COMPACTNESS RESULTS FOR NECK-STRETCHING LIMITS OF INSTANTONS

DAVID L. DUNCAN

Abstract. We prove that, under a suitable degeneration of the metric, instantons converge to holomorphic quilts. To prove the main results, we develop estimates for the Yang-Mills heat flow on surfaces and cobordisms.

Contents

1. Introduction
2. Background, notation and conventions
   2.1. Gauge theory
   2.2. The complexified gauge group
3. Compactness for products \( S \times \Sigma \)
   3.1. Small curvature connections in dimension 2
   3.2. Proof of Lemma 3.5
   3.3. Proof of Theorem 3.3
4. Compactness for cylinders \( \mathbb{R} \times Y \)
   4.1. The heat flow on cobordisms
   4.2. Uniform elliptic regularity
   4.3. Proof of Lemma 4.2
   4.4. Proof of Theorem 4.1
5. Perturbations
References

1. Introduction

In Donaldson theory one obtains invariants of 4-manifolds \( Z \) by counting instantons on a suitable bundle over \( Z \). On the other hand, in symplectic geometry one can obtain invariants of \( Z \) by counting \( J \)-holomorphic curves in a suitable symplectic manifold associated to \( Z \). For example, when \( Z = S \times \Sigma \) is a product of surfaces, then the relevant \( J \)-holomorphic curves are maps from \( S \) into the moduli space of flat connections on \( \Sigma \). One expects that these invariants are the same, an expectation that arises due to the following heuristic dating back to Atiyah [1]. Consider the case \( Z = S \times \Sigma \), and fix a metric in which the \( S \)-fibers are very large (equivalently, the \( \Sigma \)-fibers are very small). Motivated by Atiyah’s terminology, we refer to this type of metric as a ‘neck-stretching metric’, and use \( \epsilon^{-2} \) to denote the volume of \( S \). Then with respect to such a metric, the instanton equation splits into two equations; the first is essentially the holomorphic curve equation and the second shows that the curvature of the instanton is bounded by \( \epsilon^2 \) in the \( \Sigma \)-directions. In particular, in the limit \( \epsilon \to 0 \), the instanton equation formally recovers the \( J \)-holomorphic curve equation.

In this paper we formalize the heuristic of the previous paragraph in the case when \( Z = S \times \Sigma \), and also when \( Z = \mathbb{R} \times Y \), where \( Y \) has positive first Betti number. In each case, we prove that if \( \epsilon_\nu \) is a sequence converging to zero and \( A_\nu \) is an instanton with respect to the \( \epsilon_\nu \)-metric, then a subsequence of the \( A_\nu \) converges in a suitable sense to a \( J \)-holomorphic curve. The main convergence result for
each of the two cases is stated precisely in Theorem 3.3 and Theorem 4.1 respectively. Our primary interest is in the case \( Z = \mathbb{R} \times Y \), due to its relevance in the quilted Atiyah-Floer conjecture; see [8]. We also note that when \( Z = \mathbb{R} \times Y \) our main theorem extends results of Dostoglou-Salamon in [6, 7], where they consider the special case when \( Y \) is a mapping torus. The case \( Z = S \times \Sigma \) serves as a model for the technically more difficult analysis necessary for \( \mathbb{R} \times Y \). The results of this paper can be extended quite naturally to more general 4-manifolds, however we leave a full treatment of such extensions to future work.

The proofs of the main results proceed roughly as follows: As mentioned, when \( \epsilon \) is small each instanton \( A \) has curvature that is small in the \( \Sigma \)-direction. Fixing \( s \in S \), our strategy is to use the Narasimhan-Seshadri correspondence on \( \Sigma \) to map the restriction \( A|_{\{s\} \times \Sigma} \) to some nearby flat connection \( NS|_{\{s\} \times \Sigma} \) on \( \Sigma \). This correspondence preserves the equations in the sense that the limiting connection \( NS|_{\{s\} \times \Sigma} \) is holomorphic. Then the convergence result for the case \( Z = S \times \Sigma \) essentially follows from Gromov’s compactness theorem for holomorphic curves.

In the case when \( Z = \mathbb{R} \times Y \), the situation is considerably more difficult. Due to the particular choice of ‘neck-stretching metric’ considered, we expect that the limiting \( J \)-holomorphic curve will have Lagrangian boundary conditions. The problem is that, though \( NS|_{\{s\} \times \Sigma} \) continues to be holomorphic in this case, it only has approximate Lagrangian boundary conditions. The standard Gromov compactness theorem breaks down when one does not have Lagrangian boundary conditions on the nose. We get around this using the following two ingredients. First, we establish several \( C^1 \)-estimates for the map \( NS \), and these allow us to control the behavior of the holomorphic curves near the boundary. The second ingredient is provided by the Yang-Mills heat flow on 3-manifolds with boundary. This gives us candidates for what the boundary conditions should be, if they were Lagrangian. Combining this with the \( C^1 \)-estimates allows us to reprove a version of Gromov’s compactness theorem for almost Lagrangian boundary conditions.

The organization of the remainder of this paper is as follows. In the next section we introduce our notation and conventions. We begin Section 2 by precisely stating our compactness result for \( S \times \Sigma \). We then discuss the Narasimhan-Seshadri correspondence and develop the necessary estimates. We conclude Section 3 with the proof of the compactness result in the case of \( S \times \Sigma \). Section 4 opens by stating the compactness result for \( \mathbb{R} \times Y \). We then discuss the heat flow on 3-manifolds with boundary, and prove the compactness theorem for \( \mathbb{R} \times Y \).

**Acknowledgments:** The author is grateful to his thesis advisor Chris Woodward for his insight and valuable suggestions. This work was partially supported by NSF grants DMS 0904358 and DMS 1207194.

## 2. Background, notation and conventions

Throughout this section \( X \) will denote an oriented manifold. Given a fiber bundle \( F \to X \), we will use \( \Gamma(F) \) to denote the space of smooth sections. If \( V \to X \) is a vector bundle, then we will write \( \Omega^k(X, V) := \Gamma(A^k TX \otimes V) \) for the space of \( k \)-forms with values in \( V \), and we set
\[
\Omega^*(X, V) := \bigoplus_k \Omega^k(X, V).
\]

If \( V \to X \) is the trivial rank-1 bundle then we will write \( \Omega^k(X) \) for \( \Omega^k(X, V) \).

Given a metric on \( X \) and a connection on \( V \), we can define Sobolev norms on \( \Gamma(V) \) in the usual way. We will use \( W^{k,p}(X, V) \) to denote the completion of \( \Gamma(V) \) with respect to the \( W^{k,p} \)-norm. We note that the usual Sobolev embedding statements for \( W^{k,p}(X, \mathbb{R}) \) hold equally well for \( W^{k,p}(X, V) \) (if \( V \) is infinite-dimensional then we assume \( V \) is a Banach bundle). We will typically not keep track of bundle \( V \) in the notation for the Sobolev norms. For example, we will use the same symbol \( \| \cdot \|_{L^2(X)} \) for the norm on \( \Gamma(T^*X \otimes V) \) as for the norm on \( \Gamma(T^*X \otimes V) \).

Now suppose \( P \to X \) is a principal \( G \)-bundle, where \( G \) is a Lie group with Lie algebra \( \mathfrak{g} \). Given a manifold \( M \) and homomorphism \( G \to \text{Diff}(M) \), we can define the associated bundle \( P \times_G M := \)
(P × M)/G. This is naturally a fiber bundle over X with fiber M. If M has additional structure, and the image of G → Diff(M) respects this structure, then this additional structure is passed to the fiber bundle P × M/G. For example, when V is a vector space and G → GL(V) ⊂ Diff(V) is a representation, then P(V) := P × G V is a vector bundle. The most important example for us comes from the adjoint representation G → GL(g) ⊂ Diff(g). This respects the Lie algebra structure, and so the adjoint bundle P(g) := P × g g is a vector bundle with a Lie bracket on the fibers. This fiber-Lie bracket combines with the wedge product to determine a graded Lie bracket on the vector space \( \Omega^*(X, P(g)) \), and we denote this by \( \mu \otimes \nu \mapsto [\mu \wedge \nu] \). Similarly, if g is equipped with an Ad-invariant inner product \( \langle \cdot, \cdot \rangle \), then this determines an inner product on the fibers of \( P(g) \) and moreover combines with the wedge to form a graded bilinear map \( \Omega^*(X, P(g)) \otimes \Omega^*(X, P(g)) \to \Omega^*(X) \) that we denote by \( \mu \otimes \nu \mapsto \langle \mu \wedge \nu \rangle \).

2.1. Gauge theory. Let \( P \to X \) be a principal G-bundle. We will write

\[
\mathcal{A}(P) = \left\{ A \in \Omega^1(P, g) \mid (g_P)_* A = \text{Ad}(g)^{-1} A, \quad \forall g \in G \right\}
\]

for the space of connections on \( P \). Here \( g_P \) (resp. \( \xi_P \)) is the image of \( g \in G \) (resp. \( \xi \in g \)) under the map \( G \to \text{Diff}(P) \) (resp. \( g \to \text{Vec}(P) \)) afforded by the group action. It follows that \( \mathcal{A}(P) \) is an affine space modeled on \( \Omega^1(X; P(g)) \), and we denote the affine action by \( (V, A) \mapsto A + V \). In particular, \( \mathcal{A}(P) \) is a smooth (infinite dimensional) manifold with tangent space \( \Omega^1(X, P(g)) \). Each connection \( A \in \mathcal{A}(P) \) determines a covariant derivative \( d_A : \Omega^*(X, P(g)) \to \Omega^{*+1}(X, P(g)) \) and a curvature (2-form) \( F_A \in \Omega^2(X, P(g)) \). These satisfy \( d_{A+V} = d_A + [V \wedge \cdot] \) and \( F_{A+V} = F_A + d_A V + \frac{1}{2} [V \wedge V] \). We say that a connection \( A \) is irreducible if \( d_A \) is injective on 0-forms.

Given a metric on \( X \), we can define the formal adjoint \( d_A^* := (-1)^{(n-k)(k-1)} d_A^* \). Stokess’ theorem shows that this satisfies \( (d_A^* V, W)_{L^2} = (V, d_A^* W)_{L^2} \) for all compactly supported \( V, W \in \Omega^2(X, P(g)) \), where \( \langle \cdot, \cdot \rangle_{L^2} \) is the \( L^2 \)-inner product coming from the metric on \( X \).

A connection \( A \) is flat if \( F_A = 0 \). We will denote the set of flat connections on \( P \) by \( \mathcal{A}_{\text{flat}}(P) \). If \( A \) is flat then \( \text{im} (d_A) \subseteq \ker d_A \) and we can form the harmonic spaces

\[
H^k_A := H^k_A(X, P(g)) := \frac{\ker (d_A | \Omega^k(X, P(g)))}{\text{im} (d_A | \Omega^{k-1}(X, P(g)))}, \quad H^*_A := \bigoplus_k H^k_A.
\]

Suppose \( X \) is compact with (possibly empty) boundary, and let \( \partial : \Omega^*(X, P(g)) \to \Omega^*(\partial X, P(g)) \) denote the restriction. Then the Hodge isomorphism \([24]\) Theorem 6.8] says

\[
H^*_A \cong \ker (d_A + d_A^* + \partial^*), \quad \Omega^*(X, P(g)) \cong H^*_A \oplus \text{im} (d_A) \oplus \text{im} (d_A^* | \partial^*),
\]

for any flat connections \( A \) on \( X \), where the summands on the right are \( L^2 \)-orthogonal. We will treat these isomorphisms as identifications. From the first isomorphism in (1) we see that \( H^*_A \) is finite dimensional since \( d_A + d_A^* \) is elliptic. We will use \( \text{proj}_A : \Omega^*(X, P(g)) \to H^*_A \) to denote the projection; see Lemma [11] for an extension to the case where \( A \) is not flat.

**Example 2.1.** Suppose \( X = \Sigma \) is a closed, oriented surface equipped with a metric. Then the pairing \( \omega(\mu, \nu) := \int_{\Sigma} (\mu \wedge \nu) \) is a symplectic form on the vector space \( \Omega^1(X, P(g)) \). Note that changing the orientation on \( \Sigma \) replaced \( \omega \) by \( -\omega \). On surfaces, the Hodge star \( * \) squares to -1 on 1-forms and so defines a complex structure on \( \Omega^1(\Sigma, P(g)) \). It follows that the triple \( (\Omega^1(\Sigma, P(g)), *, \omega) \) is Kähler. If \( \alpha \in \mathcal{A}(P) \) is flat, then \( H^1_A \subset \Omega^1(\Sigma, P(g)) \) is a Kähler subspace.

Now suppose \( X \) is 4-manifold. Then on 2-forms the Hodge star squares to the identity, and it has eigenvalues \( \pm 1 \). Denoting by \( \Omega^\pm(X, P(g)) \) the \( \pm 1 \) eigenspace of \( * \), we have an \( L^2 \)-orthogonal decomposition \( \Omega^*(X, P(g)) = \Omega^+(X, P(g)) \oplus \Omega^-(X, P(g)) \). The elements of \( \Omega^+(X, P(g)) \) 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 \( F_A \in \Omega^-(X, P(g)) \) is an anti-self dual 2-form; that is, if \( F_A + * F_A = 0 \).
The space of connection $\mathcal{A}(P)$ admits a natural function

$$\mathcal{YM}_P : \mathcal{A}(P) \rightarrow \mathbb{R}, \quad A \mapsto \frac{1}{2} \| F_A \|_{L^2}^2$$
called the **Yang-Mills functional**. Obviously, the minimizers of this function are the flat connections, when they exist. However, in high dimensions the existence of flat connections is rare. For example, when $X$ is a closed four-manifold, one can show that the instantons minimize $\mathcal{YM}_P$, and instantons are typically not flat.

A **gauge transformation** is an equivariant bundle map $U : P \rightarrow P$ covering the identity. The set of gauge transformations on $P$ forms a Lie group, called the **gauge group**, and is denoted $\mathcal{G}(P)$. This is naturally a Lie group with Lie algebra $\Omega^0(X, P(\mathfrak{g}))$ under the map $R \mapsto \exp(-R)$, where $\exp : \mathfrak{g} \rightarrow G$ is the Lie-theoretic exponential.

The gauge group acts on the left on the space $\mathcal{A}(P)$ by pulling back by the inverse:

(2) \hspace{1cm} (U, A) \mapsto U A := (U^{-1})^* A,

for $U \in \mathcal{G}(P), A \in \mathcal{A}(P)$. In terms of a faithful matrix representation of $G$ we can write this as $U^* A = U^{-1} A U + U^{-1} dU$, where the concatenation appearing on the right is just matrix multiplication, and $dU$ is the linearization of $U$ when viewed as a $G$-equivariant map $P \rightarrow G$. The infinitesimal action of $\mathcal{G}(P)$ at $A \in \mathcal{A}(P)$ is

(3) \hspace{1cm} \Omega^0(X, P(\mathfrak{g})) \rightarrow \Omega^1(X, P(\mathfrak{g})), \quad R \mapsto d_A R

More generally, the derivative of (2) at $(U, A)$ is

(4) \hspace{1cm} (U \Omega^0(X, P(\mathfrak{g})) \times \Omega^1(X, P(\mathfrak{g})) \rightarrow \Omega^1(M, P(\mathfrak{g}))

\hspace{1cm} (UR, V) \mapsto \text{Ad}(U)d_A R + \text{Ad}(U)V = (d(U^{-1})^* \text{Ad}(UR)) U^{-1} + UVU^{-1}.

where, in writing this expression, we have chosen a faithful matrix representation $G \hookrightarrow \text{GL}(V)$. This allows us to view $\mathcal{G}(P)$ and its Lie algebra $\Omega^0(X, P(\mathfrak{g}))$ as subspaces of the same algebra $\Gamma(P(\text{End}(V)))$, and so it makes sense to identify the tangent space of $\mathcal{G}(P)$ at $U$ with $U\Omega^0(X, P(\mathfrak{g}))$, as we have done in (1).

The gauge group also acts on the left on $\Omega^*(X, P(\mathfrak{g}))$ by the pointwise adjoint action. The curvature of $A \in \mathcal{A}(P)$ transforms under $U \in \mathcal{G}(P)$ by $F_{U^* A} = \text{Ad}(U^{-1}) F_A$. This shows that $\mathcal{G}(P)$ restricts to an action on $\mathcal{A}_{\text{flat}}(P)$ and, in 4-dimensions, the instantons.

In general, we will tend to use capital letters $A, U$ to denote connections and gauge transformations on 4-manifolds $Z$, lower case letters $a, u$ to denote connections and gauge transformations on 3-manifolds $Y$, and lower case Greek letter $\alpha, \mu$ to denote gauge transformations on surfaces $\Sigma$.

We will be interested in the case $G = \text{PU}(r)$ for $r \geq 2$. We equip the Lie algebra $\mathfrak{g} \cong \mathfrak{su}(r) \subset \text{End}(\mathbb{C}^r)$ with the inner product $\langle \mu, \nu \rangle := -\kappa_r \text{tr}(\mu \cdot \nu)$; here $\kappa_r > 0$ is arbitrary, but fixed. On manifolds $X$ of dimension at most 4, the principal $\text{PU}(r)$-bundles $P \rightarrow X$ are classified, up to bundle isomorphism, by two characteristic classes $t_2(P) \in H^2(X, \mathbb{Z}_r)$ and $q_4(P) \in H^4(X, \mathbb{Z})$. These generalize the 2nd Stiefel-Whitney class and 1st Pontryagin class, respectively, to the group $\text{PU}(r)$; see [30]. When $X$ is a closed, oriented 4-manifold, there is also a Chern-Weil formula

(5) \hspace{1cm} q_4(P) = \frac{r}{4\pi^2 \kappa_r} \int_X (F_A \wedge F_A)

which holds for any connection $A \in \mathcal{A}(P)$.

Consider a principal $\text{PU}(r)$-bundle $P \rightarrow X$ where we assume $\text{dim}(X) \leq 3$. Then there are maps

$$\eta : \mathcal{G}(P) \rightarrow H^1(X, \mathbb{Z}_r), \quad \deg : \mathcal{G}(P) \rightarrow H^3(X, \mathbb{Z})$$
called the **parity** and **degree**. These detect the connected components of \(G(P)\) in the sense that if \(u\) can be connected to \(u'\) by a path if and only if \(\eta(u) = \eta(u')\) and \(\deg(u) = \deg(u')\). We denote by \(G_0(P)\) the identity component of \(G(P)\). See [9].

Suppose \(X = \Sigma\) is a closed, connected, oriented surface, and \(P \to \Sigma\) is a principal \(PU(r)\)-bundle with \(t_2(P) [\Sigma] \in \mathbb{Z}_r\), a generator. It can be shown that all flat connections on \(P\) on irreducible, and that \(G_0(P)\) acts freely on \(A_{\text{flat}}(P)\). Moreover, one has that

\[
M(P) := A_{\text{flat}}(P)/G_0(P)
\]

is a compact, simply-connected, smooth manifold. Restricting to the two boundary components induces an embedding of \(G(P)\) into \(U(M)\) here, so this will not be further identified with \(H^1_{\Sigma}\). It follows from Example [21] that \(M(P)\) is a symplectic manifold, and any metric on \(\Sigma\) determines an almost complex structure \(J_\Sigma\) on \(M(P)\) that is compatible with the symplectic form. See [27] for more details regarding these assertions.

Now suppose \(Y_{ab}\) is an oriented elementary cobordism between closed, connected, oriented surfaces \(\Sigma_a\) and \(\Sigma_b\). Fix a \(PU(r)\)-bundle \(Q_{ab} \to Y_{ab}\) with \(t_2(Q_{ab}) [\Sigma_a] \in \mathbb{Z}_r\) a generator. Then the flat connections on \(Q_{ab}\) are irreducible, and the quotient \(A_{\text{flat}}(Q_{ab})/G_0(Q_{ab})\) is a finite-dimensional, simply-connected, smooth manifold. Restricting to the two boundary components induces an embedding \(A_{\text{flat}}(Q_{ab})/G_0(Q_{ab}) \to M(\Sigma_a) \times M(\Sigma_b)\), and we let \(L(Q_{ab})\) denote the image. It follows that \(L(Q_{ab}) \subset M(\Sigma_a) \times M(\Sigma_b)\) is smooth Lagrangian submanifold, where the superscript in \(M(\Sigma_a)\) means that we have replaced the symplectic structure with its negative. See [27].

More generally, if \(G\) is a compact Lie group, and \(P \to X\) is a principal \(G\)-bundle, then we can consider the space \(A_{\text{flat}}(P)/G(P)\). (Note that we are not quotienting by the identity component here, so this will not be \(M(P)\) when \(G = PU(r)\) and \(X = \Sigma\).) This space has a natural topology, and it follows from Uhlenbeck’s compactness theorem that this space is compact when \(X\) is compact. However, it is rarely a smooth manifold.

### 2.2. The complexified gauge group

Let \(G\) be a compact, connected Lie group and fix a faithful \(G\)-invariant Hermitian inner product and a complex structure \(J_E\).

Suppose \((\Sigma, j_\Sigma)\) is a Riemann surface. We will use the \(\Omega^{k,l}(\Sigma, E)\) to denote the smooth \(E\)-valued forms of type \((k,l)\). Observe that \(j_\Sigma\) acts by the Hodge star on 1-forms. Consider the space

\[
\mathcal{C}(E) := \left\{ \mathcal{T} : \Omega^0(\Sigma, E) \to \Omega^{0,1}(\Sigma, E) \left| \begin{array}{c} D(f\xi) = f(D\xi) + (\bar{D}(\bar{f})\xi), \\
\text{for } \xi \in \Omega^0(\Sigma, E), \ f \in \Omega^0(\Sigma) \end{array} \right. \right\},
\]

of Cauchy-Riemann operators on \(E\). This can be naturally identified with the space of holomorphic structures on \(E\) (see [18] Appendix C). Each element \(D \in \mathcal{C}(E)\) has a unique extension to an operator \(\bar{D} : \Omega^{i,k}(\Sigma, E) \to \Omega^{i,k+1}(\Sigma, E)\) satisfying the Leibniz rule.

Let \(A(E)\) denote the space of \(C\)-linear covariant derivatives on \(E\):

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

There is a natural isomorphism

\[
A(E) \to \mathcal{C}(E), \quad D \mapsto \frac{1}{2} (D + J_E D \circ j_\Sigma)
\]

(6)

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 given by the formula \(\iota_X ((M \circ j_\Sigma)\xi) := \iota_{j_\Sigma(X)} (M\xi)\); here \(\iota_X\) is contraction with a vector \(X\).

Let \(P(\mathfrak{g})^C\) denote the complexification of the vector bundle \(P(\mathfrak{g})\). Then we have bundle inclusions

\[
P(\mathfrak{g}) \subset P(\mathfrak{g})^C \subset \text{End}(E),
\]
where $\text{End}(E)$ is the bundle of complex linear automorphisms of $E$ and the latter inclusion is induced by the embedding $\rho$. 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 a corresponding curvature $F_{\alpha,\rho} = d_{\alpha,\rho} \circ d_{\alpha,\rho} \in \Omega^2(\Sigma, P(\mathfrak{g}))$. Since the representation $\rho$ is faithful, we have pointwise estimates of the form

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

this allow us to discuss curvature bounds in terms of either $F_\alpha$ 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 $\Omega^1(\Sigma, P(\mathfrak{g})) \subseteq \Omega^1(\Sigma, \text{End}(E))$. In particular, restricting to the image of $\mathcal{A}(P)$ in $\mathcal{A}(E)$, the map (6) becomes an embedding

$$\mathcal{A}(P) \to \mathcal{C}(E), \quad \alpha \mapsto \omega := \frac{1}{2} (d_{\alpha,\rho} + J_E d_{\alpha,\rho} \circ j_E)$$

The image of (7) is the set of covariant derivatives that preserve the $G$-structure, and we denote it by $\mathcal{C}(P)$. See [18, Appendix C] for the case when $G = U(n)$. The space $\mathcal{C}(P)$ is an affine space modeled on $\Omega^{0,1}(\Sigma, P(\mathfrak{g})^\mathbb{C})$. Similarly, $\mathcal{A}(P)$ is an affine space modeled on $\Omega^1(\Sigma, P(\mathfrak{g}))$, and (7) intertwines these affine actions under the identification $\Omega^1(\Sigma, P(\mathfrak{g})) \cong \Omega^{0,1}(\Sigma, P(\mathfrak{g})^\mathbb{C})$ sending $\mu$ to its anti-linear part $\mu^{0,1} := \frac{1}{2} (\mu + J_E \mu \circ j_E)$. To summarize, we have a commutative diagram

$$\begin{array}{ccc}
\mathcal{A}(P) & \xrightarrow{\cong} & \mathcal{C}(P) \\
\downarrow & & \downarrow \\
\mathcal{A}(E) & \xrightarrow{\cong} & \mathcal{C}(E)
\end{array}$$

where the vertical arrows are inclusions and everything is equivariant with respect to $\Omega^1(\Sigma, P(\mathfrak{g}))$.

To define the complexified gauge group, we need to first recall some basic properties of the complexification of compact Lie groups. See [14] or [12] for more details on this material. Since $G$ is compact and connected, there is a connected complex group $G^\mathbb{C}$ and an embedding $G \to G^\mathbb{C}$ such that $G$ is a maximal compact subgroup of $G^\mathbb{C}$, and the Lie algebra $\mathfrak{g}^\mathbb{C} = \text{Lie}(G^\mathbb{C})$ is the complexification of $\mathfrak{g} = \text{Lie}(G)$. This group $G^\mathbb{C}$ is unique up to natural isomorphism and is called the complexification of $G$.

We may assume that the representation $\rho : G \to U(n)$ from above extends to an embedding $G^\mathbb{C} \to \text{GL}(\mathbb{C}^n)$, and we identify $G^\mathbb{C}$ with its image (see [12, Proof of Theorem 1.7]). Then we have $G = \{ u \in G^\mathbb{C} \mid w^* u = 1 \}$, where $w^*$ denotes the conjugate transpose on $\text{GL}(\mathbb{C}^n)$. It follows that we can write $G^\mathbb{C} = \{ g \exp(i \xi) \mid g \in G, \xi \in \mathfrak{g} \}$, 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

$$g \exp(i \xi) = \exp(i \text{Ad}(g) \xi) g$$

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

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

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

As in the real case, we may identify $\Omega^0(\Sigma, P(\mathfrak{g})^\mathbb{C})$ with the Lie algebra of $(\mathcal{G}(P)^\mathbb{C})$ via the map

$$\xi \mapsto \exp(-\xi),$$

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

(8) $g \exp(i \xi) = \exp(i \text{Ad}(g) \xi) g$
for some \(g \in \mathcal{G}(P)\) and \(\xi \in \Omega^0(\Sigma, P(\mathfrak{g}))\), and \(\mathbb{D}\) continues to hold with \(g, \xi\) interpreted as elements of \(\mathcal{G}(P), \Omega^0(\Sigma, P(\mathfrak{g}))\), respectively.

