The Yang–Mills Gradient Flow and Loop Spaces of Compact Lie Groups

Jan Swoboda

August 14, 2017

Abstract

We study the L^2 gradient flow of the Yang–Mills functional on the space of connection 1-forms on a principal G-bundle over the sphere S^2 from the perspective of Morse theory. The resulting Morse homology is compared to the heat flow homology of the space ΩG of based loops in the compact Lie group G. An isomorphism between these two Morse homologies is obtained by coupling a perturbed version of the Yang–Mills gradient flow with the L^2 gradient flow of the classical action functional on loops. Our result gives a positive answer to a question due to Atiyah.

Contents

1

1   Introduction

Let $\Sigma := S^2$ denote the unit sphere in the euclidian space $\mathbb{R}^3$. Let G be a compact Lie group, $\mathfrak{g}$ its Lie algebra (endowed with an Ad-invariant inner product), and P a principal G-bundle over Σ. In this paper we give an affirmative answer to a question raised by Atiyah relating Yang–Mills Morse homology of the space of gauge equivalence classes of $\mathfrak{g}$-valued connections on P to heat flow homology of the group ΩG of based loops in G. The Morse complexes we shall be concerned with are the complex generated by the L^2 gradient flow of the Yang–Mills functional $\mathcal{YM}$ on the one hand, and the complex generated by the L^2 gradient flow of the classical action functional on ΩG on the other. Our goal is to establish a chain isomorphism between these two Morse complexes.

Let us describe our setup. We denote by $\mathrm{ad}(P) := P \times_{\mathrm{Ad}} \mathfrak{g}$ the adjoint Lie algebra bundle over Σ, and by $\mathcal{A}(P)$ the space of $\mathfrak{g}$-valued 1-forms on P. The latter is an affine space over $\Omega^1(\Sigma, \mathrm{ad}(P))$, the space of $\mathrm{ad}(P)$-valued 1-forms on Σ. The curvature of a connection $A \in \mathcal{A}(P)$ is $F_A = dA + \frac{1}{2}[A \wedge A] \in \Omega^2(\Sigma, \mathrm{ad}(P))$. The space $\mathcal{A}(P)$ is acted on by the groups $\mathcal{G}(P)$ and $\mathcal{G}_0(P)$ of gauge, respectively based gauge transformations of P, cf. Section 2.1 for precise definitions. On $\mathcal{A}(P)$ we consider the $\mathcal{G}(P)$-invariant Yang–Mills functional

$$\mathcal{YM}\colon \mathcal{A}(P) \to \mathbb{R}, \qquad \mathcal{YM}(A) = \frac{1}{2}\int_\Sigma \big\langle F_A \wedge *F_A \big\rangle. \qquad (1)$$

The corresponding Euler-Lagrange equation is the second order partial differential equation $d_A^* F_A = 0$, called *Yang–Mills equation*. Critical points

of $\mathcal{YM}$ are degenerate (due to the gauge invariance of the functional) but satisfy the so-called Morse–Bott condition, cf. the discussion in Section 2.3. We shall be concerned with a perturbed version of the negative L^2 gradient flow equation associated with (1), which is the PDE

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

Cf. Section 2.2 below for the precise form of the perturbation $\mathcal{V}^- \colon \mathcal{A}(P) \to \mathbb{R}$. The term $d_A \Psi$ (where $\Psi \in \Omega^0(\Sigma, \mathrm{ad}(P))$) is introduced to make (2) invariant under time-dependent gauge transformations.

Let $S^1 := \mathbb{R}/2\pi\mathbb{Z}$. The *free loop group* of G is the space $\Lambda G := C^\infty(S^1, G)$, endowed with the group multiplication defined by $(x_1 x_2)(t) := x_1(t) x_2(t)$ for $x_1, x_2 \in \Lambda G$. The *based loop group* of G is the subgroup

$$\Omega G := \{ x \in \Lambda G \mid x(0) = \mathbb{1} \}$$

of ΛG. Throughout we will identify ΩG with the quotient of ΛG modulo the free action of G defined by $(h \cdot x)(t) := hx(t)$ for $h \in G$ and $x \in \Lambda G$. We endow G with the biinvariant Riemannian metric induced by the Ad-invariant inner product on $\mathfrak{g}$ and consider on ΛG the classical action functional

$$\mathcal{E} \colon \Lambda G \to \mathbb{R}, \qquad \mathcal{E}(x) = \frac{1}{2} \int_0^{2\pi} \|\partial_t x(t)\|^2 \, dt. \tag{3}$$

It descends to a functional on ΩG by biinvariance of the Riemannian metric on G. It is well-known that the critical points of $\mathcal{E}$ are precisely the closed geodesics in G. As a consequence of the invariance of the functional $\mathcal{E}$ under conjugation with elements $h \in G$, it follows that critical points of $\mathcal{E}$ are degenerate. However, also here it turns out that the Morse–Bott condition is satisfied. The (perturbed) negative L^2 gradient flow equation resulting from (3) is the PDE

$$\partial_s x - \nabla_t \partial_t x + \nabla \mathcal{V}^+(x) = 0. \tag{4}$$

For the precise form of the perturbation $\mathcal{V}^+ \colon \Omega G \to \mathbb{R}$, we refer to Section 2.2 below.

Morse homology groups for loop spaces of compact Lie groups and homogeneous spaces have been computed in a classical paper by Bott [4], which constitutes an application of Morse theoretic ideas in the context of infinite dimensional Hilbert manifolds. For further applications to the theory

of closed geodesics on general compact manifolds we refer to Klingenberg [12]. However, in both instances, Morse theory is based on a $W^{1,2}$ gradient flow, leading to an ODE in Hilbert space. In contrast, the L^2 gradient flow approach to Morse theory on loop spaces of compact Riemannian manifolds has only recently been investigated by Weber [22] and uses techniques from parabolic PDEs (cf. also Salamon and Weber [16] for an application to Floer homology of cotangent bundles). In the present work, we shall follow the latter approach and specialize some of the results in [22] to loop spaces of compact Lie groups. Complementary to heat flow homology, a Morse homology theory based on the L^2 gradient flow (2) on compact Riemann surfaces of arbitrary genus has been laid down by this author in [19]. These results are used throughout the present article.

Main results

In their seminal paper [3], Atiyah and Bott studied the Yang–Mills functional $\mathcal{YM}$ over a compact Riemann surface from a Morse–Bott theoretical point of view. This led them to the discovery of a close correspondence between the Morse theoretical picture of a stratification of the space $\mathcal{A}(P)/\mathcal{G}(P)$ into stable manifolds and certain moduli spaces of semi-stable holomorphic vector bundles, and initiated a lot of further research in algebraic geometry as documented e.g. by the review article [11] by Kirwan. In [3], Atiyah and Bott pointed out that in the genus zero case the Yang–Mills critical points correspond via a so called holonomy map to closed geodesics in G. This observation was subsequently made more explicit through works by Gravesen [9] and Friedrich and Habermann [8]. In these articles, a holonomy map $\Phi \colon \mathcal{A}(P) \to \Omega G$ is constructed which assigns to a connection A its holonomy along the greater arcs in Σ connecting the north and south pole, cf. Appendix A for details. The map Φ is equivariant with respect to the actions of $\mathcal{G}(P)$ by gauge transformations and of G by conjugation. It furthermore maps critical points of the Yang–Mills functional to closed geodesics in G (of a certain homotopy type, determined by the bundle P), preserving the Morse indices. The natural question, raised by Atiyah, whether this apparent close relation between the aforementioned sets of generators of Morse complexes extends to the full Morse theory picture has not been resolved so far. However, a formal consideration invoking an adiabatic limit of a certain deformation of the standard Riemannian metric on $\Sigma = S^2$ indicates a positive answer to his question. Namely it is suggested that for a family of Riemannian metrics on Σ which approximates a δ-impulse on the equator, connecting trajectories of the Yang–Mills gradient flow could be constructed in a bijective way from

4

from those of the heat flow (4). Such an approach, although successfully been followed in many related contexts (cf. e.g. [6, 10]), did not provide an answer to that question. However, Davies has obtained some interesting preliminary results in his unpublished PhD thesis [5] supervised by Salamon. The aim of the present paper is to settle Atiyah's question following a completely different approach. The guiding idea in our proof is to combine the L^2 gradient flows (2) and (4) by studying a so-called *hybrid moduli space problem*. For a given pair $\hat{C}^{\pm}$ of critical manifolds of the functionals $\mathcal{YM}$, respectively $\mathcal{E}$, we shall consider the space of configurations

$$
\begin{aligned}
\hat{\mathcal{M}}(\hat{\mathcal{C}}^-,&\hat{\mathcal{C}}^+) := \\
&\big\{(A, \Psi, x) \in C^{\infty}(\mathbb{R}^-, \mathcal{A}(P) \times \Omega^0(\Sigma, \mathrm{ad}(P))) \times C^{\infty}(\mathbb{R}^+, \Lambda G) \mid \\
(A, \Psi) &\text{ satisfies (2)}, \quad x \text{ satisfies (4)}, \quad x(0) = h\Phi(A(0)) \text{ for some } h \in G, \\
\lim_{s \to -\infty} &(A(s), \Psi(s)) = \big(A^-, 0\big) \in \hat{\mathcal{C}}^- \times \Omega^0(\Sigma, \mathrm{ad}(P)), \lim_{s \to +\infty} x(s) = x^+ \in \hat{\mathcal{C}}^+\big\}.
\end{aligned}
$$

Hence $\hat{\mathcal{M}}(\hat{\mathcal{C}}^-, \hat{\mathcal{C}}^+)$ is the moduli space of tuples (A, Ψ, x) such that (A, Ψ) solves the perturbed Yang–Mills gradient flow equation (2) on the negative time interval $(-\infty, 0]$, while x is a solution of the perturbed loop group gradient flow equation (4) on the positive time interval $[0, \infty)$. Both solutions are coupled under the holonomy map $\Phi \colon \mathcal{A}(P) \to \Omega G$ as introduced above. The moduli space $\mathcal{M}(\mathcal{C}^-, \mathcal{C}^+)$ then to be studied is the quotient of $\hat{\mathcal{M}}(\hat{\mathcal{C}}^-, \hat{\mathcal{C}}^+)$ modulo the actions by gauge transformations and left translations $x \mapsto hx$ (for $h \in G$).

As pointed out before, the sets of critical points of both functionals $\mathcal{YM}$ and $\mathcal{E}$ are degenerate in a Morse–Bott sense. This fact requires us to use a certain variant of Morse theory, called *Morse theory with cascades*, as introduced by Frauenfelder in [7] and described in [19, Section 8.1]. Throughout we shall work on fixed sublevel sets $\{A \in \mathcal{A}(P) \mid \mathcal{YM}(A) \leq a\}$ and $\{x \in \Lambda G/G \mid \mathcal{E}(x) \leq b\}$ (where usually $b = 4a/\pi$). As an additional datum, we fix a Morse function h on the union of critical manifolds of $\mathcal{YM}$ below the level a (respectively of $\mathcal{E}$ below the level b), the discrete set of critical points of which are the generators of two Morse complexes

$$
CM_*^a\big(\mathcal{A}(P)/\mathcal{G}_0(P), \mathcal{V}^-, h\big) \qquad \text{and} \qquad CM_*^b\big(\Lambda G/G, \mathcal{V}^+, h\big). \tag{5}
$$

Here and throughout this article we adopt the convention that $\Omega G = \Lambda G/G$ shall denote the connected component of the based loop group which contains the image of $\mathcal{A}(P)$ under the map Φ, cf. Appendix A. It is determined by the equivalence class of the principal G-bundle P. Our goal is to

set up a chain homomorphism Θ between the complexes in (5). It is defined for a pair of generators of equal Morse index by a count of elements in a certain *moduli space with cascades*. These are constructed from a suitable concatenation of elements in $\mathcal{M}(\mathcal{C}^-, \mathcal{C}^+)$ and negative gradient flow lines of the function h, cf. Definition 5.1. The key observation, allowing us to show invertibility of the homomorphism Θ, is the property of the holonomy map $\Phi \colon \mathcal{A}(P) \to \Omega G$ to decrease energy. Namely, for any connection $A \in \mathcal{A}(P)$ there holds the inequality

$$\mathcal{YM}^{\mathcal{V}^-}(A) \geq \frac{\pi}{4} \mathcal{E}^{\mathcal{V}^+}(\Phi(A)), \tag{6}$$

(cf. Lemma A.4), with equality if A is a Yang–Mills connection. Inequality (6) is not a new result and can be found in Gravesen [9]. In our context it leads directly to the proof of invertibility of Θ and thus implies the desired isomorphism in Morse homology.

Theorem 1.1 (Main result) *Let G be a compact Lie group, and P a principal G-bundle P over Σ. Let $a \geq 0$ be a regular value of $\mathcal{YM}$ and set $b := 4a/\pi$. Then, for a generic a-admissible perturbation $\mathcal{V} = (\mathcal{V}^-, \mathcal{V}^+) \in Y_a$ (cf. Definition 2.7 below) the chain homomorphism*

$$\Theta_* \colon CM_*^a\big(\mathcal{A}(P)/\mathcal{G}_0(P), \mathcal{V}^-, h\big) \to CM_*^b\big(\Lambda G/G, \mathcal{V}^+, h\big)$$

induces an isomorphism

$$\boxed{[\Theta_*] \colon HM_*^a\big(\mathcal{A}(P)/\mathcal{G}_0(P), \mathcal{V}^-, h\big) \to HM_*^b\big(\Lambda G/G, \mathcal{V}^+, h\big)}$$

of Morse homology groups.

Let us point out here that the method of defining chain maps between Morse homology complexes by utilizing hybrid moduli spaces is a fairly recent one. It has successfully been employed by Abbondandolo and Schwarz [2] in proving that Floer homology $HF_*(T^*M)$ of cotangent bundles T^*M is isomorphic to singular homology of the free loop space ΛM (M a compact manifold). In their situation, a crucial role is played by an inequality similar to (6), relating the symplectic action to the classical action via Legendre duality.

Further directions

G-equivariant Morse homology   The group G acts in a natural way on the quotient spaces $\mathcal{G}(P)/G_0(P)$ and $\Lambda G/G$. In the first case, it is given by

$g \cdot [A] = [g^*A]$ for $g \in G \cong \mathcal{G}(P)/\mathcal{G}_0(P)$. In the second case the group G acts by conjugation $g[x] = [g^{-1}xg]$. In his thesis [18] the author has worked out a G-equivariant version of Theorem 1.1. This is mainly a technical extension and requires to replace the spaces $\mathcal{A}(P)/\mathcal{G}_0(P)$ and $\Lambda G/G$ by $(\mathcal{A}(P) \times E_nG)/\mathcal{G}(P)$, respectively by $(\Lambda G \times E_nG)/(G \times G)$, for a suitable finite-dimensional approximation E_nG of the classifying space EG.

Higher genus surfaces Morse homology $HM_*^a\big(\mathcal{A}(P)/\mathcal{G}_0(P), \mathcal{V}^-, h\big)$ as considered in this article for $\Sigma = S^2$ has more generally been defined in [19] for closed Riemann surfaces of arbitrary genus. It is known from [5, 9] that Yang–Mills connections on principal G-bundles P over such surfaces correspond bijectively to certain geodesic polygons in the Lie group G. Moreover, an estimate similar to (43) relating the energy functionals $\mathcal{YM}$ and $\mathcal{E}$ continuous to hold true in this more general situation. Hence one should be able to prove a version of Theorem 1.1 for higher genus surfaces, but this is open at present.

Acknowledgements

2 Critical manifolds, Yang–Mills gradient flow lines, and Morse complexes

2.1 Preliminaries

Let $\Sigma := S^2$ be the unit sphere in $\mathbb{R}^3$, endowed with the standard round metric. Let G be a compact Lie group with Lie algebra $\mathfrak{g}$. On $\mathfrak{g}$ we fix an Ad-invariant inner product $\langle \cdot, \cdot \rangle$, which exists by compactness of G. Let P be a principal G-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 $(g, h) \mapsto g^{-1}hg$. Let $\mathrm{ad}(P)$ denote the Lie algebra bundle

associated to P via the adjoint action

$$(g, \xi) \mapsto \left. \frac{d}{dt} \right|_{t=0} g^{-1} \exp(t\xi) g \qquad (\text{for } g \in G, \ \xi \in \mathfrak{g})$$

of G on $\mathfrak{g}$. We denote the space of smooth $\mathrm{ad}(P)$-valued differential k-forms by $\Omega^k(\Sigma, \mathrm{ad}(P))$, and by $\mathcal{A}(P)$ the space of smooth connections on P. The latter is an affine space over $\Omega^1(\Sigma, \mathrm{ad}(P))$. The group $\mathcal{G}(P)$ acts on $\mathcal{A}(P)$ by gauge transformations. We call a connection $A \in \mathcal{A}(P)$ *irreducible* if the stabilizer subgroup $\mathrm{Stab}\, A \subseteq \mathcal{G}(P)$ is trivial. Otherwise it is called *reducible*. It is easy to show that $\mathrm{Stab}\, A$ is a compact Lie group, isomorphic to a subgroup of G. Let $z \in \Sigma$ be arbitrary but fixed. We let $\mathcal{G}_0(P) \subseteq \mathcal{G}(P)$ denote the group of *based gauge transformation*, i.e. those gauge transformations which leave the fibre $P_z \subseteq P$ above z pointwise fixed. It is a well-known fact that $\mathcal{G}_0(P)$ acts freely on $\mathcal{A}(P)$.

On $\mathcal{A}(P)$ we define a gauge-invariant L^2 inner product by $\langle \alpha, \beta \rangle = \int_\Sigma \langle \alpha \wedge *\beta \rangle$ for $\alpha, \beta \in \Omega^1(\Sigma, \mathrm{ad}(P))$. The curvature of the connection A is the $\mathrm{ad}(P)$-valued 2-form $F_A = dA + \frac{1}{2}[A \wedge A]$. It satisfies the Bianchi identity $d_A F_A = 0$. The covariant exterior differential induced by $A \in \mathcal{A}(P)$ is the operator

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

The formal adjoint d_A^* of it is given by $d_A = - * d_A *$. The covariant Hodge Laplacian on forms is the operator $\Delta_A := d_A^* d_A + d_A d_A^*$. The *perturbed Yang–Mills functional* $\mathcal{YM}^\mathcal{V}$ has been introduced in (1). If $\mathcal{V} = 0$, we write $\mathcal{YM}$ and call this the *unperturbed Yang–Mills functional*. The L^2 gradient of $\mathcal{YM}^\mathcal{V}$ at $A \in \mathcal{A}(P)$ is $\nabla \mathcal{YM}^\mathcal{V}(A) = d_A^* F_A + \nabla \mathcal{V}(A) \in \Omega^1(\Sigma, \mathrm{ad}(P))$. Its Hessian is the second order differential operator

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

We also make use of the notation $H_A := d_A^* d_A + *[*F_A \wedge \cdot]$.

