This is to certify that the

dissertation entitled

Torsion, TQFT, and Seiberg-Witten Invariants

of Three-Manifolds

presented by

Thomas E. Mark

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Mathematics

Ronald Fintushel
Major professor

Date April 20, 2000

TORSION, TQFT, AND SEIBERG-WITTEN INVARIANTS OF THREE-MANIFOLDS

By

Thomas E. Mark

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

2000

ABSTRACT

TORSION, TQFT, AND SEIBERG-WITTEN INVARIANTS OF THREE-MANIFOLDS

By

Thomas E. Mark

We prove a conjecture of Hutchings and Lee relating the Seiberg-Witten invariants of a closed 3-manifold with $b_1 > 1$ to an invariant that "counts" gradient flow lines—including closed orbits—of a circle-valued Morse function on the manifold. The proof is based on a method described by Donaldson for computing the Seiberg-Witten invariants of 3-manifolds by making use of a "topological quantum field theory."

ACKNOWLEDGEMENTS

I gratefully acknowledge the unflagging support and assistance of my advisor Ronald Fintushel, whose help has not only led to the completion of this work, but also contributed immeasurably to the continuation of my career. I also thank Peter Ozsváth for his generosity in sharing his great knowledge of gauge theory. Without his instruction the appendix to this work would not have been written.

Contents

1 Introduction

Recent years have seen rapid progress in the theory of three- and four-dimensional smooth manifolds, fueled in good part by the introduction of Seiberg-Witten gauge theory. One particularly important theorem in the Seiberg-Witten theory of three-manifolds is that of Meng and Taubes [11], relating the gauge-theoretic invariants to the algebraic-topological Milnor torsion. The proof outlined in [11] is based on an explicit reduction of the case of a general manifold to that of a three-manifold obtained from zero-sugery on a knot, then use of analytical cut-and-paste arguments to show the result in this case. Work of Hutchings and Lee [5, 6] has pointed out a way to prove this theorem in a manner distinct from the original: in particular, their approach relies first on proving a relationship between the Milnor torsion of a three-manifold and an invariant that they call I which "counts" gradient flow lines of a Morse function on the manifold having values in S^1, then relating I to the Seiberg-Witten invariant. The first part of this program was completed in [5, 6]. The second part, which can be seen as an analogue in three dimensions of the theorem of Taubes relating the Seiberg-Witten and Gromov invariants of a symplectic four-manifold, is stated as a conjecture in [5] and [6]; it is the goal of the present work to give a proof of a large part of that conjecture–in particular, enough to give the alternate proof of the Meng-Taubes theorem.

1.1 Background

In order to state the main theorem we will prove, we first need some preliminary definitions and notation. We begin with the notion of the torsion of an acyclic chain complex; basic references for this material include [15, 12]; see also [9].

1.1.1 Torsion

Suppose $0 \to V' \to V \to V'' \to 0$ is an exact sequence of finite-dimensional vector spaces over a field k. By a *volume element* ω for a vector space W of dimension n we mean a choice of nonzero element $\omega \in \Lambda^n W$. It is easy to show that a choice of volume element on any two of V, V', V'' induces a volume element on the third. In particular, for volumes ω' on V' and ω'' on V'', the induced volume on V will be written $\omega'\omega''$. If ω_1, ω_2 are two volume elements for V, then we can write $\omega_1 = c\omega_2$ for some nonzero element $c \in k$; we will write $c = \omega_1/\omega_2$. More generally, let $\{C_i\}_{i=1}^n$ be a complex of vector spaces with differential $\partial : C_i \to C_{i-1}$, and let us assume that C_* is acyclic, i.e., $H_*(C_*) = 0$. Suppose that each C_i comes equipped with a volume element ω_i, and choose volumes ν_i arbitrarily on each image ∂C_i. From the exact sequence

$$0 \to C_n \to C_{n-1} \to \partial C_{n-1} \to 0$$

define $\tau_{n-1} = \omega_n \nu_{n-1}/\omega_{n-1}$. For $i = 1, \ldots, n-2$ use the exact sequence

$$0 \to \partial C_{i+1} \to C_i \to \partial C_i \to 0$$

to define $\tau_i = \nu_{i+1}\nu_i/\omega_i$. We then define the *torsion* $\tau(C_*, \{\omega_i\}) \in k \setminus \{0\}$ of the (volumed) complex C_* to be

$$\tau(C_*) = \prod_{i=1}^{n-1} \tau_i^{(-1)^{i+1}} \tag{1}$$

It can be seen that this definition does not depend on the choice of ν_i. Note that in the case that our complex consists of just two vector spaces,

$$C = 0 \to C_1 \xrightarrow{\partial} C_0 \to 0,$$

we have that $\tau(C) = \det(\partial)$. We extend the definition of $\tau(C)$ to non-acyclic complexes by setting $\tau(C) = 0$ in this case.

As a slight generalization, we can allow the chain groups C_i to be finitely generated free modules over an integral domain K with fixed ordered bases rather than vector spaces with fixed volume elements, as follows. Write

$Q(K)$ for the field of fractions of K, then form the complex of vector spaces $Q(K) \otimes_K C_i$. The bases for the C_i naturally give rise to bases, and hence volumes, for $Q(K) \otimes_K C_i$. We understand the torsion of the complex of K-modules C_i to be the torsion of this latter complex, and it is therefore a nonzero element of the field $Q(K)$.

Suppose now that X is a connected, compact, oriented smooth manifold with a given CW decomposition. Following [15], suppose $\varphi : \mathbb{Z}[H_1(X;\mathbb{Z})] \to K$ is a ring homomorphism into an integral domain K. The universal abelian cover $\tilde{X}$ has a natural CW decomposition lifting the given one on X, and the action of the deck transformation group $H_1(X;\mathbb{Z})$ naturally gives the cell chain complex $C_*(\tilde{X})$ the structure of a $\mathbb{Z}[H_1(X;\mathbb{Z})]$-module. As such, $C_i(\tilde{X})$ is free of rank equal to the number of i-cells of X. We can then form the twisted complex $C^{\varphi}_*(X) = K \otimes_{\varphi} C_*(\tilde{X})$ of K-modules. We choose a sequence e of cells of $\tilde{X}$ such that over each cell of X there is exactly one element of e; this gives a basis of $C^{\varphi}_*(X)$ over K and allows us to form the torsion $\tau_{\varphi}(X, e) \in Q(K)$ relative to this basis. Note that the torsion $\tau_{\varphi}(X, e')$ arising from a different choice e' of base sequence stands in the relationship $\tau_{\varphi}(X, e) = \pm\varphi(h)\tau_{\varphi}(X, e')$ for some $h \in H_1(X;\mathbb{Z})$ (here, as is standard practice, we write the group operation in $H_1(X;\mathbb{Z})$ multiplicatively when dealing with elements of $\mathbb{Z}[H_1(X;\mathbb{Z})]$). The set of all such torsions arising from all choices of e is "the" torsion of X associated to φ and is denoted $\tau_\varphi(X)$.

We are now in a position to define the torsions we will need.

Definition 1.1 *1. For X a smooth manifold as above with $b_1(X) \geq 1$, let $\phi : X \to S^1$ be a map representing an element $[\phi]$ of infinite order in $H^1(X;\mathbb{Z})$. Let C be the infinite cyclic group generated by the formal variable t, and let $\varphi_1 : \mathbb{Z}[H_1(X;\mathbb{Z})] \to \mathbb{Z}[t,t^{-1}]$ be the map induced by the homomorphism $H_1(X;\mathbb{Z}) \to C$, $a \mapsto t^{[\phi](a)}$. Then the* Reidemeister torsion $\tau(X,\phi)$ *of X associated to ϕ is defined to be the torsion $\tau_{\varphi_1}(X)$.*

2. Write H for the quotient of $H_1(X;\mathbb{Z})$ by its torsion subgroup, and let $\varphi_2 : \mathbb{Z}[H_1(X;\mathbb{Z})] \to \mathbb{Z}[H]$ be the map induced by the projection $H_1(X;\mathbb{Z}) \to H$. The Milnor torsion $\tau(X)$ *is defined to be $\tau_{\varphi_2}(X)$.*

Remark 1.2 *1. Some authors use the term Reidemeister torsion to refer to the general torsion $\tau_\varphi(X)$; and other terms, e.g., Reidemeister-Franz-DeRham torsion, are also in use.*

2. The torsions in Definition 1.1 are defined for manifolds X of arbitrary dimension, with or without boundary. We will be concerned only with the case that X is a closed manifold of dimension 3 with $b_1(X) > 1$. In this special *case, work of Turaev [15] shows that $\tau(X)$ and $\tau(X,\phi)$, naturally subsets of $\mathbb{Q}(H)$ and $\mathbb{Q}(t)$, are actually subsets of $\mathbb{Z}[H]$ and $\mathbb{Z}[t,t^{-1}]$. We will usually think of $\tau(X,\phi)$ as an element of $\mathbb{Z}[t]$ defined up to multiplication by $\pm t^k$ and similarly $\tau(X)$ as an element of $\mathbb{Z}[H]$ defined up to translation by $\pm h$ for $h \in H$.*

1.1.2 S^1-valued Morse Theory

We briefly describe the result of Hutchings and Lee that motivate the theorem we will prove. From now on we will fix a smooth, connected, closed, oriented three-manifold X having $b_1(X) > 1$ and a function $\phi : X \to S^1$ that satisfies

- ϕ is Morse, i.e., ϕ is smooth and has only nondegenerate critical points;
- ϕ represents an element $[\phi] \in H^1(X;\mathbb{Z})$ of infinite order;
- ϕ has no critical points of index 0 or 3.

Our topological assumptions on X guarantee the existence of such functions ϕ. Given ϕ, we fix a smooth level set $\Sigma_0 = \phi^{-1}(\mathrm{pt}) \subset X$ once and for all. Upward gradient flow of ϕ defines a return map $f : \Sigma_0 \to \Sigma_0$ away from the descending manifold of the critical points. The *zeta function* of f is defined by the power series

$$\zeta(f) = \exp\left(\sum_{k\geq 1} \mathrm{Fix}(f^k)\frac{t^k}{k}\right)$$

where $\mathrm{Fix}(f^k)$ denotes the number of fixed points (counted with sign in the usual way) of the k-th iterate of f. One should think of $\zeta(f)$ as keeping track of the number of closed orbits of ϕ (which correspond to fixed points of iterates of f), as well as the "degree" of those orbits (by which we mean, for a closed orbit $\gamma \subset X$, the degree of the map $\phi : \gamma \cong S^1 \to S^1$).

To "count" gradient flows of ϕ connecting critical points, we introduce a Morse complex. Write $\mathbb{Z}[[t]]$ for the ring of Laurent series in the variable t, and let M^i denote the free $\mathbb{Z}[[t]]$-module generated by the index-i critical points of ϕ. The differential $d : M^i \to M^{i+1}$ in this complex is defined to be

$$dx_\mu = \sum_\nu a_{\mu\nu}(t) y_\nu$$

where x_μ is an index-i critical point, $\{y_\nu\}$ is the set of index-$(i+1)$ critical points, and $a_{\mu\nu}(t)$ is a series in t whose coefficient of t^n is defined to be the number of gradient flow lines of ϕ connecting x_μ with y_ν that cross Σ_0 n times. Here we count the gradient flows with sign determined by orientations on the ascending and descending manifolds of the critical points; see [6] for more details.

In our case, where X is three-dimensional and ϕ has no index 0 or index 3 critical points, there must be the same number of index 1 and index 2 critical points–i.e., that $d : M^1 \to M^2$ forms a square matrix after choosing ordered bases for the M^i. Fixing such a basis allows us to form $\det(d) \in \mathbb{Z}[[t]]$. This data is equivalent to the torsion $\tau(M^*)$ defined in the previous section.

Theorem 1.3 (Hutchings-Lee) *In this situation, we have the relation*

$$\zeta(f)\det(d) = \tau(X, \phi) \tag{2}$$

up to multiplication by $\pm t^k$.

1.2 Statement of Results

As mentioned above, it has been shown [11] that the Seiberg-Witten invariant of X (after summing over the action of the torsion subgroup of $H_1(X;\mathbb{Z})$) can be identified with the Milnor torsion $\tau(X)$. In particular, if we use our S^1-valued function ϕ to "average" with respect to Σ_0, this result implies

$$\sum_{\alpha \in \mathrm{spin}^c(X)} SW(\alpha)\, t^{c_1(\alpha).\Sigma_0/2} = t^{\chi(\Sigma_0)/2}\,\tau(X,\phi). \tag{3}$$

This statement, compared with equation (2), shows that the "counting" invariant $\zeta(f)\det(d)$ is related to the Seiberg-Witten invariant. Conversely, a proof that the left hand sides of (2) and (3) are related would provide an alternate, independent proof of the Meng-Taubes theorem. Strictly speaking, this would prove an "averaged" version of the theorem, but that (see [6]) is sufficient to recover the full relationship between the Seiberg-Witten invariant and the Milnor torsion stated in [11].

The main result of this work is a proof that the left hand side of (2) is indeed equal to $t^{-\chi(\Sigma_0)/2}$ times the left hand side of (3), as conjectured in [5, 6].

Remark 1.4 *Hutchings and Lee's conjecture is more general, in that they hypothesize that their counting invariant I of* spin^c *structures should agree with the Seiberg-Witten invariant. The present work proves this statement "modulo torsion."*

Our proof of the Hutchings-Lee conjecture is based on ideas of Donaldson for computing the Seiberg-Witten invariants of 3-manifolds. We outline Donaldson's construction here; see Section 2 below for more details. Given $\phi : X \to S^1$ a generic Morse function as above and Σ_0 the inverse image of a regular value, let $W = X \setminus nbd(\Sigma_0)$ be the complement of a small neighborhood of Σ_0. Then W is a cobordism between two copies of Σ_0 (since we assumed ϕ has no extrema—note we may also assume Σ_0 is connected);

consider the Seiberg-Witten equations on W. Note that two spin^c structures on X that differ by an element $a \in H^2(X;\mathbb{Z})$ with $a([\Sigma_0]) = 0$ restrict to the same spin^c structure on W, in particular, spin^c structures s on W are determined by their degree $\langle c_1(s), \Sigma_0\rangle$.

Now, a solution of the Seiberg-Witten equations on W restricts to a solution of the *vortex equations* on Σ_0 at each end of W (more accurately, we should complete W by adding infinite tubes $\Sigma_0 \times (-\infty, 0]$, $\Sigma_0 \times [0, \infty)$ to each end, and consider a finite-energy solution on this completed space)—see [3], [13] for example. These equations have been extensively studied, and it is known that the moduli space of solutions to the vortex equations on Σ_0 can be identified with a symmetric power of Σ_0 itself: see [2], [8]. Explicitly, if we use a spin^c structure s on W with $\langle c_1(s), \Sigma_0\rangle = 2m$ for some $m \in \mathbb{Z}$, the vortex moduli space is identified with $\mathrm{Sym}^{g-1+m}\Sigma_0$. Donaldson uses the restriction maps on the Seiberg-Witten moduli space of W to obtain a self-map κ_m of the cohomology of $\mathrm{Sym}^{g-1+m}\Sigma_0$, the alternating trace $\mathrm{Tr}\,\kappa_m$ of which is identified as the sum of Seiberg-Witten invariants of spin^c structures on X that restrict to the given spin^c structure on W. For a precise statement, see Theorem 2.2.

Our main result is the following.

Theorem 1.5 *Let X be a Riemannian 3-manifold with $b_1(X) > 1$, and fix an integer $m \in \mathbb{Z}$ as in the previous paragraph. Then we have*

$$\operatorname{Tr} \kappa_m = [\tau(M^*)\, \zeta(f)]_{g-1+m} \tag{4}$$

where Tr *denotes the alternating trace and* $[\,\cdot\,]_n$ *denotes the coefficient of* t^n *of the polynomial enclosed in brackets.*

The "translation" $(m \mapsto g - 1 + m)$ corresponds to the factor $t^{\chi(\Sigma_0)/2}$ in equation (3), and arises because a Seiberg-Witten solution on W in a spin^c structure whose determinant line has degree $2m$ restricts to a solution of the vortex equations on Σ_0 in a bundle of degree $g - 1 + m$.

The main idea of the proof is to identify both quantities in equation (4) as a Lefschetz-type intersection between the graph of a gluing map and a diagonal-like cycle in a product $\mathrm{Sym}^{g-1+m+N}\Sigma_0 \times \mathrm{Sym}^{g-1+m+N}\Sigma_0$ of symmetric products, where N is the number of index 1 critical points of ϕ. On the left-hand side this is fairly straightforward algebraic topology, thanks to further results of Donaldson; on the right-hand side it involves some combinatorial calculations and genericity arguments.

