Calc. Var. 2, 123–150 (1994)

The Yang-Mills flow in four dimensions

Michael Struwe *

ETH Zürich, Mathematik, CH-8092 Zürich, Switzerland

Received January 25, 1993/Accepted June 4, 1993

Abstract. Global existence and uniqueness is established for the Yang-Mills heat flow in a vector bundle over a compact Riemannian four-manifold for given initial connection of finite energy. Our results are analogous to those valid for the evolution of harmonic maps of Riemannian surfaces.

Mathematics Subject Classification: 35K22, 35K55, 35K65, 53C05

1. Introduction

Let M be a compact connected Riemannian 4-manifold, $\pi\colon \eta \to M$ a smooth vector bundle with fibre $\pi^{-1}(x) \cong \mathbf{R}^n$ and structure group $G \subset SO(n)$. η carries a natural metric induced by the Riemannian metric on M and the scalar product in $\mathbf{R}^n$. We consider the evolution of connections D on η to Yang-Mills connections by the L^2-gradient flow. Introducing this flow requires some terminology. The initiated reader may wish to skip the remainder of this section and go directly to the statement of our main result Theorem 2.3.

1.1. Associated bundles

There exists a cover (U_α) of M and diffeomorphisms, called local trivializations, $\sigma_\alpha\colon \eta|_{U_\alpha} \to U_\alpha \times \mathbf{R}^n$ with transition functions $S_{\alpha\beta}(x) = \sigma_\alpha\left(x, \sigma_\beta(x,\cdot)^{-1}\right) \in G$ at any point $x \in U_\alpha \cap U_\beta$ and such that

$$S_{\alpha\beta} \circ S_{\beta\gamma} \circ S_{\gamma\alpha} = id$$

on $U_\alpha \cap U_\beta \cap U_\gamma$, for all values α, β, γ.

The collection $(S_{\alpha\beta})$ defines the bundle structure in the large.

On η we have the group $\mathscr{G}$ = Aut η of gauge transformations S, where Aut $\eta = \bigcup_{x\in M}$ Aut $_x$ is the automorphism bundle associated to η. If we agree that the structure group G acts on the fibres of η by multiplication from the right, then any $S \in \mathscr{G}$ locally may be represented by a map $S\colon U_\alpha \to G$, acting on η by fibre-wise multiplication from the left. Moreover, G naturally acts on $\mathscr{G}$ by conjugation. Similarly, if we denote by ad η the adjoint bundle, whose sections $s \in \Omega^0(\text{ad}\ \eta)$ locally

* The author gratefully acknowledges the hospitality and support of Princeton University, the Institute for Advanced Study, and the École Normale Supérieure Cachan.

may be represented by maps $s: U_\alpha \to \mathfrak{g}$, the Lie algebra of G, then G acts on ad η via the adjoint action, that is, the differential of conjugation. We regard ad η as the Lie algebra of $\mathscr{G}$.

By taking tensor products with T^*M, etc., from $\Omega^0(\text{ad }\eta)$ we obtain the spaces $\Omega^1(\text{ad }\eta)$, whose sections are locally represented as $\mathfrak{g}$-valued one-forms, $\Omega^2(\text{ad }\eta)$, represented by $\mathfrak{g}$-valued 2-forms, etc., and the exterior product and Hodge star-operation extend to these spaces.

1.2. Connections

There are various equivalent ways of introducing a connection on a principal or vector bundle η; see for instance Donaldson-Kronheimer [6] or Jost [10] for a very readable exposition.

In particular, a connection is related to a covariant derivative ∇ and gives rise to an exterior differential operator D on forms taking their values in η. D is the completely anti-symmetric part of ∇. ∇ extends the Levi-Cività connection on M while D extends the exterior derivative d acting on standard differential forms with scalar coefficients. Moreover, ∇ and D naturally extend to associated bundles $\Omega^i(\text{ad }\eta)$. In particular, there holds

$$D(A \cdot s) = (DA) \cdot s - A \cdot Ds$$

for $A \in \Omega^1(\text{ad }\eta)$, $s \in \Omega^0(\text{ad }\eta)$, where we regard the fibres of ad η as subsets of the space of linear endomorphisms of the fibres of η. The change in sign in the second term on the right is due to the fact that we interchange the order of the exterior derivative and the 1-form A.

The space of connections is an affine space

$$\mathscr{D} = \left\{D = D_{\text{ref}} + A;\ A \in \Omega^1(\text{ad }\eta)\right\},$$

where D_{ref} is the exterior derivative related to some smooth reference connection ∇_{ref} which we will assume to be fixed throughout the rest of this paper.

1.3. Sobolev spaces

Using ∇_{ref} we can define Sobolev spaces of connections, etc., in the same way as for functions on a Euclidean domain. For instance, $H^{l,p}(\Omega^i(\text{ad }\eta))$ consists of i-forms A in the adjoint bundle with measurable coefficients such that

$$\|A\|_{H^{l,p}} = \left(\sum_{k=0}^{l} \|\nabla^k_{\text{ref}} A\|^p_{L^p}\right)^{\frac{1}{p}} < \infty .$$

It is easily verified that different choices of reference connection lead to the same spaces with equivalent norms. A connection D is of class $H^{1,2}$ iff $D = D_{\text{ref}} + A$, where $A \in H^{1,2}(\Omega^1(\text{ad }\eta))$.

