Max-Planck-Institut für Mathematik
Bonn

Morse homology for the Yang-Mills gradient flow

by

Jan Swoboda

Max-Planck-Institut für Mathematik
Preprint Series 2010 (119)

Morse homology for the Yang-Mills gradient flow

Jan Swoboda

Max-Planck-Institut für Mathematik
Vivatsgasse 7
53111 Bonn
Germany

MPIM 10-119

Morse Homology for the Yang-Mills Gradient Flow

Jan Swoboda (Max-Planck-Institut für Mathematik, Bonn)

January 4, 2011

Abstract

We use the Yang-Mills gradient flow on the space of connections over a closed Riemann surface to construct a Morse-Bott chain complex. The chain groups are generated by Yang-Mills connections. The boundary operator is defined by counting the elements of appropriately defined moduli spaces of Yang-Mills gradient flow lines that converge asymptotically to Yang-Mills connections.

Contents

1 Introduction

Let (Σ, g) be a closed oriented Riemann surface. Let G be a compact Lie group, $\mathfrak{g}$ its Lie algebra, and P a principal G-bundle over Σ. On $\mathfrak{g}$ we choose an ad-invariant inner product $\langle \cdot, \cdot \rangle$. The Riemannian metric induces for $k \in \{0, 1, 2\}$ the Hodge star operator $* : \Omega^k(\Sigma) \to \Omega^{2-k}(\Sigma)$ on differential k-forms. We denote by $\mathcal{A}(P)$ the affine space of $\mathfrak{g}$-valued connection 1-forms on P. The underlying vector space is the space $\Omega^0(\Sigma, \mathrm{ad}(P))$ of sections of the adjoint bundle $\mathrm{ad}(P) := P \times_{\mathrm{Ad}} \mathfrak{g}$. The curvature of a connection $A \in \mathcal{A}(P)$ is the $\mathrm{ad}(P)$-valued 2-form $F_A = dA + \frac{1}{2}[A \wedge A]$. For $A \in \mathcal{A}(P)$ we consider the *perturbed Yang-Mills functional* defined by

$$\mathcal{YM}^{\mathcal{V}}(A) = \frac{1}{2} \int_\Sigma \langle F_A \wedge *F_A \rangle + \mathcal{V}(A), \tag{1}$$

with a gauge-invariant perturbation $\mathcal{V} : \mathcal{A}(P) \to \mathbb{R}$ the precise form of which will be fixed later. The corresponding Euler-Lagrange equation is the second order partial differential equation $d_A^* F_A + \nabla \mathcal{V}(A) = 0$, called *perturbed Yang-Mills equation*. The negative L^2-gradient flow equation resulting from the Yang-Mills functional is the PDE

$$\partial_s A + d_A^* F_A + \nabla \mathcal{V}(A) = 0. \tag{2}$$

The group $\mathcal{G}(P)$ of principal fibre bundle automorphisms of P acts on the space $\mathcal{A}(P)$ by gauge transformations, i.e. as $g^* A := g^{-1} A g + g^{-1} dg$. The functional $\mathcal{YM}^{\mathcal{V}}$ in invariant under such gauge transformations, and hence are the solutions of the perturbed Yang-Mills (gradient flow) equations. The action is not free. The occuring stabilizer subgroups are subgroups of G, hence finite-dimensional. Restricting the action to the group $\mathcal{G}_0(P)$ of so-called *based gauge transformations*, i.e. those transformations which fix a prescribed fibre of P pointwise, one indeed obtains a free action. For this reason we will study solutions to the gradient flow equation (2) only up to based gauge transformations, cf. however the comment on a G-equivariant extension of the theory below.

The study of Yang-Mills connections over a Riemann surface has been initiated by Atiyah and Bott in [5]. As shown there, the gauge-equivalence classes of (unperturbed) Yang-Mills connections occur as a family of finite-dimensional submanifolds of $\mathcal{A}(P)/\mathcal{G}_0(P)$ which satisfy the so-called Morse-Bott condition. This condition asserts that, for an equivalence class $[A]$ of Yang-Mills connections, the Hessian $H_{[A]}\mathcal{YM}^{\mathcal{V}}$ is non-degenerate when restricted to the normal space at $[A]$ of the critical manifold containing $[A]$. Moreover, any Yang-Mills critical manifold is compact by a straight-forward application of Uhlenbeck's strong compactness theorem, cf. the exposition [35]. Hence the situation one encounters for the functional $\mathcal{YM}^{\mathcal{V}}$ over a Riemann surface is in precise analogy to that of finite-dimensional Morse-Bott theory. As such it has been widely studied in [5], with remarkable applications e.g. to the cohomology of moduli spaces of stable vector bundles over Σ (cf. e.g. [16] for a review of these results), however without properly developing the analysis of the underlying L^2-gradient flow (2). The aim of the present work is to introduce and work out in full detail the analytical setup for a Yang-Mills Morse homology theory over Σ.

Let us now briefly describe our setup. We shall work with so-called *abstract perturbations* $\mathcal{V}:\mathcal{A}(P)\to\mathbb{R}$, i.e. elements of a Banach space Y generated by certain gauge-invariant model perturbations $\mathcal{V}_\ell$. These are of the form

$$\mathcal{V}_\ell(A) := \rho\big(\|\alpha(A)\|_{L^2}^2\big)\langle\eta,\alpha(A)\rangle,$$

with $\rho:\mathbb{R}\to\mathbb{R}$ some cut-off function, $\eta\in\Omega^1(\Sigma,\mathrm{ad}(P))$, and $\alpha(A)=g^*A-A_0$ where $g\in\mathcal{G}(P)$ is chosen such that the local slice condition $d_{A_0}^*\alpha=0$ is satisfied. Let

$$\mathcal{P}(a) := \{A\in\mathcal{A}(P)\,|\,d_A^*F_A=0 \text{ and } \mathcal{YM}^{\mathcal{V}}(A)\leq a\}$$

denote the set of Yang-Mills connections of energy at most a. On the finite dimensional manifold $\mathcal{P}(a)$ we fix a Morse function $h:\mathcal{P}(a)\to\mathbb{R}$, i.e. a smooth function h with isolated non-degenerate critical points whose stable and unstable manifolds intersect transversally. To a critical point x of h we assign the non-negative number

$$\mathrm{Ind}(x) := \mathrm{ind}_{\mathcal{YM}^{\mathcal{V}}}(x) + \mathrm{ind}_h(x)$$

where $\mathrm{ind}_h(x)$ denotes the usual Morse index of x with respect to h and $\mathrm{ind}_{\mathcal{YM}^{\mathcal{V}}}(x)$ denotes the number of negative eigenvalues (counted with multiplicities) of the Yang-Mills Hessian $H_{[x]}\mathcal{YM}^{\mathcal{V}}$. For a regular value a of $\mathcal{YM}^{\mathcal{V}}$ we consider the $\mathbb{Z}$-module

$$CM_*^a(h) := \bigoplus_{x \in \mathrm{crit}(h) \cap \mathcal{P}(a)} \langle x \rangle$$

generated by the critical points of h of Yang-Mills energy at most a. This module is graded by the index Ind. Under certain transversality assumptions (which resemble the usual Morse-Bott transversality required in finite-dimensional Morse theory) there is a well-defined boundary operator

$$\partial_* : CM_*^a(h) \to CM_{*-1}^a(h)$$

which arises from counting so-called *cascade configurations* of (negative) L^2-gradient flow lines, i.e. tuples of solutions of

$$\partial_s A + d_A^* F_A + \nabla \mathcal{V}(A) = 0,$$

whose asymptotics as $s \to \pm\infty$ obey a certain compatibility condition. Our main result is

Theorem 1.1 *For generic perturbation $\mathcal{V} \in Y$ the map ∂_* satisfies $\partial_k \circ \partial_{k+1} = 0$ for every $k \in \mathbb{N}_0$ and thus there exist well-defined homology groups*

$$HM_k^a(\mathcal{A}(P), h) = \frac{\ker \partial_k}{\operatorname{im} \partial_{k+1}}.$$

The resulting homology $HM_*^a(h)$ is called Yang-Mills Morse homology.

In finite dimensions the construction of a Morse homology theory from the set of critical points of a Morse functions and the isolated flow lines connecting them goes back to Thom [32], Smale [28] and Milnor [19], and had later been rediscovered by Witten [37]. For a historical account we refer to the survey paper by Bott [8]. In infinite dimensions the same sort of ideas underlies the construction of Floer homology of symplectic manifolds, although the equations encountered there are of elliptic rather than parabolic type. More in the spirit of classical finite dimensional Morse homology is the so-called heat flow homology for the loop space of a compact manifold due to Weber [36], which is based on the L^2 gradient flow of the classical action functional. For another approach via the theory of ODEs on Hilbert manifolds and further references, see Abbondandolo and Majer [2]. The cascade construction of Morse homology in the presence of critical manifolds satisfying a Morse-Bott condition is due to Frauenfelder [15].

1.1 Main results

In Section 3 moduli space problem for Yang-Mills gradient flow lines with prescribed asymptotics as $s \to \pm\infty$ is put into an abstract Banach manifold setting. The moduli space $\hat{\mathcal{M}}(\mathcal{C}^-, \mathcal{C}^+)$ is exhibited as the zero set of a section $\mathcal{F}$ of a suitably defined Banach space bundle. We subsequently develop the necessary Fredholm theory for the operator $\mathcal{D}_A$ obtained by linearizing $\mathcal{F}$. The main result is as follows.

Theorem 1.2 (Fredholm theorem) *Let $A : \mathbb{R} \to \mathcal{A}(P)$, be a smooth solution of the Yang-Mills gradient flow equation* (9) *satisfying for Yang-Mills connections $A^\pm$ the asymptotic conditions*

$$\lim_{s \to \pm\infty} A(s) = A^\pm$$

in the $\mathcal{C}^\infty(\Sigma)$-topology. Then (for every $p > 1$) the operator $\mathcal{D}_A = \frac{d}{ds} + \mathcal{H}_A : \mathcal{Z}_A^{\delta,p} \to \mathcal{L}^{\delta,p}$ associated with A is a Fredholm operator of index

$$\operatorname{ind} \mathcal{D}_A = \operatorname{ind} \mathcal{H}_{A^-} - \operatorname{ind} \mathcal{H}_{A^+} - \dim \mathcal{C}^-.$$

Section 6 is concerned with compactness up to convergence to broken trajectories of the moduli space $\mathcal{M}(\mathcal{C}^-, \mathcal{C}^+)$, where we prove the following.

Theorem 1.3 (Compactness) *Let (A^ν, Ψ^ν), $\nu \in \mathbb{N}$, be a sequence of solutions to the perturbed Yang-Mills gradient flow equation*

$$\partial_s A + d_A^* F_A - d_A \Psi + \nabla \mathcal{V}(A) = 0. \tag{3}$$

Assume that there exist critical manifolds $\mathcal{C}^\pm$ such that every (A^ν, Ψ^ν) is a connecting trajectory between $\mathcal{C}^-$ and $\mathcal{C}^+$. Then for every $k < 2$ and $p < \infty$ and every compact interval $I \subset \mathbb{R}$ there exists a sequence $g^\nu \in \mathcal{G}(I \times P)$

of gauge transformations such that a subsequence of the gauge transformed sequence $(g^\nu)^*(A^\nu,\Psi^\nu)$ *converges in* $W^{k,p}(I\times\Sigma)$ *to a solution* (A^*,Ψ^*) *of* (3).

The main step here is a compactness theorem for solutions A of the Yang-Mills gradient flow equation on time intervals $I=[T_0,T_1]$ of finite length. The proof of this result relies on certain a priori L^p estimates for the curvature of A, the weak Uhlenbeck compactness theorem, and a combination of elliptic and parabolic regularity estimates. We then discuss transversality of the operator $\mathcal{F}$ at a zero $x\in\mathcal{F}^{-1}(0)$. The aim is to prove, along the usual lines using Sard's lemma, that surjectivity of the linearized operator holds for generic perturbations $\mathcal{V}\in Y$. The relevant Banach spaces of admissible perturbations are introduced in Appendix A. In Section 8.2 these results are put together and a proof of the main Theorem 1.1 is given.

1.2 Outlook

Equivariant theory

For the ease of presentation we here develop Yang-Mills Morse theory for the functional $\mathcal{YM}^{\mathcal{V}}$ on the space $\mathcal{A}(P)/G_0(P)$ of connections modulo *based* gauge equivalence. On this quotient there still acts the group G by gauge transformations as

$$g\cdot[A]=[g^*A],$$

and the functional $\mathcal{YM}^{\mathcal{V}}$ is invariant under this group action. It thus seems natural to implement this action in our setup and define a G-equivariant Morse homology, cf. [31] for details. One possible approach is to use a finite-dimensional approximation E_nG to the classifying space EG. Since the group G acts freely on E_nG we obtain a free action of the full gauge group $\mathcal{G}(P)$ on the product $\mathcal{A}(P)\times E_nG$ via

$$g^*(A,\lambda)=(g^*A,\hat{g}\lambda).$$

By extending $\mathcal{YM}^{\mathcal{V}}$ suiatably, our construction of Morse homology groups then carries over almost literally to this product manifold.

Connection with Morse homology of loop groups

The Yang-Mills Morse homology is strongly related to heat flow homology, at least in the case of the sphere $\Sigma=S^2$. This connection is due to the following result, cf. [30, 31].

Theorem 1.4 *For any compact Lie group G and any pricipal G-bundle P over Σ, the chain homomorphism*

$$\Theta : CM_*\Big(\frac{\mathcal{A}(P) \times E_nG}{\mathcal{G}(P)}\Big) \to CM_*\Big(\frac{\Omega G \times E_nG}{G}\Big)$$

induces an isomorphism

$$[\Theta] : HM_*\Big(\frac{\mathcal{A}(P) \times E_nG}{\mathcal{G}(P)}\Big) \to HM_*\Big(\frac{\Omega G \times E_nG}{G}\Big)$$

of Morse homology groups.

It would be interesting to work out a similar correspondence in the case where Σ is a Riemann surface of arbitrary genus.

Products

In finite dimensional Morse homology it is well known how to implement a module structure. In infinite dimensional situations one often encounters similar algebraic structures, like e.g. the quantum product in Floer homology or the Chas-Sullivan loop product in the Morse homology of certain loop spaces, cf. [4, 6, 9]. Using finite-dimensional Morse homology as a guiding principle, one should be able to implement a natural product structure in the setup presented here. In a subsequent step one could ask how this relates to products in loop space homology of ΩG.

Acknowledgements

This work extends the author's PhD thesis [30]. He would like to thank his supervisor Dietmar Salamon for his encouragement while working on this project. For useful comments and discussions he also thanks Werner Ballmann, Rémi Janner, Matthias Schwarz, Michael Struwe, and Joa Weber.

2 Yang-Mills functional

2.1 Preliminaries

Let (Σ, g) be a compact oriented Riemann surface. Let G be a compact Lie group with Lie algebra $\mathfrak{g}$. On $\mathfrak{g}$ we fix an ad-invariant inner product. This exists by compactness of G. Let P be a principal G-fibre bundle over Σ. A gauge transformation is a section of the bundle $\mathrm{Ad}(P) := P \times_G G$

associated to P via the action of G on itself by conjugation. Let $\mathrm{ad}(P)$ denote the Lie algebra bundle associated to P by the adjoint action of G on $\mathfrak{g}$. The spaces of smooth $\mathrm{ad}(P)$-valued differential k-forms are denoted by $\Omega^k(\Sigma, \mathrm{ad}(P))$. The space $\mathcal{A}(P)$ of smooth connections on P is an affine space over $\Omega^1(\Sigma, \mathrm{ad}(P))$. The group $\mathcal{G}(P)$ of smooth gauge transformations acts on $\mathcal{A}(P)$ and on $\Omega^k(\Sigma, \mathrm{ad}(P))$. The curvature of the connection A is $F_A = dA + \frac{1}{2}[A \wedge A] \in \Omega^2(\Sigma, \mathrm{ad}(P))$. It satisfies the Bianchi identities $d_A F_A = 0$ and $d_A^* d_A^* F_A = 0$. Covariant differentiation with respect to the Levi-Civita connection associated with the metric g and a connection $A \in \mathcal{A}(P)$ defines an operator $\nabla_A : \Omega^k(\Sigma, \mathrm{ad}(P)) \to \Omega^1(\Sigma) \otimes \Omega^k(\Sigma, \mathrm{ad}(P))$. Its antisymmetric part is the covariant exterior differentiation operator

$$d_A : \Omega^k(\Sigma, \mathrm{ad}(P)) \to \Omega^{k+1}(\Sigma, \mathrm{ad}(P)), \quad \alpha \mapsto d\alpha + [A \wedge \alpha].$$

The formal adjoints of these operators are denoted by ∇_A^* and d_A^*. The covariant Hodge Laplacian on forms is the operator $\Delta_A := d_A^* d_A + d_A d_A^*$, the covariant Bochner Laplacian on forms is $\nabla_A^* \nabla_A$. They are related through the Bochner-Weitzenböck formula

$$\nabla_A = \nabla_A^* \nabla_A + \{F_A, \cdot\} + \{R_\Sigma, \cdot\}.$$

Here the brackets $\{\cdot, \cdot\}$ denote bilinear expressions with constant coefficients. The L^2 gradient of the Yang-Mills functional $\mathcal{YM}^{\mathcal{V}}$ as in (1) at the point $A \in \mathcal{A}(P)$ is

$$\nabla \mathcal{YM}^{\mathcal{V}}(A) = d_A^* F_A + \nabla \mathcal{V}(A),$$

and its Hessian is the second order operator

$$H_A \mathcal{YM}^{\mathcal{V}} = d_A^* d_A + *[*F_A \wedge \cdot\,] + H_A \mathcal{V} : \Omega^1(\Sigma, \mathrm{ad}(P)) \to \Omega^1(\Sigma, \mathrm{ad}(P)).$$

We also use the notation $H_A := d_A^* d_A + *[*F_A \wedge \cdot\,]$. For a definition of Sobolev spaces of sections of vector bundles, of connections, and of gauge transformations we refer to [35, Appendix B]. We employ the notation $W^{k,p}(\Sigma)$ and $W^{k,p}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))$ for the Sobolev spaces of $\mathrm{ad}(P)$-valued sections, respectively $\mathrm{ad}(P)$-valued 1-forms whose weak derivatives up to order k are in L^p. Similarly, the notation $\mathcal{A}^{k,p}(P)$ indicates the Sobolev spaces of connections over Σ of class $W^{k,p}$. We also use the parabolic Sobolev spaces

$$W^{1,2;p}(I \times \Sigma, \mathrm{ad}(P)) := L^p(I, W^{2,p}(\Sigma, \mathrm{ad}(P))) \cap W^{1,p}(I, L^p(\Sigma, \mathrm{ad}(P)))$$

of $\mathrm{ad}(P)$-valued sections over $I \times \Sigma$, where $I \subseteq \mathbb{R}$ is an interval (and similarly for $\mathrm{ad}(P)$-valued 1-forms and for connections). Further notation we use frequently is $\dot{A} := \partial_s A := \frac{dA}{ds}$, etc. for derivatives with respect to time.

2.2 Perturbations

Our construction of a Banach space of perturbations is based on the following L^2-local slice theorem due to T. Mrowka and K. Wehrheim [20]. We fix $p > 2$ and let

$$\mathcal{S}_{A_0}(\varepsilon) := \{A = A_0 + \alpha \in \mathcal{A}^{0,p}(\Sigma) \,|\, d_{A_0}^* \alpha = 0, \|\alpha\|_{L^2(\Sigma)} < \varepsilon\}$$

denote the set of L^p-connections in the local slice of radius ε with respect to A_0.

Theorem 2.1 (cf. [20, Theorem 1.7]) *Let $p > 2$. For every $A_0 \in \mathcal{A}^{0,p}(\Sigma)$ there are constants $\varepsilon, \delta > 0$ such that the map*

$$\mathfrak{m} : \big(\mathcal{S}_{A_0}(\varepsilon) \times \mathcal{G}^{1,p}(\Sigma)\big) / \operatorname{Stab} A_0 \to \mathcal{A}^{0,p}(\Sigma), \qquad [(A_0 + \alpha, g)] \mapsto (g^{-1})^*(A_0 + \alpha)$$

is a diffeomorphism onto its image, which contains an L^2-ball,

$$B_\delta(A_0) := \{A \in \mathcal{A}^{0,p}(\Sigma) \,|\, \|A - A_0\|_{L^2(\Sigma)} < \delta\} \subseteq \operatorname{im} \mathfrak{m}.$$

We fix the following data.

- **(i)** A dense sequence $(A_i)_{i \in \mathbb{N}}$ of irreducible smooth connections in $\mathcal{A}(P)$.
- **(ii)** For every A_i a dense sequence $(\eta_{ij})_{j \in \mathbb{N}}$ of smooth 1-forms in $\Omega^1(\Sigma, \operatorname{ad}(P))$ satisfying $d_{A_i}^* \eta_{ij} = 0$ for all $j \in \mathbb{N}$.
- **(iii)** A smooth cutoff function $\rho : \mathbb{R} \to [0, 1]$ such that $\rho = 1$ on $[-1, 1]$ and $\operatorname{supp} \rho \subseteq [-4, 4]$ and such that $\|\rho'\|_{L^\infty} < 1$. Set $\rho_k(r) = \rho(k^2 r)$ for $k \in \mathbb{N}$.

To the triple $\ell = (i, j, k) \in \mathbb{N}^3$ we assign the map

$$\mathcal{V}_\ell : \mathcal{A}(P) \to \mathbb{R}, \qquad A \mapsto \rho_k(\|\alpha\|^2_{L^2(\Sigma)}) \langle \alpha, \eta_j \rangle, \tag{4}$$

where $g \in \mathcal{G}^{1,p}(P)$ and $\alpha \in L^p(\Sigma, \operatorname{ad}(P))$ are uniquely determined by the condition

$$g^* A - A_i = \alpha \qquad \text{and} \qquad d_{A_i}^* \alpha = 0. \tag{5}$$

It follows from Theorem 2.1 that there exists a constant $\varepsilon(A_i) > 0$ such that condition (5) is satisfied for a unique $\alpha \in L^p(\Sigma, \operatorname{ad}(P))$ whenever

$$\operatorname{dist}^2_{L^2(\Sigma)}(A, \mathcal{O}(A_i)) \leq \varepsilon(A_i).$$

Given A_i we allow only for indices $k \in \mathbb{N}$ sufficiently large such that the condition

$$\operatorname{supp} \rho_k \subseteq \Big[-\frac{\varepsilon(A_i)}{2}, \frac{\varepsilon(A_i)}{2} \Big]$$

is satisfied. This determines $\mathcal{V}_\ell$ uniquely on $\mathcal{A}(P)$. Note that $\mathcal{V}_\ell$ is invariant under gauge transformations. Given $\mathcal{V}_\ell$, we fix a constant $C_\ell > 0$ such that the following two conditions are satisfied.

(i) $\sup_{A \in \mathcal{A}(P)} |\mathcal{V}_\ell(A)| \leq C_\ell$,

(ii) $\|\nabla \mathcal{V}_\ell(A)\|_{L^p(\Sigma)} \leq C_\ell \big(1 + \|F_A\|_{L^p(\Sigma)}\big)$ for every $A \in \mathcal{A}(P)$.

The existence of the constant C_ℓ follows from Proposition A.7. The *universal space of perturbations* is the normed linear space

$$Y = \Big\{ \mathcal{V} := \sum_{\ell=1}^{\infty} \lambda_\ell \mathcal{V}_\ell \,\Big|\, \lambda_\ell \in \mathbb{R} \text{ and } \|\mathcal{V}\| := \sum_{\ell=1}^{\infty} C_\ell |\lambda_\ell| < \infty \Big\}.$$

It is a separable Banach space isomorphic to the space ℓ^1 of summable real sequences.

2.3 Critical points

In the following we shall make use of the well-known fact that the Yang-Mills functional $\mathcal{YM}$ satisfies the Palais-Smale condition in dimension 2 (which holds true also in dimension 3 but not in higher dimensions).

Definition 2.2 *A sequence $(A_i) \subseteq \mathcal{A}(P)$ is said to be a* Palais-Smale sequence *if there exists $M > 0$ such that $\|F_{A_i}\|_{L^2(\Sigma)} < M$ for all i, and*

$$\|d_{A_i}^* F_{A_i}\|_{W^{-1,2}(\Sigma)} \to 0 \qquad \text{as} \quad i \to \infty.$$

Theorem 2.3 (Equivariant Palais-Smale condition) *For any Palais-Smale sequence $(A_i) \subseteq \mathcal{A}(P)$ there exists a subsequence, which we denote (A_i), and a sequence $(g_i) \subseteq \mathcal{G}(P)$ such that $g_i^* A_i$ converges in $W^{1,2}(\Sigma)$ to a Yang-Mills connection $A_\infty \in \mathcal{A}(P)$ as $i \to \infty$.*

Proof: For a proof we refer to [22]. □

Lemma 2.4 *Let $\varepsilon, M > 0$. There exists a constant $\delta > 0$ with the following significance. If $A \in \mathcal{A}(P)$ satisfies $\|F_A\|_{L^2(\Sigma)} < M$ and*

$$\|A - A_0\|_{L^2(\Sigma)} > \varepsilon \tag{6}$$

for every Yang-Mills connection $A_0 \in \mathcal{A}(P)$ then it follows that

$$\|d_A^* F_A\|_{W^{-1,2}(\Sigma)} > \delta.$$

Proof: Assume by contradiction that there exists a sequence (A_i) satisfying $\|F_{A_i}\|_{L^2(\Sigma)} < M$ and (6) with

$$\lim_{i\to\infty} \|d_{A_i}^* F_{A_i}\|_{W^{-1,2}(\Sigma)} = 0.$$

Then by Lemma 2.4 there exists a subsequence, still denoted (A_i), a sequence of gauge transformations (g_i), and a Yang-Mills connection A_∞ with

$$\lim_{i\to\infty} g_i^* A_i = A_\infty$$

in $L^2(\Sigma)$ (even in $W^{1,2}(\Sigma)$). Therefore, for i sufficiently large, the Yang-Mills connection $g_i^* A_\infty$ satisfies $\|(g_i^{-1})^* A_\infty - A_i\|_{L^2(\Sigma)} < \varepsilon$, contradicting (6). $\square$

Theorem 2.5 *For every $\varepsilon, M > 0$ there exists a constant $\delta > 0$ with the following significance. Assume the perturbation $\mathcal{V}$ satisfies the conditions $\|\mathcal{V}\| < \delta$ and*

$$\operatorname{supp} \mathcal{V} \subseteq \mathcal{A}(P) \setminus \bigcup_{A \in \operatorname{crit}(\mathcal{YM})} B_\varepsilon(A),$$

where $B_\varepsilon(A) := \{A_1 \in \mathcal{A}(P) \,|\, \|A_1 - A\|_{L^2(\Sigma)} < \varepsilon\}$. Then the perturbed Yang-Mills functional $\mathcal{YM}^{\mathcal{V}}$ has the same set of critical points as the functional $\mathcal{YM}$ below the level M, i.e. it holds that

$$\begin{aligned}\operatorname{crit}(\mathcal{YM}^{\mathcal{V}}) \cap \{A \in \mathcal{A}(P) \,|\, \mathcal{YM}(A) < M\} \\ = \operatorname{crit}(\mathcal{YM}) \cap \{A \in \mathcal{A}(P) \,|\, \mathcal{YM}(A) < M\}.\end{aligned}$$

Proof: The inclusion $\operatorname{crit}(\mathcal{YM}) \subseteq \operatorname{crit}(\mathcal{YM}^{\mathcal{V}})$ is clear because $\mathcal{V}$ is supposed to be supported away from $\operatorname{crit}(\mathcal{YM})$. It remains to show that $\operatorname{crit}(\mathcal{YM}^{\mathcal{V}}) \cap \{A \in \mathcal{A}(P) \,|\, \mathcal{YM}(A) < M\}$ only contains points that are also critical for

$\mathcal{YM}$. Thus let $A \in \mathcal{A}(P) \setminus \cup_{A \in \mathrm{crit}(\mathcal{YM})} B_\varepsilon(A)$. Choosing $\delta > 0$ sufficiently small it follows from Lemma 2.4 and Propositions A.5 and C.3 that

$$\begin{aligned}\|\nabla\mathcal{YM}^{\mathcal{V}}(A)\|_{W^{-1,2}(\Sigma)} &\geq \|\nabla\mathcal{YM}(A)\|_{W^{-1,2}(\Sigma)} - \|\nabla\mathcal{V}(A)\|_{W^{-1,2}(\Sigma)}\\ &\geq \|\nabla\mathcal{YM}(A)\|_{W^{-1,2}(\Sigma)} - \|\nabla\mathcal{V}(A)\|_{L^2(\Sigma)} > 0.\end{aligned}$$

Hence $A \notin \mathrm{crit}(\mathcal{YM}^{\mathcal{V}})$, and this proves the remaining inclusion. □

Proposition 2.6 *There exists a constant $C(P)$ such that the estimate*

$$\sup_{z\in\Sigma} |F_A(z)|^2 \leq C(P)\mathcal{YM}(A)$$

is satisfied for every Yang-Mills connection $A \in \mathcal{A}(P)$.