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

\[
\mathcal{G}(P)^C \times \mathcal{C}(P) \longrightarrow \mathcal{C}(P), \quad (\mu, \overline{D}) \mapsto \mu \circ \overline{D} \circ \mu^{-1}
\]

Viewing \(\mathcal{G}(P)\) as a subgroup of \(\mathcal{G}(P)^C\) in the obvious way, then the identification \(\mathbb{D}\) is \(\mathcal{G}(P)\)-equivariant. We can then use \(\mathbb{D}\) and \(\mathbb{C}\) to define an action of the larger group \(\mathcal{G}(P)^C\) on \(\mathcal{A}(P)\), extending the \(\mathcal{G}(P)\)-action. We denote the action of \(\mu \in \mathcal{G}(P)^C\) on \(\alpha\) by \((\mu^{-1})^*\alpha\) or simply \(\mu \alpha\) when there is no room for confusion. Explicitly, the action on \(\mathcal{A}(P)\) takes the form

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

where the dagger is applied pointwise. In particular, the infinitesimal action at \(\alpha \in \mathcal{A}(P)\) is

\[
\Omega^0(\Sigma, P(\mathfrak{g})) \longrightarrow \Omega^1(\Sigma, P(\mathfrak{g})), \quad \xi + i\zeta \longmapsto d_{\alpha, \rho} \xi + *d_{\alpha, \rho} \zeta
\]

More generally, the derivative of the map \((\mu, \alpha) \mapsto (\mu^{-1})^*\alpha\) at \((\mu, \alpha)\) with \(\mu \in \mathcal{G}(P)\) (an element of the real gauge group) is a map \(\mu \left( \Omega^0(\Sigma, P(\mathfrak{g})) \oplus i\Omega^0(\Sigma, P(\mathfrak{g})) \right) \times \Omega^1(\Sigma, P(\mathfrak{g})) \longrightarrow \Omega^1(\Sigma, P(\mathfrak{g}))\) given by

\[
(\mu(\xi + i\zeta), \eta) \longmapsto \text{Ad}(\mu) \left( d_{\alpha, \rho} \xi + *d_{\alpha, \rho} \zeta + \eta \right) = \left\{ d_{(\mu^{-1})^*\alpha, \rho} (\mu \xi) + *d_{(\mu^{-1})^*\alpha, \rho} (\mu \zeta) \right\} \mu^{-1} + \text{Ad}(\mu) \eta
\]

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

\[
\mu^{-1} \circ F_{(\mu^{-1})^*\alpha, \rho} \circ \mu = F_{\alpha, \rho} + \overline{\partial}_\alpha \left( h^{-1} \partial_\alpha h \right),
\]

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

\[
\exp(-i\xi) \circ F_{\exp(-i\xi)^*\alpha, \rho} \circ \exp(i\xi) = *\mathcal{F}(\alpha, \xi),
\]

where we have set

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

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

\[
\mathcal{G}(E) := \Gamma(P \times_G U(n)), \quad \mathcal{G}(E)^C := \Gamma(P \times_G \text{GL}(\mathbb{C}^n)).
\]

(Note that the complexification of \(U(n)\) is \(\text{GL}(\mathbb{C}^n)\), so this terminology is consistent, and in fact motivates, the terminology above.) These are both Lie groups with Lie algebras \(\text{Lie}(\mathcal{G}(E)) = \Gamma(P \times_G U(n))\) and \(\text{Lie}(\mathcal{G}(E)^C) = \Gamma(P \times_G \text{End}(\mathbb{C}^n))\), respectively, where we are identifying \(\text{End}(\mathbb{C}^n)\) with the Lie algebra of \(\text{GL}(\mathbb{C}^n)\). We have the obvious inclusions
The space $G(E)^C$ acts on $C(E)$ by the map

\begin{equation}
G(E)^C \times C(E) \to C(E), \quad (\mu, \overline{T}) \mapsto \mu \circ \overline{T} \circ \mu^{-1}
\end{equation}

Using (9), this induces an action of $G(E)^C$ on $A(E)$ (hence an action of $G(E)$ on $A(E)$), though, neither $G(E)^C$ nor $G(E)$ restrict to actions on $A(P)$, unless $G = U(n)$.

Finally, we mention that the vector spaces $\operatorname{Lie}(G(E))$, $\operatorname{Lie}(G(E)^C)$ and $\operatorname{Lie}(G(P)^C)$ admit Sobolev completions. For example, the space $\operatorname{Lie}(G(P)^C)^{k,q}$ is the $W^{k,q}$-completion of the vector space $\Gamma(P \times_{G} P(g)^C)$. 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 $G^{k,q}(E)$, $G^{k,q}(E)^C$ and $G^{k,q}(P)^C$ by taking the $W^{k,q}$-completions of the groups of smooth functions $G(E)$, $G(E)^C$ and $G(P)^C$, which we view as lying in the vector space $\Gamma(P \times_{G} \operatorname{End}(C^\infty))^{k,q}$. The complexified gauge action extends to a smooth action of $G^{k,q}(E)^C$ on $A^{k-1,q}(E)$, and this restricts to a smooth action of $G^{k,q}(P)^C$ on $A^{k-1,q}(P)$. See [26, Appendix B] for more details regarding the Sobolev completions of these spaces.

3. Compactness for products $S \times \Sigma$

In this section we consider $Z = S \times \Sigma$, where $(S, g_S), (\Sigma, g_{\Sigma})$ are closed, connected, oriented Riemannian surfaces. We also assume that $\Sigma$ has positive genus. We will work with the metric $g = \text{proj}_{S}^* g_S + \text{proj}_{\Sigma}^* g_{\Sigma}$ on $Z$; from now on we will typically drop the projections from the notation. For $\epsilon > 0$, define a new metric by

$$g_\epsilon := \epsilon^2 g_S + g_{\Sigma}.$$  

Fix a principal $\text{PU}(r)$-bundle $P \to \Sigma$ such that $t_2(P) [\Sigma] \in \mathbb{Z}_r$ is a generator, and let $R \to Z$ be the pullback bundle under $Z \to \Sigma$.

Given orthonormal coordinates $(U, x = (s, t))$ for $S$, any connection $A$ on $R$ can be written as

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

where for each $(s, t)$, $\alpha(s, t)$ is a connection on $P$ and $\phi(s, t), \psi(s, t) \in \Omega^0(\Sigma, P(g))$. In fact, $\alpha$ can be defined in a coordinate-independent way as $x \mapsto \alpha(x) := t_2^* A \in A(P)$, where $t_2 : \Sigma = \{x\} \times \Sigma \to Z$ is the inclusion. We will say that a connection $A$ on $R$ is $\epsilon$-\textbf{ASD} or an $\epsilon$-\textbf{instanton} if it is an instanton with respect to the metric $g_\epsilon$; that is, if $F_A = - \ast_\epsilon F_A$, where $\ast_\epsilon$ is the Hodge star on $Z$ coming from $g_\epsilon$. This can be written explicitly in terms of local coordinates as

\begin{equation}
\begin{align*}
\partial_a \alpha - d_a \phi + \ast_\Sigma (\partial_a \alpha - d_a \psi) &= 0 \\
\partial_a \psi - \partial_t \phi - [\psi, \phi] + \epsilon^{-2} \ast_\Sigma F_\alpha &= 0
\end{align*}
\end{equation}

where $\ast_\Sigma$ is the Hodge star associated to the $S$-dependent metric $g_{\Sigma}$. The $\epsilon$-\textbf{energy} of $A$ is

$$E^\text{inst}_\epsilon(A) := \frac{1}{2} \int_Z |F_A|^2 d\text{vol}_\epsilon = \frac{1}{2} \int_Z \langle F_A \wedge \ast_\epsilon F_A \rangle,$$

where the norm and volume form are the ones induced by $g_\epsilon$; this is exactly the Yang-Mills functional on $A(R)$ with the metric $g_\epsilon$. If $A$ is $\epsilon$-ASD, then $E^\text{inst}_\epsilon(A) = -2\pi^2 \kappa_r \epsilon^{-1} q_4(R)$ is a topological invariant by the Chern-Weil formula (5).

On the symplectic side, we will be considering maps $v : S \to M(P)$, where $M(P) = \mathcal{A}_\text{flat}(P) / \mathcal{G}_0(P)$. Any such map $v$ has a lift $\alpha : S \to \mathcal{A}_\text{flat}(P)$ if and only if the pullback $\mathcal{G}_0(P)$-bundle $v^* \mathcal{A}_\text{flat}(P) \to S$
Proof. Suppose given any Lemma 3.2. Theorem 3.3. Fix by the Chern-Weil formula. □

representative a sequence of gauge transformations (iii) a subsequence of the (ii) a finite set (i) a sequence of gauge transformations $U_v \in \mathcal{G}_{2q}^{loc}(R)$; and (iv) a holomorphic curve representative $A_\infty \in \mathcal{A}^{1q}_{loc}(R)$.

Example 3.1. Let $(U,(s,t))$ be local orthonormal coordinates for $S$. Then $Da\|U = ds \wedge \partial_s \alpha + dt \wedge \partial_t \alpha$. Then $\chi|_U = \phi \, ds + \psi \, dt$, where $\phi, \psi : U \to \Omega^0(S, P(g))$ are the unique sections such that

$$
\partial_s \alpha(s,t) - d_{\alpha,s(t)} \phi(s,t), \quad \partial_t \alpha(s,t) - d_{\alpha,s(t)} \psi(s,t)
$$

are $\alpha(s,t)$-harmonic. Then $A_0$ is a holomorphic curve representative if and only if

$$(19) \quad \partial_t \alpha - d_{\alpha} \phi + *_{S} (d_{\alpha} \phi - d_{\alpha} \psi) = 0, \quad F_\alpha = 0,$$

where the equations should be interpreted as being pointwise in $(s,t) \in S$.

We define the energy $E^{\text{sym}}(A_0)$ of a representative $A_0$ to be the energy of the associated map $v : S \to M(P)$.

Lemma 3.2. Given any $\epsilon > 0$, the energy of a holomorphic curve representative $A_0$ is

$$
E^{\text{sym}}(A_0) = -2\pi^2 \kappa_4 r^{-1} q_4(R).
$$

Proof. Suppose $A_0$ represents $v$. Then the energy of $v$ is $\frac{1}{2} \int_S |Dv|^2 \, dvol_S$, where $Dv$ is the push-forward of $v : S \to M(P)$. In local coordinates, write $A_0 = \alpha + \phi \, ds + \psi \, dt$. Then

$$
Dv = ds \otimes \partial_s v + dt \otimes \partial_t v = ds \otimes \beta_s + dt \otimes \beta_t,
$$

where we have set $\beta_s := \partial_s \alpha - d_{\alpha} \phi$ and $\beta_t := \partial_t \alpha - d_{\alpha} \psi$. On the other hand,

$$
F_{A_0} = ds \wedge \beta_s + dt \wedge \beta_t + ds \wedge dt (\partial_s \psi - \partial_t \phi - [\psi, \phi])
$$

where we have used $F_\alpha = 0$. Then by (19) we have

$$
\frac{1}{2} \int_{\Sigma} \langle F_{A_0} \wedge F_{A_0} \rangle = \int_{\Sigma} \langle \beta_s \wedge \beta_t \rangle \wedge ds \wedge dt = \|\beta_s\|^2_{L^2(S)} \, ds \wedge dt = \frac{1}{2} |Dv|^2 \, dvol_S.
$$

Integrating the right over $S$ gives $E^{\text{sym}}(A_0)$, and integrating the left over $S$ gives $-2\pi^2 \kappa_4 r^{-1} q_4(R)$ by the Chern-Weil formula.

Now we can state the main theorem of this section.

Theorem 3.3. Fix $2 < q < \infty$ and let $R \to Z$ be as above. Suppose $(\epsilon_v)_{v \in \mathbb{N}}$ is a sequence of positive numbers converging to 0, and that, for each $v$, there is an $\epsilon_v$-ASD connection $A_v \in \mathcal{A}^{1q}_{loc}(R)$. Then there is

(i) a finite set $B \subset S$;
(ii) a subsequence of the $A_v$ (still denoted $A_v$);
(iii) a sequence of gauge transformations $U_v \in \mathcal{G}_{2q}^{loc}(R)$; and
(iv) a holomorphic curve representative $A_\infty \in \mathcal{A}^{1q}_{loc}(R)$.
such that the restrictions

$$
\sup_{x \in K} \|\epsilon_x^* (U_\nu^* A_\nu - A_\infty)\|_{C^0(\Sigma)} \rightarrow 0
$$

converge to zero for every compact $K \subset S \backslash B$. The gauge transformations $U_\nu$ can be chosen so that they restrict to the identity component $\mathcal{G}_0(P)$ on each $\{x\} \times \Sigma \subset \mathbb{Z}$. Moreover, for each $b \in B$ there is a positive integer $m_b > 0$ such that for any $\nu$,

$$
E^{\text{symp}}(A_\infty) = E^{\text{inst}}(A_\nu) - 4\pi^2 \kappa_r r^{-1} \sum_{b \in B} m_b.
$$

(20)

Throughout we will use notation of the form $A_{\{x\} \times \Sigma} = \alpha(x) + \chi(x)$ for connections, and $U_{\{x\} \times \Sigma} = \mu(x)$ for gauge transformations. Then the conclusion of the theorem says that $\mu^* \alpha_\nu$ converges to $\alpha_\infty$ in $C^0$ on compact sets in $S \backslash B$, where these are viewed as maps from $S$ to $\mathcal{A}(P)$, with the $C^0$-topology on $\mathcal{A}(P)$. Here, the action of $\mu = \mu(x)$ is the gauge action on surfaces (not 4-manifolds).

**Remark 3.4.** (a) If one allows $S$ to be a compact manifold-with-boundary, then the proof we give here remains valid, except the equality in (20) should be replaced by $\leq$. This is due to holomorphic disk bubbles occurring at the boundary – our convergence is not strong enough to show that these are non-trivial; see Remark 3.19. In Section 4 we provide tools for recovering the equality when the $\epsilon$-instantons have approximate Lagrangian boundary conditions.

(b) This theorem has a straightforward extension to the case where $S$ has cylindrical ends (assuming one knows flat connections on the ends are non-degenerate; see Section 5). In this case, there can be energy loss at the ends, and so one obtains a limiting holomorphic curve on $S$ with a finite number of broken cylindrical trajectories on the ends. Then Theorem 3.3 continues to hold provided one accounts for the energies of the broken trajectories on the left-hand side (20). See also Theorem 4.1.

(c) Suppose, for each $\nu$, we have an open set $S_\nu \subseteq S$ that is a deformation retract of $S$, and with the further property that the $S_\nu$ are increasing and exhausting: $S_\nu \subseteq S_{\nu+1}$ and $S = \bigcup_\nu S_\nu$. Then the statement of Theorem 3.3 continues to hold if we assume that $A_\nu$ is defined on $S_\nu \times \Sigma$.

(d) It is natural to ask whether the $\chi_\nu$ converge in any sense to $\chi_\infty$ (these are the components of $A_\nu$ and $A_\infty$ not addressed in the conclusion of the theorem). It is unlikely that one has $C^0$-convergence of the $\chi_\nu$ when the norm is defined with respect to the fixed metric, as we had for the $\alpha_\nu$. However, it is possible to show that the $C^0$-convergence of the $\alpha_\nu$ to $\alpha_\infty$ implies $W^{1,2}_{\infty}$-convergence of the $A_\nu$ to $A_\infty$, where $W^{1,2}_{\infty}$ is the Sobolev norm defined with respect to the $\epsilon$-dependent metric. A similar statement holds in the case of Theorem 4.1 as well. We defer the details to a future paper.

As a stepping stone to Theorem 3.3 we first prove the following lemma; the assumptions allow us to rule out bubbling a priori.

**Lemma 3.5.** Fix $2 < q < \infty$, and a submanifold $S_0 \subseteq S$ possibly with boundary. Let $R_0 := R|_{S_0 \times \Sigma}$ denote the restriction. Suppose $(\epsilon_\nu)_{\nu \in \mathbb{N}}$ is a sequence of positive numbers (not necessarily converging to zero), and suppose that for each $\nu$ there is an $\epsilon_\nu$-ASD connection $A_\nu \in A^{1,q}(R_0)$ satisfying the following conditions for each compact $K \subset S_0$.

(i) The slice-wise curvatures converge to zero: $\sup_{x \in K} \|F_{\alpha_\nu(x)}\|_{L^\infty(\Sigma)} \rightarrow 0$.

(ii) There is some constant $C$ with

$$
\sup_{\nu} \sup_{x \in K} \|\text{proj}_{\alpha_\nu(x)} \circ D_x \alpha_\nu\|_{L^2(\Sigma)} \leq C,
$$

where $\text{proj}_{\alpha_\nu(x)}$ is the harmonic projection, and $D_x \alpha_\nu : T_x S \rightarrow T_{\alpha(x)} A^{1,q}(P)$ is the push-forward.
Then there is a subsequence of the connections (still denoted $A_{\nu}$), a sequence of gauge transformations $U_{\nu}$ on $R_0$, and a holomorphic curve representative $A_{\infty}$ on $R_0$ such that

$$(21) \sup_{x \in K} \left\| \alpha_{\infty}(x) - \mu_{\nu}(x)^* \alpha_{\nu}(x) \right\|_{C^p(\Sigma)} \to 0$$

for every compact $K \subseteq S_0$, and

$$(22) \sup_{x \in K} \left\| \text{proj}_{\alpha_{\infty}(x)} \circ D_x \alpha_{\infty} - \text{Ad}(\mu_{\nu}^{-1}(x)) \text{proj}_{\alpha_{\nu}(x)} \circ D_x \alpha_{\nu} \right\|_{L^p(\Sigma)} \to 0$$

for any compact $K \subseteq \text{int} S_0$ and any $1 < p < \infty$.

The connections $A_{\nu}$ from Theorem 3.3 satisfy the same type of convergence as in (22), for compact $K \subseteq S \setminus B$. We also point out that the projection operator $\text{proj}_{\alpha_{\nu}}$ appearing in (22) can be removed by weakening the $C^p$-convergence to $L^p$-convergence; see Remark 3.10 (b).

The proofs of Lemma 3.5 and Theorem 3.3 will appear in Sections 3.2 and 3.3, respectively. First after proving Theorem 3.6 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 3.16 we establish the Narasimhan-Seshadri correspondence

$$(\alpha, F_\alpha) 
\to 
\left( \begin{array}{c} \text{proj}_{\alpha_{\infty}(x)} \circ D_x \alpha_{\infty} - \text{Ad}(\mu_{\nu}^{-1}(x)) \text{proj}_{\alpha_{\nu}(x)} \circ D_x \alpha_{\nu} \\ \text{proj}_{\alpha_{\infty}(x)} \circ D_x \alpha_{\infty} - \text{Ad}(\mu_{\nu}^{-1}(x)) \text{proj}_{\alpha_{\nu}(x)} \circ D_x \alpha_{\nu} \end{array} \right)$$

for $\mu_{\nu}, \nu \to \mu_{\nu}, \nu \to \mu_{\nu}$ such that $\mu_{\nu} \to \mu_{\nu}, \nu \to \mu_{\nu}$ in $\text{int S}_0$ and any $1 < p < \infty$.

In preparation for a boundary-value problem, we need to work in an analytic category. Consequently, we adopt an approach of Donaldson [3], and use an implicit function theorem argument to arrive at a Narasimhan-Seshadri correspondence in our setting. This allows us to establish several $C^1$ and $C^2$-estimates that will be needed for our proof of the main theorems.

3.1. Small curvature connections in dimension 2. In our proof of the various convergence results, we will encounter connections on surfaces that have small curvature. Here we develop a strategy for identifying nearby flat connections. The idea is to use the well-known fact that quotienting the subset of small curvature connections by the action of the complexified gauge group recovers the moduli space of flat connections (called a Narasimhan-Seshadri correspondence). The details of this procedure were originally carried out by Narasimhan and Seshadri [19]. They worked with unitary bundles, and this allowed them to use algebraic techniques. Later, their techniques were extended to more general structure groups by Ramanathan in his thesis [21]. (See also Kirwan’s book [15] for a finite-dimensional version.)

In preparation for a boundary-value problem, we need to work in an analytic category. Consequently, we adopt an approach of Donaldson [3], and use an implicit function theorem argument to arrive at a Narasimhan-Seshadri correspondence in our setting. This allows us to establish several $C^1$ and $C^2$-estimates that will be needed for our proof of the main theorems.

3.1.1. The analytic Narasimhan-Seshadri correspondence. The goal of this section is to define a gauge-equivariant deformation retract $NS : A^{ss} \to A_{\text{flat}}$ and establish some of its properties. Here $A^{ss}$ is a suitable neighborhood of $A_{\text{flat}}$ (the superscript stands for semistable). The relevant properties of the map $NS$ are laid out in Theorem 3.6 below. The proof will show that for each $\alpha \in A^{ss}$ there is a ‘purely imaginary’ complex gauge transformation $\mu$ such that $\mu^* \alpha$ a flat connection, and $\mu$ is unique provided it lies sufficiently close to the identity. We then define $NS(\alpha) := \mu^* \alpha$.

After proving Theorem 3.6 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 3.10 we establish the Narasimhan-Seshadri correspondence $A^{ss} / G_0^C \cong A_{\text{flat}} / G_0$, and in Proposition 3.4 and Corollary 3.18 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 3.6.** Suppose $G$ is a compact, connected Lie group, $\Sigma$ is a closed 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 $G^{2,q}(P)$-equivariant deformation retract

$$(23) \quad NS_P : \{ \alpha \in A^{1,q}(P) \mid \| F_\alpha \|_{L^q(\Sigma)} < \epsilon_0 \} \to A^{1,q}_{\text{flat}}(P)$$

that 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 3.7. 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 3.17.

Proof of Theorem 3.7. Suppose we can define \(\text{NS}_P\) on the set

\[
\left\{ \alpha \in A^{1,q}(P) \Big| \text{dist}_{W^{1,q}}(\alpha, A^{1,q}_{\text{flat}}(P)) < \epsilon_0 \right\},
\]

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

\[
\left\{ \mu^* \alpha \in A^{1,q}(P) \Big| \mu \in G^{2,q}(P), \text{dist}_{W^{1,q}}(\alpha, A^{1,q}_{\text{flat}}(P)) < \epsilon_0 \right\},
\]

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 (23), thereby reducing the problem to defining \(\text{NS}_P\) on a set of the form (24).

Claim: For any \(\bar{\epsilon}_0 > 0\), there is some \(\epsilon_0 > 0\) with

\[
\left\{ \alpha \in A^{1,q}(P) \Big| \|F_\alpha\|_{L^q} < \epsilon_0 \right\} \subseteq \left\{ \mu^* \alpha \in A^{1,q}(P) \Big| \mu \in G^{2,q}(P), \text{dist}_{W^{1,q}}(\alpha, A^{1,q}_{\text{flat}}(P)) < \bar{\epsilon}_0 \right\}.
\]

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

\[
\|\mu^* \alpha - \alpha_0\|_{W^{1,q}} \geq \bar{\epsilon}_0, \quad \forall \mu \in G^{2,q}(P), \quad \forall \alpha_0 \in A^{1,q}_{\text{flat}}(P)
\]

Then we can find a sequence of connections \(\alpha_\nu\) with \(\|F_{\alpha_\nu}\|_{L^2} \to 0\), but (25) holds with \(\alpha_\nu\) replacing \(\alpha\). By Uhlenbeck’s weak compactness theorem, there is a sequence of gauge transformations \(\mu_\nu \in G^{2,q}(P)\) such that, after possibly passing to a subsequence, \(\mu_\nu^* \alpha_\nu\) converges weakly in \(W^{1,q}\) to a limiting connection \(\alpha_0\). The condition on the curvature implies that \(\alpha_0 \in A^{1,q}_{\text{flat}}(P)\) is flat. Moreover, the embedding \(W^{1,q} \hookrightarrow L^{2q}\) is compact, so the weak \(W^{1,q}\)-convergence of \(\mu_\nu^* \alpha_\nu\) implies that \(\mu_\nu^* \alpha_\nu\) converges strongly to \(\alpha_0\) in \(L^{2q}\). By redefining \(\mu_\nu\) if necessary, we may suppose that \(\mu_\nu^* \alpha_\nu\) is in Coulomb gauge with respect to \(\alpha_0\), \(d^*_{\alpha_0}(\mu_\nu^* \alpha_\nu - \alpha_0) = 0\), and still retain the fact that \(\mu_\nu^* \alpha_\nu\) converges to \(\alpha_0\) strongly in \(L^{2q}\). This gives

\[
\|\mu_\nu^* \alpha_\nu - \alpha_0\|_{W^{1,q}}^q = \|\mu_\nu^* \alpha_\nu - \alpha_0\|_{L^q}^q + \|d_{\alpha_0} (\mu_\nu^* \alpha_\nu - \alpha_0)\|_{L^q}^q + \|d^*_{\alpha_0} (\mu_\nu^* \alpha_\nu - \alpha_0)\|_{L^{2q}}^q
\]

\[
\leq \|\mu_\nu^* \alpha_\nu - \alpha_\nu\|_{L^q}^q + \|F_{\alpha_0}\|_{L^q}^q + \frac{1}{2}\|\mu_\nu^* \alpha_\nu - \alpha_\nu\|_{L^{2q}}^q
\]

\[
\leq C \left( \|\mu_\nu^* \alpha_\nu - \alpha_\nu\|_{L^{2q}}^q + \|F_{\alpha_0}\|_{L^q}^q \right)
\]

where we have used the formula \(F_{\alpha_0 + \nu} = d_{\alpha_0}(\nu) + \frac{i}{2}[\nu \wedge \nu]\). Hence \(\|\mu_\nu^* \alpha_\nu - \alpha_0\|_{W^{1,q}} \to 0\), in contradiction to (25). This proves the claim.

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

\[
\text{NS}_P(\alpha) := \exp(i \Xi(\alpha))^* \alpha.
\]

Recall the definition of \(\mathcal{F}\) in (16). In light of (15), finding \(\Xi(\alpha)\) is equivalent to solving for \(\xi\) in \(\mathcal{F}(\alpha, \xi) = 0\), for then \(\Xi(\alpha) := -\xi\). To solve for \(\xi\) we need to pass to suitable Sobolev completions.

It follows from (14), and the Sobolev embedding and multiplication theorems, that \(\mathcal{F}\) extends to a smooth map \(A^{1,q}(P) \times \text{Lie}(G(P))^{2,q} \to \text{Lie}(\mathcal{G}(P))^{0,q}\), whenever \(q > 1\). Suppose \(\alpha_0\) is a flat connection. The linearization of \(\mathcal{F}\) at \((\alpha_0, 0)\) in the direction of \((0, \xi)\) is
\[ D_{(\alpha,0)} F(0, \xi) = 2 J_E * \partial_{\alpha,0} \xi, \]
where we have used the fact that \( d_{\alpha,0} \) commutes with \( J_E = i \) (this is because the complex structure \( J_E \) is constant and the elements of \( A(P) \) are unitary). Observe that \( J_F \) acts by the Hodge star on vectors, so

\[ d_{\alpha,0}(d_{\alpha,0} \circ j_\Sigma) = d_{\alpha,0} \circ d_{\alpha,0}, \quad (d_{\alpha,0} \circ j_\Sigma) d_{\alpha,0} = F_{\alpha,0} \circ (j_\Sigma, \text{Id}). \]

Using this and the fact that \( F_{\alpha,0} = 0 \), we have

\[ D_{(\alpha,0)} F(0, \xi) = \frac{1}{2} \Delta_{\alpha,0} \xi \]
where \( \Delta_{\alpha,0} = d_{\alpha,0}^* d_{\alpha,0} + d_{\alpha,0} d_{\alpha,0}^* \) is the Laplacian. By assumption, all flat connections are irreducible, so Hodge theory tells us that the operator \( \Delta_{\alpha,0} \) : \( \text{Lie}(G(P))^{2,q} \rightarrow \text{Lie}(G(P))^{0,q} \) is an isomorphism. Since \( \alpha_0 \) is flat, the pair \((\alpha_0,0)\) is clearly a solution to \( F(\alpha, \xi) = 0 \). It therefore follows by the implicit function theorem that there are \( \epsilon_{\alpha_0}, \epsilon'_{\alpha_0} > 0 \) such that, for any \( \alpha \in A^{1, q} \) with \( \|\alpha - \alpha_0\|_{W^{1,q}} < \epsilon_{\alpha_0} \), there is a unique \( \Xi = \Xi(\alpha) \in \text{Lie}(G(P))^{2,q} \) with \( \|\Xi(\alpha)\|_{W^{2,q}} < \epsilon'_{\alpha_0} \) and \( F(\alpha, -\Xi(\alpha)) = 0 \). The implicit function theorem also implies that \( \Xi(\alpha) \) varies smoothly in \( \alpha \) in the \( W^{1,q} \)-topology. Moreover, by the uniqueness assertion, it follows that \( \Xi(\alpha) = 0 \) if \( \alpha \) is flat.