Some care is required when we speak of gauge transformations belonging to some Sobolev space, as in this case we are dealing with a nonlinear range $G \subset SO(n)$. Regarding the latter as a subset of $\mathbf{R}^{n\times n}$, we may define

$$H^{l,p}(\mathscr{G}) = \left\{S; S|_{U_\alpha} \in H^{l,p}(U_\alpha; \mathbf{R}^{n\times n}),\ S(x) \in G \quad \text{almost everywhere}\right\} .$$

Note that this space is a manifold, modelled on $H^{l,p}(\Omega^0(\text{ad }\eta))$, and $\mathcal{G}$ is dense in $H^{l,p}(\mathcal{G})$, if $lp > \dim M = m$. Moreover, in general $\mathcal{G}$ is not dense in $H^{l,p}(\mathcal{G})$ for $lp < m$. (See Bethuel [2] for a general discussion of Sobolev spaces of maps between manifolds.)

1.4. Curvature

For $D \in \mathcal{D}$ the curvature $F = F(D)$ is a 0-order operator

$$F = D \circ D \in \Omega^2(\text{ad }\eta).$$

F satisfies the first and second Bianchi identities

$$DF = 0$$

and

$$D^*D^*F = 0.$$

1.5. Gauge transformations

The group of gauge transformations acts on connections as follows. For $S \in \mathcal{G}$, $D \in \mathcal{D}$ let

$$S^*(D) = S^{-1} \circ D \circ S$$

denote the pull-back connection under S with curvature

$$F\big(S^*(D)\big) = \big(S^*(D)\big)^2 = S^{-1} \circ D \circ D \circ S = S^{-1} \circ F \circ S.$$

For later use we note that this yields

$$\left.\frac{dS^*(D)}{dS}\right|_{S=id}(s) = -s \circ D + D \circ s = Ds \tag{1}$$

for any $s \in \Omega^0(\text{ad }\eta)$.

1.6. Yang-Mills connections

Finally, we introduce the Yang-Mills action

$$\text{YM }(D) = \frac{1}{2}\int_M |F|^2\, dx$$

of a conection D with curvature $F = F(D)$.

Because of the identity

$$F(D + ta) = (D + ta) \circ (D + ta) = F(D) + tDa + t^2 a \wedge a \tag{2}$$

for any $a \in \Omega^1(\text{ad }\eta)$, any $t > 0$, we have the relation

$$\left.\frac{d}{dt}\text{YM }(D + ta)\right|_{t=0} = (F, Da) = (D^*F, a)$$

for the first variation of YM at D in direction a. (We simultaneously use $(\cdot,\cdot)$ to denote the metric on η and the L^2-product.)

A connection D is *Yang-Mills* iff it is a critical point of YM, which then is equivalent to the equation

$$D^*F = 0\,.$$

In order to obtain Yang-Mills connections on any given bundle η, a natural approach is to try to follow the lines of steepest descent for YM , given by the evolution equation

$$\frac{d}{dt}D = -D^*F\,, \tag{3}$$

starting from any initial connection

$$D(0) = D_0\,. \tag{4}$$

This approach, standard in other problems, was first suggested for the Yang-Mills functional by Atiyah-Bott [1] but not followed. However, the flow (3) plays a fundamental role in Donaldson's work; see [6].

Remark that YM is gauge-invariant in the sense that YM $(S^*(D)) =$ YM (D) for any $S \in \mathscr{G}$, any $D \in \mathscr{D}$. Hence $\mathscr{G}$ acts on the solutions of the Yang-Mills equation which, therefore, is not elliptic, as the kernel of the linearized equation has infinite dimension.

Similarly, the evolution problem (3) is not parabolic, and the methods developed for parabolic equations cannot be directly applied to prove existence and uniqueness for the Cauchy problem.

For further background material see Donaldson-Kronheimer [6], Freed-Uhlenbeck [7], Jost [10], or Lawson [12].

2. Statement and discussion of main results

Let $\pi : \eta \to M^4$ be as above.

Definition 2.1. *A family $D = D(t)$ of connections on η is a weak solution of (3) if $D = D_{ref} + A$ with*

$$A \in L^1\Big([0,T[\,; L^2\big(\Omega^1(ad\,\eta)\big)\Big)\,,$$

$$D \circ D = F \in L^\infty\Big([0,T[\,; L^2\big(\Omega^2(ad\,\eta)\big)\Big)\,,$$

and if for any $\phi \in C^\infty\big([0,T]; \Omega^1(ad\,\eta)\big)$ vanishing near $t = 0$ and $t = T$ there holds

$$\int_0^T \left\{\left(A, \frac{d}{dt}\phi\right) - (F, D\phi)\right\} dt = 0\,.$$

Remark 2.2. In order to form $D \circ D = F$ in the distribution sense we need $A(t) \in L^2$ for almost every t. Similarly, as the product $(F, D\phi)$ involves the product of F and A, we need to require $F(D(t)) \in L^2$ almost everywhere and $A \in L^p(L^2)$, $F \in L^q(L^2)$ with conjugate exponents $1 \le p, q \le \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. The above notion of weak solution then is the weakest possible one which is compatible with the energy inequality (12) below.

Theorem 2.3. *(i) For any connection D_0 of class $H^{1,2}$ on η there is $T > 0$ and a weak solution $D = D_{ref} + A$ to the Yang-Mills evolution problem* (3), (4) *for $0 \le t < T$ such that*

$$A \in C^0\Big([0,T]; L^2\big(\Omega^1(ad\,\eta)\big)\Big) \cap H^{1,2}\Big([0,T]; L^2\big(\Omega^1(ad\,\eta)\big)\Big)\,,$$

$$F \in C^0\Big([0,T[\,; L^2\big(\Omega^2(ad\,\eta)\big)\Big)\,.$$

Moreover, D is gauge-equivalent to a smooth solution of (3) *in the following sense: There is a solution $\hat{D} = D_{\text{ref}} + \hat{A}$ of* (3) *with*

$$\hat{A} = H^{1,2}\Big([0,T]; L^2\big(\Omega^1(ad\,\eta)\big)\Big) \cap C^0\Big([0,T[\,; L^2\big(\Omega^1(ad\,\eta)\big)\Big)$$

and smooth for $0 < t < T$, and a sequence of smooth gauge transformations $\hat{S}_k \in \mathcal{G}$ and a sequence $t_k \searrow 0$ such that $\hat{S}_k \to \hat{S}_0$ in $H^{1,2}$, $\hat{S}_k^*\big(\hat{D}(t_k)\big) \to D_0$ in $H^{1,2}$, and $D = \hat{S}_0^*(\hat{D})$. D is smooth if D_0 is smooth (C^∞).