Proof: Because A is a Yang-Mills connection, its curvature satisfies $\Delta_A F_A = 0$. For the function $u := \frac{1}{2}|F_A|^2$ there thus follows the differential inequality

$$\begin{aligned}\Delta_A u &= -|\nabla_A F_A|^2 - \langle \nabla_A^*\nabla_A F_A, F_A\rangle\\ &= -|\nabla_A F_A|^2 - \langle \Delta_A F_A + \{F_A, F_A\} + \{R_\Sigma, F_A\}, F_A\rangle\\ &= -|\nabla_A F_A|^2 - \langle \{F_A, F_A\} + \{R_\Sigma, F_A\}, F_A\rangle\\ &\leq c(u + u^{\frac{3}{2}}).\end{aligned}$$

The claim now follows from the elliptic mean value inequality C.6. □

Theorem 2.7 *For every $C > 0$ and every Yang-Mills connection $A_0 \in \mathcal{A}(P)$ there exists a constant $\delta > 0$ with the following significance. If $A \in \mathcal{A}(P) \setminus \mathcal{C}(A_0)$ is a Yang-Mills connection of energy $\mathcal{YM}(A) < C$ then A has L^2-distance at least δ to the Yang-Mills critical manifold $\mathcal{C}(A_0)$ of A_0.*

Proof: Assume by contradiction that such a constant $\delta > 0$ does not exist. Applying the L^2-local slice theorem 2.1 it then follows that there exists a sequence $\alpha^\nu \in L^2(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))$ such that

$$d_{A_0}^*\alpha^\nu = 0 \qquad \text{and} \qquad d_{A_0+\alpha^\nu}^* F_{A_0+\alpha^\nu} = 0 \tag{7}$$

and $\lim_{\nu\to\infty}\alpha^\nu = 0$ in $L^2(\Sigma)$. From the second condition in (7) we obtain

$$H_{A_0}\alpha^\nu = -\frac{1}{2}d_{A_0}^*[\alpha^\nu \wedge \alpha^\nu] - *[\alpha^\nu \wedge *(F_{A_0+\alpha^\nu} - F_{A_0})]. \tag{8}$$

By the assumption $\mathcal{YM}(A) < C$ and Proposition 2.6 there exists for every $p \leq \infty$ a constant $C(p)$ such that $\|F_{A+\alpha^\nu}\|_{L^p(\Sigma)} < C(p)$ holds for all ν. This yields for the two terms on the the right-hand side of (8) and fixed $p > 2$ the following estimate,

$$\begin{aligned}
\|d_{A_0}^*[\alpha^\nu \wedge \alpha^\nu]\|_{W^{-1,p}(\Sigma)} &\leq c\|\{\nabla_{A_0}\alpha^\nu, \alpha^\nu\}\|_{W^{-1,p}(\Sigma)} \\
&\leq c\|\alpha^\nu\|_{W^{1,p}(\Sigma)}\|\alpha^\nu\|_{L^2(\Sigma)} \\
&\leq c(1 + \|F_{A_0+\alpha^\nu}\|_{L^p(\Sigma)} + \|\alpha^\nu\|_{L^2(\Sigma)})\|\alpha^\nu\|_{L^2(\Sigma)} \\
&\leq c(1 + C(p))\|\alpha\|_{L^2(\Sigma)}.
\end{aligned}$$

The second line follows from Proposition A.6, the third line is by Proposition A.7. The second term is estimated similarly,

$$\begin{aligned}
\|[\alpha^\nu \wedge *(F_{A_0+\alpha^\nu} - F_{A_0})]\|_{W^{-1,p}(\Sigma)} &\leq c\|F_{A_0+\alpha^\nu} - F_{A_0}\|_{L^p(\Sigma)}\|\alpha\|_{L^2(\Sigma)} \\
&\leq c(1 + C(p))\|\alpha\|_{L^2(\Sigma)}.
\end{aligned}$$

Hence the right-hand side of (8) converges to 0 in $W^{-1,p}(\Sigma)$ as $\nu \to \infty$. Elliptic regularity of the operator $H_{A_0} + d_{A_0}d_{A_0}^*$ then implies convergence $\|\alpha^\nu\|_{W^{1,p}(\Sigma)} \to 0$ as $\nu \to \infty$. This contradicts known results on $W^{1,p}$-separation of Yang-Mills critical orbits and proves our claim. □

3 Yang-Mills gradient flow

Definition 3.1 The *perturbed Yang-Mills gradient flow* is the nonlinear PDE

$$0 = \partial_s A + d_A^* F_A - d_A \Psi + \nabla \mathcal{V}(A) \tag{9}$$

for paths $A : s \mapsto A(s) \in \mathcal{A}(P)$ of connections and $\Psi : s \mapsto \Psi(s) \in \Omega^0(\Sigma, \mathrm{ad}(P))$ of 0-forms.

The term $-d_A\Psi$ plays the role of a gauge fixing term needed to make equation (9) invariant under time-dependent gauge transformations.

3.1 Moduli spaces

We fix a pair $(\hat{\mathcal{C}}^-, \hat{\mathcal{C}}^+)$ of critical manifolds of the perturbed Yang-Mills functional $\mathcal{YM}^\mathcal{V}$ and denote $\mathcal{C}^\pm := \frac{\hat{\mathcal{C}}^\pm}{\mathcal{G}_0(P)}$. We also fix numbers $p > 3$ and $\delta > 0$. Central to the construction of Morse homology groups (which will be carried out in Section 8) will be the *moduli space of gradient flow lines between $\mathcal{C}^-$ and $\mathcal{C}^+$*. Let us define

$$\hat{\mathcal{M}}(\hat{\mathcal{C}}^-,\hat{\mathcal{C}}^+) := \Big\{(A,\Psi)\in\mathcal{A}^{1,2;p}_\delta(P)\times W^{1,p}_\delta(\mathbb{R}\times\Sigma)\;\Big|\;(A,\Psi)\,\text{satisfies}\,(9),\\ \lim_{s\to\pm\infty}A(s)=A^\pm\,\text{for some}\,A^\pm\in\hat{\mathcal{C}}^\pm\Big\}.$$

For a definition of the Sobolev spaces $\mathcal{A}^{1,2;p}_\delta(P)$ and $W^{1,p}_\delta(\mathbb{R}\times\Sigma)$ we refer to the next section. The group $\mathcal{G}^{2,p}_\delta(\mathbb{R}\times P)$ acts freely on $\hat{\mathcal{M}}(\hat{\mathcal{C}}^-;\hat{\mathcal{C}}^+)$. The moduli space of gradient flow lines between $\mathcal{C}^-$ and $\mathcal{C}^+$ is the quotient

$$\mathcal{M}(\mathcal{C}^-,\mathcal{C}^+) := \frac{\hat{\mathcal{M}}(\hat{\mathcal{C}}^-,\hat{\mathcal{C}}^+)}{\mathcal{G}^{2,p}_\delta(\mathbb{R}\times P)}. \tag{10}$$

The aim of the next section is to reveal $\hat{\mathcal{M}}(\hat{\mathcal{C}}^-,\hat{\mathcal{C}}^+)$ as the zero set $\mathcal{F}^{-1}(0)$ of an equivariant section $\mathcal{F}$ of a suitably defined Banach space bundle $\mathcal{E}$ over a Banach manifold $\mathcal{B}$. After showing that the vertical differential $d_x\mathcal{F}$ at any such zero $x\in\mathcal{F}^{-1}(0)$ is a surjective Fredholm operator, the implicit function theorem applies and allows us to conclude that the moduli space $\mathcal{M}(\mathcal{C}^-,\mathcal{C}^+)$ is a finite-dimensional smooth manifold.

3.2 Banach manifolds

In this section we introduce the setup which will allow us to view the moduli space defined in (10) as the zero set of a Fredholm section of a certain Banach space bundle. These Banach manifolds are modeled on weighted Sobolev spaces in order to make the Fredholm theory work. We therefore choose a number $\delta>0$ and a smooth cut-off function β such that $\beta(s)=-1$ if $s<0$ and $\beta(s)=1$ if $s>1$. We define the δ-weighted $W^{k,p}$-Sobolev norm of a measurable function u over $\mathbb{R}\times\Sigma$ to be the usual $W^{k,p}$-Sobolev norm of the function $e^{\delta\beta(s)s}u$.

Let $\mathcal{A}^{1,2;p}_\delta(P)$ denote the Sobolev space of connections on the principal fibre bundle $\mathbb{R}\times P$ which are locally of class $W^{1,2;p}$ and for which there exist limiting connections $A^\pm\in\mathcal{C}^\pm$ and times $T^\pm\in\mathbb{R}$ such that the 1-forms $\alpha^\pm := A-A^\pm$ satisfy

$$\alpha^-\in W^{1,p}_\delta((-\infty,T^-],L^p(\Sigma))\cap L^p_\delta((-\infty,T^-],W^{2,p}(\Sigma)),\\ \alpha^+\in W^{1,p}_\delta([T^+,\infty),L^p(\Sigma))\cap L^p_\delta([T^+,\infty),W^{2,p}(\Sigma)).$$

Similarly, let $\mathcal{G}^{2,p}_\delta(\mathbb{R}\times P)$ denote the group of based gauge transformations which are locally of class $W^{2,p}$ and in addition satisfy the following two conditions.[1] The ad(P)-valued 1-form $g^{-1}dg$ satisfies

$$g^{-1}dg \in L^p_\delta(\mathbb{R}, W^{2,p}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))).$$

There exist limiting based gauge transformations $g^\pm \in W^{3,p}(P)$, numbers $T^\pm \in \mathbb{R}$, and ad(P)-valued 1-forms

$$\gamma^- \in W^{2,p}_\delta((-\infty, T^-] \times \Sigma) \qquad \text{and} \qquad \gamma^+ \in W^{2,p}_\delta([T^+, \infty) \times \Sigma)$$

with

$$g(s) = \exp(\gamma^-(s))g^- \quad (s \le T^-), \qquad g(s) = \exp(\gamma^+(s))g^+ \quad (s \ge T^+).$$

For a pair $(\mathcal{C}^-, \mathcal{C}^+)$ of critical manifolds and numbers $p > 3$ and $\delta > 0$ denote by $\hat{\mathcal{B}} := \hat{\mathcal{B}}(\mathcal{C}^-, \mathcal{C}^+, \delta, p)$ the Banach manifold of pairs

$$(A, \Psi) \in \mathcal{A}^{1,2;p}_\delta(P) \times W^{1,p}_\delta(\mathbb{R} \times \Sigma).$$

The group $\mathcal{G}^{2,p}_\delta(\mathbb{R} \times P)$ acts smoothly and freely on $\hat{\mathcal{B}}$ via $g \cdot (A, \Psi) = (g^*A, g^{-1}\Psi g + g^{-1}\dot{g})$. The resulting quotient space

$$\mathcal{B} := \mathcal{B}(\mathcal{C}^-, \mathcal{C}^+, \delta, p) := \frac{\hat{\mathcal{B}}(\mathcal{C}^-, \mathcal{C}^+, \delta, p)}{\mathcal{G}^{2,p}_\delta(\mathbb{R} \times P)}$$

is again a smooth Banach manifold. The tangent space at the point $[(A, \Psi)] \in \mathcal{B}$ splits naturally as a direct sum

$$T_{[(A,\Psi)]}\mathcal{B} = T^0_{[(A,\Psi)]}\mathcal{B} \oplus \mathbb{R}^{\dim \mathcal{C}^-} \oplus \mathbb{R}^{\dim \mathcal{C}^+},$$

where $T^0_{[(A,\Psi)]}\mathcal{B}$ can be identified with pairs

$$(\alpha, \psi) \in W^{1,2;p}_\delta(\mathbb{R} \times \Sigma) \oplus W^{1,p}_\delta(\mathbb{R} \times \Sigma) \tag{11}$$

which satisfy the gauge fixing condition

$$L^*_{(A,\Psi)}(\alpha, \psi) := \partial_s \psi + [\Psi, \psi] - d^*_A \alpha = 0. \tag{12}$$

Thus a tangent vector of the quotient space $\mathcal{B}$ is identified with its unique lift to $T\hat{\mathcal{B}}$ which is perpendicular to the gauge orbit. We furthermore define

[1] For a definition of Sobolev spaces of gauge transformations we refer to [35, Appendix B].

the Banach space bundle $\mathcal{E} = \mathcal{E}(\mathcal{C}^-, \mathcal{C}^+, \delta, p)$ over $\mathcal{B}$ as follows. Let $\hat{\mathcal{E}}$ be the Banach space bundle over $\hat{\mathcal{B}}$ with fibres

$$\hat{\mathcal{E}}_{(A,\Psi)} := L^p_\delta(\mathbb{R}, L^p(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))).$$

The action of $\mathcal{G}^{2,p}_\delta(\mathbb{R} \times P)$ on $\hat{\mathcal{B}}$ lifts to a free action on $\hat{\mathcal{E}}$. We denote the respective quotient space by $\mathcal{E}$. We finally define the smooth section $\mathcal{F} : \mathcal{B} \to \mathcal{E}$ by

$$\mathcal{F} : [(A, \Psi)] \mapsto [\partial_s A + d_A^* F_A - d_A \Psi + \nabla \mathcal{V}(A)]. \tag{13}$$

Note that the moduli space defined in (10) is precisely the zero set $\mathcal{F}^{-1}(0)$.

4 Exponential decay

The aim of this section is to establish exponential decay towards Yang-Mills connections for finite energy solutions of the Yang-Mills gradient flow equation (9). We shall prove the following result.

Theorem 4.1 (Exponential decay) *For a solution $A : \mathbb{R} \to \mathcal{A}(P)$ of the Yang-Mills gradient flow equation* (9) *the following statements are equivalent.*

(i) *The solution A has finite energy*

$$E(A) = \int_{-\infty}^{\infty} \|\partial_s A\|^2_{L^2(\Sigma)}\, ds.$$

(ii) *There are positive constants k and c_ℓ, $\ell \in \mathbb{N}_0$, such that the inequality*

$$\|\partial_s A\|_{C^\ell([T,\infty)\times\Sigma)} + \|\partial_s A\|_{C^\ell((-\infty,-T]\times\Sigma)} \leq c_\ell e^{-kT} \tag{14}$$

is satisfied for every $T \geq 1$.

If (i) or (ii) is satisfied, then there exist Yang-Mills connections $A^\pm \in \mathcal{A}(P)$ such that it holds exponential convergence

$$\lim_{s\to\pm\infty} A(s) = A^\pm$$

in the $C^\infty(\Sigma)$-topology.

For the proof of Theorem 4.1 we need a number of auxiliary results.

Proposition 4.2 *For every $M > 0$, $\rho > 0$, $\kappa > 0$, and $p > 1$ there exists $\varepsilon > 0$ such that the following holds. If $A : [-\rho, \rho] \to \mathcal{A}(P)$ is a solution of* (9) *with $\|F_{A(0)}\|_{L^2(\Sigma)} \leq M$ and*

$$\int_{-\rho}^{\rho} \|\partial_s A\|^2_{L^2(\Sigma)} \, ds < \varepsilon$$

then there is a Yang-Mills connection $A^\infty \in \mathcal{A}(P)$ such that

$$\|A(0) - A^\infty\|_{W^{1,p}(\Sigma)} + \|\partial_s A(0)\|_{L^\infty(\Sigma)} < \kappa. \tag{15}$$

Proof: Assume by contradiction that this is wrong for some $\rho, M > 0$ and $p > 1$. Then there exists a constant $\kappa > 0$ and a sequence $A^\nu : [-\rho, \rho] \to \mathcal{A}(P)$ of solutions of (9) with $\|F_{A^\nu(0)}\|_{L^2(\Sigma)} \leq M$ for all ν and

$$\lim_{\nu \to \infty} \int_{-\rho}^{\rho} \|\partial_s A^\nu\|^2_{L^2(\Sigma)} \, ds = 0, \tag{16}$$

but (15) fails. Hence by Theorem 6.2 there exists a sequence of gauge transformations $g^\nu \in \mathcal{G}(P)$ such that (after passing to a subsequence) $(g^\nu)^*(A^\nu, 0) = (g^\nu)^*(A^\nu, (g^\nu)^{-1}\partial_s g^\nu)$ converges in $W^{2,p}(I \times \Sigma)$ to a solution (A^∞, Ψ^∞) of (9). After modifying the sequence g^ν we may assume that $\Psi^\infty = 0$ and thus $\lim_{\nu\to\infty}(g^\nu)^{-1}\partial_s g^\nu = 0$. Hence to every sufficiently large $\nu \geq \nu_0$ we can apply a further gauge transformation to put the connection $(g^\nu)^*(A^\nu, 0)$ in temporal gauge (i.e. to achieve that $(g^\nu)^{-1}\partial_s g^\nu$ vanishes), and $(g^\nu)^* A^\nu$ still converges to A^∞. By (9) and (16) it follows that A^∞ is a Yang-Mills connection. It holds that

$$\begin{aligned}
\lim_{\nu\to\infty} \|A^\nu(0) - (g^\nu)^{-1,*}A^\infty\|_{W^{1,p}(\Sigma)} &= \lim_{\nu\to\infty} \|((g^\nu)^*A^\nu - A^\infty)(0)\|_{W^{1,p}(\Sigma)} = 0, \\
\lim_{\nu\to\infty} \|\partial_s A^\nu(0)\|_{L^\infty(\Sigma)} &= \lim_{\nu\to\infty} \|\partial_s((g^\nu)^*A^\nu)(0)\|_{L^\infty(\Sigma)} = \|\partial_s A^\infty\|_{L^\infty(\Sigma)} = 0.
\end{aligned}$$

Hence the assumption that (15) fails was wrong. This proves the Proposition. □

For a critical manifold $\mathcal{C}$ of Yang-Mills connections and any connection $A \in \mathcal{A}(P)$ sufficiently close to $\mathcal{C}$ with respect to the $W^{1,\infty}$-topology (this assumption is needed in Proposition 4.4 below) there exists a Yang-Mills connection $A_0 \in \mathcal{C}$ such that $\alpha := A - A_0 \in (T_{A_0}\mathcal{C})^\perp$. This follows from the local slice Theorem 2.1. The Morse-Bott condition implies that $\alpha \in (\ker H_{A_0})^\perp$. We decompose the Yang-Mills gradient at the point A orthogonally as $d_A^* F_A = \beta_0 + \beta_1$ with $\beta_0 \in \operatorname{im} H_{A_0}$ and $\beta_1 \in \ker H_{A_0}$.

Proposition 4.3 *For every $\varepsilon > 0$ there exists a constant $\delta > 0$ such that the term β_1 in the above decomposition satisfies*

$$\|\beta_1\|_{L^2(\Sigma)} \leq \varepsilon \|d_A^* F_A\|_{L^2(\Sigma)},$$

whenever $\|\alpha\|_{W^{1,\infty}(\Sigma)} < \delta$.

Proof: We expand $\beta_0 + \beta_1 = d_A^* F_A = d_{A_0+\alpha}^* F_{A_0+\alpha}$ as

$$\begin{aligned}\beta_0 + \beta_1 = H_{A_0}\alpha + \frac{1}{2} d_{A_0}^*[\alpha \wedge \alpha] - [*\alpha \wedge *(d_{A_0}\alpha + \frac{1}{2}[\alpha \wedge \alpha])]& \\ =: H_{A_0}\alpha + R(\alpha).& \quad (17)\end{aligned}$$

Note that there exists a constant $c > 0$ such that

$$\|R(\alpha)\|_{L^2(\Sigma)} \leq c \|\alpha\|_{W^{1,\infty}(\Sigma)} \|\alpha\|_{L^2(\Sigma)}.$$

From (17) it follows that $H_{A_0}\alpha \in (\ker H_{A_0})^\perp$ and hence

$$\|\beta_1\|_{L^2(\Sigma)} \leq \|R(\alpha)\|_{L^2(\Sigma)} \leq c\delta \|\alpha\|_{L^2(\Sigma)}.$$

Denoting by $\lambda > 0$ the smallest (in absolute value) non-zero eigenvalue of H_{A_0} it furthermore follows that $\|H_{A_0}\alpha\|_{L^2(\Sigma)} \geq \lambda \|\alpha\|_{L^2(\Sigma)}$. It now follows that (we drop subscripts after $\|\cdot\|$)

$$\begin{aligned}\frac{\|\beta_1\|}{\|d_A^* F_A\|} &\leq \frac{c\delta\|\alpha\|}{\|H_{A_0}\alpha\| - \|R(\alpha)\|} \leq \frac{c\delta\|\alpha\|}{\lambda\|\alpha\| - \|R(\alpha)\|} = \frac{c\delta}{\lambda} + \frac{c\delta\|R(\alpha)\|}{\lambda^2\|\alpha\| - \lambda\|R(\alpha)\|} \\ &\leq \frac{c\delta}{\lambda} + \frac{c\delta\|R(\alpha)\|}{\lambda^2 c^{-1}\delta^{-1}\|R(\alpha)\| - \lambda\|R(\alpha)\|} = \frac{c\delta}{\lambda} + \frac{c\delta}{\lambda^2 c^{-1}\delta^{-1} - \lambda}.\end{aligned}$$

Choose $\delta > 0$ small enough such that $\frac{c\delta}{\lambda} + \frac{c\delta}{\lambda^2 c^{-1}\delta^{-1} - \lambda} < \varepsilon$ is satisfied. The claim then follows. $\square$

Proposition 4.4 *Let $\mathcal{C} \subseteq \mathcal{A}(P)$ be a Yang-Mills critical manifold and let A and A_0 be connections as described before Proposition 4.3. Then there exists a constant $c(A_0) > 0$ such that the estimate*

$$\|\beta\|_{L^2(\Sigma)} + \|\nabla_A\beta\|_{L^2(\Sigma)} \leq c(A_0)\|H_A\beta\|_{L^2(\Sigma)} \tag{18}$$

is satisfied for every $\beta \in (\ker H_{A_0})^\perp$.

Proof: Because $\mathcal{C}$ satisfies is a non-degenerate critical manifold it follows that the restriction of the Yang-Mills Hessian H_{A_0} to $(\ker H_{A_0})^\perp$ is a bijective operator

$$\begin{aligned} H_{A_0}|_{(\ker H_{A_0})^\perp} : W^{2,2}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P)) \cap (\ker H_{A_0})^\perp \\ \to L^2(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P)) \cap (\ker H_{A_0})^\perp. \end{aligned}$$

Thus there holds the estimate

$$\begin{aligned} \|\beta\|_{L^2(\Sigma)} + \|\nabla_A\beta\|_{L^2(\Sigma)} &\leq c(A_0)\|H_{A_0}\beta\|_{L^2(\Sigma)} \\ &\leq c(A_0)\big(\|H_A\beta\|_{L^2(\Sigma)} + \|(H_A - H_{A_0})\beta\|_{L^2(\Sigma)}\big) \end{aligned} \tag{19}$$

for a constant $c(A_0) > 0$ and every $\beta \in (\ker H_{A_0})^\perp$. Now the difference $H_A - H_{A_0}$ is the operator

$$\begin{aligned} H_{A_0+\alpha} - H_{A_0} = &-d_{A_0} * [\alpha \wedge * \cdot] + [\alpha \wedge d_{A_0}^* \cdot] - *[\alpha \wedge d_{A_0} \cdot] + d_{A_0}^*[\alpha \wedge \cdot] \\ &- [\alpha \wedge *[\alpha \wedge * \cdot]] - *[\alpha \wedge *[\alpha \wedge \cdot]] + [*(d_{A_0}\alpha + \frac{1}{2}[\alpha \wedge \alpha]) \wedge \cdot], \end{aligned}$$

which converges to 0 in $\mathcal{L}(W^{1,2}(\Sigma), L^2(\Sigma))$ as $\alpha \to 0$ in $W^{1,\infty}(\Sigma)$. Thus for $\|\alpha\|_{W^{1,\infty}(\Sigma)}$ sufficiently small the term $\|(H_A - H_{A_0})\beta\|_{L^2(\Sigma)}$ can be absorbed in the left-hand side of (19). This proves the proposition. □

Lemma 4.5 (L^2-exponential decay of the gradient) *Let $s \mapsto A(s)$ with $s \in \mathbb{R}$ be a solution of the Yang-Mills gradient flow equation* (9) *such that $\lim_{s\to\pm\infty} \|\partial_s A(s)\|_{L^\infty(\Sigma)} = 0$ and such that for a constant $T > 0$ the following condition is satisfied. There exist Yang-Mills critical manifolds $\mathcal{C}^\pm$ such that the conclusion of Proposition 4.4 applies to all $A(s)$ with $|s| > T$. Then exponential decay $\|\partial_s A\|_{L^2(\Sigma)} \to 0$ for $s \to \pm\infty$ holds, i.e. there exists a constant $k > 0$ such that*

$$\|\partial_s A(s)\|^2_{L^2(\Sigma)} \leq e^{k(s+T)}\|\partial_s A(-T)\|^2_{L^2(\Sigma)} \tag{20}$$

is satisfied for all $s \leq -T$. An analogue decay estimate holds for $s \geq T$.

