $1 - |\psi| \leq \delta$ where δ is a small, but r -independent constant. Let \mathfrak{L} denote the corresponding version of (4.6). With the help of the Bochner-Weitzenboch formula for \mathfrak{L}^2 , Lemma 2.3 can be used to prove that the smallest of the absolute values of the eigenvalues of \mathfrak{L} is also no less than $\frac{1}{8}r^{1/2}$.

According to Lemma 2.3, if $|\psi|$ is nearly 1 everywhere, then so is $|\alpha|$. As a consequence (A, ψ) is gauge equivalent to (A', ψ') with $\psi' = (1, 0)$ plus a term that has norm $\mathcal{O}(\delta)$. This fact can be used to construct a path $t \rightarrow (A_t, \psi_t)$ that starts at $(A_0, (1, 0))$ and ends at (A', ψ') with no spectral flow for the corresponding path $t \rightarrow \{\mathfrak{L}_{r,(A_t,\psi_t)}\}$.

Granted such a path, there exists an r -independent lower bound to the degree of any solution of (2.1) with $1 - |\psi|$ everywhere less than δ . What with Proposition 4.2, this last conclusion implies the following answer to the third question in (2.4): Always.

As an aside, the $\mathcal{O}(r)$ lower bound to the spectrum of \mathfrak{L}_{r0}^2 allows a perturbative proof of the following [T2]: Suppose that E is isomorphic to the trivial bundle. Then there exists $\delta > 0$ and a unique gauge equivalence class of solution to any sufficiently large r version of (2.1) with $1 - |\psi| < \delta$ on the whole of M .

5. Bounds on the energy from the spectral flow

This section outlines how certain properties of the spectral flow can be used to answer the third question posed in (2.4). The story starts with an analog of Proposition 4.1 for the operator in (4.6).

5.1. Asymptotic spectral flow. The promised analog of Proposition 4.1 refers to the version of (4.6) that is defined by a triple $(r, (A, \psi))$ in the case that (A, ψ) solves (2.1). Here it is:

PROPOSITION 5.1. *Fix a configuration $(A_0, \psi_0) \in \text{Conn}(S_F) \times C^\infty(M; \mathbb{S})$ such that the $r = 1$ and $(A, \psi) = (A_0, \psi_0)$ version of (4.6) has trivial kernel. There exists constants $\delta > 0$ and $\kappa > 1$ with the following significance: Fix $r \geq 1$ and a solution (A, ψ) to (2.1). Let \mathfrak{L} denote the corresponding version of (4.6). Then $f(\mathfrak{L}_0, \mathfrak{L})$ differs from $\frac{1}{16\pi^2} \text{cs}(A)$ by no more than $\kappa r^{2-\delta}$.*

The proof of Proposition 5.1 can also be found in Section 5 of [T1]. Proposition 5.1 has the following corollary:

PROPOSITION 5.2. *Fix an integer $k \in \mathbb{Z}$ and there exists $r_k \geq 1$ with the following significance: Suppose that Θ denotes a non-zero Seiberg-Witten Floer homology class with degree k . If $r \geq r_k$ and if (A, ψ) is an irreducible solution to (2.1) whose equivalence class appears in the formal sum for Θ , then $|\mathfrak{cs}(A)| \leq r^{2-\delta}$ where $\delta > 0$ is independent of r, k and (A, ψ) .*

Proposition 5.2 is used to obtain an amiable answer to the third question in (2.4); this despite the fact that Proposition 5.2 refers solely to \mathfrak{cs} .