A key point, however, is that we can calculate $\operatorname{Tr} \kappa_m$ explicitly, given information about the intersections between ascending and descending manifolds of the critical points of ϕ. The result is stated in Theorem 3.1, and gives an entirely topological description of κ_m. The proof of this theorem, which we do not include here but refer to [3], is remarkable in the fact that one need not solve the Seiberg-Witten equations in order to obtain the result. Indeed, Donaldson shows in [3] that κ_m is determined by formal properties of the theory, using general algebraic arguments.

2 A TQFT for Seiberg-Witten Invariants

In this section we describe Donaldson's "topological quantum field theory" for computing the Seiberg-Witten invariants. We use the notation from the introduction: $\phi : X \to S^1$ is a Morse function without extrema, Σ_0 a smooth level set, and $W = X \setminus \Sigma$ is a cobordism from Σ to itself that comes with an identification $\partial_+ W \to \partial_- W$ recovering X. We will find it convenient to complete W by adding infinite cylindrical ends $\Sigma \times (-\infty, 0]$ and $\Sigma \times [0, \infty)$ to the boundary, forming a noncompact space that we denote $\hat{W}$. We take the ends to be Riemannian products of the metric on Σ_0 with the usual metric on $\mathbb{R}$.

Recall that a spin^c structure on a Riemannian 3-manifold Y is a lift of the $SO(3)$-frame bundle to a spin$^c(3) = SU(2) \times U(1)/\pm 1 = U(2)$ principal bundle. Such lifts are in 1-1 correspondence with elements of of $H^2(Y;\mathbb{Z})$. Each spin^c structure α has a determinant $U(1)$ bundle $\det(\alpha)$ defined using the obvious representation spin$^c(3) \to U(1)$, $(z, e^{i\theta}) \mapsto e^{2i\theta}$. This gives rise to a map from the set spin$^c(Y)$ of spin^c structures into $H^2(Y;\mathbb{Z})$ by $\alpha \mapsto c_1(\det(\alpha))$ that is in general not 1-1; the indeterminacy is described by the 2-torsion in $H^2(Y;\mathbb{Z})$. Note that in our case $H^2(W;\mathbb{Z}) = \mathbb{Z}$.

Now, a choice of spin^c structure α gives rise to a hermitian 2-plane bundle

S, the *spinor bundle*, using the usual representation of $U(2) = \mathrm{spin}^c(3)$. There is a map $\gamma : TY \to \mathrm{End}(S)$ satisfying the Clifford relation $\gamma(v)^2 = -|v|^2\mathbb{I}$; this is known as Clifford multiplication and we will usually write $\gamma(v)(s)$ as $v.s$.

Given a spin^c structure α with determinant line bundle L of degree $2m$, we consider the space of pairs (A, Ψ), where A is a connection on L and Ψ is a section of the spinor bundle S associated to α. The choice of A determines, together with the Levi-Civita connection on W, a connection on S that is compatible with Clifford multiplication, and an associated Dirac operator D_A. The Seiberg-Witten equations for (A, Ψ) are

$$\begin{aligned} D_A\Psi &= 0 \\ \star F_A &= i\tau(\Psi, \Psi); \end{aligned}$$

we will consider only *finite-energy* solutions on $\hat{W}$. Here $\tau(\cdot\,,\cdot) : S \otimes S \to \Omega^1(Y)$ is the adjoint of Clifford multiplication, defined by $\langle a, \tau(\phi, \psi)\rangle_{\Lambda^1} = \frac{1}{2}\langle \phi, ia.\psi\rangle_S$. The equations are invariant under the natural action of the gauge group $\mathcal{G} = \mathrm{Map}(W, S^1)$; we can then form the moduli space $\mathcal{M}_{\hat{W}}$ of solutions modulo gauge. After appropriate perturbation of the equations (a technical point that poses no difficulties for the argument to come, and which, therefore, we will ignore) and use of appropriate Sobolev norms to fit the theory into the usual Fredholm "package," the space $\mathcal{M}_{\hat{W}}$ becomes a smooth, compact manifold.

Remark 2.1 *For Y a closed manifold with $b_1(Y) \geq 1$, the moduli space $\mathcal{M}_Y$ of solutions to the Seiberg-Witten equations modulo gauge is, for generic choice of metric and perturbation, a compact 0-dimensional manifold. The Seiberg-Witten invariant of Y in the chosen* spin^c *structure is then the signed count of points in $\mathcal{M}_Y$. If $b_1(Y) > 1$, the resulting number is independent of the choice of generic metric and perturbation. The philosophy for the situation above is that solutions on $\hat{W}$ that limit to the same solutions on either end will give rise to a solution on the original closed manifold X; the topological quantum field theory described in this section is meant to be an algebraic way of counting these solutions. Some of these technical issues are addressed in the Appendix.*

On the ends of $\hat{W}$, metrically the product $\Sigma_0 \times \mathbb{R}^{\pm}$, these equations reduce to the following: for a spin structure $K^{\frac{1}{2}}$ on Σ_0 (recall that spin structures on a Riemann surface are exactly the square roots of the canonical bundle K), write the restriction $S|_{\Sigma_0}$ as $(K^{\frac{1}{2}} \oplus K^{-\frac{1}{2}}) \otimes E$ for some hermitian line bundle E; note that $E^2 = L|_{\Sigma_0}$. Then the equations are for a t-dependent section $\phi = (\alpha, \beta) \in \Gamma((K^{\frac{1}{2}} \otimes E) \oplus (K^{-\frac{1}{2}} \otimes E))$ and a t-dependent connection

B in $K^{\frac{1}{2}} \otimes E$ (compatible with the Levi-Civita connection on K), and read

$$iF_B = \tfrac{1}{2}(|\beta|^2 - |\alpha|^2)vol_\Sigma$$

$$-2i\bar{\partial}_B\alpha = \dot{\beta} \qquad 2i\bar{\partial}_B^*\beta = \dot{\alpha} \qquad \dot{B} = \alpha\bar{\beta}.$$

The general setup for this problem, which is described in detail in the Appendix, implies that a finite-energy solution (that is, an $H^k_{\delta,e}$ solution, in the notation of the Appendix) on $\hat{W}$ must approach a constant solution near the ends. For a constant solution either $\alpha = 0$ or $\beta = 0$; which of these holds is determined entirely by the degree of E. Suppose $m = \deg E < 0$: then we must have $\beta = 0$, and the equations are now

$$iF_B = -\tfrac{1}{2}|\alpha|^2 vol_\Sigma$$

$$\bar{\partial}_B\alpha = 0,$$

which are the Kähler vortex equations over Σ_0 in the bundle $K^{\frac{1}{2}}\otimes E$. In order to obtain any solutions at all to these equations (in particular, for $K^{\frac{1}{2}} \otimes E$ to admit any holomorphic sections) we must have $0 \leq \deg(K^{\frac{1}{2}} \otimes E)$, and therefore $-(g-1) \leq \deg E < 0$. In case $-(g-1) < \deg E$, we obtain a well-behaved moduli space of solutions that can, according to [8], be identified with the symmetric product $\mathrm{Sym}^n\Sigma_0$, where $n = \deg(K^{\frac{1}{2}} \otimes E) = g - 1 + m$. There is a symmetric statement for the case $\deg(E) > 0$For notational convenience, from now on we will write $\Sigma_0^{(n)}$ for $\mathrm{Sym}^n\Sigma_0$.

Writing $\Sigma_0 = \partial_- W$ and $\Sigma_1 = \partial_+ W$, we have restriction maps $\rho_i : \mathcal{M}_{\hat{W}} \to \Sigma_i^{(n)}$ for $i = 0, 1$ defined by following a finite-energy solution on $\hat{W}$ to its limiting values as $t \to \infty$ on $\Sigma_0 \times \{-t\}$, $\Sigma_1 \times \{t\}$. Note that we have an identification between Σ_0 and Σ_1. Using Poincaré duality, we then get an element

$$\begin{aligned}\kappa_m = (\rho_0 \otimes \rho_1)_*[\mathcal{M}_W] &\in H_*(\Sigma_0^{(n)}) \otimes H_*(\Sigma_1^{(n)}) \\ &\cong \mathrm{Hom}(H^*(\Sigma_0^{(n)}), H^*(\Sigma_0^{(n)})).\end{aligned}$$

(Here, as throughout the paper, we work with rational coefficients unless otherwise specified.) This is the basis for our "TQFT:" to a surface Σ we associate the cohomology of the symmetric product $\Sigma^{(n)}$, and to a cobordism W between Σ_0 and Σ_1 we assign the homomorphism κ_m.

$$\begin{aligned}\Sigma &\longleftrightarrow V_\Sigma = H^*(\Sigma^{(n)}) \\ W &\longleftrightarrow \kappa_m : V_{\Sigma_0} \to V_{\Sigma_1}\end{aligned}$$

From this point on, we will drop m from the notation, writing κ for the map $H^*(\Sigma_0^{(g-1+m)}) \to H^*(\Sigma_1^{(g-1+m)})$.

Gluing theory for Seiberg-Witten solutions (see Appendix) provides a proof of the central property of TQFTs, namely that if W_1, W_2 are two composable cobordisms then

$$\kappa_{W_1 \cup W_2} = \kappa_{W_2} \circ \kappa_{W_1}. \tag{5}$$

Theorem 2.2 (Donaldson) *The alternating trace of the map κ calculates the Seiberg-Witten invariant of X, in the following sense. If S_m denotes the set of spin^c structures α on X that have $c_1(\det \alpha) \cdot \Sigma = 2m$, then*

$$\operatorname{Tr} \kappa = \sum_{\alpha \in S_m} SW_X(\alpha),$$

where $\operatorname{Tr} \kappa = \sum_i (-1)^i \operatorname{tr}(\kappa|_{H^i(\Sigma^{(n)})})$. □

The proof rests again on gluing theory for solutions; in the current situation one finds that $\operatorname{Tr} \kappa$ is calculating a coincidence number for the two restriction maps ρ_0, ρ_1, and the theorem is a result that other gauge-theoretic results would lead one to expect.

In order to calculate the invariant, then, we use the composition rule (5) to reduce to the simplest possible situation, namely a cobordism W between a surface Σ_0 and a surface Σ_1 of one higher genus. Equivalently, the height function on W has a single index-1 critical point. It is a remarkable fact that κ is determined in this situation by the formal properties of the theory; this result will be described in the next section.

Recall (see MacDonald [10]) that the cohomology of the symmetric product $\Sigma^{(n)}$ can be expressed as follows:

$$H^*(\Sigma^{(n)}) = \bigoplus_{i=0}^{n} \Lambda^i(H^1(\Sigma)) \otimes \mathrm{Sym}^{n-i}(H^0(\Sigma) \oplus H^2(\Sigma)). \tag{6}$$

The "interesting" part of this expression is the exterior power $\Lambda^i(H^1(\Sigma))$; the symmetric part can be thought of as a fixed vector space of dimension $n - i + 1$, independent of Σ (since this part of (6) has dimension independent of the genus—see [3] for further discussion of this point). It turns out that κ is the natural extension of a map $\Lambda^*(H^1(\Sigma_0)) \to \Lambda^*(H^1(\Sigma_1))$ that is defined entirely topologically, a map we now describe.

3 The Plücker Construction

In the following two paragraphs there is no reason to restrict the cobordism W, so we return for now to the general situation of any cobordism with boundary $\partial_- W = \Sigma_0$, $\partial_+ W = \Sigma_1$. There are inclusion-induced maps $r_i : H^1(W) \to H^1(\Sigma_i)$ that we combine to produce a subspace $\Gamma_W = r_0 \oplus r_1(H^1(W)) \subset H^1(\Sigma_0) \oplus H^1(\Sigma_1)$. Recall the Plücker construction: given a subspace $S \subset V_0 \oplus V_1$ of a sum of oriented vector spaces, we can form a linear map $|S| : \Lambda^*(V_0) \to \Lambda^*(V_1)$ by wedging together all the elements of a basis for S, yielding an element of $\Lambda^*(V_0 \oplus V_1) \cong \Lambda^*(V_0) \otimes \Lambda^*(V_1) \cong \mathrm{Hom}(\Lambda^*(V_0), \Lambda^*(V_1))$ The second isomorphism uses the fact that the volume form on V_0 defines an isomorphism $(\Lambda^k V_0)^* \cong \Lambda^{\dim(V_0)-k} V_0$. The result is a linear map well-defined up to a multiplicative constant; we can reduce this to an indeterminacy of ± 1 in our situation (where $V_i = H^1(\Sigma_i)$) by using the integer lattices. Write ρ_W for the Plücker map associated to Γ_W.

This leads us, motivated by the constructions of the previous section, to a "baby TQFT:"

$$\begin{aligned} \text{surface } \Sigma \quad &\leftrightarrow \quad V_\Sigma = \Lambda^*(H^1(\Sigma)) \\ \text{cobordism } W \quad &\leftrightarrow \quad \rho_W : V_{\Sigma_0} \to V_{\Sigma_1} \end{aligned}$$

Strictly speaking this is *not* a TQFT since the composition rule fails without further assumptions on the cobordism. In the case that W is an "elementary" cobordism, i.e., contains a single critical point, the map ρ_W will be used to find κ_W. We calculate this case now.

Suppose, then, that W connects Σ_g and Σ_{g+1} in the simplest way: there is a unique critical point (of index 1) of the height function $h : W \to \mathbb{R}$, and the ascending manifold of this critical point intersects Σ_{g+1} in an essential curve that we will denote by c.

Now, c obviously bounds a disk $D \subset W$; the Poincaré dual of $[D] \in H_2(W, \partial W)$ is a 1-cocycle that we will denote $\xi_0 \in H^1(W)$. It is easy to check that ξ_0 is in the kernel of the restriction $r_1 : H^1(W) \to H^1(\Sigma_g)$, so we may complete ξ_0 to a basis $\xi_0, \xi_1, \dots, \xi_{2g}$ of $H^1(W)$ with the property that $\eta_1 := r_1(\xi_1), \dots, \eta_{2g} := r_1(\xi_{2g})$ form a basis for $H^1(\Sigma_g)$. Since the restriction $r_2 : H^1(W) \to H^1(\Sigma_{g+1})$ is injective, we know $\bar{\eta}_0 := r_2(\xi_0), \dots, \bar{\eta}_{2g} := r_2(\xi_{2g})$ are linearly independent; note that $r_2(\xi_0)$ is just c^*, the Poincaré dual of c.

The choice of basis ξ_i with its restrictions η_i, $\bar{\eta}_i$ gives rise to an inclusion $\iota : H^1(\Sigma_g) \to H^1(\Sigma_{g+1})$ in the obvious way, namely $\iota(\eta_i) = \bar{\eta}_i$. One may check that this map is independent of the choice of basis $\{\xi_i\}$ for $H^1(W)$ having ξ_0 as above.

With this understood, it is clear that $\Gamma = r_1 \oplus r_2(H^1(W)) \subset H^1(\Sigma_g) \oplus H^1(\Sigma_{g+1})$ is spanned by $\{0 \oplus c^*, \eta_1 \oplus \bar{\eta}_1, \ldots, \eta_{2g} \oplus \bar{\eta}_{2g}\}$. The reader may easily check that the Plücker map $\Lambda^* H^1(\Sigma_g) \to \Lambda^* H^1(\Sigma_{g+1})$ associated to this subspace is $\alpha \mapsto \iota(\alpha) \wedge c^*$.

Theorem 3.1 (Donaldson) *In the case that W consists of a single handle-addition as above, the map κ arising from Seiberg-Witten theory agrees with the Plücker map above on the exterior powers of $H^1(\Sigma_0)$ and is the obvious "identity" map on symmetric powers of $H^0(\Sigma_0) \oplus H^2(\Sigma_0)$.*

For proof, see [3]. It is remarkable that the proof of this theorem is entirely formal, and in particular does not rely on any analysis of the Seiberg-Witten equations. $\square$

Using the composition rule (5), we see that the map κ associated to the full cobordism arising from our original Morse function is given by the composition of the Plücker maps associated to each handle addition in the cobordism. Note that here we are using the following facts:

1. The Seiberg-Witten equations give rise to a TQFT, satisfying (5) in particular, and
2. The maps arising in this TQFT corresponding to elementary cobor-

disms agree with the maps arising from the topological construction above.

This means that if $W = W_1 \cup_{\Sigma_1} \cdots \cup_{\Sigma_{m-1}} W_m$ is a decomposition of W into elementary cobordisms and $\bar{\rho}_{W_i}$ is the extension of ρ_{W_i} to $H^*(\Sigma_i^{(n)})$, then κ_W is given by the composition $\bar{\rho}_m \circ \cdots \circ \bar{\rho}_1$.