Proof: We use Lemma C.2 to show exponential decay. By the Yang-Mills gradient flow equation (9) it follows the identity

$$\ddot{A} = -\partial_s d_A^* F_A = -d_A^* d_A \dot{A} + *[\dot{A} \wedge *F_A] = -H_A\dot{A}$$

at every large enough $|s| > T$ such that $\nabla\mathcal{V}(A(s)) = 0$. We furthermore calculate

$$\partial_s(H_A\dot{A}) = H_A\ddot{A} + d_A^*[\dot{A}\wedge\dot{A}] - *[\dot{A}\wedge *d_A\dot{A}] + *[*d_A\dot{A}\wedge\dot{A}].$$

It then follows that

$$\begin{aligned}
&\frac{d^2}{ds^2}\frac{1}{2}\|\dot{A}\|^2_{L^2(\Sigma)} = \frac{d}{ds}\langle\ddot{A},\dot{A}\rangle \\
&= \|\ddot{A}\|^2_{L^2(\Sigma)} - \langle\partial_s(H_A\dot{A}),\dot{A}\rangle \\
&= 2\|H_A\dot{A}\|^2_{L^2(\Sigma)} - \langle\dot{A}, d_A^*[\dot{A}\wedge\dot{A}]\rangle + \langle\dot{A}, *[\dot{A}\wedge *d_A\dot{A}]\rangle \\
&\quad +\langle\dot{A}, *[\dot{A}\wedge *d_A\dot{A}]\rangle \\
&= 2\|H_A\dot{A}\|^2_{L^2(\Sigma)} - 3\langle d_A\dot{A}, [\dot{A}\wedge\dot{A}]\rangle. \qquad (21)
\end{aligned}$$

We use the orthogonal decomposition of $\dot{A}$ as described before Proposition 4.3 into $\dot{A} = \beta_0 + \beta_1$ where $\beta_0 \in \operatorname{im} H_{A_0}$ and $\beta_1 \in \ker H_{A_0}$. The last term in (21) can then be estimated as

$$\begin{aligned}
&|\langle d_A\dot{A}, [\dot{A}\wedge\dot{A}]\rangle| \\
&\quad \leq \|\dot{A}\|_{L^\infty(\Sigma)}\big(\|d_A\beta_0\|^2_{L^2(\Sigma)} + \|d_A\beta_1\|^2_{L^2(\Sigma)} + \|\beta_0\|^2_{L^2(\Sigma)} + \|\beta_1\|^2_{L^2(\Sigma)}\big).
\end{aligned}$$

With β_1 satisfying $H_{A_0}\beta_1 = 0$, hence $d_{A_0}^*d_{A_0}\beta_1 = -*[*F_{A_0}\wedge\beta_1]$ we find that

$$\begin{aligned}
\|d_A\beta_1\|^2_{L^2(\Sigma)} &\leq 2\|d_{A_0}\beta_1\|^2_{L^2(\Sigma)} + 2\|[\alpha\wedge\beta_1]\|^2_{L^2(\Sigma)} \\
&\leq 2\|\beta_1\|_{L^2(\Sigma)}\|d_{A_0}^*d_{A_0}\beta_1\|_{L^2(\Sigma)} + 2\|[\alpha\wedge\beta_1]\|^2_{L^2(\Sigma)} \\
&= \|\beta_1\|_{L^2(\Sigma)}\|[*F_{A_0}\wedge\beta_1]\|_{L^2(\Sigma)} + 2\|[\alpha\wedge\beta_1]\|^2_{L^2(\Sigma)} \\
&\leq c\|\beta_1\|^2_{L^2(\Sigma)}\big(\|F_{A_0}\|_{L^\infty(\Sigma)} + \|\alpha\|^2_{L^\infty(\Sigma)}\big). \qquad (22)
\end{aligned}$$

Using Proposition 4.3 and the estimate (18) with constant $c(A_0)$, the right-hand side of (21) can now further be estimated as follows. We put $\delta :=$

$2-3c(A_0)^2\|\dot{A}\|_{L^\infty(\Sigma)}>0$. Then,

$$
\begin{aligned}
&\frac{d^2}{ds^2}\frac{1}{2}\|\dot{A}\|^2_{L^2(\Sigma)} = 2\|H_A\dot{A}\|^2_{L^2(\Sigma)} - 3\langle d_A\dot{A},[\dot{A}\wedge\dot{A}]\rangle \\
&\quad\geq\ 2\|H_A\beta_0\|^2_{L^2(\Sigma)} - 2\|H_A\beta_1\|^2_{L^2(\Sigma)} - 3\|\dot{A}\|_{L^\infty(\Sigma)}\big(\|d_A\beta_0\|^2_{L^2(\Sigma)}+\|\beta_0\|^2_{L^2(\Sigma)}\big) \\
&\quad\qquad -c\|\dot{A}\|_{L^\infty}\big(1+\|F_{A_0}\|_{L^\infty(\Sigma)}+\|\alpha\|^2_{L^\infty(\Sigma)}\big)\|\beta_1\|^2_{L^2(\Sigma)} \\
&\quad\geq\ \delta\|H_A\beta_0\|^2_{L^2(\Sigma)} - 2\|H_A\beta_1\|^2_{L^2(\Sigma)} \\
&\quad\qquad -c\|\dot{A}\|_{L^\infty(\Sigma)}\big(1+\|F_{A_0}\|_{L^\infty(\Sigma)}+\|\alpha\|^2_{L^\infty(\Sigma)}\big)\|\beta_1\|^2_{L^2(\Sigma)} \\
&\quad\geq\ \delta\|H_A\beta_0\|^2_{L^2(\Sigma)} - 2\|H_A\beta_1\|^2_{L^2(\Sigma)} - c_1(A_0)\varepsilon^2\|\dot{A}\|^2_{L^2(\Sigma)} \\
&\quad\geq\ \delta\|H_A\beta_0\|^2_{L^2(\Sigma)} - 2\|H_A-H_{A_0}\|^2_{\mathcal{L}(W^{1,2}(\Sigma),L^2(\Sigma))}\|\beta_1\|^2_{W^{1,2}(\Sigma)} \\
&\quad\qquad -c_1(A_0)\varepsilon^2\|\dot{A}\|^2_{L^2(\Sigma)}.
\end{aligned}
$$

Thanks to Proposition B.5 we can bound the term $\|\beta_1\|^2_{W^{1,2}(\Sigma)}$ as

$$
\begin{aligned}
\|\beta_1\|^2_{W^{1,2}(\Sigma)} &\leq c\big(\|d_A\beta_1\|^2_{L^2(\Sigma)}+\|d_A^*\beta_1\|^2_{L^2(\Sigma)}\big)+c\|F_A\|_{L^\infty(\Sigma)}\|\beta_1\|^2_{L^2(\Sigma)} \\
&\leq c\|\beta_1\|^2_{L^2(\Sigma)}\big(1+\|\alpha\|^2_{L^\infty(\Sigma)}+\|F_{A_0}\|^2_{L^\infty(\Sigma)}+\|F_A\|^2_{L^\infty(\Sigma)}\big) \\
&\leq c_2(A_0)\varepsilon^2\|\dot{A}\|^2_{L^2(\Sigma)}.
\end{aligned}
$$

In the second line we used (22) and the assumption $d_{A_0}^*\beta_1=0$, hence $d_A^*\beta_1 = -*[\alpha\wedge *\beta_1]$. The last line is by Proposition 4.3. Let us denote $K:=K(\alpha):=H_A-H_{A_0}$ and $\|K\|:=\|K\|_{\mathcal{L}(W^{1,2},L^2)}$. It then follows for $\varepsilon>0$ sufficiently small, for the constant $\delta_1:=\delta c(A_0)^2>0$, and with $\|\dot{A}\|^2_{L^2(\Sigma)} = \|\beta_0\|^2_{L^2(\Sigma)}+\|\beta_1\|^2_{L^2(\Sigma)}$ that

$$
\begin{aligned}
&\frac{d^2}{ds^2}\frac{1}{2}\|\dot{A}\|^2_{L^2(\Sigma)} \\
&\quad\geq\ \delta\|H_A\beta_0\|^2_{L^2(\Sigma)} - c\varepsilon^2\|K\|^2\|\dot{A}\|^2_{L^2(\Sigma)} - c_1(A_0)\varepsilon^2\|\dot{A}\|^2_{L^2(\Sigma)} \\
&\quad\geq\ \delta_1\|\beta_0\|^2_{L^2(\Sigma)} - c\varepsilon^2\|K\|^2\|\dot{A}\|^2_{L^2(\Sigma)} - c_1(A_0)\varepsilon^2\|\dot{A}\|^2_{L^2(\Sigma)} \\
&\quad\geq\ \delta_1\|\dot{A}\|^2_{L^2(\Sigma)} - \delta_1\|\beta_1\|^2_{L^2(\Sigma)} - c\varepsilon^2\|K\|^2\|\dot{A}\|^2_{L^2(\Sigma)} - c_1(A_0)\varepsilon^2\|\dot{A}\|^2_{L^2(\Sigma)} \\
&\quad\geq\ \delta_1\|\dot{A}\|^2_{L^2(\Sigma)} - \delta_1\varepsilon^2\|\dot{A}\|^2_{L^2(\Sigma)} - c\varepsilon^2\|K\|^2\|\dot{A}\|^2_{L^2(\Sigma)} - c_1(A_0)\varepsilon^2\|\dot{A}\|^2_{L^2(\Sigma)} \\
&\quad\geq\ \frac{\delta_1}{2}\|\dot{A}\|^2_{L^2(\Sigma)}.
\end{aligned}
$$

The exponential decay estimate (20) now follows from Lemma C.2. □

For solutions $A:\mathbb{R}\to\mathcal{A}(P)$ of the Yang-Mills gradient flow equation (9) satisfying the assumptions of Lemma 4.5 one easily infers L^2-convergence to limit connections $A^{\pm}$ as $s \to \pm\infty$. Namely, these limits are given by

$$A^- := A(-T) - \int_{-\infty}^{-T} \partial_s A(s)\, ds, \qquad A^+ := A(T) + \int_T^{\infty} \partial_s A(s)\, ds, \quad (23)$$

for any $T > 0$ large enough. These integrals converge due to the exponential decay of $\|\partial_s A(s)\|_{L^2(\Sigma)}$. We now turn to the proof of Theorem 4.1.
Proof: (Theorem 4.1). With

$$\mathcal{YM}^{\mathcal{V}}(A(s_0)) - \mathcal{YM}^{\mathcal{V}}(A(s_1)) \leq E(A)$$

being satisfied for all numbers $s_0 \leq s_1$ and the assumption that $E(A)$ is finite, we infer the uniform bound $\mathcal{YM}^{\mathcal{V}}(A(s)) \leq M$ for all $s \in \mathbb{R}$. The finite energy assumption furthermore yields for any $\rho > 0$, $\varepsilon > 0$ the existence of a number $T > 0$ such that

$$\int_{-s-\rho}^{-s+\rho} \|\partial_s A\|_{L^2(\Sigma)}\, ds + \int_{s-\rho}^{s+\rho} \|\partial_s A\|_{L^2(\Sigma)}\, ds < \varepsilon$$

holds for every $s \geq T$. Hence applying Proposition 4.2 we conclude estimate (15) for some constant $\kappa > 0$. For κ sufficiently small the assumptions of Lemma 4.5 are satisfied. Hence for $s \to \pm\infty$, $A(s)$ converges to limit connections $A^{\pm}$ as given by (23). It furthermore follows L^2-exponential decay of $\partial_s A$ with constant $k > 0$, i.e.

$$\|\partial_s A(s)\|^2_{L^2(\Sigma)} \leq e^{k(s-T)} \|\partial_s A(-T)\|^2_{L^2(\Sigma)} \quad (24)$$

holds for all $s \leq -T$, and similarly for all $s \geq T$. Estimate (24) now allows us to prove the asserted forward and backward exponential decay (14). Consider first backward exponential decay. As an intermediate step we claim for each $\ell \in \mathbb{N}_0$ the existence of a constant $c_\ell > 0$, which only depends on the energy $E(A)$, such that

$$\|\partial_s A\|_{W^{\ell,2}((-\infty,s]\times\Sigma)} \leq c_\ell \|\partial_s A\|_{L^2((-\infty,s+\ell]\times\Sigma)} \quad (25)$$

holds for all $s \leq -T - \ell$. This is achieved using standard parabolic bootstrapping arguments as follows. Set $\alpha := \partial_s A$. Then with $d_A^*\alpha = 0$ we find that α satisfies the linearized Yang-Mills gradient flow equation

$$(\partial_s + \Delta_A)\alpha = - * [*F_A \wedge \alpha]. \quad (26)$$

As follows from 24, α is contained in $L^2((-\infty, -T] \times \Sigma)$. Standard parabolic estimates (cf. e.g. [18, Theorem 7.13]) apply to (26) and yield via induction on ℓ the sought for estimate (25). Now fix $m \geq 0$. Then apply to each interval $(s-j, s-j+1)$, $j \in \mathbb{N}$, and for sufficiently large $\ell = \ell(m)$ the Sobolev embedding

$$W^{\ell,2}((s-j, s-j+1) \times \Sigma) \hookrightarrow C^m((s-j, s-j+1) \times \Sigma)$$

to obtain from (25) the bound

$$\|\partial_s A\|_{C^m((-\infty,s]\times\Sigma)} \leq c_{\ell,m} \|\partial_s A\|_{L^2((-\infty,s+\ell]\times\Sigma)}. \tag{27}$$

Finally, integrate estimate 24 over $(-\infty, s+\ell]$ and combine it with (27) to complete the proof of backward exponential decay. Forward exponential decay is obtained in a completely analogous manner. This proves the first implication of Theorem 4.1.

(ii) $\Rightarrow$ **(i)**. By definition of the energy $E(A)$ and the assumption on the decay of $\|\partial_s A\|_{C^0(\Sigma)}$ it follows, setting $T = 1$, that

$$\begin{aligned} E(A) &= \int_{-\infty}^{\infty} \int_\Sigma \|\partial_s A\|_{\mathfrak{g}}^2 \, ds \\ &= \int_{-1}^{1} \int_\Sigma \|\partial_s A\|_{\mathfrak{g}}^2 \, ds + \int_{-\infty}^{-1} \int_\Sigma \|\partial_s A\|_{\mathfrak{g}}^2 \, ds + \int_{1}^{\infty} \int_\Sigma \|\partial_s A\|_{\mathfrak{g}}^2 \, ds \\ &\leq C + |\Sigma| \int_{-\infty}^{-1} c_0^2 e^{2ks} \, ds + |\Sigma| \int_1^\infty c_0^2 e^{-2ks} \, ds \\ &= C + \frac{|\Sigma| c_0^2 e^{-2k}}{k}. \end{aligned}$$

This proves the second implication of Theorem 4.1. It remains to prove exponential convergence $\lim_{s\to\pm\infty} A(s) = A^\pm$ in $C^\ell(\Sigma)$. Exponential convergence in $L^2(\Sigma)$ has been shown through (23). The same argument now gives exponential convergence in $C^\ell(\Sigma)$ for all $\ell \in \mathbb{N}_0$. By what we have proved $\|\partial_s A(s)\|_{C^\ell(\Sigma)}$ converges to zero exponentially and thus

$$\|A(s_0) - A(s_1)\|_{C^\ell(\Sigma)} \leq \int_{s_0}^{s_1} \|\partial_s A\|_{C^\ell(\Sigma)} \, ds \leq \int_{s_0}^{s_1} c_\ell e^{ks} \, ds \leq \frac{c_\ell e^{ks_1}}{k}$$

holds for constants $c_\ell, k > 0$ and all $s_0 \leq s_1 \leq 0$. Letting $s_0 \to -\infty$ we obtain backward exponential decay in $C^\ell(\Sigma)$. Forward exponential decay follows similarly. Hence the proof of Theorem 4.1 is complete. $\square$

5 Fredholm theory

5.1 Yang-Mills Hessian

Here we denote by $\mathcal{H}_A$ the *augmented Yang-Mills Hessian* defined by

$$\mathcal{H}_A := \begin{pmatrix} d_A^* d_A + *[*F_A \wedge \cdot\,] & -d_A \\ -d_A^* & 0 \end{pmatrix}. \tag{28}$$

In order to find a domain which makes the subsequent Fredholm theory work, we fix a smooth connection $A \in \mathcal{A}(P)$ and decompose the space $\Omega^1(\Sigma, \mathrm{ad}(P))$ of smooth $\mathrm{ad}(P)$-valued 1-forms as the $L^2(\Sigma)$-orthogonal sum

$$\begin{aligned}\Omega^1(\Sigma, \mathrm{ad}(P)) = \quad & \ker\big(d_A^* : \Omega^1(\Sigma, \mathrm{ad}(P)) \to \Omega^0(\Sigma, \mathrm{ad}(P))\big) \\ & \oplus\ \mathrm{im}\big(d_A : \Omega^0(\Sigma, \mathrm{ad}(P)) \to \Omega^1(\Sigma, \mathrm{ad}(P))\big).\end{aligned} \tag{29}$$

Then let $W_0^{2,p}$ and $W_1^{1,p}$ denote the completions of the first component, respectively of the second component with respect to the Sobolev (k,p)-norm $(k = 1, 2)$. Now define the space

$$W_A^{2,p}(\Sigma) := W_0^{2,p} \oplus W_1^{1,p} \tag{30}$$

and endow it with the sum norm. Note that the this norm depends on the connection A. For $p > 1$ we have the operator

$$\mathcal{H}_A : W_A^{2,p}(\Sigma) \oplus W^{1,p}(\Sigma, \mathrm{ad}(P)) \to L^p(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P)) \oplus L^p(\Sigma, T^*\Sigma).$$

In the case $p = 2$ this is a densely defined symmetric operator on the Hilbert space $L^2(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P)) \oplus L^2(\Sigma, T^*\Sigma)$ with domain

$$\mathrm{dom}\,\mathcal{H}_A := W_A^{2,2}(\Sigma) \oplus W^{1,2}(\Sigma, \mathrm{ad}(P)). \tag{31}$$

We show in Proposition 5.1 below that it is self-adjoint. For the further discussion of the operator $\mathcal{D}_A$ it is convenient to also decompose each $\beta \in \mathrm{im}\,\mathcal{H}_A$ as $\beta = \beta_0 + \beta_1$ where $d_A^*\beta_0 = 0$ and $\beta_1 = d_A\omega$ for a 0-form ω. A short calculation shows that for $\alpha = \alpha_0 + d_A\varphi$ this decomposition is given by

$$\mathcal{H}_A\alpha = \beta_0 + d_A\omega,$$

where ω is a solution of

$$\Delta_A\omega = *[d_A * F_A \wedge \alpha]. \tag{32}$$

As Δ_A might not be injective due to the presence of Δ_A-harmonic 0-forms, the solution ω of (32) need not be unique. This ambiguity however is not relevant, as only $d_A\omega$ enters the definition of β_0 and β_1. With respect to the above decomposition of the space of 1-forms the augmented Hessian $\mathcal{H}_A$ takes the form

$$\mathcal{H}_A \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \psi \end{pmatrix} = \begin{pmatrix} \Delta_A\alpha_0 + *[*F_A \wedge \alpha_0] + [d_A^* F_A \wedge \varphi] - d_A\omega \\ -d_A\psi + d_A\omega \\ -d_A^*\alpha_1 \end{pmatrix}, \quad (33)$$

with $\alpha_1 = d_A\varphi$ and ω a solution of (32).

Proposition 5.1 *Let $A \in \mathcal{A}(P)$ and $p > 1$. Then the operator $\mathcal{H}_A$ with domain* $\operatorname{dom}\mathcal{H}_A$ *as defined in* (31) *is self-adjoint. It satisfies for all $(\alpha, \psi) \in$* $\operatorname{dom}\mathcal{H}_A$ *and $p > 1$ the elliptic estimate*

$$\|\alpha\|_{W_A^{2,p}} + \|\psi\|_{W^{1,p}} \leq c\big(\|\mathcal{H}_A(\alpha,\psi)\|_{L^p} + \|(\alpha,\psi)\|_{L^p}\big) \quad (34)$$

with constant $c = c(A,p)$. If A is a Yang-Mills connection, then the number of negative eigenvalues (counted with multiplicities) of $\mathcal{H}_A$ equals the Morse index of the Yang-Mills Hessian $H_A\mathcal{YM}$.

Proof: We show the elliptic estimate (34). Let $(\beta_0, \beta_1, \gamma) = \mathcal{H}_A(\alpha_0, \alpha_1, \psi)$ and assume that $(\beta_0, \beta_1, \gamma) \in L^p(\Sigma, T^*\Sigma \otimes \operatorname{ad}(P)) \oplus L^p(\Sigma, \operatorname{ad}(P))$. Then by ellipticity of the operator Δ_A on $\Omega^*(\Sigma, \operatorname{ad}(P))$ we obtain from equations (32) and the the first line of (33) the estimate

$$\|\alpha_0\|_{W^{2,p}} \leq c\big(\|\alpha_0\|_{L^p} + \|\varphi\|_{L^p} + \|\beta_0\|_{L^p} + \|\omega\|_{W^{1,p}}\big) \leq c\big(\|\alpha\|_{L^p} + \|\beta_0\|_{L^p}\big).$$

Applying d_A^* to the second and third equation in (33) the same elliptic estimate shows that

$$\begin{aligned} \|\psi\|_{W^{1,p}} &\leq c\big(\|\psi\|_{L^p} + \|\beta_1\|_{L^p} + \|\alpha_1\|_{L^p} + \|\Delta_A\omega\|_{L^p}\big) \\ &\leq c\big(\|\psi\|_{L^p} + \|\beta_1\|_{L^p} + \|\alpha\|_{L^p}\big) \end{aligned}$$

and

$$\|\alpha_1\|_{W^{1,p}} \leq c\|\varphi\|_{W^{2,p}} \leq c\big(\|\psi\|_{L^p} + \|\gamma\|_{L^p}\big).$$

Estimate (34) now follows and implies self-adjointness in the case $p = 2$. For the last statement note that by choice of the weight $\delta > 0$ the operators $\mathcal{H}_A$ and $\mathcal{H}_A^- := \mathcal{H}_A - \delta$ have the same number of negative eigenvalues. Let $(\alpha, \psi)^T$ be an eigenvector of $\mathcal{H}_A^-$ with eigenvalue $\lambda < 0$. Let $\alpha = \alpha_0 + \alpha_1$ be as above the Hodge decomposition with $d_A^*\alpha_0 = 0$ and $\alpha_1 = d_A\varphi$. Then the eigenvalue equation for $\mathcal{H}_A^-$ together with the first line of (33) gives $H_A\alpha_0 = \lambda\alpha_0$. This uses that $d_A\omega = 0$ and $d_A^*F_A = 0$ as A is a Yang-Mills connection. Hence λ is a negative eigenvalue of H_A. Conversely, any $\lambda < 0$ satisfying the eigenvalue equation $H_A\alpha_0 = \lambda\alpha_0$ is an eigenvalue of $\mathcal{H}_A$ with eigenvector $(\alpha_0, 0)^T$. $\square$

5.2 Linearized operator

We next discuss the linearization of the Yang-Mills gradient flow (9). Since any solution (A, Ψ) of the Yang-Mills gradient flow is gauge equivalent under $\mathcal{G}^\delta(\bar{P})$ to a solution satisfying $\Psi \equiv 0$, it suffices to consider the linearization along such trajectories only. Let $(A, \Psi) = (A, 0)$ be a solution of (9). For $p > 1$ we define the Banach spaces

$$\begin{aligned}
\mathcal{Z}_A^{\delta,p} := \;& \big(W_\delta^{1,p}(\mathbb{R}, L^p(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))) \cap L_\delta^p(\mathbb{R}, W_A^{2,p}(\Sigma))\big) \\
& \oplus \quad W_\delta^{1,p}(\mathbb{R} \times \Sigma, \mathrm{ad}(P)) \\
\mathcal{L}^{\delta,p} := \;& L_\delta^p(\mathbb{R} \times \Sigma, T^*\Sigma \otimes \mathrm{ad}(P)) \oplus L_\delta^p(\mathbb{R} \times \Sigma, \mathrm{ad}(P)).
\end{aligned}$$

The horizontal differential of the section $\mathcal{F}$ at (A, Ψ) is the linear operator

$$\mathcal{D}_A = \frac{d}{ds} + \mathcal{H}_A : \mathcal{Z}_A^{\delta,p} \to \mathcal{L}^{\delta,p},$$

and the *linearized Yang-Mills gradient flow equation* is the equation

$$\mathcal{D}_A \begin{pmatrix} \alpha \\ \psi \end{pmatrix} = 0. \tag{35}$$

We next show that $\mathcal{D}_A$ is a Fredholm operator and determine its index.

Remark 5.2 (i) The differential $d\mathcal{F}_A$ acts on functions which converge exponentially to constant tangent vectors $\alpha^\pm \in T_{A^\pm}\mathcal{C}^\pm$ as $s \to \pm\infty$, i.e. on the space $\mathcal{Z}_A^{\delta,p} \oplus \mathbb{R}^{\dim \mathcal{C}^-} \oplus \mathbb{R}^{\dim \mathcal{C}^-}$. It follows that

$$\operatorname{ind} d\mathcal{F}_A = \operatorname{ind} \mathcal{D}_A + \dim \mathcal{C}^- + \dim \mathcal{C}^-.$$

(To see that, we view $d\mathcal{F}_A$ as a compact perturbation of the operator $\mathcal{D}_A$, extended trivially to $\mathcal{Z}_A^{\delta,p} \oplus \mathbb{R}^{\dim \mathcal{C}^-} \oplus \mathbb{R}^{\dim \mathcal{C}^-}$.)

(ii) The operator $\mathcal{D}_A$ arises as the linearization of the *unperturbed* Yang-Mills gradient flow equation (9). The Fredholm theory in the case $\mathcal{V} \neq 0$ can be reduced to the unperturbed case because the terms involving $\mathcal{V}$ contribute only compact perturbations to the operator $\mathcal{D}_A$.