5.2. The Chern Simons functional and the energy. The functionals \mathfrak{cs} and \mathbb{E} are linked in two ways. First, both \mathfrak{cs} and \mathbb{E} appear in the function on $\text{Conn}(S_F) \times C^\infty(M; \mathbb{S})$ whose critical points give the solutions to (2.1). The latter function is

$$(5.1) \quad \mathfrak{a} = \frac{1}{2}(\mathfrak{cs} - r \mathbb{E}) + 2 \int_M \psi^\dagger D_A \psi,$$

The second link between \mathfrak{cs} and \mathbb{E} is supplied by the following lemma:

LEMMA 5.3. *There is a constant $\kappa > 1$ with the following significance: Suppose that $r \geq 1$ and that (A, ψ) is a solution to (2.1) in the case when \mathbb{S} has torsion first Chern class. Then*

$$|\mathfrak{cs}| \leq \kappa \left(r^{2/3} \mathbb{E}^{4/3} + 1 \right).$$

The proof of this last lemma uses Lemma 2.3 and the properties of a Green’s function for the operator $d + *d*$ acting on the space of smooth 1-forms on M .

5.3. An idealized scenario. To see how \mathfrak{a} enters the story, suppose for the moment that $r_0 \geq 1$ and that $r \rightarrow \mathfrak{c}(r)$ is a differentiable map from $[r_0, \infty)$ into $\text{Conn}(S_F) \times C^\infty(M; \mathbb{S})$ such that $\mathfrak{c}(r)$ obeys r ’s version of (2.1) when $r \geq r_0$. View $\mathfrak{a}(\mathfrak{c}(r))$ and also \mathfrak{cs} and \mathbb{E} on $\mathfrak{c}(r)$ as functions on $[r_0, \infty)$. Then

$$(5.2) \quad \frac{d}{dr} \mathfrak{a} = -\frac{1}{2} \mathbb{E}.$$

Indeed, this is the case because $\mathfrak{c}(r)$ is a critical point of (5.1) and so solves (2.1). Integrate this last equation so as to obtain

$$(5.3) \quad \frac{1}{r} \mathfrak{a}|_r = \frac{1}{r} \mathfrak{a}|_{r_0} + \frac{1}{2} \int_{r_0}^r \mathfrak{cs}|_s s^{-2} ds.$$

Lemma 5.3 and (5.3) imply that one of the following two assertions must hold:

- There exists $\mathcal{E} < \infty$ and an unbounded set $\{r_n\}_{n=1,2,\dots} \subset [r_0, \infty)$ such that $E(\mathfrak{c}(r_n)) < \mathcal{E}$.
- (5.4)
- There exists $\kappa \geq 1$ such that $E(\mathfrak{c}(r)) \geq \kappa^{-1}r$ and $\mathfrak{cs}(\mathfrak{c}(r)) \geq \kappa^{-1}r^2$ if r is sufficiently large.

Now let \mathfrak{L}_r denote the version of (4.6) that is defined by r and $\mathfrak{c}(r)$. Proposition 5.1 rules out the second option in (5.4) given an a priori bound on the absolute value for the spectral flow for the family $\{\mathfrak{L}_r\}_{r \geq r_0}$ as $r \rightarrow \infty$. Of course, such is the case when there is zero spectral flow for this family. For example, were there a fixed Seiberg-Witten Floer homology class Θ and were each $\mathfrak{c}(r)$ to appear in a formal sum that defines Θ , then there could be no spectral flow for the family $\{\mathfrak{L}_r\}_{r \geq r_0}$. This is how Proposition 5.1 controls the size of E .

5.4. A realistic scenario. The scenario that was just presented may not occur because there need not exist a differential map $r \rightarrow \mathfrak{c}(r)$ of the required sort. The fact that the Seiberg-Witten Floer homology classes as defined at different values of r are isomorphic does not by itself imply the existence of the required maps into $\text{Conn}(S_F) \times C^\infty(M; \mathbb{S})$. However, such maps do exist on intervals of small length where r is large. The proposition that follows elaborates.

