TAUBES'S PROOF OF THE WEINSTEIN CONJECTURE IN DIMENSION THREE

MICHAEL HUTCHINGS

Abstract. Does every smooth vector field on a closed three-manifold, for example the three-sphere, have a closed orbit? No, according to counterexamples by K. Kuperberg and others. On the other hand there is a special class of vector fields, called Reeb vector fields, which are associated to contact forms. The three-dimensional case of the Weinstein conjecture asserts that every Reeb vector field on a closed oriented three-manifold has a closed orbit. This conjecture was recently proved by Taubes using Seiberg-Witten theory. We give an introduction to the Weinstein conjecture, the main ideas in Taubes's proof, and the bigger picture into which it fits.

Taubes's proof of the Weinstein conjecture is the culmination of a large body of work, both by Taubes and by others. In an attempt to make this story accessible to nonspecialists, much of the present article is devoted to background and context, and Taubes's proof itself is only partially explained. Hopefully this article will help prepare the reader to learn the full story from Taubes's paper [62]. More exposition of this subject (which was invaluable in the preparation of this article) can be found in the online video archive from the June 2008 MSRI hot topics workshop [44], and in the article by Auroux [5].

Below, in §1–§3 we introduce the statement of the Weinstein conjecture and discuss some examples. In §4–§6 we discuss a natural strategy for approaching the Weinstein conjecture, which proves it in many but not all cases, and provides background for Taubes's work. In §7 we give an overview of the big picture surrounding Taubes's proof of the Weinstein conjecture. Readers who already have some familiarity with the Weinstein conjecture may wish to start here. In §8–§9 we recall necessary material from Seiberg-Witten theory. In §10 we give an outline of Taubes's proof, and in §11 we explain some more details of it. To conclude, in §12 we discuss some further results and open problems related to the Weinstein conjecture.

1. Statement of the Weinstein conjecture

The Weinstein conjecture asserts that certain vector fields must have closed orbits. Before stating the conjecture at the end of this section, we first outline its origins. This is discussion is only semi-historical, because only a sample of the relevant works will be cited, and not always in chronological order.

1.1. Closed orbits of vector fields.

Let Y be a closed manifold (in this article all manifolds and all objects defined on them are smooth unless otherwise stated),

---

2000 *Mathematics Subject Classification.* 57R17,57R57,53D40.

Partially supported by NSF grant DMS-0806037.

and let V be a vector field on Y. A *closed orbit* of V is a map

$$\gamma : \mathbb{R}/T\mathbb{Z} \to Y$$

for some $T > 0$, satisfying the ordinary differential equation

$$\frac{d\gamma(t)}{dt} = V(\gamma(t)).$$

Given a closed manifold Y, one can ask: Does every vector field on Y have a closed orbit? If the Euler characteristic $\chi(Y) \neq 0$, then by the Poincaré-Hopf index theorem every vector field on Y has zeroes, which give rise to constant closed orbits. In this article we will mainly focus on the case where Y is three-dimensional. Then $\chi(Y) = 0$, which implies that the signed count of zeroes of a generic vector field is zero; and it is relatively easy to cancel these in order to construct a vector field on Y with no zeroes. But understanding nonconstant closed orbits is much harder.

Although for some special three-manifolds such as the 3-torus it is easy to construct vector fields with no closed orbit, already for $Y = S^3$ the question of whether all vector fields have closed orbits is very difficult. It turns out that the answer is no. Examples of vector fields on S^3 with no closed orbit, with increasing degrees of regularity, were constructed by P. Schweizer (C^1), J. Harrison (C^2), K. Kuperberg (C^∞), G. Kuperberg and K. Kuperberg (real analytic), and G. Kuperberg (C^1 and volume preserving), see [58, 28, 41, 40, 39]. Moreover, as explained in [40], such vector fields exist on any 3-manifold. The constructions of these examples involve modifying a given vector field by inserting a suitable "plug" in a neighborhood of a point on a closed orbit, which destroys that closed orbit without creating new ones. The fact that one can do this indicates that to gain any control over closed orbits, one needs to make some additional assumption on the vector field.

1.2. Hamiltonian vector fields.

A very important class of vector fields, originating in classical mechanics, are Hamiltonian vector fields. We briefly recall the definition of these.

Let (M^{2n}, ω) be a symplectic manifold. This means that ω is a closed 2-form on M such that $\omega^n \neq 0$ everywhere, or equivalently ω defines a nondegenerate bilinear form on the tangent space T_xM at each point $x \in M$. The basic example of a symplectic manifold is $\mathbb{R}^{2n}$ with coordinates $x_1, \ldots, x_n, y_1, \ldots, y_n$ and the standard symplectic form

$$\omega_{std} = \sum_{i=1}^{n} dx_i \wedge dy_i.$$

We will drop the subscript '*std*' when it is understood. Darboux's theorem asserts that any symplectic manifold (M^{2n}, ω) is locally equivalent to $(\mathbb{R}^{2n}, \omega_{std})$. More precisely, for any $x \in M$ there exists a neighborhood U of x in M, an open set $V \subset \mathbb{R}^{2n}$, and a symplectomorphism $\phi : (U, \omega) \xrightarrow{\simeq} (V, \omega_{std})$, ie a diffeomorphism $\phi : U \xrightarrow{\simeq} V$ such that $\phi^*\omega_{std} = \omega$.

If $H : M \to \mathbb{R}$ is a smooth function, the associated *Hamiltonian vector field* is the vector field X_H on M characterized by

$$\omega(X_H, \cdot) = dH.$$

The vector field X_H exists and is unique by the nondegeneracy of ω.

For example, if $M = \mathbb{R}^{2n}$ with the standard symplectic form, then

$$X_H = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial y_i} \frac{\partial}{\partial x_i} - \frac{\partial H}{\partial x_i} \frac{\partial}{\partial y_i} \right).$$

So a trajectory $(x(t), y(t))$ of the vector field X_H satisfies the equations

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial y_i}, \qquad \frac{dy_i}{dt} = -\frac{\partial H}{\partial x_i}.$$

These are the equations of classical mechanics, where the x_i's are position coordinates, the y_i's are momentum coordinates, and H is the Hamiltonian, or energy function.

Conservation of energy in classical mechanics translates to the fact that on a general symplectic manifold, the flow of X_H preserves H, because

$$X_H(H) = dH(X_H) = \omega(X_H, X_H) = 0.$$

Therefore if $E \in \mathbb{R}$ is a regular value of H, then the level set $H^{-1}(E)$ is a $(2n-1)$-dimensional submanifold of M, which we will call a "regular level set", and X_H is a smooth vector field on it.

Must the Hamiltonian vector field X_H have a closed orbit on every regular level set? It turns out that under favorable circumstances the answer is "almost every". For example, suppose $M = \mathbb{R}^{2n}$ with the standard symplectic form. Then one has the following:

Theorem 1.1 (Hofer-Zehnder, Struwe, 1990)**.** *Let $H : \mathbb{R}^{2n} \to \mathbb{R}$ be a proper smooth function. Then the vector field X_H has a closed orbit on $H^{-1}(E)$ for almost every $E \in \mathbb{R}$ such that $H^{-1}(E) \neq \emptyset$.*

The original proof, see the exposition in [31], uses variational methods to find critical points of the "symplectic action" functional on the loop space, which are the desired closed orbits. Using Floer theory for the symplectic action functional (different versions of Floer theory will be introduced later in this article), this result can be extended to some other symplectic manifolds; for a survey of some of these developments see [25].

However there are some symplectic manifolds to which this theorem does not extend. To give a simple example due to Zehnder [75], consider the 4-torus $T^4 = (\mathbb{R}/2\pi\mathbb{Z})^4$ with coordinates x_1, x_2, x_3, x_4 and the nonstandard symplectic form

$$\omega = dx_1 dx_2 + \varepsilon dx_2 dx_3 + dx_3 dx_4,$$

where ε is an irrational constant. If $H = \sin x_4$, then

$$X_H = \cos x_4 \left(\frac{\partial}{\partial x_3} + \varepsilon \frac{\partial}{\partial x_1} \right).$$

Since ε is irrational, this vector field has no closed orbits except where $\cos x_4 = 0$. Thus no regular level set of H contains a closed orbit of X_H.

In addition, the word "almost" cannot be removed from Theorem 1.1. Namely, Ginzburg-Gürel [26] proved that there is a proper C^2 Hamiltonian H on $\mathbb{R}^4$ with a regular level set on which the vector field X_H has no closed orbit. Moreover C^∞ examples are known in $\mathbb{R}^{2n}$ for $n > 2$, see the references in [25]. So even for Hamiltonian vector fields on $\mathbb{R}^{2n}$ one needs a further assumption to guarantee the existence of closed orbits.

1.3. **Weinstein's conjecture.** To see what kind of assumption might be appropriate, let Y be a hypersurface in a symplectic manifold (M,ω) which is a regular level set of a Hamiltonian $H : M \to \mathbb{R}$. Observe that the existence of a closed orbit of X_H on Y depends only on the hypersurface Y and not on the Hamiltonian H. For suppose $K : M \to \mathbb{R}$ is another Hamiltonian which also has Y as a regular level set. Then $dK|_Y = f dH|_Y$ for some nonvanishing function $f : Y \to \mathbb{R}$, so $X_K = f X_H$ on Y. Thus the periodic orbits of X_K and X_H on Y differ only in their parametrizations. In fact one can describe the existence problem for periodic orbits on Y without referring to a Hamiltonian at all by noting that the Hamiltonian vector field on Y always lives in the rank 1 subbundle

$$L_Y := \operatorname{Ker}(\omega|_Y) \subset TY,$$

called the *charateristic foliation*. Thus the existence of a closed orbit on Y for any Hamiltonian having it as a regular level set is equivalent to the existence of a closed curve tangent to L_Y, sometimes called a *closed characteristic*.

Under what circumstances must a hypersurface in a symplectic manifold have a closed characteristic? In the late 1970's, Weinstein [72] showed that in $\mathbb{R}^{2n}$ with the standard symplectic form, if Y is a convex compact hypersurface, then Y has a closed characteristic. Rabinowitz [53] showed that the above statement holds with "convex" replaced by "star-shaped", meaning transverse to the radial vector field

$$\rho := \frac{1}{2}\sum_{i=1}^{n}(x_i\partial_{x_i} + y_i\partial_{y_i}). \tag{1.1}$$

Now the existence of a closed characteristic is invariant under symplectomorphisms of $\mathbb{R}^{2n}$, while the star-shaped condition is not. This suggests that one should look for a more general, symplectomorphism invariant condition on the hypersurface Y that might guarantee the existence of a closed characteristic. Weinstein [73] proposed such a condition as follows.

Definition 1.2. A hypersurface Y in a symplectic manifold (M,ω) is of *contact type* if there is a 1-form λ on Y such that $d\lambda = \omega|_Y$, and $\lambda(v) \neq 0$ for all nonzero $v \in L_Y$.

This condition is clearly invariant under symplectomorphisms of (M,ω).

If Y is a star-shaped hypersurface in $\mathbb{R}^{2n}$, then Y is of contact type, because one can take

$$\lambda = \frac{1}{2}\sum_{i=1}^{n}(x_i dy_i - y_i dx_i)|_Y.$$

To see why $\lambda(v) \neq 0$ for nonzero $v \in L_Y$, note that nondegeneracy of ω implies that $\omega(v,w) \neq 0$ for any vector w transverse to Y. The star-shaped condition says that ρ is transverse to Y, so $\omega(\rho, v) \neq 0$; on the other hand it follows from the above formulas that $\omega(\rho, v) = \lambda(v)$.

More generally, a *Liouville vector field* on a symplectic manifold (M,ω) is a vector field ρ such that the Lie derivative $\mathcal{L}_\rho\omega = \omega$. The radial vector field (1.1) is an example of a Liouville vector field. It turns out that a hypersurface Y in (M,ω) is of contact type if and only if there exists a Liouville vector field defined in a neighborhood of Y which is transverse to Y, see eg [45, Prop. 3.58]. In particular, given such a vector field ρ, the contact type condition is fulfilled by $\lambda = \omega(\rho,\cdot)|_Y$.

One could now conjecture that if Y is a compact hypersurface of contact type in a symplectic manifold (M, ω), then Y has a closed characteristic. This is essentially what Weinstein conjectured in [73].

1.4. **Contact forms.** In fact one can remove the symplectic manifold (M, ω) from the discussion as follows. Let Y be an oriented manifold of dimension $2n - 1$. A *contact form* on Y is a 1-form λ on Y such that

$$\lambda \wedge (d\lambda)^{n-1} > 0$$

everywhere. A contact form λ determines a vector field R on Y, called the *Reeb vector field*, characterized by

$$d\lambda(R, \cdot) = 0, \qquad \lambda(R) = 1.$$

In the definition of a contact type hypersurface, one can replace the condition "$\lambda(v) \neq 0$ for all nonzero $v \in L_Y$" by the condition "λ is a contact form". Note that for such a contact form, the associated Reeb vector field is tangent to the characteristic foliation L_Y.

Conversely, any manifold with a contact form (Y, λ) arises as a hypersurface of contact type in an associated symplectic manifold. Namely, the *symplectization* of (Y, λ) is the manifold $\mathbb{R} \times Y$ with the symplectic form $\omega = d(e^s \lambda)$, where s denotes the $\mathbb{R}$ coordinate. The slice $\{0\} \times Y$ is a hypersurface of contact type in $\mathbb{R} \times Y$, with λ fulfilling the definition of contact type.

In conclusion, the conjecture at the end of §1.3 is equivalent to the following:

Weinstein Conjecture. *Let Y be a closed oriented odd-dimensional manifold with a contact form λ. Then the associated Reeb vector field R has a closed orbit.*

The eventual goal of this article is to explain Taubes's proof of the Weinstein conjecture in dimension three:

Theorem 1.3 (Taubes). *If Y is a closed oriented three-manifold with a contact form, then the associated Reeb vector field has a closed orbit.*

1.5. **Some remarks on Taubes's proof.** Before starting on the long story which follows, let us briefly comment on what is involved in Taubes's proof, and the history leading up to it.

First of all, as we will see in §2, §3, and §6 below, methods from symplectic geometry have been used since the 1980's to prove the Weinstein conjecture in many cases, in both 3 and higher dimensions. But it is not currently known if such methods can be used to prove it in all cases.

Taubes's proof instead uses Seiberg-Witten theory. As we describe in §7, the proof uses part of a 3-dimensional version of Taubes's "Seiberg-Witten=Gromov" (SW=Gr) theorem. The SW=Gr theorem relates Seiberg-Witten invariants of symplectic 4-manifolds to counts of holomorphic curves. Its 3-dimensional counterpart relates Seiberg-Witten theory of a 3-manifold with a contact form to closed orbits of the Reeb vector field (and holomorphic curves in the 3-manifold cross $\mathbb{R}$).

The SW=Gr theorem was proved in the 1990's, and one might wonder why the 3-dimensional Weinstein conjecture was not also proved around that time using similar methods. Part of what was missing was a sufficient development of Seiberg-Witten Floer homology, the 3-dimensional counterpart of Seiberg-Witten invariants of 4-manifolds, which has since been provided by Kronheimer-Mrowka. In particular Kronheimer-Mrowka proved a nontriviality result for Seiberg-Witten

Floer homology, which is the key input on the Seiberg-Witten side in the proof of the Weinstein conjecture. Finally, as we explain in §10, the 3-dimensional analogue of SW=Gr is not just a straightforward adaptation of the 4-dimensional story, but rather involves some nontrivial new ingredients, for example a new result of Taubes estimating the spectral flow of families of Dirac operators in terms of the Chern-Simons functional.

What follows now is a somewhat leisurely tour, gradually working towards the above. The impatient reader may wish to skip ahead to §7 for an overview of the big picture surrounding Taubes's proof, or to §10 for an outline of the proof itself.

1.6. **Some terminology.** Below, to save space, we usually say "Reeb orbit" instead of "closed orbit of the Reeb vector field". Two Reeb orbits $\gamma, \gamma' : \mathbb{R}/T\mathbb{Z} \to Y$ are considered equivalent if they differ by reparametrization, ie precomposition with a translation of $\mathbb{R}/T\mathbb{Z}$. If $\gamma : \mathbb{R}/T\mathbb{Z} \to Y$ is a Reeb orbit and k is a positive integer, then the k-fold *iterate* of γ is the pullback of γ to $\mathbb{R}/kT\mathbb{Z}$. A Reeb orbit γ is embedded if and only if it is not the k-fold iterate of another Reeb orbit where $k > 1$.

2. BASIC EXAMPLES IN $2n-1$ DIMENSIONS

2.1. **Hypersurfaces in $\mathbb{R}^{2n}$.** The Weinstein conjecture for compact hypersurfaces of contact type in $\mathbb{R}^{2n}$ was proved in 1987 by Viterbo [71]. In fact the almost-existence result in Theorem 1.1 is a generalization of this. To see why, let Y be a compact hypersurface in $\mathbb{R}^{2n}$ of contact type. As discussed in §1.3, there exists a Liouville vector field ρ defined on a neighborhood of Y which is transverse to Y. Flowing the hypersurface Y along the vector field ρ for a small time ε gives another hypersurface Y_ε and a diffeomorphism $Y \simeq Y_\varepsilon$. Since ρ is transverse to Y, there exists $\delta > 0$ such that the different hypersurfaces Y_ε for $|\epsilon| < \delta$ are disjoint and sweep out a neighborhood that can be identified in the obvious way with $(-\delta, \delta) \times Y$. Also, the Liouville condition $\mathcal{L}_\rho \omega = \omega$ implies that the above diffeomorphism $Y \simeq Y_\varepsilon$ respects the characteristic foliations. Thus Y_ε has a closed characteristic for either all ε or none. On the other hand we can choose a proper Hamiltonian $H : \mathbb{R}^{2n} \to \mathbb{R}$ such that $H = \varepsilon$ on Y_ε and $|H| > \delta$ outside of our neighborhood $(-\delta, \delta) \times Y$. Then Theorem 1.1 implies that Y_ε has a closed characteristic for almost every ε. Hence there is a closed characteristic for every ε, and in particular for $\varepsilon = 0$.

2.2. **Cotangent bundles.** Another important example of a manifold with a contact form is the unit cotangent bundle of a Riemannian manifold.

To start, let Q be a smooth manifold. There is a canonical 1-form λ on the cotangent bundle T^*Q, defined as follows. Let $\pi : T^*Q \to Q$ denote the projection. If $q \in Q$, and if $p \in T_q^*Q$, then $\lambda : T_{(q,p)}T^*Q \to \mathbb{R}$ is the composition

$$T_{(q,p)}T^*Q \xrightarrow{\pi_*} T_qQ \xrightarrow{p} \mathbb{R}.$$

More explicitly, if $q_1, \ldots, q_n$ are local coordinates on a coordinate patch $U \subset Q$, then one can write any cotangent vector at a point in U as $p = \sum_{i=1}^n p_i dq_i$, and this gives local coordinates $q_1, \ldots, q_n, p_1, \ldots, p_n$ on $\pi^{-1}(U) \subset T^*Q$. In these coordinates

$$\lambda = \sum_{i=1}^n p_i dq_i.$$

It follows from this last equation that $d\lambda$ defines a symplectic form on T^*Q.

Now suppose that Q has a Riemannian metric. This induces a metric on T^*Q, and we consider the unit cotangent bundle

$$ST^*Q = \{p \in T^*Q \mid |p| = 1\}.$$

The restriction of λ to ST^*Q is a contact form. Indeed ST^*Q is a hypersurface of contact type in the symplectic manifold $(T^*Q, d\lambda)$, with transverse Liouville vector field $\rho = \sum_{i=1}^{n} p_i \partial_{p_i}$, and $\imath_\rho d\lambda = \lambda$. It turns out that the associated Reeb vector field agrees with the geodesic flow, under the identification $T^*Q = TQ$ given by the metric. Thus Reeb orbits in ST^*Q are equivalent to closed geodesics in Q. If Q is compact, then so is ST^*Q, so the Weinstein conjecture is applicable in this case, where it is equivalent to the classical Lyusternik-Fet theorem asserting that every compact Riemannian manifold has at least one closed geodesic.

More generally, the above Liouville vector field ρ shows that a hypersurface $Y \subset T^*Q$ which intersects each fiber transversely in a star-shaped subset of the fiber is of contact type. Rabinowitz proved the existence of closed characteristics for a related class of hypersurfaces in $\mathbb{R}^{2n}$. Alan Weinstein points out to me that this example and the unit cotangent bundle example were important motivation for his conjecture.

2.3. **Prequantization spaces.** Another general example of manifolds with contact forms is given by circle bundles, or "prequantization spaces". Let (Σ, ω) be a symplectic manifold of dimension $2n-2$, and suppose that the cohomology class $-[\omega]/2\pi \in H^2(\Sigma; \mathbb{R})$ is the image of an integral class $e \in H^2(\Sigma; \mathbb{Z})$. Let $p : V \to \Sigma$ be the principal S^1-bundle with first Chern class e. This means that there is a free S^1 action on V whose quotient is Σ, and $e \in H^2(\Sigma; \mathbb{Z})$ is the primary obstruction to finding a section $\Sigma \to V$. Let R denote the derivative of the S^1 action; this is a vector field on V which is tangent to the fibers. Since ω is a closed form in the cohomology class $-2\pi e$, one can find a (real-valued) connection 1-form λ on V whose curvature equals ω. These conditions mean that λ is invariant under the S^1 action, $\lambda(R) = 1$, and $d\lambda = p^*\omega$. It follows that λ is a contact form on V whose Reeb vector field is R. In particular, the Reeb orbits are the fibers (which all have period 2π) and their iterates.

The fact that Reeb orbits appear here in $(2n-2)$-dimensional smooth families is a special feature arising from the symmetry of the picture. For a "generic" contact form on a manifold Y the Reeb orbits are isolated, in the sense that if γ is a Reeb orbit of length T which goes through a point $x \in Y$, then there is no other Reeb orbit through a point close to x with length close to T. For example, on a circle bundle as above, one can get rid of most of the Reeb orbits by perturbing the contact form λ to

$$\lambda' = (1 + p^*H)\lambda,$$

where $H : \Sigma \to \mathbb{R}$ is a smooth function. This is still a contact form as long as $|H| < 1$. The new Reeb vector field is given by

$$R' = (1 + p^*H)^{-1}R + (1 + p^*H)^{-2}\widetilde{X}_H, \tag{2.1}$$

where X_H is the Hamiltonian vector field on Σ determined by H, and $\widetilde{X}_H$ denotes its horizontal lift, ie the unique vector field on V with $p_*\widetilde{X}_H = X_H$ and $\lambda(\widetilde{X}_H) = 0$. Consequently, for this new contact form, the only fibers that are Reeb orbits are the fibers over the critical points of H. On the other hand there may be additional Reeb orbits that cover closed orbits of X_H. However it follows from (2.1) that if

H and dH are small, then these Reeb orbits all have period much greater than 2π. In any case, the Weinstein conjecture here asserts that there is no way to eliminate all of the remaining Reeb orbits without introducing new ones. The Weinstein conjecture in this case can be proved using cylindrical contact homology [20, §2.9], about which we will have more to say in §6.

