Harvard University Department of Mathematics

The Seiberg-Witten Invariant of a Homology $S^1 \times S^3$

Author: Rohil Ray Prasad Advisor: Professor Clifford Taubes



This thesis is submitted in partial fulfillment of the requirements for the degree of Bachelor of Arts with Honors in Mathematics.

March 19, 2018

Contents

Introduction				2	
1	Preliminaries			5	
	1.1	The to	pology of manifolds satisfying (A1) and (A2)	5	
	1.2	The Se	eiberg-Witten invariants	8	
	1.3	The Se	eiberg-Witten invariants of manifolds satisfying (A1)	12	
2	The	The Lambda Invariant 1			
	2.1	The er	nd-periodic Dirac operator and its Fredholm theory	17	
		2.1.1	The Fourier-Laplace transform and Taubes' theorem	19	
		2.1.2	The correction term	23	
		2.1.3	The Fourier-Laplace transform along a path	23	
	2.2	The fi	rst change of index formula: excision and residue calculus	26	
		2.2.1	The excision principle	26	
		2.2.2	Reduction to the infinite cyclic cover	28	
		2.2.3	Residue calculus	29	
	2.3	The se	cond change of index formula: spectral flow	32	
	2.4	Orient	tations	40	
3	The Splitting Formula 44				
	3.1	Seiberg-Witten-Floer homology		45	
		3.1.1	Morse homology	45	
		3.1.2	Morse homology with boundary	50	
		3.1.3	The Chern-Simons-Dirac functional and its gradient flow	54	
		3.1.4	Floer homology, the exact triangle, and cobordism maps	57	
		3.1.5	The h-invariant	59	
	3.2	The splitting formula and its proof		60	
	3.3	3.3 The index calculation		68	
		3.3.1	The Dirac operator and APS boundary conditions	68	
		3.3.2	Proof of the main theorem	71	
	3.4	The m	onopole count	78	
		3.4.1	A compactness theorem	79	
		3.4.2	An application of Kronheimer-Mrowka's local gluing theorem	85	
	3.5	Comp	letion of the proof and an obstruction to positive scalar curvature	90	

Introduction

A major mathematical pursuit is that of understanding objects that have a simple *local* structure while having a non-trivial *global* structure. For example, consider the planet Earth. A human observer, perhaps standing at a street corner somewhere in Montana and looking out at the horizon, will see a flat expanse of land. It is not unreasonable given only this observation to conclude that the Earth is a flat plane. In the language of topology, the claim is that Earth is homeomorphic to the Euclidean plane \mathbb{R}^2 .

However, it is known at this point that the Earth is not a flat plane, but rather a round sphere. The error in the observer's judgement is that their field of view is limited. They see only a small patch of the land on the Earth, which indeed is homeomorphic to an open set of \mathbb{R}^2 . Therefore, referring back to the initial statement of the introduction, one can conclude that Earth is *locally* homeomorphic to \mathbb{R}^2 , but not *globally* homeomorphic to \mathbb{R}^2 .

Generalizing to higher dimensions, a topological space that is *locally* homeomorphic to the *n*-dimensional Euclidean space \mathbb{R}^n and satisfies some other technical assumptions is called a *n*-dimensional manifold or a *n*-manifold. By the above discussion, it follows that the surface of the Earth is a 2-dimensional manifold.

An important class of manifolds are those known as *smooth manifolds*. Roughly, these are manifolds that can be "smoothed out" to not have any sharp jagged points or seams or anything of the sort. The exact manner in which the manifold is "smoothed out" is called its *smooth structure*. Smooth manifolds are natural objects to study. For example, the round 2-sphere (the surface of the Earth) is a smooth 2-dimensional manifold.

Manifolds that do not have any distinguished smooth structure are generally denoted as *topological manifolds*. A major goal in topology is to classify all of the different types of *n*-dimensional topological or smooth manifolds that can occur. In practice, the manifolds are restricted to be *compact*, *oriented*, and *without boundary* to make this somewhat tractable.

This thesis discusses in particular some of the mathematics surrounding the study of smooth 4-dimensional manifolds. Smooth manifolds of dimension 4 hold an intriguing place in the classification problem described above.

It is often stated that dimension 4 is the middle ground between "rigidity" and "fluidity". For example, thanks to the work of Perelman ([MT07]), it is now known that 3-manifolds are "rigid". They can be decomposed into some finite number of geometric pieces, each of which are one of eight distinct types. Furthermore, there is no distinction between topological and smooth 3-manifolds. Any topological 3-manifold can be "smoothed out" in a unique fashion ([Hat]).

On the other hand, manifolds of dimension ≥ 5 exhibit rather "fluid" behavior. An example of the usefulness of extra dimensions is as follows. Two lines in \mathbb{R}^2 will always intersect in a point except for the very special case in which they are parallel. However, in \mathbb{R}^3 , two lines will "almost never" intersect at all! Namely, given any two intersecting lines in \mathbb{R}^3 , a small translation of one of the lines in any direction

will result in two non-intersecting lines. This "fluidity" was used by Stephen Smale to prove the Poincaré conjecture for both topological and smooth manifolds of dimension 5 or greater ([Sma61]).

In four dimensions, topological and smooth manifolds are very distinct objects and behave very differently with respect to the above framework. *Topological* 4-manifolds exhibit "fluid" behavior, and it was demonstrated by Freedman ([Fre82]) that the Poincaré conjecture holds for topological 4-manifolds as well.

The following year, Donaldson showed on the other hand that *smooth* 4-manifolds exhibit more "rigid" behavior ([Don83]). The intersection product on the middle-dimensional homology $H_2(X;\mathbb{Z})$ of a topological 4-manifold X induces a bilinear form on $H_2(X;\mathbb{Z})$ known as the *intersection form* of X. Donaldson showed that, if X is smooth and has a definite intersection form, then the intersection form must furthermore be diagonalizable over the integers. This is a rather severe restriction on the topology of smooth 4-manifolds with definite intersection form, and led to a host of examples of topological 4-manifolds that could not be "smoothed out", as their intersection form did not satisfy the required property.

The techniques used by Donaldson arise from a field of mathematics known as *gauge theory*. Roughly, gauge theory seeks to understand smooth 3- and 4-manifolds by examining geometric objects called *connections* on these manifolds. These connections are generally fixed to satisfy some first-order, elliptic partial differential equations. Furthermore, two connections are identified if they are what is known as *gauge-equivalent*, a notion that will be explicitly defined later on. The space of gauge-equivalence classes of connections satisfying these equations then often has the structure of a finite-dimensional manifold, which can be used to understand the original manifold of interest.

The equation that was used by Donaldson is called the *anti-self-dual equation*. This thesis does not discuss these equations, but rather discusses a topic in the theory of the *Seiberg-Witten equations*. These equations were introduced in 1994 by Seiberg and Witten ([SW94]). In addition to a connection, these equations involve another geometric object known as a *spinor*. The Seiberg-Witten equations are from an analytical standpoint much easier to work with than the anti-self-dual equation, and as such their introduction precipitated a flurry of works that simplified earlier results acquired using Donaldson theory, as well as proved new results.

Although it will be expanded upon later on, it is sufficient for now to know that, given a compact, oriented smooth 4-manifold X with an additional piece of data known as a *spin^c* structure, analysis of the Seiberg-Witten equations produces an integer known as the *Seiberg-Witten invariant* of X. The construction of the Seiberg-Witten invariant involves a couple of auxiliary pieces of geometric data, namely a *Riemannian metric* on X and a small *perturbation* of the equations. It is well-known, and will be briefly discussed in Chapter 1, that if X satisfies a certain topological condition then the Seiberg-Witten invariant is independent of these geometric choices, i.e. it depends only on the topology and smooth structure of X.

This thesis discusses work done on the case where X does not satisfy this topological condition. In this case, the Seiberg-Witten invariant most certainly depends on the choice of metric and perturbation on X. It is desirable to modify the construction of the Seiberg-Witten invariant somehow so that it is once again independent of the choice of metric and perturbation.

The thesis begins with some preliminary content in Chapter 1. It is known that the Seiberg-Witten invariant can be "fixed" in the sense of the above discussion when *X* satisfies two topological assumptions, known as "assumptions (A1) and (A2)". This chapter defines and discusses these two assumptions in detail, and gives a short exposition of the construction of the Seiberg-Witten invariants.

Chapter 2 presents in detail the construction of the modified Seiberg-Witten invariant (denoted by $\lambda_{SW}(X)$) for a 4-manifold X satisfying assumptions (A1) and (A2). This was done by Mrowka, Ruber-

man, and Saveliev ([MRS11]), and the exposition follows that of their paper.

A significant body of work by Mrowka, Ruberman, Saveliev, and Lin towards calculating the invariant $\lambda_{SW}(X)$ has been released since the publishing of the initial construction in [MRS11]. Several of these developments are surveyed in [RS13].

Chapter 3 presents a very recent development in this theory. In a December 2017 preprint ([LRS17]), Lin, Ruberman, and Saveliev proved a formula relating $\lambda_{SW}(X)$ to invariants arising from the theory of *Seiberg-Witten-Floer homology*, a homology theory for 3-manifolds first constructed in full generality by Kronheimer and Mrowka in 2007 ([KM07]). The chapter starts with a brief exposition of Seiberg-Witten-Floer homology and continues to give an exposition of the main theorem of [LRS17]. The chapter, and the thesis, conclude with a application of this result to the study of positive scalar curvature metrics on 4-manifolds.

Recommended background: The reader may have noticed the steady encroachment of technical content into this introduction. Indeed, there are a good number of mathematical concepts and topics that the reader is recommended to be familiar with in order to get the most out of this exposition. First of all, the language of algebraic topology is used freely without exposition. The book [Hat02] provides an encyclopedic resource for the particulars of homology and cohomology, although many more concise lecture notes can be found scattered around the internet. The book [MS74] is an excellent resource for information on characteristic classes of vector bundles, which are also mentioned at several points. Second, the reader should be familiar with the basics of smooth manifolds, differential forms, vector bundles, connections and Riemannian geometry, as in [Tau11]. Third, it is useful to understand the basic constructions of spin geometry and the Dirac operator, which are discussed in [Mor95]. Fourth, the reader should be at least somewhat familiar with Sobolev spaces and (elliptic) differential operators. The material in Chapters 5 and 6 of [Eva10] should be more than sufficient when coupled with the understanding that the notions of Sobolev spaces and elliptic differential operators extend to the more general setting of sections of vector bundles.

Chapter 1

Preliminaries

In this chapter, we introduce our main objects of study, smooth 4-manifolds with the integral homology of $S^1 \times S^3$ (known as "assumption (A1)") that have an embedded integral homology three-sphere generating the third integral homology group (known as "assumption (A2)").

Let X be a smooth 4-manifold satisfying assumptions (A1) and (A2). The first section gives the construction of a few different non-compact manifolds associated to X. The second section gives a brief overview of the Seiberg-Witten invariants of 4-manifolds. The third section then discusses specifically the Seiberg-Witten invariants of X, and how the topological assumptions on X simplify the structure of the space of solutions to the Seiberg-Witten equations.

1.1 The topology of manifolds satisfying (A1) and (A2)

In this thesis, we will explore the Seiberg-Witten theory of closed Riemannian four-dimensional manifolds *X* satisfying two conditions.

(A1) The manifold X must have the integral homology of $S^1 \times S^3$. Explicitly, the integral homology of X is

$$H_i(X;\mathbb{Z}) = \begin{cases} \mathbb{Z} & i \neq 2\\ 0 & i = 2 \end{cases}$$

(A2) For a manifold X satisfying (A1), the third homology group $H_3(X;\mathbb{Z}) \simeq \mathbb{Z}$ is generated by the fundamental class of a smoothly embedded integral homology 3-sphere. We will denote this homology 3-sphere by Y throughout the article.

We present a couple of examples below of manifolds satisfying (A1) and (A2)..

- 1. The manifold $X = S^1 \times S^3$ clearly satisfies both conditions. More generally, we can take $X = S^1 \times Y$ for any integral homology 3-sphere *Y*.
- Mapping tori of homology 3-spheres form another prominent class of examples. Given a homology 3-sphere *Y* and a self-diffeomorphism τ : *Y* → *Y*, we may form a closed 4-manifold by taking cylinder [0,1] × *Y* and glueing one end to the other by the identification (0, *y*) ~ (1, τ(*y*)). This manifold is known as the *mapping torus* of the map τ, and a standard calculation using the Mayer-Vietoris sequence verifies that it satisfies both (A1) and (A2).

If a manifold X satisfies (A1), then from standard covering space theory it has a connected infinite cyclic cover \tilde{X} . This infinite cover admits an explicit topological construction.

By the tubular neighborhood theorem, the embedded submanifold *Y* has a neighborhood $N \subset X$ that is diffeomorphic to $(-1,1) \times Y$ with *Y* identified with $\{0\} \times Y$. It follows from this local description of the manifold near *Y* that the space $W = X \setminus Y$ has the structure of a manifold with boundary $Y \cup Y$. Furthermore, it is clear from a calculation with the Mayer-Vietoris sequence that *W* is connected. In other words, *W* is a *cobordism* from *Y* to itself. Furthermore, the orientation of *X* induces an orientation on *W* that makes it an *oriented cobordism* from *Y* to itself.

The construction of such a cobordism *W* is indeed just a rigorous formulation of the action of *cutting open X along Y*, and is referred to as such in the literature. Since *W* has boundary $Y \cup Y$, we can glue *W* to itself along one of these boundary components to produce a new manifold with boundary $Y \cup Y$, denoted by $W \cup_Y W$.

We can also take \mathbb{Z} -many copies of W labeled as W_i for all $i \in \mathbb{Z}$. Then, the infinite cyclic cover \widetilde{X} is the infinite gluing

$$\cdots W_{-1} \cup_Y W_0 \cup_Y W_1 \cup_Y W_2 \cdots .$$

There is a natural transformation $T : \tilde{X} \to \tilde{X}$ that takes W_i to W_{i+1} . We recover our manifold X from \tilde{X} by quotienting out by the action of T, and furthermore the quotient projection is our desired covering map. The transformation T then generates the deck group of this covering.



Figure 1.1: The infinite cyclic cover \widetilde{X} .

Our assumption of (A2) simplifies the topology of W, and therefore that of \tilde{X} . This is described in the following theorem.

Theorem 1.1.1. Let X be a manifold satisfying both (A1) and (A2). Then the cobordism W constructed as above has vanishing first, second and fourth integral homology and $H_3(W; \mathbb{Z}) = \mathbb{Z}$.

Proof. It is immediate that $H_4(W; \mathbb{Z}) = 0$ by Poincare duality for manifolds with boundary.

Let $N \simeq (-1,1) \times Y$ be an open topological collar neighborhood of *Y*. Then $W \cup N = X$, and the Mayer-Vietoris sequence yields a long exact sequence in integral homology:

$$\cdots \to H_k(Y \cup Y) \to H_k(W) \oplus H_k(Y) \to H_k(X) \to H_{k-1}(Y \cup Y) \to \dots$$

Setting k = 1 and using reduced homology, it is found that there exists a long exact sequence of the form:

$$0 \to H_1(W) \to \mathbb{Z} \to \mathbb{Z} \oplus \mathbb{Z} \to \mathbb{Z} \to 0.$$

Any homomorphism from \mathbb{Z} into $\mathbb{Z} \oplus \mathbb{Z}$ must be either injective or equal to the zero map. Since the cokernel of this map is isomorphic to \mathbb{Z} , it must be injective and $H_1(W; \mathbb{Z}) = 0$.

Now set k = 2. Both $H_2(Y)$ and $H_2(X)$ are equal to 0, so it is immediate from the above sequence that $H_2(W) = 0$ as well.

Pick a copy of Y in W. The inclusion map $H_3(Y;\mathbb{Z}) \to H_3(X;\mathbb{Z})$ then factors through the inclusion maps $H_3(Y;\mathbb{Z}) \to H_3(W;\mathbb{Z})$ and $H_3(W;\mathbb{Z}) \to H_3(X;\mathbb{Z})$.

The inclusion map $H_3(Y;\mathbb{Z}) \to H_3(W;\mathbb{Z})$ is injective, so it follows that the inclusion map $H_3(W;\mathbb{Z}) \to H_3(X;\mathbb{Z})$ is injective as well. It follows that $H_3(W;\mathbb{Z}) \simeq \mathbb{Z}$.

Remark 1.1.2. We have shown that *W* is a *homology cobordism* from *Y* to itself. As a result, the study of manifolds satisfying (A1) and (A2) has applications to understanding the *homology cobordism group* $\Theta_{\mathbb{Z}}^3$ of integral homology 3-spheres. This is the abelian group generated by the integral homology 3-spheres, with two such spheres said to be equivalent if they are homology cobordant. Although the proof is outside the scope of this thesis, the main result of [LRS17] described in Chapter 3 is used to construct a sufficient topological condition for an element of $\Theta_{\mathbb{Z}}^3$ to have infinite order.

We can define a couple of other non-compact manifolds related to *X*. Just as the infinite cyclic cover \widetilde{X} is the \mathbb{Z} -fold gluing

$$\cdots W_{-1} \cup_Y W_0 \cup_Y W_1 \cup_Y W_2 \cdots,$$

we may define a non-compact manifold with boundary Y as the gluing

$$W_0 \cup_Y W_1 \cup_Y W_2 \cdots$$

We will denote this manifold by X_+ .

Next, pick any spin 4-manifold with boundary Y, which we will always denote by Z. Then, we may glue Z to X_+ to form a non-compact manifold that we will always denote by Z_+ .



Figure 1.2: The manifolds Z, X_+ , and Z_+ .

The triviality of the oriented spin cobordism group in three dimensions ([Sti03]) guarantees the existence of Z. We will also always assume that Z is simply-connected, which can always be achieved by performing surgery along loops in Z.

Throughout this section, we have discussed the topology of manifolds satisfying (A1) and (A2). Another interesting question is to determine whether there are manifolds satisfying (A1) that do not satisfy (A2).

In fact, as was communicated to the author by Daniel Ruberman, a variety of examples can be constructed using cyclic branched covers of knots with the appropriate Alexander polynomial. This is done by taking the mapping torus of the generator of the deck group of this branched cover. One concrete example of a branched cover that can be used in this construction is the six-fold branched cover of the trefoil knot, and more generally the pq-fold branched cover of the (p, q)-torus knot.

1.2 The Seiberg-Witten invariants

By now there are countless expositions on the Seiberg-Witten invariants. Therefore, this section has been written more towards the aim of fixing notation and discussing some important properties of the invariants.

For a compact but complete introduction to the Seiberg-Witten invariants, the author recommends [Mor95]. The book [Sal99] is far more wide-ranging in its coverage and serves as a useful reference.

Fix *X* to be a closed, oriented Riemannian 4-manifold.

Definition 1.2.1. The data of a *spin*^c *structure* on X consists of a pair of unitary rank two complex vector bundles S^+, S^- on X along with a *Clifford multiplication* action $\rho : TX \to \text{Hom}(S, S)$ where S is the direct sum $S^+ \oplus S^-$.

The Clifford multiplication ρ identifies TX isometrically with the subbundle $\mathfrak{su}(S) \subset \operatorname{Hom}(S, S)$ with the inner product $(A, B) = \operatorname{Tr}(\frac{1}{2}A^*B)$. Furthermore, Clifford multiplication is an *odd* transformation in that it sends sections of the subbundle S^+ to sections of S^- and vice versa.

We extend the Clifford multiplication to one-forms by identification with the metric. Following that, we can extend it to differential forms of any degree by the rule

$$\rho(\alpha \wedge \beta) = \rho(\alpha)\rho(\beta) - (-1)^{\deg(\alpha)\deg(\beta)}\rho(\beta)\rho(\alpha).$$

If ω is a self-dual two-form, then it follows that Clifford multiplication by ω is an endomorphism of S_+ that restricts to zero on S_- .

The bundle *S* is known as the *spinor bundle* and its sections are *spinors*. Sections of S_+ and S_- are called positive and negative spinors respectively.

Let *A* be a unitary connection on the determinant line bundle of S_+ . The connection *A* has a curvature two-form denoted by F_A . As this is a two-form on a four-dimensional manifold, it can be decomposed into self-dual and anti-self-dual parts $F_A = F_A^+ + F_A^-$ that are respectively ± 1 -eigenvectors of the Hodge star operator.

It is also a quick exercise in spin geometry to show that A induces a covariant derivative ∇_A on the spinor bundle S.

Definition 1.2.2. The *Dirac operator* associated to a spin^{*c*} connection *A* is the composition of operators

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

The Dirac operator is first-order, self-adjoint and elliptic. With respect to the bundle decomposition, it splits into two operators

$$D_A^+: C^\infty(S_+) \to C^\infty(S_-)$$

and

$$D_A^-: C^\infty(S_-) \to C^\infty(S_+)$$

that are adjoints of each other.

Fix a spin^{*c*} structure \mathfrak{s} on *X*. Let the *configuration space* $\mathcal{C}(X)$ be the space of pairs (A, φ) where *A* is a unitary connection on the determinant line bundle and $\varphi \in C^{\infty}(S_+)$ is a positive spinor.

A configuration (A, φ) is a solution to the *Seiberg-Witten equations* if

$$\frac{1}{2}F_A^+ = \rho^{-1}((\varphi \otimes \varphi^*)_0),$$
$$D_A^+ \varphi = 0.$$

The final piece of notation that we are missing is the definition of the expression $(\varphi \otimes \varphi^*)_0$. The expression $\varphi \otimes \varphi^*$ is a section of the bundle $S_+ \otimes S_+^*$ which is identified with Hom (S_+, S_+) . The subscript indicates the removal of the trace, explicitly written as

$$(\varphi \otimes \varphi^*)_0 = \varphi \otimes \varphi^* - \frac{1}{2} |\varphi|^2 \cdot \mathrm{Id}_{S_+}$$

The Seiberg-Witten equations exhibit some nontrivial symmetry arising from *gauge transformations*. In our setting, the group of gauge transformations $\mathcal{G}(X)$ is the space of smooth maps from X to S^1 . A gauge transformation $u \in \mathcal{G}$ acts on configurations by

$$u(A,\varphi) = (A - 2u^{-1}du, u\varphi).$$

One can then see that the Seiberg-Witten equations are *gauge-invariant*: If (A, φ) solves the Seiberg-Witten equations then $u(A, \varphi)$ does as well.

Write $\mathcal{B}(X) = \mathcal{C}(X)/\mathcal{G}$ for the orbit space of configurations under the action of the group of gauge transformations. Examining the action, we find a configuration (A, φ) has trivial stabilizer when φ is nonzero and a stabilizer of S^1 corresponding to the constant gauge transformations when $\varphi = 0$. Configurations satisfying the latter are called *reducible*, and prevent $\mathcal{B}(X)$ from being a Fréchet manifold. However, we see that the space $\mathcal{B}^*(X) \subset \mathcal{B}(X)$ consisting of gauge orbits of irreducible solutions is a Fréchet manifold.

The words "Fréchet manifold" may be off-putting to some. Indeed, one generally works with Sobolev completions of the spaces C(X) and $\mathcal{B}(X)$ (see Chapter 9 of [KM07]). Choosing a base connection A_0 , the affine space of unitary connections on the determinant line bundle is identified with the space of imaginary-valued one-forms on X. Then, using the metric and the standard metric connection one can define a L_k^2 -Sobolev norm on the space of connections and construct its completion. The covariant derivative ∇_{A_0} can then also be used to define an L_k^2 -Sobolev norm on the space of positive spinors. These form the configuration space $C_k(X)$, and the quotient by the action of the group of gauge transformations is $\mathcal{B}_k(X)$ with irreducible locus $\mathcal{B}_k^*(X)$.

The space $C_k(X)$ is a Hilbert space, while the space $\mathcal{B}_k^*(X)$ is a Hilbert manifold.

For the purposes of this thesis, we will omit the subscript and assume by default that we are working with Hilbert completions of sufficiently high Sobolev regularity. The word "Fréchet" will not be used henceforth, and everything will be a Hilbert space or Hilbert manifold.

Let *g* be the underlying Riemannian metric on *X*. Write $M(X,g) \subset \mathcal{B}(X)$ for the space of gauge orbits whose elements satisfy the Seiberg-Witten equations. The notation is given in this manner to emphasize the dependence on the metric. One may wonder if M(X,g) is also a manifold. From the above discussion, this is a possibility only when there are no reducible solutions.

To get rid of reducible solutions, it is often necessary to perturb the Seiberg-Witten equations. Fix ω to

be any imaginary self-dual two-form. Then the perturbed Seiberg-Witten equations are

$$\frac{1}{2}F_A^+ = \rho^{-1}((\varphi \otimes \varphi^*)_0) + \omega,$$
$$D_A^+\varphi = 0.$$

We will write $M(X, g, \omega) \subset \mathcal{B}(X)$ for the moduli space of solutions to the Seiberg-Witten equations perturbed by ω , where g is the Riemannian metric. Our introduction of the metric into this notation will become more apparent after the following discussion. It is also necessary to note that the space $M(X, g, \omega)$ is independent of the choice of Sobolev regularity. By a "folklore theorem" presented in [Mor95], any solution of the Seiberg-Witten equations is gauge-equivalent to a smooth solution.

The equations take a simpler form in the case where $\varphi = 0$, which allows one to deduce the following theorem about reducible solutions.

Theorem 1.2.3. Let $b_2^+(X)$ be the dimension of a maximal positive-definite subspace of $H^2(X; \mathbb{R})$ with respect to the intersection form. Then if $b_2^+(X) \ge 1$, there are no reducible solutions in $M(X, g, \omega)$ for all ω in the complement of a codimension $b_2^+(X)$ linear subspace of the imaginary self-dual two-forms.

Proof. Suppose there exists a reducible solution in $M(X, g, \omega)$. Choose $(A, 0) \in C(X)$ to be a representative of this solution.

Then the connection A satisfies the equation

$$\frac{1}{2}F_A^+ = \omega.$$

Let η be any closed, self-dual two-form with corresponding de Rham cohomology class [η]. The wedge product of a self-dual and an anti-self-dual two-form is always zero.

Then we have

$$\int_X \omega \wedge \eta = \int_X \frac{1}{2} F_A^+ \wedge \eta$$
$$= \frac{1}{2} \int_X F_A \wedge \eta$$
$$= \frac{\pi}{i} (c_1(S_+) \wedge [\eta]) [X].$$

This linear condition cuts out a codimension-one subspace of the space of self-dual two-forms. Putting these linear conditions together for a basis of the space of cohomology classes of closed self-dual two-forms, we obtain the theorem. \Box

Perturbing the equations also solves transversality issues. To show that $M(X, g, \omega)$ is a manifold, we apply the implicit function theorem. The moduli space $M(X, g, \omega)$ can be recast as the zero set of the first-order differential operator

$$\mathfrak{F}_{\omega}: (A,\varphi) \mapsto (\frac{1}{2}F_A^+ - \rho^{-1}((\varphi \otimes \varphi^*)_0) - \omega, D_A^+\varphi).$$

The implicit function theorem requires that the linearization $D\mathfrak{F}_{\omega}$ is surjective and Fredholm at any solution. Although it is Fredholm, it is not a priori true that it is surjective. However, it is surjective for an open, dense subset of perturbations. In this case, we say the moduli space is *regular*.

It is now a simple calculation with the Atiyah-Singer index theorem to deduce the following theorem about the Seiberg-Witten moduli space.

Theorem 1.2.4. If there are no reducible solutions and the moduli space is regular, then $M(X, g, \omega)$ is a compact, finite-dimensional manifold of dimension

$$d = (c_1(det(S_+))^2 - 2\chi(X) - 3\sigma(X))/4$$

where c_1 denotes the first Chern class, χ denotes the Euler characteristic, and σ denotes the signature of the intersection form.

The manifold $M(X, g, \omega)$ can also be oriented in a natural way given a homology orientation on *X*. The theorem below is discussed and proved quite explicitly in [Sal99].

Theorem 1.2.5. Suppose the manifold X is equipped with a homology orientation, a selection of a line in

$$\Lambda^{max}H_1(X;\mathbb{R})\otimes\Lambda^{max}H_2^+(X;\mathbb{R})$$

where $H_2^+(X;\mathbb{R}) \subset H_2(X;\mathbb{R})$ is a maximal positive-definite subspace for the intersection form.

Then the manifold $M(X, g, \omega)$ in the ambient space $\mathcal{B}^*(X)$ has a natural orientation and accompanying fundamental class $[M(X, g, \omega)] \subset H_d(\mathcal{B}^*(X)).$

The S^1 -action of the group of gauge transforamtions on the spinorial part of a configuration gives rise to a natural principal S^1 -bundle

$$P \to \mathcal{B}^*(X).$$

This principal bundle has a Chern class $u = c_1(P) \in H^2(\mathcal{B}^*(X))$.

Definition 1.2.6. Suppose *d* is even. The *Seiberg-Witten invariant* $SW(X, g, \omega)$ is the pairing

$$\langle u^{d/2}, [M(X, g, \omega)] \rangle \in \mathbb{Z}.$$

If *d* is not even, then set $SW(X, g, \omega) = 0$.

Note that if d = 0, then the Seiberg-Witten invariant is simply a signed count of solutions to the Seiberg-Witten equations.

Furthermore, although we have suppressed it here, the Seiberg-Witten invariant certainly depends on the choice of spin^{*c*} structure. We would also like to know whether the Seiberg-Witten invariant is independent of the choice of metric and perturbation.

Definition 1.2.7. The pair (g, ω) of a metric and self-dual perturbation is called *regular* if the moduli space $M(X, g, \omega)$ is regular and contains no reducibles.

Let (g_0, ω_0) and (g_1, ω_1) be two regular pairs. The space of Riemannian metrics is contractible, so there exists a path of pairs $(g_t, \omega_t)_{t \in [0,1]}$ between them.

If $b_2^+(X) \ge 1$, then we are assured that a generic pair (g, ω) is regular. However, if $b_2^+(X) \ge 2$, then a generic *path* (g_t, ω_t) will satisfy the following property.

Define the parameterized moduli space by

$$PM(X, g_t, \omega_t) = \bigcup_{t \in [0,1]} M(X, g_t, \omega_t).$$

We may view $PM(X, g_t, \omega_t)$ as the zero set of the *parameterized Seiberg-Witten operator*:

$$P\mathfrak{F}: (A,\varphi,t) \mapsto (\frac{1}{2}(F_A)^{+_t} - \rho_t((\varphi \otimes \varphi^*)_0) - \omega_t, D_A^+\varphi)$$

where the symbols " $+_t$ " and ρ_t reflect the dependence of the self-dual projection and Clifford multiplication on the metric g_t . Then, for a generic choice of path (g_t, ω_t) , it follows that the parameterized moduli space is a compact manifold itself, and furthermore a *cobordism* from $M(X, g_0, \omega_0)$ to $M(X, g_1, \omega_1)$. The following theorem is now immediate.

Theorem 1.2.8. If $b_2^+(X) \ge 2$, then the Seiberg-Witten invariants of X are independent of the choice of metric and perturbation.

In the case where $b_2^+(X) = 1$, the space of pairs (g, ω) admitting reducible solutions have "codimension one" in the sense of Theorem 1.2.3 within the space of all pairs. The effect of this is the division of the space of pairs into "chambers", where the Seiberg-Witten invariants are the same within one chamber but can change when crossing from one chamber to another. These changes are quantified by the "wall-crossing formula", published independently by [LL95] and [OT96].

1.3 The Seiberg-Witten invariants of manifolds satisfying (A1)

This section will conclude our preliminary discussion. We specialize our discussion of the previous section to a manifold X satisfying (A1), that is, an integral homology $S^1 \times S^3$.

First, assumption (A1) allows us to describe our perturbation forms in a manner that will prove to be useful later.

Lemma 1.3.1. ([MRS11], Lemma 2.1) For any self-dual two-form $\omega \in \Omega^2_+(X)$, there is a unique one-form $\beta \in \Omega^1(X)$ such that $d^+\beta = \omega$, $d^*\beta = 0$, and $\beta \perp \mathcal{H}^1(X)$ where $\mathcal{H}^1(X)$ denotes the space of harmonic 1-forms.

Proof. By assumption (A1), $H^2(X) = 0$. Therefore, the Hodge decomposition implies that $\omega = d\alpha + d^*\gamma$.

Write $\beta = *\gamma$, so this expression simplifies to $\omega = d\alpha + *d\beta$. Then since $\omega = *\omega$, we also have the equality $\omega = *d\alpha + d\beta$. By the uniqueness of the Hodge decomposition, $\alpha = \beta$ and therefore $\omega = d^+\beta$.

The map d^+ vanishes on $\operatorname{im} d$, so we can projection β onto $(\operatorname{im} d)^{\perp} = \ker d^*$. Following this, we can project β onto the space of co-exact 1-forms so that it lies in the orthogonal complement of $\mathcal{H}^1(X)$.

All of the following notation will be altered to reflect the consequence of this lemma.

Recall (see Theorem 5.8 of [Sal99]) that spin^{*c*} structures on *X* are in one-to-one correspondence with integral lifts of the second Stiefel-Whitney class $w_2(X) \in H^2(X; \mathbb{Z}/2)$. Since $b_2(X) = 0$, such a lift is unique and *X* has a unique spin^{*c*} structure. Therefore, all discussion can be made independently of the choice of spin^{*c*} structure.

The fact that $b_2(X) = 0$, however, means that we cannot be assured that $M(X, g, \omega)$ contains no reducibles for a generic choice of (g, ω) .

Geometrically, the reducible solutions are *singularities* of the space $M(X, g, \omega)$ and its ambient space $\mathcal{B}(X)$. Taking inspiration from algebraic geometry, one way to get rid of these singularities is to perform a *blow-up* at the singular locus. Indeed, this is the approach taken by [KM07] to properly construct Seiberg-Witten-Floer homology.

The *blown-up configuration space* $C^{\sigma}(X)$ is the (Hilbert completion of the) space of tuples (A, s, φ) where A is a unitary connection on the determinant spinor line bundle, s is a nonnegative real number, and $\varphi \in C^{\infty}(S^+)$ is a positive spinor such that $||\varphi||_{L^2} = 1$. Tuples with s = 0 are called *reducible*.

Note that while C(X) is a Hilbert space, the space $C^{\sigma}(X)$ is a *Hilbert manifold with boundary*. The boundary is the locus $\{s = 0\}$.

The group of gauge transformations \mathcal{G} acts on $\mathcal{C}^{\sigma}(X)$ by

$$u(A, s, \varphi) = (A - 2u^{-1}du, s, u\varphi).$$

The quotient space $C^{\sigma}(X)/\mathcal{G}$ is denoted by $\mathcal{B}^{\sigma}(X)$. The moduli space $\mathcal{M}(X, g, \beta) \subset \mathcal{B}^{\sigma}(X)$ is the space of tuples satisfying the (*perturbed*) *blown-up Seiberg-Witten equations*:

$$\frac{1}{2}F_A^+ = s^2 \rho^{-1}((\varphi \otimes \varphi^*)_0) + d^+\beta,$$
$$D_A^+\varphi = 0.$$

Here β is a 1-form in the space ker $d^* \cap (\mathcal{H}^1(X))^{\perp}$ introduced in Lemma 1.3.1.

The blown-up spaces come equipped with a projection called the "blow-down" map. This map is realized as

$$\pi: \mathcal{C}^{\sigma}(X) \to \mathcal{C}(X)$$
$$(A, s, \varphi) \mapsto (A, s\varphi)$$

in the case of configuration space and induces identical maps $\pi : \mathcal{B}^{\sigma}(X) \to \mathcal{B}(X)$ and $\mathcal{M}(X, g, \beta) \to \mathcal{M}(X, g, \beta)$. The map is a diffeomorphism outside of the reducible locus.

The following result about the moduli space holds when X satisfies (A1).

Theorem 1.3.2. For a generic choice of pair (g, β) , the moduli space $\mathcal{M}(X, g, \beta)$ is a closed, oriented zero-dimensional manifold.

Proof. The proof will proceed in a couple of steps, and amounts to using the implicit function theorem twice in the proper setting. It follows the line of argument in Lemma 27.1.1 of [KM07].

Let $\mathcal{B}^{\sigma}(X)$ be the double of $\mathcal{B}^{\sigma}(X)$. This can be explicitly described as the space of gauge-equivalence classes of tuples (A, s, φ) with φ having unit length, but the restriction $s \ge 0$ removed.

The space $\widetilde{\mathcal{B}}^{\sigma}(X)$ is a smooth Banach manifold without boundary, and $\mathcal{B}^{\sigma}(X)$ can be identified with its quotient under the involution $(A, s, \varphi) \mapsto (A, -s, \varphi)$.

Let $\widetilde{\mathcal{Z}} \subset \widetilde{\mathcal{B}}^{\sigma}(X)$ be the space of tuples satisfying $D_A^+(\varphi) = 0$.

First, we will show this is a Hilbert submanifold of $\tilde{\mathcal{B}}^{\sigma}(X)$. By the implicit function theorem, it suffices to show that the linearized operator

$$Q: (b,\psi) \mapsto \rho(b)\varphi + D_A^+\psi$$

on the tangent bundle of the double $\tilde{C}^{\sigma}(X)$ of configuration space is surjective. The domain of Q is the space of tuples (b, ψ) where b is an imaginary-valued one-form and ψ is a positive spinor in the real-orthogonal complement of φ .

Suppose some negative spinor η is orthogonal in L^2 to the image of Q. This implies $\langle D_A^-\eta, \psi \rangle_{L^2} = \langle \eta, D_A^+\psi \rangle_{L^2} = 0$ for any ψ in the real-orthogonal complement of φ . The identity $\langle D_A^-\eta, \varphi \rangle_{L^2} = \langle \eta, D_A^+\varphi \rangle_{L^2} = 0$ holds as well, so we conclude $D_A^-\eta = 0$.

By unique continuation for the Dirac operator (see Chapter 7 of [KM07]), if η or φ are nonzero then neither can vanish on an open set in *X*.

On such a sufficiently small open set there is a one-form b and a non-negative smooth function f that is supported inside this open set such that $\rho(b)\eta = f \cdot \varphi$. This can be constructed explicitly using local coordinates.