We have now seen that the 1-handle additions in W give rise to the map above that is essentially wedging with the cocore of the new handle. The 2-handles give the transpose of that simple map in a way we now describe.

Suppose V_1 and V_2 are oriented vector spaces with fixed volume forms and $\Gamma \subset V_1 \oplus V_2$ is a subspace with Plücker map $|\Gamma| : \Lambda^* V_1 \to \Lambda^* V_2$. The obvious isomorphism $V_1 \oplus V_2 \to V_2 \oplus V_1$ gives rise to a map $|\bar{\Gamma}| : \Lambda^* V_2 \to \Lambda^* V_1$, and the relationship between the two maps is given in the following

Lemma 3.2 *Whenever $a \in \Lambda^i V_1$ and $b \in \Lambda^j V_2$ are such that $|\Gamma|(a)$ and b have complementary degree in $\Lambda^* V_2$, we have*

$$a \wedge |\bar{\Gamma}|(b) = (-1)^{(\dim(\Gamma)+\dim(V_1))(i+\dim(V_1))} |\Gamma|(a) \wedge b,$$

using the volumes to identify $\Lambda^{\mathrm{top}} V_k$, $k = 1, 2$, and in particular to resolve the ± 1 indeterminacy in the definition of $|\Gamma|$.

The proof is straightforward, and we leave it to the reader. □

We are now in a position to obtain a formula for $\operatorname{Tr}\kappa$. Here and in what follows we will write Σ_0 for $\partial_- W$ and Σ_1 for $\partial_+ W$. We will also assume that we have arranged W so that under the height function h induced from the original Morse function ϕ, the index-1 critical points occur "below" the index-2 critical points—i.e., W consists of a certain number, say N, of 1-handle additions, followed by exactly N 2-handle additions. Write Σ for the "middle" of W, i.e., $\Sigma = \partial_+(\Sigma_0 \cup (\text{all 1-handles}))$. There are N distinguished curves on Σ, the ascending manifolds of the critical points, that we denote by c_i.

Note that if W_0 is the "first half" of the cobordism, i.e., $W_0 = \Sigma_0 \cup (\text{all 1-handles})$, and W_1 is the "second half," then W_0 and W_1 are topologically identical. It will be convenient to assume that W_0 and W_1 are in fact two copies of the *same space*, with a fixed identification between them, and that the original cobordism W is obtained from these two by means of an orientation-reversing diffeomorphism $A : \Sigma = \partial_+ W_0 \to \Sigma = \partial_- W_1$. (See Figure 1.)

A word about orientations. We should think of W_1 as having the opposite orientation from W_0; loosely if W_0 is written together with an orientation o_0 as (W_0, o_0), then (W_1, o_1) is taken to be $(W_0, -o_0)$. Now, o_0 induces an

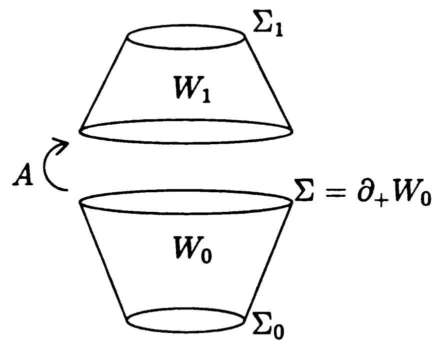


Figure 1: Decomposition of W

orientation on $\partial_+ W_0 = \Sigma$, which we will also write $\mathfrak{o}_0$: $\partial_+ W_0 = (\Sigma, \mathfrak{o}_0)$. In the same way $-\mathfrak{o}_0$ induces an orientation on $\partial_- W_1 = \Sigma$, namely $\partial_- W_1 = (\Sigma, -\mathfrak{o}_0)$ (here $\partial_+ W_0$ and $\partial_- W_1$ are taken to be orented as boundary components of their respective manifolds). Thus the orientation-reversing map $\partial_+ W_0 \to \partial_- W_1$ in fact preserves the orientation $\mathfrak{o}_0$ of Σ induced by the identification $\Sigma = \partial_+ W_0$.

Clearly A contains all information about W; therefore A will play a central role in the calculations that follow. We note that we may make certain transversality assumptions on A (for example, that $A(c_i)$ meets c_j transversely for all i and j), but as these issues will come up again later we do not dwell on them here.

To avoid confusion, when referring to Σ we will always mean $\partial_+ W_0$, so that $\{c_i\}$ are parts of the ascending manifolds of the index-1 critical points.

In this scheme, then, the descending manifolds of the index-2 critical points intersect Σ in the curves $A^{-1}(c_i)$.

With the situation now standardized, let us also fix an ordering of the critical points in W_0, and hence also an ordering of the curves c_i. Each critical point gives rise to a map $P_i : H^*(\Sigma_{g+i-1}^{(n+i-1)}) \to H^*(\Sigma_{g+i}^{(n_i)})$ where Σ_{g+i} is the genus $g+i$ surface obtained as $\partial_+(\Sigma_0 \cup \{\text{the first } i \text{ 1-handles}\})$. We will denote by P the composition of Plücker maps arising from the first N handle additions: that is $P = P_N \circ \cdots \circ P_1 : H^*(\Sigma_0^{(n)}) \to H^*(\Sigma^{(n+N)})$ is given by $P(\alpha) = \iota(\alpha) \wedge c_1^* \wedge \cdots \wedge c_N^*$ where ι is the inclusion induced in the cohomology of the symmetric products from the composition of inclusions $H^1(\Sigma_g) \to H^1(\Sigma_{g+1})$ defined previously. In particular we extend the inclusion to the exterior algebras of $H^1(\cdot)$, then to the full cohomology of the symmetric product via the identity on the $\mathrm{Sym}^k(H^0 \oplus H^2)$ factor (compare equation (6)).

Simlarly, we write $\bar{P}$ for the composition of "transposes" of the Plücker maps, in the notation of Lemma 3.2; that is, $\bar{P} = \bar{P}_1 \circ \cdots \circ \bar{P}_N$. In what follows we will omit the map ι from the notation, implicitly using it to identify $H^*(\Sigma_0^{(n)})$ as a subspace of $H^*(\Sigma^{(n+N)})$. We can write κ as the composition $\kappa = \bar{P} \circ A^{-1*} \circ P : H^*(\Sigma_0^{(n)}) \to H^*(\Sigma^{(n+N)}) \to H^*(\Sigma^{(n+N)}) \to H^*(\Sigma_1^{(n)}) = H^*(\Sigma_0^{(n)})$; our objective is to compute the alternating trace of κ.

In what follows it will be convenient to fix a basis $\{\alpha\}$ for $H^*(\Sigma_0^{(n)})$ and its cup dual basis $\{\alpha^\circ\}$—so for basis elements α and β we have $\alpha^\circ \cup \beta = \delta_{\alpha\beta}$. MacDonald [10] shows that we may take the α and α° to be "monomials," i.e., each α lies in a single summand in expression (6).

Proposition 3.3 *We have*

$$\operatorname{Tr}\kappa = \sum_{\alpha}(-1)^{|\alpha|+N|\alpha|+N(N-1)/2}(A^*(\alpha)\wedge A^*c_N^*\wedge\cdots\wedge A^*c_1^*)\cup(\alpha\wedge c_N^*\wedge\cdots\wedge c_1^*), \tag{7}$$

where $|\alpha|$ denotes the degree of α, and c_i^* is the Poincaré dual of the curve c_i on Σ.

For the proof, note that repeated application of Lemma 3.2 shows that if ω_1 and $\bar{P}(\omega_2)$ are of complementary degree in $\Lambda^*(H^1(\Sigma_0))$ then

$$\begin{aligned}\omega_1 \wedge \bar{P}(\omega_2) &= (-1)^{N|\omega_1|+N(N-1)/2}P(\omega_1)\wedge\omega_2 \\ &= (-1)^{N|\omega_1|+N(N-1)/2}(\omega_1\wedge c_1^*\wedge\cdots\wedge c_N^*)\wedge\omega_2.\end{aligned}$$

In the calculation that follows we will apply the above fact to elements ω_1, ω_2 in the cohomology of a symmetric product, and the relevant degree $|\omega_1|$ is not the degree of ω_1 as a cohomology class, but rather the degree of the part of ω_1 that lies in the exterior algebra (we will sum only over a basis of monomials, so the latter degree is defined in that case). However, these two are clearly equal modulo two, so we are justified in using the above calculation in our situation. To calculate the trace we write

$$\begin{aligned}
\operatorname{Tr}\kappa &= \sum_{\alpha}(-1)^{|\alpha|}\alpha^{\circ}\cup\bar{P}\circ A^{-1*}\circ P(\alpha) \\
&= \sum_{\alpha}(-1)^{|\alpha|+N|\alpha|+\frac{N(N-1)}{2}}(\alpha^{\circ}\wedge c_1^*\wedge\cdots\wedge c_N^*)\cup A^{-1*}(\alpha\wedge c_1^*\wedge\cdots\wedge c_N^*) \\
&= \sum_{\alpha}(-1)^{|\alpha|+N|\alpha|+\frac{N(N-1)}{2}}A^*(\alpha^{\circ}\wedge c_1^*\wedge\cdots\wedge c_N^*)\cup(\alpha\wedge c_1^*\wedge\cdots\wedge c_N^*).
\end{aligned}$$

□

Remark 3.4 *Our ordering of the critical points of ϕ determines an ordering of the c_i used above. However,* $\operatorname{Tr}(\kappa)$ *is independent of this ordering.*

Our objective is to see that (7) agrees with the invariant defined by Hutchings and Lee. We next recall the definition of that invariant.

4 Spin^c Structures and Gradient Flows

We begin by recalling an association between spin^c structures on the closed 3-manifold X and elements of the first homology. Following [6], let X' denote $X\backslash$(critical points), and write $H(\phi) \subset H_1(X', \partial X')$ for those 1-cycles whose boundary consists of the sum of all index 2 critical points minus the sum of all index 1 critical points.

Proposition 4.1 (Hutchings–Lee) *The set of spin^c structures on X is in one-to-one correspondence with $H(\phi)$ by the map sending a spin^c structure α to the dual of the first Chern class of the determinant line of α.*

For proof we refer to [6]. □

More explicitly, given α we form $c_1(\det(\alpha))$, restrict this to X', then apply Poincaré duality $H^2(X') \to H_1(X', \partial X')$.

The invariant we are interested in will be an element of (a subring of) the ring of power series with variables from $H_1(X', \partial X')$, and will have nonzero coefficients only for elements of $H(\phi)$. To be more precise, let Λ be the Novikov ring of functions $\xi : H_1(X', \partial X') \to \mathbb{Z}$ satisfying the condition that the set $\{a \in H_1(X', \partial X') | \xi(a) \neq 0 \text{ and } a.\Sigma \leq k\}$ is finite for every k. We may think of elements of Λ as series in $H_1(X', \partial X')$ (in particular the product is the usual product of series), and will write them as such. To write down the invariant I of Hutchings and Lee, define an analogue B of the Morse boundary as follows: let Q_1 be the rational vector space spanned by the index 1 critical points and Q_2 that spanned by the index 2 critical points. Then define the linear map $B : Q_1 \to Q_2 \otimes \Lambda$ on a basis $\{x\}$ by

$$B(x) = \sum_y y \otimes \sum_{\gamma \in \mathcal{F}(x,y)} [\gamma]$$

where $\mathcal{F}(x, y)$ denotes the space of flows between the index 1 critical point x and the index 2 critical point y. Letting $\mathcal{O}$ denote the set of closed orbits of the gradient flow of ϕ, we set

$$I = \prod_{\gamma \in \mathcal{O}} (1 - [\gamma])^{-\epsilon(\gamma)} \det B.$$

Here $\epsilon(\gamma) = \pm 1$ is a sign to be defined below. One checks that I has nonzero coefficients only for elements of $H(\phi)$; it is Hutchings and Lee's conjecture that as a function on spin^c structures, I agrees with the Seiberg-Witten invariant.

As mentioned in the introduction, we will prove an "averaged" version of this statement, in the following sense. Let $\rho : H_1(X', \partial X') \to \mathbb{Z}[t]$ be the map that sends $a \in H_1(X', \partial X')$ to $t^{a \cdot \Sigma}$; then we consider the invariant $\rho(I)$.

To give an explicit computation of $\rho(I)$, we recall some further definitions. To the Morse function $\phi : X \to S^1$ there is associated a "Morse complex" (M^*, d) in the following way: choose an orientation on each ascending and descending manifold of each critical point, in such a way that the ascending manifold $\mathcal{A}(x)$ and descending manifold $\mathcal{D}(x)$ of a critical point x have intersection number $\mathcal{A}(x).\mathcal{D}(x) = +1$. Then M^i is defined to be the $\mathbb{Q}(t)$-vector space spanned by the index i critical points of ϕ, and $d : M^i \to M^{i+1}$ is defined by

$$dx_i = \sum_y a_{ij}(t) y_j$$

where $a_{ij}(t)$ is a polynomial whose n-th coefficient is the number (counted with the sign determined by the intersection of the relevant ascending and descending manifolds) of gradient flows between x_i and y_j.

The Morse function ϕ also determines a return map $f : \Sigma \to \Sigma$ via the upward gradient flow, defined away from the descending manifolds of the critical points. Similarly, the iterates of f are defined away from codimension-1 subsets of Σ, though we will not keep track of this in the notation. In order to count the closed orbits of ϕ, we introduce the zeta function $\zeta(f)$ of f:

$$\zeta(f) = \exp\left(\sum_{k>0} \mathrm{Fix}(f^k)\frac{t^k}{k}\right),$$

where $\mathrm{Fix}(f^k)$ is the number of fixed points of the iterate f^k, counted with the sign associated to the corresponding intersection point of the graph of f^k with the diagonal in $\Sigma \times \Sigma$.

It can be seen (see [7, 14]) that

$$\zeta(f) = \prod_{\gamma \in \mathcal{O}} (1 - t^{\gamma \cdot \Sigma})^{-\epsilon(\gamma)}$$

where, as above, $\mathcal{O}$ denotes the collection of closed orbits, and $\epsilon(\gamma)$ is the sign of the fixed point of the iterate of f corresponding to γ.

It is not hard to see that $\rho(I)$, introduced above, is simply $\zeta(f) \cdot \det(d : M^1 \to M^2)$. In the current situation, where the Morse complex M^* has nonzero terms only in degree 1 and 2, the determinant of the differential is exactly the torsion of the complex (c.f. equation (1.3)). Our task, then, has come down to explicitly calculating the n-th coefficient of $\zeta(f)\det(d)$ and comparing this to $\mathrm{Tr}\,\kappa$. That calculation and comparison is the content of the next section.

5 Combinatorial Calculations

Our proof that $\operatorname{Tr}\kappa$ is the nth coefficient of $\zeta(f)\det(d)$ will consist in showing that both quantities are given by a Lefschetz-like intersection of the graph of the map induced by A on $\Sigma^{(n+N)}$ with a cycle D that can be thought of as a diagonal, modified to include information about the c_i. Indeed, in the case that our Morse function has no critical points (i.e., $\phi : X \to S^1$ is a fiber bundle), D will be exactly the diagonal and the intersection gives the Lefschetz number of A on $\Sigma^{(n)}$. This can be seen immediately to agree with our formula for $\operatorname{Tr}\kappa$ in this case.