3. More about contact geometry in three dimensions

We now restrict attention to the three-dimensional case. Much more is known about contact geometry in three dimensions than in higher dimensions, and to gain an appreciation for Taubes's result and its proof it will help to review some of the basics of this. For much more about this subject we refer the reader to [21] and [24].

3.1. Contact structures in three dimensions.

Recall that a contact form on a closed oriented three-manifold Y is a 1-form λ such that $\lambda \wedge d\lambda > 0$ everywhere. The associated *contact structure* is the 2-plane field $\xi = \operatorname{Ker}(\lambda)$. This has an orientation induced from the orientation of Y and the direction of the Reeb vector field. In general one defines a contact structure[1] to be an oriented 2-plane field which is the kernel of some contact form as above. A contact structure is a "totally nonintegrable" 2-plane field, which means that in a sense it as far as possible from being a foliation: the kernel of λ is a foliation if and only if $\lambda \wedge d\lambda \equiv 0$.

Different contact forms can give rise to the same contact structure. To be precise, if λ is a contact form, then λ' is another contact form giving rise to the same contact structure if and only if $\lambda' = f\lambda$ where $f : Y \to \mathbb{R}$ is a positive smooth function. For a given contact structure, the Reeb vector field depends on the choice of contact form, but it is always transverse to the contact structure.

A (three-dimensional closed) *contact manifold* is a pair (Y, ξ) where Y is a closed oriented three-manifold and ξ is a contact structure on Y. Two contact manifolds (Y, ξ) and (Y, ξ') are *isomorphic*, or *contactomorphic*, if there is an orientation-preserving diffeomorphism $\phi : Y \to Y'$ such that ϕ_* sends ξ to ξ' preserving the orientiations. Two contact structures ξ and ξ' on Y are *isotopic* if there is a one-parameter family of contact structures $\{\xi_t \mid t \in [0,1]\}$ on Y such that $\xi_0 = \xi$ and $\xi_1 = \xi'$. Gray's stability theorem asserts that ξ and ξ' are isotopic if and only if there is a contactomorphism between them which is isotopic to the identity.

A version of Darboux's theorem asserts that any contact structure on a 3-manifold is locally isomorphic to the "standard contact structure" on $\mathbb{R}^3$, which is the kernel of the contact form

$$\lambda = dz - y\,dx. \tag{3.1}$$

In fact any contact form is locally diffeomorphic to this one. The contact structure defined by (3.1) is invariant under translation in the z direction. The contact planes are horizontal along the x axis, but rotate as one moves in the y direction; the total rotation angle as y goes from $-\infty$ to $+\infty$ is π. The Reeb vector field associated to λ is simply

$$R = \partial_z.$$

---

[1] Sometimes this is called a "co-oriented contact structure". There is also a notion of unoriented contact structure in which ξ is not assumed to be oriented and is only required to be locally the kernel of a contact form.

In particular there are no Reeb orbits. Of course this does not contradict the Weinstein conjecture since $\mathbb{R}^3$ is not compact, but it does indicate that any proof of the Weinstein conjecture will need to use global considerations.

3.2. Tight versus overtwisted.

Define an *overtwisted disk* in a contact 3-manifold (Y, ξ) to be a smoothly embedded closed disk $D \subset Y$ such that for each $y \in \partial D$ we have $T_y D = \xi_y$. A contact 3-manifold is called *overtwisted* if it contains an overtwisted disk; otherwise it is called *tight*.

An example of an overtwisted contact structure on $\mathbb{R}^3$ is the kernel of the contact form given in cylindrical coordinates by

$$\lambda = \cos r\, dz + r \sin r\, d\theta.$$

This contact structure is invariant under translation in the z direction. Here the contact planes are horizontal on the z axis, but rotate infinitely many times as one moves out from the z axis along a horizontal ray. An overtwisted disk is given by a horizontal disk of radius r, where r is a positive number such that $\sin r = 0$.

On the other hand, the standard contact structure defined by (3.1) is tight, although this is less trivial to prove. More generally, (Y, ξ) is called (strongly symplectically) *fillable* if there is a compact symplectic 4-manifold (X, ω) with boundary $\partial X = Y$ as oriented manifolds, and a contact form λ for (Y, ξ), such that $d\lambda = \omega|_Y$. (In particular Y is a hypersurface of contact type in X, and an associated Liouville vector field points out of the boundary of X.) A fundamental theorem in the subject asserts that any fillable contact structure is tight [19].

3.3. Simple examples.

To get some examples of contact forms on 3-manifolds, recall from §1.3 that any star-shaped hypersurface Y in $\mathbb{R}^4$ has a contact form

$$\lambda = \frac{1}{2}\sum_{i=1}^{2}(x_i dy_i - y_i dx_i)|_Y.$$

The resulting contact structure on $Y \simeq S^3$ is tight, because it is filled by the solid region that Y bounds in $\mathbb{R}^4$. A theorem of Eliashberg asserts that all tight contact structures on S^3 are isotopic to this one.

If Y is the unit sphere in $\mathbb{R}^4$, then the Reeb vector field is tangent to the fibers of the Hopf fibration $S^3 \to S^2$. In particular, there is a family of Reeb orbits parametrized by S^2. This is in fact a special case of the circle bundle example that we considered previously, and in particular it is not “generic”. If one replaces the sphere with the ellipsoid

$$\frac{x_1^2 + y_1^2}{a_1} + \frac{x_2^2 + y_2^2}{a_2} = 1 \tag{3.2}$$

where a_1, a_2 are positive real numbers, then the Reeb vector field is given by

$$R = 2\sum_{i=1}^{2} a_i^{-1}\left(x_i \partial_{y_i} - y_i \partial_{x_i}\right).$$

This vector field rotates in the x_i, y_i plane at angular speed $2a_i^{-1}$. Thus if a_1/a_2 is irrational, then there are just two embedded Reeb orbits, namely the circles $x_1 = y_1 = 0$ and $x_2 = y_2 = 0$. The Weinstein conjecture says that we can not further modify the contact form to eliminate these two remaining orbits without introducing new ones.

We next consider some examples of contact forms on the 3-torus T^3. Write $T^3 = (\mathbb{R}/2\pi\mathbb{Z})^3$ with coordinates x, y, z. For each positive integer n, define a contact form λ_n on T^3 by

$$\lambda_n := \cos(nz)dx + \sin(nz)dy. \tag{3.3}$$

The associated Reeb vector field is given by

$$R_n = \cos(nz)\partial_x + \sin(nz)\partial_y. \tag{3.4}$$

We can regard T^3 as a T^2-bundle over S^1, where z is the coordinate on S^1 and x, y are the fiber coordinates. The Reeb vector field is then a linear vector field tangent to each fiber, whose slope rotates as z increases. Whenever the slope is rational, the fiber is foliated by Reeb orbits. That is, there is a circle of embedded Reeb orbits for each z such that $\tan(nz) \in \mathbb{Q} \cup \{\infty\}$. Again this is a non-generic situation, and it turns out for any such z, one can perturb the contact form so that the corresponding circle of Reeb orbits disintegrates into just two Reeb orbits. We will see in §6 below that the contact structures $\xi_n := \operatorname{Ker}(\lambda_n)$ are pairwise non-contactomorphic. Also they are all tight. Note that ξ_1 is isomorphic to the canonical contact structure on the unit contangent bundle of T^2 with a flat metric.

Remark 3.1. One can use the above example to illustrate that Reeb vector fields are somewhat special. To see how, note that in general if R is the Reeb vector field associated to a contact form λ on a 3-manifold, then R is volume preserving with respect to the volume form $\lambda \wedge d\lambda$, because the definition of Reeb vector field implies that the Lie derivative $\mathcal{L}_R\lambda = 0$. But not every volume-preserving vector field is a Reeb vector field. In fact the Reeb vector field (3.4) can easily be perturbed to a volume-preserving vector field with no closed orbits (which by the Weinstein conjecture cannot be a Reeb vector field). Namely, consider the vector field

$$V = \cos(nz)\partial_x + (\sin(nz) + \varepsilon_1)\partial_y + \varepsilon_2\partial_z$$

where $\varepsilon_1, \varepsilon_2 \neq 0$ and $\varepsilon_1/\varepsilon_2$ is irrational. Suppose that $(x(t), y(t), z(t))$ is a trajectory of V. Then $z(t) = z(0) + \varepsilon_2 t$, so if this is a closed orbit then the period must be $2\pi k/\varepsilon_2$ for some positive integer k. But the path $(x(t), y(t))$ moves in the sum of a circular motion with period $2\pi/(n\varepsilon_2)$ and a vertical motion of speed ε_1, so we have $(x(2\pi k/\varepsilon_2), y(2\pi k/\varepsilon_2)) = (0, 2\pi k\varepsilon_1/\varepsilon_2)$. Thus our assumption that $\varepsilon_1/\varepsilon_2$ is irrational implies that there is no closed orbit.

3.4. **Classification of overtwisted contact structures.**

A contact structure is a particular kind of oriented 2-plane field, and if two contact structures are isotopic then the corresponding oriented 2-plane fields are homotopic. A remarkable theorem of Eliashberg implies that for overtwisted contact structures, the converse is true:

Theorem 3.2 (Eliashberg [18]). *For any closed oriented 3-manifold Y, the inclusion of the set of overtwisted contact structures on Y into the set of oriented 2-plane fields on Y is a homotopy equivalence.*

A detailed exposition of the proof may be found in [24]. In particular, this theorem implies that overtwisted contact structures modulo isotopy are equivalent to homotopy classes of oriented 2-plane fields. Note that the latter always exist[2], because an oriented 3-manifold has trivial tangent bundle.

---

[2] In higher dimensions, not as much is known about which manifolds admit contact structures. For example it was only in 2002 that odd-dimensional tori were shown to admit contact structures,

Remark 3.3. To prepare for the discussion of Seiberg-Witten theory later, it is worth saying a bit more here about what the set of homotopy classes of oriented 2-plane fields on a given closed oriented (connected) Y looks like. First note that a homotopy class of oriented 2-plane fields on Y is equivalent to a homotopy class of nonvanishing sections of TY. In particular, if ξ and ξ' are two oriented 2-plane fields on Y, then the primary obstruction to finding a homotopy between them is an element of $H^2(Y;\mathbb{Z})$. If the primary obstruction vanishes, then it turns out that the remaining obstruction lives in $\mathbb{Z}/d$, where d denotes the divisibility of $c_1(\xi) = c_1(\xi')$ in $H^2(Y;\mathbb{Z})$ mod torsion, which is always an even integer.

The classification of tight contact structures is more complicated, and a subject of ongoing research. In particular, the map from tight contact structures to homotopy classes of oriented 2-plane fields is in general neither injective nor surjective. Failure of injectivity is illustrated by the contact structures ξ_n on T^3 in §3.3, which are pairwise non-contactomorphic even though they all represent the same homotopy class of oriented 2-plane fields. Failure of surjectivity follows for example from a much stronger theorem of Colin-Giroux-Honda [14], which asserts that on any given closed oriented 3-manifold there are only finitely many homotopy classes of oriented 2-plane fields that contain tight contact structures (even though there are always infinitely many homotopy classes of oriented 2-plane fields, as follows from Remark 3.3 above).

3.5. **Open book decompositions.**

There is a useful classification of all contact three-manifolds, not just the overtwisted ones, in terms of open book decompositions.

Let Σ be a compact oriented connected surface with nonempty boundary. Let $\phi : \Sigma \to \Sigma$ be an orientation-preserving diffeomorphism which is the identity near the boundary. One can then define a closed oriented three-manifold

$$\begin{aligned} Y_\phi &:= [0,1]\times\Sigma/\sim, \\ (1,x) &\sim (0,\phi(x)) \quad \forall x\in\Sigma, \\ (t,x) &\sim (t',x) \qquad \forall x\in\partial\Sigma,\ t,t'\in[0,1], \end{aligned}$$

called an *open book*. The image of a set $\{t\}\times\Sigma$ in Y_ϕ is called a *page*. The boundaries of the different pages are all identified with each other, to an oriented link in Y_ϕ called the *binding*. The map ϕ is called the *monodromy* of the open book. An *open book decomposition* of a closed oriented three-manifold Y is a diffeomorphism of Y with an open book Y_ϕ as above.

Definition 3.4. A contact structure ξ on a closed three-manifold Y is *compatible* with, or *supported* by, an open book decomposition $Y \simeq Y_\phi$ if ξ is isotopic to a contact structure given by a contact form λ such that:

- The Reeb vector field is tangent to the binding (oriented positively).
- The Reeb vector field is transverse to the interior of each page, intersecting positively.

---

by Bourgeois [8]. In dimensions greater than three, there is a homotopy-theoretic obstruction: a closed oriented $(2n-1)$ manifold Y admits a (cooriented) contact structure only if the structure group of the tangent bundle TY reduces to $U(n-1)$. I am not aware of any further known obstructions. For some positive results in the 5-dimensional case see eg [24, §8].

Example 3.5. The standard contact structure on the unit sphere described in §3.3 is compatible with an open book decomposition of S^3 in which Σ is a disk and ϕ is the identity map.

A short argument by Thurston-Winkelnkemper shows that every open book decomposition has a compatible contact structure. In fact the compatible contact structure is determined up to isotopy by the open book. Moreover, a theorem of Giroux asserts that every contact structure can be obtained in this way, and two open books determine isotopic contact structures if and only if they are related to each other by "positive stabilizations". For more about this see [27, 22].

3.6. **Some previous results on the 3d Weinstein conjecture.** There is a long history of work proving the Weinstein conjecture for various classes of contact three-manifolds. Often one can further show that Reeb orbits with certain properties must exist. To give just a few examples here:

Theorem 3.6 (Hofer [29])**.** *Let (Y,ξ) be a closed contact 3-manifold in which ξ is overtwisted or $\pi_2(Y)\neq 0$. Then for any contact form with kernel ξ, there exists a* contractible *Reeb orbit.*

Theorem 3.7 (Abbas-Cieliebak-Hofer [2])**.** *Suppose (Y,ξ) is supported by an open book in which the pages have genus zero. Then for any contact form with kernel ξ, there exists a nonempty finite collection of Reeb orbits $\{\gamma_i\}$ with*

$$\sum_i [\gamma_i] = 0 \in H_1(Y).$$

We remark that the above theorems find Reeb orbits as ends of punctured holomorphic spheres in $\mathbb{R}\times Y$ with only positive ends, cf §6.2. In the first theorem there is only one puncture, while in the second theorem the number of punctures can be any positive integer.

Colin and Honda used linearized contact homology (see §6 below) to prove the Weinstein conjecture for contact three-manifolds supported by open books in which the monodromy is periodic, and in many cases where the monodromy is pseudo-Anosov. See [15] for the precise statement. In fact, for many contact structures supported by open books with pseudo-Anosov monodromy, they proved a much stronger statement: that for any contact form, there are infinitely many free homotopy classes of loops that contain an embedded Reeb orbit.

4. Some strategies for proving the Weinstein conjecture

One naive strategy for proving the Weinstein conjecture might be to try the following:

- Define some kind of count of Reeb orbits with appropriate signs.
- Show that this count is a topological invariant.
- Calculate this invariant and show that it is nonzero.

This strategy is too simple for at least two reasons. First of all, often there are actually infinitely many embedded Reeb orbits (see §12.1), so it is not clear how to obtain a well-defined count of them. Second, even if the above difficulty can be overcome, the signed count might be zero, despite the existence of some Reeb orbits with opposite signs that cannot be eliminated. In general, as one deforms the contact form, pairs of Reeb orbits can be created or destroyed (more complicated bifurcations such as period-doubling are also possible), and one needs some way of keeping track of when this can happen.

A more refined strategy, which avoids the above two problems, is as follows:

- Define some kind of chain complex which is generated by Reeb orbits, such that, roughly speaking, there are differentials between pairs of Reeb orbits that can potentially be destroyed in a bifurcation.
- Show that the homology of this chain complex is a topological invariant.
- Compute this homology and show that it is nontrivial.

It turns out that there does exist a chain complex along these lines which is sufficient to prove the Weinstein conjecture. However it is not the first chain complex that one might think of, and the proof that it works uses Seiberg-Witten theory. We will now attempt to explain this story.

5. PROTOTYPE FOR A CHAIN COMPLEX: MORSE HOMOLOGY

The prototype for the type of chain complex we want to consider is Morse homology, which we now review. We will not give any proofs, as these require a fair bit of analysis; details can be found in [56]. There is also an interesting history of the development of the Morse complex, for which we refer the reader to [7].

5.1. Morse functions.

Let X be an n-dimensional closed smooth manifold and let $f : X \to \mathbb{R}$ be a smooth function. A *critical point* of f is a point $p \in X$ such that $0 = df_p : T_pX \to \mathbb{R}$. The basic goal of Morse theory is to relate the critical points of f to the topology of X. A first question is, on a given X, what is the minimum number of critical points that a smooth function f can have?

To make this question easier, we can require that the critical points of f be "generic", in a sense which we now specify. If $p \in X$ is a critical point of $f : X \to \mathbb{R}$, define the *Hessian*

$$H(f,p) : T_pX \otimes T_pX \longrightarrow \mathbb{R} \tag{5.1}$$

as follows. Let $\psi : X \to T^*X$ denote the section corresponding to df. Then $H(f,p)$ is the composition

$$T_pX \xrightarrow{d\psi_p} T_{(p,0)}T^*X = T_pX \oplus T_p^*X \xrightarrow{\pi} T_p^*X,$$

where π denotes the projection onto the second factor. To be more explicit, if $(x_1, \ldots, x_n)$ are local coordinates on X centered at p, then

$$H(f,p)\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right) = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

In particular, $H(f,p)$ is a symmetric bilinear form. If one chooses a metric on X, then the Hessian can be identified with a self-adjoint operator

$$H(f,p) : T_pX \longrightarrow T_pX. \tag{5.2}$$

The critical point p is said to be *nondegenerate* if the Hessian pairing (5.1) is nondegenerate, or equivalently the Hessian operator (5.2) does not have zero as an eigenvalue, or equivalently the graph of df in T^*X is transverse to the zero section at $(p,0)$. In particular, nondegenerate critical points are isolated in X.

We say that f is a *Morse function* if all of its critical points are nondegenerate. One can show that "generic" smooth functions are Morse. More precisely, the set of Morse functions is open and dense in the set of all smooth functions $f : X \to \mathbb{R}$, with the C^∞ topology. We can now ask, what is the minimum possible number of critical points of a Morse function on X?

One could start by counting the critical points with signs. If p is a (nondegenerate) critical point of f, define the *index* of p, denoted by $\operatorname{ind}(p)$, to be the maximal dimension of a subspace on which the Hessian pairing (5.1) is negative definite, or equivalently the number of negative eigenvalues, counted with multiplicity, of the Hessian operator (5.2). For example, a local minimum has index 0, and a local maximum has index n. It turns out that an appropriate sign with which to count an index i critical point is $(-1)^i$, and the signed count is the Euler characteristic of X. That is, if $c_i(f)$ denotes the number of critical points of index i, then

$$\sum_{i=0}^{n}(-1)^i c_i(f) = \chi(X).$$

One can prove this by choosing a metric on X, applying the Poincaré-Hopf index theorem to the resulting gradient vector field ∇f, and checking that the sign of the zero of ∇f at an index i critical point is $(-1)^i$. In particular, the number of critical points is at least $|\chi(X)|$. But if $\chi(X) = 0$ then this tells us nothing.

5.2. **The Morse complex.** One way to obtain better lower bounds on the number of critical points is to consider "gradient flow lines" between critical points and package these into a chain complex. Here is how this works.

Choose a metric g on X and use it to define the gradient vector field ∇f. If p and q are critical points, a (downward) *gradient flow line* from p to q is a map $\gamma : \mathbb{R} \to X$ such that

$$\frac{d\gamma(s)}{ds} = \nabla f(\gamma(s))$$

and $\lim_{s\to+\infty}\gamma(s) = p$ and $\lim_{s\to-\infty}\gamma(s) = q$. Let $\mathcal{M}(p, q)$ denote the set of downward gradient flow lines from p to q.

If the metric g is generic, then $\mathcal{M}(p, q)$ is naturally a manifold of dimension

$$\dim \mathcal{M}(p, q) = \operatorname{ind}(p) - \operatorname{ind}(q). \tag{5.3}$$

To see why, for $s \in \mathbb{R}$ let $\psi_s : X \to X$ denote the time s flow of the vector field ∇f. The *descending manifold* of a critical point p is the set

$$\mathcal{D}(p) := \{x \in X \mid \lim_{s\to+\infty}\psi_s(x) = p\}.$$

Informally, this is the set of points in X that "can be reached by downward gradient flow starting at p". Similarly, the *ascending manifold* of a critical point q is

$$\mathcal{A}(q) := \{x \in X \mid \lim_{s\to-\infty}\psi_s(x) = q\}.$$

One can show that the descending manifold $\mathcal{D}(p)$ is a smoothly embedded open ball in X of dimension $\operatorname{ind}(p)$, and also $T_p\mathcal{D}(p)$ is the negative eigenspace of the Hessian (5.2). Likewise, $\mathcal{A}(q)$ is a smoothly embedded open ball of dimension $n - \operatorname{ind}(q)$.

It follows from the definitions that there is a bijection

$$\mathcal{M}(p, q) \xrightarrow{\simeq} \mathcal{D}(p) \cap \mathcal{A}(q)$$

sending a flow line γ to the point $\gamma(0) \in X$. Consequently, if $\mathcal{D}(p)$ is transverse to $\mathcal{A}(q)$, then equation (5.3) follows by dimension counting. The pair (f, g) is said to be *Morse-Smale* if $\mathcal{D}(p)$ is transverse to $\mathcal{A}(q)$ for every pair of critical points p, q. One can show that for a given Morse function f, for generic metrics g the pair

(f, g) is Morse-Smale. Henceforth we assume by default that this condition holds. Observe also that $\mathbb{R}$ acts on $\mathcal{M}(p, q)$ by precomposition with translations, and if $p \neq q$ then this action is free, so that

$$\dim(\mathcal{M}(p,q)/\mathbb{R}) = \operatorname{ind}(p) - \operatorname{ind}(q) - 1. \tag{5.4}$$

In particular, if $\operatorname{ind}(q) \geq \operatorname{ind}(p)$ then $\mathcal{M}(p, q)$ is empty (except when $p = q$), and if $\operatorname{ind}(q) = \operatorname{ind}(p) - 1$ then $\mathcal{M}(p, q)/\mathbb{R}$ is discrete.

We now define the *Morse complex* $C_*^{\mathrm{Morse}}(X, f, g)$ as follows. The chain module in degree i is the free $\mathbb{Z}$-module generated by the index i critical points:

$$C_i^{\mathrm{Morse}}(X, f, g) := \mathbb{Z}\{p \in X \mid df_p = 0,\ \operatorname{ind}(p) = i\}.$$

The differential $\partial : C_i^{\mathrm{Morse}}(X, f, g) \to C_{i-1}^{\mathrm{Morse}}(X, f, g)$ is defined by counting gradient flow lines as follows: if p is an index i critical point, then

$$\partial p := \sum_{\operatorname{ind}(p)-\operatorname{ind}(q)=1} \# \frac{\mathcal{M}(p,q)}{\mathbb{R}} \cdot q.$$

Here '#' denotes a signed count. We wil not say more about the signs here except to note that the signs are determined by choices of orientations of the descending manifolds of the critical points, but the chain complexes resulting from different sign choices are canonically isomorphic to each other.