We need to show that \( \epsilon_{\alpha_0} \) and \( \epsilon'_{\alpha_0} \) can be chosen to be independent of \( \alpha_0 \in A_{\text{flat}, G}(P) \). Since the moduli space \( A_{\text{flat}}/G \) of flat connections on \( P \) is compact, it suffices to show that \( \epsilon_{\alpha_0} = \epsilon_{\mu^* \alpha_0} \), for all real gauge transformations \( \mu \in G^{2,q}(P) \), and likewise for \( \epsilon'_{\alpha_0} \). Fix \( \mu \in G^{2,q}(P) \) and a connection \( W^{1,q} \)-close to \( \alpha_0 \), then find \( \Xi(\alpha) \) as above. By [3] and the statement following (10) we have

\[ \exp(i \Xi(\alpha)) \mu = \mu \exp(i \text{Ad}(\mu^{-1}) \Xi(\alpha)). \]

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

\[ 0 = \text{Ad}(\mu^{-1}) F_{\exp(i \Xi)^* \alpha} = F_{(\exp(i \Xi) \mu)^* \alpha} = F_{\exp(i \text{Ad}(\mu^{-1}) \Xi)^* (\mu^* \alpha)} \]

so \( \Xi(\mu^* \alpha) = \text{Ad}(\mu^{-1}) \Xi(\alpha) \) since \( \Xi(\mu^* \alpha) \) is uniquely defined by \( F_{\exp(i \Xi(\mu^* \alpha))} (\mu^* \alpha) = 0 \). It follows immediately that \( \epsilon_{\mu^* \alpha_0} = \epsilon_{\alpha_0} \) and \( \epsilon'_{\mu^* \alpha_0} = \epsilon'_{\alpha_0} \), so we can take \( \epsilon_0 \) to be the minimum of

\[ \inf_{[\alpha_0] \in A_{\text{flat}}/G} \epsilon_{\alpha_0} > 0 \quad \text{and} \quad \inf_{[\alpha_0] \in A_{\text{flat}}/G} \epsilon'_{\alpha_0} > 0. \]

This argument also shows that \( \text{NS}_P \) is \( G^{2,q}(P) \)-equivariant.

Finally, we show that \( \text{NS}_P(\alpha) \) depends smoothly on \( \alpha \) in the \( L^p \)-topology for \( p > 2 \). It suffices to show that \( \alpha \mapsto \Xi(\alpha) \) extends to a map \( A^{0,p}(P) \rightarrow \text{Lie}(G(P))^{1,p} \) that is smooth with respect to the specified topologies. To see this, note that \( F \) from (10) is well-defined as a map

\[ A^{0,p}(P) \times \text{Lie}(G(P))^{1,p} \rightarrow \text{Lie}(G(P))^{-1,p}. \]

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

\[ A^{0,p}_{\text{flat}}(P) \rightarrow \text{Lie}(G(P))^{1,p}, \quad \alpha \mapsto \Xi(\alpha) \]

is smooth. The uniqueness of \( \Xi(\alpha) \) and \( \Xi(\alpha) \) ensures that the former is indeed an extension of the latter. \( \square \)
Remark 3.8. Let $\Pi : A^{1,q}_{\text{flat}}(P) \to A^{1,q}_{\text{flat}}(P)/G^{2,q}(P)$ denote the projection. The above proof shows that the composition $\Pi \circ \text{NS}$ is invariant under a small neighborhood of $G^{2,q}(P)$ in $G^{2,q}(P)^C$. Indeed, $\alpha$ and $\exp(i\xi)^{\alpha}$ both map to the same flat connection under NS whenever they are both in the domain of NS.

3.1.2. 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 $d_{\alpha}$ on 0-forms.

Lemma 3.9. Suppose $G$ is a compact Lie group, $\Sigma$ 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 $\alpha \in A^{1,q}(P)$ with $\|F_{\alpha}\|_{L^q(\Sigma)} < \epsilon_0$, or $\alpha \in A^{0,q}(P)$ with $\|\alpha - \alpha_0\|_{L^2(\Sigma)} < \epsilon_0$ for some $\alpha_0 \in A^{0,2q}(P)$. Then the map $d_{\alpha} : W^{1,q}(P(\mathfrak{g})) \to L^q(P(\mathfrak{g}))$ is a Banach space isomorphism onto its image. Moreover, for all $f \in W^{1,q}(P(\mathfrak{g}))$ the following holds

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

(ii) For all $\alpha \in A^{1,q}(P)$ with $\|F_{\alpha}\|_{L^q(\Sigma)} < \epsilon_0$, the Laplacian $d_{\alpha}^* d_{\alpha} : W^{2,q}(P(\mathfrak{g})) \to L^q(P(\mathfrak{g}))$ is a Banach space isomorphism. Moreover, for all $f \in W^{2,q}(P(\mathfrak{g}))$ the following holds

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

Proof. This is basically the statement of [Lemma 7.6], but adjusted a little to suit our situation. We prove (ii), the proof of (i) is similar. The assumption that all flat connections $\alpha_0$ are irreducible implies that the kernel and cokernel of the elliptic operator $d_{\alpha}^* d_{\alpha} : W^{2,q}(P(\mathfrak{g})) \to L^q(P(\mathfrak{g}))$ are trivial. In particular, we have an estimate $\|f\|_{W^{2,q}} \leq C\|d_{\alpha} \ast d_{\alpha} f\|_{L^q}$ for all $f \in W^{2,q}(P(\mathfrak{g}))$, so the statement of the lemma holds when $\alpha = \alpha_0$ is flat.

Next, fix $\alpha \in A^{1,q}(P)$ and $\alpha_0 \in A^{1,q}_{\text{flat}}(P)$. Then, by the above discussion, and the relation $d_{\alpha_0} f = d_{\alpha} f + [\alpha_0 - \alpha, f]$, we have that $\|f\|_{W^{2,q}}$ is bounded by

$$C\|d_{\alpha_0} \ast d_{\alpha} f\|_{L^q} \leq \left\{ \|d_{\alpha} \ast d_{\alpha} f\|_{L^q} + \|d_{\alpha} [s(\alpha - \alpha_0), f]\|_{L^q} + \|[\alpha - \alpha_0, \ast (\alpha - \alpha_0), f]\|_{L^q} \right\} \leq C\left\{ \|d_{\alpha_0} \ast d_{\alpha} f\|_{L^q} + C' \|f\|_{W^{2,q}} \left( \|d_{\alpha} [\ast (\alpha - \alpha_0), f]\|_{L^q} + \|[\alpha - \alpha_0, \ast (\alpha - \alpha_0)]\|_{L^q} \right) \right\},$$

for all $f \in W^{2,q}(P(\mathfrak{g}))$, where we have used the embeddings $W^{2,q} \hookrightarrow W^{1,q}$ and $W^{2,q} \hookrightarrow L^\infty$ in the last step. Now suppose that $[\alpha - \alpha_0]_{L^2} < 1/2C'$ is small. Then by composing $\alpha_0$ with a suitable gauge transformation, we may suppose $\alpha$ is in Coulomb gauge with respect to $\alpha_0$, and still retain the fact that $[\alpha - \alpha_0]_{L^2} < 1/2C'$. Then the above gives

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

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

Now we prove the lemma. Suppose (ii) in the statement of the lemma does not hold. Then there is some sequence of connections $\alpha_\nu$ with $\|F_{\alpha_\nu}\|_{L^q} \to 0$, but the estimate (29) does not hold for any $C > 0$. By Uhlenbeck’s weak compactness theorem, after possibly passing to a subsequence, there is some sequence of gauge transformations $u_{\nu}$, and a limiting flat connection $\alpha_0$, such that $\|\alpha_\nu - u_{\nu}^* \alpha_0\|_{L^2} \to 0$. So the discussion of the previous paragraph shows that, for $\nu$ sufficiently large, the estimate (29) holds with $\alpha$ replaced by $\alpha_0$. This is a contradiction, and it proves the lemma. □

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

$$W^{k,q}(T^*\Sigma \otimes \mathcal{P}(\mathfrak{g})) = H^{1}_{\alpha} \oplus (\text{im } d_{\alpha})^{k,q} \oplus (\text{im } d_{\alpha}^{*})^{k,q},$$

for any flat connection $\alpha$. Here $H^{1}_{\alpha}$ is finite dimensional (this dimension is independent of $\alpha \in \mathcal{A}_{0}$), and so is equal to its own $W^{k,q}$-closure. Furthermore, the direct sum in (30) is $L^{2}$-orthogonal, even though the spaces need not be complete in the $L^{2}$-metric. We have a similar situation whenever $\alpha$ has small curvature, as the next lemma shows.

**Lemma 3.10.** Assume that $P \to \Sigma$ satisfies the conditions of Lemma 3.9, 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 $\|F_{\alpha}\|_{L^{r}(\Sigma)} < \epsilon_{0}$, then

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

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

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

and so we have a direct sum decomposition

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

In particular, the $L^{2}$-orthogonal projection

$$\text{proj}_{\alpha} : W^{k,q}(T^*\Sigma \otimes \mathcal{P}(\mathfrak{g})) \to H^{1}_{\alpha}$$

is well-defined.

**Remark 3.11.** It follows by elliptic regularity that the space $H^{1}_{\alpha} = (\ker d_{\alpha})^{k,q} \cap (\ker d_{\alpha}^{*})^{k,q}$ consists of smooth forms. Moreover, 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

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

and this map is onto by dimensionality. Hence, the definition of $H^{1}_{\alpha}$ is independent of the choice of Sobolev constants $k, q$.

**Proof of Lemma 3.10.** We first show that, when $\|F_{\alpha}\|_{L^{r}(\Sigma)}$ is sufficiently small, we have a direct sum decomposition

$$W^{k,q}(T^*\Sigma \otimes \mathcal{P}(\mathfrak{g})) = H^{1}_{\alpha} \oplus (\text{im } d_{\alpha})^{k,q} \oplus (\text{im } d_{\alpha}^{*})^{k,q}.$$ 

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

$$[F_{\alpha}, f] = d_{\alpha} \ast d_{\alpha} g, \quad [F_{\alpha}, g] = -d_{\alpha} \ast d_{\alpha} f.$$ 

A priori, $d_{\alpha} \ast d_{\alpha} g$ and $d_{\alpha} \ast d_{\alpha} f$ 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 3.9 and the embedding $W^{2,q} \hookrightarrow L^{\infty}$, it follows that, whenever $\|F_{\alpha}\|_{L^{r}}$ is sufficiently small, we have
\[ \|f\|_{L^\infty} \leq C \|d_\alpha \ast d_\alpha f\|_{L^q} = C \| [F_\alpha, g] \|_{L^q} \leq 2C \|F_\alpha\|_{L^q} \|g\|_{L^\infty}. \]

Similarly, \( \|g\|_{L^\infty} \leq 2C \|F_\alpha\|_{L^q} \|g\|_{L^\infty}, \) and hence

\[ \|f\|_{L^\infty} \leq 4C^2 \|F_\alpha\|^2_{L^q} \|f\|_{L^\infty}. \]

If \( \|F_\alpha\|^2_{L^q} < (2C)^{-2} \), then this can happen only if \( f = g = 0 \). This establishes the direct sum (31).

Now we prove that the dimension of \( H^1_\alpha \) is finite and equals that of \( H^1_{\alpha_0} \) for any flat connection \( \alpha_0 \). It is well-known that the operator \( \circ \) is elliptic, and hence Fredholm, whenever \( \alpha_0 \) is flat. The irreducibility condition implies that it has trivial kernel, and so has index given by \( -\dim(H^1_\alpha) \), which is a constant independent of \( \alpha_0 \). Then for any other connection \( \alpha \), the operator \( d_\alpha \oplus *d_\alpha \) differs from \( d_{\alpha_0} \oplus *d_{\alpha_0} \) by the compact operator

\[ \mu \mapsto [\alpha - \alpha_0, \mu] + [\alpha - \alpha_0, \mu], \]

and so \( d_\alpha \oplus *d_\alpha \) is Fredholm with the same index \( -\dim(H^1_{\alpha_0}) \). It follows from Lemma 3.14 that the (bounded) operator

\[ d_\alpha \oplus *d_\alpha : W^{k+1,q} (P(g)) \oplus W^{k+1,q} (P(g)) \to W^{k,q} (T^*\Sigma \otimes P(g)) \]

is injective whenever \( \|F_\alpha\|_{L^q(\Sigma)} \) is sufficiently small, and hence the cokernel has finite dimension \( \dim(H^1_{\alpha_0}) \), so this finishes the proof of Lemma 3.10.

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

**Proposition 3.12.** Suppose that \( P \to \Sigma \) and \( \epsilon_0 > 0 \) are as in Lemma 3.10, and let \( 1 < q < \infty \). Then the assignment \( \alpha \mapsto \text{proj}_\alpha \) is affine-linear and bounded

\[ \|\text{proj}_\alpha - \text{proj}_{\alpha'}\|_{L^q(\Sigma)} \leq C \|\alpha - \alpha'\|_{L^q(\Sigma)}, \]

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

**Proof.** We will see that defining equations for \( \text{proj}_\alpha \) 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:

\[ W^{k,q}(\Omega^i) := W^{k,q}(\Lambda^i T^*\Sigma \otimes P(g)), \quad L^q(\Omega^i) := W^{0,q}(\Omega^i). \]

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

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

**Property B:** \( \langle \text{proj}_\alpha \mu, d_\alpha a + d_\alpha^* b \rangle = 0, \) for all \( (a, b) \in W^{1,\ast q}(\Omega^0) \oplus W^{1,q}(\Omega^2), \)

where \( q^\ast \) is the Sobolev dual to \( q: 1/q + 1/q^\ast = 1 \). Here and below, we use the notation \( \langle \mu, v \rangle \) to denote the \( L^2 \)-pairing on forms. Note that by Lemma 3.9 the operators \( d_\alpha, d_\alpha^* \) are injective on 0- and 2-forms, respectively, so any pair \( (u, v) \) satisfying Property A is unique.

Consider the map

\[ (A^0_q \oplus L^q(\Omega^1)) \times (L^q(\Omega^1) \oplus W^{1,q}(\Omega^0) \oplus W^{1,q}(\Omega^0)) \to (W^{1,q}(\Omega^0)^* \oplus (W^{1,q}(\Omega^2)^* \oplus L^q(\Omega^1)) \]
defined by

\[(\alpha, \mu; \nu, u, v) \mapsto \left( (\nu, d_\alpha(\cdot)), \langle \nu, d^*_\alpha(\cdot) \rangle, \mu - \nu - d_\alpha u - d^*_\alpha v \right)\]

The key point is that a tuple \((\alpha, \mu; \nu, u, v)\) maps to zero under \((34)\) if and only if this tuple satisfies Properties A and B above. By the identification \((W^{k,q})^* = W^{-k,q}\), we can equivalently view \((34)\) as a map

\[
(\mathcal{A}^{0,q} \times L^q(\Omega^1)) \times (L^q(\Omega^1) \times W^{1,q}(\Omega^0) \times W^{1,q}(\Omega^0)) \rightarrow W^{-1,q}(\Omega^0) \oplus W^{-1,q}(\Omega^2) \oplus L^q(\Omega^1)
\]

defined by

\[(\alpha, \mu; \nu, u, v) \mapsto (d^*_\alpha \nu, d_\alpha \nu, \mu - \nu - d_\alpha u - d^*_\alpha v).\]

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

Claim 1: The map \((35)\) is bounded affine linear in the \(A^{0,q}\)-variable, and bounded linear in the other 4 variables.

Claim 2: The linearization at \((\alpha, 0; 0, 0, 0)\) of \((35)\) in the last 3-variables is a Banach space isomorphism, provided \(\|\alpha - \alpha_b\|_{L^q}\) is sufficiently small for some flat connection \(\alpha_b\).

Before proving the claims, we describe how they prove the lemma. Observe that \((\alpha, 0; 0, 0, 0)\) is clearly a zero of \((35)\) for any \(\alpha\). Claim 1 implies that \((35)\) is smooth, and so by Claim 2 we can use the implicit function theorem to show that, for each pair \((\alpha, \mu) \in A^{0,q} \oplus L^q(\Omega^1)\), with \(\|\alpha - \alpha_b\|_{L^q}\) 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 \((35)\). (A priori this only holds for \(\mu\) in a small neighborhood of the origin, but since \((35)\) is linear in that variable, it extends to all \(\mu\).) It will then follow that \(\nu = \text{proj}_\alpha \mu\) depends smoothly on \(\alpha\) in the \(L^q\) metric. In fact, \((35)\) is affine linear in \(\alpha\) and linear in the other variables, so the uniqueness assertion of the implicit function theorem implies that \(\text{proj}_\alpha\) depends affine-linearly on \(\alpha\), and so it follows that \(\|\text{proj}_\alpha - \text{proj}_{\alpha_b}\|_{\text{op}, L^q}\) is bounded by

\[
\inf_{\|\mu\|_{L^q} = 1} \|\bigl(\text{proj}_\alpha - \text{proj}_{\alpha_b}\bigr) \mu\|_{L^q} \leq C \inf_{\|\mu\|_{L^q} = 1} \|\alpha - \alpha_b\|_{L^q} \|\mu\|_{L^q} = C \|\alpha - \alpha_b\|_{L^q}.
\]

This proves the lemma for all \(\alpha\) sufficiently \(L^q\)-close to \(\mathcal{A}_{\text{flat}}\). To extend it to all \(\alpha\) with \(\|F_\alpha\|_{L^q}\) sufficiently small, one argues by contradiction as in the proof of Lemma\((34)\) using Uhlenbeck’s compactness theorem. 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

\[
(\mathcal{A}^{0,q} \times L^q(\Omega^1)) \rightarrow W^{-1,q}(\Omega^0), \quad (\alpha, \nu) \mapsto d^*_\alpha \nu
\]

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 \((36)\) is bounded in each of the two coordinates separately. Fix \(\alpha\) and a flat connection \(\alpha_b\). Then

\[
\|d_\alpha \nu\|_{W^{-1,q}} \leq \|d_\alpha \nu\|_{W^{-1,q}} + \|\alpha - \alpha_b \wedge \nu\|_{W^{-1,q}} \\
\leq \|d_\alpha \nu\|_{W^{-1,q}} + 2 \|\alpha - \alpha_b\|_{L^q} \|\nu\|_{L^q} \\
\leq C(1 + \|\alpha - \alpha_b\|_{L^q}) \|\nu\|_{L^q}
\]

which shows that the map is bounded in the variable \(\nu\), with \(\alpha\) fixed. Next, fix \(\nu\) and write
First note that this is just the standard elliptic regularity result if $\eta = 0$. This shows the first component of (38) is bounded. The other two components are similar.

**Proof of Claim 2:** The linearization of (38) at $(\alpha, 0; 0, 0, 0)$ in the last three variables is the map

$$L^q(\Omega^1) \times W^{1,q}(\Omega^0) \times W^{1,q}(\Omega^0) \rightarrow W^{-1,q}(\Omega^0) \oplus W^{-1,q}(\Omega^2) \oplus L^q(\Omega^1)$$

$$(\nu, u, v) \mapsto (d^*_\alpha \nu, d_\alpha \nu, -\nu - d_\alpha u - d^*_\alpha v)$$

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

$$(d^*_\alpha \nu, d_\alpha \nu, -\nu - d_\alpha u - d^*_\alpha v) = (0, 0, 0).$$

Then by Lemma 3.10 we can write $\nu$ uniquely as $\nu = \nu_H + d_\alpha a + d^*_\alpha b$ for $\nu_H \in H^1_\alpha \cap \ker d^*_\alpha$, 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 (37), 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 3.9 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

$$(f, g, \eta) \in \left(W^{-1,q}(\Omega^0) \oplus W^{-1,q}(\Omega^2) \oplus L^q(\Omega^1)\right)^* = W^{1,q}(\Omega^0) \oplus W^{1,q}(\Omega^2) \oplus L^q(\Omega^1)$$

with

$$0 = \langle f, d^*_\alpha \nu \rangle, \quad 0 = \langle g, d_\alpha \nu \rangle, \quad 0 = \langle \eta, \nu + d_\alpha u + d^*_\alpha v \rangle$$

for all $(\nu, u, v)$. The first two equations imply

$$0 = \langle d_\alpha f, \nu \rangle, \quad 0 = \langle d^*_\alpha g, \nu \rangle$$

for all $\nu$. This implies $d_\alpha f = 0$ and $d^*_\alpha g = 0$, and so $f = 0$ and $g = 0$ by Lemma 3.9. For the third equation, take $(u, v) = (0, 0)$ and we get $0 = \langle \eta, \nu \rangle$ for all $\nu$. But this can only happen if $\eta = 0$, which is a contradiction to the tuple $(f, g, \eta)$ being non-zero. \qed

We end this preparatory section by establishing the analogue of Lemma 3.9 for 1-forms.

**Lemma 3.13.** Assume that $P \rightarrow \Sigma$ satisfies the conditions of Lemma 3.7 and let $1 < q < \infty$. Then there are constants $C > 0$ and $c_0 > 0$ such that

$$(38) \quad \|\eta - \text{proj}_\alpha \eta\|_{W^{1,q}(\Sigma)} \leq C \left(\|d_\alpha \eta\|_{L^q(\Sigma)} + \|d_\alpha \ast \eta\|_{L^q(\Sigma)}\right)$$

for all $\eta \in W^{1,q}(T^* \Sigma \otimes P(g))$ and all $\alpha \in A^{1,q}(\Sigma)$ with $\|F_\alpha\|_{L^q(\Sigma)} < c_0$.

**Proof.** First note that this is just the standard elliptic regularity result if $\alpha = \alpha_0$ is flat. To prove the lemma, suppose the conclusion does not hold. Then there is sequence of connections $\alpha_\nu$ with $\|F_\alpha\|_{L^q} \rightarrow 0$, and a sequence of 1-forms $\eta_\nu \in \text{im } d_{\alpha_\nu} \oplus \text{im } d_{\alpha_\nu}$ with $\|\eta_\nu\|_{W^{1,q}} = 1$ and

$$\|d_{\alpha_\nu} \eta_\nu\|_{L^q} + \|d_{\alpha_\nu} \ast \eta_\nu\|_{L^q} \rightarrow 0.$$
By applying suitable gauge transformations to the $\alpha_\nu$, and by passing to a subsequence, it follows from Uhlenbeck compactness that the $\alpha_\nu$ converge strongly in $L^{2q}$ to a limiting flat connection $\alpha_\flat$. So we have

$$
\|d_{\alpha_\nu} \eta_\nu\|_{L^q} \leq \|d_{\alpha_\nu} \eta_\nu\|_{L^q} + \|\alpha_\nu - \alpha_\flat \wedge \eta_\nu\|_{L^q} \leq \|d_{\alpha_\nu} \eta_\nu\|_{L^q} + 2\|\alpha_\nu - \alpha_\flat\|_{L^{2q}} \|\eta_\nu\|_{L^{2q}}
$$

$$
\leq \|d_{\alpha_\nu} \eta_\nu\|_{L^q} + C_0 \|\alpha_\nu - \alpha_\flat\|_{L^{2q}} \|\eta_\nu\|_{W^{1,q}} \to 0,
$$

where in the last inequality we have used the embedding $W^{1,q} \hookrightarrow L^{2q}$ for $q > 1$. Similarly $\|d_{\alpha_\nu} * \eta_\nu\|_{L^q} \to 0$. By the elliptic estimate for the flat connection $\alpha_\flat$, we have

$$
(39) \quad \|\eta_\nu - \text{proj}_{\alpha_\nu} \eta_\nu\|_{W^{1,q}} \leq C_1 (\|d_{\alpha_\nu} \eta_\nu\|_{L^q} + \|d_{\alpha_\nu} * \eta_\nu\|_{L^q}) \to 0.
$$

On the other hand, by Proposition 3.12, the projection operator $\text{proj}_{\alpha_\nu}$ is converging in the $L^q$ operator norm to $\text{proj}_{\alpha_\flat}$. In particular,

$$
\|\text{proj}_{\alpha_\nu} (\eta_\nu)\|_{W^{1,q}} \leq C_2 \|\text{proj}_{\alpha_\nu} (\eta_\nu)\|_{L^q} = C_2 \|\text{proj}_{\alpha_\nu} (\eta_\nu) - \text{proj}_{\alpha_\nu} (\eta_\nu)\|_{L^q}
$$

$$
\leq C_2 \|\alpha_\nu - \alpha_\flat\|_{L^q} \|\eta_\nu\|_{L^q} \to 0,
$$

where the first inequality holds because $H^1_{\alpha_\flat}$ is finite-dimensional (and so all norms are equivalent), the second inequality holds by Proposition 3.12 and the convergence to zero holds since

$$
\|\eta_\nu\|_{L^q} \leq C_3 \|\eta_\nu\|_{W^{1,q}} = C_3
$$

is bounded. Combining this with (39) gives

$$
1 = \|\eta_\nu\|_{W^{1,q}} \leq \|\eta_\nu - \text{proj}_{\alpha_\nu} \eta_\nu\|_{W^{1,q}} + \|\text{proj}_{\alpha_\nu} (\eta_\nu)\|_{W^{1,q}} \to 0,
$$

which is a contradiction, proving the lemma. □

3.1.3. 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 3.14.** Let $\text{NS}_P$ be the map (22), and $3/2 \leq q < \infty$. Then there are constants $C > 0$ and $\epsilon_0 > 0$ such that

$$
(40) \quad \|\text{NS}_P(\alpha) - \alpha\|_{W^{1,q}(\Sigma)} \leq C \|F_\alpha\|_{L^{2q}(\Sigma)}
$$

for all $\alpha \in A^{1,q}(P)$ with $\|F_\alpha\|_{L^{2q}(\Sigma)} < \epsilon_0$.

**Proof.** The basic idea is that $\text{NS}_P(\alpha) = \exp(i \Xi(\alpha))^* \alpha$ can be expressed as a power series, with lowest order term given by $\alpha$ (see the proof of Theorem 4.6 for the definition of $\Xi(\alpha)$). 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 $\text{Lie}(G(E)^C)^{2,q}$ can be viewed as the $W^{2,q}$-completion of the vector space $\Gamma(P \times_G \text{End}(C^n))$. Since $q > 1$, we are in the range in which pointwise matrix multiplication is well-defined, and $\text{Lie}(G(E)^C)^{2,q}$ becomes a Banach algebra. Then for any $\xi \in \text{Lie}(G(E)^C)^{2,q}$, the power series

$$
\sum_{k=0}^\infty \frac{\xi^k}{k!} \in \text{Lie}(G(E)^C)^{2,q}
$$

converges, where $\xi^k$ is $k$-fold matrix multiplication on the values of $\xi$. 
Remark 3.15. Note that we may not have $\xi^k \in \text{Lie}(\mathcal{G}(P)_{\mathbb{C}})^{2,q}$, even when $\xi \in \text{Lie}(\mathcal{G}(P)_{\mathbb{C}})^{2,q}$ (however, it is always the case that $\xi^k \in \text{Lie}(\mathcal{G}(E)_{\mathbb{C}})^{2,q}$). In particular, the infinitesimal action of $\xi^k$ on $A^{1,q}(P)$ need not lie in the tangent space to $A^{1,q}(P)$, though it will always lie in the tangent space to $A^{1,q}(E)$.

As with finite-dimensional Lie theory, this power series represents the exponential map

$$\exp : \text{Lie}(\mathcal{G}(E)_{\mathbb{C}})^{2,q} \rightarrow \mathcal{G}^{2,q}(E)_{\mathbb{C}},$$

where we are using the inclusion $\mathcal{G}^{2,q}(E)_{\mathbb{C}} \subset W^{2,q}(P \times_G \text{End}(\mathbb{C}^n))$. The power series defining $\exp$ continues to hold on the restriction

$$\exp : \text{Lie}(\mathcal{G}(P)_{\mathbb{C}})^{2,q} \rightarrow \mathcal{G}^{2,q}(P)_{\mathbb{C}}.$$ 

Similarly, the usual power series definitions of $\sin$ and $\cos$ hold in this setting:

$$\sin, \cos : \text{Lie}(\mathcal{G}(P)_{\mathbb{C}})^{2,q} \rightarrow W^{2,q}(P \times_G \text{End}(\mathbb{C}^n)),$$

and we have the familiar relation $\exp(i\xi) = \cos(\xi) + i\sin(\xi)$. Note that by [12], for any real $\xi \in \text{Lie}(\mathcal{G}(E)_{\mathbb{C}})^{2,q}$ and $\alpha \in A^{1,q}(P)$, we have

(41) $\exp(i\xi)^* \alpha - \alpha = - \left\{ d_\alpha \left( \cos(\xi) - 1 \right) + *d_\alpha \left( \sin(\xi) \right) \right\} \in T_\alpha A^{1,q}(P) \subset W^{1,q}(T^* \Sigma \otimes P \times_G \text{End}(\mathbb{C}^n))$,

where the action of $d_\alpha$ 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 $\alpha \in A^{1,q}(P)$ satisfies $\|F_\alpha\|_{L^q(\Sigma)} < \epsilon_0$ and $\|\Xi(\alpha)\|_{W^{2,q}} < \epsilon_0$, then

$$\|\text{NS}_P(\alpha) - \alpha\|_{W^{1,q}} \leq C\|F_\alpha\|_{L^{2,q}}.$$ 

This is exactly the statement of the proposition, except we have an additional assumption on $\Xi(\alpha)$. However, we know $\Xi(\alpha)$ depends continuously on $\alpha$ and vanishes when $\alpha$ is flat, so this additional assumption is superfluous.