Returning to the general case, we can describe the diagonal-like class D as follows. Write points of symmetric products as sums of points on the original surface; then $D \in H_{2n+2N}(\Sigma^{(n+N)} \times \Sigma^{(n+N)})$ will be defined by

$$D = \left[\{\textstyle\sum_{i=1}^N p_i + \sum_{j=1}^n q_j, \sum_{i=1}^N p_i' + \sum_{j=1}^n q_j | p_i, p_i' \in c_i \text{ for each } i\}\right]. \quad (8)$$

Thus D consists of diagonal points of $\Sigma^{(n)} \times \Sigma^{(n)}$ together with pairs of points on the c_i, one pair for each i. To obtain a well-defined homology class, we take D to be the image under the symmetrization map $\sigma : (\Sigma^N \times \Sigma^n) \times (\Sigma^n \times \Sigma^N) \to \Sigma^{(n+N)} \times \Sigma^{(n+N)}$ of the class $c_N \times \cdots \times c_1 \times \Delta_n \times c_1 \times \cdots \times c_N$, where Δ_n is the diagonal in $\Sigma^n \times \Sigma^n$. Then the Poincaré dual of D is given by the image of

$$c_N^* \times \cdots \times c_1^* \times \left(\sum_\beta (-1)^{|\beta|} \beta^\circ \times \beta \right) \times c_1^* \times \cdots \times c_N^*$$

under symmetrization, where $\{\beta\}$ is a basis for $H^*(\Sigma^n)$ and $\{\beta^\circ\}$ its cup dual basis. It is not hard to see that this is

$$PD(D) = \sum_\alpha (-1)^{|\alpha|} (c_N^* \wedge \cdots \wedge c_1^* \wedge \tilde{\alpha}^\circ) \times (\tilde{\alpha} \wedge c_1^* \wedge \cdots \wedge c_N^*)$$

for $\{\tilde{\alpha}\}$ a basis for $H^*(\Sigma^{(n)})$, $\{\tilde{\alpha}^\circ\}$ the cup dual basis. In fact we may take the sum to run over only the basis $\{\alpha\}$ for $H^*(\Sigma_0^{(n)}) \subset H^*(\Sigma^{(n)}) \subset H^*(\Sigma^{(n+N)})$ we used previously, since if some $\tilde{\alpha}$ is complementary to the subspace spanned by $\{\alpha\}$ then either $\tilde{\alpha}$ or $\tilde{\alpha}^\circ$ contains at least one c_i^* (we take our basis elements to be monomials in a basis for $H^0(\Sigma), H^1(\Sigma)$, and $H^2(\Sigma)$ as in MacDonald [10]) and hence the corresponding term in the sum vanishes.

If $\{\xi\}$, $\{\xi^\circ\}$ are cup dual bases for $H^*(\Sigma^{(n+N)})$ extending $\{\tilde{\alpha}\}$, $\{\tilde{\alpha}^\circ\}$ above, we can write the Poincaré dual of the graph $\Gamma_{A^{(n+N)}}$ in cohomology as

$$PD(\Gamma_{A^{(n+N)}}) = \sum_\xi (-1)^{|\xi|} \xi^\circ \times A^*(\xi).$$

We also assume that $\{\xi\}$ includes all elements of the form $\alpha \wedge c_{i_1}^* \wedge \cdots \wedge c_{i_k}^*$, $i_1 < \cdots < i_k$. Finally, we can calculate:

$$D.\Gamma_{A^{(n+N)}} = PD(\Gamma_{A^{(n+N)}}) \cup PD(D)$$

$$= \left(\sum_{\xi}(-1)^{|\xi|}\xi^{\circ} \times A^{*}(\xi)\right) \cup$$

$$\left(\sum_{\alpha}(-1)^{|\alpha|+N|\alpha|+\frac{N(N-1)}{2}}(\alpha^{\circ} \wedge c_1^{*} \wedge \cdots \wedge c_N^{*}) \times (\alpha \wedge c_1^{*} \wedge \cdots \wedge c_N^{*})\right)$$

$$= \sum_{\alpha,\xi}(-1)^{\epsilon}(\xi^{\circ} \cup (\alpha^{\circ} \wedge c_1^{*} \wedge \cdots \wedge c_N^{*})) \times (A^{*}(\xi) \cup \alpha \wedge c_1^{*} \wedge \cdots \wedge c_N^{*}),$$

where $\epsilon = |\xi| + |\alpha| + N|\alpha| + |\xi|(|\alpha| + N) + N(N-1)/2$. The first term in the last cross product is zero unless $\xi = \alpha^{\circ} \wedge c_1^{*} \wedge \cdots \wedge c_N^{*}$, and in that case it is 1. Carrying out the sum over ξ and comparing to Proposition 7 therefore gives the following:

Proposition 5.1 $\operatorname{Tr} \kappa = D.\Gamma_{A^{(n+N)}}$. □

On the other hand, we have:

Proposition 5.2 *The intersection from the previous proposition can be calculated as*

$$D.\Gamma_{A^{(N+n)}} = \sum_{k=0}^{n} \operatorname{Fix}(A^{(k)})\langle A_{\circ}^{n-k}(c_1 \wedge \cdots \wedge c_N), A_{*}^{-1}c_1 \wedge \cdots \wedge A_{*}^{-1}c_N\rangle, \quad (9)$$

where $A_{\circ}$ is the action of A on $H_1(\Sigma)$ extended as a derivation to a map on the exterior algebra, and the angle brackets refer to a certain determinant-like product of pairwise intersections of $A_{\circ}^{p}(c_i)$ with $A^{-1}(c_j)$. We take $A_{\circ}^{0} = \mathbb{I}$ and $\operatorname{Fix}(A^{(0)}) = 1$.

To be more explicit, for $\{a_i, b_i\}_{i=1}^k \in H_1(\Sigma)$ we define the notation $\langle a_1 \wedge \cdots \wedge a_k, b_1 \wedge \cdots \wedge b_k \rangle$ to mean

$$\sum_{\sigma \in S_k} (-1)^{\operatorname{sign}\sigma} (a_1.b_{\sigma(1)}) \cdots (a_k.b_{\sigma(k)}),$$

where $a_i.b_{\sigma(i)}$ is the intersection pairing on Σ. The expression in the proposition is to be evaluated by distributing $A_\circ^{n-k}$ across $c_1 \wedge \cdots \wedge c_N$ as an (unsigned) derivation, then using the above definition for each of the resulting terms.

To prove the proposition, we will count in a combinatorial fashion the points in the intersection $D \cap \Gamma_{A(N+n)}$, then see that the points are assigned the appropriate orientation. In what follows we will write points of symmetric products $\Sigma^{(k)}$ as sums of points of Σ.

Suppose, then, that $P = (\sum_{i=1}^N p_i + \sum_{j=1}^n q_j, \sum_{i=1}^N A(p_i) + \sum_{j=1}^n A(q_j))$ is a point of $\Gamma_{A(N+n)}$ that also lies on D. In this expression we take $p_i \in c_i$ for $i = 1, \ldots, N$; compare (8). It does not follow that $A(p_i) \in c_i$; rather, we know merely that N of the points on the "vertical" side of P (that is, the points making up the second coordinate of P, namely $\sum A(p_i) + \sum A(q_j)$) are on the c_k, in fact precisely one on each c_k by the definition of D. Once we have determined, given P, which N members of the list $A(p_1), \ldots, A(p_N), A(q_1), \ldots, A(q_n)$ are on $\bigcup c_k$, we know that the remaining n points, together with the q_j on the "horizontal" side form a point on the diagonal of $\Sigma^{(n)} \times \Sigma^{(n)}$.

Suppose to begin with that each $A(p_i)$ is on $\bigcup_k c_k$: then the remaining points q_j together with their images $A(q_j)$, lying as they do on Δ_n, form a fixed point of $A^{(n)}$. If M_0 is the (unsigned) number of points $\sum_i p_i \in \Sigma^{(N)}$ such that $p_i \in c_i$ and $A(p_i) \in \bigcup c_k$, we see that the number of points P of this form is

$$|\mathrm{Fix}|(A^{(n)}) \cdot M_0 \tag{10}$$

where by $|\mathrm{Fix}|$ we mean the "raw" unsigned count of fixed points. To calculate M_0, proceed as follows. For the first point p_1, we have $\sum_{j=1}^{N} \#(A(c_1) \cap c_j)$ possibilities, since $A(p_1)$ may be on any of the c_k. Given a choice for p_1, suppose that $A(p_1) \in c_1$. Then the number of possibilities for p_2 is $\sum_{j=2}^{N} \#(A(c_2) \cap c_j)$—note that $A(p_2)$ may not be on c_1 since P must be on D. And so forth: let us write I_{ij} for $\#A(c_i) \cap c_j$; then for the final calculation of M_0 we get

$$M_0 = \sum_{\sigma \in S_N} I_{1,\sigma(1)} I_{2,\sigma(2)} \cdots I_{N,\sigma(N)}.$$

Note that we could just as well set $I_{ij} = \#(c_i \cap A^{-1}(c_j))$.

Remark 5.3 *This assumes that $c_i \cap A^{-1}(c_j)$ consists of finitely many points. In Section 6 below we will prove a transversality result that justifies that assumption.*

Suppose now that P is as above, but on the "vertical" side the points $A(p_2), \ldots, A(p_N), A(q_1)$ are the ones that lie on $\bigcup c_k$. Then $A(p_1)$ must be in the diagonal part of P, and hence $A(p_1)$ must appear in the list $q_1, \ldots, q_n$ on the horizontal side. If $A(p_1) = q_1$, then $A^2(p_1) \in \bigcup c_k$, and the remaining diagonal points $(q_2, \ldots, q_n, A(q_2), \ldots, A(q_n))$ form a fixed point of $A^{(n-1)}$. To count points of this form, write $I_{ij}^{(1)} = \#(A^2(c_i) \cap c_j) = \#(A(c_i) \cap A^{-1}(c_j))$; then we see that the number of points P with $A(p_2), \ldots, A(p_N), A(q_1)$ on $\bigcup c_k$ and $A^2(p) = q_1$ is given by

$$|\mathrm{Fix}|(A^{(n-1)}) \sum_{\sigma \in S_N} I_{1,\sigma(1)}^{(1)} I_{2,\sigma(2)} \cdots I_{N,\sigma(N)}.$$

If we wish to allow any of the p_i to play the role of p_1 above (note that though they are constituents of a point in a symmetric product, the p_i are not interchangeable since we have assumed $p_i \in c_i$), we must allow the superscript (1) to appear on any of the I in the above expression and take another sum. Then we would obtain the number M_1 of points $\sum_i p_i \in \Sigma^{(N)}$ satisfying

- $p_i \in c_i$ for each i;

- there is a unique index j for which $A^2(p_j)$ is on $\bigcup c_k$;

- for all $i \neq j$ we have $A(p_i) \in \bigcup c_k$.

Finally we have an expression analogous to (10) for the number of points of the form under consideration, namely

$$|\mathrm{Fix}|(A^{(n-1)}) \cdot M_1.$$

One could continue in this vein, but instead of doing so we reorganize our count slightly. Instead of counting based on the number of $A(p_i)$ that are on $\bigcup c_k$, we note that by arguments similar to the elementary one above, for each p_i there exists a nonnegative integer α_i such that $A^{\alpha_i+1}(p_i) \in \bigcup c_k$. We arrange our count according to the quantity $\alpha = \alpha_1 + \cdots + \alpha_N$; above we have calculated the number of P with $\alpha = 0$ and with $\alpha = 1$. Note that since there are exactly $N + n$ points on the vertical side of P, N of which must be on the c_k and not contribute to a diagonal point, we must have $\alpha \leq n$. Likewise, given α, we see that there are exactly $n - \alpha$ of the q_j on the horizontal side not "used up" as the images of p_i under iterates of A—hence the remaining q_j and their images $A(q_j)$ form a fixed point of $A^{(n-\alpha)}$ Writing

$$I_{i,j}^{(r)} := \#(A^{r+1}(c_i) \cap c_j) = \#(A^r(c_i) \cap A^{-1}(c_j)),$$

we have that the total number of points P with fixed $\alpha = n - k$, say, is given by

$$|\mathrm{Fix}|(A^{(k)}) \left(\sum_{\substack{\alpha_1+\cdots+\alpha_N=n-k \\ \sigma\in S_N}} I_{1,\sigma(1)}^{(\alpha_1)} \cdots I_{N,\sigma(N)}^{(\alpha_N)} \right). \tag{11}$$

Finally, to complete the count of $\#(D \cap \Gamma_{A^{(N+n)}})$, we sum the above expression over all k from 0 to n.

It is now clear that the proposition will hold so long as the sign attached to P as an intersection point of D and $\Gamma_{A^{(N+n)}}$ is the following one: each p_i gives rise to an intersection point of $A^{\alpha_i}(c_i)$ with $A^{-1}(c_{\sigma(i)})$ for some permutation σ, and this intersection point has an associated sign ϵ_i. The permutation has a sign of its own, $(-1)^{|\sigma|}$, and the fixed point of $A^{(n-\alpha)}$ arising as above has a sign ϵ_{N+1} in the usual way. Associating the sign $(-1)^{|\sigma|}\epsilon_1 \cdots \epsilon_{N+1}$ to each of the points counted by (11) gives exactly the quantity on the right hand side of (9) (after performing the sum over k). We leave it to the reader to check that the sign of P that arises by considering P as an intersection point of D and $\Gamma_{A^{(N+n)}}$ is the one described here. $\square$

Proposition 5.4 *The n-th coefficient of $\rho(I) = \zeta(f)\det(d)$ is given by*

$$[\rho(I)]_n = \sum_{k=0}^{n} \mathrm{Fix}(f^{(k)})\langle f_\circ^{n-k}(c_1 \wedge \cdots \wedge c_N), A^{-1}(c_1) \wedge \cdots \wedge A^{-1}(c_N)\rangle. \quad (12)$$

Here $f^{(k)}$ is the induced action of f on the symmetric product $\Sigma^{(k)}$, and the angle brackets are defined as in proposition 5.2.

The proof is another direct calculation. First, it is not hard to calculate from the definition

$$\zeta(f) = \exp\left(\sum_{n=1}^{\infty} \mathrm{Fix}(f^n)\frac{t^n}{n}\right)$$

that the kth coefficient $[\zeta(f)]_k$ is given by

$$\sum_{\substack{k_1+\cdots+k_j=k \\ j=1,\ldots,k}} \frac{1}{j!}\left(\tfrac{1}{k_1}\mathrm{Fix}(f^{k_1})\right)\cdots\left(\tfrac{1}{k_j}\mathrm{Fix}(f^{k_j})\right),$$

One can see by an easy combinatorial argument that the above expression is simply $\mathrm{Fix}(f^{(k)})$.

Next we must calculate an arbitrary coefficient of the differential $d : M_1 \to M_2$ of the Morse complex associated to ϕ. This differential is represented by a matrix of rank N, and the entry in the (i,j)-th position is the series $\sum_{k\geq 0}(f^k(c_i).A^{-1}(c_j))t^k$ (note that the coefficient of the polynomial $a_{ij}(t)$ defined in section 4 is simply the intersection number used here). Some algebra shows that the k-th coefficient of the determinant of d is given by

$$\sum_{\substack{\sigma\in S_N \\ k_1+\cdots k_N=k}} (-1)^{|\sigma|}\left(f^{k_1}(c_1).(A^{-1}(c_{\sigma(1)})\right)\cdots\left(f^{k_N}(c_N).A^{-1}(c_{\sigma(N)})\right),$$

which is precisely equal to $\langle f_\circ^k(c_1\wedge\cdots\wedge c_N), A^{-1}(c_1)\wedge\cdots\wedge A^{-1}(c_N)\rangle$. Since $[\zeta(f)\det(d)]_n = \sum[\zeta(f)]_k[\det(d)]_{n-k}$, we are done. $\square$

6 Final Results

We have now proved the bulk of the main theorem (restated below); together with the TQFT results from section 2, in particular Theorem 2.2, the following result completes the proof of (the averaged version of) Hutchings' and Lee's conjecture.

Theorem 6.1 $\operatorname{Tr}\kappa = [\rho(I)]_n$.

From the work of previous sections we have:

$$\begin{aligned}
\operatorname{Tr}\kappa &= \sum_{\alpha}(-1)^{|\alpha|+N|\alpha|}(A^*c_N^* \wedge \cdots \wedge A^*c_1^* \wedge A^*(\alpha)) \cup (c_N^* \wedge \cdots \wedge c_1^* \wedge \alpha) \\
&= D.\Gamma_{A^{(N+n)}} \\
&= \sum_{k=0}^{n} \operatorname{Fix}(A^{(k)})\langle A_\circ^{n-k}(c_1 \wedge \cdots \wedge c_N), A_*^{-1}c_1 \wedge \cdots \wedge A_*^{-1}c_N\rangle \\
&= \sum_{k=0}^{n} \operatorname{Fix}(f^{(k)})\langle f_\circ^{n-k}(c_1 \wedge \cdots \wedge c_N), A^{-1}(c_1) \wedge \cdots \wedge A^{-1}(c_N)\rangle \qquad (13) \\
&= [\rho(I)]_n.
\end{aligned}$$

The only part of this left to prove is (13). We will see that the replacing of A by f is legal in two steps: first we show that the right hand side of formula (9) is unchanged if we use a certain non-generic flow map $\bar{A} : \Sigma \to \Sigma$ in place of the diffeomorphism A. That is, we will see that $\Gamma_{A^{(N+n)}}.D = [\rho(I_{\bar{A}})]_n$, where $[\rho(I_{\bar{A}})]_n$ is given by a formula analogous to (12) above.