To show that this chain complex is well-defined, one must prove that ∂ is well-defined, ie $\mathcal{M}(p, q)/\mathbb{R}$ is finite whenever $\operatorname{ind}(p)-\operatorname{ind}(q) = 1$, and one must also prove that $\partial^2 = 0$. The first step is to prove a compactness theorem which asserts that given critical points $p \neq q$, any sequence $\{\gamma_n\}_{n=1}^{\infty}$ in $\mathcal{M}(p, q)/\mathbb{R}$ has a subsequence which converges in an appropriate sense to a "k-times broken flow line" from p to q. This is a tuple $(\widehat{\gamma}_0, \ldots, \widehat{\gamma}_k)$ for some $k \geq 0$, such that there are critical points $p = r_0, r_1, \ldots, r_{k+1} = q$ for which $\widehat{\gamma}_i \in \mathcal{M}(r_i, r_{i+1})/\mathbb{R}$ is a nonconstant flow line.

If $\operatorname{ind}(p) - \operatorname{ind}(q) = 1$, then the Morse-Smale condition implies that there are no k-times broken flow lines with $k > 0$, so $\mathcal{M}(p, q)/\mathbb{R}$ is compact, and hence finite, so ∂ is well-defined.

If $\operatorname{ind}(p) - \operatorname{ind}(q) = 2$, then $\mathcal{M}(p, q)/\mathbb{R}$ is not necessarily compact and may contain a sequence converging to a once-broken flow line. However we can add in these broken flow lines to obtain a compactification $\overline{\mathcal{M}(p, q)/\mathbb{R}}$ of $\mathcal{M}(p, q)/\mathbb{R}$. it turns out that this is a compact oriented 1-manifold with boundary, whose boundary as an oriented manifold is

$$\partial\overline{\mathcal{M}(p,q)/\mathbb{R}} = \bigcup_{\operatorname{ind}(p)-\operatorname{ind}(r)=1} \frac{\mathcal{M}(p,r)}{\mathbb{R}} \times \frac{\mathcal{M}(r,q)}{\mathbb{R}}. \tag{5.5}$$

The main ingredient in the proof of this is a gluing theorem asserting that each broken flow line $(\widehat{\gamma}_0, \widehat{\gamma}_1)$ with $\widehat{\gamma}_0 \in \mathcal{M}(p, r)/\mathbb{R}$ and $\widehat{\gamma}_1 \in \mathcal{M}(r, q)/\mathbb{R}$ can be "patched" to an unbroken flow line in a unique end of $\mathcal{M}(p, q)/\mathbb{R}$. (One also has to show that $\mathcal{M}(p, q)/\mathbb{R}$ can be oriented so that the orientations on both sides of (5.5) agree.)

It follows from the boundary equation (5.5) that $\partial^2 = 0$. Namely, counting the points on both sides of (5.5) with signs gives

$$\#\partial\overline{\mathcal{M}(p,q)/\mathbb{R}} = \sum_{\operatorname{ind}(p)-\operatorname{ind}(r)=1} \langle \partial p, r\rangle \langle \partial r, q\rangle. \tag{5.6}$$

Here $\langle \partial p, r\rangle \in \mathbb{Z}$ denotes the coefficient of r in ∂p. Thus the right hand side of (5.6) is, by definition, the coefficient $\langle \partial^2 p, q\rangle$. On the other hand since a compact oriented 1-manifold has zero boundary points counted with signs, the left hand side of (5.6) is zero.

5.3. **Morse homology.** The *Morse homology* $H_*^{\text{Morse}}(X, f, g)$ is the homology of the above chain complex.

Example 5.1. Consider a Morse function $f : S^2 \to \mathbb{R}$ with two index 2 critical points x_1, x_2, one index 1 critical point y, and one index 0 critical point z. One can visualize f as the height function on a "heart-shaped" sphere embedded in $\mathbb{R}^3$. Pick any metric g on S^2; it turns out that (f, g) will automatically be Morse-Smale in this example. There is (up to reparametrization) a unique downward gradient flow line from each x_i to y. There are two gradient flow lines from y to z. The latter turn out to have opposite signs, and so for suitable orientation choices the Morse complex is given by

$$C_2^{\text{Morse}} = \mathbb{Z}\{x_1, x_2\}, \quad C_1^{\text{Morse}} = \mathbb{Z}\{y\}, \quad C_0^{\text{Morse}} = \mathbb{Z}\{z\},$$
$$\partial x_1 = y, \quad \partial x_2 = -y, \quad \partial y = 0.$$

Thus $H_2^{\text{Morse}} \simeq \mathbb{Z}$, generated by $x_1 + x_2$; $H_1^{\text{Morse}} = 0$; and $H_0^{\text{Morse}} = \mathbb{Z}$, generated by z.

The above example illustrates a fundamental theorem in the subject:

Theorem 5.2. *There is a canonical isomorphism between Morse homology and singular homology,*

$$H_*^{\text{Morse}}(X, f, g) \simeq H_*(X). \tag{5.7}$$

An immediate corollary is that there must be enough critical points to generate a chain complex whose homology is $H_*(X)$. In particular:

Corollary 5.3. *If f is a Morse function on a closed smooth manifold X, then* $c_i(f) \geq \operatorname{rank}(H_i(X))$.

5.4. **Continuation maps.** We will later construct analogues of Morse homology which generally do not have interpretations in terms of previously known invariants such as singular homology. What is most important here as a model for these later constructions is that the Morse homology $H_*^{\text{Morse}}(X, f, g)$ is a topological invariant of X which does not depend on f or g. One can prove this directly, without making the comparison with singular homology, as follows.

Let (f_0, g_0) and (f_1, g_1) be two Morse-Smale pairs. Let $\{(f_s, g_s) \mid s \in \mathbb{R}\}$ be a smooth family of pairs of functions and metrics on X such that $(f_s, g_s) = (f_0, g_0)$ for $s \leq 0$ and $(f_s, g_s) = (f_1, g_1)$ for $s \geq 1$. We do not (and in general cannot) assume that the pair (f_s, g_s) is Morse-Smale for all s. One now defines a map

$$\Phi : C_*^{\text{Morse}}(X, f_1, g_1) \longrightarrow C_*^{\text{Morse}}(X, f_0, g_0),$$

called the *continuation map*, as follows. If p_0 is an index i critical point of f_0 and p_1 is an index i critical point of f_1, then $\langle \Phi(p_1), p_0 \rangle$ is a signed count of maps $\gamma : \mathbb{R} \to X$ satisfying

$$\frac{d\gamma(s)}{ds} = \nabla f_s(\gamma(s))$$

and $\lim_{s \to -\infty} \gamma(s) = p_0$ and $\lim_{s \to +\infty} \gamma(s) = p_1$. Here the gradient of f_s is computed using the metric g_s. Similarly to the proof that ∂ is well defined and $\partial^2 = 0$, one can show that if the family of metrics $\{g_s\}$ is generic then Φ is a well-defined chain map. For example, if the family $\{(f_s, g_s)\}$ is constant then Φ is the identity map. One can also show that up to chain homotopy, Φ depends only on the homotopy class of the path $\{(f_t, g_t)\}$ rel endpoints, which since the space of pairs (f, g) is contractible means that up to chain homotopy Φ depends only on (f_0, g_0) and (f_1, g_1). Finally, related considerations show that if

$$\Phi' : C_*^{\mathrm{Morse}}(X, f_2, g_2) \longrightarrow C_*^{\mathrm{Morse}}(X, f_1, g_1)$$

is the continuation map induced by a generic path from (f_1, g_1) to (f_2, g_2), then the composition $\Phi\Phi'$ is chain homotopic to the continuation map induced by a path from (f_0, g_0) to (f_2, g_2). If $(f_2, g_2) = (f_0, g_0)$, then it follows that $\Phi\Phi'$ and $\Phi'\Phi$ are chain homotopic to the respective identity maps. Thus Φ induces an isomorphism on homology.

It follows from the above homotopy properties of continuation maps that the Morse homologies $H_*^{\mathrm{Morse}}(X, f, g)$ for different Morse-Smale pairs (f, g) are canonically isomorphic to each other via continuation maps. It turns out that the isomorphism (5.7) commutes with these continuation maps.

5.5. **Spectral flow.** We now recall a more analytical way to understand the dimension formula (5.4), which is used in the infinite dimensional variants of Morse theory to be considered later.

Given critical points p and q, let $\mathcal{P}$ denote the space of smooth paths $\gamma : \mathbb{R} \to X$ satisfying $\lim_{s\to+\infty} \gamma(s) = p$ and $\lim_{s\to-\infty} \gamma(s) = q$. An element $\gamma \in \mathcal{P}$ is a gradient flow line if and only if it satisfies the equation

$$\frac{d\gamma(s)}{ds} - \nabla f(\gamma(s)) = 0. \tag{5.8}$$

Note that the left hand side of (5.8) is a section of the pullback bundle γ^*TX over $\mathbb{R}$.

Now assume that γ is a flow line. To understand flow lines near γ, we consider the linearization of the equation (5.8). This could be regarded a linear operator D_γ which sends the tangent space to $\mathcal{P}$ at γ (namely the space of smooth sections of γ^*TX which converge to 0 as $s \to \pm\infty$) to γ^*TX. But in order to apply the tools of functional analysis one wants to work with suitable Banach space completions, for example to regard D_γ as an operator from the Sobolev space of L_1^2 sections of γ^*TX to the space of L^2 sections of γ^*TX. To write this operator more explicitly, choose a trivialization of γ^*TX which converges to some fixed trivializations of T_pX and T_qX as $s \to \pm\infty$. Then with respect to this trivialization, the operator D_γ has the form

$$D_\gamma = \partial_s + A_s : L_1^2(\mathbb{R}, \mathbb{R}^n) \longrightarrow L^2(\mathbb{R}, \mathbb{R}^n), \tag{5.9}$$

where A_s is an $n \times n$ matrix depending on $s \in \mathbb{R}$. Moreover as $s \to \pm\infty$, the matrix A_s converges to the (negative) Hessians,

$$\lim_{s\to+\infty} A_s = -H(f, p), \qquad \lim_{s\to-\infty} A_s = -H(f, q). \tag{5.10}$$

The first step in the analytic treatment is to show that if f is Morse then the operator D_γ is Fredholm; and if the metric g is generic, then for all flow lines γ the operator D_γ is surjective, and $\mathcal{M}(p, q)$ is a smooth manifold with $T_\gamma\mathcal{M}(p, q) = \operatorname{Ker}(D_\gamma)$. In particular $\dim \mathcal{M}(p, q) = \operatorname{ind}(D_\gamma)$.

To compute the index of D_γ, there is a general principle that if $\{A_s\}$ is a family of operators on a Hilbert space satisfying appropriate technical hypotheses, then the index of $\partial_s + A_s$ from L^2_1 to L^2 is the *spectral flow* of the family $\{A_s\}$, which roughly speaking is the number of eigenvalues of A_s that cross from negative to positive as s goes from $-\infty$ to $+\infty$, minus the number of eigenvalues that cross from positive to negative. For some theorems realizing this principle in different situations see eg [55, 38]. For the operator (5.9), no additional technical hypotheses are necessary and the spectral flow is simply the number of positive eigenvalues of $\lim_{s\to+\infty} A_s$ minus the number of positive eigenvalues of $\lim_{s\to-\infty} A_s$. Using (5.10) one obtains

$$\dim \mathcal{M}(p,q) = \operatorname{ind}(D_\gamma) = -\operatorname{ind}(p) - (-\operatorname{ind}(q)),$$

which recovers (5.4).

More generally we will need to apply the "index=spectral flow" principle to

$$D = \partial_s + A_s : L^2_1(\mathbb{R}\times Y, E) \longrightarrow L^2(Y,E)$$

where $\{A_s\}$ is a family of elliptic first-order differential operators on a vector bundle E over a manifold Y parametrized by $s \in \mathbb{R}$, which converge as $s \to \pm\infty$ to self-adjoint operators with zero kernel. The operator (5.9) corresponds to the case where Y is a point. Later in this article Y will be a circle (for cylindrical contact homology) or a three-manifold (for Seiberg-Witten Floer homology).

6. First attempt at a chain complex: cylindrical contact homology

Let Y be a closed oriented 3-manifold and let λ be a contact form on Y. We would like to define an analogue of the Morse complex on the loop space of Y, which is generated by Reeb orbits, and whose differential counts an appropriate notion of "flow lines" between them. Although the analogy with Morse homology breaks down somewhat, this idea leads naturally to the definition of cylindrical contact homology. This theory can be used to prove the Weinstein conjecture in many cases. Although cylindrical contact homology can be defined for contact manifolds of any odd dimension, for definiteness we stick to the three-dimensional case.

6.1. **Nondegenerate Reeb orbits.** We now explain the appropriate analogue of nondegenerate critical point in this context.

Let $\gamma : \mathbb{R}/T\mathbb{Z} \to Y$ be a Reeb orbit. Let $\psi_T : Y \to Y$ denote the diffeomorphism obtained by flowing along the Reeb vector field for time T. This preserves the contact form, because by the definition of Reeb vector field, the Lie derivative $\mathcal{L}_R\lambda = 0$. It follows that for any $t \in \mathbb{R}/T\mathbb{Z}$, we have a symplectic linear map

$$P_\gamma := d\psi_T : (\xi_{\gamma(t)}, d\lambda) \longrightarrow (\xi_{\gamma(t)}, d\lambda).$$

This map is called the *linearized return map*.

Another way to describe this map is as follows. Let D be a small embedded disk in Y centered at $\gamma(t)$ and transverse to γ, such that $T_{\gamma(t)}D = \xi_{\gamma(t)}$. For $x \in D$ close to the center, there is a unique point in D which is reached by following the Reeb flow for a time close to T. This gives a partially defined "return map" $\phi : D \to D$ which is defined near the origin. The derivative of this map at the origin is the linearized return map P_γ.

We say that the Reeb orbit γ is *nondegenerate* if P_γ does not have 1 as an eigenvalue. This condition does not depend on the choice of $t \in \mathbb{R}/T\mathbb{Z}$, because the linearized return maps for different t are conjugate to each other. If the Reeb orbit γ is nondegenerate then it is isolated, because Reeb orbits close to γ give rise to fixed points of the map ϕ, and the condition that $1 - d\phi$ is invertible at the origin implies that ϕ has no fixed points near the origin.

One can show that for a given contact structure ξ, for generic contact forms λ, all Reeb orbits are nondegenerate. We will always assume unless otherwise stated that all Reeb orbits are nondegenerate.

One can classify (nondegenerate) Reeb orbits into three types, according to the eigenvalues λ, λ^{-1} of the linearized return map:

- *elliptic*: $\lambda, \lambda^{-1} = e^{\pm 2\pi i\theta}$.
- *positive hyperbolic*: $\lambda, \lambda^{-1} > 0$.
- *negative hyperbolic*: $\lambda, \lambda^{-1} < 0$.

6.2. **Holomorphic cylinders.** The appropriate analogue of "gradient flow line" in this context is a certain kind of holomorphic cylinder in $\mathbb{R} \times Y$. We now explain what these are.

In general, recall that a *complex structure* on an even-dimensional real vector bundle $E \to X$ is a bundle map $J : E \to E$ satisfying $J^2 = -1$. An *almost complex structure* on an even-dimensional manifold X is a complex structure J on the tangent bundle TX. A *holomorphic curve*[3] in (X, J) is a map $u : \Sigma \to X$ where Σ is a surface with an almost complex structure j (ie a Riemann surface), and $J \circ du = du \circ j$. Two holomorphic curves $u : (\Sigma, j) \to X$ and $u' : (\Sigma', j') \to X$ are considered equivalent if there is a biholomorphic map $\phi : (\Sigma, j) \to (\Sigma', j')$ with $u = u' \circ \phi$. If u is an embedding then the equivalence class of u is determined by its image. That is, an embedded holomorphic curve in (X, J) is just a 2-dimensional submanifold $C \subset X$ such that $J(TC) = TC$.

Returning now to the situation of interest:

Definition 6.1. Let Y be a three-manifold with a contact form λ. An almost complex structure J on the 4-manifold $\mathbb{R} \times Y$ is *admissible* if:

1. J sends ξ to itself, rotating ξ positively with respect to the orientation of ξ given by $d\lambda$.
2. If s denotes the $\mathbb{R}$ coordinate on $\mathbb{R} \times Y$, then $J(\partial_s) = R$.
3. J is invariant under the $\mathbb{R}$ action on $\mathbb{R} \times Y$ that translates s.

Note that the space of such J is nonempty and contractible. Indeed, the choice of such a J is equivalent to the choice of a complex structure on the 2-plane bundle ξ over Y which rotates positively with respect to $d\lambda$. Fix an admissible almost complex structure J on Y below.

Observe that if γ is an embedded Reeb orbit, then $\mathbb{R} \times \gamma$ is an embedded holomorphic cylinder in $\mathbb{R} \times Y$. This follows from condition (2) above. More generally, we can study holomorphic curves in $\mathbb{R} \times Y$ that are asymptotic to such $\mathbb{R}$-invariant cylinders, or covers thereof, as the $\mathbb{R}$ coordinate goes to plus or minus infinity. To define what we mean by this, consider a "half-cylinder" $[0, \infty) \times S^1$ or $(-\infty, 0] \times S^1$ with coordinates s, t, with the almost complex structure j sending ∂_s to ∂_t. Let $\pi_{\mathbb{R}} : \mathbb{R} \times Y \to \mathbb{R}$ and $\pi_Y : \mathbb{R} \times Y \to Y$ denote the two projections. If $u : \Sigma \to \mathbb{R} \times Y$ is a holomorphic curve and if γ is a Reeb orbit (not necessarily embedded), we define a *positive end* of u at γ to be an end of Σ which can be parametrized as $[0, \infty) \times S^1$ with the almost complex structure j as above, such that $\lim_{s \to \infty} \pi_{\mathbb{R}}(s, \cdot) = \infty$, and

---

[3] Often these are instead called "pseudoholomorphic curves" or "J-holomorphic curves", in order to emphasize the fact that we are working with *almost* complex geometry, as opposed to complex manifolds.

$\lim_{s\to\infty}\pi_Y(s,\cdot)$ is a reparametrization of γ. Likewise, a *negative end* of u at γ is an end of Σ which can be parametrized as $(-\infty,0]\times S^1$, with the almost complex structure j as above, such that $\lim_{s\to-\infty}\pi_{\mathbb{R}}(s,\cdot)=-\infty$, and $\lim_{s\to-\infty}\pi_Y(s,\cdot)$ is a reparametrization of γ.

If γ_+ and γ_- are two Reeb orbits, define $\mathcal{M}(\gamma_+,\gamma_-)$ to be the set of holomorphic cylinders in $\mathbb{R}\times Y$ that have a positive end at γ_+ and a negative end at γ_-. It turns out that these holomorphic cylinders are the appropriate "gradient flow lines" from γ_+ to γ_-. Note that there is an $\mathbb{R}$ action on $\mathcal{M}(\gamma_+,\gamma_-)$ given by translating the $\mathbb{R}$ coordinate on the target[4] space $\mathbb{R}\times Y$. This action is free except on the $\mathbb{R}$-invariant cylinders $\mathbb{R}\times\gamma$ in $\mathcal{M}(\gamma,\gamma)$.

6.3. **The action functional.** Holomorphic cylinders in $\mathcal{M}(\gamma_+,\gamma_-)$ can be regarded as "gradient flow lines" of the *symplectic action* functional on the loop space of Y defined by

$$\mathcal{A}(\gamma):=\int_{S^1}\gamma^*\lambda \tag{6.1}$$

for $\gamma: S^1\to Y$. Without trying to make this analogy precise, let us just note the following essential lemma:

Lemma 6.2. *Suppose there exists a holomorphic cylinder $u\in\mathcal{M}(\gamma_+,\gamma_-)$. Then*

$$\mathcal{A}(\gamma_+)\ge\mathcal{A}(\gamma_-),$$

with equality if and only if $\gamma_+=\gamma_-$ and the image of u is an $\mathbb{R}$-invariant cylinder.

Proof. Let $u:\mathbb{R}\times S^1\to\mathbb{R}\times Y$ be a holomorphic cylinder in $\mathcal{M}(\gamma_+,\gamma_-)$. By Stokes' theorem,

$$\mathcal{A}(\gamma_+)-\mathcal{A}(\gamma_-)=\int_{\mathbb{R}\times S^1}u^*d\lambda.$$

(The integral on the right converges because of the asymptotics of u.) By condition (1) in the definition of admissible almost complex structure, $u^*d\lambda\ge 0$ at each point in $\mathbb{R}\times S^1$, with equality only where u is tangent to $\mathbb{R}$ cross the Reeb direction. ☐

Later the symplectic action will play a key role in Taubes's proof of the Weinstein conjecture.

6.4. **The chain complex.** We can now define an analogue of the Morse complex in this setting. We will give a "quick and dirty" definition to save space; for the more general context into which this definition fits, see [20].

To start, for reasons we will explain below, one must discard certain "bad" Reeb orbits for the construction to work:

Definition 6.3. A Reeb orbit γ is said to be *bad* if it is the k-fold iterate of a negative hyperbolic orbit with k even. Otherwise γ is said to be *good*.

Now fix $\Gamma\in H_1(Y)$. Define $CC(Y,\lambda,\Gamma)$ to be the free $\mathbb{Q}$-module generated by the good Reeb orbits γ representing the homology class Γ. One then defines a differential

$$\partial: CC(Y,\lambda,\Gamma)\longrightarrow CC(Y,\lambda,\Gamma)$$

---

[4] This is not to be confused with the $\mathbb{R}\times S^1$ action on the set of holomorphic maps $\mathbb{R}\times S^1\to\mathbb{R}\times Y$ given by compositions with translations of the domain, which we have already modded out by in our definition of holomorphic curve.

as follows. Fix a generic admissible almost complex structure J on $\mathbb{R}\times Y$. If γ_+ is a good Reeb orbit, then

$$\partial\gamma_+ := \sum_{\gamma_-} k_{\gamma_-} n(\gamma_+,\gamma_-)\gamma_-. \tag{6.2}$$

Here the sum is over good Reeb orbits γ_-, and k_γ denotes the unique positive integer such that γ is the k_γ-fold iterate of an embedded Reeb orbit. Meanwhile, $n(\gamma_+,\gamma_-)\in\mathbb{Q}$ is a signed count of holomorphic cylinders in $\mathcal{M}(\gamma_+,\gamma_-)/\mathbb{R}$ that live in zero-dimensional moduli spaces. Multiply covered cylinders are counted with weight ±1 divided by the covering multiplicity. (We will not explain the signs here.) The homology of this chain complex, when defined (see below), is called *cylindrical contact homology*, and we denote it by $CH(Y,\xi,\Gamma)$.

The following is a special case of a result to be proved in [11], see [15, §3.2] for the statement, asserting that a more general theory called "linearized contact homology" is well-defined.