Throughout we will use Sobolev spaces of sections of vector bundles and Banach manifolds modeled on such Sobolev spaces, like e.g. various groups of gauge transformations. A detailed account of this subject is given in the book [20, Appendix B]. We therefore keep the discussion of these matters short. Let $1 \le p \le \infty$ and $k \ge 0$ an integer. We fix a smooth reference connection $A \in \mathcal{A}(P)$. It determines a covariant derivative ∇_A on $\Omega^*(\Sigma, \mathrm{ad}(P))$. We employ the notation $W^{k,p}(\Sigma, \mathrm{ad}(P))$ and $W^{k,p}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))$ for the

Sobolev spaces of $\mathrm{ad}(P)$-valued 0- and 1-forms whose weak derivatives (with respect to ∇_A) up to order k are in L^p. These spaces are independent of the choice of A. However, for $k \geq 1$, the corresponding norms depend on this choice. The standard Sobolev embedding and Rellich–Kondrachov compactness theorems apply to these spaces. The affine (k, p)-Sobolev space of connections on P is defined as

$$\mathcal{A}^{k,p}(P) := A + W^{k,p}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P)).$$

This definition is again independent of the choice of smooth reference connection A. To define Sobolev spaces of gauge transformations we need to assume $kp > \dim \Sigma = 2$. Then let $\mathcal{G}^{k,p}(P)$ denote the set of equivariant maps $P \to G$ which are of the form $g = g_0 \exp(\varphi)$, where g_0 is a smooth such map and $\varphi \in W^{k,p}(\Sigma, \mathrm{ad}(P))$. (Here we view φ as an equivariant map $P \to \mathfrak{g}$). The space $\mathcal{G}^{k,p}(P)$ is a Banach manifold modeled on $W^{k,p}(\Sigma, \mathrm{ad}(P))$. As a well-known fact we remark that $\mathcal{G}^{k,p}(P)$ is a group with smooth group multiplication and inversion. It acts smoothly on $\mathcal{A}^{k-1,p}(P)$ by gauge transformations. Let I be a finite or infinite interval. We often make use of the parabolic Sobolev space

$$W^{1,2;p}(I \times \Sigma, \mathrm{ad}(P)) := L^p(I, W^{2,p}(\Sigma, \mathrm{ad}(P))) \cap W^{1,p}(I, L^p(\Sigma, \mathrm{ad}(P))) \quad (7)$$

of $\mathrm{ad}(P)$-valued 0-forms admitting one time and two space derivatives in L^p (and similarly for $\mathrm{ad}(P)$-valued 1-forms). The parabolic Sobolev space $\mathcal{A}^{1,2;p}(P)$ of connections is defined analogously, with $W^{2,p}(\Sigma, \mathrm{ad}(P))$ and $L^p(\Sigma, \mathrm{ad}(P))$ in (7) replaced by $\mathcal{A}^{2,p}(P)$, respectively $\mathcal{A}^{0,p}(P)$. Note that when there is no danger of confusion, we for ease of notation write $L^p(\Sigma)$ instead of $L^p(\Sigma, \mathrm{ad}(P))$ or $L^p(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))$ (and similarly for the other Sobolev spaces). Further notation frequently used is $\dot{A} := \partial_s A := \frac{dA}{ds}$, etc. for derivatives with respect to the time parameter s.

2.2 Banach spaces of abstract perturbations

Perturbations of the Yang–Mills funcional

Our construction of a Banach space of perturbations is based on the following L^2 local slice theorem due to Mrowka and Wehrheim [13]. We fix $p > 2$ and let

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

denote the set of L^p-connections in the local slice of radius ε with respect to the reference connection $A_0 \in \mathcal{A}^{0,p}(P)$.

Theorem 2.1 (L^2 **local slice theorem**) *Let $p > 2$. For every $A_0 \in \mathcal{A}^{0,p}(P)$ there are constants $\varepsilon, \delta > 0$ such that the map*

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

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

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

Proof: For a proof we refer to [13, Theorem 1.7]. $\qquad\qquad\square$

    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 each $i \in \mathbb{N}$ 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 $\rho \colon \mathbb{R} \to [0,1]$ such that $\rho = 1$ on $[-1, 1]$, $\operatorname{supp} \rho \subseteq [-4, 4]$, and $\|\rho'\|_{L^\infty(\mathbb{R})} < 1$. Set $\rho_k(r) := \rho(k^2 r)$ for $k \in \mathbb{N}$.
We fix $i \in \mathbb{N}$ and a constant $\varepsilon_i > 0$ such that the conclusion of Theorem 2.1 applies for $A_0 := A_i$ and this constant ε_i. Note that by assumption, $\operatorname{Stab} A_i = \mathbb{1}$. Theorem 2.1 thus implies that the map

$$\mathfrak{m}_i \colon \mathcal{S}_{A_i}(\varepsilon_i) \times \mathcal{G}^{1,p}(P) \to \mathcal{A}^{0,p}(P), \quad (A_i + \alpha, g) \mapsto (g^{-1})^*(A_i + \alpha)$$

is a diffeomorphism onto its image. Hence

$$\alpha_i \colon \operatorname{im} \mathfrak{m}_i \to L^p(\Sigma, \operatorname{ad}(P)), \quad A \mapsto (\operatorname{pr}_1 \circ \mathfrak{m}^{-1})(A) - A_i$$

(with $\operatorname{pr}_1 \colon \mathcal{S}_{A_i}(\varepsilon_i) \times \mathcal{G}^{1,p}(P) \to \mathcal{S}_{A_i}(\varepsilon_i)$ denoting projection) is a well-defined smooth map with image being contained in $\mathcal{S}_{A_i}(\varepsilon_i) - A_i$. We extend α_i to a map on $\mathcal{A}(P)$ by setting $\alpha_i(A) = 0$ for $A \in \mathcal{A}^{0,p}(P) \setminus \operatorname{im} \mathfrak{m}_i$. Hence

$$\mathcal{V}_\ell^- \colon \mathcal{A}(P) \to \mathbb{R}, \qquad A \mapsto \rho_k(\|\alpha_i(A)\|^2_{L^2(\Sigma)}) \langle \eta_{ij} + \alpha_i(A), \eta_{ij} \rangle \qquad (8)$$

is a well-defined map for every triple $\ell = (i, j, k) \in \mathbb{N}^3$. Note also that $\mathcal{V}_\ell^-$ is invariant under the action of $\mathcal{G}^{1,p}(P)$ by gauge transformations.

Proposition 2.2 *For every $A \in \operatorname{im} \mathfrak{m}_i$ there exists a unique $g \in \mathcal{G}^{1,p}(P)$ such that $g^*A - A_i = \alpha_i(A)$ and $d^*_{A_i} \alpha_i(A) = 0$.*

Proof: For a proof we refer to [19, Proposition 2.6]. $\qquad\qquad\square$

10

Proposition 2.3 *Let $\ell = (i, j, k) \in \mathbb{N}^3$ such that $k > \frac{10}{\delta_i}$. Then the map $\mathcal{V}_\ell^- : \mathcal{A}^{0,p}(P) \to \mathbb{R}$ defined in (8) is smooth.*

Proof: For a proof we refer to [19, Proposition 2.7]. □

Proposition 2.4 *Let $A_0 \in \mathcal{A}(P)$ and $p > 2$. There exist constants $c(A_0)$, $c(A_0, p)$ and $\delta(A_0, p)$ such that the estimates*

$$\|\alpha(A)\|_{W^{1,p}(\Sigma)} \le c(A_0, p)\big(1 + \|F_A\|_{L^p(\Sigma)}\big),$$
$$\|\nabla\mathcal{V}^-(A)\|_{C^0(\Sigma)} \le c(A_0)\big(1 + \|F_A\|_{L^3(\Sigma)}\big),$$
$$\|d_A\nabla\mathcal{V}^-(A)\|_{L^p(\Sigma)} \le c(A_0, p)\big(1 + \|F_A\|_{L^p(\Sigma)} + \|\alpha(A)\|_{L^{2p}(\Sigma)}^2\big)$$

are satisfied for all $A \in \mathcal{A}^{0,p}(P)$ with $\|\alpha(A)\|_{L^2(\Sigma)} < \delta(A_0, p)$.

Proof: For a proof we refer to [19, Proposition A.4]. □

Remark 2.5 In the following we shall admit only those perturbations $\mathcal{V}_\ell^-$ which meet the assumptions of Propositions 2.3 and 2.4. These are precisely satisfied for triples $\ell = (i, j, k) \in \mathbb{N}^3$ such that $k > \max\{\frac{10}{\delta_i}, \frac{2}{\delta(A_i, p)}\}$ where δ_i denotes the constant of Proposition 2.3, and $\delta(A_i, p)$ is as in Proposition 2.4. Moreover, the Cauchy–Schwarz inequality implies that $\langle \eta_{ij} + \alpha(A), \eta_{ij} \rangle \ge 0$ if $\|\alpha(A)\|_{L^2(\Sigma)} \le \|\eta_j\|_{L^2(\Sigma)}$, and hence the map $\mathcal{V}_\ell^-$ is non-negative for sufficiently large indices k. For the remainder of this article we allow only for triples $(i, j, k) \in \mathbb{N}^3$ such that k satisfies these conditions and renumber the subset of such triples by integers $\ell \in \mathbb{N}$.

Given $\ell \in \mathbb{N}$, we fix a constant $C_\ell > 0$ such that the following conditions are satisfied.
  (i) $\sup_{A \in \mathcal{A}(P)} |\mathcal{V}_\ell^-(A)| \le C_\ell$,
  (ii) $\sup_{A \in \mathcal{A}(P)} \|\nabla\mathcal{V}_\ell^-(A)\|_{L^2(\Sigma)} \le C_\ell$,
  (iii) $\|\nabla\mathcal{V}_\ell^-(A)\|_{C^0(\Sigma)} \le C_\ell(1 + \|F_A\|_{L^3(\Sigma)})$ for all $A \in \mathcal{A}(P)$.
  (iv) $\|H_A\mathcal{V}_\ell^-\beta\|_{L^p(\Sigma)} \le C_\ell(1 + \|F_A\|_{L^3(\Sigma)})\|\beta\|_{L^p(\Sigma)}$ for all $A \in \mathcal{A}(P)$, $\beta \in \Omega^1(\Sigma, \mathrm{ad}(P))$, and $1 < p < \infty$.
Here the notation $\nabla\mathcal{V}_\ell^- : \mathcal{A}(P) \to \Omega^1(\Sigma, \mathrm{ad}(P))$ and $H_A\mathcal{V}_\ell^- : \Omega^1(\Sigma, \mathrm{ad}(P)) \to \Omega^1(\Sigma, \mathrm{ad}(P))$ refers to the L^2 gradient and Hessian (at the point $A \in \mathcal{A}(P)$) of the map $\mathcal{V}_\ell^-$. The existence of the constant C_ℓ has been shown in [19,

11

Proposition A.6]. The *universal space of perturbations* is the normed linear space

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

It is a separable Banach space isomorphic to the space ℓ^1 of absolutely summable real sequences. Further relevant properties of the perturbations $\mathcal{V}_\ell^-$ are discussed in [19, Appendix A].

Perturbations of the loop group energy functional

We shall follow closely Salamon and Weber [16] in our construction of a Banach space Y^+ of perturbations of the loop group energy functional $\mathcal{E}$. Let us fix the following data.
  (i) A dense sequence $(x_i)_{i \in \mathbb{N}}$ of points in ΩG.
  (ii) For each $i \in \mathbb{N}$ a dense sequence $(\eta_{ij})_{j \in \mathbb{N}}$ in $T_{x_i}(\Omega G)$.
  (iii) A smooth cut-off function $\rho \colon \mathbb{R} \to [0,1]$ supported in $[-4, 4]$, and satisfying $\rho = 1$ on $[-1, 1]$ and $\|\rho'\|_{L^\infty(\mathbb{R})} < 1$. For $k \in \mathbb{N}$, set $\rho_k(r) := \rho(k^2 r)$.
   Denote by $\iota > 0$ the injectivity radius of the compact Riemannian manifold G. Fix a further cut-off function β supported in $[-\iota^2, \iota^2]$ such that $\beta = 1$ on $\left[-\frac{\iota^2}{4}, \frac{\iota^2}{4} \right]$. For $x_i \in \Omega G$ as in (i) above and $q \in G$ within distance ι of $x_i(t)$, let $\xi_q^i(t) \in T_{x_i(t)} G$ be uniquely determined by $q = \exp_{x_i(t)} \xi_q^i(t)$. For each multiindex $\ell = (i, j, k) \in \mathbb{N}^3$ we define the smooth map

$$\mathcal{V}_\ell^+ \colon \Omega G \to \mathbb{R}, \quad x \mapsto \rho_k \left(\|x - x_i\|_{L^2(S^1)}^2 \right) \int_0^1 V_{ij}(t, x(t)) \, dt,$$

where

$$V_{ij}(t, q) := \begin{cases} \beta(|\xi_q^i(t)|^2) \langle \eta_{ij}(t) + \xi_q^i(t), \eta_{ij}(t) \rangle, & \text{if } |\xi_q^i(t)| < \iota, \\ 0, & \text{else.} \end{cases}$$

   The L^2 distance appearing in the argument of ρ_k above refers to the L^2 distance induced after isometrically embedding the manifold G in some ambient euclidian space $\mathbb{R}^N$. Note that $\mathcal{V}_\ell^+$ extends uniquely to a map $\mathcal{V}_\ell^+ \colon \Lambda G \to \mathbb{R}$ on the free loop group, which is invariant under the free action $h \cdot x \mapsto hx$ of G on ΛG.

Remark 2.6 Because $\langle \eta_{ij} + \xi_q^i, \eta_{ij} \rangle \geq 0$ if $\|\xi_q^i\|_{L^2(S^1)} \leq \|\eta_{ij}\|_{L^2(S^1)}$, it follows that the map $\mathcal{V}_\ell$ is non-negative for sufficiently large indices k (for given pair

(i, j)). We henceforth consider only those multiindices $\ell = (i, j, k)$ which satisfy this condition, and renumber the set of such triples (i, j, k) by integers $\ell \in \mathbb{N}$.

Let Y^+ denote the real vector space spanned by the maps $\mathcal{V}_\ell^+$, $\ell \in \mathbb{N}$. We endow it with a norm analogously to that of the space Y^- (cf. the previous section), turning it into a separable Banach space isomorphic to the space ℓ^1 of absolutely summable real sequences (cf. [22, Section 7.1] for details).

2.3   Critical manifolds

We introduce some further notation concerning the set of critical points of the functionals $\mathcal{YM}$ and $\mathcal{E}$. Let

$$\operatorname{crit}(\mathcal{YM}) := \{A \in \mathcal{A}^{1,p}(P) \mid d_A^* F_A = 0\}$$

denote the set of critical points of $\mathcal{YM}$, the equation $d_A^* F_A = 0$ being understood in the weak sense. Similarly, the notation $\operatorname{crit}(\mathcal{YM}^{\mathcal{V}^-})$ refers to the set of critical points of the perturbed Yang–Mills functional. We furthermore let

$$\operatorname{crit}(\mathcal{E}) := \{x \in W^{1,p}(S^1, G) \mid \nabla_t \partial_t x = 0\}$$

denote the set of critical points of $\mathcal{E}$, where we again interpret the equation $\nabla_t \partial_t x = 0$ in the weak sense. The set of critical points of the perturbed energy functional is denoted by $\operatorname{crit}(\mathcal{E}^{\mathcal{V}^+})$. We subsequently make use of the holonomy map $\Phi \colon \mathcal{A}(P) \to \Lambda G$ as given by (40). A discussion of its properties is postponed to A. Let $\widehat{\mathcal{CR}}(\mathcal{YM})$ and $\widehat{\mathcal{CR}}(\mathcal{E})$ denote the set of connected components of $\operatorname{crit}(\mathcal{YM})$, respectively of $\operatorname{crit}(\mathcal{E})$. The group $\mathcal{G}_0(P)$ of based gauge transformations acts freely on $\mathcal{A}(P)$, hence on $\operatorname{crit}(\mathcal{YM})$. We define $\mathcal{CR}(\mathcal{YM})$ as the set of connected components of $\operatorname{crit}(\mathcal{YM})/\mathcal{G}_0(P)$ in $\mathcal{A}(P)/\mathcal{G}_0(P)$. It is a standard fact that every such connected component is a finite-dimensional submanifold of $\mathcal{A}(P)/\mathcal{G}_0(P)$ diffeomorphic to some homogeneous space G/H. Likewise, the group G acts freely on ΛG and on $\widehat{\mathcal{CR}}(\mathcal{E})$. We denote by $\mathcal{CR}(\mathcal{E}) = \widehat{\mathcal{CR}}(\mathcal{E})/G$ the set of connected components of $\operatorname{crit}(\mathcal{E})/G$ in $\Lambda G/G \cong \Omega G$. Furthermore, the holonomy map Φ induces a bijection between $\mathcal{CR}(\mathcal{YM})$ and $\mathcal{CR}(\mathcal{E})$ which preserves the action filtration given on both these sets, cf. Theorem A.2. For the remainder of this section we fix a regular value a of $\mathcal{YM}$ and put $b := 4a/\pi$, which is a regular value of $\mathcal{E}$. Furthermore, we denote

$$\operatorname{crit}^a(\mathcal{YM}) := \{A \in \operatorname{crit}(\mathcal{YM}) \mid \mathcal{YM}(A) \le a\},$$
$$\operatorname{crit}^b(\mathcal{E}) := \{x \in \operatorname{crit}(\mathcal{E}) \mid \mathcal{E}(x) \le b\}.$$

13

We introduce the notation $\widehat{\mathcal{CR}}^a(\mathcal{YM})$, $\mathcal{CR}^a(\mathcal{YM})$, $\widehat{\mathcal{CR}}^b(\mathcal{E})$, and $\mathcal{CR}^b(\mathcal{E})$ for the intersection of $\widehat{\mathcal{CR}}(\mathcal{YM})$ etc. as defined above with the sublevel set $\mathrm{crit}^a(\mathcal{YM})$, respectively $\mathrm{crit}^b(\mathcal{E})$. As a consequence of Theorem A.5 the holonomy map Φ induces a bijection between $\widehat{\mathcal{CR}}^a(\mathcal{YM})$ and $\mathcal{CR}^b(\mathcal{E})$. For each critical manifold $\hat{\mathcal{C}} \in \widehat{\mathcal{CR}}^a(\mathcal{YM})$ we fix a closed L^2 neighborhood $U_{\hat{\mathcal{C}}}$ of $\hat{\mathcal{C}}$ such that $U_{\hat{\mathcal{C}}_1} \cap U_{\hat{\mathcal{C}}_2} = \emptyset$ whenever $\hat{\mathcal{C}}_1 \neq \hat{\mathcal{C}}_2$. Because the set $\widehat{\mathcal{CR}}^a(\mathcal{YM})$ is finite (cf. [19] for a proof) it follows that such a choice is possible. We next choose for each $\mathcal{C} \in \mathcal{CR}^b(\mathcal{E})$ a sufficiently small closed L^2 neighborhood $U_{\mathcal{C}}$ of $\mathcal{C}$ such that $U_{\mathcal{C}_1} \cap U_{\mathcal{C}_2} = \emptyset$ if $\mathcal{C}_1 \neq \mathcal{C}_2$, and $\Phi(U_{\hat{\mathcal{C}}}) \cap U_{\mathcal{C}} = \emptyset$ for all $\hat{\mathcal{C}} \in \widehat{\mathcal{CR}}^a(\mathcal{YM})$ with $\Phi(\hat{\mathcal{C}}) \neq \mathcal{C}$.

Definition 2.7 Let a be the regular value of $\mathcal{YM}$ as fixed above. We call a perturbation $\mathcal{V}^- = \sum_{\ell=1}^{\infty} \lambda_\ell \mathcal{V}_\ell^- \in Y^-$ *a-admissible* if it satisfies

$$\mathrm{supp}\,\mathcal{V}_\ell^- \cap U_{\hat{\mathcal{C}}} \neq \emptyset \quad \text{for some} \quad \hat{\mathcal{C}} \in \widehat{\mathcal{CR}}^a(\mathcal{YM}) \quad \Longrightarrow \quad \lambda_\ell = 0.$$

For the regular value $b = 4a/\pi$ of $\mathcal{E}$ we analogously define the subspace of *b-admissible* perturbations of Y^+. We let $Y_a^- \times Y_b^+ =: Y_a \subseteq Y$ denote the subspace of pairs $(\mathcal{V}^-, \mathcal{V}^+)$ where $\mathcal{V}^-$ is a-admissible and $\mathcal{V}^+$ is b-admissible.

It is straightforward to show that the spaces Y_a^-, respectively Y_b^+ are closed subspaces of the Banach spaces $Y^\pm$, and hence Y_a is a closed subspace of Y. The following proposition shows that adding a small a-admissible perturbation to $\mathcal{YM}$ leaves the set of critical points below level a unchanged.

Proposition 2.8 *Let a be the regular value of $\mathcal{YM}$ as fixed above. Then there is a constant $\delta = \delta(a) > 0$ with the following significance. Assume $\mathcal{V}^- \in Y_a^-$ and $\|\mathcal{V}^-\| < \delta$. Then it holds,*

$$\mathrm{crit}(\mathcal{YM}^{\mathcal{V}^-}) \cap \{A \in \mathcal{A}^{1,p}(P) \mid \mathcal{YM}(A) < a\}$$
$$= \mathrm{crit}(\mathcal{YM}) \cap \{A \in \mathcal{A}^{1,p}(P) \mid \mathcal{YM}(A) < a\}.$$

Proof: For a proof we refer to [19, Proposition 2.9]. $\qquad\square$

The analogous statement holds true for the energy functional $\mathcal{E}$.