Second, we show that the invariant $\rho(I)$ computed using the flow $\bar{A}$ is the same as that computed using a generically perturbed flow—then the fact, whose proof was sketched by Hutchings and Lee and set down in final form in [4], that I is a topological invariant will complete the proof.

We define $\bar{A}$ as follows. Our cobordism W is made up of two (identical) parts $W_1 \cup_A W_2$; on W_1 we take the map $F_0 : \Sigma_0 \to \Sigma$ induced by gradient flow of the original Morse function $\phi : Y \to S^1$ (we are being sloppy here: F_0 is only defined away from the descending manifolds of the index 1 critical points). On W_2 we take the identical flow—that is, we consider W_2 simply as another copy of W_1 and use the same flow we used in that case. This latter flow map we denote by $F_1 : \Sigma \to \Sigma_1$; note that after identifying $W_1 = W_2$ we have $F_1 \circ F_0 = id = F_0 \circ F_1$ where the compositions are defined. We then set $\bar{A} = F_0 F_1 A : \Sigma \to \Sigma$; apparently $\bar{A}$ agrees with the restriction of A to the complement of the collection of curves $A^{-1}(c_i)$, the descending manifolds of index-2 critical points intersected with Σ.

To see now that $\Gamma_{A(N+n)}.S = \sum_{k=0}^{n} \mathrm{Fix}(\bar{A}^{(k)})\langle \bar{A}_\circ^{n-k}(c_1 \wedge \cdots \wedge c_N), A_*^{-1}c_1 \wedge \cdots \wedge A_*^{-1}c_N\rangle$, it will suffice to show that

i. $\mathrm{Fix}(A^{(m)}) = \mathrm{Fix}(\bar{A}^{(m)})$ for all $m \leq n$, and

ii. $A^m(c_i).A^{-1}(c_j) = \bar{A}^m(c_i).A^{-1}(c_j)$ for all i and j and all m, $0 \leq m \leq n$.

We will in fact show that these conditions are true after perturbation of A:

Lemma 6.2 *We can adjust A so that*

a) All fixed points of A^k, $1 \le k \le n$, are isolated and all intersections $A^k(c_i) \cap A^{-1}(c_j)$, $1 \le k \le n$, are transverse.

b) No fixed point of A^k occurs on $\bigcup c_i$ for any $1 \le k \le n$.

c) All fixed points of A^k occur in the domain of $\bar{A}^k$, $1 \le k \le n$.

d) All intersections $c_i \cap A^{-k-1}(c_j)$ occur in the domain of $\bar{A}^k$, for $k \le n$.

From this, it is clear that (i) and (ii) above hold.

Proof: That we can arrange for part (a) to hold we take as obvious: this is a standard transversality assumption.

The remaining arguments are all similar in flavor; each involves considering a point p that fails to meet the criterion in question and giving an explicit modification of A to repair the defect. At each stage, we must check that the modifications we make to A preserve the properties we obtained in previous steps.

For part (b), consider a point p that is a fixed point of A^k, with $p \in c_i$. We assume that k is minimal, i.e., p is not fixed by any smaller power of A, and therefore we may find neighborhoods U_j of the iterates $A^j(p)$ that are all disjoint from one another. The curve $A^{-k}(c_i)$ passes through p, and by (a) we may take this intersection to be transverse. Let X be a vector field that is supported in a small neighborhood $V \subset U_0$ of p and whose time-1 flow φ_X moves p off of c_i parallel to $A^{-k}(c_i)$: in particular, we can arrange that $\varphi_X(c_i) \cap A^{-k}(c_i)$ consists of the single point $\varphi(p)$. We replace A by $A' = A\varphi_X$; clearly (since φ_X is supported on U_0) we have $A'^k = A^k\varphi$ on U_0.

It is clear that the perturbation $A \to A'$ cannot introduce new fixed points of order less than k by choice of the neighborhood in which the perturbation is supported. We must see, then, that A'^k has no fixed points on c_i. By construction of φ_X, we have $\#(A'^k(c_i) \cap c_i) = \#(A^k\varphi_X(c_i) \cap c_i) = \#(\varphi_X(c_i) \cap A^{-k}(c_i)) = 1$; suppose this intersection q is in fact a fixed point of A'^k. Then $q = A'^k(c_i) \cap c_i = A^k(\varphi_X(c_i) \cap A^{-k}(c_i)) = A^k(\varphi_X(p))$, so that $p = (A^k\varphi_X)^{-1}(q) = q$. This is a contradiction, for p is clearly not fixed by A'^k. Hence there are no fixed points of A'^k on $c_i \cap U_0$.

By repeating this argument for all fixed points of A^k in succession (these are isolated and therefore the above argument may be carried out independently for each), we conclude that A may be modified so that all fixed points of A^k occur off of the c_i, without introducing new fixed points of lower order. Hence the same argument shows that we can modify A to have no fixed points of order less than $n+1$ that lie on $\bigcup c_i$.

Claim (c) is immediate from (b): if p is a fixed point of A^k, then since the domain of $\bar{A}$ is the complement of $\bigcup c_i$ we see that for p to fail to be in the domain of $\bar{A}^k$ it must be the case that $A^j(p) \in \bigcup c_i$ for some $j < k$. Naturally $q = A^j(p)$ is fixed under A^k, and hence is a fixed point of A^k lying on $\bigcup c_i$, in contradiction to part (b).

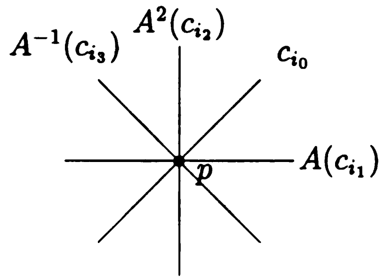
Finally, we must prove that (d) can be arranged without disturbing the previous work. Another way to state the proposition is that we would like to ensure that all the intersections

$$I_{\alpha\beta} = (\bigcup c_i) \cap (A^\alpha(\bigcup c_i)) \cap \left(A^\beta(\bigcup c_i)\right)$$

are empty, for $\alpha < \beta \leq n$. By (a), we may assume that all the $I_{\alpha\beta}$ consist of a finite number of points. We proceed inductively, ordering the $I_{\alpha\beta}$ first by β, then by α: that is, $I_{\alpha_1\beta_1} < I_{\alpha_2\beta_2}$ if $\beta_1 < \beta_2$ or $\beta_1 = \beta_2$ and $\alpha_1 < \alpha_2$. Begin, then, with $I_{1,2}$, and suppose p is in this intersection—then $p = A^2(q_2) = A(q_1)$, with, say, $q_1 \in c_{i_1}$, $q_2 \in c_{i_2}$, and $p \in c_{i_0}$. We take a neighborhood U of p such that:

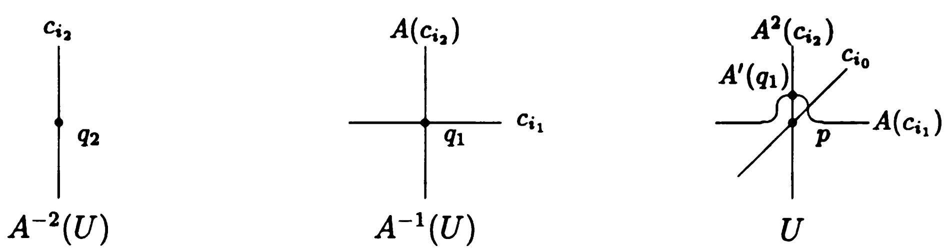
1. If $A(p) \in c_{i_3}$ for some i_3 then the configuration of curves $A^2(c_{i_2})$, $A(c_{i_1})$,

c_{i_0}, and $A^{-1}(c_{i_3})$ appear in U as four arcs intersecting transversely at the single point p:



2. If $A(p) \notin \bigcup c_i$ then the curves $A^2(c_{i_2})$, $A(c_{i_1})$, and c_{i_0} appear in U as three arcs intersecting transversely at p, as above.

Perturb A in a small neighborhood of q_1 using a flow parallel to $A(c_{i_2})$ so that under the perturbed map A' the point q_1 misses p. The sequence of neighborhoods $A^{-2}(U)$, $A^{-1}(U)$, U now appear as



Apparently p is no longer in $I_{1,2}$, but if case (1) above holds then $A(p)$ is in $I_{1,3}$. This is not a problem, since $I_{1,3}$ is still to be dealt with in our inductive procedure (if case (2) holds, then p can at worst give rise to a point of $I_{r,r+2}$

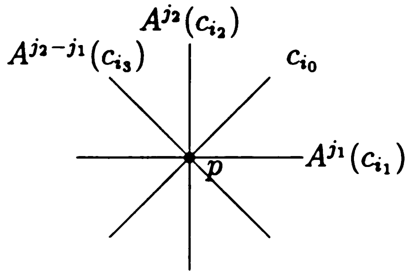
for $r > 1$, in the case that $A^r(p)$ is on $\bigcup c_i$—this set will also be dealt with later). Similarly, we have two more points $p' = A'(q)$ and $p'' = A'(c_{i_1}) \cap c_{i_0}$ as labeled above that may give rise to points in $I_{\alpha\beta}$. Indeed, if $A'^s(p') \in \bigcup c_i$ then $A'^s(p') \in I_{s+1,s+2}$, while if $A'^t(p'') \in \bigcup c_i$ then $A'^t(p'') \in I_{t,t+1}$. Note, however, that by our choice of U we must have $s, t \geq 2$, so these points are also in $I_{\alpha\beta}$ that will be considered later in the inductive process. Finally, if case (1) above holds, then we have a third new point $p''' = A'^2(c_{i_1}) \cap c_{i_3}$ that can at worst give rise to a point of $I_{q,q+2}$ for some $q \geq 1$, in the case that $A'^q(p''') \in \bigcup c_i$. Again this set will be addressed later.

The above argument shows that if $I_{1,2}$ is nonempty, we may perturb A in a neighborhood of a preimage of each of its points and by doing so empty out the set $I_{1,2}$, at the cost of possibly adding points to other $I_{\alpha\beta}$ that will be considered later. Note that this process cannot interfere with previous work since we are working near points in the orbit of a point on $\bigcup c_i$, which cannot be the same as points in the orbit of a fixed point by part (b).

The inductive step is essentially the same: suppose $I_{\alpha\beta}$ is empty for all $\beta < j_2$, and let $p \in I_{j_1 j_2}$ with $p = A^{j_2}(q_2) = A^{j_1}(q_1)$ for $q_1 \in c_{i_1}$, $q_2 \in c_{i_2}$ and $p \in c_{i_0}$. It may be the case that some image $A^k(p) \in c_{i_3}$ for some i_3, but since if this were true then $A^k(p) \in I_{k,k+j_1}$, we must have $k \geq j_2 - j_1$

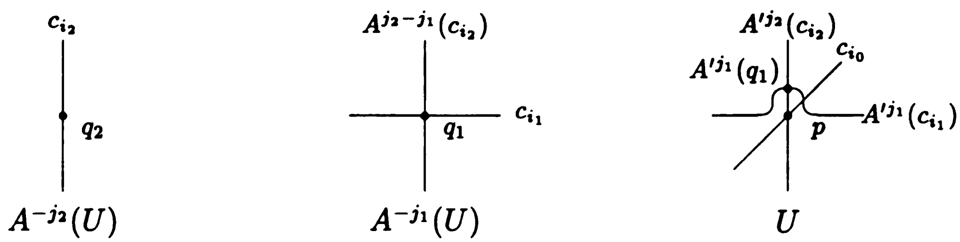
by the inductive hypothesis. Cases (1) and (2) above for the choice of the neighborhood U of p now read

1. If $A^{j_2-j_1}(p) \in c_{i_3}$ for some i_3 then the configuration of curves $A^{j_2}(c_{i_2})$, $A^{j_1}(c_{i_1})$, c_{i_0}, and $A^{j_1-j_2}(c_{i_3})$ appear in U as four arcs intersecting transversely at the single point p:



2. If $A^{j_1-j_2}(p) \notin \bigcup c_i$ then the curves $A^{j_2}(c_{i_2})$, $A^{j_1}(c_{i_1})$, and c_{i_0} appear in U as three arcs intersecting transversely at p, as above.

Again we perturb A to A' in a neighborhood of q_1 to obtain a similar perturbed picture as above:



In the worst cases, then, we see that we will have $A'^s(p) \in I_{s,s+j_2}$ for some $s \geq j_2 - j_1 + 1$, or $A'^r(p') \in I_{r+j_1,r+j_2}$ for some $r \geq j_2 - j_1 + 1$, or $A'^t(p'') \in I_{t,t+j_1}$ for some $t \geq j_2 - j_1 + 1$. Finally, in case (1), we may also have a point $p''' = A'^{j_2}(c_{i_1}) \cap c_{i_3}$ with $A'^q(p''') \in I_{q,q+j_2}$. All four of these possibilities are addressed later in the inductive process, so again we may empty $I_{j_1 j_2}$ without disturbing the status of $I_{\alpha\beta}$ previously emptied, nor the work of part (b) and (c). □

Finally, we must show that we can perturb $\bar{A}$ to a generic flow and obtain the same invariant I. It is apparent that the reason $\bar{A}$ is non-generic is that it connects the upward flows from the index-2 critical points, $\mathcal{A}(y)$, to the downward flows from the index-1 critical points $\mathcal{D}(x)$. To make it generic, therefore, we need only perturb the flow in a neighborhood of $\Sigma_0 = \Sigma_1$, in such a way that these two types of flow miss one another. In fact, we may choose to modify the flow in a small neighborhood of the N points in which $\mathcal{A}(y)$ intersect Σ_0. Such a perturbation, however, cannot affect either the number of fixed points of $\bar{A}^k$ for $k \leq n$, or the number of intersections between $\bar{A}^k(c_i)$ and c_j—simply because we can take the support of the perturbation to be away from any of the corresponding flow lines of $\bar{A}$, there being finitely many such flows. Therefore, $[\rho(I_{\bar{A}})]_n$, which is calculated as a combination of these two quantities, is unchanged by this perturbation. This completes the last step in the proof of (13).

APPENDIX

A Appendix: Technical Results

It is the purpose of this final section to provide some justification for some of the results used in previous parts of the work. In particular, the results of this section should suffice to prove both the fact that the Seiberg-Witten map κ_W satisfies the composition law (5) for topological quantum field theories, and Theorem 2.2 showing the trace of κ_W gives the Seiberg-Witten invariant for the closed manifold X. We do not, however, provide complete proofs of these facts here.

To begin with, we review the Seiberg-Witten equations on an oriented Riemannian 3-manifold X. We consider the Seiberg-Witten equations perturbed by a 2-form $\eta \in \Omega^2(\mathbb{R} \times \Sigma)$:

$$\begin{aligned} D_A \Psi &= 0 \\ \star(F_A + i\eta) &= i\tau(\Psi, \Psi) \end{aligned} \tag{14}$$

where the notation is the same as in Section 2. For convenience, let us fix a base connection A_0 on the line bundle L. We can think of the solutions of (14) as $\mathfrak{S}^{-1}(0,0)$ where $\mathfrak{S}$ is defined by

$$\begin{aligned} \mathfrak{S} : \Gamma(S) \oplus \Omega^1(X; i\mathbb{R}) &\longrightarrow \Gamma(S) \oplus \Omega^1(X; i\mathbb{R}) \\ (\Psi, a) &\longmapsto (D_{A_0}\Psi + a.\Psi,\ \star(F_{A_0} + da + i\eta) - i\tau(\Psi, \Psi)) \end{aligned}$$

By way of notation, we will sometimes write $\mathcal{C} = \mathcal{C}(X) = \Gamma(S) \oplus \Omega^1(X; i\mathbb{R})$ for the space of Seiberg-Witten configurations on X, and $\mathcal{C}^* = \mathcal{C}^*(X)$ for those configurations (Ψ, a) with Ψ not identically zero—that is, $\mathcal{C}^*$ is the space of *irreducible* configurations. The linearization $L_{(\Psi,a)}$ of $\mathfrak{S}$ at a point (Ψ, a) is given by

$$\begin{aligned} L_{(\Psi,a)} : \Gamma(S) \oplus \Omega^1(X; i\mathbb{R}) &\rightarrow \Gamma(S) \oplus \Omega^1(X; i\mathbb{R}) \\ L_{(\Psi,a)}(\phi, b) &= (D_{A_0}\phi + a.\phi + b.\Psi, \star db - 2i\mathrm{Re}\, \tau(\Psi, \phi)) \end{aligned}$$

where $\mathrm{Re}\, \tau(\Psi, \phi)$ denotes the element $\frac{1}{2}(\tau(\Psi, \phi) + \tau(\phi, \Psi))$ of $\Omega^1(X) \subset \Omega^1_{\mathbb{C}}(X)$. It can be shown that for generic metric and η the moduli space of solutions modulo gauge is smoothly cut out by the equations—that is, the linearization is always onto.