PROPOSITION 5.4. *Fix an integer k and a non-zero Seiberg-Witten Floer homology class in degree k or greater. There exists $r_k \geq 1$, an increasing set $\{\rho_m\}_{m=1,2,\dots} \subset [r_k, \infty)$ and for each m , a smooth map $r \rightarrow \mathfrak{c}(r) \in \text{Conn}(S_F) \times C^\infty(M; \mathbb{S})$ that is defined for $r \in [\rho_m, \rho_{m+1}]$ and is such that the gauge equivalence class of $\mathfrak{c}(r)$ appears in the formal sum for the class Θ . Moreover, the function $r \rightarrow \mathfrak{a}(\mathfrak{c}(r))$ that is defined on $[r_k, \infty) - \{\rho_m\}_{m=1,1,\dots}$ extends to $[r_k, \infty)$ so as to be continuous, and piecewise differentiable.*

Granted that (5.3) holds with $r_0 = \rho_m$ and $r \in [\rho_m, \rho_{m+1}]$ for each m , then the conclusions given in (5.4) still follow. As the spectral flow between any two r and $\mathfrak{c}(r)$ versions of (4.6) is zero, so the second option in (5.4) is again ruled out. Thus, the second question in (2.4) has a very amiable answer: If \mathbb{S} has torsion first Chern class, then the answer is: Always.

Proposition 5.4 is proved in Section 4 of [T1].

6. Hutchings embedded contact homology

Embedded contact homology is a very clever and subtle variant of the sort of contact homology/symplectic field theory that was introduced by Eliashberg, Givental and Hofer [EGH]. (As with Seiberg-Witten Floer homology, contact homology owes much to the pioneering work of Floer [F].)

Hutchings describes his embedded contact homology theory in [HS] and also [HT]. See also [H]. What follows is a very rough synopsis of what is

involved. To start, the homology theory assigns to each contact 3-manifold and each $\text{Spin}_{\mathbb{C}}$ structure, a vector space over \mathbb{Z} . This vector space is defined from a larger vector space, this the vector space of *chains*. It is denoted in what follows by \mathcal{C} . This vector space \mathcal{C} is freely generated over \mathbb{Z} by equivalence classes of a certain sort of finite, ordered set. An element in any such set is an ordered pair (γ, m) where $\gamma \subset M$ is a closed orbit of the Reeb vector field, where m is a positive integer subject to certain restrictions.

The $\text{Spin}_{\mathbb{C}}$ structure enters the story through one of these restrictions. To elaborate, introduce $[\gamma]$ to denote the class in the first homology of M that is generated by γ with its orientation defined by the restriction of a . Let Θ denote a set of pairs as just described. Then $2 \sum_{(\gamma, m) \in \Theta} m[\gamma]$ must be Poincaré dual to the sum of the first Chern class of the associated \mathbb{C}^2 bundle \mathbb{S} and the first Chern class of the bundle K .

A typical element in the vector space of \mathcal{C} can be written as a finite, formal sum

$$(6.1) \quad c = \sum z_{\Theta} \Theta,$$

where Θ is as just described and where $z_{\Theta} \in \mathbb{Z}$. The equivalence relation involves the ordering the elements of Θ , and it changes the sign of z_{Θ} under suitable circumstances when the ordering is changed.

The vector spaces for embedded contact homology are defined using a certain homomorphism $\delta : \mathcal{C} \rightarrow \mathcal{C}$ with $\delta^2 = 0$. Given that $\delta^2 = 0$, the kernel of δ contains the image of δ and so it makes sense to consider the quotient of the former by the latter. The embedded contact homology space is just this quotient of $\text{kernel}(\delta)$ by $\text{image}(\delta)$.