Proposition 2.9 *Let b be the regular value of $\mathcal{E}$ as fixed above. Then there is a constant $\delta = \delta(b) > 0$ with the following significance. Assume $\mathcal{V}^+ \in Y_b^+$ and $\|\mathcal{V}^+\| < \delta$. Then,*

$$\mathrm{crit}(\mathcal{E}^{\mathcal{V}^+}) \cap \{x \in \Lambda G \mid \mathcal{E}(x) < b\} = \mathrm{crit}(\mathcal{E}) \cap \{x \in \Lambda G \mid \mathcal{E}(x) < b\}.$$

Proof: For a proof we refer to [22, Lemma 7.4]. $\qquad\square$

2.4 Gradient flow lines

In this section we fix perturbations $\mathcal{V}^\pm \in Y^\pm$. As discussed in [19] it is convenient to introduce into the $\mathcal{G}(P)$-invariant Yang–Mills gradient flow equation an additional *gauge fixing term* $-d_A\Psi$, $\Psi \in \Omega^0(\Sigma, \mathrm{ad}(P))$. Solutions of this new flow equation are then invariant under *time-dependent* gauge transformations. This point requires to introduce some further notation. For an interval $I \subseteq \mathbb{R}$ we denote by $\hat{P}_I := I \times P$ the trivial extension of the principle G-bundle P to the base manifold $I \times \Sigma$, and set $\hat{P} := \hat{P}_{\mathbb{R}^-}$. We let $\mathcal{G}(\hat{P})$ denote the group of smooth gauge transformations of the principle G-bundle $\hat{P}$ and call this the group of time-dependent gauge-transformations (and similarly for $\mathcal{G}(\hat{P}_I)$). A pair $(A, \Psi) \in C^\infty(\mathbb{R}, \mathcal{A}(P) \times \Omega^0(\Sigma, \mathrm{ad}(P)))$ can naturally be identified with the connection $\mathbb{A} = A + \Psi\, ds \in \mathcal{A}(\hat{P})$. The action of the group $\mathcal{G}(\hat{P})$ on $\mathcal{A}(\hat{P})$ by gauge transformations is given by

$$g^*\mathbb{A} = g^*A + (g^{-1}\Psi g + g^{-1}\partial_s g) \wedge ds. \tag{9}$$

Definition 2.10 Let $\mathcal{V} \in Y$ be a perturbation. The $\mathcal{G}(\hat{P})$-*invariant, perturbed Yang–Mills gradient flow* is the nonlinear PDE

$$0 = \partial_s A + d_A^* F_A - d_A\Psi + \nabla\mathcal{V}^-(A) \tag{10}$$

for connections $\mathbb{A} = A + \Psi\, ds \in \mathcal{A}(\hat{P})$.

One easily checks that with $\mathbb{A} \in \mathcal{A}(\hat{P})$, also $g^*\mathbb{A}$ is a solution of (10), for every $g \in \mathcal{G}(\hat{P})$.

Definition 2.11 The *perturbed loop group gradient flow equation* is the nonlinear PDE

$$0 = \partial_s x - \nabla_t \partial_t x + \nabla\mathcal{V}^+(x) \tag{11}$$

for a smooth path $x\colon s \mapsto x(s) \in \Lambda G$ of free loops.

Equation (11) is clearly invariant under the action of the group G on ΛG via $(h\cdot x)(t) = hx(t)$. The proper analytical setup for a study of the perturbed Yang–Mills and loop group gradient flow equations will be introduced in Section 3.1.

2.5 Morse homologies for the Yang–Mills and heat flows

Morse homology theories based on the perturbed Yang–Mills and loop group gradient flows have been developed by Weber, respectively the author in [22, 19]. We here only give a brief description of both these Morse homology theories. Let $a \geq 0$ be a regular value of $\mathcal{YM}$. We fix an a-admissible perturbation $\mathcal{V}^- \in Y_a^-$ (cf. Definition 2.7). Let $h \colon \mathrm{crit}^a(\mathcal{YM})/\mathcal{G}_0(P) \to \mathbb{R}$ be a smooth Morse–Smale function (with respect to some fixed Riemannian metric on the finite-dimensional smooth manifold $\mathrm{crit}^a(\mathcal{YM})/\mathcal{G}_0(P)$). We let

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

denote the complex generated as a $\mathbb{Z}_2$ vector space by the set $\mathrm{crit}(h)$ of critical points of h. To each $x \in \mathrm{crit}(h)$ we assign the index $\mathrm{Ind}(x)$ to be the sum of the Morse indices of x as a critical point of $\mathcal{YM}$ and of the function h. For $x^-, x^+ \in \mathrm{crit}(h)$ we call the set $\mathcal{M}(x^-, x^+)$ as in [19, Section 8.2] the moduli space of Yang–Mills gradient flow lines with cascades from x^- to x^+.

Lemma 2.12 *For generic, a-admissible perturbation $\mathcal{V}^- \in Y_a^-$, Morse function h, and all $x^-, x^+ \in \mathrm{crit}(h)$, the set $\mathcal{M}(x^-, x^+)$ is a smooth manifold (with boundary) of dimension*

$$\dim \mathcal{M}(x^-, x^+) = \mathrm{Ind}(x^-) - \mathrm{Ind}(x^+) - 1.$$

Proof: For a proof we refer to [19, Lemma 8.3]. □

For $k \in \mathbb{N}$ we define the Morse boundary operator

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

to be the linear extension of the map

$$\partial_k^{\mathcal{YM}} x := \sum_{\substack{x' \in \mathrm{crit}(h) \\ \mathrm{Ind}(x') = k-1}} n(x, x') x', \tag{12}$$

where $x \in \mathrm{crit}(h)$ is a critical point of index $\mathrm{Ind}(x) = k$. The numbers $n(x, x')$ are given by counting modulo 2 the flow lines with cascades (with respect to $\mathcal{YM}^{\mathcal{V}^-}$ and h) from x to x', i.e.

$$n(x, x') := \# \mathcal{M}(x^-, x^+) \pmod 2.$$

Theorem 2.13 (Yang–Mills Morse homology) *Let $a \geq 0$ be a regular value of $\mathcal{YM}$. For any Morse function $h\colon \mathrm{crit}^a(\mathcal{YM})/\mathcal{G}_0(P) \to \mathbb{R}$ and generic, a-admissible perturbation $\mathcal{V}^- \in Y_a^-$, the map $\partial_*^{\mathcal{YM}}$ satisfies $\partial_k^{\mathcal{YM}} \circ \partial_{k+1}^{\mathcal{YM}} = 0$ for all $k \in \mathbb{N}_0$ and thus there exist well-defined homology groups*

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

This Yang–Mills Morse homology is independent of the choice of a-admissible perturbation $\mathcal{V}^-$ and Morse function h.

Proof: For a proof we refer to [19, Theorem 1.1]. □

Weber's heat flow homology for the loop space ΩM of a closed manifold M is based on a similar construction of a chain complex and a boundary operator. One of his main results is the following theorem (which he only states for the case where the action functional $\mathcal{E}^{\mathcal{V}}$ is Morse, the adaption to the present case of a Morse–Bott situation being straight-forward).

Theorem 2.14 (Heat flow homology) *Let $b \geq 0$ be a regular value of $\mathcal{E}^{\mathcal{V}}$. For any Morse function $h\colon \mathrm{crit}^b(\mathcal{E})/G \to \mathbb{R}$ and generic, b-admissible perturbation $\mathcal{V}^+ \in Y_b^+$, the map $\partial_*^{\mathcal{E}}$ (defined in analogy to (12)) satisfies $\partial_k^{\mathcal{E}} \circ \partial_{k+1}^{\mathcal{E}} = 0$ for all $k \in \mathbb{N}_0$ and thus there exist well-defined homology groups*

$$HM_k^b(\Omega M, \mathcal{V}^+, h) = \frac{\ker \partial_k^{\mathcal{E}}}{\mathrm{im}\, \partial_{k+1}^{\mathcal{E}}}.$$

*The homology $HM_*^b(\Omega M, \mathcal{V}^+, h)$ is called **heat flow homology**. It is independent of the choice of b-admissible perturbation $\mathcal{V}^+$ and Morse function h.*

Proof: For a proof we refer to [22, Theorem 1.14]. □

2.6 Hybrid moduli spaces

The main idea in comparing Yang–Mills Morse homology to heat flow homology consists in the following construction of a so-called *hybrid moduli*

space. For abbreviation we set $A_0 := A(0)$, $x_0 := x(0)$, and $x_{A_0} := \Phi(A(0))$. For critical manifolds $\hat{\mathcal{C}}^- \in \widehat{\mathcal{CR}}(\mathcal{YM})$ and $\hat{\mathcal{C}}^+ \in \widehat{\mathcal{CR}}(\mathcal{E})$ let us define

$$\hat{\mathcal{M}}(\hat{\mathcal{C}}^-, \hat{\mathcal{C}}^+) := \tag{13}$$
$$\big\{ (A, \Psi, x) \in C^\infty(\mathbb{R}^-, \mathcal{A}(P) \times \Omega^0(\Sigma, \mathrm{ad}(P))) \times C^\infty(\mathbb{R}^+, \Lambda G) \mid$$
$$(A, \Psi) \text{ satisfies (10)}, \quad x \text{ satisfies (11)}, \quad x_0 = h x_{A_0} \text{ for some } h \in G,$$
$$\lim_{s \to -\infty} (A(s), \Psi(s)) = (A^-, 0) \in \hat{\mathcal{C}}^- \times \Omega^0(\Sigma, \mathrm{ad}(P)), \lim_{s \to +\infty} x(s) = x^+ \in \hat{\mathcal{C}}^+ \big\}.$$

By construction, the group $\mathcal{G}(\hat{P}) \times G$ acts freely on the space $\hat{\mathcal{M}}(\hat{\mathcal{C}}^-, \hat{\mathcal{C}}^+)$. Let us denote $\mathcal{C}^- := \hat{\mathcal{C}}^-/\mathcal{G}_0(P)$ and $\mathcal{C}^- := \hat{\mathcal{C}}^+/G$. The moduli space we shall study further on is the quotient

$$\mathcal{M}(\mathcal{C}^-, \mathcal{C}^+) := \frac{\hat{\mathcal{M}}(\hat{\mathcal{C}}^-, \hat{\mathcal{C}}^+)}{\mathcal{G}(\hat{P}) \times G}. \tag{14}$$

We show subsequently that $\hat{\mathcal{M}}(\hat{\mathcal{C}}^-, \hat{\mathcal{C}}^+)$ arises as the zero set $\mathcal{F}^{-1}(0)$ of an equivariant (with respect to $\mathcal{G}(\hat{P}) \times G$) section $\mathcal{F}$ of a suitably defined Banach space bundle $\mathcal{E}$ over a Banach manifold $\mathcal{B}$. After proving that the horizontal differential $d_x \mathcal{F}$ at any such zero $x \in \mathcal{F}^{-1}(0)$ is a surjective Fredholm operator, it will follow from the implicit function theorem that the moduli space $\mathcal{M}(\mathcal{C}^-, \mathcal{C}^+)$ is a finite-dimensional smooth manifold.

3 Fredholm theory and transversality

3.1 The nonlinear setup

The Banach manifolds we shall use are modeled on weighted Sobolev spaces in order to make the Fredholm theory work. To define these, we choose numbers $\delta > 0$ and $p > 3$, and a smooth function β such that $\beta(s) = -1$ if $s < 0$ and $\beta(s) = 1$ if $s > 1$. We define the δ-weighted (k, p)-Sobolev norm (for $1 \leq p \leq \infty$ and an integer $k \geq 0$) of a measurable function (respectively, a measurable section of a vector bundle) u over $\mathbb{R}^- \times \Sigma$ or $\mathbb{R}^+ \times S^1$ to be the usual (k, p)-Sobolev norm of the function (or section) $e^{\delta \beta(s) s} u$.

Recall the definition of parabolic Sobolev spaces at the end of Section 2.1. We fix numbers $\delta > 0$, $p > 3$. Let $\mathcal{A}_\delta^{1,2;p}(P)$ denote the space of time-dependent connections on P which are locally of class $W^{1,2;p}$ and for which there exists a limiting connection $A^- \in \mathrm{crit}(\mathcal{YM})$ and a number $T^- \leq 0$ such that the time-dependent 1-form $\alpha^- := A - A^-$ satisfies

$$\alpha^- \in W_\delta^{1,p}((-\infty, T^-], L^p(\Sigma, \mathrm{ad}(P))) \cap L_\delta^p((-\infty, T^-], W^{2,p}(\Sigma, \mathrm{ad}(P))). \tag{15}$$

18

Similarly, let $\mathcal{G}_\delta^{2,p}(\hat{P})$ denote the group of gauge transformations of $\hat{P}$ which are locally of class $W^{2,p}$ and in addition satisfy the following two conditions. First, the time-dependent $\mathrm{ad}(P)$-valued 1-form $g^{-1}dg$ satisfies

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

(this condition being necessary to make sure that $g^*A \in L_\delta^p(\mathbb{R}, \mathcal{A}^{2,p}(P))$ for every A with this property). Second, there exists a limiting gauge transformation $g^- \in \mathcal{G}_0^{2,p}(P)$, a number $T^- \leq 0$, and a bundle valued 0-form

$$\gamma^- \in W_\delta^{2,p}((-\infty, T^-] \times \Sigma, \mathrm{ad}(\hat{P}_{(-\infty,T^-]}))$$

with $g(s) = g^- \exp(\gamma^-(s))$ for $s \leq T^-$.

Let $\hat{\mathcal{C}}^- \in \widehat{\mathcal{CR}}(\mathcal{YM})$ be a critical manifold. We denote by $\hat{\mathcal{B}}^- := \hat{\mathcal{B}}^-(\hat{\mathcal{C}}^-, \delta, p)$ the Banach manifold of pairs

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

such that $\lim_{s \to -\infty} A(s) = A^-$ holds in the sense of (15) for some $A^- \in \hat{\mathcal{C}}^-$. We identify such pairs as before with connections $\mathbb{A} = A + \Psi\,ds$ on $\hat{P}$. The action of the group $\mathcal{G}_\delta^{2,p}(\hat{P})$ on $\hat{\mathcal{B}}$ by gauge transformations as in (9) is smooth. It is free by our requirement that $\lim_{s \to -\infty} g(s) = g^-$ be a based gauge transformation. The resulting quotient space

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

is again a smooth Banach manifold. Let $\hat{\mathcal{C}}^+ \in \widehat{\mathcal{CR}}(\mathcal{E})$. We define $\hat{\mathcal{B}}^+(\hat{\mathcal{C}}^+, \delta, p)$ to be the Banach manifolds of maps $x \colon \mathbb{R}^+ \times S^1 \to G$ such that the condition

$$x - x^+ \in L_\delta^p(\mathbb{R}^+, W^{2,p}(S^1, G)) \cap W_\delta^{1,p}(\mathbb{R}^+, L^p(S^1, G))$$

is satisfied for some $x^+ \in \hat{\mathcal{C}}^+$. To make sense of the difference $x - x^+$ and of the Sobolev spaces involved in this definition, we consider the Lie group G as being isometrically embedded in some euclidian space $\mathbb{R}^N$. Now put

$$\mathcal{B}^+ := \mathcal{B}^+(\mathcal{C}^+, \delta, p) := \frac{\hat{\mathcal{B}}^+(\hat{\mathcal{C}}^+, \delta, p)}{G},$$

and define $\mathcal{B} := \mathcal{B}^- \times \mathcal{B}^+$. We define the Banach space bundle $\mathcal{E} = \mathcal{E}(\mathcal{C}^-, \mathcal{C}^+, \delta, p)$ over $\mathcal{B}$ in the following way. Let $\hat{\mathcal{E}}^-$ be the trivial Banach space bundle over

$\hat{\mathcal{B}}^-$ with fibres $\hat{\mathcal{E}}^-_{(A,\Psi)} := L^p_\delta(\mathbb{R}^-, L^p(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P)))$, and $\hat{\mathcal{E}}^+$ be the trivial Banach space bundle over $\hat{\mathcal{B}}^+$ with fibres $\hat{\mathcal{E}}^+_x := L^p_\delta(\mathbb{R}^+, L^p(S^1, \mathfrak{g}))$. We set

$$\hat{\mathcal{E}} := \hat{\mathcal{E}}(\hat{\mathcal{C}}^-, \hat{\mathcal{C}}^+, \delta, p) := \hat{\mathcal{E}}^- \times \hat{\mathcal{E}}^+ \times \Lambda G.$$

The action of $\mathcal{G}^{2,p}_\delta(\hat{P}) \times G$ on $\hat{\mathcal{B}}^- \times \hat{\mathcal{B}}^+$ lifts to a free action on $\hat{\mathcal{E}}^- \times \hat{\mathcal{E}}^+ \times \Lambda G$ via

$$(g,h) \cdot (A, \Psi, \alpha, x, \xi, x_1) := (g^*A, g^{-1}\Psi g + g^{-1}\partial_s g, g^{-1}\alpha g, hx, \xi, hx_1).$$

Let $\mathcal{E}$ denote the respective quotient space and define the smooth section $\mathcal{F}: \mathcal{B} \to \mathcal{E}$ of $\mathcal{E}$ by

$$\mathcal{F}: [(A, \Psi, x)] \mapsto \left[\begin{pmatrix} \partial_s A + d^*_A F_A - d_A \Psi + \nabla \mathcal{V}^-(A) \\ x^{-1}(\partial_s x - \nabla_t \partial_t x + \nabla \mathcal{V}^+(x)) \\ x(0)x^{-1}_{A(0)} \end{pmatrix} \right]. \tag{16}$$

3.2 Yang–Mills Hessian and linearized Yang–Mills gradient flow

For $A \in \mathcal{A}(P)$, we let $\mathcal{H}_A$ denote the *augmented Yang–Mills Hessian* defined by

$$\mathcal{H}_A := \begin{pmatrix} d^*_A d_A + *[*F_A \wedge \cdot] + H_A \mathcal{V}^- & -d_A \\ -d^*_A & 0 \end{pmatrix}:$$
$$\Omega^1(\Sigma, \mathrm{ad}(P)) \oplus \Omega^0(\Sigma, \mathrm{ad}(P)) \to \Omega^1(\Sigma, \mathrm{ad}(P)) \oplus \Omega^0(\Sigma, \mathrm{ad}(P)). \tag{17}$$

Here $H_A \mathcal{V}^-$ denotes the Hessian of the map $\mathcal{V}^-: \mathcal{A}(P) \to \mathbb{R}$. In order to find a domain for $\mathcal{H}_A$ which makes the subsequent Fredholm theory work, we fix an irreducible smooth reference connection $A_0 \in \mathcal{A}(P)$ 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

$$\Omega^1(\Sigma, \mathrm{ad}(P)) = \ker \left(d^*_{A_0}: \Omega^1(\Sigma, \mathrm{ad}(P)) \to \Omega^0(\Sigma, \mathrm{ad}(P)) \right)$$
$$\oplus \, \mathrm{im} \left(d_{A_0}: \Omega^0(\Sigma, \mathrm{ad}(P)) \to \Omega^1(\Sigma, \mathrm{ad}(P)) \right).$$

Let $W^{2,p}_0$ and $W^{1,p}_1$ denote the completions of the first component, respectively of the second component with respect to the (k,p)-Sobolev norm ($k = 1, 2$). We set $\mathcal{W}^p(\Sigma) := W^{2,p}_0 \oplus W^{1,p}_1$ and endow this space with the sum norm. It was shown in [19, Proposition 5.1] that it is independent of the

choice of reference connection A_0. We let $p > 1$ and consider the augmented Yang–Mills Hessian as an operator

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

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

$$\mathrm{dom}\,\mathcal{H}_A := \mathcal{W}^2(\Sigma) \oplus W^{1,2}(\Sigma, \mathrm{ad}(P)). \tag{18}$$

It was shown in [19, Proposition 5.2] that it is self-adjoint.

