Annals of Mathematics, **139** (1994), 581–640

Self-dual instantons and holomorphic curves

By Stamatis Dostoglou and Dietmar A. Salamon*

Table of Contents

Introduction

A gradient flow of a Morse function on a compact Riemannian manifold is said to be of Morse-Smale type if the stable and unstable manifolds of any two critical points intersect transversally. For such a Morse-Smale gradient flow there is a chain complex generated by the critical points and graded by the Morse index. The boundary operator has as its (x, y)-entry the number of gradient flow lines running from x to y counted with appropriate signs whenever the difference of the Morse indices is 1. The homology of this chain complex agrees with the homology of the underlying manifold M and this can be used to prove the Morse inequalities (cf. [33], [26]).

Around 1986, Floer generalized this idea to infinite dimensional variational problems in which every critical point has infinite Morse index but the *moduli spaces* of connecting orbits form finite dimensional manifolds for every

*This research has been partially supported by the SERC.

pair of critical points. The dimensions of these spaces give rise to a *relative Morse index* and the boundary operator is defined by counting connecting orbits when the relative Morse index is 1. The resulting *Floer homology* groups have played an inportant role in symplectic geometry (cf. [15]) and in 3 and 4 dimensional topology (cf. [14]).

The Floer homology groups of a compact oriented 3-manifold M are generated by the irreducible representations of the fundamental group in SO(3). These can be thought of as flat connections on a principal SO(3)-bundle bundle $Q \to M$ and they appear as the critical points of the Chern-Simons functional on the infinite dimensional configuration space of connections on this bundle modulo gauge equivalence. The gradient flow lines of the Chern-Simons functional are the self-dual Yang-Mills instantons on the 4-manifold $M \times \mathbb{R}$ and they determine the boundary operator of the Floer homology groups $\mathrm{HF}_*^{\mathrm{inst}}(M, Q)$. This construction requires that all flat connections be nondegenerate. If this is not the case then a suitable perturbation of the Chern-Simons functional will lead to only nondegenerate critical points. A more serious restriction is that every flat connection on Q (except for the 0-connection in the case of the trivial bundle) must be a regular point for the action of the identity component $\mathcal{G}_0(Q)$ of the group of gauge transformations. This condition is satisfied if either M is a homology-3-sphere (cf. [14]) or Q restricts to a nontrivial SO(3)-bundle over some oriented embedded Riemann surface $\Sigma \subset M$ (cf. [16]).

A special case is where the bundle $Q = P_f$ is the mapping cylinder of a nontrivial SO(3)-bundle $P \to \Sigma$ over a Riemann surface Σ for an automorphism $f : P \to P$. The underlying 3-manifold is the mapping cylinder $M = \Sigma_h$ of Σ for the diffeomorphism $h : \Sigma \to \Sigma$ induced by f. The flat connections on P_f correspond naturally to the fixed points of the symplectomorphism $\phi_f : \mathcal{M}(P) \to \mathcal{M}(P)$ induced by f on the moduli space $\mathcal{M}(P)$ of flat connections on the bundle P. This moduli space is a compact symplectic manifold (without singularities) of dimension $6k - 6$ where $k \geq 2$ is the genus of Σ. It is well known that this manifold is connected and simply connected and $\pi_2(\mathcal{M}_F(P)) = \mathbb{Z}$ (cf. [2]). For any symplectomorphism $\phi : \mathcal{M} \to \mathcal{M}$ of such a symplectic manifold there are Floer homology groups $\mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}, \phi)$. In this theory the critical points are the fixed points of ϕ and the connecting orbits are pseudoholomorphic curves $u : \mathbb{R}^2 \to \mathcal{M}$ which satisfy $u(s+1, t) = \phi(u(s,t))$ and converge to fixed points $x^\pm$ of ϕ as t tends to $\pm\infty$. The Euler characteristic of $\mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}, \phi)$ is the Lefschetz number of ϕ. If ϕ is the time-1-map of a time-dependent Hamiltonian flow then the Floer homology groups are naturally isomorphic to the homology of the underlying symplectic manifold $\mathcal{M}$ (cf. [15], [27]).

Hence for every automorphism $f : P \to P$ there are two Floer homology groups $\mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}(P), \phi_f)$ and $\mathrm{HF}_*^{\mathrm{inst}}(\Sigma_h, P_f)$. Both arise from the same chain complex which is generated by the flat connections on P_f respectively the fixed points of ϕ_f. In [10] it is shown that the relative Morse indices agree and hence the chain complex carries the same grading in both theories. The main result of the present paper asserts that there is a natural isomorphism of Floer homologies

$$\mathrm{HF}_*^{\mathrm{inst}}(\Sigma_h; P_f) = \mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}(P), \phi_f).$$

In particular, when $f = \mathrm{id}$,

$$\mathrm{HF}_*^{\mathrm{inst}}(\Sigma \times S^1; P \times S^1) = \mathrm{H}_*(\mathcal{M}(P), \mathbb{Z})$$

The proof requires a comparison of the boundary operators. Think of the mapping cylinder $P_f = P_f(\varepsilon)$ as the product $P \times [0, 1/\varepsilon]$ and identify $P \times 1/\varepsilon$ with $P \times 0$ via the automorphism f. In the limit $\varepsilon \to 0$ the self-dual instantons on $P_f(\varepsilon) \times \mathbb{R}$ will become holomorphic curves in the moduli space $\mathcal{M}(P)$. In other words it follows from an implicit function theorem that near every holomorphic curve $u(s,t)$ in $\mathcal{M}(P)$ there is a self-dual instanton $a_\varepsilon(t)$ on $P_f(\varepsilon) \times \mathbb{R}$ for ε sufficiently small. Conversely, it follows from Uhlenbeck's compactness that every such family of self-dual instantons $a_\varepsilon(t)$ will converge to a holomorphic curve in $\mathcal{M}(P)$ as ε tends to zero. The details will be carried out in sections 4-10. In sections 1-3 we discuss the necessary background about Floer homology and flat connections on P.

We would like to thank the referee for pointing out gaps in some of the proofs in earlier versions of this paper. The problem treated here was suggested to us by Andreas Floer during a visit to the ETH Zürich in 1990. He was a great mathematician and a wonderful human being. We dedicate this paper to his memory.

1. Floer homology for 3-manifolds

Let $Q \to M$ be a principal bundle over a compact oriented 3-manifold with structure group $G = \mathrm{SO}(3)$ which restricts to a nontrivial bundle over some oriented embedded Riemann surface $\Sigma \subset M$. Denote by $\mathcal{A}(Q)$ the space of connections and by $\mathcal{G}_0(Q)$ the identity component of the space of gauge transformations. Associated to Q is the bundle $\mathfrak{g}_Q \to M$ via the adjoint action of G on its Lie algebra $\mathfrak{g} = \mathfrak{so}(3)$. Think of $\mathcal{A}(Q)$ as an affine subspace of the space of 1 forms on Q with values in $\mathfrak{g}$ whose parallel vector space is $\Omega^1(\mathfrak{g}_Q)$.

A gauge transformation $g : Q \to G$ is called **even** if it lifts to a map $\tilde{g} : Q \to \mathrm{SU}(2)$. The subgroup of even gauge transformations is denoted by $\mathcal{G}^{\mathrm{ev}}(Q)$. The **degree** of a gauge transformation is the integer $\deg(g) \in$

$\mathbb{Z}$ determined by the induced map on homology $\mathrm{H}_3(M) = \mathbb{Z} \to \mathrm{H}_3(G) = \mathbb{Z}$. Every even gauge transformation is of even degree and under the above assumption there exists a gauge transformation of degree 1 (cf. [16], [9]).

The perturbed Chern-Simons functional $\mathcal{CS}_H : \mathcal{A}(Q)/\mathcal{G}_0(Q) \to \mathbb{R}$ is defined by

$$\mathcal{CS}_H(a_0+\alpha) = \tfrac{1}{2}\int_M (\langle \mathrm{d}_{a_0}\alpha \wedge \alpha\rangle + \tfrac{1}{3}\langle[\alpha\wedge\alpha]\wedge\alpha\rangle) - H(a_0+\alpha)$$

for $\alpha \in \Omega^1(\mathfrak{g}_Q)$ and a fixed flat connection $a_0 \in \mathcal{A}_{\mathrm{flat}}(Q)$. Here $\langle\, ,\, \rangle$ denotes the invariant inner product on $\mathfrak{g}$ given by minus the Killing form (in the case $G = \mathrm{SO}(3)$ this is 4 times the trace). The covariant differential $\mathrm{d}_a : \Omega^1(\mathfrak{g}_Q) \to \Omega^2(\mathfrak{g}_Q)$ is defined by $\mathrm{d}_a\alpha = \mathrm{d}\alpha + [a\wedge\alpha]$ for $a \in \mathcal{A}(Q)$ and $\alpha \in \Omega^1(\mathfrak{g}_Q)$ The perturbation $H : \mathcal{A}(Q) \to \mathbb{R}$ is a function of the SU(2)-valued holonomy of the connection along finitely many *thickened* loops in M. Thus H is invariant under the action of $\mathcal{G}^{\mathrm{ev}}(Q)$. (cf. [14], [30], [9] for a precise definition of H). The Chern-Simons functional satisfies the identity

$$\mathcal{CS}_H(a) - \mathcal{CS}_H(g^*a) = 8\pi^2 \deg(g) \tag{1.1}$$

(cf. [3]). The differential of $\mathcal{CS}_H$ is given by

$$\mathrm{d}\mathcal{CS}_H(a)\alpha = \int_M \langle (F_a - Y(a))\wedge\alpha\rangle,$$

where F_a is the curvature and $Y : \mathcal{A}(Q) \to \Omega^2(\mathfrak{g}_Q)$ represents the differential of H. The function Y is smooth with respect to the L^2-topology on $\mathcal{A}(Q)$. In particular, $Y(a)$ does not depend on the derivatives of a. In other words the curvature F is a nonlinear first order operator on $\mathcal{A}(Q)$ and Y is a zeroth order (and therefore compact) perturbation. Both operators are equivariant with respect to the action of $\mathcal{G}^{\mathrm{ev}}(Q)$ on $\mathcal{A}(Q)$ and $\Omega^2(\mathfrak{g}_Q)$. The critical points of $\mathcal{CS}_H$ are called **H-flat** connections. They satisfy $F_a = Y(a)$ and the set of such connections is denoted by $\mathcal{A}_{\mathrm{flat}}(Q,H)$. The perturbation H can be chosen such that every H-flat connection a is a **nondegenerate** critical point of $\mathcal{CS}_H$ (cf. [14], [8], [30]).

The gradient flow of $\mathcal{CS}_H$ takes the form

$$\dot{a} + *F_a - *Y(a) = 0. \tag{1.2}$$

With $Y = 0$ this is the self-duality equation on $Q\times\mathbb{R}$. If a satisfies (1.2) and has finite Yang-Mills energy then $a(t)$ converges to H-flat connections on Q as t dends to $\pm\infty$ (cf. [14], [22], [30]). Fix $a^{\pm} \in \mathcal{A}_{\mathrm{flat}}(Q,H)$ and denote by $\mathcal{M}(a^-,a^+)$ the *moduli space* of gauge equivalence classes $[a]$ of solutions of (1.2) which satisfy the limit condition

$$\lim_{t\to\pm\infty} a(t) = g_{\pm}^* a^{\pm} \in \mathcal{A}_{\mathrm{flat}}(Q,H) \tag{1.3}$$

for some $g_\pm \in \mathcal{G}_0(Q)$. Here $[a_1] \equiv [a_2]$ if and only if $a_2(t) = g^* a_1(t)$ for some $g \in \mathcal{G}_0(Q)$. The solutions of (1.2) minimize the perturbed Yang-Mills action

$$\begin{aligned}\mathcal{Y}_H(a) &= \tfrac{1}{2}\int_{-\infty}^{\infty}\Big(\|F_a - Y(a)\|^2_{\mathrm{L}^2(M)} + \|\dot a\|^2_{\mathrm{L}^2(M)}\Big)\,\mathrm{d}t \\ &= \mathcal{CS}_H(a^-) - \mathcal{CS}_H(a^+)\end{aligned}$$

subject to the limit condition (1.3). The second equality holds only for $a \in \mathcal{M}(a^-, a^+)$. For a generic perturbation H the space $\mathcal{M}(a^-, a^+)$ is a smooth manifold of dimension

$$\dim \mathcal{M}(a^-, a^+) = \mu(a^+) - \mu(a^-)$$

where $\mu(a) = \frac{1}{2}\eta(\mathrm{D}_a) - \mathcal{CS}_H(a)/2\pi^2$ (cf. [8], [14], [22], [30]). Here D_a denotes the extended Hessian of $\mathcal{CS}_H$. It is a self-adjoint operator on $\Omega^1(\mathfrak{g}_Q) \oplus \Omega^0(\mathfrak{g}_Q)$ given by

$$\mathrm{D}_a = \begin{pmatrix} *\mathrm{d}_a - *\mathrm{d}Y(a) & \mathrm{d}_a \\ \mathrm{d}_a^* & 0 \end{pmatrix}$$

and $\eta(\mathrm{D}_a)$ denotes its eta-invariant (cf. [3]). In particular,

$$\mu(g^*a) - \mu(a) = \frac{1}{2\pi^2}\left(\mathcal{CS}_H(a) - \mathcal{CS}_H(g^*a)\right) = 4\deg(g) \tag{1.4}$$

for $g \in \mathcal{G}(Q)$ (cf. [3]).

The solutions of (1.2) determine a boundary operator on the chain complex

$$C_k = \bigoplus_{\substack{[a] \in \mathcal{A}_{\mathrm{flat}}(Q)/\mathcal{G}_0(Q) \\ \mu_H(a) - \mu_H(a_0) = k}} \mathbb{Z}[a].$$

Choose coherent orientations of the moduli spaces $\mathcal{M}(a^-, a^+)$ as in [14], [18]. Whenever $a \in \mathcal{M}(a^-, a^+)$ with $\mu(a^+) - \mu(a^-) = 1$ define $\nu(a) = \pm 1$ according to whether the natural flow orientation of $a(t)$ (given by time shift) agrees with this coherent orientation or not. The (a^+, a^-)-entry of the boundary operator

$$\partial : C_{k+1} \to C_k$$

is defined by taking the sum of the numbers $\nu(a)$ over all instantons $[a] \in \mathcal{M}(a^-, a^+)/\mathbb{R}$ whenever $\mu(a^+) - \mu(a^-) = 1$. In [14] Floer proved that this number is finite and that ∂ is a boundary operator, i.e. $\partial^2 = 0$. The homology groups of this chain complex are called the **Floer homology** of the pair (M, Q) and they are denoted by

$$\mathrm{HF}_*^{\mathrm{inst}}(M, Q) = \ker\partial/\mathrm{im}\,\partial$$

The Floer homology groups are independent of the metric on M and the perturbation H used to define them (cf. [14], [16]). This means that different choices of metric and perturbation give rise to natural isomorphisms. Since there exists a gauge transformation of degree 1 it follows from (1.4) that the Floer homology groups are graded modulo 4. The Euler characteristic is Casson's invariant of the pair (M, Q).

Remark 1.1. If H can be chosen invariant under all gauge transformations (not just the even ones) then the group Γ of components of the space of degree-0 gauge transformations acts on $\mathrm{HF}_k^{\mathrm{inst}}(M, Q)$ for every k. This requires an equivariant perturbation theory which takes account of the action of a finite group.

2. Floer homology for symplectic fixed points

Let $(\mathcal{M}, \omega)$ be a $2n$-dimensional symplectic manifold and $\phi : \mathcal{M} \to \mathcal{M}$ be a symplectomorphism. This means that ω is a nondegenerate closed 2 form and $\phi^*\omega = \omega$. The tangent bundle of any symplectic manifold admits an almost complex structure $J : T\mathcal{M} \to T\mathcal{M}$ which is compatible with ω in the sense that $\langle v, w \rangle = \omega(v, Jw)$ defines a Riemannian metric. Thus $T\mathcal{M}$ is a complex vector bundle and, since the space $\mathcal{J}(\mathcal{M}, \omega)$ of all almost complex structures which are compatible with ω is connected, the first Chern class $c_1 \in \mathrm{H}^2(\mathcal{M}, \mathbb{Z})$ of $T\mathcal{M}$ is uniquely determined by ω (cf. [20], [24]). The symplectic manifold $(\mathcal{M}, \omega)$ is called **monotone** if there exists a positive constant $\lambda > 0$ such that

$$\int_{S^2} v^*\omega = \lambda \int_{S^2} v^* c_1$$

for every smooth map $v : S^2 \to \mathcal{M}$. We shall assume throughout that $(\mathcal{M}, \omega)$ is simply connected and monotone. Under this assumption there are Floer homology groups $\mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}, \phi)$ whose Euler characteristic is the Lefschetz number of ϕ. Since this is an extension of Floer's original work in [15] (to the case where $\phi \neq \mathrm{id}$), we summarize the main points of the construction.

Let $\mathbb{R} \times \mathcal{M} \to \mathbb{R} : (s, p) \mapsto H_s(p)$ be a smooth time-dependent Hamiltonian function such that $H_s = H_{s+1} \circ \phi$. The symplectomorphisms $\psi_s : \mathcal{M} \to \mathcal{M}$ generated by H are defined by

$$\frac{\mathrm{d}}{\mathrm{d}s}\psi_s = X_s \circ \psi_s, \qquad \psi_0 = \mathrm{id}, \qquad \iota(X_s)\omega = \mathrm{d}H_s.$$

They satisfy

$$\psi_{s+1} \circ \phi_H = \phi \circ \psi_s,$$

where $\phi_H := {\psi_1}^{-1} \circ \phi$. For a generic Hamiltonian H the fixed points of ϕ_H are all nondegenerate. (See [21] for the case $\phi = \mathrm{id}$. The general case is similar.)

They can be represented as the critical points of the perturbed symplectic action functional on the space of smooth paths

$$\Omega_\phi = \{\gamma : \mathbb{R} \to \mathcal{M} : \gamma(s+1) = \phi(\gamma(s))\}.$$

Since $\mathcal{M}$ is simply connected the fundamental group of Ω_ϕ is $\pi_1(\Omega_\phi) = \pi_2(\mathcal{M})$. The perturbed symplectic action functional $a_H : \Omega_\phi \to \mathbb{R}/\lambda\mathbb{Z}$ is defined as a function whose differential is given by

$$\mathrm{d}a_H(\gamma)\xi = \int_0^1 \omega(\dot{\gamma} - X_s(\gamma), \xi)\, \mathrm{d}s.$$

So the critical points of a_H are the paths of the form $x(s) = \psi_s(x_0)$ such that $x(s+1) = \phi(x(s))$. These are in one-to-one correspondence with the fixed points of ϕ_H.

Now choose a smooth map $\mathbb{R} \to \mathcal{J}(\mathcal{M}, \omega) : s \mapsto J_s$ such that $J_s = \phi^* J_{s+1}$. Such a structure determines a metric on Ω_ϕ. The gradient flow lines of a_H with respect to this metric are the solutions $u : \mathbb{R}^2 \to \mathcal{M}$ of the partial differential equation

$$\bar{\partial}_{J,H}(u) = \partial_t u + J_s(u)\left(\partial_s u - X_s(u)\right) = 0 \tag{2.1}$$

with boundary condition

$$u(s+1, t) = \phi(u(s,t)). \tag{2.2}$$

In the case where $X_s = 0$ these are Gromov's pseudoholomorphic curves (cf. [20]). If the fixed points of ϕ_H are all nondegenerate then it follows from Gromov's compactness that any solution of (2.1) and (2.2) with finite energy

$$E(u) = \tfrac{1}{2}\int_{-\infty}^{\infty}\int_0^1 \left(|\partial_s u - X_s(u)|^2 + |\partial_t u|^2\right) \mathrm{d}s\mathrm{d}t < \infty$$

has limits

$$\lim_{t\to\pm\infty} u(s,t) = \psi_s(x^\pm), \qquad x^\pm = \phi_H(x^\pm) \tag{2.3}$$

(cf. [15], [20], [24], [26]). Given any two fixed points $x^\pm$ of ϕ_H denote by $\mathcal{M}(x^-, x^+)$ the space of all solutions u of (2.1), (2.2), and (2.3). The local structure of the space $\mathcal{M}(x^-, x^+)$ can be examined by linearizing equation (2.1). This gives rise to the perturbed Cauchy-Riemann operator $\mathcal{D}_u : \mathrm{W}^{1,p}_\phi(u^*T\mathcal{M}) \to \mathrm{L}^p_\phi(u^*T\mathcal{M})$ defined by

$$\mathcal{D}_u\xi = \nabla_t\xi + J_s(u)(\nabla_s\xi - \nabla_\xi X_s(u)) + \nabla_\xi J_s(u)(\partial_s u - X_s(u)).$$

Here ∇ denotes the covariant derivative with respect to the s-dependent metric $\langle v, w\rangle_s = \omega(v, J_s w)$. Moreover, $\mathrm{L}^p_\phi(u^*T\mathcal{M})$ (respectively $\mathrm{W}^{1,p}_\phi(u^*T\mathcal{M})$) denote the completions of the space of smooth vector fields $\xi(s,t) \in T_{u(s,t)}\mathcal{M}$

along u, which satisfy $\xi(s+1,t) = d\phi(u(s,t))\xi(s,t)$ and have compact support on $S^1 \times \mathbb{R}$, with respect to the L^p-norm (respectively $\mathrm{W}^{1,p}$-norm) on $S^1 \times \mathbb{R}$. If $x^\pm$ are nondegenerate fixed points of ϕ_H and u satisfies satisfies (2.2) and (2.3) then $\mathcal{D}_u$ is a Fredholm operator and its index is given by the Maslov class of u:

$$\mathrm{index}\mathcal{D}_u = \mu(u)$$

(cf. [10] and [27]). The Maslov class $\mu(u)$ is invariant under homotopy, additive for catenations, and satisfies

$$\mu(u\#v) = \mu(u) - 2c_1(v) \tag{2.4}$$

for any sphere $v : S^2 \to \mathcal{M}$ (cf. [10] and [27]).

If $\mathcal{D}_u$ is onto then $\mathcal{M}(x^-, x^+)$ is a finite dimensional manifold near u. This follows from an implicit function theorem. For later reference we shall state here a version of that theorem. Fix a reference function $u_0 : \mathbb{R}^2 \to \mathcal{M}$ which satisfies (2.2) and $u_0(s,t) = \psi_s(x^+)$ for $t \geq 1$ and $u_0(s,t) = \psi_s(x^-)$ for $t \leq -1$.

Theorem 2.1. *Let $p > 2$ and $1/p + 1/q = 1$. Then for every constant $c_0 > 0$ there exist constants $\delta > 0$ and $c > 0$ such that the following holds. If $\xi_0 \in \mathrm{W}^{1,p}_\phi(u_0^*T\mathcal{M})$ such that*

$$\|\xi_0\|_{\mathrm{W}^{1,p}} \leq c_0, \qquad \left\|\bar\partial_{J,H}(u)\right\|_{\mathrm{L}^p} \leq \delta$$

where $u = \exp_{u_0}(\xi_0)$ and

$$\|\eta\|_{\mathrm{L}^q} \leq c_0 \left\|\mathcal{D}_u^*\eta\right\|_{\mathrm{L}^q} \tag{2.5}$$

for every $\eta \in \mathrm{W}^{1,q}(u^*T\mathcal{M})$[1] then there exists a unique section $\xi = \mathcal{D}_u^*\eta \in \mathrm{W}^{1,p}_\phi(u^*T\mathcal{M})$ such that

$$\bar\partial_{J,H}(\exp_u(\xi)) = 0, \qquad \|\xi\|_{\mathrm{W}^{1,p}} \leq c\left\|\bar\partial_{J,H}(u)\right\|_{\mathrm{L}^p}.$$

Proof. The proof is an application of the implicit function theorem for the map $\mathcal{F} : \mathrm{W}^{1,p}_\phi(u^*T\mathcal{M}) \to \mathrm{L}^p_\phi(u^*T\mathcal{M})$ defined by

$$\mathcal{F}(\xi) = \Phi_\xi(\bar\partial_{J,H}(\exp_u(\xi))) \tag{2.6}$$

where $\Phi_\xi : \mathrm{L}^p_\phi(\exp_u(\xi)^*T\mathcal{M}) \to \mathrm{L}^p_\phi(u^*T\mathcal{M})$ denotes parallel transport along the geodesic $\tau \mapsto \exp_u(\tau\xi)$. The map $\mathcal{F}$ is smooth and its derivatives are controlled by the $W^{1,p}$-norm of ξ_0. The differential at zero is given by $d\mathcal{F}(0) = \mathcal{D}_u$ and the condition (2.5) guarantees that this operator is onto and has a

[1]The formal adjoint operator $\mathcal{D}_u^*$ is obtained from $\mathcal{D}_u$ by replacing ∇_t with $-\nabla_t$.

right inverse. In fact, there is a constant $c_1 > 0$ depending only on c_0 such that

$$\|\mathcal{D}_u^*\eta\|_{\mathrm{W}^{1,p}} \le c_1 \|\mathcal{D}_u\mathcal{D}_u^*\eta\|_{\mathrm{L}^p}$$

for every $\eta \in \mathrm{W}_\phi^{1,p}(u^*T\mathcal{M})$ such that $\mathcal{D}_u^*\eta \in \mathrm{W}_\phi^{1,p}(u^*T\mathcal{M})$.[2] This estimate is proved by arguments similar to those in the proof of Lemma 4.5 below. It follows that the operator $\zeta \mapsto \mathcal{D}_u^*(\mathcal{D}_u\mathcal{D}_u^*)^{-1}\zeta$ is the required right inverse of $\mathcal{D}_u$. □

A Hamiltonian function H is called **regular** if the fixed points of ϕ_H are all nondegenerate and the operator $\mathcal{D}_u$ is onto for every $u \in \mathcal{M}(x^-, x^+)$ and any two fixed points $x^\pm$ of ϕ_H. As in Floer's papers [11] and [15] it can be proved that the set $\mathcal{H}^{\mathrm{reg}} = \mathcal{H}^{\mathrm{reg}}(J)$ of regular Hamiltonians is generic in the sense of Baire with respect to a suitable C_ε^∞-topology (see also [27] and [24]). It follows from Theorem 2.1 that for $H \in \mathcal{H}^{\mathrm{reg}}$ and $x^\pm = \phi_H(x^\pm)$ the space $\mathcal{M}(x^-, x^+)$ is a manifold whose local dimension near u is the Maslov class $\mu(u)$. By (2.4) the Maslov class determines a map $\mu : \mathrm{Fix}(\phi_H) \to \mathbb{Z}_{2N}$ (defined up to an additive constant) such that

$$\mu(u) = \mu(x^-) - \mu(x^+) \,(\mathrm{mod}\ 2N)$$

for every solution u of (2.2) and (2.3). Here the integer N is the **minimal Chern number** defined by $c_1(\pi_2(\mathcal{M})) = N\mathbb{Z}$. The additive constant can be chosen such that

$$(-1)^{\mu(x)} = \operatorname{sign}\det(\mathbb{1} - d\phi_H(x)) \tag{2.7}$$

for $x \in \mathrm{Fix}(\phi_H)$.

As in Floer's original work (cf. [15] for the case $\phi = \mathrm{id}$) the *moduli spaces* $\mathcal{M}(x^-, x^+)$ *of connecting orbits* can be used to construct a chain complex.

$$C_k = \bigoplus_{\substack{x=\phi_H(x)\\ \mu(x)=k(\mathrm{mod}\ 2N)}} \mathbb{Z}\,x.$$

The boundary operator $\partial : C_{k+1} \to C_k$ is defined by taking the sum of the numbers $\nu(u)$ over all 1-dimensional components of $\mathcal{M}(x^-, x^+)$. These numbers are defined by comparing the flow orientation of u with the coherent orientation of $\mathcal{M}(x^-, x^+)$ as in [18]. In [15] Floer proved in the case $\phi = \mathrm{id}$ that ∂ is well defined and satisfies $\partial^2 = 0$. His arguments carry over to the case $\phi \ne \mathrm{id}$. The **Floer homology** groups of ϕ are are defined as the homology of this chain complex

$$\mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}, \phi, H, J) = \ker\partial/\mathrm{im}\,\partial.$$

[2]Warning: This need not be the space $\mathrm{W}_\phi^{2,p}(u^*T\mathcal{M})$ since u is only assumed to be of class $\mathrm{W}^{1,p}$.

It can be proved as in [15] and [27] that the Floer homology groups are independent of the almost complex structures J_s and and the perurbation H used to define them. They depend on ϕ only up to Hamiltonian isotopy. In other words, there is a natural isomorphism

$$\mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}, \phi^\alpha, H^\alpha, J^\alpha) \to \mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}, \phi^\beta, H^\beta, J^\beta)$$

whenever ϕ^α and ϕ^β are related by a Hamiltonian isotopy (cf. [9]). By (2.4) the Floer homology groups are graded modulo $2N$. By (2.7) the Euler characteristic is the Lefschetz number of ϕ

$$\chi(\mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}, \phi)) = \sum_{x=\phi_H(x)} \mathrm{sign}\,\det(\mathbb{1} - d\phi_H(x)) = L(\phi).$$

Remark 2.2. A similar construction works for some classes of compact symplectic manifolds $\mathcal{M}$ which are neither monotone nor simply connected. In this case there are Floer homology groups for every component of Ω_ϕ and they are modules over a suitable Novikov ring as in [21].

Remark 2.3. If $\phi = \mathrm{id}$ then the Floer homology groups are naturally isomorphic to the homology of the underlying symplectic manifold $\mathcal{M}$:

$$\mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}, \mathrm{id}) \simeq \mathrm{H}_*(\mathcal{M}, \mathbb{Z}).$$

If $\mathcal{M}$ is not simply connected then this continues to hold for the component of contractible loops on $\mathcal{M}$ and this implies the Arnold conjecture (cf. [15]).

