COMPACTNESS RESULTS FOR THE QUILTED ATIYAH-FLOER CONJECTURE

BY DAVID LEE DUNCAN

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Mathematics

Written under the direction of

Chris Woodward

and approved by

New Brunswick, New Jersey

May, 2013

ABSTRACT OF THE DISSERTATION

Compactness results for the quilted Atiyah-Floer conjecture

by David Lee Duncan

Dissertation Director: Chris Woodward

Given a closed connected oriented Riemannian 3-manifold Y equipped with a non-homotopically trivial function $Y \to S^1$, one can define an instanton Floer cohomology group as well as a quilted Lagrangian Floer cohomology group. We develop compactness results relating the boundary operators.

Acknowledgements

First and foremost, I am grateful to my advisor Chris Woodward for his insight and valuable suggestions. So much of this work became possible due to his supervision. I would also like to thank Penny Smith, Katrin Wehrheim, Dietmar Salamon, Fabian Ziltener and Peter Ozsvath for enlightening discussions. Finally, thanks are certainly due to my fiancée Claudette D'Souza for her help with early versions of the manuscript, and for her support along the way.

Dedication

This thesis is dedicated to my mother, for her constant support and generosity. Everything I am I owe to her.

Table of Contents

Chapter 1

Introduction

In a sequence of papers [14] [15] [18], Floer introduced an invariant for closed connected oriented 3-manifolds, called *instanton Floer cohomology*. To describe this, we fix such a manifold Y and a principal G-bundle $Q \to Y$. The space $\mathcal{A}(Q)$ of connections on Q admits a function $\mathcal{CS} : \mathcal{A}(Q) \to \mathbb{R}$, called the *Chern-Simons functional.* Instanton Floer cohomology can be viewed as the Morse cohomology of the Chern-Simons functional; that is, the cohomology of a chain complex $(CF_{\mathrm{inst}}(Q), \partial_{\mathrm{inst}})$ where $CF_{\mathrm{inst}}(Q)$ is the abelian group freely generated by the critical points of $\mathcal{CS}$, and the boundary operator ∂_{inst} counts isolated gradient flow lines between critical points.[1] These critical points are precisely the flat connections on Q, and the gradient trajectories can be viewed as connections on the 4-manifold $\mathbb{R} \times Y$ which satisfy the instanton equation.

There are a variety of issues encountered when trying to define this invariant. The first issue is that, as with finite-dimensional Morse theory, the analysis requires that all critical points are non-degenerate. This, however, is never the case for the Chern-Simons functional due to the presence of an infinite dimensional group, called the *gauge group*, which acts naturally on the space of connections and (up to a constant) preserves the Chern-Simons functional. So we should really be working modulo gauge equivalence. Even modulo the action of the gauge group, we still may have degenerate critical points. However, at this point, the degeneracy issue is only a finite-dimensional problem, and non-degeneracy can be achieved by a suitable perturbation of the Chern-Simons functional, just as in finite-dimensional Morse theory.

As mentioned, to avoid certain analytic difficulties, we consider gauge equivalence

---

[1]Though it will not be discussed much in this thesis, there is an underlying grading on the chain complex $(CF_{\mathrm{inst}}(Q), \partial_{\mathrm{inst}})$. Given our choices, the boundary operator we discuss here *increases* the degree, hence Floer *co*homology.

classes of flat connections. This brings us to the second issue: the gauge group often does not act freely on the space of flat connections, which leads to the obvious differential geometric issues when we quotient by the action. For example, if Q is a trivial bundle, then the trivial connection is never a free point for the action and so its gauge equivalence class is always a singular point of the quotient space. A connection is called *irreducible* if the gauge group acts locally freely at that connection, and *reducible* otherwise. Following Casson (see [1]), Floer's original approach [14] resolved the issue of reducibles by assuming that the 3-manifold Y is a homology 3-sphere (i.e., $H_1(Y, \mathbb{Z}) = 0$), and by taking $Q := Y \times \mathrm{SU}(2)$ to be the trivial principal SU(2)-bundle. The crucial observation is that, in this case, the trivial connection is the only reducible one, and can easily be excluded. Floer was then able to show that, for a suitable metric and up to a perturbation of the Chern-Simons functional, there are only finitely many gauge equivalence classes of flat connections, and the boundary operator squares to zero, $\partial_{\mathrm{inst}}^2 = 0$. In particular, the cohomology of the chain complex $(CF_{\mathrm{inst}}(Q), \partial_{\mathrm{inst}})$ is well-defined in this case. This cohomology is precisely instanton Floer cohomology, and we denote it by $HF_{\mathrm{inst}}(Y)$. Furthermore, $HF_{\mathrm{inst}}(Y)$ only depends on Y, and not the underlying metric.

At about the same time, Floer was working on a similar program in the symplectic category [16] [17]. To describe this, consider a smooth symplectic manifold (M, ω) equipped with an almost complex structure J, and two Lagrangian submanifolds $L_0, L_1 \subset M$. In this set-up, the objective is to define the Morse cohomology of a function called the *symplectic action functional*, and thereby arrive at a chain complex $(CF_{\mathrm{symp}}(L_0, L_1), \partial_{\mathrm{symp}})$. Under suitable hypotheses on the data, the symplectic action is defined on the space of paths in M terminating at the L_i. The critical points (i.e., the generators of $CF_{\mathrm{symp}}(L_0, L_1)$) are the intersection points of the two Lagrangians, and the gradient flow lines are maps from the unit strip in $\mathbb{C}$ to M which satisfy the J-holomorphic equation and have Lagrangian boundary conditions. Here, non-degeneracy can be achieved by perturbing the Lagrangians so they intersect transversely. Floer's theorem says that in this situation, and for a suitable choice of J, the boundary operator squares to zero. (This theorem was extended by Oh [33] to include a wide class of symplectic manifolds which will be relevant to us.) The resulting cohomology $HF_{\text{symp}}(L_0, L_1)$, called *Lagrangian intersection Floer cohomology*, then depends only on M, ω, and the L_j.

One relationship between these two Floer cohomology theories was outlined by Atiyah in [2]. To describe this relationship, we continue to assume that Y is closed, connected and oriented. Fix a function $f : Y \to \mathbb{R}$ together with a regular value r in the image of f, so $\Sigma := f^{-1}(r)$ is a compact smooth surface. Then $Y^+ := f^{-1}([r, \infty))$ and $Y^- := f^{-1}((-\infty, r])$ are 3-manifolds with boundary Σ, and their union is Y.

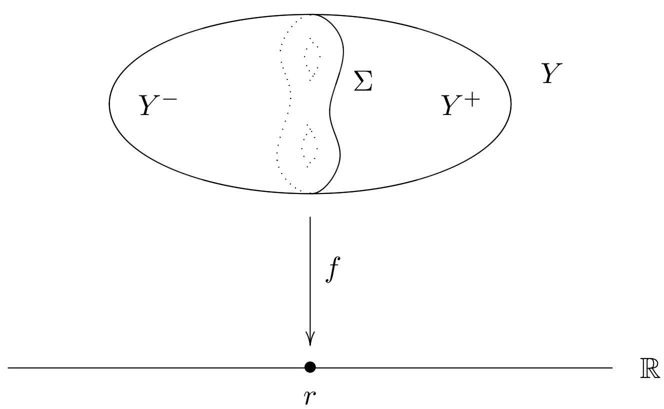


Figure 1.1: Using a Morse function to decompose a manifold Y into handle-bodies $Y^{\pm}$.

The trivial SU(2)-bundle over Y restricts to bundles over Σ, and the $Y^{\pm}$. As shown by Atiyah and Bott [3], the gauge equivalence classes of flat connections on Σ form a finite-dimensional symplectic space $M(\Sigma)$, and any choice of metric on Σ induces an almost complex structure on $M(\Sigma)$. Similarly, we let $L(Y^{\pm})$ denote the gauge equivalence classes of flat connections on $Y^{\pm}$. Then restricting to the boundary provides maps $L(Y^{\pm}) \to M(\Sigma)$, and the image of each is Lagrangian in $M(\Sigma)$. Furthermore, the intersection points of the Lagrangians are exactly the gauge equivalence classes of flat connections on Y. Assuming, for now, that both of the Floer chain complexes $(CF_{\text{inst}}(Y), \partial_{\text{inst}})$ and $(CF_{\text{symp}}(L(Y^-), L(Y^+)), \partial_{\text{symp}})$ are defined, this is saying that there is a natural isomorphism $CF_{\text{inst}}(Y) \cong CF_{\text{symp}}(L(Y^-), L(Y^+))$ of abelian groups.

Recall that the resulting Floer cohomologies do not depend on the choice of metric (instanton case) or almost complex structure (Lagrangian case). Atiyah [2] pointed out that if one begins with a metric and modifies it in such a way that a fixed neighborhood

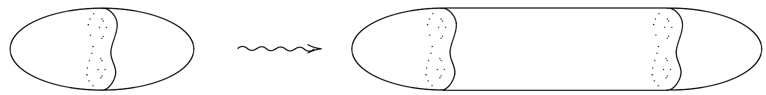


Figure 1.2: An illustration of the process of 'stretching the neck' for a 3-manifold decomposed into handle-bodies.

of Σ in Y is becoming increasingly long, then the instanton equation defining ∂_{inst} approximates the J-holomorphic curve equation defining ∂_{symp}. He called this procedure 'stretching the neck'. In principle, for suitable metrics, these two boundary operators should be counting the same thing. Atiyah and Floer were consequently led to make the following conjecture:

Atiyah-Floer Conjecture. *There is a natural isomorphism*

$$HF_{\text{inst}}(Y) \cong HF_{\text{symp}}(L(Y^-), L(Y^+)). \tag{1.1}$$

There has been much work towards a resolution of this conjecture in the past two decades. For example, Taubes [40] was able to show that the Euler characteristics of the two cohomologies agree. We describe various other approaches momentarily. It should be noted, however, that there is a real difficulty present even in the statement of the Atiyah-Floer conjecture. Namely, the Lagrangian Floer cohomology on the right-hand side of (1.1) is not well-defined. This is due to the presence of reducible connections on Σ, which lead to singularities in the symplectic space $M(\Sigma)$ (as well as in the Lagrangians). This is a serious difficulty from the symplectic perspective, and part of the conjecture is identifying what exactly should go on the right-hand side of (1.1).

One possible resolution of this issue lies in a program of Salamon and Wehrheim [36] [45] [46] [47] [49] [39], where they introduce a third Floer group which counts instantons which have (infinite-dimensional) Lagrangian boundary conditions. See [50] for a nice survey of this approach.

A second resolution, and the one we adopt in this thesis, is to work with a non-trivial bundle $Q \to Y$ (in which case $H_1(Y, \mathbb{Z})$ is necessarily non-zero). This idea was put forward by Floer in [15], where he takes Q to be a non-trivial SO(3)-bundle. The non-triviality of the bundle resolves the problem of reducibles on the instanton side, and essentially the same proof as in the SU(2) case shows that $HF_{\text{inst}}(Q)$ is well-defined (this group may now depend on the particular isomorphism class of Q). Using this set-up, and Atiyah's idea to 'stretch-the-neck', Dostoglou and Salamon [10] [11] [12] [13] were able to prove the conjecture in the case where

$$Y = Y_\Phi := ([0,1] \times \Sigma) / \Phi$$

is the mapping torus of some surface diffeomorphism $\Phi : \Sigma \to \Sigma$. They assume the bundle has been chosen so that it restricts to a non-trivial bundle on Σ. Then there are no reducibles over Σ and $M(\Sigma)$ is a *smooth* symplectic manifold. Pullback by Φ induces a symplectomorphism $\Phi^* : M(\Sigma) \to M(\Sigma)$, and Dostoglou and Salamon show that $HF_{\text{inst}}(Q) \cong HF_{\text{symp}}(\text{Graph}(\Phi^*), \Delta)$. Here, $\text{Graph}(\Phi^*) \subset M(\Sigma) \times M(\Sigma)$ is the graph of Φ^*, $\Delta \subset M(\Sigma) \times M(\Sigma)$ is the diagonal, and both of these are Lagrangian submanifolds of the product $M(\Sigma) \times M(\Sigma)$.

The approach we take in this dissertation is to generalize the Dostoglou-Salamon approach in the following fashion: Suppose Y is a closed connected oriented 3-manifold equipped with a map $f : Y \to S^1$ which is not homotopically trivial. (The mapping tori considered by Dostoglou and Salamon are examples, but there are many more than just these.) Just as in the Dostoglou-Salamon case, it is possible to choose a principal SO(3)-bundle Q for which instanton Floer cohomology is well-defined. As we will see in the next section, after a suitable refinement of f, the data (Y, f) provides a decomposition of Y into a finite union of elementary cobordisms $\left\{Y_{i(i+1)}\right\}_{i=1}^N$ which pairwise intersect along surfaces $\{\Sigma_i\}_{i=1}^N$.

It is possible to choose Q so its restriction to each Σ_i is non-trivial, hence the $M(\Sigma_i)$ are smooth symplectic manifolds. Restricting to each of the two boundary components of the elementary cobordisms defines two smooth Lagrangian submanifolds $L^{(0)}, L^{(1)}$ of the product $M(\Sigma_1) \times \ldots \times M(\Sigma_N)$. As a consequence, $HF_{\text{symp}}(L^{(0)}, L^{(1)})$ is well-defined. In this dissertation we take this as a candidate for the right-hand side of (1.1). It follows almost immediately from the definitions that there is a natural isomorphism of abelian groups $CF_{\text{inst}}(Q) \cong CF_{\text{symp}}(L^{(0)}, L^{(1)})$ (see Proposition 2.4.3 below). The

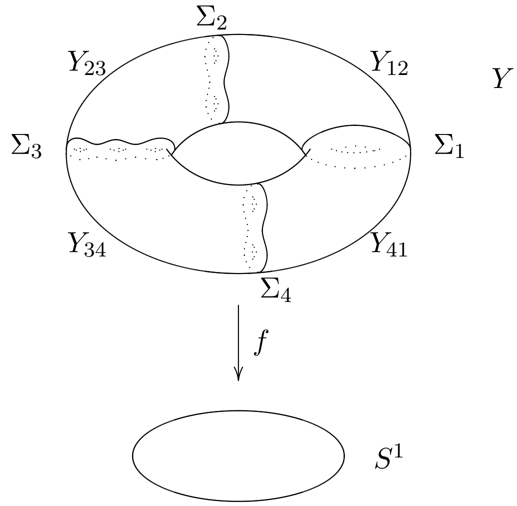


Figure 1.3: A decomposition of Y into a cyclic union of cobordisms using a circle-valued Morse function.

main result of this thesis is the following theorem, which says that, for a suitable metric, every instanton trajectory counted by the boundary operator ∂_{inst} is close to some holomorphic trajectory counted by the boundary operator ∂_{symp}.

Theorem A. *Let (Y, f) and $Q \to Y$ be as above, and assume that all flat connections on Q are non-degenerate. For each $i \in \{1, \ldots, N\}$, fix a tubular neighborhood $[0,1] \times \Sigma_i \hookrightarrow Y$ of $\Sigma_i \subset Y$. Then there is a family of metrics $\{g_\epsilon\}_{\epsilon > 0}$ on Y with the following significance:*

Let a^-, a^+ be flat connections on Q, and suppose $(\epsilon_\nu)_{\nu \in \mathbb{N}}$ is a sequence of positive numbers converging to zero. Fix $q > 2$ and suppose that, for each ν, there is a connection A_ν on $\mathbb{R} \times Q$ of Sobolev class $W^{1,q}$ which is an instanton with respect to the metric $ds^2 + g_{\epsilon_\nu}$, and lies in the zero-dimensional component of the moduli space of instantons limiting to $a^\pm$ at $\pm\infty$. Then there is a subsequence (still denoted $(A_\nu)_\nu$), a sequence of gauge transformations $(u_\nu)_\nu$, a sequence of real numbers $(s_\nu)_\nu$, and a continuous connection A_∞ on $\mathbb{R} \times Q$ of Sobolev class $W^{1,q}_{loc}$, such that:

(i) A_∞ represents a holomorphic strip with Lagrangian boundary conditions,

(ii) $A_\infty|_{\{s\} \times Y}$ converges to $a^\pm$ as $s \to \pm\infty$, and

(iii) for each $i \in \{1, \ldots, N\}$, the Σ_i-components in the sequence $(u_\nu^* \tau_{s_\nu}^* A_\nu)_\nu$ converge to the Σ_i-component of A_∞ in C^0 on compact subsets of $\mathbb{R} \times I \times \Sigma_i$.

Here, τ_s^* denotes $\mathbb{R}$-translation by s.

Theorem A is an immediate corollary of the Main Theorem 2.4.1, below. We follow the 'stretch-the-neck' approach of Atiyah, as is suggested by the family of metrics g_ϵ (although in our case, there are actually multiple 'necks', one for each of the Σ_i). To outline this proof in a little more detail, we first mention that the instanton equation is conformally invariant, so 'stretching the neck' is equivalent to keeping the neck a fixed length, but shrinking the volumes of the surface fibers Σ_i. Consequently, we shrink the volumes of the elementary cobordisms $Y_{i(i+1)}$ at the same rate. An instanton with respect to such a metric can be viewed as a tuple of maps $\left(\{\alpha_i\}, \left\{a_{i(i+1)}\right\}\right)$, where α_i is a map from the 'neck' $\mathbb{R} \times [0, 1]$ into the space of connections on Σ_i, and $a_{i(i+1)}$ is a map from $\mathbb{R}$ into the space of connections on $Y_{i(i+1)}$. It is sufficient to consider instantons with fixed small energy (this is effectively what is implied by the condition in Theorem A that the instantons A_ν lie in the zero-dimensional component of the moduli space). The small volumes imply that either (i) the curvatures of the α_i and $a_{i(i+1)}$ are close to zero, or (ii) there is a non-trivial amount of energy localized at a point. Case (ii) is ruled out by a bubbling analysis, where we show that it contradicts the assumption that the energy is small and fixed. It therefore suffices to consider case (i), where the curvatures are close to zero. By means of a Narasimhan-Seshadri correspondence on the surfaces, and a gradient flow on the cobordisms (referred to below as the *Yang-Mills heat flow*), the tuple $\left(\{\alpha_i\}, \left\{a_{i(i+1)}\right\}\right)$ descends to an honest holomorphic strip in $M(\Sigma_1) \times \ldots \times M(\Sigma_N)$ with Lagrangian boundary conditions given by $L^{(0)}$ and $L^{(1)}$, and we are able to show that this holomorphic strip is close to the original instanton.

Remark 1.0.1. *(a) In the case where there are no elementary cobordisms, we find ourselves in exactly the situation considered by Dostoglou and Salamon. The technique considered here should be contrasted with the strategy of [12, Theorems 8.1, 9.1] and*

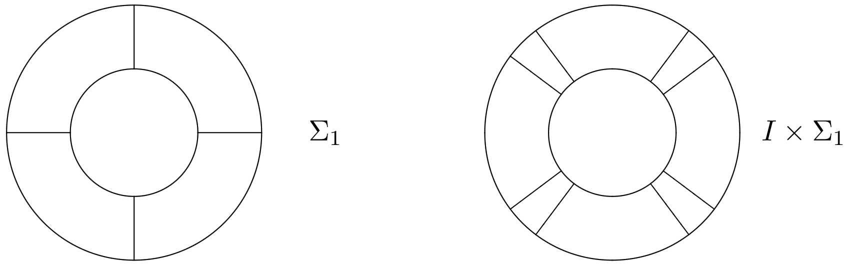


Figure 1.4: This is an illustration of the manifold Y when viewed from the top. The vertical and horizontal lines in the picture on the left represent the surface fibers Σ_j. In the picture on the right these have been thickened to cylinders $I \times \Sigma_j$, which is more convenient for performing the adiabatic limit.

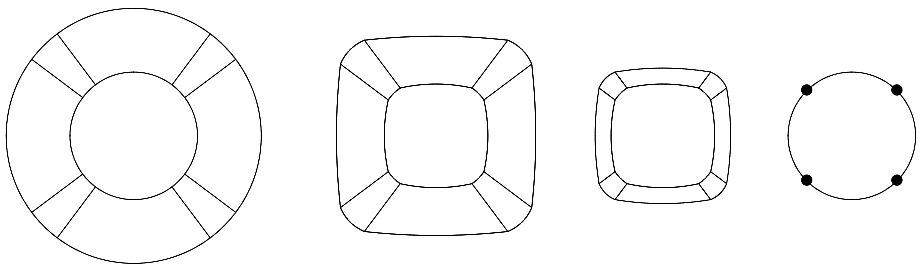


Figure 1.5: Moving from left to right, the metric on the manifold Y is being deformed in such a way that the volume of the Σ_i and the $Y_{i(i+1)}$ are going to zero. However, the volume in the I-direction (the 'neck') is remaining fixed. This process is conformally equivalent to 'stretching the neck', wherein the volumes of the Σ_i and $Y_{i(i+1)}$ are fixed, but the volume on the neck parameter I is increased.

[13, Theorem 3.1], where, for our case, having established various properties of the Narasimhan-Seshadri correspondence and Yang-Mills heat flow (Theorems 3.1.1 and 3.2.3 below), we can avoid much of the analysis on the instantons themselves. This is particularly useful in our more general setting, since an analytic approach in the case of elementary cobordisms requires additional boundary estimates which are difficult to obtain for our particular choice of degenerating metrics. More concretely, Dostoglou-Salamon's [12, Lemma 8.2] should be viewed as a first order approximation of the Narasimhan-Seshadri correspondence of Theorem 3.1.1. The usefulness of the sharper Theorem 3.1.1 becomes apparent in, for example, Theorem 4.2.1 appearing in this thesis. On the other hand, the Dostoglou and Salamon's result [13, Theorem 3.3] gives a slightly stronger result.

(b) It should be mentioned that in [20] and [21], Fukaya describes an approach very similar to the one presented here. Though quite similar, there is one striking difference between his approach and ours: Fukaya deals with a fixed smooth (albeit not everywhere positive definite) metric, whereas we deal with a sequence of metrics which degenerate to a singular metric, and these metrics are not smooth. *It seems this may be a distinction that is only really relevant at the level of the analysis, but since neither program has been successfully carried out to its desired conclusion, at the moment it is difficult to make any precise statement to this effect.*

(c) Recently M. Lipyanskiy has developed compactness results for quilts consisting of mixed patches of instantons and holomorphic curves. The motivation is that these quilts may define a chain map which interpolates between the trajectories defining each of the two Floer theories.

(d) Though the results here can be described in terms of Lagrangian intersection Floer cohomology, geometrically it is more natural to use the language of holomorphic quilts as developed by Wehrheim and Woodward [52] [53] [54]. In section 2.3, we recall the definitions from quilted Floer cohomology which will be used.

(e) Roughly speaking, the significance of the Lie group $\mathrm{SU}(2)$ *in the gauge theory described above stems from the following properties: 1)* $\mathrm{SU}(2)$ *is compact and simply-connected; 2)* $\mathrm{SU}(2)$ *is simple with a discrete center; 3)* $\mathrm{SU}(2)$*-bundles are well-understood. When passing to the quotient* $\mathrm{SO}(3) = \mathrm{SU}(2)/Z(SU(2))$*, these properties descend to properties which are desirable from the Floer-theoretic perspective. It turns out that this can be generalized from* $\mathrm{SU}(2)$ *to* $\mathrm{SU}(r)$*, for* $r \geq 2$*, where* $\mathrm{PSU}(r) = \mathrm{SU}(r)/Z(\mathrm{SU}(r))$ *now plays the role that* $\mathrm{SO}(3)$ *did previously.*

The basic outline for the remainder of this paper is as follows. Section 2 begins with a definition of the types of 3-manifolds being considered, and a discussion of the particular metrics that are used for our adiabatic (i.e., 'stretch-the-neck') limit. The remainder of section 2 is dedicated to supplying the relevant background. We end section 2 with precise statement of the Main Theorem 2.4.1 as well as a proof that the two Floer cohomology theories have the same generators. Section 4 contains the proof of the Main Theorem, and section 3 develops the geometric results about the Narasimhan-Seshadri correspondence and Yang-Mills heat flow which are used in section 4.

Finally, we remark that a full proof of the Atiyah-Floer conjecture, along the lines we present here, would require showing that near each holomorphic strip there is a unique instanton, up to gauge equivalence, and for a suitably stretched neck. That is, one would need to establish an injection $\mathcal{F}$ from trajectories of holomorphic curves to instanton trajectories. The Main Theorem we present here can be viewed as saying that such a map would be onto, assuming it exists. Using an implicit function theorem argument, Dostoglou and Salamon [12] were able to show its existence in the case where $Y = Y_\Phi$ is a mapping torus. However, the already delicate analysis used in their proof becomes even more fickle in our situation, due to the presence of the non-trivial cobordisms $Y_{i(i+1)}$. At the point of writing, the existence of such an $\mathcal{F}$ in our setting is an active research project.

Chapter 2

Background on the Atiyah-Floer conjecture

2.1 Smooth manifold theory

In this section we describe the particular decompositions of the 3-manifolds we are considering. This will allow us to define the family of metrics used to 'stretch-the-neck'. However, to do this in a meaningful way, we will need to impose a special family of smooth structures on the 3-manifold. Finally, we introduce our notation for bundles and metrics which will be used throughout the remainder of this thesis.

2.1.1 Broken fibrations of 3-manifolds

We begin with a closed connected oriented manifold Y, and we want to describe when Y admits a non-homotopically trivial map to the circle. Since S^1 is the Eilenberg-MacLane space for the group $\mathbb{Z}$, there is a group isomorphism between the first cohomology of Y and the space of homotopy classes of maps to the circle:

$$H^1(Y, \mathbb{Z}) \cong \left[Y, S^1\right].$$

(The group structure on the right is induced from the group structure on S^1.) This isomorphism can be realized explicitly as follows: Any $b \in H^1(Y, \mathbb{Z})$ can integrated over paths in Y with one fixed basepoint. This integral depends only on the homotopy class of the path, and so we get a map $Y \to \mathbb{R}$, which is well-defined up to an overall constant given by a generator n of

$$\{b(\gamma) | \gamma \in \pi_1(Y)\} \subseteq \mathbb{Z}.$$

So b determines a well-defined map

$$f : Y \longrightarrow \mathbb{R}/n\mathbb{Z},$$

and hence a group homomorphism

$$H^1(Y,\mathbb{Z}) \longrightarrow \left[Y, S^1\right].$$

It is immediate that $b = 0$ if and only if $n = 0$ if and only if f is homotopic to a point, so this homomorphism is injective. Conversely, given a smooth $f : Y \to S^1$, the pullback of the positive generator of $H^1(S^1,\mathbb{Z})$ determines a class in $H^1(Y,\mathbb{Z})$. This can equivalently be realized by differentiating f and passing to cohomology, $b := [df] \in H^1(Y,\mathbb{Z})$. If f and b are related in this way, then we will say that f is a **representative** of b. As mentioned in the introduction, we are interested in 3-manifolds Y equipped with non-homotopically trivial maps $f : Y \to S^1$. It follows that such maps exist if and only if $H^1(Y,\mathbb{Z})$ is non-zero. By standard homological considerations, this happens (for closed oriented 3-manifolds) if and only if Y has positive first Betti number.

It is a well-known fact from Morse theory that each class $b \in H^1(Y,\mathbb{Z})$ has a representative f whose critical points are all non-degenerate (i.e., f is locally a Morse function), and that these critical points have distinct critical values [29, Theorem 2.7]. Now suppose Y is a 3-manifold. Then critical points of index 0 and 3 correspond to *local* 'maxima' and 'minima'. If we assume $b \neq 0$, then any representative $f : Y \to S^1$ has no *global* maxima or minima, and so any index 0 or 3 critical points can be homotoped away, see [29] or [22]. In particular, if $b \neq 0$, then we may further assume that each critical point of f has index 1 or 2.

Fix such a function $f : Y \to S^1$ and let N denote the number of critical points. Assume, for now, that $N > 0$ is positive. Identify $S^1 \cong \mathbb{R}/C\mathbb{Z}$ for some $C > 0$. Find regular values $r_i \in S^1$ such that, for each $1 \leq i \leq N$, there is exactly one critical value $c_{i(i+1)}$ with $r_i + \delta < c_{i(i+1)} < r_{i+1} - \delta$, for some fixed $\delta > 0$ (this is possible because we have assumed the critical points have distinct critical values). Here and below we work with i modulo N. We may assume the circumference C is large enough to take $\delta = 1/2$. Define

$$\Sigma_i := f^{-1}(r_i - 1/2), \qquad Y_{i(i+1)} := f^{-1}\left([r_i + 1/2, r_{i+1} - 1/2]\right),$$

which are closed surfaces, and compact cobordisms, respectively. The orientation on Y and the data of f provide canonical orientations on the Σ_i. The $Y_{i(i+1)}$ inherit orientations as well.

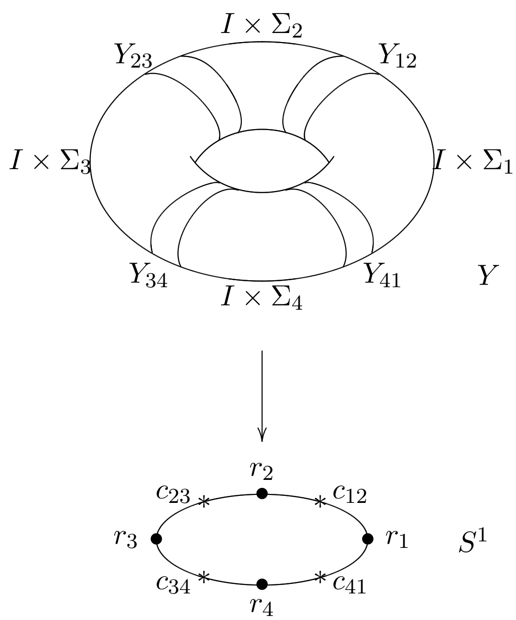


Figure 2.1: An illustration of the correspondence between the critical and regular values, and the $Y_{i(i+1)}$ and Σ_i.

Fix a metric g on Y and we assume that g is suitably generic in a sense that we will make precise later. We refer to g, or its restriction to any submanifold of Y, as the **fixed metric**. Note that there are no critical values between $r_i - 1/2$ and $r_i + 1/2$, so $V := \nabla f / |\nabla f|$ is well-defined on $f^{-1}([r_i - 1/2, r_i + 1/2])$. The time-1 gradient flow of V provides an identification

$$f^{-1}([r_i - 1/2, r_i + 1/2]) \cong I \times \Sigma_i, \tag{2.1}$$

where we have set $I := [0,1]$. This also provides an identification of $f^{-1}(t)$ with Σ_i for $t \in [r_i - 1/2, r_i + 1/2]$. So the function f together with the metric g allow us to view Y as the composition of cobordisms

$$Y_{N1} \cup_{\Sigma_1} (I \times \Sigma_1) \cup_{\Sigma_1} Y_{12} \cup_{\Sigma_2} \ldots \cup_{\Sigma_{N-1}} Y_{(N-1)N} \cup_{\Sigma_N} (I \times \Sigma_N) \cup_{\Sigma_N} \tag{2.2}$$

Note that this is *cyclic* in the sense that the cobordism $I \times \Sigma_N$ on the right is glued to the cobordism Y_{N1} on the left, reflecting the fact that f maps to the circle. By construction, each $Y_{i(i+1)}$ is an **elementary cobordism** since it admits an I-valued Morse function with no more than one critical point, and this Morse function preserves the cobordism structure between $Y_{i(i+1)}$ and I.

Since each critical point has index either 1 or 2, it follows from standard Morse theory considerations [29, Theorem 3.14] that the genus of Σ_i differs from that of Σ_{i+1} by one. For example, if a critical point $c_{i(i+1)}$ has index 1, then Σ_{i+1} is obtained from Σ_i by attaching a single 1-cell, and $f|_{Y_{i(i+1)}}$ contains the data for how this is done. If the index is 2, then the situation between Σ_{i+1} and Σ_i is reversed since dimensions 1 and 2 are dual for 3-manifolds. Hence, as we traverse along the circle, the genus of the fiber remains constant, except for at the critical points where it changes by ± 1. However when we return to where we started the genus must have had a net change of zero, and we conclude that there are just as many critical points of index 1 as there are of index 2. In particular, N is necessarily an *even* number. This also shows that each of the cobordisms in (2.2) all have the same number of connected components. Moreover, if f has k connected components in each fiber, then there exists a function $\widetilde{f} : Y \to S^1$ with $f(y) = \widetilde{f}(y)^k$, the kth power. So there exists a representative f with connected fibers if and only if $[df]$ is not a multiple of any other integer cohomology class. In particular, this is always the case when Y has positive first Betti number. Circle-valued functions f which are locally Morse with connected fibers were considered by Lekili [28], where he called them **broken fibrations**. From now on we assume $f : Y \to S^1$ is a broken fibration. It follows that, in this case, the manifolds Σ_i and $Y_{i(i+1)}$, described above, are connected.

This set-up can be modified to include the case with no critical points, $N = 0$, as follows: Let $r \in \mathbb{R}/\mathbb{Z}$ be any point (necessarily a regular value) and $\Sigma_0 := f^{-1}(r)$. By performing a suitable homotopy to the Morse function f we can replace f by a Morse function with two critical points (this is the opposite procedure to the cancellation of critical points [29], and is the Morse-theoretic version of stabilization for handlebody decompositions). This does not change the underlying manifold, only the particular choice of circle-valued Morse function in its homotopy class.

2.1.2 ϵ-Dependent smooth structures

As a matter of notational convenience, we set

$$\Sigma_\bullet := \bigsqcup_i \Sigma_i, \qquad Y_\bullet := Y \backslash \left(I \times \Sigma_\bullet \right).$$

We will refer to the connected components of the boundary $\partial Y_\bullet$ as the **seams**, and we will use the letter t to denote the coordinate variable on the interval I. In particular, $dt \in \Omega^1(I \times \Sigma)$ is identified with $df/|df| \in \Omega^1\left(\cup_i f^{-1}\left([r_i - 1/2, r_i + 1/2]\right)\right)$ under the identification (4.4).

Over $I \times \Sigma_\bullet$ the metric g has the form

$$dt^2 + g_\Sigma$$

where g_Σ is a path of metrics on $\Sigma_\bullet$. We assume that g has been chosen so that g_Σ is a *constant* path, which can always be achieved using the decomposition in (2.2) and a bump function. For $\epsilon > 0$ define a new metric

$$g_\epsilon := \begin{cases} dt^2 + \epsilon^2 g_\Sigma & \text{on } I \times \Sigma_\bullet \\ \epsilon^2 g & \text{on } Y_\bullet \end{cases}$$

We will be interested in taking the limit as ϵ approaches 0.

Let $\mathcal{S}_1$ denote the smooth structure on Y (i.e. the smooth structure in which g and f are smooth). We call this the **standard smooth structure**. It is important to note that when $\epsilon \neq 1$, the metric g_ϵ is *not* smooth with the standard smooth structure. For example, take $V = \nabla f / |\nabla f|$, where the norm and gradient are taken with respect to $g = g_1$. Then V is smooth on $(Y, \mathcal{S}_1)$, but

$$g_\epsilon(V, V) = \begin{cases} 1 & \text{on } I \times \Sigma_\bullet \\ \epsilon^2 & \text{on } Y_\bullet \end{cases} \tag{2.3}$$

is not even continuous, so g_ϵ cannot be continuous on $(Y, \mathcal{S}_1)$.

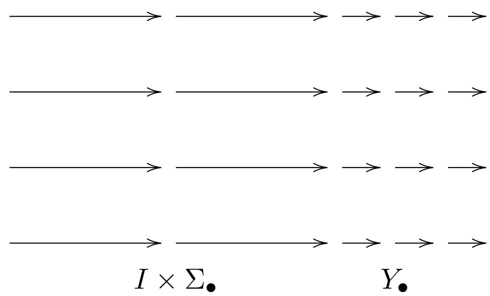


Figure 2.2: This vector field V is not continuous with respect to the topology on TY defined using the standard smooth structure. However, it is continuous with respect to the topology defined using the ϵ-dependent smooth structure.

However, there is a *different* smooth structure $\mathcal{S}_\epsilon$ in which g_ϵ is smooth, and $(Y, \mathcal{S}_\epsilon)$ is *diffeomorphic* to $(Y, \mathcal{S}_1)$. This can be seen as follows: View Y as the topological manifold obtained from the cobordisms $Y_{i(i+1)}$ and $I \times \Sigma_i$ by using the identity map on Σ_i to glue along the seams $\{0,1\} \times \Sigma_i$, as in (2.2). Following Milnor [29, Theorem 1.4], any choice of collar neighborhoods determines a smooth structure, and any two choices are given by isotopic data. These isotopies determine a diffeomorphism between the smooth structures. The smooth structure $\mathcal{S}_1$ can be viewed as arising in this way by choosing collar neighborhoods of the seams determined by the time-δ gradient flow of f, and then using the identity to glue these neighborhoods on the overlap. Here $\delta > 0$ is arbitrary, but fixed. On the other hand, the smooth structure $\mathcal{S}_\epsilon$ arises by taking the time-δ gradient flow on the $Y_{i(i+1)}$ side of the seam $\{1\} \times \Sigma_i$, but the time-$\delta\epsilon$ gradient flow on the $[0,1] \times \Sigma_i$ side of the seam, and then gluing using the map

$$(t, \sigma) \mapsto (\epsilon t, \sigma). \tag{2.4}$$

So Milnor tells us that there is a diffeomorphism

$$F_\epsilon : (Y, \mathcal{S}_1) \longrightarrow (Y, \mathcal{S}_\epsilon) \tag{2.5}$$

and, in fact, there is a canonical choice given by taking the obvious straight line homotopy from the map (2.4) to the identity. The pullback metric $F_\epsilon^* g_\epsilon$ is smooth on $(Y, \mathcal{S}_1)$. Indeed, it is just $\epsilon^2 g$. In cases when it is necessary to remember the smooth structure we will write Y^ϵ for $(Y, \mathcal{S}_\epsilon)$. However, we will often abuse notation slightly and refer to g_ϵ as a metric on Y. When a function, section of a bundle, connection, etc. is smooth in the smooth structure $\mathcal{S}_\epsilon$ we will say that it is **ϵ-smooth**. In particular, if a is a function, section, etc., then a is ϵ-smooth if and only if the pullback $F_\epsilon^* a$ is smooth in the standard smooth structure.

Remark 2.1.1. *The discontinuity in (2.3) shows that the identity map* $\mathrm{Id} : Y \to Y$*, which is a homeomorphism, does* not *give a diffeomorphism from* $(Y, \mathcal{S}_1)$ *to* $(Y, \mathcal{S}_\epsilon)$*. This also illustrates the subtle way in which the* topology *of the tangent bundle* TM *to a smooth manifold* M *depends on the* smooth structure *of* M*: Letting* TY^ϵ *and* T^*Y^ϵ *denote tangent and cotangent bundles to* Y *with smooth structure* $\mathcal{S}_\epsilon$*, we have that* g_ϵ *is a continuous (in fact, smooth) section of the symmetric product bundle* $\mathrm{Sym}^2 T^*Y^\epsilon \to Y^\epsilon$ *and* V *is a continuous (in fact, smooth) section of* $TY^1 \to Y^1$*. However, (2.3) shows that* $g_\epsilon(V, V) = g_\epsilon(\mathrm{Id}_* V, \mathrm{Id}_* V)$ *is not continuous, so* $\mathrm{Id}_* : TY^1 \to TY^\epsilon$ *is not continuous, even though* $\mathrm{Id} : Y \to Y$ *is. (Though, of course,* $(F_\epsilon)_* : TY^1 \to TY^\epsilon$ *is smooth.)*

We can also see that, though g_ϵ *fails to be smooth on* Y^1*, it only fails to do so at the seams where the* $I \times \Sigma_i$ *glue to the* $Y_{i(i+1)}$*, and even there* g_ϵ *is smooth in the directions parallel to the seam. So the discontinuity illustrated in (2.3) is the only type of thing that goes wrong. The same holds for any function, section of a bundle, connection, etc. which is smooth with the smooth structure* $\mathcal{S}_\epsilon$*.*

Moreover, by passing to local coordinates, it is straightforward to show that, for $1 \leq p \leq \infty$*, every ϵ-smooth* function *on* Y^ϵ *is of Sobolev class* $W^{1,p}_{loc}(Y)$ *with respect to the standard smooth structure. This is because the underlying topologies are identical, so the function is continuous on* Y^1 *and, on the complement of the seams, it is 1-smooth with bounded derivative. However, in general, ϵ-smooth* forms *will only be in* L^p_{loc} *with respect to the standard smooth structure. Any form which is non-zero in directions transverse to the seam will necessarily have a jump discontinuity, and so taking a derivative transverse to the seam will introduce a delta function. This applies to connections as well.*

2.1.3 Bundles, metrics, and the Hodge star

Let X be a smooth manifold (possibly with boundary). Given a fiber bundle $E \to X$, we denote the space of smooth sections by $\Gamma(E)$.

Now suppose E is a vector bundle. Then we write

$$\Omega^\bullet(X,E) := \bigoplus_k \Omega^k(X,E), \qquad \Omega^k(X,E) := \Gamma(\Lambda^k T^*X \otimes E),$$

for the space of smooth E-valued forms on X. If E is equipped with a fiber-wise inner product $\langle\cdot,\cdot\rangle : E \otimes E \to \mathbb{R}$, then this combines with the wedge to form a bilinear map

$$\begin{array}{rcl} \Omega^j(X,E) \otimes \Omega^k(X,E) & \longrightarrow & \Omega^{j+k}(X) \\ \mu \otimes \nu & \longmapsto & \langle \mu \wedge \nu \rangle \end{array}$$

If X is compact and oriented then integrating defines a non-degenerate bilinear pairing on forms of dual degree:

$$\begin{array}{rcl} \Omega^k(X,E) \otimes \Omega^{n-k}(X,E) & \longrightarrow & \mathbb{R} \\ \mu \otimes \nu & \longmapsto & \int_X \langle \mu \wedge \nu \rangle, \end{array} \tag{2.6}$$

where $n := \dim(X)$. If, in addition, X is equipped with a metric, then this combines with the orientation to induce a Hodge star $* : \Omega^k(M,E) \to \Omega^{n-k}(M,E)$, which satisfies

$$** = (-1)^{k(n-k)} : \Omega^k(X,E) \longrightarrow \Omega^k(X,E). \tag{2.7}$$

Sticking $*$ in the second slot of (2.6) defines the following L^2-inner product on the vector space $\Omega^k(X,E)$:

$$(\mu,\nu) := \int_X \langle \mu \wedge *\nu \rangle$$

for $\mu,\nu \in \Omega^k(X,E)$. We then set

$$\|\mu\|_{L^2(E)} := \sqrt{(\mu,\mu)}.$$

Similarly, for $p \geq 1$ we can define the L^p norm by

$$\|\mu\|^p_{L^p(E)} := \int_X |\mu|^p \, d\mathrm{vol},$$

where we have set

$$|\mu| := (*\langle\mu \wedge *\mu\rangle)^{1/2}.$$

Remark 2.1.2. *When the bundle E is clear from context we will often write L^p instead of $L^p(E)$. Similarly, it will be convenient to abuse notation and write $L^p(X)$ instead of $L^p(E)$ in cases where it is important to emphasize the underlying base manifold.*

Below we will be interested in various scalings of product metrics, so it will be useful to establish a few formulas. Suppose that $X = M \times N$ is a product, and M, N are oriented manifolds equipped with metrics g_M, g_N. Assume X is equipped with the product metric $g = g_M \oplus g_N$ and product orientation (our convention is to use the 'left-to-right convention': vectors on M come first, then vectors on N). The metrics and orientations on M and N induce Hodge stars $*_M$ and $*_N$, respectively, and these satisfy

$$*(\mu \wedge \nu) = (-1)^{k(\dim(M)-j)} *_M \mu \wedge *_N \nu,$$

where $\mu \in \Omega^j(M)$ and $\nu \in \Omega^k(N)$.

Example 2.1.3. *Suppose $X = \mathbb{R} \times N$, and let ds denote the standard 1-form on $\mathbb{R}$. Then*

$$*\nu = (-1)^k ds \wedge *_N \nu, \qquad * \, ds = *_N(1), \qquad * \, (ds \wedge \nu) = *_N \nu,$$

where $\nu \in \Omega^k(N)$ is a k-form on N. If N is equipped with a vector bundle $E \to N$ then the same formula holds for $\nu \in \Omega^k(N, E)$.

For a real number $c > 0$, let $*_c$ denote the Hodge star associated to the conformally scaled metric $c^2 g$ on X. Then

$$*_c\mu = c^{\dim(X)-2k} * \mu. \tag{2.8}$$

for $\mu \in \Omega^k(X, E)$.

Let G be a Lie group with Lie algebra $\mathfrak{g}$, and $\pi : P \to X$ a principal G-bundle. Given any matrix representation $\rho : G \to \mathrm{GL}(V)$, where V is a vector space, we can form the **associated bundle**

$$P(V) := P \times_G V = (P \times V)/G.$$

This is naturally equipped with the structure of vector bundle $P(V) \to X$. Of particular interest is the case when $V = \mathfrak{g}$ is the Lie algebra of G, and ρ is the adjoint representation. We will assume that $\mathfrak{g}$ is equipped with an Ad-invariant inner product $\langle \cdot, \cdot \rangle$. This is always the case if G is compact or simple. The Ad-invariance implies that the inner product ascends to a well-defined fiber-wise inner product on the vector bundle $P(\mathfrak{g})$.

The Lie bracket $[\cdot, \cdot] : \mathfrak{g} \otimes \mathfrak{g} \to \mathbb{R}$ is Ad-invariant, and so combines with the wedge to define a graded-commutative bilinear map

$$\begin{array}{rcl} \Omega^j(X, P(\mathfrak{g})) \otimes \Omega^k(X, P(\mathfrak{g})) & \longrightarrow & \Omega^{j+k}(X, P(\mathfrak{g})) \\ \mu \otimes \nu & \longmapsto & [\mu \wedge \nu] \end{array}$$

thereby equipping $\Omega^\bullet(X, P(\mathfrak{g}))$ with the structure of a graded algebra. Pullback by $\pi : P \to X$ induces a linear map

$$\pi^* : \Omega^\bullet(X, P(\mathfrak{g})) \longrightarrow \Omega^\bullet(P, \mathfrak{g}).$$

The Lie bracket defines a graded algebra structure on the codomain, and with this structure π^* is a graded algebra homomorphism. Moreover, π^* is injective with image given by the **basic** forms

$$\Omega^\bullet(P, \mathfrak{g})_{\mathrm{basic}} := \left\{ \mu \in \Omega^\bullet(P, \mathfrak{g}) \;\middle|\; \begin{array}{ll} (g_P)^*\mu = \mathrm{Ad}(g^{-1})\mu, & \forall g \in G \\ \iota_{\xi_P}\mu = 0, & \forall \xi \in \mathfrak{g} \end{array} \right\}.$$

Here g_P (resp. ξ_P) is the image of $g \in G$ (resp. $\xi \in \mathfrak{g}$) under the map $G \to \mathrm{Diff}(P)$ (resp. $\mathfrak{g} \to \mathrm{Vect}(P)$) afforded by the group action. We will often use π^* to identify the spaces $\Omega^\bullet(X, P(\mathfrak{g}))$ and $\Omega^\bullet(P, \mathfrak{g})_{\mathrm{basic}}$.

2.2 Gauge theory

Denote by

$$\mathcal{A}(P) = \left\{ A \in \Omega^1(P, \mathfrak{g}) \,\middle|\, \begin{array}{ll} (g_P)^* A = \mathrm{Ad}(g^{-1})A, & \forall g \in G \\ \iota_{\xi_P} A = \xi, & \forall \xi \in \mathfrak{g} \end{array} \right\}$$

the space of **connections** on P. The space of connections $\mathcal{A}(P)$ is an affine space modeled on $\Omega^1(X; P(\mathfrak{g}))$, and the affine action is given by:

$$\begin{array}{rcl} \Omega^1(X, P(\mathfrak{g})) \times \mathcal{A}(P) & \longrightarrow & \mathcal{A}(P) \\ (\mu, A) & \longmapsto & A + \pi^*\mu. \end{array}$$

In particular, $\mathcal{A}(P)$ is a smooth (infinite dimensional) manifold with tangent space $\Omega^1(X, P(\mathfrak{g}))$. Each connection $A \in \mathcal{A}(P)$ determines a covariant derivative

$$\begin{array}{rcl} d_A : \Omega^\bullet(X, P(\mathfrak{g})) & \longrightarrow & \Omega^{\bullet+1}(X, P(\mathfrak{g})) \\ \mu & \longmapsto & (\pi^*)^{-1}\left(d\left(\pi^*\mu\right) + [A \wedge \pi^*\mu]\right) \end{array}$$

where d is the trivial connection on the trivial bundle $P \times \mathfrak{g}$. Composing d_A with itself we obtain a degree two map $F_A := d_A \circ d_A : \Omega^\bullet(X, P(\mathfrak{g})) \to \Omega^\bullet(X, P(\mathfrak{g}))$, called the **curvature** of A. This can be computed as follows

$$F_A = (\pi^*)^{-1}\left(dA + \frac{1}{2}[A \wedge A]\right) \in \Omega^2(X, P(\mathfrak{g}))$$

Here and below we use the convention that the algebra $\Omega^\bullet(X, P(\mathfrak{g}))$ acts on itself by left multiplication. So, for example, the statement $d_A \circ d_A = F_A$ means

$$d_A(d_A(\mu)) = [F_A \wedge \mu], \qquad \mu \in \Omega^\bullet(X, P(\mathfrak{g})).$$

As noted above, if $A \in \mathcal{A}(P)$ and $\mu \in \Omega^1(X, P(\mathfrak{g}))$ then $A + \pi^*\mu \in \mathcal{A}(P)$ is another connection. However, in identifying $\Omega^\bullet(X, P(\mathfrak{g}))$ with $\Omega^\bullet(P, \mathfrak{g})_{\text{basic}}$ we will typically drop π^* from the notation. The covariant derivative and curvature satisfy the following

$$\begin{aligned} d_{A+\mu} &= d_A + \mu \\ F_{A+\mu} &= F_A + d_A\mu + \tfrac{1}{2}\left[\mu \wedge \mu\right]. \end{aligned}$$

Another useful formula is the Bianchi identity, which says

$$d_A F_A = 0.$$

Given a connection $A \in \mathcal{A}(P)$ with covariant derivative d_A, we define the **formal adjoint**

$$\begin{aligned} d_A^* : \Omega^k(X, P(\mathfrak{g})) &\longrightarrow \Omega^{k-1}(X, P(\mathfrak{g})) \\ \mu &\longmapsto -(-1)^{(n-k)(k-1)} * d_A * \mu. \end{aligned}$$

Stokes' theorem shows that this satisfies

$$(d_A\mu, \nu) = (\mu, d_A^*\nu)$$

when X is closed and $\mu \in \Omega^k(X, P(\mathfrak{g})), \nu \in \Omega^{k+1}(X, P(\mathfrak{g}))$. By (2.8) it follows that

$$d_A^{*,c} = c^{-2} d_A^*,$$

where $d_A^{*,c}$ is the adjoint defined with respect to the metric $c^2 g$; that is, using $*_c$ in place of $*$.

Suppose X is an oriented Riemannian 4-manifold. Then (2.7) shows that on 2-forms the Hodge star squares to the identity, and it has eigenvalues ± 1. Denoting by $\Omega^{2,\pm}(X, P(\mathfrak{g}))$ the ± 1 eigenspace of $*$, then we have an L^2-orthogonal decomposition

$$\Omega^2(X, P(\mathfrak{g})) = \Omega^{2,+}(X, P(\mathfrak{g})) \oplus \Omega^{2,-}(X, P(\mathfrak{g})),$$

with the orthogonal projection to $\Omega^{2,\pm}(X, P(\mathfrak{g}))$ given by $\mu \mapsto \frac{1}{2}(1 \pm *)\mu$. The elements of $\Omega^{2,-}(X, P(\mathfrak{g}))$ are called **anti-self dual** 2-forms. A connection $A \in X$ is said to be **anti-self dual (ASD)** or an **instanton** if its curvature $F_A \in \Omega^{2,-}(X, P(\mathfrak{g}))$ is an anti-self dual 2-form; that is, if

$$F_A + *F_A = 0.$$

In dimensions 2 and 3, we will be interested in the **flat** connections A. By definition, these satisfy $F_A = 0$, and we will denote the set of flat connections on P by $\mathcal{A}_{\mathrm{flat}}(P)$. If A is flat then $\mathrm{Im}\, d_A \subseteq \ker d_A$ and we can form the **harmonic spaces**

$$H_A^k := H_A^k(X, P(\mathfrak{g})) := \frac{\ker\left(d_A|\Omega^k(X, P(\mathfrak{g}))\right)}{\mathrm{Im}\,(d_A|\Omega^{k-1}(X, P(\mathfrak{g})))}.$$

We say that A is **irreducible** if $H_A^0 = 0$. By the formula (2.14) it is clear that if A is irreducible, then the gauge action at A is locally free. A similar condition that we will be interested in is when X is 3-dimensional. In that case we say that a connection A is **non-degenerate** if $H_A^1 = 0$. (The importance of this condition comes in that if X is closed and orientable then A is non-degenerate as a flat connection if and only if it is a non-degenerate critical point of the Chern-Simons functional. We will have more to say about this below.)