Next we consider the linearization of the Yang–Mills gradient flow (10). Since every solution (A, Ψ) of the Yang–Mills gradient flow is gauge equivalent under $\mathcal{G}(\hat{P})$ to a solution satisfying $\Psi \equiv 0$ (cf. [18, Proposition 3.3]), it suffices to consider the linearization along such trajectories only. We define for $p > 1$ the Banach spaces

$$\mathcal{Z}^{\delta,p,-} := \big(W^{1,p}_\delta(\mathbb{R}, L^p(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))) \cap L^p_\delta(\mathbb{R}, \mathcal{W}^p(\Sigma))\big) \oplus W^{1,p}_\delta(\mathbb{R} \times \Sigma, \mathrm{ad}(P)),$$

$$\mathcal{L}^{\delta,p,-} := L^p_\delta(\mathbb{R}^- \times \Sigma, T^*\Sigma \otimes \mathrm{ad}(P)) \oplus L^p_\delta(\mathbb{R}^- \times \Sigma, \mathrm{ad}(P)),$$

where the number $\delta > 0$ refers to the weight function fixed at the beginning of Section 3.1. In the following we shall be concerned with the linear operator

$$\mathcal{D}_A := \frac{d}{ds} + \mathcal{H}_A \colon \mathcal{Z}^{\delta,p,-} \to \mathcal{L}^{\delta,p,-} \tag{19}$$

for a smooth path $s \mapsto A(s) \in \mathcal{A}(P)$, where $s \in \mathbb{R}^-$. It arises as the linearization of the Yang–Mills gradient flow (10) at a solution $(A, \Psi) = (A, 0)$. Some of its properties are collected in B.

3.3 Linearized loop group gradient flow

We discuss the linearized loop group gradient flow, following Weber [22]. The Hessian of the energy functional $\mathcal{E}$ at the loop $x \in \Lambda G$ is the linear operator

$$H_x \colon \xi \mapsto \nabla_t \nabla_t \xi + R(\xi, \partial_t x)\partial_t x + H_x \mathcal{V}^+ \xi$$

for vector fields ξ along x. Here R denotes the Riemannian curvature tensor of G, and $H_x \mathcal{V}^+$ denotes the Hessian of the map $\mathcal{V}^+ \colon \Lambda G \to \mathbb{R}$. Let $x \colon \mathbb{R}^+ \times S^1 \to G$ be a smooth map. We define for $p > 1$ the Banach spaces

$$\mathcal{Z}^{\delta,p,+} := W^{1,p}_\delta(\mathbb{R}^+, L^p(S^1, x^*TG)) \cap W^p_\delta(\mathbb{R}^+, W^{2,p}(S^1, x^*TG)),$$

$$\mathcal{L}^{\delta,p,+} := L^p_\delta(\mathbb{R}^+ \times S^1, x^*TG).$$

21

Note that these depend on x, which is suppressed in our notation. For short we will often drop x^*TG and simply write $L^p(S^1)$ etc. The number $\delta > 0$ refers to the weight function fixed at the beginning of Section 3.1. We denote

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

for a smooth path $s \mapsto x(s) \in \Lambda G$, where $s \in \mathbb{R}^+$. Note that the operator $\mathcal{D}_x$ arises as the linearization of the loop group gradient flow (11). We discuss some of its properties in B.

3.4 Linearized moduli space problem

We fix numbers $p > 3$, $\delta > 0$ and regular values $a > 0$ and $b := \frac{4a}{\pi}$ of $\mathcal{YM}$, respectively of $\mathcal{E}$. Throughout we let $\mathcal{V} = (\mathcal{V}^-, \mathcal{V}^+) \in Y_a$ be an a-admissible perturbation with $\|\mathcal{V}\|$ sufficiently small such that Propositions 2.8 and 2.9 apply. Let $\hat{\mathcal{C}}^- \in \mathcal{CR}^a(\mathcal{YM})$ and $\hat{\mathcal{C}}^+ \in \mathcal{CR}^b(\mathcal{E})$ be critical manifolds. Recall the definition of $\hat{\mathcal{M}}(\hat{\mathcal{C}}^-, \hat{\mathcal{C}}^+)$ in (13). For $u = (A, \Psi, x) \in \hat{\mathcal{M}}(\hat{\mathcal{C}}^-, \hat{\mathcal{C}}^+)$ we define the Banach spaces

$$\mathcal{Z}^{\delta,p} := \mathcal{Z}^{\delta,p,-} \oplus \mathcal{Z}^{\delta,p,+}, \qquad \mathcal{L}^{\delta,p} := \mathcal{L}^{\delta,p,-} \oplus \mathcal{L}^{\delta,p,+} \oplus L^2(S^1, \mathfrak{g}).$$

Put $A_0 := A(0)$ (and analogously for α_0, x_0, and ξ_0). We use the notation

$$D\Phi_{A_0} := \Phi(A_0)^{-1} \, d\Phi(A_0) \colon \Omega^1(\Sigma, \mathrm{ad}(P)) \to C^\infty(S^1, \mathfrak{g}),$$

where Φ is the holonomy map as in (40). We define the linear operator $\mathcal{D}_u \colon \mathcal{Z}^{\delta,p} \to \mathcal{L}^{\delta,p}$ by

$$(\alpha, \psi, \xi) \mapsto (\mathcal{D}_A(\alpha, \psi), \mathcal{D}_x \xi, N_A(\alpha, \xi)), \tag{20}$$

where we denote $N_A(\alpha, \xi) := x_0^{-1}\xi_0 - D\Phi_{A_0}\alpha_0$. The following remark clarifies the relation between the operator $\mathcal{D}_u$ and the linearization of the section $\mathcal{F}$ in (16).

Remark 3.1   (i) From the definition of $\mathcal{F}$ as a section $\mathcal{F} \colon \mathcal{B} \to \mathcal{E}$ (cf. (16)) it follows that its linearization $d\mathcal{F}(u)$, where $u = (A, \Psi, x)$, acts on the space of pairs (α, ψ, ξ) where $\alpha(s)$ converges exponentially to some $\alpha^- \in T_{A^-}\hat{\mathcal{C}}^-$ as $s \to -\infty$, and likewise $\xi(s) \to \xi^+ \in T_{x^+}\hat{\mathcal{C}}^+$ as $s \to \infty$. This asymptotic behaviour is in slight contrast to that required for elements of $\mathcal{Z}^{\delta,p}$, the domain of the operator $\mathcal{D}_u$. However, it is easy

to see that $d\mathcal{F}(u)$ is Fredholm if and only if this property holds for $\mathcal{D}_u$, and that the Fredholm indices are related via the formula

$$\operatorname{ind} d\mathcal{F}(u) = \operatorname{ind} \mathcal{D}_u + \dim \mathcal{C}^- + \dim \mathcal{C}^+.$$

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

(ii) The operator $\mathcal{D}_u$ arises as the linearization of the *unperturbed* Yang–Mills gradient flow equation (10). The Fredholm theory for general perturbations $\mathcal{V} \in Y$ can be reduced to the unperturbed case because the terms involving $\mathcal{V}$ contribute only compact perturbations to the operator $\mathcal{D}_u$.

Because the zero eigenspaces of the Hessians $\mathcal{H}_{A^-}$ and H_{x^+} are in general (i.e. if $\dim \mathcal{C}^\pm \geq 1$) non-trivial, we cannot directly refer to standard theorems on the spectral flow to prove Theorem 3.3 below. As an intermediate step, we instead use the Banach space isomorphisms

$$\nu_1^- : \mathcal{Z}^{\delta,p,-} \to \mathcal{Z}^{0,p,-} =: \mathcal{Z}^{p,-}, \qquad \nu_2^- : \mathcal{L}^{\delta,p,-} \to \mathcal{L}^{0,p,-} =: \mathcal{L}^{p,-},$$
$$\nu_1^+ : \mathcal{Z}^{\delta,p,+} \to \mathcal{Z}^{0,p,+} =: \mathcal{Z}^{p,+}, \qquad \nu_2^+ : \mathcal{L}^{\delta,p,+} \to \mathcal{L}^{0,p,+} =: \mathcal{L}^{p,+}$$

given by multiplication with the weight function $e^{\delta\beta(s)s}$ (with $\beta \colon \mathbb{R} \to \mathbb{R}$ the function introduced at the beginning of Section 3.1). In the sequel we assume that the weight δ satisfies $0 < \delta < \delta_0(\mathcal{C}^-, \mathcal{C}^+)$ where the positive constant $\delta_0(\mathcal{C}^-, \mathcal{C}^+)$ is defined to be the infimum of the set

$$\big\{ |\lambda| \in \mathbb{R} \,\big|\, \lambda \neq 0 \text{ is eigenvalue of } \mathcal{H}_A \text{ for some } A \in \hat{\mathcal{C}}^- \text{ or}$$
$$\lambda \neq 0 \text{ is eigenvalue of } H_x \text{ for some } x \in \hat{\mathcal{C}}^+ \big\}.$$

We furthermore denote

$$\mathcal{D}_A^\delta := \nu_2^- \circ \mathcal{D}_A \circ (\nu_1^-)^{-1} \qquad \text{and} \qquad \mathcal{D}_x^\delta := \nu_2^+ \circ \mathcal{D}_x \circ (\nu_1^+)^{-1}.$$

We use the notation $\mathcal{Z}^p := \mathcal{Z}^{p,-} \times \mathcal{Z}^{p,+}$ and $\mathcal{L}^p := \mathcal{L}^{p,-} \times \mathcal{L}^{p,+} \times L^2(S^1, \mathfrak{g})$, and set $\mathcal{D}_u^\delta := (\mathcal{D}_A^\delta, \mathcal{D}_x^\delta, N_A) \colon \mathcal{Z}^p \to \mathcal{L}^p$. It is easy to check that the operator $\mathcal{D}_u$ is Fredholm if and only if this holds for $\mathcal{D}_u^\delta$, in which case both Fredholm indices coincide. Note also that the operator $\mathcal{D}_A^\delta$ takes the form

$$\mathcal{D}_A^\delta = \frac{d}{ds} + \mathcal{H}_A - (\beta + \beta' s)\delta, \tag{21}$$

23

and hence, if $\delta > 0$ is chosen sufficiently small, the operator family $s \mapsto \mathcal{H}_A - (\beta + \beta' s)\delta$ converges to the invertible operator $\mathcal{H}_{A^-} + \delta$ as $s \to -\infty$. Analogously, we have that

$$\mathcal{D}_x^\delta = \frac{d}{ds} + H_x - (\beta + \beta' s)\delta. \tag{22}$$

Here the operator family $s \mapsto H_x - (\beta + \beta' s)\delta$ converges to the invertible operator $H_{x^+} - \delta$ as $s \to \infty$.

3.5   Fredholm theorem

For short, we use notation like $L^p(I) := L^p(I, L^p(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))$ to denote the L^p space of $\mathrm{ad}(P)$-valued 1-forms over $I \times \Sigma$, where I is some interval.

Lemma 3.2 *Let $u = (A, \Phi, x) \in \hat{\mathcal{M}}(\hat{\mathcal{C}}^-, \hat{\mathcal{C}}^+)$. There exist positive constants $c(u)$ and $T(u)$ such that the estimate*

$$\|(\alpha, \psi, \xi)\|_{\mathcal{Z}^p} \leq c(u)\big(\|\mathcal{D}_u^\delta(\alpha, \psi, \xi)\|_{\mathcal{L}^p} + \|R(\alpha, \psi)\|_{\mathcal{L}^{p,-}}$$
$$+ \|(\alpha, \psi)\|_{L^p([-T(u),0])} + \|\xi\|_{L^p([0,T(u)])}\big) \tag{23}$$

is satisfied for all $(\alpha, \psi, \xi) \in \mathcal{Z}^p$. Here R denotes the compact operator as in Lemma B.2. As a consequence, the operator $\mathcal{D}_u^\delta$ has finite-dimensional kernel and closed range.

Proof: Inequality (23) is invariant under gauge transformations in $\mathcal{G}(\hat{P})$, and thus it suffices to prove it for $\Phi = 0$. Set $\zeta := (\alpha, \psi)$. For $T = T(A) > 0$ large enough we choose a smooth cut-off function $\chi: (-\infty, 0] \to \mathbb{R}$ with support in $[-T, 0]$ and such that $\chi(s) = 1$ for $s \in [-T + 1, 0]$. Then with R as in Lemma B.2 it follows for some constant $c(A, p)$ that

$$\|\zeta\|_{\mathcal{Z}^{p,-}} \leq \|\chi\zeta\|_{\mathcal{Z}^{p,-}} + \|(1 - \chi)\zeta\|_{\mathcal{Z}^{p,-}}$$
$$\leq c(A, p)\big(\|\mathcal{D}_A^\delta\zeta\|_{L^p([-T,0])} + \|\zeta\|_{L^p([-T,0])} + \|\mathcal{D}_A^\delta\zeta\|_{\mathcal{L}^{p,-}} + \|R\zeta\|_{\mathcal{L}^{p,-}}\big). \tag{24}$$

Here we used (48) to bound the term $\|\chi\zeta\|_{\mathcal{Z}^{p,-}}$. The estimate for the term $\|(1 - \chi)\zeta\|_{\mathcal{Z}^{p,-}}$ follows from Lemma B.2, which remains valid for any $\mathcal{D}_A^\delta$ sufficiently close (in operator norm) to $\mathcal{D}_{A^-}^\delta$. This property holds by Lemma B.1. We now apply Lemma B.4 to bound the term ξ. Let $T = T(x) > 0$ be as in the lemma. This yields for some constant $c = c(x, p, \delta) > 0$ that

$$\|\xi\|_{\mathcal{Z}^{p,+}} \leq c\big(\|\mathcal{D}_x^\delta\xi\|_{\mathcal{L}^{p,+}} + \|\xi\|_{L^p([0,T],L^2(S^1))} + \|\xi(0)\|_{L^2(S^1)}\big). \tag{25}$$

From the definition of N_A and boundedness of the operator $d\Phi(A_0) \colon W^{1,2}(\Sigma) \to W^{1,2}(S^1)$ cf. [18, Lemma A.1], we can further estimate

$$\|\xi(0)\|_{L^2(S^1)} \leq c\big(\|N_A(\alpha, \xi)\|_{L^2(S^1)} + \|\alpha(0)\|_{W^{1,2}(\Sigma)}\big). \tag{26}$$

To control the term $\|\alpha(0)\|_{W^{1,2}(\Sigma)}$ we use Lemma B.8 with $H = W^{1,2}(\Sigma)$, $V = W^{2,2}(\Sigma)$, and $V^* = L^2(\Sigma)$. Choosing $\varepsilon > 0$ in that lemma sufficiently large this implies for a constant $c(A, p) > 0$ the estimate

$$
\begin{aligned}
\|\alpha(0)\|_{W^{1,2}(\Sigma)}^p &\leq \|\zeta(0)\|_{W^{1,2}(\Sigma)}^p \\
&\leq c(A,p)\Big(\int_{-\infty}^0 \|\mathcal{D}_A^\delta \zeta\|_{L^2(\Sigma)}^p \, ds + \int_{-\infty}^0 \|\zeta\|_{W^{1,2}(\Sigma)}^p \, ds\Big) \\
&\leq c(A,p)\big(\|\mathcal{D}_A^\delta \zeta\|_{\mathcal{L}^{p,-}}^p + \|\zeta\|_{\mathcal{Z}^{p,-}}^p\big).
\end{aligned} \tag{27}
$$

The last line follows, as by definition the norm of $\mathcal{L}^{p,-}$ is stronger than that of $L^p(\mathbb{R}^-, L^2(\Sigma))$, and the norm of $\mathcal{Z}^{p,-}$ is stronger than that of $L^p(\mathbb{R}^-, W^{1,2}(\Sigma))$. Combining estimates (25), (26), (27) with (24) yields the claimed inequality (23) (with $T(u) := \max\{T(A), T(x)\}$). To finish the proof, we note that the operator $\mathcal{D}_u^\delta = (\mathcal{D}_A^\delta, \mathcal{D}_x^\delta, N_A)$ is bounded and that the operator R and the inclusion maps $\mathcal{Z}^{p,-}([-T(u), 0]) \hookrightarrow L^p([-T(u), 0])$ and $\mathcal{Z}^{p,+}([0, T(u)]) \hookrightarrow L^p([0, T(u)])$ are compact (by Rellich's theorem). Hence the assertions on the kernel and the range follow from the abstract closed range lemma (cf. [15, p. 14]). The proof is complete. $\qquad\square$

We now state and prove the main result concerning the linear operator $\mathcal{D}_u$.

Theorem 3.3 (Fredholm theorem) *The operator $\mathcal{D}_u$ is a Fredholm operator of index*

$$\operatorname{ind} \mathcal{D}_u = \operatorname{ind} A^- - \operatorname{ind} x^+ - \dim \mathcal{C}^+.$$

(Here $\operatorname{ind} A^-$, respectively $\operatorname{ind} x^+$, denotes the number of negative eigenvalues of $\mathcal{H}_{A^-}$ and H_{x^+}, counted with multiplicities).

Proof: As pointed out after Remark 3.1 it suffices to prove the assertion for the operator $\mathcal{D}_u^\delta$, for any sufficiently small weight $\delta > 0$. That the operator $\mathcal{D}_u^\delta$ has finite-dimensional kernel and closed range is part of Lemma 3.2. It remains to establish the formula for the index. Let $H := L^2(S^1, x^*TG)$ and denote

$$S := \{\xi_0 \in H \mid \exists \xi \in \mathcal{Z}^{p,+} \text{ s.t. } \mathcal{D}_x^\delta \xi = 0 \text{ and } \xi(0) = \xi_0\}.$$

Note that S is a closed subspace of H. Let T be the orthogonal complement of S in H. We set $K := \{(\alpha(0), \psi(0)) \mid (\alpha, \psi) \in \ker \mathcal{D}^\delta_{(A,\Psi)}\}$. It follows that the kernel of the operator $\mathcal{D}^\delta_u$ has dimension

$$\dim \ker \mathcal{D}^\delta_u = \dim \ker d\Phi|_K + \dim(N(K, 0) \cap S). \qquad (28)$$

On the other hand, because the operators $\mathcal{D}^\delta_A$ and $\mathcal{D}^\delta_x$ are surjective, the dimension of $\operatorname{coker} \mathcal{D}^\delta_u$ equals the codimension of the affine space

$$W = \left\{ N(\alpha(0), \xi(0)) \mid \exists (\alpha, \psi, \xi) \in \mathcal{Z}^p \text{ s.t. } (\mathcal{D}^\delta_A(\alpha, \psi), \mathcal{D}^\delta_x \xi) = (\beta, \omega, \eta) \right\}$$

for arbitrary but fixed $(\beta, \omega, \eta) \in \mathcal{L}^p$. Let (α, ψ) vary over the space $\ker \mathcal{D}^\delta_A$ to see that this codimension is given by

$$\operatorname{codim} W = \dim T - \dim K + \dim \ker d\Phi|_K + \dim(N(K, 0) \cap S). \qquad (29)$$

Combining (28) and (29) and using that $\dim K = \operatorname{ind} \mathcal{H}_{A^-}$ (by Lemma B.3) and $\dim T = \operatorname{ind} x^+ + \dim \mathcal{C}^+$ (by Lemma B.5), the asserted index formula follows. $\qquad \square$

Remark 3.4 In view of Remark 3.1 and Theorem 3.3 we obtain the formula

$$\operatorname{ind} d\mathcal{F}(u) = \operatorname{ind} A^- - \operatorname{ind} x^+ + \dim \mathcal{C}^- \qquad (30)$$

for the Fredholm index of the linearization of $\mathcal{F}$.

3.6 Transversality

Throughout we fix a regular value $a \geq 0$ of $\mathcal{YM}$. Our aim here is to show that for every pair $(\mathcal{C}^-, \mathcal{C}^+) \in \mathcal{CR}^a(\mathcal{YM}) \times \mathcal{CR}^b(\mathcal{E})$ (where $b = 4a/\pi$) and a residual subset of a-admmissible perturbations $\mathcal{V} = (\mathcal{V}^-, \mathcal{V}^+) \in Y_a$ (cf. Definition 2.7) the linearized section $d\mathcal{F}(u)$ is surjective, for all $u \in \mathcal{M}(\mathcal{C}^-, \mathcal{C}^+)$. Recall the definition of the Banach manifold $\mathcal{B}$, the Banach space bundle $\mathcal{E}$, and the section $\mathcal{F}$ (cf. Section 3.1). We now change our notation slightly and let $\mathcal{M}(\mathcal{C}^-, \mathcal{C}^+; \mathcal{V})$ indicate the moduli space as in (14), i.e. defined for a fixed perturbation $\mathcal{V} \in Y_a$. Throughout the rest of this section we also replace the notation $\mathcal{F}$ by $\mathcal{F}_\mathcal{V}$. Let then $\hat{\mathcal{F}} : \mathcal{B} \times Y_a \to \mathcal{E}$ denote the section of the Banach space bundle $\mathcal{E}$ defined by

$$\hat{\mathcal{F}} : [(A, \Psi, x, \mathcal{V})] \mapsto \mathcal{F}_\mathcal{V}([(A, \Psi, x)]). \qquad (31)$$

We call the zero set $\mathcal{M}^{\text{univ}}(\mathcal{C}^-, \mathcal{C}^+) := \{w \in \mathcal{B} \times Y_a \mid \hat{\mathcal{F}}(w) = 0\}$ the *universal moduli space*. Thus the perturbation $\mathcal{V} \in Y_a$ which had previously been kept fixed is now allowed to vary over the Banach space Y_a. In this section, our main result is the following.