Remark 2.4. For every symplectomorphism ψ there is a natural isomorphism of Floer homologies $\mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}, \phi) = \mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}, \psi^{-1} \circ \phi \circ \psi)$. (To see this consider the function $v(s,t) = \psi^{-1}(u(s,t))$ where $u(s,t)$ is a solution of (2.1) and (2.2).) Donaldson has suggested the construction of a homomorphism

$$\mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}, \psi) \otimes \mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}, \phi) \to \mathrm{HF}_*^{\mathrm{symp}}(\mathcal{M}, \psi \circ \phi)$$

using moduli spaces of J-holomorphic curves with three cylindrical ends (the *pair-of-pants construction*). If $\psi = \mathrm{id}$ then this determines an action of the homology of $\mathcal{M}$ on the Floer homology groups of ϕ. If $\phi = \psi = \mathrm{id}$ then this should agree with the deformed cup-product of Witten.

We close this section with an existence theorem for solutions of (2.1) which is based on Theorem 2.1.

Theorem 2.5. *Assume $H \in \mathcal{H}^{\mathrm{reg}}$. Let $x^\pm$ be fixed points of ϕ_H and let A denote a homotopy class of maps $u : \mathbb{R}^2 \to \mathcal{M}$ which satisfy (2.2) and (2.3) with $\mu(u) = 1$. Then for every $c_0 > 0$ and $p > 2$ there exist constants $\delta > 0$ and $c > 0$ such that the following holds. If $u : \mathbb{R}^2 \to \mathcal{M}$ satisfies (2.2)* *and (2.3) and represents the class A such that*

$$|\partial_t u(s,t)| \le \frac{c_0}{1+t^2}, \qquad \left\|\bar\partial_{J,H}(u)\right\|_{\mathrm{L}^p} \le \delta,$$

then there exists a unique section $\xi = \mathcal{D}_u^*\eta \in \mathrm{W}^{1,p}_\phi(u^*T\mathcal{M})$ such that

$$\bar\partial_{J,H}(\exp_u(\xi)) = 0, \qquad \|\xi\|_{\mathrm{W}^{1,p}} \le c\left\|\bar\partial_{J,H}(u)\right\|_{\mathrm{L}^p}.$$

Remark 2.6. The function $f(t) = c_0/(1+t^2)$ has been chosen because it is integrable, p-integrable, and is the derivative of functions $F^\pm$ which are p-integrable on $\mathbb{R}^\pm$, respectively. Any function with these properties will do.

LEMMA 2.7. *Assume $H \in \mathcal{H}^{\mathrm{reg}}$ and let $u_0 \in \mathcal{M}(x^-,x^+)$ and $p > 2$. Then for every constant $c_0 > 0$ there exists a constant $c > 0$ such that if $\xi_0 \in \mathrm{W}^{1,p}_\phi(u^*T\mathcal{M})$ with $\|\xi_0\|_{\mathrm{W}^{1,p}} \le c_0$ then*

$$\|\xi_0\|_{\mathrm{W}^{1,p}} \le c\left(\left\|\bar\partial_{J,H}(\exp_{u_0}(\xi_0))\right\|_{\mathrm{L}^p} + \|\xi_0\|_{\mathrm{L}^p} + \|\xi_0\|_{\mathrm{L}^\infty}\right).$$

Proof. Consider the map $\mathcal{F} : \mathrm{W}^{1,p}_\phi(u_0^*T\mathcal{M}) \to \mathrm{L}^p_\phi(u_0^*T\mathcal{M})$ defined by (2.6) with u replaced by u_0. Since $\mathcal{F}(0) = 0$ and $\mathrm{d}\mathcal{F}(0) = \mathcal{D}_{u_0}$ there is a quadratic estimate

$$\|\mathcal{F}(\xi_0) - \mathcal{D}_{u_0}\xi_0\|_{\mathrm{L}^p} \le c_1\|\xi_0\|_{\mathrm{W}^{1,p}}\|\xi_0\|_{\mathrm{L}^\infty}.$$

Here the constant c_1 depends only on u_0 and c_0. Now the statement follows from the elliptic estimate $\|\xi_0\|_{\mathrm{W}^{1,p}} \le c_2\left(\|\mathcal{D}_{u_0}\xi_0\|_{\mathrm{L}^p} + \|\xi_0\|_{\mathrm{L}^p}\right)$. □

Proof of Theorem 2.5. If the assertion of the theorem were false then there would exist a sequence u_ν representing the class A such that

$$|\partial_t u_\nu(s,t)| \le \frac{c_0}{1+t^2}, \qquad \lim_{\nu\to\infty}\left\|\bar\partial_{J,H}(u_\nu)\right\|_{\mathrm{L}^p} = 0,$$

but the conclusion of the theorem is not satisfied with $u = u_\nu$ and $c = \nu$. Now the derivatives of u_ν are uniformly bounded in the L^∞-norm and $E(u_\nu) \le c_1$. Hence, passing to a subsequence, we may assume without loss of generality that u_ν converges uniformly on compact sets to a solution u_0 of (2.1) with finite energy. Since $\mu(u_\nu) = 1$ it follows as in the usual compactness argument for Morse-Smale gradient flows that $u_0 \in \mathcal{M}(x^-,x^+)$ (cf. [26]). By the uniform decay estimate u_ν converges uniformly on $\mathbb{R}^2$. Moreover, $u_\nu = \exp_{u_0}(\xi_\nu)$ where

$$\sup_\nu \|\xi_\nu\|_{\mathrm{W}^{1,p}} < \infty, \qquad \|\xi_\nu\|_{\mathrm{L}^p} \to 0, \qquad \|\xi_\nu\|_{\mathrm{L}^\infty} \to 0.$$

Hence, by Lemma 2.7, ξ_ν converges to zero in the $\mathrm{W}^{1,p}$-norm. This implies that the operators $\mathcal{D}_{u_\nu}$ satisfy (2.5) for ν sufficiently large with a uniform constant c. Hence it follows from Theorem 2.1 that for ν sufficiently large the functions u_ν do satisfy the conclusion of the theorem with a uniform constant c. This contradicts our assumption. Hence the theorem is proved. □

3. Flat connections over a Riemann surface

Let $\pi : P \to \Sigma$ be a nontrivial SO(3)-bundle over a compact oriented Riemann surface of genus $k \geq 2$. Denote by $\mathcal{G}_0(P)$ the component of $\mathbb{1}$ in the space of gauge transformations and by

$$\mathcal{M}(P) = \mathcal{A}_{\text{flat}}(P)/\mathcal{G}_0(P)$$

the moduli space of flat connections. This space is a compact manifold of dimension $6k - 6$. Its tangent space at an equivalence class $[A]$ of a connection $A \in \mathcal{A}_{\text{flat}}(P)$ can be identified with the twisted deRham cohomology $T_{[A]}\mathcal{M}(P) = \mathrm{H}^1_A = \ker \mathrm{d}_A/\mathrm{im}\,\mathrm{d}_A$. Here $\mathrm{d}_A : \Omega^k(\mathfrak{g}_P) \to \Omega^{k+1}(\mathfrak{g}_P)$ denotes the covariant derivative defined by $\mathrm{d}_A\alpha = \mathrm{d}\alpha + [A \wedge \alpha]$. Given a conformal structure on Σ we may identify H^1_A with the space $\ker \mathrm{d}_A \cap \ker \mathrm{d}_A{}^*$ of harmonic forms. Here $\mathrm{d}_A{}^* = - * \mathrm{d}_A *$ denotes the L^2-adjoint of d_A. The moduli space $\mathcal{M}(P)$ carries a natural symplectic structure

$$\omega(\alpha, \beta) = \int_\Sigma \langle \alpha \wedge \beta \rangle$$

for $\alpha, \beta \in \mathrm{H}^1_A$. Every conformal structure on Σ determines a complex structure on $\mathcal{M}(P)$ (cf. [2]). The corresponding almost complex structure on $T\mathcal{M}(P)$ is given by the the Hodge-$*$-operators on the spaces H^1_A of harmonic forms and is compatible with ω.

The second homotopy group of $\mathcal{M}(P)$ is given by

$$\pi_2(\mathcal{M}(P)) = \pi_1(\mathcal{G}_0(P)) = \mathbb{Z}.$$

More precisely, a sphere in $\mathcal{M}(P)$ can be represented by a smooth map $A : D \to \mathcal{A}_{\text{flat}}(P)$ such that $A(e^{2\pi i\theta}) = g(\theta)^* A_0$ where $D = \{z \in \mathbb{C} : |z| \leq 1\}$ is the unit disc, $A_0 \in \mathcal{A}_{\text{flat}}(P)$ is a flat connection on P, and $g(\theta) = g(\theta + 1) \in \mathcal{G}_0(P)$ is a loop of gauge transformations. Any such loop has a degree $\deg(g) \in \mathbb{Z}$ and the integrals of ω and $c_1 = c_1(T\mathcal{M}(P))$ over A are given by

$$\langle c_1, A \rangle = \frac{1}{4\pi^2}\langle [\omega], A \rangle = 2\deg(g)$$

(cf. [10]). In particular, $\mathcal{M}(P)$ is monotone in the sense of section 2 with $\lambda = 4\pi^2$.

Every orientation preserving automorphism $f : P \to P$ determines a symplectomorphism $\phi_f : \mathcal{M}(P) \to \mathcal{M}(P)$ defined by $[A] \mapsto [f^*A]$. If f_0 and f_1 are isotopic then $\phi_{f_0} = \phi_{f_1}$ and hence the correspondence $f \mapsto \phi_f$ determines a symplectic action $\mathrm{Aut}^+(P)/\mathrm{Aut}_0(P) \to \mathrm{Diff}(\mathcal{M}(P), \omega)$ where

$\mathrm{Aut}_0(P)$ denoptes the component of the identity in the group $\mathrm{Aut}^+(P)$ of orientation preserving automorphisms. Note that the quotient $\mathrm{Aut}^+(P)/\mathrm{Aut}_0(P)$ is a finite extension of the mapping class group $\mathrm{Diff}^+(\Sigma)/\mathrm{Diff}_0(\Sigma)$ by $\pi_0(\mathcal{G}(P)) \cong \mathbb{Z}^{2k}$.

The automorphism f also determines a principal bundle $P_f \to \Sigma_h$ where $h : \Sigma \to \Sigma$ is the diffeomorphism induced by f and P_f and Σ_h denote the mapping cylinders. A connection $a \in \mathcal{A}(P_f)$ is a 1-form $a = A + \Phi\, \mathrm{d}s$ where $A(s) \in \mathcal{A}(P)$, $\Phi(s) \in \Omega^0(\mathfrak{g}_P)$ and

$$A(s+1) = f^*A(s), \qquad \Phi(s+1) = \Phi(s) \circ f.$$

The group $\mathcal{G}(P_f)$ of gauge transformations of P_f consists of smooth maps $g : \mathbb{R} \to \mathcal{G}(P)$ such that $g(s+1) = g(s) \circ f$. It acts on $\mathcal{A}(P_f)$ by

$$g^*a = g^*A + \left(g^{-1}\dot{g} + g^{-1}\Phi g\right) \mathrm{d}s.$$

Here the notation g^* is used ambiguously: g^*a denotes the action of $g \in \mathcal{G}(P_f)$ on $a \in \mathcal{A}(P_f)$ whereas g^*A denotes the pointwise action of $g(s) \in \mathcal{G}(P)$ on $A(s) \in \mathcal{A}(P)$.

The space of paths

Fix a 1-parameter family of conformal structures on Σ such that the associated Hodge-$*$-operators satisfy $*_{s+1} \circ f^* = f^* \circ *_s$. Denote by $\mathcal{A}_\Sigma(P_f)$ the subspace of those connections $a = A(s)+\Phi(s)\,\mathrm{d}s$ which satisfy $F_A = 0$ and $d_A *_s (\dot{A} - \mathrm{d}_A\Phi) = 0$. For any such connection and every s the section $\Phi(s) \in \Omega^0(\mathfrak{g}_P)$ is uniquely determined by A since $\mathrm{d}_A *_s \mathrm{d}_A\Phi = \mathrm{d}_A *_s \dot{A}$. The 1-form $\dot{A} - \mathrm{d}_A\Phi$ is the projection of $\dot{A}$ onto the space of harmonic forms with respect to the s-metric. Also denote by $\mathcal{G}_\Sigma(P_f)$ the subgroup of those $g \in \mathcal{G}(P_f)$ such that $g(s) \in \mathcal{G}_0(P)$ for all s. It follows from results in [2] and [5] that the space $\mathcal{A}_{\mathrm{flat}}(P)$ is simply connected and hence $\mathcal{A}_\Sigma(P_f)$ is connected. There is a natural bijection $\Omega_{\phi_f} = \mathcal{A}_\Sigma(P_f)/\mathcal{G}_\Sigma(P_f)$ and the quotient $\mathcal{A}_\Sigma(P_f)/\mathcal{G}_0(P_f)$ is the universal cover of Ω_{ϕ_f}. The second homotopy group of $\mathcal{M}(P)$ is the fundamental group of Ω_{ϕ_f} and can be identified with the second homotopy group of $\mathcal{M}(P)$ (cf. [10]):

$$\pi_1(\Omega_{\phi_f}) = \pi_2(\mathcal{M}(P)) = \mathcal{G}_\Sigma(P_f)/\mathcal{G}_0(P_f) = \mathbb{Z}.$$

The Chern-Simons functional

The Chern-Simons functional $\mathcal{CS} : \mathcal{A}(P_f) \to \mathbb{R}$ is given by

$$\mathcal{CS}(a) = \int_0^1 \int_\Sigma \left(\tfrac{1}{2}\langle \dot{A} \wedge (A - A_0)\rangle + \langle F_A \wedge \Phi\rangle\right) \mathrm{d}s$$

for $a = A + \Phi\, \mathrm{d}s$ where $A_0 = f^*A_0$ is a fixed flat connection. The restriction of the Chern-Simons functional to $\mathcal{A}_\Sigma(P_f)$ agrees with the symplectic action on Ω_{ϕ_f}. The critical points are the flat connections on P_f and they satisfy

$$F_A = 0, \qquad \dot{A} - \mathrm{d}_A \Phi = 0.$$

Gauge equivalence classes of such connections are in one-to-one correspondence with the fixed points of ϕ_f. If ϕ_f has degenerate fixed points choose a gauge invariant holonomy perturbation $H_s : \mathcal{A}(P) \to \mathbb{R}$ as in [10] such that

$$H_{s+1}(f^* A) = H_s(A).$$

Think of H_s as a Hamiltonian function with corresponding Hamiltonian vector field $X_s : \mathcal{A}(P) \to \Omega^1(\mathfrak{g}_P)$ defined by

$$\mathrm{d}H_s(A)\alpha = \int_\Sigma \langle X_s(A) \wedge \alpha \rangle.$$

Since $H_s : \mathcal{A}(P) \to \mathbb{R}$ is invariant under $\mathcal{G}_0(P)$ the vector field X_s satisfies

$$X_s(g^* A) = g^{-1} X_s(A) g, \qquad \mathrm{d}_A X_s(A) = 0 \tag{3.1}$$

for $g \in \mathcal{G}_0(P)$ and $A \in \mathcal{A}(P)$. The corresponding Hamiltonian symplectomorphisms $\psi_s : \mathcal{A}(P) \to \mathcal{A}(P)$ are equivariant under the action of $\mathcal{G}_0(P)$ and the curvature is constant along the flow. Moreover, $\psi_{s+1} \circ \phi_{f,H} = \phi_f \circ \psi_s$ where $\phi_{f,H} = \psi_1^{-1} \circ \phi_f$.

The perturbed Chern-Simons functional $\mathcal{CS}_H : \mathcal{A}(P_f) \to \mathbb{R}$ is defined by

$$\mathcal{CS}_H(A + \Phi\, \mathrm{d}s) = \mathcal{CS}(A + \Phi\, \mathrm{d}s) - \int_0^1 H_s(A(s))\, \mathrm{d}s.$$

A connection $A + \Phi\, \mathrm{d}s$ on P_f is a critical point of $\mathcal{CS}_H$ iff

$$F_A = 0, \qquad \dot{A} - \mathrm{d}_A \Phi - X_s(A) = 0.$$

Denote the space of such critical points by $\mathcal{A}_{\mathrm{flat}}(P_f, H)$.

Remark 3.1. There is a bijection $\mathcal{A}_{\mathrm{flat}}(P_f, H)/\mathcal{G}_\Sigma(P_f) \simeq \mathrm{Fix}(\phi_{f,H})$. Moreover, an H-flat connection $A + \Phi\, \mathrm{d}s$ is nondegenerate as a critical point of $\mathcal{CS}_H$ if and only if $A(0)$ represents a nondegenerate fixed point of $\phi_{f,H}$ (cf. [10]). The perturbation H can be chosen such that the symplectomorphism $\phi_{f,H} : \mathcal{M}(P) \to \mathcal{M}(P)$ has only nondegenerate fixed points (cf. [21]).

Instantons and holomorphic curves

Fix two nondegenerate H-flat connections $a^\pm = A^\pm + \Phi^\pm\, ds \in \mathcal{A}_{\mathrm{flat}}(P_f, H)$ and choose smooth functions $A : \mathbb{R}^2 \to \mathcal{A}_{\mathrm{flat}}(P)$ and $\Phi, \Psi : \mathbb{R}^2 \to \Omega^0(\mathfrak{g}_P)$ which satisfy

$$\begin{gathered} A(s+1,t) = f^* A(s,t), \\ \Phi(s+1,t) = \Phi(s,t) \circ f, \quad \Psi(s+1,t) = \Phi(s,t) \circ f. \end{gathered} \tag{3.2}$$

$$\lim_{t\to\pm\infty} A(s,t) = A^{\pm}(s), \quad \lim_{t\to\pm\infty} \Phi(s,t) = \Phi^{\pm}(s), \quad \lim_{t\to\pm\infty} \Psi(s,t) = 0. \tag{3.3}$$

Now choose a smooth family of conformal structures on Σ depending on a real parameter s such that $*_{s+1} \circ f^* = f^* \circ *_s$. Then the perturbed Cauchy-Riemann equations (2.1) take the form

$$\partial_t A - \mathrm{d}_A \Psi + *_s \left(\partial_s A - X_s(A) - \mathrm{d}_A \Phi\right) = 0. \tag{3.4}$$

For solutions of (3.4) the functions Φ and Ψ are uniquely determined by A. In other words the function $A : \mathbb{R}^2 \to \mathcal{A}_{\mathrm{flat}}(s,t)$ is an anti-holomorphic curve if and only if the harmonic part of $\partial_t A + *_s(\partial_s A - X_s(A))$ vanishes. If this is the case then there exist unique functions $\Phi, \Psi : \mathbb{R}^2 \to \Omega^0(\mathfrak{g}_P)$ such that (3.4) is satisfied.

Now think of $\Xi = A + \Phi\,\mathrm{d}s + \Psi\,\mathrm{d}t$ as a connection on $P_f \times \mathbb{R}$. Here the connections $A(s,t)$ are no longer required to be flat. The perturbed self-duality equations on $P_f \times \mathbb{R}$ take the form

$$\begin{array}{rcl} \partial_t A - \mathrm{d}_A \Psi + *_s(\partial_s A - X_s(A) - \mathrm{d}_A \Phi) &=& 0, \\ \partial_t \Phi - \partial_s \Psi - [\Phi, \Psi] + \varepsilon^{-2} *_s F_A &=& 0. \end{array} \tag{3.5}$$

The first equation in (3.5) agrees with (3.4) while the second equation replaces the condition on $A(s,t)$ to be flat. The factor $1/\varepsilon^2$ arises from conformally rescaling the metric on Σ by the factor ε^2. The Hodge-$*$-operator on 1-forms (the middle dimension) is invariant under conformal rescaling while the Hodge-$*$-operator on 2-forms rescales by $1/\varepsilon^2$. Alternatively, equation (3.5) can be obtained by considering a solution $\widetilde{\Xi} = \widetilde{A} + \widetilde{\Phi}\mathrm{d}s + \widetilde{\Psi}\mathrm{d}t$ of the self-duality equation with $\varepsilon = 1$ and $*_s$ replaced by $*_{\varepsilon s}$ on the domain $0 \le s \le 1/\varepsilon$. Then $A(s,t) = \widetilde{A}(s/\varepsilon, t/\varepsilon)$, $\Phi(s,t) = \widetilde{\Phi}(s/\varepsilon, t/\varepsilon)/\varepsilon$, and $\Psi(s,t) = \widetilde{\Psi}(s/\varepsilon, t/\varepsilon)/\varepsilon$ satisfy (3.5). This is a modification of Atiyah's idea to *stretch the neck* for Heegard splittings (cf. [1]).

If Ξ satisfies (3.5) then the perturbed Yang-Mills action of Ξ with respect to the rescaled metric on Σ is given by

$$\mathcal{Y}_H^{\varepsilon}(\Xi) = \int_{-\infty}^{\infty} \int_0^1 \left(\|\partial_t A - \mathrm{d}_A \Psi\|^2_{L^2(\Sigma, *_s)} + \varepsilon^{-2} \|F_A\|^2_{L^2(\Sigma, *_s)} \right) \mathrm{d}s\mathrm{d}t. \tag{3.6}$$

If Ξ satisfies (3.2), (3.3), and (3.5) then

$$\mathcal{Y}_H^{\varepsilon}(\Xi) = \mathcal{CS}_H(a^-) - \mathcal{CS}_H(a^+).$$

If instead Ξ satisfies (3.2), (3.3), and (3.4) with $F_A = 0$ then the the right hand side of (3.6) is not the Yang-Mills action but the energy of the anti-holomorphic curve represented by A.

Index theorem

The Sobolev space $W^{k,p}_f(\mathbb{R}^2 \times T^*\Sigma \otimes \mathfrak{g}_P)$ is the completion of the space of compactly supported smooth maps $\alpha : \mathbb{R}^2 \to \Omega^1(\mathfrak{g}_P)$ which satisfy $\alpha(s+1,t) = f^*\alpha(s,t)$ with respect to the $W^{k,p}$-norm on $\mathbb{R}^2 \times \Sigma_h$. The space $W^{k,p}_f(\mathbb{R}^2 \times \mathfrak{g}_P)$ is defined similarly. For any smooth map $A : \mathbb{R}^2 \to \mathcal{A}_{\mathrm{flat}}(P)$ which satisfies (3.2) the closed linear subspace $W^{k,p}_f(H_A) \subset W^{k,p}_f(\mathbb{R}^2 \times T^*\Sigma \otimes \mathfrak{g}_P)$ consists of those α_0 such that $\alpha_0(s,t)$ is a harmonic 1-forms on $\mathfrak{g}_P$ with respect to the s-metric and the connection $A(s,t)$.

Linearizing (3.4) gives rise to the Fredholm operator

$$\mathcal{D}_0 = \mathcal{D}_0(\Xi) : W^{1,p}_f(H_A) \to L^p_f(H_A)$$

defined by

$$\mathcal{D}_0\alpha_0 = \pi_A\left(\nabla_t\alpha_0 + *_s\nabla_s\alpha_0 - *_s dX_s(A)\alpha_0\right).$$

Here $\nabla_s = \partial_s + \Phi$, $\nabla_t = \partial_t + \Psi$ and $\pi_{A(s,t)}(\alpha)$ denotes the harmonic part of the 1-form α with respect to the connection $A(s,t)$ and the s-metric on Σ. Note that H_A is the pullback tangent bundle of $\mathcal{M}(P)$ under the map $[A] : \mathbb{R}^2 \to \mathcal{M}(P)$. In view of the periodicity condition (3.2) this is a complex vector bundle over $S^1 \times \mathbb{R}$ and $\mathcal{D}_0$ is a perturbed Cauchy-Riemann operator on this bundle. (See [10] for more details.)

Abbreviate

$$\xi = (\alpha, \phi, \psi) \in W^{k,p}_f = W^{k,p}_f(\mathbb{R}^2 \times T^*\Sigma \otimes \mathfrak{g}_P \oplus \mathfrak{g}_P \oplus \mathfrak{g}_P).$$

Linearizing (3.5) gives rise to the Fredholm operator

$$\mathcal{D}_\varepsilon = \mathcal{D}_\varepsilon(\Xi) : W^{1,p}_f \to L^p_f$$

defined by

$$\mathcal{D}_\varepsilon = \nabla_t + \begin{pmatrix} *_s\nabla_s & 0 & 0 \\ 0 & 0 & -\nabla_s \\ 0 & *_s\nabla_s *_s & 0 \end{pmatrix} - \begin{pmatrix} *_s dX_s(A) & *_s d_A & d_A \\ -\varepsilon^{-2} *_s d_A & 0 & 0 \\ -\varepsilon^{-2} *_s d_A *_s & 0 & 0 \end{pmatrix}.$$

The notation $\mathcal{D}_\varepsilon(\Xi)$ indicates the dependence of the operator $\mathcal{D}_\varepsilon$ on the connection $\Xi = A + \Phi ds + \Psi dt$. The following theorem was proved in [10].

THEOREM 3.2. *For any pair $a^\pm$ of nondegenerate H-flat connections on P_f and any connection $\Xi = A + \Phi\, ds + \Psi\, dt$ on $P_f \times \mathbb{R}$ which satisfies (3.2), (3.3), and $F_A = 0$ we have*

$$\mathrm{index}\,\mathcal{D}_0 = \mathrm{index}\mathcal{D}_\varepsilon = \mu_H(a^-, a^+).$$

The index formula of section 1 shows that for any connection $\Xi = A + \Phi\, ds + \Psi\, dt$ which satisfies (3.2) and (3.3)

$$\begin{aligned}\mu_H(a^-, a^+) &= \frac{1}{2}\eta(a^+) - \frac{1}{2}\eta(a^-)\\ &\quad - \frac{1}{2\pi^2}\int_{-\infty}^{\infty}\int_0^1\int_\Sigma \langle F_A \wedge (\partial_t\Phi - \partial_s\Psi - [\Phi,\Psi])\rangle\, \mathrm{d}s\mathrm{d}t\\ &\quad - \frac{1}{2\pi^2}\int_{-\infty}^{\infty}\int_0^1\int_\Sigma \langle(\partial_s A - X_s(A) - \mathrm{d}_A\Phi)\wedge(\partial_t A - \mathrm{d}_A\Psi)\rangle\, \mathrm{d}s\mathrm{d}t\end{aligned}$$

where $\eta(a) = \eta(\mathrm{D}_a)$ is the eta-invariant of the extended Hessian D_a as in section 1. In particular, if Ξ is a solution of (3.2), (3.3), and (3.4) with $F_A = 0$ then

$$\mu_H(a^-, a^+) = \frac{1}{2}\eta(a^+) - \frac{1}{2}\eta(a^-) + \frac{1}{2\pi^2}\int_{-\infty}^{\infty}\int_0^1 \|\partial_t A - \mathrm{d}_A\Psi\|^2_{L^2(\Sigma)}\, \mathrm{d}s\mathrm{d}t. \tag{3.7}$$

Moduli spaces

Given $a^\pm = A^\pm + \Phi^\pm\,\mathrm{d}s \in \mathcal{A}_\Sigma(P_f, H)$ choose a smooth connection $\Xi_0 = A_0 + \Phi_0\,\mathrm{d}s + \Psi_0\,\mathrm{d}t$ which satisfies (3.2) and (3.3) and is locally independent of t for $|t| \geq 1$. For $p > 2$ denote $\mathcal{A}^{1,p}(a^-, a^+) = \{\Xi_0 + \xi \,:\, \xi \in \mathrm{W}^{1,p}_f\}$. The associated space of gauge transformations $\mathcal{G}^{2,p}$ is the completion of the space of all smooth gauge transformations $g : P_f \times \mathbb{R} \to G$ such that $g(s,t) = \mathbb{1}$ for $|t|$ sufficiently large with respect to the $\mathrm{W}^{1,p}_f$-norm of $\xi = g^{-1}\mathrm{d}g + g^{-1}\partial_s g\,\mathrm{d}s + g^{-1}\partial_t g\,\mathrm{d}t$. This space acts on $\mathcal{A}^{1,p}(a^-, a^+)$ via

$$g^*\Xi = g^*A + \left(g^{-1}\partial_s g + g^{-1}\Phi g\right)\mathrm{d}s + \left(g^{-1}\partial_t g + g^{-1}\Psi g\right)\mathrm{d}t.$$

Consider the space

$$\mathcal{A}^{1,p}_0(a^-, a^+, H) = \left\{\Xi \in \mathcal{A}^{1,p}(a^-, a^+) \,:\, F_A = 0,\ (3.4)\right\}$$

of holomorphic curves in $\mathcal{A}_{\mathrm{flat}}(P)$ connecting a^- to a^+ and the space

$$\mathcal{A}^{1,p}_\varepsilon(a^-, a^+, H) = \left\{a \in \mathcal{A}^{1,p}(a^-, a^+) \,:\, (3.5)\right\}$$

of self-dual instantons on $P_f \times \mathbb{R}$ connecting a^- to a^+. The corresponding moduli spaces

$$\mathcal{M}_0(a^-, a^+, H) = \mathcal{A}^{1,p}_0(a^-, a^+, H)/\mathcal{G}^{2,p}$$
$$\mathcal{M}_\varepsilon(a^-, a^+, H) = \mathcal{A}^{1,p}_\varepsilon(a^-, a^+, H)/\mathcal{G}^{2,p}$$

are finite dimensional manifolds of the same dimension $\mu_H(a^-, a^+)$ provided that the operators $\mathcal{D}_0$ and $\mathcal{D}_\varepsilon$ are onto for all relevant connections Ξ. Thus denote by $\mathcal{H}^{\mathrm{reg}}_0$ the set of all perturbations H such that the fixed points of $\phi_{f,H}$

are nondegenerate and the operator $\mathcal{D}_0$ is onto for all $a^\pm \in \mathcal{A}_\Sigma(P_f, H)$ and all $\Xi \in \mathcal{A}_0^{1,p}(a^-, a^+, H)$. Likewise, denote by $\mathcal{H}_\varepsilon^{\mathrm{reg}}$ the set of all perturbations H such that the H-flat connections on P_f are nondegenerate and the operator $\mathcal{D}_\varepsilon$ is onto for all $a^\pm \in \mathcal{A}_\Sigma(P_f, H)$ and all $\Xi \in \mathcal{A}_\varepsilon^{1,p}(a^-, a^+, H)$. Both sets $\mathcal{H}_0^{\mathrm{reg}}$ and $\mathcal{H}_\varepsilon^{\mathrm{reg}}$ are of the second category in the sense of Baire (countable intersections of open and dense sets) in a suitable Banach space of smooth perturbations. (See [15], [24], [27] for the symplectic case and [8], [14], [16] for the instanton case.)