(ii) If D is irreducible in the sense of (26) for all t, then D is unique.

(iii) The maximal existence time T is characterized by

$$T = \sup\left\{\bar{t} > 0; \exists R > 0 : \sup_{\substack{x_0\in M\\ 0\le t\le \bar{t}}} \left(\int_{B_R(x_0)} |F(t)|^2\,dx\right) < \epsilon_0\right\} \tag{5}$$

where $\epsilon_0 = \epsilon_0(\eta) > 0$. At $\bar{t}_1 = T$, curvature concentrates in at most finitely many points $\bar{x}_1^j$, $j = 1, \dots, J_1$ in the sense that

$$\forall R > 0\colon \limsup_{t\nearrow \bar{t}_1} \int_{B_R\left(\bar{x}_1^j\right)} |F(t)|^2\,dx \ge \epsilon_0\,.$$

Concerning the long-time behavior we claim the following.

Theorem 2.4. *(i) At each $\bar{x}_1^j = \bar{x}$ a non-trivial Yang-Mills connection over S^4 separates in the sense that for sequences $R_k \searrow 0$, $x_k \to \bar{x}$, $t_k \nearrow \bar{t}_1$ the rescaled connections*

$$A_k(x) := R_k A(x_k + R_k x, t_k) \to \bar{A} \text{ in } H^{2,2}_{loc}(\mathbf{R}^4; \mathfrak{g}),$$

as $k \to \infty$, where $d + A(\cdot, t)$ is the expression of $D(t)$ in a local trivialization of η near $\bar{x}$, and $\bar{D} = d + \bar{A}$ extends to a smooth, non-trivial, finite-energy Yang-Mills connection on $\mathbf{R}^4 \cup \{\infty\} \cong S^4$.

Moreover, in a suitable gauge $D(t)$ converges in

$$H^{1,2}_{loc}\left(M \setminus \left\{\bar{x}_1^1, \dots, \bar{x}_1^{J_1}\right\}\right)$$

to a limiting connection $D_1 \in H^{1,2}$ on a G-bundle η_1 over M.

(ii) By iteration, the solution D may be extended uniquely for all time t, having the properties listed in Theorem 2.3 for all but finitely many times $\bar{t}_k$, $k = 1, \dots, K$, and with curvature concentrating in at most finitely many points $\left(\bar{x}_k^j, \bar{t}_k\right)$, $1 \le j \le J_k$, $1 \le k \le K$, $\sum_k J_k \le YM\,(D_0)\epsilon_0^{-1}$.

(iii) As $t \to \infty$ suitably, $D(t)$ converges to a limiting connection D_∞ in a suitable gauge and possibly away from finitely many concentration points $(\bar{x}_\infty^j)$, $1 \le j \le J_\infty$, $J_\infty \le C\,YM\,(D_0)$, where non-trivial Yang-Mills connections over S^4 separate. D_∞ extends to a smooth Yang-Mills connection on a limit bundle η_∞ over M.

Remarks: (i) The Yang-Mills evolution problem resembles the evolution problem for harmonic maps between Riemannian manifolds. The four-dimensional case is critical for the Yang-Mills problem in the same way as the two-dimensional case is critical for harmonic maps. For smooth initial data local existence was proved by Donaldson. Moreover, in the case of a holomorphic vector bundle, Donaldson was able to show global existence of solutions to (3), and their asymptotic convergence to a limiting Yang-Mills connection, if the underlying bundle was stable. Finally, he asserted that on a general bundle over a Kähler surface uniform smallness of the local L^2-norms of the curvature should imply that a solution to (3) can be globally extended. See Donaldson-Kronheimer [6], p. 236.

In the absence of a holomorphic structure, however, in general blow-up in finite time might be expected, as in the analogous case of the evolution of harmonic maps of surfaces; see Chang-Ding-Ye [4]. Thus, one is naturally led to consider initial data of class $H^{1,2}$ – that is, of finite energy – and corresponding weak solutions. For harmonic maps, the analogue of Theorem 2.3 and 2.4 was obtained by Struwe [18].

Using ideas from [18], together with the results of Uhlenbeck [22] and Sedlaceck [17] on the weak compactness properties of connections, the asymptotic behavior of solutions $D(t)$ to (3) claimed above was obtained by Chen-Shen [5], assuming the existence of a smooth, global solution $D(t)$ to (3) on $0 \leq t < \infty$.

(ii) The Yang-Mills flow over 2- or 3-dimensional manifolds was recently studied by Råde [16], and he obtained global existence and uniqueness of solutions for initial data $D(0) \in H^{1,2}$ and their exponential convergence to Yang-Mills connections as $t \to \infty$.

Råde's method consists in writing (3) as a system

$$\begin{aligned} \frac{d}{dt}D &= -D^*\Omega \\ \frac{d}{dt}\Omega &= -(DD^* + D^*D)\Omega \end{aligned} \tag{6}$$

for D and $\Omega = F(D)$; compare (10) below. If $\dim M \leq 3$, this system can be solved by linear methods. The approach, however, seems to fail for $\dim M \geq 4$.

(iii) In dimensions larger than 4, by analogy with the evolution of harmonic maps in dimensions ≥ 3 uniqueness in the energy class of weak solutions cannot be expected; moreover, blow up in finite time actually occurs; see Naito [14]. In fact, the singular set might be quite large. The recent results of Nakajima [15] on an analogue of Sedlaceks's [17] weak compactness result and the results of Chen-Shen [5], Hamilton [8] and Naito [14] on a monotonicity formula for (3) analogous to this author's [19] monotonicity formula for the harmonic map flow, however, are promising first results towards a better understanding also of the higher dimensional case.