Suppose now that X is closed. Then the Hodge isomorphism [44, Theorem 6.8] says

$$H_A^\bullet \cong \ker(d_A \oplus d_A^*), \qquad \Omega^\bullet(X, P(\mathfrak{g})) \cong H_A^\bullet \oplus \mathrm{Im}(d_A) \oplus \mathrm{Im}(d_A^*), \tag{2.9}$$

for any flat connections A on X. Here the decomposition is orthogonal with respect to the L^2 inner product defined above. We will treat these isomorphisms as identifications. From the first isomorphism in (2.9) we see that $H_A^\bullet$ is finite dimensional since $d_A \oplus d_A^*$ is elliptic (locally its leading order term is $d \oplus d^*$). Furthermore, it is clear that $* : H_A^\bullet \to H_A^{n-\bullet}$ restricts to an isomorphism on the harmonic spaces, and so the pairing (2.6) continues to be non-degenerate when restricted to the harmonic spaces.

Example 2.2.1. *Suppose $X = \Sigma$ is a closed, oriented surface equipped with a metric. Then on 1-forms the pairing (2.6) is anti-symmetric and $*$ squares to -1. This data defines a complex structure on $\Omega^1(\Sigma, P(\mathfrak{g}))$. Furthermore, the pairing in (2.6) determines* *a symplectic structure* ω *on* $\Omega^1(\Sigma, P(\mathfrak{g}))$ *and the triple* $(\Omega^1(\Sigma, P(\mathfrak{g})), *, \omega)$ *is Kähler.*

As mentioned above, for any flat connection A *on* X *the Hodge star and non-degenerate pairing (2.6) restrict to the harmonic spaces* $H_A^\bullet$*. In particular,* $(H_A^1, *, \omega)$ *is Kähler and finite dimensional.*

We end this section with a brief discussion of the **Yang-Mills functional**, which is defined to be the map

$$\begin{aligned} \mathcal{YM} : \mathcal{A}(P) &\longrightarrow \mathbb{R} \\ A &\longmapsto \frac{1}{2}\|F_A\|_{L^2}^2 \end{aligned}$$

We will refer to the value as the **energy** of a connection. When X is closed, the critical points of this functional are precisely the connections A satisfying

$$d_A^* F_A = 0.$$

These connections are called the **Yang-Mills connections**. Clearly if a flat connection exists, then it is Yang-Mills and has minimum energy among all connections. However, in dimensions greater than 3 there need not exist any flat connections. For example, in dimension 4 the Bianchi identity shows that the ASD connections are Yang-Mills as well, and it is straight-forward from the definitions that the ASD connections are those with the minimum energy. In particular, the existence of a non-trivial instanton on a bundle over a closed 4-manifold precludes the existence of a flat connection on that bundle.

A **gauge transformation** is an equivariant bundle map $u : P \to P$ covering the identity. The set of gauge transformations on P forms a Lie group, called the **gauge group**, and is denoted $\mathcal{G}(P)$. There are two other equivalent ways of viewing gauge transformations. The first as the set $\mathrm{Map}(P, G)^G$ of G-equivariant maps $P \to G$. Here G acts on itself by conjugation. The relationship with the first definition is given by sending $u : P \to P$ to the map $g_u : P \to G$ defined by the following formula

$$u(p) = g_u(p) \cdot p$$

The third equivalent formulation is as the set of section $\Gamma(P \times_G G)$ of the group bundle $P \times_G G \to X$, which is formed by taking the action of G on itself by conjugation. It is an easy exercise to show $\mathrm{Map}(P, G)^G = \Gamma(P \times_G G)$.

The exponential map $\exp : \mathfrak{g} \to G$ intertwines the adjoint action with conjugation, and so defines a map $\Omega^0(X, P(\mathfrak{g})) \to \Gamma(P \times_G G) = \mathcal{G}(P)$, where the image of $\xi \in \Omega^0(X, P(\mathfrak{g}))$ is the section

$$x \longmapsto \exp(-\xi(x))$$

of $P \times_G G$. (The minus sign here is a convention used to make (2.14) free of signs.) Viewing $\mathcal{G}(P)$ as G-equivariant bundle maps $P \to P$, the element $\exp(-\xi) : P \to P$ acts by sending

$$p \longmapsto p \cdot \exp(-\xi(p)),$$

where we are now thinking of ξ as an element in $\Omega^0(P, \mathfrak{g})_{\text{basic}} \cong \Omega^0(X, P(\mathfrak{g}))$. The exponential map allows us to define a smooth structure on $\mathcal{G}(P)$ making $\mathcal{G}(P)$ into a Lie group with Lie algebra $\Omega^0(X, P(\mathfrak{g}))$.

Remark 2.2.2. *Just as with the theory of finite-dimensional Lie groups, there are various times when it is convenient to consider the gauge group and its Lie algebra as subsets of the same space. This can be achieved by fixing a matrix representation* $G \to \mathrm{GL}(\mathbb{C}^n) \subset \mathrm{End}(\mathbb{R})$. *This induces a Lie algebra representation* $\mathfrak{g} \to \mathrm{End}(\mathbb{R})$. *Then* $\Omega^0(X, P(\mathfrak{g}))$ *and* $\mathcal{G}(P)$ *both map into* $\Gamma(P \times_G \mathrm{End}(\mathbb{R}))$. *If, in addition, the matrix representation is faithful, then these are embeddings.*

As an example of the usefulness of this vantage point, note that in the ambient space $\Gamma(P \times_G \mathrm{End}(\mathbb{R}))$ *it makes sense to multiply Lie group elements with Lie algebra elements. Hence, we can identify the tangent space* $T_u\mathcal{G}(P)$ *with the translation of the Lie algebra by* u

$$u\Omega^0(X, P(\mathfrak{g})).$$

This is entirely analogous to the finite-dimensional case.

The gauge group acts on the left on the space $\Omega^\bullet(P, \mathfrak{g})$ by pulling back by the inverse:

$$(u, A) \longmapsto uA := (u^{-1})^* A, \tag{2.10}$$

for $u \in \mathcal{G}(P)$, $A \in \Omega^\bullet(P, \mathfrak{g})$. Here the star indicates the pullback map induced by the bundle map $u^{-1} : P \to P$. Sometimes it will be useful to alternatively write $u(A)$ for uA. The action of $\mathcal{G}(P)$ on $\Omega^\bullet(P, \mathfrak{g})$ restricts to actions on $\Omega^\bullet(P, \mathfrak{g})_{\text{basic}}$ and $\mathcal{A}(P)$. Viewing a gauge transformation as a map $u : P \to G$ we can write this action as

$$(u^{-1})^* A = \mathrm{Ad}(g) A + D\left(\mathrm{L}_u\right) D(u^{-1}), \tag{2.11}$$

where $D\left(\mathrm{L}_g\right) : TG \to TG$ is the pushforward of the map $\mathrm{L}_g : G \to G$ given by left multiplication by $g \in G$, and $D(u^{-1})$ represents the pushforward of the map $P \to G$ given by $p \mapsto u^{-1}(p)$. That is, on the right-hand side of (2.11) we are viewing the gauge transformation u as a map $P \to G$ and the formula should be treated pointwise on the values of u. For example, $Du : TP \to TG$, and so $D\left(\mathrm{L}(u)\right)\ D(u^{-1}) \in \mathfrak{g}$. From the perspective of covariant derivatives the action of $\mathcal{G}(P)$ takes the form

$$d_{(u^{-1})^* A} = \mathrm{Ad}(u) d_A + D\left(\mathrm{L}_u\right) D(u^{-1}), \tag{2.12}$$

Similarly to the situation in Remark 2.2.2, fixing a matrix representation of G allows us to write this formula in a less notation-heavy way:

$$(u^{-1})^* A = uAu^{-1} + uD(u^{-1}), \tag{2.13}$$

where now the concatenation appearing on the right is just matrix multiplication. If the representation is faithful then there is no information lost in expressing the action (2.11) as we have in (2.13).

The infinitesimal action of $\mathcal{G}(P)$ at $A \in \mathcal{A}(P)$ takes the form

$$\begin{aligned} \Omega^0(X, P(\mathfrak{g})) &\longrightarrow \Omega^1(X, P(\mathfrak{g})) \\ \xi &\longmapsto d_A\xi \end{aligned} \tag{2.14}$$

More generally, the derivative of (2.10) at (u, A) is

$$\begin{aligned} u\Omega^0(X, P(\mathfrak{g})) \times \Omega^1(X, P(\mathfrak{g})) &\longrightarrow \Omega^1(M, P(\mathfrak{g})) \\ (u\xi, \eta) &\longmapsto \mathrm{Ad}(u)d_A\xi + \mathrm{Ad}(u)\eta \\ &= \left(d_{u^*A}(u\xi)\right)u^{-1} + u\eta u^{-1}. \end{aligned} \tag{2.15}$$

where, in writing this expression, we have chosen a faithful matrix representation of G as in Remark 2.2.2.

The gauge group also acts on the left on $\Omega^\bullet(X, P(\mathfrak{g}))$ by the pointwise adjoint action (so we are viewing $\mathcal{G}(P)$ as sections of $P \times_G G$), and $\pi^* : \Omega^\bullet(X, P(\mathfrak{g})) \to \Omega^\bullet(P, \mathfrak{g})$ intertwines the two actions. The curvature of $A \in \mathcal{A}(P)$ transforms under $u \in \mathcal{G}(P)$ by

$$F_{(u^{-1})^*A} = \mathrm{Ad}(u)F_A.$$

This shows that $\mathcal{G}(P)$ restricts to an action on $\mathcal{A}_{\text{flat}}(P)$ and, in 4-dimensions, the instantons.

2.2.1 Topology aspects of principal $\mathrm{PSU}(r)$-bundles

We first review some basic facts about $\mathrm{PSU}(r)$. By definition we have

$$\mathrm{PSU}(r) = \mathrm{SU}(r)/\mathbb{Z}_r = \mathrm{U}(r)/\mathrm{U}(1),$$

where $\mathbb{Z}_r$ and $\mathrm{U}(1)$ are the centers of $\mathrm{SU}(r)$ and $\mathrm{U}(r)$, respectively. It follows that $\mathrm{PSU}(r)$ has trivial center, and is connected and compact. Furthermore,

$$\pi_1(\mathrm{PSU}(r)) \cong \mathbb{Z}_r,$$

since $\mathrm{SU}(r)$ is simply-connected. Being the quotient of $\mathrm{SU}(r)$ by a discrete set, we have a Lie algebra isomorphism $\mathfrak{psu}(r) \cong \mathfrak{su}(r)$. Hence $\mathrm{PSU}(r)$ is simple and so there is a canonical choice of Ad-invariant metric $\langle \cdot, \cdot \rangle$ on $\mathrm{PSU}(r)$ given by declaring the highest coroot to have norm $\sqrt{2}$. We will always assume $\mathfrak{psu}(r)$ is equipped with this inner product. Explicitly, this is given by

$$\langle \mu, \nu \rangle = \frac{1}{4\pi^2} \mathrm{tr}(\mu \cdot \nu^*) = -\frac{1}{4\pi^2} \mathrm{tr}(\mu \cdot \nu),$$

where the trace is the one induced from the identification $\mathfrak{psu}(r) \cong \mathfrak{su}(r) \subset \mathrm{End}(\mathbb{C}^r)$. Having fixed an inner product, the adjoint can be viewed as a representation of the form $\mathrm{Ad} : \mathrm{PSU}(r) \to \mathrm{SO}(\mathfrak{psu}(r))$, and this is faithful. Finally, consider the action of $\mathrm{U}(r)$ on itself by conjugation. The center $\mathrm{U}(1)$ fixes every point in $\mathrm{U}(r)$, and so this action descends to an action of $\mathrm{PSU}(r)$ on $\mathrm{U}(r)$. This $\mathrm{PSU}(r)$ action fixes the subgroup $\mathrm{SU}(r) \subset \mathrm{U}(r)$.

In [56], L.M. Woodward exploited the adjoint representation to classify the principal $\mathrm{PSU}(r)$-bundles over spaces of dimension ≤ 4. This classification scheme assigns cohomology classes

$$t_2(P) \in H^2(X, \mathbb{Z}_r), \qquad q_4(P) \in H^4(X, \mathbb{Z})$$

to each principal $\mathrm{PSU}(r)$-bundle $P \to X$. For example, q_4 is the second Chern class of the complexified adjoint bundle $P(\mathfrak{g})_\mathbb{C} := P(\mathfrak{g}) \otimes \mathbb{C}$,

$$q_4(P) = c_2\left(P(\mathfrak{g})_\mathbb{C}\right). \tag{2.16}$$

where $\mathfrak{g} = \mathfrak{psu}(r)$. The class t_2 is defined as the mod r reduction of a suitable first Chern class. We will be mostly interested in the case where X is a smooth manifold, but these classes are defined for CW complexes as well.

Example 2.2.3. *When $r = 2$ we have $\mathrm{PSU}(r) = \mathrm{SO}(3)$, and the classes t_2 and q_4 are exactly the 2nd Stiefel-Whitney class and 1st Pontryagin class, respectively.*

We summarize the properties of these classes which we will need.

- If $\dim(X) \leq 4$ and X is a manifold, then two bundles P and P' over X are isomorphic if and only if $t_2(P) = t_2(P')$ and $q_4(P) = q_4(P')$;

- The class $t_2(P)$ is zero if and only if the structure group of P can be lifted to $\mathrm{SU}(r)$; that is, $P = \overline{P} \times_{\mathrm{SU}(r)} \mathrm{PSU}(r)$ for some principal $\mathrm{SU}(r)$-bundle $\overline{P} \to X$ [56, p. 517]. Here the action of $\mathrm{SU}(r)$ on $\mathrm{PSU}(r)$ is by left multiplication via the projection $\mathrm{SU}(r) \to \mathrm{PSU}(r)$.

- In the case of manifolds of dimension 2 or 3, the class q_4 is always zero, and t_2 determines a bijection between isomorphism classes of $\mathrm{PSU}(r)$-bundles and the space $H^2(X, \mathbb{Z}_r)$. In particular, if X is a closed oriented surface or an oriented elementary cobordism between two such surfaces, then t_2 is a bijection

$$t_2 : \left\{ \begin{array}{c} \text{isomorphism classes of} \\ \mathrm{PSU}(r)\text{-bundles over } X \end{array} \right\} \xrightarrow{\cong} \mathbb{Z}_r$$

- These classes satisfy

$$q_4(P) = t_2(P)^2 \qquad \mod r \tag{2.17}$$

- The classes t_2 and q_4 are functorial in the sense that they commute with pullback by maps $f : X' \to X$:

$$t_2(f^*P) = f^*t_2(P), \qquad q_4(f^*P) = f^*q_4(P).$$

As an application, we use these characteristic classes to study the components of the gauge group $\mathcal{G}(P)$ for a principal $\mathrm{PSU}(r)$-bundle $P \to X$. Donaldson notes the following fact. We include a proof here for convenience.

Proposition 2.2.4. *[8, Section 2.5.2] Let G be a compact Lie group, X a smooth manifold and $P \to G$ a principal G-bundle. Then there is a bijection between $\pi_0(\mathcal{G}(P))$ and isomorphism classes of principal G-bundles over $S^1 \times X$ which restrict to P on a* *fiber. This bijection is induced from the map which sends a gauge transformation u to the bundle*

$$Q_u := [0,1] \times P/(0,p) \sim (1, u(p)), \tag{2.18}$$

given by the mapping torus.

Proof. Well-defined: The isomorphism class of the bundle Q_u depends only on the path component of u in $\mathcal{G}(P)$. The gauge group is locally path-connected (it is locally modeled on the vector space consisting of sections of $P(\mathfrak{g})$), so the connected components are the path components, and the map (2.18) descends to a give a well-defined map from $\pi_0(\mathcal{G}(P))$ to the isomorphism classes of principal G-bundles over $S^1 \times X$ which restrict to P on a fiber.

Surjective: Suppose we are given any bundle $Q \to S^1 \times X$ with $Q|_{\{1\}\times X} = P$, and consider the obvious projection $\pi : Q \to S^1$. Since G is compact, Q admits a G-invariant metric. (This can be obtained by first choosing any metric $(\cdot, \cdot)$ and then declaring

$$(v, w)_{\mathrm{inv}} := \frac{1}{\mathrm{vol}(G)} \int_G (DR_g v, DR_g w) \, d\mathrm{vol}_G,$$

where DR_g is the pushforward of multiplication by $g \in G$ and we are using an invariant Haar measure to define the integral on G.) Let $\Phi_t : Q \to Q$ denote the time-t gradient flow of π, normalized so Φ_1 maps each fiber to itself (this is just saying the circle has length 1). The G-invariance implies that Φ is G-equivariant. Then $u := \Phi_1|_{\pi^{-1}(1)} : P \to P$ is the desired gauge transformation.

Injective: Suppose there is some $u \in \mathcal{G}(P)$ with $\Psi : Q_u \overset{\cong}{\to} S^1 \times P$. Let $\Phi_\bullet : I \times Q_u \to Q_u$ be the gradient flow as constructed in the previous paragraph, and $\pi : S^1 \times P \to P$ the projection. Then consider the composition

$$\pi \circ \Psi \circ \Phi_\bullet \circ \Psi^{-1}| : I \times P \longrightarrow P.$$

where, in the domain, we have set $P = \{1\} \times P \subset Q$. Since everything is equivariant, this is a path in $\mathcal{G}(P)$ from u to the identity map. $\square$

Now we combine this with L.M. Woodward's classification. Fix a principal PSU(r)-bundle $P \to X$ and assume X is a manifold of dimension ≤ 3. Let $u \in \mathcal{G}(P)$, define $Q_u \to S^1 \times X$ as above, and consider the classes

$$t_2(Q_u) \in H^2(S^1 \times X, \mathbb{Z}_r), \qquad q_4(Q_u) \in H^4(S^1 \times X, \mathbb{Z}_r).$$

By the Künneth formula, we have an isomorphism

$$H^k(S^1 \times X, R) \cong H^k(X, R) \oplus H^{k-1}(X, R),$$

where $R = \mathbb{Z}$ or $\mathbb{Z}_r$. The image of $t_2(Q_u)$ in $H^2(X, \mathbb{Z}_r)$ is exactly $t_2(P)$, so the dependence of $t_2(Q_u)$ on the gauge transformation u is contained entirely in the projection of $t_2(Q_u)$ to $H^1(X, \mathbb{Z}_r)$. We denote this projection by

$$\eta(u) \in H^1(X, \mathbb{Z}_r),$$

and call this the **parity** of u.

Now consider the class $q_4(Q_u)$. Note that the relation (2.17) gives

$$q_4(Q_u) = 2t_2(P) \smile ds \smile \eta(u), \qquad \mathrm{mod}\, r, \tag{2.19}$$

where $ds \in H^1(S^1, \mathbb{Z}_r)$ is the generator. This follows because $t_2(P)^2$ is a 4-form on a 3-manifold, and $ds^2 = 0$. In particular, $q_4(Q_u)$ is always even. Due to the dimensional restrictions on X, there is an isomorphism $H^4(S^1 \times X, \mathbb{Z}) \cong H^3(X, \mathbb{Z})$. If X is a closed connected oriented 3-manifold, then $H^3(X, \mathbb{Z}) \cong \mathbb{Z}$. In this case, declare $\deg(u) \in \mathbb{Z}$ to be the image of $q_4(Q_u)/2$ under this isomorphism; we call this the **degree** of u.

Proposition 2.2.4, and the properties of t_2 and q_4 immediately imply that the parity and degree detect the components of the gauge group:

- If $\dim X \leq 2$, or if $\dim X = 3$ and X is compact with non-empty boundary, then there is an injection

$$\eta : \pi_0(\mathcal{G}(P)) \quad \hookrightarrow \quad H^1(X, \mathbb{Z}_r).$$

- If $\dim X = 3$ and X is closed, connected and oriented, then there is an injection

$$(\eta, \deg) : \pi_0(\mathcal{G}(P)) \hookrightarrow H^1(X, \mathbb{Z}_r) \times \mathbb{Z}.$$

See [11], [15] for alternative realizations of the parity and degree.

Proposition 2.2.5. *Fix a gauge transformation $u : P \to \mathrm{PSU}(r)$. Then the following are equivalent:*

- $\eta(u) = 0$*;*
- $u : P \to \mathrm{PSU}(r)$ *lifts to a* $\mathrm{PSU}(r)$*-equivariant map* $\widetilde{u} : P \to \mathrm{SU}(r)$*;*
- *When restricted to the 1-skeleton of X, u is homotopic to the identity map.*

Moreover, if X is a compact connected oriented 3-manifold and $\eta(u) = 0$, then $\deg(u)$ *is divisible by r.*

Proof. Let $\gamma : X_1 \to X$ be a continuous map, where X_1 is a CW complex. Then any $u \in \mathcal{G}(P)$ defines a pullback gauge transformation $\gamma^* u \in \mathcal{G}(\gamma^* P)$. Moreover, it is immediate from the functoriality of t_2 that

$$\eta(\gamma^* u) = \gamma^* \eta(u) \in H^1(X_1, \mathbb{Z}_r).$$

Now suppose $X_1 \subset X$ is the 1-skeleton, and γ is the inclusion. Then $\gamma^* : H^1(X, \mathbb{Z}_r) \to H^1(X_1, \mathbb{Z}_r)$ is an isomorphism and so $\eta(u) = 0$ if and only if $\eta(\gamma^* u) = 0$. Since $\gamma^* u$ is a gauge transformation over a 1-dimensional space, this is equivalent to saying that $\gamma^* u \in \mathcal{G}_0(\gamma^* P)$ lies in the identity component. This shows the equivalence of the first and second items.

Set $G = \mathrm{PSU}(r)$ and consider the group bundles

$$P \times_G G \to X, \qquad P \times_G \mathrm{SU}(r) \to X,$$

where, in both cases, G is acting on the second factor by conjugation. There is a residual free action of the center $\mathbb{Z}_r \subset \mathrm{SU}(r)$ on the second bundle, and the projection

$$P \times_G \mathrm{SU}(r) \longrightarrow (P \times_G \mathrm{SU}(r))\,/\mathbb{Z}_r = P \times_G G,$$

is a principal $\mathbb{Z}_r$-bundle, and hence a normal covering space. The normal covering spaces of $P \times_G G$ correspond to normal subgroups of the fundamental group of $P \times_G G$. In particular, the $\mathbb{Z}_r$-covering space $P \times_G \mathrm{SU}(r)$ corresponds to the kernel of a homomorphism $\mu : \pi_1(P \times_G G) \to \mathbb{Z}_r$, which we view as an element of $H^1(P \times_G G, \mathbb{Z}_r)$.

Any gauge transformation u can be viewed as a map $X \to P \times_G G$. Then an equivariant lift of u, as described in the first item of the proposition, is exactly a lift $\widetilde{u}$ to $P \times_G \mathrm{SU}(r)$:

$$\begin{tikzcd} & P \times_G \mathrm{SU}(r) \arrow[d] \\ X \arrow[ur, "\widetilde{u}"] \arrow[r, "u"] & P \times_G G \end{tikzcd}$$

By the lifting property for covering spaces, u lifts if and only if the pullback $u^*\mu = 0$ vanishes as an element of $H^1(X, \mathbb{Z}_r)$. This happens if and only if $u_*\pi_1(X) \subset \pi_1(P \times_G G)$ lies in the image of $\pi_1(P \times_G \mathrm{SU}(r))$, and this happen if and only if the pulled back gauge transformation

$$\gamma^* u : X_1 \longrightarrow \gamma^* P \times_G G$$

lifts to a section

$$X_1 \longrightarrow \gamma^* P \times_G \mathrm{SU}(r).$$

The fibers of the bundle $\gamma^* P \times_G \mathrm{SU}(r) \to X_1$ are simply connected, so any section is homotopic to the identity. So u lifts exactly when its restriction to the 1-skeleton $\gamma^* u : X_1 \to \mathrm{PSU}(r)$ is homotopic to the identity. This shows the equivalence of the second and third items. The final assertion is immediate from (2.19).

$\square$

It will be useful to have an alternate characterization of the degree of a gauge transformation. To set this up, we first note the following Chern-Weil formula, which holds for a closed oriented 4-manifold X,[1]

$$q_4(P) = -r \int_X \langle F_A \wedge F_A \rangle \in \mathbb{Z}. \tag{2.20}$$

Here, A is any connection on $P \to X$; it follows from the Bianchi identity that this is independent of the choice of A. This formula is immediate the definition of q_4, together with the usual Chern-Weil formula for the second Chern class on $\mathrm{SU}(r)$ bundles.

Now suppose X is a closed connected oriented 3-manifold. Fix a connection $a_0 \in \mathcal{A}(P)$ and a gauge transformation $u \in \mathcal{G}(P)$. Let $a : I \to \mathcal{A}(P)$ be any path from a_0 to $u^* a_0$. This defines a connection A on $I \times P \to I \times X$ by declaring $A|_{\{s\}\times X} = a(s)$. Moreover, A descends to a connection on Q_u, so we have

$$\deg(u) \quad = \quad -\frac{r}{2} \int_{I\times X} \langle F_A \wedge F_A \rangle.$$

The curvature decomposes into components as $F_A = F_{a(s)} + ds \wedge \partial_s a(s)$, and so

$$\deg(u) = -r \int_I \left(\int_X \langle F_{a(s)} \wedge \partial_s a(s) \rangle \right). \tag{2.21}$$

We conclude this section by defining the bundles we will be considering in the sequel.

Proposition 2.2.6. *Let Y and $f : Y \to S^1$ be as in section 2. For each $d \in \mathbb{Z}_r$ there is a principal* $\mathrm{PSU}(r)$*-bundle $Q \to Y$ such that*

$$t_2(Q|_{f^{-1}(r)}) = d$$

[1] Here the particular normalization in (2.20) depends on our choice of inner product on $\mathfrak{g}$. For example, suppose $r = 2$, so $q_4 = p_1$ is the first Pontryagin class. In [9], Donaldson and Kronheimer use the Frobenius inner product $\langle \cdot, \cdot \rangle_{\mathrm{DK}} = -\mathrm{tr}(\cdot\,\cdot)$, and their formula (2.1.40, 41) reads

$$p_1(P) = -\frac{1}{2\pi^2} \int_X \langle F_A \wedge F_A \rangle_{\mathrm{DK}}.$$

Similarly, Dostoglou and Salamon use $\langle \cdot, \cdot \rangle_{\mathrm{DS}} = -4\mathrm{tr}(\cdot\,\cdot)$, so with this convention the first Pontryagin class takes the form

$$p_1(P) = -\frac{1}{8\pi^2} \int_X \langle F_A \wedge F_A \rangle_{\mathrm{DS}}.$$

Regardless of the normalization, the classes p_1 and q_4 are, of course, integral.

for any regular value $r \in S^1$*. Furthermore,* Q *depends only on* d *and the homotopy class of* f*, in the sense that it is independent of all other choices up to bundle isomorphism covering the identity on* Y*.*

Proof. This is basically just a patching argument: By the classification of $\mathrm{PSU}(r)$-bundles there are bundles $Q_{i(i+1)} \to Y_{i(i+1)}$ and $I \times P_i \to I \times \Sigma_i$ each restricting to bundles of the specified class d on each boundary component. Since d uniquely characterizes $\mathrm{PSU}(r)$-bundles on surfaces, up to isomorphism, for each i there are gauge transformations

$$Q_{i(i+1)}|_{\Sigma_i} \to \{1\} \times \Sigma_i, \qquad Q_{i(i+1)}|_{\Sigma_{i+1}} \to \{0\} \times \Sigma_{i+1},$$

which we use to glue all of these bundles together to form the desired bundle $Q \to Y$. $\square$

Remark 2.2.7. *We will be interested in bundles* Q *as in Proposition 2.2.6, where* d *is a generator of* $\mathbb{Z}_r$ *(see Theorems 2.2.15 and 2.2.16, below). We will use*

$$Q_{i(i+1)} := Q|_{Y_{i(i+1)}}, \qquad P_i := Q|_{\Sigma_i}$$

to denote the restrictions.

Let $F_\epsilon : Y^1 \to Y^\epsilon$ *be the diffeomorphism from (2.5), and define*

$$Q^\epsilon := (F_\epsilon^{-1})^* Q,$$

which is a smooth $\mathrm{PSU}(r)$*-bundle over* Y^ϵ *enjoying the same properties as* Q*. Since* F_ϵ *is the identity on* $Y_\bullet$ *and* $\{0\} \times \Sigma_\bullet$ *it follows that the restrictions*

$$Q_\bullet := Q^\epsilon|_{Y_\bullet}, \qquad P_\bullet := Q^\epsilon|_{\{0\}\times\Sigma_\bullet}$$

do not depend on ϵ*.*

2.2.2 Compactness results and gauge fixing

This section begins with a review of some basic facts about Sobolev spaces. We then move on to describe two foundational compactness results pertaining to connections with bounded curvature. The section ends with a discussion of the Coulomb and temporal gauge-fixing conditions. We refer the reader to [51, Appendix B] for more details and proofs of the various assertions.

Let X be an oriented Riemannian n-manifold, and $E \to X$ a vector bundle equipped with a connection ∇. Let $k \geq 0$ be an integer, and $p \geq 1$ a real number. Then we denote by $W^{k,p}(E)$ the closure of the space of compactly supported smooth sections in $\Gamma(E)$ with respect to the norm

$$\|e\|^p_{W^{k,p}(E)} := \sum_{0 \leq j \leq k} \|\nabla^j e\|^p_{L^p(E)}.$$

The vector space $W^{k,p}(E)$ equipped with this norm is a Banach space. Note that $W^{0,p}(E) = L^p(E)$. When $k < 0$ is a negative integer, we define

$$W^{k,p}(E) := \left(W^{-k,p^*}(E)\right)^*$$

to be the dual Banach space to $W^{-k,p^*}(E)$, where p^* is the conjugate exponent: $1/p^* + 1/p = 1$. When the bundle E is clear from context, we will write $W^{k,p}$ for $W^{k,p}(E)$. In situations where it is particularly important to emphasize the particular Sobolev space or the underlying base manifold, we will write $W^{k,p}$ or $W^{k,p}(X)$ instead of $W^{k,p}(E)$. See Remark 2.1.2.

As a vector space $W^{k,p}(E)$ is independent of the choice of connection ∇, and any two choices determine equivalent norms. The Sobolev embedding theorem states that if X is compact then $W^{k,p}(E)$ embeds as a topological vector space into $W^{k',p'}(E)$ whenever

$$k - \frac{n}{p} \geq k' - \frac{n}{p'}, \qquad \text{and} \qquad k > k' \geq 0. \tag{2.22}$$

Furthermore, this embedding is compact whenever the first inequality in (2.22) is strict; this means that any sequence which is bounded $W^{k,p}(E)$ has a subsequence which converges in $W^{k',p'}$. In the above we have assumed p and p' are both real numbers, however this result has an extension to $p' = \infty$, in which case we need to assume that the first inequality in (2.22) is strict. When this is the case, the image of $W^{k,p}(E) \hookrightarrow$

$W^{k',\infty}(E)$ lies in $C^{k'}(E) \subset W^{k',\infty}(E)$, the subspace of sections whose k'th derivative is continuous.

We will also be interested in understanding how these Sobolev spaces interact with multiplication. So we assume that the fibers of E are equipped with the structure of an algebra, which defines a smooth bundle map

$$E \otimes E \longrightarrow E, \qquad (f, g) \longmapsto fg.$$

In our applications, this will be given by the Lie bracket. Suppose $k \geq 0$ is a positive integer, and $1 \leq p, r, s < \infty$ satisfy

$$r, s \geq p, \qquad\qquad \frac{1}{r} + \frac{1}{s} < \frac{k}{n} + \frac{1}{p}.$$

Then fiberwise multiplication defines a bounded bilinear map

$$W^{k,r}(E) \otimes W^{k,s}(E) \longrightarrow W^{k,p}(E).$$

Now suppose $P \to X$ is a principal G-bundle such that $\mathfrak{g}$ admits an Ad-invariant inner product. This data determines a fiber-wise norm $|\cdot|$ on each vector bundle $\Lambda^k T^*X \otimes P(\mathfrak{g})$, and any choice of smooth reference connection $A_{\mathrm{ref}} \in \mathcal{A}(P)$ combines with the Levi-Civita connection to allow us to define the space $W^{k,p}\left(\Lambda^j T^*X \otimes P(\mathfrak{g})\right)$ of Sobolev class $W^{k,p}$ j-forms with values in $P(\mathfrak{g})$. Define

$$\mathcal{A}^{k,p}(P) := A_{\mathrm{ref}} + W^{k,p}\left(T^*X \otimes P(\mathfrak{g})\right).$$

Using the pullback π^* we have that the smooth connections $\mathcal{A}(P) \subset \mathcal{A}^{k,p}(P)$ form a dense subspace. Furthermore, the space $\mathcal{A}^{k,p}(P)$ is independent of the choice of $A_{\mathrm{ref}} \in \mathcal{A}(P)$, and the norm only depends on the choice of reference connection A_{ref} up to norm equivalence. The Sobolev embedding theorem carries over directly to the space $\mathcal{A}^{k,p}(P)$ of connections.

When X is compact the assignment $A \mapsto F_A$ is bounded as a map $\mathcal{A}^{k,p}(P) \to W^{k-1,p}(\Lambda^2 T^*X \otimes P(\mathfrak{g}))$ provided $p \geq n/(k+1)$ and $k \geq 1$. Likewise, if A is continuous the exterior derivative d_A defines a bounded linear map

$$d_A : W^{k,p}(\Lambda^\bullet T^*X \otimes P(\mathfrak{g})) \longrightarrow W^{k-1,p}(\Lambda^{\bullet+1}T^*X \otimes P(\mathfrak{g})).$$

When $A \in \mathcal{A}(P)$ is flat the isomorphisms (2.9) continue to hold in the $W^{k,p}$-completions of the relevant spaces. The direct sum decomposition remains L^2-orthogonal, even though the spaces may not be complete in the L^2 norm.

Remark 2.2.8. *Suppose X is closed (compact with no boundary) and $p \geq \dim X$. The curvature map $A \mapsto F_A$ extends to a bounded linear map $\mathcal{A}^{0,p}(P) \to W^{-1,p}(\Lambda^2 T^* \otimes P(\mathfrak{g}))$. To see this, fix a smooth reference connection $\mathcal{A}_{\mathrm{ref}}$. Then for $A \in \mathcal{A}^{0,p}(P)$, define*

$$F_A := F_{A_{\mathrm{ref}}} + d_{A_{\mathrm{ref}}}(A - A_{\mathrm{ref}}) + \frac{1}{2}\left[A - A_{\mathrm{ref}} \wedge A - A_{\mathrm{ref}}\right].$$

The first term is smooth and so clearly in $W^{-1,p}$. The derivative term $d_{A_{\mathrm{ref}}}(A - A_{\mathrm{ref}})$ should be interpreted distributionally (i.e., acting on W^{1,p^*}, where $1/p + 1/p^* = 1$), and so also lies in $W^{-1,p}$. The final term is the product of two L^p forms and so is an $L^{p/2}$ form. This latter space embeds into $W^{-1,p}$ whenever $p \geq \dim X$ (this is an easy consequence of Sobolev embedding for $k \geq 0$ mentioned above).

Note that if $A \in \mathcal{A}^{k,p}(P)$ with $k < 0$, then one runs into difficulty defining the curvature by this formula. This is due to the presence of the quadratic term, which acts by pointwise function multiplication, but the elements of $\mathcal{A}^{k,p}(P)$ are not all functions when $k < 0$.

When we have an embedding $G \subseteq \mathrm{U}(r)$, then we can define $\mathcal{G}^{k,p}(P)$ to be the subset of functions in $W^{k,p}(\mathrm{End}(\mathbb{C}^r))$ whose images lie in $G \subset \mathrm{U}(r) \subset \mathrm{End}(\mathbb{C}^r)$. Note that under such an embedding G necessarily has measure zero in $\mathrm{End}(\mathbb{C}^r)$; nonetheless, this is a meaningful definition whenever we are in the continuous range for Sobolev embedding (e.g. $kp \geq \dim X$ and $k \geq 2$, or $kp > 2$ and $k = 1$). If X is non-compact, then we write $\mathcal{A}^{k,p}_{loc}(P)$ and $\mathcal{G}^{k,p}_{loc}(P)$ for the locally $W^{k,p}$ sections.

The space $\mathcal{G}^{k,p}(P)$ forms a group when we are in the continuous range, and the group operations are smooth, making $\mathcal{G}^{k,p}(P)$ a Banach Lie group. Moreover, when this is the case, the group $\mathcal{G}^{k,p}(P)$ acts smoothly on $\mathcal{A}^{k-1,p}(P)$. See [51, Appendix B] for more details.

Below we include statements of two compactness theorems for connections. They are originally due to Uhlenbeck [43], however [51] is an excellent reference as well. For the second assertion in Uhlenbeck's Strong Compactness Theorem, we direct the reader to [8, Proposition 2.1]. In both theorems, we assume X is a Riemannian n-manifold and $P \to X$ is a principal G-bundle with G compact. We allow X to have boundary unless otherwise specified. If X is non-compact, then we assume there exists a sequence of compact subsets $X_\nu \subseteq X$ with $X_\nu \subseteq X_{\nu+1}$, $\cup X_\nu = X$ and each X_ν is a deformation retract of X.

Theorem 2.2.9. *(Weak Compactness) Suppose $P \to X$ be as above, and $1 < p < \infty$ is such that $p > n/2$. Let $(A_\nu)_{\nu \in \mathbb{N}} \subset \mathcal{A}^{1,p}_{loc}(P)$ be a sequence of connections with $\sup_\nu \|F_{A_\nu}\|_{L^p(X)} < \infty$. Then there is a subsequence (still denoted by $(A_\nu)_{\nu\in\mathbb{N}}$) and a sequence of gauge transformations $u_\nu \in \mathcal{G}^{2,p}_{loc}(P)$ such that $u_\nu^* A_\nu$ converges weakly in $W^{1,p}$ on compact sets to a limiting connection in $\mathcal{A}^{1,p}(P)$.*

Theorem 2.2.10. *(Strong Compactness) Suppose $P \to X$ is as above, but assume X has empty boundary. Let $1 < p < \infty$ be such that $p > n/2$, and we suppose in addition that $p > 4/3$ when $n = 2$. If $(A_\nu)_{\nu\in\mathbb{N}} \subset \mathcal{A}^{1,p}_{loc}(P)$ is a sequence of Yang-Mills connections with $\sup_\nu \|F_{A_\nu}\|_{L^p(X)} < \infty$, then there is a subsequence (still denoted by $(A_\nu)_{\nu\in\mathbb{N}}$) and a sequence of gauge transformations $u_\nu \in \mathcal{G}^{2,p}_{loc}(P)$ such that $u_\nu^* A_\nu$ converges in C^∞ on compact sets to a limiting smooth Yang-Mills connection.*

Furthermore, if $\dim X = 4$, each A_ν is ASD and $\sup_\nu \|F_{A_\nu}\|_{L^2(X)} < \infty$, then there is a finite set of points $X_0 \subset X$ such that, after passing to a subsequence and applying gauge transformations, the A_ν converge in C^∞ on compact subsets of $X \backslash X_0$. Moreover, there is a constant $\delta_0 > 0$ such that if $\|F_{A_\nu}\|_{L^2(X)} < \delta_0$ for all but finitely many ν, then X_0 is empty.

There are several places where we will find it convenient to choose a particular *gauge* for a connection A. That is, we replace A by u^*A, for some gauge transformation u, where u^*A now satisfies some desirable property.

The first example of this that we discuss is *temporal gauge.* For this, we suppose we are working on a product manifold $\mathbb{R} \times X$, and that there is a bundle $P \to X$. Then we equip $\mathbb{R} \times X$ with the pullback bundle $\mathbb{R} \times P$. Let s denote the coordinate variable on $\mathbb{R}$, and

$$\frac{\partial}{\partial s} \in \mathrm{Vect}(\mathbb{R} \times P)$$

the obvious vector field. Then a connection $A \in \Omega^1(\mathbb{R} \times P, \mathfrak{g})$ is in **temporal gauge** if its contraction with $\partial/\partial s$ vanishes

$$\iota_{\partial/\partial s} A = 0.$$

This terminology is justified, since, for each connection A there is a gauge transformation $u \in \mathcal{G}(\mathbb{R} \times P)$ with u^*A in temporal gauge. Indeed, viewing elements of $\mathcal{G}(\mathbb{R} \times P)$ maps $\mathbb{R} \to \mathcal{G}(P)$, then u can be taken to be the unique solution of the ODE

$$-\left(\iota_{\partial/\partial s} A\right) u = \frac{\partial}{\partial s} u, \qquad u(0) = \mathrm{Id},$$

where we have chosen a faithful representation of G to write this using matrix notation (see Remark 2.2.2). It follows that u can be taken to lie in the identity component $\mathcal{G}_0(\mathbb{R} \times P)$. It is useful to note that any connection $A \in \mathcal{A}(\mathbb{R} \times P)$ can be written in the form

$$A = a(s) + p(s)\, ds$$

for unique $a : \mathbb{R} \to \mathcal{A}(P)$ and $p : \mathbb{R} \to \Omega^0(X, P(\mathfrak{g}))$, where $ds \in \Omega^1(\mathbb{R} \times Y)$ is the obvious 1-form. Then A is in temporal gauge if and only if $p = 0$.

Next, we discuss *Coulomb gauge.* Fix a Riemannian manifold X and a principal G-bundle $P \to X$. Given connections $A, A_0 \in \mathcal{A}(P)$, we say that A is in **Coulomb gauge with respect to** A_0 if

$$d_{A_0}^*(A - A_0) = 0.$$

The theorem that is useful to us is the following (see [51, Theorem 8.1] for a proof):

Theorem 2.2.11. *(Coulomb Gauge) Suppose X is a closed Riemannian n-manifold and $P \to X$ is a principal G-bundle with G compact. Let $1 < p \leq q < \infty$ be such that*

$$p > \frac{n}{2}, \qquad \frac{1}{n} > \frac{1}{q} > \frac{1}{p} - \frac{1}{n}.$$

For any $A_0 \in \mathcal{A}^{1,p}(P)$ and $c_0 > 0$, there exist $\delta > 0$ and $C > 0$ such that the following holds: For every $A \in \mathcal{A}^{1,p}(P)$ with

$$\|A - A_0\|_{L^q} \leq \delta, \qquad \|A - A_0\|_{W^{1,p}} \leq c_0,$$

there exists a gauge transformation $u \in \mathcal{G}^{2,p}(P)$ with u^*A in Coulomb gauge with respect to A_0, and satisfying the following estimates

$$\|u^*A - A_0\|_{L^q} \leq C\|A - A_0\|_{L^q}, \qquad \|u^*A - A_0\|_{W^{1,p}} \leq C\|A - A_0\|_{W^{1,p}}.$$

2.2.3 Moduli spaces of flat connections

This section introduces various moduli spaces of flat connections which are the building blocks for the quilted Floer cohomology of our 3-manifold Y. It turns out that these moduli spaces are finite-dimensional and have natural symplectic structures.

Before defining the moduli spaces themselves, we review the relevant symplectic geometry. See [30] and [33] for more details on this material. A **symplectic manifold** is a pair $M = (M, \omega)$, where M is a smooth manifold, and $\omega \in \Omega^2(M)$, called the **symplectic form**, is a 2-form which is closed and non-degenerate. It follows that M is necessarily even-dimensional. A diffeomorphism $\varphi : M_0 \to M_1$ between two symplectic manifolds (M_0, ω_0) and (M_1, ω_1) is called a **symplectomorphism** if $\varphi^*\omega_1 = \omega_0$.

Example 2.2.12. *1. If $M = (M, \omega)$ is a symplectic manifold, then $M^- := (M, -\omega)$ is also a symplectic manifold.*

2. If (M_j, ω_j) are symplectic manifolds for $j = 0, 1$, then $M_0 \times M_1$ is symplectic with symplectic form given by $\pi_0^*\omega_0 + \pi_1^*\omega_1$, where $\pi_j : M_0 \times M_1 \to M_j$ is the projection.

An **almost complex structure** on a symplectic manifold M is a map $J \in \mathrm{End}(TM)$ which squares to minus the identity $J^2 = -\mathrm{Id}$ (so J is a complex structure on the tangent bundle TM). An almost complex structure J is said to be **compatible** with the symplectic form ω if $\omega(Jv, Jw) = \omega(v, w)$ for all $v, w \in T_pM$ and $\omega(v, Jv) > 0$ whenever $v \neq 0$. When this is the case, the assignment

$$v \otimes w \longmapsto \omega(v, Jw)$$

defines a metric on TM. We denote the induced norm on vectors by $|\cdot|_M$. The space of compatible almost complex structures is always non-empty and contractible. It follows that the first Chern class $c_1(M) := c_1(TM, J) \in H^2(M, \mathbb{Z})$ associated to the complex vector bundle (TM, J) is well-defined, and independent of the choice of compatible J.

The **minimal Chern number** of (M, ω) is defined to be

$$N := \inf \{k > 0 \,|\, c_1(A) = k, \qquad \text{for some } A \in \pi_2(M)\}.$$

In our applications this will be finite. We say that (M, ω) is **monotone** if there is a constant $\tau > 0$, called the **monotonicity constant**, such that

$$[\omega]\,(A) = \tau c_1(A), \qquad \text{for } A \in \pi_2(M),$$

where $[\omega]$ denotes the cohomology class of the closed form ω. For example, this is given by

$$[\omega]\,(A) = \int_{S^2} u^*\omega,$$

where $u : S^2 \to M$ represents $A \in \pi_2(M)$. The key point of these properties is that if M is monotone with monotonicity constant τ and if M has finite minimal Chern number N, then

$$[\omega]\,(A) \in (\tau N)\mathbb{Z} \tag{2.23}$$

for all $A \in \pi_2(M)$.

Given a symplectic manifold (M, ω), an embedded submanifold $L \hookrightarrow M$ is called **Lagrangian** if $\dim L = \frac{1}{2} \dim M$ and the symplectic form vanishes on L.

Example 2.2.13. *Suppose $\varphi : M_0 \to M_1$ is a symplectomorphism. Then its graph* $\mathrm{Graph}(\varphi) \subset M_0^- \times M_1$ *is Lagrangian. In particular, if $M_0 = M_1$ and φ is the identity, then this shows that the diagonal in $M_0^- \times M_0$ is Lagrangian.*

Suppose $L \subset M$ is Lagrangian. Given a map $u : (D^2, \partial D^2) \to (M, L)$, we can find a trivialization

$$u^* TM \cong D^2 \times T_{u(0)} M$$

that restricts to a symplectomorphism on the fibers. Then restricting to the boundary provides a loop

$$S^1 = \partial D^2 \longrightarrow \mathrm{Lag}(T_{u(0)} M) \tag{2.24}$$

into the Grassmannian of Lagrangian subspaces of $T_{u(0)} M$. It is well-known that

$$\pi_1(\mathrm{Lag}(T_{u(0)} M)) = \mathbb{Z},$$

and we declare $I(u) \in \mathbb{Z}$ to be the degree of the map (2.24). It follows that $I(u)$ is independent of all choices and depends only on the homotopy class of u. We therefore obtain a map

$$I : \pi_2(M, L) \longrightarrow \mathbb{Z}$$

called the **Maslov index**. There is a second map

$$\mathcal{A}_L : \pi_2(M, L) \longrightarrow \mathbb{R},$$

called the **(symplectic) action**, which is given by

$$\mathcal{A}_L(u) := \int_{D^2} u^* \omega.$$

Then the Lagrangian L is called **monotone** if it satisfies

$$2\mathcal{A}_L = \tau I$$

for some $\tau > 0$. Here τ is called the **monotonicity constant**. If L is monotone with monotonicity constant τ, then M must be monotone as well, with the *same* monotonicity constant τ. Moreover, if L is simply-connected and M is monotone, then L is automatically monotone. See [54, Lemma 4.1.1].

With this background in hand, we are at a place where we can introduce and discuss the various properties of the moduli space of flat connections. Fix a principal $\mathrm{PSU}(r)$-bundle $P \to X$, and we assume $\dim(X) \leq 3$. For $2 \leq q < \infty$, define

$$M(P) := \mathcal{A}^{1,q}_{\text{flat}}(P)/\mathcal{G}^{2,q}_0(P),$$

where $\mathcal{G}^{2,q}_0(P) \subseteq \mathcal{G}^{2,q}(P)$ is the identity component of the gauge group (it follows from section 2.2.1 that, in dimensions 2 and 3, a gauge transformation u lies in the identity component $\mathcal{G}^{2,q}_0(P)$ if and only if $\eta(u) = 0$ and $\deg(u) = 0$). Let

$$\Pi : \mathcal{A}^{1,q}_{\text{flat}}(P) \longrightarrow M(P)$$

be the quotient map. Then Π and $M(P)$ are independent of the choice of $2 \leq q < \infty$ in the sense that the diagram commutes

$$\begin{array}{ccc}
\mathcal{A}^{1,q}_{\text{flat}}(P) & \longrightarrow & \mathcal{A}^{1,2}_{\text{flat}}(P) \\
\Pi \big\downarrow & & \big\downarrow \Pi \\
\mathcal{A}^{1,q}_{\text{flat}}(P)/\mathcal{G}^{2,q}_0(P) & \longrightarrow & \mathcal{A}^{1,2}_{\text{flat}}(P)/\mathcal{G}^{2,2}_0(P)
\end{array}$$

and the bottom line is an isomorphism.

In favorable cases, the space $M(P)$ inherits the structure of a smooth finite-dimensional manifold. This is stated in Theorem 2.2.15 for surfaces and Theorem 2.2.16 for cobordisms. See [3] or [26] for more details in the case of surfaces. Over a closed manifold, this smooth structure is obtained, roughly, as follows (for notational convenience we are suppressing Sobolev exponents): The tangent space to $\mathcal{A}_{\text{flat}}(P)$ at α is given by $\ker(d_\alpha) \subset \Omega^1(X, P(\mathfrak{g}))$. By the Hodge decomposition (2.9), we can write

$$\ker(d_\alpha) = \mathrm{Im}(d_\alpha) \oplus H^1_\alpha,$$

where H^1_α is the (finite-dimensional) harmonic space. Infinitesimally, the gauge group acts by d_α, so at the linear level the quotient $\mathcal{A}_{\text{flat}}(P)/\mathcal{G}_0(P)$ looks like

$$\ker(d_\alpha)/\mathrm{Im}(d_\alpha) = H^1_\alpha.$$

The Coulomb gauge condition, Theorem 2.2.11, says that this infinitesimal description at the linear level carries through to the local level, thereby providing charts for $M(P)$. The only thing one needs to worry about is whether the gauge group acts freely. By choosing $P \to X$ to be a suitably non-trivial bundle (e.g., $\dim(X) = 2$, $G = \mathrm{PSU}(r)$ and $t_2(P) \in \mathbb{Z}_r$ is a generator), we have that $H^0_\alpha = 0$, which means that the infinitesimal gauge action is free. By analyzing the local behavior of this gauge action for $G = \mathrm{PSU}(r)$, it follows that the action of the identity component $\mathcal{G}_0(P)$ is free (the full gauge group does not act freely). We record this for later use.