4. Elliptic estimates

Fix two nondegenerate H-flat connections $a^\pm = A^\pm + \Phi^\pm \, \mathrm{d}s \in \mathcal{A}_\Sigma(P_f, H)$, let $p > 4$, and assume throughout that $\Xi = A + \Phi \, \mathrm{d}s + \Psi \, \mathrm{d}t \in \mathcal{A}^{1,p}(a^-, a^+)$ with $F_{A(s,t)} = 0$ for all s and t. Write $\mathcal{D}_0 = \mathcal{D}_0(\Xi)$ and $\mathcal{D}_\varepsilon = \mathcal{D}_\varepsilon(\Xi)$.

It is convenient to use the ε-dependent Banach space norms

$$\|\xi\|_{0,p,\varepsilon}^p = \int_{-\infty}^{\infty} \int_0^1 \left(\|\alpha\|_{\mathrm{L}^p(\Sigma)}^p + \varepsilon^p \|\phi\|_{\mathrm{L}^p(\Sigma)}^p + \varepsilon^p \|\psi\|_{\mathrm{L}^p(\Sigma)}^p \right) \mathrm{d}s\mathrm{d}t$$

on L_f^p,

$$\|\xi\|_{\infty,\varepsilon} = \|\alpha\|_{\mathrm{L}^\infty} + \varepsilon \|\phi\|_{\mathrm{L}^\infty} + \varepsilon \|\psi\|_{\mathrm{L}^\infty}$$

on L_f^∞, and

$$\begin{aligned} \|\xi\|_{1,p,\varepsilon}^p \quad &= \int_{-\infty}^{\infty} \int_0^1 \left(\|\alpha\|_{\mathrm{W}^{1,p}(\Sigma)}^p + \varepsilon^p \|\nabla_s \alpha\|_{\mathrm{L}^p(\Sigma)}^p + \varepsilon^p \|\nabla_t \alpha\|_{\mathrm{L}^p(\Sigma)}^p \right) \mathrm{d}s\mathrm{d}t \\ &+ \int_{-\infty}^{\infty} \int_0^1 \left(\varepsilon^p \|\phi\|_{\mathrm{W}^{1,p}(\Sigma)}^p + \varepsilon^{2p} \|\nabla_s \phi\|_{\mathrm{L}^p(\Sigma)}^p + \varepsilon^{2p} \|\nabla_t \phi\|_{\mathrm{L}^p(\Sigma)}^p \right) \mathrm{d}s\mathrm{d}t \\ &+ \int_{-\infty}^{\infty} \int_0^1 \left(\varepsilon^p \|\psi\|_{\mathrm{W}^{1,p}(\Sigma)}^p + \varepsilon^{2p} \|\partial_s \psi\|_{\mathrm{L}^p(\Sigma)}^p + \varepsilon^{2p} \|\partial_t \psi\|_{\mathrm{L}^p(\Sigma)}^p \right) \mathrm{d}s\mathrm{d}t \end{aligned}$$

on $\mathrm{W}_f^{1,p}$. Here $\nabla_s = \partial_s + \Phi$, $\nabla_t = \partial_t + \Psi$, $\|\phi\|_{\mathrm{W}^{1,p}(\Sigma)}^p = \|\phi\|_{\mathrm{L}^p(\Sigma)}^p + \|\mathrm{d}_A \phi\|_{\mathrm{L}^p(\Sigma)}^p$, and

$$\|\alpha\|_{\mathrm{W}^{1,p}(\Sigma)}^p = \|\alpha\|_{\mathrm{L}^p(\Sigma)}^p + \|\mathrm{d}_A \alpha\|_{\mathrm{L}^p(\Sigma)}^p + \|\mathrm{d}_A *_s \alpha\|_{\mathrm{L}^p(\Sigma)}^p .$$

Thus the $1, p, \varepsilon$-norm depends on the connection Ξ and the $1, p, \varepsilon$-norm of $g^{-1}\xi g$ with respect to $g^*\Xi$ agrees with the $1, p, \varepsilon$-norm of ξ with respect to Ξ.

Lemma 4.1. *For $p > 4$ there exists a constant $c = c(p) > 0$ such that*

$$\|\xi\|_{\infty,\varepsilon} \le c\varepsilon^{-2/p} \|\xi\|_{1,p,\varepsilon}$$

for $\xi \in \mathrm{W}_f^{1,p}$ and $0 < \varepsilon \le 1$.

Proof. Denote $\tilde{\xi} = \tilde{\alpha} + \tilde{\phi}\, ds + \tilde{\psi}\, dt$ where $\tilde{\alpha}(s,t) = \alpha(\varepsilon s, \varepsilon t)$, $\tilde{\phi}(s,t) = \varepsilon\phi(\varepsilon s, \varepsilon t)$, and $\tilde{\psi}(s,t) = \varepsilon\psi(\varepsilon s, \varepsilon t)$ for $0 \le s \le \varepsilon^{-1}$ and $t \in \mathbb{R}$. Then $\|\xi\|_{1,p,\varepsilon} = \varepsilon^{2/p}\|\tilde{\xi}\|_{W^{1,p}}$. Hence the statement follows from the usual Sobolev estimate for $\tilde{\xi}$. □

Lemma 4.2. *There exist constants $\varepsilon_0 > 0$ and $c > 0$ such that*

$$\|\xi\|_{1,p,\varepsilon} \le c\left(\varepsilon\,\|\mathcal{D}_\varepsilon\xi\|_{0,p,\varepsilon} + \|\pi_A(\xi)\|_{\mathrm{L}^p}\right),$$

$$\|\xi - \pi_A(\xi)\|_{1,p,\varepsilon} \le c\varepsilon\left(\|\mathcal{D}_\varepsilon\xi\|_{0,p,\varepsilon} + \|\pi_A(\xi)\|_{\mathrm{L}^p}\right)$$

for $\xi \in \mathrm{W}^{1,p}_f$ and $0 < \varepsilon \le \varepsilon_0$. The constants ε_0 and c depend continuously on Ξ (with respect to the C^∞-topology) and they are independent of ε.

Proof. We prove the statement only for $p = 2$. Moreover, it suffices to consider the case $X_s = 0$. Throughout all the norms are L^2-norms on $\Sigma_h \times \mathbb{R}$. Let $\xi = (\alpha, \phi, \psi) \in W^{1,2}_f$ and denote $\tilde{\xi} = (\tilde{\alpha}, \tilde{\phi}, \tilde{\psi}) = \mathcal{D}_\varepsilon\xi$. We shall prove the estimate

$$\begin{aligned}
&\|\mathrm{d}_A\alpha\|^2 + \|\mathrm{d}_A *_s \alpha\|^2 + \varepsilon^2\,\|\nabla_s\alpha\|^2 + \varepsilon^2\,\|\nabla_t\alpha\|^2 \\
&\quad + \varepsilon^2\,\|\mathrm{d}_A\phi\|^2 + \varepsilon^4\,\|\nabla_s\phi\|^2 + \varepsilon^4\,\|\nabla_t\phi\|^2 \\
&\quad + \varepsilon^2\,\|\mathrm{d}_A\psi\|^2 + \varepsilon^4\,\|\nabla_s\psi\|^2 + \varepsilon^4\,\|\nabla_t\psi\|^2 \\
&\quad \le c\varepsilon^2\left(\|\tilde{\xi}\|^2_{0,2,\varepsilon} + \|\alpha\|^2\right).
\end{aligned} \tag{4.1}$$

To see this assume that ξ is smooth (C^∞) and consider the identity

$$\mathcal{D}^*_\varepsilon\tilde{\xi} = \mathcal{D}^*_\varepsilon\mathcal{D}_\varepsilon\xi. \tag{4.2}$$

Here the operator $\mathcal{D}^*_\varepsilon$ is obtained from $\mathcal{D}_\varepsilon$ by replacing ∇_t with $-\nabla_t$. It is the formal adjoint operator of $\mathcal{D}_\varepsilon$ with respect to the Hilbert space structure induced by the L^2_ε-norm. The first component of (4.2) can be written as

$$\begin{aligned}
&-\nabla_t\tilde{\alpha} - \mathrm{d}_A\tilde{\psi} + *_s\nabla_s\tilde{\alpha} - *_s\mathrm{d}_A\tilde{\phi} \\
&\quad = -\nabla_t\nabla_t\alpha + *_s\nabla_s *_s \nabla_s\alpha + \varepsilon^{-2}\mathrm{d}_A{}^*\mathrm{d}_A\alpha + \varepsilon^{-2}\mathrm{d}_A\mathrm{d}_A{}^*\alpha \\
&\qquad - *_s\,[C \wedge \alpha] + [(B_t - *_s B_s) \wedge \psi] + [(*_s B_t + B_s) \wedge \phi] \\
&\qquad - \mathrm{d}_A *_s \dot{*}_s\phi - *_s\dot{*}_s\mathrm{d}_A\phi
\end{aligned}$$

where $\dot{*}_s$ denotes the derivative of the Hodge-$*$-operator $*_s$ with respect to s and

$$B_s = \partial_s A - \mathrm{d}_A\Phi, \qquad B_t = \partial_t A - \mathrm{d}_A\Psi, \qquad C = \partial_t\Phi - \partial_s\Psi - [\Phi, \Psi].$$

Take the inner product with α to obtain

$$
\begin{aligned}
&\|\nabla_s\alpha\|^2 + \|\nabla_t\alpha\|^2 + \varepsilon^{-2}\|\mathrm{d}_A\alpha\|^2 + \varepsilon^{-2}\|\mathrm{d}_A *_s \alpha\|^2 \\
&\quad = \langle \nabla_t\alpha + *_s\nabla_s\alpha, \tilde{\alpha}\rangle + \langle *_s\mathrm{d}_A\alpha, \tilde{\phi}\rangle - \langle \mathrm{d}_A{}^*\alpha, \tilde{\psi}\rangle \\
&\qquad + \langle \alpha, *_s[C\wedge\alpha]\rangle - \langle \alpha, [(B_t - *_sB_s)\wedge\psi]\rangle - \langle\alpha, [(*_sB_t + B_s)\wedge\phi]\rangle \\
&\qquad + \langle \alpha, \mathrm{d}_A *_s \dot{*}_s\phi\rangle + \langle *_s\dot{*}_s\alpha, \mathrm{d}_A\phi\rangle.
\end{aligned}
\tag{4.3}
$$

Similarly, the second component of (4.2) can be written in the form

$$
\begin{aligned}
&-\nabla_t\tilde{\phi} - \nabla_s\tilde{\psi} + \varepsilon^{-2} *_s \mathrm{d}_A\tilde{\alpha} \\
&\quad = -\nabla_t\nabla_t\phi - \nabla_s *_s \nabla_s *_s \phi + \varepsilon^{-2}\mathrm{d}_A{}^*\mathrm{d}_A\phi \\
&\qquad + [(C - \varepsilon^{-2} *_s F_A)\wedge\psi] - \varepsilon^{-2} *_s [B_s\wedge *_s\alpha] - \varepsilon^{-2} *_s [B_t\wedge\alpha] \\
&\qquad - \varepsilon^{-2}\dot{*}_s\mathrm{d}_A *_s \alpha - \varepsilon^{-2} *_s \mathrm{d}_A\dot{*}_s\alpha.
\end{aligned}
$$

Take the inner product with ϕ and use $F_A = 0$ and

$$
\langle \phi, \nabla_s *_s \nabla_s *_s \phi\rangle = -\|\nabla_s\phi + *_s\dot{*}_s\phi\|^2
$$

to obtain

$$
\begin{aligned}
&\|\nabla_s\phi\|^2 + \|\nabla_t\phi\|^2 + \varepsilon^{-2}\|\mathrm{d}_A\phi\|^2 \\
&\quad = \langle \nabla_t\phi, \tilde{\phi}\rangle + \langle \nabla_s\phi + *_s\dot{*}_s\phi, \tilde{\psi}\rangle - \varepsilon^{-2}\langle *_s\mathrm{d}_A\phi, \tilde{\alpha}\rangle \\
&\qquad - 2\langle \nabla_s\phi, *_s\dot{*}_s\phi\rangle - \| *_s \dot{*}_s\phi\|^2 - \langle \phi, [C\wedge\psi]\rangle \\
&\qquad + \varepsilon^{-2}\langle \phi, *_s[B_s\wedge *_s\alpha]\rangle + \varepsilon^{-2}\langle\phi, *_s[B_t\wedge\alpha]\rangle \\
&\qquad + \varepsilon^{-2}\langle \alpha, \mathrm{d}_A *_s \dot{*}_s\phi\rangle + \varepsilon^{-2}\langle *_s\dot{*}_s\alpha, \mathrm{d}_A\phi\rangle.
\end{aligned}
\tag{4.4}
$$

The third component of (4.2) can be written in the form

$$
\begin{aligned}
&-\nabla_t\tilde{\psi} + *_s\nabla_s *_s \tilde{\phi} + \varepsilon^{-2} *_s \mathrm{d}_A *_s \tilde{\alpha} \\
&\quad = -\nabla_t\nabla_t\psi - *_s\nabla_s *_s \nabla_s\psi + \varepsilon^{-2}\mathrm{d}_A{}^* d_A\psi \\
&\qquad - [(C - \varepsilon^{-2} *_s F_A)\wedge\phi] + \varepsilon^{-2} *_s [B_s\wedge\alpha] - \varepsilon^{-2} *_s [B_t\wedge *_s\alpha].
\end{aligned}
$$

Take the inner product with ψ and use $F_A = 0$ to obtain

$$
\begin{aligned}
&\|\nabla_s\psi\|^2 + \|\nabla_t\psi\|^2 + \varepsilon^{-2}\|\mathrm{d}_A\psi\|^2 \\
&\quad = \langle \nabla_t\psi, \tilde{\psi}\rangle - \langle \nabla_s\psi, \tilde{\phi}\rangle - \varepsilon^{-2}\langle \mathrm{d}_A\psi, \tilde{\alpha}\rangle \\
&\qquad + \langle \psi, [C\wedge\phi]\rangle - \varepsilon^{-2}\langle \psi, *_s[B_s\wedge\alpha]\rangle + \varepsilon^{-2}\langle\psi, *_s[B_t\wedge *_s\alpha]\rangle.
\end{aligned}
\tag{4.5}
$$

The estimate (4.1) follows from (4.3), (4.4), and (4.5). This proves the lemma for $p = 2$. For general p the estimate can be reduced to the L^p-estimate for Laplace's equation via the rescaling argument of Lemma 4.1. □

Lemma 4.3. *There exists a constant* $c = c(\Xi, p) > 0$ *such that*

$$\|\pi_A(\mathcal{D}_\varepsilon\xi) - \mathcal{D}_0\pi_A(\xi)\|_{\mathrm{L}^p} \le c\,\|\xi - \pi_A(\xi)\|_{0,p,\varepsilon}$$

for $\xi \in \mathrm{W}^{1,p}_f$.

Proof. Let $\xi = (\alpha, \phi, \psi) \in \mathrm{W}^{1,p}_f$ and write $\alpha = \pi_A(\alpha) + d_A\zeta + *_s d_A\eta$. Then a simple calculation shows that

$$\pi_A(\mathcal{D}_\varepsilon\xi) = \mathcal{D}_0\pi_A(\xi) + \pi_A(\theta) \tag{4.6}$$

where $\theta = [B \wedge \zeta] + *_s[B \wedge \eta] + *_s\dot{*}_s\mathrm{d}_A\eta - *_s\mathrm{d}X_s(A) *_s d_A\eta - [X_s(A) \wedge \eta]$ and $B = B_t + *_s(B_s - X_s(A))$. Hence $\|\theta\|_{\mathrm{L}^p} \le c\,\|\xi - \pi_A(\xi)\|_{0,p,\varepsilon}$ as required. □

If Ξ satisfies (3.4), the Hodge-$*$-operator $*_s = *$ is independent of s, and the perturbation X_s vanishes then $\pi_A \circ \mathcal{D}_\varepsilon = \mathcal{D}_0 \circ \pi_A$. So in this case the projection onto the harmonic part determines an isomorphism of the kernel of $\mathcal{D}_\varepsilon$ with the kernel of $\mathcal{D}_0$ provided that both operators are onto (see Lemma 4.5 below).

Lemma 4.4. *Assume*

$$\|\alpha_0\|_{\mathrm{L}^p} \le c_0\,\|\mathcal{D}_0\alpha_0\|_{\mathrm{L}^p}$$

for all $\alpha_0 \in \mathrm{W}^{1,p}_f(\mathrm{H}_A)$ *and some constant* $c_0 > 0$. *Then there exist constants* $\varepsilon_0 > 0$ *and* $c > 0$ *such that*

$$\|\xi\|_{1,p,\varepsilon} \le c\left(\varepsilon\,\|\mathcal{D}_\varepsilon\xi\|_{0,p,\varepsilon} + \|\pi_A(\mathcal{D}_\varepsilon\xi)\|_{\mathrm{L}^p}\right),$$

$$\|\xi - \pi_A(\xi)\|_{1,p,\varepsilon} \le c\varepsilon\,\|\mathcal{D}_\varepsilon\xi\|_{0,p,\varepsilon}$$

for $0 < \varepsilon \le \varepsilon_0$ *and* $\xi \in \mathrm{W}^{1,p}_f$.

Proof. For every $\xi \in \mathrm{W}^{1,p}_f$

$$\begin{aligned}
\|\pi_A(\xi)\|_{\mathrm{L}^p} &\le c_0\,\|\mathcal{D}_0\pi_A(\xi)\|_{\mathrm{L}^p} \\
&\le c_0\,\|\pi_A(\mathcal{D}_\varepsilon\xi)\|_{\mathrm{L}^p} + c_1\,\|\xi - \pi_A(\xi)\|_{0,p,\varepsilon} \\
&\le c_0\,\|\pi_A(\mathcal{D}_\varepsilon\xi)\|_{\mathrm{L}^p} + c_2\varepsilon\left(\|\mathcal{D}_\varepsilon\xi\|_{0,p,\varepsilon} + \|\pi_A(\xi)\|_{\mathrm{L}^p}\right)
\end{aligned}$$

The first inequality follows from the assumption of the lemma, the second from Lemma 4.3, and the last from Lemma 4.2. For $c_2\varepsilon_0 < 1$ this implies

$$\|\pi_A(\xi)\|_{\mathrm{L}^p} \le c_3\left(\varepsilon\,\|\mathcal{D}_\varepsilon\xi\|_{0,p,\varepsilon} + \|\pi_A(\mathcal{D}_\varepsilon\xi)\|_{\mathrm{L}^p}\right).$$

Hence the statement follows from Lemma 4.2. □

LEMMA 4.5. *Assume $\mathcal{D}_0$ is onto. Then there exist constants $\varepsilon_0 > 0$ and $c > 0$ such that $\mathcal{D}_\varepsilon$ is onto for $0 < \varepsilon < \varepsilon_0$ and*

$$\|\mathcal{D}_\varepsilon^*\xi\|_{1,p,\varepsilon} \le c\left(\varepsilon\,\|\mathcal{D}_\varepsilon\mathcal{D}_\varepsilon^*\xi\|_{0,p,\varepsilon} + \|\pi_A(\mathcal{D}_\varepsilon\mathcal{D}_\varepsilon^*\xi)\|_{\mathrm{L}^p}\right),$$

$$\|\mathcal{D}_\varepsilon^*\xi - \pi_A(\mathcal{D}_\varepsilon^*\xi)\|_{1,p,\varepsilon} \le c\varepsilon\,\|\mathcal{D}_\varepsilon\mathcal{D}_\varepsilon^*\xi\|_{0,p,\varepsilon}$$

for $\xi \in \mathrm{W}_f^{2,p}$.

Proof. The proof is in four steps.

Step 1. Choose $q > 1$ such that $1/p + 1/q = 1$. Then, since $\mathcal{D}_0$ is onto, there is a constant $c_0 > 0$ such that for $\alpha_0 \in \mathrm{W}_f^{1,p}(\mathrm{H}_A)$

$$\|\alpha_0\|_{\mathrm{L}^q} \le c_0\,\|\mathcal{D}_0^*\alpha_0\|_{\mathrm{L}^q}.$$

Step 2. There is a constant $c_1 > 0$ such that for $\alpha_0 \in \mathrm{W}_f^{1,p}(\mathrm{H}_A)$

$$\|\mathcal{D}_0^*\alpha_0\|_{\mathrm{L}^p} \le c_1 \sup_{\beta_0} \frac{\langle \mathcal{D}_0^*\alpha_0, \mathcal{D}_0^*\beta_0\rangle}{\|\mathcal{D}_0^*\beta_0\|_{\mathrm{L}^q}}$$

Choose an L^2-orthonormal basis $\alpha_1, \dots, \alpha_m$ of $\ker \mathcal{D}_0$. Now choose $\beta \in \mathrm{L}_f^q(\mathrm{H}_A)$ such that

$$\langle \beta, \mathcal{D}_0^*\alpha_0\rangle = \|\mathcal{D}_0^*\alpha_0\|_{\mathrm{L}^p}, \qquad \|\beta\|_{\mathrm{L}^q} = 1.$$

Since $\mathcal{D}_0$ is onto there exists a (unique) $\beta_0 \in \mathrm{W}^{1,q}(\mathrm{H}_A)$ such that

$$\beta = \mathcal{D}_0^*\beta_0 + \sum_{j=1}^m \langle \beta, \alpha_j\rangle \alpha_j.$$

It follows that

$$\begin{aligned}
\|\mathcal{D}_0^*\alpha_0\|_{\mathrm{L}^p} &= \langle \beta, \mathcal{D}_0^*\alpha_0\rangle \\
&= \langle \mathcal{D}_0^*\beta_0, \mathcal{D}_0^*\alpha_0\rangle \\
&= \left\|\beta - \sum_{j=1}^m \langle \beta, \alpha_j\rangle\alpha_j\right\|_{\mathrm{L}^q} \frac{\langle \mathcal{D}_0^*\beta_0, \mathcal{D}_0^*\alpha_0\rangle}{\|\mathcal{D}_0^*\beta_0\|_{\mathrm{L}^q}} \\
&\le \left(1 + \sum_{j=1}^m \|\alpha_j\|_{\mathrm{L}^p}\,\|\alpha_j\|_{\mathrm{L}^q}\right) \frac{\langle \mathcal{D}_0^*\beta_0, \mathcal{D}_0^*\alpha_0\rangle}{\|\mathcal{D}_0^*\beta_0\|_{\mathrm{L}^q}}.
\end{aligned}$$

Step 3. For every $\varepsilon > 0$ and every $\xi \in \mathrm{W}_f^{1,p}$

$$\|\pi_A(\mathcal{D}_\varepsilon^*\xi)\|_{\mathrm{L}^p} \le (1 + c_1)\,\|\pi_A(\mathcal{D}_\varepsilon^*\xi) - \mathcal{D}_0^*\pi_A(\xi)\|_{\mathrm{L}^p} + c_0c_1\,\|\mathcal{D}_0\pi_A(\mathcal{D}_\varepsilon^*\xi)\|_{\mathrm{L}^p}.$$

For all $\xi, \zeta \in \mathrm{W}^{1,p}_f$ we have

$$\begin{aligned}
\frac{\langle \mathcal{D}_0^*\pi_A(\zeta), \mathcal{D}_0^*\pi_A(\xi)\rangle}{\|\mathcal{D}_0^*\pi_A(\zeta)\|_{\mathrm{L}^q}} &= \frac{\langle \mathcal{D}_0^*\pi_A(\zeta), \mathcal{D}_0^*\pi_A(\xi) - \pi_A(\mathcal{D}_\varepsilon^*\xi)\rangle}{\|\mathcal{D}_0^*\pi_A(\zeta)\|_{\mathrm{L}^q}} \\
&\quad + \frac{\langle \pi_A(\zeta), \mathcal{D}_0\pi_A(\mathcal{D}_\varepsilon^*\xi)\rangle}{\|\mathcal{D}_0^*\pi_A(\zeta)\|_{\mathrm{L}^q}} \\
&\le \|\mathcal{D}_0^*\pi_A(\xi) - \pi_A(\mathcal{D}_\varepsilon^*\xi)\|_{\mathrm{L}^p} \\
&\quad + \|\mathcal{D}_0\pi_A(\mathcal{D}_\varepsilon^*\xi)\|_{\mathrm{L}^p}\frac{\|\pi_A(\zeta)\|_{\mathrm{L}^q}}{\|\mathcal{D}_0^*\pi_A(\zeta)\|_{\mathrm{L}^q}} \\
&\le \|\mathcal{D}_0^*\pi_A(\xi) - \pi_A(\mathcal{D}_\varepsilon^*\xi)\|_{\mathrm{L}^p} + c_0\|\mathcal{D}_0\pi_A(\mathcal{D}_\varepsilon^*\xi)\|_{\mathrm{L}^p}.
\end{aligned}$$

The last inequality follows from step 1. Now it follows from step 2 that

$$\|\mathcal{D}_0^*\pi_A(\xi)\|_{\mathrm{L}^p} \le c_1\|\mathcal{D}_0^*\pi_A(\xi) - \pi_A(\mathcal{D}_\varepsilon^*\xi)\|_{\mathrm{L}^p} + c_0c_1\|\mathcal{D}_0\pi_A(\mathcal{D}_\varepsilon^*\xi)\|_{\mathrm{L}^p}.$$

Step 4. We prove the lemma.

By step 3 and Lemma 4.3 we have

$$\begin{aligned}
\|\pi_A(\mathcal{D}_\varepsilon^*\xi)\|_{L^p} &\le c_2\|\pi_A(\mathcal{D}_\varepsilon^*\xi) - \mathcal{D}_0^*\pi_A(\xi)\|_{\mathrm{L}^p} + c_2\|\mathcal{D}_0\pi_A(\mathcal{D}_\varepsilon^*\xi)\|_{\mathrm{L}^p} \\
&\le c_3\|\xi - \pi_A(\xi)\|_{0,p,\varepsilon} \\
&\quad + c_3\|\mathcal{D}_\varepsilon^*\xi - \pi_A(\mathcal{D}_\varepsilon^*\xi)\|_{0,p,\varepsilon} + c_2\|\pi_A(\mathcal{D}_\varepsilon\mathcal{D}_\varepsilon^*\xi)\|_{\mathrm{L}^p} \\
&\le c_4\varepsilon\|\mathcal{D}_\varepsilon^*\xi\|_{0,p,\varepsilon} \\
&\quad + c_4\varepsilon\|\mathcal{D}_\varepsilon\mathcal{D}_\varepsilon^*\xi\|_{0,p,\varepsilon} + c_2\|\pi_A(\mathcal{D}_\varepsilon\mathcal{D}_\varepsilon^*\xi)\|_{\mathrm{L}^p}.
\end{aligned}$$

The first inequality follows from step 3. The second inequality follows from Lemma 4.3 applied to both $\mathcal{D}_\varepsilon$ and $\mathcal{D}_\varepsilon^*$. The last inequality follows from Lemma 4.4 for the operator $\mathcal{D}_\varepsilon^*$ and from Lemma 4.2 for $\mathcal{D}_\varepsilon$. By Lemma 4.2,

$$\begin{aligned}
\|\mathcal{D}_\varepsilon^*\xi\|_{0,p,\varepsilon} &\le c_5\varepsilon\|\mathcal{D}_\varepsilon\mathcal{D}_\varepsilon^*\xi\|_{0,p,\varepsilon} + c_5\|\pi_A(\mathcal{D}_\varepsilon^*\xi)\|_{\mathrm{L}^p} \\
&\le c_6\varepsilon\|\mathcal{D}_\varepsilon^*\xi\|_{0,p,\varepsilon} + c_6\varepsilon\|\mathcal{D}_\varepsilon\mathcal{D}_\varepsilon^*\xi\|_{0,p,\varepsilon} + c_6\|\pi_A(\mathcal{D}_\varepsilon\mathcal{D}_\varepsilon^*\xi)\|_{\mathrm{L}^p}.
\end{aligned}$$

With $c_6\varepsilon < 1$ this proves the required estimate. □

5. Approximation of holomorphic curves by self-dual instantons

Let $p > 4$, $H \in \mathcal{H}_0^{\mathrm{reg}}$ be a regular perturbation and fix two H-flat connections $a^\pm = A^\pm + \Phi^\pm\,\mathrm{d}s \in \mathcal{A}_\Sigma(P_f, H)$. We shall prove that every holomorphic curve $\Xi_0 \in \mathcal{A}_0^{1,p}(a^-, a^+, H)$ can be *approximated* by a family of self-dual instantons $\Xi_\varepsilon \in \mathcal{A}_\varepsilon^{1,p}(a^-, a^+, H)$. This requires a refinement of the implicit function theorem with constants independent of ε.

Theorem 5.1. *Assume $\mu_H(a^-, a^+) < 4$. Then for ε sufficiently small there is a smooth map $\mathcal{T}_\varepsilon : \mathcal{A}_0^{1,p}(a^-, a^+, H) \to \mathcal{A}_\varepsilon^{1,p}(a^-, a^+, H)$ such that $\Xi_\varepsilon = \mathcal{T}_\varepsilon(\Xi_0)$ satisfies*

$$\mathrm{d}_{\Xi_0}^{*_\varepsilon}(\Xi_\varepsilon - \Xi_0) = 0, \qquad \Xi_\varepsilon - \Xi_0 \in \mathrm{range}\,\mathcal{D}_\varepsilon(\Xi_0)^*. \tag{5.1}$$

for every $\Xi_0 \in \mathcal{A}_0^{1,p}(a^-, a^+, H)$. *Moreover,* Ξ_ε *satisfies an estimates*

$$\|\Xi_\varepsilon - \Xi_0\|_{1,p,\varepsilon} \le c\varepsilon^2. \tag{5.2}$$

Here the $1, p, \varepsilon$*-norm is the one determined by* Ξ_0 *and the constant* $c > 0$ *can be chosen independent of* Ξ_0 *and* ε.