Theorem 3.5 *There exists a residual subset $Y_a^{\mathrm{reg}} \subseteq Y_a$ of perturbations such that for every $\mathcal{V} \in Y_a^{\mathrm{reg}}$ and every $(\mathcal{C}^-, \mathcal{C}^+) \in \mathcal{CR}^a(\mathcal{YM}) \times \mathcal{CR}^b(\mathcal{E})$ the moduli space $\mathcal{M}(\mathcal{C}^-, \mathcal{C}^+; \mathcal{V})$ is a Banach submanifold of $\mathcal{M}^{\mathrm{univ}}(\mathcal{C}^-, \mathcal{C}^+)$.*

Proof: Let $(\mathcal{C}^-, \mathcal{C}^+) \in \mathcal{CR}^a(\mathcal{YM}) \times \mathcal{CR}^b(\mathcal{E})$. As shown in Theorem 3.6 below, the linearized operator $d\hat{\mathcal{F}}(w)$ is surjective, for every $w \in \mathcal{M}^{\mathrm{univ}}(\mathcal{C}^-, \mathcal{C}^+)$. It hence follows from the implicit function theorem that $\mathcal{M}^{\mathrm{univ}}(\mathcal{C}^-, \mathcal{C}^+)$ is a smooth Banach manifold. The claim now follows from the Sard–Smale theorem for Fredholm maps between Banach manifolds. Namely, by standard results (cf. [21, Proposition 3.3]) it follows that the projection map $\pi \colon \mathcal{M}^{\mathrm{univ}}(\mathcal{C}^-, \mathcal{C}^+) \to Y_a$ is a smooth Fredholm map between Banach manifolds (with index equal to that of $\mathcal{D}_A$). Hence by [1, Theorem 3.6.15] the set of regular values

$$Y_a^{\mathrm{reg}}(\mathcal{C}^-, \mathcal{C}^+) := \left\{ \mathcal{V} \in Y_a \,\middle|\, d\pi(w) \text{ is surjective for all } w \in \mathcal{M}(\mathcal{C}^-, \mathcal{C}^+; \mathcal{V}) \right\}$$

of π is residual in Y_a. Again by the implicit function theorem, it follows that $\mathcal{M}(\mathcal{C}^-, \mathcal{C}^+; \mathcal{V})$ is a Banach submanifold of $\mathcal{M}^{\mathrm{univ}}(\mathcal{C}^-, \mathcal{C}^+)$ for every $\mathcal{V} \in Y_a^{\mathrm{reg}}(\mathcal{C}^-, \mathcal{C}^+)$. Now the set

$$Y_a^{\mathrm{reg}} := \bigcap_{(\mathcal{C}^-, \mathcal{C}^+) \in \mathcal{CR}^a(\mathcal{YM}) \times \mathcal{CR}^b(\mathcal{E})} Y_a^{\mathrm{reg}}(\mathcal{C}^-, \mathcal{C}^+)$$

is the intersection of finitely many residual subsets, hence residual in Y_a. For this set Y_a^{reg}, the assertions of the theorem are satisfied. $\qquad\square$

Theorem 3.6 (Transversality) *The horizontal differential $d\hat{\mathcal{F}}(w)$ of the map $\hat{\mathcal{F}}$ as in (31) is surjective, for every pair $(\mathcal{C}^-, \mathcal{C}^+) \in \mathcal{CR}^a(\mathcal{YM}) \times \mathcal{CR}^b(\mathcal{E})$ and every $w \in \mathcal{M}^{\mathrm{univ}}(\mathcal{C}^-, \mathcal{C}^+)$.*

Proof: The theorem follows, combining Lemmata 3.8 and 3.9 below. $\qquad\square$

To prove Theorem 3.6 we distinguish the cases where $\mathcal{C}^+ = \Phi(\mathcal{C}^-)$ (stationary case) or $\mathcal{C}^+ \neq \Phi(\mathcal{C}^-)$ (non-stationary case). We are naturally led to this distinction by our definition of a-admissible perturbations which are assumed to be supported away from the critical manifolds.

Transversality at stationary flow lines

Throughout we fix a pair $(\mathcal{C}^-, \mathcal{C}^+) \in \mathcal{CR}^a(\mathcal{YM}) \times \mathcal{CR}^b(\mathcal{E})$ such that $\mathcal{C}^+ = \Phi(\mathcal{C}^-)$ is satisfied. In this case, we show that transversality of the section $\mathcal{F}$ holds automatically (i.e. for the perturbation $\mathcal{V} = 0$).

Proposition 3.7 *Let* $[u] \in \mathcal{M}(\mathcal{C}^-, \mathcal{C}^+; \mathcal{V})$ *where* $u \equiv (A, 0, x)$ *for some* $A \in \mathcal{C}^-$ *and* $x = \Phi(A)$. *Then* $\ker \mathcal{D}_u^\delta$ *is trivial.*

Proof: Let $(\alpha, \psi, \xi) \in \ker \mathcal{D}_u^\delta$ and consider the maps

$$\varphi^- : \mathbb{R}^- \to \mathbb{R}, \quad s \mapsto \|(\alpha(s), \psi(s))\|_{L^2(\Sigma)}^2,$$

$$\varphi^+ : \mathbb{R}^+ \to \mathbb{R}, \quad s \mapsto \|\xi(s)\|_{L^2(S^1)}^2.$$

As by assumption $\dot\zeta + \mathcal{H}_A \zeta = 0$ is satisfied by $\zeta := (\alpha, \psi)$ it follows that

$$\dot\varphi^-(s) = -2\langle \zeta, \mathcal{H}_A \zeta \rangle, \qquad \ddot\varphi^-(s) = 4\langle \mathcal{H}_A \zeta, \mathcal{H}_A \zeta \rangle \geq 0. \tag{32}$$

The inequality in (32) shows that φ^- is convex. Because $\lim_{s \to -\infty} \varphi^-(s) = 0$ it thus follows that φ^- vanishes identically or $\dot\varphi^- > 0$. Assume by contradiction the second case. Then the first identity in (32) shows that $\langle \zeta(0), \mathcal{H}_A \zeta(0) \rangle < 0$ and from Proposition A.6 it follows that

$$\langle \xi_0, H_x \xi_0 \rangle < 0, \tag{33}$$

where we denote $\xi_0 := \xi(0) = \mathrm{d}\Phi(A)\zeta(0)$. Likewise, from the assumption that ξ satisfies $\dot\xi + H_x \xi = 0$ it follows that

$$\dot\varphi^+(s) = -2\langle \xi, H_x \xi \rangle, \qquad \ddot\varphi^+(s) = 4\langle H_x \xi, H_x \xi \rangle \geq 0,$$

and so the map φ^+ is convex. Because $\lim_{s \to \infty} \varphi^+(s) = 0$ it follows that $\dot\varphi^+ \leq 0$ and hence in particular $\dot\varphi^+(0) = -2\langle \xi_0, H_x \xi_0 \rangle \leq 0$. This contradicts (33) and shows that our assumption was wrong. Hence φ^- and therefore ζ vanish identically. Now $\xi_0 = \mathrm{d}\Phi(A)\zeta(0) = 0$, and convexity of φ^+ shows that also ξ vanishes identically. Hence $\ker \mathcal{D}_u^\delta$ is trivial, as claimed. $\qquad\square$

Lemma 3.8 *The horizontal differential* $d\mathcal{F}([u])$ *is surjective, for every* $[u] \in \mathcal{M}(\mathcal{C}^-, \mathcal{C}^+; \mathcal{V})$.

Proof: By applying a suitable gauge transformations we may assume that $[u]$ is represented by u of the form $u = (A, 0, x)$. Because $\mathcal{C}^+ = \Phi(\mathcal{C}^-)$ by assumption, it follows from the gradient flow property that u does not depend on s and therefore satisfies the assumptions of Proposition 3.7. In this case it follows from (3.4) that $\mathrm{ind}\, d\mathcal{F}(u) = \dim \mathcal{C}^-$. We check that this number equals the dimension of $\ker d\mathcal{F}(u)$, which will imply the claim. First, the map $[\alpha_0] \mapsto (\alpha, \psi, \xi)$ where $(\alpha(s), \psi(s)) := (\alpha_0, 0)$ for $s \in \mathbb{R}^-$ and $\xi(s) := \mathrm{d}\Phi(A)\alpha_0$ for $s \in \mathbb{R}^+$ is an inclusion of $T_{[A]}\mathcal{C}^-$ into $\ker d\mathcal{F}(u)$. Second, this map is surjective because otherwise the kernel of the map $\mathcal{D}_u^\delta$ could not be trivial, in contradiction to Proposition 3.7. The claim follows. $\qquad\square$

Transversality in the non-stationary case

Let $(\mathcal{C}^-, \mathcal{C}^+) \in \mathcal{CR}^a(\mathcal{YM}) \times \mathcal{CR}^b(\mathcal{E})$, where $b = 4a/\pi$. We show surjectivity of the linearized operator in the case where $\mathcal{C}^+ \neq \Phi(\mathcal{C}^-)$. Under this assumption, the following result holds true.

Lemma 3.9 *The horizontal differential $d\hat{\mathcal{F}}(w)$ is surjective, for every $w = [(A, \Psi, x, \mathcal{V})] \in \mathcal{M}^{\mathrm{univ}}(\mathcal{C}^-, \mathcal{C}^+)$.*

The proof is based on the following auxiliary result.

Proposition 3.10 *Let $w = [(A, \Psi, x, \mathcal{V})] \in \mathcal{M}^{\mathrm{univ}}(\mathcal{C}^-, \mathcal{C}^+)$ and denote $\mathcal{V} = (\mathcal{V}^-, \mathcal{V}^+)$. Then the map $d\hat{\mathcal{F}}(w)$ is onto if one of the following two conditions is satisfied. (i) The linear operator*

$$\hat{\mathcal{D}}_{(A, \mathcal{V}^-)} \colon \{(\alpha, \psi) \in \mathcal{Z}^{p,-} \mid (\alpha(0), \psi(0)) = 0\} \times Y_a^- \to \mathcal{L}^{p,-},$$
$$(\alpha, \psi, v^-) \mapsto \mathcal{D}_A(\alpha, \psi) + \nabla v^-(A)$$

is surjective. (ii) The linear operator

$$\hat{\mathcal{D}}_{(x, \mathcal{V}^+)} \colon \{\xi \in \mathcal{Z}^{p,+} \mid \xi(0) = 0\} \times Y_b^+ \to \mathcal{L}^{p,+}, \quad (\xi, v^+) \mapsto \mathcal{D}_x \xi + \nabla v^+(x)$$

is surjective.

Proof: Assume case (ii). Let $(\beta, \eta, \nu) \in \mathcal{L}^{p,-} \times \mathcal{L}^{p,+} \times L^2(S^1, \mathfrak{g})$ be given. By Lemma B.3 the equation $\hat{\mathcal{D}}_{(A, \mathcal{V}^-)}(\alpha, \psi, v^-) = \beta$ admits a solution (with e.g. $v^- = 0$). Assumption (ii) implies that the equation $\hat{\mathcal{D}}_{(x, \mathcal{V}^+)}(\xi, v^+) = \eta$ can be solved for arbitrary $\xi(0)$, in particular for $\xi(0) = x(0)(\nu + D\Phi_{A(0)}\alpha(0))$, cf. (20) regarding the notation. For this $\xi(0)$, the condition $N_A(\alpha, \xi) = \nu$ is satisfied. It follows that $(\hat{\mathcal{D}}_{(A, \mathcal{V}^-)}, \hat{\mathcal{D}}_{(x, \mathcal{V}^+)})(\alpha, \psi, \xi, v^-, v^+) = (\beta, \eta, \nu)$ and hence $d\hat{\mathcal{F}}(w)$ is onto. Assuming case (i) we may argue analogously, using Lemma B.5 at the place of Lemma B.3. $\square$

Proof: (Lemma 3.9) From our initial assumption that $\mathcal{C}^+ \neq \Phi(\mathcal{C}^-)$ it follows that the gradient flow lines (A, Ψ) or x are not stationary. In this situation, the transversality results [19, Theorem 7.1] and [22, Proposition 7.5] apply (with minor modifications to the present situation where the linear operators $\hat{\mathcal{D}}_{(A, \mathcal{V}^-)}$ and $\hat{\mathcal{D}}_{(x, \mathcal{V}^+)}$ are defined on half-infinite intervals). In the first case this yields surjectivity of $\hat{\mathcal{D}}_{(A, \mathcal{V}^-)}$, in the second case surjectivity of $\hat{\mathcal{D}}_{(x, \mathcal{V}^+)}$. The claim now follows from Proposition 3.10. $\square$

4 Compactness

For a given pair $(\mathcal{C}^-, \mathcal{C}^+) \in \mathcal{CR}^a(\mathcal{YM}) \times \mathcal{CR}^b(\mathcal{E})$ (where $a > 0$ and $b = \frac{4a}{\pi}$), we aim to show compactness of the moduli spaces $\mathcal{M}(\mathcal{C}^-, \mathcal{C}^+)$ (as defined in (14)) up to so-called *convergence to broken trajectories*. Let us first introduce this notion, following the book by Schwarz [17, Definition 2.34].

Definition 4.1 A subset $K \subseteq \mathcal{M}(\mathcal{C}^-, \mathcal{C}^+)$ is called *compact up to broken trajectories of order* $\mu = (\mu^-, \mu^+) \in \mathbb{N}_0^2$ if for any sequence $[u^\nu] = [(A^\nu, \Psi^\nu, x^\nu)]$ in K the following alternative holds. Either $[u^\nu]$ possesses a C^∞ convergent subsequence, or there exist the following:

(i) numbers $0 \leq \lambda^\pm \leq \mu^\pm$ and critical manifolds

$$\mathcal{C}_0^- = \mathcal{C}^-, \ldots, \mathcal{C}_{\lambda^-}^- \subseteq \mathcal{CR}(\mathcal{YM}) \qquad \text{and} \qquad \mathcal{C}_0^+ = \mathcal{C}^+, \ldots, \mathcal{C}_{\lambda^+}^+ \subseteq \mathcal{CR}(\mathcal{E});$$

(ii) for each $0 \leq j \leq \lambda^- - 1$ a connecting trajectory $(A_j, \Psi_j) \in \hat{\mathcal{M}}(\hat{\mathcal{C}}_j^-, \hat{\mathcal{C}}_{j+1}^-)$, a sequence of gauge transformations $(g_{j,\nu})_{\nu \in \mathbb{N}} \subseteq \mathcal{G}(\hat{P})$ and a sequence of reparametrization times $(\tau_{j,\nu}^-)_{\nu \in \mathbb{N}} \subseteq [0, \infty)$;

(iii) a triple $(A^*, \Psi^*, x^*) \in \hat{\mathcal{M}}(\hat{\mathcal{C}}_{\lambda^-}^-, \hat{\mathcal{C}}_{\lambda^+}^+)$ and a sequence of gauge transformations $(g_{\lambda^-, \nu})_{\nu \in \mathbb{N}} \subseteq \mathcal{G}(\hat{P})$;

(iv) for each $1 \leq j \leq \lambda^+$ a connecting trajectory $x_j \in \hat{\mathcal{M}}(\hat{\mathcal{C}}_j^+, \hat{\mathcal{C}}_{j-1}^+)$ and a sequence of reparametrization times $(\tau_{j,\nu}^+)_{\nu \in \mathbb{N}} \subseteq [0, \infty)$, with the following significance:

There exists a subsequence (again labeled by ν) such that, as $\nu \to \infty$,

$$g_{j,\nu}^*(A^\nu(\cdot - \tau_{j,\nu}^-), \Psi^\nu(\cdot - \tau_{j,\nu}^-)) \to (A_j, \Psi_j) \qquad \text{for every } 0 \leq j \leq \lambda^- - 1,$$

$$x^\nu(\cdot + \tau_{j,\nu}^+) \to x_j \qquad \text{for every } 0 \leq j \leq \lambda^+ - 1,$$

$$g_{\lambda^-, \nu}^*(A^\nu, \Psi^\nu) \to (A^*, \Psi^*), \qquad x^\nu \to x^*$$

in C^∞ on all compact domains $I \times \Sigma$, respectively $I \times S^1$, where $I \subseteq \mathbb{R}^-$ (respectively $I \subseteq \mathbb{R}^+$) is a compact interval.

Here the notation $\hat{\mathcal{M}}(\hat{\mathcal{C}}_j^\pm, \hat{\mathcal{C}}_{j\mp 1}^\pm)$ refers to the moduli spaces of connecting trajectories for the gradient flows of $\mathcal{YM}^\mathcal{V}$, respectively of $\mathcal{E}^\mathcal{V}$ as introduced in [19] and [22]. As we show next, the moduli space $\mathcal{M}(\mathcal{C}^-, \mathcal{C}^+)$ is compact in the sense of Definition 4.1.

Theorem 4.2 (Compactness of moduli spaces) *For every pair* $(\mathcal{C}^-, \mathcal{C}^+) \in \mathcal{CR}(\mathcal{YM}) \times \mathcal{CR}(\mathcal{E})$, *the moduli space* $\mathcal{M}(\mathcal{C}^-, \mathcal{C}^+)$ *is empty or compact up to convergence to broken trajectories of order* $\mu = (\mu^-, \mu^+)$, *where*

$$\mu^- + \mu^+ = \operatorname{ind} A^- + \dim \mathcal{C}^- - \operatorname{ind} x^+. \tag{34}$$

(The integers $\operatorname{ind} A^-$ *and* $\operatorname{ind} x^+$ *denote the number of negative eigenvalues of* $\mathcal{H}_{A^-}$, *respectively of* H_{x^+} *(for* $A^- \in \hat{\mathcal{C}}^-$ *and* $x^+ \in \hat{\mathcal{C}}^+$*).*

To prove the theorem we need the following two lemmata, the first one being due to the author [19] and the second one due to Weber [22].

Lemma 4.3 *Let* $\mathbb{A}^\nu = A^\nu + \Psi^\nu\, ds$, $\nu \in \mathbb{N}$, *be a sequence of solutions of the perturbed Yang–Mills gradient flow equation* (10) *on* $\mathbb{R}^- \times \Sigma$. *Assume there exists a critical manifold* $\hat{\mathcal{C}}^- \in \hat{\mathcal{CR}}(\mathcal{YM})$ *such that* $\mathbb{A}^\nu(s)$ *converges to* $\hat{\mathcal{C}}^-$ *as* $s \to -\infty$, *for every* $\nu \in \mathbb{N}$. *Then there exists a sequence* $g^\nu \in \mathcal{G}(\hat{P})$ *of gauge transformations such that a subsequence of the gauge transformed sequence* $(g^\nu)^* \mathbb{A}^\nu$ *converges uniformly on compact sets* $I \times \Sigma$ *to a solution* $\mathbb{A}^*$ *of* (10).

Proof: For a proof we refer to [19, Theorem 6.1]. $\qquad\qquad\square$

Lemma 4.4 *Let* x^ν, $\nu \in \mathbb{N}$, *be a sequence of solutions of the gradient flow equation* (11) *on the interval* $\mathbb{R}^+$. *Assume the uniform energy bound*

$$\sup_{s \in \mathbb{R}^+} \mathcal{E}^{\mathcal{V}^+}(x^\nu(s)) \leq C \tag{35}$$

holds for some constant C *and all* $\nu \in \mathbb{N}$. *Then there exists a solution* x^* *of* (11) *on* $\mathbb{R}^+ \times S^1$ *such that a subsequence of* x^ν *converges uniformly to* x^* *on compact sets* $I \times S^1$.