When we consider the Seiberg-Witten equations on a cylinder $\mathbb{R} \times \Sigma$ and in appropriate gauge, the above linearization becomes a 0-th order perturbation of an operator of Atiyah-Patodi-Singer (APS) type. Recall that a linear first order elliptic differential operator $D : \Gamma(X; E) \rightarrow \Gamma(X; E)$ for $X = \mathbb{R} \times \Sigma$ is said to be of APS type if it can be written as $\frac{\partial}{\partial t} + A$ for a self-adjoint elliptic operator $A : \Gamma(\Sigma; E) \rightarrow \Gamma(\Sigma; E)$. We now recall the theory we will need.

A.1 Cylinders and Cylindrical Ends: Linear Theory

To begin with, we work on $\mathbb{R} \times \Sigma$ with a product metric. We are in the situation of the previous work, i.e., we have a hermitian bundle E_0 over Σ whose pullback over $\mathbb{R} \times \Sigma$ will be denoted by E, and an operator D on sections of E over $\mathbb{R} \times \Sigma$ that is of APS type. Assume E is also endowed with a hermitian connection ∇. Fix once and for all a C^∞ function $\tau(t,x)$ that agrees with $|t|$ for $|t| > 1$, and is within some small distance of $|t|$ in the C^0-norm on $[-1,1] \times \Sigma$. Since A is elliptic and self-adjoint, its spectrum is discrete. Letting $\{\lambda\}$ denote the set of eigenvalues of A, there exists a number $\lambda_0 = \min_{\lambda \neq 0}(|\lambda|)$. Fix a real number $0 < \delta < 2\lambda_0$. For $0 \leq k \in \mathbb{Z}$, define the weighted L^2-norm $||\cdot||_{k,\delta}$ on sections of E by

$$||\sigma||_{k,\delta}^2 = \int_{\mathbb{R}\times\Sigma} \left(\sum_{l=0}^{k} |\nabla^l \sigma|^2 \right) e^{\delta\tau} \mathrm{d}vol_{\mathbb{R}\times\Sigma}.$$

We then let H^k_δ denote the closure of the space of smooth, compactly supported sections of E under this norm.

We say a section σ of E is an *extended* H^k_δ section if it is locally in H^k_δ, and if for $|t|$ sufficiently large σ can be written in the form

$$\sigma(t,x) = \sigma_0(t,x) + \pi^*\hat{\sigma}(t,x)$$

where $\sigma_0 \in H^k_\delta$, $\pi : \mathbb{R} \times \Sigma \to \Sigma$ is the projection, and $\hat{\sigma} \in \ker(A)$. We denote the space of extended H^k_δ sections by $H^k_{\delta,e}$. The following result is due to Atiyah-Patodi-Singer [1].

Lemma A.1 *1. The operator D above induces a Fredholm map $H^k_{\delta,e} \to H^{k-1}_\delta$.*

2. D admits a right inverse $P : H^{k-1}_\delta \to H^k_{\delta,e}$.

Sketch of Proof: We begin by applying the spectral theorem to A to obtain a complete orthonormal basis of eigenfunctions $\phi_\lambda \in \Gamma(\Sigma; E_0)$. Then any section σ of E can be written as $\sigma(t,x) = \sum\limits_\lambda \sigma_\lambda(t)\phi_\lambda(x)$ and the operator $D = \frac{\partial}{\partial t} + A$ appears as

$$D\sigma = \sum_\lambda \left(\frac{\mathrm{d}\sigma_\lambda}{\mathrm{d}t} + \lambda\right)\phi_\lambda.$$

Following [1], we solve the equation $D\sigma = \rho$ by the explicit formulas

$$P_\lambda\rho_\lambda(t) = \begin{cases} \displaystyle\int_{-\infty}^t e^{\lambda(s-t)}\rho_\lambda(s)\mathrm{d}s & (\lambda \geq 0) \\ \displaystyle -\int_t^\infty e^{\lambda(s-t)}\rho_\lambda(s)\mathrm{d}s & (\lambda < 0) \end{cases}$$

Equivalently, if we write $\epsilon(t) = \chi_{[0,\infty)}(t)$ for the characteristic function of the right half-line, and set

$$f_\lambda(t) = \begin{cases} e^{-\lambda t}\epsilon(t) & (\lambda \geq 0) \\ -e^{-\lambda t}\epsilon(-t) & (\lambda < 0) \end{cases}$$

we define $P_\lambda \rho_\lambda = f_\lambda * \rho_\lambda$ where $*$ denotes convolution:

$$(g * h)(t) = \int_{-\infty}^{\infty} g(t-s)h(s)\mathrm{d}s.$$

To invert D, define $P\rho = \sum_\lambda (P_\lambda \rho_\lambda)\phi_\lambda$. One easily checks that $\sigma = P\rho$ formally solves $D\sigma = \rho$; it remains to determine the mapping properties of P. Now for $\lambda > 0$,

$$\|P_\lambda \rho_\lambda\|_{0,\delta}^2 = \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} e^{-\lambda(t-s)} \epsilon(t-s)\rho_\lambda(s)\mathrm{d}s \right|^2 e^{\delta\tau(t)}\mathrm{d}t,$$

where $\tau(t) \approx |t|$. Ignoring a small error factor, we replace $\tau(t)$ by $|t|$, and write

$$\begin{aligned}
\|P_\lambda \rho_\lambda\|_{0,\delta}^2 &= \int_0^{\infty} \left| \int_{-\infty}^{\infty} e^{\frac{\delta}{2}t + \lambda s - \lambda t} \epsilon(t-s)\rho_\lambda(s)\mathrm{d}s \right|^2 \mathrm{d}t \\
&\quad + \int_{-\infty}^{0} \left| \int_{-\infty}^{\infty} e^{-\frac{\delta}{2}t + \lambda s - \lambda t} \epsilon(t-s)\rho_\lambda(s)\mathrm{d}s \right|^2 \mathrm{d}t \\
&= \int_0^{\infty} \left| \int_{-\infty}^{\infty} e^{(\frac{\delta}{2} - \lambda)(t-s)} \epsilon(t-s) e^{\frac{\delta}{2}s}\rho_\lambda(s)\mathrm{d}s \right|^2 \mathrm{d}t \\
&\quad + \int_{-\infty}^{0} \left| \int_{-\infty}^{\infty} e^{-(\frac{\delta}{2} + \lambda)(t-s)} \epsilon(t-s) e^{-\frac{\delta}{2}s}\rho_\lambda(s)\mathrm{d}s \right|^2 \mathrm{d}t \\
&\leq \left\| e^{(\frac{\delta}{2} - \lambda)t}\epsilon(t) * e^{\frac{\delta}{2}t}\rho_\lambda(t) \right\|_{L^2}^2 + \left\| e^{-(\frac{\delta}{2} + \lambda)t}\epsilon(t) * e^{-\frac{\delta}{2}t}\rho_\lambda(t) \right\|_{L^2}^2
\end{aligned}$$

Applying Young's inequality $\|f * g\|_{L^2} \leq \|f\|_{L^1}\|g\|_{L^2}$, we see that the above is no more than

$$\tfrac{1}{(\frac{\delta}{2} - \lambda)^2}\|\rho_\lambda\|_{0,\delta}^2 + \tfrac{1}{(\frac{\delta}{2} + \lambda)^2}\|\rho_\lambda\|_{0,\delta}^2$$

which finally gives the bound

$$\|P_\lambda \rho_\lambda\|_{0,\delta}^2 \leq \frac{2}{(\frac{\delta}{2} - |\lambda|}) ^2 \|\rho_\lambda\|_{0,\delta}^2 \tag{15}$$

for $\lambda > 0$. An entirely analogous calculation gives an identical bound for $\|P_\lambda \rho_\lambda\|_{0,\delta}$ with $\lambda < 0$. For the case $\lambda = 0$, we have

$$P_\lambda \rho_\lambda(t) = \int_{-\infty}^{\infty} \epsilon(t-s) \rho_\lambda(s) \mathrm{d}s.$$

Since $\rho_\lambda \in H_\delta^0$, this integral has a well-defined limiting value k as $t \to \infty$. In fact, if we choose a smooth cutoff function $0 \leq \nu(t) \leq 1$ with $\nu(t) = 0$ for $t < -T \ll 0$ and $\nu(t) = 1$ for $t > T \gg 0$, the section

$$P_\lambda \rho_\lambda \phi_\lambda = (P_\lambda \rho_\lambda - \nu k) \phi_\lambda + \nu k \phi_\lambda$$

is in $H_{\delta,e}^0$. Moreover, another calculation of the type made above leads to estimates for the "non-extended part" of $P_\lambda \rho_\lambda$ near the ends of the cylinder, of the form

$$\|P_\lambda \rho_\lambda - \nu k\|_{H_\delta^0([T,\infty))} \leq \tfrac{2}{\delta} \|\rho_\lambda\|_{0,\delta} \tag{16}$$

and similar on $(-\infty, -T]$.

Remark A.2 *It is at this point that the use of the weighted Sobolev norm $\|\cdot\|_{k,\delta}$ is required: without the exponential decay guaranteed by the condition*

$\rho_\lambda \in H^0_\delta$, we do not obtain a bound even on the non-extended part of $P_\lambda \rho_\lambda$ for $\lambda = 0$, as is apparent by letting $\delta \to 0$ in the above estimate.

Our object is to see that P maps $H^0_\delta \to H^1_{\delta,e}$. First, we note that since D is a first order elliptic operator, we may define the H^1_δ norm using D rather than ∇: explicitly,

$$\|f\|^2_{1,\delta} = \|f\|^2_{0,\delta} + \|\frac{\mathrm{d}}{\mathrm{d}t} f\|^2_{0,\delta} + \|Af\|^2_{0,\delta}.$$

Then, for a section $\rho = \sum_{\lambda \neq 0} \rho_\lambda \phi_\lambda$ that does not involve the 0-eigenvectors, we have

$$\begin{aligned}
\|P\rho\|^2_{1,\delta} &= \|\sum P_\lambda \rho_\lambda\|^2_{0,\delta} + \|\sum \frac{\mathrm{d}}{\mathrm{d}t} P_\lambda \rho_\lambda\|^2_{0,\delta} + \|\sum A(P_\lambda \rho_\lambda \phi_\lambda)\|^2_{0,\delta} \\
&\leq \frac{2}{(\frac{\delta}{2} - |\lambda|)^2} \|\rho_\lambda\|^2_{0,\delta} + \sum \left(\|\rho_\lambda\|^2_{0,\delta} + \lambda^2 \|P_\lambda \rho_\lambda\|^2_{0,\delta} \right) \\
&\quad + \sum \lambda^2 \|P_\lambda \rho_\lambda\|^2_{0,\delta} \\
&\leq 2 \sum \frac{1 + \lambda^2}{(\frac{\delta}{2} - |\lambda|)^2} \|\rho_\lambda\|^2_{0,\delta} \\
&\leq C(\delta, \lambda_0) \sum (1 + \frac{1}{\lambda^2}) \|\rho_\lambda\|^2_{0,\delta},
\end{aligned}$$

for a constant $C(\delta, \lambda_0)$ depending only on δ and the number λ_0, being the minimum of $|\lambda|$ for nonzero eigenvalues λ. The sum converges since $\rho \in H^0_\delta$ and the eigenvalues λ of a self-adjoint elliptic operator A may not accumulate at 0.

One makes a similar calculation using the estimate (16) to bound the H^1_δ norm of $(P_\lambda \rho_\lambda - \nu k)\phi_\lambda$ with $\lambda = 0$ on the ends of the cylinder, and we infer that

$$P : H^0_\delta \to H^1_{\delta,e}$$

as required. Given the estimates above, we can use the same argument as in [1] to see that P extends to a continuous map $H^{k-1}_\delta \to H^k_{\delta,e}$ for general k. The existence of the right inverse P and the observation that $\ker(D) \cong \ker(A)$ is finite-dimensional completes the proof of the lemma. □

We turn now to the case of a cylindrical-end manifold. Explicitly, let W be a compact Riemannian manifold with boundary $\partial W = \Sigma$, and complete W by adding a cylinder $[0, \infty) \times \Sigma$ to the boundary. We denote the completed space by $\hat{W}$. For convenience, we will assume that Σ has a collar neighborhood $[-1, 0] \times \Sigma \subset W$, although this restriction is not really necessary (see [1]). We suppose that there is a bundle E over W restricting to $E_0 \to \Sigma$ over $\partial W = \Sigma$, and extend E over $\hat{W}$ in the natural manner. We consider an operator D on sections of E that is of APS type over $[-1, \infty) \times \Sigma$. The function $\tau(t)$ from above is defined on $[0, \infty) \times \Sigma$ and we may assume that defining $\tau(w) = 0$ for $w \in W$ extends τ smoothly to a function on $\hat{W}$, which we also denote by τ. Using this function, we may define the H^k_δ norm on $\hat{W}$

just as we did for the cylinder and again consider the space $H^k_{\delta,e}$ of extended H^k_δ sections of E.

Lemma A.3 *On $\hat{W}$, D induces a Fredholm operator $H^k_\delta \to H^{k-1}_{\delta,e}$.*

Sketch of Proof We must produce a parametrix for D on $\hat{W}$; that is an operator P so that DP and PD differ from the identity by a compact operator. The argument is a fairly standard "patching" technique, as follows. First, note that $D|W$ extends naturally to an operator D_1 on the double $\bar{W} = W \cup_\Sigma W$. Since D_1 is an elliptic operator on a closed manifold, it admits a parametrix P_1. Let us write D_2 for the operator D on $\mathbb{R} \times \Sigma$; it was proved above that D_2 has a parametrix (indeed, a right inverse) P_2. Now choose a partition of unity $\{\phi_1, \phi_2\}$ subordinate to the covering $\{U_1, U_2\}$ of $\hat{W}$ given by $U_1 = W$, $U_2 = [-1, \infty) \times \Sigma$. For notational convenience, we choose ϕ_1, ϕ_2 in such a way that $\phi_1^2 + \phi_2^2 \equiv 1$. Now define

$$P = \phi_1 P_1 \phi_1 + \phi_2 P_2 \phi_2,$$

thinking of ϕ_i as multiplication operators and omitting the obvious restrictions, extensions, and identifications necessary for this statement to make sense—e.g., for a section σ on $\hat{W}$, $\phi_1 \sigma$ vanishes on $[0, \infty) \times \Sigma$ and therefore may be thought of as a section on the double $\bar{W}$ by identifying W with one side of $\bar{W}$ and extending $\phi_1\sigma$ by zero to all of $\bar{W}$. Likewise, we can "transfer" ϕ_1 to $\bar{W}$ by extension by 0, and then for a section ρ over $\bar{W}$, $\phi_1\rho$ vanishes on one half of $\bar{W}$ and thus can be "transferred" back to $\hat{W}$ in the obvious way.