Set $\Xi = \Xi(\alpha)$ and $\eta := \text{NS}_P(\alpha) - \alpha$. Using the power series expansion of $\exp$, we have

$$\eta = \exp(i\Xi)^* \alpha - \alpha = - *d_{\alpha,\rho} \Xi + \frac{(\Xi(d_{\alpha,\rho} \Xi) + (d_{\alpha,\rho} \Xi))}{2} + \ldots$$

where the $n$th term in the sum on the right has the form

$$-\frac{1}{n!} \sum_{k=0}^{n} \Xi \ldots \Xi(d_{\alpha,\rho} \Xi) \Xi \ldots \Xi$$

with $n$ copies of $\Xi$ appearing before $d_{\alpha,\rho} \Xi$, and $n - k - 1$ copies after. By assumption $2q > 2$, and so

$$\left\| \frac{1}{n!} \sum_{k=0}^{n} \Xi \ldots \Xi(d_{\alpha,\rho} \Xi) \Xi \ldots \Xi \right\|_{W^{1,q}} \leq \|d_{\alpha,\rho} \Xi\|_{W^{1,2q}} \left( \frac{C_1^{2n}}{n!} \sum_{k=0}^{n} \|\Xi\|_{W^{1,2q}}^{n-1} \right) \leq C_1^{2n} \frac{1}{(n-1)!} \|\Xi\|_{W^{1,2q}}^{n-1},$$

where $C_1$ is the constant from the Sobolev multiplication theorem. This gives

$$\|\eta\|_{W^{1,q}} \leq C_2 \|\eta\|_{W^{1,2q}} \leq C_2 \|\Xi\|_{W^{2,q}} \sum_{n=1}^{\infty} \frac{C_1^{2n}}{(n-1)!} \|\Xi\|_{W^{1,2q}}^{n-1}.$$ 

Whenever $\|\Xi\|_{W^{1,2q}} \leq 1$, we therefore have
Now use the elliptic estimate from Lemma 3.9 (ii):

\[ k_d(43) \]  

curves in \( M \) coming from the identity \( 0 = \) 

where the second inequality is (43), and in the last inequality we used 

\[ \parallel \text{following lemma will be used to show that holomorphic curves in} \]

\( A \)

and \( \Pi \) are smooth. We will be interested in estimating the derivative of the composition \( \Pi \)

\[ (42) \]

\[ \parallel \text{is controlled by that of} \]

\( F \)

\[ \parallel \text{so we can drop the subscript} \]

\( \rho \)

\[ \parallel \text{small we have} \]

\[ \parallel \text{where we have used the embedding} \]

\( W \)

\[ \parallel \text{and} \]

\( F \)

\[ \parallel \text{NS} \]

\[ \parallel \text{by requiring that} \]

\( 10 \)

\[ \parallel \text{and} \]

\( 1 \)

\[ \parallel \text{this gives} \]

\[ \parallel \text{for} \]

\( \parallel \text{and} \]

\( 2 \)

\[ \parallel \text{where we have used the embedding} \]

\( W \)

\[ \parallel \text{which holds provided} \]

\( q \geq 3/2 \)

This gives 

\[ \parallel \text{where} \]

\( \parallel \text{and} \]

\( \parallel \text{completes the proof since we can ensure that} \]

\( \parallel \text{is sufficiently small (when} \]

\( \Xi = 0 \), it follows that \( \eta = 0 \), and everything is continuous in these 

Throughout the remainder of this section, we assume that \( \Sigma \) is closed, connected and orientable, 

\( G = \text{PU}(r) \), and \( P \to \Sigma \) is a bundle for which \( t_2(P) \) \( \Sigma \) is a generator. Let \( \Pi : A_{\text{flat}}^1(P) \to M(P) = A_{\text{flat}}^1(P)/G_0^2(P) \) denote the quotient map. The assumptions on \( G \) and \( P \) imply that \( M(P) \) and \( \Pi \) are smooth. We will be interested in estimating the derivative of the composition \( \Pi \circ \text{NS}\). 

Any choice of orientation and metric on \( \Sigma \) determines complex structures on the tangent bundles \( T_A^1(P) \) and \( TM(P) \) that is induced by the Hodge star on 1-forms. Denote by 

\[ D_A^k (\Pi \circ \text{NS}) : T_{\alpha}A^1(P) \otimes \ldots \otimes T_{\alpha}A^1(P) \to T_{\text{NS}(\alpha)}M(P) \]  

the \( k \)th derivative of \( \Pi \circ \text{NS} \) at \( \alpha \), defined with respect to the \( W^1,q \)-topology on the domain. The following lemma will be used to show that holomorphic curves in \( A^1,P \) descend to holomorphic curves in \( M(P) \).
Lemma 3.16. Suppose $G = PU(r)$, $\Sigma$ is a closed, connected, oriented Riemannian surface, and $P \to \Sigma$ is a principal $G$-bundle with $t_2(P)[\Sigma] \in \mathbb{Z}_r$, a generator. Let $1 < q < \infty$ and suppose $\alpha$ is in the domain of $\text{NS}_P$. Then the linearization $D_\alpha(\Pi \circ \text{NS}_P)$ is complex-linear:

\[ *D_\alpha(\Pi \circ \text{NS}_P) = D_\alpha(\Pi \circ \text{NS}_P) * . \]

Proof. We refer to the notation of Section 2.2. The complex gauge group $\mathcal{G}(P)^C$ 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 $\mu \in \mathcal{G}^2_q(P)^C$, $\alpha \in \mathcal{A}^{1,q}(P)$ and $\eta \in W^{1,q}(T^*\Sigma \otimes P(g))$. Then by (1) we have

\[
\frac{d}{dt}|_{t=0} \mu \circ \overline{\partial}_{\alpha+\tau \xi} \circ \mu^{-1} = \frac{d}{dt}|_{t=0} \mu \circ \overline{\partial}_o \circ \mu^{-1} + \tau \mu \circ (\ast \eta^{0,1}) \circ \mu^{-1} = * (\mu \circ \eta^{0,1} \circ \mu^{-1}),
\]

which shows the infinitesimal action of the complex gauge group is complex-linear.

Let $\mathcal{G}_0^{2,q}(P)^C \subseteq \mathcal{G}^{2,q}(P)^C$ denote the identity component. This can be described as

\[
\mathcal{G}_0^{2,q}(P)^C = \left\{ \mu \exp(i\xi) \mid \mu \in \mathcal{G}_0^{2,q}(P), \xi \in W^{2,q}(P(g)) \right\}.
\]

It follows from (1) and Lemma 3.5 that $\mathcal{G}_0^{2,q}(P)^2$ acts freely on the space of connections. Moreover, by Remark 3.5 the map $\text{NS}_P$ is equivariant under a neighborhood of $\mathcal{G}_0^{2,q}(P)$ in $\mathcal{G}_0^{2,q}(P)^C$. These two facts imply that $\text{NS}_P$ has a unique $\mathcal{G}_0^{2,q}(P)^C$-equivariant extension to the flow-out

\[
\mathcal{A}^s(P) := \left( \mathcal{G}_0^{2,q}(P)^C \right)^* \left\{ \alpha \in \mathcal{A}^{1,q}(P) \mid \| F_\alpha \|_{L^q} < \epsilon \right\}
\]

of the domain of $\text{NS}_P$. Furthermore, the group $\mathcal{G}_0^{2,q}(P)^C$ restricts to a free action on $\mathcal{A}^s(P)$.

Consider the projection

\[
\Pi^C : \mathcal{A}^s(P) \longrightarrow \mathcal{A}^s(P)/\mathcal{G}_0^{2,q}(P)^C.
\]

Using $\text{NS}_P$, we have an identification $\mathcal{A}^s(P)/\mathcal{G}_0^{2,q}(P)^C \cong M(P)$, and hence a commutative diagram

\[
\begin{array}{ccc}
\mathcal{A}^s(P) & \xrightarrow{\text{NS}_P} & \mathcal{A}^{1,q}_{\flat}(P) \\
\Pi^C \downarrow & & \downarrow \Pi \\
\mathcal{A}^s(P)/\mathcal{G}_0^{2,q}(P)^C & \cong & \mathcal{A}^{1,q}_{\flat}(P)/\mathcal{G}_0^{2,q}(P) = M(P)
\end{array}
\]

As we saw above, the infinitesimal action of $\mathcal{G}_0^{2,q}(P)$ is complex linear. This implies that the linearization of $\Pi^C : \mathcal{A}^s(P) \to M(P)$ is complex-linear, but $\Pi^C = \Pi \circ \text{NS}_P$, so this finishes the proof.

\[ \square \]

Lemma 3.17. Let $P \to \Sigma$ be as in the statement of Lemma 3.16 and $1 < q < \infty$. Assume $\alpha$ is in the domain of $\text{NS}_P$. Then the space $\text{im}(d_\alpha) \oplus \text{im}(d_\alpha^* \ast)$ lies in the kernel of $D_\alpha^k(\Pi \circ \text{NS}_P)$ in the sense that

\[
D_\alpha^k(\Pi \circ \text{NS}_P) (d_\alpha \xi + \ast d_\alpha \zeta, \cdot, \cdot, \cdot) = 0
\]

for all 0-forms $\xi, \zeta \in W^{2,q}(P(g))$. Moreover, there is an estimate

\[
\left| D_\alpha^k(\Pi \circ \text{NS}_P)(\eta_1, \ldots, \eta_k) \right|_{M(P)} \leq C \| \eta_1 \|_{L^q(\Sigma)} \cdots \| \eta_k \|_{L^q(\Sigma)}
\]

for all tuples $\eta_1, \ldots, \eta_k \in W^{1,q}(T^*\Sigma \otimes P(g))$ of 1-forms. Here $| \cdot |_{M(P)}$ is any norm on $TM(P)$. 

(45)
Proof. We prove the lemma for $k = 1$. The cases for larger $k$ are similar. By the proof of Lemma 3.16 the map $\Pi \circ \text{NS}_P$ is invariant under gauge transformations of the form $\exp(\xi + i\zeta)$ where $\xi, \zeta \in W^{2, q}(P(g))$ are real. In particular,  

$$ 0 = \frac{d}{dr} \bigg|_{r=0} \Pi \circ \text{NS}_P(\exp(\tau(\xi + i\zeta))^\star) = -D_\alpha(\Pi \circ \text{NS}_P)(d_\alpha \xi + *d_\alpha \zeta). $$

This proves the first assertion.

To prove the estimate (45), we note that by Lemma 3.10 there is a decomposition

$$ T_\alpha A^{1, q}(P) = H^1_\alpha \oplus (\text{im } d_\alpha \oplus \text{im } d_\alpha^*), $$

whenever $\alpha$ has sufficiently small curvature. Moreover, the first summand is $L^2$-orthogonal. Denote by $\text{proj}_\alpha : T_\alpha A^{1, q}(P) \to H^1_\alpha$ the projection to the $d_\alpha$-harmonic space, and note that this is continuous with respect to the $L^2$-norm on the domain and codomain. We claim that the operator

$$ D_\alpha (\Pi \circ \text{NS}_P) : T_\alpha A^{1, q} \longrightarrow H_{\text{NS}_P(\alpha)} $$

can be written as a composition

$$ T_\alpha A^{1, q} \longrightarrow H_\alpha \overset{M_\alpha}{\longrightarrow} H_{\text{NS}_P(\alpha)} $$

for some bounded linear map $M_\alpha$, where the first map is $\text{proj}_\alpha$. Indeed, by the first part of the lemma it follows that

$$ D_\alpha (\Pi \circ \text{NS}_P) (\mu) = D_\alpha (\Pi \circ \text{NS}_P) (\text{proj}_\alpha \mu) $$

since the difference $\mu - \text{proj}_\alpha \mu$ lies in $\text{im } d_\alpha \oplus \text{im } *d_\alpha$. So the claim follows by taking

$$ M_\alpha := D_\alpha (\Pi \circ \text{NS}_P) |_{H_\alpha} $$

to be the restriction. Since $M_\alpha$ is a linear map between finite-dimensional spaces, it is bounded with respect to any norm. We take the $L^2$-norm on these harmonic spaces. Then $D_\alpha (\Pi \circ \text{NS}_P)$ is the composition of two functions that are continuous with respect to the $L^2$ norm:

$$ |D_\alpha (\Pi \circ \text{NS}_P) \mu|_{M(P)} = C \| D_\alpha (\Pi \circ \text{NS}_P) \mu \|_{L^2} = C \| M_\alpha \circ \text{proj}_\alpha \mu \|_{L^2} \leq C \| \mu \|_{L^2}. $$

That this constant can be taken independent of $\alpha$, for $\alpha$ sufficiently small, follows using an Uhlenbeck compactness argument similar to the one carried out at the beginning of the proof of Theorem 3.16. Here one needs to use the fact that $D_\alpha (\Pi \circ \text{NS}_P) = \text{proj}_\alpha$ when $\alpha$ is a flat connection, and so this has norm 1 (which is clearly independent of $\alpha$). \qed

Corollary 3.18. Suppose $1 < q < \infty$, and let $P \to \Sigma$ be as in the statement of Lemma 3.16. 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 $\| F_\alpha \|_{L^2(\Sigma)} < \epsilon_0$, the following estimate holds

$$ \| \text{proj}_\alpha \eta - D_\alpha (\Pi \circ \text{NS}_P) \eta \|_{L^2(\Sigma)} \leq f(\alpha) \| \text{proj}_\alpha \eta \|_{L^2(\Sigma)} $$

for all $\eta \in L^q(T^* \Sigma \otimes P(g))$, where $\text{proj}_\alpha$ is the map (32). Furthermore, $f$ can chosen so that $f(\alpha) \to 0$ as $\| F_\alpha \|_{L^2(\Sigma)} \to 0$.

Proof. Consider the operator $\text{proj}_\alpha - D_\alpha (\Pi \circ \text{NS}_P)$. It is clear from Lemma 3.17 that its kernel contains $\text{im}(d_\alpha) \oplus \text{im}(\star d_\alpha)$, and so we have
Theorem 3.6 (i.e., since \( C \) is finite-dimensional space where we have set \( \nu \) sufficiently large. Let NS be the map constructed in Theorem 3.6 for the bundle \( \nu \) holomorphic, \( Dv \). It suffices to prove the claim in local orthonormal coordinates \( x = (s, t) \) on \( K \). Since \( v_\nu \) is holomorphic, \( Dv_\nu(x) \) is controlled by \( \partial_s v_\nu(x) \). Suppressing the point \( x \) we have

\[
|\partial_s v_\nu|_{M(P)} = \|\partial_s (\Pi \circ \text{NS}(\alpha_\nu))\|_{L^2(\Sigma)} = \|D\alpha_\nu(\Pi \circ \text{NS})(\partial_s \alpha_\nu)\|_{L^2(\Sigma)} \leq \|D\alpha_\nu(\Pi \circ \text{NS})(\text{proj}_\alpha \partial_s \alpha_\nu)\|_{L^2(\Sigma)}.
\]

where the last equality holds by Lemma 3.17 since \( \partial_s \alpha_\nu \) and \( \text{proj}_\alpha \partial_s \alpha_\nu \) differ by an element of \( \text{im} d_{\alpha_\nu} \oplus \text{im} * d_{\alpha_\nu} \). By Lemma 3.17 we have

\[
\|D\alpha_\nu(\Pi \circ \text{NS})(\text{proj}_\alpha \partial_s \alpha_\nu)\|_{L^2(\Sigma)} \leq C_0 \|\text{proj}_\alpha \partial_s \alpha_\nu\|_{L^2(\Sigma)}.
\]

By assumption (ii) in the statement of the Lemma 3.3, this last term is controlled by a constant \( C \), and this proves the claim.

It follows from Claim 1 that \( \{v_\nu\} \) is a \( C^1 \)-bounded sequence of maps \( K \to M(P) \) for each compact \( K \subset S_0 \). By the compactness of the embedding \( C^1(K) \hookrightarrow C^0(K) \), there is a subsequence, still
denoted by \( \{v_\nu\} \), which converges weakly in \( C^1 \), and strongly in \( C^0 \), to some limiting holomorphic map \( v_\infty : K \to M(P) \). By repeating the above with a sequence \( K_n \) of compact sets that exhaust \( S_0 \), and by taking a diagonal subsequence, one can show that \( v_\infty \) is defined on all of \( S_0 \) and the \( v_\nu \) converge to \( v_\infty \) in \( C^0 \) on compact subsets of \( S_0 \).

**Remark 3.19.** 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 \( v_\nu \) converge to \( v_\infty \) in \( C^\infty \) on compact subsets of the interior of \( S_0 \) (see [18, Theorem 4.1.1]).

**Claim 2:** There exists a smooth lift \( \alpha_\infty : S_0 \to \mathcal{A}_{\text{flat}}(P) \) of \( v_\infty : S_0 \to M(P) \).

This is only non-trivial when \( S_0 \) is closed, and in this case the result basically follows because (i) each \( v_\nu^* \mathcal{A}_{\text{flat}}(P) \to S_0 \) is trivializable (\( v_\nu \) admits a lift \( \alpha_\nu \)), and (ii) the \( v_\nu \) converge to \( v_\infty \) in \( C^0 \) so \( v_\infty^* \mathcal{A}_{\text{flat}}(P) \) is trivializable. To see this explicitly, we will show that, for large enough \( \nu \), there is a map \( \Gamma : I \times S_0 \to M(P) \) that restricts to \( v_\infty \) on \( \{0\} \times S_0 \) and to \( v_\nu \) on \( \{1\} \times S_0 \). Then the result will follow since the pullback over an elementary cobordism is trivial if and only if it is trivial over one of the boundary components. To construct \( \Gamma \), pick \( \nu \) large enough so \( \text{dist}(v_\nu(x), v_\infty(x)) \) is smaller than the injectivity radius of \( M(P) \). Let \( \gamma_\nu(x) \) be the geodesic from \( v_\infty(x) \) to \( v_\nu(x) \) parametrized so that it has length 1. Then defining \( \Gamma(t, x) := \gamma_t(x) \) proves the claim.

Now we use Claim 2 to translate the convergence of Claim 1 into a statement about \( \alpha_\nu \) and \( \alpha_\infty \). Note that, because \( M(P) \) is finite-dimensional, we can choose any metric we want. At this point it is convenient to choose the metric on the tangent space induced from the \( C^0 \)-norm on the harmonic spaces. In particular, the \( C^0 \)-convergence in the \( S \)-directions immediately implies that, for each \( x \in S_0 \), there are gauge transformations \( \mu_\nu(x) \in G^2_S(P) \) such that

\[
\sup_{x \in K} \|\alpha_\infty(x) - \mu_\nu(x)^* \text{NS}(\alpha_\nu(x))\|_{C^0(\Sigma)} \to 0.
\]

By perturbing the gauge transformations, we may suppose that each \( \mu_\nu(x) \) is smooth in \( x \). This gives

\[
\left\| \alpha_\infty - \mu_\nu^* \alpha_\nu \right\|_{C^0(K \times \Sigma)} \leq \sup_{x \in K} \left\{ \left\| \mu_\nu^* \alpha_\nu - \mu_\nu^* \text{NS}(\alpha_\nu) \right\|_{C^0(\Sigma)} + \left\| \alpha_\infty - \mu_\nu^* \text{NS}(\alpha_\nu) \right\|_{C^0(\Sigma)} \right\}
\]

\[
\leq C \sup_{x \in K} \left\{ \left\| F_{\alpha_\nu} \right\|_{C^0(\Sigma)} + \left\| \alpha_\infty - \mu_\nu^* \text{NS}(\alpha_\nu) \right\|_{C^0(\Sigma)} \right\}
\]

where the last inequality follows from Proposition 3.13. This last term goes to zero by assumption (i) and [18]. This proves (21).

Now we prove (22). We begin with \( p = 2 \). We will work in local coordinates \( x = (s, t) \) on \( \text{int} S_0 \). By the holomorphic and \( \epsilon \)-ASD conditions, it suffices to prove that

\[
\sup_{x \in K} \left\| \text{proj}_{\alpha_\infty} \partial_s \alpha_\infty - \text{Ad}(\mu_\nu^{-1}) \text{proj}_{\alpha_\nu} \partial_s \alpha_\nu \right\|_{L^2} \to 0
\]

for each compact \( K \subset \text{int} S_0 \), where here and below the \( L^2 \)-norms are on \( \Sigma \). The key ingredient is that \( \partial_s v_\nu \) converges to \( \partial_s v_\infty \) in \( C^0 \) on \( K \); this is coming from Remark 3.19. We first translate this to a statement about the connections: The appropriate lift of \( \partial_s v_\infty \) to \( T_{\alpha_\infty} \mathcal{A}(P) \) is the harmonic projection \( \text{proj}_{\alpha_\nu} \partial_s \alpha_\nu \). Similarly, the appropriate lift of \( \partial_s v_\nu(x) \) is the harmonic projection of the linearization \( D_{\alpha_\nu} \text{NS} (\partial_s \alpha_\nu(x)) \). This harmonic projection is exactly \( D_{\alpha_\nu}(\Pi \circ \text{NS}) (\partial_s \alpha_\nu) \), since \( D_\alpha \Pi = \text{proj}_\alpha \) whenever \( \alpha \) is flat. Then the \( C^0 \) convergence \( \partial_s v_\nu \to \partial_s v_\infty \) implies that

\[
\sup_{K} \left\| \text{proj}_{\alpha_\infty} \partial_s \alpha_\infty - \text{Ad}(\mu_\nu^{-1}) D_{\alpha_\nu}(\Pi \circ \text{NS}) (\partial_s \alpha_\nu) \right\|_{L^2} \to 0.
\]
The gauge transformations appear here are exactly the ones from the previous paragraph; this is due to the fact that since the $\mu^v_\nu \alpha_v$ converge to $\alpha_\infty$, the harmonic spaces $\text{Ad}(\mu^{-1}_\nu(x))H_{\alpha_v(x)}$ converge to $H_{\alpha_\infty(x)}$. For each $x \in K$, the triangle inequality gives

$$
\|\text{proj}_{\alpha_\infty} \partial_x \alpha_\infty - \text{Ad}(\mu^{-1}_\nu)\text{proj}_{\alpha_\nu} \partial_x \alpha_v\|_{L^2} \leq \|\text{proj}_{\alpha_\infty} \partial_x \alpha_\infty - \text{Ad}(\mu^{-1}_\nu)D_{\alpha_v}(\Pi \circ \text{NS})(\partial_x \alpha_v)\|_{L^2} + \|D_{\alpha_v}(\Pi \circ \text{NS})(\partial_x \alpha_v)\|_{L^2} 
\leq \|\text{proj}_{\alpha_\infty} \partial_x \alpha_\infty - \text{Ad}(\mu^{-1}_\nu)D_{\alpha_v}(\Pi \circ \text{NS})(\partial_x \alpha_v)\|_{L^2} + f(\alpha_v)\|\text{proj}_{\alpha_\nu} \partial_x \alpha_v\|_{L^2}
$$

where the last inequality is Corollary 3.18. It follows from (49) that the first term on the right goes to zero. For the second term, note that assumption (ii) implies that $\|\text{proj}_{\alpha_\nu} \partial_x \alpha_v\|_{L^2}$ is bounded by some constant $C$, uniformly in $x \in K$. Then Corollary 3.18 combines with assumption (i) in the statement of Lemma 3.5 to give that $f(\alpha_v) \to 0$ uniformly in $x \in K$. This finishes the proof of (22) with $p = 2$.

Now we prove (22) for arbitrary $1 < p < \infty$. To simplify notation, replace $\alpha_v$ by $\mu^v_\nu \alpha_v$. Note that by Proposition 3.12 for any $1$-forms $\eta, \eta'$ on $\Sigma$ we have

$$
\|\text{proj}_{\alpha_\infty} \eta - \text{proj}_{\alpha_\nu} \eta'\|_{L^p} \leq C\|\alpha_\infty - \alpha_v\|_{L^\infty} \|\eta\|_{L^p} + \|\text{proj}_{\alpha_\nu} \eta - \text{proj}_{\alpha_\nu} \eta'\|_{L^p},
$$

where all norms are on $\Sigma$ and we are working pointwise on $S_0$. The harmonic space $H^1_{\alpha_v}$ is finite-dimensional and converging to $H^1_{\alpha_\infty}$, so the $L^p$-norm on $H^1_{\alpha_v}$ is equivalent to the $L^2$-norm by a constant $C'$ that is independent of $\nu$. We therefore have

$$
\|\text{proj}_{\alpha_\infty} \eta - \text{proj}_{\alpha_\nu} \eta'\|_{L^p} \leq C\|\alpha_\infty - \alpha_v\|_{L^\infty} \|\eta\|_{L^p} + C'\|\text{proj}_{\alpha_\nu} \eta - \text{proj}_{\alpha_\nu} \eta'\|_{L^p} \leq C\|\alpha_\infty - \alpha_v\|_{L^\infty} \|\eta\|_{L^p} + C'\|\text{proj}_{\alpha_\infty} \eta - \text{proj}_{\alpha_\nu} \eta'\|_{L^2} + C'\|\text{proj}_{\alpha_\infty} \eta - \text{proj}_{\alpha_\nu} \eta'\|_{L^2}
$$

where we used Proposition 3.12 again. Apply this with $\eta = \text{proj}_{\alpha_\infty} \partial_x \alpha_\infty$ (which is obviously uniformly bounded in $L^p(\Sigma)$), and $\eta' = \text{proj}_{\alpha_\nu} \partial_x \alpha_v$ (which is bounded uniformly on $K$ in $L^2(\Sigma)$) by assumption (ii) in the statement of Lemma 3.5. Since $\|\alpha_\infty - \alpha_v\|_{L^\infty}$ and $\|\text{proj}_{\alpha_\infty} \eta - \text{proj}_{\alpha_\nu} \eta'\|_{L^2}$ converge to zero uniformly on $K$, it follows that $\|\text{proj}_{\alpha_\infty} \eta - \text{proj}_{\alpha_\nu} \eta'\|_{L^p}$ does as well.

3.3. Proof of Theorem 3.3. In light of Lemma 3.5 we consider each of the following cases:

- **Case 1:** $\|F_{\alpha_v}\|_{L^\infty(Z)} \to \infty$;
- **Case 2:** $\|F_{\alpha_v}\|_{L^\infty(Z)} \to \Delta > 0$;
- **Case 3:** $\|F_{\alpha_v}\|_{L^\infty(Z)} \to 0$, and $\sup_{x \in S} \|\text{proj}_{\alpha_v} \circ D_{\alpha_v}\|_{L^2(\Sigma)} \to \infty$.

To prove the theorem, we will show that each case leads to energy quantization. That is, we will show that there is a positive constant $\kappa > 0$, depending only on the group $PU(r)$, with the following significance: We will show that each case above implies there is some bubbling point $x \in S$ and a set $T_x \subset Z$ ($T_x$ will be either a point in $\{x\} \times \Sigma$, or the whole fiber $\{x\} \times \Sigma$), such that for every neighborhood $U$ of $T_x$ the energy $E^\text{inst}_{v}\alpha_v \geq \kappa$ is uniformly bounded from below for all $\nu$. To see why this implies the theorem, let $B$ denote the set of exceptional points $x \in S$. Since each $A_\nu$ is an $\epsilon_\nu$-instanton on the bundle $R$, it follows that $\epsilon_\nu$-energies are equal to $-2\pi^2 N \kappa_{1}^{-1} \eta_1(R) \geq 0$. This implies that $B$ must be a finite set. Then hypotheses of Lemma 3.5 hold on $S_0 := S \setminus B$, so the convergence result in Theorem 3.3 follows from the lemma. That $A_\infty$ extends over the bubbling set $B$ follows from the removable singularities theorem for holomorphic curves: The set $B$ projects to a finite set in $S$, now apply the removable singularities theorem to $v_\infty$ in the proof of Lemma 3.5 and lift using Claim 2 appearing in that proof. The verification of the bound (20) is standard.