Proof: For a proof we refer to [22, Proposition 4.14]. $\qquad\qquad\square$

Proof: (Theorem 4.2) Let $u^\nu = (A^\nu, \Psi^\nu, x^\nu)$, $\nu \in \mathbb{N}$, be a sequence in $\hat{\mathcal{M}}(\hat{\mathcal{C}}^-, \hat{\mathcal{C}}^+)$. The sequence x^ν satisfies condition (35). Namely, thanks to the energy inequality (44) it follows for a constant $C(\mathcal{C}^-)$ and all ν that

$$\sup_{s \in \mathbb{R}^+} \mathcal{E}^{\mathcal{V}^+}(x^\nu(s)) = \mathcal{E}^{\mathcal{V}^+}(x^\nu(0)) \leq \frac{4}{\pi} \mathcal{YM}^{\mathcal{V}^-}(A^\nu(0)) \leq C(\mathcal{C}^-).$$

The last inequality follows from the assumption that $A^\nu(s)$ converges to $\hat{\mathcal{C}}^-$ as $s \to -\infty$. Hence Lemmata 4.3 and 4.4 apply and show that there exists a subsequence of u^ν (which we still label by ν) and a sequence g^ν of gauge transformations such that $(g^\nu)^*(A^\nu, \Psi^\nu)$ and x^ν converge uniformly on compact sets $I \times \Sigma$, respectively $I \times S^1$. Furthermore, the limit as $\nu \to \infty$ of $(g^\nu)^*|_{I \times \Sigma}(A^\nu|_{I \times \Sigma}, \Psi^\nu|_{I \times \Sigma})$ (respectively of $x^\nu|_{I \times S^1}$) is the restriction of a solution of (10) or (11) of finite energy at most $C(\mathcal{C}^-)$. Then

31

the exponential decay results [19, Theorem 4.1] and [22, Theorem 1.8] imply that every such finite energy solution is contained in some moduli space $\hat{\mathcal{M}}(\hat{\mathcal{C}}_0^-, \hat{\mathcal{C}}_1^-)$, $\hat{\mathcal{M}}(\hat{\mathcal{C}}_0^+, \hat{\mathcal{C}}_1^+)$, or $\hat{\mathcal{M}}(\hat{\mathcal{C}}_0^-, \hat{\mathcal{C}}_0^+)$ for some $\hat{\mathcal{C}}_j^- \in \widehat{\mathcal{CR}}^a(\mathcal{YM})$, respectively $\hat{\mathcal{C}}_j^+ \in \widehat{\mathcal{CR}}^b(\mathcal{E})$, where $j = 0, 1$ and $a = \mathcal{YM}(\hat{\mathcal{C}}^-)$, $b = 4a/\pi$. (Here we use the notation of Section 2.3). Convergence after reparametrization as required in Definition 4.1 and the relation (34) then follow from standard arguments as in [17, Proposition 2.35]. $\qquad\qquad\qquad\qquad\qquad\square$

5   Chain isomorphism of Morse complexes

5.1   The chain map

Let $a \geq 0$ be a regular value of $\mathcal{YM}$ and set $b := 4a/\pi$. Throughout this section we fix an admissible perturbation $\mathcal{V} = (\mathcal{V}^-, \mathcal{V}^+) \in Y_a^{\mathrm{reg}}$ (with $Y_a^{\mathrm{reg}} \subseteq Y_a$ as in Theorem 3.5) such that the conditions of Theorem A.5 are satisfied. Let $h\colon \mathrm{crit}^a(\mathcal{YM})/\mathcal{G}_0(P) \to \mathbb{R}$ be a smooth Morse–Smale function, i.e. a smooth Morse function such that for all $x, y \in \mathrm{crit}(h)$ the stable and unstable manifolds $W_h^s(x)$ and $W_h^u(y)$ of h intersect transversally. We let

$$CM_*^{a,-} := CM_*^a\big(\mathcal{A}(P)/\mathcal{G}_0(P), \mathcal{V}^-, h\big), \qquad CM_*^{b,+} := CM_*^b\big(\Lambda G/G, \mathcal{V}^+, h\big)$$

denote the Morse–Bott complexes as in Section 2.5. The corrsponding Morse homology groups will for short be denoted by

$$HM_*^{a,-} := HM_*^a\big(\mathcal{A}(P)/\mathcal{G}_0(P), \mathcal{V}^-, h\big), \qquad HM_*^{b,+} := HM_*^b\big(\Lambda G/G, \mathcal{V}^+, h\big).$$

The construction of the moduli space $\mathcal{M}(\mathcal{C}^-, \mathcal{C}^+)$ in Section 2.6 gives rise to a chain map $\Theta\colon CM_*^{a,-} \to CM_*^{b,+}$ as we shall describe next.

Definition 5.1 Fix critical points $x^-, x^+ \in \mathrm{crit}(h)$ and integers $m^-, m^+ \geq 1$. A *hybrid flow line from x^- to x^+ with m^- upper and m^+ lower cascades* is a tuple

$$\begin{aligned}(\underline{x}^-, \underline{x}^0, \underline{x}^+, T^-, T^+) \\ = ((x_j^-)_{j=1,\dots,m^-}, \underline{x}^0, (x_j^+)_{j=1,\dots,m^+}, (t_j^-)_{j=1,\dots,m^-}, (t_j^+)_{j=0,\dots,m^+-1}),\end{aligned}$$

where for each j, $x_j^-\colon \mathbb{R}^- \to \mathcal{A}(P)/\mathcal{G}_0(P)$ is a nonconstant solution of the Yang–Mills gradient flow equation (10), $x_j^+\colon \mathbb{R}^+ \to \Lambda G/G$ is a nonconstant solution of the loop group gradient flow equation (11), $t_j^\pm \in \mathbb{R}^+$, and the following conditions are satisfied.

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

(ii) There exist $p^- \in W_h^u(x^-)$ and $p^+ \in W_h^s(x^+)$ such that $\lim_{s \to -\infty} x_1^-(s) = p^-$ and $\lim_{s \to \infty} x_m^+(s) = p^+$.

(iii) There exist $\mathcal{C}^- \in \mathcal{CR}(\mathcal{YM})$ and $\mathcal{C}^+ \in \mathcal{CR}(\mathcal{E})$ such that $\underline{x}^0 = [(u^-, u^+)] \in \mathcal{M}(\mathcal{C}^-, \mathcal{C}^+)$. Furthermore, there exist solutions $y_{m^-}^-$ and y_0^+ of the gradient flow equations $\dot{y}_{m^-}^- = -\nabla h(y_{m^-}^-)$ and $\dot{y}_0^+ = -\nabla h(y_0^+)$ satisfying the conditions

$$\lim_{s \to \infty} x_{m^-}^-(s) = y_{m^-}^-(0), \qquad \lim_{s \to -\infty} u^-(s) = y_{m^-}^-(t_{m^-}^-),$$

$$\lim_{s \to \infty} u^+(s) = y_0^+(0), \qquad \lim_{s \to -\infty} x_1^+(s) = y_0^+(t_0^+).$$

A *hybrid flow line with $m^- = 0$ upper cascades and $m^+ \geq 1$ lower cascades* is a tuple $(\underline{x}^0, \underline{x}^+, T^+) = (\underline{x}^0, (x_j^+)_{j=1,\dots,m^+}, (t_j^+)_{j=0,\dots,m^+ - 1})$ satisfying conditions (i-iii) above with the following adjustment in (ii). Here we require the existence of $p^- \in W_h^u(x^-)$ such that $\lim_{s \to -\infty} u^-(s) = p^-$. Conditions involving flow lines x_j^- and times t_j^- are empty in this case. A *hybrid flow line with $m^- \geq 0$ upper cascades and $m^+ = 0$ lower cascades* is defined analogously. Conditions involving flow lines x_j^+ and times t_j^+ are then empty.

For $(m^-, m^+) \in \mathbb{N}_0^2$ we denote by $\hat{\mathcal{M}}_{(m^-,m^+)}^{\mathrm{hybr}}(x^-, x^+)$ the set of hybrid flow lines from x^- to x^+ with m^- upper and m^+ lower cascades. The group $\mathbb{R}^{m^-} \times \mathbb{R}^{m^+}$ acts on tuples $(\underline{x}^-, \underline{x}^+)$ by time-shifts. The resulting quotient space is called the *moduli space of hybrid flow lines from x^- to x^+ with m^- upper and m^+ lower cascades*. We call

$$\mathcal{M}^{\mathrm{hybr}}(x^-, x^+) := \bigcup_{(m^-, m^+) \in \mathbb{N}_0^2} \mathcal{M}_{(m^-,m^+)}^{\mathrm{hybr}}(x^-, x^+)$$

the moduli space of *hybrid flow lines with cascades from x^- to x^+*.

Lemma 5.2 *The dimension of $\mathcal{M}^{\mathrm{hybr}}(x^-, x^+)$ is given by the formula*

$$\dim \mathcal{M}^{\mathrm{hybr}}(x^-, x^+) = \mathrm{Ind}(x^-) - \mathrm{Ind}(x^+).$$

(The nonnegative integers $\mathrm{Ind}(x^{\pm})$ denote the sum of the Morse indices of x^- as a critical point of $\mathcal{YM}$ and of h, respectively the sum of the Morse indices of x^+ as a critical point of $\mathcal{E}$ and of h).
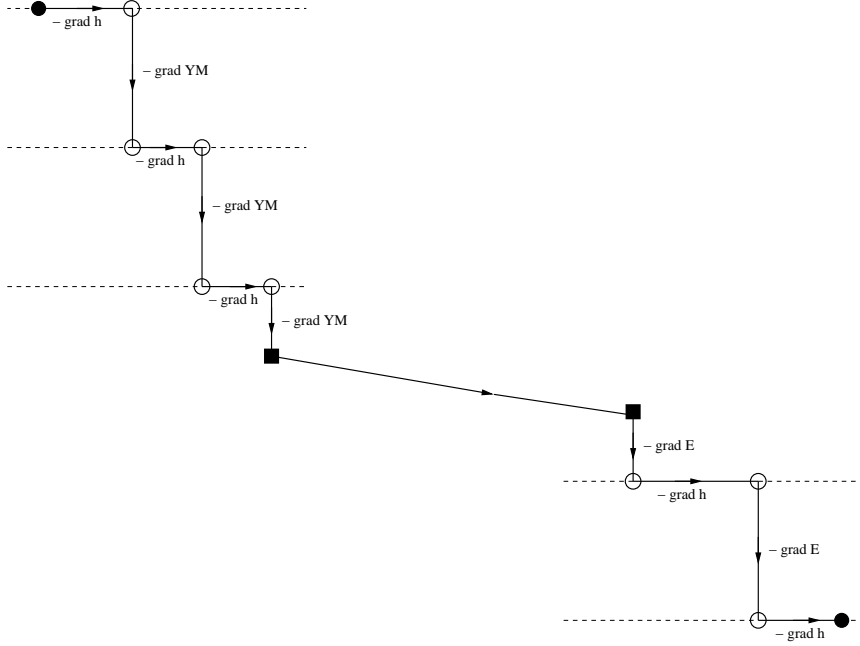
Figure 1: A hybrid flow line with 2 upper and 1 lower cascades.

Figure 2: A hybrid flow line with 2 upper and 1 lower cascades.

Proof: The assertion follows from the following formula for the dimension of $\hat{\mathcal{M}}^{\text{hybr}}_{(m^-, m^+)}(x^-, x^+)$. Namely, by the result [7, Corollary C.15] and the index formula (30) it follows that

$$\dim \hat{\mathcal{M}}^{\text{hybr}}_{(m^-, m^+)}(x^-, x^+) = \text{Ind}(x^-) - \text{Ind}(x^+) + m - 1,$$

where m denotes the total number of cascades of the elements of $\hat{\mathcal{M}}^{\text{hybr}}_{(m^-, m^+)}(x^-, x^+)$. Adapting the arguments of [7, Corollary C.15] to the present situation it follows that $m = m^- + m^+ + 1$ (we here have to take into account the additional cascade coming from the configuration $\underline{x}_0$). The asserted formula for the quotient space modulo the action of $\mathbb{R}^{m^-} \times \mathbb{R}^{m^+}$ then follows. $\qquad \square$

We obtain a chain map between the Morse–Bott complexes $CM^{a,-}_*$ and $CM^{b,+}_*$ by counting the number elements of $\mathcal{M}^{\text{hybr}}(x^-, x^+)$ for appropriate pairs (x^-, x^+) of generators.

34

Definition 5.3 For a pair $(x^-, x^+) \in \text{crit}(h) \times \text{crit}(h)$ with $\text{Ind}(x^-) = k$ let

$$\Theta_k(x^-) := \sum_{\substack{x^+ \in \text{crit}(h) \\ \text{Ind}(x^+) = k}} \# \mathcal{M}^{\text{hybr}}(x^-, x^+) \cdot x^+,$$

where $\# \mathcal{M}^{\text{hybr}}(x^-, x^+)$ denotes the number (counted modulo 2) of elements of $\mathcal{M}^{\text{hybr}}(x^-, x^+)$. We define the homomorphism $\Theta_k \colon CM_k^{a,-} \to CM_k^{b,+}$ of abelian groups accordingly by linear continuation, and set $\Theta := (\Theta_k)_{k \in \mathbb{N}_0}$.

Theorem 5.4 (Chain map) *The map Θ is a chain homomorphism between the Morse complexes $CM_*^{a,-}$ and $CM_*^{b,+}$. Thus for each $k \in \mathbb{N}_0$ it holds $\Theta_k \circ \partial_{k+1}^{\mathcal{YM}} = \partial_{k+1}^{\mathcal{E}} \circ \Theta_{k+1} \colon CM_{k+1}^{a,-} \to CM_k^{b,+}$.*

Proof: From the compactness Theorem 4.2 it follows that the number of elements of $\mathcal{M}^{\text{hybr}}(x^-, x^+)$ is finite and hence the homomorphism Θ is well-defined. The proof is then completed by standard arguments as e.g. carried out in [2]. $\qquad\square$

5.2 Proof of the main theorem

The aim of this final section is to prove Theorem 1.1. Thanks to Theorem 5.4, the map Θ induces for each $k \in \mathbb{N}_0$ a homomorphism $[\Theta_k] \colon HM_k^{a,-} \to HM_k^{b,+}$ of abelian groups. It remains to show that these homomorphisms are in fact isomorphisms. In our proof we follow closely the line of argument employed by Abbondandolo and Schwarz in [2].

Proof: (Theorem 1.1) It suffices to prove that each of the chain homomorphisms Θ_k is in fact an isomorphism and hence induces an isomorphism in homology. Let $k \in \mathbb{N}_0$ and fix a set $\underline{p} := (p_1, \ldots, p_m)$ of generators of $CM_k^{a,-}$. We order the entries p_j of the tuple $\underline{p}$ such that the following two conditions are met. First, we require that

$$i \leq j \qquad \Longrightarrow \qquad \mathcal{YM}^{\mathcal{V}^-}(p_i) \leq \mathcal{YM}^{\mathcal{V}^-}(p_j) \tag{36}$$

holds for all $1 \leq i, j \leq m$. Secondly, in case where p_i and p_j lie on the same critical manifold $\mathcal{C}^- \in \mathcal{CR}(\mathcal{YM})$ we choose our ordering such that

$$i \leq j \qquad \Longrightarrow \qquad h(p_i) \leq h(p_j). \tag{37}$$

Set $q_j := \Phi(p_j)$ and define $\underline{q} := (q_1, \ldots, q_m)$. The m-tuple $\underline{q}$ generates $CM_k^{b,+}$ which follows from the facts that Φ induces a bijection $\mathcal{CR}^a(\mathcal{YM}) \to \mathcal{CR}^b(\mathcal{E})$

which by Theorem A.2 preserves the Morse indices, and that on both sets $\mathcal{CR}^a(\mathcal{YM})$ and $\mathcal{CR}^b(\mathcal{E})$ we use the same Morse function h. By Theorem A.5 we have for each $1 \leq j \leq m$ the identity

$$\mathcal{YM}^{\mathcal{V}^-}(p_j) = \frac{\pi}{4}\mathcal{E}^{\mathcal{V}^+}(q_j), \tag{38}$$

which by our choice of the ordering of $\underline{p}$ implies that the tuple $\underline{q}$ is ordered by non-decreasing $\mathcal{E}^{\mathcal{V}^+}$ action. Furthermore, if $\mathcal{E}^{\mathcal{V}^+}(q_i) = \mathcal{E}^{\mathcal{V}^+}(q_j)$ for some $i \leq j$ then either q_i and q_j lie on different critical manifolds in $\mathcal{CR}^b(\mathcal{E})$ or otherwise $h(q_i) \leq h(q_j)$. Let us represent the homomorphism Θ_k with respect to the ordered bases $\underline{p}$ and $\underline{q}$ by the matrix $(\Theta_{ij}^k)_{1 \leq i,j \leq m} \in \mathbb{Z}_2^{m \times m}$. The following two observations are now crucial. Both are a consequence of the energy inequality

$$\mathcal{YM}^{\mathcal{V}^-}(A) \geq \frac{\pi}{4}\mathcal{E}^{\mathcal{V}^+}(\Phi(A)) \tag{39}$$

for all $A \in \mathcal{A}(P)$ as in Theorem A.5. First, $\Theta_{ii}^k = 1$ for all $1 \leq i \leq m$ because the moduli space $\mathcal{M}^{\text{hybr}}(p_i, q_i)$ consists of precisely one point. It is represented by the hybrid flow line with $(m^-, m^+) = (0,0)$ upper and lower cascades and configuration $\underline{x}^0 = (p_i, q_i) = [(u^-, u^+)]$ with stationary flow lines u^- of $\mathcal{YM}^{\mathcal{V}}$, respectively u^+ of $\mathcal{E}^{\mathcal{V}}$. Note that $(m^-, m^+) \neq (0,0)$ is not possible in this case as this would contradict (38). Secondly, if $i > j$ then $\Theta_{ij}^k = 0$ because in this case $\mathcal{M}^{\text{hybr}}(p_j, q_i) = \emptyset$. Assume by contradiction that $\mathcal{M}^{\text{hybr}}(p_j, q_i)$ contains at least one element. Let $p_j \in \mathcal{C}^-$ and $q_i \in \mathcal{C}^+$ for critical manifolds $\mathcal{C}^{\pm}$. The gradient flow property and (39) imply that $\mathcal{YM}^{\mathcal{V}^-}(p_j) \geq \frac{\pi}{4}\mathcal{E}^{\mathcal{V}^+}(q_i)$. This inequality must in fact be an equality because otherwise there would be a contradiction to condition (36) and identity (38). Hence $\mathcal{C}^+ = \Phi(\mathcal{C}^-)$, and the gradient flow property of the function h implies that $h(q_j) = h(\Phi(p_j)) \geq h(q_i)$. With $i > j$, condition (37) and again identity (38) show that this can only be the case if $h(q_j) = h(q_i)$. It follows that $q_j = q_i$ because h is monotone decreasing. Thus $i = j$, which is a contradiction. These two observations imply that the matrix $(\Theta_{ij}^k)_{1 \leq i,j \leq m}$ takes the form

$$(\Theta_{ij}^k)_{1 \leq i,j \leq m} = \begin{pmatrix} 1 & * & \cdots & \cdots & * \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & * \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \in \mathbb{Z}_2^{m \times m},$$

and thus is invertible in $\mathbb{Z}_2^{m \times m}$. This completes the proof of the theorem. $\quad\square$

A Holonomy map and critical manifolds

We let $\Sigma = S^2 \subseteq \mathbb{R}^3$ be the unit sphere, endowed with the standard round metric induced from the ambient euclidian space $\mathbb{R}^3$. Fix the pair $z^{\pm} = (0, 0, \pm 1)$ of antipodal points and set $D^{\pm} := \Sigma \setminus \{z^{\mp}\}$. We parametrize the hypersurfaces $D^{\pm} \subseteq \mathbb{R}^3$ by the maps

$$u^{\pm} \colon [0, \pi) \times [0, 2\pi) \to D^{\pm}, \quad (r, t) \mapsto \big(\pm \cos(t) \sin(r), \pm \sin(t) \sin(r), \pm \cos(r) \big).$$

The inverses of $u^{\pm}$ give rise to coordinate charts on Σ. Set $\lambda \colon [0, \pi) \to \mathbb{R}$, $r \mapsto \sin(r)$. The Riemannian metric in these coordinates is $g^{\pm} = dr^2 + \lambda^2 \, dt^2$ and the volume form is $\mathrm{dvol}(D^{\pm}) = \lambda \, dr \wedge dt$. The Hodge star operator acts on differential forms on $D^{\pm}$ as

$$*1 = \lambda \, dr \wedge dt, \qquad * \, dr \wedge dt = \lambda^{-1}, \qquad * \, dr = \lambda \, dt, \qquad * \, dt = -\lambda^{-1} \, dr.$$

We define the family γ_t ($0 \leq t \leq 2\pi$) of paths between z^+ and z^- as

$$\gamma_t \colon [0, \pi] \to \Sigma, \qquad \gamma_t(r) := (\cos(t) \sin(r), \sin(t) \sin(r), \cos(r)).$$

We fix two points $p^{\pm} \in P_{z^{\pm}}$ in the fibre $P_{z^{\pm}} \subseteq P$ above $z^{\pm}$. For a connection $A \in \mathcal{A}(P)$, let $\bar{\gamma}_t^A$ denote the horizontal lift of the path γ_t with respect to A, starting at p^+.