However, $\langle \rho(b)\eta, \varphi \rangle_{L^2} = \langle \eta, \rho(b)\varphi \rangle_{L^2} = 0$ by our initial orthogonality assumption, so we require η to vanish on an open set, which in turn implies $\eta = 0$. This tells us that Q has dense image. It is also elliptic and so has closed range. It follows that Q is surjective as desired and \tilde{Z} is a Hilbert submanifold.

Then, let $\widetilde{\mathcal{M}}(X, g, \beta) \subset \widetilde{\mathcal{Z}}$ be the space of tuples satisfying

$$\chi(A, s, \varphi) = \frac{1}{2} F_A^+ - s^2 \rho^{-1} ((\varphi \otimes \varphi^*)_0) = d^+ \beta$$

The operator χ has Fredholm linearization. Applying the implicit function theorem and the Sard-Smale theorem ([Sma65]), there is a residual set of perturbations β such that $\widetilde{\mathcal{M}}(X, g, \beta) = \chi^{-1}(\beta)$ is a manifold with dimension $\operatorname{ind}(\chi)$.

The tangent space to $\widetilde{\mathcal{Z}}$ at a tuple (A, s, φ) is the space of tuples (b, r, ψ) where *b* is an imaginary-valued one-form, *r* is a real number, and ψ is a positive spinor satisfying the equations

$$\begin{split} \rho(b)\varphi + D_A^+\psi &= 0\\ \langle \varphi, \psi \rangle_{L^2} &= 0\\ -d^*b + irs \mathrm{Re} \langle i\varphi, \psi \rangle &= 0. \end{split}$$

The last two conditions are the conditions defining the tangent space of $\mathcal{B}^{\sigma}(X)$ (see Chapter 9 of [KM07]). If we remove zeroth order terms, the index of the linearization of $\chi \chi$ at a point (A, s, φ) is equal to the index of the linear operator

$$S: (b, r, \psi) \mapsto (D^+_A \psi, d^+ b, d^* b)$$

on the space of tuples (b, r, ψ) satisfying $\langle \phi, \psi \rangle_{L^2} = 0$.

The index of this operator is the sum of the indices of the Dirac operator and the operator $d^+ \oplus d^*$.

By the Atiyah-Singer index theorem [AS68], the (real) index of the spin^c Dirac operator D_A^+ is equal to the quantity

$$\frac{1}{4}(c_1(S^+)^2 - \sigma(X)) = 0.$$

We can also show the operator $d^+ \oplus d^*$ has index zero on *X* without using the index theorem. First, suppose *b* satisfies $d^+b = d^*b = 0$.

By the Hodge decomposition, *b* decomposes into an orthogonal sum $\beta + d^*\gamma$, where β is harmonic. Since $d^+b = 0$, the two-form $db = dd^*\gamma$ is anti-self-dual. Taking the Hodge star, we find $db = -*db = -*dd^*\gamma = -*d*d*\gamma = -*d*d*\gamma = -d^*d(*\gamma)$. Therefore, *db* is both exact and co-exact, which implies it is equal to zero and $b = \beta$.

Therefore, the kernel of $d^+ \oplus d^*$ is exactly the harmonic 1-forms, the space of which has dimension $b_1(X)$.

The adjoint of $d^+ \oplus d^*$ is the operator $d^* \oplus d$. The kernel of this operator consists of all pairs (γ, f) of coclosed self-dual two-forms and closed smooth functions, respectively. There are in bijection with harmonic self-dual two-forms and harmonic functions by a similar Hodge decomposition argument to above, and it follows that the kernel has dimension $b_0(X) + b_2^+(X)$ in this case.

We conclude that the index of $d^+ \oplus d^*$ is $b_1(X) - b_0(X) - b_2^+(X) = 0$.

The restriction of χ to the locus $\{s = 0\}$ has index -1. This follows from the fact that the linearization of χ vanishes on the normal bundle of the locus $\{s = 0\}$, as the fibers consist of tuples of the form (0, r, 0). Therefore, the intersection of $\widetilde{\mathcal{M}}(X, g, \beta)$ with the locus $\{s = 0\}$ is empty for generic β .

The proofs of compactness and orientability of the moduli space $\widetilde{\mathcal{M}}(X, g, \beta)$ are quite similar to the regular setting and will be omitted.

We conclude the proof by identifying $\mathcal{M}(X, g, \beta)$ with the subset of $\widetilde{\mathcal{M}}(X, g, \beta)$ with nonnegative *s*-coordinate, and its boundary with the intersection $\widetilde{\mathcal{M}}(X, g, \beta) \cap \{s = 0\}$.

Note that this theorem requires working in the blown-up setting, as we can only apply the implicit function theorem to a function mapping out of a smooth manifold.

As before, we will call a pair (g, β) *regular* if $\mathcal{M}(X, g, \beta)$ is a closed manifold. The result above implies that there is a residual set of regular pairs.

However, since $b_2^+(X) = 0$, it is not possible in general to choose a generic path (g_t, β_t) such that the parameterized moduli space $PM(X, g_t, \beta_t)$ does not have any singularities. The best that we can do is the following result.

Theorem 1.3.3. Let $(g_t, \beta_t)_{t \in [0,1]}$ be a path of pairs such that the endpoints (g_0, β_0) and (g_1, β_1) are both regular. *The manifold*

$$P\mathcal{M}(X, g_t, \beta_t) = \bigcup_{t \in [0,1]} \mathcal{M}(X, g_t, \beta_t)$$

is a compact, oriented one-dimensional manifold with boundary equal to the disjoint union

$$\mathcal{M}(X, g_0, \beta_0) \sqcup \mathcal{M}(X, g_1, \beta_1) \sqcup \mathcal{M}^0$$

where \mathcal{M}^0 is the zero-dimensional manifold consisting of gauge-equivalence classes of reducible solutions.

Proof. The proof is analogous to the previous theorem. One applies instead the implicit function theorem to parameterized versions of the operators in the proof of that theorem. \Box



Figure 1.3: The parameterized moduli space $\mathcal{M}(X, g_t, \beta_t)$ with boundary components labelled.

Letting # denote the oriented count of points again, we find an expression for the difference in the Seiberg-Witten invariants associated to (g_0, β_0) and (g_1, β_1) :

$$#\mathcal{M}(X, g_1, \beta_1) - #\mathcal{M}(X, g_0, \beta_0) = #\mathcal{M}^0.$$

As we have seen, the Seiberg-Witten invariants are not as well-behaved as we would like in the case where $b_2^+(X) = 0$. However, the assumption (A1) makes the situation somewhat easier to deal with by making all of the moduli spaces zero-dimensional and providing us with a generic set of regular pairs.

The strategy going forward will be to add a "correction term" to the Seiberg-Witten invariants that will make them independent of the choice of metric and perturbation. This correction term is in part the index of a Dirac operator on the manifold Z_+ introduced in Section 1.1. The explicit definition of the correction term and the proof of the metric and perturbation independence of the corrected Seiberg-Witten invariant will be the subject of the next chapter.

Chapter 2

The Lambda Invariant

In this chapter, we present the construction of the modified Seiberg-Witten invariant $\lambda_{SW}(X)$ for a manifold X satisfying assumption (A1). The invariant $\lambda_{SW}(X)$ is independent of the metric and perturbation chosen on X, and its construction and proof of independence was first carried out in the paper [MRS11] by Mrowka, Ruberman, and Saveliev. All results in this chapter are attributed to their work.

The exposition is largely structured like that of [MRS11], but some of the proofs have been re-written and extra explanation has been added wherever deemed necessary.

Section 1 begins with an overview of a particular differential operator associated to *X* known as the end-periodic Dirac operator. Such operators have associated "Fourier-Laplace" transforms that may be used to understand their Fredholm theory. We will also define the Fourier-Laplace transform and prove some theoretical results. Finally, at the end of the section a rigorous definition of the invariant $\lambda_{SW}(X)$ is given.

The object of interest of the next two sections is the index of the end-periodic Dirac operator. This operator depends on a choice of a metric and perturbation pair as well, and Section 2 applies the theory of the previous section to prove a preliminary formula regarding the change of index between different pairs of metric and perturbation. This culminates in the work of Section 3, which relates the change of index of the end-periodic Dirac operator between two regular pairs to the *spectral flow* of their Fourier-Laplace transforms. In Section 4, a correspondence between the oriented count of reducible solutions along a path of pairs (g_t , β_t) is equated with the spectral flow of Section 3, which proves the invariance of $\lambda_{SW}(X)$ under metric and perturbation pairs.

2.1 The end-periodic Dirac operator and its Fredholm theory

Let *X* be a closed, oriented 4-manifold satisfying assumptions (A1) equipped with a Riemannian metric *g* and a homology orientation. Let *Y* be an embedded compact 3-manifold representing the generator of $H_3(X;\mathbb{Z})$ determined by the homology orientation.

Since $H_2(X; \mathbb{Z}) = 0$, in addition to a spin^{*c*} structure *X* also admits some *spin* structure *s*. Furthermore, the spinor bundles $S^{\pm} \rightarrow X$ associated to this spin bundle are the same as the spinor bundles for the spin^{*c*} structure \mathfrak{s} .

Recall our construction of the following associated manifolds in Section 1.1:

• *W*, the homology cobordism from *Y* to *Y* constructed by cutting open *X* along *Y*.

- \widetilde{X} , the infinite cyclic cover of X equal to a \mathbb{Z} -fold end-to-end gluing of copies of W.
- X_+ , the "positive end" of \tilde{X} .
- Z_+ , the gluing of a compact spin^c 4-manifold Z bounding Y to X_+ .

These constructions were done in Section 1.1 in the case where Y was a homology 3-sphere, but they are well-defined even when Y is any compact 3-manifold. Note that W is also still a homology cobordism from Y to itself, which is clear from modification of the proof of Theorem 1.1.1.

The spin structure naturally pulls back to the infinite cyclic cover \tilde{X} . The spinor bundles are the pullbacks of the spinor bundles over X along the projection $\pi : \tilde{X} \to X$.

The spin structure and spinor bundles can then be defined on X_+ by restriction. Finally, to define a spin structure on Z_+ we must pick an extension of the spin structure to the manifold Z. The spinor bundles on the manifolds \tilde{X} , X_+ , and Z_+ will be all be denoted by S^{\pm} as well, but the base manifold will always be clear in any use of this notation.

The same process works to extend geometric data such the metric g or a perturbation one-form β to these three noncompact manifolds.

To give some context to the section title, the manifold Z_+ is an example of an *end-periodic manifold with end modeled on* \tilde{X} . These objects were first introduced in the 1987 paper of Taubes [Tau87]. In this paper, Taubes constructed the Donaldson invariants for end-periodic 4-manifolds and used them to prove the existence of a continuum of \mathbb{R}^4 with exotic smooth structure. The Fredholm theory for end-periodic operators that was developed in the course of doing so serves as a foundation for all of the analysis that we will describe here.

Given the induced spin structure on Z_+ , there is a *spin* Dirac operator

$$D^+(Z_+): L^2_1(Z_+; S^+) \to L^2(Z_+; S^-).$$

Given some 1-form β , we can also consider the perturbed Dirac operator

$$D^+(Z_+,\beta) = D^+(Z_+) + \rho(\beta) : L_1^2(Z_+;S^+) \to L^2(Z_+;S^-).$$

Recall the infinite cyclic cover $\widetilde{X} \to X$ comes equipped with a covering transformation T. Concretely, T is the map that sends a point $x \in W_i$ to its corresponding point in W_{i+1} . This transformation is also well-defined on the periodic end $X_+ \subset Z_+$.

The Dirac operator on Z_+ commutes with this translation on the end. For any spinor φ , there is a pointwise equality find $D^+(\varphi \circ T) = D^+(\varphi) \circ T$ on the end X_+ . Therefore, the Dirac operator is an example of an *end-periodic operator* in the sense of [Tau87].

It is not a priori clear that this end-periodic Dirac operator is Fredholm, and indeed it is not always true that it is.

Before we discuss this Fredholm theory, however, it is necessary to widen our consideration to a whole class of weighted Sobolev spaces on the manifolds \tilde{X} and Z_+ .

Let $E \to X$ be an abstract vector bundle and $\tilde{E} \to \tilde{X}$ its pullback to the infinite cyclic cover. Let $\tilde{E} \to Z_+$ be some vector bundle constructed by extending the restriction $\tilde{E}|_{X_+}$ over Z_+ .

Pick a smooth map $f : X \to S^1$. We will suppose that the closed one-form $f^*(d\theta)$ represents the generator of $H^1(X;\mathbb{Z})$ fixed by the homology orientation.s

The map f has a natural lift $\tilde{f}: \tilde{X} \to \mathbb{R}$ such that $\tilde{f}(T(x)) = \tilde{f} + 1$. Fixing a weight $\delta \in \mathbb{R}$, restrict \tilde{f} to X_+ and pick an extension h of $\delta \cdot \tilde{f}$ to Z_+ .

Definition 2.1.1. For any fixed weight $\delta \in \mathbb{R}$, the Hilbert space of sections $L^2_{k,\delta}(Z_+; \widetilde{E})$ is the space of all sections φ such that $e^h \varphi \in L^2_k(Z_+; \widetilde{E})$. More compactly,

$$L^2_{k,\delta}(Z_+;\widetilde{E}) = e^{-\delta h} L^2_k(Z_+;\widetilde{E}).$$

We can define similar spaces on the cyclic cover using two weights $\delta_1, \delta_2 \in \mathbb{R}$. In this setting, we pick a smooth function $\delta : \widetilde{X} \to \mathbb{R}$ such that $\delta(x) = \delta_1$ for $x \in W_i$ for $i \leq -1$, and $\delta(x) = \delta_2$ for $x \in W_i$ for $i \geq 1$.

Definition 2.1.2. The space $L^2_{k,\delta_1,\delta_2}(\widetilde{X};\widetilde{E})$ is given by

$$L^2_{k,\delta_1,\delta_2}(\widetilde{X};\widetilde{E}) = e^{-\delta \widetilde{f}} L^2_k(\widetilde{X};\widetilde{E}).$$

If $\delta_1 = \delta_2 = \delta$ for some $\delta \in \mathbb{R}$, the weighting function is the constant function to δ and we denote the resulting weighted Sobolev space by $L^2_{k,\delta}(\widetilde{X}; \widetilde{E})$.

Note that when $\delta = 0$ or $\delta_1 = \delta_2 = 0$, we have

$$L^2_{k,\delta}(Z_+;\widetilde{E}) \simeq L^2_k(Z_+;\widetilde{E})$$

and

$$L^2_{k,\delta_1,\delta_2}(\widetilde{X};\widetilde{E}) \simeq L^2_k(\widetilde{X};\widetilde{E})$$

respectively.

It is immediate that the (perturbed) Dirac operator defines a differential operator

$$D^+(Z_+,\beta): L^2_{k\,\delta}(Z_+;\widetilde{E}) \to L^2_{k-1\,\delta}(Z_+;\widetilde{E})$$

as well.

2.1.1 The Fourier-Laplace transform and Taubes' theorem

The Fredholm properties of the Dirac operator are tied to a family of operators known as its *Fourier-Laplace transform*, introduced by Taubes in [Tau87].

Now take a pair of abstract bundles $E \to X$ and $\tilde{E} \to \tilde{X}$, we observe that E can be identified with the quotient of \tilde{E} by the action of pullback by the covering transformation T. Therefore, the sections of E are in correspondence with translation-invariant sections of \tilde{E} .

The Fourier-Laplace transform is an operation that takes a section of \tilde{E} and transforms it to a family of translation-invariant sections of \tilde{E} , and therefore a family of sections of E. Fix weights $\delta_1, \delta_2 \in \mathbb{R}$ and let $\tilde{f}, \delta : \tilde{X} \to \mathbb{R}$ be functions defined as in the previous subsection.

Definition 2.1.3. Fix $\mu \in \mathbb{C}$. For any section $\varphi \in L^2_{k,\delta_1,\delta_2}(X; \widetilde{E})$, its *Fourier-Laplace transform* is defined to be the sum

$$\hat{\varphi}_{\mu} = e^{\mu \tilde{f}} \sum_{n=-\infty}^{\infty} e^{\mu n} (\varphi \circ T^n).$$

We have

$$\hat{\varphi}_{\mu}(Tx) = e^{\mu \tilde{f}(Tx)} \sum_{n=-\infty}^{\infty} e^{\mu n} (\varphi \circ T^{n})(Tx)$$
$$= e^{\mu (\tilde{f}(x)+1)} \sum_{n=-\infty}^{\infty} e^{\mu n} (\varphi \circ T^{n+1})(x)$$
$$= e^{\mu \tilde{f}(x)} \sum_{n=-\infty}^{\infty} e^{\mu (n+1)} (\varphi \circ T^{n+1})(x)$$
$$= \hat{\varphi}_{\mu}(x)$$

so the Fourier-Laplace transform is indeed translation-invariant and defines a section of *E* whenever the sum in the definition is well-defined. Note by definition, whenever the Fourier-Laplace transform is well-defined, the family of sections $\{\hat{\varphi}_{\mu}\}$ is *holomorphic* in the variable μ .

We can characterize when the sum is well-defined quite easily.

Lemma 2.1.4. Let $\varphi \in L^2_{k,\delta_1,\delta_2}(\widetilde{X};\widetilde{E})$ be a section that is supported on X_+ . Then the family $\{\hat{\varphi}_{\mu}\}$ is holomorphic for all μ satisfying $Re(\mu) < \delta_2$.

Proof. It suffices to uniformly bound $|\hat{\varphi}_{\mu}(x)|_{L^2}^2$ for any $x \in W_0$. Note that since φ is supported in X_+ , $(\varphi \circ T^n)(x) = 0$ for any n < 0.

By the triangle inequality and the Cauchy-Schwarz inequality, we can write

$$\begin{aligned} |\hat{\varphi}_{\mu}|_{L^{2}}^{2} &\leq \sum_{n=-\infty}^{\infty} |e^{\mu(\tilde{f}+n)}(\varphi \circ T^{n})(x)|_{L^{2}}^{2} \\ &= \sum_{n \geq 0} |e^{\mu(\tilde{f}+n)}(\varphi \circ T^{n})(x)|_{L^{2}} \\ &= \sum_{n \geq 0} |e^{(\mu-\delta_{2})(\tilde{f}+n)}(e^{\delta_{2}}(\tilde{f}+n)(\varphi \circ T^{n}(x)))|_{L^{2}}^{2} \\ &\leq (\sum_{n=\geq 0} |e^{(\mu-\delta_{2})(\tilde{f}+n)}|)^{2} (\sum_{n \geq 0} ||e^{\delta_{2}\tilde{f}}(\varphi \circ T^{n})(x)||_{L^{2}}^{2}). \end{aligned}$$

The second term in the product is bounded by $||e^{\delta_2 \tilde{f}} \varphi||_{L^2}^2$, which is in turn bounded since $\varphi \in L^2_{k,\delta_1,\delta_2}(\tilde{X};\tilde{E})$. The first term is bounded if and only if $\operatorname{Re}(\mu) < \delta_2$.

The same argument works for a section supported in the other end of \tilde{X} . By applying a cutoff function, we arrive at the following corollary.

Corollary 2.1.5. For any section $\varphi \in L^2_{k,\delta_1,\delta_2}(\widetilde{X};\widetilde{E})$, the family $\{\hat{\varphi}_{\mu}\}$ is holomorphic for all μ satisfying $\delta_1 < Re(\mu) < \delta_2$.

As is the case with its namesake, the Fourier-Laplace transform is invertible. Furthermore, the data of the entire family $\hat{\varphi}_{\mu}$ is not required. Instead, we only need to know the sections $\hat{\varphi}_{\mu}$ for μ in an interval $I(\nu) = \{\nu + 2\pi i \alpha \mid \alpha \in [0, 1]\}$. Given this, the inverse Fourier-Laplace transform is defined by

$$\varphi = \frac{1}{2\pi i} \int_{I(\nu)} e^{-\mu \tilde{f}} \hat{\varphi}_{\mu} \, d\mu$$

The proof that this is an honest inverse is a bit of elementary complex analysis.

In addition to Fourier-Laplace transforms of sections, we can also consider Fourier-Laplace transforms of differential operators. Let E and F be two bundles over X and \tilde{E} , \tilde{F} their respective pullbacks over \tilde{X} . Let $D: L^2_{1,\delta}(\tilde{X}; \tilde{E}) \to L^2_{\delta}(\tilde{X}; \tilde{F})$ be a first-order differential operator.

Then, we can define a holomorphic family of differential operators

$$\hat{D}_{\mu}: L^2_1(X; E) \to L^2(X; F)$$

by requiring that the following diagram commute:

The vertical arrows denote the Fourier-Laplace transform at μ .

The Fredholm-ness of D was shown by Taubes to rely quite explicitly on the family $\{\hat{D}_{\mu}\}$.

Lemma 2.1.6. ([Tau87], Lemma 4.3) The operator D is Fredholm if and only if the operators \hat{D}_{μ} are invertible for all μ with $Re(\mu) = \delta$.

We can apply this lemma to the situation of the end-periodic Dirac operator $D^+(\tilde{X}, g, \beta)$. First, we calculate the Fourier-Laplace transform.

Lemma 2.1.7. The Fourier-Laplace transform of $D^+(\widetilde{X}, g, \beta)$ at $\mu \in \mathbb{C}$ is the operator

$$D^+_{\mu}(X,g,\beta) = D^+(X,g,\beta) - \mu \cdot \rho(f^*(d\theta))$$

where $d\theta$ is the generator of $H^1(S^1; \mathbb{Z})$ corresponding to its standard orientation.

Proof. Set $D = D^+(\widetilde{X}, g, \beta)$.

For any spinor $\varphi \in L^2_1(\widetilde{X}; S^+)$, note that $D(e^{\mu \widetilde{f}} \varphi) = e^{\mu \widetilde{f}} (\mu \cdot \rho(d\widetilde{f}) \varphi + D\varphi)$.

To conclude the lemma, we must show that the operators $D^+_{\mu}(X, g, \beta)$ satisfy the identity

$$D^+_\mu(X, g, \beta)\hat{\varphi}_\mu = \widehat{D\varphi}_\mu.$$

To calculate the left-hand side, we identify $\hat{\varphi}_{\mu}$ with a translation-invariant section of S^+ . Then, by definition of the end-periodic Dirac operator, the section $D^+(X, g, \beta)\hat{\varphi}_{\mu}$ of S^+ and the translation-invariant section $D\hat{\varphi}_{\mu}$ are identified. Under this identification, it is also the case that $\rho(f^*(d\theta))$ and $\rho(d\tilde{f})$ are identified.

Therefore, we may identify the left-hand side with

$$\begin{split} (D - \mu \cdot \rho(d\tilde{f}))\hat{\varphi}_{\mu} &= \sum_{n \in \mathbb{Z}} (D - \mu \cdot \rho(d\tilde{f}))(e^{\mu(\tilde{f}+n)}(\varphi \circ T^n)) \\ &= \sum_{n \in \mathbb{Z}} e^{\mu(\tilde{f}+n)} D(\varphi \circ T^n) \\ &= \sum_{n \in \mathbb{Z}} e^{\mu(\tilde{f}+n)} (D\varphi \circ T^n) \\ &= \widehat{D\varphi}_{\mu}. \end{split}$$

The second equality uses our commutation formula for the Dirac operator established earlier, while the third equality uses the fact that the Dirac operator commutes with translation. \Box

The proof here extends immediately to the case of any first-order differential operator D between vector bundles E and F over X. Denote by \tilde{E} and \tilde{F} the respective pullbacks of E and F by the covering map from \tilde{X} to X. The operator D induces an end-periodic operator \tilde{D} from \tilde{E} to \tilde{F} , which has an associatd family of Fourier-Laplace transforms $\{\hat{D}_{\mu}\}$. Letting $\sigma(D, -)$ be the symbol of the operator, we obtain

$$\hat{D}_{\mu} = D - \mu \cdot \sigma(D, d\tilde{f}).$$

Finally, we are in a position to prove our first major result about the Fredholm theory of the end-periodic Dirac operator.

Theorem 2.1.8. Let (g, β) be a regular pair. Then the Dirac operator

$$D^+(\widetilde{X}, g, \beta) : L^2_{1,\delta}(\widetilde{X}; S^+) \to L^2_{\delta}(\widetilde{X}; S^-)$$

is Fredholm for any $\delta \in \mathbb{R}$ *.*

Proof. Suppose first that $\delta = 0$.

By Lemma 2.1.6 and Lemma 2.1.7, the operator is Fredholm if and only if the family of operators

$$\hat{D}_{\mu} = D^+(X, g, \beta) - \mu \cdot \rho(f^*(d\theta))$$

are invertible for all μ such that $\operatorname{Re}(\mu) = 0$.

Taking the canonical spin connection as a base, the one-forms $\beta - \mu \cdot f^*(d\theta)$ are identified with a family of spin^{*c*} connections A_{μ} whose corresponding Dirac operators satisfy

$$D_{A_{\mu}}^{+} = \hat{D}_{\mu}.$$

Furthermore, we have $F_{A_{\mu}}^{+} = d^{+}\beta$ for any μ since the form $f^{*}(d\theta)$ is closed.

If (g,β) is a regular pair, then the operators $D_{A_{\mu}}^{+}$ must be invertible. If $D_{A_{\mu}}^{+}$ is not invertible for some μ with $\text{Re}(\mu) = 0$, then there is a spinor φ of L^{2} norm 1 in its kernel.

Therefore, the tuple $(A_{\mu}, 0, \varphi)$ forms a reducible solution to the blown-up Seiberg-Witten equations perturbed by β . We have assumed that there are no reducible solutions, so we arrive at a contradiction and all of the operators are invertible.

To extend to the case of arbitrary $\delta \in \mathbb{R}$, define our connections A_{μ} as before.

Suppose $\varphi \in L^2_{1,\delta}(\widetilde{X}; S^+)$ has unit L^2 norm and is in the kernel of $D^+_{A_{\mu}}$ for some μ with $\operatorname{Re}(\mu) = \delta$. Then recall that $e^{\delta \widetilde{f}} \varphi \in L^2_1(\widetilde{X}; S^+)$ and also has unit L^2 norm.

By our commutation formula for the Dirac operator, the equation $D^+_{A_{\mu}}(\varphi) = 0$ is equivalent to the identity $D^+_{A_{\mu-\delta}}(e^{\delta \hat{f}}\varphi) = 0$.

Therefore, the tuple $(A_{\mu-\delta}, 0, e^{\delta \tilde{f}} \varphi)$ is a reducible solution and we again arrive at a contradiction. \Box

As we can see from this proof, the Fourier-Laplace transform connects reducible solutions of the blownup Seiberg-Witten equations to the Fredholm theory of the end-periodic Dirac operator. This serves as a rough motivation for the definition of the correction term in $\lambda_{SW}(X)$, and further analysis will show that this definition is indeed the correct choice. The theoretical results can also be transported directly to the case of an end-periodic Dirac operator

$$D^+(Z_+, g, \beta) : L^2_{1,\delta}(Z_+; S^+) \to L^2_{\delta}(Z_+; S^-).$$

This is a direct consequence of Lemma 4.1 in [Tau87].

Lemma 2.1.9. The Dirac operator $D^+(Z_+, g, \beta)$ is Fredholm if and only if the corresponding Dirac operator $D^+(\tilde{X}, g, \beta)$ is Fredholm.

Note that the construction of the Dirac operator and the weighted Sobolev spaces on Z_+ involve the additional choices of extensions of the function \tilde{f} , the metric g, the perturbation one-form β and the spin structure to Z_+ . However, the index is invariant under these choices, a result that is proved in Section 2.2.1 using the *excision principle*.

2.1.2 The correction term

Now we introduce the correction term and the full, rigorous definition of the invariant $\lambda_{SW}(X)$.

Definition 2.1.10. A path of metrics (g_t, β_t) for $t \in [0, 1]$ is a *special path* if it has regular endpoints and for any $\tau \in [0, 1]$ such that the Seiberg-Witten equations for (g_τ, β_τ) have reducible solutions then there is some $\varepsilon > 0$ such that the metric is constant on $(\tau - \varepsilon, \tau + \varepsilon)$, i.e. $g_{\tau'} = g_{\tau}$ for any $\tau' \in (\tau - \varepsilon, \tau + \varepsilon)$.

From now on, any path of metrics (g_t, β_t) between regular pairs (g_0, β_0) and (g_1, β_1) that we discuss will be assumed to be special. Most of the subsequent will not require that path be special, but rather that it has regular endpoints. However, it is a useful assumption for simplifying the local structure of the parameterized moduli space. The justification for the ability to fix a special path comes from Appendix A of [MRS11], which proves the following theorem:

Theorem 2.1.11. Any path of metrics (g_t, β_t) with regular endpoints admits an endpoint-preserving homotopy to a special path.

Fix a regular pair (g,β) . Then, for any end-periodic manifold Z_+ with end modeled on \widetilde{X} , define the quantity

$$w(X, g, \beta) = \operatorname{ind}_{\mathbb{C}} D^+(Z_+, g, \beta) - \operatorname{sign}(Z)/8$$

The operator $D^+(Z_+, g, \beta)$ is the usual end-periodic Dirac operator between unweighted Sobolev spaces of sections. The signature term is added to make $w(X, g, \beta)$ independent of the choice of Z_+ . This independence will also be shown later using the excision principle.

Then, define for any regular pair (g, β) the rational invariant

$$\lambda_{SW}(X) = \#\mathcal{M}(X, g, \beta) - w(X, g, \beta).$$

The objective, as stated before, is to prove the following main theorem.

Theorem 2.1.12. $\lambda_{SW}(X)$ is invariant of the choice of regular pair (g, β) used in its definition.

2.1.3 The Fourier-Laplace transform along a path

Let (g_t, β_t) for $t \in [0, 1]$ be a special path.

By Theorem 1.3.3, calculating the change in Seiberg-Witten invariants from (g_0, β_0) to (g_1, β_1) is equivalent to a signed count of the reducible solutions to the blown-up Seiberg-Witten equations across all pairs (g_t, β_t) in the path.

The appearance of a reducible solution in turn corresponds to some Fourier-Laplace transform $D_{\mu}(X, g_t, \beta_t)$ ceasing to be invertible, and this is the device through which we may count reducible solutions.

Definition 2.1.13. For any pair (g, β) , its spectral set $\Sigma(g, \beta) \subset \mathbb{C}$ is the set of $\mu \in \mathbb{C}$ such that $\hat{D}^+_{\mu}(X, g, \beta)$ is not invertible.

There is an possibility that the spectral set $\Sigma(g,\beta)$ may depend on the function $f: X \to S^1$, but this is not the case.

Lemma 2.1.14. The spectral set $\Sigma(g,\beta)$ is independent of the choice of function $f: X \to S^1$.

Proof. Any two choices $f, f' : X \to S^1$ satisfy the property that the pullbacks $f^*(d\theta)$ and $(f')^*(d\theta)$ represent the same element in $H^1(X;\mathbb{Z})$. Therefore, the difference $(f')^*(d\theta) - f^*(d\theta)$ is an exact form dh, and the corresponding Dirac operators differ by $\mu \cdot \rho(dh)$.

Let $\hat{D}^+_{\mu}(X, g, \beta)$ be the Dirac operator for the function f. Then by our commutation identity,

$$(\hat{D}^+_{\mu}(X,g,\beta) - \mu \cdot \rho(dh))\varphi = e^{\mu h}\hat{D}^+_{\mu}(X,g,\beta)e^{-\mu h}\varphi$$

for any spinor φ , and the statement of the lemma is immediate.

Using some technical results from functional analysis, we can achieve a better understanding of the spectral sets along a special path (g_t, β_t) for $t \in [0, 1]$.

Lemma 2.1.15. For any fixed time τ , there exists some $\mu_0 \in \mathbb{C}$ such that $\hat{D}^+_{\mu_0}(X, g_\tau, \beta_\tau)$ is invertible.

Proof. Since (g_t, β_t) is a special path, the space of reducible solutions \mathcal{M}^0 associated to this path consists of finitely many points, so there will be a finite number of reducible solutions in $\mathcal{M}(X, g_{\tau}, \beta_{\tau})$ for any $\tau \in [0, 1]$.

Suppose there is some unit-length spinor φ and $\mu \in i\mathbb{R}$ such that φ is in the kernel of $\hat{D}^+_{\mu}(X, g_{\tau}, \beta_{\tau})$. Let A denote the connection associated to the imaginary-valued one-form $\beta + \mu \cdot f^*(d\theta)$.

Then, just like in the proof of 2.1.8, one finds that $(A, 0, \varphi)$ is a reducible solution for the pair (g_{τ}, β_{τ}) . Considering this over all $\mu \in i\mathbb{R}$, it follows that there can only be finitely many μ such that $\hat{D}^+_{\mu}(X, g_{\tau}, \beta_{\tau})$ has a nonzero kernel. Otherwise, there would be infinitely many reducible solutions for the pair (g_{τ}, β_{τ}) , which contradicts the finiteness of \mathcal{M}^0 .

Since all of these operators are of index zero, this is sufficient to prove the lemma.

This lemma, combined with the standard spectral theory of compact operators, yields the following result on the spectral set at any fixed time τ .

Lemma 2.1.16. For any $t \in [0,1]$, the spectral set $\Sigma(g_t, \beta_t)$ is a discrete subset of \mathbb{C} , and the inverse of $\hat{D}^+_{\mu}(X, g_t, \beta_t)$ *is a meromorphic function of* μ *.*

Proof. As in the previous lemma, we write the family $\hat{D}^+_{\mu}(X, g_t, \beta_t)$ as $T + \mu A$, where $T = D^+(X, g_t, \beta_t)$ and $A = -\rho(f^*(d\theta)).$

We know that $T + \mu_0 A$ is invertible for some $\mu_0 \in \mathbb{C}$.

Therefore, the operator $T + \mu A$ is invertible if and only if

$$(T + \mu A)(T + \mu_0 A)^{-1} = 1 + (\mu - \mu_0)A(T + \mu_0 A)^{-1}$$

is invertible.

By definition, this is true if and only if $-(\mu - \mu_0)^{-1}$ is in the spectrum of the operator $A(T + \mu_0 A)^{-1}$. Since *A* is compact and $(T + \mu_0 A)^{-1}$ is bounded, the operator $K = A(T + \mu_0 A)^{-1}$ is compact. Now we can apply the spectral theory of compact operators.

By Theorem 6.26 of Chapter III of [Kat66], the spectrum of *K* is countable and can only accumulate at zero. We have shown that $\Sigma(g_t, \beta_t)$ is in bijection with the spectrum of *K* via the map $\mu \mapsto -(\mu - \mu_0)^{-1}$, which is sufficient to show that it is discrete.

From the discussion in Section 5 of Chapter III in [Kat66], we also have that the resolvent $R(\lambda) = (K - \lambda I)^{-1}$ is meromorphic in λ away from the accumulation point $\lambda = 0$, which shows that $(T + \mu A)^{-1}$ is meromorphic in μ .

This lemma, combined with Lemma 2.1.6, gives an alternate proof of Theorem 3.1 of [Tau87] in the case of the Dirac operator.

Corollary 2.1.17. The end-periodic operator

$$D^+(Z_+, g, \beta) : L^2_{1,\delta}(Z_+; S^+) \to L^2_{\delta}(Z_+; S^-)$$

is Fredholm for all $\delta \in \mathbb{R}$ *outside of some discrete subset with no accumulation point.*

Next, we can examine the parameterized spectral set

$$\Sigma_P(g_t, \beta_t) = \bigcup_{t \in [0,1]} \Sigma(g_t, \beta_t) \subset \mathbb{C}.$$

The following theorem will make use of another technical result from [Kat66], which requires the following setup.

Let $\tau \in [0,1]$ be such that the corresponding spectral set $\Sigma(g_{\tau}, \beta_{\tau})$ is nonempty. Then, for any $\eta \in \Sigma(g_{\tau}, \beta_{\tau})$, we can pick a small circle *L* around it that does not contain any other element of the spectral set. Then, the following operator is well-defined:

$$P_{\eta} = \int_{L} (\hat{D}_{\mu}(X, g_{\tau}, \beta_{\tau}))^{-1} d\mu$$

Now, we can state our theorem.

Theorem 2.1.18. Let (g_t, β_t) be a path with regular endpoints as above and suppose there is a time $\tau \in [0, 1]$ and a point $\eta \in \Sigma(g_\tau, \beta_\tau)$ such that the rank of the operator P_η is one. Then, there is an open neighborhood U of η and some real $\varepsilon > 0$ such that the intersection

$$\bigcup_{|\tau-s|<\varepsilon} \Sigma(g_s,\beta_s) \cap U$$

is an embedded curve in the complex plane.

Proof. As before, set $T = D^+(X, g_\tau, \beta_\tau)$ and $A = \rho(f^*(d\theta))$. Pick $\mu_0 \in \mathbb{C}$ such that $T + \mu_0 A$ is invertible, and define the compact operator $K = A(T + \mu_0 A)^{-1}$.

Write

$$P_{\eta} = \int_{L} (\hat{D}_{\mu}(X, g_{\tau}, \beta_{\tau}))^{-1} d\mu$$

= $\int_{L} (T + \mu A)^{-1} d\mu$
= $\int_{L} (T + \mu_0 A)^{-1} (1 + (\mu - \mu_0)K)^{-1} d\mu$
= $(T + \mu_0 A)^{-1} \int_{L} (1 + (\mu - \mu_0)K)^{-1} d\mu.$

Changing variables to $\zeta = -(\mu - \mu_0)^{-1}$, one finds

$$\int_{L} (1 + (\mu - \mu_0)K)^{-1} d\mu = \int_{L'} \zeta^{-1} (K - \zeta)^{-1} d\zeta$$

for *L*' a small loop around $\xi = -(\eta - \mu_0)^{-1}$.

It follows that $P_{\eta} = 2\pi i (T + \mu_0 A)^{-1} \xi^{-1} \Pi_{\xi}$ by elementary complex analysis. It follows that Π_{ξ} has rank one as well.

Applying the local description of the spectral set around τ from Theorem 1.8 of Chapter VII of [Kat66], the theorem follows.