Lemma 2.2.14. *[53] Fix $2 \leq q < \infty$. Let Σ be a closed, connected, oriented surface, and suppose $P \to \Sigma$ is a principal $\mathrm{PSU}(r)$-bundle, with $t_2(P) \in \mathbb{Z}_r$ a generator. Then all flat connections on P are irreducible. Moreover, for every flat connection $A \in \mathcal{A}^{1,q}_{\text{flat}}(P)$ the stabilizer of A in $\mathcal{G}^{2,q}_0(P)$ is trivial:*

$$\left\{ u \in \mathcal{G}^{2,q}_0(P) \,\middle|\, u^* A = A \right\} = \{e\}\,.$$

Any natural properties exhibited by the harmonic space H^1_α are then expected to be enjoyed by the moduli space $M(P)$. For example, in dimension 2, the harmonic space H^1_α is a symplectic vector space and so we expect $M(P)$ to be a symplectic manifold. The next theorem states that this is indeed the case. For a proof, see the first two paragraphs in the proof of [53, Theorem 3.3.2], together with the last two paragraphs in the proof of [53, Proposition 3.2.4].[2]

[2] In [53], the authors work with $G = \mathrm{U}(r)$ and the space of *central curvature connections with fixed determinant*, rather than the space of flat $\mathrm{PSU}(r)$ connections, as we consider here. Their theorems carry over verbatim to our situation. In fact, the space of central curvature $\mathrm{U}(r)$-connections with fixed

Theorem 2.2.15. *Let Σ be a closed, connected, oriented surface, and $P \to \Sigma$ a principal* $\mathrm{PSU}(r)$*-bundle with $t_2(P) \in \mathbb{Z}_r$ a generator.*

1. *[53, Prop. 3.2.4, Thm. 3.3.2] If Σ has genus $g(\Sigma) \geq 1$, then the moduli space $M(P)$ is a nonempty compact, symplectic manifold of dimension $(2g(\Sigma) - 2)(r^2 - 1)$, with even minimal Chern number, and with monotonicity constant $1/2r$. The tangent space at $[A] \in M(P)$ isomorphic to H^1_A and the symplectic form $\omega_{M(P)}$ is given by restricting the pairing (2.6). If $g(\Sigma) = 0$, then $M(P) = \emptyset$. Moreover, $M(P)$ is always connected and simply-connected.*[3]

2. *[53, Lemma 3.3.5] If $P' \to \Sigma$ is a second principal* $\mathrm{PSU}(r)$*-bundle then any* $\mathrm{PSU}(r)$*-equivariant bundle isomorphism $\psi : P \to P'$ covering the identity induces a symplectomorphism $\psi^* : M(P') \to M(P)$ by pullback. Furthermore, if $\phi : P \to P'$ is a second bundle map then $\psi^* = \phi^*$, so the moduli spaces $M(P)$ and $M(P')$ are canonically symplectomorphic.*

An immediate corollary of the second part of Theorem 2.2.15 is that $M(P)$ depends (up to canonical symplectomorphism) only on characteristic class $t_2(P) \in \mathbb{Z}_r$.

Since Σ is assumed to be oriented, any choice of metric on Σ induces a Hodge star, $*$. Then $*$ descends to a compatible complex structure on the tangent bundle $TM(P)$ (still denoted by $*$).

Given a symplectic manifold $M = (M, \omega)$, let M^- denote the symplectic manifold $(M, -\omega)$. Then reversing the orientation of the surface X in the previous theorem changes $M(P)$ to $M(P)^-$. Now suppose Y is a 3-manifold with non-empty boundary $\partial Y = \overline{\Sigma^-} \sqcup \Sigma^+$, having two connected components. Let $Q \to Y$ be a principal G-bundle. Then restriction to each boundary component induces a $\mathcal{G}(Q)$-equivariant map

$$\rho : \mathcal{A}(Q) \longrightarrow \mathcal{A}(Q|_{\Sigma^-}) \times \mathcal{A}(Q|_{\Sigma^+})$$

---

determinant is isomorphic to the space of flat $\mathrm{PSU}(r)$-connections, and this isomorphism intertwines the actions given by fixed determinant gauge transformations on the former space and $\mathcal{G}_0(P)$ on the latter (see [53, Lemma 3.2.5]).

[3] *The particular monotonicity constant depends on our choice of* Ad*-invariant metric on the Lie algebra to* $\mathrm{PSU}(r)$*. See section 2.2.1.*

that preserves the flat connections ($\mathcal{G}(Q)$ acts on the right by the restriction homomorphisms $\mathcal{G}(Q) \to \mathcal{G}(Q|_{\Sigma^-}) \times \mathcal{G}(Q|_{\Sigma^+})$). In particular, the map ρ descends to a map, still denoted by ρ, at the level of moduli spaces. This allows us to define

$$L(Q) := \rho(M(Q)) \subset M(Q|_{\Sigma^-})^- \times M(Q|_{\Sigma^+}).$$

Note that if Y is an elementary cobordism, then restriction to the boundary component $\Sigma^\pm$ provides an isomorphism $H^2(Y, \mathbb{Z}_r) \cong H^2(\Sigma^\pm, \mathbb{Z}_r) = \mathbb{Z}_r$. Due to the functoriality of the characteristic class t_2, it follows that for any $d \in \mathbb{Z}_r$ there is a principal $\mathrm{PSU}(r)$-bundle $Q \to Y$ with $t_2(Q) = t_2(Q|_{\Sigma^+}) = t_2(Q|_{\Sigma_-}) = d$.

Theorem 2.2.16. *Let Y be an elementary cobordism between closed connected oriented surfaces $\Sigma^\pm$, equipped with bundles $P^\pm \to \Sigma^\pm$ with the same characteristic class $t_2(P^-) = t_2(P^+) \in \mathbb{Z}_r$, which we assume is also a generator of $\mathbb{Z}_r$. Let $Q \to Y$ be a principal* $\mathrm{PSU}(r)$ *bundle that restricts to $P^\pm$ on $\Sigma^\pm$.*

1. [53, Theorem 3.4.1] The map $\rho : M(Q) \to L(Q) \subset M(P^-)^- \times M(P^+)$ is a Lagrangian embedding. Furthermore, $L(Q)$ is compact, oriented, simply-connected, and spin. In particular, $L(Q)$ is monotone.

2. [53, Lemma 3.4.4] The Lagrangian $L(Q)$ is independent of the choice of Q under the canonical symplectomorphisms of Theorem 2.2.15. If $Y = I \times \Sigma^-$ is a product cobordism from $\Sigma^- = \Sigma^+$ to itself, then $L(Q) \subset M(P^-)^- \times M(P^-)$ is the diagonal.

It follows from the second assertion in Theorem 2.2.16 that given flat connections on $P^\pm$ there is at most one flat connection on Q restricting to these connections, up to the action of $\mathcal{G}_0(Q)$.

Remark 2.2.17. *The moduli space of flat connections has an alternative description which is often quite useful. See, for example, [3] for more details. For simplicity, let X be a closed, connected manifold, and $P \to X$ a principal G-bundle with G compact. Fix a basepoint $x_0 \in X$. Then given a flat connection $\alpha \in \mathcal{A}_{\mathrm{flat}}(P)$, the based holonomy is a group homomorphism*

$$\pi_1(X, x_0) \longrightarrow G.$$

This defines a map

$$\mathcal{A}_{\text{flat}}(P) \longrightarrow \hom(\pi_1(X, x_0), G), \tag{2.25}$$

which intertwines the action of the gauge group on the left, and the action of G on the right (this latter action is given by conjugating the image values). So (2.25) descends to a map of the form

$$\mathcal{A}_{\text{flat}}(P)/\mathcal{G}(P) \longrightarrow \hom(\pi_1(X, x_0), G)/G. \tag{2.26}$$

Consider $\pi_1(X, x_0)$ equipped with the discrete topology. This endows

$$\hom(\pi_1(X, x_0), G)/G$$

with a topology for which (2.26) is continuous. Moreover, the map (2.26) is an injective local homeomorphism mapping onto a union of connected components. (Conversely, every connected component of $\hom(\pi_1(X, x_0), G)/G$ lies in the image of (2.26) for a suitably chosen bundle P.) Since X is compact, it follows that $\pi_1(X, x_0)$ is finitely generated, and so $\hom(\pi_1(X, x_0), G)/G$ is compact. In particular,

$$\mathcal{A}_{\text{flat}}(P)/\mathcal{G}(P) \tag{2.27}$$

is compact. (This whole discussion carries through with suitable Sobolev completions as well.)

Note that (2.27) is not *$M(P) := \mathcal{A}_{\text{flat}}(P)/\mathcal{G}_0(P)$, since in $M(P)$ is defined by only modding out by identity component of the gauge group. However, the projection*

$$M(P) \longrightarrow \mathcal{A}_{\text{flat}}(P)/\mathcal{G}(P),$$

is a principal $\pi_0(\mathcal{G}(P))$-bundle (if H is any topological group, then its set of connected components is naturally a group as well). Hence,

$$M(P)/\pi_0(\mathcal{G}(P)) \cong \mathcal{A}_{\text{flat}}(P)/\mathcal{G}(P)$$

is compact.

2.3 Floer cohomology

This section defines the instanton and quilted Floer cohomology groups associated to our 3-manifold Y. We begin by describing the former. Consider a principal G-bundle $Q \to Y$ over a closed oriented Riemannian 3-manifold Y, with G compact. Equip $\mathfrak{g}$ with an Ad-invariant inner product $\langle \cdot, \cdot \rangle$. There is a natural 1-form defined on the space of connections sending $v \in T_a\mathcal{A}(Q) = \Omega^1(Y, Q(\mathfrak{g}))$ to

$$\lambda_a(v) := \int_Y \langle F_a \wedge v \rangle.$$

This form is closed, and since $\mathcal{A}(Q)$ is contractible, it follows that λ is exact. Indeed, fixing a reference connection $a_0 \in \mathcal{A}(Q)$, we obtain a well-defined function

$$\mathcal{CS}_{a_0}(a_1) := \int_I \lambda_{a(s)}(\partial_s a(s))\, ds$$

where $a : I \to \mathcal{A}(Q)$ is any path with $a(j) = a_j$ for $j = 0, 1$. The function $\mathcal{CS}_{a_0}$: $\mathcal{A}(Q) \to \mathbb{R}$ is called the **Chern-Simons functional** and a computation shows

$$\mathcal{CS}_{a_0}(a_0 + v) := \frac{1}{2}\int_Y 2\langle F_{a_0} \wedge v \rangle + \langle d_{a_0} v \wedge v \rangle + \frac{1}{3}\langle [v \wedge v] \wedge v \rangle.$$

Moreover, $\mathcal{CS}_{a_0}$ only depends on a_0 up to an overall constant. This same discussion carries over when $\mathcal{A}(Q)$ is replaced by its $W^{1,p}$-completion $\mathcal{A}^{1,q}(Q)$, at least for q sufficiently large e.g., $q \geq 2$. It is convenient to take a_0 to be flat, though this is not strictly necessary. For $G = \mathrm{PSU}(r)$ we pick the canonical inner product described in section 2.2.1. Then it follows from the definition of $\mathcal{CS}$, and the formula (2.21), that

$$\mathcal{CS}_{a_0}(a) - \mathcal{CS}_{a_0}(u^* a) = \frac{1}{r}\deg(u)$$

for $a \in \mathcal{A}(Q)$ and $u \in \mathcal{G}(Q)$. Note that if $\eta(u) = 0$, then the right-hand side is an integer by Proposition 2.2.5.

From the definition of λ, it is clear the critical points of $\mathcal{CS}_{a_0}$ are precisely the flat connections. As described in section 2.2, any metric g on Y induces a metric on the space of connections given by the L^2 inner product. Then the gradient of $\mathcal{CS}_{a_0}$ with respect to this metric is the vector field $a \mapsto *F_a$.

Instanton Floer cohomology $HF_{\text{inst}}(Q)$ can be viewed as the Morse cohomology, modulo gauge, of the Chern-Simons functional. Moreover, when it is defined, $HF_{\text{inst}}(Q)$ is the cohomology associated to a chain complex $(CF_{\text{inst}}(Q), \partial_{\text{inst}})$. Here

$$CF_{\text{inst}}(Q) := \bigoplus_{[a] \in M(Q)} \mathbb{Z}_2 \langle [a] \rangle$$

is generated over $\mathbb{Z}_2$ by the gauge equivalence classes of flat connections on Q. (We have chosen to work with $\mathbb{Z}_2$ to avoid a discussion of orientations, which will not arise in this thesis.) The desirable cases are when all of the critical points have non-degenerate Hessians. In general, a flat connection a is non-degenerate (i.e. $H^1_a = 0$) if and only if the Hessian, when viewed as an operator on $\Omega^1(Y, Q(\mathfrak{g}))/\mathrm{Im}(d_a)$, is a non-degenerate quadratic form.

The boundary operator ∂_{inst} is given by a mod-2 count of isolated negative gradient trajectories of the Chern-Simons functional. To define this precisely, we need to digress a bit to discuss moduli spaces of instantons. The negative gradient trajectories of the Chern-Simons functional are solutions $a : \mathbb{R} \to \mathcal{A}(Q)$ to

$$\partial_s a = - * F_a. \tag{2.28}$$

This equation is plainly gauge invariant, so we consider solutions modulo $\mathcal{G}_0(Q)$. Alternatively, the path of connections $s \mapsto a(s)$ can be viewed as a single connection $A = a(s)$ on the bundle $\mathbb{R} \times Q \to \mathbb{R} \times Y$, in which case (2.28) is just the ASD equation from section 2.2 with the metric $ds^2 + g$ (hence the *instanton* in 'instanton Floer cohomology').

More generally, every connection on $\mathbb{R} \times Q$ has the form $A = a(s) + p(s)\, ds$ where $a : \mathbb{R} \to \mathcal{A}(Q)$ and $p : \mathbb{R} \to \Omega^0(Y, Q(\mathfrak{g}))$. The curvature decomposes into components as

$$F_A = F_a - (\partial_s a - d_a p) \wedge ds,$$

so the ASD equations for A take the form

$$\partial_s a - d_a p = - * F_a, \tag{2.29}$$

which reduces to (2.28) when $A = a(s) + p(s) = a(s)$ is in temporal gauge. So solutions to (2.28) modulo $\mathcal{G}_0(Q)$ are identical to solutions of (2.29) modulo $\mathcal{G}_0(\mathbb{R} \times Q)$. Furthermore, the following conditions are equivalent for an instanton $A = a(s) + p(s)$ on $\mathbb{R} \times Q$:

(i) The connection A has finite energy:

$$\mathcal{YM}(A) = \frac{1}{2} \int_Y |F_A|^2 < \infty.$$

(ii) The slice-wise curvature $F_{a(s)}$ decays exponentially to zero as $s \to \pm\infty$:

$$\|F_{a(s)}\|_{L^2(Y)} \leq C e^{-\kappa|s|}$$

for some constants $C, \kappa > 0$.

(iii) The connection A converges to flat connections $a^{\pm} \in \mathcal{A}_{\text{flat}}(Q)$ at $\pm\infty$:

$$\lim_{s \to \pm\infty} a(s) = a^{\pm}, \ \lim_{s \to \pm\infty} p(s) = 0;$$

here the convergence is in C^∞ on Y.

The proof proceeds roughly as follows (for more details see [8, Chapter 4]). For (iii) $\Rightarrow$ (i), it follows by direct computation that

$$\mathcal{YM}(A) = -\frac{1}{2} \int_{\mathbb{R} \times Y} \langle F_A \wedge F_A \rangle = \mathcal{CS}_{a_0}(a^-) - \mathcal{CS}_{a_0}(a^+) \tag{2.30}$$

for any reference connection a_0. Conversely, one considers the sequence $A_\nu := A|_{[\nu, \nu+1] \times Y}$, which we view as instantons on $[0, 1] \times Q$. Since A has finite energy it follows that $\|F_{A_\nu}\|_{L^2} \to 0$ as $\nu \to \pm\infty$, and so (iii) follows by Uhlenbeck's Strong Compactness Theorem 2.2.10. Clearly (ii) implies (i), and for the converse, one shows the function

$$J(S) := \int_S^\infty \|F_{a(s)}\|^2_{L^2(Y)}\, ds$$

satisfies

$$\frac{dJ}{dS} \leq -\kappa J + C_0 J^{3/2}$$

for some constants $\kappa, C_0 > 0$. This implies that J decays exponentially at $+\infty$, and then one uses elliptic estimates to deduce that the quantity $\|F_A\|^2_{L^2((S,S+1)\times Y)}$ does as well.

For $[a^\pm] \in M(Q)$, define

$$\mathcal{M}_Q([a^-], [a^+]) := \left\{ A = a(s) + p(s)\, ds \;\middle|\; \begin{array}{c} \partial_s a - d_a p = - * F_a \\ \lim_{s\to\pm\infty} A|_{\{s\}\times Q} \in [a^\pm] \end{array} \right\} \Big/ \mathcal{G}_0(\mathbb{R}\times Q)$$

to be the set of gauge equivalence classes of finite energy ASD connections limiting to $a^\pm$, modulo $\mathcal{G}_0(Q)$, at $\pm\infty$. The space $\mathcal{M}_Q([a^-], [a^+])$ admits an action of $\mathbb{R}$ by translation, and we set

$$\widehat{\mathcal{M}}_Q([a^-], [a^+]) := \mathcal{M}_Q([a^-], [a^+])/\mathbb{R}.$$

The next theorem, originally due to Floer, is the mechanism that makes instanton cohomology well-defined:

Theorem 2.3.1. *Fix $2 \leq q < \infty$. Let $Q \to Y$ be a principal $\mathrm{PSU}(r)$-bundle over a closed connected oriented 3-manifold Y. Assume all flat connections are non-degenerate. Suppose in addition that there is an embedding $\iota : \Sigma \to Y$ of a closed oriented connected surface Σ such that the characteristic class $t_2(\iota^* Q) \in \mathbb{Z}_r$ is a generator. Then the following hold.*

1) The set $M(Q) = \mathcal{A}^{1,q}_{\text{flat}}(Q)/\mathcal{G}^{2,q}_0(Q)$ is finite.

2) For a comeager subset of metrics on Y (in the C^∞ topology), the moduli space $\mathcal{M}_Q([a^-], [a^+])$ is a smooth finite dimensional manifold for every pair $[a^-], [a^+] \in M(Q)$. Denote by $\mu_{\text{inst}}(a^-, a^+) := \dim \mathcal{M}_Q([a^-], [a^+])$ the dimension of this space.

3) Whenever $\mu_{\text{inst}}(a^-, a^+) = 1$, the action of $\mathbb{R}$ on $\mathcal{M}_Q([a^-], [a^+])$ is free and the quotient space

$$\widehat{\mathcal{M}}_Q([a^-], [a^+]) := \mathcal{M}_Q([a^-], [a^+])/\mathbb{R}$$

is compact, zero-dimensional, and hence finite.

The proof of Theorem 2.3.1 is almost identical to the one given by Floer in [15], where he considers non-trivial PSU(2)-bundles (see also [14] for similar results). The additional information needed to generalize Floer's result to $G = \mathrm{PSU}(r)$ for $r > 2$ can be expressed in the form of the following lemma. (An explicit appearance of the analogous statement for PSU(2)-bundles can be found in [11, Lemma 2.5].)

Lemma 2.3.2. *[11] Fix $2 \leq q < \infty$. Let X be a closed, orientable 3-manifold and $Q \to X$ a principal $\mathrm{PSU}(r)$-bundle. Suppose, in addition, that there is an embedding of a closed, connected, oriented surface $\iota : \Sigma \to X$ such that $t_2(\iota^* Q) \in \mathbb{Z}_r$ is a generator. Then all of the flat connections on Q are irreducible. Moreover, for every flat connection $A \in \mathcal{A}^{1,q}_{\text{flat}}(Q)$ the stabilizer of A in $\mathcal{G}^{2,q}_0(Q)$ is trivial:*

$$\left\{ u \in \mathcal{G}^{2,q}_0(Q) \,\middle|\, u^* A = A \right\} = \{e\}\,.$$

The proof of this lemma follows almost exactly as in the proof of [11, Lemma 2.5], with Lemma 2.2.14 replacing [11, Lemma 4.1].

Remark 2.3.3. *(a) In general, it need not be the case that all flat connections on Q are non-degenerate. Consequently the above theorem is vacuous in such cases. To have a non-trivial theorem, one needs to perturb the Chern-Simons functional as follows (here we follow [14], but another good reference is [8, Section 5.5]): Consider $\bigvee_{i=1}^m S^1$, the wedge sum of m circles, which we view as embedded in $\mathbb{R}^3$ in such a way that their only intersection point is at the origin, and they all have the same tangent vector there. Given an embedding $\gamma : (\bigvee_{i=1}^m S^1) \times D^2 \to Y$, and a point $z \in D^2$, define $\gamma_z := \gamma(\cdot, z)$ to be the restriction. The holonomy around the circles in the image of γ_z provides a map $\mathcal{A}(Q) \to G^m$, which we denote by $\gamma_z(a)$. Fix a non-negative 2-form μ on D^2*

with compact support in the interior, and which integrates to 1. Then, given any Ad-*invariant map* $h : G^m \to G$, *one can define*

$$\begin{aligned} h_\gamma : \mathcal{A}(Q) &\longrightarrow \mathbb{R} \\ a &\longmapsto \int_{D^2} h(\gamma_z(a))\, \mu, \end{aligned}$$

which is gauge invariant. We consider perturbations of the Chern-Simons functional that have the form $\mathcal{CS}_{a_0} + h_\gamma : \mathcal{A}(Q) \to \mathbb{R}$. *One then repeats the Morse-theoretic discussion above by considering the chain complex generated by the critical points of* $\mathcal{CS}_{a_0} + h_\gamma$, *and the boundary operator defined by counting gradient trajectories of this perturbed function (these satisfy a perturbed ASD equation) which have finite (perturbed) energy. Floer's theorem [14, Theorem 1c.1] says that, for a dense set of choices* (h, γ), *the conclusions of Theorem 2.3.1 continue to hold. Likewise, the conclusions of Theorem 2.3.4, below, continue to hold under this hypothesis as well. Furthermore, any two choices of such perturbations yield canonically isomorphic cohomology groups [14, Theorem 2].*

(b) Suppose we are in the case of Theorem 2.3.1. When it is non-empty, the dimension of the moduli space $\mathcal{M}_Q([a^-], [a^+])$ *can be computed using the following index formula [4]:*

$$\mu_{\text{inst}}(a^-, a^+) \;=\; \tfrac{1}{2}\left(\eta_{a^+} - \eta_{a^-}\right) + C_{\mathrm{PSU}(r)} \int \langle F_A \wedge F_A \rangle \tag{2.31}$$

for any connection A *on* $\mathbb{R} \times Y$ *limiting to* $a^\pm$ *at* $\pm\infty$. *Here,* η_a *is the* η*-invariant of the operator*

$$\begin{pmatrix} *d_a & d_a \\ d_a^* & 0 \end{pmatrix}$$

associated to $\mathcal{CS}_{a_0}$ *(c.f. [4], [11]), and* $C_{\mathrm{PSU}(r)} > 0$ *is a constant depending only on* $\mathrm{PSU}(r)$ *and the choice of invariant metric on the Lie algebra.*

When $\mu_{\text{inst}}(a^-, a^+) = 1$, the set $\widehat{\mathcal{M}}_Q([a^-], [a^+])$ is finite and we define $\#_Q(a^-, a^+)$

to be the mod-2 count of its elements. With these preliminaries out of the way, we define $\partial_{\text{inst}} : CF_{\text{inst}}(Q) \to CF_{\text{inst}}(Q)$ by

$$\partial_{\text{inst}}\langle [a^-] \rangle = \sum_{\substack{[a^+] \in M(Q) \\ \mu_{\text{inst}}(a^-, a^+) = 1}} \#_Q(a^-, a^+)\langle [a^+] \rangle.$$

Now we can state Floer's main theorem.

Theorem 2.3.4. *[14] Suppose $Q \to Y$ satisfies the conditions of Theorem 2.3.1, and a metric has been chosen from the comeager subset described in Theorem 2.3.1. Then $\partial^2_{\text{inst}} = 0$, and so*

$$HF_{\text{inst}}(Q) := \frac{\ker \partial_{\text{inst}}}{\operatorname{Im} \partial_{\text{inst}}}$$

is well-defined. Furthermore, $HF_{\text{inst}}(Q)$ is independent, up to isomorphism, of the choice of metric.

Unless otherwise stated, whenever we are discussing these moduli spaces or $HF_{\text{inst}}(Q)$ we assume we are in the realm where these satisfy the conclusions of the above theorems (e.g., the metric is generically chosen).

Remark 2.3.5. *Theorem 2.3.4 states that $HF_{\text{inst}}(Q)$ depends only on the bundle Q (it is not clear whether this is an invariant of Y since there may be multiple non-isomorphic bundles over Y satisfying the hypotheses). The proof of the independence of the metric proceeds roughly as follows (see [8, Section 5.3] for more details): Let Y be as in the theorem, and g_0, g_1 metrics in the open dense set of Theorem 2.3.1 (2). Consider $\mathbb{R} \times Y$ equipped with any metric G which is of the form $ds^2 + g_0$ at $-\infty$ and $ds^2 + g_1$ at $+\infty$. Counting isolated instantons on $\mathbb{R} \times Y$ with respect to G defines a map $\zeta_G : HF_{\text{inst}}(Q, g_0) \to HF_{\text{inst}}(Q, g_1)$. Here, the group $HF_{\text{inst}}(Q, g)$ is just $HF_{\text{inst}}(Q)$, except we are now remembering the metric. The count of isolated instantons is always an integer depending continuously on the metric G, so $\zeta = \zeta_G$ must be independent of G. Reversing the roles of g_0 and g_1 produces an inverse ζ^{-1}, and so ζ is the canonical isomorphism of Theorem 2.3.4.*

In fact, $HF_{\mathrm{inst}}(Q)$ depends only on the bundle equivalence class of Q. This is immediate since bundle isomorphisms are just elements of the gauge group, and we have quotiented out by gauge everywhere. See Remark 2.3.10 for similar statements in the setting of quilted Floer cohomology.

Now we specialize to the bundle $Q^\epsilon \to Y^\epsilon$ constructed from Remark 2.2.7. Observe that the hypothesis appearing in Theorems 2.3.1 and 2.3.4 regarding the embedded surface is clearly satisfied by considering any regular fiber of f. As for the perturbation of the Chern-Simons functional, we will ignore this for now and come back to it in the last section when we consider perturbations in general. In particular, we may assume that the moduli spaces $\mathcal{M}_{Q^\epsilon}([a^-],[a^+])$ satisfy the conclusions of these theorems for all $a^\pm \in \mathcal{A}_{\mathrm{flat}}(Q^\epsilon)$.

Recall that Y^ϵ is obtained by gluing the cobordisms $Y_\bullet$ and $I \times \Sigma_\bullet$ using appropriate ϵ-dependent collar neighborhoods. We have seen that every connection A on $\mathbb{R} \times Y$ can be written in the form $A = a(s) + p(s)\, ds$. Similarly, over $\mathbb{R} \times I \times \Sigma_\bullet$ we have

$$A|_{\{(s,t)\}\times\Sigma_\bullet} = \alpha(s,t) + \phi(s,t)\, ds + \psi(s,t)\, dt,$$

for some $\alpha : \mathbb{R}\times I \to \mathcal{A}(P_\bullet)$ and $\phi, \psi : \mathbb{R}\times I \to \Omega^0(\Sigma_\bullet, P_\bullet(\mathfrak{g}))$. We also have $a(s)|_{\{t\}\times\Sigma} = \alpha(s,t) + \psi(s,t)\, dt$ and $p(s)|_{\{t\}\times\Sigma} = \phi(s,t)$. The curvature F_A on the four-manifold can be written in terms of component as

$$F_A = F_\alpha - (\partial_s \alpha - d_\alpha \phi) \wedge ds - (\partial_t \alpha - d_\alpha \psi) \wedge dt + (\partial_s \psi - \partial_t \phi - [\psi, \phi])\, ds \wedge dt.$$

We can therefore view ∂_{inst} as counting $\mathcal{G}_0(\mathbb{R} \times Q^\epsilon)$-equivalence classes of connections $A \in \mathcal{A}(\mathbb{R} \times Q^\epsilon)$ satisfying

- (ϵ-ASD on $\mathbb{R} \times I \times \Sigma_\bullet$)
$$\begin{aligned} \partial_s \alpha - d_\alpha \phi + *(\partial_t \alpha - d_\alpha \psi) &= 0 \\ \partial_s \psi - \partial_t \phi - [\psi, \phi] &= -\epsilon^{-2} * F_\alpha \end{aligned}$$

- (ϵ-ASD on $\mathbb{R} \times Y_\bullet$)
$$\partial_s a - d_a p = -\epsilon^{-1} * F_a$$

- (finite energy)
$$\begin{aligned} \int_{\mathbb{R} \times I \times \Sigma_\bullet} \epsilon^{-2} |F_\alpha|^2 + |\partial_s \alpha - d_\alpha \phi|^2 &< \infty \\ \int_{\mathbb{R} \times Y_\bullet} \epsilon |\partial_s a - d_a p|^2 &< \infty \end{aligned}$$

- (limits at infinity)
$$\begin{aligned} \lim_{s \to \pm\infty} a(s) &= (u^\pm)^* a^\pm \\ \lim_{s \to \pm\infty} p(s) &= 0 \end{aligned} \tag{2.32}$$

for some $u^\pm \in \mathcal{G}_0(Q^\epsilon)$. Here the norms and Hodge stars are all with respect to the fixed metric g.

Remark 2.3.6. *For a connection*

$$A = \begin{cases} \alpha + \phi\, ds + \psi\, dt & \text{on } \mathbb{R} \times I \times \Sigma_\bullet \\ a + p\, ds & \text{on } \mathbb{R} \times Y_\bullet \end{cases}$$

it will be notationally convenient to write

$$\beta_s := \partial_s \alpha - d_\alpha \phi, \qquad \beta_t := \partial_t \alpha - d_\alpha \psi, \qquad \gamma := \partial_s \psi - \partial_t \phi - [\psi, \phi]$$

$$b_s := \partial_s a - d_a p.$$

So, for example, we have

$$F_A = \begin{cases} F_\alpha - \beta_s \wedge ds - \beta_t \wedge dt + \gamma\, ds \wedge dt & \text{on } \mathbb{R} \times I \times \Sigma_\bullet \\ F_a - b_s \wedge ds & \text{on } \mathbb{R} \times Y_\bullet \end{cases}$$

and the ϵ-ASD equations over $\mathbb{R} \times I \times \Sigma_\bullet$ can be written as

$$\beta_s + *\beta_t = 0, \qquad \epsilon^2 \gamma = - * F_\alpha.$$

Now we move on to discuss quilted Floer theory. We refer the reader to [53, Section 4] and [54] for more details. Let (M_i, ω_i) be symplectic manifolds for $1 \leq i \leq N$, with $N > 0$. A **cyclic generalized Lagrangian correspondence** is a tuple $\underline{L} = (L_{12}, L_{23}, \ldots, L_{N1})$ with

$$L_{i(i+1)} \subset M_i^- \times M_{i+1}$$

a Lagrangian submanifold in the usual sense, where M_i^- is the manifold M_i equipped with the negative symplectic form $-\omega_i$.

Here we study

$$CF_{\text{symp}}(\underline{L}) := \bigoplus_{\underline{e} \in \mathcal{I}(\underline{L})} \mathbb{Z}\langle \underline{e} \rangle,$$

where

$$\mathcal{I}(\underline{L}) := \left\{ \underline{e} = (m_1, \ldots, m_N) \in M_1 \times \ldots \times M_N \,\middle|\, (m_i, m_{i+1}) \in L_{i(i+1)} \right\}$$

are the **generalized intersection points**. These are analogous to the flat connections in the instanton theory. Rather than instantons, in this situation we are interested in **pseudoholomorphic quilts**, which are tuples $\underline{v} = (v_1, \ldots, v_N)$, where v_i is a map $\mathbb{R} \times I \to M_i$ satisfying

$$\partial_s v_i + J_i \partial_t v_i = 0.$$

For each i we have fixed an almost complex structure $J_i \in \text{End}(TM_i)$ which is ω_i-compatible. We require that these satisfy the following **Lagrangian seam conditions**

$$(v_i, v_{i+1}) \in L_{i(i+1)}.$$

The relevant notion of **energy** (i.e., the analogue of YM) here is the quantity $E(\underline{v}) := \sum_i E(v_i)$, where

$$E(v_i) := \frac{1}{2} \int_{\mathbb{R} \times I} |\partial_s v_i|^2_{M_i}.$$

Here the norm is the one on M_i given by combining the symplectic structure ω_i with the almost complex structure J_i. Similarly to the instanton case, it follows that $\underline{v}$ has finite energy if and only if it converges exponentially to generalized intersection points at $\pm\infty$ [37, Proposition 1.21].

We therefore consider the moduli spaces

$$\mathcal{M}_{\underline{L}}(\underline{e}^-, \underline{e}^+) := \left\{ \underline{v} = (v_i : \mathbb{R} \times I \to M_i)_{1 \le i \le N} \;\middle|\; \begin{array}{c} \partial_s v_i + J_i \partial_t v_i = 0 \\ (v_i, v_{i+1}) \in L_{i(i+1)} \\ \lim\limits_{s \to \pm\infty} v_i(s, \cdot) = m_i^{\pm} \end{array} \right\},$$

where $\underline{e}^{\pm} = (m_1^{\pm}, \ldots, m_N^{\pm}) \in \mathcal{I}(\underline{L})$. Notice that these spaces also admit and action of $\mathbb{R}$ given by translation, so we set

$$\widehat{\mathcal{M}}_{\underline{L}}(\underline{e}^-, \underline{e}^+) := \mathcal{M}_{\underline{L}}(\underline{e}^-, \underline{e}^+)/\mathbb{R}.$$

The elements of this space can be viewed as (equivalence classes of) *quilted pseudo-holomorphic cylinders.* These are maps from the decorated cylinder in Figure 8 to the relevant spaces indicated by the labels.

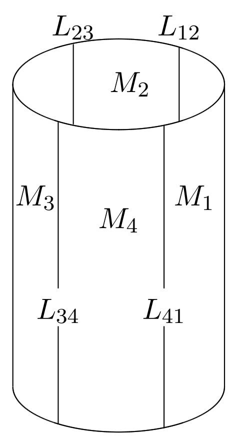


Figure 2.3: Here is the quilted cylinder. The labels indicate where the components are mapped. For example, the strip labeled M_4 is mapped to the symplectic manifold M_4, while the vertical line labeled with L_{41} is mapped to the Lagrangian $L_{41} \subset M_4 \times M_1$.

The next two results are also originally due to Floer [16], and Oh [33], though [54] and [55] are good references for the quilted set-up.

Theorem 2.3.7. *Let (M_i, ω_i) be symplectic manifolds for $1 \leq i \leq N$, and $\underline{L} = (L_{12}, L_{23}, \ldots, L_{N1})$ a cyclic generalized Lagrangian correspondence. Suppose*

- *each symplectic manifold (M_i, ω_i) is simply-connected, compact, has even minimal Chern number, and is monotone with positive monotonicity constant independent of i;*
- *each Lagrangian $L_{i(i+1)}$ is oriented, simply-connected (hence monotone), and spin;*

Then the following properties hold:

1) For a generic set of Hamiltonian perturbations of $\underline{L}$, the set $\mathcal{I}(\underline{L})$ is finite.

2) For a generic (element in a comeager set) tuple of compatible almost complex structures $(J_1, \ldots, J_N)$, the moduli space $\mathcal{M}_{\underline{L}}(\underline{e}^-, \underline{e}^+)$ is a smooth finite dimensional orientable manifold for every pair $\underline{e}^-, \underline{e}^+ \in \mathcal{I}(\underline{L})$. Let $\mathcal{M}_{\underline{L},0}(\underline{e}^-, \underline{e}^+)$ denote the dimension zero component of $\mathcal{M}_{\underline{L}}(\underline{e}^-, \underline{e}^+)$.

3) The action of $\mathbb{R}$ on $\mathcal{M}_{\underline{L},0}([a^-], [a^+])$ is free whenever this space is non-empty, and the quotient space

$$\widehat{\mathcal{M}}_{\underline{L},0}(\underline{e}^-, \underline{e}^+) := \mathcal{M}_{\underline{L},0}(\underline{e}^-, \underline{e}^+)/\mathbb{R}$$

is compact, zero-dimensional, and hence finite.

Denote by $\#_{\underline{L}}(\underline{e}^-, \underline{e}^+)$ the mod-2 count of the elements in the set $\widehat{\mathcal{M}}_{\underline{L},0}(\underline{e}^-, \underline{e}^+)$. Then we define $\partial_{\text{symp}} : CF_{\text{symp}}(\underline{L}) \to CF_{\text{symp}}(\underline{L})$ by

$$\partial_{\text{symp}}\langle \underline{e}^- \rangle = \sum_{\underline{e}^+ \in \mathcal{I}(\underline{L})} \#_{\underline{L}}(\underline{e}^-, \underline{e}^+)\langle \underline{e}^+ \rangle.$$

Theorem 2.3.8. *Suppose (M_i, ω_i) $1 \leq i \leq N$, and $\underline{L}$ satisfy the conditions of Theorem 2.3.7. Also suppose that Hamiltonian perturbations and almost complex structures are chosen as in that theorem. Then $\partial_{\text{symp}}^2 = 0$, and so*

$$HF_{\text{symp}}(\underline{L}) := \frac{\ker \partial_{\text{symp}}}{\operatorname{Im} \partial_{\text{symp}}}$$

is well-defined. Furthermore, $HF_{\mathrm{symp}}(\underline{L})$ is independent, up to isomorphism, of the choice of almost complex structures and Hamiltonian perturbations.

Unless otherwise stated, whenever we are discussing the moduli space or $HF_{\mathrm{symp}}(\underline{L})$ we assume we are in the realm where these satisfy the conclusions of the above theorems (e.g., the compatible almost complex structures are generically chosen).

Remark 2.3.9. *(a) When N is even, quilted Floer cohomology can be equivalently viewed as Lagrangian intersection Floer cohomology $HF(L_{(0)}, L_{(1)})$ for the pair of Lagrangians*

$$\begin{aligned} L_{(0)} &:= L_{12} \times L_{34} \times \ldots \times L_{(N-1)N} \\ L_{(1)} &:= (L_{23} \times L_{45} \times \ldots \times L_{N1})^T \end{aligned} \tag{2.33}$$

in the symplectic manifold $M := M_1^- \times M_2 \times M_3^- \times \ldots \times M_N$, where $Z \mapsto Z^T$ is the map

$$M_2^- \times M_3 \times \ldots \times M_N^- \times M_1 \longrightarrow M_1^- \times M_2 \times M_3^- \times \ldots \times M_N$$

transposing the last factor to the front and changing each symplectic form to its negative. This is also true when N is odd, but one needs to include a diagonal and shuffle the indices in (2.33). In both cases, there is a natural identification $\mathcal{I}(\underline{L}) \cong L_{(0)} \cap L_{(1)}$.

(b) The proof of the independence of $HF_{\mathrm{symp}}(\underline{L})$ from the underlying data follows the same basic schematic as outlined in Remark 2.3.5. For example, one connects two compatible almost complex (a.c.) structures by a path. This path can equivalently be viewed as a time-dependent a.c. structure. By counting strips in M which are holomorphic with respect to this time-dependent a.c. structure, one obtains a canonical isomorphism between the Lagrangian Floer cohomology groups. This count depends continuously on the underlying data, and so is independent of the choice of path. See, for example, [30] for more details on this type of argument in the symplectic setting.

(c) Similarly to the instanton theory, there is a map $\mu_{\mathrm{symp}}(\underline{e}^-, \underline{e}^+) : \mathcal{M}_{\underline{L}}(\underline{e}^-, \underline{e}^+) \to \mathbb{Z}$, which measures the local dimension of the moduli space $\mathcal{M}_{\underline{L}}(\underline{e}^-, \underline{e}^+)$. In fact, each

component of $\mathcal{M}_{\underline{L}}(\underline{e}^-, \underline{e}^+)$ has the same dimension mod $2N$, where N is the minimum Chern number. In particular, $\mu_{\mathrm{symp}}(\underline{e}^-, \underline{e}^+)$ defines a unique element of $\mathbb{Z}_{2N}$.

Now we specialize: Let $Q \to Y$ be the bundle from Remark 2.2.7, and take

$$Q_{i(i+1)} := Q|_{Y_{i(i+1)}}, \qquad P_i := Q|_{\{0\}\times\Sigma_i}$$

$$Q_\bullet = Q|_{Y_\bullet}, \qquad P_\bullet = Q|_{\{0\}\times\Sigma_\bullet}$$

as above. By Theorems 2.2.15 and 2.2.16, the manifolds $M(P_i)$ and $L(Q_{i(i+1)})$ satisfy the conditions of Theorems 2.3.7 and 2.3.8, except possibly for the transversality condition in Theorem 2.3.7. This can always be achieved by replacing $\underline{L}(Q)$ with a generic (open dense) Hamiltonian perturbation. We ignore this detail for now and come back to it later when where we consider all of the necessary perturbations at once. The compatible almost complex structures are provided by the Hodge stars associated to g_Σ, and choosing g suitably generically ensures that these almost complex structures belong to the comeager set guaranteed by Theorem 2.3.7. It follows that the quilted Floer cohomology $HF_{\mathrm{symp}}(\underline{L}(Q))$ is well-defined, where

$$\underline{L}(Q) := (L(Q_{12}), L(Q_{23}), \ldots, L(Q_{N1})) .$$

Remark 2.3.10. *Theorem 2.3.8 implies that $HF_{\mathrm{symp}}(\underline{L}(Q))$ depends only on the data of the bundle $Q \to Y$ together with the choice of $f : Y \to S^1$. In fact, Wehrheim and Woodward have shown that this cohomology group depends only on the homotopy class of f [53] (the bundle equivalence class of Q is determined uniquely by the homotopy class of f). This proceeds roughly as follows: Gay and Kirby have shown that any pair of homotopic Morse functions $f, f' : Y \to S^1$ can be connected through a path of functions which are Morse except at a finite number of points, where a critical point birth, death or switch occurs [22]. As a consequence, Woodward and Wehrheim only need to show that Lagrangian Floer cohomology is unchanged, up to canonical isomorphism, under isotopy, and critical point birth, death and switch. At the level of generators this is* *fairly straight-forward, and is perhaps most evident from the realization of the moduli spaces of flat connections as (equivalence classes of) representations of the fundamental group. On the other hand, the equivalence of the chain maps associated to f and f' follows from a technical 'strip-shrinking' analysis [52], which is a quilted analogue of the 'stretch-the-neck' analysis considered here.*

In particular, $HF_{\mathrm{symp}}(\underline{L}(Q))$ depends only on the bundle equivalence class of Q. This is consistent with the predictions of the Atiyah-Floer conjecture since $HF_{\mathrm{inst}}(Q)$ depends only on this equivalence class as well. See Remark 2.3.5.

We need to understand $HF_{\mathrm{symp}}(\underline{L}(Q))$ in terms of the underlying connections. Unraveling the definitions, one finds that $\mathcal{I}(\underline{L}(Q))$ is the set of tuples

$$\underline{e} = \left([a_{12}], [a_{23}], \ldots, [a_{N1}] \right),$$

of $\mathcal{G}_0(Q_{i(i+1)})$-equivalence classes of $a_{i(i+1)} \in \mathcal{A}_{\mathrm{flat}}(Q_{i(i+1)})$ satisfying

$$\left[a_{i(i+1)} \Big|_{\Sigma_i} \right] = \left[a_{(i-1)i} \Big|_{\Sigma_i} \right], \tag{2.34}$$

where each $a_{i(i+1)}$ is a flat connection on $Q_{i(i+1)}$, the bracket $\left[a_{i(i+1)} \Big|_{\Sigma_i} \right]$ denotes the $\mathcal{G}_0(P_i)$-equivalence class. (Here we are working with smooth connections and gauge transformations to simplify notation, but one could equally well work with their Sobolev completions.)

Similarly, each moduli space $\mathcal{M}_{\underline{L}(Q)}(\underline{e}^-, \underline{e}^+)$ consists of $\mathcal{G}_0(P_1) \times \ldots \times \mathcal{G}_0(P_N)$-equivalence classes of tuples $(\alpha_1, \ldots, \alpha_N)$ where each $\alpha_i : \mathbb{R} \times I \to \mathcal{A}_{\mathrm{flat}}(P_i)$ satisfies the following conditions:

- (holomorphic) $\partial_s \alpha_i + *\partial_t \alpha_i \in \left(H^1_{\alpha_i}\right)^\perp = \mathrm{Im}\left(d_{\alpha_i}\right) \oplus \mathrm{Im}\left(d^*_{\alpha_i}\right)$

- (Lagrangian seam) $\left(\left[\alpha_i\left(s,1\right)\right], \left[\alpha_{i+1}\left(s,0\right)\right]\right) \in L(Q_{i(i+1)})$

- (finite energy) $\int_{\mathbb{R}\times I} \left\|\mathrm{proj}_{\alpha_i} \partial_s \alpha_i\right\|^2_{L^2(\Sigma_i)} < \infty$

- (limits at infinity) $\lim_{s\to\pm\infty} \alpha_i(s,\cdot) = \left(u_i^\pm\right)^* a^\pm_{(i-1)i}\Big|_{\Sigma_i}$ for some $u_i^\pm \in \mathcal{G}_0(P_i)$,

where $\mathrm{proj}_{\alpha_i} : T_{\alpha_i}\mathcal{A}(P_i) \to H^1_{\alpha_i}$ is the orthogonal projection to the harmonic space and

$$\underline{e}^\pm = \left(\left[a^\pm_{12}\right], \ldots, \left[a^\pm_{N1}\right]\right).$$

It will be useful to express these conditions in a more explicit fashion. The holomorphic condition is equivalent to $\partial_s \alpha_i + *\partial_t \alpha_i = d_{\alpha_i}\phi_i + *d_{\alpha_i}\psi_i$ for some $\phi_i, \psi_i : \mathbb{R} \times I \to \Omega^0(\Sigma_i, P_i(\mathfrak{g}))$. In fact, ϕ_i, ψ_i are uniquely determined by this equation since d_{α_i} is injective on 0-forms (recall all of the flat connections are irreducible). Equivalently, differentiating $F_{\alpha_i} = 0$ gives $d_{\alpha_i}\partial_s\alpha_i = 0$ (resp. $d_{\alpha_i}\partial_t\alpha_i = 0$), so the expression $\partial_s\alpha_i - d_{\alpha_i}\phi_i$ (resp. $\partial_t\alpha_i - d_{\alpha_i}\psi_i$) can be viewed as the orthogonal projections of $\partial_s\alpha_i$ (resp. $\partial_t\alpha_i$) onto the harmonic space. Furthermore, ϕ_i and ψ_i are as smooth in s, t as α_i. We record this for later use.

Lemma 2.3.11. *Let $P \to X$ is a bundle over a compact manifold, and assume all flat connections are irreducible. Suppose $\alpha : \mathbb{R} \to \mathcal{A}^{1,q}_{\mathrm{flat}}(P)$ is a smooth path of flat connections. Then there is a unique smooth path $\phi : \mathbb{R} \to W^{2,q}(P(\mathfrak{g}))$ such that*

$$\partial_s\alpha(s) - d_{\alpha(s)}\phi(s) \in H^1_{\alpha(s)}$$

for each $s \in \mathbb{R}$. Moreover, if $\alpha(s) \in \mathcal{A}_{\mathrm{flat}}(P)$ is smooth for each s, then the 0-form $\phi(s) \in \Omega^0(X, P(\mathfrak{g}))$ is smooth for each s.

As we noted above, the *finite energy* condition automatically implies that the α_i converge at $\pm\infty$ to an element of $\mathcal{I}_{\mathrm{symp}}$, and that the integrand converges exponentially to zero. Then the *limit at infinity* condition just says we obtain the specified limiting values. It also follows that

$$\begin{aligned} \lim_{s\to\pm\infty} \phi_i(s,\cdot) &= 0 \\ \lim_{s\to\pm\infty} \psi_i(s,\cdot) &= \psi_i^\pm \end{aligned} \tag{2.35}$$

where, modulo gauge, $\psi_i^\pm \in \Omega^0(\Sigma_i, P_i(\mathfrak{g}))$ is the dt-component of the limiting connection $a^\pm_{(i-1)i}$.

The *Lagrangian seam* condition implies that there is some path $a_{i(i+1)} : \mathbb{R} \to \mathcal{A}_{\text{flat}}(Q_{i(i+1)})$ with $(\alpha_i\,(s,1)\,,\alpha_{i+1}\,(s,0)) = \left(a_{i(i+1)}(s)\Big|_{\Sigma_i}, a_{i(i+1)}(s)\Big|_{\Sigma_{i+1}} \right)$. By Lemma 2.3.11, the irreducibility of flat connections on $Q_{i(i+1)}$ implies that there is a unique $p_{i(i+1)} : \mathbb{R} \to \Omega^0(Y_{i(i+1)}, Q_i(\mathfrak{g}))$ with

$$\partial_s a_{i(i+1)} - d_{a_{i(i+1)}} p_{i(i+1)} \in H^1_{a_{i(i+1)}}.$$

It will be notationally convenient to write α (resp. a) for the connection on $\Sigma_\bullet$ (resp. $Y_\bullet$) that restricts to α_i on Σ_i ($a_{i(i+1)}$ on $Y_{i(i+1)}$). Likewise, we define ϕ, ψ which are forms on $\Sigma_\bullet$, and p which is a form on $Y_\bullet$.

To summarize, the boundary operator ∂_{symp} counts isolated $\mathcal{G}_0$-equivalence classes of tuples $(\alpha, \phi, \psi, a, p)$ satisfying

- (holomorphic)
$$\begin{aligned} \partial_s \alpha - d_\alpha \phi + *(\partial_t \alpha - d_\alpha \psi) &= 0 \\ F_\alpha &= 0 \end{aligned}$$

- (Lagrangian seam)
$$\begin{aligned} (\alpha\,(s,1)\,,\alpha\,(s,0)) &= \left(a(s)|_{\partial_1 Y_\bullet}\,,\, a(s)|_{\Sigma_{\partial_s Y_\bullet}} \right) \\ F_a &= 0 \end{aligned}$$

- (finite energy)
$$\int_{\mathbb{R}\times I\times\Sigma_\bullet} |\partial_s \alpha - d_\alpha \phi|^2 < \infty$$

- (limits at infinity)
$$\begin{aligned} \lim_{s\to\pm\infty} \alpha(s,\cdot) &= (u^\pm)^*\, a^\pm|_{\partial_1 Y_\bullet} \\ \lim_{s\to\pm\infty} \phi(s,\cdot) &= 0 \\ \lim_{s\to\pm\infty} \psi(s,\cdot) &= \psi^\pm \\ \lim_{s\to\pm\infty} a(s,\cdot) &= (u^\pm)^*\, a^\pm \\ \lim_{s\to\pm\infty} p(s,\cdot) &= 0, \end{aligned} \tag{2.36}$$

for some identity component gauge transformations $u^\pm \in \mathcal{G}_0(Q_{i(i+1)})$.

The form p provides seam conditions for the forms ϕ and ψ as follows: Note that $Y_\bullet$ is a (disconnected) cobordism from $\Sigma_\bullet$ to itself. We let $\partial_1, \partial_2 : Y_\bullet \to \Sigma_\bullet$ be the restriction to the first, second copies of $\Sigma_\bullet$, respectively. The uniqueness of ϕ (Lemma 2.3.11) combines with the Lagrangian seam condition to imply that $p(s)$ restricts to $\phi(s,1)$ and $\phi(s,0)$ on $\partial_1 Y_\bullet$ and $\partial_2 Y_\bullet$. Similarly, restricting $a(s)$ to $\partial_1 Y_\bullet$ or $\partial_2 Y_\bullet$, and then taking the normal component recovers $\psi(s,1)$ or $\psi(s,0)$.

It is important to note that though the metric on the $\Sigma_\bullet$ makes an appearance in the definition of the boundary operator (through the Hodge star), the resulting cohomology group is independent of the choice of metric.