Definition A.1 For the parameter $0 \leq t \leq 2\pi$ we define $g_A(t) \in G$ by the condition $\bar{\gamma}_t^A(\pi) = p^- \cdot g_A(t)$. Therefore $t \mapsto g_A(t)$ is a loop in ΛG. Now we define the map

$$\Phi \colon \mathcal{A}(P) \to \Omega G, \quad A \mapsto x_A := g_A^{-1}(0) g_A. \tag{40}$$

The map Φ is called the *holonomy map* of the principal G-bundle P.

Note that Φ is independent of the choice of p^-, while replacing p^+ by $p^+ \cdot h$ for some $h \in G$ results in the conjugate holonomy map $h \Phi h^{-1}$. Similarly, replacing A by $g^* A$ for a gauge transformation $g \in \mathcal{G}(P)$ with $g(p^+) = h \in G$ yields $\Phi(g^* A) = h \Phi(A) h^{-1}$. Thus for the subgroup of gauge transformations $\mathcal{G}_0(P)$ based at z^+, i.e. for

$$\mathcal{G}_0(P) := \{g \in \mathcal{G}(P) \mid g(p^+) = \mathbb{1}\},$$

it follows that Φ descends to a G-equivariant map $\Phi \colon \mathcal{A}(P) / \mathcal{G}_0(P) \to \Omega G$, again denoted Φ and named holonomy map. The next theorem gives an

explicit description of the set of Yang–Mills connections on the principal G-bundle P. Recall that the set of isomorphism classes of principal G-bundles is in bijection with the elements of the fundamental group $\pi_1(G)$, as any such bundle P is determined up to isomorphism by the homotopy class of the transition map $D^+ \cap D^- \to G$ of a trivialization of P over the open sets $D^\pm$.

Theorem A.2 (Correspondence between critical points) *Let P be a principal G-bundle over Σ of topological type $\alpha \in \pi_1(G)$. Then the map $\Phi\colon A \mapsto x_A$ induces a bijection between the set of gauge equivalence classes of Yang–Mills connections on P and the set of conjugacy classes of closed, based geodesics x on G of homotopy class $[x] = \alpha$.*

Proof: For a proof we refer to [8, Theorem 2.1]. □

Theorem A.3 *Let $[A] \in \mathcal{A}(P)/\mathcal{G}_0(P)$ be a based gauge equivalence class of Yang–Mills connections, and let $x_A = \Phi(A) \in \Omega G$ denote the corresponding closed geodesic. Then the Hessians of $\mathcal{YM}$ at $[A]$ and of $\mathcal{E}$ at x_A have the same index and nullity.*

Proof: For a proof we refer to [8, Theorem 2.2]. □

Let $i^\pm\colon D^\pm \to \Sigma$ denote the inclusion maps and $P^\pm := (i^\pm)^*P$ the corresponding pull-back bundles. Because $D^\pm \subseteq \Sigma$ is contractible, the bundles $P^\pm$ are trivial and thus admit sections over $D^\pm$. Given $A \in \mathcal{A}(P)$ we can choose sections $s^\pm\colon D^\pm \to P^\pm$ in such a way that the pull-back of the connections $(i^\pm)^*A \in \mathcal{A}(P^\pm)$ under $s^\pm$ are of the form $u^\pm\,dt$ for maps $u^\pm \in C^\infty(D, \mathfrak{g})$ satisfying

$$u^-(r,t) = x_A(t)u^+(\pi - r,t)x_A^{-1}(t) - \partial_t x_A(t)x_A^{-1}(t) \tag{41}$$

for $0 < r < \pi$. Because the connections $(i^\pm)^*A$ are well-defined near $r = 0$ it follows that $\lim_{r\to 0} u^\pm(r,t) = 0$. Following Gravesen [9] we set

$$\ell(r) := \frac{1}{2}(1 - \cos(r)) \qquad \text{and} \qquad \xi_A := x_A^{-1}\partial_t x_A \in \Omega\mathfrak{g},$$

and split the map u^+ as

$$u^+(r,t) = \ell(r)\xi_A(t) + m_A(r,t). \tag{42}$$

($m_A \in C^\infty(D^+, \mathfrak{g})$ being defined through this equation). Because $u^+(r,t) \to 0$ as $r \to 0$ this definition implies that also $\lim_{r \to 0} m_A(r,t) = 0$. Inserting (42) into (41) yields

$$u^-(r,t) = \ell(\pi - r)x_A(t)\xi_A(t)x_A^{-1}(t) + x_A^{-1}(t)m_A(\pi - r, t)x_A(t) - \partial_t x_A(t)x_A^{-1}(t).$$

Because $\ell(\pi) = 1$ and $u^-(r,t) \to 0$ as $r \to 0$ it follows that $\lim_{r \to \pi} m(r,t) = 0$. The following energy identity is due to Gravesen [9, Section 2]. There is also a generalization to higher genus surfaces, cf. Davies [5, Section 4.2].

Lemma A.4 (Energy identity) *For every $A \in \mathcal{A}(P)$, the identity*

$$\mathcal{YM}(A) = \frac{\pi}{4}\mathcal{E}(x_A) + \frac{1}{2}\|\lambda^{-1}\partial_r m_A(r,t)\|^2_{L^2(\Sigma, \mathrm{dvol}(\Sigma))}. \tag{43}$$

is satisfied.

Proof: We identify $A \in \mathcal{A}(P)$ locally on $D^\pm$ with 1-forms $u^\pm \in \Omega^1(D^\pm, \mathfrak{g})$ as in (41) and make use of the decomposition (42). It follows that the curvature F_A is identified on $D^\pm$ with the 2-form $\partial_r u^\pm \, dr \wedge dt$. Recall also the formula $*(dr \wedge dt) = \lambda^{-1}$ for the Hodge star operator on 2-forms. In addition, we use that $\partial_r \ell(r) = \frac{1}{2}\lambda(r)$. Hence it follows that

$$
\begin{aligned}
\mathcal{YM}(A) &= \frac{1}{2}\int_\Sigma \langle F_A \wedge *F_A \rangle \\
&= \frac{1}{2}\int_{D^+} \lambda^{-1}\langle \partial_r u^+, \partial_r u^+ \rangle \, dr \wedge dt \\
&= \frac{1}{2}\int_{D^+} \lambda^{-1}(r)\left\langle \frac{1}{2}\lambda(r)\xi_A(t) + \partial_r m_A(r,t), \frac{1}{2}\lambda(r)\xi_A(t) + \partial_r m_A(r,t) \right\rangle \, dr \wedge dt \\
&= \frac{1}{2}\int_{D^+} \left\langle \frac{1}{2}\xi_A(t) + \lambda^{-1}(r)\partial_r m_A(r,t), \frac{1}{2}\xi_A(t) + \lambda^{-1}(r)\partial_r m_A(r,t) \right\rangle \lambda(r) \, dr \wedge dt \\
&= \frac{\pi}{8}\int_0^{2\pi} \langle \xi_A(t), \xi_A(t) \rangle \, dt + \frac{1}{2}\int_{D^+} \left|\lambda^{-1}(r)\partial_r m_A(r,t)\right|^2 \lambda(r) \, dr \wedge dt \\
&\quad + \frac{1}{2}\int_{D^+} \partial_r \langle \xi_A(t), m_A(r,t) \rangle \, dr \wedge dt \\
&= \frac{\pi}{4}\mathcal{E}(x_A) + \frac{1}{2}\|\lambda^{-1}\partial_r m_A(r,t)\|^2_{L^2(D^+, \mathrm{dvol}(D^+))}.
\end{aligned}
$$

The term in the second but last line vanishes as follows from the above stated property $\lim_{r \to 0} m_A(r,t) = \lim_{r \to \pi} m_A(r,t) = 0$. $\qquad\square$

Theorem A.5 (Energy inequality) *Let $a \geq 0$ and $\mathcal{V} = (\mathcal{V}^-, \mathcal{V}^+)$ with $\mathcal{V}^- = \sum_{\ell=1}^{\infty} \lambda_\ell^- \mathcal{V}_\ell^- \in Y^-$ and $\mathcal{V}^+ = \sum_{\ell=1}^{\infty} \lambda_\ell^+ \mathcal{V}_\ell^+ \in Y^+$ be an a-admissible perturbation (cf. Definition 2.7). Assume that the coefficients λ_ℓ^- are non-negative, and the coefficients λ_ℓ^+ are non-positive. Then for every $A \in \mathcal{A}(P)$ there holds the inequality*

$$\mathcal{YM}^{\mathcal{V}^-}(A) \geq \frac{\pi}{4} \mathcal{E}^{\mathcal{V}^+}(x_A), \tag{44}$$

with equality if A is Yang–Mills. In this case, the loop x_A is a geodesic.

Proof: In view of Remarks 2.5 and 2.6 and the assumptions on $\lambda_\ell^{\pm}$ it follows that $\mathcal{V}^- \geq 0$ and $\mathcal{V}^+ \leq 0$. Hence it suffices to prove the inequality in the case of vanishing perturbations $\mathcal{V}^{\pm} = 0$, where it follows from the energy identity (43) (note that the last term in (43) is non-negative). Equality in the case where A is a Yang–Mills connection and the assertion that then x_A is a closed geodesic follow from the discussion in [8, p. 236]. $\square$

The following is an infinitesimal version of the energy identity (43).

Proposition A.6 *Let $A \in \mathcal{C}^-$ and $x = \Phi(A) \in \mathcal{C}^+$. Let $\alpha \in \Omega^1(\Sigma, \mathrm{ad}(P))$ such that $\langle \alpha, \mathcal{H}_A \alpha \rangle < 0$ and $\beta := \mathrm{d}\Phi(A)\alpha$. Then it follows that $\langle \beta, H_x \beta \rangle < 0$.*

Proof: The claim is an immediate consequence of the energy identity (43) which implies that for sufficiently small $\varepsilon > 0$ the map $\varepsilon \mapsto \mathcal{E}(x_{A+\varepsilon\alpha})$ is strictly monotone decreasing. $\square$

B   A priori estimates

Estimates involving $\mathcal{D}_A^\delta$

For $A \in \mathcal{A}(P)$ we let $\mathcal{H}_A$ denote the augmented Yang–Mills Hessian as in (17). Unless otherwise stated we assume in the following that the term $\mathcal{V}^-$ in $\mathcal{H}_A$ vanishes.

Lemma B.1 *For every $A^- \in \mathcal{A}(P)$ there exists $c(A^-) > 0$ such that the operator $\mathcal{D}_A$ in (19) satisfies the estimate*

$$\|\mathcal{D}_A - \mathcal{D}_{A^-}\|_{\mathcal{L}(\mathcal{Z}^{\delta,p,-}, \mathcal{L}^{\delta,p,-})} \leq c(A^-) \|\alpha\|_{C^0(\mathbb{R}^-, C^1(\Sigma))}$$

for all time-dependent connections $A = A^- + \alpha$. Similar estimates hold for the operator $\mathcal{D}_A^\delta$ in (21) and for domains with $\mathbb{R}^-$ being replaced by some interval $I \subseteq \mathbb{R}^-$.

Proof: Consider the upper left entry in $\mathcal{H}_A - \mathcal{H}_{A^-}$, i.e.

$$- *[\alpha \wedge *d_{A^-+\alpha} \cdot] + d_{A^-}^*[\alpha \wedge \cdot] + *[*(d_{A^-}\alpha + \frac{1}{2}[\alpha \wedge \alpha] \wedge \cdot].$$

It clearly admits an estimate as claimed. The other terms follow similarly. $\square$

The following is a basic estimate involving the operator $\mathcal{D}_A^\delta$ in (21), here for a stationary path $A(s) \equiv A$ (for all $s \in \mathbb{R}^-$).

Lemma B.2 *Let $A \in \mathcal{A}(P)$ and assume for some $\delta > 0$ that the operator $\mathcal{H}_A + \delta$ is injective. Then for every $p \geq 2$ there exists a constant $c(A, p, \delta)$ and a compact operator $R \colon \mathcal{Z}^{p,-} \to \mathcal{L}^{p,-}$ such that the estimate*

$$\|(\alpha, \psi)\|_{\mathcal{Z}^{p,-}} \leq c(A, p, \delta)\big(\|\mathcal{D}_A^\delta(\alpha, \psi)\|_{\mathcal{L}^{p,-}} + \|R(\alpha, \psi)\|_{\mathcal{L}^{p,-}}\big) \qquad (45)$$

holds for all $(\alpha, \psi) \in \mathcal{Z}^{p,-}$. Moreover, the operator $\mathcal{D}_A^\delta \colon \mathcal{Z}^{p,-} \to \mathcal{L}^{p,-}$ is surjective and has finite-dimensional kernel of dimension

$$\dim \ker \mathcal{D}_A^\delta = \operatorname{ind} A.$$

(The integer $\operatorname{ind} A$ denoting the number of negative eigenvalues of $\mathcal{H}_A + \delta$).

Proof: The proof follows the lines of [22, Theorem 8.5] and consists of four steps. Throughout we set $\zeta := (\alpha, \psi)$ and denote $H := L^2(\Sigma, T^*\Sigma \otimes \operatorname{ad}(P)) \oplus L^2(\Sigma, \operatorname{ad}(P))$.

Step 1 *The statement on surjectivity and the kernel is true in the case $p = 2$.*

The operator $\mathcal{H}_A + \delta$ with domain as in (18) is an unbounded self-adjoint operator on the Hilbert space H, cf. [19, Proposition 5.1]. Denote by E^- and E^+ its negative, respectively positive eigenspaces. Since $\mathcal{H}_A + \delta$ is assumed to be injective, H splits as an orthogonal sum $H = E^- \oplus E^+$. Let $P^\pm$ denote the projections onto $E^\pm$ and set $\mathcal{H}^\pm = (\mathcal{H}_A + \delta)|_{E^\pm}$. As $\mathcal{H}^-$ and $-\mathcal{H}^+$ are negative-definite self-adjoint operators, it follows from the Hille–Yosida theorem (cf. for instance [14, Section X.8]) that they generate strongly continuous contraction semigroups $s \mapsto e^{s\mathcal{H}^-}$ on E^- respectively $s \mapsto e^{-s\mathcal{H}^+}$ on E^+, both defined for $s \geq 0$. This allows us to define the map $K \colon \mathbb{R} \to \mathcal{L}(H)$ by

$$K(s) := \begin{cases} -e^{-s\mathcal{H}^-}P^- & \text{for } s \leq 0, \\ e^{-s\mathcal{H}^+}P^+ & \text{for } s > 0. \end{cases} \qquad (46)$$

As one easily checks, K is strongly continuous in $\mathbb{R} \setminus \{0\}$ and its pointwise operator norm satisfies

$$\|K(s)\|_{\mathcal{L}(H)} \le e^{-\delta_0 |s|}$$

for $\delta_0 > 0$ the smallest (in absolute value) eigenvalue of $\mathcal{H}_A + \delta$. Now consider the operator $Q \colon \mathcal{L}^{2,-} \to \mathcal{Z}^{2,-}$ defined by

$$(Q\eta)(s) := \int_{-\infty}^{0} K(s - \sigma)\eta(\sigma) \, d\sigma. \tag{47}$$

It satisfies

$$
\begin{aligned}
\frac{d}{ds}(Q\eta)(s) &= \frac{d}{ds} \int_{-\infty}^{s} e^{-(s-\sigma)\mathcal{H}^+} P^+ \eta(\sigma) \, d\sigma - \frac{d}{ds} \int_{s}^{0} e^{-(s-\sigma)\mathcal{H}^-} P^- \eta(\sigma) \, d\sigma \\
&= P^+ \eta(s) - \int_{-\infty}^{s} \mathcal{H}^+ e^{-(s-\sigma)\mathcal{H}^+} P^+ \eta(\sigma) \, d\sigma \\
&\quad + P^- \eta(s) + \int_{s}^{0} \mathcal{H}^- e^{-(s-\sigma)\mathcal{H}^-} P^- \eta(\sigma) \, d\sigma \\
&= \eta(s) - (\mathcal{H}_A + \delta)(Q\eta)(s).
\end{aligned}
$$

From this calculation we see that

$$\mathcal{D}_A^\delta Q\eta = \frac{d}{ds}(Q\eta) + (\mathcal{H}_A + \delta)(Q\eta) = \eta,$$

so Q is a right-inverse of $\mathcal{D}_A^\delta$. This proves surjectivity of $\mathcal{D}_A^\delta$. Now let $\zeta \in \ker \mathcal{D}_A^\delta$ and assume that $(\mathcal{H}_A + \delta)\zeta(0) = \lambda\zeta(0)$ for some $\lambda \in \mathbb{R}$. Then $\zeta(s) = e^{-\lambda s}\zeta(0)$, which is contained in $\mathcal{L}^{2,-}$ if and only if $\lambda < 0$. Therefore $\zeta \in \mathcal{Z}^{2,-}$ satisfies $\mathcal{D}_A^\delta \zeta = 0$ if and only if $\zeta(0) \in \mathcal{H}^-$. This shows that $\mathcal{H}^-$ and $\ker \mathcal{D}_A^\delta$ are isomorphic to each other.

Step 2 *For every $p \ge 2$ there exists a constant $c_1(A, p)$ such that the following holds. If $\zeta \in \mathcal{Z}^{2,-}$ and $\mathcal{D}_A^\delta \zeta \in \mathcal{L}^{p,-}$, then $\zeta \in \mathcal{Z}^{p,-}$ and*

$$\|\zeta\|_{\mathcal{Z}^{p,-}} \le c_1(A, p)\big(\|\mathcal{D}_A^\delta \zeta\|_{\mathcal{L}^{p,-}} + \|\zeta\|_{L^p(\mathbb{R}^-, H)}\big).$$

The claim follows from standard arguments based on the linear estimate

$$\|\zeta\|_{\mathcal{Z}^{p,-}([-1,0])} \le c(A, p)\big(\|\mathcal{D}_A^\delta \zeta\|_{L^p([-2,0])} + \|\zeta\|_{L^p([-2,0])}\big), \tag{48}$$

cf. [18, Proposition A.6]. Full details are given in [18, Lemma 3.20].

Step 3 *The operator $Q\colon \mathcal{L}^{p,-} \to L^p(\mathbb{R}^-, H)$ defined through (47) is bounded, for every $p \geq 2$. (In the following, we let $c_2(A, p)$ denote its operator norm.)*

The claim follows from Young's convolution inequality with

$$\begin{aligned}
\|Q\eta\|_{L^p(\mathbb{R}^-,H)} &= \|K * \eta\|_{L^p(\mathbb{R}^-,H)} \\
&\leq \|K\|_{L^1(\mathbb{R}^-,\mathcal{L}(H))} \|\eta\|_{L^p(\mathbb{R}^-,H)} \leq \frac{1}{\delta_0} \|\eta\|_{\mathcal{L}^{p,-}}.
\end{aligned} \tag{49}$$

In the last step we used that for $p \geq 2$ the $\mathcal{L}^{p,-}$ norm dominates the $L^p(\mathbb{R}^-, H)$ norm.