5.3 Fredholm theorem

Theorem 5.3 (Fredholm theorem) *Let $A : \mathbb{R} \to \mathcal{A}(P)$, be a smooth solution of the Yang-Mills gradient flow equation* (9) *satisfying for Yang-Mills connections $A^{\pm}$ the asymptotic conditions*

$$\lim_{s\to\pm\infty} A(s) = A^{\pm}$$

in the $\mathcal{C}^{\infty}(\Sigma)$-topology. Then (for every $p > 1$) the operator $\mathcal{D}_A = \frac{d}{ds} + \mathcal{H}_A : \mathcal{Z}_A^{\delta,p} \to \mathcal{L}^{\delta,p}$ associated with A is a Fredholm operator of index

$$\operatorname{ind} \mathcal{D}_A = \operatorname{ind} \mathcal{H}_{A^-} - \operatorname{ind} \mathcal{H}_{A^+} - \dim \mathcal{C}^-.$$

Weighted theory

As the Hessians $\mathcal{H}_{A^{\pm}}$ have non-trivial zero eigenspaces, we cannot apply directly known theorems on the spectral flow to prove Theorem 5.3. As an intermediate step we therefore use the Banach space isomorphisms

$$\nu_1 : \mathcal{Z}_A^{\delta,p} \to \mathcal{Z}_A^{0,p} =: \mathcal{Z}_A^{p} \qquad \text{and} \qquad \nu_2 : \mathcal{L}^{\delta,p} \to \mathcal{L}^{0,p} =: \mathcal{L}^{p}$$

given by multiplication with the weight function $e^{\delta\beta(s)s}$, where β denotes the cut-off function introduced before. Then the assertion of Theorem 5.3 is equivalent to the analogous one for the operator

$$\mathcal{D}_A^{\delta} := \nu_2 \circ \mathcal{D}_A \circ \nu_1^{-1} : \quad \mathcal{Z}_A^{p} \to \mathcal{L}^{p},$$

which we shall prove now.

Case $p = 2$

We first show this theorem in the case $p = 2$, where it follows from well-known results on the spectral flow for families $B(s) : \operatorname{dom} B(s) \to H$, $s \in \mathbb{R}$, of self-adjoint operators in Hilbert space H, cf. [23]. Since the operators we are concerned with have time-varying domains, we need an extension of this theory as outlined in [26, Appendix A]. The case of general Sobolev exponents $p > 1$ will afterwards be reduced to the Hilbert space case.

Set $H := L^2(\Sigma, T^*\Sigma \otimes \operatorname{ad}(P))$. For the following it is convenient to use the Hodge decomposition

$$\begin{aligned} H &= X_0(s) \oplus X_1(s) \\ &:= \{\alpha \mid d_{A(s)}^*\alpha = 0\} \oplus \{\alpha = d_{A(s)}\varphi \text{ for some } \varphi \in \Omega^0(\Sigma, \operatorname{ad}(P))\} \end{aligned}$$

with respect to the connection $A(s) \in \mathcal{A}(P)$. Recall from (31) that the domain of the operator $\mathcal{H}_{A(s)}$ is given by

$$\operatorname{dom} \mathcal{H}_{A(s)} = W^{2,2}_{A(s)}(\Sigma) \oplus W^{1,2}(\Sigma, \operatorname{ad}(P)) =: W(s) \oplus W^{1,2}(\Sigma, \operatorname{ad}(P)).$$

We fix $s_0 \in \mathbb{R}$ and set $A_0 := A(s_0)$. In the following we let $\beta(s) := A(s) - A_0$. Let $H = X_0 \oplus X_1$ be the Hodge decomposition corresponding to A_0 and denote $W_0 := W(s_0)$. For $s \in \mathbb{R}$ sufficiently close to s_0 we define the map $Q(s) : H \to H$ as follows. Let $\alpha \in H$ be decomposed as $\alpha = \alpha_0 + \alpha_1 \in X_0 \oplus X_1$. Then set

$$Q(s)\alpha := \operatorname{pr}_{X_0(s)} \alpha_0 + \operatorname{pr}_{X_1(s)} \alpha_1. \tag{36}$$

A short calculation shows that

$$Q(s)\alpha = \alpha_0 + d_{A(s)}\delta, \tag{37}$$

where δ solves the elliptic equation

$$\Delta_{A(s)}\delta = \Delta_{A_0}\varphi_0 + *[\beta(s) \wedge *(\alpha_0 - \alpha_1)], \tag{38}$$

with φ_0 such that it satisfies $d_{A_0}\varphi_0 = \alpha_1$. A solution δ exists and is unique up to adding elements of $\ker d_{A(s)}$.

Lemma 5.4 *There exists $\varepsilon > 0$ such that the map $Q(s) : H \to H$ has the following properties for every $s \in (s_0 - \varepsilon, s_0 + \varepsilon)$.*

(i) $Q(s)$ is a Hilbert space isomorphism.

(ii) $Q(s)$ preserves the Hodge decomposition of H, i.e. it holds that $Q(s)X_0 = X_0(s)$ and $Q(s)X_1 = X_1(s)$.

(iii) The restriction of $Q(s)$ to W_0 yields an isomorphism $Q(s) : W_0 \to W(s)$.

Proof: Note that for $\beta = 0$ the map Q defined in (37) is the identity map on H. Thus $Q(s)$ is bijective for all s with $\beta(s)$ sufficiently small.[1] To show (ii) we first observe that $Q(s)X_i \subseteq X_i(s)$, $i = 1, 2$, holds by definition of

[1] One easily checks that $\|Q(s) - Q(s_0)\|_{\mathcal{L}(H)} \leq c\|\beta(s)\|_{C^1(\Sigma)}$ which goes to 0 as $s \to s_0$ because $A \in W^{2,p}(I \times \Sigma)$ for all $p < \infty$.

$Q(s)$ in (36). Now let $\alpha \in H$ be given and assume that $\hat{\alpha} := Q(s)\alpha$ satisfies $d_{A(s)}^* \hat{\alpha} = 0$. It then follows that

$$\begin{aligned} 0 &= d_{A(s)}^* \alpha_0 + \Delta_{A(s)} \delta \\ &= -*[\beta(s) \wedge *\alpha_0] + \Delta_{A_0} \varphi_0 + *[\beta(s) \wedge (*\alpha_0 - *d_{A_0}\varphi_0] \\ &= \Delta_{A_0} \psi_0 - *[\beta(s) \wedge *d_{A_0}\psi_0]. \end{aligned}$$

Now the kernel of the operator $\varphi \mapsto \Delta_{A_0}\varphi - *[\beta(s) \wedge *d_{A_0}\varphi]$ contains $\ker d_{A(s)}$ and for $\|\beta(s)\|_{C^0(\Sigma)}$ sufficiently small it is not larger. It thus follows that $\alpha_1 = d_{A_0}\varphi_0 = 0$ and hence $\alpha = \alpha_0$. This shows surjectivity of the map $Q(s) : X_0 \to X_0(s)$. It follows similarly that also $Q(s) : X_1 \to X_1(s)$ is surjective and completes the proof of (ii). To prove (iii) we introduce the notation $X_0^{2,2}(s) := \mathrm{pr}_0(W(s))$ (with pr_0 denoting projection onto the first summand of $X_0(s) \oplus X_1(s)$), and $X_0^{2,2} := X_0^{2,2}(s_0)$. We have to verify that $Q(s)$ maps the space $X_0^{2,2}$ bijectively to $X_0^{2,2}(s)$. Thus let $\alpha \in X_0^{2,2}$. Then $\alpha_1 = d_{A_0}\varphi_0 = 0$ and $\delta \in W^{4,2}(\Sigma, \mathrm{ad}(P))$ as follows from (38) by elliptic regularity. Hence $Q(s)\alpha = \alpha + d_{A(s)}\delta \in X_0^{2,2}(s)$, as claimed. The opposite inclusion $Q^{-1}(s)X_0^{2,2}(s) \subseteq X_0^{2,2}$ follows similarly. □

Proof: (Theorem 5.3 in the case $p = 2$). Lemma 5.4 shows that the disjoint union $\bigsqcup_{s\in\mathbb{R}} W(s)$ is a locally trivial Hilbert space subbundle of $\mathbb{R} \times H$ in the sense of [26, Appendix A]. Moreover, for $s \to \pm\infty$ there holds convergence $A(s) \to A^\pm$ in $C^\infty(\Sigma)$. Hence the operators $Q(s)$ can be chosen near the ends in such a way that $Q(s) \to Q^\pm$ in $\mathcal{L}(H)$ as $s \to \pm\infty$ for appropriate Hilbert space isomorphisms $Q^\pm : H \to H$. One now easily checks that the operators

$$\begin{aligned} &(Q(s) \oplus \mathrm{id})^{-1} \circ \mathcal{H}_{A(s)} \circ (Q(s) \oplus \mathrm{id}) : \\ &\qquad\qquad W_0 \oplus W^{1,2}(\Sigma, \mathrm{ad}(P)) \to H \oplus L^2(\Sigma, \mathrm{ad}(P)) \end{aligned}$$

converge in $\mathcal{L}(W_0, H)$, as $s \to \pm\infty$, to the invertible operators $(Q^\pm \oplus \mathrm{id})^{-1} \circ \mathcal{H}_{A^\pm} \circ (Q^\pm \oplus \mathrm{id})$. Now Theorem A.4 of [26] applies and yields the claim. □

Case $p > 1$

The Fredholm property in this case follows from standard arguments. Full details can be found in [31]. The estimate (72) and bijectivity of the operator $\mathcal{D}_A$ for a stationary path A of connections yield the inequality

$$\|\xi\|_{\mathcal{Z}_A^p} \leq c(A)\big(\|\mathcal{D}_A\xi\|_{\mathcal{L}^p} + \|\xi\|_{L^p(I\times\Sigma)}\big)$$

for a constant $c(A)$ and a compact interval $I \subseteq \mathbb{R}$. Hence it follows from the abstract closed range lemma (cf. e.g. [31, 36]) that the operator $\mathcal{D}_A$ has finite-dimensional kernel and closed range. Similarly, one can show that $\operatorname{coker}\mathcal{D}_A$ is also finite-dimensional, and that the dimensions of the kernel and cokernel do not depend on p. This proves Theorem 5.3 in the general case.

6 Compactness

Throughout this section we identify a pair (A, Ψ) with the connection $A + \Psi\, ds$ over the 3-dimensional manifold $\mathbb{R} \times \Sigma$. Its curvature is thus given by $F_{(A,\Psi)} = F_A + (d_A\Psi - \partial_s A) \wedge ds$. Denote by $\bar{P} := \mathbb{R} \times P$ the trivial extension of the bundle P to the base manifold $\mathbb{R} \times \Sigma$. We use the symbols $\hat{*}$, $\hat{d}_{(A,\Psi)}$ and $\hat{d}^*_{(A,\Psi)}$ for operators acting on $\Omega^*(\mathbb{R} \times \Sigma, \operatorname{ad}(\bar{P}))$. In particualar, $\hat{d}_{(A,\Psi)}$ as an operator on 1-forms $\alpha + \psi\, ds$ is given by

$$\hat{d}_{(A,\Psi)}(\alpha + \psi\, ds) = d_A\alpha + d_A\psi \wedge ds + (\partial_s\alpha + [\Psi, \alpha]) \wedge ds. \tag{39}$$

The Laplace operator $\hat{\Delta}_{(A,\Psi)}$ on functions $\psi \in \Omega^0(\mathbb{R} \times \Sigma, \operatorname{ad}(\bar{P}))$ is

$$\hat{\Delta}_{(A,\Psi)}\psi = \hat{d}^*_{(A,\Psi)}\hat{d}_{(A,\Psi)}\psi = (\Delta_{A^\infty} - \partial_s^2)\psi - \partial_s[\Psi, \psi]. \tag{40}$$

The local slice condition for a connection $A + \Psi\, ds \in \mathcal{A}(\bar{P})$ with respect to the reference connection $\hat{A} \in \mathcal{A}(\bar{P})$ is

$$d^*_{\hat{A}}(A - \hat{A}) - \partial_s\Psi = 0. \tag{41}$$

Remark 6.1 We use the notation $W^{k,l;p}(I \times \Sigma)$ for the parabolic Sobolev spaces as defined in [17, Chap. 1.3, 2.2].

The aim of this section is to prove the following compactness theorem.

Theorem 6.2 (Compactness) *Let (A^ν, Ψ^ν), $\nu \in \mathbb{N}$, be a sequence of solutions to the perturbed Yang-Mills gradient flow equation*

$$\partial_s A + d_A^* F_A - d_A\Psi + \nabla\mathcal{V}(A) = 0. \tag{42}$$

Assume that there exist critical manifolds $\mathcal{C}^\pm$ such that every (A^ν, Ψ^ν) is a connecting trajectory between $\mathcal{C}^-$ and $\mathcal{C}^+$. Then for every $k < 2$ and $p < \infty$ and every compact interval $I \subseteq \mathbb{R}$ there exists a sequence $g^\nu \in \mathcal{G}(I \times P)$ of gauge transformations such that a subsequence of the gauge transformed sequence $(g^\nu)^*(A^\nu, \Psi^\nu)$ converges in $W^{k,p}(I \times \Sigma)$ to a solution (A^*, Ψ^*) of (42).

We start with the following weaker statement.

Theorem 6.3 *Let (A^ν, Ψ^ν) be a sequence of solutions of the perturbed Yang-Mills gradient flow equation (42). Assume that every (A^ν, Ψ^ν) is a connecting trajectory between some fixed pair $\mathcal{C}^\pm$ of critical manifolds. Then for every $r < \infty$ and every compact interval $I \subseteq \mathbb{R}$ there exists a constant $C(I, r)$, a smooth connection $\mathbb{A} = A^\infty + \Psi^\infty \wedge ds \in \mathcal{A}(I \times P)$, and a sequence $(g^\nu) \subset \mathcal{G}(I \times P)$ of gauge transformations such that the difference*

$$\alpha^\nu := (g^\nu)^* A^\nu - A^\infty, \qquad \psi^\nu := (g^\nu)^* \Psi^\nu - \Psi^\infty$$

satisfies the uniform bound

$$\|\alpha^\nu\|_{W^{1,2;r}(I\times\Sigma)} + \|\psi^\nu\|_{W^{2,r}(I\times\Sigma)} \leq C(I, r).$$

Proof: The proof, which we divide into three steps, is based on Uhlenbeck's weak compactness theorem, cf. the exposition [35] for details.

Step 1 *Let $1 < p < 4$. There exists a constant $C(p)$ such that there holds the uniform curvature bound*

$$\|F_{(A^\nu,\Psi^\nu)}\|_{L^p(I\times\Sigma)} \leq C(p).$$

Since the estimate is invariant under gauge transformations it suffices to prove it for $\Psi^\nu = 0$. Then the curvature is given by $F_{(A^\nu,0)} = F_{A^\nu} + (d_{A^\nu}^* F_{A^\nu} + \nabla \mathcal{V}(A^\nu))\, ds$. Uniform L^p-bounds for the terms F_{A^ν} and $d_{A^\nu}^* F_{A^\nu}$ hold by Lemmata B.7 and B.10. A uniform estimate for $\nabla\mathcal{V}(A^\nu))$ is given by (56).

Step 2 *Let $3 < p < 4$ and choose $\varepsilon > 0$. There exists a sequence $g^\nu \in \mathcal{G}^{2,p}(\bar{P})$ of gauge transformations and a smooth reference connection (A^∞, Ψ^∞) such that (up to extraction of a subsequence) the sequence $(g^\nu)^*(A^\nu, \Psi^\nu)$ satisfies the following three conditions.*

(i) Each connection $(g^\nu)^*(A^\nu, \Psi^\nu)$ is in local slice with respect to (A^∞, Ψ^∞).

(ii) The difference $(\beta^\nu, \psi^\nu) := (g^\nu)^*(A^\nu, \Psi^\nu) - (A^\infty, \Psi^\infty)$ is uniformly bounded in $W^{1,p}(I \times \Sigma)$.

(iii) The sequence (β^ν, ψ^ν) satisfies the uniform bound

$$\|\beta^\nu\|_{C^0(I\times\Sigma)} + \|\psi^\nu\|_{C^0(I\times\Sigma)} < \varepsilon.$$

The sequence (A^ν, Ψ^ν) satisfies a uniform L^p-curvature bound by Step 1. Hence Uhlenbeck's weak compactness theorem (cf. [35, Theorem 7.1]) yields a sequence $g^\nu \in \mathcal{G}^{2,p}(\bar{P})$ of gauge transformations such that a subsequence of $(g^\nu)^*(A^\nu, \Psi^\nu)$ converges weakly in $W^{1,p}(I \times \Sigma)$ to a limit (A', Ψ'). It is in particular uniformly bounded in $W^{1,p}(I \times \Sigma)$ and contains a subsequence which converges in $C^0(I \times \Sigma)$ to (A', Ψ'). We again label this subsequence by ν. It hence follows from the local slice theorem (cf. [35, Theorem 8.1]) that every $(g^\nu)^*(A^\nu, \Psi^\nu)$ with ν large enough can be put in local slice with respect to any smooth reference connection (A^∞, Ψ^∞) sufficiently close in $W^{1,p}(I \times \Sigma)$ to (A', Ψ'). Therefore condition (i) is satisfied. Moreover, the local slice theorem asserts that this can be done preserving the uniform bound in $W^{1,p}(I\times\Sigma)$ and the uniform bound (with constant ε) in $C^0(I\times\Sigma)$. Thus also conditions (ii) and (iii) are satisfied.

Step 3 *Proof of the theorem.*

In the subsequent calculations we drop the index ν. Expanding d_A, d_A^* and F_A as

$$d_A = d_{A^\infty} + [\beta \wedge \cdot\,], \qquad d_A^* = d_{A^\infty}^* - *[\beta \wedge * \cdot\,],$$
$$F_A = F_{A^\infty} + d_{A^\infty}\beta + \frac{1}{2}[\beta \wedge \beta],$$

equation (42) reads

$$\begin{aligned}0 = \partial_s A^\infty + \partial_s\beta + d_{A^\infty}^* F_{A^\infty} - *\big[\beta \wedge *\big(F_{A^\infty} + d_{A^\infty}\beta + \frac{1}{2}[\beta \wedge \beta]\big)\big] + d_{A^\infty}^* d_{A^\infty}\beta \\ + \frac{1}{2} d_{A^\infty}^*[\beta \wedge \beta] - d_{A^\infty}\psi - [\beta \wedge \psi] - d_{A^\infty}\Psi^\infty - [\beta \wedge \Psi^\infty] + \nabla\mathcal{V}(A).\end{aligned} \tag{43}$$

We combine this equation with the local slice condition (41) to obtain for β the parabolic PDE

$$\begin{aligned}\partial_s\beta + \Delta_{A^\infty}\beta = -\partial_s A^\infty - d_{A^\infty}^* F_{A^\infty} + d_{A^\infty}\partial_s\Psi^\infty - \frac{1}{2} d_{A^\infty}^*[\beta \wedge \beta] + d_{A^\infty}\partial_s\psi \\ + *\big[\beta\wedge*\big(F_{A^\infty} + d_{A^\infty}\beta + \frac{1}{2}[\beta\wedge\beta]\big)\big] + d_{A^\infty}\psi + [\beta\wedge\psi] + d_{A^\infty}\Psi^\infty + [\beta\wedge\Psi^\infty] - \nabla\mathcal{V}(A).\end{aligned} \tag{44}$$

Applying $d_{A^\infty}^*$ to both sides of equation (43) and substituting

$$d_{A^\infty}^*\partial_s\beta = \partial_s^2\psi + \partial_s^2\Psi^\infty + *[\partial_s A^\infty \wedge *\beta]$$

according to (41), and using that

$$d_{A^\infty}^* \nabla\mathcal{V}(A) = d_A^* \nabla\mathcal{V}(A) + *[\beta \wedge *\nabla\mathcal{V}(A)] = *[\beta \wedge *\nabla\mathcal{V}(A)]$$

yields for ψ the elliptic PDE

$$\begin{aligned}\hat{\Delta}_{A^\infty}\psi = d_{A^\infty}^*\partial_s A^\infty + \partial_s^2\Psi^\infty - \Delta_{A^\infty}\Psi^\infty + *[\partial_s A^\infty \wedge *\beta] + [*F_{A^\infty} \wedge *d_{A^\infty}\beta] \\ - \frac{1}{2}[*F_{A^\infty} \wedge *[\beta\wedge\beta]] - *d_{A^\infty}\big[\beta \wedge *\big(F_{A^\infty} + \frac{1}{2}[\beta\wedge\beta]\big)\big] + *\big[\beta \wedge *d_{A^\infty}^* d_{A^\infty}\beta\big] \\ - d_{A^\infty}^*[\beta\wedge\psi] - d_{A^\infty}^*[\beta\wedge\Psi^\infty] + *[\beta \wedge *\nabla\mathcal{V}(A)]. \quad (45)\end{aligned}$$

Let $p > 1$ arbitrary. From equation (44) it follows by standard parabolic regularity theory that

$$\begin{aligned}c^{-1}\|\beta\|_{W^{1,2;p}} \leq 1 + \|\beta\|_{L^p} + \|\{\beta,[\beta\wedge\beta]\}\|_{L^p} + \|\{\nabla_{A^\infty}\beta,\beta\}\|_{L^p} \\ + \|\{\beta,\psi\}\|_{L^p} + \|d_{A^\infty}\psi\|_{L^p} + \|d_{A^\infty}\partial_s\psi\|_{L^p} + \|\nabla\mathcal{V}(A)\|_{L^p}. \quad (46)\end{aligned}$$

From equation (45) and elliptic regularity we obtain the estimate

$$\begin{aligned}c^{-1}\|\psi\|_{W^{2,p}} \leq 1 + \|\beta\|_{L^p} + \|\{\beta,\beta\}\|_{L^p} + \|\nabla_{A^\infty}\beta\|_{L^p} + \|\{d_{A^\infty}\beta,\beta,\beta\}\|_{L^p} \\ + \|\{\beta, d_{A^\infty}^* d_{A^\infty}\beta\}\|_{L^p} + \|\nabla_{A^\infty}[\beta\wedge\psi]\|_{L^p} + \|\psi\|_{L^p} + \|[\beta \wedge *\nabla\mathcal{V}(A)]\|_{L^p}. \\ (47)\end{aligned}$$

Now let $3 < p < 4$. By Step 2 there holds a uniform bound for $\|\beta\|_{C^0}$ and $\|\beta\|_{W^{1,p}}$. Applying Proposition A.7 we find that the term $\|\nabla\mathcal{V}(A)\|_{L^p}$ is controlled by $\|F_A\|_{L^p} \leq c\big(1 + \|d_{A_0}\beta\|_{L^p} + \|[\beta\wedge\beta]\|_{L^p}\big)$. It thus follows that each term on the right-hand side of (46), except the term $\|d_{A^\infty}\partial_s\psi\|_{L^p}$, is uniformly bounded. This term $\|d_{A^\infty}\partial_s\psi\|_{L^p}$ can now be estimated using (47). Namely, the expression

$$\|\{\beta, d_{A^\infty}^* d_{A^\infty}\beta\}\|_{L^p} \leq c\|\beta\|_{C^0}\|d_{A^\infty}^* d_{A^\infty}\beta\|_{L^p}$$

appearing in (47) becomes absorbed by the left-hand side of (46) after fixing ε in condition (iii) of Step 2 sufficiently small. Hence it follows that ψ is uniformly bounded in $W^{2,p}$ and β is uniformly bounded in $W^{1,2;p}$. As a consequence we obtain for $\nabla_{A^\infty}\beta$ the uniform bound (with $p_1 = \frac{5}{2} + r > 4$)

$$\begin{aligned}\|\nabla_{A^\infty}\beta\|_{L^{p_1}(I\times\Sigma)}^{p_1} &\leq \int_I \|\nabla_{A^\infty}\beta\|_{L^3(\Sigma)}^5 + \|\nabla_{A^\infty}\beta\|_{L^{6r}(\Sigma)}^{2r} \\ &\leq \int_I \|\nabla_{A^\infty}\beta\|_{L^3(\Sigma)}^5 + c\|\nabla_{A^\infty}^2\beta\|_{L^2(\Sigma)}^{2r}.\end{aligned}$$

In the first line we used Hölder's inequality, and the second line follows from the Sobolev embedding $W^{1,2}(\Sigma) \hookrightarrow L^{6r}(\Sigma)$. Both terms in the last line are uniformly bounded for $r < 2$. For the last one this follows by what we have already shown, while for the first one we use interpolation (cf. Lemma C.9) and Sobolev embedding to obtain

$$\begin{aligned}\|\nabla_{A^\infty}\beta\|_{L^5(I,L^3(\Sigma))} \leq \|\nabla_{A^\infty}\beta\|_{W^{\frac{1}{3},2}(I,W^{\frac{1}{3},2}(\Sigma))} &\leq \|\beta\|_{W^{\frac{1}{3},2}(I,W^{\frac{4}{3},2}(\Sigma))} \\ &\leq \|\beta\|_{W^{1,2;2}(I\times\Sigma)}.\end{aligned}$$

Thus indeed $\nabla_{A^\infty}\beta$ is uniformly bounded in $L^{p_1}(I \times \Sigma)$ for $p_1 < \frac{9}{2}$ and we can repeat the previous line of arguments with p replaced by p_1 to get uniform bound for ψ in W^{2,p_1} and for β in $W^{1,2;p_1}$. Repeating this argument a finite number of times, we inductively obtain uniform such bounds for every $p < \infty$. This completes the proof. $\square$

Proposition 6.4 *Let (A^ν, Ψ^ν) be a sequence of connections satisfying the hypothesis of Theorem 6.3. In the notation employed there, assume $\beta^\nu \in W^{1,2;p}(I \times \Sigma)$ and $\psi^\nu \in W^{2,p}(I \times \Sigma)$ for all $p < \infty$. Then for every $r > 1$ and sufficiently large $p = p(r) < \infty$ there exists a constant $c(I, r)$ such that the estimate*

$$\|F_{A^\nu}\|_{W^{1,3;r}(I\times\Sigma)} \leq c(I,r)\big(1 + \|\beta^\nu\|_{W^{1,2;p}(I\times\Sigma)} + \|\psi^\nu\|_{W^{2,p}(I\times\Sigma)}\big)$$

holds for all $\nu \in \mathbb{N}$.