6.1. Pseudoholomorphic subvarieties. The homomorphism δ is defined using surfaces in the manifold $\mathbb{R} \times M$ with appropriate limits at $\pm\infty$ along the \mathbb{R} factor. To say more about the relevant surfaces, use v in what follows to denote the Reeb vector field on M . Fix a homomorphism, J , on the tangent space of $\mathbb{R} \times M$ with the following properties:

- $J^2 = -1$.
 - J is invariant under the \mathbb{R} action of $\mathbb{R} \times M$ that is induced by translations on the \mathbb{R} factor.
- (6.2) • Let $\frac{\partial}{\partial s}$ denote the unit length tangent vector along the \mathbb{R} factor.
- Then $J \frac{\partial}{\partial s} = v$.
- J maps the kernel of a in TM to itself, and $da(\cdot, J(\cdot))$ defines a symmetric, positive definite quadratic form on the kernel of a .

A homomorphism with $J^2 = -1$ is said to be an *almost complex structure*. One that obeys all of (6.2) is deemed a *compatible almost complex structure*.

A closed subset $\Sigma \subset \mathbb{R} \times M$ is said to be a *pseudoholomorphic subvariety* when it has the following properties:

- *The complement of a finite set of points is a smooth submanifold with J -invariant tangent space.*
- (6.3)
 - $\int_{\Sigma_s} da < \infty$. Here $\Sigma_s \subset \Sigma$ denotes the set of smooth points with the orientation defined locally by any ordered pair (w, Jw) with $w \in T\Sigma - 0$.

It follows from the first point that the restriction of da to $T\Sigma_s$ is a non-negative multiple of the induced area form. Note that the notion of pseudoholomorphic subvarieties depends on the choice of a compatible almost complex structure J . There are many available.

Hofer [Hof1] proved a remarkable theorem about pseudoholomorphic subvarieties. In what follows, s denotes the Euclidean coordinate on the \mathbb{R} factor of $\mathbb{R} \times M$. Here is Hofer’s theorem:

THEOREM 6.1. *Fix a pseudoholomorphic subvariety and there exists $R \geq 1$ such that the $|s| \geq R$ portion of the subvariety is a finite, disjoint union of embedded cylinders to which s restricts without critical points. Moreover, any constant s slice of any such cylinder defines a non-trivial braid in a tubular neighborhood of some closed integral curve of v . Finally, the s -dependent family of such braids converges pointwise to this integral curve as $|s| \rightarrow \infty$.*

For those not familiar with the terminology, the notions of tubular neighborhood and braid are as follows: There is a disk $D \subset \mathbb{C}$ with center at the origin and a diffeomorphism $\varphi : S^1 \times D \rightarrow M$ that maps $S^1 \times 0$ onto the given integral curve. The image of $S^1 \times D$ is a tubular neighborhood. Viewed from $S^1 \times D$, a braid is the image of a smooth map from S^1 into $S^1 \times D$ that sends $t \in S^1 = \mathbb{R}/(2\pi\mathbb{Z})$ to $(qt, z(t))$ where z is a smooth map to $\mathbb{C} - \{0\}$ and where q is a positive integer. The integer q is called the multiplicity of the braid.

The set of large $|s|$ component cylinders of a given pseudoholomorphic subvariety are its *ends*. Any given end has its associated limit closed integral curve of v , and its associated braid integer q . An end that lies where $s \ll -1$ is deemed to be a negative end, and an end where $s \gg 1$ is deemed to be a positive end. When γ is a closed integral curve of v , use $\mathcal{E}_{\Sigma_{\gamma-}}$ to denote the set of negative ends of Σ that appear as a braid in a tubular neighborhood of γ . When $\mathcal{E}_{\Sigma_{\gamma-}} \neq \emptyset$, set $m_{\gamma-}$ to denote the sum of the multiplicities of the braids that come from the ends in $\mathcal{E}_{\Sigma_{\gamma-}}$. Set $m_{\gamma} = 0$ when $\mathcal{E}_{\Sigma_{\gamma-}} = \emptyset$. Define $\mathcal{E}_{\Sigma_{\gamma+}}$ and $m_{\gamma+}$ in an analogous fashion using the positive ends of Σ .