It remains to prove that the three cases above lead to energy quantization. For the first two cases we follow (23), so we only sketch the details; see also Section 4.1.3.

**Case 1.** (Instantons on $S^4$) For each $\nu$ identify a point $z_\nu \in Z$ where $|F_{\alpha_v}(z_\nu)|$ is maximized. To simplify notation, we may assume $z_\nu = z_\infty$ is fixed for all $\nu$ (this is approximately true by compactness of $Z$). Restrict attention to a small fixed neighborhood of $z_\infty$. Rescale each $A_\nu$ by
some \( \epsilon_\nu \), but only in the \( S \)-directions (the \( \Sigma \)-direction remain unscaled). Then we obtain a sequence of connections \( \tilde{A}_\nu \) on increasing and exhausting subsets of \( C \times \Sigma \) that are instantons with respect to a fixed metric, and have curvature \( |F_{\tilde{A}_\nu}| \) blowing up at \( z_\infty \in C \times \Sigma \). Now we rescale again, this time in all four directions and by the maximum of \( |F_{\tilde{A}_\nu}| \). By Uhlenbeck’s strong compactness theorem to the twice-rescaled sequence of connections, these converge to a non-trivial instanton on a bundle \( R_\infty \) over \( S^4 \). The energy of such instantons is \( \hbar = -2\pi^2 \kappa_r r^{-1} q_4(R_\infty) \). Note that since \( t_2(R_\infty) \in H^2(S^4, \mathbb{Z}_r) \) obviously vanishes, \( q_4(R_\infty) \) is divisible by \( 2r \) (see, e.g., [9] Equation (4)), and so \( \hbar = 4\pi^2 \kappa_r r^{-1} \).

**Case 2. (Instantons on \( C \times P \))** Exactly as in the previous case, we rescale by \( \epsilon_\nu \) in the \( S \)-directions to obtain a sequence \( \tilde{A}_\nu \) of instantons on exhausting subsets of \( C \times \Sigma \) with curvature maximized at some \( z_\infty \in C \times \Sigma \). If these maxima \( |F_{\tilde{A}_\nu}| \) diverge, then we repeat the analysis of Case 1 and get an instanton on \( S^4 \). Otherwise, the curvatures are \( L^\infty \) bounded and we can apply Uhlenbeck’s strong compactness theorem directly to the \( \tilde{A}_\nu \) to obtain a non-flat finite-energy instanton on \( C \times P \rightarrow C \times \Sigma \). We therefore need to show that there is a minimum allowable energy \( \frac{1}{2} \int_{C \times \Sigma} |F_A|^2 \geq \hbar > 0 \) for all non-flat instantons \( A \) on bundles over the domain \( C \times \Sigma \). The basic idea is to introduce polar coordinates on the \( C \)-component in \( C \times \Sigma \), the cylinder \( \mathbb{R} \times S^4 \times \Sigma \). The finite energy instanton \( A \) limits to flat connections on the cylindrical ends \( S^4 \times \Sigma \), and the energy of \( A \) is given by the difference of the Chern-Simons functional applied to each of these limiting flat connections. Then it is shown in [9] that this difference is given by \( 4\pi^2 \kappa_r r^{-1} \text{deg}(u) \), where \( u \) is a suitable gauge transformation on \( S^4 \times P \) and necessarily has non-zero degree (otherwise \( A \) would be flat). We can therefore take \( \hbar = 4\pi^2 \kappa_r r^{-1} \). See also [23] for a similar discussion along these lines.

**Case 3. (Holomorphic spheres in \( M(P) \))** In this case, we rescale around a blow-up point to find that a holomorphic sphere bubbles off in the moduli space of flat connections. 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 theorem, but this theorem is too weak to conclude that the limiting bubbles are non-constant. Dostoglou and Salamon resolved this issue [7] by developing several intricate estimates for these types of \( c \)-instantons. We present an alternative argument using the heat flow (the key ingredient is [22]). Our argument has the additional bonus that it applies even in cases where there are boundary conditions, as we will encounter in Theorem 4.1.

Set \( \epsilon_\nu := \|\text{proj}_{\alpha_\nu(x_\nu)} \circ D_{x_\nu} \alpha_\nu\|_{L^2(\Sigma)} \) where \( x_\nu \in S \) is chosen to maximize the right-hand side. This diverges to \( \infty \), by assumption. By compactness we may assume \( x_\nu \) converges to some \( x_\infty \in S \), and we fix a small ball \( B_{2\delta}(x_\infty) \subset S \) with holomorphic coordinates \( x = (s, t) \). We may also assume \( \nu \) is large enough so that \( x_\nu \) is within \( \delta \) of \( x_\infty \). Write \( A_\nu |_{U \times \Sigma} = \alpha_\nu + \phi_\nu \, ds + \psi_\nu \, dt \) and define a rescaled connection \( \tilde{A}_\nu \) in terms of its components by

\[
\tilde{\alpha}_\nu(x) := \alpha(c_\nu^{-1}x + x_\nu), \quad \tilde{\phi}_\nu(x) := c_\nu^{-1}\phi(c_\nu^{-1}x + x_\nu), \quad \tilde{\psi}_\nu(x) := c_\nu^{-1}\psi(c_\nu^{-1}x + x_\nu).
\]

View these as being defined on subsets \( C \times \Sigma \). By Hofer’s Lemma [18] Lemma 4.6.4], we can refine the choice of \( x_\nu \) to ensure that these subsets are increasing and exhaust \( C \times \Sigma \). Write \( F_{\tilde{A}_\nu} = \tilde{F}_{\tilde{A}_\nu} - \tilde{\beta}_{s,\nu} \wedge ds - \tilde{\beta}_{t,\nu} \wedge dt + \tilde{\gamma}_{\nu} ds \wedge dt \) in terms of its components. These satisfy \( \tilde{\beta}_{s,\nu} + * \tilde{\beta}_{t,\nu} = 0 \) and \( \tilde{\gamma} = -\tilde{c}_\nu^2 + \tilde{F}_{\tilde{A}_\nu} \), where \( \tilde{c}_\nu := c_\nu \epsilon_\nu \). It may not be the case that \( \epsilon_\nu \) is decaying to zero; this is replaced by the assumption that the slice-wise curvatures are going to zero in \( L^\infty(C \times \Sigma) \):

\[
\|F_{\tilde{A}_\nu}\|_{L^\infty} = \|F_{\alpha_\nu}\|_{L^\infty} \rightarrow 0.
\]

We also have

\[
\|\text{proj}_{\tilde{\beta}_{s,0}} \partial_s \tilde{\alpha}_\nu(0)\|_{L^2(\Sigma)} = \frac{1}{2c_\nu} \|\text{proj}_{\alpha_\nu(x_\nu)} \circ D_{x_\nu} \alpha_\nu\|_{L^2(\Sigma)} = \frac{1}{2},
\]
where we have used the $\epsilon_\nu$-ASD equation in the first equality. Then the $\hat{A}_\nu$ satisfy the hypotheses of Lemma \ref{Hodgkin} so (after possibly passing to a subsequence) there exists a sequence of gauge transformations $U_\nu \in G^2_{\text{loc}}(\mathbb{C} \times P)$, and a limiting holomorphic curve representative $A_\infty \in \mathcal{A}^{1,q}_{\text{loc}}(\mathbb{C} \times P)$ such that

\[
\sup_K ||\text{proj}_{\Sigma_j} \partial_{\alpha} \partial_{\alpha} - \text{Ad}(\mu^{-1}_\nu)\text{proj}_{\Sigma_j} \partial_{\alpha} \partial_{\alpha}||_{L^2(\Sigma)} \to 0,
\]

for all compact $K \subset \mathbb{C}$. Then this descends to a finite-energy holomorphic curve $v_\infty : \mathbb{C} \to M(P)$, which is non-trivial by \cite{2}. The removal of singularities theorem \cite{18} Theorem 4.1.2 (ii) implies that $v_\infty$ extends to a holomorphic sphere $v_\infty : S^2 \to M(P)$. We have energy quantization with $\hbar = 4\pi^2 \kappa_r^{-1}$ for non-constant holomorphic spheres \cite{4} Corollary 6.3, which takes care of Case 3.

4. Compactness for cylinders $\mathbb{R} \times Y$

Fix a broken circle fibration $f : Y \to S^1$ as in \cite{8}. This means that $Y$ is a closed, connected, oriented 3-manifold, and $f$ is a Morse function with connected fibers. We will assume that the number $N$ of critical points of $f$ is positive (otherwise we are essentially in the case of Section \ref{case3}). We also assume the critical points of $f$ have distinct critical values. This implies that $N$ is even, and also allows us to view $Y$ as a composition of cobordisms:

\[
y_{01} \cup_{\Sigma_j} (I \times \Sigma_1) \cup_{\Sigma_2} Y_{12} \cup_{\Sigma_3} \cdots \cup_{\Sigma_{N-1}} Y_{(N-1)0} \cup_{\Sigma_0} (I \times \Sigma_0) \cup_{\Sigma_0} \]

where $I := [0, 1]$ is the unit interval, each $\Sigma_j \subset Y$ is a fixed regular fiber of $f$, and each $Y_{j(j+1)} \subset Y$ is a cobordism from $\Sigma_j$ to $\Sigma_{j+1}$ such that $f|_{Y_{j(j+1)}}$ has exactly one critical point. Note that \cite{22} is cyclic in the sense that the cobordism $I \times \Sigma_0$ on the right is glued to the cobordism $Y_{01}$ on the left, reflecting the fact that $f$ maps to the circle. We set $Y_* := \bigcup_j Y_{j(j+1)}$ and $\Sigma_* := \bigcup_j \Sigma_j$.

In this section we consider the 4-manifold $Z := \mathbb{R} \times Y$. We equip $Z$ with a product metric $g = ds^2 + g_Y$, where $g_Y$ is a metric on $Y$. To simplify the exposition, we assume $g_Y$ has been chosen so that $g_Y|_{I \times \Sigma_*} = dt^2 + g_{\Sigma_*}$, where $g_{\Sigma_*}$ is a metric on $\Sigma_*$ that is independent of $t \in I$. Then we define a metric $g_\epsilon$ by setting $g_\epsilon|_{\mathbb{R} \times Y_*} := ds^2 + \epsilon^2 g_Y|_{\mathbb{R} \times Y_*}$, $g_\epsilon|_{\mathbb{R} \times I \times \Sigma_*} := ds^2 + dt^2 + \epsilon^2 g_{\Sigma_*}$.

When $\epsilon \neq 1$, $g_\epsilon$ is not smooth on $Z$ with respect to the standard smooth structure. However, there is a different smooth structure on $Z$ in which $g_\epsilon$ is smooth. We call this the $\epsilon$-dependent smooth structure, and say a function, tensor, connection, etc. is $\epsilon$-smooth if it is smooth with respect to the $\epsilon$-dependent smooth structure. See \cite{8} Section 2.1 for more details.

We take $R \to Z$ to be the pullback of a PU($r$)-bundle $Q \to Y$ such that $t_2(Q)$ is Poincaré dual to $d[\gamma] \in H_1(Y, \mathbb{Z}_r)$, where $d \in \mathbb{Z}_r$ is a generator, and $\gamma : S^1 \to Y$ is a section of $f : Y \to S^1$. Set

\[
Q_{j(j+1)} := Q|_{Y_{j(j+1)}}, \quad Q_* := \bigcup_j Q_{j(j+1)}, \quad P_j := Q|[0] \times \Sigma_j, \quad P_* := \bigcup_j P_j.
\]

The assumption on $Q$ ensures $t_2(Q_{j(j+1)})[Q_{j(j+1)}] = t_2(P_j)[P_j] = d$. In \cite{9} we show that the fibers $\Sigma_* \subset Y$ determine a connected component in $\mathcal{G}(Q)$ consisting of degree 1 gauge transformations. We let $G_\Sigma \subset \mathcal{G}(Q)$ be the subgroup generated by this component.

We say that a connection on $R$ is $\epsilon$-ASD or an $\epsilon$-instanton if it is ASD with respect to $g_\epsilon$. In coordinates over $\mathbb{R} \times I \times \Sigma_* \subset Z$, this takes the form \cite{18}. Over $\mathbb{R} \times Y_*$, any connection can be written as $a(s) + p(s) ds$, where $a(s)$ is a path of connections on $Q$, and $p(s)$ is a path of $Q(g)$-valued 0-forms on $Y$. Then the $\epsilon$-ASD condition on $\mathbb{R} \times Y_*$ takes the form

\[
\partial_s a(s) - da(s)p(s) + \epsilon^{-1} *_Y F_{a(s)} = 0,
\]

where $*_Y$ is the Hodge star coming from $g_Y$. As before, the $\epsilon$-energy is $E_{\epsilon}^{\text{inst}}(A) := \frac{1}{2} \int_Z (F_A \wedge *_Y F_A)$ (this makes sense for any connection on $R$ that is $W^{1,2}$ with respect to the $\epsilon$-smooth structure). If $A$ is $\epsilon$-ASD, then the $\epsilon$-energy is finite and if only if $A$ limits to flat connections at $\pm \infty$ in the sense that there are flat connections $a^\pm \in \mathcal{A}_{\text{flat}}(Q)$ such that $\lim_{s \to \pm \infty} a(s) = a^\pm$, and $\lim_{s \to \pm \infty} p(s) = 0$.\[28\]
When this is the case, \( E^\text{inst}_\varepsilon(A) = CS(a^-) - CS(a^+) \), where \( CS \) is the Chern-Simons functional. In particular, this is independent of \( \varepsilon \) and \( A \).

Now we introduce the relevant holomorphic curve equations; see [27] and [8] for more details. Restricting to each of the two boundary components of \( Y_{(j+1)} \) determines a Lagrangian embedding \( L(Q_{j(j+1)}) \to M(P_j)^{-} \times M(P_{j+1}) \). In the language of [29] it follows that \( L(Q) := (L(Q_{j(j+1)})) \) determines a cyclic Lagrangian correspondence in the product symplectic manifold \( M \). The Hodge star from \( \alpha \) essentially follows because the symplectic action functional for \( A \) is finite if and only if \( \varepsilon \) is finite. This is concocted to that it recovers the energy of the quilted cylinder that satisfies (19).

When this is the case, \( E^\text{inst}_\varepsilon(A) = CS(a^-) - CS(a^+) \), where \( CS \) is the Chern-Simons functional. In particular, this is independent of \( \varepsilon \) and \( A \).

We are interested in \textit{quilted holomorphic cylinders} in \( M \); the seams are given by \( \mathbb{R} \times f(\text{crit}(f)) \subset \mathbb{R} \times S^1 \). For a full discussion of quilts, we refer to [29]. For our purposes, it suffices to discuss \textit{quilted holomorphic cylinder representatives}, which we define to be a connection \( A_0 \in \mathcal{A}(Z) \) satisfying the following: Over \( \mathbb{R} \times I \times \Sigma_s \), \( A_0 \) has the form \( \alpha + \phi ds + \psi dt \), where \( \alpha : \mathbb{R} \times I \to \mathcal{A}_\text{flat}(\Sigma_s, P_s(g)) \), and \( \phi, \psi : \mathbb{R} \times I \to \Omega^1(\Sigma_s, P_s(g)) \) are 0-forms defined so that \( \partial_s \alpha - d_{\alpha} \phi \) and \( \partial_t \alpha - d_{\alpha} \psi \) are harmonic; see Example 3.1. Over \( \mathbb{R} \times Y_s \), \( A_0 \) has the form \( a + p ds \), where \( a : \mathbb{R} \to \mathcal{A}_\text{flat}(Q_s) \) is determined uniquely (up to gauge) by the condition that \( a(s)|_{\partial I} = a(s, \cdot)|_{\partial I \times \Sigma_s} \), and \( p : \mathbb{R} \to \Omega^1(Y_s, Q_s(g)) \) is determined by the condition that \( \partial_s a(s) - d_{a(s)} p(s) = a(s) \)-harmonic.

It follows that any quilted cylinder representative \( A_0 \) is a smooth cylinder on \( Z \) that is \( W^{1,p} \) with respect to the \( \varepsilon \)-dependent smooth structure for any \( \varepsilon > 0 \). We will often write \( A_0 = a + p ds \) on all of \( Z \), and so \( a|_{\mathbb{R} \times I \times \Sigma_s} = \alpha + \psi dt \). We define the \textbf{energy} by the formula

\[
E^{\text{symp}}(A_0) := \frac{1}{2} \int_{\mathbb{R} \times I \times \Sigma_s} |\partial_s \alpha - d_{\alpha} \phi|^2 + |\partial_t \alpha - d_{\alpha} \psi|^2 \text{dvol}.
\]

(This is concocted to that it recovers the energy of the quilted cylinder that \( A_0 \) represents.) We will say that \( A_0 \) is a \textit{holomorphic quilted cylinder representative} if, over \( \mathbb{R} \times I \times \Sigma_s \), its components satisfy \( (19) \). When \( A_0 \) is a holomorphic quilted cylinder representative, then the energy \( E^{\text{symp}}(A_0) \) is finite if and only if \( A_0 \) limits to flat connections \( a^\pm \) at \( \pm \infty \); see [8]. In this case we have

\[
E^{\text{symp}}(A_0) = CS(a^-) - CS(a^+)
\]

and so the energy is again a topological quantity only depending on the limiting connections (this essentially follows because the symplectic action functional for \( M(P) \) is given by the Chern-Simons functional of representatives; see also Lemma 3.2).

For \( s_0 \in \mathbb{R} \), and \( A \in \mathcal{A}(R) \), let \( \tau^s_{s_0} A \in \mathcal{A}(R) \) be the connection defined by translating \( \tau^s_{s_0} A|_{\{s\} \times Y} := A_{|s+s_0} \mid_{\{s\} \times Y} \). Given \( x \in \mathbb{R} \times I \), we will use \( \iota_x : \Sigma_s \to \{x\} \times \Sigma_s \to Z \) to denote the inclusion, and so the pullback \( \alpha(x) := \iota_x^* A \) can be viewed as an \( \mathbb{R} \times I \)-dependent connection on \( \Sigma_s \). Similarly, if \( s \in \mathbb{R} \), then \( \iota_s : \{s\} \times Y \to Z \) is the inclusion, and the restriction \( a(s) := \iota_s^* A \) can be viewed as an \( \mathbb{R} \)-dependent connection on \( Q \to Y \). Now we can state the main theorem.

**Theorem 4.1.** Fix \( 2 < q < \infty \), and let \( R \to Z \) be as above. Assume all flat connections on \( Q \) are non-degenerate, and for flat connections \( a^\pm \in \mathcal{A}_\text{flat}(Q) \). Suppose \( (\epsilon_n)_{n \in \mathbb{N}} \) is a sequence of positive numbers converging to 0, and that, for each \( \nu \), there is an \( \epsilon_n \)-ASD connection \( A_{\nu} \in \mathcal{A}^1_{\text{loc}}(R) \) that limits to \( a^\pm \) at \( \pm \infty \). Then there is

(i) a finite set \( B \subset \mathbb{R} \times I \);

(ii) a subsequence of the \( A_{\nu} \) (still denoted \( A_{\nu} \));

(iii) a sequence of gauge transformations \( U_{\nu} \in \mathcal{G}^2_{\text{loc}}(R) \);

(iv) a finite sequence of flat connections \( \{a^0 = a^-; a^1, \ldots, a^{J-1}, a^J = a^+\} \subset \mathcal{A}_\text{flat}(Q) \);

(v) for each \( j \in \{1, \ldots, J\} \), a holomorphic quilted cylinder representative \( A^j \in \mathcal{A}^1_{\text{loc}}(R) \) limiting to \( u_{j-1}^+ a^{j-1} \) at \( -\infty \) and \( u^+_{j} a^j \) at \( +\infty \), for some \( u_{j-1}, u_j \in G_{\Sigma} \), possibly depending on \( A^j \);

(vi) for each \( j \), a sequence \( s_{j}^\nu \in \mathbb{R} \)
such that for each \( j \in \{1, \ldots, J\} \), the restrictions

\[
\sup_{x \in K} \left\| \tau_x^* \left( U_{\nu}^* \tau_{s \nu}^* A_{\nu} - A_j^j \right) \right\|_{C^0(S)} \overset{\nu}{\rightarrow} 0
\]

converge to zero for every compact \( K \subset \mathbb{R} \times I \setminus B \). The gauge transformations \( U_{\nu} \) can be chosen so that they restrict to the identity component \( G_0(P) \) on each \( \{x\} \times \Sigma_t \subset Z \). Moreover, for each \( b \in B \) there is a positive integer \( m_b > 0 \) such that for any \( \nu \),

\[
\sum_{j=1}^J E^{\text{symp}}(A_j^j) \leq E^{\text{inst}}(A_{\nu}) - 2\pi^2 \kappa_{\nu} r^{-1} \sum_{b \in B} m_b.
\]

Finally, \( E^{\text{symp}}(A_j^j) > 0 \) for each \( j \).

To simplify the exposition, we have assumed that all flat connections \( a \in \mathcal{A}_{\text{flat}}(Q) \) are non-degenerate, meaning that \( d_a \) is injective on 1-forms. In general, this need not be the case. However, non-degeneracy can always be achieved by first performing a suitable perturbation to the defining equations. See Section \[.\]

For \( x \in \mathbb{R} \times I \), we will continue to use the notation \( \alpha_{\nu}(x) = \iota_x^* A_{\nu}, \alpha^j(x) = \iota_x^* A_j^j \) and \( \mu(x) = \iota_x^* U \). Then the conclusion of the theorem says that for each \( j \) and compact \( K \subset \mathbb{R} \times I \setminus B \), the sequence \( \sup_{(s,t) \in K} \| (s,t)^* \alpha_{\nu}(s - s_{\nu}^j(t) - \alpha^j(s,t)) \|_{C^0(S)} \) converges to zero; see Remark \[.\](d). As in Section \[ it is convenient to first prove a modified version that a priori excludes bubbling.

**Lemma 4.2.** Let \( S_0 := \mathbb{R} \times I \setminus B \), where \( B \) is any finite set, and let \( 2 < q < \infty \). Suppose \( (\epsilon_{\nu})_{\nu \in \mathbb{N}} \) is a sequence of positive numbers \( \epsilon_{\nu} \) (not necessarily converging to zero), and that for each \( \nu \) there is an \( \epsilon_{\nu}-\text{ASD} \) connection \( A_{\nu} \in A^{1,q}(R) \) with uniformly bounded \( \epsilon_{\nu} \)-energy, and satisfying the following conditions:

(i) For each compact \( K \subset S_0 \), the slice-wise curvatures on \( \Sigma_{\bullet} \) converge to zero:

\[
\sup_{x \in K} \| F_{\alpha(x)} \|_{L^\infty(\Sigma_{\bullet})} \overset{\nu}{\rightarrow} 0.
\]

(ii) For each compact \( L \subset \mathbb{R} \) with \( L \times \{0,1\} \cap B = \emptyset \), the slice-wise curvatures on \( Y_{\bullet} \) converge to zero:

\[
\sup_{s \in L} \| F_{\alpha(s)} \|_{L^\infty(Y_{\bullet})} \overset{\nu}{\rightarrow} 0, \quad \alpha_{\nu}(s) := \iota_s^* A_{\nu}.
\]

(iii) For each compact \( K \subset S_0 \), there is some constant \( C \) with

\[
\sup_{x \in K} \sup_{\nu} \| \text{proj}_{\alpha_{\nu}(x)} \circ D_x \alpha_{\nu} \|_{L^2(\Sigma_{\bullet})} \leq C,
\]

where \( \text{proj}_{\alpha_{\nu}(x)} \) is the harmonic projection, and \( D_x \alpha_{\nu} : T_x S_0 \rightarrow T_{\alpha(x)} A^{1,q}(P_{\bullet}) \) is the push-forward.

Then there is a subsequence of the connections (still denoted \( A_{\nu} \)), a sequence of gauge transformations \( U_{\nu} \) on \( R|_{S_0 \times \Sigma} \), and a holomorphic quilted cylinder representative \( A_{\infty} \in A^{1,q}_{\text{loc}}(R) \) such that

\[
\sup_{x \in K} \left\| \alpha_{\infty}(x) - \mu_{\nu}(x)^* \alpha_{\nu}(x) \right\|_{C^0(\Sigma_{\bullet})} \overset{\nu}{\rightarrow} 0,
\]

\[
\sup_{x \in K} \left\| \text{proj}_{\alpha_{\infty}(x)} \circ D_x \alpha_{\infty} - \text{Ad}(\mu_{\nu}^{-1}(x)) \text{proj}_{\alpha_{\nu}(x)} \circ D_x \alpha_{\nu} \right\|_{L^p(\Sigma_{\bullet})} \overset{\nu}{\rightarrow} 0,
\]

for any compact \( K \subset S_0 \) and any \( 1 < p < \infty \). Here \( \alpha_{\infty}(x) := \iota_x^* A_{\infty} \), for \( x \in \mathbb{R} \times I \).
The key technical point of this lemma is that the convergence in (56) holds even for $K$ that intersect the boundary of $\mathbb{R} \times I$; compare with (22). We also point out that the connections $\tau^*_aA_\nu$ from Theorem 4.3 satisfy the same type of convergence as in (56), for compact $K \subset \mathbb{R} \times I \setminus B$. Finally, the operator $\text{proj}_{a_\nu}$ in (56) can be removed by weakening the convergence; see Remark 4.10(b).

In the next section we review the heat flow on 3-manifolds. This will be used to obtain Lagrangian boundary conditions for the limiting holomorphic curve representatives appearing in these theorems. We then develop several estimates that allow us to obtain the convergence in (56) in the case when $K$ intersects the boundary of $\mathbb{R} \times I$. The proofs of Lemma 4.2 and Theorem 4.1 will appear in Sections 4.3 and 4.4 respectively.

4.1. The heat flow on cobordisms. Suppose $Q$ is principal $G$-bundle over a Riemannian 3-manifold $Y$. In his thesis [20], Råde studied the Yang-Mills heat flow; that is, he studied solutions $\tau \mapsto a(\tau) \in \mathcal{A}(Q)$ to the gradient flow of the Yang-Mills functional

\[(57) \quad \frac{d}{d\tau}a(\tau) = -d^*_{a(\tau)}F_a(\tau), \quad a(0) = a,\]

where $a \in \mathcal{A}(Q)$ is an initial condition. Specifically, Råde proved the following:

**Theorem 4.3.** Suppose $G$ is compact and $Y$ is a closed, oriented manifold of dimension 3. Let $a \in \mathcal{A}^{1,2}(Q)$. Then (57) has a unique solution $\{\tau \mapsto a(\tau)\} \in C^0([0, \infty), \mathcal{A}^{1,2}(Q))$, with the further property that $F_a(\tau) \in C^0_{\text{loc}}([0, \infty), L^2) \cap L^2_{\text{loc}}([0, \infty), W^{1,2})$. Moreover, 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 $\tau$ and using (57) shows that $\mathcal{YM}_Q(a(\tau))$ decreases in $\tau$. Moreover, it follows from Uhlenbeck’s compactness theorem together with [20] Proposition 7.2 that the critical values of the Yang-Mills functional are discrete. Combining these two facts, it follows that there is some $\tilde{\epsilon}_Q > 0$ such that if $\mathcal{YM}_Q(a) < \tilde{\epsilon}_Q$, then the associated limiting connection $\lim_{\tau \to \infty} a(\tau)$ is flat. The flow therefore defines a continuous gauge equivariant deformation retract

\[(58) \quad \text{Heat}_Q : \{a \in \mathcal{A}^{1,2}(Q) \mid \mathcal{YM}_Q(a) < \tilde{\epsilon}_Q\} \to \mathcal{A}_{\text{flat}}^{1,2}(Q)\]

whenever $Y$ is a closed 3-manifold.