Proof: For simplicity we drop the index ν. Let $A^\infty \in \mathcal{A}(\mathbb{R} \times P)$ be the smooth reference connection as in the proof of Theorem 6.3, i.e. $\beta = A - A^\infty$, and let $\mathcal{L}_{A^\infty} := \frac{d}{ds} + \Delta_{A^\infty}$ be the heat operator induced by A^∞. From (9) it follows that F_A satisfies the evolution equation

$$\begin{aligned}\mathcal{L}_{A^\infty}F_A &= d_A\dot{A} + \Delta_{A^\infty}F_A \\ &= -d_Ad_A^*F_A + d_A\Psi - d_A\nabla\mathcal{V}(A) + \Delta_{A^\infty}F_A \\ &= [d_{A^\infty} * \beta \wedge *F_A] + [\beta \wedge *[\beta \wedge *F_A]] + d_A\Psi - d_A\nabla\mathcal{V}(A).\end{aligned}$$

By our assumptions, the right-side of the equation is uniformly bounded in $W^{0,1;r}(I \times \Sigma)$. To see that this holds for the term F_A, we expand it as $F_A = F_{A^\infty} + d_{A^\infty}\beta + \frac{1}{2}[\beta \wedge \beta]$ and use the assumption on β. The required bound on $d_A\nabla\mathcal{V}(A)$ is satisfied by Proposition A.8. The claim then follows from standard parabolic regularity results, cf. for instance [17, Chap. 5.1]. $\square$

Proposition 6.5 *Let (A^ν, Ψ^ν) be a sequence of solutions of the perturbed Yang-Mills gradient flow equation* (42) *satisfying the assumptions of Theorem 6.3. Then for every $r < \infty$ and every compact interval $I \subseteq \mathbb{R}$ there exists a constant $C(I, r)$ such that the uniform curvature bound*

$$\|F_{(A^\nu,\Psi^\nu)}\|_{W^{1,r}(I\times\Sigma)} \leq C(I, r)$$

holds for all $\nu \in \mathbb{N}$.

Proof: For simplicity we drop the index ν. After applying a suitable gauge transformation we may assume that $\Psi = 0$. Then we have that $F_{(A^\nu,\Psi^\nu)} = F_A - (d_A^* F_A + \nabla\mathcal{V}(A))$. From Proposition 6.4 it follows that $\|F_A\|_{W^{1,3;p}(I\times\Sigma)}$ is uniformly bounded. This immediately implies the required uniform estimate for F_A and for $d_A^* F_A$. The uniform bound for $\nabla\mathcal{V}(A)$ follows from Proposition A.7. □

Proof: (Theorem 6.2) By Proposition 6.5 the sequence (A^ν, Ψ^ν) has curvature uniformly bounded in $W^{1,p}(I \times \Sigma)$ for every $p < \infty$. Hence Uhlenbeck's weak compactness theorem (with one derivative more, cf. [21]) applies and shows that after modifying the sequence by suitable gauge transformations and passing to a subsequence, there holds weak convergence

$$\lim_{\nu\to\infty} (A^\nu, \Psi^\nu) = (A^*, \Psi^*)$$

in $W^{2,p}(I \times \Sigma)$ for some limiting connection (A^*, Ψ^*), as claimed. □

7 Transversality

7.1 Universal moduli spaces

For the construction of the relevant Banach space Y of perturbations we refer to Appendix 2.2. We fix a pair of critical manifolds $\mathcal{C}^\pm$ and consider the smooth Banach space bundle

$$\mathcal{E}(\delta, \mathcal{C}^-, \mathcal{C}^+) \to \mathcal{B}(\delta, \mathcal{C}^-, \mathcal{C}^+) \times Y,$$

cf. Section 3.2 for definitions. We define the smooth section $\mathcal{F}$ of $\mathcal{E}$ by

$$\mathcal{F} : [(A, \Psi, \mathcal{V})] \mapsto [\partial_s A + d_A^* F_A - d_A\Psi + \nabla\mathcal{V}(A)].$$

Its zero set $\mathcal{M}^{\mathrm{univ}}(\mathcal{C}^-, \mathcal{C}^+) := \mathcal{F}^{-1}(0)$ is the so-called the *universal moduli space*. Thus the perturbation $\mathcal{V} = (\mathcal{V}, \mathcal{V}^+)$ which had been kept fixed so far is now allowed to vary over the Banach space Y.

Theorem 7.1 *The horizontal differential $d_u\mathcal{F}$ of the map $\mathcal{F}$ is surjective for every $u \in \mathcal{F}^{-1}(0)$.*

We give the proof in the next section. Assuming Theorem 7.1 it follows from the implicit function theorem that the universal moduli space $\mathcal{M}^{\mathrm{univ}}(\mathcal{C}^-,\mathcal{C}^+)$ is a smooth Banach manifold. Let $\pi : \mathcal{M}^{\mathrm{univ}}(\mathcal{C}^-,\mathcal{C}^+) \to Y$ denote the projection onto the second factor. It is a smooth Fredholm map whose index is given by the Fredholm index of $\mathcal{D}_A$. Hence we may apply to π the Sard-Smale theorem for Fredholm maps between Banach manifolds, cf. [1, Theorem 3.6.15], from which it follows that the set of regular values

$$\mathcal{R} := \left\{\mathcal{V} \in Y \,\middle|\, d_u\pi \text{ is surjective for all } u \in \mathcal{M}(\mathcal{C}^-,\mathcal{C}^+;\mathcal{V})\right\} \subset Y$$

is residual in Y. Hence in particular there exists a regular value $\mathcal{V}_0$ in every arbitrarily small ball $B_\varepsilon(0)$ (with respect to the norm on Y) around zero. For such a $\mathcal{V}_0$, the moduli space $\mathcal{M}(\mathcal{C}^-,\mathcal{C}^+;\mathcal{V}_0)$ is a submanifold of $\mathcal{M}^{\mathrm{univ}}(\mathcal{C}^-,\mathcal{C}^+)$ of dimension equal to $\operatorname{ind}\mathcal{D}_A$.

7.2 Surjectivity of linearized operators

Let $(A,\Psi,\mathcal{V})$ be a smooth solution of the perturbed Yang-Mills gradient flow equation (63). After applying a suitable gauge transformation we may assume $\Psi = 0$. The setup for the discussion of the linearization of the section $\mathcal{F}$ parallels the one introduced in Section 5.2. It is given by the linear operator

$$\hat{\mathcal{D}}_{(A,\mathcal{V})} : \mathcal{Z}_A^{\delta,p} \times Y \to \mathcal{L}^{\delta,p}, \quad (\alpha,\psi,v) \mapsto \mathcal{D}_A(\alpha,\psi) + \nabla v(A).$$

Note that $\hat{\mathcal{D}}_{(A,\mathcal{V})}$ is the sum of the Fredholm operator $\mathcal{D}_A$ and the bounded operator $v \mapsto \nabla v(A)$, and therefore has closed range. The Fredholm property of $\mathcal{D}_A$ has been shown in Theorem 5.3. The assertion on boundedness follows from

$$
\begin{aligned}
\|\nabla v(A)\|^p_{L^p(\mathbb{R}\times\Sigma)} &= \int_{-T}^{T} \|\nabla v(A)\|^p_{L^p(\Sigma)}\, ds \\
&\leq \int_{-T}^{T} \sum_{\ell=1}^{\infty} |c_\ell|^p \cdot \|\nabla v_\ell(A)\|^p_{L^p(\Sigma)}\, ds \\
&\leq \int_{-T}^{T} \sum_{\ell=1}^{\infty} |c_\ell|^p \cdot \|v_\ell\|^p \cdot (1 + \|\alpha(A)\|^p_{L^p(\Sigma)})\, ds \\
&= \sum_{\ell=1}^{\infty} |c_\ell|^p \cdot \|v_\ell\|^p \cdot \int_{-T}^{T} (1 + \|\alpha(A)\|^p_{L^p(\Sigma)})\, ds \\
&\leq c \int_{-T}^{T} (1 + \|F_A\|^p_{L^p(\Sigma)})\, ds \cdot \|v\|^p.
\end{aligned}
$$

The first line holds for some constant $T < \infty$ because $A(s)$ is contained in the support of v only for some finite time interval (by construction, $\operatorname{supp} v$ is contained in the complement of some L^2-neighbourhood of $\mathcal{C}^- \cap \mathcal{C}^+$). The last estimate is by Proposition A.7 and the definition of $\|v\|$. Because $\hat{\mathcal{D}}_{(A,\mathcal{V})}$ has closed range, the statement of Theorem 7.1 reduces to that of the following Proposition.

Proposition 7.2 *The image of the operator* $\hat{\mathcal{D}}_{(A,\mathcal{V})} : \mathcal{Z}^{\delta,p}_A \times Y \to \mathcal{L}^{\delta,p}$ *is dense in* $\mathcal{L}^{\delta,p}$*, for every smooth solution* $(A, 0, \mathcal{V})$ *of* (63).

Proof: Density of the range is equivalent to triviality of its annihilator. This means that, given $\eta \in (\mathcal{L}^{\delta,p})^* = \mathcal{L}^{-\delta,q}$ (where $p^{-1} + q^{-1} = 1$) with

$$
\langle \mathcal{D}_{A,\mathcal{V}}(\alpha, \psi, v), \eta\rangle_{\mathbb{R}\times\Sigma} = 0 \qquad \text{for all } (\alpha, \psi, v) \in \mathcal{Z}^{\delta,p}_A, \tag{48}
$$

then $\eta = 0$. Condition (48) is equivalent to

$$
\langle \mathcal{D}_A(\alpha, \psi), \eta\rangle_{\mathbb{R}\times\Sigma} = 0, \tag{49}
$$

$$
\langle \nabla v(A), \eta\rangle_{\mathbb{R}\times\Sigma} = 0 \tag{50}
$$

for all $(\alpha, \psi, v) \in \mathcal{Z}^{\delta,p}_A$. Assume by contradiction that there exists $0 \neq \eta \in \mathcal{L}^{-\delta,q}$ which satisfies both (49) and (49). Then it follows from the identity

$$
0 = \langle \eta, \mathcal{D}_A \xi\rangle_{\mathbb{R}\times\Sigma} = \int_{-\infty}^{\infty} \langle \mathcal{D}^*_{A(s)} \eta(s), \xi(s)\rangle\, ds
$$

that $\mathcal{D}_A^*\eta = 0$, where $\mathcal{D}_A^* := -\frac{d}{ds} + \mathcal{H}_A$. Hence Proposition 7.5 below applies and yields a contradiction to (50). This shows that $\eta = 0$ and proves the proposition. □

For the proof of Proposition 7.5 we need the following auxiliary results.

Proposition 7.3 (Slicewise orthogonality) *Assume $\eta \in \mathcal{L}^{-\delta,q}$ satisfies $\mathcal{D}_A^*\eta = 0$ on $\mathbb{R}\times\Sigma$. Then the relation $\langle \dot{A}(s), \eta(s)\rangle = 0$ holds for all $s \in \mathbb{R}$.*

Proof: Set $\beta(s) := \langle \dot{A}(s), \eta(s)\rangle$. Because $\dot{A}$ satisfies the linearized Yang-Mills gradient flow equation $\mathcal{D}_A\dot{A} = 0$ it follows that

$$\dot{\beta} = \langle \dot{A}, \dot{\eta}\rangle + \langle \ddot{A}, \eta\rangle = \langle \dot{A}, \mathcal{H}_A\eta\rangle + \langle -\mathcal{H}_A\dot{A}, \eta\rangle = 0.$$

Thus β is constant. Since $\lim_{s\to-\infty}\dot{A} = 0$ it follows that β vanishes identically. □

Proposition 7.4 (No return) *For every $\delta > 0$ there exists a constant $\varepsilon > 0$ with the following significance. Assume A is a solution* (63) *on $\mathbb{R}\times\Sigma$ and there exists $s_0 \in \mathbb{R}$ such that $A_0 := A(s_0)$ is not a Yang-Mills connection. Then it holds that*

$$\operatorname{dist}_{L^2}(A(s), \mathcal{O}(A_0)) < 3\varepsilon \qquad \Rightarrow \qquad s \in (s_0 - \delta, s_0 + \delta).$$

Proof: Assume by contradiction that there is a sequence of positive numbers $\varepsilon_i \to 0$ as $i \to \infty$ and a sequence $(s_i) \subseteq \mathbb{R}$ such that

$$\operatorname{dist}_{L^2}(A(s_i), \mathcal{O}(A_0)) < 3\varepsilon_i \tag{51}$$

and $s_i \notin (s_0 - \delta, s_0 + \delta)$. Denote $A^\pm := \lim_{s\to\pm\infty} A(s) \in \mathcal{C}^\pm$. By the L^2-local slice theorem (cf. Theorem 2.1) and the assumption that A_0 is not Yang-Mills it follows that

$$\operatorname{dist}_{L^2}(A^\pm, \mathcal{O}(A_0)) > 3\varepsilon_i \tag{52}$$

holds for all sufficiently large $i \geq i_0$. Assume first that the sequence (s_i) is unbounded. Hence we can choose a subsequence (without changing notation) such that s_i converges to $-\infty$ or to $+\infty$. It follows that (for one sign $\pm$)

$$A(s_i) \stackrel{L^2(\Sigma)}{\longrightarrow} A^\pm \qquad \text{as} \quad i \to \pm\infty.$$

This contradicts (52). Therefore the sequence (s_i) has an accumulation point $s_* \in \mathbb{R}$ and there exists a subsequence (s_i) with $\lim_{i\to\infty} s_i = s_*$.

By continuity of the gradient flow line A (as a map $\mathbb{R} \to L^2(\Sigma)$) it follows that $\lim_{i\to\infty} A(s_i) = A(s_*)$ in $L^2(\Sigma)$. From (51) we hence infer that $\mathrm{dist}_{L^2}(A(s_*), \mathcal{O}(A_0)) = 0$. Again by the L^2-local slice theorem this implies that $A(s_*) \in \mathcal{O}(A_0)$ and $\mathcal{YM}^{\mathcal{V}}(A(s_*)) = \mathcal{YM}^{\mathcal{V}}(A_0)$. As $\mathcal{YM}^{\mathcal{V}}(A(s))$ is strictly monotone decreasing in s it follows that $s_* = s_0$, which contradicts $s_i \notin (s_0 - \delta, s_0 + \delta)$. Hence the assumption was wrong and the claim follows. $\square$

Proposition 7.5 *Let A, A_0, s_0 be as in Proposition 7.4. Assume $\eta \in \mathcal{L}^{-\delta,q}$ satisfies $\eta(s_0) \neq 0$. Then there exists a constant $\varepsilon > 0$ and a gauge-invariant smooth map $\mathcal{V}_0 : \mathcal{A}(P) \to \mathbb{R}$ such that*

(i) $\mathrm{supp}\, \mathcal{V}_0 \subseteq \{A \in \mathcal{A}(P) \mid \mathrm{dist}_{L^2}(A, \mathcal{O}(A_0)) \leq 2\varepsilon\}$,

(ii) $\langle \nabla \mathcal{V}_0(A_0), \eta_0 \rangle_\Sigma = \|\eta_0\|^2_{L^2(\Sigma)}$,

(iii) $\langle \nabla \mathcal{V}_0(A), \eta \rangle_{\mathbb{R}\times\Sigma} > 0$.

Before entering the proof of the proposition we remark the following. Denote $\alpha'(s) := d\alpha(A(s))\eta(s)$, where $\alpha : \mathcal{A}(P) \to L^2(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))$ is defined as in (5). Then it follows from continuous differentiability of the path $A : \mathbb{R} \to L^2(\Sigma)$ and continuity of the map $\eta : \mathbb{R} \to L^2(\Sigma)$ that there exists a constant $\delta > 0$ with the following significance. For all $s \in (s_0 - \delta, s_0 + \delta)$ we have that

(A) $\|\eta(s)\|_{L^2} \leq 2\|\eta_0\|_{L^2}$,

(B) $\langle \alpha'(s), \eta_0 \rangle \geq \frac{1}{2}\|\eta_0\|^2_{L^2} > 0$,

(C) and with $\mu := \|\partial_s A(s_0)\|_{L^2} > 0$ that

$$\frac{1}{2}\mu \leq \frac{\mathrm{dist}_{L^2}(A(s) - \mathcal{O}(A_0))}{|s - s_0|} \leq \frac{3}{2}\mu.$$

Proof: (Proposition 7.5) Let $\varepsilon > 0$ be such that the following two conditions are satisfied. First, any $A \in \mathcal{A}(P)$ with $\mathrm{dist}_{L^2}(A, \mathcal{O}(A_0)) < 2\varepsilon$ can be put in local slice with respect to A_0. Second, the condition

$$\mathrm{dist}_{L^2}(A(s), \mathcal{O}(A_0)) < 3\varepsilon \qquad \Rightarrow \qquad s \in (s_0 - \delta, s_0 + \delta)$$

holds for all $s \in \mathbb{R}$. The existence of such an ε follows from the L^2-local slice theorem and Proposition 7.4. Now let $\rho : \mathbb{R} \to [0, 1]$ be a smooth cut-off function with support in $[-4,4]$ and satisfying $\|\rho'\|_{L^\infty} < 1$ and $\rho(r) = 1$ for $r \in [-1,1]$. Define $\rho_\varepsilon(r) := \rho(\varepsilon^{-2} r)$. Note that $\|\rho_\varepsilon'\|_{L^\infty} < \varepsilon^{-2}$. Define

$$\mathcal{V}_0(A) := \rho_\varepsilon(\|\alpha\|_{L^2}^2)\langle \eta_0 + \alpha, \eta_0\rangle$$

with $\alpha = \alpha(A)$ as in (5). The perturbation $\mathcal{V}_0$ clearly satisfies condition (i). Furthermore, it follows from Proposition A.3 (using $\rho_\varepsilon(0) = 1$, $\rho_\varepsilon'(0) = 0$, and $\alpha(s_0) = 0$) that

$$d\mathcal{V}_0(A_0)\eta_0 = \langle \eta_0 - d_A T_{A_0,\alpha}\eta_0, \eta_0\rangle = \|\eta_0\|_{L^2}^2,$$

so that condition (ii) is satisfied. It remains to show property (iii). We fix constants $\sigma_1, \sigma_2, s_1, s_2$ with $\sigma_1 < s_1 < s_0 < s_2 < \sigma_2$ as follows. Let s_2 be such that $\operatorname{dist}_{L^2}(A(s_2), \mathcal{O}(A_0)) = \varepsilon$ and $\operatorname{dist}_{L^2}(A(s), \mathcal{O}(A_0)) < \varepsilon$ for all $s \in (s_0, s_2)$, and similarly for s_1. Let σ_2 be such that $\operatorname{dist}_{L^2}(A(\sigma_2), \mathcal{O}(A_0)) = 2\varepsilon$ and $\operatorname{dist}_{L^2}(A(s), \mathcal{O}(A_0)) > 2\varepsilon$ for all $s \in (\sigma_2, s_0 + \delta)$, similarly for σ_1. It then follows that

$$\begin{aligned}\langle \nabla \mathcal{V}_0(A), \eta\rangle_{\mathbb{R}\times\Sigma} &= \int_{s_0-\delta}^{s_0+\delta} d\mathcal{V}_0(A(s))\eta(s)\, ds = \int_{s_0-\delta}^{s_0+\delta} \rho_\varepsilon(\|\alpha\|^2)\langle \alpha', \eta_0\rangle\, ds \\ &+ 2\|\eta_0\|^2 \int_{s_0-\delta}^{s_0+\delta} \rho_\varepsilon'(\|\alpha\|^2)\langle \alpha, \alpha'\rangle\, ds + 2\int_{s_0-\delta}^{s_0+\delta} \rho_\varepsilon'(\|\alpha\|^2)\langle \alpha, \alpha'\rangle\langle \alpha, \eta_0\rangle\, ds.\end{aligned}$$

We estimate the last three terms separately. For the first one we obtain

$$\begin{aligned}&\int_{s_0-\delta}^{s_0+\delta} \rho_\varepsilon(\|\alpha\|^2)\langle \alpha', \eta_0\rangle\, ds \geq \int_{s_1}^{s_2} 1 \cdot \langle \alpha', \eta_0\rangle\, ds \geq \frac{1}{2}(s_2 - s_1)\|\eta_0\|^2 \\ &\geq \frac{1}{6\mu}\big(\operatorname{dist}_{L^2}(A(s_1), \mathcal{O}(A_0)) + \operatorname{dist}_{L^2}(A(s_2), \mathcal{O}(A_0))\big)\|\eta_0\|^2 = \frac{1}{3\mu}\|\eta_0\|^2\varepsilon.\end{aligned}$$

The second inequality uses property (B), the third one property (C) above. We define functions $f, g : \mathbb{R} \to \mathbb{R}$ by

$$f(s) := \langle \alpha(s), \alpha'(s)\rangle \qquad \text{and} \qquad g(s) := \langle \alpha(s), \eta_0\rangle.$$

As $\alpha(s_0) = 0$ it follows that $f(s_0) = g(s_0) = 0$. By Proposition A.3 we have that $\dot\alpha(s_0) = \dot A(s_0)$ and $\alpha'(s_0) = \eta_0$. Using Proposition 7.3 it follows that

$$\dot f(s_0) = \langle \dot\alpha(s_0), \alpha'(s_0)\rangle + \langle \alpha(s_0), \partial_s(\alpha')(s_0)\rangle = \langle \dot A(s_0), \eta_0\rangle = 0,$$

and similarly that

$$\dot g(s_0) = \langle \dot\alpha(s_0), \eta_0\rangle = \langle \dot A(s_0), \eta_0\rangle = 0.$$

Hence there exists a constant $C(A)$ (independent of η) such that

$$|f(s)| \le C(A)\|\eta\|_{C^0((s_0-\delta,s_0+\delta),L^2(\Sigma))}(s-s_0)^2 \le 2C(A)\|\eta_0\|_{L^2(\Sigma)}(s-s_0)^2$$

(here we use property (A) in the second inequality) and

$$|g(s)| \le C(A)\|\eta_0\|_{L^2(\Sigma)}(s-s_0)^2$$

holds for all $s \in (s_0-\delta, s_0+\delta)$. For the second term we therefore obtain the estimate

$$\begin{aligned}
&2\|\eta_0\|^2 \int_{s_0-\delta}^{s_0+\delta} \rho_\varepsilon'(\|\alpha\|^2)\langle\alpha,\alpha'\rangle\,ds = 2\|\eta_0\|^2 \int_{\sigma_1}^{\sigma_2} \rho_\varepsilon'(\|\alpha\|^2)\langle\alpha,\alpha'\rangle\,ds \\
\ge\ & -2\varepsilon^{-2}\|\eta_0\|^2 \int_{\sigma_1}^{\sigma_2} |\langle\alpha,\alpha'\rangle|\,ds \\
\ge\ & -4\varepsilon^{-2}C(A)\|\eta_0\|^3 \int_{\sigma_1}^{\sigma_2} (s-s_0)^2\,ds \\
=\ & -\frac{4}{3}\varepsilon^{-2}C(A)\|\eta_0\|^3\big(|\sigma_2-s_0|^3+|\sigma_1-s_0|^3\big) \\
\ge\ & -\frac{4}{3}\varepsilon^{-2}C(A)\|\eta_0\|^3\cdot\big(\frac{2}{3\mu}\big)^3\big(\operatorname{dist}_{L^2}(A(\sigma_2),\mathcal{O}(A_0))^3 \\
& +\operatorname{dist}_{L^2}(A(\sigma_1),\mathcal{O}(A_0))^3\big) \\
=\ & -\frac{64}{3}\big(\frac{2}{3\mu}\big)^3 C(A)\|\eta_0\|^3\varepsilon.
\end{aligned}$$

The last inequality is by property (C). The third term is estimated as follows.

$$\begin{aligned}
&2\int_{s_0-\delta}^{s_0+\delta} \rho_\varepsilon'(\|\alpha\|^2)\langle\alpha,\alpha'\rangle\langle\alpha,\eta_0\rangle\,ds \\
=\ & 2\int_{\sigma_1}^{\sigma_2} \rho_\varepsilon'(\|\alpha\|^2)\langle\alpha,\alpha'\rangle\langle\alpha,\eta_0\rangle\,ds \\
\ge\ & -8\varepsilon^{-2}C(A)^2\|\eta_0\|^2 \int_{\sigma_1}^{\sigma_2} (s-s_0)^4\,ds \\
=\ & -\frac{8}{5}\varepsilon^{-2}C(A)^2\|\eta_0\|^2\big(|\sigma_1-s_0|^5+|\sigma_2-s_0|^5\big) \\
\ge\ & -\frac{8}{5}\varepsilon^{-2}C(A)^2\|\eta_0\|^2\cdot\big(\frac{2}{3\mu}\big)^5\big(\operatorname{dist}_{L^2}(A(\sigma_2),\mathcal{O}(A_0))^5 \\
& +\operatorname{dist}_{L^2}(A(\sigma_1,\mathcal{O}(A_0))^5\big) \\
=\ & -\frac{512}{5}C(A)^2\|\eta_0\|^2\cdot\big(\frac{2}{3\mu}\big)^5\varepsilon^3.
\end{aligned}$$

The last inequality follows again from property (C). Combining these estimates it follows that

$$\langle \nabla \mathcal{V}_0(A), \eta \rangle_{\mathbb{R}^- \times \Sigma} \geq \\ \frac{1}{3\mu}\|\eta_0\|^2\varepsilon - \frac{64}{3}\Big(\frac{2}{3\mu}\Big)^3 C(A)\|\eta_0\|^3\varepsilon - \frac{512}{5}\Big(\frac{2}{3\mu}\Big)^5 C(A)^2\|\eta_0\|^2\varepsilon^3.$$

The last expression becomes strictly positive after choosing $\|\eta_0\|$ and ε still smaller if necessary (which does not affect the argumentation so far). This shows property (iii) and completes the proof. □

8 Yang-Mills Morse homology

8.1 Morse-Bott theory

We briefly recall Frauenfelder's *cascade construction* of Morse homology for Morse functions with degenerate critical points satisfying the Morse-Bott condition (cf. [14, Appendix C]). Let (M, g) be a Riemannian (Banach) manifold . A smooth function $f : M \to \mathbb{R}$ is called *Morse-Bott* if the set $\operatorname{crit}(f) \subset M$ of its critical points is a finite-dimensional submanifold of M and if for each $x \in \operatorname{crit}(f)$ the *Morse-Bott condition* $T_x \operatorname{crit}(f) = \ker H_x f$ is satisfied. As an additional datum, we fix a Morse function $h : \operatorname{crit}(f) \to \mathbb{R}$ which satisfies the Morse-Smale condition, i.e. the stable and unstable manifolds $W_h^s(x)$ and $W_h^u(y)$ of any two critical points $x, y \in \operatorname{crit}(h)$ intersect transversally. We assign to a critical point $x \in \operatorname{crit}(h) \subset \operatorname{crit}(f)$ the index

$$\operatorname{Ind}(x) := \operatorname{ind}_f(x) + \operatorname{ind}_h(x). \tag{53}$$

Definition 8.1 Let $x^-, x^+ \in \operatorname{crit}(h)$ and $m \in \mathbb{N}$. A *flow line from x^- to x^+ with m cascades* is a tuple $(\mathtt{x}, T) := (x_1, \ldots, x_m, t_1, \ldots, t_{m-1})$ with $x_j \in C^\infty(\mathbb{R}, M)$ and $t_j \in \mathbb{R}^+$ such that the following conditions are satisfied.

(i) Each x_j is a nonconstant solution of the gradient flow equation $\partial_s x_j = -\nabla f(x_j)$.