2.2 The first change of index formula: excision and residue calculus

Suppose $\delta \in \mathbb{R}$ is chosen such that the end-periodic operator $D^+(Z_+, g, \beta)$ is Fredholm on the corresponding weighted Sobolev spaces. We write $\operatorname{ind}_{\delta} D^+(Z_+, g, \beta)$ for its index.

In this section, we derive a formula comparing the indices $\operatorname{ind}_{\delta} D^+(Z_+, g, \beta)$ for different values of $\delta \in \mathbb{R}$. The excision principle for the index allows us to reduce the problem to computing the index of a Dirac operator on the infinite cyclic cover. We then use the Fourier-Laplace transform to derive an explicit formula for this index.

2.2.1 The excision principle

Let A_1, B_1, A_2, B_2 be oriented 4-manifolds with boundary. We have $\partial A_1 = \partial A_2 = Y$ and $\partial B_1 = \partial B_2 = \overline{Y}$ where *Y* is a compact oriented 3-manifold. Suppose there are Fredholm operators

$$D_1 : L_1^2(A_1 \cup_Y B_1) \to L^2(A_1 \cup_Y B_1)$$
$$D_2 : L_1^2(A_2 \cup_Y B_2) \to L^2(A_2 \cup_Y B_2)$$

such that $D_1 = D_2$ on Y. We can also define operators

$$\overline{D}_1 : L_1^2(A_1 \cup_Y B_2) \to L^2(A_1 \cup_Y B_2)$$
$$\overline{D}_2 : L_1^2(A_2 \cup_Y B_1) \to L^2(A_2 \cup_Y B_2)$$

by

$$\overline{D}_1 = \begin{cases} D_1 & \text{on } A_1 \\ D_2 & \text{on } B_2 \end{cases} \text{ and } \overline{D}_2 = \begin{cases} D_2 & \text{on } A_2 \\ D_1 & \text{on } B_1 \end{cases}.$$

If \overline{D}_1 and \overline{D}_2 are Fredholm operators, then the four operators D_1 , D_2 , \overline{D}_1 , \overline{D}_2 have indices related by the *excision principle*.

Theorem 2.2.1. Let D_1 , D_2 , \overline{D}_1 , and \overline{D}_2 be Fredholm operators constructed in the manner above. Then

$$\operatorname{ind}(D_1) + \operatorname{ind}(D_2) = \operatorname{ind}(\overline{D}_1) + \operatorname{ind}(\overline{D}_2).$$

The standard excision principle requires the manifolds to be compact, and can be found in Chapter 7 of [DK90]. However, it was extended by Gromov and Lawson in [GJ83] to the case of non-compact manifolds.

Using the excision principle, we can prove the promised statements about the invariance of the index of $D^+(Z_+, g, \beta)$ and the further invariance of the correction term $w(X, g, \beta)$.

Lemma 2.2.2. The index $\operatorname{ind}_{\delta} D^+(Z_+, g, \beta)$ is independent of the ways in which the metric, spin structure, perturbation form, and function $\delta \cdot \tilde{f}$ are extended to Z.

Proof. Consider two different such extensions. The data of these extensions are two tuples $(g_1, \mathfrak{s}_1, \beta_1, h_1)$ and $(g_2, \mathfrak{s}_2, \beta_2, h_2)$ of metric, spin structure, perturbation form, and extension of \tilde{f} respectively.

Let \overline{Z}_+ be the orientation reversal of Z_+ . Then define $D_1 = D^+(Z_+, g_1, \beta_1) - \rho(dh_1)$ on Z_+ and $D_2 = D^+(\overline{Z}_+, g_2, \beta_2) + \rho(dh_2)$ on \overline{Z}_+ . Note that to apply the excision principle exactly as stated above, we are required to work with operators on *unweighted* Sobolev spaces.

We will apply the excision principle to these two operators. Set A_1 , A_2 , B_1 , B_2 to be Z, \overline{X}_+ , X_+ , and \overline{Z} respectively. The manifold \overline{X}_+ is the orientation reversal of X_+ .

By the commutation formula for the Dirac operator, the following diagram commutes:

$$L^{2}_{1}(Z_{+};S^{+}) \xrightarrow{D_{1}} L^{2}(Z_{+};S^{-})$$

$$\downarrow_{e^{-h}} \qquad \qquad \qquad \downarrow_{e^{-h}}$$

$$L^{2}_{1\ \delta}(Z_{+};S^{+}) \xrightarrow{D^{+}(Z_{+},g_{1},\beta_{1})} L^{2}_{\delta}(Z_{+};S^{-})$$

The vertical maps are isometries, so we conclude the operator D_1 acting on the unweighted Sobolev spaces has index $\operatorname{ind}_{\delta}(D^+(Z_+, g_1, \beta_1))$.

To calculate the index of D_2 , observe

$$D_{2} = D^{+}(\overline{Z}_{+}, g_{2}, \beta_{2}) + \rho(dh_{2})$$
$$= D^{-}(Z_{+}, g_{2}, \beta_{2}) + \rho(dh_{2})$$
$$= (D^{+}(Z_{+}, g_{2}, \beta_{2}) - \rho(dh_{2}))^{*}$$

Therefore, $ind(D_2) = -ind_{\delta}(D^+(Z_+, g_2, \beta_2)).$

The operator \overline{D}_1 is the Dirac operator D^+ plus some zero-order terms on the closed manifold $Z \cup_Y \overline{Z}$. Therefore, it has index zero.

The operator \overline{D}_2 also has index zero. Let ι be the orientation-reversing involution on $\overline{X}_+ \cup_Y X_+$. Reversing the orientation on $\overline{X}_+ \cup_Y X_+$ switches the spinor bundles. Therefore, the pullback of \overline{D}_2 by ι is a

differential operator from negative spinors to positive spinors, which is easy to see is the adjoint of \overline{D}_2 . It follows that

$$\operatorname{ind} \overline{D}_2 = -\operatorname{ind}(\iota^* \overline{D}_2)$$

However, the kernel and cokernel of the operators $\overline{D}_2 \circ \iota$ and \overline{D}_2 are identified via the involution as well, so they both must have index zero.

Lemma 2.2.3. The correction term

 $w(X, g, \beta) = \operatorname{ind}_{\mathbb{C}} D^+(Z_+, g, \beta) - \operatorname{sign}(Z)/8$

is independent of the choice of manifold Z bounding Y.

Proof. Let Z_1 , Z_2 be two choices of manifolds bounding Y.

We can apply the excision principle again to the Dirac operator for A_1 , B_1 , A_2 , and B_2 equal to Z_1 , X_+ , \overline{Z}_2 , \overline{X}_+ , respectively.

This indicates that the difference $\operatorname{ind}_{\mathbb{C}} D^+((Z_1)_+, g, \beta) - \operatorname{ind}_{\mathbb{C}} D^+((Z_2)^+, g, \beta)$ is equal to the index of the Dirac operator on $Z_1 \cup_Y \overline{Z}_2$.

By the application of the Atiyah-Singer index theorem to the Dirac operator, this index is equal to $p_1/24$, where p_1 is the Pontryagin number of $Z_1 \cup_Y \overline{Z}_2$.

Applying the Hirzebruch signature theorem, we have the identity

$$\operatorname{sign}(Z_1) - \operatorname{sign}(Z_2) = \operatorname{sign}(Z_1 \cup_Y \overline{Z}_2) = p_1/3.$$

Rearranging, we conclude

$$\operatorname{ind}_{\mathbb{C}} D^+((Z_1)_+, g, \beta) - \operatorname{sign}(Z_1)/8 = \operatorname{ind}_{\mathbb{C}} D^+((Z_2)_+, g, \beta) - \operatorname{sign}(Z_2)/8$$

as desired.

2.2.2 Reduction to the infinite cyclic cover

Pick $\delta_1, \delta_2 \in \mathbb{R}$ such that the operator $D^+(Z_+, g, \beta)$ is Fredholm when acting on the respective weighted Sobolev spaces.

It will be useful to us to compute the difference in indices

$$\operatorname{ind}_{\delta_1} D^+(Z_+, g, \beta) - \operatorname{ind}_{\delta_2} D^+(Z_+, g, \beta).$$

For the Dirac operator $D^+(\widetilde{X}, g, \beta)$ on the infinite cyclic cover \widetilde{X} , let $\operatorname{ind}_{\delta_1, \delta_2} D^+(\widetilde{X}, g, \beta)$ denote its index as an operator

$$D^+(\widetilde{X},g,\beta): L^2_{1,\delta_1,\delta_2}(\widetilde{X};S^+) \to L^2_{\delta_1,\delta_2}(\widetilde{X};S^-).$$

This theorem is the first step in our explicit computation of the change of index.

Theorem 2.2.4. For any $\delta_1, \delta_2 \in \mathbb{R}$ such that the end-periodic Dirac operator is Fredholm on the respective weighted Sobolev spaces,

$$\operatorname{ind}_{\delta_2} D^+(Z_+, g, \beta) - \operatorname{ind}_{\delta_1} D^+(Z_+, g, \beta) = \operatorname{ind}_{\delta_1, \delta_2} D^+(X, g, \beta).$$

Proof. Let $X_{-} = \widetilde{X} \setminus X_{+}$ be the other end of the cover, equipped with the induced orientation from \widetilde{X} .

Apply the excision principle with A_1 , B_1 , A_2 , B_2 equal to Z, X_+ , \overline{Z} , X_- . Recall the function $\delta \cdot \tilde{f}$ on X_+ that defines the weighted Sobolev space with weights δ_1 , δ_2 . It is clear that any other such function h on X_+ agreeing with $\delta \cdot \tilde{f}$ outside of $W_0 \subset X_+$ defines the same weighted Sobolev space.

Now let $h_1 : Z_- \to \mathbb{R}$ be the extension of $\delta_1 \cdot \tilde{f}$ to Z_- , and $h_2 : Z_+ \to \mathbb{R}$ be the extension of $\delta_2 \cdot \tilde{f}$ to Z_+ . Set $D_1 = D^+(Z_+, g, \beta) - dh_2$ and $D_2 = D^+(Z_-, g, \beta) - dh_1$.

The Dirac operator on the manifold $Z \cup_Y \overline{Z}$ has index zero. If we write $Z_- = \overline{Z} \cup_Y X_-$, the excision principle then yields the identity

$$\operatorname{ind}_{\delta_2} D^+(Z_+, g, \beta) + \operatorname{ind}(D^+(Z_-, g, \beta) - dh_1) = \operatorname{ind}_{\delta_1, \delta_2} D^+(\tilde{X}, g, \beta).$$

It remains to show the identity

$$\operatorname{ind}_{\delta_1} D^+(Z_-, g, \beta) = -\operatorname{ind}_{\delta_1} D^+(Z_+, g, \beta),$$

To see this, apply the excision principle again with A_1 , B_1 , A_2 , B_2 equal to Z, X_+ , \overline{Z} , X_- . Pick an extension $h'_1 : Z_+ \to \mathbb{R}$ of $\delta_1 \cdot \tilde{f}$ to Z_+ .

This time, use the operators $D_1 = D^+(Z_+, g, \beta) - dh'_1$, $D_2 = D^+(Z_-, g, \beta) - dh_1$ for the excision principle. Then, one finds

$$\operatorname{ind}_{\delta_1} D^+(Z_+, g, \beta) + \operatorname{ind}_{\delta_1} D^+(Z_-, g, \beta) = \operatorname{ind}_{\delta_1} D^+(X, g, \beta)$$

The right-hand side is zero, as will be implied by the index formula proved independently in the next subsection. \Box

2.2.3 Residue calculus

Now that we have reduced the problem to calculating $\operatorname{ind}_{\delta_1,\delta_2} D^+(\widetilde{X}, g, \beta)$, it remains to explicitly calculate the kernel and the cokernel of this operator.

First, pick some element φ in the kernel of this operator. Let $\xi : \widetilde{X} \to \mathbb{R}$ be a smooth function supported on X_+ that is equal to 1 on the subset

$$W_1 \cup_Y W_2 \cup_Y \cdots \subset X_+.$$

Then write $u = \xi \cdot \varphi$, $v = (1 - \xi) \cdot \varphi$.

We calculate $D^+(\widetilde{X},g,\beta)u = \rho(d\xi) \cdot \varphi$, $D^+(\widetilde{X},g,\beta)v = -\rho(d\xi) \cdot \varphi$.

Define $w = \rho(d\xi) \cdot \varphi$. Applying the Fourier transform to the two equations above, we find

$$D^+_\mu(X,g,\beta)\hat{u}_\mu = \hat{w}_\mu$$

and

$$\hat{D}^+_\mu(X,g,\beta)\hat{v}_\mu = -\hat{w}_\mu.$$

By Lemma 2.1.4, the Fourier transforms \hat{u}_{μ} and \hat{v}_{μ} are only holomorphic in μ for $\text{Re}(\mu) \leq \delta_1$ and $\text{Re}(\mu) \geq \delta_2$ respectively. If, say, $\delta_2 \geq \delta_1$, then we cannot immediately recover φ using the inverse Fourier-Laplace transform.

However, we can analytically continue \hat{u}_{μ} and \hat{v}_{μ} to meromorphic functions on the entire plane.

Let R_{μ} be the inverse of $\hat{D}^{+}_{\mu}(X, g, \beta)$. By Lemma 2.1.16, R_{μ} is meromorphic in μ . Furthermore, the spinor w is supported in W_0 , so its Fourier-Laplace transform \hat{w}_{μ} is holomorphic for all $\mu \in \mathbb{C}$. Write

$$\hat{u}_{\mu} = R_{\mu}\hat{w}_{\mu}$$

and

$$\hat{v}_{\mu} = -R_{\mu}\hat{w}_{\mu}.$$

The right-hand side is a meromorphic function on the entire plane, which yields the desired extension of \hat{u}_{μ} and \hat{v}_{μ} .

Now we can apply the inverse Fourier-Laplace transform to recover φ .

Since the Dirac operator is Fredholm, Lemma 2.1.6 shows that the inverse R_{μ} does not have any poles on the lines {Re(μ) = δ_1 } or {Re(μ) = δ_2 }. Pick $\nu_1 \in \mathbb{C}$ with Re(ν_1) = δ_1 , and $\nu_2 = \nu_1 + \delta_2 - \delta_1$ such that Re(ν_2) = δ_2 .

Recall the definition of the intervals $I(\nu_j)$: They are equal to $\{\nu_j + 2\pi i\alpha \mid \alpha \in [0,1]\} \subset \mathbb{C}$ for j = 1, 2. We apply the inverse Fourier-Laplace transform to get the following formula for φ :

$$\varphi = \frac{1}{2\pi i} \int_{I(\nu_1)} e^{-\mu \tilde{f}} R_\mu \hat{w}_\mu \, d\mu - \frac{1}{2\pi i} \int_{I(\nu_2)} e^{-\mu \tilde{f}} R_\mu \hat{w}_\mu$$

This formula is reminiscent of a contour integral, which can be calculated using the residue theorem. For any $\nu \in \mathbb{C}$, define the horizontal interval $H(\nu) = \{\nu + (\delta_2 - \delta_1)\alpha \mid \alpha \in [0, 1]\} \subset \mathbb{C}$.

The four intervals $I(\nu_1)$, $I(\nu_2)$, $H(\nu_1)$, $H(\nu_1 + 2\pi i)$ form a closed rectangle that will be labeled as *C*. The two terms in the formula above are represented in the clockwise contour integral around *C*:

$$\int_{C} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu = \int_{I(\nu_{1})} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu - \int_{I(\nu_{2})} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu - \int_{H(\nu_{1})} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu - \int_{I(\nu_{2})} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu - \int_{H(\nu_{1})} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu - \int_{H(\nu_{1})} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu - \int_{H(\nu_{1})} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\nu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\mu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\mu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\mu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\mu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\mu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\mu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\mu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\mu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\mu_{1}+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu + \int_{H(\mu_{1}+$$

Note by definition that the Fourier-Laplace transform satisfies $\hat{w}_{\mu+2\pi i} = \hat{w}_{\mu}$, which further implies $\hat{D}^+_{\mu+2\pi i}(X,g,\beta) = \hat{D}^+_{\mu}(X,g,\beta)$ and $R_{\mu+2\pi i} = R_{\mu}$. Then, it is immediate by a change of variables that

$$\int_{H(\nu_1+2\pi i)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu = \int_{H(\nu_1)} e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} \, d\mu.$$

This shows that

$$\varphi = \frac{1}{2\pi i} \int_C e^{-\mu \tilde{f}} R_\mu \hat{w}_\mu \, d\mu.$$

Let Sing_C be the set of poles of the function $e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu}$ in the interior of the rectangle C. By the residue theorem, it follows that

$$\varphi = \sum_{\mu \in \operatorname{Sing}_C} \operatorname{Res}(\mu).$$

Before calculating the residues, one may recall that the cokernel of $D^+(\widetilde{X}, g, \beta)$ still needs to be examined.

This task is avoided entirely by the following trick.

Recall the diagram

$$L^{2}_{1}(\widetilde{X}; S^{+}) \xrightarrow{D^{+}-\rho(dh)} L^{2}(\widetilde{X}; S^{-})$$

$$\downarrow^{e^{-h}} \qquad \qquad \downarrow^{e^{-h}}$$

$$L^{2}_{1,\delta_{1},\delta_{2}}(\widetilde{X}; S^{+}) \xrightarrow{D^{+}} L^{2}_{\delta_{1},\delta_{2}}(\widetilde{X}; S^{-})$$

commutes. The vertical arrows are isometries, so this identifies the cokernel of D^+ with the cokernel of $D^+ - \rho(dh)$ acting on L^2 .

The operator $D^+ - \rho(dh)$ has formal L^2 adjoint $D^- + \rho(dh)$, and its cokernel is identified with the kernel of this operator. The diagram

$$L_1^2(\widetilde{X}; S^-) \xrightarrow{D^- + \rho(dh)} L^2(\widetilde{X}; S^+)$$
$$\downarrow_{e^h} \qquad \qquad \downarrow_{e^h}$$
$$L_{1,-\delta_1,-\delta_2}^2(\widetilde{X}; S^-) \xrightarrow{D^-} L_{-\delta_1,-\delta_2}^2(\widetilde{X}; S^+)$$

Therefore, the kernel of $D^- + dh$ is identified with the kernel of the negative Dirac operator acting on the weighted space $L^2_{1,-\delta_1,-\delta_2}$.

We can use an identical method to the one detailed for the kernel of D^+ to compute the kernel of D^- , and therefore the cokernel of D^+ .

Our residue formula simplifies this even further. From Lemma 2.1.4, if $\delta_2 \leq \delta_1$ then the functions \hat{u}_{μ} and \hat{v}_{μ} will be holomorphic inside the rectangle *C*, so $\varphi = 0$ and the kernel of D^+ vanishes.

On the other hand, if $\delta_1 \leq \delta_2$, then we find the cokernel vanishes due to the above trick. Since the only objective is to calculate the difference in indices

$$\operatorname{ind}_{\delta_1} D^+ - \operatorname{ind}_{\delta_2} D^+,$$

it suffices to assume that $\delta_1 \leq \delta_2$.

Let η be a pole of $R_{\mu}\hat{w}_{\mu}$. Consider the Laurent expansion at η :

$$R_{\mu}\hat{w}_{\mu} = \sum_{k=-d}^{\infty} \psi_k (\mu - \eta)^k.$$

Then we can apply the Fourier-Laplace transform $\hat{D}^+_{\mu}(X, g, \beta) = D^+(X, g, \beta) - \mu \cdot \rho(f^*(d\theta))$ to this equation to get

$$\begin{split} \hat{w}_{\mu} &= \sum_{k=-d}^{\infty} \hat{D}_{\mu}^{+}(\tilde{X}, g, \beta) \psi_{k}(\mu - \eta)^{k} \\ &= \sum_{k=-d}^{\infty} (D^{+}(X, g, \beta) - \mu \rho(f^{*}(d\theta))) \psi_{k}(\mu - \eta)^{k} \\ &= \sum_{k=-d}^{\infty} (D^{+}(X, g, \beta) - \eta \rho(f^{*}(d\theta))) \psi_{k}(\mu - \eta)^{k} - \rho(f^{*}(d\theta)) \psi_{k}(\mu - \eta)^{k+1} \\ &= \sum_{k=-d}^{\infty} (\hat{D}_{\eta}^{+}(X, g, \beta) \psi_{k} - \rho(f^{*}(d\theta)) \psi_{k-1})(\mu - \eta)^{k} \end{split}$$

However, \hat{w}_{μ} is an entire function, so all of its negative-order Laurent series coefficients vanish. This shows that the spinors ψ_k satisfy

$$\hat{D}_{\eta}^{+}(X,g,\beta)\psi_{k} = \rho(f^{*}(d\theta))\psi_{k-1}$$

for all $-d + 1 \leq k \leq -1$ and

$$\hat{D}_{\eta}^{+}(X,g,\beta)\psi_{-d} = 0.$$

Denote this system of equations by system (S1).

The residue of $e^{-\mu \hat{f}} R_{\mu} \hat{w}_{\mu}$ at η can be calculated explicitly in terms of the coefficients $\{\psi_k\}$. The Laurent expansion at η is

$$e^{-\mu \tilde{f}} R_{\mu} \hat{w}_{\mu} = \left(\sum_{j=0}^{\infty} \frac{1}{j!} (-\tilde{f})^{j} e^{-\eta \tilde{f}} (\mu - \eta)^{j}\right) \left(\sum_{k=-d}^{\infty} \psi_{k} (\mu - \eta)^{k}\right)$$
$$= \sum_{p=-d}^{\infty} \left(\sum_{j+k=p} \frac{\psi_{k}}{j!} (-\tilde{f})^{j}\right) e^{-\eta \tilde{f}} (\mu - \eta)^{p}.$$

The residue is the p = -1 term, given by

$$\operatorname{Res}(\eta) = e^{-\eta \widetilde{f}} \sum_{j=0}^{d-1} \frac{\psi_{-j-1}}{j!} (-\widetilde{f})^j.$$

Let $d(\eta)$ denote the dimension of the vector space of tuples of sections $\{\psi_k\}_{k=-d}^{-1}$ satisfying system (S1).

This formula indicates that $\hat{D}_{\eta}^+(X, g, \beta)$ is invertible if and only if $d(\eta) = 0$. By the Mittag-Leffler theorem, we also find that the kernel of $D^+(\tilde{X}, g, \beta)$ is determined by the negative Laurent coefficients at each pole of $R_{\mu}\hat{w}_{\mu}$. The dimension of the kernel will therefore be the sum of $d(\eta)$ across all of the poles. More generally, we can sum up across the entire rectangle *C* to get

$$\operatorname{ind}_{\delta_1} D^+(Z_+, g, \beta) - \operatorname{ind}_{\delta_2} D^+(Z_+, g, \beta) = \sum_{\eta \in C} d(\eta).$$

2.3 The second change of index formula: spectral flow

The formula given at the end of the previous section is rather explicit, but still not directly applicable to the ultimate purpose of proving Theorem 2.1.12. In particular, the difference

ind
$$D^+(Z_+, g_0, \beta_0) - \operatorname{ind} D^+(Z_+, g_1, \beta_1)$$

remains a mystery.

A re-organization of the previously introduced notation, along with the application of the previously demonstrated theoretical results regarding the Fourier-Laplace transform, will be used to build onto the first change of index formula and show that this change of index is equal to a type of *spectral flow*. This idea is most famously encountered in the work of Atiyah-Patodi-Singer [APS75] on index theorems for manifolds with boundary.

As a simple model of spectral flow, let $\{D_t\}_{t \in [0,1]}$ be a family of self-adjoint operators on \mathbb{C}^n . Each operator D_t has finite spectrum $\Sigma(D_t) \subset \mathbb{C}$, and with some work and some small endpoint-preserving perturbation of the family $\{D_t\}$ one can find that the union $\cup_{t \in [0,1]} \Sigma(D_t)$ is a union of simple embedded

curves in the complex plane that intersect the positive imaginary axis transversely. The *spectral flow* of this family is the oriented intersection number of these curves with the imaginary axis.

An intersection point is counted positively if the respective curve is traveling from the left half plane to the right half, e.g. the eigenvalue is going from "negative to positive". It is counted negatively if the curve is traveling in the other direction.



Figure 2.1: Depiction of spectral flow for two eigenvalues. The top curve goes from right to left and so contributes negatively to the spectral flow, while the bottom curve contributes positively.

The terminology is used in a slightly different manner in our scenario.

First, we will change the notation from the last section a bit. Recall the formula for the Fourier-Laplace transform of a spinor:

$$\hat{\varphi}_{\mu} = e^{\mu \tilde{f}} \sum_{n \in \mathbb{Z}} e^{\mu n} (\varphi \circ T^n).$$

This formula relies not on the value of μ but rather on the value of $z = e^{\mu}$. Thus, an equivalent definition of the Fourier-Laplace transform uses a parameter $z \in \mathbb{C} - \{0\}$, written as

$$\hat{\varphi}_z = z^{\widetilde{f}} \sum_{n \in \mathbb{Z}} z^n (\varphi \circ T^n)$$

This choice of notation was not made from the beginning because it renders the Fourier-Laplace transform of the Dirac operator less transparent. It is written as

$$\hat{D}_z^+(X,g,\beta) = D^+(X,g,\beta) - \log z \cdot \rho(f^*(d\theta)).$$

This expression is not a priori well-defined as the complex logarithm is multi-valued. However, any value of $\log z$ is the same modulo addition of an integer multiple of $2\pi i$. It has already been established that the operators $\hat{D}^+_{\mu+2\pi i}(X,g,\beta)$ and $\hat{D}^+_{\mu}(X,g,\beta)$ are isomorphic for any $\mu \in \mathbb{C}$, so the operator $\hat{D}^+_z(X,g,\beta)$ is well-defined.

This change of variables, however, slightly simplifies the residue calculus carried out in the previous section. In order to calculate the index of the Dirac operator $D^+(\tilde{X}, g, \beta)$, we were required to choose a rectangle *C* with horizontal edges arbitrarily chosen to avoid any poles of the inverse map R_{μ} .

Applying the exponential map, we are freed from this arbitrary consideration. The strip $\{\delta_1 < \operatorname{Re}(\mu) < \delta_2\}$ is mapped onto the annulus $\{e^{\delta_1} < |z| < e^{\delta_2}\}$, to which we can apply the residue theorem. It is also the case that $d(\eta + 2\pi i) = d(\eta)$ for any $\eta \in \mathbb{C}$, so $d(\eta)$ only depends on the value of e^{η} as well. Therefore, we can write d(z) for the quantity $d(\log z)$.

The change of index formula then translates to

$$\operatorname{ind}_{\delta_1} D^+(Z_+, g, \beta) - \operatorname{ind}_{\delta_2} D^+(Z_+, g, \beta) = \sum_{e^{\delta_1} < |z| < e^{\delta_2}} d(z).$$

Now we define the spectral flow in our situation. Let (g_0, β_0) and (g_1, β_1) be regular pairs and (g_t, β_t) a special path between them as usual.

Define the cylinder $C \subset (\mathbb{C} - \{0\}) \times [0,1]$ to be the set of all pairs (z,t) with |z| = 1. Within the space $(\mathbb{C} - \{0\}) \times [0,1]$, we can also define the space S to be the set of tuples (z,t) such that the Dirac operator $\hat{D}_z^+(X, g_t, \beta_t)$ is not invertible. The space S admits the following description.

Theorem 2.3.1. In a small neighborhood of C, the subset S is a disjoint union of finitely many smooth embedded curves that intersect C transversely.

In addition, write $d_t(z)$ for the quantity d(z) with respect to the metric and perturbation at time t. Then, at any $p = (z, \tau) \in S \cap C$, the quantity $d_t(w)$ at most 1 for any pair (w, t) sufficiently close to p.

Assume for now that this theorem holds. It will be proven later.

It follows from Theorem 2.3.1 that $S \cap C$ is a finite, discrete set of points. Then, the *spectral flow* $SF(g_t, \beta_t)$ for the path (g_t, β_t) is defined to be an oriented count of the intersection $S \cap C$. For any such $p \in S \cap C$, there is a small neighborhood $p \in U \subset (\mathbb{C} - \{0\}) \times [0, 1]$ such that $S \cap U$ is a smooth embedded curve γ . Then, p is counted with sign +1 if γ is *exiting* the cylinder C and -1 if it is *entering* the cylinder.



Figure 2.2: Depiction of spectral flow through the cylinder *C*. The first two curves are entering and are counted with sign -1, while the final curve is counted with sign 1.

Pick some point $p = (z, \tau) \in S \cap C$. First suppose that S does not intersect C at time τ except at p. For any $\delta > 0$, we can pick $\varepsilon > 0$ sufficiently small such that the following two statements hold:

- 1. Write $a = \tau \varepsilon$, $b = \tau + \varepsilon$. The intersection of S with the slice $\{(z, t) \in (\mathbb{C} \{0\}) \times [0, 1] | t \in [a, b]\}$ is an embedded curve γ intersecting C transversely at p.
- 2. The embedded curve γ (and therefore S) does not intersect the cylinder slice $\{(z,t) \in (\mathbb{C} \{0\}) \times [0,1] | |z| = e^{\delta}, t \in [a,b]\}$.

By Theorem 2.1.8, it follows that the operator

$$D^+(Z_+, g_t, \beta_t) : L^2_{1,\delta}(Z_+; S^+) \to L^2_{\delta}(Z_+; S^-)$$

is Fredholm for all $t \in [a, b]$.

They are isomorphic to a continuous path of Fredholm operators in the standard L^2 space, so it follows that

$$\operatorname{ind}_{\delta} D^+(Z_+, g_a, \beta_a) = \operatorname{ind}_{\delta} D^+(Z_+, g_b, \beta_b).$$

. Then we have the following two change of index formulae:

$$\operatorname{ind} D^{+}(Z_{+}, g_{a}, \beta_{a}) - \operatorname{ind}_{\delta} D^{+}(Z_{+}, g_{a}, \beta_{a}) = \sum_{1 < |z| < e^{\delta}} d_{a}(z),$$
$$\operatorname{ind} D^{+}(Z_{+}, g_{b}, \beta_{b}) - \operatorname{ind}_{\delta} D^{+}(Z_{+}, g_{b}, \beta_{b}) = \sum_{1 < |z| < e^{\delta}} d_{b}(z).$$

It follows that

$$\operatorname{ind} D^+(Z_+, g_b, \beta_b) - \operatorname{ind} D^+(Z_+, g_a, \beta_a) = \sum_{1 < |z| < e^{\delta}} d_b(z) - \sum_{1 < |z| < e^{\delta}} d_a(z)$$

By Theorem 2.3.1, any point (z,t) on the embedded curve γ satisfies $d_t(z) = 1$. Furthermore, by definition any pair (z,t) for $1 < |z| < e^{\delta}$ and $t \in [a,b]$ has $d_t(z) > 0$ (and therefore $d_t(z) = 1$) if and only if $(z,t) \in \gamma$.

Then, since γ is an embedded smooth curve, by making δ and ε sufficiently small it can be fixed such that the projection of γ onto the time coordinate is injective.

It follows that, for any $t \in [a, b]$, the sum $\sum_{1 < |z| < e^{\delta}} d_t(z)$ is equal to 1 if γ is in the region $\{1 < |z| < e^{\delta}\}$ at time a, and 0 otherwise.

If γ is *entering* the cylinder, then it is in this region at time *a* and outside of this region (i.e. inside the cylinder) at time *b*. Therefore, in this case,

ind
$$D^+(Z_+, g_b, \beta_b)$$
 - ind $D^+(Z_+, g_a, \beta_a) = -1$.

On the other hand, if γ is *exiting* the cylinder, then it follows that

ind
$$D^+(Z_+, g_b, \beta_b) - \text{ind } D^+(Z_+, g_a, \beta_a) = 1.$$

If there are multiple points p_1, \ldots, p_k in the intersection of S and C at time τ , then since they are isolated the proof proceeds in an identical manner. The only difference is that each point p_i will contribute ± 1 to the change of index from time a to b depending on the direction of the embedded curve around it.

If $p = (z, \tau)$, $q = (z', \tau')$ are two points in $S \cap C$ such that $\tau' > \tau$ and S does not intersect C at any time in the interval (τ, τ') , then the operators $D^+(Z_+, g_t, \beta_t)$ for all $t \in (\tau, \tau')$ form a continuous path of Fredholm operators. It follows that

ind
$$D^+(Z_+, g_b, \beta_b) = \operatorname{ind} D^+(Z_+, g_a, \beta_a)$$

for any $a, b \in (\tau, \tau')$.

Adding up all of the contributions to the change of index on the interval [0,1], we have proven the following change of index formula.

Theorem 2.3.2. For any two regular pairs (g_0, β_0) and (g_1, β) and a special path (g_t, β_t) between them, the indices
of the respective Dirac operators differ by the spectral flow:

ind
$$D^+(Z_+, g_1, \beta_1)$$
 - ind $D^+(Z_+, g_0, \beta_0) = SF(g_t, \beta_t)$.

It remains to prove Theorem 2.3.1.

Recall from the proof of Theorem 2.1.8 that, for $z \in \mathbb{C} - \{0\}$ with |z| = 1, elements of the kernel of $\hat{D}_z^+(X, g, \beta)$ correspond to the reducible solutions for the blown-up Seiberg-Witten equations for the pair (g, β) . This yields the following additional proposition:

Proposition 2.3.3. The points of $S \cap C$ are in bijection with the space \mathcal{M}^0 of reducible solutions from Theorem 1.3.3.

This correspondence will be useful in proving the theorem. In particular, understanding the local structure of the parameterized moduli space $P\mathcal{M}(X, g_t, \beta_t)$ near the reducible boundary points will be crucial, as it will be used to shed light on the corresponding local structure of the spectral set S near the cylinder C.

Recall parameterized moduli space that was briefly mentioned in the first chapter. The space $PZ \subset \mathcal{B}^{\sigma}(X) \times [0,1]$ was defined as the subset of tuples (A, s, φ, t) such that $D_A^+\varphi = 0$, where the Dirac operator of course depends on the metric g_t .

This space PZ is a Hilbert manifold with boundary ∂Z equal to the intersection of PZ with the locus $\{s = 0\}$. Then, $P\mathcal{M}(X, g_t, \beta_t)$ is equal to $P\chi^{-1}(0)$, where $P\chi$ is the map from PZ into the space of imaginary-valued self-dual two-forms defined by

$$(A, s, \varphi, t) \mapsto F_A^+ - s^2 \rho(\varphi \varphi^*)_0 - d^+ \beta_t.$$

The space of reducible solutions $\mathcal{M}^0 \subseteq P\mathcal{M}(X, g_t, \beta_t)$ is in turn equal to $\partial \chi^{-1}(0)$, where $\partial \chi$ is the restriction of $P\chi$ to $\partial \mathcal{Z}$.

Since the path (g_t, β_t) has regular endpoints, the space $P\mathcal{M}(X, g_t, \beta_t)$ is a manifold (see Theorem 1.3.3). This implies that the operator $P\chi$ is transverse to the zero section at any reducible point $(A, 0, \varphi, \tau) \in \mathcal{M}^0$. This in turn implies that its linearization \mathcal{D} at any such point is surjective with kernel equal to the tangent space of \mathcal{M}^0 at that point. However, the space \mathcal{M}^0 is zero-dimensional, so \mathcal{D} has trivial kernel.

We have arrived at the primary reason for choosing a special path (g_t, β_t) . The fact that the path (g_t, β_t) has a constant metric term near any reducible solution gives the linearization of $\partial \mathcal{Z}$ a particularly simple form, since it does not depend on any derivatives of the path of metrics.

The tangent space of ∂Z at $(A, 0, \varphi, \tau)$ is a subset of the space of tuples of the form (b, ψ, u) , where *b* is an imaginary-valued one-form, ψ is a unit-length positive spinor, and *u* is a real number. This tuple is required to satisfy the equations

$$\begin{split} \rho(b)\varphi + D_A^+\psi &= 0\\ \langle \varphi, \psi \rangle_{L^2} &= 0\\ d^*b &= 0. \end{split}$$

The L^2 inner product taken in the second equation is with respect to the metric g_{τ} .

The linearization $\ensuremath{\mathcal{D}}$ is the map sending

$$(b, \psi, u) \mapsto d^+b - ud^+]\dot{\beta}$$

where $\dot{\beta}$ is a shorthand for the time derivative of the path β_t at $t = \tau$.

For algebraic convenience, change variables and set $a = b - u\dot{\beta}$. Combined with the other equations defining the tangent space, it follows that a tuple (a, ψ, u) is in the kernel of the linearized operator \mathcal{D} if and only if

$$\begin{aligned} d^*a &= 0\\ d^+a &= 0\\ \langle \varphi, \psi \rangle_{L^2} &= 0\\ \rho(a)\varphi + u \cdot \rho(\dot{\beta})\varphi + D^+_A \psi &= 0. \end{aligned}$$

Since the kernel is empty, this system of equations only has the solution (0, 0, 0). Denote this system of equations by (S2).

It will also now be required that, for each metric g_t , the defining map $f_t : X \to S^1$ is *harmonic* for the metric g_t . A harmonic map $f : M \to N$ between Riemannian manifolds M and N with respective metrics g and h is a smooth map that is a critical point for the *energy functional*:

$$E(f) = \int_M ||df||^2 \mathrm{dvol}_M.$$

The general theorem for existence and uniqueness of harmonic maps is a theorem of Eells and Lemaire ([EL78]), although the theorem below for $N = S^1$ is a consequence of basic Hodge theory.

Theorem 2.3.4. For any metric g on a manifold X satisfying assumption (A1), there exists a unique harmonic map $f: X \to S^1$ up to translation by the action of S^1 such that the pullback $f^*(d\theta)$ represents a generator of $H^1(X; \mathbb{Z})$.

All of the theory above will be used now to prove a couple of technical lemmas, which will then finish off the proof of Theorem 2.3.1.

Lemma 2.3.5. For any (g_{τ}, β_{τ}) in the path, when |z| = 1, the quantity $N = \dim_{\mathbb{C}}(\ker \hat{D}_z^+(X, g_{\tau}, \beta_{\tau}))$ is at most one.

Proof. Recall the proof of Theorem 2.1.8. Pick $z \in \mathbb{C} - \{0\}$ such that the operator $\hat{D}_z^+(X, g_\tau, \beta_\tau)$ is not invertible.