**Remark 4.4.** 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 $\text{NS}_P$ and $\text{Heat}_P$ are both maps of the form $\{a \in \mathcal{A}^{1,2}(P) \mid \mathcal{YM}_P(a) < \epsilon_P\} \to \mathcal{A}_{\text{flat}}^{1,2}(P)$, for some $\epsilon_P > 0$. It turns out these are equal, up to a gauge transformation. That is,

\[(59) \quad \Pi \circ \text{NS}_P = \Pi \circ \text{Heat}_P,\]

where $\Pi : \mathcal{A}_{\text{flat}}^{1,2}(P) \to \mathcal{A}_{\text{flat}}^{1,2}(P)/\mathcal{G}_0^{2,2}(P)$ is the quotient map. Though we will not use this fact in this paper, 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 4.3 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 not all $W^{1,2}$-sections are continuous. This 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 $a \in \mathcal{A}(Q)$ that satisfy...
on some neighborhood $U$ of $\partial Q$ ($U$ may depend on $a$). Here $\partial_n \in \Gamma(TQ)$ is a fixed extension of the outward pointing unit normal to $\partial Q$; we may assume that the set $U$ is always contained in the region in which $\partial_n$ is non-zero. Use the normalized gradient flow of $\partial_n$ to write $U = [0, \epsilon) \times \partial Y$. Let $t$ denote the coordinate on $[0, \epsilon)$. Then in these coordinates we can write any connection as $a|_{[0, \epsilon) \times \partial Y} = a(t) + \psi(t) \, dt$. Then (60) is equivalent to requiring $\psi(t) = 0$.

Set $A_{\text{flat}}(Q, \partial Q) := A^{1,2}(Q, \partial Q) \cap A_{\text{flat}}^{1,2}(Q)$. Both of the spaces $A^{1,2}(Q, \partial Q)$ and $A_{\text{flat}}^{1,2}(Q, \partial Q)$ 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 $\partial Q$. (We are purposefully only working with the smooth gauge transformations here.)

**Theorem 4.5.** Let $G$ be a compact, connected Lie group, and $Q \to Y$ be a principal $G$-bundle over a compact, connected, oriented Riemannian 3-manifold $Y$ with non-empty boundary.

1. There is some $\epsilon_Q > 0$ and a continuous strong deformation retract

$$\text{Heat}_Q : \{ a \in A^{1,2}(Q, \partial Q) \mid \mathcal{YM}_Q(a) < \epsilon_Q \} \longrightarrow A_{\text{flat}}^{1,2}(Q, \partial Q).$$

Furthermore, $\text{Heat}_Q$ intertwines the action of $\mathcal{G}(Q, \partial Q)$.

2. Suppose $\Sigma \subset Y$ is an embedded surface that is closed and oriented. Suppose further that either $\Sigma \subset \text{int} \, Y$, or $\Sigma \subset \partial Y$. Then for every $\epsilon > 0$, there is some $\delta > 0$ such that if $a \in A^{1,2}(Q, \partial Q)$ satisfies $||F_a||_{L^2(Y)} < \delta$, then $||((\text{Heat}_Q(a) - a)|_\Sigma||_{L^2(\Sigma)} < \epsilon$, for every $1 \leq q \leq 4$.

**Remark 4.6.** Recently, Charalambous [2] has proven similar results for manifolds with boundary.

**Proof.** Consider the double $Y^{(2)} := \overline{Y} \cup_{\partial Y} Y$, which is a closed 3-manifold. Denote by $i_Y : Y \to Y^{(2)}$ the inclusion of the second factor. We will identify $Y$ with its image under $i_Y$. 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 $i_Y$ smooth and $\sigma$ a diffeomorphism (this is just the smooth structure obtained by choosing the same collar on each side of $\partial Y$). Clearly the map $\sigma$ is orientation-reversing, satisfies $\sigma^2 = \text{Id}$ and has fixed point set equal to $\partial Y$. Similarly, we can form $Q^{(2)} := Q \cup_{\partial Q} Q$ and an involution $\tilde{\sigma} : Q^{(2)} \to Q^{(2)}$. Then $Q^{(2)}$ is naturally a principal $G$-bundle over $Y^{(2)}$ and $\tilde{\sigma}$ is a bundle map covering $\sigma$. Furthermore, $\tilde{\sigma}$ commutes with the $G$-action on $Q^{(2)}$.

Though $\tilde{\sigma}$ is not a gauge transformation (it does not cover the identity), it behaves as one in many ways. For example, since $\tilde{\sigma}$ is a bundle map, the space of connections $\mathcal{A}(Q^{(2)})$ is invariant under pullback by $\tilde{\sigma}$. The action on covariant derivatives takes the form $d_{\tilde{\sigma} \ast} a = \sigma^* \circ d_a \circ \sigma^*$, where $\sigma^* : \Omega(Y^{(2)}, Q^{(2)}(g)) \to \Omega(Y^{(2)}, Q^{(2)}(g))$ is pullback by $\sigma$. The induced action on the tangent space $T_a \mathcal{A}(Q^{(2)}) = \Omega^1(Y^{(2)}, Q^{(2)}(g))$ is given by pullback by $\sigma$. Likewise, the curvature satisfies $F_{\tilde{\sigma} \ast} a = \sigma^* F_a$. In particular, the flow equation (67) on the double $Y^{(2)}$ is invariant under the action of $\tilde{\sigma}$. We set $\epsilon_Q := \epsilon_{Q^{(2)}}/2$, where $\epsilon_{Q^{(2)}} > 0$ is as in (58).

Now suppose $a \in A^{1,2}(Q, \partial Q)$ has $\mathcal{YM}_Q(a) < \epsilon_Q$. Then $a$ has a unique extension $a^{(2)}$ to all of $Q^{(2)}$, satisfying $\tilde{\sigma}^* a^{(2)} = a^{(2)}$. We call $a^{(2)}$ the double of $a$, and we claim that $a^{(2)} \in A^{1,2}(Q^{(2)})$. 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 $\partial Q$. 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: $\lim_{y \to \partial Y} \partial_n a = - \lim_{y \to \partial Y} \partial_n \tilde{\sigma}^* a$, unless they are both zero, and this latter condition is not imposed by our boundary conditions.) More generally, every $a \in A^{1,2}(Q, \partial Q)$ is a $W^{1,2}$-limit of smooth connections $a_j$ whose normal component vanishes in a neighborhood of the boundary. By the linearity of the integral it is immediate that the doubles of the $a_j$ converge to $a^{(2)}$ in $W^{1,2}$, and this proves the claim.
By assumption, we have $\mathcal{YM}_{Q(2)}(a^{(2)}) < \bar{c}_{Q(2)}$, so by the discussion at the beginning of this section, there is a unique solution $a^{(2)}(\tau)$ to the flow equation \((67)\) on the closed 2-manifold $Y^{(2)}$, with initial condition $a^{(2)}(0) = a^{(2)}$. Furthermore, the limit $\text{Heat}_{Q(2)}(a^{(2)}):= \lim_{\tau \to \infty} a^{(2)}(\tau)$ exists and is flat. Since \((67)\) is $\bar{c}$-invariant, the uniqueness assertion guarantees that $\overline{\sigma} \cdot a^{(2)}(\tau) = a^{(2)}(\tau)$ for all $\tau$. In particular,

$$
(61) \quad \overline{\sigma} \cdot \text{Heat}_{Q(2)}(a^{(2)}) = \text{Heat}_{Q(2)}(a^{(2)}).
$$

Define $\text{Heat}_Q(a) := \text{Heat}_{Q(2)}(a^{(2)})|_Q$. Then \((61)\) shows that

$$
\iota_{\partial_n} \text{Heat}_Q(a)|_{\partial Y} = 0,
$$

so $\text{Heat}_Q$ does map into $A^{1,2}_{\text{flat}}(Q, \partial Q)$. Similarly, gauge transformation $u \in G(Q, \partial Q)$ has a unique extension to a $\overline{\sigma}$-invariant gauge transformation in $G(Q^{(2)})$. In particular, $\text{Heat}_Q(u^*a) = u^*\text{Heat}_Q(a)$ follows from the $G(Q^{(2)})$-equivariance of $\text{Heat}_{Q(2)}$. This finishes the proof of 1.

To prove 2, we will assume $\Sigma \subset \text{int} Y$. The remaining case $\Sigma \subset \partial Y$ follows by replacing $Y$ with its double, for then we have $\Sigma \subset \text{int} Y^{(2)}$ and the analysis carries over directly. For sake of contradiction, suppose there is some sequence $a_\nu \in A^{1,2}(Q)$ with $||F_{a_\nu}||_{L^2} \to 0$, but

$$
(62) \quad c_0 \leq ||(\text{Heat}_Q(a_\nu) - a_\nu)|_{\Sigma}||_{L^4(\Sigma)}
$$

for some fixed $c_0 > 0$. By Uhlenbeck’s weak compactness theorem, there is a sequence of gauge transformations $u_\nu \in G^{2,2}$ such that $u_\nu^*a_\nu$ converges weakly in $W^{1,2}$ (hence strongly in $L^4$) to a limiting connection $a_\infty \in A^{1,2}(Q)$, after possibly passing to a subsequence. Then $a_\infty$ is necessarily flat. Be redefining $u_\nu$, if necessary, we may assume that each $u_\nu^*a_\nu$ is in Coulomb gauge with respect to $a_\infty$, and still retain the fact that $u_\nu^*a_\nu$ converges to $a_\infty$ strongly in $L^4$. Then

$$
||u_\nu^*a_\nu - a_\infty||_{W^{1,2}}^2 = ||u_\nu^*a_\nu - a_\infty||_{L^2}^2 + ||d_{a_\infty}(u_\nu^*a_\nu - a_\infty)||_{L^2}^2 \leq C_1(||u_\nu^*a_\nu - a_\infty||_{L^2}^2 + ||F_{a_\nu}||_{L^2}^2 + ||u_\nu^*a_\nu - a_\infty||_{L^4}^4)
$$

for some constant $C_1$. Observe that the right-hand side is going to zero, so $a_\nu$ is converging in $W^{1,2}$ to the space of flat connections, and so

$$
(63) \quad ||a_\nu - (u_\nu^{-1})^*a_\infty||_{W^{1,2}} \to 0.
$$

On the other hand, by the trace theorem [26, Theorem B.10], we have

$$
(64) \quad c_0 \leq ||\text{Heat}_Q(a_\nu) - a_\nu||_{L^4(\Sigma)} \leq C_2||\text{Heat}_Q(a_\nu) - a_\nu||_{W^{1,2}(Y)}
$$

for some $C_2$ depending only on $Y$ and $1 \leq q \leq 4$ (the inequality on the left is \((63)\)). Since $\text{Heat}_Q$ 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 $a_\nu$ is within $\epsilon'$ of the space of flat connections, then

$$
C_2||\text{Heat}_Q(a_\nu) - a_\nu||_{W^{1,2}(Y)} \leq \frac{c_0}{2}.
$$

By \((63)\) $a_\nu$ is within $\epsilon'$ of $A_{\text{flat}}(Q)$ for $\nu$ large, and so we have a contradiction to \((64)\).

The next lemma states that we can always put a connection $a \in A(Q)$ in a gauge so that it is an element of $A(Q, \partial Q)$. This is basically just a variation on the fact that connections can be put in temporal gauge. We state a version with an additional $\mathbb{R}$ parameter, since this is the context in which the lemma will be used. 

\[\square\]
Lemma 4.7. Let $R \to Z$ be as in the introduction to Section 4. Then for every $A \in \mathcal{A}^{1,2}(R)$ there is an identity-component gauge transformation $U \in \mathcal{G}^{1,2}_{loc}(R)$ with

$$U^*A|_{\{s\} \times Y_{\alpha(s)}^{(i)}} \in \mathcal{A}^{1,2}(Q_{\alpha(s)}^{(i)}), \quad \forall i \in \{0, \ldots, N-1\}, \forall s \in \mathbb{R}.$$  

Furthermore, if $A$ is smooth then $U^*A$ is smooth as well.

Proof of Remark 4.4. We suppress the Sobolev exponents, unless they are relevant. By definition, (65)

$$\frac{d}{d\tau} \tilde{\alpha}(\tau) = -d^*_{\tilde{\alpha}(\tau)} F_{\tilde{\alpha}(\tau)} + d_{\tilde{\alpha}(\tau)} \xi(\tau), \quad \tilde{\alpha}(0) = \alpha$$

has a unique solution $\tau \mapsto \tilde{\alpha}(\tau)$ for all $0 \leq \tau < \infty$. The solution has the further property that it takes the form $\tilde{\alpha}(\tau) = \mu(\tau)^* \alpha$, for some path of complex gauge transformations $\mu(\tau) \in G(P)^C$ starting at the identity. It is then immediate that $\alpha(\tau) := \exp \left( \int_0^\tau \xi \right)^* \tilde{\alpha}(\tau)$ solves (57), and so

$$\text{Heat}_P(\alpha) = \lim_{\tau \to \infty} \exp \left( \int_0^\tau \xi \right)^* \mu(\tau)^* \alpha.$$ 

Clearly $\exp \left( \int_0^\tau \xi \right)^* \mu(\tau)^* \alpha$ lies in the complex gauge orbit of $\alpha$ for all $\tau$, and so $\text{Heat}_P(\alpha)$ must as well. Now $\text{NS}_P(\alpha)$ lies in the complex gauge orbit by definition, so there is some complex gauge transformation $\tilde{\mu} \in G(P)^C$ (possibly depending on $\alpha$) with $\tilde{\mu}^* \text{NS}(\alpha) = \text{Heat}(\alpha)$. We will be done if we can show that $\tilde{\mu}$ is a real gauge transformation that lies in the identity component. The former statement is equivalent to showing $\bar{h} := \tilde{\mu}^* \tilde{\mu} = \text{Id}$. By (14) we must have that $\bar{h}$ is a solution to

(66) $$\bar{F}(h) := i\partial_{\text{Heat}(\alpha)} (h^{-1} \partial_{\text{Heat}(\alpha)} h) = 0.$$ 

Clearly the identity gauge transformation is a solution as well. It suffices to show that (66) has a unique solution, at least for $\alpha$ close to the space of flat connections. The map $\bar{F}$ is defined on (the $W^{2,2}$-completion of) $G(P)^C$, we can take its codomain to be (the $L^2$-completion of) $\Omega^0(\Sigma, P(\mathfrak{g})^C)$. Similarly to our analysis of $F$ in the proof of Theorem 3.5, the derivative of $\bar{F}$ at the identity is

$$W^{2,2}(P(\mathfrak{g})^C) \longrightarrow L^2(P(\mathfrak{g})^C), \quad \eta \mapsto \frac{1}{2} \Delta_{\text{Heat}(\alpha)} \eta.$$ 

which is invertible. So by the inverse function theorem $\bar{F}$ is a diffeomorphism in a neighborhood of the identity. This is the uniqueness we are looking for, provided we can arrange so that $\bar{h}$ lies in a suitably small neighborhood of the identity. However, this is immediate since the gauge transformation $\bar{h}$ depends continuously on $\alpha$ in the $W^{1,2}$-topology, and $\bar{h} = \text{Id}$ if $\alpha$ is flat.

To finish the proof of (57), we need to show that $\tilde{\mu} \in G(P)$ is actually in the identity component $G_0(P)$. However, this is also immediate from the continuous dependence of $\tilde{\mu}$ on $\alpha$. Indeed, pick any path from $\alpha$ to $A_{\text{flat}}(P)$ that never leaves a suitably small neighborhood of $A_{\text{flat}}(P)$. Applying the above construction to the values of this path of connections provides a path of real gauge transformations from $\tilde{\mu}$.

4.2. Uniform elliptic regularity. We establish an elliptic estimate that will be used in the proof of Lemma 4.2 below. Throughout we will use the notation from the introduction to Section 4. We also introduce the $\epsilon$-dependent norm

$$\|\eta\|^2_{L^2(U),\epsilon} := \int_U \langle \eta \wedge * \eta \rangle,$$

for subsets $U \subseteq \mathbb{R} \times Y$, where $*_{\epsilon}$ is the Hodge star on $\mathbb{R} \times Y$ determined by the metric $g_{\epsilon}$. 


Proposition 4.8. For any \( \rho > 0 \) and compact \( K \subset (-\rho, \rho) \times Y \), there is some \( C = C(K, \rho) > 0 \) with

\[
\|\nabla_s F_A\|_{L^2(K), \epsilon} + \|\nabla_s^2 F_A\|_{L^2(K), \epsilon} \leq C \|F_A\|_{L^2((-\rho, \rho) \times Y), \epsilon}
\]

for all \( \epsilon > 0 \) and all \( \epsilon \)-ASD connections \( A \). Here \( \nabla_s = \partial_s + [p, \cdot] \), where \( p \) is the \( ds \)-component of \( A \).

We will use Proposition 4.8 in the following capacity.

Corollary 4.9. Given any compact \( K \subset \mathbb{R} \times I \times \Sigma_* \), there is some \( C = C(K) > 0 \) with the following significance. Let \( \epsilon > 0 \) and suppose \( A \) is an \( \epsilon \)-ASD connection that limits to flat connections \( a^\pm \in \mathcal{A}_{flat}(Q) \) at \( \pm \infty \). Let \( -\beta_s \) denote the \( ds \)-component of \( F_A|_{\mathbb{R} \times I \times \Sigma_*} \). Then

\[
\|\nabla_s \beta_s\|_{L^2(K)}^2 + \|\nabla_s^2 \beta_s\|_{L^2(K)}^2 \leq C \left( \mathcal{CS}(a^-) - \mathcal{CS}(a^+) \right).
\]

Proof of Corollary 4.9. Let \( \eta \) be a 1-form on \( \Sigma_* \). Then by the conformal scaling of 1-forms on surfaces, we have \( \|\eta \wedge ds\|_{L^2(K), \epsilon} = \|\eta\|_{L^2(K)} \) is independent of \( \epsilon \). Let \( A \) be as in the statement of the corollary. Then \( ds \wedge \nabla_s \beta_s \) is a component of \( \nabla_s F_A \), so we have

\[
\|\nabla_s \beta_s\|_{L^2(K)}^2 = \|ds \wedge \nabla_s \beta_s\|_{L^2(K), \epsilon} \leq \|\nabla_s F_A\|_{L^2(K), \epsilon}^2.
\]

By Proposition 4.8 this is bounded by a constant times \( E^\text{inst}_{\epsilon}(A) = \mathcal{CS}(a^-) - \mathcal{CS}(a^+) \). The same computation holds with \( \nabla_s \) replaced by \( \nabla_s^2 \).

Proof of Proposition 4.8. We first prove

\[
\|\nabla_s F_A\|_{L^2(K), \epsilon} \leq C \|F_A\|_{L^2(\Omega), \epsilon}.
\]

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 \( (-\rho, \rho) \times Y \). There is a constant \( C_0 \), depending only on \( h \) (and consequently only on \( \rho \) and \( K \)), with

\[
|\partial_s h| + |\partial_s^2 h| \leq C_0 h.
\]

Set \( \Omega := (-\rho, \rho) \times Y \). Then we have

\[
\|\nabla_s F_A\|_{L^2(K), \epsilon}^2 \leq \|h \nabla_s F_A\|_{L^2(\Omega), \epsilon}^2 = \int_{\Omega} h^2 (\nabla_s F_A \wedge \nabla_s F_A) = - \int_{\Omega} h^2 (\nabla_s F_A \wedge \nabla_s F_A),
\]

where we have used the \( \epsilon \)-ASD condition and the commutativity relation \( \nabla_s \ast_s = \ast_s \nabla_s \). Write \( A = a(s) + p(s) \, ds \). Then we have \( F_A = F_{a(s)} - b_s \wedge ds \), where \( b_s := \partial_s a - d_A p \). This gives

\[
\nabla_s F_A = \partial_s F_A + [p, F_A] = d_A \partial_s A - d_A (d_A p) = d_A (\partial_s A - d_A p).
\]

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 get

\[
\nabla_s F_A = d_A b_s,
\]

where we are viewing \( b_s \) as a 1-form on the 4-manifold \( \mathbb{R} \times Y \). Using (71) and Stokes’ theorem, we obtain
\[
\|\nabla_s F_A\|_{L^2(K),\epsilon}^2 \leq \|h\nabla_s F_A\|_{L^2(\Omega),\epsilon}^2 = -\int_{\Omega} h^2 (d_A b_s \wedge d_A b_s) \\
= \int_{\Omega} 2h \partial_1 h \, ds \wedge \langle b_s \wedge d_A b_s \rangle - \int_{\Omega} h^2 \langle b_s \wedge [F_A \wedge b_s] \rangle \\
= \int_{\Omega} 2h \partial_1 h \, ds \wedge \langle b_s \wedge d_A b_s \rangle - \int_{\Omega} h^2 \langle b_s \wedge [b_s \wedge b_s] \rangle \wedge ds,
\]
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
\[
2ab \leq 2a^2 + \frac{1}{2}b^2
\]
with \( \|F_A\|_{L^2(\Omega),\epsilon} \) and the identity \( \equiv \) on the first term on the right
\[
\|h\nabla_s F_A\|_{L^2(\Omega),\epsilon}^2 \leq 2C_0 \|hb_s\|_{L^2(\Omega),\epsilon}^2 + \frac{1}{2}\|h\nabla_s F_A\|_{L^2(\Omega),\epsilon}^2 - \int_{\Omega} h^2 \langle b_s \wedge [b_s \wedge b_s] \rangle \wedge ds.
\]
Subtract the term \( \frac{1}{2}\|h\nabla_s F_A\|_{L^2(\Omega),\epsilon}^2 \) from both sides to get
\[
\frac{1}{2}\|\nabla_s F_A\|_{L^2(\Omega),\epsilon}^2 \leq \frac{1}{2}\|h\nabla_s F_A\|_{L^2(\Omega),\epsilon}^2 \leq 2C_0 \|hb_s\|_{L^2(\Omega),\epsilon}^2 - \int_{\Omega} h^2 \langle b_s \wedge [b_s \wedge b_s] \rangle \wedge ds.
\]
We record the following for later use:
\[
0 \leq 2C_0 \|hb_s\|_{L^2(\Omega),\epsilon}^2 - \int_{\Omega} h^2 \langle b_s \wedge [b_s \wedge b_s] \rangle \wedge ds.
\]
We want to estimate the right-hand side of (73) in terms of \( \|F_A\|_{L^2(\Omega),\epsilon} = 2\|b_s\|_{L^2(\Omega),\epsilon} \). Set
\[
g(s) := 2C_0 \|hb_s\|_{L^2(Y),\epsilon}^2 - \int_Y h^2 \langle b_s \wedge [b_s \wedge b_s] \rangle
\]
(so if we integrate over \((-\rho, \rho)\) then we recover the right-hand side of (73), since \( \Omega = (-\rho, \rho) \times Y \)). Then \( \equiv \equiv \), and the fact that \( h \) vanishes outside of \((-\rho, \rho)\), gives
\[
\int_{-\rho}^{\rho} g(s) \, ds \geq 0, \quad \text{and} \quad g(s) = 0 \quad \text{for} \quad s \notin [-\rho, \rho].
\]
Next, set \( e(s) := \frac{1}{2}\|hb_s\|_{L^2(Y),\epsilon}^2 \).

**Claim:** There are constants \( D_0, D_1 > 0 \), depending only on \( K, \rho \), with \( e''(s) + D_0 e(s) \geq D_1 g(s) \).

Then \( \equiv \equiv \) follows from the claim and Lemma \([1, 1]\) below (the latter is where we use (75)):
\[
\|\nabla_s F_A\|_{L^2(K),\epsilon}^2 \leq 2g(s) \leq C \int_{[-\rho-1, \rho+1]} \int_Y e = C \int_{[-\rho, \rho]} \int_Y e \leq \frac{1}{2}C \|b_s\|_{L^2(\Omega),\epsilon}^2 \quad = \quad C\|F_A\|_{L^2(\Omega),\epsilon}^2,
\]
where the first inequality is (73), the second inequality is the assertion of Lemma \([1, 1]\) and the other estimates follow from the definitions.

To prove the claim, we have \( e''(s) = \|\nabla_s (hb_s)\|_{L^2(Y),\epsilon}^2 + \int_Y \langle \nabla_s \nabla_s (hb_s) \wedge s^Y hb_s \rangle \), where \( s^Y \) is the Hodge star on \( Y \) coming from \( g_s \). We also have
\[
\nabla_s^2 (hb_s) = (\partial_s^2 h) \ b_s + 2 (\partial_s h) \nabla_s b_s + (d_a^* \epsilon d_a b_s - h \ s^Y [b_s \wedge b_s]),
\]
where \( d_a^* \epsilon = s^Y d_a s^Y \) on 1- and 2-forms on \( Y \). Then
\[
\nabla_s^2 (hb_s) = \epsilon \ b_s + 2 (\partial_s h) \nabla_s b_s + (d_a^* \epsilon d_a b_s - h \ s^Y [b_s \wedge b_s]),
\]
where \( d_a^* \epsilon = s^Y d_a s^Y \) on 1- and 2-forms on \( Y \). Then
We begin with the second term by applying (69):

\[
\int_Y \langle (\partial_s^2 h) b_s \rangle \geq -C_1 e.
\]

For the third term in (76) we do the same, except we also use (72):

\[
2 \int_Y \langle (\partial_s h) \nabla b_s \rangle \geq -2e - \frac{1}{2} \| h \nabla b_s \|_{L^2(Y), \epsilon}^2.
\]

For the fourth term in (76) we integrate by parts and use (72):

\[
\int_Y h^2 (d_s^* d_s b_s \wedge \ast_e b_s) = \int_Y h^2 (d_s \ast_e d_s b_s \wedge b_s) = - \int_Y 2h dh \wedge (\ast_e d_s b_s \wedge b_s) + \| h d_s b_s \|_{L^2(Y), \epsilon}^2
\]

\[
\geq -2e - \frac{1}{2} \| h d_s b_s \|_{L^2(Y), \epsilon}^2 + \| h d_s b_s \|_{L^2(Y), \epsilon}^2 = -2e + \frac{1}{2} \| h d_s b_s \|_{L^2(Y), \epsilon}^2.
\]

Putting this all together, and using \( d_s b_s = -\ast_e \nabla b_s \), we get

\[
e''(s) \geq \| \nabla_s (h b_s) \|^2_{L^2(Y), \epsilon} - C_2 e - \int_Y h^2 (\ast_e [b_s \wedge b_s] \wedge \ast_e b_s)
\]

\[
\geq -C_2 e - \int_Y h^2 (\ast_e [b_s \wedge b_s] \wedge \ast_e b_s).
\]

By adding \((C_2 + 2C_0) e\) to both sides we recover the claim and this finishes the proof of the first derivative bound (68).

To finish the proof of the proposition, we need to prove \( \| \nabla^2 F_A \|_{L^2(K), \epsilon} \leq C \| F_A \|_{L^2(\Omega), \epsilon} \). This is very similar to the proof of (68), so we only sketch the main points. The analogue of (72) for this case is

\[
0 \leq \frac{1}{2} \| \nabla^2 F_A \|^2_{L^2(K), \epsilon} \leq 2C_0 \| h \nabla b_s \|^2_{L^2(\Omega), \epsilon} + \int_{-\rho}^{0} \bar{g}(s) ds,
\]

where we have set \( g_1(s) := -3 \int_Y h^2 (\nabla b_s \wedge [\nabla b_s \wedge b_s]) + \int_Y h^2 ([b_s \wedge b_s] \wedge \nabla^2 b_s). \) Then (75) continues to hold with \( g \) replaced by \( g_1 \). Set \( \varepsilon_1(s) := \frac{1}{2} \| h \nabla b_s \|^2_{L^2(Y), \epsilon} \) and, just as before, one can show \( e''_1(s) + D_0 \varepsilon_1(s) \geq D_1 g_1(s) \). Then the result follows from Lemma 1.11.