THEOREM 5.2. *Assume* $\mu_H(a^-, a^+) < 4$ *and let* $\Xi_0 \in \mathcal{A}_0^{1,p}(a^-, a^+, H)$. *Then there exist constants* $\delta > 0$ *and* $\varepsilon_0 > 0$ *such that if* $\Xi \in \mathcal{A}_\varepsilon^{1,p}(a^-, a^+, H)$ *satisfies (5.1) and*

$$\|\Xi - \Xi_0\|_{0,p,\varepsilon} + \varepsilon^{2/p} \|\Xi - \Xi_0\|_{\infty,\varepsilon} \le \delta\varepsilon^{2/p+1/2} \tag{5.3}$$

with $0 < \varepsilon \le \varepsilon_0$ *then* $\Xi = \mathcal{T}_\varepsilon(\Xi_0)$.

Remark 5.3. Here $d_\Xi^{*\varepsilon}$ denotes the formal adjoint operator of d_Ξ with respect to the ε inner product. The operator d_Ξ represents the infinitesimal action of $\mathcal{G}^{2,p}$ on $\mathcal{A}^{1,p}(a^-, a^+)$ and is given by

$$\mathrm{d}_\Xi \eta = \mathrm{d}_A \eta + \nabla_s \eta ds + \nabla_t \eta dt$$

for $\eta \in W_f^{2,p}(\mathbb{R}^2 \times \mathfrak{g}_P)$. Its formal adjoint is

$$\mathrm{d}_\Xi^{*\varepsilon} \xi = - *_s \mathrm{d}_A *_s \alpha - \varepsilon^2 *_s \nabla_s *_s \phi - \varepsilon^2 \nabla_t \psi$$

for $\xi = (\alpha, \phi, \psi) \in W_f^{1,p}$. This operator agrees with the third component of $\mathcal{D}_\varepsilon$ up to a scalar factor.

Remark 5.4. The estimate (5.2) shows that A_ε, $\partial A_\varepsilon/\partial s$, $\partial A_\varepsilon/\partial t$, Φ_ε, and Ψ_ε converge in the L^p-norm as ε tends to zero while $\partial \Phi_\varepsilon/\partial s$, $\partial \Phi_\varepsilon/\partial t$, $\partial \Psi_\varepsilon/\partial s$, and $\partial \Psi_\varepsilon/\partial t$ remain bounded, uniformly in ε. In particular, the curvature F_{A_ε} converges to zero in the L^p-norm like ε^2.

Remark 5.5. The map $\mathcal{T}_\varepsilon$ is equivariant under the action of $\mathcal{G}^{2,p}$

$$\mathcal{T}_\varepsilon(g^* \Xi_0) = g^* \mathcal{T}_\varepsilon(\Xi_0).$$

The induced map of the moduli spaces will also be denoted by $\mathcal{T}_\varepsilon$.

Remark 5.6. In the case $\mu_H(a^-, a^+) \ge 4$ the proof shows that for sufficiently small ε the map $\mathcal{T}_\varepsilon$ can be defined on any compact subset of $\mathcal{M}_0(a^-, a^+, H)$. It does not show whether there is a uniform $\varepsilon > 0$ which works simultaneously for all Ξ_0.

Proof of Theorem 5.1. We prove that with a suitable constant c and $\varepsilon > 0$ sufficiently small there exists a unique solution $\Xi_\varepsilon \in \mathcal{A}^{1,p}_\varepsilon(a^-, a^+, H)$ of (5.1) and (5.2) for every $\Xi_0 \in \mathcal{A}^{1,p}_0(a^-, a^+, H)$.

Throughout let $c_0, c_1, c_2, \ldots$ denote constants which are independent of Ξ_0 and ε. Since $H \in \mathcal{H}^{\mathrm{reg}}_0$ the operator $\mathcal{D}_0 = \mathcal{D}_0(\Xi_0)$ is onto for every $\Xi_0 \in \mathcal{A}^{1,p}_0(a^-, a^+, H)$. Since $\mu_H(a^-, a^+) < 4$ there is no bubbling and it follows from Floer's glueing construction [12] that the operators $\mathcal{D}_0$ satisfy the estimate of Lemma 4.5 with a constant which is independent of Ξ_0. This is obvious in the case $\mu = 1$ since there are only finitely many connecting orbits and the estimate is invariant under gauge transformations. In the case $\mu = 2$ Floer's glueing construction shows that the estimate holds with a uniform constant for all orbits near a catenation of two connecting orbits with index difference 1. This takes care of the ends of the moduli space $\mathcal{M}(a^-, a^+)$ and the complement of the ends is compact. A similar argument works for $\mu = 3$. But for $\mu \geq 4$ bubbling may occur.

Now it follows from Lemma 4.5 that there exist constants $c_1 > 0$ and $\varepsilon_1 > 0$ such that for every $\Xi_0 \in \mathcal{A}^{1,p}_0(a^-, a^+)$ and for $0 < \varepsilon \leq \varepsilon_1$ the operator $\mathcal{D}_\varepsilon = \mathcal{D}_\varepsilon(\Xi_0)$ satisfies

$$\|\mathcal{D}_\varepsilon^*\eta\|_{1,p,\varepsilon} + \varepsilon^{2/p}\|\mathcal{D}_\varepsilon^*\eta\|_{\infty,\varepsilon} \leq c_1\left(\varepsilon\|\mathcal{D}_\varepsilon\mathcal{D}_\varepsilon^*\eta\|_{0,p,\varepsilon} + \|\pi_A(\mathcal{D}_\varepsilon\mathcal{D}_\varepsilon^*\eta)\|_{\mathrm{L}^p}\right). \tag{5.4}$$

Here the L^∞ estimate follows from Lemma 4.1. The above argument shows that c_1 can be chosen independent of Ξ_0.

The left hand side of (3.5) defines a smooth map $\mathcal{F}_\varepsilon : \mathcal{A}^{1,p} \to L^p_f$ given by

$$\begin{pmatrix} A \\ \Phi \\ \Psi \end{pmatrix} \mapsto \begin{pmatrix} \partial_t A - \mathrm{d}_A\Psi + *_s(\partial_s A - X_s(A) - \mathrm{d}_A\Phi) \\ \partial_t\Phi - \partial_s\Psi - [\Phi, \Psi] + \varepsilon^{-2} *_s F_A \\ 0 \end{pmatrix}.$$

Now Ξ_0 is an approximate zero of $\mathcal{F}_\varepsilon$ in the sense that

$$\|\mathcal{F}_\varepsilon(\Xi_0)\|_{0,p,\varepsilon} \leq c_0\varepsilon \tag{5.5}$$

We shall construct the solution Ξ_ε of $\mathcal{F}_\varepsilon(\Xi_\varepsilon) = 0$ by Newton's iteration. The first step is to define

$$\Xi_1 = \Xi_0 + \xi_0, \qquad \xi_0 = \mathcal{D}_\varepsilon^*\eta_0, \qquad \mathcal{D}_\varepsilon\mathcal{D}_\varepsilon^*\eta_0 = -\mathcal{F}_\varepsilon(\Xi_0). \tag{5.6}$$

In particular, this means that the third component of $\mathcal{D}_\varepsilon\xi_0$ vanishes. The last equation in (5.6) has a unique solution $\eta_0 \in W^{2,p}_f(\mathbb{R}^2 \times \mathfrak{g}_P)$. Since the first component of $\mathcal{F}_\varepsilon(\Xi_0)$ vanishes it follows from (5.4) that

$$\|\xi_0\|_{1,p,\varepsilon} + \varepsilon^{2/p}\|\xi_0\|_{\infty,\varepsilon} \leq c_1\varepsilon\|\mathcal{F}_\varepsilon(\Xi_0)\|_{0,p,\varepsilon} \leq c_0c_1\varepsilon^2 \tag{5.7}$$

Since the first two components of the operator $\mathcal{D}_\varepsilon$ are the linearization of $\mathcal{F}_\varepsilon$ the last equation in (5.6) can be written as $d\mathcal{F}_\varepsilon(\Xi_0)\xi_0 = -\mathcal{F}_\varepsilon(\Xi_0)$. Hence

$$\begin{aligned} \mathcal{F}_\varepsilon(\Xi_1) &= \mathcal{F}_\varepsilon(\Xi_0+\xi_0) - \mathcal{F}_\varepsilon(\Xi_0) - \mathrm{d}\mathcal{F}_\varepsilon(\Xi_0)\xi_0 \\ &= \begin{pmatrix} -[\alpha_0\wedge\psi_0] - *_s[\alpha_0\wedge\phi_0] - *_s\tilde{\alpha}_0 \\ -[\phi_0\wedge\psi_0] + 2^{-1}\varepsilon^{-2} *_s [\alpha_0\wedge\alpha_0] \\ 0 \end{pmatrix} \end{aligned}$$

where $\xi_0 = (\alpha_0, \phi_0, \psi_0)$ and $\tilde{\alpha}_0 = X_s(A_0+\alpha_0) - X_s(A_0) - \mathrm{d}X_s(A_0)\alpha_0$. Hence

$$\|\mathcal{F}_\varepsilon(\Xi_1)\|_{0,p,\varepsilon} \le c_2\varepsilon^{-1}\|\xi_0\|_{0,p,\varepsilon}\|\xi_0\|_{\infty,\varepsilon} \le c_0c_1c_2\varepsilon^{1-2/p}\|\xi_0\|_{0,p,\varepsilon}.$$

The last inequality follows from (5.7).

Now assume $\Xi_\nu \in \mathcal{A}^{1,p}$ has been constructed for $\nu \ge 1$ and define $\Xi_{\nu+1}$ by

$$\Xi_{\nu+1} = \Xi_\nu + \xi_\nu, \qquad \xi_\nu = \mathcal{D}_\varepsilon^*\eta_\nu, \qquad \mathcal{D}_\varepsilon\mathcal{D}_\varepsilon^*\eta_\nu = -\mathcal{F}_\varepsilon(\Xi_\nu).$$

We shall prove by induction that

$$\begin{aligned} \|\xi_\nu\|_{1,p,\varepsilon} + \varepsilon^{2/p}\|\xi_\nu\|_{\infty,\varepsilon} &\le 2c_1\|\mathcal{F}_\varepsilon(\Xi_\nu)\|_{0,p,\varepsilon} \\ \|\xi_\nu\|_{1,p,\varepsilon} + \varepsilon^{2/p}\|\xi_\nu\|_{\infty,\varepsilon} &\le 2^{-\nu}c_0c_1\varepsilon^2, \\ \|\mathcal{F}_\varepsilon(\Xi_{\nu+1})\|_{0,p,\varepsilon} &\le c_3\varepsilon^{1-2/p}\|\xi_\nu\|_{0,p,\varepsilon} \end{aligned} \tag{5.8}$$

for $0 < \varepsilon < \varepsilon_0$ provided that $\varepsilon_0 > 0$ is sufficiently small. The first inequality in (5.8) follows from (5.4). (The first component of $\mathcal{F}_\varepsilon(\Xi_\nu)$ no longer vanishes.) The second inequality in (5.8) follows from the first and from the previous induction steps:

$$\begin{aligned} \|\xi_\nu\|_{1,p,\varepsilon} + \varepsilon^{2/p}\|\xi_\nu\|_{\infty,\varepsilon} &\le 2c_1\|\mathcal{F}_\varepsilon(\Xi_\nu)\|_{0,p,\varepsilon} \\ &\le 2c_1c_3\varepsilon^{1-2/p}\|\xi_{\nu-1}\|_{0,p,\varepsilon} \\ &\le 2^{-1}\|\xi_{\nu-1}\|_{0,p,\varepsilon} \\ &\le 2^{-\nu}\|\xi_0\|_{0,p,\varepsilon} \\ &\le 2^{-\nu}c_0c_1\varepsilon^2. \end{aligned}$$

Here we have used (5.7) and chosen ε_0 such that $2c_1c_3\varepsilon_0^{1-2/p} \le 1/2$. This implies

$$\|\Xi_\nu - \Xi_0\|_{1,p,\varepsilon} \le \sum_{j=0}^{\nu-1}\|\xi_j\|_{1,p,\varepsilon} \le 2c_0c_1\varepsilon^2. \tag{5.9}$$

A similar argument shows that

$$\|\Xi_\nu - \Xi_0\|_{\infty,\varepsilon} \le 2c_0c_1\varepsilon^{2-2/p}. \tag{5.10}$$

To prove the third inequality in (5.8) note that

$$\begin{aligned}\mathcal{F}_\varepsilon(\Xi_{\nu+1}) &= \mathcal{F}_\varepsilon(\Xi_\nu+\xi_\nu) - \mathcal{F}_\varepsilon(\Xi_\nu) - \mathrm{d}\mathcal{F}_\varepsilon(\Xi_0)\xi_\nu \\ &= \mathcal{F}_\varepsilon(\Xi_\nu+\xi_\nu) - \mathcal{F}_\varepsilon(\Xi_\nu) - \mathrm{d}\mathcal{F}_\varepsilon(\Xi_\nu)\xi_\nu \\ &\quad +(\mathrm{d}\mathcal{F}_\varepsilon(\Xi_\nu) - \mathrm{d}\mathcal{F}_\varepsilon(\Xi_0))\xi_\nu\end{aligned}$$

and hence

$$\begin{aligned}\|\mathcal{F}_\varepsilon(\Xi_{\nu+1})\|_{0,p,\varepsilon} &\le \|\mathcal{F}_\varepsilon(\Xi_\nu+\xi_\nu) - \mathcal{F}_\varepsilon(\Xi_\nu) - \mathrm{d}\mathcal{F}_\varepsilon(\Xi_\nu)\xi_\nu\|_{0,p,\varepsilon} \\ &\quad + \|(\mathrm{d}\mathcal{F}_\varepsilon(\Xi_\nu) - \mathrm{d}\mathcal{F}_\varepsilon(\Xi_0))\xi_\nu\|_{0,p,\varepsilon} \\ &\le c_4\varepsilon^{-1}\left(\|\xi_\nu\|_{\infty,\varepsilon} + \|\Xi_\nu - \Xi_0\|_{\infty,\varepsilon}\right)\|\xi_\nu\|_{0,p,\varepsilon} \\ &\le 3c_0c_1c_4\varepsilon^{1-2/p}\|\xi_\nu\|_{0,p,\varepsilon}.\end{aligned}$$

The last inequality follows from (5.10) and finishes the induction.

By (5.8) the sequence Ξ_ν converges in $\mathcal{A}^{1,p}(a^-,a^+)$ and by (5.9) the limit $\Xi_\varepsilon = \lim_{\nu\to\infty}\Xi_\nu$ satisfies $\|\Xi_\varepsilon - \Xi_0\|_{1,p,\varepsilon} \le 2c_0c_1\varepsilon^2$. The second inequality in (5.8) shows that $\mathcal{F}_\varepsilon(\Xi_\varepsilon) = 0$ and this proves the theorem. □

Proof of Theorem 5.2. Assume $\Xi \in \mathcal{A}^{1,p}_\varepsilon(a^-,a^+,H)$ satisfies (5.1) and (5.3). We shall first prove that if δ and ε are sufficiently small then

$$\|\Xi - \Xi_0\|_{1,p,\varepsilon} \le c\varepsilon^2. \tag{5.11}$$

for some constant $c > 0$ which is independent of Ξ and ε. To see this note that, by (5.1), the last component of $\mathcal{D}_\varepsilon(\Xi - \Xi_0) = (\alpha,\phi,0)$ vanishes. The first two are given by

$$\begin{aligned}\alpha &= [A - A_0 \wedge \Psi - \Psi_0] + *_s[A - A_0 \wedge \Phi - \Phi_0] \\ &\quad + *_s\left(X_s(A) - X_s(A_0) - \mathrm{d}X_s(A_0)(A - A_0)\right), \\ \phi &= [\Phi - \Phi_0 \wedge \Psi - \Psi_0] - \tfrac{1}{2}\varepsilon^{-2} *_s [A - A_0 \wedge A - A_0] - C_0\end{aligned}$$

where $C_0 = \partial_t\Phi_0 - \partial_s\Psi_0 - [\Phi_0,\Psi_0]$. Moreover, again by (5.1), $\Xi - \Xi_0 \in \mathrm{range}\mathcal{D}^*_\varepsilon$. Hence, by Lemma 4.5,

$$\begin{aligned}\|\Phi - \Phi_0\|_{L^p} + \|\Psi - \Psi_0\|_{L^p} &\le c_1\|\mathcal{D}_\varepsilon(\Xi - \Xi_0)\|_{0,p,\varepsilon} \\ &\le c_2\left(\|\alpha\|_{L^p} + \varepsilon\|\phi\|_{L^p}\right) \\ &\le c_3\|\Xi - \Xi_0\|_{\infty,\varepsilon}\left(\|\Phi - \Phi_0\|_{L^p} + \|\Psi - \Psi_0\|_{L^p}\right) \\ &\quad + c_3\left(\varepsilon + \varepsilon^{-1}\|A - A_0\|_{L^\infty}\|A - A_0\|_{L^p}\right) \\ &\le c_3\delta\varepsilon^{1/2}\left(\|\Phi - \Phi_0\|_{L^p} + \|\Psi - \Psi_0\|_{L^p}\right) \\ &\quad + c_3(\varepsilon + \delta^2\varepsilon^{2/p}).\end{aligned}$$

The last inequality follows from (5.3). With $c_3\delta\varepsilon^{1/2} \le 1/2$ we obtain

$$\|\Phi - \Phi_0\|_{L^p} + \|\Psi - \Psi_0\|_{L^p} \le 2c_3(\varepsilon + \delta^2\varepsilon^{2/p}).$$

Since $\|A - A_0\|_{L^p} \le \delta\varepsilon^{2/p+1/2}$ it follows that

$$\|\Xi - \Xi_0\|_{L^p} \le \delta\varepsilon^{2/p}$$

provided that ε and δ are sufficiently small. Now use Lemma 4.5 again to obtain

$$\begin{aligned}
\|\Xi - \Xi_0\|_{1,p,\varepsilon} &\le c_4\left(\varepsilon\,\|\mathcal{D}_\varepsilon(\Xi - \Xi_\varepsilon)\|_{0,p,\varepsilon} + \|\pi_{A_0}(\mathcal{D}_\varepsilon(\Xi - \Xi_\varepsilon))\|_{L^p}\right) \\
&\le c_5\left(\|\alpha\|_{L^p} + \varepsilon^2\,\|\phi\|_{L^p}\right) \\
&\le c_6\left(\|\Xi - \Xi_0\|_{L^p}\,\|\Xi - \Xi_0\|_{\infty,\varepsilon} + \varepsilon^2\right) \\
&\le c_6\delta\varepsilon^{2/p}\,\|\Xi - \Xi_0\|_{\infty,\varepsilon} + c_6\varepsilon^2 \\
&\le c_7\delta\,\|\Xi - \Xi_0\|_{1,p,\varepsilon} + c_6\varepsilon^2.
\end{aligned}$$

With $c_7\delta < 1$ we obtain (5.11).

Now denote $\Xi_\varepsilon = \mathcal{T}_\varepsilon(\Xi_0)$. Then both Ξ and Ξ_ε satisfy the estimate (5.11) and (5.1). By (5.1) the third component of $\mathcal{D}_\varepsilon(\Xi - \Xi_\varepsilon) = (\alpha, \phi, 0)$ vanishes. The first two are given by

$$\begin{aligned}
\alpha &= [A - A_\varepsilon \wedge \Psi - \Psi_0] + [A_\varepsilon - A_0 \wedge \Psi - \Psi_\varepsilon] \\
&\quad + *_s\,[A - A_\varepsilon \wedge \Phi - \Phi_0] + *_s[A_\varepsilon - A_0 \wedge \Phi - \Phi_\varepsilon] \\
&\quad + *_s\,(X_s(A) - X_s(A_\varepsilon) - \mathrm{d}X_s(A_0)(A - A_\varepsilon)), \\
\phi &= [\Phi - \Phi_\varepsilon \wedge \Psi - \Psi_0] + [\Phi_\varepsilon - \Phi_0 \wedge \Psi - \Psi_\varepsilon] \\
&\quad -\varepsilon^{-2} *_s [\tfrac{1}{2}(A + A_\varepsilon) - A_0 \wedge A - A_\varepsilon].
\end{aligned}$$

Moreover, again by (5.1), $\Xi - \Xi_\varepsilon \in \mathrm{range}\mathcal{D}_\varepsilon^*$. Hence, by Lemma 4.5,

$$\begin{aligned}
\|\Xi - \Xi_\varepsilon\|_{0,p,\varepsilon} &\le c_8\left(\varepsilon\,\|\mathcal{D}_\varepsilon(\Xi - \Xi_\varepsilon)\|_{0,p,\varepsilon} + \|\pi_{A_0}(\mathcal{D}_\varepsilon(\Xi - \Xi_\varepsilon))\|_{L^p}\right) \\
&\le c_9\left(\|\alpha\|_{L^p} + \varepsilon^2\|\phi\|_{L^p}\right) \\
&\le c_{10}\varepsilon^{-1}\left(\|\Xi - \Xi_0\|_{\infty,\varepsilon} + \|\Xi_\varepsilon - \Xi_0\|_{\infty,\varepsilon}\right)\|\Xi - \Xi_\varepsilon\|_{0,p,\varepsilon} \\
&\le c_{11}\varepsilon^{1-2/p}\,\|\Xi - \Xi_\varepsilon\|_{0,p,\varepsilon}\,.
\end{aligned}$$

The third inequality follows by examining α and ϕ term by term. With $c_{11}\varepsilon^{1-2/p} < 1$ this implies $\Xi = \Xi_\varepsilon$. □

Proposition 5.7. *If $\mu_H(a^-, a^+) = 1$ then the map $\mathcal{T}_\varepsilon : \mathcal{M}_0(a^-, a^+, H) \to \mathcal{M}_\varepsilon(a^-, a^+, H)$ of Theorem 5.1 is injective for $\varepsilon > 0$ sufficiently small.*

Proof. Suppose not. Then, since $\mathcal{M}_0(a^-, a^+, H)/\mathbb{R}$ is a finite set, there exist connections $\Xi_0, \Xi_0' \in \mathcal{A}_0^{1,p}(a^-, a^+, H)$ which are not gauge equivalent and a sequence $\varepsilon_\nu \to 0$ such that $\Xi_\nu = \mathcal{T}_{\varepsilon_\nu}(\Xi_0)$ and $\Xi_\nu' = \mathcal{T}_{\varepsilon_\nu}(\Xi_0')$ are gauge equivalent: $\Xi_\nu' = g_\nu^*\Xi_\nu$. The usual compactness argument as in [7, pp.64,65] shows that the sequence g_ν has a weakly converging subsequence and hence Ξ_0 and Ξ_0' must be gauge equivalent. □

The statement of the previous proposition should remain valid without the assumption of index difference 1 but this is not needed for the proof of our main theorem. It remains to show that $\mathcal{T}_\varepsilon$ is onto for small ε when the index difference is 1. The proof will occupy the next four sections.

6. Relative Coulomb gauge

The uniqueness theorem 5.2 requires that $\Xi - \Xi_0$ be in the kernel of the operator $d_{\Xi_0}^{*_\varepsilon}$ and in the range of the operator $\mathcal{D}_\varepsilon^* = \mathcal{D}_\varepsilon(\Xi_0)^*$. The first condition can be achieved by a suitable gauge transformation and the second by a suitable time shift provided that the relative Morse index is 1. Given $\tau \in \mathbb{R}$ denote

$$\Xi \circ \sigma_\tau(s,t) = \Xi(s, t+\tau).$$

We shall prove the following

Theorem 6.1. *Let $\Xi_0 \in \mathcal{A}_0^{1,p}(a^-, a^+, H)$ with $\mu(a^-, a^+) = 1$. Then there exist constants $\varepsilon_0 > 0$ and $\delta > 0$ such that the following holds. If $0 < \varepsilon \le \varepsilon_0$ and $\Xi \in \mathcal{A}_\varepsilon^{1,p}(a^-, a^+, H)$ such that*

$$\|\Xi - \Xi_0\|_{1,p,\varepsilon} \le \delta \varepsilon^{2/p+1/2}$$

then there exist $\tau \in \mathbb{R}$ and $g \in \mathcal{G}^{2,p}$ such that $g^*(\Xi \circ \sigma_\tau) = \mathcal{T}_\varepsilon(\Xi_0)$.

This result is an immediate consequence of Theorem 5.2 and the next two propositions.

Proposition 6.2. *Assume $q \ge p > 2$, $q > 4$, and $qp/(q-p) > 4$. Let $\Xi_0 \in \mathcal{A}^{1,p}(a^-, a^+)$ such that $F_{A(s,t)} = 0$ for all s and t. Then for every constant $c_0 > 0$ there exist constants $\delta > 0$ and $c > 0$ such that the following holds for $0 < \varepsilon \le 1$. If $\Xi \in \mathcal{A}^{1,p}(a^-, a^+)$ satisfies*

$$\left\| \mathrm{d}_{\Xi_0}^{*_\varepsilon}(\Xi - \Xi_0) \right\|_{\mathrm{L}^p} \le c_0 \varepsilon^{2/p}, \qquad \|\Xi - \Xi_0\|_{0,q,\varepsilon} \le \delta \varepsilon^{2/q},$$

then there exists a gauge transformation $g \in \mathcal{G}^{2,p}$ such that $\mathrm{d}_{\Xi_0}^{*_\varepsilon}(g^*\Xi - \Xi_0) = 0$ and

$$\|g^*\Xi - \Xi\|_{1,p,\varepsilon} \le c\left(1 + \varepsilon^{-2/p}\|\Xi - \Xi_0\|_{1,p,\varepsilon}\right) \left\| \mathrm{d}_{\Xi_0}^{*_\varepsilon}(\Xi - \Xi_0) \right\|_{\mathrm{L}^p}.$$

Proposition 6.3. *Let $\Xi_0 \in \mathcal{A}_0^{1,p}(a^-, a^+, H)$ with $\mu(a^-, a^+) = 1$. Then there exist constants $\varepsilon_0 > 0$, $\delta > 0$, and $c > 0$ such that the following holds. If $0 < \varepsilon < \varepsilon_0$ and $\Xi \in \mathcal{A}^{1,p}(a^-, a^+)$ such that*

$$\|\Xi - \Xi_0\|_{1,p,\varepsilon} \leq \delta\varepsilon^{1/p+1/2} \tag{6.1}$$

then there exist $\tau \in \mathbb{R}$ and $g \in \mathcal{G}^{2,p}$ such that $\Xi_\varepsilon = g^*(\Xi \circ \sigma_\tau)$ satisfies (5.1) and $\|\Xi_\varepsilon - \Xi_0\|_{1,p,\varepsilon} \leq c\,\|\Xi - \Xi_0\|_{1,p,\varepsilon}$.

Note that the assumptions of Proposition 6.2 are satisfied whenever $\|\Xi - \Xi_0\|_{1,p,\varepsilon} \leq \delta\varepsilon^{2/p}$ with $\delta > 0$ sufficiently small. The proof of both propositions relies on the following three lemmata.

Lemma 6.4. *Let $\xi \in \mathrm{W}_f^{1,p}$ and $0 < \varepsilon \leq 1$. Then there exists a unique $\eta \in \mathrm{W}_f^{2,p}(\mathbb{R}^2 \times \mathfrak{g}_P)$ such that*

$$\mathrm{d}_{\Xi_0}^{*_\varepsilon}\mathrm{d}_{\Xi_0}\eta = \mathrm{d}_{\Xi_0}^{*_\varepsilon}\xi.$$

This solution satisfies estimates

$$\|\eta\|_{2,p,\varepsilon} \leq c\,\|\mathrm{d}_{\Xi_0}^*\xi\|_{\mathrm{L}^p}, \qquad \|\eta\|_{1,q,\varepsilon} \leq c\,\|\xi\|_{0,q,\varepsilon}$$

for $0 < \varepsilon \leq 1$ where the constant $c = c(p) > 0$ is independent of ε.

Proof. For $p = 2$ the second estimate follows from $\|d_{\Xi_0}\eta\|_{0,2,\varepsilon}^2 = \langle d_{\Xi_0}\eta, \xi\rangle_\varepsilon$. For general p both estimates follow by rescaling as in Lemma 4.1. The first estimate shows that the operator $\mathrm{d}_{\Xi_0}^{*_\varepsilon}\mathrm{d}_{\Xi_0} : \mathrm{W}_f^{2,p}(\mathbb{R}^2\times\mathfrak{g}_P) \to \mathrm{L}_f^p(\mathbb{R}^2\times\mathfrak{g}_P)$ is injective and has a closed range. Its cokernel is the kernel of $\mathrm{d}_{\Xi_0} : \mathrm{W}_f^{1,r}(\mathbb{R}^2\times\mathfrak{g}_P) \to \mathrm{L}_f^r$ with $1/p+1/r = 0$.[3] Hence the aforementioned operator $\mathrm{d}_{\Xi_0}^{*_\varepsilon}\mathrm{d}_{\Xi_0}$ is bijective and this proves the existence statement. □

Given $\eta \in \mathrm{W}_f^{1,q}(\mathbb{R}^2 \times \mathfrak{g}_P)$ and $\xi \in \mathrm{W}_f^{1,p}$ denote $\mathrm{ad}(\eta)\xi = [\eta \wedge \xi]$.

Lemma 6.5. *Assume $q \geq p > 2$, $q > 4$, and $qp/(q-p) > 4$. Then there exists a constant $c > 0$ such that*

$$\left\|\mathrm{ad}(\eta)^k\xi\right\|_{1,p,\varepsilon} \leq c^k\varepsilon^{-2/q}\,\|\eta\|_{\mathrm{L}^\infty}^{k-1}\left(\|\xi\|_{0,q,\varepsilon}\|\eta\|_{2,p,\varepsilon} + \|\xi\|_{1,p,\varepsilon}\|\eta\|_{1,q,\varepsilon}\right),$$

$$\left\|\mathrm{d}_{\Xi_0}^{*_\varepsilon}\mathrm{ad}(\eta)^k\xi\right\|_{\mathrm{L}^p} \leq c^k\varepsilon^{-2/q}\,\|\eta\|_{\mathrm{L}^\infty}^{k-1}\left(\|\xi\|_{0,q,\varepsilon}\|\eta\|_{2,p,\varepsilon} + \left\|\mathrm{d}_{\Xi_0}^{*_\varepsilon}\xi\right\|_{\mathrm{L}^p}\|\eta\|_{1,q,\varepsilon}\right)$$

for $\xi \in \mathrm{W}_f^{1,p}$, $\eta \in \mathrm{W}_f^{2,p}(\mathbb{R}^2 \times \mathfrak{g}_P)$, $0 < \varepsilon \leq 1$, and $k = 1, 2, 3, \ldots$.