Let A_0 be the base spin connection. For any z such that |z| = 1 and $\varphi \in \ker \hat{D}_z^+(X, g_\tau, \beta_\tau)$ of unit L^2 norm, the tuple

$$(A_z = A_0 + \beta_\tau - \log z \cdot f_\tau^*(d\theta), 0, \varphi)$$

is a reducible solution to the blown-up Seiberg-Witten equations for the metric (g_{τ}, β_{τ}) .

Therefore, before quotienting out by gauge equivalence, the reducible solutions for a fixed z form a sphere of dimension 2N - 1. After quotienting out by gauge equivalence, this becomes a manifold of dimension 2N - 2.

By Theorem 1.3.3, the moduli space of reducible solutions has dimension zero. It follows that N must be equal to one in this case, and at most one overall.

The following lemma extends this statement for operators where |z| is very close to 1.

Lemma 2.3.6. Fix $z \in \mathbb{C}$ such that |z| = 1 and $\tau \in [0, 1]$ such that $\hat{D}_z^+(X, g_\tau, \beta_\tau)$ has kernel of complex dimension one. Then there exists a neighborhood U of (z, τ) in $\mathbb{C} - \{0\} \times [0, 1]$ such that the operator $\hat{D}_{z'}^+(X, g_{\tau'}, \beta_{\tau'})$ has kernel of complex dimension at most one for any $(z', \tau') \in U$.

Proof. Since $\hat{D}_z^+(X, g_\tau, \beta_\tau)$ has kernel of dimension one, by the proof of Lemma 2.3.5 this implies that there is a reducible solution to the Seiberg-Witten equations for the pair (g_τ, β_τ) .

Since the path is a special path, one can then fix $\gamma > 0$ such that $g_{\tau'} = g_{\tau}$ for all $\tau' \in (\tau - \gamma, \tau + \gamma)$. Restrict τ' to this interval.

Pick some $z' \in \mathbb{C} - \{0\}$. Fix $\mu \in \mathbb{C}$ such that $e^{\mu} = z$, and $\mu' \in \mathbb{C}$ such that $e^{\mu'} = z'$.

Then the operator $\hat{D}_{z'}^+(X, g_{\tau'}, \beta_{\tau'})$ is equal to $\hat{D}_z^+(X, g_{\tau}, \beta_{\tau}) + (\mu' - \mu)\rho(f^*(d\theta))$. By choosing z' in an arbitrarily small neighborhood of z, the quantity $|\mu - \mu'|$ can be made arbitrarily small. It follows that $\hat{D}_{z'}^+(X, g_{\tau'}, \beta_{\tau'})$ is a small perturbation of $\hat{D}_z^+(X, g_{\tau}, \beta_{\tau})$.

By Theorem 5.17 of Chapter IV of [Kat66] (the standard stability result that has been mentioned a couple of times), this implies that the kernel of $\hat{D}_{z'}^+(X, g_{\tau'}, \beta_{\tau'})$ has dimension less than or equal to that of $\hat{D}_z^+(X, g_{\tau}, \beta_{\tau})$ as desired.

Recall our definition of the operator P_{μ} from 2.1.18. This operator also satisfies $P_{\mu+2\pi i} \simeq P_{\mu}$, so it is well-defined to write P_z for the operator $P_{\log z}$ for any choice of $\log z$.

Lemma 2.3.7. For any (g_{τ}, β_{τ}) in the path and $z \in \mathbb{C} - \{0\}$ such that |z| = 1 and $\hat{D}_{z}^{+}(X, g_{\tau}, \beta_{\tau})$ is not invertible, $d(z) = \operatorname{rank} P_{z} = 1$.

Proof. By the previous lemma, the kernel of $\hat{D}_z^+(X, g_\tau, \beta_\tau)$ has dimension one.

It follows by definition that $d(z) \ge 1$. Suppose d(z) > 1.

Recall that d(z) is the dimension of the vector space of solutions to the system (S1):

$$\hat{D}_z^+(X, g_\tau, \beta_\tau)(\psi_k) = \rho(f_\tau^*(d\theta))(\psi_{k-1})$$

for k from -d + 1 to -1 and

$$\hat{D}_z^+(X, g_\tau, \beta_\tau)(\psi_{-d}) = 0.$$

It follows that the sections ψ_k for k from -d + 2 to -1 are nonzero only if ψ_{-d+1} is nonzero. Therefore, given that d(z) > 1, there must exist a solution to system (S1) with both ψ_{-d} and ψ_{-d+1} nonzero.

Without loss of generality, normalize ψ_{-d} to unit L^2 length. Replace ψ_{-d+1} with the orthogonal projection $\psi_{-d+1} - \langle \psi_{-d+1}, \psi_{-d} \rangle_{L^2} \psi_{-d}$. Then ψ_{-d+1} is orthogonal to ψ_{-d} and the two spinors still satisfy the last two equations in system (S1), although the rest of the system may not be satisfied anymore. Furthermore, since the function f_{τ} is harmonic, the pullback $f^*(d\theta)$ is a harmonic one-form.

Again let $A_z = A_0 + \beta_\tau + \log z \cdot f^*(d\theta)$. Then the tuple $(A_z, 0, \psi_{-d}, \tau)$ is a reducible solution to the Seiberg-Witten equations as before. It follows that the tuple $(f^*_\tau(d\theta), \psi_{-d+1}, 0)$ is in the kernel of the linearized Dirac operator \mathcal{D} at the point $(A_z, 0, \psi_{-d}, \tau)$ (with the algebraic rearrangement carried out above).

Since the linearized Dirac operator is injective, this implies that ψ_{-d+1} must be zero. However, the initial assumption required it to be nonzero, so we arrive at a contradiction and d(z) = 1. It remains to show P_z has rank 1. Reverting for a moment to the old notation, pick $\eta \in \mathbb{C}$ such that $e^{\eta} = z$. Then $d(\eta) = 1$, and the problem is to show that

$$P_{\eta} = \int_{L} R_{\mu} d\mu$$

has rank 1, where R_{μ} is the inverse of $\hat{D}^+_{\mu}(X, g_t, \beta_t)$.

The inverse R_{μ} is meromorphic on the complex plane, so its Laurent expansion at η can be written as

$$R_{\eta} = \sum_{k=-d}^{\infty} A_k (\mu - \eta)^k$$

where A_k are differential operators.

By the residue theorem, the operator A_{-1} is equal to $\frac{1}{2\pi i}P_{\eta}$. Then, the negative Laurent coefficients of $\hat{D}_{\mu}^{+}(X, g_{\tau}, \beta_{\tau})R_{\mu} = 1$ vanish. The operators A_{k} solve the system

$$D^+(X, g_\tau, \beta_\tau)A_k = \eta \cdot \rho(f_\tau^*(d\theta))A_{k-1}$$

for k = -d + 1 to -1 and

$$D^+(X, g_\tau, \beta_\tau)A_{-d} = 0.$$

Take a set of spinors $\psi_k \in \text{im } A_k$ for every k from -d to -1. Then these spinors solve system (S1). It follows by definition that

$$1 = d(\eta) \ge \sum_{k=-d}^{-1} \operatorname{rank} A_k \ge \operatorname{rank} P_{\eta}.$$

This also shows immediately that all but one of the operators A_k must be equal to zero.

Suppose P_{η} is equal to zero. Then, suppose that A_k has rank one for some $k \neq -1$. For any nonzero spinor ψ in the image of A_k , it follows from the system of equations (S1) that ψ is both in the kernel of the Dirac operator and the operator $\rho(f^*(d\theta))$. Note that in the case k = -1, one would only have ψ being in the kernel of the Dirac operator.

However, $\rho(f^*(d\theta))$ maps the kernel of the Dirac operator injectively to the orthogonal complement of its image. Since ψ is nonzero, we arrive at a contradiction. Therefore, it follows that P_{η} has rank one as desired.

Combining Lemma 2.3.7 and Theorem 2.1.18, it follows that the intersection of S with a sufficiently small neighborhood of a point in $S \cap C$ is an embedded curve.

At any point $p = (z, \tau) \in S \cap C$, choose some such embedded curve around p and project onto the complex coordinate to obtain a smooth embedded curve $\gamma : (\tau - \varepsilon, \tau + \varepsilon) \rightarrow \mathbb{C} - \{0\}$.

Define the smooth family of Dirac operators D^+_{γ} over $(\tau - \varepsilon, \tau + \varepsilon)$ by

$$D_{\gamma}^{+}(t) = \hat{D}_{\gamma(t)}^{+}(X, g_t, \beta_t) = D^{+}(X, g_t, \beta_t) - \gamma(t)\rho(f_t^{*}(d\theta)).$$

By making ε sufficiently small and combining Lemma 2.3.5 and Lemma 2.3.6, the kernel $D_{\gamma}^{+}(t)$ is onedimensional for all $t \in (\tau - \varepsilon, \tau + \varepsilon)$. Write ker D_{γ}^{+} for the subset of $(\tau - \varepsilon, \tau + \varepsilon) \times L_{1}^{2}(X; S^{+})$ equal to the union of the subsets $\{t\} \times \ker D_{\gamma}^{+}(t)$ for all $t \in (\tau - \varepsilon, \tau + \varepsilon)$. This space inherits a topology from its ambient space, and projection onto the time coordinate gives a natural map ker $D_{\gamma}^{+} \to (\tau - \varepsilon, \tau + \varepsilon)$.

The following lemma is well-known given the fact that ker $D_{\gamma}^{+}(t)$ has constant dimension for every t.

Lemma 2.3.8. The space ker D_{γ}^+ is a vector bundle over $(\tau - \varepsilon, \tau + \varepsilon)$.

Now we can show that the spectral curve γ is transverse to the cylinder *C*.

Lemma 2.3.9. The time derivative $\dot{\gamma}$ of $\gamma(t)$ at $t = \tau$ is nonzero.

Proof. The linearization of this family at $t = \tau$, using the fact that the path is special, is equal to the operator $\rho(\dot{\beta} - \dot{\gamma}f^*(d\theta))$.

Suppose for the sake of contradiction that $\dot{\gamma} = 0$. By the previous lemma, choosing a local non-vanishing section of the vector bundle ker D_{γ}^+ around τ and normalizing produces a family of unit-length spinors $\{\varphi_t\}$ parameterized by $t \in (\tau - \varepsilon', \tau + \varepsilon')$ for some small $\varepsilon > \varepsilon' > 0$ such that $D_{\gamma(t)}^+ \varphi_t = 0$.

Differentiating this equation at $t = \tau$ and applying the fact that $\dot{\gamma} = 0$, it follows that

$$\rho(\dot{\beta})\varphi_{\tau} + D_{\gamma}^{+}(\tau)\dot{\varphi} = 0.$$

Since the spinors φ_t have unit length, $\dot{\varphi}$ is L^2 orthogonal to φ_{τ} .

However, then the tuple $(0, \dot{\varphi}, 1)$ satisfies the system of equations for the linearized operator at the reducible solution $(A_0 + \beta_\tau - \log z \cdot f^*(d\theta), 0, \varphi_\tau, \tau) \in \mathcal{M}^0$, so we arrive at a contradiction and $\dot{\gamma}$ cannot be equal to zero.

The lemma immediately implies that the embedded curve around p has transverse intersection with the cylinder.

It remians to show the final statement of Theorem 2.3.1. This is immediate by Lemma 2.3.6.

2.4 Orientations

There is a correspondence between points in the cylinder $S \cap C$ and reducible solutions \mathcal{M}^0 .

The difference $#\mathcal{M}(X, g_0, \beta_0) - #\mathcal{M}(X, g_1, \beta_1)$ is an oriented count $#\mathcal{M}^0$.

Meanwhile, the difference $\operatorname{ind}_{\mathbb{C}} D^+(Z_+, g_0, \beta_0) - \operatorname{ind}_{\mathbb{C}} D^+(Z_+, g_1, \beta_1)$ is an oriented count of points in the cylinder, totaled up to the spectral flow $\operatorname{SF}(g_t, \beta_t)$.

The only remaining obstacle is to show that these oriented counts are the same: A point in \mathcal{M}^0 will be counted with the same sign as its corresponding point in $S \cap C$ is counted in the spectral flow, and vice versa.

To prove this, it is of course necessary to understand exactly how the orientation is constructed on the moduli space \mathcal{M}^0 .

Pick a point $(A, 0, \varphi, \tau) \in \mathcal{M}^0$. Write \mathcal{D}_1 for the linearized operator of $\partial \chi$ at $(A, 0, \varphi, \tau)$, defined by

$$(a, \psi, u) \mapsto (d^+a, d^*a, \rho(a)\varphi + u \cdot \rho(\beta)\varphi + D^+_A\psi)$$

with domain consisting of tuples (a, ψ, u) for a an imaginary-valued one-form, ψ a positive spinor in the L^2 orthogonal complement of φ , and u a real number.

An orientation of the point $(A, 0, \varphi, \tau)$ is equivalent to an orientation of the kernel of \mathcal{D}_1 , but a priori there is no canonical way to orient the kernel of this operator. However, there is a canonical way to orient the kernel of a related operator.

The operator \mathcal{D}_0 is then defined by

$$\mathcal{D}_0: (a, \psi, u) \mapsto (d^+a, d^*a, D^+_A \psi).$$

From the proof of Lemma 2.3.5, it was discovered that, at any reducible solution, the Dirac operator D_A^+ has kernel of dimension one. Therefore, if $D_A^+\psi = 0$, it follows that ψ must be a scalar multiple of φ .

It follows that the kernel of \mathcal{D}_0 are elements of the form (a, 0, u) for a a harmonic, imaginary-valued oneform and u any real number. The top exterior power of ker \mathcal{D}^0 is isomorphic to det $H^1(X; \mathbb{R})$. Therefore, the homology orientation of X determines an orientation of ker \mathcal{D}_0 .

In addition, coker $\mathcal{D}_0 \simeq \operatorname{coker} D_A^+$, which is canonically oriented by the complex structure on the spinor bundle.

For a parameter $s \in [0, 1]$, and define the path of operators \mathcal{D}_s by

$$\mathcal{D}_s: (a, \psi, u) \mapsto (d^+a, d^*a, s \cdot \rho(a)\varphi + us \cdot \rho(\dot{\beta})\varphi + D^+_A\psi).$$

This is a continuous path of Fredholm operators connecting \mathcal{D}_0 to \mathcal{D}_1 .

Suppose for a moment that the kernels of \mathcal{D}_s have the same dimension for all $s \in [0, 1]$. It follows that the union of the kernels form a vector bundle over [0, 1] (see Lemma 2.3.8). Moreover, any vector bundle over [0, 1] is trivial. Therefore, an orientation on the kernel of \mathcal{D}_0 canonically induces an orientation on every other fiber of the vector bundle (it can be checked that this is independent of the choice of trivialization).

However, this situation is not guaranteed. The dimension of the kernel of D_s is not necessarily a continuous function in *s*, i.e. the dimension of the kernel may "jump" as *s* changes.

This situation can be salvaged, however, if the dimension of ker D_s is sufficiently well-behaved. For our purposes, it suffices to consider the situation where the dimension of ker D_s is zero except in a finite, discrete subset of [0, 1].

Outside of these "jumps", the kernels of the operators D_s are locally trivial and orientation can be transported in the manner described above.

Pick $s \in [0, 1]$ such that the kernel of \mathcal{D}_s has *real* dimension one. Write \mathcal{D}_s^* for the L^2 formal adjoint of \mathcal{D}_s , and write Δ_s for the operator $\mathcal{D}_s^*\mathcal{D}_s$. Then, the orientation across the "jump" *s* is transported using the *spectral flow* of the family Δ_s . This idea is based on the original work done for Cauchy-Riemann operators on Riemann surfaces by Quillen in [Qui85]. Note that Δ_s is self-adjoint, and therefore has a real spectrum.

For any $s \in [0, 1]$, the "formal determinant" det $(\Delta_s) \in \mathbb{R}$ can be defined as follows. Let z be a complex number with real part greater than 1. Then, define the zeta function

$$\zeta(z) = \sum_\lambda \lambda^{-z}$$

where the sum is over all nonzero eigenvalues of Δ_s .

This function can be continued to a meromorphic function on the complex plane that is holomporphic at z = 0. Then, set

$$\det(\Delta_s) = \exp(-\zeta'(0))$$

if Δ_s has no kernel and

$$\det(\Delta_s) = 0$$

if it has a kernel.

As a consequence of the work of Quillen, there is a holomorphic line bundle *L* over [0, 1] and a canonical section σ of *L* equipped with a metric such that

$$||\sigma(s)||^2 = \det(\Delta_s)$$

for every $s \in [0, 1]$.

Therefore, it follows that $det(\Delta_s)$ is a smooth real function of s that vanishes exactly when \mathcal{D}_s is not invertible. Pick some $s \in [0, 1]$ such that \mathcal{D}_s is not invertible.

For $s' \neq s$, an orientation of the kernel of \mathcal{D}_s is simply a choice of an element of the two-element set $\{-1, 1\}$. Suppose a uniform orientation $o \in \{-1, 1\}$ has been chosen for all $s' \in (s - \varepsilon, s)$. Then, any $s'' \in (s, s + \varepsilon)$ is oriented in the following manner. For a choice of small enough $\varepsilon > 0$, the determinant function det (Δ_t) is either crossing from negative to positive or from positive to negative for $t \in (s - \varepsilon, s + \varepsilon)$. In the former case, the kernel of $\mathcal{D}_{s''}$ is oriented by o, while in the latter case it is oriented by -o.

Observe for $s \neq 0$, the operator \mathcal{D}_s is injective. Otherwise, if (a, ψ, u) were a non-zero element in the kernel of \mathcal{D}_s , then $(a/s, \psi, u/s)$ would be a non-zero solution of system (S2). However, no such solutions exist.

Therefore, it suffices to calculate the spectral flow contribution at 0. Combined with the canonical orientation on the kernel of \mathcal{D}_0 , this induces a canonical orientation on \mathcal{D}_s for all s > 0.

Although this is a very natural way to define the change of orientation along a path, it is not immediately amenable to calculation. Equation (1.5.9) of [Nic00] provides a convenient way to calculate this spectral flow contribution. Fix a point $(A, 0, \varphi, \tau) \in \mathcal{M}^0$.

Write $\dot{D} = \frac{d}{ds}|_{s=0}D_s$. Restricting this operator to the kernel of D_0 and composing with projection onto the cokernel of D_0 , this produces a linear isomorphism

$$R_0: \ker \mathcal{D}_0 \to \operatorname{coker} \mathcal{D}_0.$$

Let sign $(R_0) \in \{-1, 1\}$ denote whether R_0 is orientation-reversing or orientation-preserving respectively.

Then the spectral flow contribution at 0 is in fact equal to $sign(R_0)$, although the proof of this fact is outside the scope of this thesis.

It remains to find $sign(R_0)$.

The operator $\hat{\mathcal{D}}$ can be written as the map

$$(a, \psi, u) \mapsto (0, 0, \rho(a)\varphi + u\rho(\dot{\beta})\varphi).$$

Letting Π be the projection from the space of negative spinors onto the cokernel of D_A^+ , the operator $R_0 : \ker \mathcal{D}_0 \to \operatorname{coker} D_A^+$ can be written as the map

$$(a, 0, u) \mapsto \Pi(\rho(a)\varphi + u\rho(\dot{\beta})\varphi).$$

Any harmonic imaginary-valued one-form a is an imaginary multiple of the harmonic form $f^*(d\theta)$. It is a quick calculation to show that $\Pi(\rho(a)\varphi) \neq 0$. If it were equal to zero, then there is a spinor ψ orthogonal to φ such that $D_A^+\psi = \rho(a)\varphi$. However, then $(a, 0, \psi)$ is a non-trivial solution to system (S2), which yields a contradiction.

Therefore, the spinors $\Pi(\rho(a)\varphi)$ and $i\Pi(\rho(a)\varphi)$ form a positively-oriented basis of coker D_A^+ .

Write $f_{\tau} = f$. Recall that the connection A is of the form $A_0 + \beta_{\tau} - \log z \cdot f^*(d\theta)$ for A_0 the background spin connection and |z| = 1. The point (z, τ) lies in the spectral set $S \cap C$ discussed in the previous section, so by Theorem 2.3.1 there is an embedded curve $(\gamma(t), t)$ on the interval $(\tau - \varepsilon, \tau + \varepsilon)$ for some small $\varepsilon > 0$ intersecting C transversely at (z, τ) .

Write A_t for the corresponding connection $A_0 + \beta_t - \log \gamma(t) \cdot f^*(d\theta)$. Differentiating at $t = \tau$, we get the

one-form

$$\dot{A} = \dot{\beta} - \partial_t (\log \gamma(t)) \cdot f^*(d\theta).$$

Let φ_t be a path of spinors such that $\varphi_t \in \ker(D_{A_t}^+)$ for all $t \in (\tau - \varepsilon', \tau + \varepsilon')$ for some small $\varepsilon > \varepsilon' > 0$. Taking the derivative at $t = \tau$, it follows that

$$D_A^+ \dot{\varphi} + \rho(\dot{A})\varphi = 0.$$

Therefore, $D_A^+ \dot{\varphi} = \rho(\partial_t (\log \gamma(t)) \cdot f^*(d\theta) - \dot{\beta})\varphi$. In particular, this implies

$$\Pi(\rho(\dot{\beta})\varphi) = \Pi(\rho(\partial_t(\log\gamma(t)) \cdot f^*(d\theta))\varphi)$$

Write $\log \gamma(t) = x(t) + iy(t)$, and $\partial_t (\log \gamma(t)) = \dot{x} + i\dot{y}$.

Then we can write

$$\Pi(\rho(\dot{\beta})\varphi) = \dot{x}\Pi(\rho(f^*(d\theta))\varphi) + i\dot{y}\Pi(\rho(f^*(d\theta))\varphi).$$

Recall from above that $v_1 = \rho(f^*(d\theta)\varphi)$ and $v_2 = i\rho(f^*(d\theta)\varphi)$ form a positively oriented basis of coker D_A^+ .

Pick the positively-oriented basis $e_1 = (0, 0, 1)$, $e_2 = (if^*(d\theta), 0, 0)$ of ker \mathcal{D}_0 .

Then R_0 is the linear map $e_1 \mapsto \dot{x}v_1 + \dot{y}v_2$, $e_2 \mapsto v_2$.

Then sign(R_0) is just equal to the sign of det(R_0), which by definition is equal to \dot{x} . We conclude that the point ($A, 0, \varphi, \tau$) is oriented by the sign of

$$\frac{d}{dt}|_{t=\tau} \operatorname{Re}\log\gamma(t).$$

It is clear that $\dot{x} \neq 0$. If it were equal to 0, then the curve $\gamma(t)$ would not intersect the cylinder *C* transversely by definition.

Therefore, p has sign 1 if $\operatorname{Re} \log \gamma(t)$ is increasing, which is equivalent to $\gamma(t)$ *exiting* the cylinder C. It has sign -1 if $\operatorname{Re} \log \gamma(t)$ is decreasing, which is equivalent to $\gamma(t)$ *entering* the cylinder C. It follows that the sign of p is the same as the contribution of (z, τ) to the spectral flow.

We conclude the following theorem.

Theorem 2.4.1. $\#M^0 = SF(g_t, \beta_t)$.

Combining this with Theorem 1.3.3 and Theorem 2.3.2 immediately implies Theorem 2.1.12, that

$$\lambda_{SW}(X) = \#\mathcal{M}(X, g, \beta) - \operatorname{ind}_{\mathbb{C}} D^+(Z_+, g, \beta) + \operatorname{sign}(Z)/8$$

is independent of the choice of regular pair (g, β) .

Chapter 3

The Splitting Formula

The results of Chapter 2 required only that X is a manifold satisfying assumption (A1). Now, suppose that X also satisfies (A2).

Let *Y* be an embedded integral homology 3-sphere generating $H_3(X;\mathbb{Z})$. Then, the second Stiefel-Whitney class of *Y* vanishes and so it certainly admits a spin structure. Given some spin 4-manifold *Z* bounding *Y* with spin structure extending the one on *Y*, the *Rokhlin invariant* $\rho(Y) \in \mathbb{Z}/2$ of *Y* is defined to be the reduction modulo 2 of the integer sign(*Z*)/8.

It is an immediate consequence of the Hirzebruch signature theorem that $\rho(Y)$ does not depend on the choice of manifold *Z*. Furthermore, for any two choices *Y*, *Y'* of integral homology 3-spheres generating $H_3(X;\mathbb{Z})$, $\rho(Y) = \rho(Y')$. To see this, let \tilde{X} be the infinite cyclic cover constructed by cutting open along *Y*. Then there is a lift of *Y'* embedded inside this cover not intersecting some copy of *Y*, so there exists some homology cobordism between *Y* and *Y'*. The Rokhlin invariant is invariant under homology cobordism, so $\rho(Y) = \rho(Y')$.

Therefore, the Rokhlin invariant is an invariant of the manifold *X*, denoted by $\rho(X)$.

It is proven in Section 9 of [MRS11] that $\lambda_{SW}(X)$ reduces modulo 2 to the Rokhlin invariant $\rho(X)$. It is also well known (see [Sav02]) that the Rokhlin invariant is the modulo 2 reduction of the *Casson invariant* of Y, which is in turn related to the *instanton Floer homology* of Y (see [Tau90] for some important work in this direction). Therefore, one can in some sense consider $\lambda_{SW}(X)$ to be a Seiberg-Witten version of the Casson invariant, and the Floer homology theory that it should be related to is instead the *Seiberg-Witten-Floer homology* of Y.

The only "additional data" needed to reconstruct *X* from the homology 3-sphere *Y* is the cobordism *W* from *Y* to itself. This loosely motivates the possibility that $\lambda_{SW}(X)$ can be split into a sum of two invariants: one arising from the cobordism *W*, and another arising from the Seiberg-Witten-Floer homology of *Y*.

A splitting formula of this exact form has been recently derived by Lin, Ruberman, and Saveliev in the paper [LRS17]. This chapter will be an exposition of this derivation, and all results are attributed to the aforementioned paper unless explicitly stated otherwise.

The first section will be a brief overview of Seiberg-Witten-Floer homology and its main properties. This will provide the notation and definitions necessary to state the splitting formula in Section 2. Following this, some basic homological algebra and additional technical inputs will be used to reduce the proof of the splitting formula to two distinct calculations. The first calculation, which will take up Section 3, is the calculation of the index of the end-periodic Dirac operator on the manifold X with a "stretched neck".

The second calculation, which will take up Section 4, will be a count of the solutions to the Seiberg-Witten equations on this "stretched neck" version of X with a carefully chosen metric and perturbation. The final section will discuss an application towards an obstruction to positive scalar curvature on manifolds satisfying (A1) and (A2).

An integral part of this count of solutions is a technical eigenvalue estimate proven in Section 7 of [LRS17], and its subsequent application to the compactness theorem proved in the first subsection of Section 4 of this chapter. This chapter will focus more on the direct applications of the underlying theory of Seiberg-Witten-Floer homology to proving the splitting formula, and as a result will not discuss the proof of this estimate.

3.1 Seiberg-Witten-Floer homology

In this section, we present a brief exposition of Seiberg-Witten-Floer homology and its main properties. The actual construction of this Floer homology, as with others, is quite technical and cannot be done justice in the limited space of this thesis. Therefore, we will defer the reader to [KM07] for a wonderfully written, rigorous exposition.

Our goal here is rather to clear the quickest path possible to understanding the splitting formula for λ_{SW} and its proof, while still maintaining a coherent and logical progression of ideas. Most proofs are omitted or only briefly outlined.

3.1.1 Morse homology

Kronheimer and Mrowka's construction of Seiberg-Witten-Floer homology can be thought of as an infinitedimensional version of Morse homology on a manifold with boundary. To motivate our exposition of the former, we will sketch here the construction of the latter.

First, we explain the construction of Morse homology for a closed, oriented Riemannian manifold *B*. A rigorous treatment can be found in [AD14].

For a smooth function $f : B \to \mathbb{R}$, a point $x \in B$ is a *critical point* of f if $(df)_x = 0$.

Under the isomorphism of the tangent and cotangent bundles using the metric, the one-form df is identified with the *gradient vector field* grad f. In other words, grad f is the vector field satisfying

$$\langle \operatorname{grad} f, V \rangle = df(V)$$

for any other vector field V.

Taking the Levi-Civita derivative of grad f, we obtain the *Hessian* H(f). The Hessian is a linear map taking vector fields to vector fields. It is a quick exercise to show further that it is a *self-adjoint* linear map.

Now let $x \in B$ be a critical point of the function f. We say x is a *non-degenerate* critical point if the Hessian at x is non-singular.

The function *f* is a *Morse function* if all of its critical points are nondegenerate. Among other things, this ensures that the critical points are a discrete (and therefore finite) set of points in *B*.

As an example, the "height functions" on S^2 or the 2-torus standing on its end are both Morse functions. The critical points are depicted below.

Let *x* be our non-degenerate critical point as above. Since the Hessian is symmetric, it is diagonalizable. Furthermore, it is non-degenerate and so its spectrum does not contain zero. It follows that there exists a



Figure 3.1: The critical points (green) of the height functions on S^2 (left) and T^2 (right).

splitting

$$T_x B = K_x^+ \oplus K_x^-$$

of the tangent space into the positive and negative eigenspaces of the Hessian. The *index* of x, denoted ind(x), is the dimension of K_x^- . We present our two examples again, now with the indices of the critical points labeled.



Figure 3.2: The critical points with indices labelled.

Morse homology is constructed by counting *gradient flow lines* between critical points. Let x, y be two critical points. The space of flow lines from x to y, denoted by M(x, y), is the space of (non-constant) smooth maps $u : \mathbb{R} \to B$ such that:

- $\frac{du}{dt}(t_0) = \operatorname{grad} f(u(t_0))$ for every $t_0 \in \mathbb{R}$.
- $\lim_{t \to -\infty} u(t) = x.$
- $\lim_{t\to\infty} u(t) = y.$

Some examples of gradient flow lines for the height functions on S^2 and the torus are drawn below.

It is instructive to examine the gradient flow lines of the height function on the torus and the sphere. First, observe that there are no flow lines going from a critical point of lower index to a critical point of higher index. Second, there seem to be finitely many flow lines between points of adjacent index, but infinitely many otherwise. Third, the flow lines from the top of S^2 to the bottom of S^2 can be identified



Figure 3.3: Some gradient flow lines (blue) for the height functions on S^2 and T^2 .

with S^1 , a 1-dimensional compact manifold, by taking the midpoint of each flow line. Therefore, we may expect that flow lines between critical points of higher difference in indices form smooth manifolds.

If u(t) is a flow line, then the translation $v(t) = u(t + t_0)$ has the exact same image as u in the manifold B. Therefore, the space of flow lines that corresponds with our topological intuition is the *unparameterized* space of flow lines

$$\check{M}(x,y) = M(x,y)/\mathbb{R}$$

where \mathbb{R} acts freely by translation, defined explicitly as $(t_0 \cdot u)(t) = u(t + t_0)$ for $t_0 \in \mathbb{R}$.

Nearly all of our observations about flow lines are true in general, but not necessarily for any choice of Morse function. The main issue is one of transversality.

Define the *unstable manifold* U_x of a critical point as follows. The gradient of f has a well-defined flow everywhere on the manifold B. Therefore, any point $x' \in B$ is part of a unique flow line u. Then U_x consists of all the points x' such that their associated flow line u satisfies

$$\lim_{t \to -\infty} u(t) = x.$$

Similarly, the stable manifold of a critical point is the set of all points whose flow lines satisfy

$$\lim_{t \to \infty} u(t) = x.$$

In less formal terms, U_x consists of all points that flow *backwards* and settle at x under the gradient flow, while S_x consists of all points that flow *forwards* and settle at x.

Basic Morse theory assures as that the stable and unstable manifolds are indeed manifolds. Furthermore, U_x has dimension ind(x), while S_x has dimension dim(B) - ind(x) (see [AD14] for a proof of this). By definition, we can for two critical points $x \neq y$ write M(x, y) as the intersection of the unstable manifold of x and the stable manifold of y:

$$M(x,y) = U_x \cap S_y.$$

Note that if x = y, then this identity does not hold true since M(x, x) only consists of *non-constant* flow lines, and is therefore empty.

If this intersection were transverse, then from our discussion above it would be immediate that M(x, y) is a manifold of dimension ind(x) - ind(y) when this number is positive, and empty when it is negative.

Applying the free \mathbb{R} -action, we find that the dimension of M(x, y) is one lower.

We can now state our full theorem on the structure of the spaces of flow lines.

Theorem 3.1.1. There is a residual set of Morse functions f such that the spaces $\check{M}(x, y)$ are manifolds for any pair of critical points $x, y \in B$. Furthermore, $\check{M}(x, y)$ has dimension $\operatorname{ind}(x) - \operatorname{ind}(y) - 1$ if $\operatorname{ind}(x) > \operatorname{ind}(y)$, and is empty if $\operatorname{ind}(x) \leq \operatorname{ind}(y)$.

A Morse function f in this residual set is said to satisfy the Morse-Smale condition.

In the process of the proof of this theorem, it is also shown that the parameterized spaces M(x, y) are manifolds of dimension ind(x) - ind(y) if ind(x) - ind(y) if ind(x) > ind(y), and are empty if $ind(x) \le ind(y)$.

Now we have partially confirmed our intuition from before. If ind(x) - ind(y) = 1, then the space of flow lines $\check{M}(x, y)$ is a manifold of dimension zero, and therefore a discrete set of points. However, we do not know yet if it is finite. More generally, we do not know immediately whether the spaces $\check{M}(x, y)$ are compact.

In fact, they are not compact in most cases. One must construct an appropriate compactification, which in this case will be the space of broken flow lines. A *broken flow line* from a critical point x to a critical point y is defined by a sequence of critical points $x = a_0, a_1, \ldots, a_k = y$ and a sequence of flow lines $\gamma_1, \ldots, \gamma_k$ such that $\gamma_i \in \check{M}(a_{i-1}, a_i)$ for every i.

The *space of broken flow lines* from x to y is denoted by $M^+(x, y)$. It can be given a rather natural topology such that the following two important theorems are true.

Theorem 3.1.2. The space $\check{M}^+(x, y)$ is compact.

Theorem 3.1.3. The embedding $\check{M}(x, y) \hookrightarrow \check{M}^+(x, y)$ is continuous. If ind(x) - ind(y) = 2, then $\check{M}^+(x, y)$ is a compact manifold with boundary consisting of exactly the broken trajectories:

$$\partial \check{M}^+(x,y) = \bigcup_{\text{crit. points } z} \check{M}(x,z) \times \check{M}(z,y).$$

Theorem 3.1.2 confirms that the space $\check{M}(x,y)$ is finite when $\operatorname{ind}(x) - \operatorname{ind}(y) = 1$ since $\check{M}(x,y) = \check{M}^+(x,y)$.

Theorem 3.1.3 is the prototypical example of what is known in the Floer theory literature as a *gluing* theorem. This comes roughly from the equivalent idea that two pieces of a broken flow line meeting at a critical point can be "glued" together, i.e. there is some arbitrarily small deformation into a continuous flow line.

We are now ready to define Morse homology, at least over $\mathbb{Z}/2$. The full definition of \mathbb{Z} requires orienting our spaces of flow lines and is discussed in [AD14].

Let *f* be our Morse function satisfying the Morse-Smale condition. Define the $\mathbb{Z}/2$ -vector space $C_k = C_k(B, f; \mathbb{Z}/2)$ to be the space generated by formal sums with coefficients in $\mathbb{Z}/2$ of the critical points of index *k*.

There are natural boundary maps $\partial_k : C_k \to C_{k-1}$ given by counting flow lines. If M(x, y) has dimension zero, then let N(x, y) be the number of points in M(x, y) modulo 2. For some critical point $[x] \in C_k$, we then define the boundary map by

$$\partial_k[x] = \sum_{\text{crit points } y \text{ of index } k - 1} N(x, y)[y].$$



Figure 3.4: Gradient flow lines from x to y (left) converging to a broken trajectory from x to z to y (right).

The map $\partial^2 : C_k \to C_{k-2}$ can be written as

$$\partial^2[x] = \sum_y (\sum_z N(x,z)N(z,y))[y]$$

where the sum is taken over all critical points y of index k - 2 and critical points z of index k - 1.

For a fixed critical point y of index k - 2, the sum $\sum_{z} N(x, z)N(z, y)$ is exactly a modulo 2 count of the broken trajectories from x to y, i.e. the number of points in the zero-dimensional compact manifold

$$\bigcup_{\text{crit. points } z} \check{M}(x,z) \times \check{M}(z,y)$$

By Theorem 3.1.3, this manifold is the boundary of the compact 1-dimensional manifold $\check{M}^+(x, y)$. As the boundary of a 1-dimensional manifold, it must have an even number of points. We conclude:

Corollary 3.1.4. $\partial^2 = 0$.

The set of vector spaces $\{C_k(B, f; \mathbb{Z}/2)\}$ along with the boundary maps $\{\partial_k\}$ therefore a *chain complex*, and the homology of this chain complex is the *Morse homology* $H_*^{Morse}(B, f; \mathbb{Z}/2)$.

There are a few more verifications that we will not make here. First, Morse homology is *independent* of the choice of Morse function f, so we can write Morse homology as a single sequence of groups $H_*^{Morse}(B; \mathbb{Z}/2)$.

We can compute Morse homology explicitly for the sphere S^2 and the torus T^2 with the height function as the Morse function.

There are two critical points on S^2 for the height function. One is of index 2, and one is of index zero. Therefore, the maps in the Morse chain complex are all the zero map and we conclude

$$H_k^{Morse}(S^2; \mathbb{Z}/2) = \begin{cases} \mathbb{Z}/2 & k = 0, 2\\ 0 & \text{otherwise} \end{cases}$$

On the other hand, the torus has four critical points. Two are of index 1, one is of index zero, and one is of index two. Now it is necessary to count flow lines, but from examining Figure 3.3, one sees that there are an even number of flow lines between critical points of adjacent index. Therefore, the maps in the Morse

chain complex are again the zero map and we conclude

$$H_k^{Morse}(T^2; \mathbb{Z}/2) = \begin{cases} \mathbb{Z}/2 & k = 0, 2\\ \mathbb{Z}/2 \oplus \mathbb{Z}/2 & k = 1\\ 0 & \text{otherwise} \end{cases}$$

As these initial computations may begin to indicate, there is an isomorphism of Morse homology with the standard singular homology:

$$H^{Morse}_*(B;\mathbb{Z}/2) \simeq H_*(B;\mathbb{Z}/2).$$

3.1.2 Morse homology with boundary

Now we will let *B* be a compact Riemannian manifold with nonempty boundary ∂B .