Remark 2.3.12. *(a) The holomorphic and Lagrangian seam conditions in (2.36) earn their name because they become honest holomorphic and Lagrangian seam conditions when we descend to the finite dimensional symplectic moduli spaces. However, they can be viewed as holomorphic and Lagrangian seam conditions on infinite dimensional* *spaces of connections in the following way. The Hodge star, viewed as an operator from $T_\alpha \mathcal{A}(P_i) = \Omega^1(\Sigma_i, P_i(\mathfrak{g}))$ to itself, provides a holomorphic structure on each $\mathcal{A}(P_i)$. However, it does not restrict to a holomorphic structure on $\mathcal{A}_{\mathrm{flat}}(P_i)$ since it carries $\ker(d_\alpha) = T_\alpha \mathcal{A}_{\mathrm{flat}}(P_i)$ to $\ker(d_\alpha^*)$ it is never the case that $\ker(d_\alpha^*)$ is contained in $\ker(d_\alpha)$ (indeed, the intersection is always the harmonic space, which is finite dimensional by elliptic theory, but $\ker(d_\alpha^*)$ is always infinite dimensional). Similarly wedging and integration determine a symplectic form on $\mathcal{A}_{\mathrm{flat}}(P_i)$. (This form is closed because $T_\alpha \mathcal{A}_{\mathrm{flat}}(P_i) = \ker(d_\alpha)$, and hence this same formula does not provide a closed form on all of $\mathcal{A}(P_i)$.) In particular, we can discuss holomorphic maps into $\mathcal{A}(P_i)$ and Lagrangian subspaces of $\mathcal{A}_{\mathrm{flat}}(P_i)$.*

This being said, the holomorphic condition in (2.36) is exactly the perturbed holomorphic equation for the curve $\alpha : \mathbb{C} \to \mathcal{A}(P_i)$, whose image happens to lie in $\mathcal{A}_{\mathrm{flat}}(P_i)$. It is notable that the perturbation, given by $d_{\alpha_i}\phi_i + *d_{\alpha_i}\psi_i$, lies in $\mathrm{Im}(d_{\alpha_i}) \oplus \mathrm{Im}(d_{\alpha_i}^*)$, which is why the perturbed holomorphic equation descends to an honest holomorphic equation when we pass to the moduli space.

Similarly, the Lagrangian boundary condition in (2.36) are exactly boundary conditions for the curve α_i given by the embedded Lagrangian submanifold $\mathcal{A}_{\mathrm{flat}}(Q_{i(i+1)}) \hookrightarrow \mathcal{A}_{\mathrm{flat}}(P_i)^- \times \mathcal{A}_{\mathrm{flat}}(P_{i+1})$, where the inclusion is given by restriction.

(b) Observe that, in the small ϵ limit, the equations in (2.32) reduce to those in (2.36). This thesis is a step in the direction of showing that we can identify the moduli spaces in these two situations, at least for small ϵ and for suitable limits at $\pm\infty$.

2.4 Statement of the Main Theorem

We write the components of a connection A as

$$A = \begin{cases} \alpha(s,t) + \phi(s,t)\,ds + \psi(s,t)\,dt & \text{on } \{(s,t)\} \times \Sigma_\bullet \\ a(s) + p(s)\,ds & \text{on } \{s\} \times Y_\bullet. \end{cases}$$

For $s_0 \in \mathbb{R}$, we have a map $\tau_{s_0} : \mathbb{R} \times Y \to \mathbb{R} \times Y$, given by translating by s_0. This ascends to a pullback map on connections and forms in the usual way. For example,

$$\tau_{s_0}^* : \mathcal{A}^{1,q}(\mathbb{R} \times Q) \longrightarrow \mathcal{A}^{1,q}(\mathbb{R} \times Q)$$

is given by

$$\left(\tau_{s_0}^* A\right)|_{\{s\}\times Y} := A|_{\{s+s_0\}\times Y}$$

and this preserves the ϵ-ASD connections. Now we are at a place where we can state the main result.

Main Theorem 2.4.1. *Let $Q \to Y$ be as in Remark 2.2.7, and suppose all flat connections on Q are non-degenerate. Fix $q > 2$ and two flat connections $a^{\pm} \in \mathcal{A}_{\text{flat}}(Q)$. Let $(\epsilon_\nu)_{\nu\in\mathbb{N}}$ be a sequence of positive numbers converging to 0. Suppose that for each ν there is an ϵ_ν-ASD connection $A_\nu \in \mathcal{A}^{1,q}_{loc}(\mathbb{R} \times Q)$, which descends to an element of the zero-dimensional moduli space $\widehat{\mathcal{M}}_{Q^{\epsilon_\nu},0}([a^-],[a^+])$. Then there is a continuous connection*

$$A_\infty = \begin{cases} \alpha_\infty(s,t) + \phi_\infty(s,t)\,ds + \psi_\infty(s,t)\,dt & \text{on } \{(s,t)\}\times\Sigma_\bullet \\ a_\infty(s) + p_\infty(s)\,ds & \text{on } \{s\}\times Y_\bullet \end{cases}$$

in $\mathcal{A}^{1,q}_{loc}(\mathbb{R}\times Q)$, which

(i) is holomorphic:

$$\partial_s\alpha_\infty - d_{\alpha_\infty}\phi_\infty + *\left(\partial_t\alpha_\infty - d_{\alpha_\infty}\psi_\infty\right) = 0, F_{\alpha_\infty} = 0;$$

(ii) has Lagrangian boundary conditions:

$$F_{a_\infty} = 0;$$

(iii) and converges to the flat connections $a^{\pm}$ at $\pm\infty$:

$$\lim_{s\to\pm\infty}\left(u^{\pm}\right)^* A_\infty|_{\{s\}\times Y} = a^{\pm}, \textit{for some } u^{\pm} \in \mathcal{G}_0^{2,q}(Q).$$

Furthermore, there is a subsequence (still denoted by A_ν) which converges to A_∞ in the following sense: There is a sequence of gauge transformations $u_\nu \in \mathcal{G}^{2,q}_{loc}(\mathbb{R} \times Q)$, and a sequence of times $s_\nu \in \mathbb{R}$, with

$$\left\| \alpha_\infty - u_\nu^* \tau_{s_\nu}^* \alpha_\nu \right\|_{C^0(K \times \Sigma_\bullet)} \xrightarrow{\nu} 0$$

for every compact $K \subset \mathbb{R} \times I$. Here the action of u_ν on α_ν is the (s,t)-pointwise action of $u_\nu(s,t) \in \mathcal{G}^{2,q}(P_\bullet)$ on connections over the surface $\Sigma_\bullet$.

Remark 2.4.2. *(a) Theorem A, from the introduction, is an immediate consequence of Theorem 2.4.1.*

(b) Included in the statement of Theorem 2.4.1 (as well as in a few other places in this thesis) is the hypothesis that all flat connections on Q are non-degenerate. In general, this need not be the case, and so Theorem 2.4.1 is, strictly speaking, vacuous in such cases. To account for this, one would need to replace the Chern-Simons functional by a suitable perturbation. See Remark 2.3.3. For the most part this is a fairly standard general position argument (see, e.g., [8, Proposition 5.17], [11]), though one still needs to verify that the perturbations in both Floer theories are compatible in some sense. We ignore most of these details here, relegating a complete discussion of perturbations to a future paper.

Section 4 is dedicated to the proof of the Main Theorem. We conclude the current section by showing that the generators of the two Floer theories agree, up to the action of the group

$$H_\eta := \ker\left(\eta : \mathcal{G}(Q) \to \mathbb{Z}_r\right)/\mathcal{G}_0(Q) \cong \mathbb{Z},$$

where the isomorphism is given by the degree. The gauge group action descends to an action of H_η on the moduli space of flat connection $M(Q)$. Observe that if $u \in \ker \eta$, then the restriction $u|_{Y_{i(i+1)}}$ has parity zero and so is in the identity component of $\mathcal{G}(Q_{i(i+1)})$. In particular, if $a \in \mathcal{A}_{\text{flat}}(Q)$, then the restrictions $a|_{Y_{i(i+1)}}$ and $u^*a|_{Y_{i(i+1)}}$ are $\mathcal{G}_0(Q_{i(i+1)})$-gauge equivalent, and so the map

$$\begin{aligned}\Psi : M(Q)/H_\eta \quad &\longrightarrow \quad \mathcal{I}(\underline{L}(Q)) \\ [a] \quad &\longmapsto \quad \underline{e}_a := \left(\left[a|_{Y_{12}} \right], \left[a|_{Y_{23}} \right], \ldots, \left[a|_{Y_{N1}} \right] \right),\end{aligned} \tag{2.37}$$

is well-defined. Here the brackets denote the $\mathcal{G}_0$-equivalence class on the relevant 3-manifold.

Proposition 2.4.3. *The map Ψ is a set bijection. Moreover, Ψ is natural in the sense that if $Q \to Q'$ is any* $\mathrm{PSU}(r)$*-bundle map covering the identity, then the diagram commutes:*

$$\begin{array}{ccc} M(Q)/H_\eta & \longrightarrow & M(Q')/H_\eta \\ {\scriptstyle\Psi}\big\downarrow & & \big\downarrow{\scriptstyle\Psi} \\ \mathcal{I}(\underline{L}(Q)) & \longrightarrow & \mathcal{I}(\underline{L}(Q')) \end{array}$$

Proof. The naturality is immediate since bundle maps covering the identity are gauge transformations, and everything is gauge invariant. It therefore suffices to show Ψ is a bijection. Throughout we will denote by $\iota : \Sigma_i \hookrightarrow Y_{i(i+1)}$ and $\iota' : \Sigma_{i+1} \hookrightarrow Y_{(i-1)i}$ the inclusion of the boundary components. For simplicity in notation we do not keep track of the index i in the notation of ι and ι.

To prove that Ψ is injective we first show that any two $a, a' \in \mathcal{A}_{\mathrm{flat}}(Q)$ that restrict (modulo gauge) to the same connection on $Y_\bullet$, also restrict (modulo gauge) to the same connection on $I \times \Sigma_\bullet$. This is essentially a consequence of the injectivity of the restriction

$$M(I \times P) \hookrightarrow M\left(\{0\} \times P\right) \times M\left(\{1\} \times P\right)$$

from Theorem 2.2.16 for bundles $I \times P \to I \times \Sigma$ over product cobordisms. Indeed, that a and a' restrict to the same connection mod gauge on $Y_\bullet$, in particular, means that they restrict to the same connection (modulo gauge) on each Σ_i. So $\left[a|_{I\times\Sigma_i} \right]$ and $\left[a'|_{I\times\Sigma_i} \right]$ have the same image in $M(\Sigma_i) \times M(\Sigma_i)$ and hence these are equal $\left[a|_{I\times\Sigma_i} \right] = \left[a'|_{I\times\Sigma_i} \right]$. So there are gauge transformations $u_i \in \mathcal{G}_0(I \times P_i)$ and $u_{i(i+1)} \in \mathcal{G}_0(Q_{i(i+1)})$ such that $u_i^* a|_{I\times\Sigma_i} = a'|_{I\times\Sigma_i}$ and $u_{i(i+1)}^* a\Big|_{Y_{i(i+1)}} = a'|_{Y_{i(i+1)}}$. The data $\left\{ u_i, u_{i(i+1)} \right\}_{i=1,\ldots,N}$ patch together to form a global (possibly discontinuous) gauge transformation u. We will have proven injectivity if we can show that u is in fact smooth (it is contained in $\ker \eta$ by construction).

We first claim that u is continuous. To see this, note that

$$\begin{aligned} \left(u_i|_{\{1\}\times\Sigma}\right)^* \left(a|_{\{1\}\times\Sigma_i}\right) &= a'|_{\{1\}\times\Sigma_i} \\ &= a'|_{\iota(\Sigma_i)} \\ &= \left(u_{i(i+1)}\Big|_{\iota(\Sigma_i)}\right)^* \left(a|_{\iota(\Sigma_i)}\right). \end{aligned}$$

The functoriality of the characteristic classes from section 2.2.1 imply that the map $\iota^* : \mathcal{G}(Q_{i(i+1)}) \to \mathcal{G}(P_i)$ induced by ι restricts to a map $\mathcal{G}_0(Q_{i(i+1)}) \to \mathcal{G}_0(P_i)$. This shows that $\left(u_{i(i+1)}\Big|_{\iota(\Sigma_i)}\right)^{-1} u_i|_{\{1\}\times\Sigma_i}$ is an element of $\mathcal{G}_0(\Sigma_i)$ and fixes $a|_{\{1\}\times\Sigma_i} = a|_{\iota(\Sigma_i)}$. By Lemma 2.2.14 we must have that this is the identity gauge transformation, and so $u_{i(i+1)}\Big|_{\iota(\Sigma_i)} = u_i|_{\{1\}\times\Sigma_i}$. A similar argument shows $u_{i(i+1)}\Big|_{\iota'(\Sigma_{i+1})} = u_{i+1}|_{\{0\}\times\Sigma_{i+1}}$. This proves the claim.

To see that u is actually smooth, we use the following trick from [9, Chapter 2.3.7] to bootstrap: As in the (2.13), by choosing a faithful matrix representation we can write the action of u on a as

$$u^*a = u^{-1}au - u^{-1}du,$$

where on the left we are viewing the gauge transformation as a map $u : Q \to G$, and the concatenation is matrix multiplication. Rearranging this and using $u^*a = a'$ gives

$$du = au + ua'.$$

The right-hand side is $\mathcal{C}^0$, so u is of differentiability class $\mathcal{C}^1$. Repeatedly bootstrapping in this way shows that u is in $\mathcal{C}^\infty$, and hence $u \in \mathcal{G}(Q)$.

Now we show that $u \in \mathcal{G}_0(Q)$ is in the component of the identity. We have that each $u_{i(i+1)} : Y_{i(i+1)} \to Q_{i(i+1)} \times_G G$ and $u_i : I \times \Sigma_i \to I \times P_i \times_G G$ are all homotopic to the identity. We can assume these homotopies agree at the boundary (i.e. the homotopy for u_1 restricted to $\{1\} \times \Sigma_1$ equals the homotopy for u_{12} when restricted to $\Sigma_1 \subset \partial Y_{12}$). Then these homotopies patch to give a homotopy for u to the identity.

Now we must prove that (2.37) is surjective. Fix some $([a_{12}], [a_{23}], \ldots, [a_{N1}]) \in \mathcal{I}(\underline{L}(Q))$. We need to show that there is some $a \in \mathcal{A}_{\text{flat}}(Q)$ with $a|_{Y_{i(i+1)}} \in \left[a_{i(i+1)}\right]$. Choose a representative $a_{i(i+1)} \in \mathcal{A}_{\text{flat}}(Q_{i(i+1)})$ for each $\left[a_{i(i+1)}\right]$. These give us an obvious definition of a over the $Y_\bullet$, however we need to define a over $I \times \Sigma_\bullet$ as well. To do this, let $\alpha_i := a_{(i-1)i}\Big|_{\iota'(\Sigma_i)}$. Then $\alpha_i \in \mathcal{A}_{\text{flat}}(P_i)$ is flat, and therefore so is $\text{proj}^*\alpha_i$ where $\text{proj} : I \times \Sigma_i \to \Sigma_i$ is the projection. The boundary condition (2.34) implies that there is some $\mu_i \in \mathcal{G}_0(P_i)$ with $\mu_i^*\alpha_i = a_{i(i+1)}\Big|_{\iota(\Sigma_i)}$. By definition, $\mathcal{G}_0(\Sigma_i)$ is path-connected so there is some path $u_i : I \to \mathcal{G}_0(P_i)$ connecting the identity to μ_i. We can equivalently view the path $u_i \in \mathcal{G}(I \times P_i)$ as a gauge transformation over the cylinder $I \times \Sigma_i$. Then $u_i^* (\text{proj}^*\alpha_i)$ is a flat connection on $I \times P_i$ connecting $a_{(i-1)i}$ to $a_{i(i+1)}$. Define a to be $a_{i(i+1)}$ over $Y_{i(i+1)}$ and to be $u_i^* (\text{proj}^*\alpha_i)$ over $I \times \Sigma_i$. Then a is continuous, flat and restricts to the desired connections over the $Y_{i(i+1)}$. By choosing the path $u_i : I \to \mathcal{G}_0(P_i)$ to extend smoothly to be constant at the endpoints we can also ensure that a is smooth. This completes the proof of Proposition 2.4.3.

$\square$

Chapter 3

Small curvature connections in dimensions 2 and 3

In our proof of the main result, we will encounter connections on surfaces and 3-manifolds, and these connections will have uniformly small curvature. We want a uniform way of identifying nearby flat connections. In the case of surfaces, this can be made precise by the *complexified gauge group*, which acts freely on the subset of connections having sufficiently small curvature. It is well-known to experts that quotienting this subset by the action of the complexified gauge group (called a **Narasimhan-Seshadri correspondence**) recovers the moduli space of flat connections. This was originally carried out for unitary bundles on surfaces by Narasimhan and Seshadri [32], using algebraic techniques. Later, it was extended to more general structure groups by Ramanathan in his thesis [35]. (See also Kirwan's book [27] for a finite-dimensional version.) Our approach is more in the spirit of Donaldson [7], where he works in an analytic category and uses an implicit function theorem argument. In section 3.1, we develop precise C^1 and C^2-estimates associated to this quotient, which will be needed for our proof of the Main Theorem 2.4.1.

In the case of 3-manifolds it is not clear how to set up an analogous implicit function theorem argument. To obtain similar results, we instead appeal to the Yang-Mills heat flow, which was worked out by Råde [34] for closed 3-manifolds. In section 3.2, we extend Råde's result to compact manifolds with boundary. The lack of an implicit function theorem means that we only obtain C^0 estimates, however this is sufficient for our purposes.

3.1 Semistable connections over a surface

The goal of this section is to define a gauge-equivariant deformation retract

$$\mathrm{NS} : \mathcal{A}^{ss} \to \mathcal{A}_{\mathrm{flat}},$$

and establish some of its properties. Here $\mathcal{A}^{ss}$ is a neighborhood of $\mathcal{A}_{\mathrm{flat}}$ (the superscript stands for *semistable*). The relevant properties of the map NS are laid out in Theorem 3.1.1, below. After stating the theorem, we will define the complexified gauge group $\mathcal{G}^{\mathbb{C}}$ and its action on $\mathcal{A}$ (this is only used in the proofs appearing in this section, and will not be used in the rest of the thesis). The proof of the theorem will show that for each $\alpha \in \mathcal{A}^{ss}$, there is a 'purely imaginary' complex gauge transformation u such that $u^*\alpha$ a flat connection, and u is unique provided it lies sufficiently close to the identity. We then define $\mathrm{NS}(\alpha) := u^*\alpha$.

After proving Theorem 3.1.1 below, where the map NS is formally defined, we spend the remainder of this section establishing useful properties and estimates for NS. For example, in the proof of Lemma 3.1.11 we establish the Narasimhan-Seshadri correspondence

$$\mathcal{A}^{ss}/\mathcal{G}_0^{\mathbb{C}} \cong \mathcal{A}_{\mathrm{flat}}/\mathcal{G}_0,$$

and in Proposition 3.1.9 and Corollary 3.1.13 we show that, to first order, the map NS is the identity plus the L^2 orthogonal projection to the tangent space of flat connections.

Theorem 3.1.1. *Suppose G is a compact connected Lie group, Σ is a closed oriented Riemannian surface, and $P \to \Sigma$ is a principal G-bundle such that all flat connections are irreducible. Then for any $1 < q < \infty$, there are constants $C > 0$ and $\epsilon_0 > 0$, and a $\mathcal{G}^{2,q}(P)$-equivariant deformation retract*

$$\mathrm{NS}_P : \left\{ \alpha \in \mathcal{A}^{1,q}(P) \,\middle|\, \|F_\alpha\|_{L^q(\Sigma)} < \epsilon_0 \right\} \longrightarrow \mathcal{A}^{1,q}_{\mathrm{flat}}(P) \tag{3.1}$$

which is smooth with respect to the $W^{1,q}$-topology on the domain and codomain. Moreover, the map NS_P is also smooth with respect to the L^p-topology on the domain and codomain, for any $2 < p < \infty$.

Remark 3.1.2. *The restriction in the second part of the theorem to $2 < p < \infty$*

is merely an artifact of our proof, and it is likely that the conclusion holds for, say, $1 < p \leq 2$ as well. See Lemma 3.1.12.

Our strategy for the proof of Theorem 3.1.1 is to work in the setting of holomorphic structures on a complex vector bundle, rather than on the space of connections themselves. Geometrically this is nothing more than a change of perspective, but this perspective has the advantage that it makes clear how the gauge group action extends to an action of the *complexified* gauge group. Before proving Theorem 3.1.1 we recall the definition and various properties of the complexified gauge group.

Set G be a compact, connected Lie group and fix a faithful Lie group embedding $\rho : G \hookrightarrow \mathrm{U}(n)$ for some n. We identify G with its image in $\mathrm{U}(n)$. Define $E := P \times_G \mathbb{C}^n$, and equip it with the G-invariant Hermitage structure induced by the embedding $G \subset \mathrm{U}(n)$. Let J_E denote the complex structure on E induced from the standard complex structure on $\mathbb{C}^n$. The metric and orientation on Σ determine a complex structure j_Σ, and we will use the notation $\Omega^{k,l}(\Sigma, E)$ to denote the smooth E-valued forms of type (k, l). Observe that j_Σ acts by the Hodge star on 1-forms.

Consider the space

$$\mathcal{C}(E) := \left\{ \overline{D} : \Omega^0(\Sigma, E) \to \Omega^{0,1}(\Sigma, E) \;\middle|\; \begin{array}{c} \overline{D}(f\xi) = f(\overline{D}\xi) + (\overline{\partial} f)\xi, \\ \text{for } \xi \in \Omega^0(\Sigma, E), f \in \Omega^0(\Sigma) \end{array} \right\},$$

of Cauchy-Riemann operators on E. This can be naturally identified with the space of holomorphic structures on E (see [30, Appendix C]). Each element $\overline{D} \in \mathcal{C}(E)$ has a unique extension to an operator $\overline{D} : \Omega^{j,k}(\Sigma, E) \to \Omega^{j,k+1}(\Sigma, E)$ satisfying the Leibniz rule.

Let $\mathcal{A}(E)$ denote the space of $\mathbb{C}$-linear covariant derivatives on E:

$$\mathcal{A}(E) := \left\{ D : \Omega^0(\Sigma, E) \to \Omega^1(\Sigma, E) \;\middle|\; \begin{array}{c} D(f\xi) = f(D\xi) + (df)\xi, \\ \text{for } \xi \in \Omega^0(\Sigma, E), f \in \Omega^0(\Sigma) \end{array} \right\}$$

There is a natural isomorphism

$$\begin{array}{ccc} \mathcal{A}(E) & \longrightarrow & \mathcal{C}(E) \\ D & \longmapsto & \frac{1}{2}\left(D + J_E D \circ j_\Sigma\right) \end{array} \tag{3.2}$$

Here and below we are using the symbol $\circ$ to denote composition of operators. For example, if $M : \Omega(\Sigma, E) \to \Omega(\Sigma, E)$ is a derivation we define $M \circ j_\Sigma : \Omega(\Sigma, E) \to \Omega(\Sigma, E)$ to be the derivation of the same degree given by

$$\iota_X\left((M \circ j_\Sigma)\,\xi\right) = \iota_{j_\Sigma X}\left(M\xi\right).$$

Let $P(\mathfrak{g})^{\mathbb{C}}$ denote the exemplification of the vector bundle $P(\mathfrak{g})$. Then we have the bundle inclusions $P(\mathfrak{g}) \subset P(\mathfrak{g})^{\mathbb{C}} \subset \mathrm{End}(E)$, where $\mathrm{End}(E)$ is the bundle of complex linear anthropomorphism of E and the latter inclusion is induced by the embedding ρ. Each connection $\alpha \in \mathcal{A}(P)$ induces a covariant derivative $d_{\alpha,\rho} : \Omega^k(\Sigma, E) \to \Omega^{k+1}(\Sigma, E)$, and its corresponding curvature $F_{\alpha,\rho} = d_{\alpha,\rho} \circ d_{\alpha,\rho} \in \Omega^2(\Sigma, P(\mathfrak{g}))$. Since the representation ρ is faithful, we have pointwise estimates of the form

$$c|F_{\alpha,\rho}| \leq |F_\alpha| \leq C|F_{\alpha,\rho}|,$$

which allow us to discuss curvature bounds in terms of F_α or $F_{\alpha,\rho}$. Furthermore, the map $\mathcal{A}(P) \to \mathcal{A}(E)$ is an embedding of $\Omega^1(\Sigma, P(\mathfrak{g}))$-spaces. Here $\Omega^1(\Sigma, P(\mathfrak{g}))$ acts on $\mathcal{A}(E)$ via the inclusion $\Omega^1(\Sigma, \mathrm{End}(E))$. In particular, restricting to the image of $\mathcal{A}(P)$ in $\mathcal{A}(E)$, the map (3.2) becomes an embedding

$$\begin{array}{ccc} \mathcal{A}(P) & \longrightarrow & \mathcal{C}(E) \\ \alpha & \longmapsto & \overline{\partial}_\alpha := \frac{1}{2}\left(d_{\alpha,\rho} + J_E d_{\alpha,\rho} \circ j_\Sigma\right) \end{array} \tag{3.3}$$

The image of (3.3) is the set of covariant derivatives which preserve the G-structure, and we denote it by $\mathcal{C}(P)$. See [30, Appendix C] for the case when $G = \mathrm{U}(n)$. The space $\mathcal{C}(P)$ is an affine space modeled on $\Omega^{0,1}(\Sigma, P(\mathfrak{g})^{\mathbb{C}})$. Similarly, we have seen that $\mathcal{A}(P)$ is an affine space modeled on $\Omega^1(\Sigma, P(\mathfrak{g}))$. The mapping (3.3) is affine under the identification $\Omega^1(\Sigma, P(\mathfrak{g})) \cong \Omega^{0,1}(\Sigma, P(\mathfrak{g})^{\mathbb{C}})$ sending μ to its anti-linear part $\mu^{0,1} := \frac{1}{2}\left(\mu + J_E \mu \circ j_\Sigma\right)$. To summarize, we have a commutative diagram

$$\begin{tikzcd} \mathcal{A}(P) \arrow[r, "\cong"] \arrow[d] & \mathcal{C}(P) \arrow[d] \\ \mathcal{A}(E) \arrow[r, "\cong"] & \mathcal{C}(E) \end{tikzcd}$$

As mentioned above, the *complexified gauge group* acts on $\mathcal{A}(P)$. To describe this, we need to first recall some basic properties of the complexification of compact Lie groups. See [25] or [23] on this material. Since G is compact and connected, there is a connected complex group $G^{\mathbb{C}}$ with $G \subset G^{\mathbb{C}}$ a maximal compact subgroup, and with Lie algebra $\mathfrak{g}^{\mathbb{C}}$, the complexification of $\mathfrak{g}$. This group $G^{\mathbb{C}}$ is unique up to natural isomorphism and is called the **complexification** of G. We may further assume that the representation ρ extends to an embedding $G^{\mathbb{C}} \hookrightarrow \mathrm{GL}(\mathbb{C}^n)$, and we identify $G^{\mathbb{C}}$ with its image (see [23, Proof of Theorem 1.7]). Then we have

$$G = \left\{ u \in G^{\mathbb{C}} \middle| u^\dagger u = \mathrm{Id} \right\},$$

where $u^\dagger$ denotes the conjugate transpose on $\mathrm{GL}(\mathbb{C}^n)$. It follows that we can write

$$G^{\mathbb{C}} = \left\{ g\exp(i\xi) \,|\, g \in G, \xi \in \mathfrak{g} \right\},$$

and this decomposition is unique. The same holds true if we replace $g\exp(i\xi)$ by $\exp(i\xi)g$. It is then immediate that

$$g\exp(i\xi) = \exp(i\mathrm{Ad}(g)\xi)g \tag{3.4}$$

for all $g \in G$ and all $\xi \in \mathfrak{g}$.

We can now define the **complexified gauge group on** P to be

$$\mathcal{G}(P)^{\mathbb{C}} := \Gamma(P \times_G G^{\mathbb{C}}).$$

Similarly to the real case, we may identify $\Omega^0(\Sigma, P\left(\mathfrak{g}\right)^{\mathbb{C}})$ with the Lie algebra $\mathrm{Lie}\left(\mathcal{G}(P)^{\mathbb{C}}\right)$ via the map

$$\xi \mapsto \exp(-\xi), \tag{3.5}$$

(compare this with (2.14)) hence the Lie group theoretic exponential map on $\mathcal{G}(P)^{\mathbb{C}}$ is given pointwise by the exponential map on $G^{\mathbb{C}}$. It follows by the analogous properties of $G^{\mathbb{C}}$ that each element of $\mathcal{G}(P)^{\mathbb{C}}$ can be written uniquely in the form

$$g\exp(i\xi) \tag{3.6}$$

for some $g \in \mathcal{G}(P)$ and $\xi \in \Omega^0(\Sigma, P(\mathfrak{g}))$, and (3.4) continues to hold with g, ξ interpreted as elements of $\mathcal{G}(P), \Omega^0(\Sigma, P(\mathfrak{g}))$, respectively.

The complexified gauge group acts on $\mathcal{C}(P)$ by

$$\begin{array}{rcl} \mathcal{G}(P)^{\mathbb{C}} \times \mathcal{C}(P) & \longrightarrow & \mathcal{C}(P) \\ (u, \overline{D}) & \longmapsto & u \circ \overline{D} \circ u^{-1} \end{array} \tag{3.7}$$

Viewing $\mathcal{G}(P)$ as a subgroup of $\mathcal{G}(P)^{\mathbb{C}}$ in the obvious way, then the identification (3.3) is $\mathcal{G}(P)$-equivariant. We can then use (3.3) and (3.7) to define an action of the larger group $\mathcal{G}(P)^{\mathbb{C}}$ on $\mathcal{A}(P)$, extending the $\mathcal{G}(P)$-action. We denote the action of $u \in \mathcal{G}(P)^{\mathbb{C}}$ on α by $u^*\alpha$. Explicitly, the action on $\mathcal{A}(P)$ takes the form

$$d_{u^*\alpha,\rho} = (u^\dagger)^{-1} \circ \partial_\alpha \circ u^\dagger + u \circ \overline{\partial}_\alpha \circ u^{-1}.$$

where the dagger is applied point-wise. Of particular note is that the infinitesimal action at $\alpha \in \mathcal{A}(P)$ is given by

$$\begin{array}{rcl} \Omega^0\left(\Sigma, P(\mathfrak{g})^{\mathbb{C}}\right) & \longrightarrow & \Omega^1\left(\Sigma, P(\mathfrak{g})\right) \\ \xi + i\zeta & \longmapsto & d_{\alpha,\rho}\xi + *d_{\alpha,\rho}\zeta \end{array} \tag{3.8}$$

More generally, the derivative of the map $(u, \alpha) \mapsto u^*\alpha$ at (u, α) with $u \in \mathcal{G}(P)$ (an element of the *real* gauge group) is a map

$$u\left(\Omega^0(\Sigma, P(\mathfrak{g})) \oplus i\Omega^0(\Sigma, P(\mathfrak{g}))\right) \times \Omega^1(\Sigma, P(\mathfrak{g})) \longrightarrow \Omega^1(\Sigma, P(\mathfrak{g}))$$

given by

$$\begin{array}{rcl} (u(\xi + i\zeta), \eta) & \longmapsto & \mathrm{Ad}(u)\left(d_\alpha\xi + *d_\alpha\zeta + \eta\right) \\ & = & \left\{d_{u^*\alpha}(u\xi) + *d_{u^*\alpha}(u\zeta)\right\} u^{-1} + \mathrm{Ad}(u)\eta \end{array} \tag{3.9}$$

Compare with (2.15). Here we are using the fact that $\mathcal{G}(P)^{\mathbb{C}}$ and its Lie algebra both embed into the space $\Gamma(P \times_G \mathrm{End}(\mathbb{C}^n))$, and so it makes sense to multiply Lie group and Lie algebra elements. See Remark 2.2.2. The curvature transforms under $u \in \mathcal{G}(P)^{\mathbb{C}}$ by

$$u^{-1} \circ F_{u^*\alpha,\rho} \circ u = F_{\alpha,\rho} + \overline{\partial}_\alpha \left(h^{-1}\partial_\alpha h\right), \tag{3.10}$$

where we have set $h = u^\dagger u$. We will mostly be interested in this action when $u = \exp(i\xi)$ for $\xi \in \Omega^0(\Sigma, P(\mathfrak{g}))$, in which case the action can be written as

$$\exp(-i\xi) \circ F_{\exp(i\xi)^*\alpha,\rho} \circ \exp(i\xi) \quad = \quad *\mathcal{F}(\alpha,\xi), \tag{3.11}$$

where we have set

$$\mathcal{F}(\alpha,\xi) := *\left(F_{\alpha,\rho} + \overline{\partial}_\alpha\left(\exp(-2i\xi)\partial_\alpha \exp(2i\xi)\right)\right). \tag{3.12}$$

It will be useful to define the **(real) gauge group on** E and the **complexified gauge group on** E by, respectively,

$$\mathcal{G}(E) := \Gamma(P \times_G \mathrm{U}(n)), \qquad \mathcal{G}(E)^{\mathbb{C}} := \Gamma(P \times_G \mathrm{GL}(\mathbb{C}^n)).$$

(Note that the complexification of $\mathrm{U}(n)$ is $\mathrm{GL}(\mathbb{C}^n)$, so this terminology is consistent, and in fact motivates, the terminology above.) These are both Lie groups with Lie algebras

$$\mathrm{Lie}(\mathcal{G}(E)) = \Gamma(P \times_G \mathfrak{u}(n)), \qquad \mathrm{Lie}(\mathcal{G}(E)) = \Gamma(P \times_G \mathrm{End}(\mathbb{C}^n)),$$

where we are identifying $\mathrm{End}(\mathbb{C}^n)$ with the Lie algebra of $\mathrm{GL}(\mathbb{C}^n)$. We have the obvious inclusions

$$\begin{array}{ccc} \mathcal{G}(P) & \longrightarrow & \mathcal{G}(P)^{\mathbb{C}} \\ \downarrow & & \downarrow \\ \mathcal{G}(E) & \longrightarrow & \mathcal{G}(E)^{\mathbb{C}} \end{array}$$

The space $\mathcal{G}(E)^{\mathbb{C}}$ acts on $\mathcal{C}(E)$ by the map

$$\begin{aligned} \mathcal{G}(E)^{\mathbb{C}} \times \mathcal{C}(E) &\longrightarrow \mathcal{C}(E) \\ (u, \overline{D}) &\longmapsto u \circ \overline{D} \circ u^{-1} \end{aligned} \tag{3.13}$$

Using (3.2), this induces an action of $\mathcal{G}(E)^{\mathbb{C}}$ on $\mathcal{A}(E)$ (hence an action of $\mathcal{G}(E)$ on $\mathcal{A}(E)$), though, neither $\mathcal{G}(E)^{\mathbb{C}}$ nor $\mathcal{G}(E)$ restrict to actions on $\mathcal{A}(P)$, unless $G = \mathrm{U}(n)$.

Finally, we mention that the vector spaces

$$\mathrm{Lie}(\mathcal{G}(E)), \qquad \mathrm{Lie}(\mathcal{G}(E)^{\mathbb{C}}), \qquad \text{and} \qquad \mathrm{Lie}(\mathcal{G}(P)^{\mathbb{C}}),$$

admit Sobolev completions. For example, the space $\mathrm{Lie}(\mathcal{G}(P)^{\mathbb{C}})^{k,q}$ is the $W^{k,q}$-completion of the vector space $\Gamma(P \times_G P(\mathfrak{g})^{\mathbb{C}})$. When we are in the continuous range for Sobolev embedding (e.g., when $kq > 2$) then these are Banach Lie algebras. Similarly, when we are in the continuous range we can form the Banach Lie groups

$$\mathcal{G}^{k,q}(E), \qquad \mathcal{G}^{k,q}(E)^{\mathbb{C}}, \qquad \text{and} \qquad \mathcal{G}^{k,q}(P)^{\mathbb{C}},$$

by taking the $W^{k,q}$-completions of the groups of smooth functions

$$\mathcal{G}(E), \qquad \mathcal{G}(E)^{\mathbb{C}}, \qquad \text{and} \qquad \mathcal{G}(P)^{\mathbb{C}}$$

which we view as lying in the vector space $\Gamma(P \times_G \mathrm{End}(\mathbb{C}^n))^{k,q}$ (compare with section 2.2.2). The complexified gauge action extends to a smooth action of $\mathcal{G}^{k,q}(E)^{\mathbb{C}}$ on $A^{k-1,q}(E)$, and this restricts to a smooth action of $\mathcal{G}^{k,q}(P)^{\mathbb{C}}$ on $A^{k-1,q}(P)$.

Proof of Theorem 3.1.1. Set $G = \mathrm{PSU}(r)$. Suppose we can define NS_P on the set

$$\left\{ \alpha \in \mathcal{A}^{1,q}(P) \,\middle|\, \mathrm{dist}_{W^{1,q}} \left(\alpha, \mathcal{A}^{1,q}_{\mathrm{flat}}(P) \right) < \epsilon_0 \right\}, \tag{3.14}$$

for some $\epsilon_0 > 0$, and show that it satisfies the desired properties on this smaller domain. Then the $\mathcal{G}^{2,q}$-equivariance will imply that it extends uniquely to the flow-out by the real gauge group:

$$\left\{ u^*\alpha \in \mathcal{A}^{1,q}(P) \,\middle|\, u \in \mathcal{G}^{2,q}(P), \qquad \mathrm{dist}_{W^{1,q}}\left(\alpha, \mathcal{A}^{1,q}_{\text{flat}}(P)\right) < \epsilon_0 \right\},$$

and continues to have the desired properties on this larger domain. The next claim shows that this flow-out contains a neighborhood of the form appearing in the domain in (3.1), thereby reducing the problem to defining NS_P on a set of the form (3.14).

Claim: For any $\widetilde{\epsilon}_0 > 0$, there is some $\epsilon_0 > 0$ with

$$\{\alpha \in \mathcal{A}^{1,q}(P) |\, \|F_\alpha\|_{L^q} < \epsilon_0\}$$

$$\subseteq \left\{ u^*\alpha \in \mathcal{A}^{1,q}(P) \,\middle|\, u \in \mathcal{G}^{2,q}(P), \qquad \mathrm{dist}_{W^{1,q}}\left(\alpha, \mathcal{A}^{1,q}_{\text{flat}}(P)\right) < \widetilde{\epsilon}_0 \right\}.$$

For sake of contradiction, suppose that for all $\epsilon > 0$ there is a connection α with $\|F_\alpha\|_{L^q} < \epsilon$, but

$$\|u^*\alpha - \alpha_0\|_{W^{1,q}} \geq \widetilde{\epsilon}_0, \qquad \forall u \in \mathcal{G}^{2,q}(P), \qquad \forall \alpha_0 \in \mathcal{A}^{1,q}_{\text{flat}}(P) \tag{3.15}$$

So we may find a sequence of connections α_ν with $\|F_{\alpha_\nu}\|_{L^2} \to 0$, but (3.15) holds with α_ν replacing α. By Uhlenbeck's Weak Compactness Theorem 2.2.9, there is a sequence of gauge transformations $u_\nu \in \mathcal{G}^{2,q}(P)$ such that, after possibly passing to a subsequence, $u_\nu^*\alpha_\nu$ converges weakly in $W^{1,q}$ to a limiting connection $\alpha_\flat$. The condition on the curvature implies that $\alpha_\flat \in \mathcal{A}^{1,q}_{\text{flat}}(P)$ is flat. Moreover, the embedding $W^{1,q} \hookrightarrow L^{2q}$ is compact, so the weak $W^{1,q}$-convergence of $u_\nu^*\alpha_\nu$ implies that $u_\nu^*\alpha_\nu$ converges strongly to $\alpha_\flat$ in L^{2q}. By redefining u_ν if necessary, we may suppose that $u_\nu^*\alpha_\nu$ is in Coulomb gauge with respect to $\alpha_\flat$ (section 2.2.2):

$$d^*_{\alpha_0}(u_\nu^*\alpha_\nu - \alpha_\flat) = 0,$$

and still retain the fact that $u_\nu^*\alpha_\nu$ converges to α_∞ strongly in L^{2q} (see Theorem 2.2.11, use $p = \frac{1}{2}(1+q)$ when $1 < q \leq 2$). This gives

$$\begin{aligned}\|u_\nu^*\alpha_\nu - \alpha_\flat\|_{W^{1,q}}^q &= \|u_\nu^*\alpha_\nu - \alpha_\flat\|_{L^q}^q + \|d_{\alpha_\flat}\left(u_\nu^*\alpha_\nu - \alpha_\flat\right)\|_{L^q}^q + \|d_{\alpha_\flat}^*\left(u_\nu^*\alpha_\nu - \alpha_\flat\right)\|_{L^q}^q \\ &\leq \|u_\nu^*\alpha_\nu - \alpha_\flat\|_{L^q}^q + \|F_{\alpha_\nu}\|_{L^q}^q + \tfrac{1}{2}\|u_\nu^*\alpha_\nu - \alpha_\flat\|_{L^{2q}}^q \\ &\leq C\left(\|u_\nu^*\alpha_\nu - \alpha_\flat\|_{L^{2q}}^q + \|F_{\alpha_\nu}\|_{L^q}^q\right)\end{aligned}$$

where we have used the formula

$$F_{\alpha_\flat+\mu} = d_{\alpha_\flat}(\mu) + \frac{1}{2}\left[\mu \wedge \mu\right].$$

Hence

$$\|u_\nu^*\alpha_\nu - \alpha_\flat\|_{W^{1,q}}^q \longrightarrow 0,$$

in contradiction to (3.15). This proves the claim.

To define NS_P, it therefore suffices to show that for α sufficiently $W^{1,q}$-close to $\mathcal{A}_{\mathrm{flat}}(P)$ there is a unique $\Xi(\alpha) \in \Omega^0(\Sigma, P(\mathfrak{g}))$ close to 0, with $F_{\exp(i\Xi(\alpha))^*\alpha,\rho} = 0$. Once we have shown this, then we will define

$$\mathrm{NS}_P(\alpha) := \exp(i\Xi(\alpha))^*\alpha.$$

In light of (3.10) finding $\Xi(\alpha)$ is equivalent to solving for ξ in $\mathcal{F}(\alpha,\xi) = 0$. To do this we need to pass to suitable Sobolev completions.

It follows from the formula (3.10) and the Sobolev embedding and multiplication theorems that $\mathcal{F}$, defined in (3.12), extends to a map

$$\mathcal{A}^{1,q}(P) \times \mathrm{Lie}(\mathcal{G}(P))^{2,q} \longrightarrow \mathrm{Lie}(\mathcal{G}(P))^{0,q},$$

whenever $q > 1$. Suppose α_0 is a flat connection. The linearization of $\mathcal{F}$ at $(\alpha_0, 0)$ in the direction of $(0,\xi)$ is

$$D_{(\alpha_0,0)}\mathcal{F}(0,\xi) = 2J_E * \overline{\partial}_{\alpha_0}\partial_{\alpha_0}(\xi),$$

where we have used the fact that $d_{\alpha,\rho}$ commutes with $J_E = i$ (this is because the complex structure J_E is constant and the elements of $\mathcal{A}(P)$ are unitary). Observe that j_Σ acts by the Hodge star on vectors, so

$$d_{\alpha_0,\rho}(d_{\alpha_0,\rho} \circ j_\Sigma) = d_{\alpha_0,\rho} * d_{\alpha_0,\rho}, \qquad (d_{\alpha_0,\rho} \circ j_\Sigma) d_{\alpha_0,\rho} = F_{\alpha_0,\rho} \circ (j_\Sigma, \mathrm{Id}). \tag{3.16}$$

Using this and the fact that $F_{\alpha_0,\rho} = 0$, we have

$$D_{(\alpha_0,0)}\mathcal{F}(0,\xi) = \frac{1}{2}\Delta_{\alpha_0,\rho}\xi$$

where $\Delta_{\alpha_0,\rho} = d^*_{\alpha_0,\rho} d_{\alpha_0,\rho} + d_{\alpha_0,\rho} d^*_{\alpha_0,\rho}$ is the Laplacian. By assumption, all flat connections are irreducible, so Hodge theory tells us that the operator $\Delta_{\alpha_0,\rho} : \mathrm{Lie}(\mathcal{G}(P))^{2,q} \to \mathrm{Lie}(\mathcal{G}(P))^{0,q}$ is an isomorphism. Since α_0 is flat, the pair $(\alpha_0, 0)$ is clearly a solution to $\mathcal{F}(\alpha,\xi) = 0$. It therefore follows by the implicit function theorem that there are $\epsilon_{\alpha_0}, \epsilon'_{\alpha_0} > 0$ such that, for any $\alpha \in \mathcal{A}^{1,q}$ with $\|\alpha - \alpha_0\|_{W^{1,q}} < \epsilon_{\alpha_0}$, there is a unique

$$\Xi = \Xi(\alpha) \in \mathrm{Lie}(\mathcal{G}(P))^{2,q}$$

with $\|\Xi(\alpha)\|_{W^{2,q}} < \epsilon'_{\alpha_0}$ and

$$\mathcal{F}(\alpha, \Xi(\alpha)) = 0.$$

The implicit function theorem also implies that $\Xi(\alpha)$ varies smoothly α in the $W^{1,q}$-topology. Moreover, by the uniqueness assertion, it follows that $\Xi(\alpha) = 0$ if α is flat.

We need to show that ϵ_{α_0} and ϵ'_{α_0} can be chosen to be independent of $\alpha_0 \in \mathcal{A}^{1,q}_{\mathrm{flat}}(P)$. Since the moduli space of flat connections is compact, it suffices to show that $\epsilon_{\alpha_0} = \epsilon_{u^*\alpha_0}$, for all real gauge transformations $u \in \mathcal{G}^{2,q}(P)$, and likewise for ϵ'_{α_0}. Fix $u \in \mathcal{G}^{2,q}(P)$ and α a connection $W^{1,q}$-close to α_0, then find $\Xi(\alpha)$ as above. By (3.4) and the statement following (3.6) we have

$$u \exp(i\Xi(\alpha)) = \exp(i\mathrm{Ad}(u)\Xi(\alpha))u. \tag{3.17}$$

Since the curvature is $\mathcal{G}^{2,q}(P)$-equivariant, we also have

$$\begin{aligned} 0 &= \mathrm{Ad}(u) F_{\exp(i\Xi)^*\alpha} \\ &= F_{(u\exp(i\Xi))^*\alpha} \\ &= F_{\exp(i\mathrm{Ad}(u)\Xi)^*(u^*\alpha)} \end{aligned}$$

so $\Xi(u^*\alpha) = \mathrm{Ad}(u)^*\Xi(\alpha)$ since this is the defining property of $\Xi(u^*\alpha)$. It follows immediately that $\epsilon_{u^*\alpha_0} = \epsilon_{\alpha_0}$ and $\epsilon'_{u^*\alpha_0} = \epsilon'_{\alpha_0}$, so we can take ϵ_0 to be the minimum of $\inf_{[\alpha_0]\in M(P)} \epsilon_{\alpha_0} > 0$ and $\inf_{[\alpha_0]\in M(P)} \epsilon'_{\alpha_0} > 0$. This also shows that NS_P is $\mathcal{G}^{2,q}(P)$-equivariant.

Finally, we show that $\mathrm{NS}_P(\alpha)$ depends smoothly on α in the L^p-topology for $p > 2$. It suffices to show that $\alpha \mapsto \Xi(\alpha)$ extends to a map $\mathcal{A}^{0,p} \to \mathrm{Lie}(\mathcal{G}(P))^{1,p}$ which is smooth with respect to the specified topologies. To see this, note that $\mathcal{F}$ from (3.12) is well-defined as a map

$$\mathcal{A}^{0,p}(P) \times \mathrm{Lie}(\mathcal{G}(P))^{1,p} \longrightarrow \mathrm{Lie}(\mathcal{G}(P))^{-1,p}.$$

and is smooth with respect to the specified topologies (the restriction to $p > 2$ is required so that Sobolev multiplication is well-defined, see Section 2.2.2). Then the implicit function theorem argument we gave above holds verbatim to show that for each α sufficiently L^p-close to $\mathcal{A}^{0,p}_{\mathrm{flat}}(P)$, there is a unique $W^{1,p}$-small $\widetilde{\Xi}(\alpha) \in \mathrm{Lie}(\mathcal{G}(P))^{1,p}$ such that $\exp(i\widetilde{\Xi}(\alpha))^*\alpha$ is flat. Moreover, the assignment

$$\mathcal{A}^{0,p}_{\mathrm{flat}}(P) \longrightarrow \mathrm{Lie}(\mathcal{G}(P))^{1,p}, \qquad \alpha \longmapsto \widetilde{\Xi}(\alpha)$$

is smooth. The uniqueness of $\widetilde{\Xi}(\alpha)$ and $\Xi(\alpha)$ ensures that the former is indeed an extension of the latter. $\square$

Remark 3.1.3. *Let $\Pi : \mathcal{A}^{1,q}_{\mathrm{flat}}(P) \to M(P)$ denote the projection. The above proof shows that the composition $\Pi \circ \mathrm{NS}_P$ is invariant under a small neighborhood of $\mathcal{G}^{2,q}(P)$ in $\mathcal{G}^{2,q}(P)^{\mathbb{C}}$. Indeed, α and $\exp(i\xi)^*\alpha$ both map to the same flat connection under NS_P whenever they are both in the domain of NS_P.*

3.1.1 Analytic properties of almost flat connections

This section is of a preparatory nature. The results extend several elliptic properties, which are standard for flat connections, to connections with small curvature. The following lemma addresses elliptic regularity for the operator d_α on 0-forms.

Lemma 3.1.4. *Suppose G is a compact Lie group, Σ is a closed oriented Riemannian surface and $P \to \Sigma$ is a principal G-bundle such that all flat connections are irreducible. Let $1 < q < \infty$. Then there are constants $C > 0$ and $\epsilon_0 > 0$ with the following significance.*

(i) Suppose that either $\alpha \in \mathcal{A}^{1,q}(P)$ with $\|F_\alpha\|_{L^q(\Sigma)} < \epsilon_0$, or $\alpha \in \mathcal{A}^{0,q}(P)$ with $\|\alpha - \alpha_\flat\|_{L^{2q}(\Sigma)} < \epsilon_0$ for some $\alpha_\flat \in \mathcal{A}^{0,2q}_{\mathrm{flat}}(P)$. Then the map

$$d_\alpha : W^{1,q}(P(\mathfrak{g})) \longrightarrow L^q(P(\mathfrak{g}))$$

is a Banach space isomorphism onto its image. Moreover, the following estimate holds

$$\|f\|_{W^{1,q}(\Sigma)} \leq C\|d_\alpha f\|_{L^q(\Sigma)} \tag{3.18}$$

for all $f \in W^{1,q}(P(\mathfrak{g}))$.

(ii) For all $\alpha \in \mathcal{A}^{1,q}(P)$ with $\|F_\alpha\|_{L^q(\Sigma)} < \epsilon_0$, the Laplacian

$$d_\alpha^* d_\alpha : W^{2,q}(P(\mathfrak{g})) \longrightarrow L^q(P(\mathfrak{g}))$$

is a Banach space isomorphism. Moreover, the following estimate holds

$$\|f\|_{W^{2,q}(\Sigma)} \leq C\|d_\alpha * d_\alpha f\|_{L^q(\Sigma)} \tag{3.19}$$

for all $f \in W^{2,q}(P(\mathfrak{g}))$.

Proof. This is basically the statement of [12, Lemma 7.6], but adjusted a little to suit our situation. We prove (ii), the proof of (i) is similar.

The assumptions on the bundle imply that all flat connections $\alpha_\flat$ are irreducible, and so the kernel and cokernel of the elliptic operator

$$d^*_{\alpha_\flat} d_{\alpha_\flat} : W^{2,q}(P(\mathfrak{g})) \longrightarrow L^q(P(\mathfrak{g}))$$

are trivial. In particular, we have an estimate

$$\|f\|_{W^{2,q}} \leq C \|d_{\alpha_\flat} * d_{\alpha_\flat} f\|_{L^q}$$

for all $f \in W^{2,q}(P(\mathfrak{g}))$, and the statement holds when $\alpha = \alpha_\flat$ is flat.