(ii) For each $1 \leq j \leq m-1$ there exists a solution $y_j \in C^\infty(\mathbb{R}, \operatorname{crit}(f))$ of the gradient flow equation $\partial_s y_j = -\nabla h(y_j)$ such that $\lim_{s\to\infty} x_j(s) = y_j(0)$ and $\lim_{s\to-\infty} x_{j+1}(s) = y_j(t_j)$.

(iii) There exist points $p^- \in W_h^u(x^-) \subset \operatorname{crit}(f)$ and $p^+ \in W_h^s(x^+) \subset \operatorname{crit}(f)$ such that $\lim_{s\to-\infty} x_1(s) = p^-$ and $\lim_{s\to\infty} x_m(s) = p^+$.

A *flow line with $m = 0$ cascades* simply is an ordinary flow line of $-\nabla h$ on crit(f) from x^- to x^+.

Denote by $\mathcal{M}_m(x^-, x^+)$ the set of flow lines from x^- to x^+ with m cascades (modulo the action of the group $\mathbb{R}^m$ by time-shifts on tuples $(x_1, \ldots, x_m)$) . We call

$$\mathcal{M}(x^-, x^+) := \bigcup_{m \in \mathbb{N}_0} \mathcal{M}_m(x^-, x^+)$$

the set of *flow lines with cascades from x^- to x^+*. In analogy to usual Morse theory (where the Morse function is required to have only isolated non-degenerate critical points), a sequence of broken flow lines with cascades may converge to a limit configuration which is a connected chain of such flow lines with cascades. This limiting behaviour is captured in the following definition.

Definition 8.2 Let $x^-, x^+ \in \mathrm{crit}(h)$. A *broken flow line with cascades from x^- to x^+* is a tuple $\mathtt{v} = (v_1, \ldots, v_\ell)$ where each v_j, $j = 1, \ldots, \ell$, consists of a flow line with cascades from $x^{(j-1)}$ to $x^{(j)} \in \mathrm{crit}(h)$ such that $x^{(0)} = x^-$ and $x^{(\ell)} = x^+$.

Theorem 8.3 *Let $x^-, x^+ \in \mathrm{crit}(h)$. Under suitable transversality assumptions (as specified in [14, Appendix C]) the set $\mathcal{M}(x^-, x^+)$ is a smooth manifold with boundary of dimension* $\dim \mathcal{M}(x^-, x^+) = \mathrm{Ind}(x^-) - \mathrm{Ind}(x^+) - 1$. *It is compact up to convergence to broken flow lines with cascades.*

Proof: For a proof we refer to [14, Theorems C.10, C.11]. □

We denote by $CM_*(M, f, h)$ the chain complex generated (as a $\mathbb{Z}$-module) by the critical points of h and graded by the index Ind. Thanks to Theorem 8.3 we may define a boundary operator $\partial_k : CM_k(M, f, h) \to CM_{k-1}(M, f, h)$ by linear extension of

$$\partial_k x := \sum_{\mathrm{Ind}(x')=k-1} n(x, x') x'$$

for $x \in \mathrm{crit}(h)$ with $\mathrm{Ind}(x) = k$. Here $n(x, x')$ denotes the (oriented) count of elements in the zero dimensional moduli space $\mathcal{M}(x, x')$. As was shown in [14] the maps ∂_k give rise to a boundary operator ∂_* satisfying $\partial_*^2 = 0$. We define the *Morse-Bott homology* $HM_*(M, f, h)$ of (M, f, h) by

$$HM_k(M, f, h) := \frac{\ker \partial_k}{\mathrm{im}\, \partial_{k+1}} \qquad (k \in \mathbb{N}_0).$$

8.2 Yang-Mills Morse complex

Let $a \geq 0$ be a regular value of $\mathcal{YM} : \mathcal{A}(P) \to \mathbb{R}$, $\mathcal{V}$ be a regular perturbation, and $h : \mathcal{P}(a) \to \mathbb{R}$ be a smooth Morse function. Define

$$\mathcal{P}(a) := \{A \in \mathcal{A}(P) \,|\, d_A^* F_A = 0 \text{ and } \mathcal{YM}^{\mathcal{V}}(A) \leq a\}$$

to be the set of Yang-Mills connections of energy at most a. We denote by

$$CM_*^a(\mathcal{A}(P), \mathcal{V}, h)$$

be the complex generated (as a $\mathbb{Z}$-module) by the set $\mathrm{crit}(h) \subseteq \mathcal{P}(a)$ of critical points of h. We define the Morse boundary operator

$$\partial_k : CM_k^a(\mathcal{A}(P), \mathcal{V}, h) \to CM_{k-1}^a(\mathcal{A}(P), \mathcal{V}, h)$$

for $k \in \mathbb{N}_0$ as the linear extension of the map

$$\partial_k x := \sum_{\substack{x' \in \mathcal{P}(a) \\ \mathrm{Ind}(x') = k-1}} n(x, x') x',$$

where $x \in \mathcal{P}(a)$ is a critical point of index $\mathrm{Ind}(x) = k$. The numbers $n(x, x')$ are given by counting oriented flow lines with cascades between x and x'.

Theorem 8.4 (Morse homology) *For any Morse function $h : \mathcal{P}(a) \to \mathbb{R}$ and generic perturbation $\mathcal{V} \in Y$, the map ∂_* satisfies $\partial_k \circ \partial_{k+1} = 0$ for all $k \in \mathbb{N}_0$ and thus there exist well-defined homology groups*

$$HM_k^a(\mathcal{A}(P)) = \frac{\ker \partial_k}{\mathrm{im}\, \partial_{k+1}}.$$

The homology $HM_*^a(\mathcal{A}(P))$ is called Yang-Mills Morse homology. *It is independent of the choice of perturbation $\mathcal{V}$ and Morse function h.*

Proof: For critical points $x^\pm \in \mathrm{crit}(h)$ let $\mathcal{M}(x^-, x^+)$ denote the moduli space of flow lines with cascades from x^- to x^+. From compactness of the moduli space $\mathcal{M}(\mathcal{C}^-, \mathcal{C}^+)$ (cf. Theorem 6.2) it follows that $\mathcal{M}(x^-, x^+)$ is compact up to convergence to broken flow lines with cascades. The proof that $\mathcal{M}(x^-, x^+)$ is a smooth manifold for regular perturbation $\mathcal{V}$ follows the standard routine by writing $\mathcal{M}(x^-, x^+)$ as the zero set of a Fredholm section $\hat{\mathcal{F}}$ of a suitable Banach space bundle, and then applying the implicit function theorem. For this we remark that $\mathcal{M}(x^-, x^+)$ can be written as a subset of products of moduli spaces $\mathcal{M}(\mathcal{C}_i^-, \mathcal{C}_j^+)$ for suitable pairs $(\mathcal{C}_i^-, \mathcal{C}_j^+)$ of critical manifolds (cf. [14, Theorem C.13]). Hence the Fredholm theory for $\mathcal{M}(x^-,x^+)$ reduces to that for the moduli spaces $\mathcal{M}(\mathcal{C}_i^-,\mathcal{C}_j^+)$ as discussed in Section 5. Surjectivity of the horizontal differential $d\hat{\mathcal{F}}$ is then achieved by the same perturbation arguments as in Theorem 7.1 and does not require any new arguments. From homotopy arguments standardly used in Floer theory (cf. [13, 24]) it follows that the Yang-Mills Morse homology groups do not depend on the regular perturbation $\mathcal{V}$ or on the Morse function h. $\Box$

A Properties of the perturbations

We introduce the following operators (for $A \in \mathcal{A}(P)$ and $\alpha \in \Omega^1(\Sigma, \mathrm{ad}(P))$).

$$\begin{aligned}
&L_{A,\alpha} : \Omega^0(\Sigma,\mathrm{ad}(P)) \to \Omega^0(\Sigma,\mathrm{ad}(P)), \quad L_{A,\alpha}\lambda := \Delta_A\lambda + *[*\alpha\wedge d_A\lambda],\\
&R_{A,\alpha} := L_{A,\alpha}^{-1} : \Omega^0(\Sigma,\mathrm{ad}(P)) \to \Omega^0(\Sigma,\mathrm{ad}(P)),\\
&M_\alpha : \Omega^1(\Sigma,\mathrm{ad}(P)) \to \Omega^0(\Sigma,\mathrm{ad}(P)), \quad M_\alpha\xi := *[\alpha\wedge *\xi],\\
&T_{A,\alpha} := R_{A,\alpha}\circ M_\alpha : \Omega^1(\Sigma,\mathrm{ad}(P)) \to \Omega^0(\Sigma,\mathrm{ad}(P)).
\end{aligned}$$

Proposition A.1 *Fix $A \in \mathcal{A}(P)$ smooth and $\alpha \in L^2(\Sigma, T^*\Sigma\otimes \mathrm{ad}(P))$. Then the operator*

$$L_{A,\alpha} : L^2(\Sigma,\mathrm{ad}(P)) \to L^2(\Sigma,\mathrm{ad}(P))$$

is a densely defined self-adjoint operator with domain $W^{2,2}(\Sigma,\mathrm{ad}(P))$. Its inverse $R_{A,\alpha}$ is a bounded operator

$$R_{A,\alpha} : L^2(\Sigma,\mathrm{ad}(P)) \to W^{2,2}(\Sigma,\mathrm{ad}(P)).$$

Furthermore, there exists a constant $c_1(A)$ such that for every α with $\|\alpha\|_{L^2(\Sigma)} < c_1(A)$ the estimate

$$\|R_{A,\alpha}\varphi\|_{W^{2,2}(\Sigma)} \leq c_2(A)\|\varphi\|_{L^2(\Sigma)}$$

holds for a constant $c_2(A)$ and all $\varphi \in L^2(\Sigma,\mathrm{ad}(P))$.

Proof: Symmetry of the operator $L_{A,\alpha}$ follows from an easy computation. As the Laplace operator Δ_A is self-adjoint with domain $W^{2,2}(\Sigma)$ the same holds true by the Kato-Rellich theorem for the perturbation $L_{A,\alpha}$ of Δ_A. That $R_{A,\alpha} : L^2(\Sigma,\mathrm{ad}(P)) \to W^{2,2}(\Sigma,\mathrm{ad}(P))$ is bounded follows from elliptic regularity. The assumption that A is a regular connection implies that

$\Delta_A = L_{A,0}$ is injective and therefore (by self-adjointness) bijective. Thus the bounded inverse theorem yields a constant $c(A)$ such that

$$\|R_{A,0}\varphi\|_{W^{2,2}(\Sigma)} \leq c(A)\|\varphi\|_{L^2(\Sigma)}$$

holds for all $\varphi \in L^2(\Sigma, \mathrm{ad}(P))$. It follows from this that for all $\lambda \in W^{2,2}(\Sigma)$

$$\begin{aligned}\|\lambda\|_{W^{2,2}(\Sigma)} \leq c(A)\|L_{A,0}\lambda\|_{L^2(\Sigma)} &\leq c(A)\big(\|L_{A,\alpha}\lambda\|_{L^2(\Sigma)} + \|[*\alpha \wedge d_A\lambda]\|_{L^2(\Sigma)}\big)\\ &\leq c'(A)\big(\|L_{A,\alpha}\lambda\|_{L^2(\Sigma)} + \|\alpha\|_{L^2(\Sigma)}\|\lambda\|_{W^{2,2}(\Sigma)}\big).\end{aligned}$$

As bijectivity is preserved under small perturbations (with respect to the operator norm) we can put $\lambda = R_{A,\alpha}\varphi$ for some $\varphi \in L^2(\Sigma, \mathrm{ad}(P))$. Then with $\|\alpha\|_{L^2(\Sigma)} < c_1(A)$ sufficiently small it follows that

$$\|R_{A,\alpha}\varphi\|_{W^{2,2}(\Sigma)} \leq c_2(A)\|\varphi\|_{L^2(\Sigma)},$$

as claimed. □

Proposition A.2 *Assume $A_0 \in \mathcal{A}(P)$ smooth and $\alpha \in L^2(\Sigma, \mathrm{ad}(P))$. Then the operator*

$$T^*_{A_0,\alpha} : L^2(\Sigma, \mathrm{ad}(P)) \to L^2(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))$$

is bounded with norm $\|T^*_{A_0,\alpha}\| \leq c(A_0)$.

Proof: By definition, $T^*_{A_0,\alpha} = M^*_\alpha \circ R_{A,\alpha}$ with $M^*_\alpha : \varphi \mapsto [\alpha, \varphi]$. Let $\xi \in L^2(\Sigma, \mathrm{ad}(P))$ be given. Then it follows from Proposition A.1 and the Sobolev embedding $W^{2,2}(\Sigma) \hookrightarrow C^0(\Sigma)$ that

$$\begin{aligned}\|T^*_{A_0,\alpha}\xi\|_{L^2(\Sigma)} = \|[\alpha, R_{A,\alpha}\xi]\|_{L^2(\Sigma)} &\leq c\|\alpha\|_{L^2(\Sigma)}\|R_{A,\alpha}\xi\|_{C^0(\Sigma)}\\ &\leq c\|\alpha\|_{L^2(\Sigma)}\|\xi\|_{L^2(\Sigma)},\end{aligned}$$

and thus the claim follows. □

Proposition A.3 *The map $\mathcal{V}_\ell : \mathcal{A}(P) \to \mathbb{R}$ has the following properties.*

(i) (We denote $A_0 := A_i$, $\eta := \eta_{ij}$, and $\rho := \rho_k$.) Its differential and L^2-gradient are given by

$$\begin{aligned}d\mathcal{V}_\ell(A)\xi &= 2\rho'(\|\alpha\|^2_{L^2})\langle\alpha, \hat{\xi}\rangle\langle\alpha, \eta\rangle + \rho(\|\alpha\|^2_{L^2})\langle\hat{\xi} - d_{g^*A}T_{A_0,\alpha}\hat{\xi}, \eta\rangle,\\ g^{-1}\nabla\mathcal{V}(A)g &= 2\rho'(\|\alpha\|^2_{L^2})\langle\alpha, \eta\rangle\alpha + \rho(\|\alpha\|^2_{L^2})(\eta + T^*_{A_0,\alpha}(*[\alpha \wedge *\eta]),\end{aligned}$$

with $\hat{\xi} := g^{-1}\xi g$. Here we assume that $\xi \in \Omega^1(\Sigma, \mathrm{ad}(P))$ satisfies $d^*_A\xi = 0$.

(ii) *We have that*

$$d_A \nabla \mathcal{V}(A) = \rho(\|\alpha\|_{L^2}^2)\big(d_A T^*_{A_0,\alpha}(*[\alpha \wedge *\eta]) + d_A\eta\big) + 2\rho'(\|\alpha\|_{L^2}^2)\langle \alpha, \eta\rangle d_A\alpha$$

(iii) *Let* $\beta \in \Omega^1(\Sigma, \mathrm{ad}(P))$ *such that* $d_A^*\beta = 0$ *and set* $\gamma := \beta - d_{A_0} T_{A_0,\alpha}\beta$. *Then the Hessian of* $\mathcal{V}(A)$ *is the map* $H_A\mathcal{V} : \Omega^1(\Sigma, \mathrm{ad}(P)) \to \Omega^1(\Sigma, \mathrm{ad}(P))$ *given by the formula*

$$\begin{aligned}
&g^{-1}(H_A\mathcal{V}\beta)g + [g^{-1}\nabla\mathcal{V}(A)g, \lambda] = \\
&\qquad \rho(\|\alpha\|_{L^2}^2)\left(S^*_{A_0,\alpha,\gamma}(*[\alpha \wedge *\eta]) + T^*_{A_0,\alpha}(*[\gamma \wedge *\eta])\right) \\
&\qquad +2\rho'(\|\alpha\|_{L^2}^2)\left(\langle\alpha,\gamma\rangle(\eta + T^*_{A_0,\alpha}(*[\alpha \wedge *\eta]) + \langle \eta,\gamma\rangle\alpha + \langle\alpha,\eta\rangle\gamma\right) \\
&\qquad +4\rho''(\|\alpha\|_{L^2}^2)\langle\alpha,\gamma\rangle\langle\alpha,\eta\rangle\alpha.
\end{aligned}$$

Here we denote

$$S_{A_0,\alpha,\gamma} := R_{A_0,\alpha} \circ M_\gamma \circ (\mathbf{1} - d_A \circ R_{A_0,\alpha}) : \Omega^1(\Sigma, \mathrm{ad}(P)) \to \Omega^0(\Sigma, \mathrm{ad}(P)).$$

Proof:

(i) Let $A(t) = A + t\xi$. Assume $A(t)$, $\alpha(t)$ and $g(t)$ satisfy condition (5) for $t \in (-\varepsilon, \varepsilon)$ with $\varepsilon > 0$ sufficiently small. Set $\dot\alpha := \frac{d}{dt}\big|_{t=0}\alpha(t)$ and $\lambda := g^{-1}(0)\,\frac{d}{dt}\big|_{t=0} g(t)$. Differentiating the equation $d^*_{A_0}\alpha = 0$ at $t = 0$ yields

$$\begin{aligned}
0 &= d^*_{A_0}(g^{-1}\xi g + d_{g^*A}\lambda) \\
&= g^{-1}(d_{(g^{-1})^*A_0}\xi)g + d^*_{A_0}d_{A_0}\lambda + d^*_{A_0}[\alpha \wedge \lambda] \\
&= g^{-1}(d_{A-g\alpha g^{-1}}\xi)g + \Delta_{A_0}\lambda + d^*_{A_0}[\alpha \wedge \lambda] \\
&= -g^{-1} * [g\alpha g^{-1} \wedge *\xi]g + \Delta_{A_0}\lambda + d^*_{A_0}[\alpha \wedge \lambda] \\
&= - * [\alpha \wedge *g^{-1}\xi g] + \Delta_{A_0}\lambda + d^*_{A_0}[\alpha \wedge \lambda] \\
&= -M_\alpha\hat\xi + L_{A_0,\alpha}\lambda.
\end{aligned}$$

Hence $\lambda = T_{A_0,\alpha}\hat\xi$ by definition of $T_{A_0,\alpha}$, and $\dot\alpha = \hat\xi + d_{g^*A}T_{A_0,\alpha}\hat\xi$. From this we obtain

$$\begin{aligned}
d\mathcal{V}(A)\xi &= \frac{d}{dt}\bigg|_{t=0} \rho(\|\alpha(t)\|^2)\langle\alpha(t), \eta\rangle \\
&= 2\rho'(\|\alpha\|^2)\langle\alpha, \dot\alpha\rangle\langle\alpha,\eta\rangle + \rho(\|\alpha\|^2)\langle\dot\alpha, \eta\rangle \\
&= 2\rho'(\|\alpha\|^2)\langle\alpha, \hat\xi + d_{g^*A}T_{A_0,\alpha}\hat\xi\rangle\langle\alpha,\eta\rangle + \rho(\|\alpha\|^2)\langle\hat\xi + d_{g^*A}T_{A_0,\alpha}\hat\xi, \eta\rangle \\
&= 2\rho'(\|\alpha\|^2)\langle\alpha, \hat\xi\rangle\langle\alpha,\eta\rangle + \rho(\|\alpha\|^2)\langle\hat\xi - d_{g^*A}T_{A_0,\alpha}\hat\xi, \eta\rangle.
\end{aligned}$$

In the last line we used that $d_{g^*A}^*\alpha = 0$. The formula for $\nabla\mathcal{V}(A)$ follows from this by taking adjoints and using that $d_{g^*A}^*\eta = d_{A_0}^*\eta - *[\alpha\wedge *\eta] = - * [\alpha \wedge *\eta]$.

(ii) This follows by direct calculation.

(iii) The formula follows from differentiating the expression for $g^{-1}\nabla\mathcal{V}(A)g$ in (i), and using the definition of the Hessian,

$$H_A\mathcal{V}\beta = \frac{d}{dt}\Big|_{t=0}\nabla\mathcal{V}(A+t\beta).$$

The operator $S_{A_0,\alpha,\gamma}$ arises from differentiating

$$\frac{d}{dt}\Big|_{t=0} T_{A_0,\alpha} = \frac{d}{dt}\Big|_{t=0} L_{A_0,\alpha}^{-1}\circ M_\alpha = -L_{A_0,\alpha}^{-1}\dot{L}_{A_0,\alpha}L_{A_0,\alpha}^{-1} + R_{A_0,\alpha}\circ\dot{M}_\alpha,$$

where $\dot{L}_{A_0,\alpha} = *[*\gamma\wedge d_{A_0}\cdot\,]$ and $\dot{M}_\alpha = M_\gamma$.

$\square$

Proposition A.4 *Assume $A_0, A \in \mathcal{A}(P)$ smooth. Then $\nabla\mathcal{V}(A)$ satisfies the estimate*

$$\|\nabla\mathcal{V}(A)\|_{L^2(\Sigma)} \le c(A_0)$$

for a constant $c(A_0)$ independent of A.

Proof: From Proposition A.2 and the formula for $\nabla\mathcal{V}(A)$ stated in Proposition A.3 it follows that

$$\|\nabla\mathcal{V}(A)\|_{L^2(\Sigma)} \le c(1+\|\alpha\|_{L^2(\Sigma)} + \|T_{A_0,\alpha}^*(*[\alpha\wedge *\eta]\|_{L^2(\Sigma)}) \le c(A_0).$$

$\square$

Proposition A.5 *For every $\varepsilon > 0$ there exists a constant $\delta > 0$ with the following significance. Assume the perturbation $\mathcal{V}$ satisfies $\|\mathcal{V}\| < \delta$. Then for all $A \in \mathcal{A}(P)$ there holds the estimate*

$$\|\nabla\mathcal{V}(A)\|_{L^2(\Sigma)} < \varepsilon.$$

Proof: From Proposition A.2 and the expression for $\nabla\mathcal{V}(A)$ given in Proposition A.3 it follows the estimate $\|\nabla\mathcal{V}(A)\|_{L^2(\Sigma)} < c(A_0)\|\eta\|_{L^2(\Sigma)}$ for a constant $c(A_0)$ and all $A \in \mathcal{A}(P)$. Now choose the terms $\eta(A_0)$ accordingly. $\square$

Proposition A.6 *For every $p > 2$ there exists a constant $c(p)$ such the estimate*

$$\|uv\|_{W^{-1,p}(\Sigma)} \leq c(p)\|u\|_{L^2(\Sigma)}\|v\|_{L^p(\Sigma)} \tag{54}$$

is satisfied for all functions $u \in L^2(\Sigma)$ and $v \in L^p(\Sigma)$.

Proof: Let $q < 2$ denote the dual Sobolev exponent of p. Let $r := \frac{2p}{2+p} < 2$ and $s := \frac{2q}{2-q} > 2$, i.e. $\frac{1}{r} + \frac{1}{s} = 1$. Then the Sobolev embedding $W^{1,q}(\Sigma) \hookrightarrow L^{\frac{2q}{2-q}}(\Sigma)$ implies the dual embedding $L^r(\Sigma) \hookrightarrow W^{-1,p}(\Sigma)$. Hence for a constant $c(p)$ it follows that

$$\|uv\|_{W^{-1,p}(\Sigma)} \leq c(p)\|uv\|_{L^r(\Sigma)},$$

and Hölder's inequality (with exponents $\ell = \frac{2}{r} > 1$ and $\ell' = \frac{2}{2-r} > 1$) then implies that

$$\|uv\|^r_{L^r(\Sigma)} \leq \Big(\int_\Sigma |u|^2\Big)^{\frac{r}{2}}\Big(\int_\Sigma |v|^{r\ell'}\Big)^{\frac{1}{\ell'}} = \|u\|^r_{L^2(\Sigma)}\|v\|^r_{L^p},$$

as claimed. $\square$

Proposition A.7 *Assume $A_0 \in \mathcal{A}(P)$ smooth and let $p > 2$. Then there exists constants $c(A_0, p)$ and $\delta(A_0, p)$ such that the estimates*

$$\|\alpha\|_{W^{1,p}(\Sigma)} \leq c(A_0,p)\big(1 + \|F_A\|_{L^p(\Sigma)}\big), \tag{55}$$

$$\|\nabla\mathcal{V}(A)\|_{L^p(\Sigma)} \leq c(A_0,p)\big(1 + \|F_A\|_{L^p(\Sigma)}\big), \tag{56}$$

$$\|d_A\nabla\mathcal{V}(A)\|_{L^p(\Sigma)} \leq c(A_0,p)\big(1 + \|F_A\|_{L^p(\Sigma)} + \|\alpha\|^2_{L^{2p}(\Sigma)}\big) \tag{57}$$

are satisfied for all $A \in \mathcal{A}(P)$ such that $\|\alpha\|_{L^2(\Sigma)} < \delta(A_0, p)$.