Rather than working with a completely general Morse function on *B*, it will be easier to work with Morse functions with gradient vector field tangent to the boundary.

We will follow the setup in Chapter 2 of Kronheimer and Mrowka ([KM07]). Fix a Morse function f such that grad f is always tangent to the boundary. More precisely, for any $x \in \partial B$, we have

$$\operatorname{grad} f(x) \subset T_x \partial B \subset T_x B.$$

The critical points of *f* split into three different types. The first type is critical points in the interior of *B*.

Critical points on ∂B are classified into two categories. Let ν denote the unit normal vector field of ∂B . At any boundary critical point $x \in \partial B$, since the gradient of f vanishes and is contained within the tangent bundle of ∂B , it is evident that $T_x \partial B$ is an invariant subspace of the Hessian. It follows that ν must be an eigenvector of the Hessian, and since the Hessian is nondegenerate it has either a positive or negative eigenvalue. In the former case, the critical point x is called *boundary-stable*, while in the latter case it is called *boundary-unstable*.

Boundary-stable and boundary-unstable critical points admit opposite topological descriptions of their associated stable and unstable manifolds. If x is boundary-stable, then its unstable manifold U_x is contained in the boundary ∂B , while its stable manifold S_x may intersect the interior of B. If x is boundary-unstable, then its unstable manifold may intersect the interior of B, while its stable manifold is contained in the boundary ∂B .

This leads to a new transversality problem that did not arise in the case of empty boundary. If x is boundary-stable and y is boundary-unstable, then the intersection $M(x, y) = U_x \cap S_y$ cannot hope to be transverse in B, since the tangent bundles of U_x and S_y are both contained in the tangent bundle of ∂B . We call such a space of flow lines *boundary-obstructed*.

This issue is addressed instead by requiring that the intersection is transverse in ∂B . In this case, a dimension count indicates that the expected dimension of M(x, y) will be ind(x) - ind(y) + 1 rather than ind(x) - ind(y). With this modification, an analogous theorem to Theorem 3.1.1 holds:

Theorem 3.1.5. There is a residual set of Morse functions f such that the spaces M(x, y) are manifolds for any pair of critical points $x, y \in B$.

If M(x, y) is not boundary-obstructed, it has dimension ind(x) - ind(y) - 1 if ind(x) > ind(y), and is empty if $ind(x) \le ind(y)$.

If $\tilde{M}(x, y)$ is boundary-obstructed, it has dimension $\operatorname{ind}(x) - \operatorname{ind}(y)$ if $\operatorname{ind}(x) > \operatorname{ind}(y)$, and is empty if $\operatorname{ind}(x) \leq \operatorname{ind}(y)$.

If *x* and *y* are two general boundary critical points, then M(x, y) can intersect the interior of *B* if and only if *x* is boundary-unstable and *y* is boundary-stable. In this case, the intersection $M^{\partial}(x, y) = M(x, y) \cap \partial B$ consisting of flow lines that lie on the boundary is a codimension one submanifold of M(x, y). The same holds true for the unparameterized version $\check{M}^{\partial}(x, y) = \check{M}(x, y) \cap \partial B$.

As before, we can define the space of broken flow lines $\check{M}^+(x, y)$ and state the following theorem:

Theorem 3.1.6. $\check{M}^+(x, y)$ is compact.

Gluing is a good deal more complicated in the presence of a boundary. The boundary of $\check{M}^+(x, y)$ depends on the types of x and y. We state the full theorem below, and discuss one of the cases in a little more depth right after. It is not necessary for the rest of this thesis to know the statement of this theorem, but thinking about the topology behind each of the cases is helpful for understanding the analogous gluing theorem in Seiberg-Witten-Floer homology.

Theorem 3.1.7. 1. If x and y are both interior critical points, x has index k, and y has index k - 2 then $\check{M}^+(x, y)$ has boundary components of the form:

- $\check{M}(x,z) \times \check{M}(z,y)$, with z an interior critical point of index k-1.
- *M*(*x*, *a*) × *M*(*a*, *b*) × *M*(*b*, *y*), with *a* a boundary-stable critical point of index *k* − 1, and *b* a boundaryunstable critical point of index *k* − 1 (so *M*(*a*, *b*) is boundary-obstructed).
- 2. If x is an interior critical point and y is a boundary-stable critical point, x has index k and y has index k 2 then $\check{M}^+(x, y)$ has boundary components of the form:
 - $\check{M}(x,z) \times \check{M}(z,y)$, with z an interior critical point of index k-1.
 - $\check{M}(x, a) \times \check{M}(a, y)$, with a a boundary-stable critical point of index k 1.
 - *M*(*x*, *a*) × *M*(*a*, *b*) × *M*(*b*, *y*) with *a a* boundary-stable critical point of index k − 1, and b a boundary-unstable critical point of index k − 1.
- 3. If x is a boundary-unstable critical point and y is an interior critical point, x has index k and y has index k 2 then $\check{M}^+(x, y)$ has boundary components of the form:
 - $\check{M}(x,z) \times \check{M}(z,y)$, with z an interior critical point of index k-1.
 - $\check{M}(x,b) \times \check{M}(b,y)$, with b a boundary-unstable critical point of index k-1.
 - *M*(*x*, *a*) × *M*(*a*, *b*) × *M*(*b*, *y*) with *a* a boundary-stable critical point of index *k* − 1, and *b* a boundaryunstable critical point of index *k* − 1.
- 4. If x is a boundary-unstable critical point, y is a boundary-stable critical point, x has index k and y has index k 2 then $\check{M}^+(x, y)$ has boundary components of the form:
 - $\check{M}^{\partial}(x, y)$, the space of flow lines from x to y along ∂B .
 - $\check{M}(x,z) \times \check{M}(z,y)$, with z an interior critical point of index k-1.
 - $\check{M}(x, a) \times \check{M}(a, y)$, with a a boundary-stable critical point of index k 1.

- $\check{M}(x,b) \times \check{M}(b,y)$, with b a boundary-unstable critical point of index k-1.
- $\check{M}(x,a) \times \check{M}(a,b) \times \check{M}(b,y)$ with a a boundary-stable critical point of index k-1, and b a boundaryunstable critical point of index k-1.
- 5. If x and y are any boundary critical points, x has index k and y has index k 2 then $\check{M}^+(x, y)$ (replaced with $\check{(M^{\partial})}^+(x, y)$ in the case where x and y satisfy (4)) has boundary components of the form
 - $\check{M}(x,z) \times \check{M}(z,y)$ with z any boundary critical point of index k-1.



Figure 3.5: Interior trajectories converging to a three-component broken trajectory, with the middle component on the boundary.

Consider, for example, the situation in part (1) of Theorem 3.1.7, which is illustrated in Figure 3.5. A sequence of trajectories between two interior critical points can still converge to a two-component trajectory between three interior critical points. However, the addition of boundary points complicates the matter and a sequence of trajectories could also split into a three-component trajectory with one component running along the boundary ∂B .

The complexity of the boundaries of the compactified spaces of flow lines is represented in the construction of the Morse chain complexes. As before, we will ignore the issue of orientation and work in characteristic 2. For an integer k, let C_k^o , C_k^s , and C_k^u be the vector spaces over $\mathbb{Z}/2$ of formal sums of interior, boundary-stable, and boundary-unstable critical points of index k respectively. Then, define the chain groups

$$\overline{C}_k = C_k^s \oplus C_{k+1}^u,$$

$$\check{C}_k = C_k^o \oplus C_k^s,$$

$$\hat{C}_k = C_k^o \oplus C_k^u.$$

Kronheimer and Mrowka constructed differentials between these chain groups out of eight distinct operators that count flow lines. For critical points x, y such that $\check{M}(x,y)$ is zero-dimensional, let n(x,y) be the modulo two count of points in $\check{M}(x,y)$.

The first four operators (with k arbitrary) are denoted by

$$\begin{split} &\partial_o^o: C_k^o \to C_{k-1}^o, \\ &\partial_s^o: C_k^o \to C_{k-1}^s, \\ &\partial_o^u: C_k^u \to C_{k-1}^o, \\ &\partial_o^u: C_k^u \to C_{k-1}^o, \\ &\partial_s^u: C_k^u \to C_{k-1}^s. \end{split}$$

The operator ∂_o^o is defined by

$$\partial_o^o([x]) = \sum_y n(x,y)[y],$$

where the sum is taken across all interior critical points y of index k - 1.

The operators ∂_s^o , ∂_o^u , ∂_s^u are defined similarly.

The last four operators are denoted by

$$\begin{split} \overline{\partial}_s^s &: C_k^s \to C_{k-1}^s, \\ \overline{\partial}_u^s &: C_k^s \to C_k^u, \\ \overline{\partial}_s^u &: C_k^u \to C_{k-2}^s, \\ \overline{\partial}_u^u &: C_k^u \to C_{k-1}^u. \end{split}$$

To define these operators, consider the case where x and y are boundary critical points. Then, let $\check{M}^{\partial}(x,y) = \check{M}(x,y) \cap \partial B$ be the manifold of flow lines from x to y that are contained in the boundary. Note that this coincides with $\check{M}(x,y)$ except possibly in the case where x is boundary-unstable and y is boundary-stable.

If $\check{M}^{\partial}(x, y)$ is zero-dimensional, let $\overline{n}(x, y)$ be the modulo two count of points in this space. In the boundary-obstructed case, this is true when the index of *y* equals the index of *x*. In the case where *x* is boundary-unstable and *y* is boundary-stable, it is true when the index of *y* is *two* less than the index of *x*. In the other two cases, it is true when the index of *y* is *one* less than the index of *x*.

With all of these considerations in mind, the operators $\overline{\partial}_*^*$ are defined in an analogous way to the operators ∂_*^* , counting boundary flow lines using $\overline{n}(x, y)$ instead of all flow lines using n(x, y).

Note that there are two operators, ∂_s^u and $\overline{\partial}_s^u$ counting flow lines from a boundary-unstable critical point and a boundary-stable critical point. This is due to the separate contribution of the space of boundary flow lines $\check{M}^{\partial}(x, y)$ to the boundary of the compactified space.

We can package these eight operators into differentials for the chain complexes \overline{C}_* , \check{C}_* , and \hat{C}_* :

$$\begin{split} \overline{\partial} &= \begin{pmatrix} \overline{\partial}_s^s & \overline{\partial}_s^u \\ \overline{\partial}_u^s & \overline{\partial}_u^u \end{pmatrix}, \\ \check{\partial} &= \begin{pmatrix} \partial_o^o & \partial_o^u \overline{\partial}_u^s \\ \partial_s^o & \overline{\partial}_s^s + \partial_s^u \overline{\partial}_u^s \\ \partial_e^o & \overline{\partial}_u^u + \overline{\partial}_e^s \overline{\partial}_u^u \end{pmatrix}, \\ \hat{\partial} &= \begin{pmatrix} \partial_o^o & \partial_u^o \\ \overline{\partial}_u^s \partial_s^o & \overline{\partial}_u^u + \overline{\partial}_u^s \partial_s^u \end{pmatrix}. \end{split}$$

A direct consequence of our gluing theorem are the following identities for the eight operators.

Theorem 3.1.8. 1. $\partial_o^o \partial_o^o + \partial_o^u \overline{\partial}_u^s \partial_o^o = 0.$ 2. $\partial_s^o \partial_o^o + \overline{\partial}_s^s \partial_s^o + \partial_s^u \overline{\partial}_u^s \partial_s^o = 0.$ 3. $\partial_o^o \partial_o^u + \partial_s^u \overline{\partial}_u^u + \partial_o^u \overline{\partial}_u^s \partial_s^u = 0.$ 4. $\overline{\partial}_s^u + \partial_s^o \partial_o^u + \overline{\partial}_s^o \overline{\partial}_s^u + \overline{\partial}_s^u \overline{\partial}_u^u + \partial_s^u \overline{\partial}_u^s \partial_s^u = 0.$ 5. $\overline{\partial}^2 = 0.$

Proof. Each of 1 to 5 follows directly from the corresponding statement in Theorem 3.1.7 and the fact that a compact 1-dimensional manifold with boundary has an even number of boundary points. \Box

Applying identities 1-4 from this theorem and a bit of algebra then gives us:

Corollary 3.1.9. $\check{\partial}^2 = 0$ and $\hat{\partial}^2 = 0$.

Taking the homologies of these three chain complexes, we obtain three sequences of homology groups:

$$\overline{H}^{Morse}_{*}(B), \check{H}^{Morse}_{*}(B), \text{ and } \widehat{H}^{Morse}_{*}(B).$$

Just like the case of Morse homology without boundary, these three Morse homologies are isomorphic to three singular homologies associated to the manifold *B*:

$$\begin{aligned} \overline{H}^{Morse}_{*}(B) &\simeq H_{*}(\partial B), \\ \breve{H}^{Morse}_{*}(B) &\simeq H_{*}(B), \\ \widehat{H}^{Morse}_{*}(B) &\simeq H_{*}(B, \partial B). \end{aligned}$$

3.1.3 The Chern-Simons-Dirac functional and its gradient flow

Now we can motivate the construction of Seiberg-Witten-Floer homology by direct analogy with Morse theory on a manifold with boundary.

Let *Y* be a closed, oriented Riemannian integral homology 3-sphere. The construction of Seiberg-Witten-Floer homology in [KM07] works for all closed, oriented Riemannian 3-manifolds but it is a bit simpler for the case of integral homology 3-spheres and this is the only case we will need.

We can define a spin^{*c*} structure on a 3-manifold in a similar manner to the definition of 4-manifolds. The data of a spin^{*c*} structure on *Y* is a rank two complex vector bundle $S \rightarrow Y$ equipped with a Clifford multiplication

$$\rho: TY \to \operatorname{Hom}(S, S).$$

As before, ρ is a bundle isometry between *TY* and $\mathfrak{su}(S) \subset \operatorname{Hom}(S, S)$.

Given a spin^{*c*} structure on *Y*, we can define the configuration space C(Y), the group of gauge transformations \mathcal{G} , and the quotient space $\mathcal{B}(Y)$ identically to the four-manifold case.

The blown-up configuration space $C^{\sigma}(Y)$ is the space of tuples (B, r, ψ) such that B is a unitary connection on the determinant line bundle det(S), r is a nonnegative real number, and ψ is a unit-length spinor (a section of S).

Seiberg-Witten-Floer homology is the "infinite-dimensional Morse homology" derived from counting flow lines of the gradient of a functional on the space $\mathcal{B}^{\sigma}(Y) = \mathcal{C}(Y)/\mathcal{G}$, which is a Hilbert manifold with boundary.

The other ingredient to a Morse theory is, of course, the Morse function. This is given by the *Chern-Simons-Dirac functional*. Pick a base connection B_0 on the determinant line bundle. Then for any other such connection B, the operator $B - B_0$ is an imaginary-valued one-form on Y.

Definition 3.1.10. The Chern-Simons-Dirac (CSD) functional

$$\mathcal{L}:\mathcal{C}(Y)\to\mathbb{R}$$

is defined by

$$\mathcal{L}(B,\psi) = -\frac{1}{8} \int_Y (B - B_0) \wedge (F_B + F_{B_0}) \operatorname{dvol} + \frac{1}{2} \int_Y \langle D_B \psi, \psi \rangle \operatorname{dvol}.$$

The CSD functional extends to a functional

$$\mathcal{L}^{\sigma}: \mathcal{C}^{\sigma}(Y) \to \mathbb{R}$$

by setting

$$\mathcal{L}^{\sigma}(B, r, \psi) = \mathcal{L}(B, r\psi).$$

The CSD functional is only guaranteed to be gauge-invariant when the first Chern class of the chosen spin^{*c*} structure is torsion, which is true in the setting of this paper (see Lemma 4.1.3 in [KM07] for the non-torsion case).

The CSD functional is now well-defined as a map $\mathcal{L}^{\sigma} : \mathcal{B}^{\sigma}(Y) \to \mathbb{R}$. For the sake of concreteness we will generally treat it as a map out of $\mathcal{C}^{\sigma}(Y)$, but it is understood (and important!) that everything proceeds smoothly upon passing to the quotient space $\mathcal{B}^{\sigma}(Y)$. The motivation behind the particular choice of the CSD functional as a Morse function lies in the special form taken on by its gradient.

It is a routine computation to find

$$\operatorname{grad}\mathcal{L}^{\sigma}(B,r,\psi) = (-*F_B - 2r^2\rho^{-1}(\psi\psi^*)_0, -\Lambda(B,r,\psi)r, -D_B\psi + \Lambda(B,r,\psi)\psi)$$

where $\Lambda(B, r, \psi) = \operatorname{Re}\langle \psi, D_B \psi \rangle_{L^2}$.

Let $\gamma(t)$ be a flow line for the gradient flow of \mathcal{L}^{σ} , i.e. a continuous path of tuples $(B(t), r(t), \psi(t)) \in \mathcal{C}^{\sigma}(Y)$ for all $t \in \mathbb{R}$ satisfying

$$\frac{d}{dt}\gamma(t) = \operatorname{grad} \mathcal{L}^{\sigma}(B, r, \psi).$$

The path $\gamma(t)$ defines an element (A, s, φ) of a the blown-up configuration space denoted by $C^{\sigma}(Z)$, where $Z = \mathbb{R} \times Y$ is the metric cylinder. To lend some rigor to the work below, we must clarify the spin^{*c*} structure on *Z*. The spin^{*c*} structure on *Z* is defined by spinor bundles $S^+, S^- \simeq S$ with $S_Z = S \oplus S$, and Clifford multiplication ρ_Z is defined by

$$\rho_Z(\partial/\partial t) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$
$$\rho_Z(w) = \begin{pmatrix} 0 & -\rho(w)^* \\ \rho(w) & 0 \end{pmatrix}$$

Let t, y denote the \mathbb{R} - and Y-coordinates on the cylinder Z.

Then the connection *A* is defined by its covariant derivative:

$$\nabla_{A;\partial_t} = \frac{d}{dt} + \nabla_{B(t);\partial_t}$$
$$\nabla_{A;\partial_y} = \nabla_{B(t);\partial_y}.$$

Note that not all connections on Z take this form. We say connections like A are in *temporal gauge*.

The scalar *s* is the L^2 norm of the path $\psi(t)$ over the cylinder: $||\psi(t)||_{L^2(Z)}$. The spinor φ is the rescaled path of spinors $\psi(t)/||\psi(t)||_{L^2(Z)}$.

Now suppose the path $(B(t), r(t), \psi(t))$ is a flow line for the gradient flow of \mathcal{L}^{σ} . Let (A, s, φ) be the associated element in $\mathcal{C}^{\tau}(Z)$.

The following computation is not too difficult but fundamental, so we will write it out here.

The curvature of *A* is given by

$$F_A = dt \wedge \frac{d}{dt}B + F_B$$

and its self-dual projection is then equal to

$$F_{A}^{+} = \frac{1}{2}(F_{A} + *F_{A})$$

= $\frac{1}{2}(dt \wedge \frac{d}{dt}B + F_{B} + *_{3}\frac{d}{dt}B + dt \wedge *_{3}F_{B})$
= $\frac{1}{2}(dt \wedge (\frac{d}{dt}B + *_{3}F_{B}) + F_{B} + *_{3}\frac{d}{dt}B),$

where $*_3$ denotes the three-dimensional Hodge star.

Now we plug in the gradient flow equation and apply Clifford multiplication to get

$$\frac{1}{2}\rho_Z(F_A^+) = s^2(\varphi\varphi^*)_0.$$

It is also immediate that

$$D_A^+\varphi = (\frac{d}{dt} + D_B)(\psi(t)) = 0$$

by the gradient flow equation.

These equations satisfied by (A, s, φ) are quite familiar to us by now. We have verified one direction of the following theorem:

Theorem 3.1.11. *Gradient flow lines of the CSD functional on* $\mathcal{B}^{\sigma}(Y)$ *are in one-to-one correspondence with gauge-equivalence classes of solutions to the blown-up Seiberg-Witten equations on the cylinder Z.*

This theorem provides the jumping off point for understanding the Morse theory of \mathcal{L}^{σ} and constructing Floer homology. The construction of the moduli spaces of flow lines and the relevant compactness and gluing theorems all require a large amount of detailed analysis on the Seiberg-Witten equations on the cylinder. Many of the arguments in the upcoming sections have their basis in this analysis.

Before we move on, the last additional piece of vocabulary that we must present is perturbations. Just like in the finite dimensional case, there are often transversality issues and our moduli spaces of flow lines are not necessarily manifolds. To fix this, we perturb the CSD functional.

Let $f : \mathcal{C}(Y) \to \mathbb{R}$ be some gauge-invariant smooth function. Then

$$\operatorname{grad}(\mathcal{L}+f) = \operatorname{grad}\mathcal{L} + \mathfrak{q}$$

for a vector field $q = \operatorname{grad} f$.

Since the gradient is the main object of interest here, we widen our class of possible perturbations a little to vector fields q that are the *formal gradient* of some gauge-invariant function f, i.e. along any path $\gamma : [0,1] \rightarrow C(Y)$, we have

$$\int_0^1 \langle \mathfrak{q}(\gamma(t)), \dot{\gamma}(t) \rangle dt = f(\gamma(1)) - f(\gamma(0)).$$

A perturbation q then induces a perturbation on the blown-up space $C^{\sigma}(Y)$, denoted by q^{σ} . Through the correspondence between gradient flow lines and solutions to the Seiberg-Witten equations, this defines a perturbation \hat{q}^{σ} to the Seiberg-Witten equations on the cylinder *Z*. These constructions and further details can be found in Chapter 10 of [KM07].

The following (somewhat vague) theorem will be all that we need. Recall that *Y* is fixed to be an integral homology sphere. A perturbation \mathfrak{q} is said to be *nice* if \mathfrak{q}^{σ} vanishes at the reducible configurations.

Furthermore, we say a critical point [a] is *non-degenerate* if the Hessian of the Chern-Simons-Dirac functional at [a] is surjective when considered as a map on tangent spaces.

Theorem 3.1.12. There is a Banach space \mathcal{P} of nice perturbations and a residual subset $U \subset \mathcal{P}$ such that all of the critical points are non-degenerate and all of the spaces of flow lines for the Chern-Simons-Dirac functional perturbed by any element of U are smooth, finite-dimensional manifolds of the expected dimension.

There is some missing subtlety with regards to what the "expected dimension" is and so on, but we may regard this as an analogue of Theorem 3.1.1.

3.1.4 Floer homology, the exact triangle, and cobordism maps

Assuming all goes well, we can define three Floer chain complexes by counting flow lines and obtain three Floer homology groups for our 3-manifold *Y*:

$$\overline{HM}_*(Y)$$
, $\widetilde{HM}_*(Y)$, and $\widehat{HM}_*(Y)$.

These are pronounced "HM-bar", "HM-to", and "HM-from" respectively.

These three Floer homologies fit into an exact triangle:



The image of the map j_* is denoted by $HM_*(Y)$, known as the *reduced Floer homology* of Y.

The other important property of Seiberg-Witten-Floer homology is that it is functorial with respect to cobordisms. Let Y_0 and Y_1 be two closed, oriented Riemannian manifolds and W an cobordism with oriented boundary $-Y_0 \cup Y_1$. Then W induces maps

$$\overline{HM}_{*}(W) : \overline{HM}_{*}(Y_{0}) \to \overline{HM}_{*}(Y_{1}),$$

$$\overline{HM}_{*}(W) : \overline{HM}_{*}(Y_{0}) \to \overline{HM}_{*}(Y_{1}),$$

$$\widehat{HM}_{*}(W) : \widehat{HM}_{*}(Y_{0}) \to \widehat{HM}_{*}(Y_{1}).$$

Furthermore, if Y_0 , Y_1 , Y_2 are three 3-manifolds with cobordisms $W_0 : Y_0 \to Y_1$, $W_1 : Y_1 \to Y_2$ then the composite cobordism $W_1 \circ W_0 : Y_0 \to Y_2$ satisfies

$$\overline{HM}_*(W_1 \circ W_0) = \overline{HM}_*(W_1) \circ \overline{HM}_*(W_0)$$

and so on for the other two types of maps.

The cobordism maps are in fact defined on the chain level. They arise from counting solutions to the Seiberg-Witten equations on the non-compact manifold

$$W_{\infty} = (-\infty, 0] \times Y_0) \cup W \cup ([0, \infty) \times Y_1)$$

given by attaching tubular ends to the boundary of *W*, depicted in the figure below.

A solution to the Seiberg-Witten equations on W^* when restricted to the tubular end $(-\infty, 0] \times Y_0$ can be regarded as a half flow-line for the CSD functional on $\mathcal{B}^{\sigma}(Y_0)$. The same is true for the restriction to the tubular end $[0, \infty) \times Y_1$.

For a critical point $[\alpha_0] \in \mathcal{B}(Y_0)$ and a critical point $[\alpha_1] \in \mathcal{B}(Y_1)$, we can define the "moduli spaces of flow lines" from $[\alpha_0]$ to $[\alpha_1]$ to be the solutions to the Seiberg-Witten equations on W^* whose restrictions to $(-\infty, 0] \times Y_0$ and $[0, \infty) \times Y_1$ tend to $[\alpha_0]$ and $[\alpha_1]$ respectively.



Figure 3.6: The manifold W^* with a Seiberg-Witten trajectory (blue) from $[\mathfrak{a}_0]$ to $[\mathfrak{a}_1]$

For example, if we denote $C^o_*(Y_i)$ to be the abelian group generated by formal sums of the irreducible critical points in $\mathcal{B}^{\sigma}(Y_i)$ for i = 0, 1, then the cobordism *W* defines a map

$$m_o^o: C^o_* \to C^o_*$$

The cobordism maps also commute with the exact triangles on Y_0 and Y_1 . The relevant diagram is:

As a consequence, we find that *W* induces a map

$$HM_*(W): HM_*(Y_0) \to HM_*(Y_1).$$

One may notice that everything defined here has a "star" subscript. This is to serve as an indication that all of these objects are graded. However, the grading is more complex than in the case of Morse homology. In that setting, we could take the index of a critical point to be the number of negative eigenvalues of the Hessian of the Morse function at that point. However, the Hessian of the CSD functional has infinitely many positive and negative eigenvalues, so this approach does not work as is.

3.1.5 The h-invariant

Before defining the *h*-invariant of *Y*, we will discuss the reducible critical points of *Y* and their gradings. Fix a nice perturbation q as in Theorem 3.1.12.

Lemma 3.1.13. There is exactly one reducible critical point $[\theta] \in C(Y)$ of the Chern-Simons-Dirac functional with perturbation q.

Proof. Since q is nice, it suffices to show this holds true in the absence of a perturbation.

By our computation of the gradient, a reducible configuration (B, 0) is a critical point if and only if the connection *B* is flat. Therefore, it suffices to show any two flat unitary connections on the determinant line bundle on *Y* are gauge-equivalent.

Let *B* and *B'* be two such flat connections. The difference b = B - B' is an imaginary-valued one-form on *Y*. Since the connections are flat, it follows that $db = F_B - F_{B'} = 0$, so *b* is closed. However, since $H^1(Y;\mathbb{Z}) = 0$, it is also exact and so $b = d\xi$ for some imaginary-valued smooth function ξ .

Apply the gauge transformation $u = e^{\xi/2}$ to *B* to get $u(B) = B - d\xi = B'$ as desired.

Expand $[\theta] = [(B_0, 0)]$. Let \mathfrak{q}^1 denote the spinorial part of the perturbation \mathfrak{q} over $\mathcal{C}(Y)$. By examining the blown-up gradient flow, we obtain the following corollary:

Corollary 3.1.14. The reducible critical points in $\mathcal{B}^{\sigma}(Y)$ of the Chern-Simons-Dirac functional with perturbation \mathfrak{q}^{σ} are the equivalence classes represented by tuples of the form $(B_0, 0, \psi)$. Here B_0 is the connection part of a representative of $[\theta]$ and ψ is an eigenvector of the operator $\phi \mapsto D_B \phi + \mathcal{D}_{(B_0,0)}\mathfrak{q}^1(0, \phi)$.

The operator in question is a small perturbation of the self-adjoint Dirac operator D_B , so it has a discrete, real spectrum with trivial kernel. We can label the eigenvalues as the sequence $\{\lambda_i\}_{i\in\mathbb{Z}}$. We require $\lambda_i > 0$ if $i \ge 0$, $\lambda_i < 0$ if i < 0, and $\lambda_{i+1} > \lambda_i$ for every $i \in \mathbb{Z}$.

The critical point in $\mathcal{B}^{\sigma}(Y)$ corresponding to the eigenvalue λ_i will be labeled henceforth as $[\mathfrak{a}_i]$. This construction also gives the most direct classification of boundary critical points. The critical points $[\mathfrak{a}_i]$ for $i \ge 0$ are the *boundary-stable* critical points and the critical points $[\mathfrak{a}_i]$ for i < 0 are the *boundary-unstable* critical points.

As mentioned at the end of the previous section, the critical points admit a grading. This grading takes values in $\frac{m}{n} + \mathbb{Z}$ for some rational number $\frac{m}{n}$. The \mathbb{Q} -grading of a critical point [b] is written as $\operatorname{gr}^{\mathbb{Q}}([\mathfrak{b}])$.

The only necessary information about the \mathbb{Q} -grading that will be used in the following exposition is the grading of the reducible critical points. Pick any spin manifold Z bounding Y and extending its spin structure. Then, attach a cylindrical end $[0, \infty) \times Y$ to Z to create a compact spin^{*c*} manifold denoted by Z_{∞} . Letting $D^+(Z_{\infty})$ be the Dirac operator, the number

$$n(Y) = \operatorname{ind} D^+(Z_{\infty}) - \operatorname{sign}(Z)/8$$

can be shown to be independent of the choice of *Z* by an excision argument. Note that n(Y) depends on the spin structure on *Y*, although this notation has been omitted.

Lemma 3.1.15. The critical points $[\mathfrak{a}_i]$ for $i \ge 0$ satisfy $gr^{\mathbb{Q}}([\mathfrak{a}_i]) = -2n(Y) + 2i$.

Proof. This follows from unrolling Definition 28.3.1. of [KM07].

There is an additional grading taking values in $\mathbb{Z}/2$, denoted by $\operatorname{gr}^{\mathbb{Z}/2}([\mathfrak{b}])$. For $i \in \mathbb{Z}$, $\operatorname{gr}^{\mathbb{Z}/2}([\mathfrak{a}_i]) = 0$.

Both gradings interact as one may expect with the exact triangle and cobordism maps. The maps i_* and j_* in the exact triangle are degree zero maps, while the map p_* has degree -1. The cobordism maps preserve gradings.

The *h*-invariant h(Y) is a rational number derived from the Q-gradings of the reducible critical points. It was first introduced by Frøyshov (and so is often called the "Frøyshov invariant") in [Frø10] for a possibly different version of Seiberg-Witten-Floer homology, and adapted to Kronheimer and Mrowka's version in Chapter 39 of [KM07].

Recall one of the component maps of the exact triangle:

$$i_*: \overline{HM}_*(Y) \to \widecheck{HM}_*(Y).$$

The boundary-stable critical points lie within the Floer chain complexes for both of these groups. Therefore, the following is well-defined.

Definition 3.1.16. The *h*-invariant h(Y) is the integer such that -2h(Y) is the lowest grading of a nonzero class in $im(i_*) \subseteq HM_*(Y)$ represented by a boundary-stable critical point.

Accompanying this definition, Frøyshov also showed in his paper that, for a 4-manifold X satisfying (A1) and (A2), any embedded integral homology three-sphere generating $H_3(X;\mathbb{Z})$ has the same *h*invariant.

We are now, at long last, armed with all the proper definitions and vocabulary to properly understand and discuss the splitting formula for $\lambda_{SW}(X)$.

3.2 The splitting formula and its proof

Let *X* be a manifold satisfying assumptions (A1) and (A2), *Y* an embedded integral homology 3-sphere generating $H_3(X;\mathbb{Z})$, and *W* the cobordism from *Y* to *Y* given by cutting open *X* along *Y*. As discussed before, *W* induces a map on the reduced Floer homology of *Y*:

$$HM(W)_* : HM_*(Y) \to HM_*(Y).$$

Furthermore, the map HM(W) preserves the mod 2 grading on $HM_*(Y)$. Therefore, there is a well-defined *Lefschetz number* with respect to this grading, defined as

$$\operatorname{Lef}(HM(W)_*) = \sum_{i \in \mathbb{Z}/2} (-1)^i \operatorname{Tr}(HM(W)_i : HM_i(Y) \otimes \mathbb{Q} \to HM_i(Y) \otimes \mathbb{Q})$$

Denote by h(Y) the *h*-invariant of *Y*, discussed at the end of the last section. The splitting formula is stated as follows.

Theorem 3.2.1. $\lambda_{SW}(X) \equiv \text{Lef}(HM(W)_*) + h(Y) \text{ modulo } 2.$

This splitting formula is a concrete realization of the vague idea mentioned at the start of the chapter. The invariant $\lambda_{SW}(X)$ decomposes into an invariant depending on the cobordism W (Lef($HM(W)_*$)) and one depending on the integral homology 3-sphere Y (the *h*-invariant). Furthermore, we must mention that this formula is the modulo 2 reduction of the integral splitting formula presented in [LRS17]. Owing to technical reasons regarding orientations, we will only describe the proof of this version.

The first part of the proof of Theorem 3.2.1 relies on a couple of natural exact sequences in Seiberg-Witten-Floer homology, coupled with the fact that Lefschetz numbers are additive on such exact sequences.

Lemma 3.2.2. Let C_* , D_* , and E_* be chain complexes of vector spaces such that there is a short exact sequence

$$0 \to C_* \to D_* \to E_* \to 0.$$

Let

$$0 \longrightarrow C_* \longrightarrow D_* \longrightarrow E_* \longrightarrow 0$$
$$\downarrow^{T_C} \qquad \downarrow^{T_D} \qquad \downarrow^{T_E}$$
$$0 \longrightarrow C_* \longrightarrow D_* \longrightarrow E_* \longrightarrow 0$$

be an endomorphism of this short exact sequence.

Then the Lefschetz numbers of the three vertical maps satisfy the identity

$$\operatorname{Lef}(T_D) = \operatorname{Lef}(T_C) + \operatorname{Lef}(T_E).$$

Proof. As the chain complexes C_* , D_* , and E_* fit in a short exact sequence, there is a corresponding long exact sequence of their homology groups, depicted compactly as an exact triangle of graded objects where the map δ has degree -1 and the others have degree 0:



This long exact sequence splits into short exact sequences of the form

$$0 \to \operatorname{coker}(i) \to H_*(D_*) \to \operatorname{im}(p) \to 0.$$

By this, the additivity of Lefschetz numbers is reduced to the assertion that, for any short exact sequence of finite-dimensional vector spaces

$$0 \to U \to V \to W \to 0$$

and endomorphism

the traces of the vertical maps satisfy the identity

$$\operatorname{Tr}(T_V) = \operatorname{Tr}(T_U) \oplus \operatorname{Tr}(T_W).$$

This latter identity is simple linear algebra. One proof, although rather non-canonical, is to use the fact that short exact sequences of finite-dimensional vector spaces split. \Box

Fix a perturbation q for the Seiberg-Witten equations on *Y* from the residual subspace *U* of the Banach space \mathcal{P} as described in Theorem 3.1.12. Given this perturbation, *Y* admits exactly one reducible critical point [θ] in the blown-down space, and furthermore all of the critical points are non-degenerate and all of the relevant spaces of flow lines are smooth, finite-dimensional manifolds of the expected dimension.

The first short exact sequence that we will make use of is rather simple. Let C_*^s be the vector space of (gauge-equivalence classes of) boundary-stable critical points, C_*^o the vector space of interior critical points, and $\check{C}_* = C_*^s \oplus C_*^o$. Then, there is a split exact sequence

$$0 \to C^s_* \to \check{C}_* \to C^o_* \to 0$$

In order to apply Lemma 3.2.2 to some endomorphism, the spaces C_*^s and C_*^o must have finite dimension. This is clearly not true for the former, but it is for the latter.

Lemma 3.2.3. There are finitely many gauge-equivalence classes of interior critical points on Y.

Proof. Assume for the moment that the perturbation q is equal to zero. Since we have not discussed the analytic properties of the perturbations, we will only prove this lemma in the un-perturbed cse and take it for granted that the proof extends to the perturbed case. For details, see Chapter 10 of [KM07].

A blown-down configuration (B, ψ) with $\psi \neq 0$ is a critical point of the perturbed Chern-Simons-Dirac functional if and only if it satisfies the equations

$$*F_B - \rho^{-1}(\psi \otimes \psi^*)_0 = 0$$
$$D_B \psi = 0$$

The standard compactness theorem works in this case (see [Mor95]) as well, so the space of gaugeequivalence classes of solutions is compact. The linearized operator is Fredholm of index zero, so it follows that the space of solutions is a finite set of points, which implies there are finitely many interior critical points. \Box

The space C_*^s can be made finite by truncation. By the previous lemma, there exists $N \ge 0$ sufficiently large such that the \mathbb{Q} -grading of any interior critical point is less than or equal to N.

Denote by C_n^s the space of boundary-stable critical points with \mathbb{Q} -grading n, with a similar definition for C_n^o and \check{C}_n . Then $\bigoplus_{n \leq N} C_n^o = C_*^o$ and the sequence

$$0 \to \bigoplus_{n \leqslant N} C_n^s \to \bigoplus_{n \leqslant N} \check{C}_n \to C_*^o \to 0$$

is a short exact sequence of finite-dimensional vector spaces.

Setting *N* to be equal to -2n(Y) + 2k for some integer *k*, it follows that $\bigoplus_{n \leq N} C_n^s$ is a finite-dimensional vector space of dimension k + 1.