**Remark 4.10.** (a) Write \( Y = Y_1 \cup I \times \Sigma \). The proof we give here proves the proposition with \( \nabla \) replacing \( \nabla_s \), but only for compact sets lying in the open set \( \mathbb{R} \times (0, 1) \times \Sigma \). It fails to provide an \( \epsilon \)-independent bound for compact sets \( K \) that intersect \( \mathbb{R} \times Y \). However, if \( K \subset \mathbb{R} \times [0, 1] \times \Sigma \), intersects the seams, then the proof we give here shows the weaker estimate

\[
e^{1/2} \| \nabla_t F_A \|_{L^2(K), \epsilon} + e^{1/2} \| \nabla_s \nabla_t F_A \|_{L^2(K), \epsilon} + \epsilon \| \nabla^2 F_A \|_{L^2(K), \epsilon} \leq C \| F_A \|_{L^2((-\rho, \rho) \times Y), \epsilon}.
\]

To see this, multiply \( \nabla_t \) by a bump function \( h \) supported in a small neighborhood of \( K \). Then the only modification to the proof is that the estimate (68) needs to be replaced by

\[
\epsilon |\partial_t h| + \epsilon |\partial_s \partial_t h| + \epsilon^2 |\partial_t^2 h| \leq C_0 h.
\]
(b) The projection operator \( \text{proj}_{\alpha^r} \) appearing in (22) and (50) is not very natural, since the connection \( \alpha^r \) is not flat. However, it can be removed at the cost of weakening the sup norm to an \( L^r \)-norm for \( 1 < r < \infty \). We argue how this can be done in the framework of Lemma 4.2; the framework of Lemma 3.3 is similar, but the estimates are easier to obtain.

We will show that under the hypotheses of Lemma 4.2, we have

\[
\int_K \| \text{proj}_{\alpha^r} \circ D_\alpha \omega - \text{Ad}(\mu^{-1}_r)D_\alpha \omega \|_{L^2(\Sigma^r)}^r \, \text{dvol}_S \to 0
\]

for each compact \( K \subset \mathbb{R} \times I \) and \( 1 < r < \infty \). Set \( \beta_{s,\infty} := \partial_s \alpha_{s,\infty} - d_{s,\infty} \phi_{s,\infty} \), where we have written \( A_{s,\infty}|_{\mathbb{R} \times I} = \alpha_{s,\infty} + \phi_{s,\infty} \psi \). Define \( \beta_{t,\infty} \) and \( \beta_{s,t} \) similarly. Since \( \alpha_{s,\infty} \) is flat, it follows that \( \beta_{s,\infty} \) (resp. \( \beta_{t,\infty} \)) is the \( \alpha_{s,\infty} \)-harmonic projection of \( \partial_s \alpha_{s,\infty} \) (resp. \( \partial_t \alpha_{s,\infty} \)), and so

\[
\text{proj}_{\alpha^r}(x) \circ D_\alpha \omega = ds \otimes \beta_{s,\infty} + dt \otimes \beta_{t,\infty}.
\]

On the other hand, since \( \alpha^r \) is not flat the situation is not as simple. However, we can write

\[
\text{proj}_{\alpha^r}(x) \circ D_\alpha \omega = ds \otimes \text{proj}_{\alpha^r} \beta_{s,t} + dt \otimes \text{proj}_{\alpha^r} \beta_{s,t},
\]

since, e.g., \( \beta_{s,t} \) and \( \text{proj}_{\alpha^r} \beta_{s,t} \) differ by an element of \( \text{im} \partial_s \alpha_{s,\infty} \oplus \text{im} \partial_t \alpha_{s,\infty} \). By the holomorphic/\( \epsilon \)-ASD equations, the behavior of the ds-component is identical to the behavior of the dt-component, and so to prove (77) it suffices to restrict attention to the ds-component of the object in the \( L^2(\Sigma^r) \)-norm. Then with this notation, (50) implies

\[
\int_K \| \beta_{s,\infty} - \text{Ad}(\mu^{-1}_r(x))\text{proj}_{\alpha^r}(x) \beta_{s,t} \|^2_{L^2(\Sigma^r)} \, \text{dvol}_S \to 0
\]

for every compact \( K \subset S_0 \) and (77) would follow if we could show

\[
\int_K \| \beta_{s,\infty} - \text{Ad}(\mu^{-1}_r(x))\text{proj}_{\alpha^r}(x) \beta_{s,t} \|^2_{L^2(\Sigma^r)} \, \text{dvol}_S \to 0.
\]

By the triangle inequality, it therefore suffices to show \( \int_K \| \beta_{s,t} - \text{proj}_{\alpha^r} \beta_{s,t} \|^2_{L^2(\Sigma^r)} \to 0 \). Working pointwise on \( S_0 \), (73) gives

\[
\| \beta_{s,t} - \text{proj}_{\alpha^r} \beta_{s,t} \|^2_{L^2(\Sigma^r)} \leq C \left( \| d_{s,t} \beta_{s,t} \|^2_{L^2(\Sigma^r)} + \| d_{s,t} \text{proj}_{\alpha^r} \beta_{s,t} \|^2_{L^2(\Sigma^r)} \right)
\]

where in the equality we have used the identities \( \nabla_{s,t} F_{\alpha_{s,t}} = d_{s,t} \beta_{s,t} \) and \( \nabla_{s,t} F_{\alpha_{s,t}} = d_{s,t} \beta_{s,t} \). Squaring and integrating over \( K \subset S_0 \) already proves the result for \( r = 2 \) because Proposition 4.8 and Remark 4.10 (a) give

\[
\| \nabla_{s,t} F_{\alpha_{s,t}} \|^2_{L^2(K \times \Sigma^r)} + \| \nabla_{s,t} F_{\alpha_{s,t}} \|^2_{L^2(K \times \Sigma^r)} \leq \epsilon^{1/2} C R E_{\text{inst}}(A_{s,t}),
\]

where we have used the scaling property \( \| \eta \|^2_{L^2(\Sigma^r)} = \| \eta \|^2_{L^2(\Sigma^r)} \) for 2-forms \( \eta \) on \( \Sigma \). Since the \( A_{s,t} \) have uniformly bounded \( \epsilon \)-energy, this last term is going to zero and proves (77) with \( r = 2 \).

For arbitrary \( r \), one needs to work a little harder. Consider the two sequences of maps \( S_0 \to L^2(P_{\Sigma^r}(g)) \)

\[
(s,t) \to \ast_{\Sigma^r} \nabla_{s,t} F_{\alpha_{s,t}} , \quad \ast_{\Sigma^r} \nabla_{s,t} F_{\alpha_{s,t}}
\]

where \( L^2(P_{\Sigma^r}(g)) \) is the \( L^2 \)-completion of \( \Omega^0(\Sigma^r, P_{\Sigma^r}(g)) \). By Remark 4.10 (a) each sequence is uniformly bounded in \( W^{1,2}(K) \) for compact \( K \subset S_0 \). It follows by Sobolev embedding that, for each sequence, a subsequence converges in \( L^r(K) \) for \( 1 < r < \infty \). As in the case \( r = 2 \), it is not hard to see that these subsequences converge to zero, and this combines with (78) to prove (77).
Lemma 4.11. Consider functions $e, f, g : B_{R+r} \to \mathbb{R}$, where $B_x := (-x,x) \subset \mathbb{R}$. Assume $e \geq 0$ is $C^2$, $f \geq 0$ is $C^0$, and $g$ is $C^0$ and satisfies $g(s) \geq 0$ for all $s \in B_{R+r}\setminus B_R$ and $\int_{B_R} g \geq 0$. If $g(s) \leq f(s) + e''(s)$ for all $s \in B_{R+r}$, then
\[
\int_{B_R} g \leq \int_{B_{R+r}} f + \frac{4}{\nu^2} \int_{B_{R+r}\setminus B_R} e.
\]

Proof. This is a variation of [13, Lemma 9.2]. By considering the rescaled functions
\[
e(s) := e(rs), \quad \tilde{f}(s) := r^2 f(rs), \quad \tilde{g}(s) := r^2 g(rs),
\]
(which satisfy $\tilde{g} \leq \tilde{f} + \tilde{e}'$), it suffices to assume $r = 1$. The positivity conditions on $f$ and $g$ give
\[
\int_{-R}^{R} g - \int_{-R-1}^{R+1} f \leq \int_{-R-s}^{R+s} g - f = \int_{-R-s}^{R+s} \frac{d^2}{ds^2} e = e'(R + s) - e'(-R - s),
\]
for all $s \in [0,1]$. (If we know that $e' = 0$ on $B_{R+1}\setminus B_R$, as we do in the proof of Proposition 1.8 then the Lemma is proved at this point.) Note also that we have $\frac{2}{\nu} (e(R + s) + e(-R - s)) = e'(R + s) - e'(0 - s)$. In particular, integrating in $s$ from $1/2$ to $t$ gives
\[
\frac{1}{2} \left( \int_{-R}^{R} g - \int_{-R-1}^{R+1} f \right) \leq e(R + t) + e(-R - t) - e(R + 1/2) - e(-R - 1/2)
\leq e(R + t) + e(-R - t)
\]
by the positivity of $e$. Now integrate in $t$ from $1/2$ to $1$:
\[
\frac{1}{4} \left( \int_{-R}^{R} g - \int_{-R-1}^{R+1} f \right) \leq e(R + t) + e(-R - t) = \int_{B_{R+1}\setminus B_{R+1/2}} e.
\]
\[\square\]

4.3. Proof of Lemma 4.12. Throughout we write $A_{\nu} |_{\{s\} \times \Sigma} = a_{\nu}(s) + p(s) \, ds$, $A_{\nu} |_{\{s,t\} \times \Sigma_{\nu}} = \alpha_{\nu}(s,t) + \phi_{\nu}(s,t) \, ds + \psi(s,t) \, dt$, $\beta_{s,\nu} = \partial_s \alpha_{\nu} - d_{\alpha_{\nu}} \phi_{\nu}$, and $\beta_{t,\nu} = \partial_t \alpha_{\nu} - d_{\alpha_{\nu}} \psi_{\nu}$. We use similar notation for $A_{\infty}$.

By the assumption on the curvature of the $\alpha_{\nu}$, it follows that Theorem 3.6 applies to $\alpha_{\nu}(s,t)$ for each $s,t$ and $\nu$ sufficiently large. Let $NS_{P_{1,\nu}}$ be the map constructed in Theorem 3.6 for the bundle $P_{1,\nu} \to \Sigma_{\nu}$. Set $M := M(P_0)^- \times M(P_1) \times M(P_2)^- \times \cdots \times M(P_{N-1})$, and define a map
\[
v_{\nu} : \mathbb{R} \times I \to M,
\]
\[
(s,t) \mapsto (\Pi \circ NS_{P_0} (\alpha_{\nu}(s,1-t)|\Sigma_0), \ldots, \Pi \circ NS_{P_{N-1}} (\alpha_{\nu}(s,t)|\Sigma_{N-1}));
\]
the even terms have $1 - t$, the odd terms have $t$. Then Lemmas 3.16 and 3.17 imply that $v_{\nu}$ is holomorphic with respect to the complex structure $\sum_{i=0}^{N-1} (-1)^{i+1} \Sigma_i$ on $M$. (The map $v_{\nu}$ is just a folded up quilt in $M = (M(P_i))$, as in [29], except it will not have Lagrangian boundary conditions at the seams.) We will denote by $| \cdot |_M$ the norm induced by this almost complex structure. Since $S_0 = \mathbb{R} \times I \setminus B$, and $B$ is a finite set, it follows from the removal of singularities theorem for holomorphic maps [18, Theorem 4.1.2 (ii)] that each $v_{\nu}$ extends to a holomorphic map defined on all of the interior $\mathbb{R} \times (0,1)$ (we do not have Lagrangian boundary conditions, so we it may not extend over $B \cap \mathbb{R} \times \{0,1\}$.

It follows exactly as in Claim 1 appearing in Section 3.2 that for each compact $K \subset S_0$ there is a uniform bound of the form $\sup_K |\partial_\nu v_{\nu}|^2_M \leq C_K$. In particular, there is a subsequence, still denoted by $\{v_{\nu}\}$, that converges weakly in $C^1$, and strongly in $C^0$, on compact subsets of $S_0$, including the boundary. Let $v_{\infty} \in C^1(S_0,M)$ denote the limiting holomorphic curve. As with the $v_{\nu}$, $v_{\infty}$ extends to $\mathbb{R} \times (0,1)$ and is $C^\infty$ in this region [18, Theorem B.4.1].
Remark 4.12. We can actually say quite a bit more: The uniform energy bound implies that, after possibly passing to a further subsequence, we have that the \( v_\nu : \mathbb{R} \times I \to M \) converge to \( v_\infty \) in \( \mathcal{C}^\infty \) on compact subsets of the interior, \( S_0 \cap \mathbb{R} \times (0, 1) \) (see [18, Theorem 4.1.1]). In particular, this automatically proves [57] for \( K \subset S_0 \cap \mathbb{R} \times (0, 1) \). However, for applications we will need this for \( K \) including the boundary \( \partial S_0 \); we address this in Claim 2, below.

Claim 1 below states that \( v_\infty \) actually does have Lagrangian boundary conditions. This will follow because the \( v_\nu \) have approximate Lagrangian boundary conditions. To state this precisely, consider the product

\[
L(0) := L(Q_{01}) \times L(Q_{23}) \times \ldots \times L(Q_{(N-2)(N-1)}) \subset M,
\]

where \( L(Q_{i(i+1)}) \subset M(R_i) \times M(R_{i+1}) \) is the Lagrangian coming from the flat connections on \( Q_{i(i+1)} \) as in Section 2. Similarly, we obtain a second Lagrangian \( L(1) \subset M \) coming from the product of the \( L(Q_{i(i+1)}) \) for \( i \) odd. By construction, there are canonical bijections \( L(0) \cap L(1) \cong L(Q) \cong \mathcal{A}_{\text{flat}}(Q)/\mathcal{G}_{\Sigma} \), where \( \mathcal{G}_{\Sigma} \) is defined in the beginning of Section 4.

Let \( B \) be as in the statement of Lemma 4.2 and let \( B_\mathbb{R} \) denote the set of \( s \in \mathbb{R} \) such that \( (s, 0), (s, 1) \notin B \).

Claim 1: For \( j \in \{0, 1\} \) and for each \( s \in B_\mathbb{R} \), the sequence \( (v_\nu(s,j))_\nu \) converges (in the metric on \( M \)) to a point in \( L(j) \). In particular, since \( B_\mathbb{R} \) is finite \( v_\infty \), extends to a map defined on all of \( \mathbb{R} \times I \) with Lagrangian boundary conditions: \( v_\infty(\cdot, j) : \mathbb{R} \to L(j) \).

By applying suitable gauge transformations, we may assume each \( A_\nu \) satisfies the conclusions of Lemma 4.2. Fix a compact \( K \subset \mathbb{R} \setminus B_\mathbb{R} \). Consider the hypothesis in Lemma 4.2 stating that the norms \( \|F_{A_\nu,\nu}(s)\|_{L^\infty} \) decay to zero uniformly on \( K \). This implies that Theorem 4.3 applies to each \( a_\nu(s) \) for each \( s \in K \) and for all \( \nu \) sufficiently large (exactly how large will depend on \( K \)). For each \( i \), let \( \Pi \circ \text{Heat}_{Q_{i(i+1)}}(a_\nu(s)) \), be the assignment \( s \mapsto \Pi \circ \text{Heat}_{Q_{i(i+1)}}(a_\nu(s)) \) and because \( \Pi \circ \text{Heat}_{Q_{i(i+1)}}(a_\nu(s)) \) determines a map \( \ell_{\nu,0} : K \to L(0) \). Similarly, we obtain a map \( \ell_{\nu,1} : K \to L(1) \) by using \( \Pi \circ \text{Heat}_{Q_{i(i+1)}}(a_\nu(s)) \) and with \( i \) odd. We will show

\[
(81) \quad \sup_{s \in K} \text{dist}_M(\ell_{\nu,j}(s), v_\nu(s,j)) \xrightarrow{\nu} 0.
\]

Then Claim 1 follows by repeating this for a sequence of compact \( K \) that exhaust \( \mathbb{R} \setminus B_\mathbb{R} \).

The proof of \((81)\) is just a computation. Fix \( s \in K \); we keep track of the bundle \( P \) in the notation for the projection \( \Pi_P : \mathcal{A}_{\text{flat}}(P) \to M(P) \). By the definition of distance on \( M \), we have that \( \text{dist}_M(\ell_{\nu,j}(s), v_\nu(s,j)) \) is given by

\[
\sum_{i} \text{dist}_M(\Pi_{Q_{i(i+1)}}(a_\nu(s), |_{Y_{i(i+1)}}), \Pi_{P_{i+j}} \circ \text{NS}_{P_{i+j}}(a_\nu(s), |_{|_{\Sigma_{i+j}}}))^2
\]

\[
= \sum_{i} \text{dist}_M(\Pi_{P_{i+j}}(\Pi_{Q_{i(i+1)}}(a_\nu(s), |_{Y_{i(i+1)}}), |_{|_{\Sigma_{i+j}}}), \Pi_{P_{i+j}}(\Pi_{Q_{i(i+1)}}(a_\nu(s), |_{Y_{i(i+1)}}), |_{|_{\Sigma_{i+j}}}), \Pi_{P_{i+j}}(\Pi_{Q_{i(i+1)}}(a_\nu(s), |_{Y_{i(i+1)}}), |_{|_{\Sigma_{i+j}}}))^2
\]

\[
\leq \sum_{i} \|\text{Heat}_{Q_{i(i+1)}}(a_\nu(s), |_{Y_{i(i+1)}}), |_{|_{\Sigma_{i+j}}}) - \text{NS}_{P_{i+j}}(a_\nu(s), |_{|_{\Sigma_{i+j}}})\|^2_{L^2(|_{\Sigma_{i+j}})}.
\]

The first 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 \( M(P) \), and because \( \Pi_{P_{i+j}} \) has operator norm equal to one. Taking the supremum over \( s \in K \) and using the triangle inequality, we can continue this to get that \( \text{sup}_s \text{dist}_M(\ell_{\nu,j}(s), v_\nu(s,j))^2 \) is bounded by
\[
\sup_s \sum_i \left\{ \left\| \text{Heat}_{Q(i+1)} a_{\nu}(s) Y_{i(i+1)} \right\|_{\Sigma_{i+j}}^2 - \left\| a_{\nu}(s) Y_{i(i+1)} \right\|_{\Sigma_{i+j}}^2 \right\} \\
+ \left\| a_{\nu}(s) Y_{i(i+1)} \right\|_{\Sigma_{i+j}}^2 - \left\| \text{NS}_{P,j} \left( \alpha_{\nu} \left( s, j \right) \right) \right\|_{L^2(\Sigma_{i+j})}^2 \right\} \\
= \sup_s \sum_i \left\{ \left\| \text{Heat}_{Q(i+1)} a_{\nu}(s) Y_{i(i+1)} \right\|_{\Sigma_{i+j}}^2 - \left\| a_{\nu}(s) Y_{i(i+1)} \right\|_{\Sigma_{i+j}}^2 \right\} \\
+ \left\| \alpha_{\nu} \left( s, j \right) \right\|_{\Sigma_{i+j}}^2 - \left\| \text{NS}_{P,j} \left( \alpha_{\nu} \left( s, j \right) \right) \right\|_{L^2(\Sigma_{i+j})}^2 \right\}.
\]

The equality holds because \( \alpha_{\nu} \) agrees with \( a_{\nu} \) at the boundary of \( Y \). The second part of Theorem 4.3 shows that the first term in the summand goes to zero as \( v \to \infty \), since \( F_{a_{\nu}} \) converges to zero in \( L^\infty \) (uniformly in \( s \in K \)). Similarly, the second term in the summand goes to zero by Proposition 3.14. This verifies (81) and proves Claim 1.

Just as in Claim 2 appearing in Section 3.2, the holomorphic strip \( v_\infty : \mathbb{R} \times I \to M \) lifts to a smooth \( \alpha_\infty : \mathbb{R} \times I \to A_{\text{flat}}(P) \), and hence determines a quilted holomorphic cylinder trajectory \( A_\infty \in \mathcal{A}_{\text{loc}}^{L^q}(\mathbb{R} \times Q) \) (this follows because \( \mathbb{R} \times S^1 \) retracts to its 1-skeleton). Then the convergence statement in (68) follows exactly as the proof of (21) in Lemma 3.3. It remains to prove (69).

Claim 2: For each compact \( K \subset S_0 \), we have \( \sup_v \| v_\nu \|_{W^{3,\infty}(K)} \leq C_K \).

Before proving Claim 2, we show how it is used to prove (69). First of all, by the compact Sobolev embedding \( W^{3,2} \hookrightarrow C^1 \) on surfaces, Claim 2 implies that, after passing to a subsequence, the \( v_{\nu} \) converge to \( v_\infty \) in \( C^1 \) on compact subsets of \( S_0 \). In particular, \( \partial_s v_{\nu} \) converges to \( \partial_s v_\infty \) in \( C^0 \) on compact subsets, and the proof of (69) follows exactly as the proof of (22) in Lemma 3.3.

To prove Claim 2, and thereby complete the proof of Lemma 4.2, we need to bound all mixed partial derivatives of \( v_{\nu} \) up to degree 3. Since \( v_{\nu} \) is holomorphic, this reduces to finding uniform bounds for the first, second and third s-derivatives of \( v_{\nu} \). We will use the relation \( v_{\nu} = \Pi \circ \text{NS}_P(\alpha_{\nu}) \) to translate derivatives on \( v_{\nu} \) into derivatives on \( \alpha_{\nu} \in A^{1,q}(P) \). Note that we have the freedom to choose a convenient representative in the gauge equivalence class of \( A_{\nu} \).

We begin with the first s-derivative. The components of \( \partial_s v_{\nu} \) in \( M = M(P_0) \times \cdots \times M(P_{N-1}) \) are given by \( D_{\alpha_{\nu}} (\Pi \circ \text{NS}_P) (\partial_s \alpha_{\nu}) = D_{\alpha_{\nu}} (\Pi \circ \text{NS}_P) (\beta_{s,v}) \), where the equality holds because \( \partial_s \alpha_{\nu} - \beta_{s,v} \) is an exact 1-form and so lies in the kernel of the linearization \( D_{\alpha_{\nu}} (\Pi \circ \text{NS}_P) \). By Theorem 3.6 (iii), the \( L^2 \)-operator norms of the operators \( D_{\alpha_{\nu}} (\Pi \circ \text{NS}_P) \) are uniformly bounded, so to bound \( \| \partial_s v_{\nu} \|_{L^2} \) it suffices to bound \( \| \beta_{s,v} \|_{L^2} \). However, this is a component of the energy of \( A_{\nu} \), which we have assumed is uniformly bounded:

\[
\| \beta_{s,v} \|_{L^2(K \times \Sigma_{\nu})}^2 \leq \| F_{A_{\nu}} \|_{L^2(\mathbb{R} \times Y), \nu}^2 = 2L_{\text{inst}}(A_{\nu}).
\]

For the second s-derivative, the product rule implies that the second s-derivative of the components of \( v_{\nu} \) are controlled by

\[
D^2_{\alpha_{\nu}} (\Pi \circ \text{NS}_P) (\beta_{s,v}, \partial_s \alpha_{\nu}) + D_{\alpha_{\nu}} (\Pi \circ \text{NS}_P) (\partial_s \beta_{s,v}) \\
= D^2_{\alpha_{\nu}} (\Pi \circ \text{NS}_P) (\beta_{s,v}, \beta_{s,v}) + D_{\alpha_{\nu}} (\Pi \circ \text{NS}_P) (\partial_s \beta_{s,v})
\]

where the equality holds because \( \Pi \circ \text{NS}_P \) is gauge-equivariant, and so \( \text{im}(d_{\alpha}) \) lies in the kernel of the Hessian \( D^2_{\alpha} (\Pi \circ \text{NS}_P) \). Consider the Hessian term in (82). Lemma 3.14 says that the \( L^2 \)-operator norms of the operators \( D^2_{\alpha_{\nu}} (\Pi \circ \text{NS}_P) \) are uniformly bounded, and so bounding the Hessian terms reduces to bounding \( \| \beta_{s,v} \|_{L^2(K \times \Sigma_{\nu})} \), which we have already done in (69). To bound the second s-derivative of \( v_{\nu} \), it therefore suffices to bound \( D_{\alpha_{\nu}} (\Pi \circ \text{NS}_P) (\partial_s \beta_{s,v}) \). As before, since \( D_{\alpha_{\nu}} (\Pi \circ \text{NS}_P) \) is uniformly \( L^2 \)-bounded as an operator, it suffices to bound \( \| \partial_s \beta_{s,v} \|_{L^2} \). For this, we exploit the gauge freedom and assume that \( \phi_{\nu} = 0 \) (i.e., \( A_{\nu} \) is in temporal gauge). Then \( \| \partial_s \beta_{s} \|_{L^2(K \times \Sigma_{\nu})} = \| \nabla_{\Sigma} \beta_{s} \|_{L^2(K \times \Sigma_{1})} \), and the result follows by Corollary 3.9.
The bound for the third $s$-derivatives is similar to the bound for the second $s$-derivative. This completes the proof of Lemma [122].

4.4. Proof of Theorem [141]. We follow the proof of Theorem [163] quite closely, with Lemma [142] used in place of Lemma [143]. In particular, we assume there is some compact $K \subset \mathbb{R}$ for which one of the following cases holds

Case 1: $\|F_{\alpha_r}\|_{L^\infty(K \times I \times \Sigma_j)} + \|F_{a_r}\|_{L^\infty(K \times Y)} \to \infty$;
Case 2: $\|F_{\alpha_r}\|_{L^\infty(K \times I \times \Sigma_j)} + \|F_{a_r}\|_{L^\infty(K \times Y)} \to \Delta > 0$;
Case 3: $\sup_{x \in K \times I} \|\text{proj}_{\alpha_r} \circ D_x \alpha_r\|_{L^2(\Sigma)} \to \infty$, and for all compact $K' \subset Z$ we have $\|F_{\alpha_r}\|_{L^\infty(K')} + \|F_{a_r}\|_{L^\infty(K')} \to 0$.

Most of the work is in showing that each case leads to energy quantization. Supposing we have shown this, it would then follow (from Lemma [122] applied to the complement of the bubbling set) that a subsequence of the $A_r$ converges in the sense of the statement of Theorem [111] (with $J = 1$ and $s_r = 0$) to a limiting holomorphic quilted cylinder representative $A^1$ on $R \to Z$. The exception is that one does not know that $A^1$ has non-zero energy. For example, it could be the case that all of the energy has escaped to $\infty$, and so $A^1$ is a flat connection. To rectify this, one incorporates time translations, and we defer this until after of our case analysis for energy quantization.

Case 1. (Instantons on $S^4$) By passing to a subsequence, we may assume the $L^\infty$-norm of each curvature is always achieved on $\Sigma_{j+1}$ or $Y_{j(j+1)}$ for some $j$:

\begin{align}
(84) & \quad \|F_{\alpha_r}\|_{L^\infty(K \times I \times \Sigma_{j+1})} \geq \max \{\|F_{\alpha_r}\|_{L^\infty(K \times I \times \Sigma_j)}, \|F_{a_r}\|_{L^\infty(K \times Y)}\} \quad \forall \nu, \text{ or} \\
(85) & \quad \|F_{\alpha_r}\|_{L^\infty(K \times I \times \Sigma_{j+1})} \geq \max \{\|F_{\alpha_r}\|_{L^\infty(K \times I \times \Sigma_j)}, \|F_{a_r}\|_{L^\infty(K \times Y)}\} \quad \forall \nu.
\end{align}

Find points $(s_\nu, t_\nu) \in K \times I$ if (S4) holds, or $s_\nu \in K$ if (S5) holds, with

$$
\|F_{\alpha_r(s_\nu, t_\nu)}\|_{L^\infty(\Sigma_{j+1})} = \|F_{\alpha_r}\|_{L^\infty(K \times I \times \Sigma_{j+1})}, \\
\text{or} \quad \|F_{\alpha_r(s_\nu)}\|_{L^\infty(Y_{j(j+1)})} = \|F_{\alpha_r}\|_{L^\infty(K \times Y_{j(j+1)})}.
$$

By passing to a subsequence, we may suppose the $s_\nu$ converge to some element of $K \subset \mathbb{R}$. Similarly, we may assume $t_\nu \to t_\infty \in I$ converges. Strictly speaking, we need to distinguish between whether $t_\infty$ lies in the interior of $I$, or on the boundary. However, the analysis for when $t_\infty$ lies in the boundary can be incorporated to the analysis for when (S5) holds. Precisely, we find ourselves considering the following subcases:

Subcase 1: (S4) holds and $t_\infty \neq 0, 1$;
Subcase 2: (S5) holds, or (S2) holds and $t_\infty = 0, 1$.

Without loss of generality, we may suppose $j = 0$ and $t_\infty \in [0, 1)$.