Proof. To prove the first estimate for $k = 1$ note that

$$\|[\eta \wedge \xi]\|_{1,p,\varepsilon} \leq c_1\left(\|\xi\|_{0,q,\varepsilon}\|\eta\|_{1,r,\varepsilon} + \|\xi\|_{1,p,\varepsilon}\|\eta\|_{\mathrm{L}^\infty}\right)$$

[3] Since the kernel consists of smooth sections the choice of the Sobolev norm is not important.

where $1/q + 1/r = 1/p$. Since $q > 4$ we have $r = qp/(q-p) < 4p/(4-p)$. Hence there are inclusions $\mathrm{W}^{1,q} \hookrightarrow \mathrm{L}^\infty$ and $W^{2,p} \hookrightarrow \mathrm{W}^{1,r}$ and the estimate follows from Lemma 4.1. To prove the second estimate for $k = 1$ note that

$$\mathrm{d}_{\Xi_0}^{*_\varepsilon}[\eta \wedge \xi] = [\eta, \mathrm{d}_{\Xi_0}^{*_\varepsilon}\xi] - *_s[d_{A_0}\eta \wedge *_s\alpha] - \varepsilon^2[\nabla_s\eta, \phi] - \varepsilon^2[\nabla_t\eta, \psi]$$

for $\xi = \alpha + \phi\,\mathrm{d}s + \psi\,\mathrm{d}t$. Hence

$$\left\|\mathrm{d}_{\Xi_0}^{*_\varepsilon}[\eta \wedge \xi]\right\|_{\mathrm{L}^p} \le c_2\left(\|\xi\|_{0,q,\varepsilon}\|\eta\|_{1,r,\varepsilon} + \left\|\mathrm{d}_{\Xi_0}^{*_\varepsilon}\xi\right\|_{\mathrm{L}^p}\|\eta\|_{\mathrm{L}^\infty}\right).$$

Now for general k both estimates follow by induction. □

Lemma 6.6. *Assume $q \ge p > 2$, $q > 4$, and $pq/(q-p) > 4$. Given $c_0 > 0$ there exists a constant $c > 0$ a such that, if $\|\eta\|_{\mathrm{L}^\infty} \le c_0$ and $g = \exp(\eta)$, then*

$$\begin{aligned}\left\|\mathrm{d}_{\Xi_0}^{*_\varepsilon}(g^*\Xi - \Xi - d_\Xi\eta)\right\|_{\mathrm{L}^p} \quad &\le \quad c\varepsilon^{-2/q}\left(\|\eta\|_{1,q,\varepsilon} + \|\Xi - \Xi_0\|_{0,q,\varepsilon}\right)\|\eta\|_{2,p,\varepsilon} \\ &\quad + c\varepsilon^{-2/q}\left\|\mathrm{d}_{\Xi_0}^{*_\varepsilon}(\Xi - \Xi_0)\right\|_{\mathrm{L}^p}\|\eta\|_{1,q,\varepsilon},\end{aligned}$$

and if $\|\eta\|_{1,q,\varepsilon} + \|\Xi - \Xi_0\|_{0,q,\varepsilon} \le c_0\varepsilon^{2/q}$, then

$$\|g^*\Xi - \Xi\|_{0,q,\varepsilon} \le c\|\eta\|_{1,q,\varepsilon},$$

$$\|g^*\Xi - \Xi\|_{1,p,\varepsilon} \le c\left(\|\eta\|_{2,p,\varepsilon} + \varepsilon^{-2/q}\|\Xi - \Xi_0\|_{1,p,\varepsilon}\|\eta\|_{1,q,\varepsilon}\right).$$

Proof. Use Lemma 6.5 and the identity

$$g^*\Xi - \Xi = \sum_{k=0}^{\infty}\frac{(-1)^k}{(k+1)!}\mathrm{ad}(\eta)^k\mathrm{d}_\Xi\eta$$

for $g = \exp(\eta)$. □

Proof of Proposition 6.2. The proof is based on a Newton type iteration. Denote $\Xi_1 = \Xi$ and for $\nu \ge 2$ define Ξ_ν inductively by

$$\Xi_{\nu+1} = g_\nu^*\Xi_\nu, \qquad g_\nu = \exp(\eta_\nu), \qquad d_{\Xi_0}^{*_\varepsilon}(d_{\Xi_0}\eta_\nu + \Xi_\nu - \Xi_0) = 0.$$

By Lemma 6.4 $\eta_\nu \in \mathrm{W}_f^{2,p}(\mathbb{R}^2 \times \mathfrak{g}_P)$ satisfies estimates

$$\begin{aligned}\|\eta_\nu\|_{2,p,\varepsilon} + \varepsilon^{2/p-2/q}\|\eta_\nu\|_{1,q,\varepsilon} \quad &\le \quad c_1\left\|\mathrm{d}_{\Xi_0}^{*}(\Xi_\nu - \Xi_0)\right\|_{\mathrm{L}^p}, \\ \|\eta_\nu\|_{1,q,\varepsilon} \quad &\le \quad c_1\|\Xi_\nu - \Xi_0\|_{0,q,\varepsilon}.\end{aligned} \tag{6.2}$$

We shall prove by induction that there exist constants c_2, c_3, and c_4 such that

$$\|\Xi_\nu - \Xi_0\|_{0,q,\varepsilon} \le c_2\|\Xi - \Xi_0\|_{0,q,\varepsilon}, \tag{6.3}$$

$$\left\|\mathrm{d}_{\Xi_0}^{*_\varepsilon}(\Xi_\nu - \Xi_0)\right\|_{\mathrm{L}^p} \le c_3\varepsilon^{-2/q}\|\Xi_{\nu-1} - \Xi_0\|_{0,q,\varepsilon}\left\|\mathrm{d}_{\Xi_0}^{*_\varepsilon}(\Xi_{\nu-1} - \Xi_0)\right\|_{\mathrm{L}^p}, \tag{6.4}$$

$$\left\|\mathrm{d}_{\Xi_0}^{*_\varepsilon}(\Xi_\nu - \Xi_0)\right\|_{\mathrm{L}^p} \le 2^{1-\nu}\left\|\mathrm{d}_{\Xi_0}^{*_\varepsilon}(\Xi - \Xi_0)\right\|_{\mathrm{L}^p}, \tag{6.5}$$

$$\|\eta_\nu\|_{1,q,\varepsilon} \le c_4 2^{-\nu} \|\Xi - \Xi_0\|_{0,q,\varepsilon}. \tag{6.6}$$

For $\nu = 1$ the inequalities (6.3) and (6.5) are obvious, (6.6) follows from (6.2), and (6.4) is empty. For $\nu \ge 2$ it follows from the previous induction steps with δ sufficiently small that $\|\eta_j\|_{1,q,\varepsilon} + \|\Xi_j - \Xi_0\|_{0,q,\varepsilon} \le \varepsilon^{2/q}$ for $j = 1, \cdots, \nu - 1$. Hence, by Lemma 6.6, with a suitable constant $c_5 > 0$

$$\|\Xi_{j+1} - \Xi_j\|_{0,q,\varepsilon} \le c_5 \|\eta_j\|_{1,q,\varepsilon} \le c_4 c_5 2^{-j} \|\Xi - \Xi_0\|_{0,q,\varepsilon}.$$

Here we have used $c_2\delta \le 1$. Now use this inequality for $j = 1, \ldots, \nu - 1$ to obtain (6.3) with $c_2 = 1 + c_4 c_5$. To prove the estimate (6.4) note that, by the previous induction step, $\|\eta_{\nu-1}\|_{L^\infty} \le 1$ provided that δ is sufficiently small. Moreover,

$$\mathrm{d}_{\Xi_0}^{*_\varepsilon}(\Xi_{\nu+1} - \Xi_0) = d_{\Xi_0}^{*_\varepsilon}(g_\nu^*\Xi_\nu - \Xi_\nu - \mathrm{d}_{\Xi_\nu}\eta_\nu) + d_{\Xi_0}^{*_\varepsilon}[\Xi_\nu - \Xi_0 \wedge \eta_\nu]$$

and hence (6.4) follows from the Lemmata 6.5 and 6.6. Now (6.5) follows immediately from (6.4) and (6.3) with $c_2 c_3 \delta \le 1/2$. Finally, we prove (6.6). For $\nu = 2$ it follows from (6.2) and (6.4) that

$$\|\eta_2\|_{1,q,\varepsilon} \le c_0 c_1 c_3 \|\Xi - \Xi_0\|_{0,q,\varepsilon}.$$

Hence in this case (6.6) holds with $c_4 = 4c_0c_1c_3$. For $\nu \ge 3$ we can use (6.4) twice and one checks easily that

$$\|\eta_\nu\|_{1,q,\varepsilon} \le 8 c_0 c_1 c_2^2 c_3^2 \delta 2^{-\nu} \|\Xi - \Xi_0\|_{0,q,\varepsilon}.$$

So in this case (6.6) holds with $c_4 = 1$ provided that δ is sufficiently small. This completes the induction. Note that the conditions $c_4 \ge 4c_0c_1c_3$ and $c_2 = 1 + c_4c_5$ are compatible.

Now it follows from Lemma 6.6 that

$$\|\Xi_{\nu+1} - \Xi_\nu\|_{1,p,\varepsilon} \le c_6 \left(1 + \varepsilon^{-2/p} \|\Xi_\nu - \Xi_0\|_{1,p,\varepsilon}\right) \left\|\mathrm{d}_{\Xi_0}^{*_\varepsilon}(\Xi_\nu - \Xi_0)\right\|_{L^p}$$

and, by induction,

$$\|\Xi_\nu - \Xi_0\|_{1,p,\varepsilon} \le \varepsilon^{2/p} + 2\|\Xi - \Xi_0\|_{1,p,\varepsilon}$$

provided that δ is sufficiently small. Hence, by (6.5) the sequence Ξ_ν converges in $\mathcal{A}^{1,p}(a^-, a^+)$ and the limit connection $\Xi_\varepsilon = \lim_{\nu\to\infty}$ satisfies $\mathrm{d}_{\Xi_0}^{*_\varepsilon}(\Xi_\varepsilon - \Xi_0) = 0$ and the required estimate. Moreover $\Xi_\nu = h_\nu^*\Xi$ where $h_\nu = g_1 g_2 \cdots g_\nu$ converges in $\mathcal{G}^{2,p}$. This proves the proposition. □

Proof of Proposition 6.3. Let $\Xi \in \mathcal{A}^{1,p}(a^-, a^+)$ satisfy (6.1) with ε and δ sufficiently small. For every τ we have

$$\begin{aligned} \|\Xi \circ \sigma_\tau - \Xi_0\|_{1,p,\varepsilon} &\le \|\Xi - \Xi_0\|_{1,p,\varepsilon} + \|\Xi_0 \circ \sigma_\tau - \Xi_0\|_{1,p,\varepsilon} \\ &\le \|\Xi - \Xi_0\|_{1,p,\varepsilon} + |\tau| \, \|\partial_t \Xi_0\|_{1,p,\varepsilon}. \end{aligned}$$

Hence it follows from Proposition 6.2 that for $|\tau| \le \delta\varepsilon^{2/p}$ there exists a gauge transformation $g_\tau \in \mathcal{G}^{2,p}$ such that $\Xi_\tau = g_\tau^*(\Xi \circ \sigma_\tau)$ satisfies

$$d_{\Xi_0}^{*_\varepsilon}(\Xi_\tau - \Xi_0) = 0.$$

Moreover, we have

$$\|\Xi_\tau - \Xi_0\|_{1,p,\varepsilon} \le c_1 \left(|\tau| + \|\Xi - \Xi_0\|_{1,p,\varepsilon} \right) \tag{6.7}$$

with a suitable constant $c_1 > 0$. Assume without loss of generality that $g_0 = 1\!\!1$.

We shall prove that there exists a number τ such that

$$g_\tau^*(\Xi \circ \sigma_\tau) - \Xi_0 \in \mathrm{range}\mathcal{D}_\varepsilon^*, \qquad |\tau| \le c_2 \|\Xi - \Xi_0\|_{1,p,\varepsilon}. \tag{6.8}$$

To see this note that the operator $\mathcal{D}_0 = \mathcal{D}_0(\Xi_0)$ is onto and of index 1. Hence its kernel is spanned by

$$\xi_0 = \partial_t \Xi_0 \in \mathrm{W}_f^{1,p}.$$

By Theorem 3.2 the operator $\mathcal{D}_\varepsilon$ has also index 1 and by Lemma 4.5 it is onto. Its kernel is spanned by the vector

$$\xi_\varepsilon = \xi_0 - \mathcal{D}_\varepsilon^*(\mathcal{D}_\varepsilon\mathcal{D}_\varepsilon^*)^{-1}\mathcal{D}_\varepsilon\xi_0.$$

The harmonic part of $\mathcal{D}_\varepsilon\xi_0$ vanishes and hence, by Lemma 4.5,

$$\begin{aligned}\|\xi_\varepsilon - \xi_0\|_{1,p,\varepsilon} &\le c_1\varepsilon \|\mathcal{D}_\varepsilon\xi_0\|_{0,p,\varepsilon} \\ &= c_1\varepsilon^2 \left(\|\nabla_t\phi_0 - \nabla_s\psi_0\|_{\mathrm{L}^p}^p + \|\nabla_t\psi_0 + *_s\nabla_s *_s \phi_0\|_{\mathrm{L}^p}^p \right)^{1/p} \\ &\le c_2\varepsilon^2 \|\pi_A(\xi_0)\|_{L^p}.\end{aligned}$$

The last inequality follows from the basic regularity estimate for $\mathcal{D}_0$.

Now consider the function

$$\theta(\tau) = \theta_{\varepsilon,\Xi}(\tau) = \langle \xi_\varepsilon, \Xi_\tau - \Xi_0 \rangle_\varepsilon$$

where the expression $\langle\, , \rangle_\varepsilon$ abbreviates the ε-pairing between L_f^q and L_f^p with $1/p + 1/q = 1$. Then equation (6.8) can be written as $\theta(\tau) = 0$. We shall prove that there exist constants $\delta_0 > 0$, $\varepsilon_0 > 0$, and $\rho_0 > 0$ such that

$$|\tau| + \|\Xi - \Xi_0\|_{1,p,\varepsilon} \le \delta_0\varepsilon^{1/p+1/2}, \quad 0 < \varepsilon < \varepsilon_0 \qquad \Longrightarrow \qquad \theta'(\tau) \ge \rho_0. \tag{6.9}$$

Then the existence of a zero follows from the fact that

$$|\theta(0)| = |\langle \xi_\varepsilon, \Xi - \Xi_0 \rangle_\varepsilon| \le \|\xi_\varepsilon\|_{0,q,\varepsilon} \|\Xi - \Xi_0\|_{0,p,\varepsilon} \le c_3\delta\varepsilon^{1/p+1/2}.$$

In fact, if $c_3\delta < \frac{1}{2}\delta_0\rho_0$ and $\delta \le \frac{1}{2}\delta_0$ then $\|\Xi - \Xi_0\|_{1,p,\varepsilon} \le \frac{1}{2}\delta_0\varepsilon^{1/p+1/2}$ and, by (6.9), there exists a number $\tau \in \mathbb{R}$ with $|\tau| \le |\theta(0)|/\rho_0 \le \frac{1}{2}\delta_0\varepsilon^{1/p+1/2}$ such that $\theta(\tau) = 0$. This number τ satisfies (6.8) as required.

To prove (6.9) define

$$\eta_\tau = {g_\tau}^{-1}\left(\partial_\tau g_\tau - \partial_t g_\tau\right)$$

and observe that

$$\theta'(\tau) = \langle \xi_\varepsilon, \partial_t \Xi_\tau + d_{\Xi_\tau}\eta_\tau \rangle_\varepsilon.$$

Now differentiate the identity $d_{\Xi_0}^{*_\varepsilon}(\Xi_\tau - \Xi_0) = 0$ with respect to τ to obtain

$$d_{\Xi_0}^{*_\varepsilon} d_{\Xi_0}\eta_\tau + d_{\Xi_0}^{*_\varepsilon}[\Xi_\tau - \Xi_0 \wedge \eta_\tau] + d_{\Xi_0}^{*_\varepsilon}\partial_t \Xi_\tau = 0.$$

By Lemma 6.4 this equation has a unique solution $\eta_\tau \in \mathrm{W}_f^{2,p}(\mathbb{R}^2 \times \mathfrak{g}_P)$ whenever $\varepsilon^{-2/p}\left\|\Xi_\tau - \Xi_0\right\|_{1,p,\varepsilon}$ is sufficiently small. Moreover, η_τ satisfies

$$\left\|\eta_\tau\right\|_{1,p,\varepsilon} \le c_4 \left\|\partial_t \Xi_\tau\right\|_{0,p,\varepsilon} \le c_5\left(1 + \varepsilon^{-1}\left\|\Xi_\tau - \Xi_0\right\|_{1,p,\varepsilon}\right).$$

Since $d_{\Xi_0}^{*_\varepsilon}\xi_\varepsilon = 0$ we obtain

$$\begin{aligned}
|\langle \xi_\varepsilon, d_{\Xi_\tau}\eta_\tau\rangle_\varepsilon| &= |\langle \xi_\varepsilon, [\Xi_\tau - \Xi_0 \wedge \eta_\tau]\rangle_\varepsilon| \\
&\le c_6 \left\|\Xi_\tau - \Xi_0\right\|_{\infty,\varepsilon} \left\|\eta_\tau\right\|_{0,p,\varepsilon} \\
&\le c_7 \varepsilon^{-2/p}\left\|\Xi_\tau - \Xi_0\right\|_{1,p,\varepsilon}\left\|\eta_\tau\right\|_{1,p,\varepsilon} \\
&\le c_5 c_7 \varepsilon^{-2/p}\left\|\Xi_\tau - \Xi_0\right\|_{1,p,\varepsilon}\left(1 + \varepsilon^{-1}\left\|\Xi_\tau - \Xi_0\right\|_{1,p,\varepsilon}\right) \\
&\le c_8 \delta_0.
\end{aligned}$$

In the last inequality we have used the fact that $\varepsilon > 0$ is sufficiently small and, by (6.7), $\left\|\Xi_\tau - \Xi_0\right\|_{1,p,\varepsilon} \le c_1(|\tau| + \left\|\Xi - \Xi_0\right\|_{1,p,\varepsilon}) \le c_1\delta_0\varepsilon^{1/p+1/2}$. Moreover,

$$\langle \xi_\varepsilon, \partial_t \Xi_\tau\rangle_\varepsilon = \langle \partial_t \xi_\varepsilon, \Xi_0 - \Xi_\tau\rangle_\varepsilon + \langle \xi_\varepsilon, \partial_t \Xi_0\rangle_\varepsilon$$

Since $\partial_t A_0 \ne 0$ we have

$$\left\|\partial_t \Xi_0\right\|_{0,2,\varepsilon} \ge 3\rho_0 > 0$$

for some constant $\rho_0 > 0$ and hence

$$\langle \xi_\varepsilon, \partial_t \Xi_0\rangle_\varepsilon \ge 2\rho_0$$

for $\varepsilon > 0$ sufficiently small. This implies

$$\langle \xi_\varepsilon, \partial_t \Xi_\tau + \mathrm{d}_{\Xi_\tau}\eta_\tau\rangle_\varepsilon > \rho_0$$

for $|\tau| + \left\|\Xi - \Xi_0\right\|_{1,p,\varepsilon} < \delta_0\varepsilon^{1/p+1/2}$ provided δ_0 and ε are sufficiently small. Thus we have proved (6.9) and this finishes the proof of the proposition. □

7. Estimates on the curvature

This section is of a preparatory nature. We prove estimates on the derivatives of the curvature for ε-self-dual connections with bounded curvature. We also establish uniform exponential decay of the curvature as t tends to $\pm\infty$. These results are used in the next section to prove a compactness theorem for ε-self-dual connections with ε converging to zero.

The curvature of the connection $\Xi = A + \Phi\, ds + \Psi\, dt$ is given by

$$F_\Xi = F_A - B_s\, ds - B_t\, dt - C\, ds \wedge dt$$

where

$$B_s = \partial_s A - \mathrm{d}_A \Phi, \qquad B_t = \partial_t A - \mathrm{d}_A \Psi, \qquad C = \partial_t \Phi - \partial_s \Psi - [\Phi, \Psi]. \tag{7.1}$$

The Bianchi identity takes the form

$$\nabla_s F_A = \mathrm{d}_A B_s, \qquad \nabla_t F_A = \mathrm{d}_A B_t, \qquad \nabla_s B_t - \nabla_t B_s = \mathrm{d}_A C \tag{7.2}$$

where $\nabla_s = \partial_s + \Phi$ and $\nabla_t = \partial_t + \Psi$. The curvature terms B_s, B_t, and C also appear as commutators

$$\nabla_s \mathrm{d}_A - \mathrm{d}_A \nabla_s = B_s, \qquad \nabla_t \mathrm{d}_A - \mathrm{d}_A \nabla_t = B_t, \qquad \nabla_t \nabla_s - \nabla_s \nabla_t = C.$$

The perturbed self-duality equation (3.5) can be written in the form

$$B_t + *_s(B_s - X_s(A)) = 0, \qquad C + \varepsilon^{-2} *_s F_A = 0. \tag{7.3}$$

If Ξ satisfies these equations with $X_s = 0$ then F_Ξ is harmonic with respect to the ε-dependent Laplacian. This implies the following estimate for connections with L^∞ bounds on the curvature.

THEOREM 7.1. *Let $\Omega \subset \mathbb{C}$ be an open set, $Q \subset \Omega$ be a compact subset, and $c_0 > 0$. Then there exist constants $c > 0$ and $\varepsilon_0 > 0$ such that the following holds. If $\Xi = A + \Phi\, ds + \Psi\, dt$ satisfies (7.3) for $s + it \in \Omega$ with $0 < \varepsilon \le \varepsilon_0$ and*

$$\|B_t\|_{\mathrm{L}^\infty(\Omega\times\Sigma)} + \varepsilon\, \|C\|_{\mathrm{L}^\infty(\Omega\times\Sigma)} \le c_0$$

then for $2 \le p \le \infty$

$$\begin{aligned}
&\varepsilon^{2/p}\, \|B_t\|_{\mathrm{L}^p(Q\times\Sigma)} + \|\mathrm{d}_A B_t\|_{\mathrm{L}^p(Q\times\Sigma)} + \|\mathrm{d}_A *_s B_t\|_{\mathrm{L}^p(Q\times\Sigma)} \\
&\quad + \varepsilon\, \|\nabla_s B_t\|_{\mathrm{L}^p(Q\times\Sigma)} + \varepsilon\, \|\nabla_t B_t\|_{\mathrm{L}^p(Q\times\Sigma)} \\
&\quad + \varepsilon\, \|\mathrm{d}_A C\|_{\mathrm{L}^p(Q\times\Sigma)} + \varepsilon^2\, \|\nabla_s C\|_{\mathrm{L}^p(Q\times\Sigma)} + \varepsilon^2\, \|\nabla_t C\|_{\mathrm{L}^p(Q\times\Sigma)} \\
&\quad \le c\varepsilon^{2/p} \left(\|B_t\|_{\mathrm{L}^2(\Omega\times\Sigma)} + \varepsilon\, \|C\|_{\mathrm{L}^2(\Omega\times\Sigma)} \right).
\end{aligned} \tag{7.4}$$

Remark 7.2. The estimate (7.4) is standard for $\varepsilon = 1$ (or if the constant is allowed to depend on ε): the $W^{2,p}$-norm of a harmonic function on a compact set can be estimated above by the L^p-norm on a neighbourhood of this set. For the standard Laplacian this follows from the mean value property.

Proof of Theorem 7.1. Consider the nonnegative function $u_0 : \Omega \to \mathbb{R}$ defined by

$$u_0(s,t) = \tfrac{1}{2}\left(\|B_t(s,t)\|^2_{L^2(\Sigma,*_s)} + \varepsilon^2 \|C(s,t)\|^2_{L^2(\Sigma,*_s)}\right).$$

The Laplacian of u_0 is given by

$$\begin{aligned}\Delta u_0 &= \|\nabla_t B_t\|^2 + \|\nabla_s B_t\|^2 + \varepsilon^2\|\nabla_t C\|^2 + \varepsilon^2\|\nabla_s C\|^2 \\ &\quad + \langle \nabla_t\nabla_t B_t + \nabla_s\nabla_s B_t, B_t\rangle + \varepsilon^2\langle \nabla_t\nabla_t C + *_s\nabla_s\nabla_s *_s C, C\rangle \\ &\quad -2\langle \nabla_s B_t, *_s\dot{*}_s B_t\rangle - \tfrac{1}{2}\langle B_t, *_s\ddot{*}_s B_t\rangle - \tfrac{1}{2}\varepsilon^2\langle C, *_s\ddot{*}_s C\rangle.\end{aligned}$$

Here all norms and inner products are L^2-norms and L^2-inner products on Σ induced by the $*_s$-metric. Now we have

$$\begin{aligned}\varepsilon^2\nabla_s *_s C &= -d_A *_s B_t, \\ \varepsilon^2\nabla_t C &= - *_s d_A B_t, \\ \nabla_t B_t + *_s\nabla_s B_t &= *_s dX_s(A)B_t + *_s d_A C,\end{aligned}$$

and

$$\begin{aligned}\varepsilon^2\left(\nabla_t\nabla_t C + *_s\nabla_s\nabla_s *_s C\right) &= - *_s d_A *_s d_A C \\ &\quad -2 *_s [B_t\wedge B_t] + *_s[*_s X_s(A)\wedge B_t] \\ &\quad - *_s d_A *_s dX_s(A)B_t - *_s d_A\dot{*}_s B_t, \\ \nabla_t\nabla_t B_t + \nabla_s\nabla_s B_t &= -\varepsilon^{-2} *_s d_A *_s d_A B_t - \varepsilon^{-2} d_A *_s d_A *_s B_t \\ &\quad +3 *_s [B_t\wedge C] + [X_s(A)\wedge C] \\ &\quad + dX_s(A)\nabla_s B_t + *_s dX_s(A)\nabla_t B_t \\ &\quad + d\dot{X}_s(A)B_t + \dot{*}_s\nabla_t B_t - d_A *_s \dot{*}_s C \\ &\quad + d^2X_s(A)(*_s B_t + X_s(A), B_t) \\ &\quad + *_s d^2X_s(A)(B_t, B_t).\end{aligned}$$

Hence $\Delta u_0 = 2v_0 + f_0$ where

$$\begin{aligned}v_0 &= \tfrac{1}{2}\Big(\varepsilon^{-2}\|d_A B_t\|^2 + \varepsilon^{-2}\|d_A *_s B_t\|^2 \\ &\quad + \|\nabla_s B_t\|^2 + \|\nabla_t B_t\|^2 + \|d_A C\|^2 \\ &\quad + \varepsilon^2\|\nabla_s C\|^2 + \varepsilon^2\|\nabla_t C\|^2\Big)\end{aligned}$$

and

$$\begin{aligned} f_0 &= 5\langle B_t, *_s[B_t \wedge C]\rangle + 2\langle B_t, [X_s(A) \wedge C]\rangle \\ &\quad +\langle \mathrm{d}_A C, \mathrm{d}X_s(A)B_t\rangle - \langle \mathrm{d}_A C, *_s\dot{*}_s B_t\rangle + \langle \mathrm{d}_A *_s B_t, \dot{*}_s C\rangle \\ &\quad -2\langle \nabla_s B_t, *_s\dot{*}_s B_t\rangle - \tfrac{1}{2}\langle B_t, *_s\ddot{*}_s B_t\rangle - \tfrac{1}{2}\varepsilon^2\langle C, *_s\ddot{*}_s C\rangle \\ &\quad +\langle B_t, \mathrm{d}X_s(A)\nabla_s B_t + *_s\mathrm{d}X_s(A)\nabla_t B_t\rangle + \langle B_t, \mathrm{d}\dot{X}_s(A)B_t + \dot{*}_s\nabla_t B_t\rangle \\ &\quad +\langle B_t, \mathrm{d}^2X_s(A)(*_s B_t + X_s(A), B_t) + *_s\mathrm{d}^2X_s(A)(B_t, B_t)\rangle. \end{aligned}$$

It follows from the L^∞ estimate on the curvature that $|f_0| \le v_0 + c_1 u_0$ with a suitable constant $c_1 > 0$. In particular, $\|F_A\|_{\mathrm{L}^\infty(\Sigma)} \le \varepsilon c_0$ and $X_s(A)$, $\mathrm{d}X_s(A)$ and $\mathrm{d}^2X_s(A)$ are uniformly bounded in this domain. Hence

$$\Delta u_0 \ge v_0 - c_1 u_0 \tag{7.5}$$

and, by Lemma 7.3 below,

$$\sup_Q u_0 + \int_Q v_0 \le c_2 \int_\Omega u_0$$

with a suitable constant $c_2 > 0$. This proves the proposition for $p = 2$. We shall now prove the estimate for $p = \infty$ and then the general case will follow by interpolation.