As discussed in the previous section, the cobordism *W* induces a map

$$m_o^o: C^o_* \to C^o_*$$

preserving the mod-2 grading.

It also induces a map

$$\overline{m}_s^s:\bigoplus_{n\leqslant N}C_n^s\to\bigoplus_{m\leqslant N}C_n^s$$

and a map

$$\check{m}:\bigoplus_{n\leqslant N}\check{C}_n\to\bigoplus_{n\leqslant N}\check{C}_n$$

such that the diagram

commutes.

The fact that the cobordism map induced by W preserves the mod-2 grading is, of course not trivial. It is a consequence of Proposition 25.4.3 of [KM07], which specializes in the situation of this thesis to the statement that W preserves the mod-2 grading if and only if the quantity

$$\iota(W) = \frac{1}{2}(\chi(W) + \operatorname{sign}(W))$$

is even, where χ denotes the Euler characteristic.

However, both the Euler characteristic and the signature of W vanish, which implies that W preserves the mod-2 grading.

The fact that this diagram commutes is also not at all obvious, but it is a matter of reading off the definition of \check{m} from Definition 25.3.3 of [KM07] and using the fact that Y is a homology 3-sphere to show that the appropriate differentials or cobordism chain maps vanish.

Taking Lefschetz numbers with respect to the mod-2 grading and applying Lemma 3.2.2, it follows that

$$\operatorname{Lef}(\check{m}) = \operatorname{Lef}(m_o^o) + \operatorname{Lef}(\overline{m}_s^s).$$

Lemma 3.2.4. The map \overline{m}_s^s is the identity map. Since the boundary-stable critical points all have even grading, this in turn implies that

$$\operatorname{Lef}(\overline{m}_s^s) = \dim(\bigoplus_{n \leq N} C_n^s) = \frac{N + 2n(Y)}{/2} + 1$$

Proof. This is proven in Proposition 39.1.2 of [KM07].

It follows that

$$\operatorname{Lef}(\check{m}) = \operatorname{Lef}(m_o^o) + \operatorname{Lef}(\overline{m}_s^s)$$

The definition of a "Lefschetz number" specializes naturally to the definition given before the statement of Theorem 3.2.1 for a graded endomorphism of a $\mathbb{Z}/2$ -graded vector space. This can be seen by considering the vector space as a chain complex with trivial differentials. In particular, Lemma 3.2.2 still holds for short exact sequences of $\mathbb{Z}/2$ -graded vector spaces.

There is a short exact sequence

$$0 \to \operatorname{im}(i_*) \xrightarrow{i_*} \widetilde{HM}_*(Y) \to HM_*(Y) \to 0.$$

Considered with real coefficients, these are vector spaces graded by either the \mathbb{Q} -grading or the mod-2 grading. Define $\widetilde{HM}_n(Y)$ to be the subspace of linear combinations of homology classes represented by critical points of \mathbb{Q} -grading n, with $\operatorname{im}(i_n)$ and $HM_n(Y)$ defined identically. Then, we obtain another short exact sequence of finite-dimensional, $\mathbb{Z}/2$ -graded vector spaces as before:

$$0 \to \bigoplus_{n \leqslant N} \operatorname{im}(i_n) \xrightarrow{i_*} \bigoplus_{n \leqslant N} \widecheck{H\!M}_n(Y) \to \bigoplus_{n \leqslant N} H\!M_n(Y) \to 0.$$

For sufficiently large *n*, the map $i_n : \overline{HM}_n(Y) \to \widetilde{HM}_n(Y)$ is surjective.

This is a consequence of the discussion in Section 39.1 of [KM07].

The differentials of the complex \overline{C}_* actually vanish when Y is a homology three-sphere. Therefore, it is identified with the graded vector space $\mathbb{R}[U, U^{-1}]$ where $1 \in \mathbb{R}[U, U^{-1}]$ has degree zero and multiplication by U is a map of degree -2. It is a graded module over the ring $S = \mathbb{R}[[U]]$ of formal power series in U.

Then ker (i_*) is identified with a proper *S*-submodule of $\mathbb{R}[U, U^{-1}]$. These submodules are exactly of the form U^hS for some $h \in \mathbb{Z}$, which is the submodule of $\mathbb{R}[U, U^{-1}]$ consisting of the closure of all elements with grading higher than -2h. It follows that the image of i_* has its degree bounded above, from which it can be deduced that i_n is surjective for large n.

Pick the number N from before to be large enough so that this happens, resulting in the identification

$$\bigoplus_{n \leqslant N} HM_n(Y) = HM_*(Y).$$

The cobordism maps commute with the exact triangle, so they commute with the short exact sequence:

Applying Lemma 3.2.2, it follows that

$$\operatorname{Lef}(\widetilde{HM}_*(W)) = \operatorname{Lef}(\operatorname{im}(i_*)) + \operatorname{Lef}(HM_*(W))$$

where "Lef(im(i_*))" is interpreted to be the Lefschetz number of $\widetilde{HM}_*(W)$ restricted to the truncated image

of i_* .

By definition,

$$\operatorname{Lef}(\check{m}) = \operatorname{Lef}(\check{HM}_*(W)).$$

By another application of Proposition 39.1.2 of [KM07], the cobordism map on $\overline{HM}_*(Y)$ is the identity map. Since the cobordism maps commute with the exact triangle, it follows that it is the identity map on $\operatorname{im}(i_*)$ as well.

Finally, $im(i_*)$ by definition consists of all of the boundary unstable critical points with grading greater than or equal to -2h(Y). Therefore, it follows that $Lef(im(i_*)) = \frac{N+2h(Y)}{2} + 1$.

Combining the two identities shows that

$$\frac{N+2n(Y)}{2} + 1 + \text{Lef}(m_o^o) = \frac{N+2h(Y)}{2} + 1 + \text{Lef}(HM_*(W))$$

so

$$n(Y) + \operatorname{Lef}(m_o^o) = h(Y) + \operatorname{Lef}(HM_*(W)).$$

To show Theorem 3.2.1, it remains to show that $\lambda_{SW}(X) \equiv n(Y) + \text{Lef}(m_o^o) \mod 2$. Given a regular pair (g, β) , $\lambda_{SW}(X)$ decomposes as

$$\lambda_{SW}(X) = \#\mathcal{M}(X, g, \beta) - \operatorname{ind} D^+(Z_+, g, \beta) + \operatorname{sign}(Z)/8.$$

Recall

$$n(Y) = \operatorname{ind} D^+(Z_{\infty}) - \operatorname{sign}(Z)/8.$$

This is rather similar to the correction term in $\lambda_{SW}(X)$. It would certainly prove the theorem if the equalities

$$\operatorname{ind} D^+(Z_+, g, \beta) = \operatorname{ind}_{\mathbb{C}} D^+(Z_\infty)$$

and

$$#\mathcal{M}(X,g,\beta) = -\operatorname{Lef}(m_o^o)$$

held modulo 2.

This is not a priori true, but will work if we "stretch the neck" of *X* around *Y*.

Suppose the metric g takes the form $dt^2 + h$ for some metric h on Y in a bicollar neighborhood $[-\varepsilon, \varepsilon] \times Y \subset X$.

This assumption gives a "neck" around *Y* that can be stretched.

Given this, define for any R > 0 the manifold X_R constructed by cutting out $[-\varepsilon, \varepsilon] \times Y$ and gluing in the cylinder $[-R, R] \times Y$. The metric g on $X \setminus ([-\varepsilon, \varepsilon] \times Y)$ glues together with the cylindrical metric $dt^2 + h$ on $[-R, R] \times Y$ to form a metric g_R . Similarly, define W_R to be the cobordism from Y to itself obtained by cutting open X_R along $\{0\} \times Y$. Equivalently, W_R is equal to the gluing

$$[-R,0] \times Y \smile W \smile [0,R] \times Y.$$

Set Z_R to be the gluing

$$Z \smile W_{R,0} \smile W_{R,1} \smile \ldots$$

with $W_{R,i} = W_R$ for every *i*.

Metrics and spin structures are induced on W_R and Z_R from g_R in the expected way.

The manifold X_R is clearly diffeomorphic to X, so it follows that $\lambda_{SW}(X) = \lambda_{SW}(X_R)$ for any R.

Denote by W_{∞} the manifold obtained by attaching two tubular ends $[0, \infty) \times Y$ to the boundary components of Y, oriented in the expected manner. Since the metric g is cylindrical near Y, it follows that the metric on W is cylindrical near the boundary, from which it follows that there is a natural metric on W_{∞} as well.

As $R \to \infty$, the manifold X_R is "asymptotic" to W_{∞} . This can be seen as another motivation for the splitting formula. The solutions to the Seiberg-Witten equations on X_R should (and will, given some care) for large R behave like solutions to the Seiberg-Witten equations on W_{∞} that are asymptotic to the same critical point on both ends. If the critical point is not reducible, then the appropriately oriented count of such solutions on W_{∞} is exactly the Lefschetz number Lef (m_{α}^o) .

Similarly, as $R \to \infty$, the end-periodic manifold Z_R is "asymptotic" to the manifold Z_{∞} . In this case, one would expect the indices of the relevant Dirac operators to agree for sufficiently large R. Putting this together with the above would yield the splitting formula for $\lambda_{SW}(X_R)$ for sufficiently large R.

Making this intuition precise, however, requires a host of technical assumptions on the metrics and perturbations used.

First, we will require that the metric g satisfies two additional assumptions:

1. The Dirac operator $D^+(Y)$ with respect to the metric *h* has no kernel.

2. The Dirac operator $D^+(W_{\infty})$ with respect to the metric induced by g is invertible.

The existence of generic metrics satisfying both of these properties in addition to the first assumption is shown in Section 10 of [LRS17].

Next, observe that there are two distinct ways that one can perturb the Seiberg-Witten equations on the manifold X_R . The first way is the one described in the last chapter, perturbing the self-dual part of the curvature of a solution by $d^+\beta$ for an imaginary-valued one-form β .

The second way is a perturbation in the "Floer homology" style described in the previous subsection. Identify a collar neighborhood of the boundary of the cobordism W with $(-\varepsilon, 0) \times (\overline{Y} \cup Y)$. Then, take two nice perturbations q and \mathfrak{p}_0 for the Chern-Simons-Dirac functional as in Theorem 3.1.12.

As mentioned briefly before the introduction of Theorem 3.1.12, these induce perturbations \hat{q} and \hat{p}_0 for the Seiberg-Witten equations on the collar neighborhood.

Now let β be a smooth function on W that is equal to 1 on the boundary and vanishes outside of the collar neighborhood, and let β_0 be a smooth function on W that has compact support inside the collar neighborhood. Then, one can form the perturbation

$$\hat{\mathfrak{p}} = \beta \hat{\mathfrak{q}} + \beta_0 \hat{\mathfrak{p}}_0.$$

By definition, \hat{p} can be considered a perturbation of the Seiberg-Witten equations on *W* that vanishes outside of the collar neighborhood.

Now the manifold X_R is equal to the gluing of W with the cylinder $[-R, R] \times Y$. Therefore, one can form a perturbation to the Seiberg-Witten equations on X_R by gluing \hat{p} with the perturbation \hat{q} on $[-R, R] \times Y$. This perturbation is denoted by \mathfrak{p}_R .

For sufficiently small perturbations, the moduli space of solutions

$$\mathcal{M}(X_R, g_R, \hat{p}_R)$$

to the blown-up Seiberg-Witten equations is a compact, finite-dimensional manifold. The argument for this is outlined in Section 9 of [MRS11].

Recall that $\mathcal{Z} \subset \mathcal{B}^{\sigma}(X_R)$ is the space of all configurations (A, s, φ) such that $D_A^+ \varphi = 0$. Then, the solutions to the *unperturbed* Seiberg-Witten equations are the zero set of the operator $\chi : \mathcal{Z} \to \Omega^2_+(X_R; i\mathbb{R})$ sending (A, s, φ) to $F_A^+ - s^2 \rho^{-1}(\varphi \otimes \varphi^*)_0$.

This can be phrased differently. Let \mathcal{V} be the trivial bundle over \mathcal{Z} with fiber $\Omega^2_+(X_R; i\mathbb{R}) \oplus L^2_k(X_R; S^-)$ at every point. Then χ paired with the zero section on the second summand in the fiber is a section of \mathcal{V} . We may perturb χ by some small section γ such that χ is transverse to the zero section, so the resulting moduli space

$$\mathcal{M}(X_R, g_R, \gamma) = (\chi + \gamma)^{-1}(0)$$

is a manifold.

Both the perturbations β and \mathfrak{p}_R are special cases of these types of perturbations. In particular, the former is simply a constant section of \mathcal{V} . If both are sufficiently small, however, then it is shown in [MRS11] that they produce the same spaces of solutions.

Lemma 3.2.5. ([MRS11], Lemma 9.4) For a generic metric g_R and sufficiently small generic perturbations β and $\hat{\mathfrak{p}}_R$, the moduli spaces $\mathcal{M}(X_R, g_R, \beta)$ and $\mathcal{M}(X_R, g_R, \hat{\mathfrak{p}}_R)$ are in bijective correspondence.

Lemma 9.4 of [MRS11] only considers the case where a perturbation γ takes values of zero in its spinorial part, but the exact same proof works for the more general case under consideration.

The other point that must be addressed is the index-theoretic term in the definition of $\lambda_{SW}(X_R)$. We will only be able to calculate the index of the *unperturbed* Dirac operator $D^+(Z_R)$, but the index-theoretic term in the invariant $\lambda_{SW}(X_R)$ is the index of a *perturbed* end-periodic Dirac operator. Luckily, it is well-known that the Fredholm index is stable under bounded perturbations of small operator norm (see Theorem 5.17 in Chapter IV of [Kat66] for a general statement).

Therefore, if the metric is chosen properly such that this operator is Fredholm, which we are assuming to be the case, then $D^+(Z_R) + \rho(\beta)$ will be Fredholm with index equal to that of $D^+(Z_R)$.

This allows us to re-define the invariant $\lambda_{SW}(X_R)$ in the desired manner. First, pick a regular pair (g_R, β) such that g_R satisfies all the required assumptions and β is sufficiently small. Then one can write

$$\lambda_{SW}(X_R) = \#\mathcal{M}(X_R, g_R, \beta) - \operatorname{ind} D^+(Z_R, g_R, \beta) + \operatorname{sign}(Z)/8.$$

However, because β is small, we have the equalities

$$#\mathcal{M}(X_R, g_R, \beta) = #\mathcal{M}(X_R, g_R, \hat{\mathfrak{p}}_R)$$

for small $\hat{\mathfrak{p}}_R$ and

ind
$$D^+(Z_R, g_R, \beta) = \operatorname{ind} D^+(Z_R, g_R)$$
.

It follows that

$$\lambda_{SW}(X_R) = \#\mathcal{M}(X_R, g_R, \mathfrak{p}_R) - \operatorname{ind} D^+(Z_R, g_R) + \operatorname{sign}(Z)/8$$

Now, the proof outline for the splitting formula given above rests on more rigorous footing.

3.3 The index calculation

In this section, we will show the following theorem. Note that the proof is rather technical, so the reader may wish to skip this section on a first reading.

Theorem 3.3.1. For sufficiently large R > 0,

$$\operatorname{ind} D^+(Z_R) = \operatorname{ind} D^+(Z_\infty).$$

There are two ways that this index is computed in [LRS17]. The first method, which is the one we will discuss, is to repeatedly apply a lemma that allows one to add and remove "redundant" parts of a Fredholm operator while preserving its index. The second method uses the index theorem for end-periodic operators derived in [cite this paper], which is analogous to the work of [APS75] in that it expresses the indices of these operators as a topological term corrected by an "end-periodic eta invariant". It is shown that, as $R \to \infty$, the end-periodic eta invariant approaches the standard Atiyah-Patodi-Singer eta invariant. Since both quantities are integers, it follows that they become equal for sufficiently large *R*.

The central lemma of this section is the following.

Lemma 3.3.2. Let \mathcal{H}_1 , \mathcal{H}_2 , and \mathcal{H}_3 be Hilbert spaces.

Let $F : \mathcal{H}_1 \to \mathcal{H}_2$ and $G : \mathcal{H}_1 \to \mathcal{H}_3$ be operators such that G is surjective and the operator $F \oplus G : \mathcal{H}_1 \to \mathcal{H}_2 \oplus \mathcal{H}_3$ is Fredholm.

Then the restriction $F|_{\ker G}$: ker $G \to \mathcal{H}_2$ is Fredholm, and ind $F = \operatorname{ind} F \oplus G$.

Proof. Write $\widetilde{F} = F|_{\ker G}$.

First, we can identify the kernel of \tilde{F} with that of $F \oplus G$. It is clear that any $v \in \ker \tilde{F}$ lies in $\ker(F \oplus G)$. On the other hand, if $v \in \ker(F \oplus G)$, that implies both F(v) and G(v) = 0, so $v \in \ker \tilde{F}$ as desired.

Next, we can identify the kernel of the adjoint \tilde{F}^* with that of $(F \oplus G)^*$. If $\tilde{F}^*(v) = 0$, then it is certainly the case that

$$(F \oplus G)^*(v,0) = F^*(v,0) = F^*(v) = 0$$

On the other hand, if $(F \oplus G)^*(v, w) = 0$, this is equivalent to $\tilde{F}^*(v) + G^*(0, w) = 0$. Since $\tilde{F}^*(v)$ lies in $\ker(G)$, it is orthogonal to $G^*(0, w)$. It follows that both vectors vanish. However, since G is surjective, G^* is injective, and so w = 0. This is sufficient to identify the kernels of the adjoint operators.

Essentially, this lemma states that one can remove the "surjective" part of any Fredholm operator while retaining the same index.

The proof of Theorem 3.3.1 will follow from repeated application of this lemma, but also depends heavily on the assumptions made on the metric that were described in the last section. It is necessary to have a good understanding of the various implications of these assumptions.

3.3.1 The Dirac operator and APS boundary conditions

Consider the half-cylinder $\mathbb{R}_{\leq 0} \times Y$ with the metric $dt^2 + h$. The spinor bundle on the half-cylinder can be identified with the direct sum of two copies of the spinor bundle $S \to Y$, pulled back by the projection onto the *Y* factor. Therefore, a spinor on the half-cylinder can be regarded as a time-dependent spinor on *Y* over the time interval $[0, \infty)$.

Under this identification, the Dirac operator $D^+(\mathbb{R}_{\leq 0} \times Y)$ takes the form of the differential operator $\frac{d}{dt} + D(Y)$. For a path of spinors $\varphi(t)$ on $\mathbb{R}_{\leq 0} \times Y$, one writes

$$D^{+}(\mathbb{R}_{\leq 0} \times Y)(\varphi(t)) = \frac{d\varphi}{dt}(t) + D(Y)(\varphi(t)).$$

Suppose we are trying to solve the problem

$$D^+(\mathbb{R}_{\leq 0} \times Y)(\varphi(t)) = \psi(t)$$

for smooth, compactly supported spinors $\varphi(t)$, $\psi(t)$.

Recall that the operator D(Y) has no kernel. Furthermore, it is self-adjoint, so it has a discrete real spectrum $\operatorname{Spec}(D(Y))$ and the space of spinors on Y admits an L^2 -orthonormal basis of eigenvectors $\{\varphi_{\lambda}\}_{\lambda \in \operatorname{Spec}(D(Y))}$. The eigenvector φ_{λ} has eigenvalue $\lambda \in \operatorname{Spec}(D(Y))$.

It follows that there are tuples of smooth functions $\{f_{\lambda}\}, \{g_{\lambda}\}$ on the half-cylinder such that

$$\varphi(t) = \sum_{\lambda \in \operatorname{Spec}(D(Y))} f_{\lambda}(t) \varphi_{\lambda}$$

and

$$\psi(t) = \sum_{\lambda \in \operatorname{Spec}(D(Y))} g_{\lambda}(t) \varphi_{\lambda}.$$

Expanding out the Dirac operator, it follows that

$$D^{+}(\mathbb{R}_{\leq 0} \times Y)(\varphi(t)) = \sum_{\lambda} D^{+}(\mathbb{R}_{\leq 0} \times Y)(f_{\lambda}(t)\varphi_{\lambda})$$
$$= \sum_{\lambda} (\frac{d}{dt} + D(Y))(f_{\lambda}(t)\varphi_{\lambda})$$
$$= \sum_{\lambda} (\frac{df_{\lambda}}{dt}(t) + \lambda f_{\lambda}(t))\varphi_{\lambda}.$$

Equating the coefficients of φ_{λ} on both sides, this equation reduces to a collection of ODEs of the form

$$\frac{df_{\lambda}}{dt}(t) + \lambda f_{\lambda}(t) = g_{\lambda}(t).$$

If $\lambda > 0$, this can be solved explicitly by the reader's favorite method:

$$f_{\lambda}(t) = \int_{-\infty}^{t} e^{\lambda(s-t)} g_{\lambda}(s) ds.$$

However, if $\lambda \leq 0$, this function may blow up as *t* approaches zero. Instead, one defines

$$f_{\lambda}(t) = -\int_{t}^{0} e^{\lambda(s-t)} g_{\lambda}(s) ds.$$

Note that this solution by default fixes the value $f_{\lambda}(0) = 0$ for any $\lambda \leq 0$. In other words, the spinor $\varphi(t)$ satisfies a *spectral boundary condition*, which in this case means its restriction to the boundary of the half-cylinder must lie in the span of the *negative* eigenspaces of D(Y).

Therefore, one sees that when working with these types of problems involving differential operators on manifolds with boundary, it is necessary to fix boundary conditions to make the problem well-posed. Without boundary conditions on its domain, the Dirac operator on the half-cylinder does not admit a left inverse.

The work of Atiyah-Patodi-Singer in [APS75] extended this type of work to (and beyond) the more general situation of the Dirac operator on a Riemannian manifold with a cylindrical neighborhood around its boundary. In our case, the manifold for which we would like to understand better the Dirac operator is the cobordism W, with oriented boundary $\overline{Y} \cup Y$. The Dirac operator $D^+(W)$ is not a priori Fredholm: it has a finite-dimensional cokernel, but an infinite-dimensional kernel. This can be regarded as a generalization of the failure of the Dirac operator of the half-cylinder to admit a left inverse without boundary conditions.

Therefore, one must impose boundary conditions. Identify a neighborhood of the boundary of W with the cylinder $(-\varepsilon, 0] \times (\overline{Y} \cup Y)$. As is convention, this identification takes $\{0\} \times (\overline{Y} \cup Y)$ to the boundary of W. The Dirac operator on \overline{Y} is given by $D(\overline{Y}) = -D(Y)$.

In the toy example above, we examined the Dirac operator between spaces of smooth sections. However, in our situation, the Dirac operator acts between L^2 Hilbert spaces of sections. In this case, sections still have "boundary values", but some Sobolev regularity is lost.

Theorem 3.3.3. ([KM07], Theorem 17.1.1) There is a continuous, surjective restriction map

$$L^{2}_{1}(W; S^{+}) \rightarrow L^{2}_{1/2}(\overline{Y} \cup Y; S) = L^{2}_{1/2}(\overline{Y}; S) \oplus L^{2}_{1/2}(Y; S)$$

It is a clear corollary that restriction to either of the boundary components is also continuous and surjective. For a spinor $\varphi \in L^2_1(W; S^+)$, its restriction to \overline{Y} or Y will be denoted by $r^W_-(\varphi)$ and $r^W_+(\varphi)$ respectively.

Since the Dirac operator D(Y) is self-adjoint and has no kernel, there is an orthogonal splitting

$$L^2_{1/2}(Y;S) = \mathcal{K}_+(Y) \oplus \mathcal{K}_-(Y)$$

where $\mathcal{K}_{\pm}(Y)$ denote the spans of the positive and negative eigenspaces respectively.

Write $\Pi_{\pm} : L^2_{1/2}(Y;S) \to \mathcal{K}_{\pm}(Y)$ for the orthogonal projections onto these eigenspaces. These act as projections on the space $L^2_{1/2}(\overline{Y};S)$ since it is canonically identified with $L^2_{1/2}(Y;S)$. However, the Dirac operator switches sign when the orientation of *Y* is reversed, so $\mathcal{K}_+(Y)$ is the span of the *negative* eigenspaces of $D(\overline{Y})$ and vice versa.

It is with this, along with the previous discussion in mind that we introduce the *Dirac operator with boundary conditions*:

$$(D^{+}(W), \Pi_{+} \circ r_{-}^{W}, \Pi_{-} \circ r_{+}^{W}) : L^{2}_{1}(W; S^{+}) \to L^{2}(W; S^{-}) \oplus \mathcal{K}_{+}(Y) \oplus \mathcal{K}_{-}(Y).$$

This is an example of the type of operator that is discussed in the first part of [APS75]. As such, the theoretical results in this paper apply here, and the following theorem is a direct consequence of Proposition 3.11 in the first paper of [APS75], as well as the fact that D(Y) has no kernel.

Theorem 3.3.4. The kernel and cokernel of the Dirac operator with boundary conditions are isomorphic to the kernel and cokernel of $D^+(W_{\infty})$. In other words, the operator $(D^+(W), \Pi_+ \circ r_-^W, \Pi_- \circ r_+^W)$ is an isomorphism.

It will also be useful to understand the Dirac operator on a finite cylinder $C = [-R, R] \times Y$. Write \overline{Y} for the slice $\{-R\} \times Y$ and Y for $\{R\} \times Y$. Here the Sobolev restriction maps are denoted by r_+^C , and are still

surjective.

In this situation, one can solve explicitly for the kernel of $D^+(C)$ using the method above. If $\varphi \in \ker D^+(C)$, then one may decompose $\varphi(t)$ as a time-dependent linear combination of eigenvectors $\sum_{\lambda} f_{\lambda}(t)\varphi_{\lambda}$.

Then the equation $D^+(C)(\varphi)=0$ is equivalent to a system of ODEs

$$\left(\frac{d}{dt} + \lambda\right)f_{\lambda}(t) = 0.$$

Choose for each f_{λ} an initial value $c_{\lambda} = f_{\lambda}(0)$. Then, the ODE has the unique solution

$$f_{\lambda}(t) = c_{\lambda} e^{-t\lambda}$$

and so

$$\varphi(t) = \sum_{\lambda} c_{\lambda} e^{-t\lambda} \varphi_{\lambda}.$$

This derivation implies the following theorem.

Theorem 3.3.5. The Dirac operator $D^+(C): L^2_1(C; S^+) \to L^2(C; S^-)$ is surjective. Furthermore, let

$$(D^+(C), \Pi_+ \circ r_-^C, \Pi_- \circ r_+^C) : L^2_1(C; S^+) \to L^2(C; S^-) \oplus \mathcal{K}_+(Y) \oplus \mathcal{K}_-(Y)$$

be the Dirac operator with boundary conditions on C. Then, the restriction

$$(\Pi_+ \circ r^C_-, \Pi_- \circ r^C_+) : \ker D^+(C) \to \mathcal{K}_+(Y) \oplus \mathcal{K}_-(Y)$$

is an isomorphism.

Proof. The kernel and cokernel of the Dirac operator with boundary conditions are identified with that of the Dirac operator on the cylinder $\mathbb{R} \times Y$, again using Proposition 3.11 of the first paper of [APS75] and the fact that D(Y) has no kernel. This operator is certainly surjective, as our construction of a right inverse for the operator on the half-cylinder at the beginning of the section works in this case as well.

This implies the first assertion, and the fact that $(\Pi_+ \circ r_-^C, \Pi_- \circ r_+^C)$ is surjective with closed range.

The second assertion then follows from our explicit construction of the kernel of $D^+(C)$ given sets of initial values $\{c_{\lambda}\}$. As long at least one of the c_{λ} is nonzero, it follows that the boundary values $\varphi(-R)$ and $\varphi(R)$ are nonzero, which shows that the map $(\Pi_+ \circ r_-^C, \Pi_- \circ r_+^C)$ is injective.

3.3.2 **Proof of the main theorem**

We are now ready to prove the main theorem. The presentation here is almost verbatim the presentation given in [LRS17], with some changes in notation and explanations given at various steps in the process. The calculation will require some careful book-keeping of the different operators involved, as well as technical input from the results of the previous subsection. Let W_i denote the cobordism W for every i > 0. Let C_i denote the cylinder $[-R, 0] \times Y$ for i = 0 and $[-R, R] \times Y$ for i > 0.

The manifold X_R is by definition the gluing $W \cup [-R, R] \times Y$.

Therefore, gluing along boundary components, one may write

$$Z_R = Z \smile_{Y_{0,+}} C_0 \smile_{Y_{1,-}} W_1 \smile_{Y_{1,+}} C_1 \smile_{Y_{2,-}} W_2 \smile_{Y_{2,+}} \dots$$
Here, the notation is set such that $Y_{i,-}$ and $Y_{i,+}$ denote the boundary components corresponding to \overline{Y} and Y in W_i for all $i \ge 1$. The manifold $Y_{0,+}$ is the boundary of Z, and the cylinders C_i have boundary $Y_{i,+}$ and $Y_{i+1,-}$ for all $i \ge 0$. For any of the W_i , write r_{\pm}^W for the restrictions to $Y_{i,\pm}$. For the C_i , write r_{-}^C for the restriction to the "left" boundary component $Y_{i,+}$ and r_{+}^C for the restriction to the "right" boundary component $Y_{i,+}$.

The restriction from Z to its boundary is similarly denoted by $r_+^Z : L_1^2(Z; S^+) \to L_{1/2}^2(Y_{0,+}; S)$. The Dirac operator $D^+(Z_R)$ is an operator from $L_1^2(Z_R; S^+)$ to $L^2(Z_R; S^-)$.

The latter space decomposes into a direct sum of the spaces of spinors over its components:

$$L^2(Z_R;S^-) \simeq L^2(Z;S^-) \oplus (\bigoplus_{i \ge 0} L^2(C_i;S^-)) \oplus (\bigoplus_{j \ge 1} L^2(W_j;S^-)).$$

The former space decomposes into tuples of spinors on each of the components, but these spinors are now required to agree along the manifolds Y_i^{\pm} . That is, $L_1^2(Z_R; S^+)$ is isomorphic to the space of elements

$$\alpha \oplus (\psi_i)_{i \ge 0} \oplus (\varphi_j)_{j \ge 1} \in L^2_1(Z; S^+) \oplus (\bigoplus_{i \ge 0} L^2_1(C_i; S^+)) \oplus (\bigoplus_{j \ge 1} L^2_1(W_j; S^+))$$

such that

$$\begin{aligned} r^{Z}_{+}(\alpha) &= r^{C}_{-}(\psi_{0}), \\ r^{C}_{+}(\psi_{i}) &= r^{W}_{-}(\varphi_{i+1}) \\ r^{C}_{-}(\psi_{j})(Y^{+}_{i}) &= r^{W}_{+}(\varphi_{j}) \end{aligned}$$

for every $i \ge 0$, $j \ge 1$. A quick proof of such a statement may be deduced from Lemma 3 of [Man07]. A consequence of this lemma is that any element of this space must lie in $L^2_{1,loc}(Z_R; S^+)$. Therefore, this space is isomorphic to the subspace of elements in $L^2_{1,loc}(Z_R; S^+)$ with finite L^2_1 -norm, which is exactly the space $L^2_1(Z_R; S^+)$.

It follows that $L_1^2(Z_R; S^+)$ is isomorphic to the kernel of the map

$$\begin{aligned} \mathbf{R} : L_1^2(Z; S^+) \oplus (\bigoplus_{i \ge 0} L_1^2(C_i; S^+)) \oplus (\bigoplus_{j \ge 1} L_1^2(W_j; S^+)) \\ \to L_{1/2}^2(Y_{0,+}; S) \oplus (\bigoplus_{i \ge 1} L_{1/2}^2(Y_{i,+}; S)) \oplus (\bigoplus_{j \ge 1} L_{1/2}^2(Y_{j,-}; S)) \end{aligned}$$

that sends

$$\alpha \oplus (\psi_i)_{i \ge 0} \oplus (\varphi_j)_{j \ge 1}$$

to

$$(r_{+}^{Z}(\alpha) - r_{-}^{C}(\psi_{0})) \oplus (r_{+}^{C}(\psi_{i}) - r_{-}^{W}(\varphi_{i+1}))_{i \ge 0} \oplus (r_{-}^{C}(\psi_{j}) - r_{+}^{W}(\varphi_{j}))_{j \ge 0}$$

This map is certainly surjective by application of Theorem 3.3.3.

The Dirac operator $D^+(Z_R)$ is isomorphic to the restriction to ker **R** of the operator

$$D_0: L_1^2(Z; S^+) \oplus (\bigoplus_{i \ge 0} L_1^2(C_i; S^+)) \oplus (\bigoplus_{j \ge 1} L_1^2(W_j; S^+)) \to L^2(Z; S^-) \oplus (\bigoplus_{i \ge 0} L^2(C_i; S^-))$$
$$\oplus (\bigoplus_{j \ge 1} L^2(W_j; S^-))$$

that sends

$$\alpha \oplus (\psi_i)_{i \ge 0} \oplus (\varphi_j)_{j \ge 1}$$

to

$$D^+(Z)(\alpha) \oplus (D^+(C_i)(\psi_i))_{i \ge 0} \oplus (D^+(W_j)(\varphi_j))_{j \ge 1}$$

Define D_1 to be the operator $D_0 \oplus \mathbf{R}$. Applying Lemma 2.3.7, it follows that D_1 is Fredholm and has index equal to that of $D^+(Z_R)$.

Note that D_1 is a direct sum of the operators

$$\alpha \oplus (\psi_i)_{i \ge 0} \oplus (\varphi_j)_{j \ge 1} \mapsto (D^+(C_i)(\psi_i))_{i \ge 0}$$

and the operator

$$\begin{split} D_2: L^2_1(Z; S^+) \oplus (\bigoplus_{i \ge 0} L^2_1(C_i; S^+)) \oplus (\bigoplus_{j \ge 1} L^2_1(W_j; S^+)) \to L^2(Z; S^-) \oplus (\bigoplus_{j \ge 1} L^2(W_j; S^-)) \\ & \oplus L^2_{1/2}(Y_{0,+}; S) \\ & \oplus (\bigoplus_{i \ge 1} L^2_{1/2}(Y_{i,+}; S)) \\ & \oplus (\bigoplus_{j \ge 1} L^2_{1/2}(Y_{i,+}; S)) \end{split}$$

sending

$$\alpha \oplus (\psi_i)_{i \ge 0} \oplus (\varphi_j)_{j \ge 1}$$

to

$$D^{+}(Z)(\alpha) \oplus (D^{+}(W_{j})(\varphi_{j}))_{j \geq 1} \oplus (r^{Z}_{+}(\alpha) - r^{C}_{-}(\psi_{0}))$$
$$\oplus (r^{C}_{+}(\psi_{i}) - r^{W}_{-}(\varphi_{i+1}))_{i \geq 0}$$
$$\oplus (r^{C}_{-}(\psi_{j}) - r^{W}_{+}(\varphi_{j}))_{j \geq 0}.$$

By Theorem 3.3.5, the first operator is surjective, so we can apply Lemma 2.3.7 once more. Restrict D_2 to the kernel of the first operator. This restriction will also be written as D_2 , with domain and range specified appropriately:

$$\begin{split} D_2: L_1^2(Z; S^+) \oplus (\bigoplus_{i \ge 0} \ker(D^+(C_i))) \oplus (\bigoplus_{j \ge 1} L_1^2(W_j; S^+)) \to L^2(Z; S^-) \oplus (\bigoplus_{j \ge 1} L^2(W_j; S^-)) \\ & \oplus L_{1/2}^2(Y_{0,+}; S) \\ & \oplus (\bigoplus_{i \ge 1} L_{1/2}^2(Y_{i,+}; S)) \\ & \oplus (\bigoplus_{j \ge 1} L_{1/2}^2(Y_{j,-}; S)) \end{split}$$

Then, Lemma 2.3.7 states that D_2 is Fredholm and has index equal to that of D_1 . Next, we apply the second assertion of 3.3.5. This requires first projecting the boundary values onto the two spectral subspaces

of D(Y), casting D_2 as the isomorphic operator

$$D_{3}: L_{1}^{2}(Z; S^{+}) \oplus (\bigoplus_{i \ge 0} \ker(D^{+}(C_{i}))) \oplus (\bigoplus_{j \ge 1} L_{1}^{2}(W_{j}; S^{+})) \to L^{2}(Z; S^{-}) \oplus (\bigoplus_{j \ge 1} L^{2}(W_{j}; S^{-}))$$
$$\oplus \mathcal{K}_{+}(Y_{0,+}) \oplus \mathcal{K}_{-}(Y_{0,+})$$
$$\oplus (\bigoplus_{i \ge 1} \mathcal{K}_{+}(Y_{i,+}) \oplus \mathcal{K}_{-}(Y_{i,+}))$$
$$\oplus (\bigoplus_{j \ge 1} \mathcal{K}_{+}(Y_{j,-}) \oplus \mathcal{K}_{-}(Y_{j,-}))$$

sending

$$\alpha \oplus (\psi_i)_{i \ge 0} \oplus (\varphi_j)_{j \ge 1}$$

to

$$D^{+}(Z)(\alpha) \oplus (D^{+}(W_{j})(\varphi_{j}))_{j \ge 1} \oplus (\Pi_{+}(r_{+}^{Z}(\alpha) - r_{-}^{C}(\psi_{0}))) \\ \oplus (\Pi_{-}(r_{+}^{Z}(\alpha) - r_{-}^{C}(\psi_{0}))) \\ \oplus (\Pi_{+}(r_{+}^{C}(\psi_{i}) - r_{-}^{W}(\varphi_{i+1})))_{i \ge 0} \\ \oplus (\Pi_{-}(r_{+}^{C}(\psi_{i}) - r_{-}^{W}(\varphi_{i+1})))_{i \ge 0} \\ \oplus (\Pi_{+}(r_{-}^{C}(\psi_{j}) - r_{+}^{W}(\varphi_{j})))_{j \ge 1} \\ \oplus (\Pi_{-}(r_{-}^{C}(\psi_{j}) - r_{+}^{W}(\varphi_{j})))_{j \ge 1}.$$

Recall the calculation of the kernel of the Dirac operator on the finite cylinder *C*. For an element φ in this kernel, write down its eigenfunction decomposition $\varphi(t) = \sum_{\lambda} f_{\lambda}(t)\varphi_{\lambda}$. Then one finds

$$\varphi(-R) = \sum_{\lambda} e^{\lambda R} f_{\lambda}(0) \varphi_{\lambda}$$

and

$$\varphi(R) = \sum_{\lambda} e^{-\lambda R} f_{\lambda}(0) \varphi_{\lambda}.$$

In general, one finds for any times $t_1, t_2 \in [-R, R]$ with $t_1 \leq t_2$ that

$$\varphi(t_2) = \sum_{\lambda} e^{-\lambda(t_2 - t_1)} f_{\lambda}(0) \varphi_{\lambda}.$$

A common shorthand for this kind of expression is

$$\varphi(t_2) = e^{(t_2 - t_1)D(Y)}\varphi(t_1).$$

The operator $e^{(t_2-t_1)D(Y)}$ commutes with the projections Π_{\pm} , and so preserves the eigenspaces $\mathcal{K}_{\pm}(Y)$. In particular, to reconstruct the entire spinor $\varphi(t)$, one needs only know $\Pi_+(\varphi(-R))$ and $\Pi_-(\varphi(R))$. Then $\varphi(t)$ is written as

$$\varphi(t) = e^{(t+R)D(Y)} \Pi_{+}(\varphi(-R)) + e^{(t-R)D(Y)} \Pi_{-}(\varphi(R)).$$

This implies the map

$$\mathcal{K}_+(Y) \oplus \mathcal{K}_-(Y) \to \ker(D^+(C))$$

defined by

$$(\eta_+,\eta_-)\mapsto\varphi(t)=e^{(t+R)D(Y)}\eta_++e^{(t-R)D(Y)}\eta_-$$

is an isomorphism, as it forms an inverse to the isomorphism

$$\ker(D^+(C)) \to \mathcal{K}_+(Y) \oplus \mathcal{K}_-(Y)$$

defined in Theorem 3.3.5.