(iv) Our proof of uniqueness apparently can be adapted to show local uniqueness for Hamilton's Ricci flow, provided the initial metric is 'irreducible' in a suitable sense.

3. Preliminaries

3.1. Weizenböck formulae and consequences

Let ∇ be a covariant derivative operator with corresponding exterior differential operator D. Associated with D we define the Hodge Laplacian

$$\Delta = \Delta^H = D^* D + DD^* .$$

Similarly we define the rough (or crude) Laplacian $\nabla^* \nabla$.

Acting on bundle-valued forms ϕ, these operators differ by a curvature term

$$\nabla^* \nabla \phi = \Delta \phi + F \# \phi + \mathrm{Rm} \, \# \phi , \tag{7}$$

where $F = F(D)$ and Rm ist the Riemannian curvature on the base manifold; see for instance Lawson [12], Appendix. Here and in the following, # denotes any multi-linear map with smooth coefficients.

As a consequence of the Weizenböck formula (7) we obtain the following linear estimates.

Lemma 3.1. *Let $D = D_{ref} + A$, $A \in C^1(\Omega^1(ad\ \eta))$. There exist constants $C_1 = C_1(\eta)$, $C_2 = C_2(\eta, \|A\|_{C^1})$ such that for any $\phi \in H^{2,2}(\Omega^i(ad\ \eta))$ there holds*

$$\|\phi\|^2_{H^{2,2}} \le C_1 \|\Delta \phi\|^2_{L^2} + C_2 \|\phi\|^2_{L^2} .$$

We sketch the proof for completeness.

Proof. From (7) we have

$$\|\Delta \phi\|_{L^2} = \|\nabla^* \nabla \phi + F \# \phi + \mathrm{Rm} \, \# \phi\|_{L^2} .$$

Suppose first that $A = 0$, $D = D_{\text{ref}}$, $\Delta = \Delta_{\text{ref}}$. By Minkowski's inequality then we have

$$\|\Delta \phi\|_{L^2} \ge \|\nabla^* \nabla \phi\|_{L^2} - C \|\phi\|_{L^2} ,$$

where $C = \|F\|_{L^\infty} + \|\mathrm{Rm}\ \|_{L^\infty} = C(\eta)$.

Finally, integrating by parts and interchanging derivatives (which introduces a further curvature term) we obtain

$$\|\nabla^* \nabla \phi\|^2_{L^2} \ge \|\nabla^2 \phi\|^2_{L^2} - C(\eta) \|\nabla \phi\|^2_{L^2} \ge \frac{1}{2} \|\nabla^2 \phi\|^2_{L^2} - C(\eta) \|\phi\|^2_{L^2} ,$$

where the last inequality results from interpolation between the L^2- and $H^{2,2}$-norms.

In the general case we estimate

$$\Delta - \Delta_{\text{ref}} = \nabla_{\text{ref}} \# A + A \# A , \tag{8}$$

whence

$$\begin{aligned} \|\Delta \phi\|_{L^2} &\ge \|\Delta_{\text{ref}} \phi\|_{L^2} - \|A \# \nabla_{\text{ref}} \phi\|_{L^2} - \|\nabla_{\text{ref}} A \# \phi\|_{L^2} - \|A \# A \# \phi\|_{L^2} \\ &\ge C(\eta)^{-1} \|\phi\|_{H^{2,2}} - C(\eta, \|A\|_{L^\infty}, \|\nabla A\|_{L^\infty}) \|\phi\|_{L^2} . \end{aligned}$$

□

Lemma 3.1 implies analogous linear estimates for evolution equations. For ease of notation for any $T > 0$ we introduce the space

$$V = V_T(\Omega^i(\text{ad } \eta)) = L^2\Big([0, T]; H^{2,2}(\Omega^i(\text{ad } \eta))\Big) \cap H^{1,2}\Big([0, T]; L^2(\Omega^i(\text{ad } \eta))\Big) .$$

Note that by the relations

$$\frac{1}{2}\frac{d}{dt}(\phi,\phi) = \left(\frac{d}{dt}\phi,\phi\right) \le \left\|\frac{d}{dt}\phi\right\|_{L^2} \|\phi\|_{L^2} \le \frac{1}{2}\left(\left\|\frac{d}{dt}\phi\right\|_{L^2}^2 + \|\phi\|_{L^2}^2\right)$$

and (for $\nabla = \nabla_{\rm ref}$)

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}(\nabla\phi,\nabla\phi) &= \left(\nabla\frac{d}{dt}\phi,\nabla\phi\right) = \left(\frac{d}{dt}\phi,\nabla^*\nabla\phi\right)\\ &\le \left\|\frac{d}{dt}\phi\right\|_{L^2}\|\nabla^2\phi\|_{L^2} \le \frac{1}{2}\left(\left\|\frac{d}{dt}\phi\right\|_{L^2}^2 + \|\nabla^2\phi\|_{L^2}^2\right)\end{aligned}$$

the space V continously embeds into

$$L^\infty\Big([0,T];H^{1,2}\big(\Omega^i(\text{ad }\eta)\big)\Big)$$

with

$$\sup_{0\le t\le T}\|\phi(t)\|_{H^{1,2}}^2 \le \|\phi(0)\|_{H^{1,2}}^2 + 2\,\|\phi\|_V^2 \,. \tag{9}$$

In fact, $V \hookrightarrow C^0\Big([0,T];H^{1,2}\big(\Omega^i(\text{ad }\eta)\big)\Big)$; see for instance Lions-Magenes [13], Theorem 3.1, p. 19.