Consider the functions $u_1, v_1 : \Omega \to \mathbb{R}$ defined by

$$u_1 = \tfrac{1}{2}\left(\|\nabla_s B_t\|^2 + \|\nabla_t B_t\|^2 + \varepsilon^2\|\nabla_s C\|^2 + \varepsilon^2\|\nabla_t C\|^2\right)$$

and

$$\begin{aligned} v_1 &= \tfrac{1}{2}\Big(\varepsilon^{-2}\|\mathrm{d}_A\nabla_t B_t\|^2 + \varepsilon^{-2}\|\mathrm{d}_A *_s \nabla_t B_t\|^2 \\ &\quad +\varepsilon^{-2}\|\mathrm{d}_A\nabla_s B_t\|^2 + \varepsilon^{-2}\|\mathrm{d}_A *_s \nabla_s B_t\|^2 \\ &\quad +\|\mathrm{d}_A\nabla_s C\|^2 + \|\mathrm{d}_A\nabla_t C\|^2 \\ &\quad +\|\nabla_s\nabla_s B_t\|^2 + \|\nabla_t\nabla_t B_t\|^2 + \|\nabla_s\nabla_t B_t\|^2 + \|\nabla_t\nabla_s B_t\|^2 \\ &\quad +\varepsilon^2\|\nabla_s\nabla_s C\|^2 + \varepsilon^2\|\nabla_t\nabla_t C\|^2 + 2\varepsilon^2\|\nabla_s\nabla_t C\|^2\Big). \end{aligned}$$

Here all norms are L^2-norms on Σ with respect to the s-metric. To simplify the formulae we shall now restrict ourselves to the case where the Hodge-$*$-operator $*_s = *$ is independent of t and the perturbation $X_s \equiv 0$. Then we have

$$\Delta u_1 = 2v_1 + f_1$$

where

$$\begin{aligned} f_1 &= \varepsilon^{-2}\langle \mathrm{d}_A * \mathrm{d}_A C, [B_t \wedge B_t]\rangle - 4\varepsilon^2\langle C, [\nabla_s C, \nabla_t C]\rangle \\ &\quad -3\langle C, *[\mathrm{d}_A C \wedge \mathrm{d}_A C]\rangle + 10\langle C, *[\nabla_t B_t \wedge *_s\nabla_s B_t]\rangle \\ &\quad -10\langle \nabla_s C, *[\nabla_s B_t \wedge B_t]\rangle - 10\langle \nabla_t C, *[\nabla_t B_t \wedge B_t]\rangle \end{aligned}$$

Now it follows from the the L^∞ estimates on the curvature that

$$|f_1| \le v_1 + c_3\left(\varepsilon^{-1}v_0 + \varepsilon^{-2}u_0\right)$$

In particular, the term $\varepsilon^{-2}\|d_A * d_A C\|^2 = \varepsilon^{-2}\|d_A\nabla_t B_t + d_A * \nabla_s B_t\|^2$ can be estimated by v_1. Hence

$$\Delta u_1 \ge v_1 - \varepsilon^{-1}c_3 v_0 - \varepsilon^{-2}c_3 u_0.$$

If $\varepsilon c_3 < 1/2$ then, by (7.5),

$$\begin{aligned}\Delta(u_0 + \varepsilon^2 u_1) &\ge v_0 + \varepsilon^2 v_1 - c_0 u_0 - \varepsilon c_3 v_0 - c_3 u_0 \\ &\ge \tfrac{1}{2}(v_0 + \varepsilon^2 v_1) - c_4 u_0\end{aligned}$$

where $c_4 = c_0 + c_3$. This inequality remains valid in the general case (arbitrary metric and perturbation) and it follows again from Lemma 7.3 below that

$$\sup_Q (u_0 + \varepsilon^2 u_1) \le c_5 \int_\Omega u_0.$$

Similar arguments show that

$$\sup_Q (u_0 + \varepsilon^2 u_1 + \varepsilon^4 u_2 + \varepsilon^6 u_3) \le c_6 \int_\Omega u_0.$$

where u_j is defined as above with derivatives of order j. This implies the assertion of the theorem for $p = \infty$. To see this note that (pointwise for every s and t)

$$\begin{aligned}\varepsilon^2 \|\nabla_t B_t\|^2_{L^\infty(\Sigma)} &\le c_7\varepsilon^2\Big(\|\nabla_t B_t\|^2_{L^2(\Sigma)} + \|d_A *_s d_A \nabla_t B_t\|^2_{L^2(\Sigma)} \\ &\qquad + \|d_A *_s d_A *_s \nabla_t B_t\|^2_{L^2(\Sigma)}\Big) \\ &\le c_8\left(u_0 + \varepsilon^2 u_1 + \varepsilon^4 u_2 + \varepsilon^6 u_3\right).\end{aligned}$$

The first inequality follows from arguments similar to Lemma 7.6 below. The second inequality follows from identities of the form

$$\begin{aligned}&\nabla_t\nabla_t\nabla_t B_t + \nabla_t\nabla_s\nabla_s B_t \\ &\quad = \nabla_t d_A \nabla_s C + \nabla_t *_s d_A \nabla_t C + 3\nabla_t *_s [B_t \wedge C] + \cdots \\ &\quad = d_A \nabla_t \nabla_s C + *_s d_A \nabla_t \nabla_t C + \cdots \\ &\quad = -\varepsilon^{-2} d_A *_s d_A *_s \nabla_t B_t - \varepsilon^{-2} *_s d_A *_s d_A \nabla_t B_t + \cdots.\end{aligned}$$

Thus we have proved the proposition for $p = 2$ and $p = \infty$. For general p the statement follows from the interpolation inequality $\|u\|_{L^p} \le \|u\|_{L^2}^{2/p}\|u\|_{L^\infty}^{1-2/p}$ for $2 \le p \le \infty$. □

LEMMA 7.3. *Let* $B_R = \{s + it : s^2 + t^2 \le R^2\}$, $u : B_{R+r} \to \mathbb{R}$ *be a* C^2*-function, and* $v : B_{R+r} \to \mathbb{R}$ *be continuous such that*

$$\Delta u \ge v - cu, \qquad u \ge 0, \qquad v \ge 0$$

for some constant $c > 0$. *Then*

$$\int_{B_R} v \le \left(c + \frac{4}{r^2}\right) \int_{B_{R+r}} u, \qquad \frac{\pi}{2} \sup_{B_R} u \le \left(c + \frac{4}{r^2}\right) \int_{B_{R+r}} u.$$

Proof. It suffices to prove the lemma for $r = 1$. To prove the first estimate note that

$$\int_{B_R} v - c \int_{B_{R+1}} u \le \int_{\partial B_{R+s}} \frac{\partial u}{\partial \nu} \le \frac{\mathrm{d}}{\mathrm{d}s} \int_{\partial B_{R+s}} u$$

for $1/2 \le s \le 1$. (The last inequality holds since $u \ge 0$.) Integrate this inequality from $1/2$ to t to obtain

$$\int_{B_R} v - c \int_{B_{R+1}} u \le 2 \int_{\partial B_{R+t}} u$$

for $1/2 \le t \le 1$. Integrate this inequality again from $1/2$ to 1 to obtain the first estimate for $r = 1$.

To prove the second estimate for $r = 1$ consider the function

$$f(\rho) = (1 - \rho)^2 \sup_{B_\rho} u.$$

Choose $\rho^* < 1$ to be any number at which f attains its maximum value and define $c^* = \sup_{B_{\rho^*}(0)} u = u(w^*)$ and $\delta = (1 - \rho^*)/2$. Then $u(w) \le 4c^*$ for $w \in B_{\rho^*+\delta}(0)$ and hence $\Delta u \ge -4cc^*$ in $B_\delta(w^*)$. This implies that the function $\tilde{u}(w) = u(w) + cc^*|w - w^*|^2$ is subharmonic in $B_\delta(w^*)$ and hence

$$c^* = u(w^*) \le \frac{cc^*\rho^2}{2} + \frac{1}{\pi\rho^2} \int_{B_\rho(w^*)} u, \qquad 0 < \rho \le \delta.$$

If $c\delta^2 \ge 1$ choose $\rho^2 = c^{-1} \le \delta^2$ to obtain

$$u(0) \le c^* \le \frac{2c}{\pi} \int_{B_\rho(w^*)} u.$$

If $c\delta^2 \le 1$ choose $\rho = \delta$ to obtain

$$c^*\delta^2 \le \frac{2}{\pi} \int_{B_\delta(w^*)} u$$

and use the inequality $u(0) = f(0) \le f(\rho^*) = (1-\rho^*)^2 c^* = 4\delta^2 c^*$. This proves the second estimate for $r = 1$. The general case can be reduced to the case $r = 1$ by rescaling. □

We shall now prove an exponential estimate on the curvature of a connection $\Xi \in \mathcal{A}^{1,p}_\varepsilon(a^-, a^+, H)$ for $t \to \infty$. For a fixed number $\varepsilon > 0$ such estimates are well known. Our result is quantitative and shows how the constants vary as ε tends to 0.

THEOREM 7.4. *Assume all H-flat connections $a \in \mathcal{A}_{\mathrm{flat}}(P_f, H)$ are nondegenerate. Then for every $c_0 > 0$ there exists a constants $\delta > 0$, $\varepsilon_0 > 0$, $c > 0$, and $\rho > 0$ such that the following holds. If $\Xi \in \mathcal{A}^{1,p}_\varepsilon(a^-, a^+, H)$ with $0 < \varepsilon \le \varepsilon_0$ satisfies*

$$\varepsilon^{-1} \|F_A\|_{\mathrm{L}^\infty(\Sigma_h \times \mathbb{R})} + \|\partial_t A - \mathrm{d}_A \Psi\|_{\mathrm{L}^\infty(\Sigma_h \times \mathbb{R})} \le c_0 \tag{7.6}$$

and

$$\mathcal{Y}^\varepsilon_{[0,\infty)}(\Xi) = \varepsilon^{-2} \|F_A\|^2_{\mathrm{L}^2(\Sigma_h \times [0,\infty))} + \|\partial_t A - \mathrm{d}_A \Psi\|^2_{\mathrm{L}^2(\Sigma_h \times [0,\infty))} \le \delta$$

then

$$\mathcal{Y}^\varepsilon_{[T,\infty)}(\Xi) \le c e^{-\rho T}, \qquad T \ge 0.$$

LEMMA 7.5. *Assume all H-flat connections $a \in \mathcal{A}_{\mathrm{flat}}(P_f, H)$ are nondegenerate. Then there exist a constants $\delta > 0$, $\varepsilon_0 > 0$, and $c > 0$ such that for every connection $A + \Phi\,\mathrm{d}s \in \mathcal{A}(P_f)$ with*

$$\|F_A\|_{\mathrm{L}^\infty(\Sigma_h)} + \|\partial_s A - X_s(A) - \mathrm{d}_A \Phi\|_{\mathrm{L}^\infty(\Sigma_h)} \le \delta$$

and for $0 < \varepsilon \le \varepsilon_0$ there is an estimate

$$\begin{aligned} &\|\alpha\|^2 + \varepsilon^2 \|\phi\|^2 + \varepsilon^2 \|\psi\|^2 \\ &\le c \Big(\|{*_s}\nabla_s \alpha - {*_s}\mathrm{d}X_s(A)\alpha - {*_s}\mathrm{d}_A\phi - \mathrm{d}_A\psi\|^2 \\ &\quad + \varepsilon^2 \left\|\nabla_s\psi - \varepsilon^{-2}\mathrm{d}_A\alpha\right\|^2 + \varepsilon^2 \left\|\nabla_s {*_s} \phi + \varepsilon^{-2}\mathrm{d}_A {*_s} \alpha\right\|^2 \Big) \end{aligned} \tag{7.7}$$

for $\alpha \in \mathrm{W}^{1,2}_f(\mathbb{R} \times T^*\Sigma \otimes \mathfrak{g}_P)$ and $\phi, \psi \in \mathrm{W}^{1,2}_f(\mathbb{R} \times \mathfrak{g}_P)$. Here all norms are L^2-norms on Σ_h.

Proof. Suppose not. Then there exists a sequence $\varepsilon_\nu \to 0$ and a sequence of connections $A_\nu + \Phi_\nu\,\mathrm{d}s \in \mathcal{A}(P_f)$ such that

$$\lim_{\nu\to\infty} \Big(\|F_{A_\nu}\|_{\mathrm{L}^\infty(\Sigma_h)} + \|\partial_s A_\nu - X_s(A_\nu) - \mathrm{d}_{A_\nu}\Phi_\nu\|_{\mathrm{L}^\infty(\Sigma_h)} \Big) = 0$$

and the estimate (7.7) does not hold with $c = \nu$ and $A + \Phi\,\mathrm{d}s$ replaced by $A_\nu + \Phi_\nu\,\mathrm{d}s$ and ε replaced by ε_ν. By Uhlenbeck's compactness (cf. [32]) we may assume that $A_\nu + \Phi_\nu\,\mathrm{d}s$ converges to an H-flat connection $A + \Phi\,\mathrm{d}s \in$

$\mathcal{A}_{\mathrm{flat}}(P_f)$. (If necessary, pass to a subsequence and apply a sequence of gauge transformations. Note that the estimate (7.7) is invariant under gauge transformations.) Since $A + \Phi\, \mathrm{d}s$ is nondegenerate there exists a constant $c_0 > 0$ such that the estimate

$$\|\alpha_0\|^2 \le c_0 \|\pi_{A_\nu}(\nabla_s \alpha_0 - \mathrm{d}X_s(A_\nu)\alpha_0)\|^2$$

holds for ν sufficiently large and $\alpha_0(s) \in \mathrm{H}^1_{A_\nu}(s,t)$ with $\alpha_0(s+1) = f^*\alpha_0(s)$. Hence it follows from Lemma 7.4 in [10] that there exist constants $\varepsilon_0 > 0$, $\nu_0 \in \mathbb{N}$, and $c > 0$ such that the estimate (7.7) holds with $0 < \varepsilon \le \varepsilon_0$ and $A + \Phi\, \mathrm{d}s$ replaced by $A_\nu + \Phi_\nu\, \mathrm{d}s$ where $\nu \ge \nu_0$. With $\varepsilon = \varepsilon_\nu$ and $\nu > c$ this contradicts our assumption. □

Lemma 7.6. *Let $p > 2$. Then there exist constants $\delta > 0$ and $c > 0$ such that for every connection $A \in \mathcal{A}(P)$ with*

$$\|F_A\|_{\mathrm{L}^p} \le \delta$$

there are estimates

$$\|\phi\|_{\mathrm{L}^\infty} \le c \|\mathrm{d}_A\phi\|_{\mathrm{L}^p}, \qquad \|\mathrm{d}_A\phi\|_{\mathrm{L}^\infty} \le c \|\mathrm{d}_A *_s \mathrm{d}_A\phi\|_{\mathrm{L}^p},$$

for $\phi \in C^\infty(\mathfrak{g}_P)$ and $s \in \mathbb{R}$.

Proof. Since every flat connection on P is irreducible the estimates hold when $F_A = 0$. Moreover, given a flat connection A_0, there exist constants $\delta > 0$ and $c > 0$ such that the estimates hold for every connection $A \in \mathcal{A}(P)$ with

$$\|A - A_0\|_{\mathrm{L}^p} + \|F_A\|_{\mathrm{L}^p} \le \delta.$$

Now, if the statement were false then there would exist a sequence $A_\nu \in \mathcal{A}(P)$ such that $\|F_{A_\nu}\|_{\mathrm{L}^p} \to 0$ and one of the estimates fails to hold with $c = \nu$. By Uhlenbeck's compactness theorem there exist a subsequence (still denoted by A_ν) and a sequence $g_\nu \in \mathcal{G}(P)$ such that $g_\nu^* A_\nu$ converges in the L^p-norm to a flat connection A_0. Hence the estimates hold for the connections $g_\nu^* A_\nu$ with a uniform constant c. Hence they hold for A_ν with a uniform constant c. This contradicts our assumption on A_ν and proves the lemma. □

Proof of Theorem 7.4. Consider the function

$$f(t) = \tfrac{1}{2} \int_0^1 \left(\|B_t\|^2_{\mathrm{L}^2(\Sigma,*_s)} + \varepsilon^2 \|C\|^2_{\mathrm{L}^2(\Sigma,*_s)} \right) \mathrm{d}s$$

where B_t and C are defined by (7.1). By (7.6) we have

$$f(t) \le \tfrac{1}{2}\mathrm{Vol}(\Sigma) c_0^2 = c_1$$

for $t \in \mathbb{R}$. By (7.2) and (7.3) we have

$$\varepsilon^2 \nabla_t C = - *_s \mathrm{d}_A B_t, \quad \varepsilon^2 \nabla_t \nabla_t C = - *_s \mathrm{d}_A \nabla_t B_t - *_s [B_t \wedge B_t]$$

and, since $\nabla_t X_s(A) = \mathrm{d}X_s(A)B_t$,

$$\begin{aligned}\nabla_t B_t &= -*_s\left(\nabla_s B_t - \mathrm{d}X_s(A)B_t - \mathrm{d}_A C\right),\\ \nabla_t\nabla_t B_t &= -*_s\left(\nabla_s\nabla_t B_t - \mathrm{d}X_s(A)\nabla_t B_t - \mathrm{d}_A\nabla_t C\right)\\ &\quad -2*_s[C\wedge B_t] + *_s\mathrm{d}^2X_s(A)(B_t,B_t).\end{aligned}$$

Now $f'(t) = \langle\nabla_t B_t, B_t\rangle + \varepsilon^2\langle\nabla_t C, C\rangle$ and the second derivative is given by

$$\begin{aligned}f''(t) &= \|\nabla_t B_t\|^2 + \varepsilon^2\|\nabla_t C\|^2 + \langle\nabla_t\nabla_t B_t, B_t\rangle + \varepsilon^2\langle\nabla_t\nabla_t C, C\rangle\\
&= \|\nabla_t B_t\|^2 + \varepsilon^{-2}\|\mathrm{d}_A B_t\|^2\\
&\quad -\langle *_s(\nabla_s\nabla_t B_t - \mathrm{d}X_s(A)\nabla_t B_t - \mathrm{d}_A\nabla_t C), B_t\rangle - \langle *_s\mathrm{d}_A\nabla_t B_t, C\rangle\\
&\quad -\langle 2*_s[C\wedge B_t] + *_s\mathrm{d}^2X_s(A)(B_t,B_t), B_t\rangle - \langle *_s[B_t\wedge B_t], C\rangle\\
&= \|\nabla_t B_t\|^2 + \varepsilon^{-2}\|\mathrm{d}_A B_t\|^2\\
&\quad -\langle\nabla_t B_t, *_s(\nabla_s B_t - \mathrm{d}X_s(A)B_t - \mathrm{d}_A C)\rangle - \langle\nabla_t C, *_s\mathrm{d}_A B_t\rangle\\
&\quad -3\langle C, *_s[B_t\wedge B_t]\rangle + \langle *_s\mathrm{d}^2X_s(A)(B_t,B_t), B_t\rangle\\
&= 2\|\nabla_s B_t - \mathrm{d}X_s(A)B_t - \mathrm{d}_A C\|^2 + 2\varepsilon^{-2}\|\mathrm{d}_A B_t\|^2\\
&\quad -3\langle C, *_s[B_t\wedge B_t]\rangle + \langle *_s\mathrm{d}^2X_s(A)(B_t,B_t), B_t\rangle\\
&= \|\nabla_s B_t - \mathrm{d}X_s(A)B_t - \mathrm{d}_A C\|^2 + \|\nabla_s B_t - \mathrm{d}X_s(A)B_t\|^2 + \|\mathrm{d}_A C\|^2\\
&\quad +2\varepsilon^{-2}\|\mathrm{d}_A B_t\|^2 - 2\langle\mathrm{d}_A C, \nabla_s B_t - \mathrm{d}X_s(A)B_t\rangle\\
&\quad -3\langle C, *_s[B_t\wedge B_t]\rangle + \langle *_s\mathrm{d}^2X_s(A)(B_t,B_t), B_t\rangle\\
&= \|\nabla_s B_t - \mathrm{d}X_s(A)B_t - \mathrm{d}_A C\|^2 + \|\nabla_s B_t - \mathrm{d}X_s(A)B_t\|^2 + \|\mathrm{d}_A C\|^2\\
&\quad +2\varepsilon^{-2}\|\mathrm{d}_A B_t\|^2 + 2\varepsilon^{-2}\|\mathrm{d}_A *_s B_t\|^2 + 2\langle\mathrm{d}_A C, \mathrm{d}X_s(A)B_t\rangle\\
&\quad +2\langle *_s[X_s(A)\wedge C], B_t\rangle + 2\langle\dot{*}_s C, \mathrm{d}_A *_s B_t\rangle - 2\langle\mathrm{d}_A C, *_s\dot{*}_s B_t\rangle\\
&\quad -5\langle C, *_s[B_t\wedge B_t]\rangle + \langle *_s\mathrm{d}^2X_s(A)(B_t,B_t), B_t\rangle.\end{aligned}$$

Here all norms and inner products are L^2-norms and L^2-inner products on Σ_h. The third equality follows from the fact that the operators $\alpha \mapsto *_s\nabla_s\alpha$ and $\alpha \mapsto *_s\mathrm{d}X_s(A)\alpha$ are self-adjoint. Here the "bad" terms are the ones which involve the product of B_t and C. To control these, it is necessary to isolate the term $\|d_A C\|^2$ in the above identity. For fixed ε it would have been sufficient to use the fourth expression.

By Theorem 7.1, we have

$$\varepsilon^{-1}\|F_A\|_{\mathrm{L}^\infty(\Sigma_h\times T)} + \|\partial_s A - X_s(A) - \mathrm{d}_A\Phi\|_{\mathrm{L}^\infty(\Sigma_h\times T)} \le c_2\delta$$

for $T \ge 1$. Choose $\delta > 0$ so small that Lemma 7.5 holds with δ replaced by $c_2\delta$. Choose $\varepsilon_0 > 0$ and $c_3 > 0$ to be the constants of Lemma 7.5 so that the estimate (7.7) holds with $c = c_3$, $A(s) = A(s,t)$, and $\Phi(s) = \Phi(s,t)$ provided that $\varepsilon \le \varepsilon_0$ and $t \ge 1$. Apply this estimate to $\alpha = B_t$, $\phi = C$, $\psi = 0$, and use the identity $\nabla_s *_s C + \varepsilon^{-2}\mathrm{d}_A *_s B_t = 0$, to obtain

$$\|B_t\|^2 + \varepsilon^2\|C\|^2 \le c_3\left(\|\nabla_s B_t - \mathrm{d}X_s(A)B_t - \mathrm{d}_A C\|^2 + \varepsilon^{-2}\|\mathrm{d}_A B_t\|^2\right).$$

Moreover, by Lemma 7.6, there is a constant $c_4 > 0$ such that

$$\|C\|^2 \le c_4 \|\mathrm{d}_A C\|^2$$

for $t \ge 0$ provided that ε_0 is sufficiently small. Hence the above formula for $f''(t)$ shows that there exists a constant $\rho > 0$ such that

$$f''(t) \ge \rho^2 f(t), \qquad t \ge 1.$$

This implies

$$f(t) \le e^{-\rho(t-1)} f(1) \le e^{-\rho(t-1)} c_1, \qquad t \ge 1.$$

(To see this note that the function $g(t) = e^{-\rho t}(f'(t) + \rho f(t))$ is strictly increasing. Since $f(t)$ does not converge to infinity it follows that $g(t) < 0$ and hence $e^{\rho t} f(t)$ is decreasing.) Hence

$$\mathcal{Y}^{\varepsilon}_{[T,\infty)}(\Xi) = \int_T^\infty f(t)\,\mathrm{d}t \le \rho^{-1} c_1 e^{-\rho(T-1)} = c_5 e^{-\rho T}$$

for $T \ge 1$. With $c_5 \ge \delta e^{\rho}$ the theorem follows. □

8. Compactness with bounded curvature

In this section we shall prove that every sequence of ε_ν-self-dual instantons Ξ_ν connecting a^- to a^+ with $\varepsilon_\nu \to 0$ has a subsequence which converges, modulo gauge transformation and time shift, to a holomorphic curve Ξ_0. We shall also prove that the convergence is sufficiently fast (with the rate $\varepsilon_\nu^{1+2/p}$) so that, by Theorem 6.1, Ξ_ν is in the range of the operator $\mathcal{T}_{\varepsilon_\nu} : \mathcal{A}_0^{1,p}(a^-, a^+, H) \to \mathcal{A}_{\varepsilon_\nu}^{1,p}(a^-, a^+, H)$ of Theorem 5.1 for ν sufficiently large.

Theorem 8.1. *Assume $H \in \mathcal{H}_0^{\mathrm{reg}}$. Then for every constant $c_0 > 0$ there exists a constant $\varepsilon_0 > 0$ such that the following holds. If $a^\pm \in \mathcal{A}_{\mathrm{flat}}(P_f, H)$ with $\mu(a^-, a^+) = 1$ and $\Xi \in \mathcal{A}_\varepsilon^{1,p}(a^-, a^+, H)$ with $0 < \varepsilon \le \varepsilon_0$ and*

$$\varepsilon^{-2} \|F_A\|_{\mathrm{L}^\infty} + \|\partial_t A - d_A \Psi\|_{\mathrm{L}^\infty} \le c_0 \tag{8.1}$$

then there exists a connection $\Xi_0 \in \mathcal{A}_0^{1,p}(a^-, a^+, H)$ such that

$$\Xi = \mathcal{T}_\varepsilon(\Xi_0).$$

Lemma 8.2. *Let $p > 2$. Then there exist constants $\delta > 0$ and $c > 0$ such that the following holds. For every connection $A \in \mathcal{A}(P)$ with*

$$\|F_A\|_{\mathrm{L}^p} \le \delta$$

there exists a unique section $\eta \in C^\infty(\mathfrak{g}_P)$ such that

$$F_{A+*\mathrm{d}_A\eta} = 0, \qquad \|\mathrm{d}_A \eta\|_{\mathrm{L}^\infty} \le c \|F_A\|_{\mathrm{L}^p}.$$

Proof. The condition $F_{A+*\mathrm{d}_A\eta} = 0$ is equivalent to

$$F_A + \mathrm{d}_A * \mathrm{d}_A\eta + \tfrac{1}{2}[\mathrm{d}_A\eta \wedge \mathrm{d}_A\eta] = 0.$$

Hence the result follows from Lemma 7.6 and the implicit function theorem. More explicitly, use a Newton type iteration argument by constructing a sequence $\eta_{\nu+1} = \eta_\nu + \zeta_\nu$ where $\eta_1 = 0$ and ζ_ν is the unique solution of the elliptic equation

$$\mathrm{d}_A * \mathrm{d}_A\zeta_\nu + F_{A+*\mathrm{d}_A\eta_\nu} = 0.$$

This solution exists by Lemma 7.6 and the sequence η_ν converges to the required solution η. The details of this argument are left to the reader. □

Proof of Theorem 8.1. Assume that the statement were false. Then there exist H-flat connections $a^\pm \in \mathcal{A}_{\mathrm{flat}}(P_f, H)$, a sequence $\varepsilon_\nu \to 0$, and a sequence $\Xi_\nu \in \mathcal{A}^{1,p}_{\varepsilon_\nu}(a^-, a^+, H)$ such that (8.1) holds with $\Xi = \Xi_\nu$ and $\varepsilon = \varepsilon_\nu$ but Ξ_ν is not in the range of $\mathcal{T}_{\varepsilon_\nu}$. Hence $g^*(\Xi_\nu \circ \sigma_\tau)$ is not in the range of $\mathcal{T}_{\varepsilon_\nu}$ for every $g \in \mathcal{G}^{2,p}$ and every $\tau \in \mathbb{R}$. Applying a suitable time shift we may assume without loss of generality that

$$\mathcal{CS}_H(A_\nu(s,0) + \Phi_\nu(s,0)\,ds) = \tfrac{1}{2}(\mathcal{CS}_H(a^-) + \mathcal{CS}_H(a^+))$$

Applying a suitable gauge transformation we may also assume that $\Psi_\nu(s,t) = 0$ for $|t| \geq T_0$. We shall prove in seven steps that Ξ_ν is in the range of $\mathcal{T}_{\varepsilon_\nu}$ for some ν in contradiction to our assumption.

Step 1: *There exist constants $c > 0$ and $\rho > 0$ such that*

$$\mathcal{Y}^{\varepsilon_\nu}_{[T,\infty)}(\Xi_\nu) \leq ce^{-\rho T}$$

for $T \geq 0$ and similarly for $T \leq 0$.

In view of Theorem 7.4 it suffices to prove that

$$\lim_{T\to\infty} \inf_\nu \mathcal{Y}^{\varepsilon_\nu}_{[-T,T]}(\Xi_\nu) = \mathcal{CS}_H(a^-) - \mathcal{CS}_H(a^+).$$

We prove this by contradiction. If this equation would not hold then there would exist a number $\delta > 0$, a subsequence (still denoted by Ξ_ν), and a sequence $T_\nu \to \infty$ such that

$$\mathcal{Y}^{\varepsilon_\nu}_{[-T_\nu,T_\nu]}(\Xi_\nu) \leq \mathcal{CS}_H(a^-) - \mathcal{CS}_H(a^+) - \delta.$$

Now the curvature of Ξ_ν satisfies a uniform L^∞ estimate. Hence, by Uhlenbeck's weak compactness theorem (cf. [32]), we may choose a further subsequence and a sequence $g_\nu \in \mathcal{G}^{2,p}$ such that $g_\nu^*\Xi_\nu$ converges to $\Xi_0 = A_0 + \Phi_0\,\mathrm{d}s + \Psi_0\,\mathrm{d}t$, uniformly on compact sets and weakly in $\mathrm{W}^{1,p}$ on compact sets. The limit connection Ξ_0 satisfies the holomorphic curve equation (3.4) and has finite energy

$$E(\Xi_0) = \int_{-\infty}^{\infty}\int_0^1 \|\partial_t A_0 - \mathrm{d}_{A_0}\Psi_0\|^2_{\mathrm{L}^2(\Sigma,*_s)}\,\mathrm{d}s\mathrm{d}t \le \mathcal{CS}_H(a^-) - \mathcal{CS}_H(a^+) - \delta.$$

Hence the limits of Ξ_0 for $t \to \pm\infty$ exist and are H-flat connections on P_f. Since $\mu(a^-, a^+) = 1$ the limits of Ξ_0 agree with those of Ξ_ν. Otherwise it would follow from the usual arguments in Floer homology or in finite dimensional Morse theory (see e.g. [26]) that there exist H-flat connections $a_0, a_1, \ldots, a_\ell \in \mathcal{A}_\Sigma(P_f, H)$ with $\ell > 1$ such that $a_0 = a^-$, $a_\ell = a^+$, and $\mathcal{M}_0(a_j, a_{j+1}, H) \neq \emptyset$ for every j. Since $H \in \mathcal{H}_0^{\mathrm{reg}}$ this would imply that

$$\mu_H(a_j, a_{j+1}) \ge 1$$

for every j and hence

$$\mu_H(a^-, a^+) = \sum_{j=0}^{\ell-1} \mu_H(a_j, a_{j+1}) \ge 2$$

in contradiction to the assumption $\mu_H(a^-, a^+) = 1$. This shows that $\Xi_0 \in \mathcal{A}_0^{1,p}(a^-, a^+, H)$ and hence

$$E(\Xi_0) = \mathcal{CS}_H(a^-) - \mathcal{CS}_H(a^+).$$

This contradicts the above inequality and proves step 1.