Step 4 *We prove the lemma.*

The estimates of Step 2 and Step 3 imply that

$$\begin{aligned}
\|\zeta\|_{\mathcal{Z}^{p,-}} &\leq c_1(A, p)\big(\|\mathcal{D}_A^\delta \zeta\|_{\mathcal{L}^{p,-}} + \|\zeta\|_{L^p(\mathbb{R}^-,H)}\big) \\
&\leq c_1(A, p)\big(\|\mathcal{D}_A^\delta \zeta\|_{\mathcal{L}^{p,-}} + \|Q\mathcal{D}_A^\delta \zeta\|_{L^p(\mathbb{R}^-,H)} + \|\zeta - Q\mathcal{D}_A^\delta \zeta\|_{L^p(\mathbb{R}^-,H)}\big) \\
&\leq c_1(A, p)\big((1 + c_2(A, p))\|\mathcal{D}_A^\delta \zeta\|_{\mathcal{L}^{p,-}} + \|\zeta - Q\mathcal{D}_A^\delta \zeta\|_{L^p(\mathbb{R}^-,H)}\big).
\end{aligned}$$

This shows (45) because the operator

$$R := \mathbb{1} - Q\mathcal{D}_A^\delta \colon \mathcal{Z}^{p,-} \to L^p(\mathbb{R}^-, H)$$

has finite rank (of dimension equal to $\dim \ker \mathcal{D}_A^\delta$) and therefore is compact. To prove surjectivity we note that the operator $\mathcal{D}_A^\delta$ has closed range by (45) and the abstract closed range lemma (cf. [15, p. 14]). Hence it suffices to show that $\operatorname{ran} \mathcal{D}_A^\delta$ is dense in $\mathcal{L}^{p,-}$. Let $\eta \in \mathcal{L}^{p,-} \cap \mathcal{L}^{2,-}$ be given. The latter is a dense subspace of $\mathcal{L}^{p,-}$ because it contains all compactly supported smooth functions. By Step 1 there exists some $\zeta \in \mathcal{Z}^{2,-}$ such that $\mathcal{D}_A \xi = \eta$. From (45) and the assumption $\eta \in \mathcal{L}^{p,-}$ it follows that $\xi \in \mathcal{Z}^{p,-}$ which implies surjectivity. Again (45) together with finiteness of the rank of R shows that the kernels of the operators $\mathcal{D}_A^\delta \colon \mathcal{Z}^{2,-} \to \mathcal{L}^{2,-}$ and $\mathcal{D}_A^\delta \colon \mathcal{Z}^{p,-} \to \mathcal{L}^{p,-}$ coincide, for all $p \geq 2$. Hence the assertion on the kernel follows from Step 1. This finishes the proof of the lemma. $\qquad \square$

Lemma B.3 *Let $s \mapsto A(s)$, $s \in \mathbb{R}^-$, be a smooth solution of (10) such that for a Yang–Mills connection A^- the asymptotic condition $\lim_{s \to -\infty} A(s) = A^-$ is satisfied in $C^1(\Sigma)$. Let $\delta > 0$ be such that the operator $\mathcal{H}_A + \delta$ is*

injective. Then the operator $\mathcal{D}_A^\delta \colon \mathcal{Z}^{p,-} \to \mathcal{L}^{p,-}$ *associated with A is surjective and has finite-dimensional kernel of dimension*

$$\dim \ker \mathcal{D}_A^\delta = \operatorname{ind} A^-. \tag{50}$$

(The integer $\operatorname{ind} A$ *denoting the number of negative eigenvalues of $\mathcal{H}_A + \delta$).*

Proof: The proof is divided into three steps.

Step 1 *Stationary case.*

Consider the case where the path $A \equiv A^-$ is stationary. In this case Lemma B.2 applies and yields surjectivity of $\mathcal{D}_A^\delta$ and formula (50). In particular, $\mathcal{D}_A^\delta$ is a Fredholm operator.

Step 2 *Nearby case.*

Let us assume that for some sufficiently small $\varepsilon > 0$ the condition

$$\|A - A^-\|_{\mathcal{C}^0(\mathbb{R}^-, \mathcal{C}^1(\Sigma))} < \varepsilon \tag{51}$$

is satisfied. We here consider A^- as a stationary connection over $\mathbb{R}^- \times \Sigma$. Surjectivity and the Fredholm index are preserved under small perturbations with respect to the operator norm. By Lemma B.1, the operator norm of $\mathcal{D}_A^\delta$ depends continuously on A with respect to the $C^0(\mathbb{R}^-, C^1(\Sigma))$ topology. Therefore surjectivity is implied by assumption (51). As the Fredholm indices of $\mathcal{D}_{A^-}^\delta$ and $\mathcal{D}_A^\delta$ coincide it follows that (50) holds true in the nearby case.

Step 3 *General case.*

The general case can be reduced to the nearby case by a standard argument as e.g. carried out in the proof of [22, Proposition 8.3]. $\qquad\square$

Estimates involving $\mathcal{D}_x^\delta$

The following lemma gives a basic estimate for the operator $\mathcal{D}_x^\delta$ in (22).

Lemma B.4 *Let $p \geq 2$ and $x \in C^\infty(\mathbb{R}^1 \times S^1, G)$ such that $\lim_{s \to \infty} x(s) = x^+$ holds in $C^1(S^1)$ for some $x^+ \in C^\infty(S^1, G)$. Then for every sufficiently small $\delta > 0$ there exist positive constants c and T, which depend only on x, p, and δ, such that for every $\xi \in \mathcal{Z}^{p,+}$ the estimate*

$$\|\xi\|_{\mathcal{Z}^{p,+}} \leq c\big(\|\mathcal{D}_x^\delta \xi\|_{\mathcal{L}^{p,+}} + \|\xi\|_{L^p([0,T], L^2(S^1))} + \|\xi(0)\|_{L^2(S^1)}\big) \tag{52}$$

is satisfied.

Proof: We start with the standard parabolic estimate

$$\|\xi\|_{\mathcal{Z}^{p,+}} \leq c(p,x)\big(\|\mathcal{D}_x^\delta \xi\|_{\mathcal{L}^{p,+}} + \|\xi\|_{L^p(\mathbb{R}^+,L^2(S^1))}\big) \tag{53}$$

as obtained (for any $p \geq 2$) in Step 3 of the proof of [22, Theorem 8.5]. To prove the lemma, it therefore remains to further estimate the last term in (53), which we split into integrals over $[0,T]$ and $[T,\infty)$ for sufficiently large $T > 0$. We apply Lemma B.6 to the self-adjoint operator $L := H_{x^+} - \delta$ on the Hilbert space $L^2(S^1)$, which has spectrum bounded away from 0 for sufficiently small $\delta > 0$. Defining $\mathcal{D}_{x^+}^\delta := \frac{d}{ds} + L$ this yields for a constant $c(x^+, p)$ the estimate

$$\|\xi\|_{L^p([T,\infty),L^2(S^1))} \leq c(x^+,p)\big(\|\mathcal{D}_{x^+}^\delta \xi\|_{L^p([T,\infty),L^2(S^1))} + \|\xi(T)\|_{L^2(S^1)}\big), \tag{54}$$

for all $\xi \in L^p([T,\infty), L^2(S^1))$. Denoting by $\mathcal{L}$ the space of bounded linear maps $W^{1,2}(S^1) \to L^2(S^1)$ and by $\|\cdot\|_\mathcal{L}$ the corresponding operator norm it follows for $s \geq T$ the estimate

$$\|\mathcal{D}_{x^+}^\delta \xi(s) - \mathcal{D}_{x(s)}^\delta \xi(s)\|_{L^2(S^1)} = \|H_{x^+}\xi(s) - H_{x(s)}\xi(s)\|_{L^2(S^1)}$$
$$\leq \|H_{x^+} - H_{x(s)}\|_\mathcal{L}\|\xi(s)\|_{W^{1,2}(S^1)}.$$

Thus we can further estimate the term $\mathcal{D}_{x^+}^\delta \xi$ in (54) as

$$\|\mathcal{D}_{x^+}^\delta \xi\|_{L^p([T,\infty),L^2(S^1))}$$
$$\leq \quad \|\mathcal{D}_{x^+}^\delta \xi - \mathcal{D}_x^\delta \xi\|_{L^p([T,\infty),L^2(S^1))} + \|\mathcal{D}_x^\delta \xi\|_{L^p([T,\infty),L^2(S^1))}$$
$$\leq \quad \|H_{x^+} - H_x\|_{L^\infty([T,\infty),\mathcal{L})}\|\xi\|_{L^p([T,\infty),W^{1,2}(S^1))} + \|\mathcal{D}_x^\delta \xi\|_{L^p([T,\infty),L^2(S^1))}.$$

Using the assumption $\lim_{s\to\infty} x(s) = x^+$ in $C^1(S^1)$, it can be checked that $\|H_{x^+} - H_x\|_{L^\infty([T,\infty),\mathcal{L})} \to 0$ as $T \to \infty$. Hence the term involving $H_{x^+} - H_x$ can be absorbed in the left-hand side of (52) for $T = T(x)$ sufficiently large. Here we use that the norm of $\mathcal{Z}^{p,+}$ dominates that of $L^p(\mathbb{R}^+, W^{1,2}(S^1))$. Furthermore, the term $\|\mathcal{D}_x^\delta \xi\|_{L^p([T,\infty),L^2(S^1))}$ is controlled by $\|\mathcal{D}_x^\delta \xi\|_{\mathcal{L}^{p,+}}$ as is clear from the assumption $p \geq 2$. The desired estimate now follows after applying Lemma B.7 (with $L(s) := H_{x(s)} - \delta$ and $\varepsilon > 0$ sufficienty small) to the remaining term $\|\xi(T)\|_{L^2(S^1)}$ in (54). This introduces a further term $\varepsilon\|H_x\xi\|_{L^2([0,T],L^2(S^1))}$ which can be absorbed in the left-hand side of the asserted inequality (52), and a term $\|\mathcal{D}_x^\delta \xi\|_{L^2([0,T],L^2(S^1))} + \|\xi\|_{L^2([0,T],L^2(S^1))}$ which for $p \geq 2$ is dominated by the term $\|\mathcal{D}_x^\delta \xi\|_{L^p([0,T],L^p(S^1))} + \|\xi\|_{L^p([0,T],L^2(S^1))}$ appearing on the right hand side of (52). $\qquad \square$

45

Lemma B.5 *Let $s \mapsto x(s)$, $s \in \mathbb{R}^+$, be a smooth solution of* (4) *such that for a closed geodesic x^+ the asymptotic condition $\lim_{s \to \infty} x(s) = x^+$ is satisfied in $C^1(S^1)$. Then the operator $\mathcal{D}_x^\delta \colon \mathcal{Z}^{p,+} \to \mathcal{L}^{p,+}$ associated with x is surjective and has finite-dimensional cokernel of dimension*

$$\dim \operatorname{coker} \mathcal{D}_x^\delta = \operatorname{ind} x^+.$$

(The integer $\operatorname{ind} x^+$ denoting the number of negative eigenvalues of $H_{x^+} - \delta$).

Proof: The proof of an analogous result in [22, Proposition 8.3] for the backward halfcylinder $\mathbb{R}^- \times S^1$ carries over to the present situation by taking the adjoint of $\mathcal{D}_x^\delta$ and time-reversal $s \mapsto -s$. $\qquad\square$

Further linear estimates

Lemma B.6 *Let H be a real Hilbert space and $L \colon \operatorname{dom}(L) \to H$ be the infinitesimal generator of a strongly continuous one-parameter semigroup on H. We assume that the spectrum of L is contained in $(-\infty, -\lambda] \cup [\lambda, \infty)$ for some $\lambda > 0$. Let $p \geq 1$ and $\eta \in L^p(\mathbb{R}^+, H)$. Then any solution $\xi \colon \mathbb{R}^+ \to H$ of the equation $\dot\xi + L\xi = \eta$ satisfies the estimate*

$$\|\xi\|_{L^p(\mathbb{R}^+, H)} \leq \frac{2}{\lambda} \|\eta\|_{L^p(\mathbb{R}^+, H)} + \frac{1}{\sqrt[p]{\lambda p}} \|\xi_0\|_H.$$

Here we denote $\xi_0 := \xi(0)$.

Proof: Similar to the proof of Lemma we can construct a right-inverse Q to the operator $\frac{d}{ds} + L$ by convolution with a kernel as in (46). Then $\|Q\eta\|_{L^p(\mathbb{R}^+, H)} \leq \frac{2}{\lambda} \|\eta\|_{L^p(\mathbb{R}^+, H)}$ as follows by applying Young's inequality as in (49). Furthermore, the $L^p(\mathbb{R}^+, H)$ norm of elements ξ in the kernel of $\frac{d}{ds} + L$ is bounded above by

$$\left(\int_0^\infty e^{-\lambda p s} \|\xi_0\|_H^p \, ds \right)^{\frac{1}{p}} = \frac{1}{\sqrt[p]{\lambda p}} \|\xi_0\|.$$

The claim then follows. $\qquad\square$

Lemma B.7 *Let H be a Hilbert space. Assume $\xi \colon [s_0, s_1] \to H$ satisfies the equation $\dot\xi + L\xi = \eta$ for a path $s \mapsto L(s)$ of (densely defined) linear operators on H. Then for every $\varepsilon > 0$ there holds the estimate*

$$\|\xi(s_1)\|_H^2 - \|\xi(s_0)\|_H^2$$
$$\leq (1 + \varepsilon^{-1}) \|\xi\|_{L^2([s_0, s_1], H)}^2 + \varepsilon \|L\xi\|_{L^2([s_0, s_1], H)}^2 + \|\eta\|_{L^2([s_0, s_1], H)}^2.$$

Proof: We integrate the equation

$$\frac{d}{ds}\frac{1}{2}\|\xi(s)\|_H^2 = \langle \xi(s), \eta(s)\rangle - \langle \xi(s), L(s)\xi(s)\rangle$$

over the interval $[s_0, s_1]$ and apply to the two terms on the right-hand side the Cauchy-Schwarz inequality. This immediately yields the result. □

The following is a general interpolation lemma for operators of type $D = \frac{d}{ds} + L(s)$.

Lemma B.8 (Interpolation Lemma) *Let $V \subseteq H \subseteq V^*$ be a Gelfand triple. Assume that the family $L(s)\colon V \to V^*$ ($s \in \mathbb{R}^-$) of operators satisfies for all $\xi \in V$ the uniform bound*

$$\|\xi\|_V^2 \le c_1\langle L(s)\xi, \xi\rangle + c_2\|\xi\|_H^2$$

for constants $c_1, c_2 > 0$. Then for any $\varepsilon > 0$ and $p \ge 2$ there holds the estimate

$$\frac{1}{p}\|\xi(0)\|_H^p + \Big(\frac{1}{c_1} - \frac{1}{2\varepsilon}\Big)\int_{-\infty}^0 \|\xi(s)\|_H^{p-2}\|\xi(s)\|_V^2\, ds$$

$$\le \Big(\frac{\varepsilon(p-2)}{2p} + \frac{c_2}{c_1}\Big)\|\xi\|_{L^p(\mathbb{R}^-,H)}^p + \frac{\varepsilon}{p}\|\dot\xi + L\xi\|_{L^p(\mathbb{R}^-,V^*)}^p$$

for all $\xi \in W^{1,p}(\mathbb{R}^-, V^*) \cap L^p(\mathbb{R}^-, V)$.

Proof: Set $\eta := \dot\xi + L\xi \in L^p(\mathbb{R}^-, V^*)$. For every $s \in \mathbb{R}^-$ there holds the estimate

$$\frac{1}{p}\frac{d}{ds}\|\xi(s)\|_H^p = \|\xi(s)\|_H^{p-2}\langle \dot\xi(s), \xi(s)\rangle_H$$

$$= \|\xi(s)\|_H^{p-2}\langle \eta(s) - L(s)\xi(s), \xi(s)\rangle_H$$

$$\le \|\xi(s)\|_H^{p-2}\Big(\|\eta(s)\|_{V^*}\|\xi(s)\|_V - \frac{1}{c_1}\|\xi(s)\|_V^2 + \frac{c_2}{c_1}\|\xi(s)\|_H^2\Big)$$

$$\le \Big(\frac{1}{2\varepsilon} - \frac{1}{c_1}\Big)\|\xi(s)\|_H^{p-2}\|\xi(s)\|_V^2 + \frac{\varepsilon}{2}\|\xi(s)\|_H^{p-2}\|\eta(s)\|_{V^*}^2 + \frac{c_2}{c_1}\|\xi(s)\|_H^p,$$

47

for any constant $\varepsilon > 0$. Integrating this inequality over $\mathbb{R}^-$ and applying Hölder's inequality yields

$$
\begin{aligned}
\frac{1}{p}\|\xi(0)\|_H^p &= \int_{-\infty}^0 \frac{1}{p}\frac{d}{ds}\|\xi(s)\|_H^p \, ds \\
&\leq \int_{-\infty}^0 \big(\frac{1}{2\varepsilon} - \frac{1}{c_1}\big)\|\xi(s)\|_H^{p-2}\|\xi(s)\|_V^2 + \frac{\varepsilon}{2}\|\xi(s)\|_H^{p-2}\|\eta(s)\|_{V^*}^2 + \frac{c_2}{c_1}\|\xi(s)\|_H^p \, ds \\
&\leq \int_{-\infty}^0 \big(\frac{1}{2\varepsilon} - \frac{1}{c_1}\big)\|\xi(s)\|_H^{p-2}\|\xi(s)\|_V^2 \, ds + \frac{c_2}{c_1}\int_{-\infty}^0 \|\xi(s)\|_H^p \, ds \\
&\quad + \frac{\varepsilon}{2}\Big(\int_{-\infty}^0 \|\xi(s)\|_H^p \, ds\Big)^{\frac{p-2}{p}} \cdot \Big(\int_{-\infty}^0 \|\eta(s)\|_{V^*}^p \, ds\Big)^{\frac{2}{p}}.
\end{aligned}
$$

We now apply Young's inequality to the product term in the last line. The claim then follows. $\qquad\square$

References

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

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

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

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

[5] T. J. Davies, The Yang–Mills functional over Riemann surfaces and the loop group, Ph.D. thesis. University of Warwick (1996)

[6] S. Dostoglou, D. A. Salamon, Self-dual instantons and holomorphic curves, Ann. of Math. **139**, 581-640 (1994)

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

[8] T. Friedrich, L. Habermann, Yang–Mills Equations on the Two-Dimensional Sphere, Comm. Math. Phys. **100**, 231-243 (1985)

[9] J. Gravesen, Loop groups and Yang–Mills theory in dimension two, Comm. Math. Phys. **127**, 597-605 (1990)

[10] S. Jarvis, P. Norbury, Degenerating metrics and instantons on the four-sphere, J. Geometry and Physics **27**, 79-98 (1998)

[11] 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, 239-283. American Mathematical Society, Providence (2004)

[12] W. Klingenberg, Closed geodesics on Riemannian manifolds, Regional Conference Series in Mathematics, Number 53. American Mathematical Society, Providence (1983)

[13] T. Mrowka, K. Wehrheim, L^2 topology and Lagrangians in the space of connections over a Riemann surface, Geom. Funct. Anal. **20**, 1278-1305 (2010)

[14] M. Reed, B. Simon, Methods of modern mathematical physics, II, Fourier analysis, self-adjointness. Academic Press, New York, London (1975)

[15] 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**, 143-230. American Mathematical Society, Providence (1999)

[16] D. A. Salamon, J. Weber, Floer homology and the heat flow, Geom. Funct. Anal. **16**, 1050-1138 (2006)

[17] M. Schwarz, Morse Homology. Birkhäuser Verlag, Basel (1993)

[18] J. Swoboda, The Yang–Mills Gradient Flow and Loop Groups, Diss. ETH no. 18296. ETH Zürich (2009)

[19] J. Swoboda, Morse Homology for the Yang–Mills Gradient Flow, J. Math. Pures Appl. **98**, 160-210 (2012)

[20] K. Wehrheim, Uhlenbeck Compactness, EMS Series of Lectures in Mathematics. European Mathematical Society Publishing House, Zürich (2004)

[21] J. Weber, Perturbed closed geodesics are periodic orbits: index and transversality, *Math. Z.* **241**, 45-82, 2002.

[22] J. Weber, The heat flow and the homology of the loop space, Habilitationsschrift, Humboldt-Universität zu Berlin (2010)