Next, let $\alpha \in \mathcal{A}^{1,q}(P)$ and $\alpha_\flat \in \mathcal{A}^{1,q}_{\text{flat}}(P)$. Then, by the above discussion, and the relation

$$d_{\alpha_\flat} f = d_\alpha f + [\alpha_\flat - \alpha, f]\,,$$

we have

$$\begin{aligned}
\|f\|_{W^{2,q}} &\leq C \|d_{\alpha_\flat} * d_{\alpha_\flat} f\|_{L^q} \\
&\leq C\left\{\|d_\alpha * d_\alpha f\|_{L^q} \right.\\
&\qquad \left. + \|d_\alpha\left[*(\alpha - \alpha_\flat), f\right]\|_{L^q} + \|[\alpha - \alpha_\flat \wedge [*(\alpha - \alpha_\flat), f]]\|_{L^q}\right\} \\
&\leq C\left\{\|d_\alpha * d_\alpha f\|_{L^q} \right.\\
&\qquad \left. + C'\left(\|d_\alpha * (\alpha - \alpha_\flat)\|_{L^q} + \|\alpha - \alpha_\flat\|_{L^{2q}}\right)\|f\|_{W^{2,q}}\right\},
\end{aligned}$$

for all $f \in W^{2,q}(P(\mathfrak{g}))$, where we have used the embeddings

$$W^{2,q} \hookrightarrow W^{1,q} \qquad \text{and} \qquad W^{2,q} \hookrightarrow L^\infty$$

in the last step. Now suppose that $\|\alpha - \alpha_\flat\|_{L^{2q}} < 1/2CC'$ is small. Then by composing $\alpha_\flat$ with a suitable gauge transformation, we may suppose α is in Coulomb gauge with respect to $\alpha_\flat$, and still retain the fact that $\|\alpha - \alpha_\flat\|_{L^{2q}} < 1/2CC'$. Then the above gives

$$\|f\|_{W^{2,q}} \leq C\|d_\alpha * d_\alpha f\|_{L^q} + \frac{1}{2}\|f\|_{W^{2,q}},$$

which shows that $d_\alpha^* d_\alpha$ is injective when sufficiently L^{2q}-close to the space of flat connections.

Now we prove the lemma. Suppose the result of (ii) in the lemma does not hold. Then there is some sequence of connections α_ν with $\|F_{\alpha_\nu}\|_{L^q} \to 0$, but the estimate (3.19) does not hold for any $C > 0$. By Uhlenbeck's Weak Compactness Theorem 2.2.9, after possibly passing to a subsequence, there is some sequence of gauge transformations u_ν, and a limiting flat connection $\alpha_\flat$, such that

$$\|\alpha_\nu - u_\nu^* \alpha_\flat\|_{L^{2q}} \longrightarrow 0.$$

So the discussion of the previous paragraph shows that, for ν sufficiently large, the estimate (3.19) holds with α replaced by α_ν. This is a contradiction, and it proves the lemma. $\square$

Now we move on to study the action of d_α on 1-forms. First we establish a Hodge-decomposition result for connections with small curvature. For $2 \leq q < \infty$ and $k \in \mathbb{Z}$, let $V^{k,q}$ denote the $W^{k,q}$-closure of a vector subspace $V \subseteq W^{k,q}(T^*\Sigma \otimes P(\mathfrak{g}))$. The standard Hodge decomposition (2.9) reads

$$W^{k,q}(T^*\Sigma \otimes P(\mathfrak{g})) = H^1_{\alpha_\flat} \oplus \left(\mathrm{Im}\, d_{\alpha_\flat}\right)^{k,q} \oplus \left(\mathrm{Im}\, d^*_{\alpha_\flat}\right)^{k,q}, \tag{3.20}$$

for any flat connection $\alpha_\flat$. Here $H^1_{\alpha_\flat}$ is finite dimensional (with a dimension that is independent of $\alpha_\flat \in \mathcal{A}_\flat$), and so is equal to its $W^{k,q}$-closure. Furthermore, the direct sum in (3.20) is L^2-orthogonal, even though the spaces need not be complete in the L^2-metric. We have a similar situation whenever α has small curvature, as the next lemma shows.

Lemma 3.1.5. *Assume that $P \to \Sigma$ satisfies the conditions of Lemma 3.1.4, and let $1 < q < \infty$ and $k \geq 0$. Then there are constants $\epsilon_0 > 0$ and $C > 0$ with the following significance. If $\alpha \in \mathcal{A}^{1,q}(P)$ has $\|F_\alpha\|_{L^q(\Sigma)} < \epsilon_0$, then*

$$H^1_\alpha := (\ker d_\alpha)^{k,q} \cap (\ker d^*_\alpha)^{k,q} \subseteq W^{k,q}(T^*\Sigma \otimes P(\mathfrak{g}))$$

has finite dimension equal to $\dim H^1_{\alpha_\flat}$*, for any flat connection* $\alpha_\flat$*. Furthermore, the space* H^1_α *equals the* L^2*-orthogonal complement of the image of* $d_\alpha \oplus d^*_\alpha$*:*

$$H^1_\alpha = \left((\mathrm{Im}\, d_\alpha)^{k,q} \oplus (\mathrm{Im}\, d^*_\alpha)^{k,q} \right)^\perp ,$$

and so we have a direct sum decomposition

$$W^{k,q}\left(T^*\Sigma \otimes P(\mathfrak{g})\right) = H^1_\alpha \oplus \left((\mathrm{Im}\, d_\alpha)^{k,q} \oplus (\mathrm{Im}\, *d_\alpha)^{k,q} \right). \tag{3.21}$$

In particular, the L^2*-orthogonal projection*

$$\mathrm{proj}_\alpha : W^{k,q}(T^*\Sigma \otimes P(\mathfrak{g})) \longrightarrow H^1_\alpha \tag{3.22}$$

is well-defined.

Remark 3.1.6. *It follows by elliptic regularity that the space*

$$(\ker d_\alpha)^{k,q} \cap (\ker d^*_\alpha)^{k,q}$$

consists of smooth forms. In fact, when $k - 2/q \geq k' - 2/q'$*, the inclusion* $W^{k,q} \subseteq W^{k',q'}$ *restricts to an inclusion of finite-dimensional spaces*

$$(\ker d_\alpha)^{k,q} \cap (\ker d^*_\alpha)^{k,q} \hookrightarrow (\ker d_\alpha)^{k',q'} \cap (\ker d^*_\alpha)^{k',q'} ,$$

and this map is onto by dimensionality. Hence, the definition of H^1_α *is independent of the choice of* k, q*.*

Proof of Lemma 3.1.5. We first show that, when $\|F_\alpha\|_{L^q(\Sigma)}$ is sufficiently small, we have a direct sum decomposition

$$W^{k,q}\left(T^*\Sigma \otimes P(\mathfrak{g})\right) = H^1_\alpha \oplus (\mathrm{Im}\, d_\alpha)^{k,q} \oplus (\mathrm{Im}\, *d_\alpha)^{k,q} .$$

We prove this in the case $k = 0$, the case $k > 0$ is similar but slightly easier. By definition of H^1_α, it suffices to show that the images of d_α and $*d_\alpha$ intersect trivially. Towards this end, write $d_\alpha f = *d_\alpha g$ for 0-forms f, g of Sobolev class $L^q = W^{0,q}$. Acting by d_α and then $d_\alpha *$ gives

$$[F_\alpha, f] = d_\alpha * d_\alpha g, \qquad [F_\alpha, g] = -d_\alpha * d_\alpha f.$$

A priori, $d_\alpha * d_\alpha g$ and $d_\alpha * d_\alpha f$ are only of Sobolev class $W^{-2,q}$, however, the left-hand side of each of these equations is in L^r, where $1/r = 1/q + 1/p$. So elliptic regularity implies that f and g are each $W^{2,q}$. (This bootstrapping can be continued to show that f, g are smooth, but we will see in a minute that they are both zero.) By Lemma 3.1.4 and the embedding $W^{2,q} \hookrightarrow L^\infty$ it follows that, whenever $\|F_\alpha\|_{L^q}$ is sufficiently small, we have

$$\begin{aligned} \|f\|_{L^\infty} &\leq C\|d_\alpha * d_\alpha f\|_{L^q} \\ &= C\| \, [F_\alpha, g] \, \|_{L^q} \\ &\leq C\|F_\alpha\|_{L^q}\|g\|_{L^\infty}. \end{aligned}$$

Similarly, $\|g\|_{L^\infty} \leq C\|F_\alpha\|_{L^q}\|g\|_{L^\infty}$, and hence

$$\|f\|_{L^\infty} \leq C^2\|F_\alpha\|^2_{L^q}\|f\|_{L^\infty}.$$

By requiring that $\|F_\alpha\|^2_{L^q} < C^{-2}$, this can be happen only if $f = g = 0$. This establishes the direct sum (3.21).

Now we prove that the dimension of H^1_α is finite and equals that of $H^1_{\alpha_\flat}$ for any flat connection $\alpha_\flat$. It is well-known that the operator

$$d_{\alpha_\flat} \oplus *d_{\alpha_\flat} : W^{k+1,q}(P(\mathfrak{g})) \oplus W^{k+1,q}(P(\mathfrak{g})) \longrightarrow W^{k,q}(T^*\Sigma \otimes P(\mathfrak{g}))$$

is elliptic, and hence Fredholm, whenever $\alpha_\flat$ is flat. The irreducibility condition implies that it has trivial kernel, and so has index given by $-\dim(H^1_{\alpha_\flat})$, which is a constant independent of $\alpha_\flat$. Then for any other connection α, the operator $d_\alpha \oplus *d_\alpha$ differs from $d_{\alpha_\flat} \oplus *d_{\alpha_\flat}$ by the compact operator

$$\mu \longmapsto [\alpha - \alpha_\flat, \mu] + *[\alpha - \alpha_\flat, \mu],$$

and so $d_\alpha \oplus *d_\alpha$ is Fredholm with the same index $-\dim(H^1_{\alpha_\flat})$ [30, Theorem A.1.5]. It follows from Lemma 3.1.4 that the (bounded) operator

$$d_\alpha \oplus *d_\alpha : W^{k+1,q}(P(\mathfrak{g})) \oplus W^{k+1,q}(P(\mathfrak{g})) \longrightarrow W^{k,q}(T^*\Sigma \oplus P(\mathfrak{g}))$$

is injective whenever $\|F_\alpha\|_{L^q(\Sigma)}$ is sufficiently small, and hence the cokernel has finite dimension $\dim(H^1_{\alpha_\flat})$:

$$\dim(H^1_\alpha) = \dim(H^1_{\alpha_\flat}).$$

This finishes the proof of Lemma 3.1.5. □

Next we show that the L^2-orthogonal projection to $H^1_\alpha = \ker d_\alpha \cap \ker d^*_\alpha$ depends smoothly on α in the L^q topology.

Proposition 3.1.7. *Suppose that $P \to \Sigma$ and $\epsilon_0 > 0$ are as in Lemma 3.1.5, and let $1 < q < \infty$. Then the assignment $\alpha \mapsto \mathrm{proj}_\alpha$ is affine-linear and bounded*

$$\|\mathrm{proj}_\alpha - \mathrm{proj}_{\alpha'}\|_{\mathrm{op},L^q} \leq C\|\alpha - \alpha'\|_{L^q(\Sigma)}, \tag{3.23}$$

provided $\|F_\alpha\|_{L^q}, \|F_{\alpha'}\|_{L^q} < \epsilon_0$, where $\|\cdot\|_{\mathrm{op},L^q}$ is the operator norm on the space of linear maps $L^q(T^*\Sigma \otimes P(\mathfrak{g})) \to L^q(T^*\Sigma \otimes P(\mathfrak{g}))$.

Proof. We will see that defining equations for proj_α are affine linear, and so the statement will follow from the implicit function theorem in the affine-linear setting.

First, we introduce the following shorthand:

$$W^{k,q}(\Omega^j) := W^{k,q}\left(\wedge^j T^*\Sigma \otimes P(\mathfrak{g})\right), \qquad L^q(\Omega^j) := W^{0,q}(\Omega^j).$$

Next, we note that for $\mu \in L^q(\Omega^1)$, the L^2-orthogonal projection $\mathrm{proj}_\alpha \mu$ is uniquely characterized by the following properties:

<u>Property A</u>: $\mu - \mathrm{proj}_\alpha \mu = d_\alpha u + d_\alpha^* v$, for some $(u, v) \in W^{1,q}(\Omega^0) \oplus W^{1,q}(\Omega^2)$,

<u>Property B</u>: $(\mathrm{proj}_\alpha \mu,\ d_\alpha a + d_\alpha^* b) = 0$, for all $(a, b) \in W^{1,q^*}(\Omega^0) \oplus W^{1,q^*}(\Omega^2)$,

where q^* is the Sobolev dual to q: $1/q + 1/q^* = 1$. Here and below, we are continuing to use the notation

$$(\mu, \nu) := \int_\Sigma \langle \mu \wedge * \nu \rangle$$

to denote the pairing on forms. Note that by Lemma 3.1.4, the operators d_α and d_α^* are injective on 0- and 2-forms, respectively, so any pair (u, v) satisfying Property A is unique.

Consider the map

$$\begin{aligned} &\left(\mathcal{A}^{0,q} \oplus L^q(\Omega^1)\right) \times \left(L^q(\Omega^1) \oplus W^{1,q}(\Omega^0) \oplus W^{1,q}(\Omega^0)\right) \\ &\longrightarrow \left(W^{1,q^*}(\Omega^0)\right)^* \oplus \left(W^{1,q^*}(\Omega^2)\right)^* \oplus L^q(\Omega^1) \end{aligned} \tag{3.24}$$

defined by

$$(\alpha, \mu; \nu, u, v) \longmapsto \left(\left(\nu, d_\alpha\left(\cdot\right)\right), \left(\nu, d_\alpha^*\left(\cdot\right)\right), \mu - \nu - d_\alpha u - d_\alpha^* v\right)$$

The key point is that a tuple (α, μ, ν, u, v) maps to zero under (3.24) if and only if this tuple satisfies Properties A and B above. By the identification $(W^{k,q^*})^* = W^{-k,q}$, can equivalently view (3.24) as a map

$$\begin{aligned} &\left(\mathcal{A}^{0,q} \times L^q(\Omega^1)\right) \times \left(L^q(\Omega^1) \times W^{1,q}(\Omega^0) \times W^{1,q}(\Omega^0)\right) \\ &\longrightarrow W^{-1,q}(\Omega^0) \oplus W^{-1,q}(\Omega^2) \oplus L^q(\Omega^1) \end{aligned} \tag{3.25}$$

defined by

$$(\alpha, \mu; \nu, u, v) \longmapsto (d_\alpha^* \nu,\ d_\alpha \nu,\ \mu - \nu - d_\alpha u - d_\alpha^* v)\,.$$

(The topology on each space is given by the Sobolev norm indicated in its exponent.)

Claim 1: The map (3.25) is bounded affine linear in the $\mathcal{A}^{0,q}$-variable, and bounded linear in the other 4 variables.

Claim 2: The linearization at $(\alpha, 0; 0, 0, 0)$ of (3.25) in the last 3-variables is a Banach space isomorphism, provided $\|\alpha - \alpha_\flat\|_{L^q}$ is sufficiently small for some flat connection $\alpha_\flat$.

Before proving the claims, we describe how they prove the lemma. Observe that $(\alpha, 0; 0, 0, 0)$ is clearly a zero of (3.25) for any α. Claim 1 implies that (3.25) is smooth, and so by Claim 2 we can use the implicit function theorem to show that, for each pair $(\alpha, \mu) \in \mathcal{A}^{0,q} \oplus L^q(\Omega^1)$, with $\|\alpha - \alpha_\flat\|_{L^q}$ sufficiently small, there is a unique $(\nu, u, v) \in L^q(\Omega^1) \oplus W^{1,q}(\Omega^0) \oplus W^{1,q}(\Omega^0)$ such that $(\alpha, \mu; \nu, u, v)$ is a zero of (3.25). (*A priori* this only holds for μ in a small neighborhood of the origin, but since (3.25) is linear in that variable, it extends to all μ.) It will then follow that $\nu = \mathrm{proj}_\alpha \mu$ depends smoothly on α in the L^q-metric. In fact, (3.25) is affine linear in α and linear in the other variables, so the uniqueness assertion of the implicit function theorem implies that proj_α depends affine-linearly on α, and so we get:

$$\begin{aligned}
\|\mathrm{proj}_\alpha - \mathrm{proj}_{\alpha_\flat}\|_{\mathrm{op}, L^q} &= \inf_{\|\mu\|_{L^q}=1} \left\| \left(\mathrm{proj}_\alpha - \mathrm{proj}_{\alpha_\flat}\right) \mu \right\|_{L^q} \\
&\leq C \inf_{\|\mu\|_{L^q}=1} C\|\alpha - \alpha_\flat\|_{L^q} \|\mu\|_{L^q} \\
&= C\|\alpha - \alpha_\flat\|_{L^q}.
\end{aligned}$$

This proves the lemma for all α sufficiently L^q-close to $\mathcal{A}_{\mathrm{flat}}$. To extend it to all α with $\|F_\alpha\|_{L^q}$ sufficiently small, one argues by contradiction as in the proof of Lemma 3.1.4, using Uhlenbeck's Weak Compactness Theorem 2.2.9. It therefore remains to prove the claims.

Proof of Claim 1: It suffices to verify boundedness for each of the three (codomain) components separately. The first component is the map

$$\begin{aligned} \mathcal{A}^{0,q} \times L^q(\Omega^1) &\longrightarrow W^{-1,q}(\Omega^0) \\ (\alpha, \nu) &\longmapsto d_\alpha^* \nu \end{aligned} \tag{3.26}$$

It is a standard consequence from the principle of uniform boundedness that a bilinear map is continuous if it is continuous in each variable separately. The same holds if the map is linear in one variable and affine-linear in the second, so it suffices to show that (3.26) is bounded in each of the two coordinates separately. Fix α and a flat connection $\alpha_\flat$. Then

$$\begin{aligned} \|d_\alpha \nu\|_{W^{-1,q}} &\leq \|d_{\alpha_\flat}\nu\|_{W^{-1,q}} + \|\,[\alpha - \alpha_\flat \wedge \nu]\,\|_{W^{-1,q}} \\ &\leq \|d_{\alpha_\flat}\nu\|_{W^{-1,q}} + \|\alpha - \alpha_\flat\|_{L^q}\|\nu\|_{L^q} \\ &\leq C\,(1 + \|\alpha - \alpha_\flat\|_{L^q})\,\|\nu\|_{L^q} \end{aligned}$$

which shows that the map is bounded in the variable ν, with α fixed. Next, fix ν and write

$$\begin{aligned} \|d_\alpha \nu - d_{\alpha_\flat}\nu\|_{W^{-1,q}} &= \|\,[\alpha - \alpha_\flat \wedge \nu]\,\|_{W^{-1,q}} \\ &\leq \|\nu\|_{L^q}\|\alpha - \alpha_\flat\|_{L^q} \end{aligned}$$

which shows it is bounded in the α-variable. This shows the first component of (3.25) is bounded. The other two components are similar.

Proof of Claim 2: The linearization of (3.25) at $(\alpha, 0; 0, 0, 0)$ in the last three variables is the map

$$\begin{aligned} L^q(\Omega^1) \times W^{1,q}(\Omega^0) \times W^{1,q}(\Omega^0) &\longrightarrow W^{-1,q}(\Omega^0) \oplus W^{-1,q}(\Omega^2) \oplus L^q(\Omega^1) \\ (\nu, u, v) &\longmapsto (d_\alpha^* \nu,\ d_\alpha \nu,\ -\nu - d_\alpha u - d_\alpha^* v) \end{aligned}$$

By Claim 1, this is bounded linear, so by the open mapping theorem, it suffices to show that it is bijective. Suppose

$$(d_\alpha^* \nu,\ d_\alpha \nu;\ -\nu - d_\alpha u - d_\alpha^* v) = (0, 0, 0). \tag{3.27}$$

Then by Lemma 3.1.5, we can write ν uniquely as

$$\nu = \nu_H + d_\alpha a + d_\alpha^* b$$

for $\nu_H \in H_\alpha^1 = \ker d_\alpha \cap \ker d_\alpha^*$, and $(a, b) \in W^{1,q}(\Omega^0) \times W^{1,q}(\Omega^2)$, provided $\|F_\alpha\|_{L^q}$ is sufficiently small. This uniqueness, together with the first two components of (3.27), imply that $\nu = \nu_H$. The third component then reads

$$\nu_H = -d_\alpha u - d_\alpha^* v,$$

which is only possible if $\nu_H = d_\alpha u = d_\alpha^* v = 0$. By Lemma 3.1.4, this implies $(\nu, u, v) = (0, 0, 0)$, which proves injectivity.

To prove surjectivity, suppose the contrary. Then by the Hahn-Banach theorem, there are non-zero dual elements

$$\begin{aligned} (f, g, \eta) \quad &\in \quad \left(W^{-1,q}(\Omega^0) \oplus W^{-1,q}(\Omega^2) \oplus L^q(\Omega^1)\right)^* \\ &= \quad W^{1,q^*}(\Omega^0) \oplus W^{1,q^*}(\Omega^2) \oplus L^{q^*}(\Omega^1) \end{aligned}$$

with

$$0 = (f, d_\alpha^* \nu), \qquad 0 = (g, d_\alpha \nu), \qquad 0 = (\eta, \nu + d_\alpha u + d_\alpha^* v)$$

for all (ν, u, v). The first two equations imply

$$0 = (d_\alpha f, \nu), \qquad 0 = (d_\alpha^* g, \nu)$$

for all ν. This implies $d_\alpha f = 0$ and $d^* g = 0$, and so $f = 0$ and $g = 0$ by Lemma 3.1.4. For the third equation, take $(u, v) = (0, 0)$ and we get $0 = (\eta, \nu)$ for all ν. But this can only happen if $\eta = 0$, which is a contradiction to the tuple (f, g, η) being non-zero. $\square$

We end this preparatory section by establishing the analogue of Lemma 3.1.4 for 1-forms.

Lemma 3.1.8. *Assume that $P \to \Sigma$ satisfies the conditions of Lemma 3.1.4, and let $1 < q < \infty$. Then there are constants $C > 0$ and $\epsilon_0 > 0$ such that*

$$\|\mu - \mathrm{proj}_{\alpha}\mu\|_{W^{1,q}(\Sigma)} \leq C\left(\|d_{\alpha}\mu\|_{L^q(\Sigma)} + \|d_{\alpha} * \mu\|_{L^q(\Sigma)}\right) \tag{3.28}$$

for all $\mu \in W^{1,q}(T^*\Sigma \otimes P(\mathfrak{g}))$ and all $\alpha \in \mathcal{A}^{1,q}(\Sigma)$ with $\|F_{\alpha}\|_{L^q(\Sigma)} < \epsilon_0$.

Proof. First note that this is just the standard elliptic regularity result if $\alpha = \alpha_{\flat}$ is flat. To prove the lemma, suppose the conclusion does not hold. Then there is sequence of connections α_{ν}, with curvature going to zero in L^q, and a sequence of 1-forms $\mu_{\nu} \in \mathrm{Im}\, d_{\alpha_{\nu}} \oplus \mathrm{Im}\, * d_{\alpha_{\nu}}$ with $\|\mu_{\nu}\|_{W^{1,q}} = 1$ and

$$\|d_{\alpha_{\nu}}\mu_{\nu}\|_{L^q} + \|d_{\alpha_{\nu}} * \mu_{\nu}\|_{L^q} \longrightarrow 0.$$

By applying suitable gauge transformations to the α_{ν}, and by passing to a subsequence, it follows from Uhlenbeck compactness that the α_{ν} converge strongly in L^{2q} to a limiting flat connection $\alpha_{\flat}$. So we have

$$\begin{aligned}
\|d_{\alpha_{\flat}}\mu_{\nu}\|_{L^q} &\leq \|d_{\alpha_{\nu}}\mu_{\nu}\|_{L^q} + \|\,[\alpha_{\nu} - \alpha_{\flat} \wedge \mu_{\nu}]\,\|_{L^q} \\
&\leq \|d_{\alpha_{\nu}}\mu_{\nu}\|_{L^q} + \|\alpha_{\nu} - \alpha_{\flat}\|_{L^{2q}}\|\mu_{\nu}\|_{L^{2q}} \\
&\leq \|d_{\alpha_{\nu}}\mu_{\nu}\|_{L^q} + C_0\|\alpha_{\nu} - \alpha_{\flat}\|_{L^{2q}}\|\mu_{\nu}\|_{W^{1,q}} \\
&\longrightarrow 0,
\end{aligned}$$

where in the last step we have used the embedding $W^{1,q} \hookrightarrow L^{2q}$ for $q > 1$. Similarly

$$\|d_{\alpha_{\flat}} * \mu_{\nu}\|_{L^q} \longrightarrow 0.$$

By the elliptic estimate for the flat connection $\alpha_{\flat}$, we have

$$\|\mu_\nu - \mathrm{proj}_{\alpha_\flat}\mu_\nu\|_{W^{1,q}} \leq C_1 \left(\|d_{\alpha_\flat}\mu_\nu\|_{L^q} + \|d_{\alpha_\flat} * \mu_\nu\|_{L^q}\right) \longrightarrow 0. \tag{3.29}$$

On the other hand, by Proposition 3.1.7, the projection operator $\mathrm{proj}_{\alpha_\nu}$ is converging in the L^q operator norm to $\mathrm{proj}_{\alpha_\flat}$. In particular,

$$\begin{array}{rcl} \|\mathrm{proj}_{\alpha_\flat}(\mu_\nu)\|_{W^{1,q}} & \leq & C_2\|\mathrm{proj}_{\alpha_\flat}(\mu_\nu)\|_{L^q} \\ & = & C_2\|\mathrm{proj}_{\alpha_\flat}(\mu_\nu) - \mathrm{proj}_{\alpha_\nu}(\mu_\nu)\|_{L^q} \\ \text{(Proposition 3.1.7)} & \leq & C_2\|\alpha_\flat - \alpha_\nu\|_{L^q}\|\mu_\nu\|_{L^q} \\ & \longrightarrow & 0, \end{array}$$

where the first inequality holds because $H^1_{\alpha_\flat}$ is finite-dimensional (and so all norms are equivalent) and the convergence to zero holds since $\|\mu_\nu\|_{L^q} \leq C_3\|\mu_\nu\|_{W^{1,q}} = C_3$ is bounded. Combining this with (3.29) gives

$$1 = \|\mu_\nu\|_{W^{1,q}} \leq \|\mu_\nu - \mathrm{proj}_{\alpha_\flat}\mu_\nu\|_{W^{1,q}} + \|\mathrm{proj}_{\alpha_\flat}(\mu_\nu)\|_{W^{1,q}} \longrightarrow 0,$$

which is a contradiction, proving the lemma. □

3.1.2 Analytic properties of NS

The next proposition will be used to obtain C^0 estimates for convergence of instantons to holomorphic curves. It provides a quantitative version of the statement that NS is approximately the identity map on connections with small curvature.

Proposition 3.1.9. *Let* NS_P *be the map (3.1), and* $3/2 \leq q < \infty$. *Then there are constants* $C > 0$ *and* $\epsilon_0 > 0$ *such that*

$$\|\mathrm{NS}_P(\alpha) - \alpha\|_{W^{1,q}(\Sigma)} \leq C\|F_\alpha\|_{L^{2q}(\Sigma)} \tag{3.30}$$

for all $\alpha \in \mathcal{A}^{1,q}(P)$ *with* $\|F_\alpha\|_{L^q(\Sigma)} < \epsilon_0$.

Proof. The basic idea is that $\mathrm{NS}_P(\alpha) = \exp(i\Xi(\alpha))^*\alpha$ may be expressed as a power series, with lowest order term given by α. The goal is then to bound the higher order terms using the curvature. To describe this precisely, we digress to discuss the power series expansion for the exponential.

As discussed above, the space $\mathrm{Lie}(\mathcal{G}(E)^{\mathbb{C}})^{2,q}$ can be viewed as the $W^{2,q}$-completion of the vector space $\Gamma(P \times_G \mathrm{End}(\mathbb{C}^n))$. Since $q > 1$, we are in the range in which pointwise matrix multiplication is well-defined, and this becomes a Banach algebra. Then for any $\xi \in \mathrm{Lie}(\mathcal{G}(E)^{\mathbb{C}})^{2,q}$, the power series

$$\sum_{k=0}^{\infty} \frac{\xi^k}{k!} \in \mathrm{Lie}(\mathcal{G}(E)^{\mathbb{C}})^{2,q}$$

converges, where ξ^k is k-fold matrix multiplication on the values of ξ.

Remark 3.1.10. *Note that we may not have $\xi^k \in \mathrm{Lie}(\mathcal{G}(P)^{\mathbb{C}})^{2,q}$, even when $\xi \in \mathrm{Lie}(\mathcal{G}(P)^{\mathbb{C}})^{2,q}$ (however, it is always the case that $\xi^k \in \mathrm{Lie}(\mathcal{G}(E)^{\mathbb{C}})^{2,q}$). In particular, the infinitesimal action of ξ^k on $\mathcal{A}^{1,q}(P)$ need not lie in the tangent space to $\mathcal{A}^{1,q}(P)$, though it will always lie in the tangent space to $\mathcal{A}^{1,q}(E)$.*

As with finite-dimensional Lie theory, this power series represents the exponential map

$$\exp : \mathrm{Lie}(\mathcal{G}(E)^{\mathbb{C}})^{2,q} \longrightarrow \mathcal{G}^{2,q}(E)^{\mathbb{C}},$$

where we are using the inclusion

$$\mathcal{G}^{2,q}(E)^{\mathbb{C}} \subset W^{2,q}(P \times_G \mathrm{End}(\mathbb{C}^n)).$$

The power series defining exp continues to hold on the restriction

$$\exp : \mathrm{Lie}(\mathcal{G}(P)^{\mathbb{C}})^{2,q} \longrightarrow \mathcal{G}^{2,q}(P)^{\mathbb{C}}.$$

Similarly, the usual power series definitions of sin and cos hold in this setting:

$$\sin,\ \cos : \mathrm{Lie}(\mathcal{G}(P)^{\mathbb{C}})^{2,q} \longrightarrow W^{2,q}(P \times_G \mathrm{End}(\mathbb{C}^n)),$$

and we have the familiar relation

$$\exp(i\xi) = \cos(\xi) + i\sin(\xi).$$

The infinitesimal action of $\mathrm{Lie}(\mathcal{G}(E)^{\mathbb{C}})^{2,q}$ on $\mathcal{A}^{1,q}(E)$ continues to have the form (3.8). In particular, for any real $\xi \in \mathrm{Lie}(\mathcal{G}(E))^{2,q}$ and $\alpha \in \mathcal{A}^{1,q}(P)$, we have

$$\begin{aligned} \exp(i\xi)^*\alpha - \alpha &= -\{d_\alpha(\cos(\xi) - 1) + *d_\alpha(\sin(\xi))\} \\ &\in T_\alpha\mathcal{A}^{1,q}(P) \subset W^{1,q}(T^*\Sigma \otimes P \times_G \mathrm{End}(\mathbb{C}^n)), \end{aligned} \tag{3.31}$$

where the action of d_α on each of these power series is defined term by term.

Now we prove the proposition. We will show that that there is some $\epsilon_0 > 0$ and $C > 0$ such that, if $\alpha \in \mathcal{A}^{1,q}(P)$ satisfies

$$\|F_\alpha\|_{L^q(\Sigma)} < \epsilon_0 \qquad \text{and} \qquad \|\Xi(\alpha)\|_{W^{2,q}} < \epsilon_0,$$

then

$$\|\mathrm{NS}_P(\alpha) - \alpha\|_{W^{1,q}} \leq C\|F_\alpha\|_{L^{2q}}.$$

This is exactly the statement of the proposition, except for the condition on $\Xi(\alpha)$. However, the proof of Theorem 3.1.1 shows that $\Xi(\alpha)$ depends continuously on $\alpha \in \mathcal{A}^{1,q}(P)$, and $\Xi(\alpha) = 0$ whenever α is flat. Hence, the condition on $\Xi(\alpha)$ is superfluous.

Set $\Xi = \Xi(\alpha)$ and $\eta = \mathrm{NS}_P(\alpha) - \alpha$. Using the power series expansion of exp, we have

$$\eta = \exp(i\Xi)^*\alpha - \alpha = - * d_{\alpha,\rho}\Xi + \frac{(\Xi(d_{\alpha,\rho}\Xi) + (d_{\alpha,\rho}\Xi)\Xi)}{2} + \ldots$$

where the nth term in the sum on the right has the form

$$-\frac{*^n}{n!}\sum_{k=0}^{n} \Xi \ldots \Xi(d_{\alpha,\rho}\Xi)\Xi \ldots \Xi$$

with k copies of Ξ appearing before $d_{\alpha,\rho}\Xi$, and $n - k - 1$ copies after. By assumption $2q > 2$, and so the Sobolev multiplication theorem gives

$$\left\| \frac{*^n}{n!} \sum_{k=0}^{n} \Xi \ldots \Xi (d_{\alpha,\rho}\Xi) \Xi \ldots \Xi \right\|_{W^{1,2q}} \leq \|d_{\alpha,\rho}\Xi\|_{W^{1,2q}} \left(\frac{C_1^{2n}}{n!} \sum_{k=0}^{n} \|\Xi\|_{W^{1,2q}}^{n-1} \right)$$

$$\leq \|\Xi\|_{W^{2,2q}} \left(\frac{C_1^{2n}}{(n-1)!} \|\Xi\|_{W^{1,2q}}^{n-1} \right),$$

where C_1 is the constant from the Sobolev multiplication theorem. This gives

$$\begin{aligned} \|\eta\|_{W^{1,q}} &\leq C_2 \|\eta\|_{W^{1,2q}} \\ &\leq C_2 \|\Xi\|_{W^{2,2q}} \sum_{n=1}^{\infty} \frac{C_1^{2n}}{(n-1)!} \|\Xi\|_{W^{1,2q}}^{n-1} \\ &\leq C_2 \|\Xi\|_{W^{2,2q}} \sum_{n=1}^{\infty} \frac{C_1^{2n}}{(n-1)!} \\ &= C_3 \|\Xi\|_{W^{2,2q}} \end{aligned} \tag{3.32}$$

where the third inequality holds for $\|\Xi\|_{W^{1,2q}} \leq 1$. It suffices to estimate $\|\Xi\|_{W^{2,2q}}$ in terms of η and F_α.

By (3.31) and the definition of NS_P we have

$$\begin{aligned} d_{\alpha,\rho}\eta &= d_{\alpha,\rho} \left(\exp(i\Xi)^* \alpha - \alpha \right) \\ &= -d_{\alpha,\rho} \left(* d_{\alpha,\rho}(\sin(\Xi)) + d_{\alpha,\rho}(\cos(\Xi) - 1) \right) \\ &= -d_{\alpha,\rho} * d_{\alpha,\rho}(\sin(\Xi)) + F_{\alpha,\rho}(1 - \cos(\Xi)) \end{aligned} \tag{3.33}$$

Now use the elliptic estimate from Lemma 3.1.4 (ii):

$$\begin{aligned} \|\sin(\Xi)\|_{W^{2,2q}} &\leq C_4 \|d_{\alpha,\rho} * d_{\alpha,\rho}(\sin(\Xi))\|_{L^{2q}} \\ &\leq C_5 \left\{ \|d_{\alpha,\rho}\eta\|_{L^{2q}} + \|F_{\alpha,\rho}(1 - \cos(\Xi))\|_{L^{2q}} \right\} \\ &\leq C_6 \left\{ \|\eta\|_{L^{4q}}^2 + \|F_{\alpha,\rho}\|_{L^{2q}} \left(1 + \|1 - \cos(\Xi)\|_{L^\infty} \right) \right\} \end{aligned}$$

where the second inequality is (3.33), and in the last inequality we used

$$\|d_{\alpha,\rho}\eta\|_{L^{2q}} \leq C\left(\|F_{\alpha,\rho}\|_{L^{2p}} + \|\eta\|_{L^{4q}}^2\right)$$

coming from

$$0 = F_{\mathrm{NS}_P(\alpha),\rho} = F_{\alpha,\rho} + d_{\alpha,\rho}\eta + \frac{1}{2}\left[\eta \wedge \eta\right].$$

Note also that the norm of $F_{\alpha,\rho}$ is controlled by that of F_α, so we can drop the subscript ρ by picking up another constant:

$$\|\sin(\Xi)\|_{W^{2,2q}} \leq C_7\left\{\|\eta\|_{L^{4q}}^2 + \|F_\alpha\|_{L^{2q}}\left(1 + \|1-\cos(\Xi)\|_{L^\infty}\right)\right\} \tag{3.34}$$

For $\|\Xi\|_{W^{2,q}}$ sufficiently small we have

$$\|\Xi\|_{W^{2,q}} \leq 2\|\sin(\Xi)\|_{W^{2,q}}, \qquad \|1-\cos(\Xi)\|_{L^\infty} \leq 1.$$

So (3.34) gives

$$\|\Xi\|_{W^{2,2q}} \leq C_8\left(\|\eta\|_{L^{4q}}^2 + \|F_\alpha\|_{L^{2q}}\right).$$

Returning to (3.32), we conclude

$$\begin{aligned} \|\eta\|_{W^{1,q}} &\leq C_9\left(\|\eta\|_{L^{4q}}^2 + \|F_\alpha\|_{L^{2q}}\right) \\ &\leq C_{10}\left(\|\eta\|_{W^{1,q}}^2 + \|F_\alpha\|_{L^{2q}}\right) \end{aligned}$$

where we have used the embedding $W^{1,q} \hookrightarrow L^{4q}$, which holds provided $q \geq 3/2$. This gives

$$\|\eta\|_{W^{1,q}}\left(1 - C_{10}\|\eta\|_{W^{1,q}}\right) \leq C_{10}\|F_\alpha\|_{L^{2q}},$$

which completes the proof since we can ensure that $\|\eta\|_{W^{1,q}} \leq 1/2C_{10}$ by requiring that $\|\Xi\|_{W^{2,q}}$ is sufficiently small (when $\Xi = 0$, it follows that $\eta = 0$, and everything is continuous in these norms). $\square$

Let $\Pi : \mathcal{A}^{1,q}_{\text{flat}}(P) \to M(P)$ denote the quotient map. Throughout the remainder of this section we will be interested in the derivative of the composition $\Pi \circ \mathrm{NS}_P$. We will therefore assume that Σ is closed, connected and orientable, $G = \mathrm{PSU}(r)$ and $P \to \Sigma$ is a bundle for which $t_2(P) \in \mathbb{Z}_r$ is a generator. This ensures that $M(P)$ and Π are both smooth.

Recall that any choice of orientation and metric on Σ determines complex structures on the tangent bundles $T\mathcal{A}^{1,q}(P)$ and $TM(P)$, which is induced by the Hodge star on 1-forms. Denote by

$$D^k_\alpha \left(\Pi \circ \mathrm{NS}_P\right) : T_\alpha \mathcal{A}^{1,q}(P) \otimes \ldots \otimes T_\alpha \mathcal{A}^{1,q}(P) \longrightarrow T_{\Pi \circ \mathrm{NS}_P(\alpha)} M(P)$$

the kth derivative of $\Pi \circ \mathrm{NS}_P$ at α, defined with respect to the $W^{1,q}$-topology on the domain. The following lemma will be used to show that holomorphic curves in $\mathcal{A}^{1,q}(P)$ descend to holomorphic curves in $M(P)$.

Lemma 3.1.11. *Suppose $G = \mathrm{PSU}(r)$, Σ is a closed connected oriented Riemannian surface and $P \to \Sigma$ is a principal G-bundle with $t_2(P) \in \mathbb{Z}_r$ a generator. Let $1 < q < \infty$ and suppose α is in the domain of NS_P. Then the linearization $D_\alpha(\Pi \circ \mathrm{NS}_P)$ is complex-linear:*

$$*D_\alpha(\Pi \circ \mathrm{NS}_P) = D_\alpha(\Pi \circ \mathrm{NS}_P) * .$$

Proof. The complex gauge group $\mathcal{G}(P)^{\mathbb{C}}$ acts on $\mathcal{C}(P)$, and hence $\mathcal{A}(P)$, in a way that preserves the complex structure, and this holds true in the Sobolev completions of these spaces. Indeed, let $u \in \mathcal{G}^{2,q}(P)^{\mathbb{C}}$, $\alpha \in \mathcal{A}^{1,q}(P)$ and $\eta \in W^{1,q}\left(T^*\Sigma \otimes P(\mathfrak{g})\right)$. Then by (3.7) we have

$$\begin{aligned}
\left.\tfrac{d}{d\tau}\right|_{\tau=0} u \circ \overline{\partial}_{\alpha + \tau * \eta} \circ u^{-1} &= \left.\tfrac{d}{d\tau}\right|_{\tau=0} u \circ \overline{\partial}_\alpha \circ u^{-1} + \tau u \circ \left(*\eta^{0,1}\right) \circ u^{-1} \\
&= *\left(u \circ \eta^{0,1} \circ u^{-1}\right) \\
&= * \left.\tfrac{d}{d\tau}\right|_{\tau=0} u \circ \overline{\partial}_{\alpha + \tau\eta} \circ u^{-1},
\end{aligned}$$

which shows the infinitesimal action of the complex gauge group is complex-linear.

Let $\mathcal{G}_0^{2,q}(P)^{\mathbb{C}} \subseteq \mathcal{G}^{2,q}(P)^{\mathbb{C}}$ denote the identity component. This can be described as

$$\mathcal{G}_0^{2,q}(P)^{\mathbb{C}} = \left\{ u \exp(i\xi) \middle| u \in \mathcal{G}_0^{2,q}(P), \qquad \xi \in W^{2,q}(P(\mathfrak{g})) \right\}.$$

It follows from (3.8) and Lemma 3.1.4 that $\mathcal{G}_0^{\mathbb{C}}(P)^{2,q}$ acts freely on the space of connections. Moreover, by Remark 3.1.3, the map NS_P is equivariant under a neighborhood of $\mathcal{G}_0^{2,q}(P)$ in $\mathcal{G}_0^{2,q}(P)^{\mathbb{C}}$. These two facts imply that NS_P has a unique $\mathcal{G}_0^{2,q}(P)^{\mathbb{C}}$-equivariant extension to the flow-out

$$\mathcal{A}^{ss}(P) := \left(\mathcal{G}_0^{2,q}(P)^{\mathbb{C}}\right)^* \left\{ \alpha \in \mathcal{A}^{1,q}(P) | \, \|F_\alpha\|_{L^q} < \epsilon \right\}$$

of the domain of NS_P. Furthermore, the group $\mathcal{G}_0^{2,q}(P)^{\mathbb{C}}$ restricts to a free action on $\mathcal{A}^{ss}(P)$.

Consider the projection

$$\Pi^{\mathbb{C}} : \mathcal{A}^{ss}(P) \longrightarrow \mathcal{A}^{ss}(P)/\mathcal{G}_0^{2,q}(P)^{\mathbb{C}}.$$

Using NS_P, we have an identification

$$\mathcal{A}^{ss}(P)/\mathcal{G}_0^{2,q}(P)^{\mathbb{C}} \cong M(P),$$

and hence a commutative diagram

$$\begin{tikzcd}
\mathcal{A}^{ss}(P) \arrow[r, "\mathrm{NS}_P"] \arrow[d, "\Pi^{\mathbb{C}}"'] & \mathcal{A}^{1,q}_{\text{flat}}(P) \arrow[d, "\Pi"] \\
\mathcal{A}^{ss}(P)/\mathcal{G}_0^{2,q}(P)^{\mathbb{C}} \arrow[r, "\cong"] & \mathcal{A}^{1,q}_{\text{flat}}(P)/\mathcal{G}_0^{2,q}(P) = M(P)
\end{tikzcd}$$

As we saw above, the infinitesimal action of $\mathcal{G}_0^{2,q}(P)$ is complex linear. This implies that

$$\Pi^{\mathbb{C}} : \mathcal{A}^{ss}(P) \longrightarrow M(P)$$

is complex-linear, but $\Pi^{\mathbb{C}} = \Pi \circ \mathrm{NS}_P$, so this finishes the proof. $\square$

Lemma 3.1.12. *Let $P \to \Sigma$ be as in the statement of Lemma 3.1.11 and $1 < q < \infty$. Assume α is in the domain of NS_P. Then the space $\mathrm{Im}\,(d_\alpha) \oplus \mathrm{Im}\,(d_\alpha^*)$ lies in the kernel of $D^k_\alpha(\Pi \circ \mathrm{NS}_P)$ in the sense that*

$$D^k_\alpha(\Pi \circ \mathrm{NS}_P)\,(d_\alpha \xi + *d_\alpha \zeta,\ \cdot\ , \ldots,\ \cdot\ ) = 0$$

for all 0-forms $\xi, \zeta \in W^{2,q}(P(\mathfrak{g}))$. Moreover, there is an estimate

$$|D^k_\alpha(\Pi \circ \mathrm{NS}_P)\,(\mu_1, \ldots, \mu_k)\,|_{M(P)} \leq C \|\mu_1\|_{L^q(\Sigma)} \cdots \|\mu_k\|_{L^q(\Sigma)} \tag{3.35}$$

for all tuples $\mu_1, \ldots, \mu_k \in W^{1,q}(T^*\Sigma \otimes P(\mathfrak{g}))$ of 1-forms. Here $|\cdot|_{M(P)}$ is any norm on $TM(P)$.

Proof. We prove the lemma for $k = 1$. The cases for larger k are similar. By the proof of Lemma 3.1.11, the map $\Pi \circ \mathrm{NS}_P$ is invariant under gauge transformations of the form $\exp(\xi + i\zeta)$ where $\xi, \zeta \in W^{2,q}(P(\mathfrak{g}))$ are real. In particular,

$$0 = \left.\frac{d}{d\tau}\right|_{\tau=0} \Pi \circ \mathrm{NS}_P(\exp(\tau(\xi + i\zeta))^*\alpha) = -D_\alpha(\Pi \circ \mathrm{NS}_P)(d_\alpha \xi + *d_\alpha \zeta).$$

This proves the first assertion.

To prove the estimate (3.35), we note that by Lemma 3.1.5 there is a decomposition

$$T_\alpha \mathcal{A}^{1,q}(P) = H^1_\alpha \oplus (\mathrm{Im}\, d_\alpha \oplus \mathrm{Im}\, d^*_\alpha)\,,$$

whenever α has sufficiently small curvature. Moreover, the first summand is L^2-orthogonal. Denote by

$$\mathrm{proj}_\alpha : T_\alpha \mathcal{A}^{1,q}(P) \longrightarrow H^1_\alpha$$

the projection to the d_α-harmonic space, and note that this is continuous with respect to the L^q-norm on the domain and codomain (projections are always continuous). We claim that the operator

$$D_\alpha\,(\Pi \circ \mathrm{NS}_P) : T_\alpha \mathcal{A}^{1,q} \longrightarrow H_{\mathrm{NS}_P(\alpha)}$$

can be written as a composition

$$T_\alpha \mathcal{A}^{1,q} \longrightarrow H_\alpha \xrightarrow{M_\alpha} H_{\mathrm{NS}_P(\alpha)}$$

for some bounded linear map M_α, where the first map is proj_α. Indeed, by the first part of the lemma it follows that

$$D_\alpha \left(\Pi \circ \mathrm{NS}_P\right)(\mu) = D_\alpha \left(\Pi \circ \mathrm{NS}_P\right)(\mathrm{proj}_\alpha \mu)$$

since the difference $\mu - \mathrm{proj}_\alpha \mu$ lies in $\mathrm{Im}\, d_\alpha \oplus \mathrm{Im}\, * d_\alpha$. So the claim follows by taking

$$M_\alpha := D_\alpha \left(\Pi \circ \mathrm{NS}_P\right)|_{H_\alpha}$$

to be the restriction. Since M_α is a linear map between finite-dimensional spaces, it is bounded with respect to any norm. We take the L^q-norm on these harmonic spaces. Then $D_\alpha \left(\Pi \circ \mathrm{NS}_P\right)$ is the composition of two functions which are continuous with respect to the L^q norm:

$$\begin{aligned} \left|D_\alpha \left(\Pi \circ \mathrm{NS}_P\right) \mu\right|_{M(P)} &= C \|D_\alpha \left(\Pi \circ \mathrm{NS}_P\right) \mu\|_{L^q} \\ &= C \|M_\alpha \circ \mathrm{proj}_\alpha \mu\|_{L^q} \\ &\leq C_\alpha \|\mu\|_{L^q}. \end{aligned}$$

That this constant can be taken independent of α, for F_α sufficiently small, follows using an Uhlenbeck compactness argument similar to the one carried out at the beginning of the proof of Theorem 3.1.1. Here one needs to use the fact that $D_\alpha(\Pi \circ \mathrm{NS}_P) = \mathrm{proj}_\alpha$ when α is a flat connection, and so this has norm 1 (which is clearly independent of α). Similarly, one can show that the operator norm

$$\|D_\alpha \left(\Pi \circ \mathrm{NS}_P\right)\|_{\mathrm{op}, L^q}$$

(defined using the L^q-topology on the domain) □

Corollary 3.1.13. *Suppose* $1 < q < \infty$*, and let* $P \to \Sigma$ *be as in the statement of Lemma 3.1.11. Then there is a constant* $\epsilon_0 > 0$ *and a bounded function* $f : \mathcal{A}^{0,q}(P) \to \mathbb{R}^{\geq 0}$ *such that for each* $\alpha \in \mathcal{A}^{1,q}(P)$ *with* $\|F_\alpha\|_{L^{2q}(\Sigma)} < \epsilon_0$*, the following estimate holds*

$$\|\mathrm{proj}_\alpha \mu - D_\alpha \left(\Pi \circ \mathrm{NS}_P\right) \mu\|_{L^q(\Sigma)} \leq f\left(\alpha\right) \|\mathrm{proj}_\alpha \mu\|_{L^q(\Sigma)} \tag{3.36}$$

for all $\mu \in L^q(T^*\Sigma \otimes P(\mathfrak{g}))$*, where* proj_α *is the map (3.22). Furthermore,* f *can chosen so that* $f(\alpha) \to 0$ *as* $\|F_\alpha\|_{L^q(\Sigma)} \to 0$*.*

Proof. Consider the operator

$$\mathrm{proj}_\alpha - D_\alpha \left(\Pi \circ \mathrm{NS}_P\right).$$

It is clear from Lemma 3.1.12 that its kernel contains $\mathrm{Im}(d_\alpha) \oplus \mathrm{Im}(*d_\alpha)$, and so we have[1]

$$\begin{aligned}
\|\mathrm{proj}_\alpha \mu - D_\alpha \left(\Pi \circ \mathrm{NS}_P\right) \mu\|_{L^q} &\leq C_1 \|\mathrm{proj}_\alpha \mu - D_\alpha \left(\Pi \circ \mathrm{NS}_P\right) \mu\|_{L^{2q}} \\
&= C_1 \| \left(\mathrm{proj}_\alpha - D_\alpha (\Pi \circ \mathrm{NS}_P)\right) \left(\mathrm{proj}_\alpha \mu\right) \|_{L^{2q}} \\
&\leq C_1 \left\|\left(\mathrm{proj}_\alpha - D_\alpha \left(\Pi \circ \mathrm{NS}_P\right)\right)\right\|_{\mathrm{op}, L^{2q}} \ \|\mathrm{proj}_\alpha \mu\|_{L^{2q}}
\end{aligned}$$

On the finite-dimensional space H_α, the L^q- and L^{2q}-norms are equivalent:

$$\|\mathrm{proj}_\alpha \mu\|_{L^{2q}} \leq C_2 \|\mathrm{proj}_\alpha \mu\|_{L^q}.$$

The constant C_2 is independent of $\mathrm{proj}_\alpha \mu \in H_\alpha$, however it may depend on α. Proposition 3.1.7 tells us that C_2 is independent of α provided $\|F_\alpha\|_{L^{2q}}$ is sufficiently small. So we have

$$\|\mathrm{proj}_\alpha \mu - D_\alpha \left(\Pi \circ \mathrm{NS}_P\right) \mu\|_{L^q} \leq f(\alpha) \|\mathrm{proj}_\alpha \mu\|_{L^q},$$

[1] In the first line, we replace the L^q norms by L^{2q} norms to ensure that we are in the range to use Theorem 3.1.1 (i.e., since $p := 2q > 2$). If we know that $q > 2$ then this is not necessary, and the proof simplifies a little. However, in our applications below we will need the case $q = 2$.

where we have set

$$f(\alpha) := C_1 C_2 \left\| (\mathrm{proj}_\alpha - D_\alpha (\Pi \circ \mathrm{NS}_P)) \right\|_{\mathrm{op}, L^{2q}}.$$

By Theorem 3.1.1 and Proposition 3.1.7, the function $f(\alpha)$ depends continuously on α in the L^{2q}-topology. If $\alpha = \alpha_\flat$ is flat, then $D_\alpha (\Pi \circ \mathrm{NS}_P)$ equals the projection proj_α, and so $f(\alpha_\flat) = 0$. In particular, $f(\alpha) \to 0$ as α approaches $\mathcal{A}^{1,q}_{\mathrm{flat}}(P)$ in the L^{2q}-topology. That $f(\alpha) \to 0$ as $\|F_\alpha\|_{L^q} \to 0$ follows from a contradiction argument using weak Uhlenbeck compactness (Theorem 2.2.9), just as we did at the beginning of the proof of Theorem 3.1.1. $\square$

3.2 Heat flow on cobordisms

Suppose Q is principal G-bundle over a Riemannian manifold Y of dimension 3. In his thesis [34], Råde studied the **Yang-Mills heat flow**; that is, the solution $\tau \mapsto a_\tau \in \mathcal{A}(Q)$ to the gradient flow of the Yang-Mills functional

$$\frac{d}{d\tau} a_\tau = -d^*_{a_\tau} F_{a_\tau}, \qquad a_0 = a, \tag{3.37}$$

for some fixed initial condition $a \in \mathcal{A}(Q)$. Specifically, Råde proved the following:

Theorem 3.2.1. *Suppose G is compact and Y is a closed orientable manifold of dimension 3. Let $a \in \mathcal{A}^{1,2}(Q)$. Then the equation (3.37) has a unique solution $\{\tau \mapsto a_\tau\} \in C^0_{loc}([0,\infty), \mathcal{A}^{1,2}(Q))$, with the further property that*

$$F_{a_\tau} \in C^0_{loc}\left([0,\infty), L^2\right) \cap L^2_{loc}\left([0,\infty), W^{1,2}\right).$$

Furthermore, the limit $\lim_{\tau \to \infty} a_\tau$ exists, is a critical point of the Yang-Mills functional, and varies continuously with the initial data a in the $W^{1,2}$-topology.