Step 2: *Let ρ be the constant of step 1. Let B_s^ν, B_t^ν and C^ν be defined by (7.1) with Ξ replaced by Ξ_ν. Then there exist constants $c > 0$ and $\nu_0 \in \mathbb{N}$ such that*

$$\begin{aligned}
&\varepsilon_\nu^{2/p}\,\|B_t^\nu\|_{\mathrm{L}^p(\Sigma_h\times[T,\infty))} + \|\mathrm{d}_{A_\nu} B_t^\nu\|_{\mathrm{L}^p(\Sigma_h\times[T,\infty))} + \|\mathrm{d}_{A_\nu} *_s B_t^\nu\|_{\mathrm{L}^p(\Sigma_h\times[T,\infty))}\\
&\quad +\varepsilon_\nu\,\|\nabla_s B_t^\nu\|_{\mathrm{L}^p(\Sigma_h\times[T,\infty))} + \varepsilon_\nu\,\|\nabla_t B_t^\nu\|_{\mathrm{L}^p(\Sigma_h\times[T,\infty))}\\
&\quad\quad +\varepsilon_\nu\,\|\mathrm{d}_{A_\nu} C^\nu\|_{\mathrm{L}^p(\Sigma_h\times[T,\infty))}\\
&\quad\quad\quad +\varepsilon_\nu^2\,\|\nabla_s C^\nu\|_{\mathrm{L}^p(\Sigma_h\times[T,\infty))} + \varepsilon_\nu^2\,\|\nabla_t C^\nu\|_{\mathrm{L}^p(\Sigma_h\times[T,\infty))}\\
&\quad\quad\quad\quad \le c\varepsilon^{2/p} e^{-\rho T}.
\end{aligned}$$

for $2 \le p \le \infty$, $\nu \ge \nu_0$, and $T \ge 0$. Similarly for $T \le 0$. Moreover,

$$\|C^\nu\|_{\mathrm{L}^p} \le c$$

for $2 \le p \le \infty$ and $\nu \ge \nu_0$.

The first inequality follows from step 1 and Theorem 7.1. For $p = 2$ the second inequality follows from the first and for $p = \infty$ it holds by assumption. For general p it follows by interpolation using Hölder's inequality.

Step 3: *By Lemma 8.2 choose* $\eta_\nu(s,t) \in C^\infty(\mathfrak{g}_P)$ *such that*

$$F_{A'_\nu} = 0, \qquad A'_\nu = A_\nu + *\mathrm{d}_{A_\nu}\eta_\nu.$$

Moreover, choose $\Phi'_\nu(s,t), \Psi'_\nu(s,t) \in C^\infty(\mathfrak{g}_P)$ *such that*

$$\mathrm{d}_{A'_\nu} *_s (\partial_s A'_\nu - \mathrm{d}_{A'_\nu}\Phi'_\nu - X_s(A'_\nu)) = 0, \qquad \mathrm{d}_{A'_\nu} *_s (\partial_t A'_\nu - \mathrm{d}_{A'_\nu}\Psi'_\nu) = 0.$$

Then $\Xi'_\nu \in \mathcal{A}^{1,p}(a^-, a^+)$ *and there exist constants* $c > 0$ *and* $\nu_0 \in \mathbb{N}$ *such that*[4]

$$\left\|\Xi'_\nu - \Xi_\nu\right\|_{1,p,\varepsilon_\nu,\Xi_\nu} \le c\varepsilon_\nu^{1+2/p} \tag{8.2}$$

$$\left\|\partial_t A'_\nu - \mathrm{d}_{A'_\nu}\Psi'_\nu\right\|_{\mathrm{L}^\infty(\Sigma,*_s)} \le ce^{-\rho|t|}, \tag{8.3}$$

$$\left\|\partial_t A'_\nu - \mathrm{d}_{A'_\nu}\Psi'_\nu + *_s(\partial_s A'_\nu - X_s(A'_\nu) - \mathrm{d}_{A'_\nu}\Phi'_\nu)\right\|_{\mathrm{L}^p} \le c\varepsilon_\nu^2. \tag{8.4}$$

for $\nu \ge \nu_0$ *and* $2 \le p \le \infty$.

We shall suppress the subscript ν and write A, A', ε for A_ν, A'_ν, ε_ν etc. All constants are independent of ν. By Lemma 8.2 we have an estimate

$$\left\|A' - A\right\|_{\mathrm{L}^\infty(\Sigma,*_s)} \le c_1 \left\|F_A\right\|_{\mathrm{L}^p(\Sigma,*_s)}$$

pointwise for every s and t. Hence, by step 2,

$$\left\|A' - A\right\|_{\mathrm{L}^p} \le c_2\varepsilon^2$$

for $2 \le p \le \infty$ where the L^p-norm is to be understood on the infinite cylinder $\Sigma_h \times \mathbb{R}$. It follows also from step 2 that

$$\left\|A' - A\right\|_{\mathrm{L}^\infty(\Sigma,*_s)} + \left\|F_A\right\|_{\mathrm{L}^\infty(\Sigma,*_s)} \le c_3\varepsilon e^{-\rho|t|}$$

pointwise for every s and t. Now differentiate the identity

$$F_A + \mathrm{d}_A *_s \mathrm{d}_A\eta + \tfrac{1}{2}[\mathrm{d}_A\eta \wedge \mathrm{d}_A\eta] = 0$$

with respect to t to obtain

$$\begin{aligned} \mathrm{d}_A *_s \mathrm{d}_A\nabla_t\eta &= -\mathrm{d}_A B_t - [\mathrm{d}_A\nabla_t\eta \wedge \mathrm{d}_A\eta] - [[B_t \wedge \eta] \wedge \mathrm{d}_A\eta] \\ &\quad -2[B_t \wedge *_s\mathrm{d}_A\eta] - [\mathrm{d}_A *_s B_t \wedge \eta] \end{aligned} \tag{8.5}$$

This implies

$$\left\|\mathrm{d}_A\nabla_t\eta\right\|_{\mathrm{L}^\infty(\Sigma,*_s)} \le c_4\left(\left\|\mathrm{d}_A B_t\right\|_{\mathrm{L}^p(\Sigma,*_s)} + \left\|A' - A\right\|_{\mathrm{L}^p(\Sigma,*_s)}\right),$$

[4]The subscript Ξ_ν in (8.2) indicates that the covariant derivatives in the definition of the $1,p,\varepsilon_\nu$-norm are with respect to the connection Ξ_ν.

and hence

$$\|\mathrm{d}_A\nabla_t\eta\|_{\mathrm{L}^\infty(\Sigma,*_s)} \le c_5 e^{-\rho|t|}, \qquad \|\mathrm{d}_A\nabla_t\eta\|_{\mathrm{L}^p} \le c_5\varepsilon^{2/p}. \tag{8.6}$$

A similar estimate holds for $\nabla_s\eta$.

Now denote $B_s' = \partial_s A' - \mathrm{d}_{A'}\Phi'$ and $B_t' = \partial_t A' - \mathrm{d}_{A'}\Psi'$. Then $B_t' - B_t = \mathrm{d}_{A'}(\Psi - \Psi') + \nabla_t(A' - A)$ and hence

$$B_t' - B_t = \mathrm{d}_{A'}(\Psi - \Psi') + *_s\mathrm{d}_A\nabla_t\eta + *_s[B_t \wedge \eta] \tag{8.7}$$

This implies

$$\begin{aligned} \mathrm{d}_A *_s \mathrm{d}_A(\Psi' - \Psi) &= \mathrm{d}_A *_s B_t - [A' - A \wedge \mathrm{d}_A\nabla_t\eta + [B_t \wedge \eta]] \\ &\quad -[\mathrm{d}_A B_t \wedge \eta] - [F_A \wedge \nabla_t\eta] \end{aligned} \tag{8.8}$$

Hence

$$\|\mathrm{d}_A(\Psi' - \Psi)\|_{\mathrm{L}^\infty(\Sigma,*_s)} \le c_6 e^{-\rho|t|}, \qquad \|\mathrm{d}_A(\Psi' - \Psi)\|_{\mathrm{L}^p} \le c_6\varepsilon_\nu^{2/p}$$

and, by (8.7),

$$\|B_t'\|_{\mathrm{L}^\infty(\Sigma,*_s)} \le c_7 e^{-\rho|t|}, \qquad \|B_t' - B_t\|_{\mathrm{L}^p} \le c_7\varepsilon_\nu^{2/p}$$

This proves (8.3).

It follows also from (8.7) and an analogous identity for $B_s' - B_s$ that

$$\begin{aligned} B_t' + *_s(B_s' - X_s(A')) &= *_s\dot{*}_s\mathrm{d}_A\eta - [X_s(A) \wedge \eta] - *_s(X_s(A') - X_s(A)) \\ &\quad +[A' - A \wedge \nabla_s\eta] - *_s[A' - A \wedge \nabla_t\eta] \\ &\quad -\mathrm{d}_{A'}(\Psi' - \Psi + \nabla_s\eta) - *_s\mathrm{d}_{A'}(\Phi' - \Phi - \nabla_t\eta). \end{aligned}$$

In view of (8.6) it follows that the A'-harmonic part of $B_t' + *_s(B_s' - X_s(A'))$ can be estimated by $A' - A$. This proves (8.4).

To prove the estimate (8.2) note that

$$\mathrm{d}_A(A' - A) = -F_A + \tfrac{1}{2}[A - A' \wedge A - A'], \qquad \mathrm{d}_A *_s (A' - A) = -[F_A \wedge \eta].$$

The L^p-norm of both terms can be estimated by ε^2. The L^p-norm of $\mathrm{d}_A(\Psi' - \Psi)$ has already been estimated above by $\varepsilon^{2/p}$ and for $\mathrm{d}_A(\Phi' - \Phi)$ the argument is similar. Since

$$\nabla_t(A' - A) = *_s[B_t \wedge \eta] + *_s\mathrm{d}_A\nabla_t\eta$$

it follows from (8.6) that the L^p-norm of $\nabla_t(A' - A)$ can be estimated by $\varepsilon^{2/p}$. Similarly for $\nabla_s(A' - A)$. To estimate $\nabla_t(\Psi' - \Psi)$ differentiate the identity (8.5) with respect to t to obtain an estimate for $\nabla_t\nabla_t\eta$ in terms of $\varepsilon^{2/p-1}$ Then differentiate (8.8) with respect to t. The expressions $\nabla_s(\Psi' - \Psi)$, $\nabla_t(\Phi' - \Phi)$, and $\nabla_s(\Phi' - \Phi)$ can be estimated by similar arguments the details of which are left to the reader. This proves (8.2).

Step 4: *Fix a constant $p > 4$. Then for ν sufficiently large there exists a smooth map $A''_\nu : \mathbb{R}^2 \to \mathcal{A}_{\mathrm{flat}}(P)$ which satisfies the Cauchy-Riemann equations (3.4), the boundary condition (3.2),*

$$\mathrm{d}_{A''_\nu} *_s (A''_\nu - A'_\nu) = 0, \tag{8.9}$$

and

$$\|A''_\nu - A'_\nu\|_{\mathrm{L}^p} + \|A''_\nu - A'_\nu\|_{\mathrm{L}^\infty} \le c\varepsilon_\nu^2,$$
$$\|\partial_t A''_\nu - \mathrm{d}_{A''_\nu}\Psi''_\nu - \partial_t A'_\nu + \mathrm{d}_{A'_\nu}\Psi'_\nu\|_{\mathrm{L}^p} \le c\varepsilon_\nu^2.$$

Here Ψ''_ν is chosen such that $\mathrm{d}_{A''_\nu} *_s (\partial_t A''_\nu - \mathrm{d}_{A''_\nu}\Psi''_\nu) = 0$. The constant c is independent of ν.

The assertion follows from step 3 and Theorem 2.5. Condition (8.9) means that for every s and t the connection $A''_\nu(s,t)$ minimizes the L^2-distance of the orbit of $A''_\nu(s,t)$ under $\mathcal{G}(P)$ to the connection $A'_\nu(s,t)$ with respect to the s-metric.

Step 5: *For ν sufficiently large there exists a smooth map $A^0_\nu : \mathbb{R}^2 \to \mathcal{A}_{\mathrm{flat}}(P)$ which satisfies (3.4), (3.2),*

$$\mathrm{d}_{A_\nu} *_s (A_\nu - A^0_\nu) = 0, \tag{8.10}$$

and

$$\left\|A_\nu - A^0_\nu\right\|_{\mathrm{L}^p} \le c\varepsilon_\nu^2,$$
$$\left\|\partial_t A_\nu - \mathrm{d}_{A_\nu}\Psi_\nu - \partial_t A^0_\nu + \mathrm{d}_{A^0_\nu}\Psi^0_\nu\right\|_{\mathrm{L}^p} \le c\varepsilon_\nu^{2/p}.$$

Here Ψ^0_ν is chosen such that $\mathrm{d}_{A^0_\nu} *_s (\partial_t A^0_\nu - \mathrm{d}_{A^0_\nu}\Psi^0_\nu) = 0$. The constant $c > 0$ is independent of ν.

It follows from step 3 and step 4 that

$$\|A_\nu - A''_\nu\|_{\mathrm{L}^\infty} + \|A_\nu - A''_\nu\|_{\mathrm{L}^p} \le c_1\varepsilon_\nu^2$$

and

$$\|\partial_t A_\nu - \mathrm{d}_{A_\nu}\Psi_\nu - \partial_t A''_\nu + \mathrm{d}_{A''_\nu}\Psi''_\nu\|_{\mathrm{L}^p} \le c_1\varepsilon_\nu^{2/p}$$

with a suitable constant $c_1 > 0$. Moreover, with η_ν as in step 3 we have

$$\begin{aligned}
\mathrm{d}_{A_\nu} *_s (A_\nu - A''_\nu) &= \mathrm{d}_{A_\nu} *_s (A_\nu - A'_\nu) + \mathrm{d}_{A_\nu} *_s (A'_\nu - A''_\nu) \\
&= -\mathrm{d}_{A_\nu}\mathrm{d}_{A_\nu}\eta_\nu + [A_\nu - A''_\nu \wedge *_s(A'_\nu - A''_\nu)] \\
&= -[F_{A_\nu} \wedge \eta_\nu] + [A_\nu - A'_\nu \wedge *_s(A'_\nu - A''_\nu)].
\end{aligned}$$

and

$$\mathrm{d}_{A_\nu}(A_\nu - A''_\nu) = F_{A_\nu} - \tfrac{1}{2}[A_\nu - A''_\nu \wedge A_\nu - A''_\nu].$$

Hence there is a constant $c_2 > 0$ such that for every ν

$$\sup_{s,t} \|A_\nu - A_\nu''\|_{W^{1,p}(\Sigma,*_s)} + \|A_\nu - A_\nu''\|_{L^p} \le c_2 \varepsilon_\nu^2.$$

This implies that for all s and t there exists a gauge transformation $g_\nu(s,t) \in \mathcal{G}(P)$ such that the map $A_\nu^0 = g_\nu^* A_\nu''$ satisfies (8.10) and

$$\left\| A_\nu - A_\nu^0 \right\|_{W^{1,p}(\Sigma,*_s)} \le c_3 \|A_\nu - A_\nu''\|_{W^{1,p}(\Sigma,*_s)}$$

with a suitable constant $c_3 > 0$. Since A_ν and A_ν'' are smooth so is g_ν. Hence A_ν^0 satisfies the requirements of step 5. In particular the last estimate follows from the identity

$$\partial_t A_\nu^0 + d_{A_\nu^0} \Psi_\nu^0 = g_\nu^{-1}(\partial_t A_\nu'' + d_{A_\nu''}\Psi_\nu'')g_\nu.$$

Step 6: *Choose* Φ_ν^0 *and* Ψ_ν^0 *such that*

$$d_{A_\nu^0} *_s \left(\partial_s A_\nu^0 - d_{A_\nu^0}\Phi_\nu^0 - X_s(A_\nu^0)\right) = 0, \qquad d_{A_\nu^0} *_s \left(\partial_t A_\nu^0 - d_{A_\nu^0}\Psi_\nu^0\right) = 0.$$

Then $\Xi_\nu^0 \in \mathcal{A}_0^{1,p}(a^-, a^+, H)$ *and there exist constants* $c > 0$ *and* $\nu_0 \in \mathbb{N}$ *such that*

$$\left\| \Xi_\nu - \Xi_\nu^0 \right\|_{1,p,\varepsilon_\nu,\Xi_\nu^0} \le c\varepsilon_\nu^{1+2/p}$$

for $\nu \ge \nu_0$.

Again we shall suppress the subscript ν and write A, A^0, ε for A_ν, A_ν^0, ε_ν etc. Moreover, denote $\nabla_t{}^0 = \partial_t + \Psi^0$, $\nabla_s{}^0 = \partial_t + \Phi^0$, $B_t^0 = \partial_t A^0 - d_{A^0}\Psi^0$, and $B_s^0 = \partial_s A^0 - d_{A^0}\Phi^0$. The identity

$$d_{A^0}(A - A^0) = F_A - \tfrac{1}{2}[A - A^0 \wedge A - A^0]$$

shows that

$$\left\| d_{A^0}(A - A^0) \right\|_{L^p} + \left\| d_{A^0} *_s (A - A^0) \right\|_{L^p} \le c_1 \varepsilon^2.$$

As in the proof of step 3 we have

$$d_{A^0}(\Psi - \Psi^0) = \nabla_t(A - A^0) + B_t^0 - B_t. \tag{8.11}$$

Differentiating the identity $d_{A^0} *_s (A - A^0) = 0$ with respect to t gives

$$d_{A^0} *_s \nabla_t{}^0(A - A^0) = [A - A^0 \wedge *_s B_t^0]$$

and hence

$$d_{A^0} *_s d_{A^0}(\Psi - \Psi^0) = -d_A *_s B_t + [A - A^0 \wedge *_s(B_t + B_t^0 - d_{A^0}(\Psi - \Psi^0))]. \tag{8.12}$$

This implies

$$\left\| d_{A^0}(\Psi - \Psi^0) \right\|_{L^p} \le c_2 \varepsilon^{2/p}.$$

By step 5, we have

$$\left\| B_t - B_t^0 \right\|_{\mathrm{L}^p} \le c_3 \varepsilon^{2/p}$$

and hence it follows from (8.11)

$$\left\| \nabla_t{}^0 (A - A^0) \right\|_{\mathrm{L}^p} \le c_4 \varepsilon^{2/p}.$$

To estimate $\nabla_t{}^0(\Psi - \Psi^0)$ and $\nabla_s{}^0(\Psi - \Psi^0)$ by $\varepsilon^{2/p-1}$ differentiate the identity (8.12) with respect to t and s. The expressions $\mathrm{d}_{A^0}(\Phi - \Phi^0)$, $\nabla_s{}^0(A - A^0)$, $\nabla_t{}^0(\Phi - \Phi^0)$, and $\nabla_s{}^0(\Phi - \Phi^0)$ can be estimated by similar arguments. This proves step 6.

Step 7: *For ν sufficiently large there exist $g_\nu \in \mathcal{G}^{2,p}$ and $\tau_\nu \in \mathbb{R}$ such that*

$$g_\nu^*(\Xi_\nu \circ \sigma_{\tau_\nu}) = \mathcal{T}_{\varepsilon_\nu}(\Xi_\nu^0).$$

The real numbers act on the moduli space $\mathcal{M}_0(a^-, a^+, H)$ by time shift. Since $\mu(a^-, a^+) = 1$ the quotient $\mathcal{M}_0(a^-, a^+, H)/\mathbb{R}$ consists only of finitely many points. Now the constants δ and ε_0 of Theorem 6.1 are invariant under gauge transformations and time shift. Hence step 7 follows from step 6 and Theorem 6.1. This proves Theorem 8.1. □

9. Bubbling

In this section we prove that the assumption of bounded curvature in Theorem 8.1 is necessarily satisfied when the index difference is 1 or the energy is sufficiently small.

THEOREM 9.1. *Let $a^\pm \in \mathcal{A}_\Sigma(P_f, H)$ and assume that either $\mathcal{CS}_H(a^-) - \mathcal{CS}_H(a^+) < 8\pi^2$ or $H \in \mathcal{H}_0^{\mathrm{reg}}$ and $\mu_H(a^-, a^+) \le 3$. Then there exist constants $c_0 > 0$ and $\varepsilon_0 > 0$ such that*

$$\varepsilon^{-2} \left\| F_A \right\|_{\mathrm{L}^\infty} + \left\| \partial_t A - d_A \Psi \right\|_{\mathrm{L}^\infty} \le c_0$$

for every $\Xi \in \mathcal{A}_\varepsilon^{1,p}(a^-, a^+, H)$ with $0 < \varepsilon \le \varepsilon_0$.

Theorem 5.1 asserts that every holomorphic curve $\Xi_0 \in \mathcal{A}_0^{1,p}(a^-, a^+, H)$ can be approximated by self-dual instantons $\Xi_\varepsilon \in \mathcal{A}_\varepsilon^{1,p}(a^-, a^+, H)$. The next theorem asserts that when the relative Morse index is 1 then for ε sufficiently small every self-dual instanton connecting a^- to a^+ can be obtained this way.

THEOREM 9.2. *Assume $H \in \mathcal{H}_0^{\mathrm{reg}}$ and $a^\pm \in \mathcal{A}_{\mathrm{flat}}(P_f, H)$ such that $\mu_H(a^-, a^+) = 1$. Then the map $\mathcal{T}_\varepsilon : \mathcal{A}_0^{1,p}(a^-, a^+, H) \to \mathcal{A}_\varepsilon^{1,p}(a^-, a^+, H)$ of Theorem 5.1 is onto for $\varepsilon > 0$ sufficiently small.*

Proof. Theorem 8.1 and Theorem 9.1 □

The proof of Theorem 9.1 involves a bubbling argument. Roughly speaking, a sequence of ε_ν-self-dual instantons with $\varepsilon_\nu \to 0$ may not satisfy a uniform L^∞-estimate in arbitrarily small neighborhoods of finitely many points and in this case either instantons on S^4 or instantons on $\mathbb{C} \times P$ or holomorphic spheres in $\mathcal{M}(P)$ will split off. But this cannot happen when the relative Morse index is 1. We use the following observation due to Hofer.

LEMMA 9.3. *Let M be a complete metric space and $f : M \to \mathbb{R}$ be continuous and nonnegative. Given $x \in M$ and $r > 0$ there exist $\xi \in M$ and $0 < \rho < r$ such that*

$$d(x, \xi) \le r, \qquad \sup_{B_\rho(\xi)} f \le 2f(\xi), \qquad \rho f(\xi) \ge r f(x)/2.$$

Proof of Theorem 9.1. Assume that the statement were false. Then there would exist a sequence $\Xi_\nu \in \mathcal{A}^{1,p}_{\varepsilon_\nu}(a^-, a^+, H)$ with $\varepsilon_\nu \to 0$ such that

$$\varepsilon_\nu^{-2} \|F_{A_\nu}\|_{\mathrm{L}^\infty} + \|\partial_t A_\nu - d_{A_\nu}\Psi_\nu\|_{\mathrm{L}^\infty} \to \infty. \tag{9.1}$$

We first prove that a subsequence (still denoted by Ξ_ν) satisfies the estimate

$$\sup_\nu \left(\varepsilon_\nu^{-2} \|F_{A_\nu}\|_{\mathrm{L}^\infty(K)} + \|\partial_t A_\nu - \mathrm{d}_{A_\nu}\Psi_\nu\|_{\mathrm{L}^\infty(K)}\right) < \infty \tag{9.2}$$

for every compact subset $K \subset (\mathbb{C} \setminus W) \times \Sigma$ where $W \subset \mathbb{C}$ is a discrete set (to be constructed) which intersects $[0,1] + i\mathbb{R}$ in a finite set. If (9.2) does not hold for some compact set $K \subset \mathbb{C} \times \Sigma$ then there exists a bounded sequence $w_\nu \in \mathbb{C}$ such that

$$c_\nu = c_\nu(w_\nu) = \varepsilon_\nu^{-1} \left\|F_{A_\nu(w_\nu)}\right\|^{1/2}_{\mathrm{L}^\infty(\Sigma)} + \left\|\partial_t A_\nu(w_\nu) - \mathrm{d}_{A_\nu(w_\nu)}\Psi_\nu(w_\nu)\right\|_{\mathrm{L}^\infty(\Sigma)}$$

diverges to ∞. (Pass to a subsequence if necessary.) Assume without loss of generality that w_ν converges and denote its limit by $w_0 = s_0 + it_0$. There are three cases.

Instantons on S^4

Assume that the sequence $\varepsilon_\nu c_\nu$ is unbounded. Consider the self-dual instantons

$$\widetilde{\Xi}_\nu = \widetilde{A}_\nu + \widetilde{\Phi}_\nu\, ds + \widetilde{\Psi}_\nu\, dt$$

given by

$$\widetilde{A}_\nu(w) = A_\nu(w_\nu{+}\varepsilon_\nu w), \quad \widetilde{\Phi}_\nu(w) = \varepsilon_\nu \Phi_\nu(w_\nu{+}\varepsilon_\nu w), \quad \widetilde{\Psi}_\nu(w) = \varepsilon_\nu \Psi_\nu(w_\nu{+}\varepsilon_\nu w)$$

for $w = s + it$. Passing to a subsequence we may assume that $\varepsilon_\nu c_\nu \to \infty$. Hence there exists a sequence $z_\nu \in \Sigma$ such that the norm of the curvature of $\widetilde{\Xi}_\nu$ at $(0, z_\nu)$ diverges to ∞. Assume without loss of generality that z_ν converges to z_0. Then it follows from the usual renormalization argument that an instanton on S^4 splits off near $(0, z_0)$ (cf. [29]). This implies that the energy of Ξ_ν in an arbitrarily small neighbourhood of $w_\nu \times \Sigma$ is in the limit at least $16\pi^2$. Hence an instanton on S^4 can only split off near finitely many points (w_0, z_0). Let W_1 denote the discrete set of complex numbers $w_0 \in \mathbb{C}$ such that there exists a sequence $w_\nu \to w_0$ with $\sup_\nu \varepsilon_\nu c_\nu(w_\nu) = \infty$.