Now we calculate, writing $\sigma_D : \Omega^1(\hat{W}) \to \mathrm{End}(E)$ for the symbol of D:

$$\begin{aligned} DP &= D(\phi_1 P_1 \phi_1 + \phi_2 P_2 \phi_2) \\ &= \sigma_D(\mathrm{d}\phi_1)P_1\phi_1 + \phi_1 D(P_1\phi_1) \\ &\quad + \sigma_D(\mathrm{d}\phi_2)P_2\phi_2 + \phi_2 D(P_2\phi_2) \end{aligned}$$

Since D agrees with D_1 on $W = \mathrm{supp}(\phi_1)$, and with D_2 on $[-1,\infty) \times \Sigma = \mathrm{supp}(\phi_2)$, the second term becomes $\phi_1(\mathbb{I} + K)\phi_1$ where K is a compact operator, and the fourth term is just ϕ_2^2. Thus

$$DP = \mathbb{I} + \sigma_D(\mathrm{d}\phi_1)P_1\phi_1 + \sigma_D(\mathrm{d}\phi_2)P_2\phi_2 + \phi_1(K\phi_1). \tag{17}$$

We must check that the terms following $\mathbb{I}$ form a compact operator. Now, we may assume that $|\mathrm{d}\phi_1| \leq 1$, so we calculate for a section $\rho \in H^0_\delta$ over $\hat{W}$,

$$\begin{aligned} \|\sigma_D(\mathrm{d}\phi_1)P_1\phi_1\rho\|_{L^2_1(\hat{W})} &\leq c\|P_1(\phi_1\rho)\|_{L^2_1(\bar{W})} \\ &\leq c\|\phi_1\rho\|_{L^2(\bar{W})} \\ &= c\|\phi_1\rho\|_{H^0_\delta(\hat{W})} \\ &\leq c\|\rho\|_{H^0_\delta(\hat{W})}, \end{aligned}$$

for universal constants c. Here we have used that $P_1 : L^2(\bar{W}) \to L^2_1(\bar{W})$ is continuous. Therefore, the map $\rho \mapsto \sigma_D(\mathrm{d}\phi_1)P_1\phi_1\rho$ is a bounded map $H^0_\delta \to H^1_\delta$, and hence induces a compact operator $H^0_\delta \to H^0_\delta$. The remaining terms in (17) are dealt with in a similar manner, and an analogous calculation also shows that PD is the identity modulo compact operators. This gives the lemma. $\square$

Finally, we consider the case of two cylindrical end manifolds, and the linear gluing problem associated to our elliptic operators. Let W_1, W_2 be compact Riemannian manifolds with $\partial W_1 = \partial W_2 = \Sigma$, and form

$$\hat{W}_1 = W_1 \cup ([0, \infty) \times \Sigma)$$

$$\hat{W}_2 = ((-\infty, 0] \times \Sigma) \cup W_2.$$

Assume we have bundles E_1 and E_2 over W_1 and W_2 respectively, that each restrict to E_0 over Σ and extend in the usual way to bundles over $\hat{W}_1$ and $\hat{W}_2$, also denoted E_i, $i = 1, 2$. Let D_i, $i = 1, 2$, be operators on sections of E_i that each restrict to the ends of $\hat{W}_i$ to $\frac{\mathrm{d}}{\mathrm{d}t} + A$ for fixed A as above. For T a nonnegative real number, we can form $W_T = W_1 \cup [-T, T] \times \Sigma \cup W_2$. Then W_T has a natural bundle E_T that is isomorphic to E_i over W_i and to $\pi^* E_0$ over $[-T, T] \times \Sigma$ (recall $\pi : [-T, T] \times \Sigma \to \Sigma$ is the projection). Likewise, we obtain an operator D_T on sections of E_T. To fit into our previous weighted

Sobolev picture, we define the function τ_T on W_T using the functions τ_i on the cylindrical-end manifolds $\hat{W}_i$ as follows:

$$\begin{aligned}
\tau_T(w_i) &= \tau_i(w_i) = 0 && \text{for } w_i \in W_i \\
\tau_T(t,x) &= \tau_1(t+T,x) && \text{for } t \in [-T,0],\ x \in \Sigma \\
\tau_T(t,x) &= \tau_2(t-T,x) && \text{for } t \in [0,T],\ x \in \Sigma
\end{aligned}$$

In fact we take τ_T to be a C^∞ function that is C^0-close to the one described above, which is only continuous on $\{0\} \times \Sigma$. Thus $\tau_T \equiv 0$ on W_i, $\tau_T(t,x) \approx T + t$ for $t \in [-T,0]$, and $\tau_T(t,x) \approx T - t$ for $t \in [0,T]$. Using τ_T we can define the H^k_δ norm $\|\cdot\|_{k,\delta}$ on W_T, by the formula

$$\|\sigma\|_{k,\delta} = \int_{W_T} \left(\sum_{i=0}^{k} |\nabla^i \sigma|^2 \right) e^{\delta \tau_T} dvol.$$

We define the space of extended H^k_δ sections, $H^k_{\delta,e}$, as follows. Fix a cutoff function ν_1 on $\hat{W}_1$, where $\nu = 0$ on $W_1 \backslash ([-1,0] \times \Sigma)$ and $\nu_1 = 1$ on $[0,\infty) \times \Sigma$. For any element $\hat{\psi} \in \ker(A)$, there is a natural extension of $\nu\pi^*\hat{\psi}$ to $\hat{W}_1$ which will also be written $\nu\pi^*\hat{\psi}$. A section ψ of E_1 is said to be an extended H^k_δ section, $\psi \in H^k_{\delta,e}$, if ψ can be written in the form

$$\psi = \psi' + \nu\pi^*\hat{\psi} \tag{18}$$

for some $\hat{\psi} \in \ker(A)$, and some $\psi' \in H^k_\delta(\hat{W}_1)$. Note that we have a natural map $H^k_{\delta,e} \to \ker(A)$ by $\psi \mapsto \hat{\psi}$. A similar definition using a cutoff function

ν_2 holds for $H^k_{\delta,e}(\hat{W}_2)$. We define the $H^k_{\delta,e}(\hat{W}_1)$ norm of $\psi \in H^k_{\delta,e}(\hat{W}_1)$ by

$$\|\psi\|^2_{k,\delta,e} = \|\psi'\|^2_{k,\delta} + \|\hat{\psi}\|^2_{L^2_k(\Sigma)},$$

and similarly for $\hat{W}_2$.

We are interested in "gluing" $H^1_{\delta,e}$ solutions of $D_i\psi_i = 0$ over $\hat{W}_i$ that have the same limiting value $\hat{\psi}_0 := \hat{\psi}_1 = \hat{\psi}_2$ to form solutions of $D_T\psi_T = 0$ over W_T for sufficiently large T. By way of notation, we will write V for the subspace of $\ker(D_1) \oplus \ker(D_2)$ consisting of pairs of sections having the same limiting value. Explicitly, V is the kernel of the map $\ker(D_1) \oplus \ker(D_2) \to \ker(A)$ given by $(\psi_1, \psi_2) \mapsto \hat{\psi}_1 - \hat{\psi}_2$.

Lemma A.4 *1. If* $\operatorname{coker}(D_1) = \operatorname{coker}(D_2) = 0$, *then there exists a right inverse P_T for D_T for all sufficiently large T. Furthermore, the operator norm $\|P_T\|$ defined using the H^0_δ and $H^1_{\delta,e}$ norms is bounded independent of T.*

2. In the situation of part (1) above, there exists an isomorphism

$$\ker(D_T) \cong V \subset \ker(D_1) \oplus \ker(D_2).$$

Proof Under the assumption $\operatorname{coker}(D_i) = 0$, the paramatrices P_i constructed in Lemma A.3 may be taken to be right inverses. We apply another patching argument. Let ϕ_1 be a cutoff function on W_T, with $\phi_1 \equiv 1$ on W_1, $\phi_1 \equiv 0$ on W_2. We may assume that the derivative $\mathrm{d}\phi$ is bounded in norm by a constant times $\frac{1}{T}$. Likewise, let ϕ_2 be a cutoff function with $\phi_2 \equiv 1$ on W_2 and $\phi_2 \equiv 0$ on W_1; we assume that $\phi_1^2 + \phi_2^2 = 1$ on W_T.

Define a parametrix

$$Q_T = \phi_1 P_1 \phi_1 + \phi_2 P_2 \phi_2.$$

The same calculation as in Lemma A.3 shows that

$$D_T Q_T = \mathbb{I} + \sigma_{D_1}(\mathrm{d}\phi_1) P_1 \phi_1 + \sigma_{D_2}(\mathrm{d}\phi_2) P_2 \phi_2,$$

where there is no additional term involving a compact operator K since we have assumed that P_i is a right inverse of D_i. Then we calculate

$$\begin{aligned} \|\sigma(\mathrm{d}\phi_1) P_1 \phi_1 \psi\|_{H^0_\delta(W_T)} &= \|\sigma(\mathrm{d}\phi_1) P_1 \phi_1 \psi\|_{H^0_\delta(\hat{W}_1)} \\ &\leq \|\mathrm{d}\phi_1\|_{L^\infty} \|P_1 \phi_1 \psi\|_{H^0_\delta(\hat{W}_1)} \\ &\leq C\frac{1}{T}\|P_1\| \|\phi_1 \psi\|_{H^0_\delta(\hat{W}_1)} \\ &\leq C\frac{1}{T}\|P_1\| \|\psi\|_{H^0_\delta(W_T)}, \end{aligned}$$

and obtain a similar bound on the third term in the expression for $D_T Q_T$. This tells us that $\|\mathbb{I} - D_T Q_T\| \leq C\frac{1}{T}$ for a constant C independent of T, and we recall that by a well-known geometric series argument whenever an operator is within 1 of the identity in the operator norm it is invertible. Thus for sufficiently large T, D_TQ_T is invertible, and setting

$$P_T = Q_T(D_TQ_T)^{-1}$$

we obtain the desired right inverse of D_T. To bound the norm of P_T, we note that from the formula $(\mathbb{I} - F)^{-1} = 1 + F + F^2 + \cdots$, we obtain the bound $||(\mathbb{I} - F)^{-1}|| \leq \frac{1}{1-||F||}$. Taking $F = \mathbb{I} - D_TQ_T$, this gives

$$||(D_TQ_T)^{-1}|| \leq \frac{1}{1 - ||\mathbb{I} - D_TQ_T||} \leq \frac{1}{1 - \frac{C}{T}},$$

which is bounded as $T \to \infty$. It is easy to see that $||Q_T|| \leq ||P_1|| + ||P_2||$, and these bounds on Q_T and $(D_TQ_T)^{-1}$ lead to a bound for P_T that is independent of T. This proves part 1 of the lemma.

For part 2, we construct a map $f : \ker(D_T) \to V$ as follows. For a section $\psi_T \in \ker(D_T)$, we have an expansion $\psi|[-T,T] \times \Sigma = \sum \rho_\lambda(t)\phi_\lambda$ in terms of eigenfunctions ϕ_λ for A. In fact, since $\psi_T|[-T,T] \times \Sigma \in \ker(\frac{d}{dt} + A)$, we have that $\rho_\lambda(t) = \rho_\lambda(-T)e^{-\lambda(t+T)}$. Now, the restriction $\psi_T|W_1$ solves $D_T\psi_T = D_1\psi_T = 0$ over W_1 since D_T and D_1 agree on W_1. Furthermore, $\psi_T|_{W_1}$ has a unique extension to $\hat{W}_1$ as an element ψ_1 of $\ker(D_1)$, namely

$$\psi_1 = \begin{cases} \psi_T|_{W_1} & \text{on } W_1 \\ \sum \rho_\lambda(-T)e^{-\lambda t}\phi_\lambda & \text{on } [0,\infty) \times \Sigma \end{cases} \tag{19}$$

Note that $\psi_1 \in H^1_{\delta,e}(\hat{W}_1)$, and $\hat{\psi}_1 = \nu \sum_{\lambda=0} \rho_\lambda(-T)\phi_\lambda$. We make a similar restriction and extension of ψ_T to $\hat{W}_2$ to obtain $\psi_2 \in H^1_{\delta,e}(\hat{W}_2)$, and define

$$f(\psi_T) = (\psi_1, \psi_2).$$

It is a straightforward matter to check that $f(\psi_T)$ is indeed in V.

For convenience below, we extend f to all of $H^k_\delta(W_T)$ as follows: any section ψ has an expansion of the form used above over $[-T, T] \times \Sigma$. To define ψ_1, restrict ψ over W_1, then use formula (19) to extend this section to $\hat{W}_1$. A similar construction for W_2 gives a map

$$\begin{aligned} \tilde{f} : H^k_\delta(W_T) &\rightarrow H^k_{\delta,e}(\hat{W}_1) \oplus H^k_{\delta,e}(\hat{W}_2) \\ \psi &\mapsto \tilde{f}(\psi) = (\tilde{f}_1(\psi), \tilde{f}_2(\psi)) = (\psi_1, \psi_2) \end{aligned}$$

Now define $g : V \rightarrow \ker(D_T)$ as follows. For sections $(\psi_1, \psi_2) \in V$ we have decompositons $\psi_i = \psi_i' + \nu_i \pi^* \hat{\psi}_0$ coming from the definition (18), where $\hat{\psi}_0$ is the common limiting value of ψ_1 and ψ_2. Using the partition functions ϕ_i from above, we define

$$g(\psi_1, \psi_2) = \Pi(\phi_1^2 \psi_1' + \phi_2^2 \psi_2' + \nu_1 \nu_2 \pi^* \hat{\psi}_0),$$

where $\Pi : H^1_\delta(W_T) \rightarrow \ker(D_T)$ is orthogonal projection.

Claim: The linear maps f and g satisfy

$$\|(\mathbb{I} - g \circ f)\psi_T\|_{1,\delta} \leq \frac{C}{T} \|\psi_T\|_{1,\delta}$$

Hence, for sufficiently large T, $g \circ f$ is an isomorphism.

Proof: First we note that by expanding ψ_T in a series $\sum \rho_\lambda \phi_\lambda$, the ϕ_λ being eigenfunctions for A, and where the expansion is valid over $[-T-1, T+1] \times \Sigma$, we can write $\psi_T = \psi_T' + \nu_T \pi^* \hat{\psi}_0$ where ν_T is a cutoff function supported in $[-T-1, T+1] \times \Sigma$ and identically 1 on $[-T, T] \times \Sigma$ and where $\hat{\psi}_0$ is the part of ψ_T involving 0-eigenfunctions. Then we can write

$$\psi_T = \phi_1^2(\psi_T' + \nu\pi^*\hat{\psi}_0) + \phi_2^2(\psi_T' + \nu\pi^*\hat{\psi}_0).$$

Furthermore, by definition of f we have that $\phi_i^2 f(\psi_T) = \phi_i^2 \psi_T$ either as sections over W_T or over $\hat{W}_i$. Observe that from part 1 of the lemma, D_T has no cokernel and therefore the projection Π onto its kernel is given by $\mathbb{I} - P_T D_T$. Thus for $\psi_T \in \ker(D_T)$,

$$\begin{aligned}
\|(\mathbb{I} - g\, f)\psi_T\|_{1,\delta} &= \|\psi_T - (\mathbb{I} - P_T D_T)(\phi_1^2 f_1(\psi_T)' + \phi_2^2 f_2(\psi_T)' + \nu\pi^*\hat{\psi}_0)\|_{1,\delta} \\
&\leq \|\psi_T - (\phi_1^2 f_1(\psi_T)' + \phi_2^2 f_2(\psi_T)' + \nu\pi^*\hat{\psi}_0)\|_{1,\delta} \\
&\quad + \|P_T D_T(\phi_1^2 f_1(\psi_T)' + \phi_2^2 f_2(\psi_T)' + \nu\pi^*\hat{\psi}_0)\|_{1,\delta} \\
&= \|P_T(\sigma_{D_1}(\mathrm{d}\phi_1^2)\psi_T' + \sigma_{D_2}(\mathrm{d}\phi_2^2)\psi_T')\|_{1,\delta} \\
&\leq \frac{C}{T}\|P_T\|\|\psi_T\|_{1,\delta}.
\end{aligned}$$

In the last line above we have used that $\psi_T \in \ker(D_T)$ to obtain a bound in terms of the $1, \delta$ norm of ψ_T rather than the $0, \delta$ norm. Since $\|P_T\|$ is bounded independent of T by part 1, the claim is proved.

Claim: We have

$$\|(\mathbb{I} - f \circ g)(\psi_1, \psi_2)\|_{H^1_{\delta,e}(\hat{W}_1)\oplus H^1_{\delta,e}(\hat{W}_2)} \leq \frac{C}{T}\|(\psi_1, \psi_2)\|_{H^1_{\delta,e}(\hat{W}_1)\oplus H^1_{\delta,e}(\hat{W}_2)}$$

The proof is similar to the above calculation; it is easiest in this case to use the extension $\tilde{f}$.

We have shown that for sufficiently large T both $g \circ f$ and $f \circ g$ are isomorphisms, from which it follows that each of f and g is an isomorphism. This completes the proof of Lemma A.4.

A.2 Gluing Seiberg-Witten Solutions

We turn now to the non-linear situation of the Seiberg-Witten equations. As in the linear setup, we will work with $H^k_{\delta,e}$ configurations. More precisely, we consider a cylindrical-end 3-manifold $\hat{W}$ with end isometric to $[0, \infty) \times \Sigma$ for Σ a genus g Riemann surface, and a line bundle L over $\hat{W}$ that restricts over the end to the pullback of some bundle L_0 over Σ. As in Section 2, the spinor bundle S_W restricts to a bundle over Σ of the form $(K^{\frac{1}{2}} \oplus K^{-\frac{1}{2}}) \otimes E_0$, with $E_0^2 = L_0$. We assume once and for all that there are no reducible solutions to the vortex equations on Σ, which is equivalent to the assumption that $\deg E_0 \neq 1 - g$. In the situation that we consider, namely that Σ is a regular level surface of an S^1-valued Morse function, we can always arrange for this condition to hold by introducing cancelling critical points for the Morse function and thereby artificially inflating the genus of Σ.