Theorem 6.4. *Suppose there are no contractible Reeb orbits. Then ∂ is well-defined*[5]*, $\partial^2 = 0$, and the homology $CH(Y,\xi,\Gamma)$ depends only on Y, the contact structure ξ, and the homology class Γ, and not on the contact form λ or almost complex structure J.*

A few comments are in order. First, the factors of k_γ in (6.2) are needed to make $\partial^2=0$ work, because when one glues two (not multiply covered) holomorphic cylinders along a Reeb orbit γ which is the k-fold iterate of an embedded Reeb orbit, there are k different ways to glue, compare §5.2. This is also why bad Reeb orbits need to be discarded: it turns out that these k different gluings all have the same sign when γ is good, but have cancelling signs when γ is bad. Finally, the assumption that there are no contractible Reeb orbits ensures that the necessary compactness arguments go through, by ruling out bubbling off of holomorphic planes.

6.5. **The index.** Unlike Morse homology, cylindrical contact homology is not $\mathbb{Z}$-graded. Rather, it is relatively $\mathbb{Z}/d(2c_1(\xi))$-graded, where $d(2c_1(\xi))$ denotes the divisibility of $2c_1(\xi)$ in $H^2(Y;\mathbb{Z})$ mod torsion. This means that any two generators γ_+ and γ_- have a well-defined relative grading, which can be regarded as the grading difference between γ_+ and γ_-, and which is an element of $\mathbb{Z}/d(2c_1(\xi))$. In this sense the differential ∂ decreases the grading by 1. The reason why there is no absolute grading analogous to the Morse index is that the analogue of the Hessian in this setting has infinitely many negative and infinitely many positive eigenvalues. Nonetheless it still makes sense to define the relative grading of γ_+ and γ_- to be the expected dimension of the moduli space of holomorphic cylinders $\mathcal{M}(\gamma_+,\gamma_-)$ that represent some relative homology class Z. This is given by a certain spectral flow (see §5.5), which is computed by a topological formula which we will not state here. It is only defined modulo $d(2c_1(\xi))$, because if Z' is a different relative homology class of cylinder then the corresponding expected dimensions differ by

---

[5] The expert reader may worry that even for generic J, multiply covered holomorphic cylinders might have smaller index than the cylinders that they cover, leading to failure of the compactness needed to show that ∂ is defined. It turns out that this does not happen for holomorphic cylinders in the symplectization of a contact 3-manifold. However this is an issue in defining the continuation maps and chain homotopies needed to prove the invariance statement in Theorem 6.4, for which some abstract perturbations of the moduli spaces are needed.

$\langle 2c_1(\xi), Z - Z'\rangle$, where $Z - Z' \in H_2(Y)$ denotes the difference between the two relative homology classes.

There is also a canonical absolute $\mathbb{Z}/2$-grading: a Reeb orbit has odd grading if it is positive hyperbolic, and even grading if it is elliptic or negative hyperbolic. The differential ∂ also has degree -1 with respect to this $\mathbb{Z}/2$-grading.

6.6. **Examples.** (1) Consider the contact form λ_n on T^3 defined in (3.3). Recall that all Reeb orbits represent homology classes of the form $(a, b, 0) \in H_1(T^3)$ with $(a, b) \neq (0, 0)$. As a consequence, the cylindrical contact homology $CH_*(T^3, \xi_n, \Gamma)$ is nonzero only for Γ of this form. Fix such a class $\Gamma = (a, b, 0)$. All Reeb orbits γ in the homology class Γ have symplectic action

$$\mathcal{A}(\gamma) = 2\pi\sqrt{a^2 + b^2}.$$

So by Lemma 6.2, there are no non-$\mathbb{R}$-invariant holomorphic cylinders between them. Now the cylindrical contact homology is not yet defined because λ_n is a "Morse-Bott" contact form whose Reeb orbits are not isolated but rather appear in one-parameter families. But one can show, see [9], that one can perturb λ_n to a contact form λ_n' such that each of the n circles of Reeb orbits in the homology class Γ splits into two Reeb orbits, one elliptic and one positive hyperbolic; there are no other Reeb orbits in the class Γ, except possibly for some much longer Reeb orbits which can be disregarded in the computation using a direct limit argument; and the differential on $CC_*(Y, \lambda_n', \Gamma)$ vanishes for any choice of admissible almost complex structure J. (Each S^1 of Reeb orbits is perturbed using a Morse function $f : S^1 \to \mathbb{R}$ with two critical points which become the two Reeb orbits after perturbation. There are two holomorphic cylinders from the elliptic orbit to the hyperbolic orbit after perturbation, counting with opposite signs, corresponding to the Morse complex of f on S^1.) The conclusion is that the $\mathbb{Z}/2$-graded cylindrical contact homology is given by

$$CH_{\text{even}}(Y, \xi_n, (a, b, 0)) \simeq CH_{\text{odd}}(Y, \xi_n, (a, b, 0)) \simeq \mathbb{Q}^n.$$

It now follows from Theorem 6.4 that the different contact structures ξ_n are pairwise non-contactomorphic. Also, they all satisfy the Weinstein conjecture. Because for any contact form λ with $\xi_n = \operatorname{Ker}(\lambda)$, either there is no contractible Reeb orbit, in which case the cylindrical contact homology is well-defined and nonzero, or else there is a contractible Reeb orbit.

In fact one can do a little better. There is a generalization of cylindrical contact homology, called *linearized contact homology*, which can sometimes be defined even when there are contractible Reeb orbits, by adding certain correction terms to the cylindrical contact homology differential. Linearized contact homology can be used in this example to prove that for every contact form there is a Reeb orbit in the homology class $(a, b, 0)$.

(2) Let us compute the cylindrical contact homology of the irrational ellipsoid (S^3, λ) from §3.3. Here of course we must take $\Gamma = 0$. Strictly speaking Theorem 6.4 is not applicable here because all Reeb orbits are contractible, but it turns out that the cylindrical contact homology is still defined in this example because all Reeb orbits satisfy a certain Conley-Zehnder index condition which rules out troublesome holomorphic planes.

Denote the two embedded Reeb orbits by γ_1 and γ_2. These are elliptic. Let γ_i^k denote the k-fold iterate of γ_i, see §1.6. The chain complex $CC_*(S^3, \lambda, 0)$ has a relative $\mathbb{Z}$-grading, and because $\Gamma = 0$ there is in fact a canonical way to normalize it to an absolute $\mathbb{Z}$-grading. This grading is given as follows: there are positive irrational numbers ϕ_1 and ϕ_2 with $\phi_1\phi_2 = 1$ (in the notation of equation (3.2), ϕ_1 and ϕ_2 are a_1/a_2 and a_2/a_1) such that the grading of γ_i^k is

$$|\gamma_i^k| = 2\lfloor k(1+\phi_i)\rfloor. \tag{6.3}$$

It is an exercise to deduce from (6.3) that there is one generator of each positive even grading. Hence the differential vanishes for degree reasons, and

$$CH_*(S^3,\xi,0) \simeq \begin{cases} \mathbb{Q}, & * = 2,4,\ldots, \\ 0, & \text{otherwise.} \end{cases}$$

Here ξ denotes the contact structure determined by λ, which as mentioned in §3.3 is the unique tight contact structure on S^3. See [10] for some applications of contact homology to the Reeb dynamics of other contact forms determining this contact structure.

(3) As mentioned previously, Colin-Honda used linearized contact homology to prove the Weinstein conjecture in many cases. However it is not currently known whether linearized contact homology can be used to prove the Weinstein conjecture for all tight contact three-manifolds. (It turns out that in the overtwisted case linearized contact homology is never defined, but the failure of linearized contact homology to be defined implies the existence of a contractible Reeb orbit, reproducing Hofer's Theorem 3.6.) Taubes's proof of the Weinstein conjecture for all contact three-manifolds needs Seiberg-Witten theory.

7. The big picture surrounding Taubes's proof of the Weinstein conjecture

7.1. Seiberg-Witten invariants of four-manifolds.

The Seiberg-Witten invariants (and the conjecturally equivalent Ozsvath-Szabo invariants [51]) are the most powerful tool currently available for distinguishing smooth four-manifolds. To briefly outline what these are, let X be a closed oriented connected smooth four-manifold. If $b_2^+(X) > 1$, then the Seiberg-Witten invariant of X is, after certain orientation choices have been made, a function

$$SW : \operatorname{Spin}^c(X) \longrightarrow \mathbb{Z}.$$

Here $b_2^+(X)$ denotes the maximal dimension of a subspace of $H_2(X;\mathbb{R})$ on which the intersection pairing is positive definite. Also $\operatorname{Spin}^c(X)$ denotes the set of spin-c structures on X. This is an affine space over $H^2(X;\mathbb{Z})$, which we will say more about in §8.1. Given a spin-c structure $\mathfrak{s}$, the Seiberg-Witten invariant $SW(X,\mathfrak{s})$ is defined by appropriately counting solutions to the Seiberg-Witten equations on X. (We will not write down the Seiberg-Witten equations here, but we will see a three-dimensional version of them in §8.3, and a two-dimensional version in §11.1.) The Seiberg-Witten invariants depend only on the diffeomorphism type of X, and can distinguish many pairs of smooth four-manifolds that are homeomorphic but not diffeomorphic. A detailed definition of the Seiberg-Witten invariants of four-manifolds may be found in [48], and a review of the early results in this area is given in [16].

7.2. **Taubes's "SW=Gr" theorem.** Suppose now that our four-manifold X is symplectic. A major result of Taubes from the 1990's asserts that the Seiberg-Witten invariants of X are equivalent to a certain count of holomorphic curves.

To describe this result, let ω denote the symplectic form on X. It turns out that ω defines a bijection

$$\imath_\omega : \operatorname{Spin}^c(X) \xrightarrow{\simeq} H_2(X).$$

Now let J be a generic, ω-tame almost complex structure on X; the tameness condition means that $\omega(v, Jv) > 0$ for all nonzero tangent vectors v. Taubes then defines a "Gromov invariant"

$$Gr : H_2(X) \longrightarrow \mathbb{Z}$$

roughly as follows. For each $Z \in H_2(X)$, the integer $Gr(X, Z)$ is a count of certain J-holomorphic curves C in X representing the homology class Z. The curves C that are counted are required to be embedded, except that they may include multiple covers of tori of square zero. Such a curve C is not required to be connected, but each component of C must live in a zero-dimensional moduli space. Each such holomorphic curve C is counted with a certain integer weight. (The weight is ± 1, except when C includes multiply covered tori, in which case the weight is given by a somewhat complicated recipe.) Taubes's theorem is now:

Theorem 7.1 (Taubes)**.** *Let X be a closed connected symplectic 4-manifold with $b_2^+(X) > 1$. Then for each $\mathfrak{s} \in \operatorname{Spin}^c(X)$ we have*

$$SW(X, \mathfrak{s}) = Gr(X, \imath_\omega(\mathfrak{s})).$$

The proof of this theorem is given in [61]; for an introduction to it see [34], and for a discussion from a physics perspective see [74]. The basic idea is that one deforms the Seiberg-Witten equations using a large multiple of the symplectic form, and shows that solutions to the deformed Seiberg-Witten equations "concentrate along" holomorphic curves. Taubes's proof of the Weinstein conjecture involves a three-dimensional version of this.

7.3. **Three-dimensional Seiberg-Witten theory.** Let Y be a closed oriented 3-manifold. The set of spin-c structures on Y, denoted by $\operatorname{Spin}^c(Y)$, is again an affine space over $H^2(Y; \mathbb{Z})$. A spin-c structure on Y determines a product spin-c structure on $\mathbb{R} \times Y$. There are then various ways to define topological invariants of the 3-manifold Y by studying solutions to the Seiberg-Witten equations on the noncompact 4-manifold $\mathbb{R} \times Y$.

To start, one can consider $\mathbb{R}$-invariant solutions to the Seiberg-Witten equations on $\mathbb{R} \times Y$. These are equivalent to solutions to the three-dimensional Seiberg-Witten equations on Y, which we will discuss in §8.3. When $b_1(Y) > 0$, one can count these solutions with signs to obtain the three-dimensional *Seiberg-Witten invariant*[6] $SW : \operatorname{Spin}^c(Y) \to \mathbb{Z}$. As mostly[7] shown in [17, 46] and fully shown in [69], this invariant agrees with the Turaev torsion of Y, which generalizes the Alexander polynomial of a knot, and which is explicitly computable in terms of the determinants of certain matrices of polynomials associated to a triangulation of Y.

To get more interesting invariants of Y, one observes that solutions to the Seiberg-Witten equations on Y are critical points of a certain functional $\mathcal{F}$ on a configuration space associated to Y. Moreover, solutions to the four-dimensional Seiberg-Witten equations on $\mathbb{R} \times Y$ (not necessarily $\mathbb{R}$-invariant) are equivalent to gradient flow lines of this functional. It turns out that one can then define a version of Morse homology for the functional $\mathcal{F}$, called *Seiberg-Witten Floer homology*, which we will say more about in §9. The analytical details of this construction are highly nontrivial and have been carried out by Kronheimer-Mrowka [38]. Roughly speaking, the Seiberg-Witten Floer homology is the homology of a chain complex which is generated by solutions to the Seiberg-Witten equations on Y, and whose differential counts solutions to the Seiberg-Witten equations on $\mathbb{R} \times Y$. When $b_1(Y) > 0$, for non-torsion spin-c structures, the Euler characteristic of the Seiberg-Witten Floer homology agrees with the Seiberg-Witten invariant discussed above. However Seiberg-Witten Floer homology can also be defined for torsion spin-c structures and without any assumption on $b_1(Y)$.

---

[6] When $b_1(Y) = 1$, for "torsion" spin-c structures, ie spin-c structures whose first Chern class is torsion, see §8.1, this invariant depends on the choice of one of two possible "chambers".

[7] ie after identifying spin-c structures that differ by torsion elements in $H^2(Y; \mathbb{Z})$

In fact there are two basic versions of Seiberg-Witten Floer theory that one can define, depending on how one treats "reducibles", see §8.4. The first is denoted by $\check{HM}_*(Y)$, and pronounced "HM-to"; this assigns a relatively graded homology group $\check{HM}_*(Y, \mathfrak{s})$ to each spin-c structure $\mathfrak{s}$ on Y, which is conjecturally isomorphic to the Heegaard Floer homology $HF^+_*(Y, \mathfrak{s})$ defined in [50]. The second is pronounced "HM-from", denoted by $\widehat{HM}_*(Y, \mathfrak{s})$, and conjecturally isomorphic to the Heegaard Floer homology $HF^-_*(Y, \mathfrak{s})$. For non-torsion spin-c structures, there are no reducibles and $\check{HM}_*$ and $\widehat{HM}_*$ are the same. For any spin-c structure $\mathfrak{s}$, there is a canonical isomorphism

$$\check{HM}_*(-Y, \mathfrak{s}) = \widehat{HM}^{-*}(Y, \mathfrak{s}),$$

where $-Y$ denotes Y with its orientation reversed, and $\widehat{HM}^*$ denotes the cohomological version of $\widehat{HM}_*$ obtained by dualizing the chain complex.

7.4. **Embedded contact homology.** Suppose now that our three-manifold Y is equipped with a contact form λ. Recall from §1.4 that the four-manifold $\mathbb{R} \times Y$ then has a symplectic form $d(e^s \lambda)$, where s denotes the $\mathbb{R}$ coordinate. It is natural to seek an analogue of Taubes's "SW=Gr" theorem for the noncompact symplectic four-manifold $\mathbb{R} \times Y$. That is, one would like to understand the Seiberg-Witten Floer homology of Y in terms of holomorphic curves in $\mathbb{R} \times Y$. For this purpose it is appropriate to use an admissible almost complex structure as in Definition 6.1.

The analogy suggests that the Seiberg-Witten Floer homology of Y should be isomorphic to the homology of a chain complex whose differential counts certain (mostly) embedded holomorphic curves in $\mathbb{R} \times Y$, and which is generated by certain $\mathbb{R}$-invariant holomorphic curves in $\mathbb{R} \times Y$, that is to say unions of Reeb orbits. The resulting theory is called *embedded contact homology*, or ECH for short. There is some resemblance between ECH and cylindrical contact homology; but among other differences, ECH does not require the holomorphic curves that are counted to be cylinders, while cylindrical contact homology does not require them to be embedded.

To say a bit more about what ECH is, assume as usual that all Reeb orbits are nondegenerate.

Definition 7.2. An *orbit set* is a finite set of pairs $\alpha = \{(\alpha_i, m_i)\}$ where:

- The α_i's are distinct embedded Reeb orbits.
- The m_i's are positive integers.

The homology class of the orbit set α is defined by

$$[\alpha] := \sum_i m_i[\alpha_i] \in H_1(Y).$$

The orbit set α is *admissible* if $m_i = 1$ whenever α_i is hyperbolic (see §6.1).

Given $\Gamma \in H_1(Y)$, the embedded contact homology $ECH_*(Y, \lambda, \Gamma)$ is the homology of a chain complex which is freely generated over $\mathbb{Z}$ by admissible orbit sets α with $[\alpha] = \Gamma$. The differential counts certain (mostly) embedded holomorphic curves in $\mathbb{R} \times Y$. It has a relative $\mathbb{Z}/d(c_1(\xi) + 2\,\mathrm{PD}(\Gamma))$ grading, where 'd' denotes divisibility in $H^2(Y;\mathbb{Z})$ mod torsion. For the full definition of this theory, see [33] for an overview and [32, 35] for more details.

Note that the empty set of Reeb orbits is a legitimate generator[8] of the ECH chain complex with $\Gamma = 0$.

Example 7.3. Consider again the irrational ellipsoid (S^3, λ) as discussed in §6.6. Of course we must take $\Gamma = 0$. The generators of the ECH chain complex have the form $\gamma_1^{m_1}\gamma_2^{m_2}$ where γ_1 and γ_2 are the two embedded Reeb orbits, m_1 and m_2 are nonnegative integers, and $\gamma_1^{m_1}\gamma_2^{m_2}$ is shorthand for the orbit set consisting of the pair (γ_1, m_1) (when $m_1 \neq 0$) together with the pair (γ_2, m_2) (when $m_2 \neq 0$). In this case the chain complex has a relative $\mathbb{Z}$-grading, and since $\Gamma = 0$ this has a canonical refinement to an absolute $\mathbb{Z}$-grading such that the grading of the empty set is zero. It is shown in [36] that the grading is given by

$$|\gamma_1^{m_1}\gamma_2^{m_2}| = 2\left(m_1 + m_2 + m_1m_2 + \sum_{i=1}^{2}\sum_{k=1}^{m_i}\lfloor k\phi_i\rfloor\right)$$

where ϕ_1 and ϕ_2 are as in §6.6. One can deduce from this formula, see [36], that there is one generator of each nonnegative even grading, so that

$$ECH_*(S^3, \lambda, 0) \simeq \begin{cases} \mathbb{Z}, & * = 0, 2, \ldots, \\ 0, & \text{otherwise} \end{cases}$$

In general it turns out that ECH depends only on Y, the contact structure ξ, and the homology class Γ, and not on the choice of contact form λ or admissible almost complex structure J. This follows from a much stronger result recently proved by Taubes [66, 67], which is the analogue of Gr=SW in this setting:

Theorem 7.4 (Taubes). *Let Y be a closed oriented 3-manifold with a contact form λ such that all Reeb orbits are nondegenerate. Then for each $\Gamma \in H_1(Y)$ there is an isomorphism*

$$ECH_*(Y, \lambda, \Gamma) \simeq \widehat{HM}^{-*}(Y, \mathfrak{s}_\xi + \mathrm{PD}(\Gamma)), \tag{7.1}$$

up to a grading shift[9].

---

[8] It is always a cycle in the ECH chain complex, by the argument in Lemma 6.2. Its homology class in $ECH_*(Y, \lambda, 0)$ agrees with the invariant of contact structures in Seiberg-Witten Floer homology under the isomorphism (7.1), and conjecturally also with the contact invariant in Heegaard Floer homology.

[9] In fact both sides of (7.1) have canonical absolute gradings by homotopy classes of oriented 2-plane fields, see [38, 32], and it is natural to conjecture that the isomorphism (7.1) respects these gradings. It is further shown in [68] that the isomorphism (7.1) respects some additional structures on embedded contact homology and Seiberg-Witten Floer cohomology, namely the aforementioned contact invariants, the "U maps", the actions of $H_1(Y)$ mod torsion, and the twisted versions.

Here ξ denotes the contact structure determined by λ, and $\mathfrak{s}_\xi$ is a spin-c structure determined by ξ, see Example 8.2 below.

7.5. **Significance for the Weinstein conjecture.** To prove the Weinstein conjecture, it is enough to show that the embedded contact homology is nontrivial. More precisely, if (Y, λ) were a counterexample to the Weinstein conjecture, then one would have

$$ECH(Y, \lambda, \Gamma) \simeq \begin{cases} \mathbb{Z}, & \Gamma = 0, \\ 0, & \Gamma \neq 0. \end{cases} \tag{7.2}$$

Here the $\mathbb{Z}$ corresponds to the empty set of Reeb orbits. However, by Theorem 7.4, the ECH cannot be trivial as in (7.2), because the Seiberg-Witten Floer homology is always infinitely generated:

Theorem 7.5. (Kronheimer-Mrowka [38, §35.1]) *Let Y be a closed oriented 3-manifold and let $\mathfrak{s}$ be a spin-c structure with $c_1(\mathfrak{s})$ torsion. Then $\widehat{HM}^*(Y, \mathfrak{s})$ is nonzero for infinitely many values of the grading $*$, which are bounded from above.*

Here $c_1(\mathfrak{s})$ denotes the first Chern class of the spin-c structure, which is defined in §8.1 below. In terms of the correspondence between $\mathrm{Spin}^c(Y)$ and $H_1(Y)$ in (7.1), one has

$$c_1(\mathfrak{s}_\xi + \mathrm{PD}(\Gamma)) = c_1(\xi) + 2\,\mathrm{PD}(\Gamma).$$

Since TY is trivial, one can always find a spin-c structure $\mathfrak{s}$ such that $c_1(\mathfrak{s}) = 0$, and in particular $c_1(\mathfrak{s})$ is torsion. Thus Theorems 7.4 and 7.5 imply the following version of the Weinstein conjecture:

- *Let Y be a closed oriented 3-manifold with a contact form λ* such that all Reeb orbits are nondegenerate. *Let $\Gamma \in H_1(Y)$ such that $c_1(\xi) + 2\,\mathrm{PD}(\Gamma) \in H^2(Y; \mathbb{Z})$ is torsion. (Such Γ always exist.) Then there is a nonempty* admissible *orbit set α with $[\alpha] = \Gamma$.*

Note that the fact that ECH is infinitely generated does not imply that there are infinitely many embedded Reeb orbits, as shown by the irrational ellipsoid in Example 7.3.