Granted Hofer’s theorem, a pseudoholomorphic subvariety Σ can be used to define two finite sets, $\Theta_{\Sigma-}$ and $\Theta_{\Sigma+}$, whose typical element is a pair

(γ, m) where γ is a closed integral curve of v and m is a positive integer. Set $\Theta_{\Sigma_{\pm}} = \{(\gamma, m_{\gamma_{\pm}}) : m_{\gamma_{\pm}} > 0\}$.

6.2. The definition of I . Now suppose that Θ_- and Θ_+ are two sets of the sort that are used to define the vector space \mathcal{C} . Let $\Sigma \subset \mathbb{R} \times M$ denote a pseudoholomorphic subvariety with $\Theta_{\Sigma_-} = \Theta_-$ and $\Theta_{\Sigma_+} = \Theta_+$. Hutchings associates to Σ an integer I that can be viewed as a sum of three terms. The first is the difference between a term that is defined using Θ_- and an analogous one defined using Θ_+ . The contribution to either for a given pair (γ, m) of closed integral curve of v and positive integer m depends on the choice of the tubular neighborhood map for γ . This contribution is the sum, indexed by $k \in \{1, \dots, m\}$, of the so-called Conley-Zhender index for the integral curve of v that wraps k times around γ . The second contribution to I is denoted by $c_1(K|_{\Sigma})$. This integer is obtained by pairing the fundamental class of Σ with the pull-back of the first Chern class of K to Σ via the projection from $\mathbb{R} \times M$ to M . Given that Σ is not compact, such a pairing requires a trivialization of K 's pull-back over the ends of Σ . The required trivialization is obtained from the previously chosen tubular neighborhood maps for closed integral curves of v that help define Θ_- and Θ_+ .

The third contribution to I is a relative self-intersection number for Σ . Its definition requires pushing Σ off of itself and then making a suitably weighted count of the intersections between the original and the push-off. The allowed push-offs are constrained along the ends of Σ and the reader is referred to Section 11 of [HS] for the precise instructions.

6.3. The definition of δ . Let $\mathcal{M}(\Theta_-, \Theta_+)$ denote the set of pseudoholomorphic subvarieties with $\Theta_{\Sigma_{\pm}} = \Theta_{\pm}$ and with $I = 1$. Hutchings proves the following: If J is suitably generic, then

- $\mathcal{M}(\Theta_-, \Theta_+)$ is a finite union of smooth, 1-dimensional manifolds.
- Each submanifold in $\mathcal{M}(\Theta_-, \Theta_+)$ is embedded, and any two in the same component differ by a constant translation along the \mathbb{R} factor of $\mathbb{R} \times M$.
- Each component of $\mathcal{M}(\Theta_-, \Theta_+)$ has an associated sign, either $+1$ or -1 .

The sign arises by comparing two canonical orientations that can be associated to each component of $\mathcal{M}(\Theta_-, \Theta_+)$. One orientation is that induced by the translations along the \mathbb{R} factor of $\mathbb{R} \times M$. The second is defined along lines introduced by Quillen; the definition uses a $\mathbb{Z}/(2\mathbb{Z})$ analog of spectral flow for families of Fredholm operators. This $\mathbb{Z}/(2\mathbb{Z})$ analog of spectral flow is defined as follows for two index zero operators with trivial cokernel: Take a suitably generic interpolating path of Fredholm operators and count the number (mod 2) of points on the path where the kernel is non-trivial. This count can be interpreted as a relative orientation of the kernels of the two operators. The latter interpretation generalizes to the case of positive index

operators and so gives a relative orientation to the kernels of any two such operators if they have trivial cokernel.

The operator in question in the case at hand is obtained as follows: A small distance push-off of Σ is also pseudoholomorphic if and only if it satisfies a certain non-linear, partial differential equation. The linearization of this equation defines a \bar{d} -bar operator on Σ with an \mathbb{R} -linear zero'th order term. This \mathbb{R} -linear \bar{d} -bar operator can be viewed as a Fredholm operator, and it is the latter that gives, via Quillen's construction, the second orientation to Σ 's component in $\mathcal{M}(\Theta_-, \Theta_+)$. The details are in [BM].