Proof: With α satisfying $d_{A_0}\alpha = F_A - F_{A_0} - \frac{1}{2}[\alpha\wedge\alpha]$ and hence

$$\Delta_{A_0}\alpha = d^*_{A_0}(F_A - F_{A_0}) - \frac{1}{2}d^*_{A_0}[\alpha\wedge\alpha] \tag{58}$$

we use elliptic regularity of the operator $\Delta_{A_0} : W^{1,p}(\Sigma) \to W^{-1,p}(\Sigma)$ to estimate

$$\begin{aligned}
&\|\alpha\|_{W^{1,p}(\Sigma)} \qquad (59)\\
&\leq c(A_0,p)\big(\|d_{A_0}^* F_{A_0}\|_{W^{-1,p}(\Sigma)} + \|d_{A_0}^* F_A\|_{W^{-1,p}(\Sigma)} \\
&\quad + \|d_{A_0}^*[\alpha\wedge\alpha]\|_{W^{-1,p}(\Sigma)} + \|\alpha\|_{L^p(\Sigma)}\big)\\
&\leq c(A_0,p)\big(1 + \|F_A\|_{L^p(\Sigma)} + \|\{\nabla_{A_0}\alpha,\alpha\}\|_{W^{-1,p}(\Sigma)} + \|\alpha\|_{L^p(\Sigma)}\big)\\
&\leq c(A_0,p)\big(1 + \|F_A\|_{L^p(\Sigma)} + \|\alpha\|_{L^2(\Sigma)}\|\nabla_{A_0}\alpha\|_{L^p(\Sigma)} + \|\alpha\|_{L^p(\Sigma)}\big).
\end{aligned}$$

In the last step we applied Proposition A.6. From the Sobolev embedding $W^{1,2}(\Sigma) \hookrightarrow L^p(\Sigma)$ and the second inequality of the previous estimate we furthermore obtain

$$\begin{aligned}
\|\alpha\|_{L^p(\Sigma)} &\leq c\|\alpha\|_{W^{1,2}(\Sigma)}\\
&\leq c(A_0,2)\big(1 + \|F_A\|_{L^2(\Sigma)} + \|\{\nabla_{A_0}\alpha,\alpha\}\|_{W^{-1,2}(\Sigma)} + \|\alpha\|_{L^2(\Sigma)}\big)\\
&\leq c(A_0,p)\big(1 + \|F_A\|_{L^2(\Sigma)} + \|\{\nabla_{A_0}\alpha,\alpha\}\|_{W^{-1,p}(\Sigma)}\big)\\
&\leq c(A_0,p)\big(1 + \|F_A\|_{L^2(\Sigma)} + \|\alpha\|_{L^2(\Sigma)}\|\nabla_{A_0}\alpha\|_{L^p(\Sigma)}\big).
\end{aligned}$$

Now fix $\delta(A_0,p) > 0$ sufficiently small such that $\|\alpha\|_{L^2(\Sigma)}\|\nabla_{A_0}\alpha\|_{L^p(\Sigma)} \leq \|\alpha\|_{W^{1,p}(\Sigma)}$ holds for all α with $\|\alpha\|_{L^2(\Sigma)} < \delta(A_0,p)$, to conclude estimate (55). To get (56) we note that $\|\nabla\mathcal{V}(A)\|_{L^p(\Sigma)}$ is controlled by $\|\alpha\|_{L^p(\Sigma)}$ as follows from Proposition A.3. To prove (57) let us denote $\beta := *[\alpha\wedge *\eta]$ and $\gamma := R_{A_0,\alpha}\beta$. From the expression for $d_A\nabla\mathcal{V}(A)$ as in Proposition A.3 we see that it suffices to estimate the terms $d_A\alpha$ and

$$d_A T_{A_0,\alpha}^*\beta = d_A[\alpha\wedge R_{A_0,\alpha}\beta] = [d_A\alpha\wedge\gamma] - [\alpha\wedge d_A\gamma]. \qquad (60)$$

The required bound for $d_A\alpha = d_{A_0}\alpha + [\alpha\wedge\alpha]$ follows by an estimate similar to (59) where now an additional term $\|d_{A_0}^*[\alpha\wedge\alpha]\|_{W^{-1,p}(\Sigma)}$ appears which is controlled by $\|\alpha\|_{L^{2p}(\Sigma)}^2$. From the Sobolev embedding $W^{2,2}(\Sigma) \hookrightarrow C^0(\Sigma)$ and Proposition A.1 we furthermore obtain

$$\|\gamma\|_{C^0(\Sigma)} \leq c\|\gamma\|_{W^{2,2}(\Sigma)} \leq c(A_0)\|\beta\|_{L^2(\Sigma)} \leq c(A_0)\|\alpha\|_{L^2(\Sigma)}. \qquad (61)$$

With $\|\alpha\|_{L^2(\Sigma)} \leq \delta$ for some constant δ it remains to bound the term $d_A\alpha$, which has been been done before. Finally consider the term $[\alpha\wedge d_A\gamma] = [\alpha\wedge d_{A_0}\gamma] + [\alpha\wedge[\alpha\wedge\gamma]]$ in (60). It is bounded by $\|\alpha\|_{L^{2p}(\Sigma)}^2$ as follows from the Sobolev embedding $W^{2,2}(\Sigma) \hookrightarrow W^{1,p}(\Sigma)$ and an estimate similar to (61). $\square$

Proposition A.8 *For every $p > 1$ there exists a constant $c = c(A_0, p)$ such that the estimate*

$$\|\nabla_{A_0} d_A \nabla \mathcal{V}(A)\|_{L^p(I\times\Sigma)} \leq c\big(1 + \|F_A\|^4_{L^{4p}(I\times\Sigma)} + \|\alpha\|^4_{L^{4p}(I\times\Sigma)}\big)$$

holds for all $A \in \mathcal{A}(P)$.

Proof: Denote $\gamma := R_{A_0,\alpha} * [\alpha \wedge *\eta]$. From the expression for $d_A\nabla\mathcal{V}(A)$ of Proposition A.3 we see that it suffices to estimate the expression

$$\int_{I\times\Sigma} |\nabla_{A_0} d_A\alpha|^p + |[\nabla_{A_0} d_A\alpha \wedge \gamma]|^p + |[d_A\alpha \wedge \nabla_{A_0}\gamma]|^p + |[\nabla_{A_0}\alpha \wedge d_A\gamma]|^p$$
$$+|[\alpha \wedge \nabla_{A_0} d_A\gamma]|^p.$$

After applying Hölder's inequality the terms to estimate are (apart from some lower order terms)

$$\int_{I\times\Sigma} |\nabla_{A_0} d_A\alpha|^{2p}, \quad \int_{I\times\Sigma} |\nabla_{A_0}\gamma|^{2p}, \quad \int_{I\times\Sigma} |\nabla_{A_0}\alpha|^{2p}, \quad \int_{I\times\Sigma} |\nabla_{A_0} d_A\gamma|^{2p}. \quad (62)$$

After rewritig

$$d_A\alpha = F_A - F_{A_0} - \frac{1}{2}[\alpha \wedge \alpha]$$

the estimate of the first term in (62) reduces to that of $\int_{I\times\Sigma} |F_A|^{2p} + |\nabla_{A_0}\alpha|^{4p}$. Proposition A.7 then gives the bound

$$\int_{I\times\Sigma} |\nabla_{A_0}\alpha|^{4p} \leq c(A_0, p) \int_{I\times\Sigma} 1 + |F_A|^{4p} + |\alpha|^{4p},$$

as required. The estimate for the third term in (62) follows by the same argument. The estimate

$$\int_\Sigma |\gamma|^{2p} \leq c(A_0) \int_\Sigma |\alpha|^{2p}$$

with uniform constant $c(A_0)$ gives the required bound for the second and fourth term in (62). □

B Perturbed Yang-Mills gradient flow

Let $\mathcal{V} : \mathcal{A}(P) \to \mathbb{R}$ be a perturbation as above. We call the equation

$$\partial_s A + d_A^* F_A + \nabla \mathcal{V}(A) = 0 \tag{63}$$

the *perturbed Yang-Mills gradient flow.* In the following we fix $I = [a, b]$ an interval and $I' = [a', b]$, where $a < a' < b$, a subinterval.

Proposition B.1 *There exists a constant $C = C(\mathcal{V}) \geq 0$ such that*

$$\|F_{A(s)}\|_{L^2(\Sigma)} \leq \|F_{A(a)}\|_{L^2(\Sigma)} + C$$

is satisfied for all $s \in I$, where A is a solution of (63) on $I \times \Sigma$.

Proof: The energy $\mathcal{YM}^{\mathcal{V}}(A) = \frac{1}{2}\int_\Sigma |F_A|^2 + \mathcal{V}(A)$ is monotone decreasing along flow lines, hence

$$\begin{aligned}
\frac{1}{2}\|F_{A(s)}\|^2_{L^2(\Sigma)} &\leq \mathcal{YM}^{\mathcal{V}}(A(s)) + |\mathcal{V}(A(s))| \\
&\leq \mathcal{YM}^{\mathcal{V}}(A(a)) + \sup_{A \in \mathcal{A}(P)} |\mathcal{V}(A)| \\
&= \frac{1}{2}\|F_{A(a)}\|^2_{L^2(\Sigma)} + |\mathcal{V}(A(a))| + \sup_{A \in \mathcal{A}(P)} |\mathcal{V}(A)| \\
&\leq \frac{1}{2}\|F_{A(a)}\|^2_{L^2(\Sigma)} + \frac{1}{2}C^2,
\end{aligned}$$

where $C := 2\sup_{A \in \mathcal{A}(P)} |\mathcal{V}(A)|^{\frac{1}{2}}$. The claim follows. $\square$

Let $\Delta_\Sigma = - * d * d$ denote the Laplace-Beltrami operator on Σ and let $L_\Sigma := \partial_s + \Delta_\Sigma$ be the corresponding heat operator. For the following calculations we need the Bochner-Weitzenböck formula

$$\Delta_A \alpha = \nabla_A^* \nabla_A \alpha + \{F_A, \alpha\} + \{R_\Sigma, \alpha\}, \tag{64}$$

relating the covariant Hodge Laplacian Δ_A and the connection Laplacian $\nabla_A^* \nabla_A$ on forms in $\Omega^k(\Sigma, \mathrm{ad}(P))$, cf. [21]. Here $\{F_A, \alpha\}$ etc. denote bilinear expressions in F_A and α, and R_Σ is a term involving the curvature operator of Σ. For a form $\alpha \in \Omega^k(\Sigma, \mathrm{ad}(P))$ there holds the identity

$$\Delta_\Sigma \frac{1}{2}|\alpha|^2 = -|\nabla_A \alpha|^2 + \langle \nabla_A^* \nabla_A \alpha, \alpha \rangle. \tag{65}$$

We shall make use of the commutator identity

$$[\nabla_A, \nabla_A^* \nabla_A]\alpha = \{\alpha, \nabla_A \alpha\}, \tag{66}$$

cf. [10, p. 17].

Proposition B.2 *Assume that A solves* (63) *on $I \times \Sigma$. Consider the function $u_{0,p} : I \times \Sigma \to \mathbb{R}$ defined (for $p \geq 2$) by*

$$u_{0,p}(s,z) := \frac{1}{p} |*F_{A(s)}(z)|^p.$$

Denote $u_0 := u_{0,2}$. Then the following holds.

$$\begin{aligned} L_\Sigma u_0 &= -|d_A * F_A|^2 - \langle *F_A, *d_A \nabla \mathcal{V}(A) \rangle, \\ L_\Sigma u_{0,p} &= |*F_A|^{p-2} \left(-|d_A * F_A|^2 - \langle *F_A, *d_A \nabla \mathcal{V}(A) \rangle \right), \end{aligned}$$

Proof: We calculate, using (63),

$$\frac{d}{ds} \frac{1}{2} \langle *F_A, *F_A \rangle = \langle *F_A, *d_A \dot{A} \rangle = \langle *F_A, -*\Delta_A F_A - *d_A \nabla \mathcal{V}(A) \rangle.$$

From this it follows that

$$\begin{aligned} L_\Sigma u_0 &= \frac{1}{2} \big(\partial_s - *d*d \big) \langle *F_A, *F_A \rangle \\ &= -\langle *\Delta_A F_A + *d_A \nabla \mathcal{V}(A), *F_A \rangle - *d* \langle d_A * F_A, *F_A \rangle \\ &= -\langle *\Delta_A F_A + *d_A \nabla \mathcal{V}(A), *F_A \rangle - \langle *d_A * d_A * F_A, *F_A \rangle - |d_A * F_A|^2 \\ &= -\langle *d_A \nabla \mathcal{V}(A), *F_A \rangle - |d_A * F_A|^2. \end{aligned}$$

The formula for $u_{0,p}$ follows from that for u_0 and the chain rule

$$\big(\frac{1}{p} \langle *F_A, *F_A \rangle^{\frac{p}{2}} \big)' = \langle *F_A, *F_A \rangle^{\frac{p}{2}-1} \langle F_A, F_A' \rangle = |F_A|^{p-2} \langle F_A, F_A' \rangle.$$

□

Proposition B.3 *Assume that A solves* (63) *on $I \times \Sigma$. Consider the function $u_1 : I \times \Sigma \to \mathbb{R}$ defined by*

$$u_1(s,z) := \frac{1}{2} |\nabla_{A(s)} F_{A(s)}(z)|^2.$$

It satisfies

$$L_\Sigma u_1 = -|\nabla_A \nabla_A F_A|^2 + \langle \nabla_A F_A, \{\nabla_A F_A, F_A\} + \nabla_A \{R_\Sigma, F_A\} + \{\nabla \mathcal{V}(A), F_A\} - \nabla_A d_A \nabla \mathcal{V}(A) \rangle. \quad (67)$$

Proof: We calculate

$$\begin{aligned}
\frac{d}{ds}\nabla_A F_A &= \nabla_A d_A \dot{A} + \{\dot{A}, F_A\} \\
&= \nabla_A\big(-d_A d_A^* F_A - d_A \nabla\mathcal{V}(A)\big) + \{d_A^* F_A + \nabla\mathcal{V}(A), F_A\} \\
&= \nabla_A\big(-\nabla_A^*\nabla_A F_A + \{F_A, F_A\} + \{R_\Sigma, F_A\} - d_A\nabla\mathcal{V}(A)\big) \\
&\quad + \{d_A^* F_A + \nabla\mathcal{V}(A), F_A\} \\
&= -\nabla_A^*\nabla_A\nabla_A F_A + \{\nabla_A F_A, F_A\} + \nabla_A\{F_A, F_A\} + \nabla_A\{R_\Sigma, F_A\} \\
&\quad -\nabla_A d_A \nabla\mathcal{V}(A) + \{\nabla\mathcal{V}(A), F_A\}.
\end{aligned}$$

The third line is by (64), and the last line uses (66). Inserting this expression into (65) we obtain

$$L_\Sigma u_1 = -|\nabla_A\nabla_A F_A|^2 + \big\langle \nabla_A F_A, \{\nabla_A F_A, F_A\} + \nabla_A\{R_\Sigma, F_A\} + \{\nabla\mathcal{V}(A), F_A\} \\ -\nabla_A d_A \nabla\mathcal{V}(A)\big\rangle,$$

as claimed. □

The following estimate is a consequence of the Bochner-Weitzenböck formula (64).

Proposition B.4 *There exists a constant $c_1(P)$ such that for every C^1-connection A and a further constant $c_2(A,P)$ there holds*

$$\|\varphi\|^2_{W^{2,2}(\Sigma)} \leq c_1(P)\|\Delta_A\varphi\|^2_{L^2(\Sigma)} + c_2(\|A\|_{C^1}, P)\|\varphi\|^2_{L^2(\Sigma)}$$

for every $\varphi \in \Omega^k(\Sigma, \mathrm{ad}(P))$.

Proof: For a proof we refer to Struwe [29, Lemma 3.1]. □

The following lemma is an adaption of a result by Struwe [29, Lemma 3.3].

Lemma B.5 *Let $A \in \mathcal{A}^{1,2}(P)$ be a fixed connection and let $p > 1$. There exists a constant $c = c(p)$ such that for any section $\varphi \in \Omega^k(\Sigma, \mathrm{ad}(P))$ there holds*

$$\|\varphi\|^2_{L^p} \leq c\big(\|d_A\varphi\|^2_{L^2(\Sigma)} + \|d_A^*\varphi\|^2_{L^2(\Sigma)}\big) + c\langle\{F_A, \varphi\}, \varphi\rangle.$$

Proof: The proof of [29, Lemma 3.3] applies with minor modifications. The first one is the Sobolev embedding $W^{1,2}(\Sigma) \hookrightarrow L^p(\Sigma)$ for all $p < \infty$ (instead of only $W^{1,2} \hookrightarrow L^4$ in dimension 4). The second one is that at this point we do not further estimate the term $\langle\{F_A, \varphi\}, \varphi\rangle$. □

As a consequence of the lemma we obtain the following estimate for F_A.

Proposition B.6 *Let $p > 1$ and $I = [a, b]$. There exists a constant $c = c(I, p)$ such that if A is a solution of* (63) *on $I \times \Sigma$, then*

$$\int_I \|F_{A(s)}\|^2_{L^p(\Sigma)}\, ds \le c\big(1 + \mathcal{YM}^{\mathcal{V}}(A(a))\big).$$

Proof: Integrating the estimate of Lemma B.5 with $\varphi = F_A$ and using the Bianchi identity $d_A F_A = 0$ gives

$$\begin{aligned}
&\int_I \|F_A\|^2_{L^p(\Sigma)} \\
&\le c(p) \int_I \big(\|d_A^* F_A\|^2_{L^2(\Sigma)} + \|F_A\|^3_{L^3(\Sigma)}\big) \\
&\le c(p) \int_I \big(2\|d_A^* F_A + \nabla\mathcal{V}(A)\|^2_{L^2(\Sigma)} + 2\|\nabla\mathcal{V}(A)\|^2_{L^2(\Sigma)} + \|F_A\|^3_{L^3(\Sigma)}\big) \\
&\le 2c(p)\mathcal{YM}^{\mathcal{V}}(A(a)) + 2|I| \sup_{A\in\mathcal{A}(P)} \|\nabla\mathcal{V}(A)\|^2_{L^2(\Sigma)} + c(p)\|F_A\|^3_{L^3(I\times\Sigma)}.
\end{aligned}$$

The middle term is bounded by a universal constant C. It remains to estimate the last term. Using Hölder's inequality we obtain

$$\|F_A\|^3_{L^3(I\times\Sigma)} \le \int_I \|F_A\|_{L^2(\Sigma)}\|F_A\|^2_{L^4(\Sigma)} \le \|F_A\|_{L^\infty(I,L^2(\Sigma))}\|F_A\|^2_{L^2(I,L^4(\Sigma))}. \quad (68)$$

Now choose a locally finite cover of Σ with balls $B_\varepsilon(x_i)$ of radius $\varepsilon > 0$. We can then further estimate

$$\begin{aligned}
\|F_A\|^2_{L^2(I,L^4(\Sigma))} &\le \int_I \sum_i \Big(\int_{B_\varepsilon(x_i)} |F_A|^4\Big)^{\frac12} \\
&\le \int_I \sum_i \Big(\int_{B_\varepsilon(x_i)} |F_A|^8\Big)^{\frac14} \cdot \Big(\int_{B_\varepsilon(x_i)} 1\Big)^{\frac14} \\
&= \mathrm{vol}(B_\varepsilon)^{\frac14} \int_I \sum_i \Big(\int_{B_\varepsilon(x_i)} |F_A|^8\Big)^{\frac14} \\
&\le c\varepsilon^{\frac12}\|F_A\|^2_{L^2(I,L^8(\Sigma))}.
\end{aligned}$$

For $p \ge 8$ and $\varepsilon > 0$ and sufficiently small (compared to $\|F_A\|_{L^\infty(I,L^2(\Sigma))} \le \|F_{A(a)}\|_{L^2(\Sigma)} \le C$) the last term in (68) is absorbed by $\int_I \|F_A\|^2_{L^p(\Sigma)}$. The claim thus follows. $\square$

Lemma B.7 *Let $1 < p < 4$ and $I = [a, b]$. There exists a constant $c(I, p)$ such that if A is a solution of* (63) *on $I \times \Sigma$, then*

$$\|F_A\|^2_{L^p(I\times\Sigma)} \le c(I, p)\big(1 + \mathcal{YM}^{\mathcal{V}}(A(a))\big).$$

Proof: Hölder's inequality yields for $p < 4$ the estimate

$$\|F_A\|^p_{L^p(I\times\Sigma)} \leq \|F_A\|^{p-2}_{L^\infty(I,L^2(\Sigma))}\|F_A\|^2_{L^2(I,L^{\frac{4}{4-p}}(\Sigma))}.$$

Each of the last two factors is uniformly bounded. For the first one this follows from Proposition B.1, and for the second one it follows from Proposition B.6. $\square$

A similar kind of estimate also holds for $d_A^*F_A$.

Proposition B.8 *Let $p > 1$ and $I = [a, b]$. There exists a constant $c(I, p)$ such that if A is a solution of* (63) *on $I \times \Sigma$, then*

$$\int_I \|d^*_{A(s)}F_{A(s)}\|^2_{L^p(\Sigma)}\,ds \leq c(I,p)\int_I \|d_Ad_A^*F_A\|^2_{L^2(\Sigma)}. \tag{69}$$

Proof: Integrating the estimate of Lemma B.5 with $\varphi = d_A^*F_A$ and using that $d_A^*d_A^*F_A = 0$, we obtain

$$\int_I \|d^*_{A(s)}F_{A(s)}\|^2_{L^p(\Sigma)}\,ds \leq c\int_I \|d_Ad_A^*F_A\|^2_{L^2(\Sigma)} + c\int_I\int_\Sigma \langle\{F_A, d_A^*F_A\}, d_A^*F_A\rangle.$$

Proceeding as in the proof of Proposition B.6, the last term may be absorbed by $\int_I \|d^*_{A(s)}F_{A(s)}\|^2_{L^p(\Sigma)}\,ds$. The claim then follows. $\square$

Proposition B.9 *Let $I = [a, b]$. Suppose A is a solution of* (63) *on $I \times \Sigma$. Then the function $R(s) := \frac{1}{2}\|d^*_{A(s)}F_{A(s)}\|^2_{L^2(\Sigma)}$ satisfies the estimate*

$$\begin{aligned}\sup_{a\leq s\leq b} R(s) \leq R(a) &+ \int_I \|d_A\nabla\mathcal{V}(A)\|^2_{L^2(\Sigma)} \\ &+ \int_I\int_\Sigma \big|\langle d_A^*F_A, \{d_A^*F_A, F_A\}\rangle\big| + \big|\langle d_A^*F_A, \{\nabla\mathcal{V}(A), F_A\}\rangle\big|,\end{aligned}$$

where $\{\cdot\}$ indicates a certain bilinear expression with smooth (time-independent) coefficients.

Proof: From equation (63) it follows for every $a \leq s \leq b$ that

$$\begin{aligned}\frac{d}{ds}R(s) &= \langle d_A^*F_A, d_A^*d_A\dot{A} - *[\dot{A}\wedge *F_A]\rangle \\ &= -\langle d_A^*F_A, d_A^*d_Ad_A^*F_A + d_A^*d_A\nabla\mathcal{V}(A) - *[(d_A^*F_A + \nabla\mathcal{V}(A))\wedge *F_A]\rangle \\ &= -\|d_Ad_A^*F_A\|^2 - \langle d_Ad_A^*F_A, d_A\nabla\mathcal{V}(A)\rangle + \langle d_A^*F_A, *[(d_A^*F_A + \nabla\mathcal{V}(A))\wedge *F_A]\rangle \\ &\leq \|d_A\nabla\mathcal{V}(A)\|^2 + \big|\langle d_A^*F_A, \{d_A^*F_A, F_A\}\rangle\big| + \big|\langle d_A^*F_A, \{\nabla\mathcal{V}(A), F_A\}\rangle\big|.\end{aligned}$$

Integrating this inequality over the interval $[a,s] \subseteq I$ and taking the supremum over $a \le s \le b$ yields the result. $\square$

Lemma B.10 *Let $1 < p < 4$, $I = [a,b]$ and $I' = [a_1, b]$, where $a_1 \in (a,b)$. There exists a constant $c(I,I',p)$ such that if A is a solution of* (63) *on $I \times \Sigma$, then*

$$\|d_A^* F_A\|_{L^p(I'\times\Sigma)} \le c(I,I',p)\big(\mathcal{YM}^{\mathcal{V}}(A(a)) + \|F_{A(a)}\|_{L^2(\Sigma)}^2 + \|F_{A(a)}\|_{L^2(\Sigma)}^3\big).$$

Proof: By Fubini's theorem we can find $s_0 \in (a, a_1)$ such that

$$\begin{aligned}\|d_{A(s_0)}^* F_{A(s_0)}\|_{L^2(\Sigma)}^2 \le 2(a_1-a)^{-1}\int_a^{a_1} \|d_{A(s)}^* F_{A(s)}\|_{L^2(\Sigma)}^2\, ds \\ \le c(a_1-a)^{-1}\mathcal{YM}^{\mathcal{V}}(A(a)).\end{aligned}$$

Now apply Proposition B.9 with $R(a) := R(s_0)$ to obtain the estimate

$$\begin{aligned}\sup_{s\in I'} \|d_{A(s)}^* F_{A(s)}\|_{L^2(\Sigma)}^2 \le c(a_1-a)^{-1}\mathcal{YM}^{\mathcal{V}}(A(a)) + \int_I \|d_A \nabla\mathcal{V}(A)\|_{L^2(\Sigma)}^2 \\ + \int_I\int_\Sigma \big|\langle d_A^* F_A, \{d_A^* F_A, F_A\}\rangle\big| + \big|\langle d_A^* F_A, \{\nabla\mathcal{V}(A), F_A\}\rangle\big|,\end{aligned}$$

The last term in the first line and the two terms in the second line admit the required bound as follows from Proposition B.12 (use Lemma B.7 to further estimate the integral over $|F_A|^3$). The same type of estimate holds for $\|d_A^* F_A\|_{L^2(I'(L^p(\Sigma))}$ by Proposition B.8, for any $p < \infty$. To see this we apply Proposition B.13 to bound the term $\int_I \|d_A d_A^* F_A\|_{L^2(\Sigma)}^2$ on the right-hand side of (69). Using Hölder's inequality as in the proof of Lemma B.7 then completes the proof. $\square$

The following Propositions B.11, B.12, B.13, and B.14 are auxiliary results needed in the proofs of Lemmata B.7 and B.10.

Proposition B.11 *Let A solve* (63) *on $I\times\Sigma$. Then the product $|F_A|^2|d_A\nabla\mathcal{V}(A)|$ admits the estimate*

$$\int_{I\times\Sigma} |F_A|^2 |d_A\nabla\mathcal{V}(A)| \le c(I)\big(1 + \mathcal{YM}^{\mathcal{V}}(A(a))^2\big).$$

for a constant $c(I)$.

Proof: Using the Sobolev embedding $W^{1,2}(\Sigma) \hookrightarrow L^p(\Sigma)$ for every $p < \infty$ we may combine (55) and (57) to get for any fixed $\varepsilon > 0$ the estimate

$$\|d_A \nabla \mathcal{V}(A)\|_{L^p(\Sigma)} \le c\big(1 + \|F_A\|_{L^p(\Sigma)} + \|F_A\|^2_{L^{2+\varepsilon}(\Sigma)}\big).$$

For $p = 3$ it follows that

$$\begin{aligned}
\int_{I\times\Sigma} |F_A|^2 |d_A \nabla \mathcal{V}(A)| &\le \int_I \Big(\int_\Sigma |F_A|^3\Big)^{\frac{2}{3}} \cdot \Big(\int_\Sigma |d_A \nabla \mathcal{V}(A)|^3\Big)^{\frac{1}{3}} \\
\le\ & c\int_I \Big(\int_\Sigma |F_A|^3\Big)^{\frac{2}{3}} \cdot \Big(1 + \Big(\int_\Sigma |F_A|^3\Big)^{\frac{1}{3}} + \Big(\int_\Sigma |F_A|^{2+\varepsilon}\Big)^{\frac{2}{2+\varepsilon}}\Big) \\
\le\ & c\int_I \Big(\int_\Sigma |F_A|^3\Big)^{\frac{2}{3}} + c\int_{I\times\Sigma} |F_A|^3 + c\int_I \Big(\int_\Sigma |F_A|^3\Big)^{\frac{4}{3}} \\
& + c\int_I \Big(\int_\Sigma |F_A|^{2+\varepsilon}\Big)^{\frac{4}{2+\varepsilon}} \\
\le\ & c\big(\|F_A\|^3_{L^3(I\times\Sigma))} + \|F_A\|^4_{L^4(I,L^3(\Sigma))}\big).
\end{aligned}$$

The bound for $\|F_A\|^3_{L^3(I\times\Sigma))}$ follows from Lemma B.7. By the same arguments as used there and Hölder's inequality

$$\|F_A\|^4_{L^4(I,L^3(\Sigma))} \le \|F_A\|^2_{L^\infty(I,L^2(\Sigma))} \|F_A\|^2_{L^2(I,L^6(\Sigma))}$$

we obtain the required estimate for the last term. □

Proposition B.12 *Let A solve* (63) *on $I \times \Sigma$. Then the function $|F_A| \cdot |d_A^* F_A|^2 : I \times \Sigma \to \mathbb{R}$ as in Proposition B.2 satisfies the estimate*

$$\int_{I'\times\Sigma} |F_A| \cdot |d_A^* F_A|^2 \le c(I,I')\big(1 + \mathcal{YM}^{\mathcal{V}}(A(a))^2\big)$$

for a constant $c(I,I')$.