Here and in the following, we denote

$$L^2(H^{2,2}) = L^2\Big([0,T];H^{2,2}\big(\Omega^i(\text{ad }\eta)\big)\Big)\,,\quad \text{etc.}$$

Moreover, we use double indices to denote space-time L^p-L^q-norms.

$$\|\phi\|_{L^{q,p}} = \left(\int_0^T \|\phi(t)\|_{L^p}^q\,dt\right)^{\frac{1}{q}}\,,\quad 1\le p,q<\infty\,,$$

etc. In particular, $\|\cdot\|_{L^{2,2}}$ denotes the L^2-norm over space-time. Finally, we let

$$\|\phi\|_V^2 := \left\|\frac{d}{dt}\phi\right\|_{L^{2,2}}^2 + \|\phi\|_{L^2(H^{2,2})}^2\,.$$

Lemma 3.2. *Let $D = D_{ref} + A$, $A \in C^1\big(\Omega^1(ad\,\eta)\big)$. Then there exists a constant $C_2 = C_2(\eta)$ and a number $T = T(\eta, A) > 0$ sucht that for any $\phi \in V_T$ there holds*

$$\|\phi\|_V^2 \le C_2\left\|\left(\frac{d}{dt}+\Delta\right)\phi\right\|_{L^{2,2}}^2 + C_2\,\|\phi(0)\|_{H^{1,2}}\,.$$

Proof. Compute

$$\left\|\left(\frac{d}{dt}+\Delta\right)\phi\right\|_{L^2}^2 = \left\|\frac{d}{dt}\phi\right\|_{L^2}^2 + \|\Delta\phi\|_{L^2}^2 + 2\left(\frac{d}{dt}\phi,\Delta\phi\right)$$

for almost every t. By Lemma 3.1 we have

$$\|\Delta\phi\|_{L^2}^2 \ge C(\eta)^{-1}\|\phi\|_{H^{2,2}}^2 - C(\eta,\|A\|_{C^1})\,\|\phi\|_{L^2}^2\,.$$

Moreover, by (7), (8) we have

$$
\begin{aligned}
2\left(\frac{d}{dt}\phi, \Delta\phi\right) &\geq 2\left(\frac{d}{dt}\phi, \Delta_{\text{ref}}\phi\right) - C(\eta, \|A\|_{C^1}) \left\|\frac{d}{dt}\phi\right\|_{L^2} \|\phi\|_{H^{1,2}} \\
&\geq 2\left(\frac{d}{dt}\phi, \nabla^*_{\text{ref}}\nabla_{\text{ref}}\phi\right) - C\left(\eta, \|A\|_{C^1}\right) \left\|\frac{d}{dt}\phi\right\|_{L^2} \|\phi\|_{H^{1,2}} \\
&= \frac{d}{dt}\|\nabla_{\text{ref}}\phi\|^2_{L^2} - C(\eta, \|A\|_{C^1}) \left\|\frac{d}{dt}\phi\right\|_{L^2} \|\phi\|_{H^{1,2}} \\
&\geq \frac{d}{dt}\|\nabla_{\text{ref}}\phi\|^2_{L^2} - \frac{1}{2}\left\|\frac{d}{dt}\phi\right\|^2_{L^2} - C(\eta, \|A\|_{C^1}) \|\phi\|^2_{H^{1,2}} .
\end{aligned}
$$

Thus, upon integrating in time we obtain

$$
\|\phi\|^2_V \leq C(\eta)\left(\left\|\left(\frac{d}{dt}+\Delta\right)\phi\right\|^2_{L^{2,2}} + \|\phi(0)\|^2_{H^{1,2}}\right) + C(\eta, \|A\|_{C_1})\|\phi\|^2_{L^2(H^{1,2})} .
$$

Finally we use (9) to estimate

$$
\|\phi(t)\|^2_{L^2(H^{1,2})} \leq T\,\|\phi\|^2_{L^\infty(H^{1,2})} \leq T\,\|\phi(0)\|^2_{H^{1,2}} + 2T\,\|\phi\|^2_V .
$$

Choosing $T = \frac{1}{4(1+C(\eta,\|A\|_{C^1}))}$, the lemma follows. □

The linear estimates above may fail in the borderline case where $A \in H^{1,2}$, which is the reason why the evolution problem for Yang-Mills connections over a 4-dimensional base manifold is interesting.

However, under certain circumstances L^p-estimates are still available. To obtain these estimates we first observe that for any $\phi, \psi \in \Omega^i(\text{ad}\,\eta)$ we have

$$
d(\phi, \psi) = (\nabla\phi, \psi) + (\phi, \nabla\psi),
$$

where now again $(\cdot, \cdot)$ denotes the pointwise inner product. In particular, for sections $\phi \in \Omega^0(\text{ad}\,\eta)$, where $\nabla\phi = D\phi$, we obtain Kato's inequality

$$
|d\,|\phi|\,| \leq |D\phi| .
$$

We can combine this with the Sobolev embedding theorem to obtain

$$
\|\phi\|_{L^4} \leq C(\|D\phi\|_{L^2} + \|\phi\|_{L^2})
$$

for any $\phi \in \Omega^0(\text{ad}\,\eta)$ with a uniform constant $C = C(\eta)$ independent of D.

To obtain similar estimates for forms of degree $i \geq 1$ we need to compare with full covariant derivatives. Note that on account of (7) this introduces a curvature term which, however, can be absorbed as long as the curvature does not concentrate.