Differentiating $\mathcal{YM}_Q(a_\tau)$ in τ and using (3.37) shows that $\mathcal{YM}_Q(a_\tau)$ decreases in τ. Moreover, it follows from Uhlenbeck's Compactness Theorem 2.2.10 together with, say, [34, Proposition 7.2] that the critical values of the Yang-Mills functional are discrete. Combining these two facts, it follows that there is some $\widetilde{\epsilon}_Q > 0$ such that if $\mathcal{YM}_Q(a) < \widetilde{\epsilon}_Q$, then the associated limiting connection

$$\lim_{\tau} a_\tau \in \mathcal{A}^{1,2}_{\mathrm{flat}}(Q)$$

is flat, where a_τ satisfies (3.37). This therefore defines a continuous gauge equivariant deformation retract

$$\mathrm{Heat}_Q : \left\{ a \in \mathcal{A}^{1,2}(Q) \,|\, \mathcal{YM}_Q(a) < \widetilde{\epsilon}_Q \right\} \longrightarrow \mathcal{A}^{1,2}_{\mathrm{flat}}(Q) \tag{3.38}$$

whenever Y is a closed 3-manifold.

Remark 3.2.2. *Råde's theorem continues to hold, exactly as stated, in dimension 2 as well. Given a bundle $P \to \Sigma$ over a closed connected oriented surface, we therefore have that* NS_P *and* Heat_P *are both maps of the form*

$$\left\{ \alpha \in \mathcal{A}^{1,2}(P) \,|\, \mathcal{YM}_P(\alpha) < \epsilon_P \right\} \longrightarrow \mathcal{A}^{1,2}_{\mathrm{flat}}(P),$$

for a suitably small $\epsilon_P > 0$. It turns out these are the same map, up to a gauge transformation. That is,

$$\Pi \circ \mathrm{NS}_P = \Pi \circ \mathrm{Heat}_P, \tag{3.39}$$

where $\Pi : \mathcal{A}^{1,2}_{\mathrm{flat}}(P) \to \mathcal{A}^{1,2}_{\mathrm{flat}}(P)/\mathcal{G}^{2,2}_0(P)$ is the quotient map. Though we will not use this fact in this thesis, we sketch a proof at the end of this section for completeness.

In the remainder of this section we prove a version of Råde's Theorem 3.2.1, but for bundles Q over 3-manifolds *with boundary.* The most natural boundary condition for our application is of Neumann type. This will allow us to use a reflection principle and thereby appeal directly to Råde's result for closed 3-manifolds.

Råde's result holds with the $W^{1,2}$-topology. However, on 3-manifolds $W^{1,2}$-functions (forms, connections, etc.) are not all continuous, which makes the issue of boundary conditions rather tricky. One way to get around this is to observe that, in dimension 3, restricting $W^{1,2}$-functions to codimension-1 subspaces is in fact well-defined. We take an equivalent approach by considering the space $\mathcal{A}^{1,2}(Q, \partial Q)$, which we define to be the $W^{1,2}$-closure of the set of smooth $a \in \mathcal{A}(Q)$ which satisfy

$$\iota_{\partial_n} a|_U = 0 \tag{3.40}$$

on some neighborhood U of ∂Q (U may depend on a). Here we have fixed an extension ∂_n of the outward pointing unit normal to ∂Q, and we may assume that the set U is always contained in the region in which ∂_n is non-zero. This can be described more explicitly as follows: Use the normalized gradient flow of ∂_n to write $U = [0, \epsilon) \times \partial Y$. Let t denote the coordinate on $[0, \epsilon)$. Then in these coordinates we can write $a|_{\{t\} \times \partial Y} = \alpha(t) + \psi(t)\, dt$. The condition (3.40) is equivalent to requiring $\psi(t) = 0$.

Set

$$\mathcal{A}^{1,2}_{\text{flat}}(Q, \partial Q) := \mathcal{A}^{1,2}(Q, \partial Q) \cap \mathcal{A}^{1,2}_{\text{flat}}(Q).$$

Both of the spaces $\mathcal{A}^{1,2}(Q, \partial Q)$ and $\mathcal{A}^{1,2}_{\text{flat}}(Q, \partial Q)$ admit the action of the subgroup $\mathcal{G}(Q, \partial Q) \subset \mathcal{G}(Q)$ consisting of gauge transformations that restrict to the identity in a neighborhood of ∂Q. (We are purposefully only working with the *smooth* gauge transformations here.)

Theorem 3.2.3. *Let G be a compact connected Lie group, and $Q \to Y$ be a principal G-bundle over a compact connected oriented Riemannian 3-dimensional manifold Y with boundary.*

1. There is some $\epsilon_Q > 0$ and a continuous strong deformation retract

$$\text{Heat}_Q : \left\{ a \in \mathcal{A}^{1,2}(Q, \partial Q) \,|\, \mathcal{YM}_Q(a) < \epsilon_Q \right\} \longrightarrow \mathcal{A}^{1,2}_{\text{flat}}(Q, \partial Q).$$

Furthermore, Heat_Q *intertwines the action of $\mathcal{G}(Q, \partial Q)$.*

2. Suppose $\Sigma \subset Y$ is an embedded surface which is closed and oriented. Suppose further that either $\Sigma \subset \operatorname{int} Y$, or $\Sigma \subset \partial Y$. Then for every $\epsilon > 0$, there is some $\delta > 0$ such that if $a \in \mathcal{A}^{1,2}(Q, \partial Q)$ satisfies $\|F_a\|_{L^2(Y)} < \delta$, then

$$\|(\mathrm{Heat}_Q(a) - a)\,|_\Sigma\|_{L^q(\Sigma)} < \epsilon$$

for any $1 \leq q \leq 4$.

Remark 3.2.4. *Recently, Charalambous [6] has proven similar results for manifolds with boundary.*

Proof. Consider the **double** of Y

$$Y^{(2)} := \overline{Y} \cup_{\partial Y} Y,$$

which is a closed 3-manifold, and denote by $\iota_Y : Y \hookrightarrow Y^{(2)}$ the inclusion to the second factor. We will identify Y with its image under ι_Y. There is a natural involution $\sigma : Y^{(2)} \to Y^{(2)}$ defined by switching the factors in the obvious way. Then $Y^{(2)}$ has a natural smooth structure making ι_Y smooth and σ a diffeomorphism (this is just the smooth structure obtained by choosing *the same* collar on each side of ∂Y). Clearly the map σ is orientation reversing, satisfies $\sigma^2 = \mathrm{Id}$ and has fixed point set equal to ∂Y. Similarly, we can form $Q^{(2)} := \overline{Q} \cup_{\partial Q} Q$ and an involution $\widetilde{\sigma} : Q^{(2)} \to Q^{(2)}$. Then $Q^{(2)}$ is naturally a principal G-bundle over $Y^{(2)}$ and $\widetilde{\sigma}$ is a bundle map covering σ. Furthermore, $\widetilde{\sigma}$ commutes with the G-action on $Q^{(2)}$.

Though $\widetilde{\sigma}$ is not a gauge transformation (it does not cover the identity), it behaves as one in many ways. For example, since it $\widetilde{\sigma}$ a bundle map, the space of connections $\mathcal{A}(Q^{(2)})$ is invariant under pullback by $\widetilde{\sigma}$. The action on covariant derivatives takes the form

$$d_{\widetilde{\sigma}^* a} = \sigma^* \circ d_a \circ \sigma^*, \tag{3.41}$$

where $\sigma^* : \Omega(Y^{(2)}, Q^{(2)}(\mathfrak{g})) \to \Omega(Y^{(2)}, Q^{(2)}(\mathfrak{g}))$ is pullback by σ, and this symbol $\circ$ denotes composition of operators. The induced action on the tangent space $T_a\mathcal{A}(Q^{(2)}) = \Omega^1(Y^{(2)}, Q^{(2)}(\mathfrak{g}))$ is just given by pullback by σ. Likewise, the curvature satisfies

$$F_{\widetilde{\sigma}^* a} = \sigma^* F_a. \tag{3.42}$$

In particular, the flow equation (3.37) on the double $Y^{(2)}$ is invariant under the action of $\widetilde{\sigma}$. We set $\epsilon_Q := \widetilde{\epsilon}_{Q^{(2)}}/2$, where $\widetilde{\epsilon}_{Q^{(2)}} > 0$ is as in (3.38).

Now suppose $a \in \mathcal{A}^{1,2}(Q, \partial Q)$ has $\mathcal{YM}_Q(a) < \epsilon_Q$. Then a has a unique extension $a^{(2)}$ to all of $Q^{(2)}$, satisfying $\widetilde{\sigma}^* a^{(2)} = a^{(2)}$. We claim that $a^{(2)} \in \mathcal{A}^{1,2}(Q^{(2)})$. To see this, first suppose that a is smooth. Then the boundary condition on a implies that $a^{(2)}$ is continuous on all of $Q^{(2)}$ and smooth on the complement of ∂Q. In particular, $a^{(2)}$ is $W^{1,2}$. (Note that in general $a^{(2)}$ will not be smooth, even if a is. For example, the normal *derivatives* on each side of the boundary do not agree: $\lim_{y \to \partial Y} \partial_n a = - \lim_{y \to \overline{Y}} \partial_n \widetilde{\sigma}^* a$, unless they are both zero, and this latter condition is not imposed by our boundary conditions.) More generally, every $a \in \mathcal{A}^{1,2}(Q, \partial Q)$ is the $W^{1,2}$ limit of smooth functions a_j whose normal component vanishes in a neighborhood of the boundary. But then by the linearity of the integral it is immediate that the $a_j^{(2)}$ approach $a^{(2)}$ in $W^{1,2}$, which proves the claim.

By assumption, we have

$$\mathcal{YM}_{Q^{(2)}}(a^{(2)}) = 2\mathcal{YM}_Q(a) < 2\epsilon_Q = \widetilde{\epsilon}_{Q^{(2)}},$$

so by the discussion at the beginning of this section, there is a unique solution $a_\tau^{(2)}$ to the flow equation (3.37) on the *closed* 2-manifold $Y^{(2)}$, with initial condition $a_0^{(2)} = a^{(2)}$. Furthermore, the limit $\mathrm{Heat}_{Q^{(2)}}(a^{(2)}) := \lim_{\tau \to \infty} a_\tau^{(2)}$ exists and is flat. Since (3.37) is $\widetilde{\sigma}$-invariant, the uniqueness assertion guarantees that $\widetilde{\sigma}^* a_\tau^{(2)} = a_\tau^{(2)}$ for all τ. In particular,

$$\widetilde{\sigma}^* \mathrm{Heat}_{Q^{(2)}}(a^{(2)}) = \mathrm{Heat}_{Q^{(2)}}(a^{(2)}). \tag{3.43}$$

Define

$$\mathrm{Heat}_Q(a) := \mathrm{Heat}_{Q^{(2)}}(a^{(2)})\Big|_Q .$$

Then (3.43) shows that $\iota_{\partial_n} \mathrm{Heat}_Q(a)|_{\partial Y} = 0$, so Heat_Q does map into $\mathcal{A}^{1,2}_{\text{flat}}(Q, \partial Q)$.

We already know that $\mathrm{Heat}_{Q^{(2)}}(a^{(2)})$ is $\mathcal{G}(Q^{(2)})$-equivariant. Each element $g \in \mathcal{G}(Q, \partial Q)$ has a unique extension to a $\widetilde{\sigma}$-invariant gauge transformation $g^{(2)} \in \mathcal{G}(Q^{(2)})$.

This shows

$$\begin{aligned}
\mathrm{Heat}_Q(g^*a) &= \mathrm{Heat}_{Q^{(2)}}(g^{(2),*}a^{(2)})\Big|_Q \\
&= g^{(2),*}\mathrm{Heat}_{Q^{(2)}}(a^{(2)})\Big|_Q \\
&= g^*\mathrm{Heat}_Q(a)
\end{aligned}$$

which finishes the proof of 1.

To prove 2, we will assume $\Sigma \subset \mathrm{int}\, Y$. This is a local problem bounded away from ∂Y, so the boundary will not effect our analysis. The remaining case $\Sigma \subset \partial Y$ follows by replacing Y with its double, for then we have $\Sigma \subset \mathrm{int}\, Y^{(2)}$ and the analysis carries over directly.

For sake of contradiction, suppose there is some sequence $a_\nu \in \mathcal{A}^{1,2}(Q)$ with

$$\|F_{a_\nu}\|_{L^2} \longrightarrow 0,$$

but

$$c_0 \leq \|(\mathrm{Heat}_Q(a_\nu) - a_\nu)\,|_\Sigma\|_{L^q(\Sigma)} \tag{3.44}$$

for some fixed $c_0 > 0$. By Uhlenbeck's Weak Compactness Theorem 2.2.9, there is a sequence of gauge transformations $u_\nu \in \mathcal{G}^{2,2}$ such that $u_\nu^* a_\nu$ converges weakly in $W^{1,2}$ (hence strongly in L^4) to a limiting connection $a_\infty \in \mathcal{A}^{1,2}(Q)$, after possibly passing to a subsequence. Then a_∞ is necessarily flat. Be redefining u_ν, if necessary, we may assume that each $u_\nu^* a_\nu$ is in Coulomb gauge with respect to a_∞, and still retain the fact that $u_\nu^* a_\nu$ converges to a_∞ strongly in L^4. Then

$$\begin{aligned}
\|u_\nu^* a_\nu - a_\infty\|_{W^{1,2}}^2 &= \|u_\nu^* a_\nu - a_\infty\|_{L^2}^2 + \|d_{a_\infty}(u_\nu^* a_\nu - a_\infty)\|_{L^2}^2 \\
&\leq C_1\left(\|u_\nu^* a_\nu - a_\infty\|_{L^2}^2 + \|F_{a_\nu}\|_{L^2}^2 + \|u_\nu^* a_\nu - a_\infty\|_{L^4}^4\right)
\end{aligned}$$

for some constant C_1. Observe that the right-hand side is going to zero, so a_ν is converging in $W^{1,2}$ to the space of flat connections:

$$\|a_\nu - (u_\nu^{-1})^* a_\infty\|_{W^{1,2}} \longrightarrow 0. \tag{3.45}$$

On the other hand, by the trace theorem [51, Theorem B.10], we have

$$c_0 \leq \|\mathrm{Heat}_Q(a_\nu) - a_\nu|_\Sigma\|_{L^q(\Sigma)} \quad \leq \quad C_2 \, \|\mathrm{Heat}_Q(a_\nu) - a_\nu\|_{W^{1,2}(Y)} \tag{3.46}$$

for some C_2 depending only on Y and $1 \leq q \leq 4$ (the inequality on the left is (3.44)). Since Heat_Q is continuous in the $W^{1,2}$-topology, and restricts to the identity on the space of flat connections, there is some $\epsilon' > 0$ such that if a_ν is within ϵ' of the space of flat connections, then

$$C_2 \, \|\mathrm{Heat}_Q(a_\nu) - a_\nu\|_{W^{1,2}(Y)} \leq \frac{c_0}{2}.$$

By (3.45) this condition is satisfied, and so we have a contradiction to (3.46).

□

The next lemma states that we can always put a connection $a \in \mathcal{A}(Q)$ in a gauge so that it is an element of $\mathcal{A}(Q, \partial Q)$. We state a version with an additional $\mathbb{R}$ parameter, since this is the context in which the lemma will be used.

Lemma 3.2.5. *Let Y be an oriented compact 3-manifold, possibly with boundary, and $Q \to Y$ a principal G-bundle. Suppose Σ is a closed orientable surface with an embedding $\iota : \Sigma \to Y$. Furthermore, we suppose that $\iota(\Sigma)$ lies entirely in the interior of Y (resp. is a boundary component of Y). Fix a bicollar (resp. collar) neighborhood $I \times \Sigma \hookrightarrow U \subset Y$ of Σ, and let t denote the I-coordinate induced by this embedding. Then for every $A \in \mathcal{A}^{1,2}(\mathbb{R} \times Q)$, there is a gauge transformation $u \in \mathcal{G}_0^{2,2}(\mathbb{R} \times Q)$ such that, in $\mathbb{R} \times U$, the dt component of u^*A vanishes. Furthermore, if A is smooth, then u^*A is smooth as well.*

Proof. Fix $A \in \mathcal{A}^{1,2}(\mathbb{R} \times Q)$. We may suppose $\iota(\Sigma) \subset \operatorname{int} Y$, by replacing Y with its double. Fix a bump function b with support in a small neighborhood of U and equal to 1 on U. Just as in the temporal gauge construction (section 2.2.2), there is a gauge transformation $u \in \mathcal{G}_0^{2,2}(\mathbb{R} \times Q)$ such that $\iota_{bV} u^* A = 0$. $\square$

In particular, we immediately have the following corollary.

Corollary 3.2.6. *Let $Q \to Y$ be as in Remark 2.2.7. Then for every $A \in \mathcal{A}^{1,2}(\mathbb{R} \times Q)$ there is a gauge transformation $u \in \mathcal{G}_0^{2,2}(\mathbb{R} \times Q)$ with*

$$u^* A|_{\{s\} \times Y_{i(i+1)}} \in \mathcal{A}^{1,2}(Q_{i(i+1)}, \partial Q_{i(i+1)}), \qquad \forall i.$$

Furthermore, if A is smooth then $u^* A$ is smooth as well.

Proof of Remark 3.2.2. We suppress the Sobolev exponents, for simplicity, indicating when and how they become relevant. By definition, $\mathrm{NS}_P(\alpha)$ lies in the complex gauge orbit of α. The key observation to the proof of (3.39) is that the Yang-Mills heat flow, and hence $\mathrm{Heat}_P(\alpha)$, always lies in the complexified gauge orbit of the initial condition α. Indeed, in [7] Donaldson shows that for any $\alpha \in \mathcal{A}(P)$ there is some path $\mu(\tau) \in \Omega^0(\Sigma, P(\mathfrak{g}))$ for which the equation

$$\frac{d}{d\tau} \widetilde{\alpha}_\tau = -d^*_{\widetilde{\alpha}_\tau} F_{\widetilde{\alpha}_\tau} + d_{\widetilde{\alpha}_\tau} \mu(\tau), \qquad \widetilde{\alpha}_0 = \alpha \tag{3.47}$$

has a unique solution $\tau \mapsto \widetilde{\alpha}_\tau$ for all $0 \leq \tau < \infty$. The solution has the further property that it takes the form

$$\widetilde{\alpha}_\tau = u_\tau^* \alpha$$

for some path of *complex* gauge transformations $u_\tau \in \mathcal{G}(P)^{\mathbb{C}}$ starting at the identity. It is then immediate that $\alpha_\tau := \exp\left(\int_0^\tau \mu\right)^* \widetilde{\alpha}_\tau$ solves (3.37), and so

$$\mathrm{Heat}_P(\alpha) = \lim_{\tau \to \infty} \exp\left(\int_0^\tau \mu\right)^* u_\tau^* \alpha.$$

Clearly $\exp\left(\int_0^\tau \mu\right)^* u_\tau^* \alpha$ lies in the complex gauge orbit of α for all τ, and so $\mathrm{Heat}_P(\alpha)$ must as well. Now $\mathrm{NS}_P(\alpha)$ lies in the complex gauge orbit by definition, so there is some complex gauge transformation $\widetilde{u} \in \mathcal{G}(P)^{\mathbb{C}}$ (possibly depending on α) with $\widetilde{u}^*\mathrm{NS}(\alpha) = \mathrm{Heat}(\alpha)$.

We will be done if we can show that $\widetilde{u}$ is a real gauge transformation which lies in the identity component. The former statement is equivalent to showing $\widetilde{h} := \widetilde{u}^\dagger\widetilde{u} = \mathrm{Id}$. By (3.10) we must have that $\widetilde{h}$ is a solution to

$$\widetilde{\mathcal{F}}(h) := i\overline{\partial}_{\mathrm{Heat}(\alpha)}\left(h^{-1}\partial_{\mathrm{Heat}(\alpha)}h\right) = 0.$$

Clearly the identity, Id, is a solution as well. It suffices to show that this equation has a unique solution, at least for α close to the space of flat connections. The map $\widetilde{\mathcal{F}}$ is defined on (the $W^{2,2}$-completion of) $\mathcal{G}(P)^{\mathbb{C}}$, we can take its codomain to be (the L^2 completion of) $\Omega^0\left(\Sigma, P(\mathfrak{g})^{\mathbb{C}}\right)$. Similarly to our analysis of $\mathcal{F}$ in the proof of Theorem 3.1.1, the derivative of $\widetilde{\mathcal{F}}$ at the identity is

$$\begin{aligned} W^{2,2}\left(P(\mathfrak{g})^{\mathbb{C}}\right) &\longrightarrow L^2\left(P(\mathfrak{g})^{\mathbb{C}}\right) \\ \xi &\longmapsto \tfrac{1}{2}\Delta_{\mathrm{Heat}(\alpha),\rho}\xi, \end{aligned}$$

which is invertible. So by the inverse function theorem $\widetilde{\mathcal{F}}$ is a diffeomorphism in a neighborhood of the identity, which is the uniqueness we are looking for, provided we can arrange so that $\widetilde{h}$ lies in a suitably small neighborhood of the identity. However, this is immediate since the gauge transformation $\widetilde{h}$ depends continuously on α in the $W^{1,2}$-topology, and $\widetilde{h} = \mathrm{Id}$ if α is flat.

To finish the proof of (3.39), we need to show that $\widetilde{u} \in \mathcal{G}(P)$ is actually in the identity component $\mathcal{G}_0(P)$. However, this is also immediate from the continuous dependence of $\widetilde{u}$ on α. Indeed, path from α to $\mathcal{A}_{\mathrm{flat}}(P)$ that never leaves a suitably small neighborhood of $\mathcal{A}_{\mathrm{flat}}(P)$ provides (via this construction applied to the values of this path of connections) a path of real gauge transformations from $\widetilde{u}$ to the identity. $\square$

Chapter 4

Neck-stretching limit of instantons

In this section we carry out the proof of the Main Theorem, which we split up into three parts. The first part establishes some uniform elliptic estimates which are at the heart of proving the type of convergence we seek. The second is Theorem 4.2.1, which says that the conclusions of the Main Theorem hold if we assume several hypotheses on the sequence of connections. In the third part (section 4.3) we prove that these hypotheses must hold due to energy considerations.

4.1 Uniform elliptic regularity

We establish an elliptic estimate that will be used in the proof of Theorem 4.2.1, below. To state it, we introduce the ϵ-dependent norm

$$\|\eta\|^2_{L^2(U),\epsilon} := \int_U \langle \eta \wedge *_\epsilon \eta \rangle,$$

for subsets $U \subseteq \mathbb{R} \times Y$, where $*_\epsilon$ is the Hodge star on $\mathbb{R} \times Y$ determined by the metric $ds^2 + g_\epsilon$ from section 2.1.

Proposition 4.1.1. *Let $Q \to Y$ be the bundle from Remark 2.2.7. Then for any compact $K \subset \mathbb{R} \times Y$ and $R > 0$ with $K \subset (-R, R) \times Y$, there is some constant $C = C(K, R)$ with*

$$\|\nabla_s F_A\|_{L^2(K),\epsilon} + \|\nabla_s^2 F_A\|_{L^2(K),\epsilon} \leq C \|F_A\|_{L^2((-R,R)\times Y),\epsilon}$$

for all $\epsilon > 0$ and all ϵ-ASD connections A. Here $\nabla_s = \partial_s + [p, \cdot]$, where p is the ds-component of A.

Remark 4.1.2. *Write $Y = Y_\bullet \cup I \times \Sigma_\bullet$, and let t denote the I-variable on $I \times \Sigma_\bullet$. The proof we give here proves the proposition with ∇_t replacing ∇_s, but only for compact sets lying in the open set $\mathbb{R} \times (0,1) \times \Sigma_\bullet$. It fails to provide an ϵ-independent bound for compact sets K which intersect $\mathbb{R} \times Y_\bullet$, since ∇_t is not canonically defined on $Y_\bullet$, and any open set containing K must overlap into the interior of $\mathbb{R} \times Y_\bullet$. However, if $K \subset \mathbb{R} \times [0,1] \times \Sigma_\bullet$ intersects the seams, then the proof we give here shows*

$$\epsilon^{1/2}\|\nabla_t F_A\|_{L^2(K),\epsilon} + \epsilon^{1/2}\|\nabla_s \nabla_t F_A\|_{L^2(K),\epsilon} + \epsilon\|\nabla_t^2 F_A\|_{L^2(K),\epsilon} \leq C\|F_A\|_{L^2((-R,R)\times Y),\epsilon}.$$

To see this, multiply ∇_t by a bump function h supported in a small neighborhood of K. Then the only modification to the proof is that the estimate (4.3) needs to be replaced by

$$\epsilon|\partial_t h| + \epsilon|\partial_s \partial_t h| + \epsilon^2|\partial_t^2 h| \leq C_0 h.$$

We will use Proposition 4.1.1 in the following capacity.

Corollary 4.1.3. *Let $Q \to Y$ be the bundle from Remark 2.2.7, and write $Y = Y_\bullet \cup I \times \Sigma_\bullet$, as above. Then for any compact $K \subset \mathbb{R} \times I \times \Sigma_\bullet$, there is some $C = C(K) > 0$ such that*

$$\|\nabla_s \beta_s\|^2_{L^2(K)} + \|\nabla_s^2 \beta_s\|^2_{L^2(K)} \leq C\left(\mathcal{CS}_{a_0}(a^-) - \mathcal{CS}_{a_0}(a^+)\right) \tag{4.1}$$

for any $\epsilon > 0$ and any ϵ-ASD connection A with limits

$$\lim_{s \to \pm\infty} A|_{\{s\}\times Y} = a^\pm \in \mathcal{A}_{\text{flat}}(Q).$$

Here a_0 is any fixed reference connection, and $-\beta_s$ is the ds-component of the curvature F_A over $\mathbb{R} \times I \times \Sigma_\bullet$:

$$F_A = F_\alpha - \beta_s \wedge ds - \beta_t \wedge dt + \gamma\, ds \wedge dt.$$

Proof of Corollary 4.1.3. Over $I \times \Sigma_\bullet \subset Y$, the metric g_ϵ has the form $dt^2 + \epsilon^2 g_\Sigma$, where g_Σ is a fixed metric on $\Sigma_\bullet$. Let μ be a 1-form on $\Sigma_\bullet$. Then by the scaling relation (2.8) it follows that the norm

$$\|\mu \wedge ds\|_{L^2(K),\epsilon} = \|\mu \wedge ds\|_{L^2(K)} = \|\mu\|_{L^2(K)}$$

is independent of ϵ.

Let A be as in the statement of the corollary. Then $\nabla_s \beta_s \wedge ds$ is a component of $\nabla_s F_A$, so we have

$$\begin{aligned} \|\nabla_s \beta_s\|^2_{L^2(K)} &= \|\nabla_s \beta_s \wedge ds\|^2_{L^2(K),\epsilon} \\ &\leq \|\nabla_s F_A\|^2_{L^2(K),\epsilon} \\ \text{(Proposition 4.1.1)} \quad &\leq C\mathcal{YM}(A) \\ \text{(Equation (2.30))} \quad &= C\left(\mathcal{CS}_{a_0}(a^-) - \mathcal{CS}_{a_0}(a^+)\right). \end{aligned}$$

The same computation holds with ∇_s replaced by ∇_s^2. □

Proof of Proposition 4.1.1. We first prove

$$\|\nabla_s F_A\|_{L^2(K),\epsilon} \leq C\|F_A\|_{L^2(\Omega),\epsilon}. \tag{4.2}$$

Fix a smooth bump function $h : \mathbb{R} \to \mathbb{R}^{\geq 0}$, which we view as being a Y-independent function defined on $\mathbb{R} \times Y$. Assume $h|_K = 1$ and that h has compact support in $(-R, R) \times Y$. There is a constant C_0, depending only on h (and consequently only on R and K), with

$$|\partial_s h| + |\partial_s^2 h| \leq C_0 h. \tag{4.3}$$

Set $\Omega := (-R, R) \times Y$. Then we have

$$\begin{aligned}
\|\nabla_s F_A\|^2_{L^2(K),\epsilon} &\leq \|h\nabla_s F_A\|^2_{L^2(\Omega),\epsilon} \\
&= \int_\Omega h^2 \langle \nabla_s F_A \wedge *_\epsilon \nabla_s F_A \rangle \\
&= -\int_\Omega h^2 \langle \nabla_s F_A \wedge \nabla_s F_A \rangle,
\end{aligned}$$

where we have used the ϵ-ASD condition and the commutativity relation

$$\nabla_s *_\epsilon = *_\epsilon \nabla_s.$$

Write $A = a(s) + p(s)\, ds$. Then we have

$$F_A = F_{a(s)} - b_s \wedge ds,$$

where

$$b_s := \partial_s a - d_a p.$$

Then

$$\begin{aligned}
\nabla_s F_A &= \partial_s F_A + [p, F_A] \\
&= d_A \partial_s A - d_A(d_A p) \\
&= d_A \left(\partial_s A - d_A p \right).
\end{aligned}$$

where d_A is the covariant derivative on the 4-manifold $\mathbb{R} \times Y$. We also have

$$\partial_s A - d_A p = \partial_s a - d_a p + (\partial_s p - \nabla_s p)\ ds = b_s,$$

since the ds terms cancel, and so we can write

$$\nabla_s F_A = d_A b_s, \tag{4.4}$$

where we are viewing b_s as a 1-form on the 4-manifold $\mathbb{R} \times Y$. Using (4.4) and Stokes' theorem, we obtain

$$\begin{aligned}
\|\nabla_s F_A\|^2_{L^2(K),\epsilon} &\leq \|h\nabla_s F_A\|^2_{L^2(\Omega),\epsilon} \\
&= -\int_\Omega h^2 \langle d_A b_s \wedge d_A b_s\rangle \\
&= \int_\Omega 2h\, \partial_s h\, ds \wedge \langle b_s \wedge d_A b_s\rangle - \int_\Omega h^2 \langle b_s \wedge [F_A \wedge b_s]\rangle \\
&= \int_\Omega 2h\, \partial_s h\, ds \wedge \langle b_s \wedge d_A b_s\rangle - \int_\Omega h^2 \langle b_s \wedge [b_s \wedge b_s]\rangle \wedge ds
\end{aligned}$$

where, in the last step, we used $F_A = F_a - b_s \wedge ds$ and the fact that Y does not admit non-zero 4-forms (it's a 3-manifold!). Next, use the inequality

$$2ab \leq 2a^2 + \frac{1}{2}b^2 \tag{4.5}$$

with (4.3) and the identity (4.4) on the first term on the right to get

$$\begin{aligned}
\left\{\|\nabla_s F_A\|^2_{L^2(K),\epsilon} \leq\right\} \|h\nabla_s F_A\|^2_{L^2(\Omega),\epsilon} &\leq 2C_0\|hb_s\|^2_{L^2(\Omega),\epsilon} + \tfrac{1}{2}\|h\nabla_s F_A\|^2_{L^2(\Omega),\epsilon} \\
&\quad - \int_\Omega h^2\langle b_s \wedge [b_s \wedge b_s]\rangle \wedge ds
\end{aligned}$$

Subtract the term $\frac{1}{2}\|h\nabla_s F_A\|^2_{L^2(\Omega),\epsilon}$ from both sides to get

$$\tfrac{1}{2}\|\nabla_s F_A\|^2_{L^2(K),\epsilon} \leq \tfrac{1}{2}\|h\nabla_s F_A\|^2_{L^2(\Omega),\epsilon} \leq 2C_0\|hb_s\|^2_{L^2(\Omega),\epsilon} - \int_\Omega h^2\langle b_s \wedge [b_s \wedge b_s]\rangle \wedge ds \tag{4.6}$$

In particular, we record the following:

$$0 \leq 2C_0\|hb_s\|^2_{L^2(\Omega),\epsilon} - \int_\Omega h^2\langle b_s \wedge [b_s \wedge b_s]\rangle \wedge ds. \tag{4.7}$$

We want to estimate the right-hand side of (4.6) in terms of $\|F_A\|_{L^2(\Omega),\epsilon} = 2\|b_s\|_{L^2(\Omega),\epsilon}$. Set

$$g(s) := 2C_0 \|hb_s\|^2_{L^2(Y),\epsilon} - \int_Y h^2 \langle b_s \wedge [b_s \wedge b_s] \rangle$$

(so if we integrate over $(-R, R)$ then we recover the right-hand side of (4.7), since $\Omega = (-R, R) \times Y$). Then (4.7), and the fact that h vanishes outside of $(-R, R)$, gives

$$\int_{-R}^{R} g(s)\, ds \geq 0, \qquad \text{and} \qquad g(s) = 0 \quad \text{for } s \notin [-R, R]\,. \tag{4.8}$$

Next, set

$$e(s) := \frac{1}{2} \|hb_s\|^2_{L^2(Y),\epsilon}$$

(again, only integrating over Y).

Claim: There are constants $D_0, D_1 > 0$, depending only on K, R, with

$$e''(s) + D_0 e(s) \geq D_1 g(s)$$

The bound (4.2) follows immediately from the claim and Lemma 4.1.4 below (the latter uses (4.8)):

$$\begin{array}{rcl}
\|\nabla_s F_A\|^2_{L^2(K),\epsilon} & \leq & 2g(s) \\
& \leq & C \displaystyle\int_{[-R-1,R+1]\times Y} e(s) \\
(h = 0 \text{ outside } [-R, R]) & = & C \displaystyle\int_{[-R,R]\times Y} e(s) \\
& \leq & \frac{1}{2} C \|b_s\|^2_{L^2(\Omega),\epsilon} \\
& = & C \|F_A\|^2_{L^2(\Omega),\epsilon}
\end{array}$$

where the first inequality is (4.6), the second is the assertion of Lemma 4.1.4. The last two lines are just definitions.

To prove the claim, we have

$$e''(s) \;=\; \|\nabla_s(hb_s)\|^2_{L^2(Y),\epsilon} + \int_Y \langle \nabla_s \nabla_s(hb_s) \wedge *_\epsilon hb_s \rangle$$

Next, we have

$$\nabla_s^2(hb_s) = \left(\partial_s^2 h\right)\; b_s + 2\left(\partial_s h\right)\nabla_s b_s + h\, d_a^{*_\epsilon} d_a b_s - h *_\epsilon [b_s \wedge b_s]\,,$$

where $d_a^{*_\epsilon} = *_\epsilon d_a *_\epsilon$ on 1- and 2-forms on Y. Then

$$\begin{aligned} e''(s) \;=\; & \|\nabla_s(hb_s)\|^2_{L^2(Y),\epsilon} + \int_Y \langle \left(\partial_s^2 h\right)\; b_s \wedge *_\epsilon h\, b_s \rangle \\ & +2\int_Y \langle (\partial_s h)\,\nabla_s b_s \wedge *_\epsilon h\, b_s\rangle + \int_Y h^2 \langle d_a^{*_\epsilon} d_a b_s \wedge *_\epsilon b_s \rangle \\ & -\int_Y h^2 \langle *_\epsilon [b_s \wedge b_s] \wedge *_\epsilon b_s \rangle \end{aligned} \tag{4.9}$$

The first term on the right-hand side of (4.9) is fine, but we need to estimate the remaining terms. We begin with the second term by applying (4.3):

$$\int_Y \langle \left(\partial_s^2 h\right)\; b_s \wedge *_\epsilon h\, b_s \rangle \geq -C_1 e.$$

For the third term in (4.9) do the same, except also use (4.5):

$$2\int_Y \langle (\partial_s h)\,\nabla_s b_s \wedge *_\epsilon h\, b_s\rangle \geq -2e - \frac{1}{2}\|h\nabla_s b_s\|^2_{L^2(Y),\epsilon}$$

For the fourth term in (4.9) we integrate by parts and use (4.5):

$$\begin{aligned} \int_Y h^2 \langle d_a^{*_\epsilon} d_a b_s \wedge *_\epsilon b_s \rangle \;&=\; \int_Y h^2 \langle d_a *_\epsilon d_a b_s \wedge b_s \rangle \\ &=\; -\int_Y 2h\, dh \wedge \langle *_\epsilon d_a b_s \wedge b_s \rangle + \|h\, d_a b_s\|^2_{L^2(Y),\epsilon} \\ &\geq\; -2e - \tfrac{1}{2}\|h\, d_a b_s\|^2_{L^2(Y),\epsilon} + \|h\, d_a b_s\|^2_{L^2(Y),\epsilon} \\ &=\; -2e + \tfrac{1}{2}\|h\, d_a b_s\|^2_{L^2(Y),\epsilon}. \end{aligned}$$

Putting this all together, and using $d_a b_s = - *_\epsilon \nabla_s b_s$, we get

$$\begin{aligned} e'' &\geq \|\nabla_s(hb_s)\|^2_{L^2(Y),\epsilon} - C_2 e - \int_Y h^2 \langle *_\epsilon [b_s \wedge b_s] \wedge *_\epsilon b_s \rangle \\ &\geq -C_2 e - \int_Y h^2 \langle *_\epsilon [b_s \wedge b_s] \wedge *_\epsilon b_s \rangle \end{aligned}$$

By adding $(C_2 + 2C_0)e$ to both sides we recover the claim and this finishes the proof of the first derivative bound (4.2).

To finish the proof of the proposition, we need to prove the following bound on the second derivative:

$$\|\nabla_s^2 F_A\|_{L^2(K),\epsilon} \leq C\|F_A\|_{L^2(\Omega),\epsilon}$$

This is very similar to the proof of (4.2), so we only sketch the main points. The analogue of (4.6) for this case is

$$0 \leq \frac{1}{2}\|\nabla_s^2 F_A\|^2_{L^2(K),\epsilon} \leq 2C_0\|h\nabla_s b_s\|^2_{L^2(\Omega),\epsilon} + \int_{-R}^{R} \widetilde{g}(s)\, ds,$$

where we have set

$$g_1(s) := -3\int_Y h^2 \langle \nabla_s b_s \wedge [\nabla_s b_s \wedge b_s]\rangle + \int_Y h^2 \langle [b_s \wedge b_s] \wedge \nabla_s^2 b_s \rangle.$$

Then (4.8) continues to hold with g replaced by g_1. Set

$$e_1(s) := \frac{1}{2}\|h\nabla_s b_s\|^2_{L^2(Y),\epsilon}$$

and, just as before, one can show

$$e_1''(s) + D_0 e_1(s) \geq D_1 g_1(s)$$

and the result follows from Lemma 4.1.4. □

Lemma 4.1.4. *Consider functions $e, f, g : B_{R+r} \to \mathbb{R}$, where $B_\rho := (-\rho, \rho) \subset \mathbb{R}$, and assume these satisfy*

- *$e \geq 0$ is C^2;*

- *$f \geq 0$ is C^0;*
- *g is C^0 and satisfies*

$$\int_{B_R} g \geq 0, and g(s) \geq 0 for\, s \in B_{R+r} \backslash B_R.$$

Suppose

$$g(s) \leq f(s) + e''(s), \qquad s \in B_{R+r}.$$

Then

$$\int_{B_R} g \leq \int_{B_{R+r}} f + \frac{4}{r^2} \int_{B_{R+r} \backslash B_R} e.$$

Proof. This is a variation of [24, Lemma 9.2], and is proved in essentially the same way. However, we recall the proof for convenience.

By considering the rescaled functions

$$\widetilde{e}(s) := e(rs), \qquad \widetilde{f}(s) := r^2 f(rs), \qquad \widetilde{g}(s) := r^2 g(rs),$$

(which satisfy $\widetilde{g} \leq \widetilde{f} + \widetilde{e}''$), it suffices to assume $r = 1$. The positivity conditions on f and g give

$$\begin{aligned} \int_{-R}^{R} g - \int_{-R-1}^{R+1} f \quad &\leq \quad \int_{-R-s}^{R+s} g - f \\ &= \quad \int_{-R-s}^{R+s} \frac{d^2}{ds^2} e = e'(R+s) - e'(-R-s), \end{aligned}$$

for all $s \in [0, 1]$.[1] Note also that we have

$$\frac{d}{ds} \left(e(R+s) + e(-R-s) \right) \quad = \quad e'(R+s) - e'(-R-s)$$

In particular, integrating in s from $1/2$ to t gives

[1] If we know that $e' = 0$ on $B_{R+1} \backslash B_R$, as we do in the proof of Proposition 4.1.1, then this proves the result of the Lemma.

$$\begin{aligned}\frac{1}{2}\left(\int_{-R}^{R} g - \int_{-R-1}^{R+1} f\right) &\leq e(R+t) + e(-R-t) - e(R+1/2) - e(-R-1/2) \\ &\leq e(R+t) + e(-R-t)\end{aligned}$$

by the positivity of e. Now integrate in t from 1/2 to 1:

$$\frac{1}{4}\left(\int_{-R}^{R} g - \int_{-R-1}^{R+1} f\right) \leq \int_{-1/2}^{1} e(R+t) + e(-R-t) = \int_{B_{R+1}\setminus B_{R+1/2}} e.$$

$\square$

4.2 Convergence to holomorphic strips

This section establishes a convergence result which will be used in the proof of the Main Theorem. It provides sufficient conditions for a sequence of instantons to converge, in C^0 on compact subsets of $\mathbb{R}\times I\times\Sigma_\bullet$, to a holomorphic strip with Lagrangian boundary conditions.

The simplest version of the theorem holds for connections on $\mathbb{R}\times Y$. However, for applications we will need to consider slightly more general domains. These will be of the form $(\mathbb{R}\times Y)\setminus\cup_{k=1}^{K} S_k$, where each $S_k\subset\mathbb{R}\times Y$ is either a point or a slice:

$$\begin{array}{lll}
\{(s_k, y_k)\} & \subset \mathbb{R}\times Y_\bullet & \text{point} \\
\{(s_k, t_k, \sigma_k)\} & \subset \mathbb{R}\times(0,1)\times\Sigma_\bullet & \text{point} \\
 & & \\
\{s_k\}\times Y_\bullet & \subset \mathbb{R}\times Y_\bullet & \text{slice} \\
\{(s_k, t_k)\}\times\Sigma_\bullet & \subset \mathbb{R}\times(0,1)\times\Sigma_\bullet & \text{slice}
\end{array}$$

Each S_k induces a **shadow**, $\hat{S}_k$, in $\mathbb{R}\times I$ defined as the projection to $\mathbb{R}\times I$ of the intersection $S_k\cap\mathbb{R}\times I\times\Sigma_\bullet$:

$$\hat{S}_k = \begin{cases} \{(s_k,0),(s_k,1)\} & \text{if } S_k = \{(s_k,y_k)\} \\ \{(s_k,t_k)\} & \text{if } S_k = \{(s_k,t_k,\sigma_k)\} \\ & \\ \{(s_k,0),(s_k,1)\} & \text{if } S_k = \{s_k\} \times Y_\bullet \\ \{(s_k,t_k)\} & \text{if } S_k = \{(s_k,t_k)\} \times \Sigma_\bullet. \end{cases} \tag{4.10}$$

The key point is that $\cup_k \hat{S}_k$ is always a finite set of *points*.

To state the theorem, we also recall from Lemma 3.1.5 that the L^2-orthogonal projection

$$\mathrm{proj}_\alpha : W^{1,q}\left(T^*\Sigma \otimes P(\mathfrak{g})\right) \longrightarrow \left(\mathrm{Im}\, d_\alpha \oplus \mathrm{Im}\, d_\alpha^*\right)^\perp = \ker d_\alpha \cap \ker d_\alpha^*,$$

is well-defined whenever $q > 1$ and $\|F_\alpha\|_{L^q(\Sigma)}$ is sufficiently small. Let $\tau_{s_0} : \mathbb{R} \times Y \to \mathbb{R} \times Y$ be the map given by translating by $s_0 \in \mathbb{R}$, as described in section 2.4.

Theorem 4.2.1. *Fix $2 < q < \infty$. Let Q be as in Remark 2.2.7, and suppose $\{S_k\}_k$ is a finite collection of points or slices $S_k \subset \mathbb{R} \times Y$ as above. Write $\hat{S}_k$ for the shadow of S_k as in (4.10), and set*

$$X := (\mathbb{R} \times Y) \setminus \cup_k S_k,$$

$$\hat{X} := (\mathbb{R} \times I) \setminus \cup_k \hat{S}_k.$$

Assume that all flat connections on Q are non-degenerate, and fix $a^\pm \in \mathcal{A}_{\mathrm{flat}}(Q)$.

Suppose $(\epsilon_\nu)_{\nu \in \mathbb{N}}$ is a sequence of positive numbers converging to 0. In addition, suppose that for each ν there is an ϵ_ν-ASD connection

$$A_\nu = \begin{cases} \alpha_\nu(s,t) + \phi_\nu(s,t)\, ds + \psi_\nu(s,t)\, dt & \text{on } \{(s,t)\} \times \Sigma_\bullet \\ a_\nu(s) + p_\nu(s)\, ds & \text{on } \{s\} \times Y_\bullet \end{cases}$$

in $\mathcal{A}^{1,q}_{loc}(\mathbb{R} \times Q)$, which satisfies the following conditions:

- *each A_ν limits to $a^\pm$ at $\pm\infty$*

$$\lim_{s\to\pm\infty} \left(u_\nu^\pm\right)^* A_\nu|_{\{s\}\times Y} = a^\pm, \textit{for some } u_\nu^\pm \in \mathcal{G}_0^{2,q}(Q),$$

- *for each compact subset $K \subset X$, the slice-wise curvature terms decay to zero*

$$\|F_{\alpha_\nu}\|_{L^\infty(K)} + \|F_{a_\nu}\|_{L^\infty(K)} \xrightarrow{\nu} 0.$$

- *for each compact $\hat{K} \subset \hat{X}$, there is some constant $C = C(\hat{K})$ with*

$$\sup_{\hat{K}} \|\mathrm{proj}_{\alpha_\nu} \left(\partial_s \alpha_\nu - d_{\alpha_\nu}\phi_\nu\right)\|_{L^2(\Sigma_\bullet)} \leq C.$$

Then there exist a finite sequence of flat connections

$$\left\{a^- = a^0, a^1, \ldots, a^{J-1}, a^J = a^+\right\} \subseteq \mathcal{A}_{\mathrm{flat}}(Q)$$

and, for each $j \in \{1, \ldots, J\}$, a continuous connection

$$A^j = \begin{cases} \alpha^j(s,t) + \phi^j(s,t)\, ds + \psi^j(s,t)\, dt & \textit{on } \{(s,t)\} \times \Sigma_\bullet \\ a^j(s) + p^j(s)\, ds & \textit{on } \{s\} \times Y_\bullet \end{cases}$$

in $\mathcal{A}_{loc}^{1,q}(\mathbb{R} \times Q)$, which

(i) is holomorphic:

$$\partial_s \alpha^j - d_{\alpha^j}\phi^j + * \left(\partial_t \alpha^j - d_{\alpha^j}\psi^j\right) = 0, F_{\alpha^j} = 0,$$

(ii) has Lagrangian boundary conditions:

$$F_{a^j} = 0,$$

(iii) and limits to the flat connections a^{j-1}, a^j at $\pm\infty$:

$$\lim_{s\to+\infty} \left(u_+^j\right)^* A^j|_{\{s\}\times Y} = a^j, \textit{for some } u_+^j \in \mathcal{G}_0^{2,q}(Q),$$

$$\lim_{s \to -\infty} \left(u_-^j\right)^* A^j|_{\{s\} \times Y} = a^{j-1}, \textit{for some } u_-^j \in \mathcal{G}_0^{2,q}(Q).$$

Moreover, after possibly passing to a subsequence, the A_ν converge to A^j in the following sense: There is a sequence of gauge transformations $u_\nu \in \mathcal{G}_{loc}^{2,q}(\mathbb{R} \times Q)$, and, for each j, a sequence $s_\nu^j \in \mathbb{R}$ such that

(iv) $\left\| \alpha^j - u_\nu^* \tau_{s_\nu^j}^* \alpha_\nu \right\|_{C^0(K)} \xrightarrow{\nu} 0$

(v) $\sup_{\hat{K}} \left\| (\partial_s \alpha^j - d_{\alpha^j} \phi^j) - \mathrm{Ad}(u_\nu) \tau_{s_\nu^j}^* \mathrm{proj}_{\alpha_\nu} (\partial_s \alpha_\nu - d_{\alpha_\nu} \phi_\nu) \right\|_{L^2(\Sigma_\bullet)} \xrightarrow{\nu} 0$

for all compact subsets $K \subset X$ and $\hat{K} \subset \hat{X}$. In item (iv) *above, the gauge action is the (s,t)-pointwise action of $u_\nu(s,t) \in \mathcal{G}^{2,q}(P_\bullet)$ on connections over the surface $\Sigma_\bullet$. Furthermore, the energy of this limit is bounded by the energy of the sequence of instantons:*

$$\sum_{j=1}^{J} E(A^j) \leq \sup_K \ \limsup_{\nu \to \infty} \frac{1}{2} \|F_{A_\nu}\|^2_{L^2(K), \epsilon_\nu} \tag{4.11}$$

where $\sup_K$ is the supremum over all compact $K \subset X$, and

$$E(A^j) := \frac{1}{2} \int_{\mathbb{R} \times I \times \Sigma_\bullet} |\partial_s \alpha^j - d_{\alpha^j} \phi^j|^2$$

is the energy of the holomorphic curve A^j.

Remark 4.2.2. *(a) The same proof we give here carries over with only minor notational changes to the following situation. Set*

$$X := \mathbb{R} \times \{Y \cup_{\partial Y} ([0, \infty) \times \partial Y)\}$$

or

$$X := \mathbb{C} \times \Sigma$$

for some compact connected oriented elementary cobordism Y, or closed connected oriented surface Σ. Suppose, for each ν, we have an open set $U_\nu \subseteq X$ which is a deformation retract of X, and with the further property that the U_ν are increasing and exhausting: $U_\nu \subset U_{\nu+1}$ and $X = \cup_\nu U_\nu$. Then the statement of Theorem 4.2.1 continues to hold if we assume that A_ν is defined on U_ν. In this case we cannot assume that the connections A_ν have fixed limits at ∞, so there is no analogue of item (iii) from Theorem 4.2.1 for these domains, and it is sufficient to take $s_\nu = 0$ for all ν. We also do not need to have any non-degeneracy assumptions on the flat connections on the bundle Q, since these assumptions are only used to prove item (iii).