Proof: Consider the function $u_{0,3} = \frac{1}{3}| * F_A|^3 : I \times \Sigma \to \mathbb{R}$ as defined in Proposition B.2. It satisfies

$$L_\Sigma u_{0,3} = -| * F_A|\big(|d_A * F_A|^2 + \langle *F_A, *d_A \nabla \mathcal{V}(A)\rangle\big).$$

Lemma C.8 thus yields the estimate

$$\int_{I'\times\Sigma} |F_A| \cdot |d_A^* F_A|^2 \le c(I,I') \int_{I\times\Sigma} \frac{1}{3}|F_A|^3 + |F_A| \cdot |\langle *F_A, *d_A \nabla \mathcal{V}(A)\rangle|.$$

Now apply Lemma B.7 and Proposition B.11 to obtain the result. □

Proposition B.13 *There exists a constant $c > 0$ such that if A is a solution of (63) on $I \times \Sigma$, then*

$$\int_{I' \times \Sigma} |\nabla_A \nabla_A F_A|^2 \leq c\big(\|F_{A(0)}\|^2_{L^2(\Sigma)} + \|F_{A(0)}\|^3_{L^2(\Sigma)}\big).$$

Proof: We apply the mean value inequality (cf. Lemma C.7) to (67). This yields for a universal constant $C = C(I, I')$ the estimate

$$\int_{I' \times \Sigma} |\nabla_A \nabla_A F_A|^2 \leq C\Big(\int_{I \times \Sigma} |\nabla_A F_A|^2 + |\nabla_A F_A|^2 |F_A| + |F_A|^2 + |\nabla \mathcal{V}(A)|^2 + |\nabla_A d_A \nabla \mathcal{V}(A)|^2\Big).$$

We estimate the terms on the right-hand side separately. The estimate for $|\nabla_A F_A|^2 |F_A|$ follows from Proposition B.12. The term $|\nabla_A d_A \nabla \mathcal{V}(A)|^2$ can be estimated using Proposition B.14. $\square$

Proposition B.14 *Let $p > 2$. Then there exists a constant $c = c(I, p)$ such that*

$$\int_{I \times \Sigma} |\nabla_A d_A \nabla \mathcal{V}(A)|^2 \leq c\Big(1 + \int_I \|F_A\|^4_{L^p(\Sigma)} + \int_I \|\nabla_A F_A\|^2_{L^2(\Sigma)}\Big)$$

holds for all $A \in \mathcal{A}(P)$.

Proof: We first consider the term $|\nabla_A d_A \alpha|^2$ appearing in $\nabla_A d_A \nabla \mathcal{V}(A)$, cf. Proposition A.3. With $d_A \alpha = F_A - F_{A_0} + \frac{1}{2}[\alpha \wedge \alpha]$ it suffices to estimate the term $|\nabla_A [\alpha \wedge \alpha]|^2$. Set $q^{-1} := 1 - p^{-1}$. Using Hölder's inequality and estimate (55) we find that

$$\begin{aligned}
\int_{I \times \Sigma} |\nabla_A [\alpha \wedge \alpha]|^2 &\leq c \int_{I \times \Sigma} |\nabla_A \alpha|^2 |\alpha|^2 \\
&\leq c \int_I \Big(\int_\Sigma |\nabla_A \alpha|^p\Big)^{\frac{4}{p}} + \Big(\int_\Sigma |\alpha|^q\Big)^{\frac{4}{q}} \\
&= c \int_I \|\nabla_A \alpha\|^4_{L^p(\Sigma)} + \|\alpha\|^4_{L^q(\Sigma)} \\
&\leq c\Big(1 + \int_I \|F_A\|^4_{L^p(\Sigma)}\Big).
\end{aligned}$$

Denote $\beta := *[\alpha \wedge *\eta]$ and $\gamma := R_{A_0,\alpha}\beta$. We estimate the remaining term

$$\begin{aligned}
\nabla_A d_A T^*_{A_0,\alpha}\beta &= \nabla_A d_A [\alpha \wedge R_{A_0,\alpha}\beta] = \nabla_A [d_A \alpha \wedge \gamma] - \nabla_A [\alpha \wedge d_A \gamma] \\
&= [\nabla_A d_A \alpha \wedge \gamma] + [d_A \alpha \wedge \nabla_A \gamma] - [\nabla_A \alpha \wedge d_A \gamma] - [\alpha \wedge \nabla_A d_A \gamma].
\end{aligned}$$

Estimates for the last four terms are obtained similarly as above for $\nabla_A d_A \alpha$. Here we use that $\|\gamma\|_{W^{2,2}(\Sigma)}$ and hence $\|\gamma\|_{C^0(\Sigma)}$ is bounded in terms of $\|\alpha\|_{L^2(\Sigma)}$ as follows from Proposition A.1. $\square$

C Auxiliary results

We derive an a priori estimate for the linearized Yang-Mills gradient flow along a path $s \mapsto A(s) \in \mathcal{A}(P)$ of connections. This linearization is given by the operator $\mathcal{D}_A = \frac{d}{ds} + \mathcal{H}_A$ with

$$\mathcal{H}_A \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \psi \end{pmatrix} = \begin{pmatrix} \Delta_A \alpha_0 + *[*F_A \wedge \alpha_0] + [d_A^* F_A \wedge \varphi] - d_A \omega \\ -d_A \psi + d_A \omega \\ -d_A^* \alpha_1 \end{pmatrix} \tag{70}$$

the augmented Yang-Mills Hessian as defined in (33). Here $\alpha = \alpha_0 + \alpha_1$ denotes the Hodge decomposition of α with respect to A (i.e. $d_A^* \alpha_0 = 0$ and $\alpha_1 = d_A \varphi$ for some $\varphi \in \Omega^0(\Sigma, \mathrm{ad}(P))$). The map ω has been defined as the solution of the equation

$$\Delta_A \omega = *[d_A * F_A \wedge \alpha]. \tag{71}$$

Proposition C.1 *Fix $p > 3$. For any path $A \in W^{2,p}(I \times \Sigma)$ of connections and compact intervals $I_1 \subset I \subseteq \mathbb{R}$ there exists a constant $c(A, I_1, I, p)$ such that the estimate*

$$\begin{aligned} &\|\alpha_0\|_{W^{1,2;p}(I_1 \times \Sigma)} + \|\alpha_1\|_{W^{1,p}(I_1 \times \Sigma)} + \|\psi\|_{W^{1,p}(I_1 \times \Sigma)} \\ &\qquad \leq c(A, I_1, I, p)\big(\|\mathcal{D}_A \xi\|_{L^p(I \times \Sigma)} + \|\psi\|_{L^p(I \times \Sigma)}\big) \end{aligned} \tag{72}$$

is satisfied for every $\xi = (\alpha_0, \alpha_1, \psi) \in \operatorname{dom} \mathcal{D}_A$.

Proof: We first remark that the term $d_A \omega$ appearing in the first and second line of (70) can be estimated as $\|d_A \omega\|_{L^p(I_1 \times \Sigma)} \leq c(A, I_1, p)\|\alpha\|_{L^p(I_1 \times \Sigma)}$ for some constant $c(A, I_1, p)$. This follows from (71) by elliptic regularity which yields for fixed $s \in \mathbb{R}$ the estimate

$$\begin{aligned} \|d_{A(s)} \omega(s)\|_{L^p(\Sigma)} &\leq c(A(s))\|[d_{A(s)} * F_{A(s)} \wedge \alpha(s)]\|_{W^{-1,p}(\Sigma)} \\ &\leq c(A(s))\|d_{A(s)} * F_{A(s)}\|_{L^2(\Sigma)}\|\alpha(s)\|_{L^p(\Sigma)}. \end{aligned}$$

The last inequality is by (A.6). Integrating this estimate over I_1, the result follows. The estimate for the term α_0 follows from the first line of (70) via a standard parabolic estimate for the linear heat operator $\frac{d}{ds}+\Delta_{A(s)}:W^{1,2;p}(I\times\Sigma)\to L^p(I\times\Sigma)$. It remains to estimate the terms α_0 and ψ. For this let us define the linear operator

$$B:=\begin{pmatrix} 0 & -d_A \\ -d_A^* & 0 \end{pmatrix}$$

acting on pairs $(\alpha_1,\psi)^T$ and denote by

$$M:=[B,\partial_s]=\begin{pmatrix} 0 & [\dot A\wedge\,\cdot\,] \\ -*[\dot A\wedge *\cdot\,] & 0 \end{pmatrix}$$

its commutator with ∂_s. We also set

$$L:=(-\partial_s+B)(\partial_s+B)=-\partial_s^2+B^2+M=:\mathrm{diag}(L_1,L_2)+M$$

with $L_1=-\partial_s^2+d_Ad_A^*$ and $L_2=-\partial_s^2+d_A^*d_A$. Note that L_2 equals the Laplace operator $\hat\Delta_A$ on ad-valued forms over $\mathbb{R}\times\Sigma$ as introduced in Section 6. Similarly, L_1 acts on $\alpha_1=d_A\varphi$ as

$$\begin{aligned} L_1\alpha_1 &= (-\partial_s^2+d_Ad_A^*+d_A^*d_A)\alpha_1-d_A^*d_Ad_A\varphi \\ &= \hat\Delta_A\alpha_1+[*F_A\wedge *\alpha_1]-[d_A^*F_A\wedge\varphi]. \end{aligned}$$

We consider $-\partial_s+B$ as a bounded operator $L^p(I\times\Sigma)\to W^{-1,p}(I\times\Sigma)$ and denote by K its norm. The claimed estimate then follows from elliptic regularity of the Laplace operator $\hat\Delta_A:W^{1,p}(I\times\Sigma)\to W^{-1,p}(I\times\Sigma)$,

$$\begin{aligned} &\|\alpha_1\|_{W^{1,p}}+\|\psi\|_{W^{1,p}} \\ &\quad\le c\big(\|\hat\Delta_A\alpha_1\|_{W^{-1,p}}+\|\hat\Delta_A\psi\|_{W^{-1,p}}+\|\alpha_1\|_{L^p}+\|\psi\|_{L^p}\big) \\ &\quad\le c\big(\|L_1\alpha_1\|_{W^{-1,p}}+\|[*F_A\wedge *\alpha_1]-[d_A^*F_A\wedge\varphi]\|_{W^{-1,p}}+\|L_2\psi\|_{W^{-1,p}} \\ &\qquad+\|\alpha_1\|_{L^p}+\|\psi\|_{L^p}\big) \\ &\quad\le c\big(\|L(\alpha_1,\psi)^T\|_{W^{-1,p}}+\|M(\alpha_1,\psi)^T\|_{W^{-1,p}}+(1+\|F_A\|_{L^\infty})\|\alpha_1\|_{L^p} \\ &\qquad+(1+\|d_A^*F_A\|_{L^p})\|\varphi\|_{L^\infty}\big) \\ &\quad\le c\big(K\|(\partial_s+B)(\alpha_1,\psi)^T\|_{L^p}+(1+\|F_A\|_{L^\infty})\|\alpha_1\|_{L^p} \\ &\qquad+(1+\|d_A^*F_A\|_{L^p})\|\varphi\|_{W^{1,p}}\big) \\ &\quad\le c\big(K\|\partial_s\alpha_1-d_A\psi+d_A\omega\|_{L^p}+\|d_A\omega\|_{L^p}+K\|\partial_s\psi-d_A^*\alpha_1\|_{L^p} \\ &\qquad+(1+\|F_A\|_{L^\infty})\|\alpha_1\|_{L^p}+(1+\|d_A^*F_A\|_{L^p})\|\alpha_1\|_{L^p}\big). \end{aligned}$$

The fourth inequality follows from boundedness of the operator $M:L^p\to L^p$. Using the remark on the estimate for $d_A\omega$ made initially, the proof is complete. □

Lemma C.2 *Let $f : (-\infty, T) \to \mathbb{R}$ be a bounded C^2-function such that $f \geq 0$ and the differential inequality*

$$f'' \geq c_0 f + c_1 f' \tag{73}$$

is satisfied for constants $c_0 > 0$ and $c_1 \in \mathbb{R}$. Then f satisfies the decay estimate

$$f(s) \leq e^{k(s+T)} f(T)$$

for a constant $k = k(c_0, c_1) > 0$ and all $-\infty < s \leq T$.

Proof: Set

$$\begin{aligned} k &:= -\frac{c_1}{2} + \frac{1}{2}\sqrt{4c_0 + c_1^2} > 0 \\ \lambda &:= \frac{c_1}{2} + \frac{1}{2}\sqrt{4c_0 + c_1^2} > 0. \end{aligned}$$

Assume by contradiction that $f'(s_0) - kf(s_0) < 0$ for some $s_0 \leq T$ and set $g(s) := e^{ks}(f'(s) - \lambda f(s))$. Then

$$g' = e^{ks}(f'' + (k - \lambda)f' - k\lambda f) = e^{ks}(f'' - c_1 f' - c_0 f) \geq 0,$$

so g is monotone increasing. Therefore $g(s) \leq g(s_0)$ for all $s \leq s_0$ and

$$f'(s) \leq \lambda f(s) + e^{k(s_0 - s)}(f'(s_0) - \lambda f(s_0)).$$

Because f is bounded and $f'(s_0) - kf(s_0) < 0$ it follows that $f'(s) \to -\infty$ as $s \to -\infty$. This contradicts the boundedness of f as $f(s_0) - f(s) = \int_s^{s_0} f'(\sigma)\, d\sigma$. Therefore the assumption was wrong and $f'(s) - kf(s) \geq 0$ holds for all $-\infty < s \leq T$. Then with $h(s) := e^{-ks} f(s)$ it follows that

$$g' = e^{-ks}(f' - kf) \geq 0,$$

which implies $f(s) \leq e^{k(s-T)} f(T)$ for all $s \leq T$. □

Proposition C.3 *For every 1-form $\beta \in L^2(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))$ there holds the estimate $\|\beta\|_{W^{-1,2}(\Sigma)} \leq \|\beta\|_{L^2(\Sigma)}$.*

Proof: The result follows from the estimate

$$\begin{aligned} \|\beta\|_{W^{-1,2}(\Sigma)} &= \sup_{\|\varphi\|_{W^{1,2}(\Sigma)}=1} \Big| \int_\Sigma \langle \beta, \varphi \rangle \Big| \leq \sup_{\|\varphi\|_{W^{1,2}(\Sigma)}=1} \|\varphi\|_{L^2(\Sigma)} \|\beta\|_{L^2(\Sigma)} \\ &\leq \|\beta\|_{L^2(\Sigma)} \sup_{\|\varphi\|_{W^{1,2}(\Sigma)}=1} \|\varphi\|_{W^{1,2}(\Sigma)} = \|\beta\|_{L^2(\Sigma)}, \end{aligned}$$

where we used the Cauchy-Schwarz inequality in the second step. □

Lemma C.4 *For each $\gamma < 1$ there is a continuous embedding $L^{\frac{2}{1+\gamma}}(\Sigma) \hookrightarrow W^{-\gamma,2}(\Sigma)$.*

Proof: The usual Sobolev embedding theorem asserts that $W^{\gamma,2}(\Sigma)$ embeds continuously into $L^{\frac{2}{1-\gamma}}(\Sigma)$. It thus follows for $\varphi \in C^\infty(\Sigma)$ that

$$\begin{aligned}\|\varphi\|_{W^{-\gamma,2}(\Sigma)} = \sup_{\|\psi\|_{W^{\gamma,2}(\Sigma)}=1} \langle \varphi, \psi \rangle &\leq \sup_{\|\psi\|_{W^{\gamma,2}(\Sigma)}=1} \|\varphi\|_{L^{\frac{2}{1+\gamma}}(\Sigma)} \|\psi\|_{L^{\frac{2}{1-\gamma}}(\Sigma)} \\ &\leq C\|\varphi\|_{L^{\frac{2}{1+\gamma}}(\Sigma)}.\end{aligned}$$

In the second step we used Hölder's inequality. This proves the claim. □

Lemma C.5 (Sobolev multiplication) *For every $1 < p < 2$ there is a constant $C(p)$ such that the estimate*

$$\|uv\|_{L^p(\Sigma)} \leq C(p)\|u\|_{L^2(\Sigma)}\|v\|_{W^{1,2}(\Sigma)}$$

is satisfied for all functions $u \in L^2(\Sigma)$ and $v \in W^{1,2}(\Sigma)$.

Proof: The claim follows with $q := \frac{2}{2-p}$ from the estimate

$$\int_\Sigma |uv|^p \leq \||u|^p\|_{L^{\frac{2}{p}}(\Sigma)} \||v|^p\|_{L^q(\Sigma)} = \|u\|^p_{L^2(\Sigma)}\|v\|^p_{L^{pq}(\Sigma)} \leq C\|u\|^p_{L^2(\Sigma)}\|v\|^p_{W^{1,2}(\Sigma)}.$$

In the first step we used Hölder's inequality, and the last step follows from the Sobolev embedding $W^{1,2}(\Sigma) \hookrightarrow L^r(\Sigma)$ for all $1 \leq r < \infty$. □

Lemma C.6 (Elliptic mean value inequality) *There is a constant $c > 0$ such that the following holds for all $r > 0$ smaller than the injectivity radius of M. If $a \geq 0$ and the function $u : M \to \mathbb{R}$ is of class C^2 in such that*

$$u \geq 0 \qquad \text{and} \qquad \Delta_\Sigma u \leq au,$$

then for all $x \in M$,

$$u(x) \leq \frac{c}{r^n} \int_{B_r(x)} u.$$

Proof: For the proof we refer to [24]. □

Lemma C.7 (Parabolic mean value inequality) *There is a constant $c > 0$ such that the following holds for all $r > 0$ smaller than the injectivity radius of M. If $a \geq 0$ and the function $u : \mathbb{R} \times M \to \mathbb{R}$ is of class C^1 in the s-variable and of class C^2 in the spatial variables such that*

$$u \geq 0 \qquad \text{and} \qquad (\partial_s + \Delta_\Sigma) u \leq au,$$

then for all $x \in M$,

$$u(x) \leq \frac{c e^{ar^2}}{r^{n+2}} \int_{P_r(x)} u.$$

Proof: For the proof we refer to [25, Lemma B.2]. □

Lemma C.8 *Let $R, r > 0$, $u : P_{R+r} \to \mathbb{R}$ be a C^2 function and $f, g : P_{R+r} \to \mathbb{R}$ be continuous functions such that*

$$-L_\Sigma u \geq g - f, \qquad u \geq 0, \quad f \geq 0, \quad g \geq 0.$$

Then

$$\int_{P_{R/2}} g \leq 2\left(1 + \frac{r}{R}\right)\left(\int_{P_{R+r}} f + \left(\frac{4}{r^2} + \frac{1}{Rr}\right) \int_{P_{R+r}} u\right).$$

Proof: For the proof we refer to [25, Lemma B.5]. □

Lemma C.9 (Interpolation) *For real numbers r, r', s, s' the intersection $W^{r,s;2}(I \times \Sigma) \cap W^{r',s';2}(I \times \Sigma)$ is a Banach space with norm $\big(\| \cdot \|_{W^{r,s;2}(I\times\Sigma)} + (\| \cdot \|_{W^{r',s';2}(I\times\Sigma)}\big)^{\frac{1}{2}}$. For any $\theta \in [0,1]$ there is a bounded embedding*

$$W^{r,s;2}(I \times \Sigma) \cap W^{r',s';2}(I \times \Sigma) \hookrightarrow W^{\theta r + (1-\theta) r', \theta s + (1-\theta) s';2}(I \times \Sigma).$$

Proof: For the proof we refer to [38, Lemma A.0.3]. □

Lemma C.10 (Traces) *For $r > \frac{1}{2}$ the map $u \mapsto u(0)$ on smooth functions extends to a bounded map (trace map) $W^{r,s;2}([0,T] \times \Sigma) \to W^{s,2}(\Sigma)$.*

Proof: For the proof we refer to [38, Lemma A.0.3]. □

References

[1] R. Abraham, J. E. Marsden, T. Ratiu, *Manifolds, Tensor Analysis, and Applications. Second Edition*, New York, Berlin, Heidelberg: Springer, 1988.

[2] A. Abbondandolo, P. Majer, Lectures on the Morse complex for infinite dimensional manifolds, in *Morse theoretic methods in nonlinear analysis and in symplectic topology*, 1-74, NATO Science Series II: Mathematics, Physics and Chemistry, P. Biran, O. Cornea, and F. Lalonde Eds, Springer, 2006.

[3] A. Abbondandolo, M. Schwarz, On the Floer Homology of Cotangent Bundles, *Comm. Pure and Applied Math.* **LIX**, 254-316, 2006.

[4] A. Abbondandolo, M. Schwarz, Notes on Floer homology and loop space homology, in *Morse theoretic methods in nonlinear analysis and in symplectic topology*, 1-74, NATO Science Series II: Mathematics, Physics and Chemistry, P. Biran, O. Cornea, and F. Lalonde Eds, Springer, 2006.

[5] M. F. Atiyah, R. Bott, The Yang-Mills Equations over Riemann Surfaces, *Phil. Trans. R. Soc. Lond. A* **308**, 523-615, 1982.

[6] P. Biran, O. Cornea, Quantum Structures for Lagrangian Submanifolds, *preprint*, arXiv:0708.4221, 2007.

[7] R. Bott, An application of the Morse theory to the topology of Lie groups, *Bull. Soc. Math. France* **84**, 251-281, 1956.

[8] R. Bott, Morse theory indomitable, *Publ. Math. I.H.E.S.* **68**, 99-114, 1988.

[9] R. L. Cohen, M. Schwarz, A Morse theoretic description of string topology, in *New perspectives and challenges in symplectic field theory*, 147-172, CRM Proc. Lecture Notes, **49**, Amer. Math. Soc., Providence, RI, 2009.

[10] S. K. Donaldson, Anti self-dual Yang-Mills connections over complex algebraic surfaces and stable vector bundles, *Proc. London Math. Soc.* **50**, 1-26, 1985.

[11] S. K. Donaldson, *Floer Homology Groups in Yang-Mills Theory*, Cambridge: Cambridge University Press, 2002.

[12] A. Floer, An instanton invariant for 3-manifolds, *Comm. Math. Phys.* **118**, 215-240, 1988.

[13] A. Floer, Symplectic fixed points and holomorphic spheres, *Comm. Math. Phys.* **120**, 575-611, 1989.

[14] U. Frauenfelder, *Floer Homology of Symplectic Quotients and the Arnold-Givental Conjecture*, Diss. ETH no. 14981, ETH Zürich, 2003.

[15] U. Frauenfelder, The Arnold-Givental Conjecture and Moment Floer Homology, *Int. Math. Res. Not.* **42**, 2179-2269, 2004.

[16] F. Kirwan, Moduli Spaces of Bundles over Riemann Surfaces and the Yang-Mills Stratification Revisited, In *Strings and Geometry*, edited by M. Douglas, J. Gauntlett, and M. Gross, Proceedings of the Clay Mathematics Institute 2002 Summer School, Providence: American Mathematical Society, 2004, pp 239-283.

[17] N. V. Krylov, *Lectures on Elliptic and Parabolic Equations in Sobolev Spaces*, Providence: American Mathematical Society, 2008.

[18] G. M. Lieberman, *Second Order Parabolic Differential Equations*, Singapore: World Scientific Publ., 1996.

[19] J. W. Milnor, *Morse theory*, Ann. of. Math. Studies **51**, Princeton Univ. Press, 1963.

[20] T. Mrowka, K. Wehrheim, L^2-topology and Lagrangians in the space of connections over a Riemann surface, to appear in *GAFA*, 2010.

[21] J. Råde, On the Yang-Mills heat equation in two and three dimensions, *J. reine angew. Math.* **431**, 123-163, 1992.

[22] J. Råde, Compactness theorems for invariant connections, *preprint*, arXiv:math/0004041v1, 2000.

[23] J. W. Robbin, D. A. Salamon, The spectral flow and the Maslov index, *Bull. London Math. Soc.* **27**, 1-33, 1995.

[24] D. A. Salamon, Lectures on Floer Homology, In *Symplectic Geometry and Topology*, edited by Y. Eliashberg and L. Traynor, IAS/Park City Mathematics Series, Vol **7**, Providence: American Mathematical Society, 1999, pp 143-230.

[25] D. A. Salamon, J. Weber, Floer homology and the heat flow, *GAFA* **16**, 1050-1138, 2006.

[26] D. A. Salamon, K. Wehrheim, Instanton Floer homology with Lagrangian boundary conditions, *Geometry & Topology* **12**, 747-918, 2008.

[27] M. Schwarz, *Morse Homology*, Basel: Birkhäuser Verlag, 1993.

[28] S. Smale, On gradient dynamical systems, *Ann. of Math.* **74**, 199-206, 1961.

[29] M. Struwe, The Yang-Mills flow in four dimensions, *Calc. Var.* **2**, 123-150, 1994.

[30] J. Swoboda, *The Yang-Mills Gradient Flow and Loop Groups*, Diss. ETH no. 18296, ETH Zürich, 2009.

[31] J. Swoboda, Yang-Mills Morse complex and homology of loop groups, *in preparation*, 2010.

[32] R. Thom, Sur une partition en cellules associés à une fonction sur une variété, *C. R. Acad. Sci. Paria* **228**, 973-975, 1949.

[33] K. K. Uhlenbeck, Removable singularities in Yang-Mills fields, *Comm. Math. Phys.* **83**, 11-29, 1982.

[34] K. K. Uhlenbeck, Connections with L^p-bounds on curvature, *Comm. Math. Phys.* **83**, 31-42, 1982.

[35] K. Wehrheim, *Uhlenbeck Compactness*, EMS Series of Lectures in Mathematics, Zürich: European Mathematical Society Publishing House, 2004.

[36] J. Weber, *The heat flow and the homology of the loop space*, Habilitationsschrift, HU Berlin, 2010.

[37] E. Witten, Supersymmetry and Morse theory, *J. Diff. Geom.* **17**, 661-692, 1982.

[38] C. T. Woodward, The Yang-Mills heat flow on the moduli space of framed bundles on a surface, *Amer. J. Math.* **128**, no. 2, 311–359, 2006.