Lemma 3.3. *Let $D = D_{ref} + A$, $A \in H^{1,2}$, with curvature $F = F(D) \in L^2$. There exist constants $C_3 = C_3(\eta)$, $\delta = \delta(\eta) > 0$ such that for any $\phi \in \Omega^i(ad\,\eta)$, any $0 < R < 1$, there holds*

$$
\|\phi\|^2_{L^4} + \|\nabla\phi\|^2_{L^2} \leq C_3\left(\|D\phi\|^2_{L^2} + \|D^*\phi\|^2_{L^2}\right) + C_3 R^{-2}\,\|\phi\|^2_{L^2} ,
$$

provided

$$
\sup_{x_0}\int_{B_R(x_0)} |F|^2\,dx \leq \delta .
$$

Proof. Using (7) and Sobolev's embedding, we find

$$\begin{aligned} C(\eta)^{-1}\|\phi\|_{L^4}^2 - \|\phi\|_{L^2}^2 \le \|\nabla\phi\|_{L^2}^2 &= (\nabla^*\nabla\phi,\phi) = (\Delta\phi,\phi) + (F\#\phi,\phi) + (\mathrm{Rm}\ \#\phi,\phi) \\ &= \|D\phi\|_{L^2}^2 + \|D^*\phi\|_{L^2}^2 + (F\#\phi,\phi) + (\mathrm{Rm}\ \#\phi,\phi)\,. \end{aligned}$$

Now

$$(\mathrm{Rm}\ \#\phi,\phi) \le C\,\|\phi\|_{L^2}^2\,, \quad C = C(\eta)\,.$$

To estimate the middle term we use Sobelev's embedding theorem on a suitable cover of M by balls $B_R(x_i)$:

$$\begin{aligned} (F\#\phi,\phi) &\le C\sum_i \|F\|_{L^2(B_R(x_i))}\|\phi\|_{L^4(B_R(x_i))}^2 \\ &\le C\delta\sum_i \left(\|\nabla\phi\|_{L^2(B_R(x_i))}^2 + R^{-2}\|\phi\|_{L^2(B_R(x_i))}^2\right) \end{aligned}$$

where $C = C(\eta)$. Since M is compact, there exist a constant $R_0 = R_0(M) > 0$ and a number L (independent of M) such that for $0 < R \le R_0$ there is a cover $(B_R(x_i))$ such that at most L distinct balls of this cover overlap at any point of M.

Thus, for $R \le R_0$ and with this choice of $(B_R(x_i))$ the above estimate yields

$$(F\#\phi,\phi) \le CL\delta\left(\|\nabla\phi\|_{L^2}^2 + R^{-2}\|\phi\|_{L^2}^2\right)\,.$$

Choosing $\delta \le \frac{1}{2CL}$, the claim follows. □

3.2. Evolution of curvature and energy inequality

Formula (2) for the first variation of curvature implies that the curvature $F = F(D(t))$ of a classical solution D of (3), (4) satisfies

$$\frac{d}{dt}F = \lim_{\epsilon\to 0}\frac{F\left(D+\epsilon\frac{d}{dt}D\right) - F(D)}{\epsilon} = D\left(\frac{d}{dt}D\right) = -DD^*F\,. \tag{10}$$

In view of the first Bianchi identity, thus we find that F is a solution of the heat equation

$$\left(\frac{d}{dt} + \Delta\right)F = 0 \tag{11}$$

with respect to the evolving connection.

From (11), upon multiplying by F we obtain the identity

$$0 = \frac{1}{2}\frac{d}{dt}\|F\|_{L^2}^2 + (\Delta F, F) = \frac{d}{dt}\mathrm{YM}\ (D) + \|D^*F\|_{L^2}^2\,.$$

In particular, any (classical) solution of (3) and (4) satisfies the energy inequality

$$\mathrm{YM}\ (D(T)) + \int_0^T \|D^*F\|_{L^2}^2\,dt \le \mathrm{YM}\ (D_0)\,; \tag{12}$$

in fact, equality holds. Moreover, multiplying by $F\varphi^2$, where φ is a suitable cut-off function with support in a coordinate ball $B_{2R}(x_0)$ and such that $0 \le \phi \le 1$, $\phi \equiv 1$ on $B_R(x_0)$, $|\nabla\phi| \le CR^{-1}$, we obtain

$$\frac{d}{dt}\int_{B_{2R}(x_0)} |F|^2 \varphi^2\, dx + \int_{B_{2R}(x_0)} |D^*F|^2 \varphi^2\, dx = -2(D^*F, F\varphi D\varphi)$$
$$\le \int_{B_{2R}(x_0)} |D^*F|^2 \varphi^2\, dx + \int_{B_{2R}(x_0)} |F|^2 |\nabla\varphi|^2\, dx\,,$$

which after integration in t yields the local energy inequality

$$\sup_{0\le t\le T}\int_{B_R(x)} |F(t)|^2\, dx \le \int_{B_{2R}(x_0)} |F(0)|^2\, dx + CTR^{-2}\mathrm{YM}\,(D_0) \tag{13}$$

as in [18], Lemma 3.6, for the evolution of harmonic maps.

Next, observe that by Hölder's and Young's inequalities $L^{\infty,2}\cap L^{2,4} \hookrightarrow L^{3,3}$ with

$$\|\phi\|^2_{L^{3,3}} \le \|\phi\|^{\frac{2}{3}}_{L^{\infty,2}}\, \|\phi\|^{\frac{4}{3}}_{L^{2,4}} \le \frac{1}{3}\,\|\phi\|^2_{L^{\infty,2}} + \frac{2}{3}\,\|\phi\|^2_{L^{2,4}} \tag{14}$$

for any ϕ. Combining this estimate with Lemma 3.3, (12) and the first Bianchi identity, we immediately obtain

Lemma 3.4. *Let $\delta = \delta(\eta) > 0$ be as in Lemma 3.3 and suppose D is a classical solution of (3) and (4) on $]0,T[$ with*

$$\sup_{x_0,\,0<t<T}\int_{B_R(x_0)} |F(D(t))|^2\, dx < \delta \tag{15}$$

for some $0 < R \le 1$. Then $F \in L^{3,3}$ with

$$\|F\|^2_{L^{3,3}} \le C_4(1 + TR^{-2})YM\,(D_0)\,,$$

where $C_4 = C(\eta)$.