In fact one does not need the full force of the isomorphism in Theorem 7.4 to prove the Weinstein conjecture; rather one just needs a way of passing from generators of Seiberg-Witten Floer homology to ECH generators. This is what Taubes's original proof of the Weinstein conjecture in [62] establishes, yielding a proof of the following theorem, which is slightly different than the statement above:

Theorem 7.6 (Taubes). *Let Y be a closed oriented 3-manifold with a contact form λ. Let $\Gamma \in H_1(Y)$ such that $c_1(\xi) + 2\,\mathrm{PD}(\Gamma) \in H^2(Y; \mathbb{Z})$ is torsion. (Such Γ always exist.) Then there is a nonempty orbit set α with $[\alpha] = \Gamma$.*

It is interesting to compare Theorem 7.6 with Theorems 3.6 and 3.7, which for certain contact structures produce a nonempty orbit set α with $[\alpha] = 0$. In a sequel [63] to the paper proving the Weinstein conjecture, Taubes uses more nontriviality results for Seiberg-Witten Floer homology to find nonempty orbit sets in some other homology classes.

Our goal in §8-11 below is to explain Taubes's proof of Theorem 7.6.

8. THE THREE-DIMENSIONAL SEIBERG-WITTEN EQUATIONS

To proceed further, we now need to recall the three-dimensional Seiberg-Witten equations.

8.1. Spin-c structures. Let Y be a closed oriented connected 3-manifold, and choose a Riemannian metric on Y.

Definition 8.1. A *spin-c structure* on Y is a pair $\mathfrak{s} = (\mathbb{S}, \rho)$ where $\mathbb{S}$ is a rank 2 Hermitian vector bundle on Y, called the *spinor bundle* (a section ψ of $\mathbb{S}$ is often called a *spinor*), and

$$\rho : TY \longrightarrow \operatorname{End}(\mathbb{S})$$

is a bundle map, called *Clifford multiplication*, such that:

(1) If $a, b \in T_yY$, then

$$\rho(a)\rho(b) + \rho(b)\rho(a) = -2\langle a, b\rangle.$$

(2) If e_1, e_2, e_3 is an oriented orthonormal frame for T_yY, then

$$\rho(e_1)\rho(e_2)\rho(e_3) = 1.$$

Properties (1) and (2) of Clifford multiplication are equivalent to the following: For each $y \in Y$, there is an oriented orthonormal frame e_1, e_2, e_3 for T_yY, and a basis for $\mathbb{S}_y$, in which Clifford multiplication is given by the Pauli matrices

$$\rho(e_1) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \qquad \rho(e_2) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad \rho(e_3) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{8.1}$$

Example 8.2. An oriented 2-plane field $\xi \subset TY$ determines a spin-c structure $\mathfrak{s}_\xi$ as follows. The spinor bundle is defined by

$$\mathbb{S} := \underline{\mathbb{C}} \oplus \xi, \tag{8.2}$$

where $\underline{\mathbb{C}}$ denotes the trivial complex line bundle, and ξ is regarded as a Hermitian line bundle using its orientation and the Riemannian metric on Y. Clifford multiplication is defined at a point $y \in Y$ by the equations (8.1), where e_1, e_2, e_3 are an oriented orthonormal frame for T_yY such that e_2, e_3 is an oriented orthonormal basis for ξ_y, and the basis for $\mathbb{S}_y$ is given in terms of the decomposition (8.2) by $(1, e_2)$.

Two spin-c structures $(\mathbb{S}, \rho)$ and $(\mathbb{S}', \rho')$ are isomorphic if there is a Hermitian vector bundle isomorphism $\phi : \mathbb{S} \xrightarrow{\simeq} \mathbb{S}'$ such that $\rho'(v) \circ \phi = \phi \circ \rho(v)$ for every tangent vector v. Let $\operatorname{Spin}^c(Y)$ denote the set of isomorphism classes of spin-c structures on Y. This does not depend on the choice of Riemannian metric on Y. There is an action of $H^2(Y; \mathbb{Z})$ on $\operatorname{Spin}^c(Y)$ defined as follows: Given $\alpha \in H^2(Y; \mathbb{Z})$, let L_α denote the complex line bundle on Y with $c_1(L_\alpha) = \alpha$, assign it a Hermitian metric, and define

$$\alpha \cdot (\mathbb{S}, \rho) := (\mathbb{S} \otimes L_\alpha, \rho \otimes 1).$$

It turns out that this action is free and transitive, so that $\operatorname{Spin}^c(Y)$ is an affine space over $H^2(Y; \mathbb{Z})$.

Remark 8.3. If $\mathcal{P}(Y)$ denotes the set of homotopy classes of oriented 2-plane fields on Y, then the map $\mathcal{P}(Y) \to \operatorname{Spin}^c(Y)$ defined in Example 8.2 is surjective. Two oriented 2-plane fields give rise to isomorphic spin-c structures if and only if they are homotopic over the 2-skeleton of Y (for some triangulation), compare Remark 3.3.

Here is an alternate definition of a spin-c structure, on an oriented manifold Y of any dimension $n > 1$ with a Riemannian metric. Let $F \to Y$ denote the *frame bundle*, whose fiber over $y \in Y$ is the set of orientation-preserving linear isometries $\mathbb{R}^n \overset{\simeq}{\to} Y_y$. Note that F is a principal $\mathrm{SO}(n)$-bundle over Y, where $\mathrm{SO}(n)$ acts on the right on F by precomposition with automorphisms of $\mathbb{R}^n$. Now the Lie group $\mathrm{Spin}^c(n)$ is defined by

$$\mathrm{Spin}^c(n) := \mathrm{Spin}(n) \times_{\mathbb{Z}/2} \mathrm{U}(1). \tag{8.3}$$

Here $\mathrm{Spin}(n)$ is the connected double cover of $\mathrm{SO}(n)$; and $\mathbb{Z}/2$ acts on $\mathrm{Spin}(n)$ as the nontrivial covering transformation, and on $\mathrm{U}(1)$ as multiplication by -1. A spin-c structure on Y is then defined to be a lift of F to a principal $\mathrm{Spin}^c(n)$-bundle, ie a principal $\mathrm{Spin}^c(n)$-bundle $\widetilde{F} \to Y$ together with a map $\widetilde{F} \to F$ which commutes with the group actions and the projections to Y.

When $n = 3$, this definition is equivalent to the previous one. In particular, given a lift $\widetilde{F}$ of the frame bundle, the spinor bundle and Clifford multiplication are recovered as follows. We can identify $\mathrm{Spin}(3) = \mathrm{SU}(2)$ and $\mathrm{Spin}^c(3) = \mathrm{U}(2)$. The spinor bundle is then associated to $\widetilde{F}$ via the fundamental representation of $\mathrm{U}(2)$. On the other hand the tangent bundle of Y is associated to $\widetilde{F}$ by the representation of $\mathrm{U}(2)$ on $\mathbb{R}^3$ given by the projection $\mathrm{Spin}^c(3) \to \mathrm{SO}(3)$. Clifford multiplication is then defined on these associated bundles using a model linear map $\mathbb{R}^3 \to \mathrm{End}(\mathbb{C}^2)$ that sends the standard basis vectors of $\mathbb{R}^3$ to the three Pauli matrices (8.1).

8.2. **The Dirac operator.**

Let $\mathfrak{s} = (\mathbb{S}, \rho)$ be a spin-c structure.

Definition 8.4. A *spin-c connection* on $\mathfrak{s}$ is a Hermitian connection on $\mathbb{S}$ such that the associated covariant deriviative ∇_A is compatible with Clifford multiplication in the following sense: If v is a section of TY and if ψ is a section of $\mathbb{S}$, then

$$\nabla_A(\rho(v)\psi) = \rho(\nabla v)\psi + \rho(v)\nabla_A\psi.$$

Here ∇v denotes the covariant derivative of v with respect to the Levi-Civita connection on TY.

It follows from the above definition that any two spin-c connections differ by an imaginary-valued 1-form on Y. It is not hard to show that spin-c connections exist. In fact a spin-c connection is equivalent to a Hermitian connection on the $\mathrm{U}(1)$-bundle $\det(\mathbb{S})$. One can see this by using the second definition of spin-c structure and noting from (8.3) that the Lie algebra of $\mathrm{Spin}^c(3) = \mathrm{U}(2)$ is the sum of the Lie algebras of $\mathrm{SO}(3)$ and of $\mathrm{U}(1)$. A connection on $\mathbb{S}$ is then determined by a connection on TY (which we take to be the Levi-Civita connection) and a connection on the complex line bundle associated to the determinant map $\mathrm{U}(2) \to \mathrm{U}(1)$, namely $\det(\mathbb{S})$. So in the notation ∇_A, we regard A as a connection on $\det(\mathbb{S})$. From this perspective, adding an imaginary-valued 1-form a to the connection A on $\det(\mathbb{S})$ adds $a/2$ to the associated spin-c connection ∇_A on $\mathbb{S}$.

Definition 8.5. Given a connection A on $\det(\mathbb{S})$, define the *Dirac operator* D_A to be the composition

$$C^\infty(Y, \mathbb{S}) \xrightarrow{\nabla_A} C^\infty(Y, T^*X \otimes \mathbb{S}) \xrightarrow{\rho} C^\infty(Y, \mathbb{S}).$$

Here the Clifford action is extended to cotangent vectors using the metric.

A key property of the Dirac operator is that its square is the "connection Laplacian", plus some zeroth order terms involving curvature. More precisely, it satisfies the Bochner-Lichnerowitz-Weitzenbock formula

$$D_A^2\psi = \nabla_A^*\nabla_A\psi + \frac{s}{4}\psi - \frac{1}{2}\rho(*F_A)\psi. \tag{8.4}$$

Here s denotes the scalar curvature of Y, which is a real-valued function; F_A denotes the curvature of the connection A on $\det(\mathbb{S})$, which is an imaginary-valued closed 2-form on Y; and $*$ denotes the Hodge star.

Another important property is that the Dirac operator is formally self-adjoint: if ψ_1 and ψ_2 are two spinors, then

$$\int_Y \langle D_A\psi_1, \psi_2\rangle d\operatorname{vol} = \int_Y \langle \psi_1, D_A\psi_2\rangle d\operatorname{vol}.$$

For much more about Dirac operators, see eg [6, 42].

8.3. The Seiberg-Witten equations.

Fix a spin-c structure $(\mathbb{S},\rho)$. The Seiberg-Witten equations concern a pair (A,ψ) where A is a connection on $\det(\mathbb{S})$ and ψ is a section of $\mathbb{S}$.

Definition 8.6. The (unperturbed) *Seiberg-Witten equations* for the pair (A,ψ) are

$$\begin{aligned} D_A\psi &= 0, \\ *F_A &= \langle \rho(\cdot)\psi, \psi\rangle. \end{aligned}$$

Note that it follows from the properties of Clifford multiplication that if ψ is any spinor then $\langle \rho(\cdot)\psi,\psi\rangle$ is an imaginary-valued 1-form.

Remark 8.7. Conventions for the Seiberg-Witten equations vary in the literature (and sometimes are not completely explicit). For example, one could multiply one side of the second equation by a positive constant, and the solutions to the equations would be equivalent via rescaling the spinor. However the sign is crucial: switching the sign in the second equation would ruin certain a priori estimates on the solutions, such as those in Lemma 11.2 below, which play a key role in the subject.

One also needs to consider certain perturbations of the equations. In particular, one often needs to make small perturbations in order to obtain transversality, while Taubes's proof of the Weinstein conjecture will involve a large perturbation.

Definition 8.8. Let μ be a real closed 2-form on Y. The *Seiberg-Witten equations with perturbation* μ for the pair (A,ψ) are

$$\begin{aligned} D_A\psi &= 0, \\ *F_A &= \langle \rho(\cdot)\psi, \psi\rangle + i*\mu. \end{aligned} \tag{8.5}$$

8.4. Gauge transformations.

The equations (8.5) have a large amount of symmetry. Namely, the Seiberg-Witten equations are defined on the *configuration space*

$$\mathcal{C} := \operatorname{Conn}(\det(\mathbb{S})) \times C^\infty(Y,\mathbb{S}),$$

where $\operatorname{Conn}(\det(\mathbb{S}))$ denotes the set of Hermitian connections on $\det(\mathbb{S})$. Define the *gauge group*

$$\mathcal{G} := C^\infty(Y, S^1).$$

This can be regarded as the automorphism group of the spin-c structure $(\mathbb{S},\rho)$. As such it has a natural action on the configuration space $\mathcal{C}$ defined as follows: If $g: Y \to S^1$ is in $\mathcal{G}$, then regarding S^1 as the unit circle in $\mathbb{C}$, one defines

$$g\cdot(A,\psi) := (A - 2g^{-1}dg, g\psi).$$

Lemma 8.9. *The set of solutions to the Seiberg-Witten equations* (8.5) *is invariant under the action of the gauge group* $\mathcal{G}$.

Proof. The curvature equation is invariant because the curvature of a connection is invariant under gauge transformations. The Dirac equation is invariant because

$$\begin{aligned} D_{A-2g^{-1}dg}(g\psi) &= \rho((\nabla_A - g^{-1}dg)g\psi) \\ &= \rho(dg\otimes\psi + g\nabla_A\psi - g^{-1}dg\otimes g\psi) \\ &= \rho(g\nabla_A\psi) \\ &= gD_A\psi. \end{aligned} \tag{8.6}$$

□

Two solutions to the Seiberg-Witten equations are called *gauge equivalent* if they differ by the action of an element of $\mathcal{G}$. In general, one studies solutions only modulo gauge equivalence.

Observe that the action of $\mathcal{G}$ on $\mathcal{C}$ is free, except that configurations $(A,0)$ have S^1 stabilizer given by the constant maps $Y\to S^1$. To keep track of this, a configuration (A,ψ) is called *reducible* if $\psi\equiv 0$, and *irreducible* otherwise.

Lemma 8.10.
(1) *Reducible solutions exist if and only if* $[\mu] = -2\pi c_1(\mathfrak{s})$ *in* $H^2(Y;\mathbb{R})$.
(2) *In this case the set of reducible solutions modulo gauge equivalence can be identified with the torus* $H^1(Y;\mathbb{R})/2\pi H^1(Y;\mathbb{Z})$.

Proof. Part (1) holds because if A is any connection on $\det(\mathbb{S})$ then the curvature F_A is a closed 2-form representing the cohomology class $-2\pi i c_1(\mathfrak{s})$.

To prove part (2), note that any two connections with the same curvature differ by an imaginary-valued closed 1-form. Thus the set of reducible solutions modulo gauge equivalence is an affine space over the set of imaginary closed 1-forms modulo $\{g^{-1}dg \mid g: Y\to S^1\subset\mathbb{C}\}$. To understand the latter subspace, note that any exact form idf can be written as $g^{-1}dg$ where $g = e^{if}$. On the other hand the set of homotopy classes of maps $Y\to S^1$ can be identified with $H^1(Y;\mathbb{Z})$, and the image of the homotopy class of $g: Y\to S^1$ under the map $H^1(Y;\mathbb{Z})\to H^1(Y;\mathbb{R})$ is the cohomology class of the 1-form $(2\pi i)^{-1}g^{-1}dg$. The claim follows. □

9. SEIBERG-WITTEN FLOER HOMOLOGY

We now briefly review what we need to know about Seiberg-Witten Floer homology, from [38].

9.1. **The Chern-Simons-Dirac functional.** We begin by realizing the solutions to the Seiberg-Witten equations as the critical points of a functional. The Seiberg-Witten Floer theory will then be some kind of Morse homology for this functional.

Let Y be a closed connected oriented 3-manifold with a Riemannian metric and a spin-c structure $\mathfrak{s} = (\mathbb{S},\rho)$. Fix a real closed 2-form μ for use in defining the perturbed Seiberg-Witten equations (8.5). Also fix a reference connection A_0 on

$\det(\mathbb{S})$, so that if A is any other connection on $\det(\mathbb{S})$ then $A - A_0$ is an imaginary-valued 1-form on Y.

Definition 9.1. Define the (perturbed) *Chern-Simons-Dirac functional*

$$\mathcal{F} : \mathcal{C} \to \mathbb{R}$$

by

$$\mathcal{F}(A,\psi) := -\frac{1}{8}\int_Y (A - A_0)\wedge(F_A + F_{A_0} - 2i\mu) + \frac{1}{2}\int_Y \langle D_A\psi,\psi\rangle d\,\mathrm{vol}\,. \tag{9.1}$$

Lemma 9.2. *(A,ψ) is a critical point of $\mathcal{F}$ if and only if (A,ψ) satisfies the perturbed Seiberg-Witten equations* (8.5).

Proof. Let $(\dot{A},\dot{\psi})$ be a tangent vector to the configuration space $\mathcal{C}$ at (A,ψ). This means that $\dot{A}$ is an imaginary-valued 1-form and $\dot{\psi}$ is a spinor. We compute

$$\begin{aligned} d\mathcal{F}_{(A,\psi)}(\dot{A},\dot{\psi}) = & -\frac{1}{8}\int_Y \dot{A}\wedge(F_A + F_{A_0} - 2i\mu) - \frac{1}{8}\int_Y (A - A_0)\wedge d\dot{A} \\ & + \frac{1}{4}\int_Y \langle\rho(\dot{A})\psi,\psi\rangle d\,\mathrm{vol} + \frac{1}{2}\int_Y \langle D_A\dot{\psi},\psi\rangle d\,\mathrm{vol} + \frac{1}{2}\int_Y \langle D_A\psi,\dot{\psi}\rangle d\,\mathrm{vol}\,. \end{aligned}$$

Applying Stokes' theorem to the second term, using the properties (8.1) of Clifford multiplication to manipulate the third term, and applying self-adjointness of the Dirac operator to the fourth term, we obtain

$$d\mathcal{F}_{(A,\psi)}(\dot{A},\dot{\psi}) = -\frac{1}{4}\int_Y \dot{A}\wedge(F_A - i\mu - *\langle\rho(\cdot)\psi,\psi\rangle) + \int_Y \mathrm{Re}\langle D_A\psi,\dot{\psi}\rangle d\,\mathrm{vol}\,.$$

This vanishes for all $\dot{A}$ and $\dot{\psi}$ if and only if (A,ψ) satisfy the perturbed Seiberg-Witten equations (8.5). □

We now consider the behavior of $\mathcal{F}$ under gauge transformations. Recall that the set of homotopy classes of maps $Y \to S^1$ can be identified with $H^1(Y;\mathbb{Z})$, and denote the homotopy class of a map g by $[g]$.

Lemma 9.3. *If $g : Y \to S^1$ is a gauge transformation, then*

$$\mathcal{F}(g\cdot(A,\psi)) - \mathcal{F}(A,\psi) = \pi\int_{[Y]} [g]\smile(2\pi c_1(\mathfrak{s}) + [\mu]).$$

Proof. We compute, using (8.6), that

$$\begin{aligned} \mathcal{F}(g\cdot(A,\psi)) - \mathcal{F}(A,\psi) = & -\frac{1}{8}\int_Y (-2g^{-1}dg)\wedge(F_A + F_{A_0} - 2i\mu) \\ & + \frac{1}{2}\int_Y \langle gD_A\psi, g\psi\rangle d\,\mathrm{vol} - \frac{1}{2}\int_Y \langle D_A\psi,\psi\rangle d\,\mathrm{vol}\,. \end{aligned}$$

The second line vanishes. To process the first line, recall that $g^{-1}dg$ represents the class $2\pi i[g]$, while F_A and F_{A_0} both represent the class $-2\pi i c_1(\mathfrak{s})$. The lemma follows. □

In particular, $\mathcal{F}$ is gauge invariant if and only if $[\mu] = -2\pi c_1(\mathfrak{s})$ in $H^2(Y;\mathbb{R})$.

9.2. **Seiberg-Witten Floer homology.** Roughly speaking, Seiberg-Witten Floer homology is the Morse homology of the functional $\mathcal{F}$ on $\mathcal{C}/\mathcal{G}$, where the perturbation 2-form μ is taken to be exact[10]. The detailed construction is carried out in [38]. Some basic points to keep in mind are the following:

(1) When $c_1(\mathfrak{s})$ is not torsion, the functional $\mathcal{F}$ is not gauge invariant, so it is not actually a real-valued functional on $\mathcal{C}/\mathcal{G}$. One can still define its Morse theory in this case, but we will not explain the details of this because we will only be concerned with the case where $c_1(\mathfrak{s})$ is torsion below.

(2) The quotiented configuration space $\mathcal{C}/\mathcal{G}$ on which $\mathcal{F}$ is defined is not a manifold in any natural sense, because $\mathcal{G}$ does not act freely on the reducibles. On the other hand, $\mathcal{C}/\mathcal{G}$ is the quotient of a manifold by an S^1 action. Namely, if one fixes a point $y_0 \in Y$ and defines $\mathcal{G}_0 := \{g \in \mathcal{G} \mid g(y_0) = 1\}$, then $\mathcal{G}_0$ acts freely on $\mathcal{C}$, so $\mathcal{C}/\mathcal{G}_0$ is a manifold, and $\mathcal{C}/\mathcal{G} = (\mathcal{C}/\mathcal{G}_0)/S^1$. One now wants to define some kind of S^1-equivariant Morse homology of $\mathcal{F}$ on $\mathcal{C}/\mathcal{G}_0$.

The approach taken by Kronheimer-Mrowka is, roughly speaking, to blow up the singularities of $\mathcal{C}/\mathcal{G}$, so as to obtain a manifold-with-boundary $\widetilde{\mathcal{C}/\mathcal{G}}$, where the boundary arises from the reducibles. The gradient flow of $\mathcal{F}$ induces a (partially defined) flow on $\widetilde{\mathcal{C}/\mathcal{G}}$, which is tangent to the boundary.

There is now a finite-dimensional model for how to proceed. Let X be a finite dimensional compact manifold with boundary, and let (f, g) be a Morse-Smale pair on X such that the gradient flow is tangent to the boundary. In this context, as explained in [38, §2], there are three versions of Morse homology one can define, which fit into a long exact sequence:

$$\overline{H}_*^{\mathrm{Morse}}(X,f,g) \longrightarrow \check{H}_*^{\mathrm{Morse}}(X,f,g) \longrightarrow \widehat{H}_*^{\mathrm{Morse}}(X,f,g) \longrightarrow \overline{H}_{*-1}^{\mathrm{Morse}}(X,f,g) \longrightarrow$$

Here $\overline{H}_*$ is just the Morse homology of the boundary. The version $\check{H}_*$ is the homology of a chain complex which is freely generated over $\mathbb{Z}$ by interior critical points and "boundary stable" critical points on the boundary, and whose differential counts certain configurations of flow lines between them. The version $\widehat{H}_*$ is similar but its generators include "boundary unstable" critical points on the boundary instead. The above exact sequences turns out to agree with the relative homology exact sequence

$$H_*(\partial X) \longrightarrow H_*(X) \longrightarrow H_*(X, \partial X) \longrightarrow H_{*-1}(\partial X) \longrightarrow \cdots$$

Carrying out an analogue of this construction on the blown-up configuration space now gives three versions of Seiberg-Witten Floer homology, which fit into a long exact sequence:

$$\overline{HM}_*(Y,\mathfrak{s}) \longrightarrow \check{HM}_*(Y,\mathfrak{s}) \longrightarrow \widehat{HM}_*(Y,\mathfrak{s}) \longrightarrow \overline{HM}_{*-1}(Y,\mathfrak{s}) \longrightarrow \cdots$$

The version $\overline{HM}$ comes entirely from the reducibles. Since the reducibles are described explicitly by Lemma 8.10, it is possible (although not trivial) to compute $\overline{HM}$ in terms of classical algebraic topology, specifically the triple cup product on $H^*(Y)$, see [38, §35.1]. As such, $\overline{HM}_*$ may seem less interesting than the other two versions $\check{HM}_*$ and $\widehat{HM}_*$. However the computation of $\overline{HM}_*$ is used in conjunction with the above exact sequence to prove Kronheimer-Mrowka's nontriviality result in Theorem 7.5, which plays an essential role in the proof of the Weinstein conjecture.