(b) The projection operator appearing in conclusion (v) of Theorem 4.2.1 is a little awkward. However, it can be removed at the cost of weakening the sup norm to an L^p norm. Indeed, we have

$$\int_{\hat{K}} \left\| \left(\partial_s \alpha^j - d_{\alpha^j}\phi^j\right) - \mathrm{Ad}(u_\nu)\tau^*_{s^j_\nu}\left(\partial_s\alpha_\nu - d_{\alpha_\nu}\phi_\nu\right)\right\|^p_{L^2(\Sigma_\bullet)} \stackrel{\nu}{\longrightarrow} 0$$

for any $1 < p < \infty$, after possibly passing to a further subsequence. To see this, set

$$\beta_{s,\nu} := \partial_s\alpha_\nu - d_{\alpha_\nu}\phi_\nu.$$

It suffices to show

$$\int_{\hat{K}} \|\beta_{s,\nu} - \mathrm{proj}_{\alpha_\nu}\beta_{s,\nu}\|^p_{L^2(\Sigma_\bullet)} \longrightarrow 0.$$

By (3.28), we have the following s,t-pointwise estimate:

$$\begin{aligned} \|\beta_{s,\nu} - \mathrm{proj}_{\alpha_\nu}\beta_{s,\nu}\|_{L^2(\Sigma_\bullet)} &\leq C\left(\|d_{\alpha_\nu}\beta_{s,\nu}\|_{L^2(\Sigma_\bullet)} + \|d_{\alpha_\nu} * \beta_{s,\nu}\|_{L^2(\Sigma_\bullet)}\right) \\ &= C\left(\|\nabla_{s,\nu}F_{\alpha_\nu}\|_{L^2(\Sigma_\bullet)} + \|\nabla_{t,\nu}F_{\alpha_\nu}\|_{L^2(\Sigma_\bullet)}\right). \end{aligned}$$

Consider the sequence of maps

$$\begin{array}{ccc} \mathbb{R}\times I & \longrightarrow & L^2(P(\mathfrak{g})) \\ (s,t) & \longrightarrow & *\nabla_s F_{\alpha_\nu} \end{array} \tag{4.12}$$

By Remark 4.1.2 there is a uniform bound

$$\|\nabla_s \nabla_s F_{\alpha_\nu}\|_{L^2(K\times\Sigma_\bullet)} + \|\nabla_t \nabla_s F_{\alpha_\nu}\|_{L^2(K)} \leq C\left(\mathcal{CS}(a^-) - \mathcal{CS}(a^+)\right)$$

for each compact set $K \subset \mathbb{R} \times I$. In particular, by Sobolev embedding in dimension 2, we have that the sequence (4.12) is bounded in $W^{1,2}$ and so a subsequence converges strongly in L^p on compact sets, for any $1 < p < \infty$. The same argument holds for $*\nabla_{\nu,t} F_{\alpha_\nu}$ and so we have

$$\int_K \|\beta_{s,\nu} - \mathrm{proj}_{\alpha_\nu}\beta_{s,\nu}\|^p_{L^2(\Sigma_\bullet)} \leq C \int_K \|\nabla_{s,\nu} F_{\alpha_\nu}\|^p_{L^2(\Sigma_\bullet)} + \|\nabla_{t,\nu} F_{\alpha_\nu}\|^p_{L^2(\Sigma_\bullet)} \longrightarrow 0,$$

as desired.

Proof of Theorem 4.2.1. We will begin by proving the theorem but ignoring the $\mathbb{R}$-translations. We will show that the sequence converges, in the sense of items (iv) and (v), to a single limiting connection A_∞ satisfying (i) and (ii). (This will, for example, take care of the proof of the types of domains described in Remark 4.2.2 (a).) Once we have proven this, then we will describe how to incorporate $\mathbb{R}$-translations to obtain the broken holomorphic trajectory $(A^1, \ldots, A^J)$ with the limits specified in (iii).

We use the notation

$$\beta_{s,\nu} := \partial_s \alpha_\nu - d_{\alpha_\nu}\phi_\nu, \qquad \beta_{s,\infty} = \partial_s \alpha_\infty - d_{\alpha_\infty}\phi_\infty.$$

By the assumption on the curvature of the α_ν, it follows that Theorem 3.1.1 applies to $\alpha_\nu(s,t)$ for each s, t and ν sufficiently large. Let NS_{P_i} be the map constructed in Theorem 3.1.1 for the bundle $P_i \to \Sigma_i$. Define a map

$$v_\nu : \mathbb{R} \times I \longrightarrow M := M(P_1) \times \ldots \times M(P_N)$$

in terms of its coordinates by using $\Pi \circ \mathrm{NS}_{P_i}$ to project α_ν:

$$v_\nu(s,t) := (\Pi \circ \mathrm{NS}_{P_1}\left(\alpha_\nu(s,t)|_{\Sigma_1}\right), \ldots, \Pi \circ \mathrm{NS}_{P_N}\left(\alpha_\nu(s,t)|_{\Sigma_N}\right)).$$

Then Lemmas 3.1.11 and 3.1.12 imply that v_ν is holomorphic.

Claim 1: The v_ν have uniformly bounded energy:

$$\sup_\nu \sup_{\hat{K}} |\partial_s v_\nu|^2_M < \infty \tag{4.13}$$

for each compact set $\hat{K} \subset \hat{X} = (\mathbb{R} \times I) \backslash \cup_k \hat{S}_k$.

At each $(s,t) \in \hat{K}$, we have

$$\begin{aligned} |\partial_s v_\nu(s,t)|^2_M = |\partial_s v_\nu|^2_M &= \sum_i \|\partial_s \left(\Pi \circ \mathrm{NS}_{P_i}(\alpha_\nu)\right)\|^2_{L^2(\Sigma_i)} \\ &= \sum_i \|D_{\alpha_\nu}(\Pi \circ \mathrm{NS}_{P_i})(\partial_s \alpha_\nu)\|^2_{L^2(\Sigma_i)} \\ &= \sum_i \|D_{\alpha_\nu}(\Pi \circ \mathrm{NS}_{P_i})(\beta_{s,\nu})\|^2_{L^2(\Sigma_i)} \end{aligned}$$

where the last equality holds by Lemma 3.1.12, since $\partial_s \alpha_\nu$ and $\beta_{s,\nu}$ differ by an exact 1-form. Similarly, by definition of the projection, the 1-form $\beta_{s,\nu}$ differs from $\mathrm{proj}_{\alpha_\nu}\beta_{s,\nu}$ by an element of $\mathrm{Im}\, d_{\alpha_\nu} \oplus \mathrm{Im} * d_{\alpha_\nu}$, and this difference is also in the kernel of $D_{\alpha_\nu}(\Pi \circ \mathrm{NS}_{P_i})$ by Lemma 3.1.12. Hence

$$\begin{aligned} |\partial_s v_\nu|^2_M &= \sum_i \|D_{\alpha_\nu}(\Pi \circ \mathrm{NS}_{P_i})(\mathrm{proj}_{\alpha_\nu}\beta_{s,\nu})\|^2_{L^2(\Sigma_i)} \\ \text{Lemma 3.1.12} \quad &\leq C_0 \sum_i \|\mathrm{proj}_{\alpha_\nu}\beta_{s,\nu}\|^2_{L^2(\Sigma_i)} \\ &= C_0 \|\mathrm{proj}_{\alpha_\nu}\beta_{s,\nu}\|^2_{L^2(\Sigma_\bullet)} \\ &\leq C_1, \end{aligned}$$

where the last *equality* holds because $\Sigma_\bullet = \sqcup_i \Sigma_i$ (by definition), and the last *inequality* holds with $C_1 = C_0 C$ (C is as in the statement of the theorem we are proving). The constant $C_1 = C_1(\hat{K})$ is independent of s, t, ν and so this proves Claim 1.

By the holomorphic condition, we have that $\partial_t v_\nu$ is uniformly bounded on compact subsets of $\hat{X}$ as well. In particular, $\{v_\nu\}$ is a C^1-bounded sequence of maps for each compact $\hat{K} \subset \hat{X}$. By the compactness of the embedding $C^1 \hookrightarrow C^0$ for compact sets, there is a subsequence, still denoted by $\{v_\nu\}$, which converges weakly in C^1, and strongly in C^0, on compact subsets of $\hat{X}$, including the boundary:

$$v_\nu \xrightarrow{w} v_\infty, \qquad \text{weakly in } C^1 \text{ on compact subsets of } (\mathbb{R} \times I) \backslash \cup_k \hat{S}_k$$

for some $v_\infty \in C^1(\hat{X}, M)$. By the weak C^1-convergence, it follows that v_∞ is holomorphic as well

$$\partial_s v_\infty + * \partial_t v_\infty = 0.$$

By the removal of singularities theorem for holomorphic maps [30, Theorem 4.1.2 (ii)], v_∞ extends to a holomorphic map on the interior $\mathbb{R} \times (0,1)$ (this uses the fact that the shadows $\cup_k \hat{S}_k$ form a finite set). It also follows that v_∞ is C^∞ in the interior $\mathbb{R} \times (0,1)$ [30, Theorem B.4.1].

Remark 4.2.3. *We can actually say quite a bit more: The uniform energy bound given in Claim 1 implies that, after possibly passing to a further subsequence, we have that the $v_\nu : \mathbb{R} \times I \to M$ converge to v_∞ in C^∞ on compact subsets of the* interior*, $\hat{X} \cap \mathbb{R} \times (0,1)$ (see [30, Theorem 4.1.1]). In particular, this automatically proves item (v) for $\hat{K} \subset \hat{X} \cap \mathbb{R} \times (0,1)$. However, for applications we will need to prove (v) for $\hat{K} \subset \hat{X}$ (all the way up to the boundary $\partial \hat{X} \subset \{0,1\} \times \Sigma_\bullet$), which requires a more careful analysis. (See Claim 4.)*

Claim 2 below states that v_∞ has Lagrangian boundary conditions at $\hat{X} \cap \mathbb{R} \times \{0,1\}$. This will follow because the v_ν have *approximate* Lagrangian boundary conditions. To state this precisely, consider the product $M(Q_{12}) \times M(Q_{23}) \times \ldots \times M(Q_{N1})$ of moduli spaces of flat connections on the $Q_{i(i+1)}$. For each $Q_{i(i+1)}$, the restriction to each of the two boundary components provides two maps

$$\partial_{i,0} : M(Q_{i(i+1)}) \hookrightarrow M(P_i), \qquad \partial_{i,1} : M(Q_{i(i+1)}) \hookrightarrow M(P_{i+1}).$$

Fix $j = 0$ or 1. Then the $\{\partial_{i,j}\}_i$ piece together to form an embedding

$$\partial_j := M(Q_{12}) \times M(Q_{23}) \times \ldots \times M(Q_{N1}) \hookrightarrow M$$

and the image, which we denote by $L_{(j)}$, is Lagrangian by Theorem 2.2.16. (See section (2.3) for more details on this restriction.)

Claim 2: For each $s \in \hat{X} \cap \mathbb{R} \times \{0\}$, the sequence $\{v_\nu(s,j)\}_\nu$ converges (in the standard metric on M) to a point in $L_{(j)}$, for $j \in \{0,1\}$. In particular, v_∞ has Lagrangian boundary conditions

$$v_\infty(\cdot, j) : \mathbb{R} \to L_{(j)}.$$

The hypothesis that the norms $\|F_{a_\nu}\|_{C^0}$ decay to zero on compact sets away from the S_k implies that Theorem 3.2.3 applies to each $a_\nu(s)$ for all ν sufficiently large and all $s \in \hat{X} \cap \mathbb{R} \times \{0\}$ (after applying suitable gauge transformations as in Corollary 3.2.6).[2] For each i, let $\Pi \circ \mathrm{Heat}_{Q_{i(i+1)}}$ be the map constructed in Theorem 3.2.3. Define

$$\ell_\nu : \hat{X} \cap \mathbb{R} \times \{0\} \longrightarrow M(Q_{12}) \times M(Q_{23}) \times \ldots \times M(Q_{N1})$$

by

$$\ell_\nu(s) := \left(\Pi \circ \mathrm{Heat}_{Q_{12}}\left(a_\nu(s)|_{Y_{12}}\right), \ldots, \Pi \circ \mathrm{Heat}_{Q_{N1}}\left(a_\nu(s)|_{Y_{N1}}\right)\right).$$

Composing with ∂_j, for $j = 0, 1$, gives a path in the Lagrangian $L_{(j)}$:

$$\ell_{\nu,j} := \partial_j \circ \ell_\nu : \hat{X} \cap \mathbb{R} \times \{0\} \longrightarrow L_{(j)} \subset M, \qquad j = 0, 1.$$

To prove the claim, we will show

---

[2] One may be concerned that since A_ν is only assumed to be $W^{1,q}$ with $q > 2$ possibly small, it may not be the case that the codimension-1 restrictions $a_\nu(s)$ are in $W^{1,2}$ on $Y_\bullet$, and so Theorem 3.2.3 does not apply. We can get around this as follows: Since A_ν is an ϵ_ν-ASD connection on a 4-manifold, it is gauge equivalent to an ϵ_ν-smooth connection A'_ν [51, Theorem 9.4]. The restriction $A'_\nu|_{Y_\bullet}$ is smooth (so $W^{1,2}$), and A'_ν enjoys all of the same properties we assumed of A_ν. So replacing A_ν by A'_ν does the trick.

$$\sup_{s \in \hat{K} \cap \mathbb{R} \times \{0\}} \mathrm{dist}_M \left(\ell_{\nu,j}(s), v_\nu(s,j) \right) \stackrel{\nu}{\longrightarrow} 0, \tag{4.14}$$

for every compact $\hat{K} \subset \hat{X}$, where dist_M is the distance on M given by the symplectic form and the holomorphic structure determined by the Hodge star. Similarly, $\mathrm{dist}_{M(P_i)}$ is the distance on the symplectic manifold $M(P_i)$, and these determine dist_M in the usual way. The proof of (4.14) is just a computation (for clarity, we use Π_P to denote the projection $\mathcal{A}_{\mathrm{flat}}(P) \to M(P)$):

$$\begin{aligned}
& \sup_s \ \mathrm{dist}_M \left(\ell_{\nu,j}(s),\ v_\nu(s,j) \right)^2 \\
= \quad & \sup_s \sum_i \mathrm{dist}_{M(P_{i+j})} \Big(\Big\{ \Pi_{Q_{i(i+1)}} \circ \mathrm{Heat}_{Q_{i(i+1)}} \Big(a_\nu\,(s)\,|_{Y_{i(i+1)}} \Big) \Big\} \,|_{\Sigma_{i+j}}, \\
& \qquad\qquad \Pi_{P_{i+j}} \circ \mathrm{NS}_{P_{i+j}} \Big(\alpha_\nu\,(s,j)\,|_{\Sigma_{i+j}} \Big) \Big)^2 \\
= \quad & \sup_s \sum_i \mathrm{dist}_{M(P_{i+j})} \Big(\Pi_{P_{i+j}} \Big\{ \Big(\mathrm{Heat}_{Q_{i(i+1)}} a_\nu(s)|_{Y_{i(i+1)}} \Big) \,|_{\Sigma_{i+j}} \Big\}, \\
& \qquad\qquad \Pi_{P_{i+j}} \Big\{ \mathrm{NS}_{P_{i+j}} \Big(\alpha_\nu\,(s,j)\,|_{\Sigma_{i+j}} \Big) \Big\} \Big)^2 \\
\leq \quad & \sup_s \sum_i \Big\| \Big(\mathrm{Heat}_{Q_{i(i+1)}} a_\nu(s)|_{Y_{i(i+1)}} \Big) \,|_{\Sigma_{i+j}} - \mathrm{NS}_{P_{i+j}} \Big(\alpha_\nu\,(s,j)\,|_{\Sigma_{i+j}} \Big) \Big\|^2_{L^2(\Sigma_{i+j})} .
\end{aligned}$$

The last *equality* holds because restricting a flat connection on $Q_{i(i+1)}$ to the boundary commutes with harmonic projections $\Pi_{Q_{i(i+1)}}$ and $\Pi_{P_{i+j}}$. The *inequality* holds by the definition of the distance on the $M(P_i)$, and because Π_{i+j} has operator norm equal to one. By the triangle inequality, we can continue this to get

$$
\begin{aligned}
& \sup_s \ \mathrm{dist}_M \left(\ell_{\nu,j}(s),\ v_\nu(s,j)\right)^2 \\
\leq \quad & \sup_s \sum_i \Big\{ \left\| \left(\mathrm{Heat}_{Q_{i(i+1)}} a_\nu(s)|_{Y_{i(i+1)}}\right)|_{\Sigma_{i+j}} - \left(a_\nu(s)|_{Y_{i(i+1)}}\right)|_{\Sigma_{i+j}} \right\|^2_{L^2(\Sigma_{i+j})} \\
& \qquad + \left\| \left(a_\nu(s)|_{Y_{i(i+1)}}\right)|_{\Sigma_{i+j}} - \mathrm{NS}_{P_{i+j}}\left(\alpha_\nu\left(s,j\right)|_{\Sigma_{i+j}}\right) \right\|^2_{L^2(\Sigma_{i+j})} \Big\} \\
\leq \quad & \sup_s \sum_i \Big\{ \left\| \left(\mathrm{Heat}_{Q_{i(i+1)}} a_\nu(s)|_{Y_{i(i+1)}}\right)|_{\Sigma_{i+j}} - \left(a_\nu(s)|_{Y_{i(i+1)}}\right)|_{\Sigma_{i+j}} \right\|^2_{L^2(\Sigma_{i+j})} \\
& \qquad + \left\| \alpha_\nu\left(s,j\right)|_{\Sigma_{i+j}} - \mathrm{NS}_{P_{i+j}}\left(\alpha_\nu\left(s,j\right)|_{\Sigma_{i+j}}\right) \right\|^2_{L^2(\Sigma_{i+j})} \Big\}.
\end{aligned}
$$

The second part of Theorem 3.2.3 shows that the first term in the summand goes to zero as $\nu \to \infty$, since F_{a_ν} converges to zero in C^0 (uniformly in s). Similarly, the second term in the summand goes to zero by Proposition 3.1.9. This verifies (4.14) and proves that v_∞ has Lagrangian boundary conditions away from the boundary shadows $\hat{S}_k \cap \mathbb{R} \times \{0,1\}$. Since this set is finite and away from it the map v_∞ is holomorphic with Lagrangian boundary conditions, we can apply the removal of singularities theorem again to deduce that v_∞ has Lagrangian boundary conditions on all of $\mathbb{R} \times \{0,1\}$. This proves Claim 2.

Claim 3: There exists a smooth lift $\alpha_\infty : \mathbb{R} \times I \to \mathcal{A}^{1,q}_{\mathrm{flat}}(P_\bullet)$ *of* $v_\infty : \mathbb{R} \times I \to M$.

Write $v_\infty = (v^1_\infty, \ldots, v^N_\infty) \in M(P_1) \times \ldots \times M(P_N)$. For each i, the map $\Pi : \mathcal{A}^{1,q}_{\mathrm{flat}}(P_i) \to M(P_i)$ is a principal $\mathcal{G}^{2,q}_0(P_i)$-bundle, and pullback by v^i_∞ provides a bundle over $\mathbb{R} \times I$:

$$\left(v^i_\infty\right)^* \mathcal{A}^{1,q}_{\mathrm{flat}}(P_i) \longrightarrow \mathbb{R} \times I.$$

Sections of this bundle are exactly lifts of v^i_∞. Since the base is contractible, the space of smooth sections of $\left(v^i_\infty\right)^* \mathcal{A}^{1,q}_{\mathrm{flat}}(P_i)$ is homotopic to the space of smooth sections of the $\mathcal{G}^{2,q}_0(P_i)$-bundle over a point. The latter is clearly non-empty, and so the former is non-empty as well. This proves Claim 3.

By exactly the same type of argument, the boundary conditions of v_∞, together with the injectivity of the embeddings $M(Q_{i(i+1)}) \hookrightarrow M(P_i) \times M(P_{i+1})$ (see Theorem 2.2.16) provide the data for a smooth lift $a_\infty : \mathbb{R} \to \mathcal{A}^{1,q}_{\text{flat}}(Q_\bullet)$. Then α_∞ and a_∞ are unique up to $\mathcal{G}_0^{2,q}$-gauge transformation, and it is possible to choose them so that they match up along the seam. Then by Lemma 2.3.11, the 0-forms $\phi_\infty, \psi_\infty : \mathbb{R} \times I \to W^{1,q}(P_\bullet(\mathfrak{g}))$ and $p_\infty : \mathbb{R} \to W^{1,q}(Q_\bullet(\mathfrak{g}))$ are all uniquely determined by α_∞, a_∞ via items (i) and (ii) in the conclusion of the proposition. These patch together to form a continuous connection $A_\infty \in \mathcal{A}^{1,q}_{loc}(\mathbb{R} \times Q)$. It remains to prove items (iii), (iv) and (v). We save (iii) for the end.

Item (iv) follows by transferring the C^0 convergence of the v_ν to a statement about the α_ν. To spell this out, first note that, because M is finite-dimensional, we can choose any metric we want. To prove (iv), it is convenient to choose the metric induced from the C^0 norm on the harmonic spaces (recall that the tangent space to M at $[\alpha] \in M$ can be identified with the harmonic space H^1_α). In particular, the C^0 convergence of the v_ν to v_∞ immediately implies that, for each (s,t), there are gauge transformations $u_\nu(s,t) \in \mathcal{G}_0^{2,q}(P_\bullet)$ such that

$$\sup_K \sum_i \|u_\nu^* \mathrm{NS}_{P_i}(\alpha_\nu) - \alpha_\infty\|^2_{C^0(\Sigma_i)} \longrightarrow 0, \tag{4.15}$$

for all compact $K \subset \mathbb{R} \times I$ (here the gauge action is the point-wise action of $u_\nu(s,t)$). By perturbing the gauge transformations, we may suppose that each u_ν is smooth as a map into $\mathcal{G}_0^{2,q}(P_\bullet)$. This gives

$$\begin{aligned}
\|u_\nu^*\alpha_\nu - \alpha_\infty\|^2_{C^0(K\times\Sigma_\bullet)} \quad &\leq \quad 4\ \sup_K \sum_i \Big\{ \|u_\nu^*\alpha_\nu - u_\nu^*\mathrm{NS}_{P_i}(\alpha_\nu)\|^2_{C^0(\Sigma_i)} \\
&\qquad\qquad + \|u_\nu^*\mathrm{NS}_{P_i}(\alpha_\nu) - \alpha_\infty\|^2_{C^0(\Sigma_i)} \Big\} \\
&\leq \quad C\ \sup_K \sum_i \Big\{ \|F_{\alpha_\nu}\|^2_{C^0(\Sigma_i)} \\
&\qquad\qquad + \|u_\nu^*\mathrm{NS}_{P_i}(\alpha_\nu) - \alpha_\infty\|^2_{C^0(\Sigma_i)} \Big\} \\
\text{(By (4.15))} \quad &\longrightarrow \quad 0,
\end{aligned}$$

where the last inequality follows from Proposition 3.1.9. This proves (iv).

Claim 4: For each compact $\hat{K} \subset \hat{X}$, there is a uniform $W^{3,2}$-bound

$$\sup_{\nu} \|v_\nu\|_{W^{3,2}(\hat{K})} \leq C.$$

Before proving Claim 4, we show how it is used to prove item (v) in the proposition. First of all, by the compact Sobolev embedding $W^{3,2} \hookrightarrow C^1$, Claim 4 implies that, after passing to a subsequence, the v_ν converge to v_∞ in C^1 on compact subsets of $\hat{X}$. In particular $\partial_s v_\nu$ converges to $\partial_s v_\infty$ in C^0 on compact subsets. At this point, the proof is formally much like the proof of item (iv), except we use Corollary 3.1.13 instead of Proposition 3.1.9. Explicitly, this is done as follows: The appropriate lift of $\partial_s v_\infty$ to $T_{\alpha_\infty}\mathcal{A}^{1,q}(P_\bullet)$ is $\beta_{s,\infty}$, since it is the harmonic projection of $\partial_s \alpha_\infty$ (see Lemma 2.3.11). Similarly, the appropriate lift of $\partial_s v_\nu$ is the *harmonic projection* of $D_{\alpha_\nu}\left(\mathrm{NS}_{P_i}\right)(\partial_s \alpha_\nu)$. This harmonic projection is exactly

$$D_{\alpha_\nu}\left(\Pi \circ \mathrm{NS}_{P_i}\right)(\partial_s \alpha_\nu).$$

since $D_\alpha \Pi = \mathrm{proj}_\alpha$ whenever α is flat. Moreover, the image of d_{α_ν} is in the kernel of $D_{\alpha_\nu}\left(\Pi \circ \mathrm{NS}_{P_i}\right)$, so we may equally well use

$$D_{\alpha_\nu}\left(\Pi \circ \mathrm{NS}_{P_i}\right)(\beta_{s,\nu})$$

as the harmonic lift of $\partial_s v_\nu$. Then the C^0 convergence $\partial_s v_\nu \to \partial_s v_\infty$ implies that

$$\sup_{\hat{K}} \sum_i \|\beta_{s,\infty} - \mathrm{Ad}(u_\nu) D_{\alpha_\nu}\left(\Pi \circ \mathrm{NS}_{P_i}\right)(\beta_{s,\nu})\|^2_{L^2(\Sigma_i)} \xrightarrow{\nu} 0, \tag{4.16}$$

for any fixed compact $\hat{K} \subset \hat{X}$ (we have chosen to use the metric given by the L^2 norm on Σ, though any L^p norm with $2 \leq p < \infty$ would work just as well). The gauge transformations that appear here are exactly the ones from the proof of item (iv), and arise from the fact, since $u_\nu^* \alpha_\nu$ converge to α_∞, the harmonic spaces $\mathrm{Ad}(u_\nu(s,t)) H_{\alpha_\nu(s,t)}$ converge to $H_{\alpha_\infty(s,t)}$. For each $(s,t) \in \hat{K}$ and i, the triangle inequality gives

$$\begin{aligned}
\left\|\beta_{s,\infty} - \mathrm{Ad}(u_\nu)\mathrm{proj}_{\alpha_\nu}\beta_{s,\nu}\right\|_{L^2(\Sigma_i)} &\leq \left\|\beta_{s,\infty} - \mathrm{Ad}(u_\nu) D_{\alpha_\nu}\left(\Pi \circ \mathrm{NS}_{P_i}\right)(\beta_{s,\nu})\right\|_{L^2(\Sigma_\bullet)} \\
&\quad + \left\| D_{\alpha_\nu}\left(\Pi \circ \mathrm{NS}_{P_i}\right)(\beta_{s,\nu}) - \mathrm{proj}_{\alpha_\nu}(\beta_{s,\nu})\right\|_{L^2(\Sigma_\bullet)} \\
&\leq \left\|\beta_{s,\infty} - \mathrm{Ad}(u_\nu) D_{\alpha_\nu}\left(\Pi \circ \mathrm{NS}_{P_i}\right)(\beta_{s,\nu})\right\|_{L^2(\Sigma_\bullet)} \\
&\quad + f(\alpha_\nu)\left\|\mathrm{proj}_{\alpha_\nu}(\beta_{s,\nu})\right\|_{L^2(\Sigma)}
\end{aligned}$$

where the last inequality is Corollary 3.1.13. That corollary states that $f(\alpha_\nu) \to 0$ uniformly in s, t, as $\nu \to 0$. Recall that we have assumed in the hypotheses that $\|\mathrm{proj}_{\alpha_\nu}(\beta_{s,\nu})\|_{L^2(\Sigma)}$ is uniformly bounded by some constant C, so summing over i and taking the supremum over $(s,t) \in \hat{K}$ gives

$$\begin{aligned}
&\sup_{\hat{K}} \textstyle\sum_i \left\|\beta_{s,\infty} - \mathrm{Ad}(u_\nu)\mathrm{proj}_{\alpha_\nu}\beta_{s,\nu}\right\|^2_{L^2(\Sigma_i)} \\
&\qquad \leq \sup_{\hat{K}} \textstyle\sum_i \left\|\beta_{s,\infty} - \mathrm{Ad}(u_\nu) D_{\alpha_\nu}\left(\Pi \circ \mathrm{NS}_{P_i}\right)(\beta_{s,\nu})\right\|^2_{L^2(\Sigma_\bullet)} + f(\alpha_\nu)^2 C.
\end{aligned}$$

This goes to zero by (4.16), and proves (v).

To prove Claim 4, and thereby complete the proof of item (v), we need to bound all mixed partial derivatives of v_ν up to degree 3. Since v_ν is holomorphic, this reduces to finding uniform bounds for the first, second and third s-derivatives of v_ν. Writing $v_\nu = \Pi \circ \mathrm{NS}_P(\alpha_\nu)$, we want to translate derivatives on v_ν into derivatives on $\alpha_\nu \in \mathcal{A}^{1,q}(P_\bullet)$. Note that we have the freedom to choose a convenient representative in the gauge equivalence class of A_ν. The components of $\partial_s v_\nu$ in $M(P_\bullet) = M(P_1) \times \ldots \times M(P_N)$ are given by

$$D_{\alpha_\nu}\left(\Pi \circ \mathrm{NS}_{P_i}\right)(\partial_s \alpha_\nu) = D_{\alpha_\nu}\left(\Pi \circ \mathrm{NS}_{P_i}\right)(\beta_{s,\nu}),$$

where the equality holds because $\partial_s \alpha_\nu - \beta_{s,\nu}$ is an exact 1-form and so lies in the kernel of the linearization $D_{\alpha_\nu}\left(\Pi \circ \mathrm{NS}_{P_i}\right)$. By Theorem 3.1.1 (iii), the L^2-operator norms of the operators $D_{\alpha_\nu}(\Pi \circ \mathrm{NS}_{P_i})$ are uniformly bounded, so to bound $|\partial_s v_\nu|_M$ it suffices to bound $\|\beta_{s,\nu}\|_{L^2}$. However, this is a component of the energy of A_ν and so is uniformly bounded a priori:

$$\begin{aligned}
\|\beta_{s,\nu}\|^2_{L^2(\hat{K}\times\Sigma_\bullet)} &= \|\beta_{s,\nu}\|^2_{L^2(\hat{K}\times\Sigma_\bullet),\epsilon_\nu} \\
&\leq \|F_{A_\nu}\|^2_{L^2(\hat{K}\times\Sigma_\bullet),\epsilon_\nu} \\
&\leq \|F_{A_\nu}\|^2_{L^2(\mathbb{R}\times Y),\epsilon_\nu} \\
&= 2\left(\mathcal{CS}_{a_0}(a^-) - \mathcal{CS}_{a_0}(a^+)\right).
\end{aligned} \tag{4.17}$$

(For more details, see the proof of Corollary 4.1.3.) This provides uniform bounds on the first derivatives of the v_ν.

The second derivatives are similar: By the product rule, the second s-derivative of the components of v_ν are controlled by the following term:

$$\begin{aligned}
&D^2_{\alpha_\nu}\left(\Pi\circ \mathrm{NS}_{P_i}\right)\left(\beta_{s,\nu},\partial_s\alpha_\nu\right) + D_{\alpha_\nu}\left(\Pi\circ \mathrm{NS}_{P_i}\right)\left(\partial_s\beta_{s,\nu}\right) \\
&\qquad = D^2_{\alpha_\nu}\left(\Pi\circ \mathrm{NS}_{P_i}\right)\left(\beta_{s,\nu},\beta_{s,\nu}\right) + D_{\alpha_\nu}\left(\Pi\circ \mathrm{NS}_{P_i}\right)\left(\partial_s\beta_{s,\nu}\right)
\end{aligned} \tag{4.18}$$

where the equality holds because $\Pi\circ \mathrm{NS}_{P_i}$ is gauge-equivariant, and so $\mathrm{Im}(d_\alpha)$ lies in the kernel of the Hessian $D^2_\alpha(\Pi\circ \mathrm{NS}_{P_i})$. Consider the Hessian term in (4.18). Lemma 3.1.12 says that the L^2-operator norms of the operators $D^2_{\alpha_\nu}\left(\Pi\circ \mathrm{NS}_{P_i}\right)$ are uniformly bounded, and so bounding the Hessian terms reduces to bounding

$$\|\beta_{s,\nu}\|_{L^2(\hat{K}\times\Sigma_\bullet)},$$

which we have already done in (4.17). To prove Claim 4, it therefore suffices to bound the other term in (4.18):

$$D_{\alpha_\nu}\left(\Pi\circ \mathrm{NS}_{P_i}\right)\left(\partial_s\beta_{s,\nu}\right).$$

As before, since $D_{\alpha_\nu}\left(\Pi\circ \mathrm{NS}_{P_i}\right)$ is uniformly L^2-bounded as an operator, it suffices to bound $\|\partial_s\beta_{s,\nu}\|_{L^2}$. For this, we exploit the gauge freedom and assume that $\phi_\nu = 0$ (i.e.,

A_ν is in temporal gauge). Then $\partial_s = \nabla_s$, and it suffices to bound

$$\|\partial_s \beta_s\|_{L^2(\hat{K}\times\Sigma_i)} = \|\nabla_s \beta_s\|_{L^2(\hat{K}\times\Sigma_i)}$$

That this is uniformly bounded is exactly the content of Corollary 4.1.3. So we have bounded the second s-derivatives of the v_ν. The bound for the third s-derivatives is similar. This completes the proof of Claim 4, and so also the proof of item (v) in Theorem 4.2.1.

It remains to prove item (iii) and the energy bound (4.11). Note that all of the analysis we have done so far remains valid. The only catch is that it may be the case that the limiting connection A_∞ is constant (so all of the energy in the sequence escapes to $\pm\infty$). To get around this, we need to translate by just the right values to make sure that the energy does not escape. This type of result is standard, and we include a sketch for completeness. See [38, Proposition 4.2] for more details. Before defining appropriate real numbers s_ν used for this translation, we make a few preliminary remarks.

First of all, the hypotheses on the A_ν are translation invariant. That is, we may replace A_ν by $\tau^*_{s_\nu} A_\nu$, for *any* real number $s_\nu \in \mathbb{R}$, and the proof up until this point goes through without a problem.

Second, since we have assumed that all flat connections on Q are non-degenerate, it follows from Theorem 2.3.1 that $M(Q)$, the moduli space of flat connections on $Q \to Y$, is a *finite* set of points. By Proposition 2.4.3, this moduli space maps onto the set of Lagrangian intersection points $L_{(0)} \cap L_{(1)} \subset M(P_\bullet)$ (see also Remark 2.3.9 (a)). In particular, the set $L_{(0)} \cap L_{(1)}$ is finite. Let $\epsilon_0 > 0$ be small enough so that the ϵ_0-balls centered at the p_i are mutually disjoint: $B_{\epsilon_0}(p) \cap B_{\epsilon_0}(p') = \emptyset$, for distinct $p, p' \in L_{(0)} \cap L_{(1)}$.

Lastly, note that for any $s_\nu \in \mathbb{R}$, the map NS_{P_i} commutes with translation:

$$\mathrm{NS}_{P_i}(\tau^*_{s_\nu}\alpha_\nu) = \tau^*_{s_\nu}\mathrm{NS}_{P_i}(\alpha_\nu) = \tau^*_{s_\nu} v_\nu.$$

By assumption, each A_ν converges at $\pm\infty$, modulo gauge transformation, to the flat connection $a^\pm$. Since the maps NS_{P_i} preserve flat connections, it follows that each v_ν

converges, as $s \to \pm\infty$, to the intersection point

$$p^{\pm} := \Psi([a^{\pm}]) \in L_{(0)} \cap L_{(1)}$$

associated to $a^{\pm}$, where $\Psi : M(Q)/H_{\eta} \to L_{(0)} \cap L_{(1)}$ is the map from Proposition 2.4.3 which identifies the generators of the Floer groups. The same holds true if we replace A_{ν} by $\tau_{s_{\nu}}^{*} A_{\nu}$ for any $s_{\nu} \in \mathbb{R}$. That is, $\tau_{s_{\nu}}^{*} v_{\nu}$ converges to $p^{\pm}$ as $s \to \pm\infty$, for any $s_{\nu} \in \mathbb{R}$.

With these remarks out of the way, define s_{ν} as follows:

$$s_{\nu} := \sup \left\{ s \in \mathbb{R} \,\middle|\, \mathrm{dist}_{M(P_{\bullet})} \left(p^{-},\ v_{\nu}(s,t)\right) > \epsilon_0, \qquad \text{for some } t \in I \right\}.$$

Then for each ν we have

$$\mathrm{dist}_{M(P_{\bullet})} \left(p^{-},\ \left(\tau_{s_{\nu}}^{*} v_{\nu}\right)(s,t)\right) = \epsilon_0 \qquad \forall t \in I, \forall s \leq 0, \tag{4.19}$$

and

$$\mathrm{dist}_{M(P_{\bullet})} \left(\left(\tau_{s_{\nu}}^{*} v_{\nu}\right)(0,t), p^{+}\right) = \epsilon_0 \qquad \forall t \in I. \tag{4.20}$$

Then the discussion above shows that, after passing to a subsequence, the translates $\tau_{s_{\nu}}^{*} v_{\nu}$ converge to a limiting holomorphic strip v^1, and this must satisfy

$$\lim_{s \to -\infty} v^1(s,t) = p^{-} =: p^0$$

for all t, by (4.19). The equalities expressed in (4.20) and the definition of ϵ_0, show that $v^1(0,t)$ is *not* at a Lagrangian intersection point. In particular, v^1 is non-constant. Since v^1 is a holomorphic strip, there is some intersection point $p^1 \in L_{(0)} \cap L_{(1)}$ with

$$\lim_{s \to \infty} v^1(s,t) = p^1,$$

(see section 2.3). This also means that the v_{ν} become arbitrarily close to p^1.

Next, repeat this procedure with p^1 replacing p^0. It follows that a suitable translation of the v_{ν} converge (after passing to yet another subsequence) to holomorphic strip

v^2 which limits to p^1 at $-\infty$. It must also limit to some other intersection point p^2 at $+\infty$. Moreover, due to the C^1 convergence, it follows that we have an energy bound

$$E(v^2;\hat{K}) \leq \limsup_{\nu} E(v_\nu;\hat{K}) - E(v^1;\hat{K})$$

for every compact $\hat{K} \subset \hat{X}$ (see [37, Corollary 3.4]). Here we are using $E(v,\hat{K}) := \frac{1}{2}\int_{\hat{K}} |\partial_s v|_M^2$ to denote the energy on the set $\hat{K}$.

Continuing inductively, we obtain a sequence of intersection points p^j and non-constant holomorphic strips v^j with

$$\lim_{s\to-\infty} v^j(s,t) = p^{j-1}, \qquad \lim_{s\to\infty} v^j(s,t) = p^j.$$

and

$$\sum_j E(v^j;\hat{K}) \leq \limsup_{\nu} E(v_\nu;\hat{K}) \tag{4.21}$$

for all compact $\hat{K} \subset \hat{X}$. This shows that there can only be a finite number of strips v^j: Each non-constant holomorphic strip carries a minimum allowable energy $\hbar > 0$.[3] The right-hand side of (4.21) is bounded by $\mathcal{CS}_{a_0}(a^-) - \mathcal{CS}_{a_0}(a^+)$, the total energy of the instantons, so there can only be finitely many terms on the left-hand side.

Finally, lift each intersection point p^j to a flat connection $a^j \in \mathcal{A}_{\mathrm{flat}}(Q)$, and lift each holomorphic strip v^j to a connection A^j as we did when we defined A_∞ above. The energy bound (4.11) follows immediately from (4.21).

□

4.3 Proof of the Main Theorem

In this section we carry out the proof of the Main Theorem 2.4.1. Our overall strategy is to show that the hypotheses of Theorem 4.2.1 are satisfied with each $S_k = \emptyset$ empty, at which point the Main Theorem essentially follows immediately. This will be carried

[3] By reflecting each strip one can compute that this is one-half the minimal amount for holomorphic disks and one-quarter the amount for holomorphic spheres; see [30, Proposition 4.1.4] and the proof of [30, Lemma 4.3.1 (ii)].

out in two stages. The first is to show that if the hypotheses of Theorem 4.2.1 are not satisfied, then we have *energy quantization* (defined below). The second is to show that any energy quantization is excluded a priori by the hypotheses of the Main Theorem.

Throughout this discussion, we fix the following notation: Let $Q \to Y$ be as in Remark 2.2.7, and we decompose $Y = Y_\bullet \cup (I \times \Sigma_\bullet)$ as a union of cylindrical cobordisms

$$I \times \Sigma_\bullet = I \times \cup_i \Sigma_i$$

and non-trivial cobordisms

$$Y_\bullet = \cup_i Y_{i(i+1)}.$$

We write $Q_\bullet := Q|_{Y_\bullet}$ and $P_\bullet := Q|_{\{0\}\times\Sigma_\bullet}$.

For a connection A on $\mathbb{R} \times Q$ we will use the notation

$$A = \begin{cases} \alpha + \phi\, ds + \psi\, dt & \text{on } \mathbb{R} \times I \times \Sigma_\bullet \\ a + p\, ds & \text{on } \mathbb{R} \times Y_\bullet \end{cases} \tag{4.22}$$

and

$$\beta_s = \partial_s \alpha - d_\alpha \phi, \qquad \beta_t = \partial_t \alpha - d_\alpha \psi, \qquad \gamma = \partial_s \psi - \partial_t \phi - [\psi, \phi]$$

$$b_s = \partial_s a - d_a p.$$

We use analogous notation for the components of $A_\nu, \widetilde{A}_\nu, \overline{A}_\nu$, and $\hat{A}_\nu$, which will appear later on in our analysis. See Remark 2.3.6. It will be notationally convenient to consider ϕ, ψ, p as being defined on the whole manifold $\mathbb{R} \times Y$ by extending them by zero. Likewise, we consider F_α and F_a as being defined on all of $\mathbb{R} \times Y$ as being extended by zero. Note, however, that these extensions are typically *not* continuous.[4]

[4]Though we will not be using this perspective, perhaps it is worth pointing out that we can think of these extension of F_α (resp. F_a) as being the curvature of some α' (resp. a) where $\alpha'|_{\{(s,t)\}\times\Sigma_\bullet} = \alpha(s,t)$ (resp. $\alpha'|_{\{s\}\times Y_\bullet} = a(s)$) and is a flat connection on each $\{s\} \times Y_\bullet$ (resp. $\{(s,t)\} \times \Sigma_\bullet$).

4.3.1 Energy quantization (bubbling analysis)

As described above, our goal is to show that the hypotheses of Theorem 4.2.1 are satisfied. In the current section, we show that if these hypotheses do not hold, then this implies **energy quantization**: there is a subset $S \subset \mathbb{R} \times Y$ and a positive constant $\hbar > 0$, depending only on the group $\mathrm{PSU}(r)$, such that for every neighborhood U of S, the energy

$$\frac{1}{2} \int_U \langle F_A \wedge *_\epsilon F_A \rangle \geq \hbar$$

is bounded from below. Here $*_\epsilon$ is the Hodge star defined with respect to the ϵ-dependent metric. In our cases, the set S will either be a point or a slice $\{s_0\} \times Y_{i(i+1)}$ or $\{(s_0, t_0)\} \times \Sigma_i$, as described in the beginning of section 4.2.

Fix limiting flat connections $a^\pm$ and a real number $q > 2$. For sake of contradiction, suppose the existence statement in the Main Theorem 2.4.1 does not hold. Then there is a positive number $\delta_0 > 0$, a sequence of positive numbers $\epsilon_\nu \to 0$, and, for each ν, an ϵ_ν-ASD connection A_ν descending to the zero-dimensional moduli space $\widehat{M}_{Q^{\epsilon_\nu},0}([a^-], [a^+])$, and such that

$$\|\alpha_\nu - \alpha_\infty\|_{C^0(K \times \Sigma_\bullet)} \geq \delta_0$$

for all $A_\infty \in \mathcal{A}^{1,q}(\mathbb{R} \times Q)$ satisfying (i) and (ii) from the Main Theorem 2.4.1 and for some compact $K \subset \mathbb{R} \times I$. Since $q > 2$, by applying a suitable gauge transformation in $\mathcal{G}^{2,q}_{loc}(\mathbb{R} \times Q^{\epsilon_\nu})$, we may suppose that A_ν is ϵ_ν-smooth [51, Theorem 9.4]. (Equivalently, we may assume that $(F^{\epsilon_\nu})^* A_\nu$ is smooth with the standard smooth structure on $\mathbb{R} \times Y$, where F^ϵ is the map (2.5).)

Let $\mathrm{proj}_{\alpha_\nu}$ be as in Lemma 3.1.5. Observe that if the curvatures

$$\|F_{\alpha_\nu}\|_{L^\infty(\mathbb{R} \times Y)} + \|F_{a_\nu}\|_{L^\infty(\mathbb{R} \times Y)} \longrightarrow 0$$

converge to zero, and if the term

$$\sup_{\mathbb{R}\times I} \|\mathrm{proj}_{\alpha_\nu}\beta_{s,\nu}\|_{L^2(\Sigma_\bullet)} \leq C$$

is uniformly bounded, then Theorem 4.2.1 applies (with the S_k all empty), and provides us with a contradiction. It therefore suffices to rule out each of the following cases:

Case 1 $\|F_{\alpha_\nu}\|_{L^\infty(\mathbb{R}\times Y)} + \|F_{a_\nu}\|_{L^\infty(\mathbb{R}\times Y)} \to \infty$;

Case 2 $\|F_{\alpha_\nu}\|_{L^\infty(\mathbb{R}\times Y)} + \|F_{a_\nu}\|_{L^\infty(\mathbb{R}\times Y)} \to \Delta > 0$;

Case 3 $\|F_{\alpha_\nu}\|_{L^\infty(\mathbb{R}\times Y)} + \|F_{a_\nu}\|_{L^\infty(\mathbb{R}\times Y)} \to 0$,

but $\sup_{\mathbb{R}\times I} \|\mathrm{proj}_{\alpha_\nu}\beta_{s,\nu}\|_{L^2(\Sigma_\bullet)} \to \infty$.

We rule these out by showing that each leads to energy quantization. In section 4.3.2, below, we will show that any energy quantization leads to a contradiction.

Case 1: Instantons on S^4.

In this case we identify a point in $\mathbb{R}\times Y$ where the curvature diverges (a *blow-up point*). We will then conformally rescale in a small neighborhood of this point to show that a non-trivial instanton on S^4 bubbles off. The energy of such instantons cannot be arbitrarily small, so this implies energy quantization for this case. Here are the details:

By passing to a subsequence, we may assume the L^∞-norm of each curvature is always achieved on Σ_i or $Y_{i(i+1)}$ for some i (same i for all ν):

$$\|F_{\alpha_\nu}\|_{L^\infty(\mathbb{R}\times I\times\Sigma_{i+1})} + \|F_{a_\nu}\|_{L^\infty(\mathbb{R}\times Y_{i(i+1)})} = \|F_{\alpha_\nu}\|_{L^\infty(\mathbb{R}\times Y)} + \|F_{a_\nu}\|_{L^\infty(\mathbb{R}\times Y)}.$$

Find points $(s_\nu, t_\nu) \in \mathbb{R}\times I$ with

$$\|F_{\alpha_\nu(s_\nu,t_\nu)}\|_{L^\infty(\Sigma_{i+1})} + \|F_{a_\nu(s_\nu)}\|_{L^\infty(Y_{i(i+1)})} = \|F_{\alpha_\nu}\|_{L^\infty(\mathbb{R}\times Y)} + \|F_{a_\nu}\|_{L^\infty(\mathbb{R}\times Y)}.$$

This discussion is translation invariant, so we may suppose $s_\nu = 0$. Also, by passing to a subsequence we may suppose $t_\nu \to t_\infty \in I$ converges. We need to distinguish the sub-cases when the blow-up point occurs in the cylindrical part $I \times \Sigma_i$, and when it occurs on the non-trivial part $Y_{i(i+1)}$:

Sub-case 1 For all but finitely many ν we have

$$\|F_{\alpha_\nu(s_\nu,t_\nu)}\|_{L^\infty(\Sigma_{i+1})} \geq \|F_{a_\nu(s_\nu)}\|_{L^\infty(Y_{i(i+1)})} \tag{4.23}$$

and $t_\infty \neq 0, 1$;

Sub-case 2 For all but finitely many ν we have

$$\|F_{\alpha_\nu(s_\nu,t_\nu)}\|_{L^\infty(\Sigma_{i+1})} \leq \|F_{a_\nu(s_\nu)}\|_{L^\infty(Y_{i(i+1)})} \tag{4.24}$$

or (4.23) holds and $t_\infty = 0, 1$.

Without loss of generality, we may suppose $i = 1$ and $t_\infty \in [0, 1)$.

In Sub-case 1, define rescaled connections $\widetilde{A}_\nu$ in terms of its components as follows:

$$\begin{aligned}
\widetilde{\alpha}_\nu(s,t) &:= \alpha(\epsilon_\nu s, \epsilon_\nu t + t_\nu)|_{\Sigma_2} \\
\widetilde{\phi}_\nu(s,t) &:= \epsilon_\nu \phi(\epsilon_\nu s, \epsilon_\nu t + t_\nu)|_{\Sigma_2} \\
\widetilde{\psi}_\nu(s,t) &:= \epsilon_\nu \psi(\epsilon_\nu s, \epsilon_\nu t + t_\nu)|_{\Sigma_2}
\end{aligned} \tag{4.25}$$

which we view as a connection defined on $B_{\epsilon_\nu^{-1}\eta} \times \Sigma_2 \subseteq \mathbb{C} \times \Sigma_2$, where $\eta = \frac{1}{2}\min\{t_\infty, 1 - t_\infty\}$, $B_r \subset \mathbb{C}$ is the ball of radius r centered at zero, and we assume ν is large enough so $t_\nu \leq \eta$.

In Sub-case 2 we take

$$
\begin{aligned}
\widetilde{a}_\nu(s) &:= a_\nu(\epsilon_\nu s)|_{Y_{12}} \\
\widetilde{p}_\nu(s) &:= \epsilon_\nu p_\nu(\epsilon_\nu s)|_{Y_{12}} \\
\widetilde{\alpha}_\nu(s,t) &:= \begin{cases} \alpha(\epsilon_\nu s, -\epsilon_\nu t + 1)|_{\Sigma_1} & \text{on } \Sigma_1 \\ \alpha(\epsilon_\nu s, \epsilon_\nu t)|_{\Sigma_2} & \text{on } \Sigma_2 \end{cases} \\
\widetilde{\phi}_\nu(s,t) &:= \begin{cases} \epsilon_\nu \phi(\epsilon_\nu s, -\epsilon_\nu t + 1)|_{\Sigma_1} & \text{on } \Sigma_1 \\ \epsilon_\nu \phi(\epsilon_\nu s, \epsilon_\nu t)|_{\Sigma_2} & \text{on } \Sigma_2 \end{cases} \\
\widetilde{\psi}_\nu(s,t) &:= \begin{cases} \epsilon_\nu \psi(\epsilon_\nu s, -\epsilon_\nu t + 1)|_{\Sigma_1} & \text{on } \Sigma_1 \\ \epsilon_\nu \psi(\epsilon_\nu s, \epsilon_\nu t)|_{\Sigma_2} & \text{on } \Sigma_2 \end{cases}
\end{aligned}
\tag{4.26}
$$

which we view as a connection defined on $\mathbb{R} \times Y_{12}(\epsilon_\nu^{-1})$, where we are using the following notation:[5]

$$
X(r) := X \cup_{\partial X} [0, r) \times \partial X, \qquad X^\infty := X \cup_{\partial X} [0, \infty) \times \partial X. \tag{4.27}
$$

for a smooth manifold X with boundary ∂X and for $r > 0$.