Composing this with the map

$$\ker(D^+(C)) \to \mathcal{K}_+(Y) \oplus \mathcal{K}_-(Y)$$

defined by

$$\varphi(t) \mapsto (\Pi_+(\varphi(R)), \Pi_-(\varphi(-R)))$$

yields a map

$$\mathcal{K}_+(Y) \oplus \mathcal{K}_-(Y) \to \mathcal{K}_+(Y) \oplus \mathcal{K}_-(Y)$$

defined by

$$(\eta_+, \eta_-) \mapsto (e^{2RD(Y)}\eta_+, e^{-2RD(Y)}\eta_-)$$

Now we can remove the ker $(D^+(C_i))$ -summands in the domain of D_3 , as through this isomorphism described above D_3 is itself isomorphic to an operator

$$\begin{split} D_4: L_1^2(Z; S^+) \oplus (\bigoplus_{i \ge 0} \mathcal{K}_+(Y_{i,+}) \oplus \mathcal{K}_-(Y_{i+1,-})) \oplus (\bigoplus_{j \ge 1} L_1^2(W_j; S^+)) \to L^2(Z; S^-) \oplus (\bigoplus_{j \ge 1} L^2(W_j; S^-)) \\ & \oplus \mathcal{K}_+(Y_{0,+}) \oplus \mathcal{K}_-(Y_{0,+}) \\ & \oplus (\bigoplus_{i \ge 1} \mathcal{K}_+(Y_{i,+}) \oplus \mathcal{K}_-(Y_{i,+})) \\ & \oplus (\bigoplus_{j \ge 1} \mathcal{K}_+(Y_{j,-}) \oplus \mathcal{K}_-(Y_{j,-})). \end{split}$$

The notation is somewhat difficult to parse through at this point, but this operator is can be expressed explicitly by replacing each ψ_i with a pair $(\eta_{i,+}, \eta_{i+1,-})$ and adjusting the map accordingly. The restriction $\Pi_+(r^C_-(\psi_i))$ is replaced with $\eta_{i,+}$ while $\Pi_-(r^C_-(\psi_i))$ is replaced with $e^{-2RD(Y)}\eta_{i+1,-}$. Similarly, the restriction $\Pi_+(r^C_+(\psi_i))$ is replaced with $e^{2RD(Y)}\eta_{i,+}$ while $\Pi_-(r^C_+(\psi_i))$ is replaced with $\eta_{i+1,-}$.

The operator D_4 is defined as the map sending

$$\alpha \oplus (\eta_{i,+}, \eta_{i+1,-})_{i \ge 0} \oplus (\varphi_j)_{j \ge 1}$$

$$D^{+}(Z)(\alpha) \oplus (D^{+}(W_{j})(\varphi_{j}))_{j \geq 1} \oplus (\Pi_{+}(r_{+}^{Z}(\alpha)) - \eta_{0,+}) \\ \oplus (\Pi_{-}(r_{+}^{Z}(\alpha)) - e^{-2RD(Y)}\eta_{1,-}) \\ \oplus (e^{2RD(Y)}\eta_{i,+} - \Pi_{+}(r_{-}^{W}(\varphi_{i+1})))_{i \geq 0} \\ \oplus (\eta_{i+1,-} - \Pi_{-}(r_{-}^{W}(\varphi_{i+1})))_{i \geq 0} \\ \oplus (\eta_{j,+} - \Pi_{+}(r_{+}^{W}(\varphi_{j})))_{j \geq 1} \\ \oplus (e^{-2RD(Y)}\eta_{j+1,-} - \Pi_{-}(r_{+}^{W}(\varphi_{j})))_{j \geq 1}.$$

Consider the operator sending

$$\alpha \oplus (\eta_{i,+}, \eta_{i+1,-})_{i \ge 0} \oplus (\varphi_j)_{j \ge 1}$$

to

.

$$(\Pi_{+}(r_{+}^{Z}(\alpha)) - \eta_{0,+}) \oplus (\eta_{i+1,-} - \Pi_{-}(r_{-}^{W}(\varphi_{i+1})))_{i \ge 0} \oplus (\eta_{j,+} - \Pi_{+}(r_{+}^{W}(\varphi_{j})))_{j \ge 1}$$

This is clearly surjective when restricting to the locus of tuples in the domain satisfying $\alpha = 0$, $\varphi_j = 0$ for every $j \ge 1$, so the operator is surjective overall. Therefore, we may apply Lemma 2.3.7 again to obtain an operator D_5 that is Fredholm and has the same index as D_4 . Note that our expressions have finally started becoming simpler, since restricting to the kernel of the above operator implies $\Pi_+(r_+^Z(\alpha)) = \eta_{0,+}$, $\Pi_-(r_-^W(\varphi_{i+1})) = \eta_{i,-}$ for every $i \ge 1$ and $\Pi_+(r_+^W(\varphi_j)) = \eta_{j,+}$ for every $j \ge 1$.

Then after switching some signs, D_5 is written as the map

$$D_{5}: L_{1}^{2}(Z; S^{+}) \oplus (\bigoplus_{i \ge 0} \mathcal{K}_{+}(Y_{i,+}) \oplus \mathcal{K}_{-}(Y_{i+1,-})) \oplus (\bigoplus_{j \ge 1} L_{1}^{2}(W^{j}; S^{+})) \to L^{2}(Z; S^{-}) \oplus (\bigoplus_{j \ge 1} L^{2}(W_{j}; S^{-}))$$
$$\oplus \mathcal{K}_{-}(Y_{0,+})$$
$$\oplus (\bigoplus_{i \ge 1} \mathcal{K}_{-}(Y_{i,+}))$$
$$\oplus (\bigoplus_{j \ge 1} \mathcal{K}_{+}(Y_{j,-}))$$

sending

$$\alpha \oplus (\eta_{i,+}, \eta_{i,-})_{i \ge 0} \oplus (\varphi_j)_{j \ge 1}$$

to

$$D^{+}(Z)(\alpha) \oplus (D^{+}(W_{j})(\varphi_{j}))_{j \geq 1} \oplus (\Pi_{-}(r_{+}^{Z}(\alpha)) - e^{-2RD(Y)}\Pi_{-}(r_{-}^{W}(\varphi_{1})))$$
$$\oplus (\Pi_{+}(r_{-}^{W}(\varphi_{i+1})) - e^{2RD(Y)}\Pi_{+}(r_{+}^{W}(\varphi_{i})))_{i \geq 0}$$
$$\oplus (\Pi_{-}(r_{+}^{W}(\varphi_{j})) - e^{-2RD(Y)}\Pi_{-}(r_{-}^{W}(\varphi_{j})))_{j \geq 1}.$$

Next, we observe that making R arbitrarily large allows us to "remove" the terms depending on R. Namely, define the operators

$$\begin{split} D_6: L^2_1(Z; S^+) \oplus (\bigoplus_{i \ge 0} \mathcal{K}_+(Y_{i,+}) \oplus \mathcal{K}_-(Y_{i+1,-})) \oplus (\bigoplus_{j \ge 1} L^2_1(W^j; S^+)) \to L^2(Z; S^-) \oplus (\bigoplus_{j \ge 1} L^2(W_j; S^-)) \\ & \oplus \mathcal{K}_-(Y_{0,+}) \\ & \oplus (\bigoplus_{i \ge 1} \mathcal{K}_-(Y_{i,+})) \\ & \oplus (\bigoplus_{j \ge 1} \mathcal{K}_+(Y_{j,-})) \end{split}$$

sending

$$\alpha \oplus (\eta_{i,+}, \eta_{i,-})_{i \ge 0} \oplus (\varphi_j)_{j \ge 1}$$

to

$$D^{+}(Z)(\alpha) \oplus (D^{+}(W_{j})(\varphi_{j}))_{j \geq 1} \oplus (\Pi_{-}(r^{Z}_{+}(\alpha)))$$
$$\oplus (\Pi_{+}(r^{W}_{-}(\varphi_{i+1})))_{i \geq 0}$$
$$\oplus (\Pi_{-}(r^{W}_{+}(\varphi_{j})))_{j \geq 1}.$$

and

$$\begin{split} A: L_1^2(Z; S^+) \oplus (\bigoplus_{i \ge 0} \mathcal{K}_+(Y_{i,+}) \oplus \mathcal{K}_-(Y_{i+1,-})) \oplus (\bigoplus_{j \ge 1} L_1^2(W^j; S^+)) \to L^2(Z; S^-) \oplus (\bigoplus_{j \ge 1} L^2(W_j; S^-)) \\ & \oplus \mathcal{K}_-(Y_{0,+}) \\ & \oplus (\bigoplus_{i \ge 1} \mathcal{K}_-(Y_{i,+})) \\ & \oplus (\bigoplus_{j \ge 1} \mathcal{K}_+(Y_{j,-})) \end{split}$$

sending

$$\alpha \oplus (\eta_{i,+}, \eta_{i,-})_{i \ge 0} \oplus (\varphi_j)_{j \ge 1}$$

to

$$\{0\} \oplus \{0\} \oplus (e^{-2RD(Y)}\Pi_{-}(r_{-}^{W}(\varphi_{1})))$$
$$\oplus (e^{2RD(Y)}\Pi_{+}(r_{+}^{W}(\varphi_{i})))_{i \ge 0}$$
$$\oplus (e^{-2RD(Y)}\Pi_{-}(r_{-}^{W}(\varphi_{j})))_{j \ge 1}.$$

We have $D_5 = D_6 + A$. Since D(Y) has no kernel, we may set $\mu > 0$ to be the smallest absolute value of the eigenvalues of D(Y). Let ϕ_{λ} be the eigenfunction with eigenvalue λ . If $\lambda > 0$, then $e^{-2RD(Y)}\varphi_{\lambda} = 0$

 $e^{-2R\lambda}\varphi_{\lambda}$. If $\lambda < 0$, then $e^{2RD(Y)}\varphi_{\lambda} = e^{2R\lambda}\varphi_{\lambda}$. Along with the facts that the restriction and spectral projection maps are bounded, it follows that there is some constant M > 0 independent of R such that the operator norm of A is less than $Me^{-\mu R}$. Therefore, by making R arbitrarily large, we can make the operator norm of A arbitrarily small.

Recall it was mentioned at the end of the last section that Fredholm operators are *stable* under perturbation by operators of small operator norm. In particular, if D_5 is Fredholm, then upon making R sufficiently large, D_6 is Fredholm with the same index as D_5 .

Now observe that D_6 can be expressed as the direct sum of the operators

$$D_7: \alpha \mapsto (D^+(Z)(\alpha), \Pi_-(r^Z_+(\alpha)))$$

and

$$(\varphi_j)_{j \ge 1} \mapsto (D^+(W_j)(\varphi_j), \Pi_+(r^W_-(\varphi_j)), \Pi_-(r^W_+(\varphi_j)))_{j \ge 1}$$

The last set of operators is a direct sum of the Dirac operators with APS boundary conditions on the manifold W. It was shown earlier in Theorem 3.3.4 that these operators are isomorphisms. It follows that D_6 is isomorphic to D_7 .

However, the operator D_7 is simply the Dirac operator with APS boundary conditions on the 4-manifold Z. By the exact same proof as Theorem 3.3.4, namely the application of Proposition 3.11 of the first part of [APS75], it follows that the Fredholm index of D_7 is the same as the Fredholm index of the Dirac operator on Z with a cylindrical end attached, which is exactly the manifold Z_{∞} introduced earlier. Therefore, we have constructed a sequence of equivalences that show

$$\operatorname{ind} D^+(Z_R) = \operatorname{ind} D^+(Z_\infty)$$

for sufficiently large R.

3.4 The monopole count

In this section, we will show the second theorem necessary to finish off the proof of 3.2.1.

Theorem 3.4.1. Recall the space of solutions $\mathcal{M}(X_R, g_R, \hat{\mathfrak{p}}_R)$ discussed at the end of Section 2. There is an equality

$$#\mathcal{M}(X_R, g_R, \hat{\mathfrak{p}}_R) \equiv \operatorname{Lef}(m_o^o)$$

modulo 2.

This theorem depends crucially on the foundational theory of Seiberg-Witten-Floer homology discussed in [KM07]. As we have been doing, we will freely make use of results from this book without proof.

First, we will establish some additional notation.

Recall that $X_R = W \cup_Y [-R, R] \times Y$. Similarly, define the manifold

$$W_{R'} = [-R', 0] \times Y \cup_{\{-R'\} \times Y} W \cup_{\{R'\} \in Y} [0, R'] \times Y.$$

Also, for any 0 < R' < R, write $I_{R',R}$ for the interval [-R + R', R - R']. Then, one can also write

$$X_R = W_{R'} \cup_Y I_{R',R} \times Y.$$

Write W_{∞} for the manifold

$$(-\infty, 0] \times Y \cup_Y W \cup_Y [0, \infty) \times Y.$$

Then for finite R', write $I_{R',\infty} = ((-\infty, -R'] \cup [R',\infty)) \times Y$.

The main intuition behind Theorem 3.4.1 is as follows. A solution to the Seiberg-Witten equations on W_R may be restricted to the two boundary components \overline{Y} and Y to produce two configurations, both lying in the space $\mathcal{B}^{\sigma}(Y)$. Solutions to the Seiberg-Witten equations on X_R correspond to those solutions on W_R for which these two restrictions are the same.

Therefore, as $R \to \infty$, solutions to the Seiberg-Witten equations on X_R should correspond to solutions to the Seiberg-Witten equations on W_{∞} that are asymptotic to the same critical point in $\mathcal{B}^{\sigma}(Y)$ on both ends.

3.4.1 A compactness theorem

The first step in verifying this intuition is by proving a *compactness* theorem. For any critical points $[a], [b] \in \mathcal{B}^{\sigma}(Y)$, denote by $\mathcal{M}([\mathfrak{a}], W_{\infty}, [\mathfrak{b}])$ the space of gauge-equivalence classes of solutions to the Seiberg-Witten equations on W_{∞} that are asymptotic to $[\mathfrak{a}]$ on the "negative" end and $[\mathfrak{b}]$ on the "positive" end. The metric and perturbation are induced from the ones given on X, and the perturbation is such that all of these moduli spaces are regular. For a detailed discussion of this, see Lemma 8.1 of [LRS17] and Proposition 24.4.7 of [KM07].

We will show that solutions to the Seiberg-Witten equations on X_R , as $R \to \infty$, are indeed "asymptotic" to elements in $\mathcal{M}([\mathfrak{a}], W_{\infty}, [\mathfrak{a}])$ in the appropriate sense.

We define this "appropriate sense" below. Note that from our discussion about perturbations at the end of Section 3.2, the spaces $\mathcal{M}(X_R, g_R, \mathfrak{p}_R)$ contain no reducibles. Therefore, it is well-defined to refer to elements $\alpha \in \mathcal{M}(X_R, g_R, \mathfrak{p}_R)$ by their "blown-down" versions, i.e. gauge-equivalence classes of pairs (A, φ) of a spin^{*c*} connection and spinor that satisfy the regular perturbed Seiberg-Witten equations on X_R . Similarly, for the space $\mathcal{M}([\mathfrak{a}], W_{\infty}, [\mathfrak{a}])$, if $[\mathfrak{a}]$ is an irreducible critical point, then the elements of the space are most certainly irreducible solutions of the Seiberg-Witten equations on W_{∞} .

Definition 3.4.2. Fix some integer $k \ge 3$. Let R_n be a sequence of positive real numbers such that $R_n \to \infty$. Then, a sequence of solutions $\gamma_n \in \mathcal{M}(X_{R_n}, g_{R_n}, \mathfrak{p}_{R_n})$ are said to *converge* to a solution $\gamma \in \mathcal{M}([\mathfrak{a}], W_{\infty}, [\mathfrak{a}])$ if:

- 1. There are gauge-equivalence class representatives (A_n, φ_n) of γ_n for every n, a representative (A, φ) of γ , and L^2_{k+1} gauge transformations $u_n : X_{R_n} \to S^1$ such that the sequence $u_n \cdot (A_n, \varphi_n)$ converges to (A, φ) in L^2_k on all compact sets in W_{∞} .
- 2. Let α be a gauge-equivalence representative of the irreducible critical point $[\mathfrak{a}] \in \mathcal{B}(Y)$. For any interval $I_{R',R}$, let γ_{α} be the solution to the Seiberg-Witten equations on $I_{R',R} \times Y$ given by the constant trajectory at the critical point α . Then, for any $\varepsilon > 0$, there exists a real number R > 0 and an integer N > 0 such that, for any n > N we have $R_n > R$ and an L^2_{k+1} gauge transformation $v_n : X_{R_n} \to S^1$ such that

$$||v_n \cdot (A_n, \varphi_n)|_{I_{R,R_n}} - \gamma_\alpha ||_{L^2_{k}} < \varepsilon$$

where γ_{α} is the constant trajectory on I_{R,R_n} .

The condition of "convergence on compact subsets" is well-defined here, since any compact subset of W_{∞} is identified canonically with a compact subset of X_{R_n} for sufficiently large n.

Given this, we may state the compactness theorem.

Theorem 3.4.3. Let R_n be a sequence of positive real numbers such that $R_n \to \infty$. Consider any sequence of solutions $\gamma_n \in \mathcal{M}(X_{R_n}, g_{R_n}, \mathfrak{p}_{R_n})$. Then, after passing to a subsequence, there exists some irreducible critical point $[\mathfrak{a}]$ such that the sequence (γ_n) converges in the sense of Definition 3.4.2 to an element of $\mathcal{M}([\mathfrak{a}], W_{\infty}, [\mathfrak{a}])$.

This theorem is proved in a very similar fashion to the theorem of Kronheimer and Mrowka showing that spaces of broken trajectories are compact.

However, the result is simplified by a simple dimension count, which shows that the "broken" trajectories can really only be single trajectories.

We begin with some foundational compactness theorems.

All of these depend on control of a quantity known as the *perturbed topological energy* of a solution.

For a cylinder of the form $I \times Y$ for $I = [t_1, t_2] \subset \mathbb{R}$ a finite interval, and a perturbation q, the perturbed topological energy of a connection-spinor pair (A, φ) is given by

$$\mathcal{E}_{\mathfrak{q}}^{top}(A,\varphi) = 2(\mathcal{L}_{\mathfrak{q}}((A,\varphi)|_{\{t_1\}\times Y}) - \mathcal{L}_{\mathfrak{q}}((A,\varphi)|_{\{t_2\}\times Y})).$$

For the cobordism W_R , it is defined as

$$\mathcal{E}_{\mathfrak{q}}^{top}(A,\varphi) = 2(\mathcal{L}_{\mathfrak{q}}((A,\varphi)|_{\{R\}\times Y}) - \mathcal{L}_{\mathfrak{q}}((A,\varphi)|_{\{-R\}\times Y})).$$

The following compactness theorem on the finite cylinder is shown in Theorem 10.7.1 of [KM07]. The approach is similar to the original proof of compactness of the space of solutions to the Seiberg-Witten equations on a 4-manifold without boundary. However, the presence of a boundary does add some complications. In particular, the elliptic estimate is used in the original proof to obtain convergence in arbitrarily high Sobolev norm. In the presence of a boundary, the elliptic estimate only holds on an interior region, so convergence can only be guaranteed on an interior region.

Also, for all subsequent compactness theorems, "convergence in L_k^2 " or similar statements will always be assumed to have *k* be some large integer greater than or equal to 3.

Theorem 3.4.4. Let (A_n, φ_n) be a sequence of solutions to the Seiberg-Witten equations (perturbed by \mathfrak{q}) on $[t_1, t_2] \times Y$ such that there is a constant M > 0 independent of n such that

$$\mathcal{E}^{top}_{\mathfrak{g}}(A_n,\varphi_n) \leqslant M$$

for every *n*. Then, after passing to a subsequence, there is a sequence of L^2_{k+1} gauge transformations $u_n : [t_1, t_2] \times Y \to S^1$ such that the sequence $u_n \cdot (A_n, \varphi_n)$ converges in L^2_k to some solution (A, φ) on any interior region.

Given two (gauge-equivalence classes of) critical points $[\mathfrak{a}]$, $[\mathfrak{b}]$ of the Chern-Simons-Dirac functional, let the space $\mathcal{M}([\mathfrak{a}], [\mathfrak{b}])$ denote the space of all gauge-equivalence classes of solutions to the regular (not blown-up) Seiberg-Witten equations on the infinite cylinder $\mathcal{R} \times Y$ that are asymptotic to $[\mathfrak{a}]$ on the negative end and $[\mathfrak{b}]$ on the positive end.

Theorem 3.4.5. Let $I_n = [t_{1,n}, t_{2,n}] \times Y$ be a sequence of finite intervals such that $\lim_{n\to\infty} t_{1,n} = -\infty$, $\lim_{n\to\infty} t_{2,n} = \infty$. Let (A_n, φ_n) be a sequence of connection-spinor pairs such that (A_n, φ_n) is a solution of the Seiberg-Witten equations on $I_n \times Y$. Furthermore, suppose there exists some constant M > 0 independent of n such that the topological energies satisfy the bound

$$\mathcal{E}^{top}_{\mathfrak{a}}(A_n,\varphi_n) \leqslant M$$

for every *n*. Then, after passing to a subsequence, there is a sequence of L_{k+1}^2 gauge transformations $u_n : [t_1, t_2] \times Y \rightarrow S^1$ such that the sequence $u_n \cdot (A_n, \varphi_n)$ converges in L_k^2 on any compact subset to some solution (A, φ) of the Seiberg-Witten equations on the infinite cylinder $\mathbb{R} \times Y$ whose orbit under the action of the group of gauge transformations lies in a moduli space of the form $\mathcal{M}([\mathfrak{a}], [\mathfrak{b}])$ for some critical points $[\mathfrak{a}], [\mathfrak{b}]$ in $\mathcal{B}(Y)$.

Proof. This follows from a classic diagonal argument using Theorem 3.4.4.

Wherever defined, there are a sequence of gauge transformations $u_{n,1}$ such that, after passing to a subsequence, $u_{n,1} \cdot (A_n, \varphi_n)$ converges in L_k^2 to some pair $(A^{(1)}, \varphi^{(1)})$ on $[-1, 1] \times Y$.

Denote this sequence by $(A_{n,1}, \varphi_{n,1}) = u_{n,1} \cdot (A_n, \varphi_n)$.

Repeating this again, we may take a subsequence and apply gauge transformations $u_{n,2}$ to get $u_{n,2} \cdot (A_{n,1}, \varphi_{n,1}) = (A_{n,2}, \varphi_{n,2})$ converging in L_k^2 to some pair $(A^{(2)}, \varphi^{(2)})$ on $[-2, 2] \times Y$.

The restriction of $(A^{(2)}, \varphi^{(2)})$ to $[-1, 1] \times Y$ must be gauge-equivalent to $(A^{(1)}, \varphi^{(1)})$. We can prove this by the following argument.

First, note that the space of gauge-equivalence classes of configurations on a finite cylinder is Hausdorff (Proposition 9.3.1 of [KM07]). Then, the restrictions of $(A_{n,2}, \varphi_{n,2})$ to $[-1, 1] \times Y$ are gauge-equivalent to the corresponding terms in the sequence $(A_{n,1}, \varphi_{n,1})$. It follows that their limits on $[-1, 1] \times Y$ must then also be gauge-equivalent.

Continue taking subsequences in this fashion to get sequences $(A_{n,m}, \varphi_{n,m})$ that converge in L_k^2 on $[-m,m] \times Y$ for every positive integer m. Then, take the diagonal subsequence $(A_{n,n}, \varphi_{n,n})$.

This is the desired "gauge-transformed" subsequence of (A_n, φ_n) that by definition converges in L_k^2 on any finite cylinder $[-m, m] \times Y$, and therefore any compact subset of $\mathbb{R} \times Y$.

It remains to show that the limits $(A^{(m)}, \varphi^{(m)})$ on $[-m, m] \times Y$ are gauge-equivalent to the restriction of some single solution (A, φ) on $\mathbb{R} \times Y$. This is true by a rather "formal" argument. It is clear that, after quotienting out by gauge-equivalence, the limits $(A^{(m)}, \varphi^{(m)})$ glue together to form a gauge orbit of solutions on the infinite cylinder that has uniformly bounded topological energy on any finite cylinder. This gauge orbit must then lie in one of the spaces $\mathcal{M}([\mathfrak{a}], [\mathfrak{b}])$. Following this, we simply pick a gauge representative and that is the desired limit (A, φ) .

Theorem 3.4.6. Let R_n be a sequence of positive real numbers such that $R_n \to \infty$. Let (A_n, φ_n) be a sequence of connection-spinor pairs such that (A_n, φ_n) is a solution to the Seiberg-Witten equations on W_{R_n} for every n. Furthermore, suppose there exists some constant M > 0 independent of N such that the topological energies satisfy the bound

$$\mathcal{E}^{top}_{\mathfrak{a}}(A_n,\varphi_n) \leqslant M$$

for every *n*. Then, after passing to a subsequence, there is a sequence of L_{k+1}^2 gauge transformations $u_n : W_{R_n} \to S^1$ such that the sequence $u_n \cdot (A_n, \varphi_n)$ converges in L_k^2 on all compact subsets to some solution (A, φ) of the Seiberg-Witten equations on W_∞ whose orbit under the action of the group of gauge transformations lies in a moduli space of the form $\mathcal{M}([\mathfrak{a}], W_\infty, [\mathfrak{b}])$ for some critical points $[\mathfrak{a}], [\mathfrak{b}]$ in $\mathcal{B}(Y)$.

Proof. This is essentially identical to the proof of Theorem 3.4.5. Instead of using Theorem 3.4.4, instead one uses the analogous Theorem 24.5.2 of [KM07]. \Box

Although we have not emphasized it too much as of yet, it is generally nicer analytically to work with the non-blown-up Seiberg-Witten equations. The compactness theorems above require control of the topological energy only, while in the blown-up setting additional control is required, specifically on the slicewise norm of the spinorial part of the solutions.

Next, we will prove a result regarding Seiberg-Witten trajectories that are close to the constant trajectory. A solution γ to the Seiberg-Witten equations on a finite cylinder is a constant trajectory if and only if $\mathcal{E}_{q}^{top}(\gamma) = 0$.

If γ is constant, it takes the same values on either boundary component of the cylinder and so its topological energy vanishes by definition. On the other hand, the topological energy of a trajectory γ vanishes if and only if it takes the same value on either boundary component. However, if γ solves the Seiberg-Witten equations, then it is a downward gradient flow line. If the topological energy vanishes, then γ is a downward gradient flow line that takes the same value on either boundary component. It follows that it must take the same value on every slice $\{t\} \times Y$, and so is equal to a constant trajectory.

For the proof of Theorem 3.4.3, consider a sequence of solutions on X_{R_n} . On the neck $[-R_n, R_n] \times Y$ for large enough n, it will be shown that the topological energy of the solutions are uniformly bounded independent of n. This energy behaves in an interesting manner, in that it "concentrates" inside regions of the form $J \times Y \subset [-R_n, R_n] \times Y$ for small intervals J of size independent of n. Since the energy is finite, this naturally means the restrictions of the solutions to most of the neck have very low topological energy, and so are in this sense "nearly constant". It will be shown that these restrictions are in fact "nearly constant" in the traditional sense, that is they are very close in the L_k^2 norm to the constant trajectory for some critical point.

These phenomena will become more apparent as we continue to prove these results.

- **Lemma 3.4.7.** 1. For any solution (A, φ) to the Seiberg-Witten solutions on the cylinder $I \times Y$ for I a finite interval, the topological energy $\mathcal{E}_{q}^{top}(A, \varphi)$ is nonnegative and vanishes if and only if (A, φ) is gauge-equivalent to a constant trajectory.
 - 2. There is some constant M' independent of R such that $\mathcal{E}_{q}^{top}(A, \varphi) \ge M'$ for any solution (A, φ) to the Seiberg-Witten equations on W_R .

Proof. The first statement was proven in our discussion above.

The proof of the second statement makes use of the analytic properties of the perturbation in the perturbed case, which we have not discussed. Therefore, we will assume for this exposition that the perturbation is equal to zero.

Recall the cobordism *W* has boundary components \overline{Y} and *Y* to which the cylinders $[-R, 0] \times Y$ and $[0, R] \times Y$ are attached to create W_R .

The topological energy of (A, φ) restricted to W is by definition equal to $2(\mathcal{L}((A, \varphi)|_{\overline{Y}}) - \mathcal{L}((A, \varphi)|_{Y}))$. It is then given by the first statement that

$$\mathcal{E}_{\mathfrak{q}}^{top}(A,\varphi) = \mathcal{E}_{\mathfrak{q}}^{top}((A,\varphi)|_{[-R,0]\times Y}) + 2(\mathcal{L}((A,\varphi)|_{\overline{Y}}) - \mathcal{L}((A,\varphi)|_{Y})) + \mathcal{E}_{\mathfrak{q}}^{top}((A,\varphi)|_{[0,R]\times Y})$$

$$\geq 2(\mathcal{L}((A,\varphi)|_{\overline{Y}}) - \mathcal{L}((A,\varphi)|_{Y})).$$

It remains to bound this last term below. By Proposition 4.5.2 of [KM07], one has the equality

$$2(\mathcal{L}((A,\varphi)|_{\overline{Y}}) - \mathcal{L}((A,\varphi)|_{Y})) = \frac{1}{4} \int_{W} |F_{A}|^{2} + \int_{W} |\nabla_{A}\phi|^{2} + \frac{1}{4} \int_{W} (|\phi|^{2} + (s/2))^{2} - \int_{W} s^{2}/16 e^{-\frac{1}{4}} \int_{W} |\nabla_{A}\phi|^{2} + \frac{1}{4} \int_{W} (|\phi|^{2} + (s/2))^{2} - \int_{W} s^{2}/16 e^{-\frac{1}{4}} \int_{W} |\nabla_{A}\phi|^{2} + \frac{1}{4} \int_{W} (|\phi|^{2} + (s/2))^{2} - \int_{W} s^{2}/16 e^{-\frac{1}{4}} \int_{W} |\nabla_{A}\phi|^{2} + \frac{1}{4} \int_{W} (|\phi|^{2} + (s/2))^{2} - \int_{W} s^{2}/16 e^{-\frac{1}{4}} \int_{W} |\nabla_{A}\phi|^{2} + \frac{1}{4} \int_{W} (|\phi|^{2} + (s/2))^{2} - \int_{W} s^{2}/16 e^{-\frac{1}{4}} \int_{W} |\nabla_{A}\phi|^{2} + \frac{1}{4} \int_{W} (|\phi|^{2} + (s/2))^{2} - \int_{W} s^{2}/16 e^{-\frac{1}{4}} \int_{W} |\nabla_{A}\phi|^{2} + \frac{1}{4} \int_{W} |\nabla_{A}\phi|^{2} + \frac{1}{4} \int_{W} (|\phi|^{2} + (s/2))^{2} - \int_{W} s^{2}/16 e^{-\frac{1}{4}} \int_{W} |\nabla_{A}\phi|^{2} + \frac{1}{4} \int_{W} (|\phi|^{2} + (s/2))^{2} + \frac{1}{4} \int_{W} |\nabla_{A}\phi|^{2} + \frac{1}{4} \int_{W} (|\phi|^{2} + (s/2))^{2} + \frac{1}{4} \int_{W} |\nabla_{A}\phi|^{2} + \frac{1}{4} \int_{W} |\nabla_{A$$

where s denotes the scalar curvature.

All of these terms are positive but one, the scalar curvature, which depends only on the geometry of the manifold W and is bounded below. It follows that there is a lower bound depending only on the geometry of W, and thus independent of the pair (A, φ) .

Next, pick some "separating neighborhoods" for the critical points. Specifically, for any gauge-equivalence class of (not necessarily irreducible) critical points [a], pick neighborhoods $U_{[\mathfrak{a}]}$ of $[\mathfrak{a}]$ in $\mathcal{B}(Y)$ such that $U_{[\mathfrak{a}]}$ and $U_{[\mathfrak{b}]}$ are disjoint whenever $[\mathfrak{a}] \neq [\mathfrak{b}]$. This can be done because the critical points are a finite, discrete set in $\mathcal{B}(Y)$.

Let $[\gamma_{\mathfrak{a}}]$ be the corresponding class of trajectories on $[0,1] \times Y$ gauge-equivalent to the constant trajectory at some representative of $[\mathfrak{a}]$. Choose an open neighborhood $U_{[\mathfrak{a}]}^{[0,1]}$ of $[\gamma_{\mathfrak{a}}]$ in $\mathcal{B}([0,1] \times Y)$ such that for any $\gamma \in U_{[\mathfrak{a}]}^{[0,1]}$, the restriction of γ to the slice $\{t\} \times Y$ for any $t \in [0,1]$ lies in $U_{[\mathfrak{a}]}$.

Lemma 3.4.8. There is a constant $\varepsilon_0 > 0$ such that any solution (A, φ) to the Seiberg-Witten equations on $[-1, 2] \times Y$ with topological energy $\mathcal{E}_{q}^{top}(A, \varphi) < 3\varepsilon_0$ satisfies $[(A, \varphi)|_{[0,1] \times Y}] \subset U_{[\mathfrak{a}]}^{[0,1]}$ for some critical point $[\mathfrak{a}]$. The term $[(A, \varphi)|_{[0,1] \times Y}]$ denotes the gauge orbit of the restriction of (A, φ) to $[0, 1] \times Y$.

Proof. Suppose that this is not the case.

Then, there is a sequence of solutions (A_n, φ_n) on $[-1, 2] \times Y$ with topological energy decreasing to 0 such that they do not lie in any of these neighborhoods.

Since their topological energy is bounded, we may apply gauge transformations and pass to a subsequence to find that they converge to a solution of zero topological energy on $[0,1] \times Y$. This solution, by the first statement of Lemma 3.4.7, must be gauge-equivalent to a constant solution.

This is a contradiction, because we have $[(A_n, \varphi_n)|_{[0,1] \times Y}]$ lies in some $U_{[\mathfrak{a}]}^{[0,1]}$ for sufficiently large n. \Box

Now we can begin the proof of Theorem 3.4.3.

Let (A_n, φ_n) be a sequence of solutions on X_{T_n} . To make use of the earlier compactness theorems, the first step is to show uniform upper bounds on the topological energy.

Observe that

$$\mathcal{E}_{\mathfrak{q}}^{top}((A_n,\varphi_n)|_{[-T_n,T_n]\times Y}) = 2(\mathcal{L}_{\mathfrak{q}}((A_n,\varphi_n)|_{\{-T_n\}\times Y}) - \mathcal{L}_{\mathfrak{q}}((A_n,\varphi_n)|_{\{T_n\}\times Y}))$$
$$= -\mathcal{E}_{\mathfrak{q}}^{top}((A_n,\varphi_n)|_W).$$

Now apply Lemma 3.4.7. Combined with the above, it follows that $\mathcal{E}_{\mathfrak{q}}^{top}((A_n, \varphi_n)|_{[-T_n, T_n] \times Y}) \leq -M'$ and $\mathcal{E}_{\mathfrak{q}}^{top}((A_n, \varphi_n)|_W) \leq 0$.

These immediately imply there is some uniform upper bound M on the topological energy of (A_n, φ_n) on W_T for any $0 < T < T_n$ and on $I \times Y$ for any $I \subset [-T_n, T_n]$.

Let $\varepsilon_0 > 0$ be the constant of Lemma 3.4.8. Then, for any n, there are at most M/ε_0 integers p such that $[p, p+1] \subset [-T_n, T_n]$ and

$$\mathcal{E}^{top}_{\mathfrak{q}}((A_n,\varphi_n)|_{[p,p+1]\times Y}) \ge \varepsilon_0.$$

These can be regarded as the areas where the energy of the solution concentrates. However, these may change wildly with n.

Their behavior can be made much simpler by passing to subsequences. After taking a subsequence, one can assume there is a constant number k of them, independent of n. Label them in increasing order by

$$p_1^n < p_2^n < \dots < p_k^n.$$

Also set $p_0^n = [-T_n]$ and $p_{k+1}^n = [T_n] - 1$. Now the differences $p_{i+1}^n - p_i^n$ form sequences of positive integers, and as such they have two possible modes of behavior. Either they have an infinite constant subsequence, or they diverge and grow arbitrarily large in size with n. Therefore, after passing to a subsequence, we can assume for every i that the sequence $p_{i+1}^n - p_i^n$ is either independent of n or satisfies $\lim_{n\to\infty} p_{i+1}^n - p_i^n = \infty$.