---

[10] One can also define versions of Seiberg-Witten Floer homology when μ is closed but not exact, but these have different properties.

(3) The three versions of Seiberg-Witten Floer homology above all have a relative $\mathbb{Z}/d(c_1(\mathfrak{s}))$-grading (given by the expected dimension of the moduli space of flow lines, which equals a certain spectral flow). So if $c_1(\mathfrak{s})$ is torsion then the Seiberg-Witten Floer homologies are relatively $\mathbb{Z}$-graded. As mentioned previously, there is in fact an absolute grading by homotopy classes of oriented 2-plane fields, which is compatible with the map $\mathcal{P}(Y) \to \operatorname{Spin}^c(Y)$ discussed in Remark 8.3. However for our purposes it is enough to just regard the grading as taking values in some affine space over $\mathbb{Z}/d(c_1(\mathfrak{s}))$.

(4) Heuristically the differentials in the Seiberg-Witten Floer chain complexes count gradient flow lines of $\mathcal{F}$, but in fact some abstract perturbations of the equations are required in order to obtain the transversality needed to count solutions. The perturbations are explained in detail in [38] and will be suppressed in the exposition here.

10. Outline of Taubes's proof

We now have the background in place to describe Taubes's proof of the Weinstein conjecture. This section gives an outline, and the next section explains some more details. Below we mostly follow Taubes's paper [62] and MSRI lectures [44].

10.1. **Geometric setup.** Let Y be a closed oriented connected 3-manifold with a contact form λ. Fix a Riemannian metric on Y such that $|\lambda| = 1$ and $d\lambda = 2 * \lambda$. We can do this because $\lambda \wedge d\lambda > 0$.

To be consistent with Taubes, denote[11] the oriented 2-plane field $\xi = \operatorname{Ker}(\lambda)$, regarded as a Hermitian line bundle, by K^{-1}. Recall that ξ determines a distinguished spin-c structure $\mathfrak{s}_\xi$, in which

$$\mathbb{S} = \underline{\mathbb{C}} \oplus K^{-1}. \tag{10.1}$$

Any spin-c structure $\mathfrak{s}$ is obtained from $\mathfrak{s}_\xi$ by tensoring with a Hermitian line bundle E, so that

$$\mathbb{S} = E \oplus K^{-1}E. \tag{10.2}$$

In this decomposition, E is the $+i$ eigenspace of Clifford multiplication by λ, while $K^{-1}E$ is the $-i$ eigenspace. The significance of E is that Taubes's Theorem 7.4 ultimately shows that the Seiberg-Witten Floer cohomology $\widehat{HM}^{-*}(Y, \mathfrak{s})$ is isomorphic to $ECH_*(Y, \lambda, \Gamma)$ where Γ is Poincare dual to $c_1(E)$.

For any spin-c structure as in (10.2), connections on $\det(\mathbb{S})$ can be written as $A_0 + 2A$ where A_0 is a reference connection on K^{-1} while A is a connection on E. In fact Taubes picks out a distinguished connection A_0 on K^{-1} as follows. For the distinguished spin-c structure $\mathfrak{s}_\xi$, let ψ_0 denote the spinor given by $(1, 0)$ in the decomposition (10.1).

Lemma 10.1. *There is a unique Hermitian connection A_0 on K^{-1} such that*

$$D_{A_0}\psi_0 = 0.$$

---

[11] This notation is carried over from the SW=Gr story, where K denotes the canonical bundle of a symplectic 4-manifold. For a contact 3-manifold, K is the canonical bundle of the symplectization.

Proof. Uniqueness follows from the formula

$$D_{A_0+2a}\psi_0 = D_{A_0}\psi_0 + \rho(a)\psi_0$$

and the equations (8.1).

To prove existence, let A be any Hermitian connection on K^{-1}, let ∇_A denote the associated spin-c connection, and write $\nabla_A\psi_0 = (\alpha, \beta)$ where α and β are 1-forms on Y with values in $\mathbb{C}$ and K^{-1} respectively. Since ∇_A is Hermitian, we have $\mathrm{Re}(\alpha) = 0$. It follows that there is a unique A_0 such that

$$\nabla_{A_0}\psi_0 = (-i\lambda, \beta_0) \tag{10.3}$$

for some β_0. We now show that $D_{A_0}\psi_0 = 0$.

We need the following Leibniz-type formula for the Dirac operator: If A is any spin-c connection, ω is any differential form, and ψ is any spinor, then

$$D_A(\rho(\omega)\psi) = \rho((d + d^*)\omega)\psi + \rho((1 \otimes \rho(\omega))\nabla_A\psi). \tag{10.4}$$

Here Clifford multiplication is extended to an action of Λ^*TX on $\mathbb{S}$ by the rule

$$\rho(a \wedge b) = \frac{1}{2}\Big(\rho(a)\rho(b) + (-1)^{\deg(a)\deg(b)}\rho(b)\rho(a)\Big).$$

Taking $A = A_0$, $\psi = \psi_0$, and $\omega = \lambda$ in (10.4), we obtain

$$iD_{A_0}\psi_0 = (-2i, 0) + \rho(\lambda, -i\beta_0).$$

On the other hand, applying ρ to equation (10.3) and multiplying by $-i$ gives

$$-iD_{A_0}\psi_0 = \rho(-\lambda, -i\beta_0).$$

Subtracting the above equation from the previous one and using the fact that $\rho(\lambda, 0) = (i, 0)$ gives $2iD_{A_0}\psi_0 = 0$. □

Henceforth, think of spin connections as being determined by the distinguished connection A_0 on K^{-1} together with a connection A on E.

10.2. **Taubes's perturbation.**

The idea of Taubes's proof of the Weinstein conjecture is to deform the Seiberg-Witten equations by a sequence of increasingly large perturbations, and to use a sequence of solutions to the perturbed equation, provided by the known nontriviality of Seiberg-Witten Floer homology in Theorem 7.5, to yield a nonempty collection of Reeb orbits.

The basic perturbation of the Seiberg-Witten equations considered by Taubes is

$$\begin{aligned} *F_A &= r(\langle\rho(\cdot)\psi, \psi\rangle - i\lambda) + i\overline{\omega}, \\ D_A\psi &= 0. \end{aligned} \tag{10.5}$$

These should be regarded as equations for the pair (A, ψ) that are parametrized by r. Here $r \geq 1$ (the deformation involves taking $r \to \infty$), and remember that now A denotes a connection on E, while ψ is a section of $E \oplus K^{-1}E$. Also $\overline{\omega}$ denotes the harmonic 1-form whose Hodge star represents the image of $\pi c_1(K^{-1})$ in $H^2(Y; \mathbb{R})$. Taubes's equations (10.5) are equivalent to a case of the perturbed Seiberg-Witten equations (8.5), via rescaling the spinor by a factor of $\sqrt{r}$ and taking the perturbation to be the exact 2-form

$$\mu = -rd\lambda - iF_{A_0} + 2 * \overline{\omega}. \tag{10.6}$$

Some parts of the argument involve further perturbations of the equations (10.5) in order to obtain necessary transversality. For simplicity, we will suppress these in the exposition below.

10.3. **The trivial solution.** A first observation regarding the equations (10.5) is that for the spin-c structure $\mathfrak{s}_\xi$, if the $\overline{\omega}$ term were omitted, then the pair (A_0, ψ_0) from Lemma 10.1 would be a solution for any r. In general we do not want to remove the $\overline{\omega}$ term from (10.5) because the perturbation 2-form (10.6) needs to be exact in order to obtain the correct Seiberg-Witten Floer homology. However since the $\overline{\omega}$ term is much smaller than the term with a factor of r in (10.5) when r is large, a perturbation argument can be used to prove the following:

Lemma 10.2. [63, Prop. 2.8] *For any $\delta > 0$, if r is sufficiently large then:*

- *There exists a unique (up to gauge equivalence) solution (A_{triv}, ψ_{triv}) to the equations* (10.5) *for the spin-c structure $\mathfrak{s}_\xi$ such that $1 - |\psi_{triv}| \le \delta$ on all of Y. (In fact $1 - |\psi_{triv}| = O(r^{-1/2})$.)*
- *The grading of (A_{triv}, ψ_{triv}) in the Seiberg-Witten Floer chain complex is independent of r.*

We will call (A_{triv}, ψ_{triv}) the "trivial solution". Ultimately, in the isomorphism with embedded contact homology, the trivial solution corresponds to the empty set of Reeb orbits.

10.4. **Convergence to Reeb orbits.** Taubes now proves the following theorem[12] which finds Reeb orbits from a sequence of solutions to (10.5).

Theorem 10.3. *(*[62, Thm. 2.1]*) Fix the line bundle E and let (r_n, ψ_n, A_n) be a sequence of solutions to the equations* (10.5) *with $r_n \to \infty$. Suppose that:*

(1) *There is a constant $\delta > 0$ with $\sup_Y (1 - |\psi_n|) > \delta$.*
(2) *There is a constant $C < \infty$ with $i \int_Y \lambda \wedge F_{A_n} < C$.*

Then there exists a nonempty orbit set a with $[a] = \operatorname{PD}(c_1(E))$.

We will explain the proof of this theorem in some detail in §11. For now we remark that this is a three-dimensional analogue of an earlier theorem of Taubes for symplectic four-manifolds in [60], part of the SW=Gr story, which obtains holomorphic curves from sequences of solutions to the Seiberg-Witten equations perturbed using the symplectic form. However since the dimension is one less here, subtle measure-theoretic arguments that were used in the four-dimensional case can be avoided, and the proof is considerably shorter. The basic idea is to write $\psi_n = (\alpha_n, \beta_n)$, where α_n is a section of E and β_n is a section of $K^{-1}E$, and show that one can pass to a subsequence such that $\alpha_n^{-1}(0)$ converges as a current to a nonempty orbit set. This will then, of course, represent the Poincare dual of $c_1(E)$. In fact, when n is large, $|\beta_n|$ will be close to zero everywhere, while $|\alpha_n|$ will be close to 1 except near its zero set. Also the curvature F_{A_n} will be concentrated near the zero set of α_n, and its direction will be approximately dual to the normal plane to $\alpha_n^{-1}(0)$.

Assumption (1) is needed to avoid solutions with α_n nonvanishing, which can exist when $c_1(E) = 0$, as we have seen in Lemma 10.2.

The idea of assumption (2) is that when n is large, $i \int_Y \lambda \wedge F_{A_n}$ is approximately 2π times the symplectic action (6.1) of the orbit set to which $\alpha_n^{-1}(0)$ is converging. A uniform upper bound on this integral is needed in order to obtain an orbit set of finite length. (In a more general situation without condition (2) one can still obtain some weaker conclusions, see §12.3.)

---

[12] Taubes actually proves more general versions of this theorem, and many of the other results of his that we are quoting here.

Note also that assumption (2) guarantees that (A_n, ψ_n) is irreducible when n is sufficiently large, because it follows from the equations (10.5) that if $(A, 0)$ is a reducible solution to (10.5) then $i\int_Y \lambda \wedge F_A$ is a linear, increasing function of r.

10.5. **Avoiding the empty set.** Now fix E such that $c_1(K^{-1}) + 2c_1(E)$ is torsion in $H^2(Y;\mathbb{Z})$. Let $\mathfrak{s}$ denote the corresponding spin-c structure. Kronheimer-Mrowka's Theorem 7.5 guarantees the existence of solutions to the perturbed equations (10.5) for all $r \geq 1$. To complete the proof of Theorem 7.6, which implies the Weinstein conjecture, we need to find a sequence of such solutions with $r \to \infty$ such that conditions (1) and (2) in Theorem 10.3 are satisfied.

One can achieve condition (1) using the following lemma, proved in §11:

Lemma 10.4. *If $c_1(E) \neq 0$, then there is a constant $c > 0$ such that if r is sufficiently large, and if (A, ψ) is a solution to the equations* (10.5)*, then there exist points in Y where $1 - |\psi| \geq 1 - c/\sqrt{r}$.*

This means that if $c_1(E) \neq 0$ then any sequence of solutions will automatically satisfy condition (1) (after discarding some initial terms). On the other hand, by Lemma 10.2, if $c_1(E) = 0$ then a sequence of solutions will likewise satisfy condition (1) as long as we avoid the gauge equivalence class of the trivial solution (A_{triv}, ψ_{triv}). We can easily do this since we know from Theorem 7.5 that the Seiberg-Witten Floer homology is nonzero in infinitely many gradings.

10.6. **Three functionals.** The hardest part of the proof of the Weinstein conjecture is to achieve condition (2) in Theorem 10.3. This is a new problem which does not arise in the four-dimensional SW=Gr story. (On a symplectic four-manifold (X, ω), to obtain convergence to a holomorphic curve, one needs an analogue of condition (2) in which λ replaced by the symplectic form ω; but there the quantity that needs to be bounded is constant because the symplectic form is closed.)

The first step is to write the Chern-Simons-Dirac functional $\mathcal{F}$ in (9.1), of which Seiberg-Witten Floer homology is the Morse homology, as the sum of two other functionals. To do so, fix a reference connection A_1 on E.

Definition 10.5. If A is a connection on E, define the *Chern-Simons functional* $cs(A)$ by

$$cs(A) := -\int_Y (A - A_1) \wedge (F_A + F_{A_1} - 2i * \overline{\omega}). \tag{10.7}$$

Note that this is gauge invariant thanks to our assumption that $2c_1(E) + c_1(K^{-1})$ is torsion. Also, define the *energy*

$$\mathcal{E}(A) := i\int_Y \lambda \wedge F_A. \tag{10.8}$$

This is the quantity that we want to control.

Observe now that for a given r, if in the definition of $\mathcal{F}$ we take our reference connection on $\det(\mathbb{S})$ to be $A_0 + 2A_1$, then we have

$$\mathcal{F}(A, \psi) = \frac{1}{2}\left(cs(A) - r\mathcal{E}(A)\right) + \frac{r}{2}\int_Y \langle D_A\psi, \psi\rangle d\,\mathrm{vol},$$

up to the addition of an r-dependent constant. Since adding a constant to $\mathcal{F}$ does not affect its Morse homology, we will ignore this constant and take the above equation to be the new definition of $\mathcal{F}$. In particular, if (A,ψ) is a solution to the perturbed Seiberg-Witten equations (10.5), then the three functionals in play are related by

$$\mathcal{F}(A,\psi) = \frac{1}{2}\left(cs(A) - r\mathcal{E}(A)\right). \tag{10.9}$$

10.7. **A piecewise smooth family of solutions.** The next step is:

Lemma 10.6. *(Up to the perturbations we are suppressing) one can choose for each r sufficiently large a solution $(A(r),\psi(r))$ to the equations* (10.5) *such that:*

- $(A(r),\psi(r))$ *is a piecewise smooth function of r.*
- $\mathcal{F}(A(r),\psi(r))$ *is a continuous function of r.*
- *For all r at which $(A(r),\psi(r))$ is smooth as a function of r, $(A(r),\psi(r))$ is nondegenerate*[13] *and its grading in the Seiberg-Witten Floer chain complex is independent of r.*
- $(A(r),\psi(r))$ *is not gauge equivalent to the trivial solution (A_{triv},ψ_{triv}) described in Lemma 10.2.*

The idea of the proof of Lemma 10.6 is as follows. First, one shows that for any given grading, if r is sufficiently large then all generators of the chain complex defining $\widehat{HM}_*(Y,\mathfrak{s})$ with that grading are irreducible. This is proved using a spectral flow estimate related to Proposition 10.11 below. Thus, for any given range of gradings, if r is sufficiently large then the differential in the Seiberg-Witten Floer chain complex just counts (perturbed) gradient flow lines of $\mathcal{F}$, without the subtleties arising from reducibles.

Now by Theorem 7.5, there is a nonzero class σ in the Seiberg-Witten Floer homology $\widehat{HM}_*(Y,\mathfrak{s})$, and when $c_1(E)=0$ we can assume that the grading of σ is not the same as that of (A_{triv},ψ_{triv}). Fix such a class σ.

One now sets up the perturbations (that we have suppressed in the exposition) so that the Seiberg-Witten Floer chain complex is defined for generic r. For such r, we define $h(r)\in\mathbb{R}$ to be "the minimum height of $\mathcal{F}$ needed to represent the class σ". More precisely, a chain representing the class σ can be expressed as $\sum_i n_i c_i$, where n_i is a nonzero integer and c_i is a critical point of $\mathcal{F}$ for each i in some finite set. Define $h(r)$ to be the minimum, over all chains $\sum_i n_i c_i$ representing σ, of $\max_i \mathcal{F}(c_i)$.

The idea is then to define $(A(r),\psi(r))$ to be a maximal $\mathcal{F}$ critical point in a representative of σ realizing the minimum $h(r)$. One can choose this $(A(r),\psi(r))$ to vary piecewise smoothly with r, jumping when the "Morse complex" undergoes a bifurcation involving the critical point $(A(r),\psi(r))$.

To complete the proof of Lemma 10.6, one needs to show that the function $h(r)$ defined above extends to a continuous function of all sufficiently large r. Taubes does so by explicitly studying the bifurcations that can happen in a generic one-parameter family of "Morse complexes". One can presumably also prove this by estimating that the continuation maps that relate the Seiberg-Witten Floer homologies for nearby values of r do not increase the functional $\mathcal{F}$ too much. This method has been used to prove analogous continuity results in symplectic Floer homology, see eg [57, §2.4]. Here is a model for this argument in finite dimensional Morse theory:

---

[13] A critical point (A,ψ) of $\mathcal{F}$ is "nondegenerate" if the Hessian of the functional $\mathcal{F}$ at (A,ψ) has kernel zero, so that the grading of (A,ψ) in the Seiberg-Witten Floer chain complex is well-defined.

Proposition 10.7. *Let* X *be a finite dimensional closed manifold and let* $\{(f_r, g_r)\}$ *be a generic smooth family of functions* $f_r : X \to \mathbb{R}$ *and metrics* g_r *on* X *parametrized by* $r \in [0,1]$. *Fix* $0 \neq \sigma \in H_*(X)$. *For generic* r, *such that the pair* (f_r, g_r) *is Morse-Smale, define* $h(r)$ *to be the minimum height of a representative of the class* σ. *Then* h *extends to a continuous function on all of* $[0,1]$.

Proof. Note that the pair (f_r, g_r) is Morse-Smale for all but finitely many r. Recall from §5.4 that the Morse homologies of the Morse-Smale pairs (f_r, g_r) are canonically isomorphic to each other, via continuation maps, and also to $H_*(X)$, so that σ defines a class in the Morse homology for each Morse-Smale pair (f_r, g_r) which is preserved by these continuation maps.

Now suppose that $r < r'$ and the pairs (f_r, g_r) and $(f_{r'}, g_{r'})$ are both Morse-Smale. The continuation isomorphism from r' to r is induced by a chain map

$$\Phi : C_*^{\text{Morse}}(X, f_{r'}, g_{r'}) \longrightarrow C_*^{\text{Morse}}(f_r, g_r)$$

which counts maps $f : \mathbb{R} \to X$ satisfying the equation

$$\frac{d\gamma(s)}{ds} = \nabla f_{\phi(s)}(\gamma(s)) \tag{10.10}$$

where $\phi : \mathbb{R} \to [r, r']$ is a monotone smooth function satisfying $\phi(s) = r$ for $s \leq 0$ and $\phi(s) = r'$ for $s \geq 1$. Now if γ is a solution to (10.10) then by the chain rule we have

$$\frac{d}{ds} f_{\phi(s)}(\gamma(s)) = |\nabla f_{\phi(s)}(\gamma(s))|^2 + \frac{d\phi(s)}{ds}\frac{\partial f_r(x)}{\partial r}\Big|_{r=\phi(s), x=\gamma(s)}. \tag{10.11}$$

Since $[0,1] \times X$ is compact there is a constant c such that $|\partial f_r(x)/\partial r| < c$ for all $(r, x) \in [0,1] \times X$. It then follows from (10.11) that if p' and p are critical points of $f_{r'}$ and f_r respectively with $\langle \Phi p', p \rangle \neq 0$, then

$$f_{r'}(p') \geq f_r(p) - c(r' - r).$$

That is, the continuation map from r' to r increases the height by at most $c(r'-r)$, so

$$h(r) \leq h(r') + c(r' - r).$$

The same holds for the continuation map in the other direction from r to r', so we conclude that

$$|h(r) - h(r')| \leq c|r - r'|.$$

The proposition follows. □

Accepting Lemma 10.6, we now have:

Lemma 10.8. *Let* $\{(A(r), \psi(r))\}$ *be a piecewise smooth family from Lemma 10.6. Then*

$$\frac{d}{dr}\mathcal{F}(A(r), \psi(r)) = -\frac{1}{2}\mathcal{E}(A(r)).$$

Proof. This follows from a general principle: If X is a smooth manifold (finite or infinite dimensional), if $f : \mathbb{R} \times X \to \mathbb{R}$ is a smooth function, and if $\{x(t)\}$ is a smooth family of critical points of $f_t := f(t, \cdot)$ on X defined for t in some interval, then

$$\frac{d}{dt} f_t(x(t)) = \frac{\partial f}{\partial t}(t, x(t)).$$

To prove this one uses the chain rule as in (10.11) to compute

$$\frac{d}{dt} f(t, x(t)) = \frac{\partial f}{\partial t}(t, x(t)) + df_t\left(\frac{dx(t)}{dt}\right),$$

and notes that the second term on the right vanishes because $x(t)$ is a critical point of f_t. □

10.8. **The energy dichotomy.** Let $\{(A(r), \psi(r)\}$ be a piecewise smooth family given by Lemma 10.6. To prove Theorem 7.6 and thereby the Weinstein conjecture, by Theorem 10.3 we just need to show that there is a sequence $r_n \to \infty$ such that the energy $\mathcal{E}(A(r_n))$ is bounded. The next step in Taubes's argument is to show that if this is not the case, then there is a sequence $r_n \to \infty$ such that the energy $\mathcal{E}(A(r_n))$ grows at least linearly, and the Chern-Simons functional grows quadratically. (The last step will be to show that this quadratic growth of the Chern-Simons functional leads to a contradiction.)

Lemma 10.9. *Let $\{(A(r), \psi(r)\}$ be a piecewise smooth family given by Lemma 10.6. Then at least one of the following two alternatives holds:*

(1) *There is a sequence $r_n \to \infty$ and a constant C such that $\mathcal{E}(A(r_n)) < C$ for all n.*
(2) *There is a sequence $r_n \to \infty$ and a constant $c > 0$ such that $\mathcal{E}(A(r_n)) \geq cr_n$ and $cs(A(r_n)) \geq cr_n^2$ for all n.*

The proof of Lemma 10.9 uses the following a priori estimate, which is proved in §11.