Next, we use (11) and (7) to deduce

Lemma 3.5. *Under the assumptions of Lemma 3.4 there holds*

$$D^*F \in L^2_{loc}\Big(]0,T];L^4\big(\Omega^1(ad\,\eta)\big)\Big)\,,\quad \frac{d}{dt}F \in L^2_{loc}\Big(]0,T];L^2\big(\Omega^2(ad\,\eta)\big)\Big)\,.$$

Proof. By (11), at any time $t > 0$ we have

$$\begin{aligned}\left\|\frac{d}{dt}F\right\|^2_{L^2} + \|DD^*F\|^2_{L^2} &= -2\left(DD^*F, \frac{d}{dt}F\right)\\ &= -2\left(D^*F, \frac{d}{dt}(D^*F)\right)\\ &\quad +2\left(D^*F, \left(\frac{d}{dt}D\right)\#F\right)\\ &\le -\frac{d}{dt}\,\|D^*F\|^2_{L^2} + 2\,\|F\|_{L^3}\,\|D^*F\|^2_{L^3}\,.\end{aligned}$$

By Lemma 3.3, (15), and the second Bianchi identity there holds

$$\|D^*F\|^2_{L^4} \le C\,\|DD^*F\|^2_{L^2} + CR_0^{-2}\,\|D^*F\|^2_{L^2}\,,$$

for any t.

Upon integrating in time, on any interval $[t_0, t_1] \subset [0, T]$ we obtain

$$\begin{aligned}\left\|\frac{d}{dt}F\right\|^2_{L^{2,2}} + \quad & \|D^*F(t_1)\|^2_{L^2} + \|D^*F\|^2_{L^{2,4}} \\ \leq \quad & C\,\|F\|_{L^{3,3}}\,\|D^*F\|^2_{L^{3,3}} + \|D^*F(t_0)\|^2_{L^2} + CR_0^{-2}\,\|D^*F\|^2_{L^{2,2}} \\ \leq \quad & C_5\,\|F\|_{L^{3,3}}\left(\|D^*F\|^2_{L^{\infty,2}} + \|D^*F\|^2_{L^{2,4}}\right) \\ & + \|D^*F(t_0)\|^2_{L^2} + C_5(t_1 - t_0)R_0^{-2}\,\|D^*F\|^2_{L^{\infty,2}}\ ,\end{aligned}$$

where we also used (14).

By Fubini's theorem, given $\tau > 0$ we can find $t_0 \in [0, \tau]$ such that

$$\|D^*F(t_0)\|^2_{L^2} \leq 2\tau^{-1}\int_0^\tau \|D^*F(t)\|^2_{L^2}\,dt \leq 2\tau^{-1}\mathrm{YM}\,(D_0)\,.$$

Moreover, by absolute continuity of the Lebesgue integral and Lemma 3.4, we can achieve that

$$\|F\|_{L^{3,3}} = \left(\int_{t_0}^{t_1} \|F(t)\|^3_{L^3}\,dt\right) \leq \frac{1}{4C_5}$$

uniformly in t_0 and t_1, if the difference $h = t_1 - t_0$ is sufficiently small. We may assume that $h \leq \frac{R_0^2}{4C_5}$.

Finally, for any such pair $t_0 < t_1 \leq t_0 + h$ we may choose $t_1' \in [t_0, t_1]$ such that

$$\|D^*(t_1')\|^2_{L^2} \geq \frac{2}{3}\,\|D^*F\|^2_{L^{\infty,2}}\ .$$

Hence we obtain the assertion of the lemma on $[t_0, t_1]$. Covering the interval $[\tau, T]$ with finitely many intervals of length h, we conclude. □

Lemma 3.6. *Under the assumptions of Lemma 3.4 there holds $D = D_{ref} + A$, where A extends to $A \in C_{loc}\Big(]0, T]; H^{1,2}\big(\Omega^1(ad\,\eta)\big)\Big)$.*

Proof. By Lemma 3.5 and (3) we have

$$\frac{d}{dt}A \in L^2_{\mathrm{loc}}(]0, T]; L^4)\,,$$

whence $A \in C^0(]0, T]; L^4)$.

Thus by (2) we see that

$$\frac{d}{dt}(D_{\mathrm{ref}}A) = \frac{d}{dt}F(D) + \frac{d}{dt}A\#A \in L^2_{\mathrm{loc}}\big(]0, T]; L^2\big)$$

and

$$D_{\mathrm{ref}}A \in C^0(]0, T]; L^2)\,.$$

Moreover, by the second Bianchi identity and Lemma 3.5

$$\frac{d}{dt}\left(D^*_{\mathrm{ref}}A\right) = D^*_{\mathrm{ref}}\left(\frac{d}{dt}D\right) = A\#D^*F \in L^2_{\mathrm{loc}}(]0, T]; L^2)\,,$$

whence

$$D^*_{\mathrm{ref}}A \in C^0\big(]0, T]; L^2\big)\,.$$

4. Local existence

Before we begin with the proof of Theorem 2.3 it is instructive to review Donaldson's approach to local existence in the smooth case.

First consider the action of gauge transformations on (3). Let $D = D(t)$ be a solution of (3), $S = S(t)$ a family of gauge transformations depending differentiably on t, $\bar{D} = S^*(D)$.