The homomorphism δ that defines embedded contact homology acts on a generator Θ of \mathcal{C} as

$$(6.5) \quad \delta\Theta = \sum_{\Theta' \in \mathcal{C}} \sigma(\Theta', \Theta) \Theta'$$

where $\sigma(\Theta', \Theta) \in \mathbb{Z}$ is zero when $\mathcal{M}(\Theta', \Theta) = \emptyset$ and otherwise it is the sum of the signs of that are associated via the third bullet of (6.4) to the components of $\mathcal{M}(\Theta', \Theta)$. The Eliashberg-Hofer-Givental contact homology is defined in much the same manner save for the definition of I . Hutchings version of I appears to have magical powers.

6.4. Embedded contact homology and Seiberg-Witten Floer homology. Embedded contact homology is intrinsically defined using the closed integral curves of the Reeb vector field for a contact structure on M . Seiberg-Witten Floer homology has, on the face of it, nothing to do with contact structures, but as just seen, it too sees something of the closed integral curves of the Reeb vector field. Hutchings conjectured that these two homologies are essentially the same:

HUTCHINGS' CONJECTURE. There is an isomorphism between the Seiberg-Witten Floer cohomology and the embedded contact homology as defined for any given contact 1-form and any given $\text{Spin}_{\mathbb{C}}$ -structure on any given compact, oriented 3-manifold.

The Seiberg-Witten Floer cohomology is a sort of dual to the homology.

Hutchings confirmed this conjecture for certain contact structures on the 3-sphere, on the 3-torus and on $S^1 \times S^2$. Hutchings also noted that embedded contact homology has $\mathbb{Z}/p\mathbb{Z}$ grading of the same sort as Seiberg-Witten Floer cohomology.

Since writing this article, the author has completed a proof of Hutchings' conjecture that uses many of the same ideas that enter the proof of the Weinstein conjecture [T7].

References

- [ACH] C. Abbas, K. Cielebak and H. Hofer, *The Weinstein conjecture for planar contact structures in dimension three*, preprint arXiv:math.SG/0409355v2 March 2005.
- [A] M. F. Atiyah, *Floer homology, in the Floer memorial volume*, Progr. Math 133 (1995) 77–97.