Lemma 10.10. *There is a constant c such that if (r, A, ψ) is a solution to the equations* (10.5) *with $\mathcal{E}(A) > 1$ then*

$$|cs(A)| \leq cr^{2/3}\mathcal{E}(A)^{4/3}. \tag{10.12}$$

Granted this, we can now give:

Proof of Lemma 10.9. Introduce the shorthand $cs(r) := cs(A(r))$, $\mathcal{E}(r) := \mathcal{E}(A(r))$ and $\mathcal{F}(r) := \mathcal{F}(A(r), \psi(r))$. We can assume without loss of generality that $\mathcal{E}(r) > 1$ for all r sufficiently large (since otherwise case (1) holds). Now fix $\varepsilon_0 \in (0, 1/5)$. We consider two cases.

Case A: There is a sequence $r_n \to \infty$ with

$$cs(r_n) \geq \varepsilon_0 r_n \mathcal{E}(r_n)$$

for all n. It follows in this case from the inequality (10.12) that alternative (2) holds.

Case B: For all r sufficiently large,

$$cs(r) < \varepsilon_0 r \mathcal{E}(r). \tag{10.13}$$

In this case we will show that alternative (1) holds.

To do so, define

$$v(r) := \mathcal{E}(r) - \frac{cs(r)}{r} = -\frac{2\mathcal{F}}{r}.$$

It then follows from Lemma 10.8 that

$$\frac{dv}{dr} = \frac{cs}{r^2}. \tag{10.14}$$

On the other hand, the hypothesis (10.13) is equivalent to

$$\mathcal{E} < (1-\varepsilon_0)^{-1}v. \tag{10.15}$$

It now follows from (10.14), (10.13) and (10.15) that

$$\frac{dv}{dr} < \frac{\varepsilon v}{r},$$

where $\varepsilon := (1-\varepsilon_0)^{-1}\varepsilon_0 < 1/4$. Therefore

$$v < c_1 r^{\varepsilon} \tag{10.16}$$

for some constant c_1. On the other hand, by (10.12), (10.15) and (10.16), we have

$$cs < c_2 r^{2/3+(4/3)\varepsilon}$$

for some constant c_2. Putting this back into (10.14), we get

$$\frac{dv}{dr} < c_2 r^{(4/3)(\varepsilon-1)}.$$

Since $\varepsilon < 1/4$, the exponent in the above inequality is less than -1. Consequently the above inequality can be integrated to show that v is bounded from above. Then $\mathcal{E}$ is also bounded from above by (10.15). □

10.9. **Controlling the Chern-Simons functional.** The last step in Taubes's proof of the Weinstein conjecture is the following proposition relating the Seiberg-Witten Floer grading to the Chern-Simons functional. To state it, if (A,ψ) is a solution to the perturbed Seiberg-Witten equations (10.5), let $\deg(A,\psi)$ denote its grading in the Seiberg-Witten Floer chain complex, and recall that (A_{triv},ψ_{triv}) denotes the distinguished Seiberg-Witten Floer generator given by Lemma 10.2.

Proposition 10.11. [62, Prop. 5.1] *There exists $\kappa > 0$ such that for all r sufficiently large, if (A,ψ) is a nondegenerate solution to* (10.5)*, then*

$$\left|\deg(A,\psi) - \deg(A_{triv},\psi_{triv}) + \frac{1}{4\pi^2}cs(A)\right| < \kappa r^{31/16}.$$

This is proved using a new estimate on the spectral flow of one-parameter families of Dirac operators (the latter determines the relative grading in Seiberg-Witten Floer homology). While this is a crucial new element of Taubes's proof of the Weinstein conjecture, is too much for us to explain here, so we refer the reader to [62, §5] for the spectral flow estimate that proves Proposition 10.11, and to [64] for a higher-dimensional generalization.

10.10. **Conclusion.** To prove the Weinstein Conjecture, more specifically Theorem 7.6, let $\Gamma \in H_1(Y)$ such that $c_1(\xi) + 2\,\mathrm{PD}(\Gamma)$ is torsion, and let E be the line bundle with $c_1(E) = \mathrm{PD}(\Gamma)$. Let $\{(A(r), \psi(r)\}$ be a piecewise smooth family of the perturbed Seiberg-Witten equations (10.5) given by Lemma 10.6. Alternative (2) in Lemma 10.9 is impossible by Proposition 10.11, because $\deg(A, \psi)$ and $\deg(A_{triv}, \psi_{triv})$ are independent of r. So alternative (1) in Lemma 10.9 holds. Then we have a sequence $r_n \to \infty$ such that condition (2) in Theorem 10.3 holds. Condition (1) in Theorem 10.3 also holds by Lemmas 10.2 and 10.4. Thus Theorem 10.3 applies to produce the desired nonempty orbit set.

11. More details of Taubes's proof

We now fill in some of the (more basic) details of Taubes's proof that were omitted from §10. In particular §11.4 sketches the proof of Theorem 10.3 on convergence of Seiberg-Witten solutions to Reeb orbits.

11.1. **Prelude: the vortex equations.** Before studying the three-dimensional Seiberg-Witten equations in more detail, it is useful to recall a two-dimensional version of the equations: the *vortex equations* on $\mathbb{C}$. These are equations for a pair (A, α) where A is an imaginary-valued 1-form on $\mathbb{C}$, and α is a complex-valued function on $\mathbb{C}$. One can also think of A as a Hermitian connection on the trivial line bundle over $\mathbb{C}$, and α as a section of this bundle. The equations are now

$$\begin{aligned} \overline{\partial}_A \alpha &= 0, \\ *F_A &= -i(1 - |\alpha|^2). \end{aligned} \tag{11.1}$$

Here $\overline{\partial}_A\alpha = (\overline{\partial} + A^{0,1})\alpha$, and $F_A = dA$. For example there is a "trivial solution" ($A \equiv 0, \alpha \equiv 1$). More generally, solutions to the vortex equations arise as $\mathbb{R}$-invariant solutions to the Seiberg-Witten equations on $\mathbb{R} \times \mathbb{C}$, for a suitable perturbation. The following are some basic properties of solutions to the vortex equations, see [37] and [59, §2b]:

- If $|\alpha| = 1$ at any point, then $|\alpha| \equiv 1$.
- The zeroes of α are isolated and have positive multiplicity.
- If one further assumes the "finite energy" condition
$$\int_{\mathbb{C}} (1 - |\alpha|^2) < \infty,$$
then $\int_{\mathbb{C}}(1 - |\alpha|^2) = 2\pi k$ where k is a nonnegative integer, and α has exactly k zeroes counted with multiplicity.

The first and last of the above facts imply the following basic observation which will be needed later:

Lemma 11.1. *If (A, α) is a finite energy solution to the vortex equations, and if $|\alpha| \not\equiv 1$, then α has a zero.*

For motivational purposes we now recall some additional facts about solutions to the vortex equations:

- There is a constant $c > 0$ such that if $d : \mathbb{C} \to [0, \infty)$ denotes the distance to the set where $|\alpha|^2 < 1/2$, then where $d > c^{-1}$ one has

$$1 - |\alpha|^2 \le e^{-cd}. \tag{11.2}$$

- There is a bijection from the set of solutions with $\int_{\mathbb{C}}(1-|\alpha|^2) = 2\pi k$, modulo gauge equivalence, to the k^{th} symmetric product of $\mathbb{C}$, sending (A, α) to $\alpha^{-1}(0)$. Here the gauge group $\mathcal{G} = \operatorname{Maps}(\mathbb{C}, S^1)$ acts on the set of solutions by $u \cdot (A, \alpha) := (A - u^{-1}du, u\alpha)$.

This last fact can be regarded as a two-dimensional version of the "SW=Gr" Theorem 7.1. It also generalizes to vortices over a closed surface, see [23]. For some more recent work relating this story to Seiberg-Witten theory see eg [43, 52, 70].

It is also useful to consider a variant of the vortex equations (11.1), let us call these the "r-vortex equations", in which the second equation is replaced by

$$*F_A = -ir(1 - |\alpha|^2)$$

where r is a large positive constant. A solution to the r-vortex equations is equivalent to a solution to the original vortex equations (11.1) under the rescaling $\mathbb{C} \to \mathbb{C}$ sending $z \mapsto r^{1/2}z$. So for a solution to the r-vortex equations, the exponential decay (11.2) is replaced by the stronger decay

$$1 - |\alpha|^2 \le e^{-cr^{1/2}d}.$$

The upshot is that for r large, given a solution (A, α) to the r-vortex equations, the curvature F_A is concentrated near the zero set of α, while away from the zero set A is close to flat and $|\alpha|$ is close to 1.

The picture we are now aiming for in three dimensions is that if $(A, \psi = (\alpha, \beta))$ is a solution to the three-dimensional perturbed Seiberg-Witten equations (10.5) where r is large and the energy (10.8) is bounded, then F_A is concentrated near the zero set of α, the latter is approximated by a union of Reeb orbits, and away from this zero set $|\alpha|$ is close to 1 and β is close to 0.

11.2. **The perturbed Seiberg-Witten equations, more explicitly.** If (A, ψ) is a solution to the perturbed Seiberg-Witten equations (10.5), write $\psi = (\alpha, \beta)$ where α is a section of E and β is a section of $K^{-1}E$. We now rewrite the perturbed Seiberg-Witten equations in terms of A, α and β and establish some notation which will be used below.

The curvature equation in (10.5) can be written as

$$*F_A = ir\left((|\alpha|^2 - |\beta|^2 - 1)\lambda + 2\operatorname{Im}(\alpha\beta^*)\right) + i\varpi. \tag{11.3}$$

Here $\alpha\beta^*$ denotes the $\mathbb{C}$-valued 1-form that projects a vector onto K^{-1}, pairs it with dual of β in $(K^{-1}E)^*$, and then pairs the result with α.

To rewrite the Dirac equation, let ∇_A denote the covariant derivative on E corresponding to A, and also the covariant derivative on $K^{-1}E$ induced by A together with the canonical connection A_0 on K^{-1} from Lemma 10.1. Let $\overline{\partial}_A\alpha$ denote the complex antilinear part of the restriction of $\nabla_A\alpha : TY \to E$ to $\xi = K^{-1}$. Likewise let $\partial_A\beta$ denote the complex linear part of the restriction of $\nabla_A\beta : TY \to K^{-1}E$ to K^{-1}. Let $\nabla_{A,R}$ denote the covariant derivative along the Reeb vector field. The Dirac equation in (10.5) can now be written as[14]

$$\begin{pmatrix} i\nabla_{A,R}\alpha - 2\partial_A\beta \\ -i\nabla_{A,R}\beta + 2\overline{\partial}_A\alpha \end{pmatrix} + \begin{pmatrix} f_0\alpha + f_1\beta \\ f_0'\alpha + f_1'\beta \end{pmatrix} = 0. \tag{11.4}$$

Here f_0, f_1, f_0', f_1' are bundle maps between E and $K^{-1}E$ which do not depend on r; they arise from the failure of the spin covariant derivative on $E \oplus K^{-1}E$ to agree with the covariant derivatives ∇_A on E and $K^{-1}E$ (which in turn is related to the failure of the contact geometry to be parallel with respect to the Levi-Civita connection).

---

[14] Taubes obtains a slightly different equation here. See Remark 8.7.

11.3. **A priori estimates.** The starting point for the analysis of the perturbed Seiberg-Witten equations (10.5) is the following lemma giving a priori estimates on their solutions.

Lemma 11.2. [62, Lemmas 2.2 and 2.3] *There exists a constant* c_0 *such that if* r *is sufficiently large and if* $(A, \psi = (\alpha, \beta))$ *is a solution to the equations* (10.5)*, then:*

$$|\alpha| \le 1 + \frac{c_0}{r}, \tag{11.5}$$

$$|\beta|^2 \le \frac{c_0}{r}\left|1 - |\alpha|^2\right| + \frac{c_0}{r^2}, \tag{11.6}$$

$$|\nabla_A \alpha| \le c_0 r^{1/2}, \tag{11.7}$$

$$|\nabla_A \beta| \le c_0. \tag{11.8}$$

Proof. Let $\widetilde{\nabla}_A$ denote the spin covariant derivative on $\mathbb{S}$, in order to avoid confusion with the covariant derivatives ∇_A on E and $K^{-1}E$. Putting the perturbed Seiberg-Witten equations (10.5) into the Bochner-Lichnerowitz-Weitzenbock formula (8.4), we obtain

$$0 = \widetilde{\nabla}_A^* \widetilde{\nabla}_A \psi + \frac{s}{4}\psi + \frac{r}{2}\left(|\psi|^2\psi + i\rho(\lambda)\psi\right) - \frac{i}{2}\rho(\overline{\omega})\psi. \tag{11.9}$$

Since the spin connection is compatible with the Hermitian metric we have $\frac{1}{2}d^*d|\psi|^2 = \operatorname{Re}\langle \psi, \widetilde{\nabla}_A^*\widetilde{\nabla}_A\psi\rangle - |\widetilde{\nabla}_A\psi|^2$. Also $\rho(\lambda)(\alpha,\beta) = (i\alpha, -i\beta)$. So taking the real inner product of (11.9) with ψ, we obtain

$$0 \ge \frac{1}{2}d^*d|\psi|^2 + |\widetilde{\nabla}_A\psi|^2 + \frac{r}{2}\left(|\psi|^4 - (1 + c_1 r^{-1})|\psi|^2\right),$$

where the constant c_1 depends only on the minimum of the scalar curvature s and the maximum of $|\overline{\omega}|$. At a point where $|\psi|$ is maximized, the maximum principle tells us that $d^*d|\psi|^2 \ge 0$. Therefore

$$|\psi|^2 \le 1 + c_1 r^{-1}$$

where $|\psi|$ is maximized, and consequently everywhere. This implies (11.5).

To prove (11.6), we consider separately the E and $K^{-1}E$ components of equation (11.9). These are

$$0 = \nabla_A^*\nabla_A\alpha + \frac{r}{2}(|\alpha|^2 + |\beta|^2 - 1)\alpha + f_2\alpha + f_3\beta + f_4\nabla_A\alpha + f_5\nabla_A\beta,$$
$$0 = \nabla_A^*\nabla_A\beta + \frac{r}{2}(|\alpha|^2 + |\beta|^2 + 1)\beta + f_2'\alpha + f_3'\beta + f_4'\nabla_A\alpha + f_5'\nabla_A\beta$$

where $f_2, \ldots, f_5'$ are bundle maps between E and $K^{-1}E$ depending only on the Riemannian metric on Y. Taking the inner products of these equations with α and β respectively, we obtain

$$0 \ge \frac{1}{2}d^*d|\alpha|^2 + |\nabla_A\alpha|^2 + \frac{r}{2}(|\alpha|^2 + |\beta|^2 - 1)|\alpha|^2 - c_2(|\alpha|^2 + |\alpha||\beta| + |\alpha||\nabla_A\beta|),$$
$$0 \ge \frac{1}{2}d^*d|\beta|^2 + |\nabla_A\beta|^2 + \frac{r}{2}(|\alpha|^2 + |\beta|^2 + 1)|\beta|^2 - c_2(|\beta|^2 + |\alpha||\beta| + |\nabla_A\alpha||\beta|)$$

where c_2 is a constant depending only on the Riemannian metric on Y. Now let c_3 and c_4 be large constants. Adding c_3r^{-1} times the first inequality to the second gives

$$\begin{aligned} 0 \geq \frac{1}{2}d^*d\left(|\beta|^2 - \frac{c_3}{r}(1-|\alpha|^2) - \frac{c_4}{r^2}\right) + \frac{r}{2}\left(|\beta|^2 - \frac{c_3}{r}(1-|\alpha|^2) - \frac{c_4}{r^2}\right) \\ + \frac{c_4}{2r} + \frac{c_3}{r}|\nabla_A\alpha|^2 + \frac{c_3}{2}((1-|\alpha|^2)^2 + |\alpha|^2|\beta|^2) + |\nabla_A\beta|^2 + \frac{r}{2}(|\alpha|^2+|\beta|^2)|\beta|^2 \\ -c_2(|\beta|^2 + |\alpha||\beta| + |\nabla_A\alpha||\beta|) - \frac{c_2c_3}{r}(|\alpha|^2 + |\alpha||\beta| + |\alpha||\nabla_A\beta|). \end{aligned}$$

Inspection of the above inequality using (11.5) shows that the sum of the terms on the second and third line is nonnegative provided that $r >> c_4 >> c_3 >> 0$. (The $|\nabla_A\alpha||\beta|$ term is a bit tricky and requires consideration of several cases.) The maximum principle now implies (11.6).

The estimates (11.7) and (11.8) are then obtained using elliptic regularity, as explained in [62, §6.2]. □

As a quick corollary, we have:

Proof of Lemma 10.4. Suppose $c_1(E) \neq 0$ and (A,ψ) is a solution to the equations (10.5). Since $c_1(E) \neq 0$ there exists a point where $\alpha = 0$. At this point $1-|\psi| = 1-|\beta|$, and it follows from (11.6) that this is at least $1 - cr^{-1/2}$ for some constant c if r is sufficiently large. □

We also need an a priori estimate on the connection. Recall that A_1 denotes a reference connection on E. Write a general connection on E as $A = A_1 + \hat{a}$ where $\hat{a}$ is an imaginary-valued 1-form on Y. Also recall that $\mathcal{E} = \mathcal{E}(A)$ denotes the energy functional (10.8).

Lemma 11.3. [62, Lemma 2.4] *Let* (A,ψ) *be a solution to the equations* (10.5). *Then by a gauge transformation one can arrange that*

$$|\hat{a}| \leq c(r^{2/3}|\mathcal{E}|^{1/3} + 1) \tag{11.10}$$

where c is an r-independent constant.

Proof. Recall the Hodge decomposition

$$\Omega^1 = \mathcal{H}^1 \oplus d\Omega^0 \oplus d^*\Omega^2$$

where $\mathcal{H}^1$ denotes the space of harmonic 1-forms. Let $\mathfrak{h} \in \mathcal{H}^1$ denote the harmonic component of $-i\hat{a}$. A gauge transformation $g : Y \to S^1$ subtracts $2g^{-1}dg$ from $\hat{a}$. Thus by a gauge transformation one can eliminate the $d\Omega^0$ component of $-i\hat{a}$, and one can shift the harmonic component $\mathfrak{h}$ by any element of $4\pi H^1(Y;\mathbb{Z})$. So we can apply a (unique) gauge transformation to arrange that $\hat{a}$ is co-closed, and $\mathfrak{h}$ sends every element of some chosen basis of $H_1(Y;\mathbb{Z})$ mod torsion to a number in the interval $[0,4\pi)$. In particular $|\mathfrak{h}|$ is bounded from above by a constant depending only on the Riemannian metric.

We now have

$$\hat{a} = \mathfrak{h} + (d+d^*)^{-1}(F_A - F_{A_1})$$

where $(d+d^*)^{-1}$ denotes the Green's function for $d+d^*$, and $|\mathfrak{h} - (d+d^*)^{-1}F_{A_1}|$ is bounded by an r-independent constant. Standard estimates for the Green's function then give

$$|\hat{a}(x)| \le c\left(1 + \int_Y \frac{|F_A|}{\operatorname{dist}(x,\cdot)^2}\right). \tag{11.11}$$

It follows from equation (11.3) and the a priori estimates (11.5) and (11.6) that

$$*F_A = -ir|1 - |\alpha|^2|\lambda + O(1). \tag{11.12}$$

In particular, it follows from (11.12) that the integrand in the energy functional (10.8) satisfies

$$i\lambda \wedge F_A = |F_A| + O(1). \tag{11.13}$$

(We are omitting the volume form from the notation here.)

To analyze the integral in (11.11), divide the integration domain into a ball of radius ρ centered at x, and its complement. Since $|F_A|$ is bounded by a constant multiple of r, the integral over the ball is bounded by a constant times $r\rho$. On the other hand, by (11.13) the integral over the complement of the ball is bounded by a constant times $|\mathcal{E}|\rho^{-2}$, plus some other constant times ρ^{-2}. Thus

$$|\hat{a}(x)| \le c\left(1 + r\rho + \frac{|\mathcal{E}| + c}{\rho^2}\right) \tag{11.14}$$

for some new constant $c > 0$. Now take $\rho = r^{-1/3}|\mathcal{E}|^{1/3}$. Note that $\rho = O(1)$. The estimate (11.14) then implies (11.10). □

We can now give:

Proof of Lemma 10.10. Recall that we are assuming that $2c_1(E) + c_1(K^{-1})$ is torsion, so that the Chern-Simons functional is gauge invariant. Thus we can assume that the connection is the one provided by Lemma 11.3. It then follows from the definition of the Chern-Simons functional (10.7) that

$$|cs(A)| \le \int_Y c(r^{2/3}\mathcal{E}^{1/3} + 1)(|F_A| + c).$$

On the other hand, by equation (11.13) we have $\int_Y |F_A| = \mathcal{E} + O(1)$. The lemma follows. □

11.4. **Existence of a Reeb orbit.** We can now sketch the proof of Taubes's Theorem 10.3. To start, we will outline the proof of the following:

Claim. *Fix $\delta, C > 0$. Then there exists $c > 0$ such that the following holds. Suppose $r > c$. Let (A, ψ) be a solution to the perturbed Seiberg-Witten equations* (10.5) *with $\sup_Y(1 - |\psi|) > \delta$ and $\mathcal{E}(A) < C$. Then there exists $n < c$ and a finite collection of arcs $\{\gamma_i\}$ in Y indexed by $i \in \mathbb{Z}/n$ such that:*

- *Each γ_i is tangent to the Reeb vector field.*
- *The length of each arc γ_i is at most c. Also the sum of the lengths of the arcs is at least c^{-1}.*
- *The distance between the endpoint of γ_i and the starting point of γ_{i+1} is at most $cr^{-1/2}$.*

It follows from the claim that given a sequence (r_n, A_n, ψ_n) of solutions to the perturbed Seiberg-Witten equations satisfying conditions (1) and (2) in Theorem 10.3, one can pass to a subsequence such that the corresponding chains of arcs γ_i converge to a Reeb orbit.

To prove the claim, let (A,ψ) be a solution to the perturbed Seiberg-Witten equations with $\sup_Y(1-|\psi|)>\delta$ and $\mathcal{E}(A)<C$. Write $\psi=(\alpha,\beta)$ as usual.

It follows from the Dirac equation (11.4) and the a priori estimates (11.5), (11.6), and (11.8) that

$$|\nabla_{A,R}\alpha| \le c_0, \tag{11.15}$$

$$|\overline{\partial}_A\alpha| \le c_0 \tag{11.16}$$

where c_0 is an r-independent constant.

It also follows from (11.5) and (11.6) that if r is sufficiently large, then there exist points where $1-|\alpha|^2>\delta$. Assume that it is and choose such a point p. Also fix r-independent constants $\rho_1,\rho_2>0$ (to be specified more later). We can now choose an embedding of the disc D of radius $\rho_1 r^{-1/2}$ into Y mapping the center of the disc to p, such that the disc is orthogonal to the Reeb vector field at p, and the induced metric on the disc is Euclidean to first order at the origin. Denote the coordinates on the disc by $z=x+iy$. We can uniquely extend the embedding of the disc D to a map from the cylinder $D\times[0,\rho_2]$ to Y such that the derivative of the interval coordinate t maps to the Reeb vector field. We can assume that the map of the cylinder is an embedding (for $r>>\rho_1$), otherwise we already know that the claim is true. We then take the arc γ_1 to be the image of $\{0\}\times[0,\rho_2]$ under this embedding.