Next, define an equivalence relation on the set of indices $\{0, 1, ..., k + 1\}$ by saying that $m_1 \sim m_2$ if the sequence $|p_{m_1}^n - p_{m_2}^n|$ is independent of n. There are a total of d equivalence classes. The number d may depend on n, but by passing to a further subsequence we can ensure that it is independent of n again.

Pick representatives $m_1 < \cdots < m_d$ of the equivalence classes. Then, for every $1 \le i \le d$, set a_i^n to be the minimum element of the equivalence class of m_i and b_i^n to be the maximum element of the equivalence class of m_i .

With this construction, we have isolated the intervals where the energy of the solutions are concentrated. One has $b_i^n - a_i^n$ is independent of n for every i, while $a_{i+1}^n - b_i^n$ grows arbitrarily large as n goes to ∞ . Qualitatively, the regions $[a_i^n, b_i^n + 1]$ will have a large amount of energy for all n, while the intervals $[b_i^n + 1, a_{i+1}^n]$ will have a small amount of energy.

Quantitatively, any interval of the form [m - 1, m + 2] in $[b_i^n + 1, a_{i+1}^n]$ will by definition satisfy the property that, for any n,

$$\mathcal{E}_{\mathfrak{q}}^{top}((A_n,\varphi_n)|_{[m-1,m+2]\times Y}) < 3\varepsilon_0.$$

Translating and applying Lemma 3.4.8 to all of these intervals, it follows that there is some critical point $[\mathfrak{a}_i]$ such that $[(A_n, \varphi_n)|_{\{t\}\times Y}] \in U_{[\mathfrak{a}_i]}$ for all $t \in [b_i^n + 2, a_{i+1}^n - 1]$. The critical point $[\mathfrak{a}]$ may depend on n, but a reader who has been diligently following along will know the drill by now: pass to a subsequence so that $[\mathfrak{a}_i]$ is independent of n.

This implies that the sequence (A_n, φ_n) , after passing to a subsequence, converges to a "broken trajectory" in the following sense. Using Theorem 3.4.6, there are gauge transformations $u_n : X_{T_n} \to S^1$ such that (A_n, φ_n) converges on compact sets in L_k^2 to some solution on W_{∞} . Using the previous work, we can see that in fact this solution lies in $\mathcal{M}([\mathfrak{a}_{d-1}], W_{\infty}, [\mathfrak{a}_1])$.

One way to see this explicitly is by restricting (A_n, φ_n) to the manifold

$$[-T_n, a_2^n + 1] \times Y \cup_{\{-T_n\} \times Y} W \cup_{\{T_n\} \times Y} [b_{d-1}^n + 2, T_n] \times Y$$

for every *n*. The further restrictions of (A_n, φ_n) to these cylinder ends lie in the neighborhoods $U_{[\mathfrak{a}_1]}$ and $U_{[\mathfrak{a}_{d-1}]}$ (on each time slice, of course) respectively.

Taking T_n to ∞ then makes the statement clear. However, it is crucial to note that this limit is *not* in the space $\mathcal{M}([\mathfrak{a}_1], W_{\infty}, [\mathfrak{a}_{d-1}])$, but rather $\mathcal{M}([\mathfrak{a}_{d-1}], W_{\infty}, [\mathfrak{a}_1])$. Recall the metric is defined on X_{T_n, T_n} such that the neck $[-T_n, T_n] \times Y$ is a metric cylinder. Because of this, the time coordinate in the induced metric on W_{∞} is *reversed*. Therefore, the Seiberg-Witten gradient flow lines travel "backwards" on W_{∞} , going from

 $[\mathfrak{a}_{d-1}]$ to $[\mathfrak{a}_1]$ rather than the other way around.

Now take a fixed integer m between 2 and d-1. Take for any n the restriction of (A_n, φ_n) to the cylinder $[b_{m-1}^n + 2, a_{m+1}^n - 1] \times Y \subset [-T_n, T_n] \times Y$.

We have that these endpoints go to $-\infty$ and ∞ respectively as $n \to \infty$. Applying Theorem 3.4.5, after applying gauge transformations and passing to a subsequence, they converge in L_k^2 on compact sets to some solution on $\mathbb{R} \times Y$. By identical reasoning to above, this limit will lie in $\mathcal{M}([\mathfrak{a}_{m-1}], [\mathfrak{a}_m])$.

With this, we have produced a "broken trajectory". By restricting the sequence (A_n, φ_n) to different parts of the manifold X_{T_n} and taking $n \to \infty$, we have produced a trajectory in the spaces

$$\mathcal{M}([\mathfrak{a}_1], [\mathfrak{a}_2]), \ldots, \mathcal{M}([\mathfrak{a}_{d-2}], [\mathfrak{a}_{d-1}])$$

and in

$$\mathcal{M}([\mathfrak{a}_{d-1}], W_{\infty}, [\mathfrak{a}_1])$$

We now claim:

Lemma 3.4.9. The variable d must be equal to 2.

Proof. Recall all of the moduli spaces are assumed to be regular. Let $gr^{\mathbb{Q}}([\mathfrak{a}])$ denote the \mathbb{Q} -grading of a critical point $[\mathfrak{a}]$.

It follows that any moduli space of the form $\mathcal{M}([\mathfrak{a}], [\mathfrak{b}])$ has dimension equal to $\operatorname{gr}^{\mathbb{Q}}([\mathfrak{a}]) - \operatorname{gr}^{\mathbb{Q}}([\mathfrak{b}]) - k$ for some nonnegative integer k. Furthermore, by translation-invariance of gradient flow lines, the space is nonempty if and only if this dimension is at least one.

This can be determined from the fact that the relative grading is the difference in \mathbb{Q} -gradings, and the blown-down version of Proposition 14.5.7 from [KM07]. It is also crucial to note that at least one of $[\mathfrak{a}]$ or $[\mathfrak{b}]$ must be irreducible for the moduli space to be nontrivial, as *Y* is an integral homology 3-sphere so it only has one gauge-equivalence class of reducible critical points.

The fact that $\mathcal{M}([\mathfrak{a}_i], [\mathfrak{a}_{i+1}])$ is nonempty for $1 \leq i \leq d-2$ requires that the \mathbb{Q} -gradings of the critical points obey the inequality

$$\operatorname{gr}^{\mathbb{Q}}([\mathfrak{a}_i]) > \operatorname{gr}^{\mathbb{Q}}([\mathfrak{a}_{i+1}])$$

The analogous result for cobordisms, Proposition 24.4.6 of [KM07], shows that $gr^{\mathbb{Q}}([\mathfrak{a}_{d-1}]) \ge gr^{\mathbb{Q}}([\mathfrak{a}_1])$. If d > 2, this is clearly impossible.

Therefore, we must have d = 2.

We have almost arrived at the proof of Theorem 3.4.3. Substituting in d = 2, $[\mathfrak{a}] = [\mathfrak{a}_{1}] = [\mathfrak{a}_{d-1}]$, the solutions (A_n, φ_n) converge in the sense of Definition 3.4.2 to a solution in $\mathcal{M}([\mathfrak{a}], W_{\infty}, [\mathfrak{a}])$.

It remains to show that [a] is irreducible. For the proof of this, we will refer the curious reader to Lemma 8.10 of [LRS17] as, while it is a crucial result, it does not add much to the exposition to present it here.

3.4.2 An application of Kronheimer-Mrowka's local gluing theorem

In this subsection, we will apply a result of Kronheimer and Mrowka described as a "local gluing theorem" along with the compactness result of Theorem 3.4.3 to prove Theorem 3.4.1.

Specifically, we will prove the following theorem, which almost immediately implies Theorem 3.4.1.

Suppressing the metric and perturbation, let $\mathcal{M}^*(X_R)$ denote the moduli space of *irreducible* solutions to the Seiberg-Witten equations on X_R . Similarly, for any 0 < R' < R, let $\mathcal{M}^*(W_{R'})$ and $\mathcal{M}^*(I_{R',R} \times Y)$ denote the spaces of irreducible solutions on $W_{R'}$ and the finite cylinder over the interval $I_{R',R} = [-R + R', R - R']$, respectively.

Theorem 3.4.10. Let \mathfrak{C}^* be the set of all gauge-equivalence classes of critical points. For all sufficiently large R > 0, the moduli space $\mathcal{M}(X_R, g_R, \mathfrak{p}_R)$ is regular and there is a homeomorphism

$$\rho: \mathcal{M}(X_R, g_R, \mathfrak{p}_R) \to \cup_{[\mathfrak{a}] \in \mathfrak{C}^*} \mathcal{M}([\mathfrak{a}], W_{\infty}, [\mathfrak{a}]).$$

The statement about regularity is important, as we have not actually shown that $\mathcal{M}(X_R)$ is regular, which is necessary for using it as a component of the invariant $\lambda_{SW}(X_R)$!

Let $\mathcal{B}^*(Y)$ denote the space of irreducible configurations on *Y*. Then, by restricting to the boundary components, there are natural restriction maps

$$r_{R'}^-: \mathcal{M}^*(W_{R'}) \to \mathcal{B}^*(Y) \times \mathcal{B}^*(Y)$$

and

$$r^+_{R',R}: \mathcal{M}^*(I_{R',R}) \to \mathcal{B}^*(Y) \times \mathcal{B}^*(Y).$$

The latter map is also defined for $R = \infty$, that is there is a restriction map

$$r^+_{R',\infty}: \mathcal{M}^*(I_{R',\infty}) \to \mathcal{B}^*(Y) \times \mathcal{B}^*(Y).$$

The natural approach to understanding solutions on X_R are to restrict to $W_{R'}$ and $I_{R',R} \times Y$ and study the two pieces separately. As $R \to \infty$, with the appropriate choice of R' the solution will be very close to zero on $I_{R',R} \times Y$. By the "local gluing theorem", these solutions can be parameterized by small balls in $\mathcal{B}^*(Y)$ and matched up with corresponding solutions on $I_{R',\infty} \times Y$, which will show Theorem 3.4.10.

Intuitively, an irreducible solution on X_R will be equivalent to a pair consisting of an irreducible solution on $W_{R'}$ and an irreducible solution on $I_{R',R} \times Y$ such that their restrictions match up in $\mathcal{B}^*(Y) \times \mathcal{B}^*(Y)$. In other words, it is an element of the *fiber product*

$$\operatorname{Fib}(r_{R'}^{-}, r_{R',R}^{+}) = \{(\gamma_{W}, \gamma_{I}) \in \mathcal{M}^{*}(W_{R'}) \times \mathcal{M}^{*}(I_{R',R} \times Y) \mid r_{R'}^{-}(\gamma_{W}) = r_{R',R}^{+}(\gamma_{I})\}.$$

Lemma 3.4.11. The natural restriction map from $\mathcal{M}^*(X_R)$ onto $\mathcal{M}^*(W_{R'}) \times \mathcal{M}^*(I_{R',R} \times Y)$ is a homeomorphism onto $\operatorname{Fib}(r_{R'}^-, r_{R',R}^+)$.

This is certainly not an obvious statement, and a proof follows from exactly the technique used in Lemma 19.1.1 of [KM07]. Given a solution on $W_{R'}$ and a solution on $I_{R',R} \times Y$, one can apply a gauge transformation such that they are in temporal gauge in a collar neighborhood of the intersection of $W_{R'}$ and $I_{R',R} \times Y$. This implies that they satisfy a differential equation near the intersection, from which it can be deduced that the two solutions can be "connected" to form a solution on X_R .

Now, we state the local gluing theorem. Let I_R be the interval $[-R, R] \times Y$ and r_R^+ the corresponding restriction map from $\mathcal{M}^*(I_R)$ to $\mathcal{B}^*(Y) \times \mathcal{B}^*(Y)$. For $R = \infty$, set $I_\infty = ([0, \infty) \cup (-\infty, 0]) \times Y$.

Theorem 3.4.12. There exists a a constant $R_1 > 0$ such that for all $R \ge R_1$ and irreducible critical points $[\mathfrak{a}]$, there

exist smooth maps

$$u_{[\mathfrak{a}]}(R,-):B([\mathfrak{a}]) \to \mathcal{M}^*(I_R)$$
$$u_{[\mathfrak{a}]}(\infty,-):B([\mathfrak{a}]) \to \mathcal{M}^*(I_\infty)$$

which are diffeomorphisms from a fixed neighborhood $B([\mathfrak{a}])$ of $[\mathfrak{a}]$ in $\mathcal{B}^*(Y)$ not depending on R onto some neighborhood on the gauge-equivalence class of the constant trajectory $[\gamma_{\mathfrak{a}}]$. The maps

$$\mu_{[\mathfrak{a}]}(R,-) = r_R^+ \circ u_{[\mathfrak{a}]}(R,-) : B([\mathfrak{a}]) \to \mathcal{B}^*(Y) \times \mathcal{B}^*(Y)$$

are smooth embeddings for all $R \in [R_1, \infty]$ and

$$\mu_{[\mathfrak{a}]}(R,-) \to \mu[\mathfrak{a}](\infty,-)$$

in the C_{loc}^{∞} topology. There also exists a constant $\eta > 0$ independent of R such that the image of the map $u_{[\mathfrak{a}]}(R, -)$ contains every trajectory $[\gamma] \in \mathcal{M}^*(I_R)$ with gauge representative γ satisfying

$$||\gamma - \gamma_{\mathfrak{a}}||_{L^2_h} < \eta.$$

Finally, for any $R, S \in [R_1, \infty]$ and distinct critical points $[\mathfrak{a}] \neq [\mathfrak{b}] \in \mathfrak{C}^*$, the maps $u_{[\mathfrak{a}]}(R, -)$ and $u_{[\mathfrak{b}]}(S, -)$ have disjoint images, as do the maps $\mu_{[\mathfrak{a}]}(R, -)$ and $\mu_{[\mathfrak{b}]}(S, -)$.

This theorem is discussed and proved in Section 18 of [KM07]. As mentioned in the book, this theorem is motivated by the case of solving a finite-dimensional linear flow. The reader is encouraged to read Section 18.1 of [KM07] to better understand Theorem 3.4.12.

The following lemma now proves that for a solution on X_R for R large, the restriction to the cylinder $I_{R',R} \times Y$ is close to a constant trajectory. This is the main application of the compactness theorem of the previous subsection.

Lemma 3.4.13. Let $\eta > 0$ be as in Theorem 3.4.12. Then there are constants $0 < R_2 < R_3 < \infty$ such that for any $R \in [R_3, \infty]$, any element of $\mathcal{M}^*(X_R)$ has a gauge-representative (A, φ) such that

$$||(A,\varphi)|_{I_{R_2,R}\times Y} - \gamma_{\mathfrak{a}}||_{L^2_k} < \eta$$

for some irreducible critical point a.

Proof. Suppose there are no such constants R_2 and R_3 .

Then there exist constants $R_n, R'_n \to \infty$ and solutions $\gamma_n \in \mathcal{M}^*(X_{R_n})$ such that any gauge representative (A_n, φ_n) of γ_n satisfies

$$||(A_n,\varphi_n)|_{I_{R'_n,R_n}\times Y} - \gamma_{\mathfrak{a}}||_{L^2_k} \ge \eta$$

for any irreducible critical point a.

Pick such a sequence of gauge representatives (A_n, φ_n) . Then by Theorem 3.4.3, after applying gauge transformations and taking a subsequence, this sequence converges in the sense of Definition 3.4.2. However, the second condition in Definition 3.4.2 is in direct contradiction with the condition above, so the lemma follows.

The previous results allow us to represent $\mathcal{M}^*(X_T)$ as a union of fiber products.

Lemma 3.4.14. Let R_2 be as in the previous lemma. For R sufficiently large, there exists a homeomorphism

$$\rho: \mathcal{M}^*(X_R) = \bigcup_{[\mathfrak{a}] \in \mathfrak{C}^*} \operatorname{Fib}(r_{R_2}^-, \mu_{[\mathfrak{a}]}(R - R_2, -))$$

and $\mathcal{M}^*(X_R)$ is regular if and only if the maps $r_{R_2}^-$ and $\mu_{[\mathfrak{a}]}(R-R_2,-)$ are transverse in $\mathcal{B}^*(Y) \times \mathcal{B}^*(Y)$.

Proof. By Lemma 3.4.11, we can write $\mathcal{M}^*(X_R)$ as $\operatorname{Fib}(r_{R_2}^-, r_{R_2,R}^+)$ for any R.

Then by Lemma 3.4.13, the restriction of any solution in $\mathcal{M}^*(X_R)$ to $I_{R_2,R} \times Y$ has a near-constant gauge-representative, which upon application of Theorem 3.4.12 implies that it can be parameterized by $\mu_{[\mathfrak{a}]}(R - R_2, -)$ for some irreducible critical point $[\mathfrak{a}]$. The fact that the images of these parameterizations are disjoint then implies the first assertion of the lemma.

The second assertion of the lemma is proved in Theorem 19.1.4 of [KM07].

To finish off the proof of Theorem 3.4.10, the following lemma will be required. Its proof requires an explicit understanding of the parameterization map in Theorem 3.4.12, so we will defer the reader to the proof in Lemma 9.7 of [LRS17].

Lemma 3.4.15. Let $B([a]) \subset \mathcal{B}^*(Y)$ be a neighborhood as in Theorem 3.4.12. Let $x_n \in B([\mathfrak{a}])$ be a sequence of points and $T_n \to \infty$ a sequence of positive real numbers such that $\mu_{[\mathfrak{a}]}(T_n, x_n)$ converges to $\mu_{[\mathfrak{a}]}(\infty, x)$ for some $x \in B([[\mathfrak{a}])$. Then x_n converges to x in the topology of $\mathcal{B}^*(Y)$.

Now, we will prove Theorem 3.4.10. The first observation is that a similar claim to Lemma 3.4.14 for the case $R = \infty$ holds.

Lemma 3.4.16. For any irreducible critical point $[\mathfrak{a}]$, the maps $r_{R_2}^-$ and $\mu_{[\mathfrak{a}]}(\infty, -)$ intersect each other transversely and there is a homeomorphism

$$\mathcal{M}([\mathfrak{a}], W_{\infty}, [\mathfrak{a}]) = \operatorname{Fib}(r_{R_2}^-, \mu_{[\mathfrak{a}]}(\infty, -)).$$

Proof. The transversality follows by our assumption that the moduli spaces $\mathcal{M}([\mathfrak{a}], W_{\infty}, [\mathfrak{a}])$ are regular. The homeomorphism is immediate by Lemma 3.4.14.

For sufficiently large R, the local gluing theorem along with the implicit function theorem implies that points of $\operatorname{Fib}(r_{R_2}^-, \mu_{[\mathfrak{a}]}(R - R_2, -))$ and $\operatorname{Fib}(r_{R_2}^-, \mu_{[\mathfrak{a}]}(\infty, -))$ can be matched up.

Specifically, Theorem 3.4.12 states that $\mu_{[\mathfrak{a}]}(R - R_2, -)$ converges to $\mu_{[\mathfrak{a}]}(\infty, -)$ in the C_{loc}^{∞} topology. It follows that, for R sufficiently large, the map $\mu_{[\mathfrak{a}]}(R - R_2, -)$ must intersect $r_{R_2}^-$ transversely near any element of $\operatorname{Fib}(r_{R_2}^-, \mu_{[\mathfrak{a}]}(\infty, -))$. Furthermore, by the implicit function theorem, there exists large $R_4 > 0$ such that the intersection of the union of the images of the maps $\mu_{[\mathfrak{a}]}(R', -)$ for $R' \in [R_4, \infty]$ and the image of r^-R_2 can be parameterized around this element of $\operatorname{Fib}(r_{R_2}^-, \mu_{[\mathfrak{a}]}(\infty, -))$ by the interval $[R_4, \infty]$.

The consequence is that, for any $(\gamma_W, \gamma_I) \in \operatorname{Fib}(r_{R_2}^-, \mu_{[\mathfrak{a}]}(\infty, -))$ there is a neighborhood $U(\gamma_W, \gamma_I) \subset \mathcal{M}^*(W_{R_2}) \times B([\mathfrak{a}])$ such that, for sufficiently large R, there is exactly one element of $\operatorname{Fib}(r_{R_2}^-, \mu_{[\mathfrak{a}]}(R-R_2, -))$ in $U(\gamma_W, \gamma_I)$ and the intersection of $r_{R_2}^-$ and $\mu_{[\mathfrak{a}]}(R-R_2, -)$ at this point is transverse.

Therefore, we have "matched up" a point of $\mathcal{M}([\mathfrak{a}], W_{\infty}, [\mathfrak{a}])$ with a corresponding point of $\mathcal{M}^*(X_R)$ for large *R*. It remains to show that this is a surjective map, i.e. it accounts for all points of $\mathcal{M}^*(X_R)$.

Lemma 3.4.17. For any large R and any element of $\mathcal{M}^*(X_R)$ given by

$$(\gamma'_W, \gamma'_I) \in \operatorname{Fib}(r_{R_2}^-, \mu_{[\mathfrak{a}]}(R - R_2, -))$$

for some irreducible critical point $[\mathfrak{a}]$, there exists a point (γ_W, γ_I) such that $(\gamma'_W, \gamma'_I) \in U(\gamma_W, \gamma_I)$.

Proof. Suppose that this is not the case. Then, there is a sequence $R'_n \to \infty$ and a sequence of points $(A_n, \varphi_n) \in \mathcal{M}^*(X_{R'_n})$ that do not satisfy this condition.

By Theorem 3.4.3, after passing to a subsequence and applying gauge transformations, they converge in the sense of Definition 3.4.2 to an element of the form

$$(\gamma_W, \gamma_I) \in \operatorname{Fib}(r_{R_2}^-, \mu_{[\mathfrak{a}]}(\infty, -)).$$

Write (A_n, φ_n) as the point

$$(\gamma_{W,n},\gamma_{I,n}) \in \operatorname{Fib}(r_{R_2}^-,\mu_{[\mathfrak{a}_n]}(R'_n-R_2,-)).$$

Since the fiber products $\operatorname{Fib}(r_{R_2}^-, \mu_{[\mathfrak{a}_n]}(R'_n - R_2, -))$ are disjoint for distinct critical points $[\mathfrak{a}_n]$, we must have $[\mathfrak{a}_n] = [\mathfrak{a}]$ for large *n*, otherwise there would be no convergence.

By Definition 3.4.2, it is immediate that $\gamma_{W,n}$ converges to γ_W .

Fix *n* large enough. Then, one has that $\mu_{[\mathfrak{a}]}(R'_n - R_2, \gamma_{I,n}) = r_{R_2}^-(\gamma_{W,n})$, so it follows as $n \to \infty$ that $\mu_{[\mathfrak{a}]}(R'_n - R_2, \gamma_{I,n})$ converges to $\mu_{[\mathfrak{a}]}(\infty, \gamma_I)$. By Lemma 3.4.15, this implies $\gamma_{I,n}$ converges to γ_I .

Therefore, $(\gamma_{W,n}, \gamma_{I,n})$ converges to (γ_W, γ_I) in the topology of $\mathcal{M}^*(W_{R_2}) \times B([\mathfrak{a}])$. This is a contradiction, because then certainly one of these elements must lie inside $U(\gamma_W, \gamma_I)$.

We have now all but proven the theorem. By these previous results, it follows that, for large *R*,

$$\mathcal{M}^*(X_R) = \cup_{[\mathfrak{a}]} \mathcal{M}([\mathfrak{a}], W_{\infty}, [\mathfrak{a}]).$$

It remains to show that X_R does not have any reducible solutions for large R. By Lemma 7.1 of [LRS17], the Dirac operator $D^+(X_R)$ has no kernel for sufficiently large R. This immediately implies that there are no reducible solutions in $\mathcal{M}(X_R)$.

An outline of the proof of this is as follows, and the reader is referred to [LRS17] for a complete proof. Let ψ be an element of the kernel of $D^+(X_R)$. Write φ for the restriction of ψ to W. Momentarily write D for the Dirac operator with APS boundary conditions on W introduced in the previous section. Then, after a series of calculations, it is found that for sufficiently large R, φ is in the kernel of an operator written as D' + K, where D' is isomorphic to $D \oplus D^*$ and K has arbitrarily small operator norm. Since D is assumed to be an isomorphism, and the kernel of an operator is stable under small perturbations (see Theorem 5.17 in Chapter IV of [Kat66]), it follows that $\varphi = 0$.

It remains to show that the restriction of ψ to I_R is equal to zero as well. Since $\varphi = 0$, the restrictions of $\psi|_{I_R}$ to the boundary of I_R are equal to zero. By the unique continuation theorem for solutions of the Dirac equation (see Section 7 of [KM07]), it follows that $\psi|_{I_R} = 0$ as well.

Therefore, $\mathcal{M}(X_R) = \mathcal{M}^*(X_R)$ for large *R*, and the desired theorem has been proved.

3.5 Completion of the proof and an obstruction to positive scalar curvature

The only remaining part of the proof of Theorem 3.4.1 is to show the (modulo 2) equality

$$\sum_{[\mathfrak{a}]\in\mathfrak{C}^*} \#\mathcal{M}([\mathfrak{a}], W_{\infty}, [\mathfrak{a}]) = \operatorname{Lef}(m_o^o)$$

For any two irreducible critical points $[\mathfrak{a}]$, $[\mathfrak{b}]$, the dimension of the moduli space $\mathcal{M}([\mathfrak{a}], W_{\infty}, [\mathfrak{b}])$ is $\operatorname{gr}^{\mathbb{Q}}([\mathfrak{a}]) - \operatorname{gr}^{\mathbb{Q}}([\mathfrak{b}])$.

Modulo 2, the Lefschetz number of m_o^o is equal to its trace. Therefore, the moduli space is nonempty iff the Q-grading of [a] is greater than or equal to that of [b]. If we order the interior critical points in decreasing order of Q-grading, then the matrix of m_o^o with respect to this ordered basis is upper triangular, and clearly has trace equal to $\sum_{[a] \in \mathfrak{C}^*} \#\mathcal{M}([\mathfrak{a}], W_\infty, [\mathfrak{a}])$ modulo 2.

We conclude from all of the above work that the splitting formula of Theorem 3.2.1 holds modulo 2:

$$\lambda_{SW}(X) = h(Y) + \operatorname{Lef}(m_o^o).$$

The modulo 2 restriction can be removed, but requires keeping track of orientations in terms of how the points of $\mathcal{M}([\mathfrak{a}, W_{\infty}, [\mathfrak{a}])$ are counted in the calculation of the Lefschetz number. Like the last chapter, this (roughly) makes use of the machinery of determinant index line bundles, but the arguments are more technical and outside of the scope of this thesis.

To finish off this section, chapter, and thesis, we give a curious application of the invariant $\lambda_{SW}(X)$ and the splitting formula to the study of positive scalar curvature metrics on 4-manifolds.

The main question in this setting is, given a manifold X of dimension n, when does it admit a metric of positive scalar curvature? For $n \ge 5$, a complete answer to this question is given for simply-connected X by the work of Stolz [Sto90]. For n = 1, the question is trivial, while for n = 2, the answer follows from the uniformization theorem for compact surfaces. For n = 3, the question is answered by Perelman's resolution of the geometrization conjecture.

For n = 4, however, there is no conclusive answer. One of the most interesting obstructions to positive scalar curvature in four dimensions comes from Seiberg-Witten theory.

Theorem 3.5.1. Let X have dimension 4 and satisfy $b_2^+(X) \ge 2$. Then, if X admits a metric of positive scalar curvature, its Seiberg-Witten invariant for any spin^c structure vanishes.

Proof. Choose a metric of positive scalar curvature on *X*. Then, we will show that *X* does not admit any irreducible solutions.

Let (A, φ) be some solution of the Seiberg-Witten equations on X. Let ∇_A be the covariant derivative associated to the connection A. Then, the square of the Dirac operator D_A and the covariant Laplacian can be related by the formula

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

This formula is known as the *Lichnerowicz-Weitzenböck formula*, and is also important for deriving the initial estimate for the proof of compactness of the space of solutions to the Seiberg-Witten equations.

Taking the inner product with φ and then integrating on *X*, this yields the formula

$$||D_A^+\varphi||_{L^2}^2 = ||\nabla_A\varphi||_{L^2}^2 + \frac{1}{2}\int_X \langle \rho(F_A^+)\varphi,\varphi\rangle d\mathrm{vol}_X + \int_X \frac{s}{4}|\varphi|^2 d\mathrm{vol}_X$$

where s is the scalar curvature.

Using the fact that (A, φ) satisfies the Seiberg-Witten equations, it follows that

$$0 = ||\nabla_A \varphi||_{L^2}^2 + \frac{1}{2} ||\varphi||_{L^4}^4 + \int_X \frac{s}{4} |\varphi|^2 d\mathrm{vol}_X$$

If the scalar curvature *s* is positive, then the right-hand side is positive unless $\varphi = 0$, so the solution (A, φ) must be reducible if it is to exist.

Since there are no irreducible solutions, the Seiberg-Witten invariants must vanish.

However, as mentioned in the statement of the theorem, this does not work for manifolds with $b_2^+(X) = 0$.

The theory presented in this thesis allows one to derive an obstruction to positive scalar curvature for 4-manifolds X with a homology orientation that satisfy assumptions (A1) and (A2). Recall assumption (A1) states that X has the integral homology of $S^1 \times S^3$, while assumption (A2) states that X has an embedded integral homology three-sphere Y such that the fundamental class of Y generates $H_3(X;\mathbb{Z})$ and the generator agrees with the homology orientation.

Theorem 3.5.2. Any manifold X satisfying (A1) and (A2) such that $\lambda_{SW}(X) \neq h(Y)$ modulo 2 cannot admit a *metric of positive scalar curvature.*

Proof. Since *X* admits a metric of positive scalar curvature, it follows from a theorem of Schoen and Yau in [SY72] that the generator of $H_3(X;\mathbb{Z})$ can be represented by an embedded 3-manifold *M* that admits a metric of positive scalar curvature.

Then, it is immediate that any lift of M to the cyclic cover \widetilde{X} is separating, that is $\widetilde{X} - M$ has two connected components. Furthermore, it is clear that there is a copy of Y in each component of $\widetilde{X} - M$.

These two copies of *Y* are connected by a cobordism given by appending some *k* copies of *W* end-toend, denoted as W^k . The two components of W^k , denoted W^k_- and W^k_+ define cobordisms from *Y* to *M* and *M* to *Y*. Since Floer homology is functorial with respect to cobordisms, we have $HM_*(W^k) = (HM_*(W))^k$ and the following diagram commutes:



However, since *M* has positive scalar curvature, we can show that $HM_*(M) = 0$.

First, we show that there are no irreducible critical points. Any critical point a induces a constant trajectory $\gamma_{\mathfrak{a}}$ which solves the Seiberg-Witten equations on $[0,1] \times Y$, and therefore certainly induces a solution to the Seiberg-Witten equations on $S^1 \times Y$ with the standard metric on S^1 . However, this metric has positive scalar curvature, so by Theorem 3.5.1, it follows that this solution must be reducible and therefore the critical point must be reducible.

Lemma 36.1.1 of [KM07] shows that the differential map $\overline{\partial}_u^s$ counting reducible trajectories from boundarystable critical points to boundary-unstable critical points vanishes as well in this situation.

By definition, this implies the map j_* in the exact triangle vanishes and so $HM_*(M) = im(j_*) = 0$. This implies now that the operator $HM_*(W)^k = 0$, and so the operator $HM_*(W)$ is nilpotent. It follows that its Lefschetz number must be zero, and then the theorem follows given the splitting formula.

Theorem 3.5.2 was first proved in a different way by Lin in [Lin16]. This article is interesting in its own right, as in the process of proving the theorem it sets up a large amount of the analysis necessary for understanding the Seiberg-Witten equations on general end-periodic 4-manifolds.

Bibliography

- [AD14] Michèle Audin and Mihai Damian. *Morse Theory and Floer Homology*. Springer-Verlag London, 2014.
- [APS75] Michael F. Atiyah, Vinay K. Patodi, and Isadore M. Singer. "Spectral asymmetry and Riemannian geometry I, II, III". In: *Math Proc. Camb. Phil. Soc.* 77–79 (1975–1976).
- [AS68] Michael F. Atiyah and Isadore M. Singer. "The index of elliptic operators I". In: *Annals of Mathematics* 87.3 (1968), pp. 484–530.
- [DK90] Simon K. Donaldson and Peter B. Kronheimer. *The Geometry of Four-Manifolds*. Oxford University Press, 1990.
- [Don83] Simon K. Donaldson. "An application of gauge-theory to 4-dimensional topology". In: Journal of Differential Geometry 18 (1983), pp. 279–315.
- [EL78] James Eells and Luc Lemaire. "A report on harmonic maps". In: *Bull Lond. Math. Soc.* 10 (1978), pp. 1–68.
- [Eva10] Lawrence C. Evans. Partial Differential Equations. 2010.
- [Fre82] Michael Freedman. "The topologoy of four-dimensional manifolds". In: Journal of Differential Geometry 17 (1982), pp. 357–453.
- [Frø10] Kim A. Frøyshov. "Monopole Floer homology for rational homology 3-spheres". In: Duke J. Math. 155.3 (2010), pp. 519–576.
- [GJ83] Mikhail Gromov and H. Blaine Lawson Jr. "Positive scalar curvature and the Dirac operator on complete Riemannian manifolds". In: *Inst. Hautes Études Sci. Publ. Math.* 58 (1983), pp. 83–196.
- [Hat] Allen Hatcher. The classification of 3-manifolds a brief overview. https://www.math.cornell. edu/~hatcher/Papers/3Msurvey.pdf.
- [Hat02] Allen Hatcher. Algebraic Topology. Cambridge University Press, 2002.
- [Kat66] Tosio Kato. *Perturbation theory for linear operators*. Springer-Verlag Berlin Heidelberg, 1966.
- [KM07] Peter B. Kronheimer and Tomasz Mrowka. Monopoles and Three-Manifolds. Cambridge, UK: Cambridge University Press, 2007.
- [Lin16] Jianfeng Lin. *The Seiberg-Witten equations on end-periodic manifolds and positive scalar curvature metrics.* 2016. arXiv: 1603.03698 [math.GT].
- [LL95] Tian-Jun Li and Ai-Ko Liu. "General wall crossing formula". In: Mathematical Research Letters 2 (1995), pp. 797–810.

- [LRS17] Jianfeng Lin, Daniel Ruberman, and Nikolai Saveliev. A splitting theorem for the Seiberg-Witten invariant of a homology $S^1 \times S^3$. 2017. arXiv: 1702.04417 [math.GT].
- [Man07] Ciprian Manolescu. "A gluing theorem for the relative Bauer-Furuta invariants". In: *Journal of Differential Geometry* 76 (2007), pp. 117–153.
- [Mor95] John Morgan. *The Seiberg-Witten Equations and Applications to the Topology of Smooth Four-Manifolds*. Princeton, New Jersey: Princeton University Press, 1995.
- [MRS11] Tomasz Mrowka, Daniel Ruberman, and Nikolai Saveliev. "Seiberg-Witten equations, end-periodic Dirac operators, and a lift of Rohlin's invariant". In: *Journal of Differential Geometry* 88.2 (2011), pp. 333–377.
- [MS74] John W. Milnor and James D. Stasheff. Characteristic Classes. Princeton University Press, 1974.
- [MT07] John Morgan and Gang Tian. Ricci Flow and the Poincaré Conjecture. A.M.S., 2007.
- [Nic00] Liviu I. Nicolaescu. Notes on Seiberg-Witten Theory. American Mathematical Society, 2000.
- [OT96] Christian Okonek and Andrei Teleman. "Seiberg-Witten invariants for manifolds with $b_+ = 1$, and the universal wall crossing formula". In: *International Journal of Mathematics* 7.6 (1996), p. 811.
- [Qui85] Daniel Quillen. "Determinants of Cauchy-Riemann operators over a Riemann surface". In: *Funk-tsional'nyi Analiz i Ego Prilozheniya* 19 (1985), pp. 37–41.
- [RS13] Daniel Ruberman and Nikolai Saveliev. Casson-type invariants from the Seiberg-Witten equations. 2013. arXiv: 1303.2273 [math.GT].
- [Sal99] Dietmar A. Salamon. Spin Geometry and Seiberg-Witten invariants. 1999.
- [Sav02] Nikolai Saveliev. Invariants of Homology 3-Spheres. Springer-Verlag Berlin Heidelberg, 2002.
- [Sma61] Stephen Smale. "Generalized Poincaré's conjecture in dimensions greater than four". In: *Annals* of *Math.* 74 (1961), pp. 391–406.
- [Sma65] Stephen Smale. "An infinite dimensional version of Sard's Theorem". In: American Journal of Mathematics 87 (1965), pp. 861–866.
- [Sti03] András I. Stipsicz. "On the vanishing of the third spin cobordism group Ω_3^{Spin} ". In: *Journal of Mathematical Sciences* 113.6 (2003), pp. 898–905.
- [Sto90] Stephan Stolz. "Simply connected manifolds of positive scalar curvature". In: *Bull. Amer. Math. Soc.* 23.2 (1990), pp. 427–432.
- [SW94] Nathan Seiberg and Edward Witten. "Electric-magnetic duality, monopole condensation, and confinement in supersymmetric Yang-Mills theory". In: *Nuclear Phys. B.* 426 (1994), pp. 19–52.
- [SY72] Richard Schoen and Shing-Tung Yau. "On the structure of manifolds with positive scalar curvature". In: *Manuscripta Math.* 28 (1972), pp. 159–183.
- [Tau11] Clifford H. Taubes. Differential Geometry: Bundles, Connections, Metrics and Curvature. Oxford University Press, 2011.
- [Tau87] Clifford H. Taubes. "Gauge theory on asymptotically periodic 4-manifolds". In: Journal of Differential Geometry 25.3 (1987), pp. 363–430.
- [Tau90] Clifford H. Taubes. "Casson's invariant and gauge theory". In: *Journal of Differential Geometry* 31 (1990), pp. 547–599.