- [BM] F. Bourgeois and K. Mohnke, *Coherent orientations in symplectic field theory*, Math. Zeitschrift (2004), 123–146.
- [Bo1] R. Bott, *Morse theory indomitable*, Publ. IHES 68 (1988), 99–114.
- [Bo2] R. Bott, *Lectures on Morse theory, old and new*, Bull. Amer. Math. Soc. 7 (1982), 331–358.
- [CH] V. Colin and K. Honda, *Reeb vector fields and open book decompositions I: the periodic case*, preprint 2005.
- [Ch] W. Chen, *Pseudo-holomorphic curves and the Weinstein conjecture*, Comm. Anal. Geom. 8 (2000), 115–131.
- [EGH] Y. Eliashburg, A. Givental and H. Hofer, *Introduction to symplectic field theory*, GAFA 2000 (Tel Aviv 1999), Geom. Funct. Anal. 2000, Special Volume, Part II, 560–673.
- [F] A. Floer, *Morse theory for Lagrangian intersections*, J. Diff. Geom. 28 (1988), 513–547.
- [Gi] V. L. Ginzburg, *The Weinstein conjecture and the theorems of nearby and almost existence*, in The Breadth of Symplectic and Poisson Geometry. Festschrift in Honor of Alan Weinstein; J.E. Marsden and T.S. Ratiu (Eds.), Birkhuser, 2005, pp. 139–172.
- [Ga] D. Gay, *Four dimensional symplectic cobordisms containing three-handles*, Geometry and Topology 10 (2006) 1749–1759.
- [Hof1] H. Hofer, *Pseudoholomorphic curves in symplectizations with applications to the Weinstein conjecture in dimension three*, Invent. Math. 114 (1993), 515–563.
- [Hof2] H. Hofer, *Dynamics, topology and holomorphic curves*, in Proceedings of the International Congress of Mathematicians, Vol I (Berlin 1998), Doc. Math. 1998 Extra Vol. I, 255–280.
- [Hof3] H. Hofer, *Holomorphic curves and dynamics in dimension three*, in Symplectic Geometry and Topology, IAS/Part City Math Ser., 7, Amer. Math. Soc., Providence R. I. 1999, 35–101.
- [Hof4] H. Hofer, *Holomorphic curves and real three dimensional dynamics*, Geom. Funct. Anal. 2000, Special Volume, Part II 674–704.
- [Hon] K. Honda, *The topology and geometry of contact structures in dimension three*, International Congress of Mathematicians. Vol. II, 705–717, Eur. Math. Soc., Zrich, 2006.
- [H] M. Hutchings, *An index inequality for embedded pseudoholomorphic curves in symplectizations*, J. Eur. Math. Soc. (JEMS) 4 (2002) 313–361.
- [HS] M. Hutchings and M. Sullivan, *Rounding corners of polygons and embedded contact homology*, Geometry and Topology 10 (2006) 169–266.
- [HT] M. Hutchings and C. H. Taubes, *Gluing pseudoholomorphic curves along branched covered cylinders I*; Jour. Sympl. Geom., to appear; arXiv:math/0701300.
- [Ka] T. Kato, *Perturbation theory for linear operators*, Springer-Verlag, Berlin-Heidelberg-New York, 1966.
- [KM] P. Kronheimer and T. Mrowka, *Monopoles and Three-Manifolds*, Cambridge University Press; to appear.
- [Mil] J. Milnor, *Morse Theory*, Princeton University Press 1963.
- [Sm] S. Smale, *An infinite dimensional version of Sard's theorem*, Amer. J. Math. 87 (1968) 861–866.
- [T1] C. H. Taubes, *The Seiberg-Witten equations and the Weinstein conjecture*, Geometry and Topology 11 (2007) 2117–2202.
- [T2] C. H. Taubes, *The Seiberg-Witten equations and the Weinstein conjecture II: More closed integral curves for the Reeb vector field*. arXiv:math/0702366.
- [T3] C. H. Taubes, *The Seiberg-Witten and the Gromov invariants*, Math. Res. Letters 2 (1995) 221–238.

- [T4] C. H. Taubes, *Seiberg-Witten and Gromov Invariants for Symplectic 4-manifolds*, International Press, Somerville MA, 2005.
- [T5] C. H. Taubes, *SW \Rightarrow Gr: From Seiberg-Witten equations to pseudo-holomorphic curves*, in *Seiberg-Witten and Gromov Invariants for Symplectic 4-manifolds* by C. H. Taubes, International Press, Somerville MA, 2005.
- [T6] C. H. Taubes, *Asymptotic spectral flow for Dirac operators*, *Commun. Anal. Geom.* 15 (2007) 215–233.
- [T7] C. H. Taubes, *Embedded contact homology and Seiberg-Witten Floer Cohomology I–V*, preprints. Paper 1 is posted as arXiv:0811.3985v2.
- [U] K. K. Uhlenbeck, *Connections with L_p bounds on curvature*, *Comm. Math. Phys.* 83 (1982) 31–42.
- [W] A. Weinstein, *On the hypotheses of Rabinowitz's orbit theorems*, *J. Diff. Eq.* 33 (1979) 353–358.

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, CAMBRIDGE, MA 02138
E-mail address: `chtaubes@math.harvard.edu`