Now let $\chi:[0,1]\to\mathbb{R}$ be a cutoff function which is 1 on $[0,1/3]$, monotone decreasing, and 0 on $[2/3,1]$. For $t\in[0,\rho_2]$ define

$$L(t) := r\int_{D\times\{t\}}(1-|\alpha|^2)\chi(|z|/\rho_1). \tag{11.17}$$

In view of (11.12) and (11.13), this is the contribution to the energy $\mathcal{E}(A)$ from the disc $D\times\{t\}$ (weighted by the cutoff function), up to an error of order $r^{-1/2}$. Now it follows from (11.7) that $1-|\alpha|^2>\delta/2$ on a disc of radius order $r^{-1/2}$ in $D\times\{0\}$, so

$$L(0)\ge c_1 \tag{11.18}$$

for some r-independent constant $c_1>0$. Also, differentiating (11.17) with respect to t and using the a priori estimates (11.5) and (11.15) shows that

$$\left|\frac{dL(t)}{dt}\right| < c_2 \tag{11.19}$$

for some r-independent constant c_2. In particular, if $\rho_2>0$ is chosen sufficiently small, then $L(\rho_2)\ge c_1/2$, and the contribution to the energy $\mathcal{E}(A)$ from the cylinder is bounded from below by an r-independent constant.

To complete the proof of the claim, it is enough to show that the disk $D\times\{\rho_2\}$ contains a point where $1-|\alpha|^2>\delta$. We can then repeat the above argument to construct a sequence of cylinders in Y, and define the arcs γ_i to be the core intervals of these cylinders. There is an upper bound to the number of cylinders one can construct before a new cylinder overlaps an old cylinder, because each cylinder contributes at least c_2 to the energy, while the total energy is bounded from above by C. When a new cylinder overlaps an old one, we then obtain the desired chain of arcs.

In fact, if the constant ρ_1 is sufficiently large, then α must have a zero on $D\times\{\rho_2\}$. To see why, rescale $D\times\{\rho_2\}$ by the map $z\mapsto r^{1/2}z$, so as to identify $D\times\{\rho_2\}$ with the disc of radius ρ_1. The restriction of (A, α) to the latter now satisfies

$$\begin{aligned}\overline{\partial}_A\alpha &= O(r^{-1/2}),\\ *F_A &= -i(1-|\alpha|^2) + O(r^{-1/2}).\end{aligned}$$

as a result of (11.16), (11.3), (11.5) and (11.6). That is, (A, α) satisfies the vortex equations (11.1) on the disk, up to an error of order $r^{-1/2}$. Moreover, if ρ_1 is sufficiently large, then there must exist a large $\rho' < \rho_1$ such that $1 - |\alpha|^2 < \delta$ on the circle of radius ρ'; otherwise the arguments giving (11.18) and (11.19) would imply that the energy $\mathcal{E}(A) > C$. A compactness argument then shows that if r is sufficiently large, then (A, α) not only approximately solves the vortex equations, but is C^0-approximated by an actual solution (A', α') to the vortex equations on $\mathbb{C}$ with finite energy, where $|\alpha'|$ is close to 1 outside of the disc of radius r'. It now follows from Lemma 11.1 that α must have a zero in the disc of radius ρ'.

This completes the sketch of the proof of the claim, and hence the existence of a Reeb orbit. A more careful version of this argument keeping track of all of the zero set of α, see [62, §6.4], shows that in fact there exists a nonempty orbit set Poincaré dual to $c_1(E)$.

12. Beyond the Weinstein conjecture

With the three-dimensional Weinstein conjecture now proved, there are various directions in which one might try to generalize.

12.1. **Improved lower bounds.**

The Weinstein conjecture gives a lower bound of one on the number of embedded Reeb orbits for a contact form λ on a closed oriented 3-manifold Y. Can one improve this lower bound?

As we saw in §3.3, the standard contact form on an irrational ellipsoid has exactly two embedded Reeb orbits. Also one can take the quotient of the ellipsoid by a $\mathbb{Z}/p$ action on $\mathbb{C}^2$ that rotates each $\mathbb{C}$ factor, to obtain a contact form on a lens space with exactly two embedded Reeb orbits. Perhaps surprisingly, these are the only examples known to us of contact forms on closed (connected) 3-manifolds with only finitely many embedded Reeb orbits. It is shown in [30] that for a large class of contact forms on S^3 there are either two or infinitely many embedded Reeb orbits. For other 3-manifolds one can ask:

Question. *If Y is a closed oriented connected 3-manifold other than a sphere or a lens space, then does every contact form on Y have infinitely many embedded Reeb orbits?*

As mentioned in §3.6, Colin-Honda [15] used linearized contact homology to prove the existence of infinitely many embedded Reeb orbits for many cases of contact structures supported by open books with pseudo-Anosov monodromy. Also it is proved in [36], using the isomorphism between Seiberg-Witten Floer homology and embedded contact homology (see §7.4), that if Y is a closed oriented 3-manifold other than a lens space, then any contact form on Y with all Reeb orbits nondegenerate (see §6.1) has at least three embedded Reeb orbits. Nonetheless there remains a substantial gap between what we can prove and what seems to be true.

12.2. **More general vector fields.** Next one might try to prove the existence of closed orbits for somewhat more general vector fields than Reeb vector fields. For example, inspection of the proof of the Weinstein conjecture for a compact hypersurface Y in $\mathbb{R}^{2n}$ of contact type, see §2.1, shows that the contact type hypothesis can be replaced by the weaker assumption that the hypersurface is "stable". A hypersurface Y in a symplectic manifold (M, ω) is called *stable* if it has a neighborhood N with an identification $N \simeq (-\delta, \delta) \times Y$ sending Y to $\{0\} \times Y$, such that the characteristic foliations on $\{\epsilon\} \times Y$ are conjugate for all $\epsilon \in (-\delta, \delta)$. It turns out (see [13, Lem. 2.3]) that a compact hypersurface Y is stable if and only if $\omega|_Y$ is part of a "stable Hamiltonian structure" on Y. A *stable Hamiltonian structure* on a $2n-1$ dimensional oriented manifold Y is a pair (λ, ω) where λ is a 1-form on Y and ω is a closed 2-form on Y, such that $\lambda \wedge \omega^{n-1} > 0$ and $\operatorname{Ker}(\omega) \subset \operatorname{Ker}(d\lambda)$. A stable Hamiltonian structure determines a "Reeb vector field" R characterized by $\omega(R, \cdot) = 0$ and $\lambda(R) = 1$. A contact form is a special case of this in which $\omega = d\lambda$.

It is shown in [36], again using the isomorphism between Seiberg-Witten Floer homology and embedded contact homology, that for any closed oriented connected 3-manifold Y that is not a T^2-bundle over S^1, for any stable Hamiltonian structure on Y, the associated Reeb vector field has a closed orbit. The same conclusion is proved without using Seiberg-Witten theory, but under some additional hypotheses, by Rechtman [54]. In general, however, it remains unclear what exactly one needs to assume about a vector field in order to guarantee the existence of a closed orbit.

12.3. **Non-unique ergodicity.** For even more general vector fields, one can try to prove weaker statements than the existence of a closed orbit. For example let Y be a closed oriented 3-manifold with a volume form Ω, and let V be a smooth vector field on Y. The vector field V generates a 1-parameter family of diffeomorphisms $\phi_t : Y \to Y$. Assume that V is divergence free, meaning that each ϕ_t preserves the volume form, or equivalently $d(\imath_V \Omega) = 0$. A measure σ on Y is said to be "V-invariant" if $(\phi_t)_* \sigma = \sigma$ for each $t \in \mathbb{R}$. The vector field V is called *uniquely ergodic* if the only V-invariant measures on Y are real multiples of the volume form. Note that if V has a closed orbit then it is not uniquely ergodic. More generally, one could look for conditions on V that guarantee that it is not uniquely ergodic.

Taubes [65] establishes such a condition as follows. Call V "exact" if $\imath_V \Omega = d\lambda$ for some 1-form λ. In this case define the "self-linking" of V, cf [4], to be

$$s_V := \int_Y \lambda \wedge \imath_V \Omega. \tag{12.1}$$

By Stokes theorem, this does not depend on the choice of λ. For example, if $\lambda(V) > 0$ everywhere, then the integrand in (12.1) is positive, so $s_V > 0$. In this case λ is a contact form and V is a positive multiple of its Reeb vector field. More generally one has:

Theorem 12.1 (Taubes [65])**.** *Let Y be a closed oriented 3-manifold with a volume form Ω. Let V be an exact vector field on Y. Suppose that $s_V \neq 0$. Then V is not uniquely ergodic.*

Taubes proves this similarly to the Weinstein conjecture. One considers the perturbed Seiberg-Witten equations (10.5) with λ replaced by $*d\lambda$, and proves a modification of Theorem 10.3 in which hypothesis (2) in that theorem is dropped. A nontrivial V-invariant measure is obtained as a limit of a subsequence of the sequence of measures

$$\sigma_n := \mathcal{E}(A_n)^{-1} r_n (1 - |\alpha_n|^2)\Omega$$

for an appropriately chosen sequence (r_n, A_n, ψ_n).

12.4. **The Arnold chord conjecture.** There is also the following "relative" version of the Weinstein conjecture. Let (Y, ξ) be a contact 3-manifold. A *Legendrian knot* is a knot $L \subset Y$ such that $T_pL \subset \xi_p$ for every $p \in L$. Now choose a contact form λ with $\operatorname{Ker}(\lambda) = \xi$. A *Reeb chord* of L is a path $\gamma : [0, T] \to Y$ for some $T > 0$ such that $\gamma(0), \gamma(T) \in L$ and γ is a flow line of the Reeb vector field, ie $\gamma'(t) = R(\gamma(t))$. For example, any Legendrian knot L in $\mathbb{R}^3$ with the standard contact form (3.1) must have a Reeb chord, because the projection of L to the x, y plane must have area zero, so it must have a crossing. A version of the Arnold chord conjecture asserts that for any Legendrian knot in a closed 3-manifold with a contact form, there exists a Reeb chord. This has been proved for the standard contact structure on S^3 [47], and for Legendrian unknots in tight contact 3-manifolds satisfying certain assumptions [1], but it seems that not too much is known about the general case.

12.5. **Higher dimensions.** In higher dimensions, although the Weinstein conjecture is known for compact hypersurfaces of contact type in $\mathbb{R}^{2n}$ and for some other cases, in general it is wide open. The techniques used in Taubes's proof are special to three dimensions. In particular no good analogue of the Seiberg-Witten invariants is currently known in dimensions greater than four. In addition there is no obvious higher-dimensional analogue of embedded contact homology. (In higher dimensions one expects that generically all non-multiply-covered holomorphic curves of the relevant index are embedded, compare [49].) One can still use holomorphic curves in higher dimensions to define linearized contact homology and related invariants from symplectic field theory [20]. It is unclear if these invariants are sufficient to prove the Weinstein conjecture in all cases. It is shown in [3] that the Weinstein conjecture holds for "PS-overtwisted" contact structures, which are a certain higher dimensional analogue of overtwisted contact structures, cf §3.2. It is also shown in [12] that symplectic field theory invariants recover this fact. In any case we will close with:

Question. *Is there a proof of the Weinstein conjecture in three dimensions using only holomorphic curves (and no Seiberg-Witten theory)? What about in higher dimensions?*

References

1. C. Abbas, *The chord problem and a new method of filling by pseudoholomorphic curves*, Int. Math. Res. Not. **2004**, 913–927.
2. C. Abbas, K. Cieliebak, and H. Hofer, *The Weinstein conjecture for planar contact structures in dimension three*, Comment. Math. Helv. **80** (2005), 771-793.
3. P. Albers and H. Hofer, *On the Weinstein conjecture in higher dimensions*, Comment. Math. Helv. **84** (2009), 429–436.
4. V. Arnold and B. Khesin, *Topological methods in hydrodynamics*, Applied Mathematical Sciences 125, Springer Verlag, 1998.
5. D. Auroux, *La conjecture de Weinstein en dimension 3 [d'après C.H. Taubes]*, Séminaire Bourbaki, 2008-2009, no. 1002.
6. N. Berline, E. Getzler, and M. Vergne, *Heat kernels and Dirac operators*, Springer Verlag, 1990.
7. R. Bott, *Morse theory indomitable*, IHES Publ. Math. (1988) 99–114.

8. F. Bourgeois, *Odd dimensional tori are contact manifolds*, Int. Math. Res. Not. **2002**, no. 30, 1571–1574.
9. F. Bourgeois, *A Morse-Bott approach to contact homology*, Symplectic and contact topology: interactions and perspectives, 55–77, Fields Inst. Commum. 35, AMS, 2003.
10. F. Bourgeois, K. Cieliebak, and T. Ekholm, *A note on Reeb dynamics on the tight 3-sphere*, J. Modern Dynamics **1** (2007), 597–613.
11. F. Bourgeois, T. Ekholm, and Y. Eliashberg, *A Legendrian surgery long exact sequence for linearized contact homology*, in preparation.
12. F. Bourgeois and K. Niederkrüger, *Towards a good definition of algebraically overtwisted*, arXiv:0709.3415
13. K. Cieliebak and J. Mohnke, *Compactness of punctured holomorphic curves*, J. Symplectic Geom. **3** (2005), 589–654.
14. V. Colin, E. Giroux, and K Honda, *Finitude homotopique et isotopique des structures de contact tendues*, Publ. Math. IHES **109** (2009), 245–293.
15. V. Colin and K. Honda, *Reeb vector fields and open book decompositions*, arXiv:0809.5088.
16. S.K. Donaldson, *The Seiberg-Witten equations and 4-manifold topology*, Bull. AMS **33** (1996), 45-70.
17. S.K. Donaldson, *Topological field theories and formulae of Casson and Meng-Taubes*, Proceedings of the Kirbfest (Berkeley, CA, 1998), 87–102, Geom. Topol. Monogr. **2**, Geom. Topol. Publ., 1999.
18. Y. Eliashberg, *Classification of overtwisted contact structures on 3-manifolds*, Invent. Math. **98** (1989), 623–637.
19. Y. Eliashberg, *Filling by holomorphic discs and its applications*, Geometry of low-dimensional manifolds, 2 (Durham, 1989), 45–67, London Math. Soc. Lecture Note Ser. **151**, Cambridge Univ. Press, 1990.
20. Y. Eliashberg, A. Givental, and H. Hofer, *Introduction to symplectic field theory*, GAFA 2000, Special Volume, Part II, 560–673.
21. J. Etnyre, *Introductory lectures on contact geometry*, Proc. Sympos. Pure Math. 71, pp. 81–107, Amer. Math. Soc. 2003.
22. J. Etnyre, *Lectures on open book decompositions and contact structures*, Floer homology, gauge theory, and low-dimensional topology, 103–141, Clay Math. Proc. 5, AMS, 2006.
23. O. Garcia-Prada, *A direct existence proof for the vortex equations over a compact Riemann surface*, Bull. Londn Math. Soc. **26** (1994), 88–96.
24. H. Geiges, *An introduction to contact topology*, Cambridge University Press, 2008.
25. V. Ginzburg, *The Weinstein conjecture and the theorems of nearby and almost existence*, The Breadth of Symplectic and Poisson Geometry, Festschrift in Honor of Alan Wenstein, Birkhäuser (2005), 139–172.
26. V. Ginzburg and B. Gürel, *A C^2-smooth counterexample to the Hamiltonian Seifert conjecture in $\mathbb{R}^4$*, Ann. of Math. **158** (2003), 953–976.
27. E. Giroux, *Géometrie de contact: de la dimension trois vers les dimensions supérieures*, Proceedings of the ICM, Beijing, 2002, Vol. II, 405–414.
28. J. Harrison, *C^2 counterexamples to the Seifert conjecture*, Topology **27** (1998), 249–278.
29. H. Hofer, *Pseudoholomorphic curves in symplectizations with applications to the Weinstein conjecture in dimension three*, Invent. Math. **114** (1993), 515–563.
30. H. Hofer, K. Wysocki, and E. Zehnder, *Finite energy foliations of tight three-spheres and Hamiltonian dynamics*, Ann. of Math. **157** (2003), 125-255.
31. H. Hofer and E. Zehnder, *Symplectic invariants and Hamiltonian dynamics*, Birkhäuser, 1994.
32. M. Hutchings, *The embedded contact homology index revisited*, arXiv:0805.1240, to appear in the Yashafest proceedings.
33. M. Hutchings and M. Sullivan, *Rounding corners of polygons and the embedded contact homology of T^3*, Geom. Topol. **10** (2006), 269–266.
34. M. Hutchings and C.H. Taubes, *An introduction to the Seiberg-Witten equations on symplectic four-manifolds*, Symplectic geometry and topology (Park City, UT, 1997), 103–142, AMS, 1999.
35. M. Hutchings and C.H. Taubes, *Gluing pseudoholomorphic curves along branched covered cylinders I*, J. Symplectic Geom. **5** (2007), 43–137.
36. M. Hutchings and C.H. Taubes, *The Weinstein conjecture for stable Hamiltonian structures*, Geom. Topol. **13** (2009), 901–941.

37. A. Jaffe and C.H. Taubes, *Vortices and monopoles. Structure of static gauge theories*, Progress in Physics, 2. Birkhäuser, Boston, 1980.
38. P.B. Kronheimer and T.S. Mrowka, *Monopoles and 3-manifolds*, Cambridge University Press, 2008.
39. G. Kuperberg, *A volume-preserving counterexample to the Seifert conjecture*, Comment. Math. Helv. **71** (1996), 70–97.
40. G. Kuperberg and K. Kuperberg, *Generalized counterexamples to the Seifert conjecture*, Ann. of Math. **143** (1996), 547–576.
41. K. Kuperberg, *A smooth counterexample to the Seifert conjecture*, Ann. of Math. **140** (1994), 723–732.
42. H. Lawson and M. Michelson, *Spin geometry*, Princeton University Press, 1989.
43. Y-J. Lee and C.H. Taubes, *Periodic Floer homology and Seiberg-Witten Floer cohomology*, arXiv:0906.0383.
44. MSRI Hot Topics Workshop, *Contact structures, dynamics and the Seiberg-Witten equations in dimension 3*, June 2008, videos at www.msri.org.
45. D. McDuff and D. Salamon, *Introduction to symplectic topology*, 2nd edition, Oxford University Press, 1998.
46. G. Meng and C.H. Taubes, $\underline{SW}$ *=Milnor torsion*, Math. Res. Lett. **3** (1996), 661–674.
47. K. Mohnke, *Holomorphic disks and the chord conjecture*, Ann. of Math. **154** (2001), 219–222.
48. J. Morgan, *The Seiberg-Witten equations and applications to the topology of smooth four-manifolds*, Mathematical Notes 44, Princeton Univ. Press, 1996.
49. Y-G. Oh and K. Zhu, *Embedding property of J-holomorphic curves in Calabi-Yau manifolds for generic J*, arXiv:0805.3581.
50. P. Ozsváth and Z. Szabó, *Holomorphic disks and topological invariants for closed three-manifolds*, Ann. of Math. **159** (2004), 1027–1158.
51. P. Ozsváth and Z. Szabó, *Holomorphic triangles and invariants for smooth four-manifolds*, Adv. Math. **202** (2006), 326–400.
52. T. Perutz, *Lagrangian matching invariants for fibred four-manifolds: I*, Geom. Topol. **11** (2007), 759–828.
53. P. Rabinowitz, *Periodic solutions of Hamiltonian systems*, Comm. Pure Appl. Math **31** (1978), 157–184.
54. A. Rechtman, *Existence of periodic orbits for geodesible vector fields on closed 3-manifolds*, arXiv:0904.2719.
55. J. Robbin and D. Salamon, *The spectral flow and the Maslov index*, Bull. London Math. Soc. **27** (1995), 1–33.
56. M. Schwarz, *Morse homology*, Progress in Mathematics, Birkhäuser, 1993.
57. M. Schwarz, *On the action spectrum for closed symplectially aspherical manifolds*, Pac. J. Math. **193** (2000), 419–461.
58. P. Schweizer, *Counterexamples to the Seifert conjecture and opening closed leaves of foliations*, Ann. of Math. **100** (1974), 386–400.
59. C.H. Taubes, *Gr $\Rightarrow$ SW: from pseudo-holomorphic curves to Seiberg-Witten solutions*, J, Diff. Geom. **51** (1999), 203-334.
60. C.H. Taubes, *SW $\Rightarrow$ Gr: from the Seiberg-Witten equations to pseudo-holomorphic curves*, in "Seiberg-Witten and Gromov invariants for symplectic 4-manifolds", Int. Press, 2000.
61. C.H. Taubes, *Seiberg-Witten and Gromov invariants for symplectic 4-manifolds*, First International Press Lecture Series 2, International Press, 2000.
62. C.H. Taubes, *The Seiberg-Witten equations and the Weinstein conjecture*, Geom. Topol. **11** (2007), 2117–2202.
63. C.H. Taubes, *The Seiberg-Witten equations and the Weinstein conjecture II: More closed integral curves for the Reeb vector field*, Geom. Topol. **13** (2009), 1337-1417.
64. C.H. Taubes, *Asymptotic spectral flow for Dirac operators*, Comm. Anal. Geom. **15** (2007), 569–587.
65. C.H. Taubes, *An observation concerning uniquely ergodic vector fields on 3-manifolds*, arXiv:0811.3983.
66. C.H. Taubes, *Embedded contact homology and Seiberg-Witten Floer homology I*, arXiv:0811.3985.
67. C.H. Taubes, *Embedded contact homology and Seiberg-Witten Floer homology II-IV*, preprints, 2008.

68. C.H. Taubes, *Embedded contact homology and Seiberg-Witten Floer homology V*, preprint, 2008.
69. V. Turaev, *A combinatorial formulation for the Seiberg-Witten invariants of* 3*-manifolds*, Math. Res. Lett. **5** (1998), 583–598.
70. M. Usher, *Vortices and a TQFT for Lefschetz fibrations on 4-manifolds*, Algebr. Geom. Topol. **6** (2006), 1677-1743.
71. C. Viterbo, *A proof of Weinstein's conjecture in* $\mathbb{R}^{2n}$, Annales de l'Institut Henri Poincaré-Anal. Non Linéaire, **4** (1987), 337–356.
72. A. Weinstein, *Periodic orbits for convex Hamiltonian systems*, Ann. of Math. **108** (1978), 507–518.
73. A. Weinstein, *On the hypothesis of Rabinowitz' periodic orbit theorems*, J. Differential Equations **33** (1979), 353–358.
74. E. Witten, *From superconductors and four-manifolds to weak interactions*, Bull. AMS **44** (2007), 361–391.
75. E. Zehnder, *Remarks on periodic solutions on hypersurfaces*, Periodic solutions of hamiltonian systems and related topics, Reidel Publishing Co. (1987), 267–279.

Mathematics Department, 970 Evans Hall, University of California, Berkeley CA 94720

E-mail address: `hutching@math.berkeley.edu`