In Subcase 1, for each $\nu$ define a rescaled connection $\tilde{A}_\nu$ in terms of its components as follows:

\begin{align}
\tilde{\alpha}_\nu(s, t) & := \alpha(\epsilon_\nu s + s_\nu, \epsilon_\nu t + t_\nu)|_{\Sigma_j} \\
\tilde{\phi}_\nu(s, t) & := \epsilon_\nu \phi(\epsilon_\nu s + s_\nu, \epsilon_\nu t + t_\nu)|_{\Sigma_j} \\
\tilde{\psi}_\nu(s, t) & := \epsilon_\nu \psi(\epsilon_\nu s + s_\nu, \epsilon_\nu t + t_\nu)|_{\Sigma_j}.
\end{align}

We view these as connections and 0-forms defined on $B_{\epsilon_{\nu}^{-1}} \times \Sigma_j \subseteq \mathbb{C} \times \Sigma_2$, where $\eta = \frac{1}{2} \min \{t_\infty, 1 - t_\infty\}$, $B_\eta \subset \mathbb{C}$ is the ball of radius $r$ centered at zero, and we assume $\nu$ is large enough so that $t_\nu \leq \eta$.

In Subcase 2, for each $\nu$ we define a connection $\tilde{A}_\nu$ on a neighborhood of $\mathbb{R} \times Y_{01}$ as follows:

$$
\tilde{\alpha}_\nu(s) := a_\nu(\epsilon_\nu s + s_\nu)|_{Y_{01}}, \quad \tilde{\rho}_\nu(s) := \epsilon_\nu \rho_\nu(\epsilon_\nu s + s_\nu)|_{Y_{01}}.
$$
As in Case 1 from Section 3.3, this means that we have energy quantization with which we view as a connection defined on $[23]$ (see also $[18$, Theorem 4.6.1] for the closely-related case of spaces on which these bubbles form can be more exotic. Define that instantons near the blow-up point bubble off. However, this time the geometry of the underlying geometry of the underlying

$$X(r) := X \cup_{\partial X} [0, r) \times \partial X, \quad X^\infty := X \cup_{\partial X} [0, \infty) \times \partial X.$$  

Remark 4.13. There exists smooth structures on these spaces that are compatible in the sense that the inclusions

$$\mathbb{R} \times X^\infty = \left(\mathbb{R} \times X\right) \cup \left(\mathbb{R} \times \partial X\right).$$

In both Subcases the connections $\bar{A}_\nu$ are ASD with respect to the fixed metric, and have uniformly bounded energy $\frac{1}{2} \| F_{\bar{A}_\nu} \|_{L^2}^2 \leq CS(\alpha^-) - CS(\alpha^+)$; here the norm should be taken on the domain on which the connection is defined. Furthermore, the energy densities are bounded from below:

$$\| F_{\bar{A}_\nu} \|_{L^\infty} \geq \| F_{\bar{A}_\nu} \|_{L^\infty} + \| F_{\bar{A}_\nu} \|_{L^\infty} = \| F_{\bar{A}_\nu} \|_{L^\infty} + \| F_{\bar{A}_\nu} \|_{L^\infty}.$$  

In particular, the condition of Case 1 implies that $\| F_{\bar{A}_\nu} \|_{L^\infty} \to \infty$. Following the usual rescaling argument $[23]$ [9] Section 9] (see also $[13]$ 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, there is a subsequence that converges, modulo gauge, in $C^\infty$ on compact sets to a finite-energy, non-constant instanton $\bar{A}_\infty$ on $\mathbb{R}^4$. By Uhlenbeck’s removable singularities theorem this extends to a non-constant instanton, also denoted by $\bar{A}_\infty$, on a $PU(r)$-bundle $R_\infty \to S^4$. Since $\bar{A}_\infty$ is ASD and non-constant we have

$$0 < \frac{1}{2} \int_{S^4} |F_{\bar{A}_\infty}|^2 = -\frac{1}{2} \int_{S^4} \langle F_{\bar{A}_\infty} \wedge F_{\bar{A}_\infty} \rangle = -2\pi^2 \kappa r^{-1} q_4(R_\infty).$$

As in Case 1 from Section 3.3, this means that we have energy quantization with $\hbar = 4\pi^2 \kappa r$.

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. Define $\bar{A}_\nu$ exactly as in Case 1 above. Everything up to and including equation $[21]$ continues to hold. In particular, $\liminf \| F_{\bar{A}_\nu} \|_{L^\infty}$ is bounded from below by $\Delta > 0$. After possibly passing to a subsequence, we may assume the energy densities $\| F_{\bar{A}_\nu} \|_{L^\infty}$ converge to some $\Delta' \in [\Delta, \infty]$. If $\Delta' = \infty$ then we are done by precisely the
same analysis as in Case 1. So we may assume $0 < \Delta' < \infty$, in which case we can apply Uhlenbeck’s strong compactness theorem directly to the sequence $\tilde{A}_\nu$. We may therefore assume this sequence converges to a non-flat finite-energy instanton $\tilde{A}_\infty$ on a bundle over one of the spaces $\mathbb{R} \times Y_0^{\infty}$ or $\mathbb{C} \times \Sigma_1$, depending on whether we are in Subcase 1 or 2 (see the discussion above Remark 4.13 for a definition of $Y_0^{\infty}$). We show in [10] that the energy of $A_\infty$ is $2\pi^2k_\nu^{-1}k$ for some positive $k \in \mathbb{N}.$

Case 3. (Holomorphic spheres and disks in $M(P)$) Write $\beta_{s,\nu} := \partial_s \alpha_\nu - d_{\alpha_\nu} \phi_\nu$, and set $c_\nu := \sup_{j,K,I} \| \text{proj}_{\tilde{A}_\nu} \beta_{s,\nu} \|_{L^2(\Sigma_1)}$. The conditions of this case imply that $c_\nu \to \infty$. Find $j_\nu \in \{0, \ldots, N-1\}$ and points $(s_\nu, t_\nu) \in K \times I$ with $c_\nu = \| \text{proj}_{\tilde{A}_\nu} \beta_{s,\nu}(s_\nu, t_\nu) \|_{L^2(\Sigma_1)}$ (such points exist since $\beta_{s,\nu}$ decays at $\pm \infty$, due to the finite energy assumption; alternatively, one could replace $c_\nu$ by $c_\nu/2$, without changing the argument). By passing to a subsequence, we can assume that $j_\nu = 1$ for all $\nu$ and that the $(s_\nu, t_\nu)$ converge to some $(s_\infty, t_\infty) \in K \times I$. The two relevant subcases to consider are as follows:

Subcase 1: $t_\infty \in (0, 1)$
Subcase 2: $t_\infty \in \{0, 1\}$

We may assume, without loss of generality, that if Subcase 2 holds then $t_\infty = 0$. Define rescaled connections $\hat{A}_\nu$ using (80) and (87), except replace every $\epsilon_\nu$ by $\epsilon_\nu^{-1}$ (the subcases here correspond to those from Case 1 in the obvious way). Subcase 1 can be treated exactly as Case 3 from Section 4.2.

For Subcase 2 we will use an argument similar to the one appearing in that section. The argument will show that a holomorphic disk bubble appears in $M(P_0) \times M(P_1)$ with Lagrangian boundary conditions in $L(Q_{01}) \subset M(P_0) \times M(P_1)$.

The rescaling for Subcase 2 is such that we view the connections $\hat{A}_\nu$ as being defined on $\mathbb{R} \times Y_0^{\infty}(c_\nu)$ (see the discussion above Remark 4.13). The components of $F_{\hat{A}_\nu}$ satisfy

$$\hat{\beta}_{s,\nu} + \hat{\gamma} = 0, \quad \hat{\gamma} = -\hat{\epsilon}_\nu^{-2} \ast F_{\hat{A}_\nu}, \quad \hat{b}_{s,\nu} = -\hat{\epsilon}_\nu^{-1} F_{\hat{A}_\nu},$$

where $\hat{\epsilon}_\nu := c_\nu \epsilon_\nu$, and the Hodge star is the one on the surface $\Sigma_0 \cup \Sigma_1$. It may not be the case that the $\hat{\epsilon}_\nu$ are decaying to zero; this is replaced by the assumption in this case that the slice-wise curvatures converge to zero in $L^\infty$:

$$\| F_{\hat{A}_\nu} \|_{L^\infty} = \| F_{\hat{A}_\nu} \|_{L^\infty}, \quad \| F_{\hat{A}_\nu} \|_{L^\infty} = \| F_{\hat{A}_\nu} \|_{L^\infty} \to 0.$$ 

Our choice of rescaling also gives

$$\| \text{proj}_{\hat{A}_\nu} \beta_{s,\nu}(0, 0) \|_{L^2(\Sigma_1)} = 1.$$

By arguing as in Lemma 4.12 it follows that, after possibly passing to a subsequence, there exists a sequence of gauge transformations $\mu_\nu : \mathbb{H} \to \mathcal{G}(P_0 \sqcup P_1)$, and a limiting connection $\hat{A}_\infty \in \mathcal{A}(\mathbb{R} \times Q_{01}^{\infty})$ that is a holomorphic representative

$$\hat{\beta}_{s,\infty} + \ast \hat{\beta}_{l,\infty} = 0, \quad F_{\hat{A}_\infty} = 0, \quad F_{\hat{A}_\infty} = 0,$$

and satisfies

$$\sup_{K'} \| \text{Ad}(\mu_\nu^{-1}) \text{proj}_{\hat{A}_\nu} \beta_{s,\nu} - \beta_{s,\infty} \|_{L^2(\Sigma_1)} \to 0$$

for all compact $K' \subset \mathbb{H}$; here we are using [10] to, e.g., view $\alpha_\infty$ as a map defined on $\mathbb{H}$. Let $\Pi_{P_1} : \mathcal{A}_{\text{flat}}(P_1) \to M(P_1)$ and $\Pi_{Q_{01}} : \mathcal{A}_{\text{flat}}(Q_{01}) \to L(Q_{01})$ be the projections to the moduli spaces. Then

$$v_\infty := (\Pi_{P_0}(\hat{A}_\infty|_{\Sigma_0}), \Pi_{P_1}(\hat{A}_\infty|_{\Sigma_1})) : \mathbb{H} \to M(P_0) \times M(P_1)$$

is a holomorphic curve with Lagrangian boundary conditions $\mathbb{R} \to L(Q_{01}) \subset M(P_0) \times M(P_1)$ determined by $a_\infty : \mathbb{R} \to \mathcal{A}_{\text{flat}}(Q_{01})$. Furthermore, $v_\infty$ has bounded energy.
Fixing the \( v \) represented by some \( X \) by a suitable function. We describe this now, freely referring to the notation established above.

Then the case analysis above combines with Lemma 4.2 to show that, after passing to a subsequence,

\[
\int_{
\mathbb{R}^2}
|\partial_x v^\infty|^2 = \int_{
\mathbb{R}^2 \times \Sigma_1 \cup \Sigma_2}
|\hat{\beta}_{s,v}|^2 \leq \liminf_{\nu \to \infty} ||\hat{\beta}_{s,v}||_{L^2(\mathbb{R} \times Y), e_{\nu}}
\]

\[
\leq \liminf_{\nu \to \infty} \frac{1}{2}||F_{A_{\nu}}||_{L^2(\mathbb{R} \times Y), e_{\nu}} = (CS(a^-) - CS(a^+)).
\]

In particular, the removal of singularities theorem [18, Theorem 4.1.2 (ii)] applies and so \( v^\infty \) extends to a holomorphic disk \( v^\infty : \mathcal{D} \rightarrow M(P_1) \times M(P_2) \) with Lagrangian boundary conditions. Then \( \nu^j \) combines with (92) to give

\[
|\partial_x v^\infty(0,0)| \geq ||\text{proj}_{\hat{\beta}_{s,v}}(0,0)||_{L^2(\Sigma_1)} = \lim_{\nu \to \infty} ||\text{proj}_{\hat{\beta}_{s,v}}(0,0)||_{L^2(\Sigma)} = 1.
\]

In particular, \( v^\infty \) is non-constant. The minimal energy of non-constant holomorphic disks is half that of spheres, so we can take \( \hbar = 2\pi^2 \kappa_{n-1} \) in the case; see also [18, Proposition 4.1.4].

Finally, we address translations; we follow the strategy of [22]. The moduli space of flat connections on \( Q \) is canonically identified with the set of Lagrangian intersection points \( L(0) \cap L(1) \), and the non-degeneracy assumption on the elements of \( \mathcal{A}_{\text{flat}}(Q) \) implies that \( L(0) \cap L(1) \) is a finite set in \( M \); see [3, Section 4]. In particular, there is some \( \epsilon_0 > 0 \) so that \( B_{\epsilon_0}(p) \cap B_{\epsilon_0}(p') = \emptyset \), for all \( p, p' \in L(0) \cap L(1) \). Define \( v_{\nu} \) as in (90). By assumption, each \( A_{\nu} \) converges at \( \pm \infty \) to the flat connection \( a^\pm \). Since the maps \( NS_{P_1} \) preserve flat connections, it follows that each \( v_{\nu} \) converges at \( \pm \infty \) to the Lagrangian intersection point \( p^\pm \in L(0) \cap L(1) \) associated to \( a^\pm \).

Define

\[
s^j_{\nu} := \text{sup} \left\{ s \in \mathbb{R} \mid \text{dist}_{M(P_1)}(p^-, v_{\nu}(s, t)) \leq \epsilon_0, \text{ for all } t \in I \right\}.
\]

(We may assume \( p^- \neq p^+ \), otherwise all instantons are trivial and all holomorphic curves are constant; in particular, the set defining \( s^j_{\nu} \) is non-empty.) Then for each \( \nu \) we have

\[
dist_{M(P_1)} \left( p^-, \left( \tau_{s_{\nu}^j} v_{\nu} \right)(s, t) \right) \leq \epsilon_0 \quad \forall t \in I, \forall s \leq 0,
\]

\[
dist_{M(P_1)} \left( \left( \tau_{s_{\nu}^j} v_{\nu} \right)(0, t), p^+ \right) = \epsilon_0 \quad \text{for some } t \in I.
\]

Then the case analysis above combines with Lemma 4.2 to show that, after passing to a subsequence, the translates \( \tau_{s_{\nu}^j} v_{\nu} \) converge on compact sets off of a finite bubbling set to a limiting holomorphic strip \( v^1 \). This limits to some Lagrangian intersection point \( p^0 \) at \( -\infty \) and \( p^1 \) at \( \infty \). By [18] and the definition of \( \epsilon_0 \), we must have \( p^0 = p^- \). On the other hand, the equalities expressed in \( \text{[18]} \) show that \( v^1(0, t) \) is not at a Lagrangian intersection point for some \( t \in I \). In particular, \( v^1 \) is non-constant and so \( p^+ \neq p^0 \). This proves that \( v^1 \) has positive energy, and also shows that the \( \tau_{s_{\nu}^j} v_{\nu} \) become arbitrarily close to \( p^1 \).

Continue inductively with \( p^1 \) replacing \( p^0 \), etc. to obtain a sequence of limiting holomorphic strips \( v^j \) that limit to Lagrangian intersection points \( p^{j-1} \) and \( p^j \). The theorem follows by lifting the \( v^j \) and \( p^j \) to representatives, and converting the convergence of the \( \tau_{s_{\nu}^j} v_{\nu} \) to statements about the representatives, as we did in the proof of Lemma 4.2.

5. Perturbations

Theorems 3.3 and 4.1 both have extensions to the case where the ASD equations are perturbed by a suitable function. We describe this now, freely referring to the notation established above.

We begin with Theorem 3.3. Suppose \( H \) is a section of the bundle \( T^* S \otimes \text{Maps}(\mathcal{A}(P), \mathbb{R})^{G(P)} \rightarrow S \). Fixing \( x \in S \) and \( v \in T_x S \), then the differential of the contraction \( \iota_v H(x) : \mathcal{A}(P) \rightarrow \mathbb{R} \) can be represented by some \( X^H_{x,v} : \mathcal{A}(P) \rightarrow \Omega^1(\Sigma, P(g)) \) in the sense that

\[
d(\iota_v H(x))_{\alpha}(\nu) = \int_{\Sigma} \langle X^H_{x,v}(\alpha) \wedge \nu \rangle
\]
for all $\alpha \in \mathcal{A}(P)$ and all $v \in \Omega^1(\Sigma, P(g))$. We assume $H$ has been chosen so that for each $x,v$, 

(a) $\sup_{\alpha \in \mathcal{A}(P)} \left\| X^H_{x,v}(\alpha) \right\|_{L^\infty(\Sigma)} < C_{x,v}$, for some $C_{x,v}$ that depends continuously on $x,v$, and 

(b) if $\{\alpha_j\}$ is any sequence of connections on $R$ such that $\left\| \ell_* (A_{\nu} - A_0) \right\|_{L^p(Z)}$ is bounded, then $X^H_{x,v}(\ell_* A_\nu)$ has an $L^p(Z)$-convergent subsequence.

In particular, Kronheimer proves that these conditions are satisfied whenever $H$ is defined by the holonomy around loops in $\Sigma$ [10, Lemma 10].

Allowing $x$ and $v$ to vary, $X^H_{x,v}$ naturally determines a map $X = X^H : \mathcal{A}(R) \to \Omega^2(\Sigma, R(g))$ by declaring $X(A)$ to be the $2$-form defined for $v \in T_x \Sigma$, $w \in T_{\Sigma} S$ by

$$X(A)(v, w) := X^H_{x,v}(\alpha(x))(w), \quad X(A)(w, v) := -X(A)(v, w)$$

and defined to be zero otherwise. This formula combines with the $\mathcal{G}(P)$-invariance of $H$ to imply that $X$ satisfies $X(U^* A) = \text{Ad}(U^{-1})X(A)$ for all $U \in \mathcal{G}(R)$. In local coordinates $x = (s,t)$ on $S$, $X$ has the form $X(A)(\alpha) = ds \wedge F_x(\alpha) + dt \wedge G_x(\alpha)$ for $x$-dependent $F_x(\alpha), G_x(\alpha) \in \Omega^1(\Sigma, P(g))$.

The relevant perturbed $\epsilon$-instanton equation is

$$\left( F_A - X(A) \right)^+ := \frac{1}{2} \left( 1 + \ast_x \right) (F_A - X(A)) = 0.$$ 

In local coordinates, this condition has the form (18), where the zero on the right-hand side of the top equation is replaced by $F_x(\alpha) + \ast_x G_x(\alpha)$, $X$ is the symplectic side, the gauge invariance implies that $X$ determines a $1$-form on $S$ with values in the vertical bundle in $TM(P)$; we denote this $1$-form by the same symbol $X$. The relevant holomorphic curve equation is

$$X(A)(\alpha) := \frac{1}{2} \left( (dv - \text{proj}_x X(\alpha)) + \ast_x (dv - \text{proj}_x X(\alpha)) \circ j_{\Sigma} \right) = 0,$$

where $\alpha$ is a lift of $v$, and $\text{proj}_x$ is the $L^2$-orthogonal projection to the $\alpha$-harmonic space (the gauge invariance of $H$ implies that this is independent of the lift $\alpha$). Suppose $A \in \mathcal{A}(R)$ represents a solution of this equation. In local coordinates, this has the form (19), where the zero in the right-hand side of the first equation is again replaced by $F_x(\alpha) + \ast_x G_x(\alpha)$. Then Theorem 3.3 continues to hold in this $X$-perturbed setting, provided one replaces the energies $E_{\text{sym} \alpha}$ and $E_{\text{inst} \alpha}$ with the $X$-perturbed energies obtained by replacing $F_A$ with $F_A - X(A)$.

The adjustments to the proof are as follows: One should replace every appearance of $D_x \alpha_\nu$ with $D_x \alpha_\nu - X(\alpha_\nu)$, and likewise with $D_x \alpha_\infty$. Then Claim 1 in the proof of Lemma 3.3 would show $D_{\nu} - D_{\nu} (\Pi \circ NS) X(\alpha_\nu)$ is uniformly bounded on compact subsets of $S_0$ (in local coordinates, the $ds$-component of this is $ds \nu_x - D_{\alpha_x} (\Pi \circ NS) F_x(\alpha_\nu)$). It follows from assumption (a) on $H$ that the $D_{\nu}$ is uniformly bounded, and so converge weakly in $C^1$ on compact sets to a limiting $v_{\infty}$. The $v_\nu$ satisfy $\overline{\partial}_{\nu} = X_{\nu}^{(0,1)}$, where $X_{\nu}^{(0,1)}$ is the $(0,1)$-component of $D_{\alpha_x} (\Pi \circ NS) X(\alpha_\nu)$. It follows from assumption (b) on $H$ that the limiting map $v_{\infty}$ satisfies (97). Then the proofs of Lemma 3.3 and Theorem 3.3 continue as in the unperturbed case.

Now we address Theorem 4.1. Let $H$ be a function as above, with $S := \mathbb{R} \times I$ and $P := P_*$. We also assume that $X$ and its first derivatives are bounded in $C^0$ on $\mathbb{R} \times I \times \Sigma_*$ (see the next paragraph for more details). In addition to these assumptions, we assume that, for any connection $A$ over $\mathbb{R} \times I \times \Sigma_*$, the induced $2$-form $X(A)$ vanishes to all orders near $\mathbb{R} \times \{0,1\} \times \Sigma_*$ for all of these conditions are satisfied if $H$ is obtained by integrating the holonomy around thickened loops lying in $\mathbb{R} \times \{0,1\} \times \Sigma_*$; see [3, Section 5.3]. Then $X$ admits a canonical smooth extension as a map $X : \mathcal{A}(R) \to \Omega^2(\Sigma, R(g))$, by declaring $X(A)$ to be zero on $\mathbb{R} \times Y_*$. The modifications to the proofs of Theorem 4.1 and Lemma 4.2 are as follows: Of course, one needs to make the modifications discussed in the $S \times \Sigma$ case above. In addition to this, one needs to extend the analysis of Section 4.2. The only real difference from the unperturbed case shows up in (70), where we used the $\epsilon$-ASD relations. In the perturbed case there are terms of the
form \( \int h^2 \langle \nabla_s F_A \wedge \nabla_s X(A) \rangle \) and \( \int h^2 \langle \nabla_s F_A \wedge \nabla_s X(A) \rangle \) appearing on the right-hand side of (70). Working in temporal gauge, for simplicity, we have

\[
\nabla_s X(A) = ds \wedge \partial_s F(\alpha) + dt \wedge \partial_s G(\alpha) = ds \wedge (\langle \partial_s F(\alpha) + DF(\beta_s) \rangle + dt \wedge \langle \partial_s G(\alpha) + DG(\beta_s) \rangle),
\]

where \( DF \) and \( DG \) are derivatives of \( F \) and \( G \) in the \( \alpha \)-variable, and \( \partial_s F \) and \( \partial_s G \) are the derivatives in \( s \). Then these additional terms are all controlled by

\[
C \left( \| \nabla_s \beta_s \| + \| \nabla_s \beta_t \| \right) \beta_s \|, \tag{98}
\]

where the norms are \( L^2 \)-norms, and \( C \) depends linearly on the \( C^0 \) norm of the derivatives of \( X \) (which we assumed are bounded). It follows that (98) is bounded by \( \delta \| \nabla_s F_A \|^2 + \delta^{-1} C' \| F_A \|^2 \) for any \( \delta > 0 \), and so the \( \nabla_s F_A \) term can be absorbed into the left-hand side of (70). Then the proof of Lemma 4.2 goes through as before. If we assume \( H \) is independent of \( R \), then Theorem 4.1 holds with the obvious modifications made to the statement, analogous to the case of \( S \times \Sigma \) above; one also needs to replace every occurrence of the word ‘flat’ with ‘\( H \)-flat’ (see [8] for a definition).

Theorem 4.1 has an extension to the case when \( H \) does depend on \( R \) as well. In this case we assume that \( H \) is independent of \( s \) when \( |s| \) is large. One has to be a little careful working with the translations as we did at the end of Section 4.4. Let \( X_s \) be the perturbations 1-form coming from \( H \), but translated by \( s \), and let \( X_{s \pm} \) be the perturbation 1-form that is constant in the \( \mathbb{R} \)-direction and agrees with \( X_s \) for \( s \) near \( \pm \infty \). Since we do not have translational invariance, the translated holomorphic curves now satisfy a perturbed equation of the form \( \partial_s v + J\partial_t v = X_{s_j}(v) \). However, since \( H \) is independent of \( s \) for large \( |s| \), these \( \nu \)-dependent perturbations can be controlled. In particular, for each \( j \) a subsequence of \( \tau_{s_j} \nu \) converges to a curve \( \nu' \) that satisfies the perturbed equation with \( X_{s_{-}} \), \( X_{s_{1}} \), or \( X_{s_{+}} \) on the right-hand side, depending on whether \( s_j \) converges to \( -\infty \), \( s_j \in \mathbb{R} \) or \( +\infty \).

**References**

[1] M. Atiyah. New invariants of three and four dimensional manifolds. *The Math. Heritage of Hermann Weyl*, Proc. Sympos. Pure Math. 48, 285-299, 1988.

[2] Charalambous. The Yang-Mills heat semigroup on three-manifolds with boundary. 2010. arXiv:1004.1639.

[3] 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.

[4] 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.

[5] 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.

[6] S. Dostoglou, D. Salamon. Self-dual instantons and holomorphic curves. *Ann. of Math.* (2) 139 (1994), no. 3, 581-640.

[7] S. Dostoglou, D. Salamon. Corrigendum: Self-dual instantons and holomorphic curves. *Ann. of Math.* 165 (2007), 665-673.

[8] D. Duncan. An introduction to the quilted Atiyah-Floer conjecture. In preparation.

[9] D. Duncan. On the components of the gauge group for \( PU(\nu) \)-bundles. In preparation.

[10] D. Duncan. On the critical values of the Chern-Simons functional. In preparation.

[11] D. Gay, R. Kirby. Indefinite Morse 2-functions: broken fibrations and generalizations, 2011. arXiv:1102.0750.

[12] S. Grundled. Moment maps and diffeomorphisms. *Diploma Thesis*, ETH Zürich, 2005.

[13] A.R. Gaio, D. Salamon. Gromov-Witten invariants of symplectic quotients and adiabatic limits. *J. Symplectic Geom.*, Volume 3, Number 1 (2005), 55-159.

[14] V. Guillemin S. Sternberg. Geometric quantization and multiplicities of group representations. *Invent. Math.*, 67:515-538, 1982.

[15] F. Kirwan. *Cohomology of quotients in symplectic and algebraic geometry*. Mathematical Notes Series, Vol. 31, Princeton University Press, 1984. ISBN 0691083703, 9780691083704.

[16] P. Kronheimer. Four-manifold invariants from higher rank-bundles.

[17] Y. Lekili. Heegaard Floer homology of broken fibrations over the circle. arXiv:0903.1773.
[18] D. McDuff, D. Salamon. *J-holomorphic curves and symplectic topology*. American Mathematical Society Colloquium Publications, Vol. 52. American Mathematical Society, Providence, RI, 2004.

[19] M. Narasimhan, C. Seshadri. Stable and unitary vector bundles on a compact Riemann surface. *Annals of Mathematics. Second Series* 82 (3): 540-567, 1965.

[20] J. Råde. On the Yang-Mills heat equation in two and three dimensions. *J. Reine Angew.* 431, 123-163 (1992).

[21] A. Ramanathan. Stable Principal bundles on a compact Riemann surface. *Math. Ann.* 213, 29-152 (1975).

[22] D. Salamon. *Morse theory, the Conley index and Floer homology*.

[23] C. Taubes. Self-dual Yang-Mills connections on non-self-dual 4-manifolds. *J. Di. Geom.* 17 (1982), 139-170.

[24] F. Warner. *Foundations of differentiable manifolds and Lie groups*. Scott-Foreman, Glen view, Illinois, 1971.

[25] K. Wehrheim. Energy identity for anti-self dual instantons on $C \times \Sigma$. *Math. Res. Lett.* 13 (2006), 161-166.

[26] K. Wehrheim. *Uhlenbeck compactness*. EMS Series of Lectures in Mathematics, 2004.

[27] K. Wehrheim, C. Woodward. Floer field theory. Preprint.

[28] K. Wehrheim, C. Woodward. Pseudoholomorphic quilts.

[29] K. Wehrheim, C. Woodward. Quilted Floer cohomology. *Geometry and Topology*, 14:833-902.

[30] 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.