As in the introduction to the Appendix, we will fix a background connection A_0 on L in order to identify the space $\mathcal{A}(L)$ of connections on L with $\Omega^1(\hat{W}; i\mathbb{R})$; we will assume that A_0 is a product connection on the end of $\hat{W}$. Note that A_0, together with the Levi-Civita connection, induces a connection on the spinor bundle $S|_\Sigma = (K^{\frac{1}{2}} \oplus K^{-\frac{1}{2}}) \otimes E$. We will not distinguish A_0 and this induced connection in the notation. As in the previous section, we define a cutoff function ν that is equal to 1 on $[0, \infty) \times \Sigma$ and vanishes away from $[-1, \infty) \times \Sigma$ (using the collar neighborhood $[-1, 0] \times \Sigma \subset W$). A solution $(\hat{\psi}, \hat{a}) = ((\alpha, \beta), \hat{a}) \in \Gamma((K^{\frac{1}{2}} \oplus K^{-\frac{1}{2}}) \otimes E) \oplus \Omega^1(\Sigma; i\mathbb{R})$ of the vortex equations

$$iF_{A_0} + id\hat{a} = \tfrac{1}{2}(|\beta|^2 - |\alpha|^2)\mathrm{d}vol_\Sigma$$

$$\bar{\partial}_{A_0+\hat{a}}\alpha = 0 \qquad \bar{\partial}^*_{A_0+\hat{a}}\beta = 0 \qquad \alpha\bar{\beta} = 0$$

can be pulled back to $\hat{W}$ using ν, to a configuration we denote $(\nu\pi^*\hat{\psi}, \nu\pi^*\hat{a})$. An extended H^k_δ configuration (Ψ, a) on $\hat{W}$ is then one that can be written in the form

$$(\Psi, a) = (\Psi' + \nu\pi^*\hat{\psi}, a' + \nu\pi^*\hat{a})$$

for $(\Psi', a') \in H^k_\delta(\hat{W})$. We consider the function $\mathfrak{S}$ from the beginning of the Appendix to map $H^k_{\delta,e} \to H^{k-1}_\delta$ for appropriate choice of k.

Let $\hat{W}_1 = W_1 \cup [0, \infty) \times \Sigma$ and $\hat{W}_2 = W_2 \cup (-\infty, 0] \times \Sigma$ be a pair of cylindrical-end manifolds with spin^c structures $\mathfrak{t}_i$, spinor bundles S_i, and determinant lines L_i. We assume that over the ends of $\hat{W}_i$ the L_i are equal to the pullback of a bundle L_0 on Σ, i.e., that the spin^c structures "agree on the ends." Let $\mathcal{C}^*_i$ denote the space of irreducible $H^k_{\delta,e}$ configurations on $\hat{W}_i$, and let us write $\mathcal{B}_i \subset \mathcal{C}^*_i$ for the space of solutions of the Seiberg-Witten equations (14) on $\hat{W}_i$ in the spin^c structures on $\hat{W}_i$. Let $\mathcal{M}_i = \mathcal{M}_i(\mathfrak{t}_i)$ be the moduli spaces of gauge equivalence classes of solutions, i.e., $\mathcal{M}_i = \mathcal{B}_i/\mathcal{G}_i$ where $\mathcal{G}_i = \mathrm{Map}(\hat{W}_i; S^1)$ is the space of H^{k+1}_δ gauge transformations. Our object is, given a spin^c structure $\mathfrak{t}$ on $W_T = W_1 \cup [-T, T] \times \Sigma \cup W_2$ that has $\mathfrak{t}|_{W_i} = \mathfrak{t}_i|_{W_i}$, to construct a "gluing map" for large T:

$$\gamma : \mathcal{M}_1(\mathfrak{t}_1) \times_\partial \mathcal{M}_2(\mathfrak{t}_2) \to \mathcal{M}_T(\mathfrak{t})$$

where $\mathcal{M}_T = \mathcal{B}_T/\mathcal{G}_T$ is the moduli space of Seiberg-Witten solutions on W_T. The notation $\times_\partial$ represents a fiber product: $\mathcal{M}_1(\mathfrak{t}_1) \times_\partial \mathcal{M}_2(\mathfrak{t}_2)$ stands for those pairs of gauge equivalence classes $[\Psi_1, a_1]$, $[\Psi_2, a_2]$ of solutions on $\hat{W}_1$ and $\hat{W}_2$ whose limiting values $(\hat{\psi}_i, \hat{a}_i)$ agree. As a first step, we have the following.

Proposition A.5 *There exists an "approximate gluing map"*

$$\tilde{\gamma} : \mathcal{B}_1 \times_{\partial} \mathcal{B}_2 \to \mathcal{C}_T^*$$

that has

$$\mathfrak{S}(\tilde{\gamma}((\Psi_1, a_1), (\Psi_2, a_2))) = \epsilon_T = (\epsilon_T^1, \epsilon_T^2) \in H_\delta^{k-1}(S \oplus \Lambda^1(T^*W_T; i\mathbb{R}))$$

where ϵ_T is a configuration approaching 0 as $T \to \infty$.

Proof: We introduce cutoff functions ϕ_1, ϕ_2 on $\hat{W}_1$ and $\hat{W}_2$ as follows. On $\hat{W}_1$, take ϕ_1 to be supported on $W_1 \cup ([0, T] \times \Sigma)$, and to be identically 1 on $W_1 \cup ([0, T-1] \times \Sigma)$. Similarly, ϕ_2 is supported on $W_2 \cup ([-T, 0] \times \Sigma)$ and equal to 1 on $W_2 \cup ([-T+1, 0] \times \Sigma)$. Let (Ψ_i, a_i), $i = 1, 2$, be extended H_δ^k solutions over $\hat{W}_i$ both having limiting values $(\hat{\psi}_0, \hat{a}_0)$. That is, on $\hat{W}_i$ we can write

$$\Psi_i = \Psi_i' + \nu\pi^*\hat{\psi}_0 \qquad a_i = a_i' + \nu\pi^*\hat{a}_0$$

for $(\Psi_i', a_i') \in H_\delta^k$. Define multiplication operators $\tilde{\phi}_i$, $i = 1, 2$, by setting

$$\tilde{\phi}_i(\Psi_i, a_i) = (\tilde{\phi}_i\Psi_i, \tilde{\phi}_i a_i) = (\phi_i\Psi_i' + \nu\pi^*\hat{\psi}_0, \phi_i a_i' + \nu\pi^*\hat{a}_0).$$

That is, $\tilde{\phi}_i$ is a cutoff to the limiting value.

Clearly we may think of $\tilde{\phi}_i(\Psi_i, a_i)$ as configurations on W_T (after shifting the origin in the identification of the ends of $\hat{W}_i$ with $\mathbb{R}^\pm \times \Sigma$), so for $(\Psi_i, a_i) \in$

$\mathcal{B}_i$ we define

$$\tilde{\gamma}((\Psi_1, a_1), (\Psi_2, a_2)) = \begin{cases} \tilde{\phi}_1(\Psi_1, a_1) & \text{on } W_1 \cup ([-T, 0] \times \Sigma) \\ \tilde{\phi}_2(\Psi_2, a_2) & \text{on } W_2 \cup ([0, T] \times \Sigma) \end{cases}$$

By construction, this configuration is in $H^k_\delta(W_T)$ and solves the Seiberg-Witten equations on $W_1 \cup ([-T, -1] \times \Sigma) \cup ([1, T] \times \Sigma) \cup W_2$. For convenience, we will write u_i for the solutions (Ψ_i, a_i). To estimate the error term $\epsilon_T = \mathfrak{S}(\tilde{\gamma}(u_1, u_2))$, then, we need only consider the region $[-1, 1] \times \Sigma \subset W_T$. On $[-1, 0] \times \Sigma$ we have

$$\begin{aligned} \|\epsilon_T\|_{k-1,\delta} &= \|\mathfrak{S}_{\hat{W}_1}(\Psi' + \nu\pi^*\hat{\psi}_0,\ a' + \nu\pi^*\hat{a}_0)\|_{k-1,\delta} \\ &\leq \|(D_{A_0}\nu\pi^*\hat{\psi}_0 + \nu\pi^*\hat{a}_0.\nu\pi^*\hat{\psi}_0 + \mathrm{d}\phi_1.\Psi' + \phi_1 D_{A_0}\Psi' \\ &\qquad + \phi_1^2 a'.\Psi' + \phi_1(a'.\nu\pi^*\hat{\psi}_0 + \nu\pi^*\hat{a}_0.\Psi')\| \\ &\quad + \|\star(F_{A_0} + \mathrm{d}\nu\pi^*\hat{a}_0 + i\eta) - \tau(\nu\pi^*\hat{\psi}_0,\ \nu\pi^*\hat{\psi}_0) \\ &\quad\ + \star(\mathrm{d}\phi_1 \wedge a' + \phi_1 \mathrm{d}a') - \phi_1^2\tau(\Psi', \Psi') - 2\phi_1 \mathrm{Re}\tau(\Psi', \nu\pi^*\hat{\psi}_0)\| \end{aligned}$$

Now, since $\nu \equiv 1$ on $[T-1, T] \times \Sigma \subset \hat{W}_1$, we have $(\nu\pi^*\hat{\psi}_0, \nu\pi^*\hat{a}_0) = (\pi^*\hat{\psi}_0, \pi^*\hat{a}_0)$ is a constant Seiberg-Witten solution in this region. Therefore the terms involving only $\hat{\psi}_0$ and $\hat{a}_0$ drop out of the above expression. Furthermore, we know that (Ψ', a') is an H^k_δ configuration and therefore has exponential decay on the cylinder. Since the remaining terms in the expression above can all be estimated in terms of constants times the norms of Ψ' or a', we infer that $||\epsilon_T||_{k-1,\delta} \to 0$ as $T \to \infty$. This proves the proposition.$\Box$

We now must consider the difference between the image of the approximate gluing map $\tilde{\gamma}$ and the actual Seiberg-Witten solution space on W_T. First, let us suppose that $(\Theta, b) \in \mathcal{C}_T$ is a configuration on W_T. Using the notation $L_{(\Theta,b)}$ for the linearization of $\mathfrak{S}$ at (Θ, b) a calculation shows that for any variation $(\theta, \tilde{b})$ we have

$$\mathfrak{S}(\Theta + \theta, b + \tilde{b}) - \mathfrak{S}(\Theta, b) = L_{(\Theta,b)}(\theta, \tilde{b}) + (\tilde{b}.\theta, -\tau(\theta, \theta)).$$

That is, $\mathfrak{S}$ differs from its linearization by a term of the form $q((\theta, \tilde{b}), (\theta, \tilde{b}))$, where

$$q((\theta_1, b_1), (\theta_2, b_2)) = (b_1.\theta_2, -\tau(\theta_1, \theta_2)).$$

It is a straightforward matter to verify that q satisfies an inequality of the form

$$||q(u, u) - q(u, v)|| \leq C||u - v||(||u|| + ||v||). \tag{20}$$

Our object is to use a quadratic contraction mapping principle to produce an exact solution on W_T from the approximate solution $\tilde{\gamma}(u_1, u_2)$. First, note that for generic metric and perturbation η, the conditions of Lemma A.4 hold for the linearizations of the Seiberg-Witten maps $\mathfrak{S}_1$, $\mathfrak{S}_2$ on $\hat{W}_1$

and $\hat{W}_2$. Therefore, there is a right inverse P for the linearization $L_{\tilde{\gamma}(u_1,u_2)}$, whose norm is bounded independent of T (see the remark after Theorem A.7, however). Writing $\tilde{u} = \tilde{\gamma}(u_1, u_2)$ for the approximate solution, we will look for an exact solution of the form

$$u = \tilde{u} + Pv$$

for some configuration v on W_T. This means that v must satisfy

$$0 = \mathfrak{S}(u) = \mathfrak{S}(\tilde{u} + Pv) = \mathfrak{S}(\tilde{u}) + L_{\tilde{\gamma}(u_1,u_2)}(Pv) + q(Pv, Pv)$$

or

$$0 = \epsilon_T + v + q(Pv, Pv).$$

Lemma A.6 *If Q is a quadratic operator on a Banach space B, $Q : B \oplus B \to B$, satisfying an inequality of the form (20), then for all sufficiently small z there exists a solution x to the equation*

$$x = Q(x, x) + z.$$

Note that since the norm of P is bounded independent of T, the map $Q : v \mapsto q(Pv, Pv)$ satisfies the condition of the lemma.

Thus we obtain:

Theorem A.7 *For all sufficiently large T, there exists a map*

$$\gamma_T : \mathcal{M}_1 \times_\partial \mathcal{M}_2 \to \mathcal{M}_T$$

defined by

$$\gamma(u_1, u_2) = \tilde{\gamma}(u_1, u_2) + Pv$$

where P is the right inverse of the linearization $L_{\tilde{\gamma}(u_1,u_2)}$ and v is the solution of

$$0 = \mathfrak{S}(\tilde{\gamma}(u_1, u_2)) + v + q(Pv, Pv)$$

provided by Lemma A.6. Furthermore, γ_T is a diffeomorphism onto its image, which is an open subset of $\mathcal{M}_T$.

Remark A.8 *Here we are suppressing the fact that the inverse P of the linearization exists only "modulo gauge." To be more honest, we should introduce another component into the map $\mathfrak{S}$ that fixes the gauge, in which case the linearization will indeed be of APS type and the work of the previous section applies. This point is not particularly difficult to deal with, but involves more notation than we care to use here. Note, however, that the fact that the gauge action is free—i.e., we are dealing only with irreducible configurations—enters in an essential manner.*

We will end our discussion of gluing theory here, noting only that the smoothness of γ follows essentially from Lemma A.4: that $\tilde{\gamma}$ is smooth is a routine check, and the fact that the passage from $\tilde{\gamma}(u_1, u_2)$ to $\gamma(u_1, u_2)$ is smooth can be seen roughly as follows. The derivative of $\tilde{\gamma}$ provides an isomorphism with the tangent space of the image of $\tilde{\gamma}$ with $\ker L_{u_1} \times_{\theta} \ker L_{u_2}$ (the space V of Lemma A.4), while the tangent space of $\gamma(u_1, u_2)$ is identified with $\ker L_{\gamma(u_1,u_2)}$. The map between the two can be seen to be essentially the linear gluing construction of Lemma A.4.

REFERENCES

References

[1] M. F. Atiyah, V. K. Patodi, and I. M. Singer. Spectral asymmetry and Riemannian geometry, i. *Math. Proc. Camb. Phil. Soc.*, 77:43–69, 1975.

[2] S. B. Bradlow. Vortices in holomorphic line bundles over closed Kähler manifolds. *Comm. Math. Phys.*, 135(1):1–17, 1990.

[3] S. K. Donaldson. Topological field theories and formulae of Casson and Meng-Taubes. In *Proceedings of the Kirbyfest*, volume 2 of *Geometry and Topology Monographs*, 1999.

[4] M. Hutchings. Reidemeister torsion in generalized Morse theory. Preprint DG/9907066, 1999.

[5] M. Hutchings and Y.-J. Lee. Circle-valued Morse theory and Reidemeister torsion. Preprint, 1997.

[6] M. Hutchings and Y.-J. Lee. Circle-valued Morse theory, Reidemeister torsion, and Seiberg-Witten invariants of 3-manifolds. *Topology*, 38:861–888, 1999.

[7] E. Ionel and T. H. Parker. Gromov invariants and symplectic maps. *Math. Ann.*, 314(1):127–158, 1999.

[8] A. Jaffe and C. H. Taubes. *Vortices and Monopoles: Structure of Static Gauge Theories.* Birkhäuser, Boston, Mass., 1980.

[9] D. Johnson. A geometric form of Casson's invariant and its connection to Reidemeister torsion. Unpublished notes.

[10] I. G. Macdonald. Symmetric products of an algebraic curve. *Topology*, 1:319–343, 1962.

[11] G. Meng and C. H. Taubes. $\underline{\mathrm{SW}}$ = Milnor torsion. *Math. Res. Lett.*, 3:661–674, 1996.

[12] J. Milnor. A duality theorem for Reidemeister torsion. *Ann. Math.*, 76:137–147, 1962.

[13] T. Mrowka, P. Ozsváth, and B. Yu. Seiberg-Witten monopoles on Seifert fibered spaces. *Comm. Anal. Geom.*, 5(4):685–791, 1997.

[14] D. Salamon. Seiberg-Witten invariants of mapping tori, symplectic fixed points, and Lefschetz numbers. In *Proceedings of 6th Gökova Geometry-Topology Conference*, pages 117–143, 1998.

[15] V. G. Turaev. Reidemeister torsion in knot theory. *Russian Math. Surveys*, 41(1):119–182, 1986.