Remark 4.3.1. *There exists smooth structures on these spaces which are compatible in the sense that the inclusions*

$$
X(r) \subseteq X(r') \subseteq X^\infty \tag{4.28}
$$

are smooth embeddings for $r \leq r'$. If X has a metric g, then we will consider the metric on $X(r)$ and X^∞ which is given by g on X and $dt^2 + g|_{\partial X}$ on the end. In particular, the embeddings (4.28) become metric *embeddings. This will be called the* fixed metric *on the given manifold, and we denote its various norms by $|\cdot|, \|\cdot\|_{L^p}$, etc. If X is*

[5]Strictly speaking, for this case we need to only consider these connections as being defined on $\mathbb{R} \times \left(Y_{12} \cup_{\Sigma_2} \left[0, \epsilon_\nu^{-1}\right) \times \Sigma_2\right)$. However, the next case requires that we consider them on the larger space $\mathbb{R} \times Y_{12}(\epsilon_\nu^{-1})$, so we do so here in an attempt to better streamline the discussion.

equipped with a bundle $B \to X$ then we define bundles $B(r) \to X(r)$ and $B^\infty \to X^\infty$ in the obvious way. Note also that we have the following decomposition

$$\mathbb{R} \times X^\infty = (\mathbb{R} \times X) \cup (\mathbb{H} \times \partial X)\,. \tag{4.29}$$

In both Sub-cases, by construction, the connections $\widetilde{A}_\nu$ are ASD with respect to the fixed metric, and have uniformly bounded energy

$$\frac{1}{2}\left\|F_{\widetilde{A}_\nu}\right\|_{L^2}^2 \leq \mathcal{CS}(a^-) - \mathcal{CS}(a^+).$$

Here the norm should be taken on the domain on which the connection is defined. Furthermore, the energy densities are bounded from below:

$$\begin{aligned} \|F_{\widetilde{A}_\nu}\|_{L^\infty} &\geq \|F_{\widetilde{\alpha}_\nu}\|_{L^\infty} + \|F_{\widetilde{a}_\nu}\|_{L^\infty} \\ &= \|F_{\alpha_\nu}\|_{L^\infty} + \|F_{a_\nu}\|_{L^\infty}. \end{aligned} \tag{4.30}$$

In particular, the condition of Case 1 implies that

$$\|F_{\widetilde{A}_\nu}\|_{L^\infty} \longrightarrow \infty.$$

Following the usual rescaling argument [41] [12, Section 9] (see also [30, Theorem 4.6.1] for the closely-related case of J-holomorphic curves) we can conformally rescale in a small neighborhood U of the blow-up point to obtain a sequence of finite energy instantons with energy density bounded by 1 and defined on increasing balls in $\mathbb{R}^4$. By Uhlenbeck's Strong Compactness Theorem 2.2.10, there is a subsequence that converges, modulo gauge, in C^∞ in all derivatives to a finite energy non-constant instanton $\widetilde{A}_\infty$ on $\mathbb{R}^4$. By Uhlenbeck's removable singularities theorem this extends to non-constant instanton, also denoted by $\widetilde{A}_\infty$, on a PSU(r)-bundle $R \to S^4$. Since $\widetilde{A}_\infty$ is ASD and non-constant we have

$$\begin{aligned} 0 < \frac{1}{2}\int_{S^4} |F_{\widetilde{A}_\infty}|^2 &= \frac{1}{2}\int_{S^4} \langle F_{\widetilde{A}_\infty} \wedge *F_{\widetilde{A}_\infty}\rangle \\ &= -\frac{1}{2}\int_{S^4} \langle F_{\widetilde{A}_\infty} \wedge F_{\widetilde{A}_\infty}\rangle \\ &= q_4(R) \in \mathbb{Z} \end{aligned}$$

is the PSU(r)-characteristic class. So $1 \leq \|F_{\widetilde{A}_\infty}\|^2_{L^2(S^4)} = \lim_{\nu\to\infty} \|F_{A_\nu}\|^2_{L^2(U),\epsilon_\nu}$ for every neighborhood U of the blow-up point. In particular, we have energy quantization for this case with $\hbar = 1$.

Case 2: Instantons on non-compact domains.

This case is much the same as the previous, in that instantons near the blow-up point bubble off. However, this time the geometry of the underlying spaces on which these bubbles form can be more exotic. The key ingredient used to show that we have energy quantization for these domains is Proposition 4.3.4, below.

Define $\widetilde{A}_\nu$ exactly as in Case 1 above. Everything up to and including equation (4.30) continues to hold. In particular,

$$\liminf \|F_{\widetilde{A}_\nu}\|_{L^\infty} \geq \Delta > 0.$$

After possibly passing to a subsequence, the energy densities converge to some $\Delta' \in [\Delta, \infty]$:

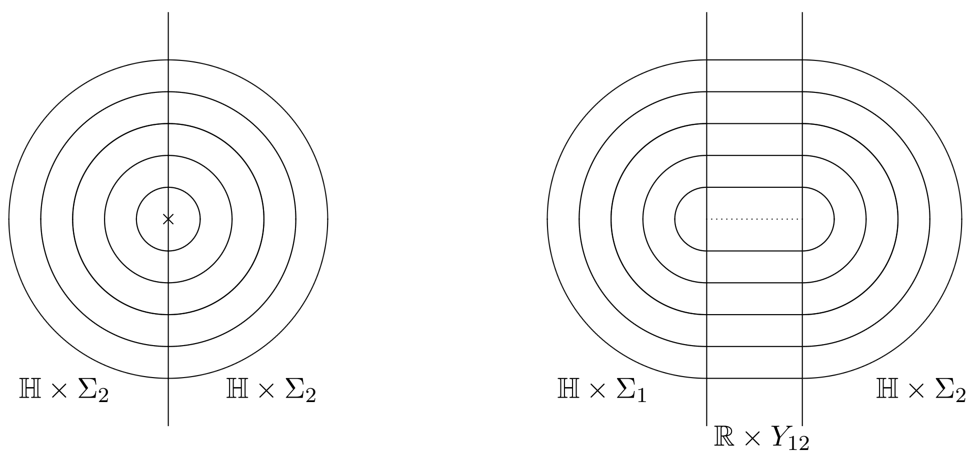
$$\|F_{\widetilde{A}_\nu}\|_{L^\infty} \longrightarrow \Delta'.$$

If $\Delta' = \infty$ then we are done by precisely the same analysis as in Case 1. So we may suppose $0 < \Delta' < \infty$, in which case we can apply Uhlenbeck's Strong Compactness Theorem 2.2.10 directly to the sequence $\widetilde{A}_\nu$, which therefore converges to a non-flat finite-energy instanton $\widetilde{A}_\infty$ on a bundle over one of the spaces $\mathbb{R} \times Y_{12}^\infty$ or $\mathbb{C} \times \Sigma_2$, depending on where they blow-up occurs (see the discussion above Remark 4.3.1 for a definition of Y_{12}^∞). We need to show that there is a minimum allowable energy

$$\int_{\mathbb{R}\times Y_{12}^{\infty}} |F_A|^2 \geq \hbar > 0$$

for all non-flat instantons A on bundles over the domains $\mathbb{C} \times \Sigma_2$ and $\mathbb{R} \times Y_{12}^{\infty}$.[6]

We begin by briefly recalling from [12] the case $\mathbb{C} \times \Sigma_2$. This motivates the basic approach for the (notationally more cumbersome) case $\mathbb{R} \times Y_{12}^{\infty}$. Fix a non-flat finite energy instanton A on a bundle over $\mathbb{C} \times \Sigma_2$. The basic idea is to introduce polar coordinates on the $\mathbb{C}$ component in $\mathbb{C} \times \Sigma_2$. This allows us to view A as being defined on the cylinder $\mathbb{R} \times S^1 \times \Sigma_2 \cong \mathbb{C}\backslash \{0\} \times \Sigma_2$. By the usual argument (see section2.3), the finite energy instanton A limits to flat connections on the cylindrical ends $S^1 \times \Sigma_2$, at $\pm\infty$, and the energy of A is given by the difference of the Chern-Simons functional applied to each of these limiting flat connections. To show the energy of non-flat instantons is bounded from zero, it suffices to show that the Chern-Simons functional only obtains discrete values. We prove this in Proposition 4.3.4, below. Also see Remark 4.3.5.



(a) Polar coordinates on $\mathbb{C} \times \Sigma_2$.

(b) Polar coordinates on $\mathbb{R} \times Y_{12}^{\infty}$

In the case of instantons on $\mathbb{R} \times Y_{12}^{\infty}$, we want to do a similar thing by identifying "polar coordinates" on $\mathbb{R} \times Y_{12}^{\infty}$, which will allow us to view this manifold as a cylinder

[6]There is nothing special about the particular subscripts in Σ_2 and Y_{12} here. Indeed, these can be replaced by any compact oriented surface-without-boundary Σ or 3-manifold-with-boundary Y, respectively.

$\mathbb{R}\times X$ for some closed oriented 3-manifold X (to do this we will need to "cut" $\mathbb{R}\times Y_{12}^{\infty}$ in a way analogous to removing the origin when we make the identification $\mathbb{C}\setminus\{0\} \cong \mathbb{R}\times S^1$). These "polar coordinates" arise by exploiting the decomposition (4.29). In particular, since $\partial Y_{12} = \overline{\Sigma_1} \sqcup \Sigma_2$ is disconnected, we can write

$$\mathbb{R} \times Y_{12}^{\infty} = \left(\mathbb{H} \times \overline{\Sigma_1}\right) \cup (\mathbb{R} \times Y_{12}) \cup (\mathbb{H} \times \Sigma_2)\,.$$

The middle slice $\mathbb{R} \times Y_{12}$ plays the same role here as $\{y\text{-axis}\} \times \Sigma_2$ did in the previous case.

Example 4.3.2. *Suppose $Y_{12} = I \times \Sigma_2$ is a cylinder. If we imagine letting the volume on I go to zero, then we recover exactly the case $\mathbb{C} \times \Sigma_2$ described above.*

The "polar coordinates" we use will restrict to the usual polar coordinates on each copy of $\mathbb{H}$, patched together along the middle strip $\mathbb{R} \times Y_{12}$. Explicitly, we define these coordinates as follows: Begin by writing a connection A in the usual Cartesian coordinates as

$$A = \begin{cases} a + p\, ds, & \mathbb{R} \times Y_{12} \\ \alpha + \phi\, ds + \psi\, dt & \mathbb{H} \times \Sigma_1 \cup \Sigma_2 \end{cases}$$

On $\mathbb{H}$ the polar coordinates are $(s,t) = (e^{\tau}\cos(\theta), e^{\tau}\sin(\theta))$, with $\tau \in \mathbb{R}$ and $\theta \in [0,\pi]$. In these coordinates the connection A takes the form

$$A|_{\mathbb{H}\times\Sigma_1\cup\Sigma_2} = \overline{\alpha}(\tau,\theta) + \overline{\phi}(\tau,\theta)\, d\tau + \overline{\psi}(\tau,\theta)\, d\theta.$$

The relationship between the two coordinate expressions can be written in matrix notation as

$$\begin{pmatrix} \overline{\alpha}(\tau,\theta) \\ \overline{\phi}(\tau,\theta) \\ \overline{\psi}(\tau,\theta) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{\tau}\cos(\theta) & e^{\tau}\sin(\theta) \\ 0 & -e^{\tau}\sin(\theta) & e^{\tau}\cos(\theta) \end{pmatrix} \begin{pmatrix} \alpha(e^{\tau}\cos(\theta), e^{\tau}\sin(\theta)) \\ \phi(e^{\tau}\cos(\theta), e^{\tau}\sin(\theta)) \\ \psi(e^{\tau}\cos(\theta), e^{\tau}\sin(\theta)) \end{pmatrix} \tag{4.31}$$

The ASD relations in polar coordinates become

$$\partial_\tau \overline{\alpha} - d_{\overline{\alpha}} \overline{\phi} + * \left(\partial_\theta \overline{\alpha} - d_{\overline{\alpha}} \overline{\psi} \right) = 0$$

$$\partial_\tau \overline{\psi} - \partial_\theta \overline{\phi} - \left[\overline{\psi}, \overline{\phi} \right] + e^{2\tau} * F_{\overline{\alpha}} = 0. \tag{4.32}$$

The energy takes the form

$$\begin{aligned} \frac{1}{2} \| F_A \|^2_{L^2(\mathbb{R} \times Y_{12}^\infty)} &= \int_{-\infty}^{\infty} \int_0^\pi \Big\{ \| \partial_\tau \overline{\alpha} - d_{\overline{\alpha}} \overline{\phi} \|^2_{L^2(\Sigma_1 \sqcup \Sigma_2)} \\ &\qquad\qquad + \tfrac{1}{2} \left(1 + e^{2\tau} \right) \| F_{\overline{\alpha}} \|^2_{L^2(\Sigma_1 \sqcup \Sigma_2)} \Big\} \, d\theta \, d\tau \\ &\quad + \int_{-\infty}^{\infty} \| F_a \|^2_{L^2(Y_{12})} \, ds \end{aligned}$$

We then view $\mathbb{R} \times Y_{12}^\infty \backslash \{0\} \times Y_{12}$ as a cylinder in the following way: The domain $\mathbb{R} \times Y_{12}^\infty \backslash \{0\} \times Y_{12}$, is foliated by a family of closed connected 3-manifolds, where each leaf of this foliation is diffeomorphic to[7]

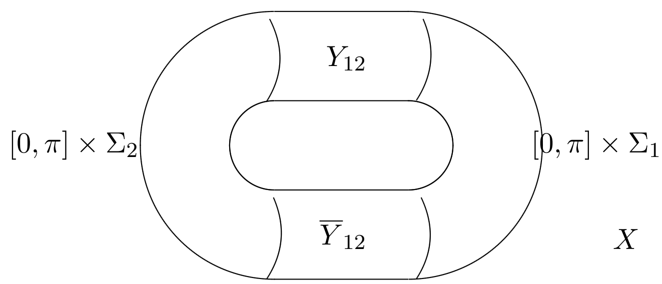
$$X := \overline{Y}_{12} \cup_{\Sigma_1} \left([0, \pi] \times \Sigma_1 \right) \cup_{\Sigma_1} Y_{12} \cup_{\Sigma_2} \left([0, \pi] \times \Sigma_2 \right) \cup_{\Sigma_2} . \tag{4.33}$$

This is a *cyclic* decomposition, so the right side wraps around and glues onto the left side, just as in (2.2). Then polar coordinates provide a map

$$\iota : \mathbb{R} \times X \hookrightarrow \mathbb{R} \times Y_{12}^\infty,$$

which is an embedding onto the complement of $\{0\} \times Y_{12}$ in $\mathbb{R} \times Y_{12}^\infty$ (the $\mathbb{R}$ in $\mathbb{R} \times X$ is the radial parameter τ, whereas the $\mathbb{R}$ in $\mathbb{R} \times Y_{12}^\infty$ is the parameter s).

[7]Strictly speaking, X is only a C^1-manifold. It is perhaps easiest to see this by reducing the dimension. So suppose $Y_{12} = I$ is an interval, which means we take the Σ_i to be points. Then X is obtained by gluing the end points of two semi-circles to the endpoints of two line segments. The result is a topological circle. Consider the tangent lines to the semi-circles near the endpoints. These tangent lines approach a horizontal line at the endpoints, but only in a C^0 sense. So we only get a C^1 manifold X. To get a C^∞ manifold, one would need to perturb the polar coordinates slightly so that these tangent lines converge in C^∞. We leave the details to the reader.



Example 4.3.3. *As in example 4.3.2, suppose $Y_{12} = I \times \Sigma_2$ is the trivial cobordism. Then $\mathbb{R} \times Y_{12}^{\infty} = \mathbb{C} \times \Sigma_2$, and $X = S^1 \times \Sigma_2$. In this case, the foliation is provided by a family of embeddings $S^1 \times \Sigma_1 \hookrightarrow \mathbb{C} \times \Sigma_1$, defined by embedding $S^1 \hookrightarrow \mathbb{C}$ as a circle of some radius $\tau > 0$.*

In our situation, Y_{12} is not a trivial cobordism, but the discussion of the previous example carries through. Namely, the polar coordinates on $\mathbb{H} \times \Sigma_1 \cup \Sigma_2$ provide a family of embeddings $\iota_\tau : X \hookrightarrow \mathbb{R} \times Y_{12}^{\infty}$. For each $\tau > 0$, the restriction of A to the image of ι_τ defines a continuous connection $\overline{a}(\tau)$ on the closed oriented 3-manifold X. In terms of the polar coordinates defined above, we can write

$$\overline{a}(\tau) = \begin{cases} a(-e^{\tau}) & \text{on } \overline{Y}_{12} \\ \overline{\alpha}(\tau, \cdot) + \overline{\psi}(\tau, \cdot)\, d\theta|_{\Sigma_1} & \text{on } [0,\pi] \times \Sigma_1 \\ a(e^{\tau}) & \text{on } Y_{12} \\ \overline{\alpha}(\tau, \cdot) + \overline{\psi}(\tau, \cdot)\, d\theta|_{\Sigma_2} & \text{on } [0,\pi] \times \Sigma_2 \end{cases}$$

Then if A is an instanton on $\mathbb{R} \times Y_{12}^{\infty}$, we have

$$\begin{aligned} \frac{1}{2}\int_{\mathbb{R}\times Y_{12}^{\infty}} |F_A|^2 &= \int_{-\infty}^{\infty} \int_X \langle F_{\overline{a}(\tau)} \wedge \partial_\tau \overline{a}(\tau)\rangle \, d\tau \\ &= \int_{-\infty}^{\infty} \Big\{ \int_X \langle d_{\overline{a}_0}(\overline{a}(\tau) - \overline{a}_0) \wedge \partial_\tau \overline{a}(\tau)\rangle \\ &\qquad + \frac{1}{2}\langle [\overline{a}(\tau) - \overline{a}_0 \wedge \overline{a}(\tau) - \overline{a}_0] \wedge \partial_\tau \overline{a}(\tau)\rangle \Big\} \, d\tau \\ &= \int_{-\infty}^{\infty} \frac{d}{d\tau} \mathcal{CS}_{\overline{a}_0}(\overline{a}(\tau)) \, d\tau \end{aligned} \tag{4.34}$$

for any fixed reference connection $\overline{a}_0$ on X, which we may assume is flat.

We want to be able to say that there are limiting flat connections $\overline{a}^{\pm}$ over X with

$$\lim_{\tau \to \pm\infty} \overline{a}(\tau) = \overline{a}^{\pm}.$$

This would be immediate if A pulled back under

$$\iota : \mathbb{R} \times X \hookrightarrow \mathbb{R} \times Y_{12}^{\infty}$$

to an instanton on $\mathbb{R} \times X$ with respect to the *usual* metric on $\mathbb{R}$. However, since A is an instanton on $\mathbb{R} \times Y_{12}^{\infty}$ with respect to a product metric $ds^2 + g$, its pullback $\iota^* A$ is an instanton on $\mathbb{R} \times X$ with respect to the pulled back metric

$$e^{2\tau} d\tau^2 + g' = \iota^*(ds^2 + g),$$

and this is enough to conclude that $\overline{a}(\tau)$ limits to a flat connection at $+\infty$. Indeed, consider the restrictions

$$\overline{A}_n := \iota^* A|_{[n,n+1] \times X}$$

for n large. View each $\overline{A}_n$ as being defined on $[0,1] \times X$. Then, with respect to the fixed metric $d\tau^2 + g'$ on $[0,1] \times X$, this satisfies

$$\begin{aligned} \|F_{\overline{A}_n}\|^2_{L^2([0,1]\times X), d\tau^2+g'} &= \|F_{\overline{a}}\|^2_{L^2([n,n+1]\times X), d\tau^2+g'} + \|\partial \overline{a} - d_{\overline{a}}\overline{p}\|^2_{L^2([n,n+1]\times X), d\tau^2+g'} \\ &\leq e^{-n}\left(\|F_{\overline{a}}\|^2_{L^2, d\tau^2+g'} + \|\partial \overline{a} - d_{\overline{a}}\overline{p}\|^2_{L^2, \iota^*(ds^2+g)} \right) \\ &\leq e^{-n} \|F_A\|_{L^2, ds^2+g}. \end{aligned}$$

Since A has finite energy, this shows $\|F_{\overline{A}_n}\|^2_{L^2([0,1]\times X), d\tau^2+g'}$ converges to zero. By Uhlenbeck's weak compactness theorem, it follows that a subsequence of the $\overline{A}_n$ converges, after composing with suitable gauge transformations, to a flat connection $\overline{A}^+$ on $[0,1] \times X$ as $n \to \infty$. Write $\overline{A}^+ = \overline{a}^+ + \overline{p}^+ d\tau$. If we assume that A is in radial gauge (i.e., $d\tau$-component equal to zero) then $\overline{p}^+ = 0$, and $\overline{a}^+$ is independent of the τ variable. By this same argument, every subsequence of $\overline{A}_n$ has subsequence which converges to a flat connection. If all flat connection on X are non-degenerate (i.e., isolated, modulo gauge), then these limiting flat connections must all be the same, and so the full sequence $\overline{A}_n$ converges to flat connection $\overline{a}^+$, after composing with a suitable sequence of gauge transformations. Since the $\overline{A}_n$ are restrictions of the same connection, these gauge transformations can be chosen to piece together to form a single continuous gauge transformation on $[0, \infty) \times X$. This shows that $\overline{a}(\tau)$ converges

$$\lim_{\tau \to \infty} \overline{a}(\tau) = \overline{a}^+$$

to some flat connection.

At first glance, this argument breaks down when one tries to repeat the process as $\tau \to -\infty$ (since $e^{-\tau} > 1$ for $\tau << 0$, our energy bound from before fails). Geometrically, however, there should be no issue at $-\infty$ because this corresponds to the cut $\{0\} \times Y_{12}$, and A extends continuously over the cut. There are various ways to deal with this issue, and we follow an argument similar to the one given by Wehrheim in [48]. The idea is to put A in a suitable gauge, and to find a suitable reference connection $\overline{a}_0$, for which

$$\lim_{\tau \to -\infty} \mathcal{CS}_{\overline{a}_0}(\overline{a}(\tau)) = 0.$$

We assume that A is in a gauge such that $\overline{\psi} = 0$ on $\mathbb{H} \times \Sigma_1 \sqcup \Sigma_2$, at least for $\tau << 0$. This is a 'θ-temporal gauge'. (It may not be possible to do this for large τ since above we needed A to be in radial (τ-temporal) gauge for large τ.) We assume that $\overline{a}_0$ satisfies a similar condition: Over $[0, \pi] \times \Sigma_1 \sqcup \Sigma_2 \subset X$ we write our flat connection in coordinates as $\overline{a}_0 = \overline{\alpha}_0 + \overline{\psi}_0 \, d\theta$. Then by acting on it with a suitable gauge transformation, we may assume $\overline{\psi}_0 = 0$.

Setting $\mu(s) = a(s) - \overline{a}_0|_{Y_{12}}$, we have

$$\begin{aligned}\mathcal{CS}_{a_0}(\overline{a}(\tau)) &= \frac{1}{2}\int_{\overline{Y}_{12}} \langle d_{\overline{a}_0}\mu(\tau)\wedge\mu(-e^{\tau})\rangle + \frac{1}{3}\langle[\mu(-e^{\tau})\wedge\mu(-e^{\tau})]\wedge\mu(-e^{\tau})\rangle \\ &+\frac{1}{2}\int_{Y_{12}} \langle d_{\overline{a}_0}\mu(e^{\tau})\wedge\mu(e^{\tau})\rangle + \frac{1}{3}\langle[\mu(e^{\tau})\wedge\mu(e^{\tau})]\wedge\mu(e^{\tau})\rangle \\ &+\int_0^{\pi}\int_{\Sigma_1\sqcup\Sigma_2} \frac{1}{2}\langle\partial_\theta\overline{\alpha}(\tau,\theta)\wedge\overline{\alpha}(\tau,\theta)-\overline{\alpha}_0(\theta)\rangle + \langle F_{\overline{\alpha}(\tau,\theta)},\overline{\psi}(\tau,\theta)\rangle\, d\theta\end{aligned}$$

The first two integrals on the right cancel as $\tau \to -\infty$, because A extends continuously over $\{0\} \times Y_{12}$. Likewise, the third term goes to zero as well since we have assumed $\overline{\psi}(\tau,\theta) = 0$, and (4.31) shows that $\partial_\theta\overline{\alpha}(\tau,\theta)$ converges to zero as τ approaches $-\infty$.

We therefore have

$$\frac{1}{2}\int_{\mathbb{R}\times Y_{12}^{\infty}} |F_A|^2 = \mathcal{CS}_{\overline{a}_0}(\overline{a}^+)$$

for some flat connections $\overline{a}_0, \overline{a}^+$. If A is not flat, then Proposition 4.3.4 below shows that there is some $\hbar > 0$ (depending only on the bundle) with

$$\frac{1}{2}\int_{\mathbb{R}\times Y_{12}^{\infty}} |F_A|^2 \geq \hbar.$$

This finishes case 2.

Proposition 4.3.4. *Let Y be a closed, connected, oriented 3-manifold, and $Q \to Y$ a principal* $\mathrm{PSU}(r)$*-bundle. Fix a reference connection $a_0 \in \mathcal{A}^{1,2}(Q)$. Then when restricted to the flat connections, the Chern-Simons functional $\mathcal{CS}_{a_0} : \mathcal{A}^{1,2}_{\text{flat}}(Q) \to \mathbb{R}$ obtains only discrete values. Moreover, the set*

$$\left\{\mathcal{CS}_{a_0}(a)\,\middle|\,a, a_0 \in \mathcal{A}^{1,2}_{\text{flat}}(Q)\right\}$$

is discrete.

Remark 4.3.5. *In the case where the group is* $\mathrm{PSU}(2)$ *and Y is the mapping torus of some surface diffeomorphism, Dostoglou and Salamon showed the stronger result that $\mathcal{CS}_{a_0}$ only takes on values in $C_{\mathrm{PSU}(2)}\mathbb{Z}$ [12]. In [48], Wehrheim extended this to other*

Lie groups (but still only for mapping tori) by showing that each flat connection can be gauge transformed to one with Chern-Simons value zero.

Proof of Proposition 4.3.4. We prove this in three steps.

Step 1: The path-connected components of $\mathcal{A}^{1,2}_{\mathrm{flat}}(Q)$ *are the connected components.*

It suffices to show that $\mathcal{A}^{1,2}_{\mathrm{flat}}(Q)$ is locally path-connected. The Yang-Mills heat flow provides a nice proof of this. For $i = 0, 1$, let $a_i \in \mathcal{A}^{1,2}_{\mathrm{flat}}(Q)$. We want to show that if a_0 and a_1 are close enough, then they are connected by a path in $\mathcal{A}^{1,2}_{\mathrm{flat}}(Q)$. Consider the straight-line path $\gamma(t) = a_0 + t(a_1 - a_0)$, which lies in $\mathcal{A}^{1,2}(Q)$. Then

$$F_{\gamma(t)} = t d_{a_0}(a_1 - a_0) + \frac{t^2}{2}\left[a_1 - a_0 \wedge a_1 - a_0\right],$$

and so

$$\begin{aligned} \|F_{\gamma(t)}\|_{L^2} &\leq \|d_{a_0}(a_1 - a_0)\|_{L^2} + \|a_1 - a_0\|^2_{L^4} \\ &\leq C\left(\|a_1 - a_0\|_{W^{1,2}} + \|a_1 - a_0\|^2_{W^{1,2}}\right) \end{aligned}$$

since $\|a_1 - a_0\|_{L^4} \leq C'\|a_1 - a_0\|_{W^{1,2}}$ by Sobolev embedding. Let ϵ_Q be the constant from (3.38), and take $\delta_Q := \min\{1, \epsilon_Q/2C\}$. Then $\gamma(t)$ is in the realm of the heat flow map, Heat_Q, whenever $\|a_1 - a_0\|_{W^{1,2}} < \delta_Q$. In particular,

$$a_t := \mathrm{Heat}_Q(\gamma(t)) \in \mathcal{A}^{1,2}_{\mathrm{flat}}(Q)$$

is a continuous path from a_0 to a_1, as desired (this uses the fact that Heat_Q is the identity on $\mathcal{A}^{1,2}_{\mathrm{flat}}(Q)$).

Step 2: The restriction of the Chern-Simons functional

$$\mathcal{CS}_{a_0} : \mathcal{A}^{1,2}_{\mathrm{flat}}(Q) \longrightarrow \mathbb{R}$$

is locally constant. In particular, it descends to a map

$$\mathcal{CS}_{a_0} : M(Q) = \mathcal{A}^{1,2}_{\mathrm{flat}}(Q)/\mathcal{G}^{2,2}_0(Q) \longrightarrow \mathbb{R}.$$

Suppose $a : I \to \mathcal{A}^{1,2}_{\text{flat}}(Q)$ is a smooth path. Set $\gamma(t) = a(t) - a_0$. Then we have

$$\begin{aligned}
\tfrac{d}{dt}\mathcal{CS}_{a_0}(a(t)) &= \frac{1}{2}\frac{d}{dt}\int_Y 2\langle F_{a_0} \wedge \gamma(t)\rangle + \langle d_{a_0}\gamma(t) \wedge \gamma(t)\rangle + \frac{1}{3}\langle[\gamma(t) \wedge \gamma(t)] \wedge \gamma(t)\rangle \\
&= \frac{1}{2}\int_Y \{2\langle F_{a_0} \wedge \gamma'(t)\rangle + \langle d_{a_0}\gamma'(t) \wedge \gamma(t)\rangle + \langle d_{a_0}\gamma(t) \wedge \gamma'(t)\rangle \\
&\qquad + \langle[\gamma(t) \wedge \gamma(t)] \wedge \gamma'(t)\rangle\} \\
&= \frac{1}{2}\int_Y \langle F_{a(t)} \wedge \gamma'(t)\rangle \\
&= 0.
\end{aligned}$$

Hence $\mathcal{CS}_{a_0}(a(0)) = \mathcal{CS}_{a_0}(a(1))$.

Step 3: For each gauge transformation $u \in \mathcal{G}^{2,2}(Q)$ (not necessarily in the identity component), and each flat connection $a \in \mathcal{A}^{1,2}_{\text{flat}}(Q)$, we have

$$\mathcal{CS}_{a_0}(u^*a) - \mathcal{CS}_{a_0}(a) = C_{\mathrm{PSU}(r)} \deg(u),$$

for some constant $C_{\mathrm{PSU}(r)}$ depending only on $\mathrm{PSU}(r)$ *and choice of metric on its Lie algebra.*

To see this, first define the bundle

$$Q_u := I \times Q/(\{0\}, q) \sim (\{1\}, u(q)).$$

This is a bundle over $S^1 \times Y$. Here we identify $S^1 = \mathbb{R}/\mathbb{Z}$. The path of connections

$$a(\cdot) : t \longmapsto a + t(u^*a - a)$$

descends to give a connection A on Q_u. From the definitions, we have

$$\begin{aligned}
\mathcal{CS}_{a_0}(u^*a) - \mathcal{CS}_{a_0}(a) &= \int_0^1 \int_Y \langle F_{a(t)} \wedge \partial_t a(t) \rangle \, dt \\
&= \int_{S^1 \times Y} \langle F_A \wedge F_A \rangle \\
&= C_{\mathrm{PSU}(r)} q_4(Q_u) \\
&= C_{\mathrm{PSU}(r)} \deg(u),
\end{aligned}$$

where characteristic class q_4 is as in section 2.2.1.

Combining steps 2 and 3 shows that $\mathcal{CS}_{a_0}$ descends to a map

$$\widetilde{\mathcal{CS}}_{a_0} : M(Q)/\pi_0(\mathcal{G}^{2,2}(Q)) \longrightarrow S^1 = \mathbb{R}/C_G\mathbb{Z}$$

which is locally constant. Therefore, it only depends on the number of connected components of $M(Q)/\pi_0(\mathcal{G}^{2,2}(Q))$. By Remark 2.2.17, the space is $M(Q)/\pi_0(\mathcal{G}^{2,2}(Q))$ is compact, so there are only finitely many components. This shows that $\widetilde{\mathcal{CS}}_{a_0}$ only obtains finitely many values, and so $\mathcal{CS}_{a_0}$ only obtains discrete values. This proves the first part of the proposition.

The second part is similar: Notice that it follows from the definition that

$$\mathcal{CS}_{a_0}(a) = -\mathcal{CS}_a(a_0),$$

and so if a_0 and a are both flat and u is a gauge transformation, then

$$\mathcal{CS}_{u^*a_0}(a) - \mathcal{CS}_{a_0}(a) = \mathcal{CS}_a(a_0) - \mathcal{CS}_a(u^*a_0) = C_{\mathrm{PSU}(r)} \deg(u).$$

So we again obtain a well-defined map

$$M(Q) \times M(Q) \longrightarrow \mathbb{R}$$

sending $([a], [a_0])$ to $\mathcal{CS}_{a_0}(a)$, and as before this descends to a circle-valued map

$$M(Q)/\pi_0(\mathcal{G}^{2,2}) \times M(Q)/\pi_0(\mathcal{G}^{2,2}) \longrightarrow S^1.$$

Since $M(Q)/\pi_0(\mathcal{G}^{2,2}) \times M(Q)/\pi_0(\mathcal{G}^{2,2})$ is compact, we are done by the same argument as before.

□

Case 3: Holomorphic spheres and disks in $M(P)$.

In this case, we rescale around the blow-up point to find that a holomorphic sphere or disk bubbles off in the moduli space of flat connections. However, these rescaled connections do not satisfy a fixed ASD equation, and so Uhlenbeck's Strong Compactness Theorem does not apply, as it did in Cases 1 and 2. One could try to use Uhlenbeck's Weak Compactness, but it is too weak to conclude that the bubbles are non-constant, as we were able to do easily in the previous two cases. The technique of Dostoglou and Salamon [13] applies here to show that the holomorphic *spheres* are non-constant. However, due to the presence of boundaries, it is not clear how to extend their analysis to apply to holomorphic disks. In our approach, the non-triviality of the holomorphic disks stems directly from item (v) in Theorem 4.2.1 (or, rather, a variation of this proposition suited to other domains, as in Remark 4.2.2 (a)).

The condition of this case implies that

$$c_\nu := \sup_{\mathbb{R}\times I} \|\mathrm{proj}_{\alpha_\nu}\beta_{s,\nu}\|_{L^2(\Sigma_i)} \longrightarrow \infty$$

for some i. Find points $(s_\nu, t_\nu) \in \mathbb{R} \times I$ with

$$\|\mathrm{proj}_{\alpha_\nu}\beta_{s.\nu}(s_\nu, t_\nu)\|_{L^2(\Sigma_i)} = c_\nu.$$

(Such points exist since $\beta_{s,\nu}$ decays at ∞, due to the finite energy; alternatively, one could replace c_ν by $c_\nu/2$, without changing the argument.) We may translate so that $s_\nu = 0$, and pass to a subsequence so that $t_\nu \to t_\infty \in I$ converges. The two relevant subcases to consider are as follows:

Sub-case 1 $t_\infty \in (0,1)$

Sub-case 2 $t_\infty \in \{0,1\}$

We may assume, without loss of generality, that $i = 2$ and, if Sub-case 2 holds, that $t_\infty = 0$. Define rescaled connections $\hat{A}_\nu$ using (4.25) and (4.26), except with every ϵ_ν replaced by c_ν^{-1}, and with the Sub-cases here replacing the ones in Case 1. We will prove we have energy quantization in Sub-case 2, by showing a holomorphic disk bubbles off. Sub-case 1 is similar, but we get a holomorphic sphere instead and we leave the details to the reader.

The rescaling for Sub-case 2 is such that we view the connections $\hat{A}_\nu$ as being defined on $\mathbb{R} \times Y_{12}(c_\nu)$ (see the discussion above Remark 4.3.1). The components of $F_{\hat{A}_\nu}$ satisfy

$$\hat{\beta}_{s,\nu} + *\hat{\beta}_{t,\nu} = 0, \qquad \hat{\gamma} = -\hat{\epsilon}_\nu^{-2} * F_{\hat{\alpha}_\nu}, \qquad \hat{b}_{s,\nu} = -\hat{\epsilon}_\nu^{-1} F_{\hat{a}_\nu},$$

where $\hat{\epsilon}_\nu := c_\nu \epsilon_\nu$. It may not be the case that $\hat{\epsilon}_\nu$ is decaying to zero; this is replaced by the assumption in this case that the slice-wise curvatures converge to zero in L^∞:

$$\|F_{\hat{\alpha}_\nu}\|_{L^\infty} = \|F_{\alpha_\nu}\|_{L^\infty}, \qquad \|F_{\hat{a}_\nu}\|_{L^\infty} = \|F_{a_\nu}\|_{L^\infty}.$$

We also have

$$\|\mathrm{proj}_{\hat{\alpha}_\nu} \hat{\beta}_{s,\nu}(0,0)\|_{L^2(\Sigma_2)} = 1. \tag{4.35}$$

Then exactly the same proof as in Theorem 4.2.1 (see Remark 4.2.2 (a)) shows that, after possibly passing to a subsequence, there exists a sequence of gauge transformations $u_\nu \in \mathcal{G}^{2,q}_{loc}(\mathbb{H} \times P_\bullet)$, and a limit connection $\hat{A}_\infty \in \mathcal{A}^{1,q}_{loc}(\mathbb{R} \times Q^\infty_{12})$ satisfying

(i) $\hat{\beta}_{s,\infty} + *\hat{\beta}_{t,\infty} = 0$

(ii) $F_{\hat{\alpha}_\infty} = 0, F_{\hat{a}_\infty} = 0$

(iii) $\sup_K \|\mathrm{Ad}(u_\nu)\mathrm{proj}_{\hat{\alpha}_\nu}\hat{\beta}_{s,\nu} - \hat{\beta}_{s,\infty}\|_{L^2(\Sigma_\bullet)} \xrightarrow{\nu} 0$

for all compact $K \subset \mathbb{H}$. Let $\Pi_{P_i} : \mathcal{A}_{\text{flat}}(P_i) \to M(P_i)$ and $\Pi_{Q_{12}} : \mathcal{A}_{\text{flat}}(Q_{12}) \to M(Q_{12})$ be the projections to the moduli spaces. Then

$$v_\infty := (\Pi_{P_1}(\hat{\alpha}_\infty|_{\Sigma_1}), \Pi_{P_2}(\hat{\alpha}_\infty|_{\Sigma_2})) : \mathbb{H} \longrightarrow M(P_1) \times M(P_2)$$

is a holomorphic curve with Lagrangian boundary conditions

$$\mathbb{R} \longrightarrow M(Q_{12}) \hookrightarrow M(P_1) \times M(P_2)$$

determined by $a_\infty : \mathbb{R} \to \mathcal{A}^{1,q}_{\text{flat}}(Q_{12})$. Furthermore, v_∞ has bounded energy

$$\begin{aligned} \int_{\mathbb{H}} |\partial_s v_\infty|^2 &= \int_{\mathbb{H}\times\Sigma_1\sqcup\Sigma_2} |\hat{\beta}_{s,\infty}|^2 \\ &\leq \liminf_\nu \|\beta_{s,\nu}\|^2_{L^2(\mathbb{R}\times Y)} \\ &= \liminf_\nu \|\beta_{s,\nu}\|^2_{L^2(\mathbb{R}\times Y),\epsilon_\nu} \\ &\leq \liminf_\nu \|F_{A_\nu}\|^2_{L^2(\mathbb{R}\times Y),\epsilon_\nu} \\ &= 2\left(\mathcal{CS}_{a_0}(a^-) - \mathcal{CS}_{a_0}(a^+)\right) \end{aligned} \tag{4.36}$$

In particular, the removal of singularities theorem [30, Theorem 4.1.2 (ii)] applies and so v_∞ extends to a holomorphic disk $v_\infty : \mathbb{D} \to M(P_1)\times M(P_2)$ with Lagrangian boundary conditions. Condition (iii) above combines with (4.35) to show that

$$\begin{aligned} |\partial_s v_\infty(0,0)| &\geq \|\mathrm{proj}_{\hat{\alpha}_\infty(0,0)}\hat{\beta}_{s,\infty}(0,0)\|_{L^2(\Sigma_\bullet)} \\ &= \lim_{\nu\to\infty} \|\mathrm{proj}_{\hat{\alpha}_\nu(0,0)}\hat{\beta}_{s,\nu}(0,0)\|_{L^2(\Sigma_\bullet)} \\ &= 1. \end{aligned}$$

In particular, v_∞ is non-constant. We have energy quantization for non-constant holomorphic disks [30, Proposition 4.1.4], which takes care of Case 3.

4.3.2 Energy quantization implies the Main Theorem

Here we show that any energy quantization leads to a contradiction. By the assumptions of the Main Theorem, we only consider connections $A = A_\nu$ which satisfy the ϵ-ASD equations and limit to fixed flat connection $a^\pm$ at $\pm\infty$. Any such connection has energy which depends only on the $a^\pm$:

$$\mathcal{YM}_\epsilon(A) = \mathcal{CS}_{a_0}(a^-) - \mathcal{CS}_{a_0}(a^+),$$

where

$$\mathcal{YM}_\epsilon(A) := \frac{1}{2}\int_{\mathbb{R}\times Y} \langle F_A \wedge *_\epsilon F_A \rangle$$

is the total energy, and a_0 is any fixed reference connection on Y. In particular, the energy is finite, so energy quantization can only occur for a finite number of points (from cases 1 and 2) or slices (from case 3). Denote the collection of these points and slices by $\{S_k\}_{k=1}^K$. On the complement of the S_k, the hypotheses of Theorem 4.2.1 are satisfied, and so there is some limiting broken trajectory $\left(A^1, \ldots, A^J\right)$ on $\mathbb{R} \times Y$, and tuple of flat connection $(a^0 = a^-, a^1, \ldots, a^J = a^+)$, where each A^j descends to a holomorphic curve with Lagrangian boundary conditions:

$$v^j : \mathbb{R} \times I \longrightarrow M(P_1) \times \ldots \times M(P_N).$$

Moreover, v^j lifts to a connection on $\mathbb{R} \times Y$ which converges to a^{j-1} and a^j at $-\infty$ and ∞, respectively. By the energy bound (4.11) and the energy quantization, we immediately have

$$\begin{aligned}
\sum_j E(v^j) &= \frac{1}{2}\int_{\mathbb{R}\times I\times\Sigma_\bullet} |\partial_s\alpha_\infty - d_{\alpha_\infty}\phi_\infty|^2 \\
&\leq -K\hbar + \limsup_\nu \mathcal{YM}(A_\nu) \\
&= -K\hbar + (\mathcal{CS}_{a_0}(a^-) - \mathcal{CS}_{a_0}(a^+))
\end{aligned} \tag{4.37}$$

To achieve the desired contradiction, we assume the conjectural index relation

$$\mu_{\mathrm{symp}}(a,a') = \mu_{\mathrm{inst}}(a,a')$$

for flat connections a, a'. Then the index formula (2.31) gives

$$\begin{aligned}
\sum_{j=1}^J \mu_{\mathrm{symp}}(a^{j-1},a^j) &= \sum_{j=1}^J \mu_{\mathrm{inst}}(a^{j-1},a^j) \\
&= \sum_{j=1}^J \frac{1}{2}(\eta_{a^j} - \eta_{a^{j-1}}) + C\left(\mathcal{CS}(a^{j-1}) - \mathcal{CS}(a^j)\right) \\
&= \sum_{j=1}^J \frac{1}{2}(\eta_{a^j} - \eta_{a^{j-1}}) + CE(v_j) \\
&\leq -KC\hbar + \frac{1}{2}(\eta_{a^+} - \eta_{a^-}) + C\mathcal{YM}(A_\nu) \\
&= -KC\hbar + \frac{1}{2}(\eta_{a^+} - \eta_{a^-}) + C\left(\mathcal{CS}(a^-) - \mathcal{CS}(a^+)\right) \\
&= -KC\hbar + \mu_{\mathrm{inst}}(a^-,a^+) \\
&= -KC\hbar + 1,
\end{aligned}$$

If the number of bubbles K is positive then, since μ_{symp} only takes on integer values, this would imply that $\mu_{\mathrm{symp}}(a^{j-1},a^j) \leq 0$ for some j. This is not possible since this would mean that v_j descends to a moduli space of negative dimension (quotienting by $\mathbb{R}$ reduces the dimension by 1). This shows that $K = 0$ and so no bubbling can occur.

Similarly, if $J \geq 2$, then $\mu_{\mathrm{symp}}(a^{j-1}, a^j) \leq 0$ for some j, which cannot happen for the same reason. This shows that the broken trajectory $(A^1, \ldots, A^J)$ consists of a single trajectory A^1, which finishes the proof of Theorem 2.4.1.

References

[1] S. Akbulut, J. McCarthy. *Casson's invariant for oriented homology 3-spheres, an exposition.* Math. Notes, No. 36, Princeton Univ. Press, Princeton, 1990.

[2] M. Atiyah. New invariants of three and four dimensional manifolds. *The Math. Heritage of Hermann Weyl, Proc. Sympos. Pure Math.* 48, 285-299, 1988.

[3] M. Atiyah, R. Bott. The Yang-Mills equations over Riemann surfaces. *Phil. Trans. Roy. Soc. London Ser. A,* 308:523-615, 1982.

[4] M.F. Atiyah, V.K. Patodi and I.M. Singer. Spectral asymmetry and Riemannian geometry III. *Math. Proc. Camb. Phil. Soc.* 79 (1976), 71-99.

[5] R. Bott, L. Tu. *Differential forms in algebraic topology.* 3rd corrected printing, Graduate Texts in Mathematics, Vol. 82, Springer, New York, 1995.

[6] Charalambous. The Yang-Mills heat semigroup on three-manifolds with boundary. 2010. arXiv:1004.1639.

[7] S. Donaldson. Anti-self dual Yang-Mills connections over complex algebraic surfaces and stable vector bundles. *Proc. London Math. Soc.* (3) 50, no. 1, 1-26, 1985.

[8] S. Donaldson. *Floer homology groups in Yang-Mills theory.* Cambridge University Press, Cambridge, 2002. With the assistance of M. Furuta and D. Kotschick.

[9] S. Donaldson, P.B. Kronheimer. *The Geometry of Four-Manifolds*, Clarendon Press, Oxford 1990.

[10] S. Dostoglou, D. Salamon. Cauchy-Riemann operators, self-duality and the spectral flow. *First European Congress of Mathematics, Vol.* I (Paris, 1992), 511-545, Progr. Math., 119, Birkhäuser, Basel, 1994.

[11] S. Dostoglou, D. Salamon. Instanton homology and symplectic fixed points. *Symplectic geometry*, 57-93, *London Math. Soc. Lecture Note Ser.*, 192, Cambridge Univ. Press, Cambridge, 1993.

[12] S. Dostoglou, D. Salamon. Self-dual instantons and holomorphic curves. *Ann. of Math.* (2) 139 (1994), no. 3, 581-640.

[13] S. Dostoglou, D. Salamon. Corrigendum: Self-dual instantons and holomorphic curves. *Ann. of Math.* 165 (2007), 665-673.

[14] A. Floer. An instanton-invariant for 3-manifolds. *Comm. Math. Phys.*, 118(2):215-240, 1988.

[15] A. Floer. Instanton homology and Dehn surgery. *Floer Memorial Volume, Progress in Mathematics, Vol.* 133, Birkhäuser Verlag, 1995. ISBN 3-7643-5044-X, Basel; ISBN 0-8176-5044-8, Boston.

[16] A. Floer. Morse theory for Lagrangian intersections. *J. Differential Geom.* 28 (1988), no. 3, 513-547.

[17] A. Floer. The unregularized gradient flow of the symplectic action. *Comm. Pure Appl. Math.* 41 (1988), 775-813.

[18] A. Floer. Witten's complex and infinite dimensional Morse theory. *J. Differential Geom.* 30 (1989), no. 1, 207-221.

[19] G. Folland. *Real analysis-modern techniques and their applications.* 2nd ed., Wiley, New York, 1999.

[20] K. Fukaya. Floer homology for 3 manifold with boundary I. 1997. Unpublished manuscript, available at `http://www.math.kyoto-u.ac.jp/~fukaya/fukaya.html`.

[21] K. Fukaya. Anti-self-dual equation on 4-manifolds with degenerate metric. *Geom. Funct. Anal.* 8, 466-528 1998.

[22] D. Gay, R. Kirby. Indefinite Morse 2-functions: broken fibrations and generalizations, 2011. arXiv:1102.0750.

[23] S. Grundel. Moment maps and diffeomorphisms. *Diploma Thesis*, ETH Zürich, 2005.

[24] A.R. Gaio, D. Salamon. Gromov-Witten invariants of symplectic quotients and adiabatic limits. *J. Symplectic Geom. Volume* 3, Number 1 (2005), 55-159.

[25] V. Guillemin S. Sternberg. Geometric quantization and multiplicities of group representations. *Invent. Math.*, 67:515-538, 1982.

[26] L. Jeffrey. Extended moduli spaces of flat connections on Riemann surfaces. *Math. Ann.* 298, 667-692 (1994).

[27] F. Kirwan. *Cohomology of quotients in symplectic and algebraic geometry.* Mathematical Notes Series, Vol. 31, Princeton University Press, 1984. ISBN 0691083703, 9780691083704.

[28] Y. Lekili. Heegaard Floer homology of broken fibrations over the circle. arXiv:0903.1773.

[29] J. Milnor. *Lectures on the h-Cobordism Theorem. Notes by L. Siebenmann and J. Sondow.* Princeton University Press, Princeton, N.J., 1965.

[30] D. McDuff, D. Salamon. *J-holomorphic curves and symplectic topology.* American Mathematical Society Colloquium Publications, Vol. 52. American Mathematical Society, Providence, RI, 2004.

[31] T. Mrowka, K. Wehrheim. L^2 topology and Lagrangians in the space of connections over a Riemann surface. *Geo. and Func. Anal.* Vol 20-5 (2010), 1278-1305.

[32] M. Narasimhan, C. Seshadri. Stable and unitary vector bundles on a compact Riemann surface. *Annals of Mathematics. Second Series* 82 (3): 540-567, 1965.

[33] Y.-G. Oh. Floer cohomology of Lagrangian intersections and pseudo-holomorphic disks I. *Comm. Pure Appl. Math.*, 46(7):949-993, 1993.

[34] J. Råde. On the Yang-Mills heat equation in two and three dimensions. *J. Reine Angew.* 431, 123-163 (1992).

[35] A. Ramanathan. Stable Principal bundles on a compact Riemann surface. *Math. Ann.* 213, 29-152 (1975).

[36] D. Salamon. Lagrangian Intersections, 3-Manifolds with Boundary, and the Atiyah-Floer Conjecture. *Proceedings of the ICM*, Zürich 1994, Vol. 1, 526-536.

[37] D. Salamon. Lectures on Floer homology. *Park City Series* 7 (1999), 145-229.

[38] D. Salamon. Morse theory, the Conley index and Floer homology. *Bul l. L.M.S.* 22 (1990), 113-140.

[39] D. Salamon, K. Wehrheim. Instanton Floer homology with Lagrangian boundary conditions. arXiv:math/0607318.

[40] C. Taubes. Casson's invariant and gauge theory. *J. Di. Geom.* 31 (1990), 547-599.

[41] C. Taubes. Self-dual Yang-Mills connections on non-self-dual 4-manifolds. *J. Di. Geom.* 17 (1982), 139-170.

[42] K. Uhlenbeck. Removable singularities in Yang-Mills fields. *Comm. Math. Phys.* 83 (1982), 11-29.

[43] K. Uhlenbeck. Connections with L^p bounds on the curvature. *Comm. Math. Phys.* 83 (1982), 31-42.

[44] F. Warner. *Foundations of differentiable manifolds and Lie groups.* Scott-Foreman, Glen view, Illinois, 1971.

[45] K. Wehrheim. Banach space valued Cauchy-Riemann equations with totally real boundary conditions. *Comm. Contemp. Math.* 6(4) (2004), 601-635.

[46] K. Wehrheim. Anti-self-dual instantons with Lagrangian boundary conditions I: Elliptic theory. *Comm. Math. Phys.* 254(1)(2005),45-89.

[47] K. Wehrheim. Anti-self-dual instantons with Lagrangian boundary conditions II: Bubbling. *Comm. Math. Phys.* 258(2) (2005), 275-315.

[48] K. Wehrheim. Energy identity for anti-self dual instantons on $\mathbb{C} \times \Sigma$. *Math. Res. Lett.* 13 (2006), 161-166.

[49] K. Wehrheim. Energy quantization and mean value inequalities for nonlinear boundary value problems. *J. Eur. Math. Soc.* 7(3) (2005), 305-318.

[50] K. Wehrheim. Lagrangian boundary conditions for anti-self-dual instantons and the Atiyah-Floer conjecture. *J. Symp. Geo.*, Vol. 3, Number 4, 703-747, 2005.

[51] K. Wehrheim. *Uhlenbeck compactness.* EMS Series of Lectures in Mathematics, 2004.

[52] K. Wehrheim, C. Woodward. Floer cohomology and geometric composition of Lagrangian correspondences. arXiv:0905.1368.

[53] K. Wehrheim, C. Woodward. Floer field theory. Preprint.

[54] K. Wehrheim, C. Woodward. Quilted Floer cohomology. *Geometry and Topology*, 14:833-902, 2010. arXiv:0905.1370.

[55] K. Wehrheim, C. Woodward. Erratum: Quilted Floer cohomology. *Geometry and Topology* 16 (2012) 127-154.

[56] L. M. Woodward. The classification of principal PU_n-bundles over a 4-complex. *J. London Math. Soc.* (2) 25 (1982), no. 3, 513-524.

Vita

David Lee Duncan

2013 Ph. D. in Mathematics, Rutgers University

2002-06 B. Sc. in Mathematics from the University of Washington

2002 Graduated from Highline High School

2010-2013 RASTL Fellow, Graduate School, Rutgers University

2011, 2012 Head teaching assistant (summer), Department of Mathematics, Rutgers University

2008-2012 Teaching assistant, Department of Mathematics, Rutgers University

2010, 2012 REU Co-mentor, Department of Mathematics, Rutgers University

2009, 2010 REU Graduate Student Coordinator, Center for Discrete Mathematics and Theoretical Computer Science, Rutgers University