Instantons on $\mathbb{C} \times P$

Let $w_0 \in \mathbb{C} \setminus W_1$ and assume that there exists a sequence $w_\nu \to w_0$ such that $\varepsilon_\nu c_\nu(w_\nu) \geq \delta > 0$. Since $w_0 \notin W_1$ there exist constants $c > 0$ and $\rho > 0$ such that

$$\sup_{|w-w_\nu|\leq\rho} \varepsilon_\nu c_\nu(w) \leq c$$

for every ν. Let $\widetilde{\Xi}_\nu = \widetilde{A}_\nu + \widetilde{\Phi}_\nu \, \mathrm{d}s + \widetilde{\Psi}_\nu \, \mathrm{d}t$ be defined as above. Then $\widetilde{\Xi}_\nu$ is a sequence of self-dual connections on $\mathbb{C} \times P$

$$\begin{aligned} \partial_t \widetilde{A}_\nu - \mathrm{d}_{\widetilde{A}_\nu} \widetilde{\Psi}_\nu + *_{s_\nu + \varepsilon_\nu s} \left(\partial_s \widetilde{A}_\nu - \mathrm{d}_{\widetilde{A}_\nu} \widetilde{\Phi}_\nu \right) &= 0 \\ \partial_t \widetilde{\Phi}_\nu - \partial_s \widetilde{\Psi}_\nu - [\widetilde{\Phi}_\nu, \widetilde{\Psi}_\nu] + *_{s_\nu + \varepsilon_\nu s} F_{\widetilde{A}_\nu} &= 0. \end{aligned}$$

(Here $s_\nu = \mathrm{Re}\, w_\nu$ and $s = \mathrm{Re}\, w$.) The curvature of $\widetilde{\Xi}_\nu$ is uniformly bounded on any compact set. Hence it follows from Uhlenbeck's compactness theorem that there exists a subsequence (still denoted by $\widetilde{\Xi}_\nu$) and a sequence of gauge transformations $g_\nu : \mathbb{C} \times P \to G$ such that $g_\nu^* \widetilde{\Xi}_\nu$ converges uniformly with all derivatives on compact sets. The limit connection

$$\widetilde{\Xi} = \widetilde{A} + \widetilde{\Phi} \, \mathrm{d}s + \widetilde{\Psi} \, \mathrm{d}t = \lim_{\nu \to \infty} \widetilde{\Xi}_\nu$$

is self-dual on $\mathbb{C} \times P$ with respect to the metric $* = *_{s_0}$ on Σ where $s_0 = \mathrm{Re}\, w_0$. Since $\varepsilon_\nu c_\nu \geq \delta > 0$ it follows that

$$\left\| F_{\widetilde{A}(0)} \right\|_{L^\infty(\Sigma)} + \left\| \partial_t \widetilde{A}(0) - \mathrm{d}_{\widetilde{A}(0)} \widetilde{\Psi}(0) \right\|_{L^\infty(\Sigma)} \neq 0. \tag{9.3}$$

Now introduce polar co-ordinates $w = s + it = e^{\tau + i\theta}$ and define $\Xi = A + \Phi \, \mathrm{d}s + \Psi \, \mathrm{d}t$ by

$$\begin{aligned} A(\theta, \tau) &= \widetilde{A}(e^{\tau + i\theta}) \\ \Phi(\theta, \tau) &= e^\tau \cos\theta \, \widetilde{\Psi}(e^{\tau + i\theta}) - e^\tau \sin\theta \, \widetilde{\Phi}(e^{\tau + i\theta}) \\ \Psi(\theta, \tau) &= e^\tau \cos\theta \, \widetilde{\Phi}(e^{\tau + i\theta}) + e^\tau \sin\theta \, \widetilde{\Psi}(e^{\tau + i\theta}) \end{aligned}$$

Then

$$\begin{aligned} \partial_\tau A - \mathrm{d}_A \Psi - * (\partial_\theta A - d_A \Phi) &= 0, \\ \partial_\tau \Phi - \partial_\theta \Psi - [\Phi, \Psi] - e^{2\tau} *_s F_A &= 0 \end{aligned} \tag{9.4}$$

and the Yang-Mills action of Ξ is finite

$$\mathcal{Y}(\Xi) = \int_{-\infty}^{\infty}\int_0^{2\pi} \left(\|\partial_\tau A - \mathrm{d}_A\Psi\|_{\mathrm{L}^2(\Sigma)}^2 + e^{2\tau}\|F_A\|_{\mathrm{L}^2(\Sigma)}^2\right) \mathrm{d}\theta\mathrm{d}\tau < \infty$$

Similar arguments as in section 7 show that the curvature decays exponentially as τ tends to ∞. Moreover, we may assume without loss of generality that Ξ is in radial gauge, i.e. $\Psi = 0$ for τ sufficiently large. This implies that the limits

$$A_\infty(\theta) = \lim_{\tau\to\infty} A(\theta,\tau), \qquad \Phi_\infty(\theta) = \lim_{\tau\to\infty} \Phi(\theta,\tau)$$

exist and define a flat connection on $P \times S^1$. Hence

$$A_\infty(\theta) = g(\theta)^* A_0, \qquad \Phi_\infty(\theta) = g(\theta)^{-1}\dot g(\theta)$$

where $g(\theta) = g(\theta + 2\pi) \in \mathcal{G}_0(P)$ and $A_0 = A_\infty(0) \in \mathcal{A}_{\mathrm{flat}}(P)$. It follows that the Yang-Mills action of Ξ is given by

$$\begin{aligned}
\mathcal{Y}(\Xi) &= \int_{-\infty}^{\infty}\int_0^{2\pi} \left(\|\partial_\tau A\|_{\mathrm{L}^2(\Sigma)}^2 + e^{2r}\|F_A\|_{\mathrm{L}^2(\Sigma)}^2\right) \mathrm{d}\theta\mathrm{d}\tau \\
&= \int_{-\infty}^{\infty}\frac{\mathrm{d}}{\mathrm{d}\tau}\int_0^{2\pi}\int_\Sigma \left(\tfrac12\langle\partial_\theta A \wedge (A - A_0)\rangle + \langle F_A \wedge \Phi\rangle\right) \mathrm{d}\theta\mathrm{d}\tau \\
&= \int_0^{2\pi}\int_\Sigma \tfrac12\langle \dot A_\infty \wedge (A_\infty - A_0)\rangle \,\mathrm{d}\theta \\
&= -8\pi^2 \deg(g).
\end{aligned}$$

By (9.3) the Yang-Mills action of Ξ is positive. Hence the Yang-Mills action of Ξ_ν in an arbitrarily small neighbourhood of $w_0 \times \Sigma$ is in the limit at least $8\pi^2$. This shows that an instanton on $\mathbb{C} \times P$ can only split off near finitely many points w_0. Let W_2 denote the discrete set of complex numbers $w_0 \in \mathbb{C}$ such that there exists a sequence $w_\nu \to w_0$ with $\varepsilon_\nu c_\nu(w_\nu) \not\to 0$.

Holomorphic spheres in $\mathcal{M}(P)$

Now let $w_0 \in \mathbb{C} \setminus W_2$ and assume that there exists a sequence $w_\nu \to w_0$ such that $c_\nu = c_\nu(w_\nu)$ diverges to ∞. By Lemma 9.3 we may assume that there exists a sequence $0 < \rho_\nu < 1/2$ such that

$$\sup_{|w - w_\nu| \le \rho_\nu} c_\nu(w) \le 2c_\nu(w_\nu), \qquad \rho_\nu c_\nu(w_\nu) \to \infty. \tag{9.5}$$

Moreover, since $w_0 \notin W_2$

$$\lim_{\nu\to\infty} \varepsilon_\nu c_\nu(w_\nu) = 0.$$

Now define $\widetilde\Xi_\nu = \widetilde A_\nu + \widetilde\Phi_\nu\,\mathrm{d}s + \widetilde\Psi_\nu\,\mathrm{d}t$ by $\widetilde A_\nu(w) = A_\nu(w_\nu + c_\nu^{-1}w)$, $\widetilde\Phi_\nu(w) = c_\nu^{-1}\Phi_\nu(w_\nu + c_\nu^{-1}w)$, and $\widetilde\Psi_\nu(w) = c_\nu^{-1}\Psi_\nu(w_\nu + c_\nu^{-1}w)$. This sequence satisfies the partial differential equation

$$\begin{aligned} \partial_t\widetilde{A}_\nu - \mathrm{d}_{\widetilde{A}_\nu}\widetilde{\Psi}_\nu + *_{s_\nu+s/c_\nu}\left(\partial_s\widetilde{A}_\nu - \mathrm{d}_{\widetilde{A}_\nu}\widetilde{\Phi}_\nu\right) &= 0, \\ \partial_t\widetilde{\Phi}_\nu - \partial_s\widetilde{\Psi}_\nu - [\widetilde{\Phi}_\nu,\widetilde{\Psi}_\nu] + \frac{1}{\varepsilon_\nu^2 c_\nu^2} *_{s_\nu+s/c_\nu} F_{\widetilde{A}_\nu} &= 0. \end{aligned}$$

Moreover, the Yang-Mills action is finite over $B_{\rho_\nu c_\nu}(0)\times\Sigma$:

$$\int_{s^2+t^2\le\rho_\nu^2 c_\nu^2} \left(\left\|\partial_t\widetilde{A}_\nu - \mathrm{d}_{\widetilde{A}_\nu}\widetilde{\Psi}_\nu\right\|^2_{\mathrm{L}^2(\Sigma)} + \frac{1}{\varepsilon_\nu^2 c_\nu^2}\left\|F_{\widetilde{A}_\nu}\right\|^2_{\mathrm{L}^2(\Sigma)}\right)\mathrm{d}s\mathrm{d}t \le c,$$

and, by (9.5), the curvature is bounded:

$$\sup_{|w|\le\rho_\nu c_\nu}\left(\frac{1}{\varepsilon_\nu^2 c_\nu^2}\left\|F_{\widetilde{A}_\nu(w)}\right\|_{\mathrm{L}^\infty(\Sigma)} + \left\|\partial_t\widetilde{A}_\nu(w) - \mathrm{d}_{\widetilde{A}_\nu(w)}\widetilde{\Psi}_\nu(w)\right\|_{\mathrm{L}^\infty(\Sigma)}\right) \le 6.$$

By Uhlenbeck's weak compactness theorem (cf. [32]) we may assume, passing to a subsequence and up to gauge equivalence, that $\widetilde{\Xi}_\nu$ converges strongly in L^∞ and weakly in $\mathrm{W}^{1,p}$ on compact sets. The limit connection Ξ_0 represents a non-constant holomorphic map $\mathbb{C}\to\mathcal{M}(P)$ with respect to the conformal structure $* = *_{s_0}$

$$\partial_t A_0 - \mathrm{d}_{A_0}\Psi_0 + *\left(\partial_s A_0 - \mathrm{d}_{A_0}\Phi_0\right), \qquad F_{A_0} = 0.$$

Since $\rho_\nu c_\nu(w_\nu)\to\infty$ this holomorphic curve has finite energy

$$E(\Xi_0) = \int_{\mathbb{C}} \|\partial_t A_0 - \mathrm{d}_{A_0}\Psi_0\|^2_{\mathrm{L}^2(\Sigma)}\,\mathrm{d}s\mathrm{d}t \le c.$$

By the removable singularity theorem Ξ_0 extends to a nonconstant holomorphic sphere $v_0 : S^2\to\mathcal{M}(P)$. The energy of such a holomorphic sphere is at least $8\pi^2$ (cf. [10]). Hence a holomorphic sphere on $\mathcal{M}(P)$ can only split off near finitely many points w_0. Thus we have proved that the set $W\subset\mathbb{C}$ of all points w_0 such that there exists a sequence $w_\nu\to w_0$ with $c_\nu(w_\nu)\to\infty$ intersects $[0,1]+i\mathbb{R}$ in a finite set. We must prove that this set is empty.

Assume, by contradiction that W is nonempty. By (9.2) and Uhlenbeck's weak compactness theorem, we may assume that Ξ_ν converges, modulo gauge equivalence, on the complement of W to a connecting orbit $\Xi_0\in\mathcal{A}_0^{1,p}(a_0,a_1,H)$ for some H-flat connections $a_0,a_1\in\mathcal{A}_{\mathrm{flat}}(P_f,H)$. We may assume without loss of generality that $a_0=a^-$ and proceed by induction as in [26] to obtain finitely many such limit trajectories

$$\Xi_j\in\mathcal{A}_0^{1,p}(a_j,a_{j+1},H), \qquad j=0,\ldots,\ell-1,$$

with $a_\ell = a^+$. Now each bubble carries energie at least $8\pi^2$. Since $W \neq \emptyset$ it follows that the limit connections have total energy

$$\sum_{j=1}^{\ell-1} E(\Xi_j) \leq E - 8\pi^2$$

where $E = \mathcal{CS}_H(a^-) - \mathcal{CS}_H(a^+)$ is the perturbed Yang-Mills action of the connections Ξ_ν. This is not possible if $E < 8\pi^2$. Hence in this case it follows that $W = \emptyset$. If $\mu_H(a^-, a^+) \leq 3$ and $W \neq \emptyset$ then the index formula (3.7) shows that

$$\begin{aligned}
\sum_{j=1}^{\ell-1} \mu_H(a_j, a_{j+1}) &= \sum_{j=1}^{\ell-1}\Big(\tfrac{1}{2}\eta(\mathrm{D}_{a_{j+1}}) - \tfrac{1}{2}\eta(\mathrm{D}_{a_j}) + \frac{1}{2\pi^2}E(\Xi_j)\Big) \\
&= \tfrac{1}{2}\eta(\mathrm{D}_{a^+}) - \tfrac{1}{2}\eta(\mathrm{D}_{a^-}) + \frac{1}{2\pi^2}\sum_{j=1}^{\ell-1}E(\Xi_j) \\
&\leq \tfrac{1}{2}\eta(\mathrm{D}_{a^+}) - \tfrac{1}{2}\eta(\mathrm{D}_{a^-}) + \frac{1}{2\pi^2}E - 4 \\
&= \mu_H(a^-, a^+) - 4 \\
&\leq -1
\end{aligned}$$

Hence $\mu_H(a_j, a_{j+1}) < 0$ for some j. But since $H \in \mathcal{H}_0^{\mathrm{reg}}$ it follows that for this value of j the set $\mathcal{A}_0^{1,p}(a_j, aj+1, H)$ must be empty. This is a contradiction and shows that $W = \emptyset$ whenever $\mu_H(a^-, a^+) \leq 3$ and $H \in \mathcal{H}_0^{\mathrm{reg}}$.

Thus we have proved in both cases that $W = \emptyset$. Hence the estimate (9.2) holds for every compact subset $K \subset \mathbb{C}$. It continues to hold when Ξ_ν is replaced by $\Xi_\nu \circ \sigma_{\tau_\nu}$ with any sequence $\tau_\nu \in \mathbb{R}$. Hence (9.2) holds for $K = \mathbb{C}$ in contradiction to (9.1). This proves the theorem. □

10. The main theorem

THEOREM 10.1. *There is a natural isomorphism of Floer homologies*

$$HF_*^{\mathrm{inst}}(\Sigma_h, P_f) = HF_*^{\mathrm{symp}}(\mathcal{M}(P), \phi_f).$$

In particular, for $f = \mathrm{id}$,

$$HF_*^{\mathrm{inst}}(\Sigma \times S^1, P \times S^1) = HF_*(\mathcal{M}(P), \mathbb{Z}).$$

Coherent orientation

We follow the line of argument in [18]. Let $H \in \mathcal{H}^{\mathrm{reg}}(0)$ and fix two H-flat connections $a^\pm \in \mathcal{A}_\Sigma(P_f, H)$. Let $p > 4$ and denote

$$\mathcal{A}_\Sigma^{1,p}(a^-, a^+) = \Big\{\Xi = A + \Phi\, ds + \Psi\, dt \in \mathcal{A}^{1,p}(a^-, a^+) \,:\, F_A = 0\Big\}.$$

This space is nonempty, connected, and simply connected. For every $\Xi \in \mathcal{A}^{1,p}_{\Sigma}(a^-, a^+)$ there are Fredholm operators $\mathcal{D}_0(\Xi)$ and $\mathcal{D}_\varepsilon(\Xi)$. Consider the determinant line bundle

$$\mathcal{L}_0 \to \mathcal{A}^{1,p}_{\Sigma}(a^-, a^+)$$

whose fibre at Ξ is the 1-dimensional real vector space

$$\det(\mathcal{D}_0(\Xi)) = \Lambda^{\max}(\ker \mathcal{D}_0(\Xi)) \otimes \Lambda^{\max}(\ker \mathcal{D}_0(\Xi)^*)$$

The bundle $\mathcal{L}_\varepsilon \to \mathcal{A}^{1,p}_{\Sigma}(a^-, a^+)$ is defined similarly. Since $\mathcal{A}^{1,p}_{\Sigma}(a^-, a^+)$ is simply connected both line bundles are orientable. For the bundle $\mathcal{L}_\varepsilon$ this also follows from the fact that it extends to a determinant line bundle over the affine space

$$\mathcal{L}_\varepsilon \to \mathcal{A}^{1,p}(a^-, a^+).$$

Denote by $\mathrm{Or}_0(a^-, a^+)$ and $\mathrm{Or}_\varepsilon(a^-, a^+)$ the spaces of orientations of $\mathcal{L}_0$ and $\mathcal{L}_\varepsilon$, respectively, each consisting of 2 elements. The spaces $\mathrm{Or}_\varepsilon(a^-, a^+)$ for different values of ε are naturally isomorphic.

PROPOSITION 10.2. *For every pair $a^\pm \in \mathcal{A}_{\mathrm{flat}}(P_f, H)$ there is a natural bijection*

$$\tau_\varepsilon(a^-, a^+) : \mathrm{Or}_0(a^-, a^+) \to \mathrm{Or}_\varepsilon(a^-, a^+).$$

Proof. If $\mathcal{D}_0(\Xi)$ is onto and $\varepsilon > 0$ is sufficiently small then, by Lemma 4.5 and Lemma 4.3, there is a linear bijection

$$\mathrm{T}_\varepsilon(\Xi) : \ker \mathcal{D}_0(\Xi) \to \ker \mathcal{D}_\varepsilon(\Xi)$$

given by

$$\mathrm{T}_\varepsilon(\Xi)\alpha_0 = \xi_0 - \mathcal{D}_\varepsilon^*(\mathcal{D}_\varepsilon \mathcal{D}_\varepsilon^*)^{-1}\mathcal{D}_\varepsilon \xi_0.$$

Here $\xi_0 = \alpha_0 + \phi_0\, \mathrm{d}s + \psi_0\, \mathrm{d}t$ and ϕ_0 and ψ_0 are determined by the requirement that

$$\mathcal{D}_0(\Xi)\alpha_0 = \nabla_t \alpha_0 - \mathrm{d}_A \psi_0 + *_s(\nabla_s \alpha_0 - \mathrm{d}X_s(A)\alpha_0 - \mathrm{d}_A \phi_0)$$

is harmonic. If $\mathcal{D}_0$ is not onto choose a number N and a linear map $\mathrm{L} : \mathbb{R}^N \to \mathrm{L}^p_f(\mathrm{H}_A)$ such that

$$\mathcal{D}_0 \oplus \mathrm{L} : \mathrm{W}^{1,p}_f(\mathrm{H}_A) \oplus \mathbb{R}^N \to \mathrm{L}^p_f(\mathrm{H}_A).$$

is onto. As in Lemma 4.5 one can show that the operator $\mathcal{D}_\varepsilon \oplus L$ is onto for ε sufficiently small and there is a linear bijection

$$\det(\mathcal{D}_0) \simeq \Lambda^{\max}(\ker(\mathcal{D}_0 \oplus \mathrm{L})) \to \Lambda^{\max}(\ker(\mathcal{D}_\varepsilon \oplus \mathrm{L})) \simeq \det(\mathcal{D}_\varepsilon).$$

The induced map $\mathrm{Or}(\det(\mathcal{D}_0)) \to \mathrm{Or}(\det(\mathcal{D}_\varepsilon))$ is independent of the extension L used to define it. □

Now for any three H-flat connections $a_0, a_1, a_2 \in \mathcal{A}_{\mathrm{flat}}(P_f, H)$ Floer's glueing construction determines a natural map

$$\sigma_0(a_0, a_1, a_2) : \mathrm{Or}_0(a_0, a_1) \oplus \mathrm{Or}_0(a_1, a_2) \to \mathrm{Or}_0(a_0, a_2).$$

A **coherent orientation** for $\mathcal{L}_0$ is a collection of orientations $\sigma_0(a^-, a^+) \in \mathrm{Or}_0(a^-, a^+)$ such that

$$\sigma_0(a_0, a_1, a_2)(\sigma_0(a_0, a_1), \sigma_0(a_1, a_2)) = \sigma_0(a_0, a_2).$$

Similarly for $\mathcal{L}_\varepsilon$. In [18] it is shown that such coherent orientations exist. The next proposition asserts that τ_ε and σ commute and hence every coherent orientation for $\mathcal{L}_0$ is mapped under τ_ε to a coherent orientation for $\mathcal{L}_\varepsilon$.

Proposition 10.3.

$$\sigma_\varepsilon(a_0, a_1, a_2) \circ \tau_\varepsilon(a_0, a_1) \oplus \tau_\varepsilon(a_1, a_2) = \tau_\varepsilon(a_0, a_2) \circ \sigma_0(a_0, a_1, a_2)$$

Proof. Let $\Xi \in \mathcal{A}^{1,p}_\Sigma(a_0, a_1)$ and $\Xi' \in \mathcal{A}^{1,p}_\Sigma(a_1, a_2)$ such that $\Psi(s,t) = \Psi'(s,t) = 0$, $A(s,t) + \Phi(s,t)\,\mathrm{d}s = a_1$ for $t \geq T$, and $A'(s,t) + \Phi'(s,t)\,\mathrm{d}s = a_1$ for $t \leq -T$. Assume without loss of generality that $\mathcal{D}_0(\Xi)$ and $\mathcal{D}_0(\Xi')$ are onto.

For $R > T$ define the catenation $\Xi''_R = \Xi \#_R \Xi' \in \mathcal{A}^{1,p}_\Sigma(a_0, a_2)$ of Ξ and Ξ' by $\Xi''(s,t) = \Xi(s, t+R)$ for $t \leq 0$ and $\Xi''_R(s,t) = \Xi'(s, t-R)$ for $t \geq 0$. Then for $R > 0$ sufficiently large there exist isomorphisms

$$\mathrm{S}_0 : \ker\, \mathcal{D}_0(\Xi) \oplus \ker\, \mathcal{D}_0(\Xi') \to \ker\, \mathcal{D}_0(\Xi''_R),$$

$$\mathrm{S}_\varepsilon : \ker\, \mathcal{D}_\varepsilon(\Xi) \oplus \ker\, \mathcal{D}_\varepsilon(\Xi') \to \ker\, \mathcal{D}_\varepsilon(\Xi''_R).$$

These maps are small perturbations of the obvious shift-overlap maps $(\xi, \xi') \mapsto \xi''(s,t) = \xi(s, t+R) + \xi'(s, t-R)$. They induce the maps $\sigma_0(a_0, a_1, a_2)$ and $\sigma_\varepsilon(a_0, a_1, a_2)$ on the spaces of orientations. Now let $\mathrm{T}_\varepsilon(\Xi)$ be defined as in the proof of Proposition 10.2. Then the linear operators

$$\mathrm{S}_\varepsilon \circ \mathrm{T}_\varepsilon(\Xi) \oplus \mathrm{T}_\varepsilon(\Xi') : \ker\, \mathcal{D}_0(\Xi) \oplus \ker\, \mathcal{D}_0(\Xi') \to \ker\, \mathcal{D}_\varepsilon(\Xi''_R),$$

$$\mathrm{T}_\varepsilon(\Xi''_R) \circ \mathrm{S}_0 : \ker\, \mathcal{D}_0(\Xi) \oplus \ker\, \mathcal{D}_0(\Xi') \to \ker\, \mathcal{D}_\varepsilon(\Xi''_R)$$

are close to each other for $\varepsilon > 0$ sufficiently small and $R > 0$ sufficiently large. This is because the maps S_0 and S_ε for large $R > 0$ are close to the shift-overlap maps while $\mathrm{T}_\varepsilon(\Xi)$ for small $\varepsilon > 0$ is close to the identity. The details are left to the reader. (For maps of the form S_0 and S_ε see [18].) □

Proof of Theorem 10.1. Choose $\varepsilon > 0$ sufficiently small and let $H \in \mathcal{H}^{\mathrm{reg}}_0 \cap \mathcal{H}^{\mathrm{reg}}_\varepsilon$. Throughout fix a sequence of H-flat connections $a_0, a_1, a_2, \ldots$

such that each equivalence class in $\mathcal{A}_{\mathrm{flat}}(P_f, H)/\mathcal{G}_0(P_f)$ is represented by precisely one member of this sequence. Both Floer homology groups are generated by the same chain complex

$$C_k = \bigoplus_{\mu_H(a_0,a_j)=k} \mathbb{Z}[a_j].$$

By Theorem 3.2 the grading of this chain complex is the same in both theories. We must prove that the boundary operators ∂^{symp} and $\partial_\varepsilon^{\mathrm{inst}}$ agree for ε sufficiently small.

Choose coherent orientations $\sigma_0(a^-, a^+)$ for $\mathcal{L}_0$ and consider the induced coherent orientations $\sigma_\varepsilon(a^-, a^+)$ for $\mathcal{L}_\varepsilon$. These determine orientations of the moduli spaces $\mathcal{M}_0$ and $\mathcal{M}_\varepsilon$ and hence of the quotient spaces $\widetilde{\mathcal{M}}_0 = \mathcal{M}_0/\mathbb{R}$ and $\widetilde{\mathcal{M}}_\varepsilon = \mathcal{M}_\varepsilon/\mathbb{R}$ of connecting orbits modulo time shift. These orientations are invariant under Floer's glueing maps

$$\mathcal{S}_0(a_j, a_k, a_\ell) : \widetilde{\mathcal{M}}_0(a_j, a_k, H) \times (R, \infty) \times \widetilde{\mathcal{M}}_0(a_k, a_\ell, H) \to \widetilde{\mathcal{M}}_0(a_j, a_\ell, H).$$

in the symplectic case and under the corresponding maps $\mathcal{S}_\varepsilon(a_j, a_k, a_\ell)$ in the instanton case. To see this note that the induced maps on the spaces of orientations are given by $\sigma_0(a_j, a_k, a_\ell)$ and $\sigma_\varepsilon(a_j, a_k, a_\ell)$, respectively. Now fix two H-flat connections $a^\pm \in \mathcal{A}_{\mathrm{flat}}(P_f, H)$ and consider the map

$$\mathcal{T}_\varepsilon : \mathcal{M}_0(a^-, a^+, H) \to \mathcal{M}_\varepsilon(a^-, a^+, H)$$

of Theorem 5.1. The induced map on the space of orientations agrees with the map $\tau_\varepsilon(a^-, a^+)$ of Proposition 10.2. Hence $\mathcal{T}_\varepsilon$ is orientation preserving.

Now assume $\mu_H(a^-, a^+) = 1$. Then, by Proposition 5.7 and Theorem 9.2, the map $\mathcal{T}_\varepsilon$ is bijective and hence induces a bijection of finite sets

$$\widetilde{\mathcal{T}}_\varepsilon : \widetilde{\mathcal{M}}_0(a^-, a^+, H) \to \widetilde{\mathcal{M}}_\varepsilon(a^-, a^+, H)$$

for $\varepsilon > 0$ sufficiently small. By the above argument, this map preserves the coherent orientations. Moreover, the differential $d\mathcal{T}_\varepsilon(\Xi_0)$ satisfies

$$d\mathcal{T}_\varepsilon(\Xi_0)\partial_t \Xi_0 = \partial_t \Xi_\varepsilon$$

where $\Xi_\varepsilon = \mathcal{T}_\varepsilon(\Xi_0)$. Since $\mu_H(a^-, a^+) = 1$ the vector $\partial_t \Xi_0$ determines the flow orientation of $\mathcal{M}_0(a^-, a^+, H)$ while $\partial_t \Xi_\varepsilon$ determines the flow orientation of $\mathcal{M}_\varepsilon(a^-, a^+, H)$. Hence $\mathcal{T}_\varepsilon$ preserves both the coherent orientation and the flow orientation. Hence it preserves the signs $\nu^{\mathrm{symp}}(\Xi_0)$ and $\nu_\varepsilon^{\mathrm{inst}}(\Xi_\varepsilon)$ which are determined by comparing both orientations. This shows that the oriented number of connecting orbits from a^- to a^+ is the same in both theories. Hence the boundary operators ∂^{symp} and $\partial_\varepsilon^{\mathrm{inst}}$ agree for $\varepsilon > 0$ sufficiently small. This proves the theorem. □

University of Warwick, Coventry CV4 7AL, Great Britain

References

[1] M.F. Atiyah, New invariants of three and four dimensional manifolds, *Proc. Symp. Pure Math.* **48** (1988).

[2] M.F. Atiyah and R. Bott, The Yang Mills equations over Riemann surfaces, *Phil. Trans. R. Soc. Lond. A* **308** (1982), 523–615.

[3] M.F. Atiyah, V.K. Patodi and I.M. Singer, Spectral asymmetry and Riemannian geometry III, *Math. Proc. Camb. Phil. Soc.* **79** (1976), 71–99.

[4] C.C. Conley and E. Zehnder, Morse-type index theory for flows and periodic solutions of Hamiltonian equations, *Commun. Pure Appl. Math.* **37** (1984), 207–253.

[5] G.D. Daskalopoulos and K.K. Uhlenbeck, An application of transversality to the topology of the moduli space of stable bundles, Preprint, 1990.

[6] S.K. Donaldson, The orientation of the Yang-Mills moduli spaces and 4-manifold topology, *J. Diff. Geom.* **26** (1987), 397–428.

[7] S. Donaldson and P.B. Kronheimer, *The Geometry of Four-Manifolds*, Clarendon Press, Oxford 1990.

[8] S. Donaldson, M. Furuta and D. Kotschick, *Floer homology groups in Yang-Mills theory*, in preparation.

[9] S. Dostoglou and D.A. Salamon, Instanton homology and symplectic fixed points, University of Warwick, preprint, 1990.

[10] S. Dostoglou and D.A. Salamon, Cauchy-Riemann operators, self-duality, and the spectral flow, University of Warwick, preprint, 1992.

[11] A. Floer, The unregularized gradient flow of the symplectic action, *Comm. Pure Appl. Math.* **41** (1988), 775–813.

[12] A. Floer, Morse theory for Lagrangian intersections, *J. Diff. Geom.* **28** (1988), 513–547.

[13] A. Floer, A relative Morse index for the symplectic action, *Commun. Pure Appl. Math.* **41** (1988), 393–407.

[14] A. Floer, An instanton invariant for 3-manifolds, *Commun. Math. Phys.* **118** (1988), 215–240.

[15] A. Floer, Symplectic fixed points and holomorphic spheres, *Commun. Math. Phys.* **120** (1989), 575–611.

[16] A. Floer, Instanton homology and Dehn surgery, Preprint 1991.

[17] A. Floer, Instanton homology for knots, Preprint 1991.

[18] A. Floer and H. Hofer, Coherent orientations for periodic orbit problems in symplectic geometry, Preprint, Ruhr-Universität Bochum, 1990.

[19] D. Freed and K. Uhlenbeck, *Instantons and Four-Manifolds*, Springer-Verlag, 1984.

[20] M. Gromov, Pseudoholomorphic curves in symplectic manifolds, *Invent. Math.* **82** (1985), 307–347.

[21] H. Hofer and D.A. Salamon, Floer homology and Novikov rings, University of Warwick, preprint, 1992.

[22] J.D.S. Jones, J.H. Rawnsley and D.A. Salamon, Instanton homology, University of Warwick, preprint, 1992.

[23] D. McDuff, Elliptic methods in symplectic geometry, *Bull. A.M.S.* **23** (1990), 311–358.

[24] D. McDuff and D.A. Salamon, *Symplectic Topology*, in preparation.

[25] J.W. Robbin, and D. Salamon, The spectral flow and the Maslov index, University of Warwick, preprint, 1992.

[26] D.A. Salamon, Morse theory, the Conley index and Floer homology, *Bull. L.M.S.* **22** (1990), 113–140.

[27] D.A. Salamon and E. Zehnder, Morse theory for periodic orbits of Hamiltonian systems and the Maslov index, *Comm. Pure Appl. Math.*, to appear.

[28] S. Smale, An infinite dimensional version of Sard's theorem, *Am. J. Math.* **87** (1973), 213–221.

[29] C.H. Taubes, Self-dual Yang-Mills connections on non-self-dual four-manifolds, *J. Diff. Geom.* **17** (1982), 139–170.
[30] C.H. Taubes, L^2 moduli spaces on 4-manifolds with cylindrical ends, Harvard University, preprint, 1991.
[31] C.H. Taubes, Casson's invariant and gauge theory, *J. Diff. Geom.* **31** (1990), 547–599.
[32] K. Uhlenbeck, Connections with L^p bounds on curvature *Commun. Math. Phys.* **83** (1982), 31–42.
[33] E. Witten, Supersymmetry and Morse theory, *J. Diff. Geom.* **17** (1982), 661–692.
[34] T. Yoshida, Floer homology and splittings of manifolds, *Ann. of Math.* **134** (1991), 277–323.

Received May 19, 1992.