Then if $S(t_0) = id$, by (1) we have

$$\frac{d}{dt}\bar{D} = \frac{d}{dt}D + D\left(\frac{d}{dt}S\right) = -D^*F + Ds$$

at $t = t_0$, where $F = F(D)$, $s = \frac{d}{dt}S \in \Omega^0(\text{ad } \eta)$. Similarly, in general we find

$$\frac{d}{dt}\bar{D} = -\bar{D}^*\bar{F} + \bar{D}s \tag{16}$$

with $\bar{F} = F(\bar{D})$, $s = S^{-1} \circ \frac{d}{dt}S \in \Omega^0(\text{ad } \eta)$, as can easily be seen by applying the time-independent gauge transformation $S(t_0)^{-1}$ to the preceding formulas.

4.1. *Donaldson's Ansatz*

Donaldson uses a version of De Turck's trick [20]. He makes the Ansatz $\bar{D} = D_0 + a$, where he determines a by solving

$$\frac{d}{dt}\bar{D} = \frac{d}{dt}a = -\bar{D}^*\bar{F} + \bar{D}(-\bar{D}^*a), \quad a(0) = 0. \tag{17}$$

Note that by (2) we have

$$\bar{F} = F(D_0) + D_0 a + a\#a = F(D_0) + \bar{D}a + a\#a.$$

Hence, for smooth initial connection D_0, equation (17) is simply a smoothly perturbed version of the heat equation on M which can easily be solved uniquely for small time $0 < t < T_*$. Through the identification

$$s = S^{-1} \circ \frac{d}{dt}S = -\bar{D}^*a,$$

moreover, the solution a generates a family of gauge transformations S that can readily be recovered by solving the initial value problem

$$\frac{d}{dt}S = S \circ s, \qquad S(0) = id. \tag{18}$$

Letting $D = \left(S^{-1}\right)^* \bar{D}$, then we obtain the desired solution D of (3), (4).

However, if we consider initial data $D_0 \in H^{1,2}$ this approach fails, because the term $\bar{D}^*F(D_0)$ in (17) only belongs to $H^{-1}\left(\Omega^1(\text{ad } \eta)\right)$, the dual of $H^{1,2}$.

Thus (17) can only generate a solution a of class $H^{1,2}$, whence $s \in L^2$ and S, defined by (18), will only be measurable (and bounded) in $x \in M$. This, however, is not sufficient to interpret $D = (S^{-1})^*\bar{D}$ even as a weak solution of (3).

The problem can be overcome if we do not attempt to fix the initial connection as background connection for all time but use a variable (and smoothed) background connection $D_{\text{bg}}(t)$ to express the evolving connection $D(t)$ as $\bar{D}(t) = D_{\text{bg}}(t) + a(t)$.

4.2. Choice of background connection

Let D_0 of class $H^{1,2}$ be given. Fix a smooth connection $D_1 \in \mathscr{D}$ and express $D_0 = D_1 + A_0$ in terms of this connection. Solve the initial value problem

$$\frac{d}{dt}A_{\text{bg}} + \Delta_1 A_{\text{bg}} = 0, \quad A_{\text{bg}}(0) = A_0\,, \tag{19}$$

where $\Delta_1 = D_1^* D_1 + D_1 D_1^*$ is the Laplace operator for D_1.

The initial value problem (19) has a unique, global solution A_{bg}. A_{bg} is smooth for $t > 0$ and, in addition, by Lemma 3.2 we have $A_{\text{bg}} \in L^2\left(H^{2,2}\right) \cap C^0\left(H^{1,2}\right) \cap H^{1,2}\left(L^2\right)$ with estimates depending on D_1 and A_0. In particular, there exist constants $C = C(\eta)$ and $T_1 = T_1(\eta, D_1) > 0$ such that on the interval $[0, T_1]$ there holds

$$\|A_{\text{bg}}\|_{L^\infty(H^{1,2})} \leq C\,\|A_0\|_{H^{1,2}}\,. \tag{20}$$

Moreover, $\|A_0\|_{H^{1,2}}$ may be chosen as small as we please. Finally we let $D_{\text{bg}} = D_1 + A_{\text{bg}}$.

4.3. Local existence for the gauge-equivalent flow

We make the Ansatz

$$\bar{D} = D_{\text{bg}} + a\,, \quad a(0) = 0\,,$$

where $\bar{D}$ solves (16) with "drift term" $s = -\bar{D}^* a$; that is

$$\frac{d}{dt}\bar{D} + \bar{D}^*\bar{F} + \bar{D}(\bar{D}^* a) = 0\,,$$

with $\bar{F} = F(\bar{D})$. Expanding

$$\bar{F} = F(\bar{D}) = F(D_{\text{bg}}) + \bar{D}a + a\#a\,,$$

we then obtain the evolution equation

$$\begin{aligned} \frac{d}{dt}a + \bar{\Delta}a &= -\bar{D}^*(F_{\text{bg}} + a\#a) - \frac{d}{dt}D_{\text{bg}} \\ &= -D_{\text{bg}}^* F_{\text{bg}} - \frac{d}{dt}A_{\text{bg}} + a\#F_{\text{bg}} - \bar{D}^*(a\#a) \end{aligned} \tag{21}$$

for a, where $\bar{\Delta}$ is the Laplace operator for $\bar{D}$ and $F_{\text{bg}} = F(D_{\text{bg}})$.

Note that

$$\begin{aligned} F_{\text{bg}} &= F(D_1 + A_{\text{bg}}) \\ &= F(D_1) + D_1 A_{\text{bg}} + A_{\text{bg}}\#A_{\text{bg}} \in L^2(H^{1,2}) \cap C^0(L^2)\,, \\ D_{\text{bg}}^* F_{\text{bg}} &= D_1^* F_{\text{bg}} + A_{\text{bg}}\#F_{A_{\text{bg}}} \in L^{2,2} \end{aligned}$$

and are smooth for $t > 0$. (Here and in the following we use Sobolev's embedding $H^{1,2} \hookrightarrow L^4$ and Hölder's inequality.)

Moreover $\bar{D}^*(a\#a) = \nabla_{\text{bg}} a\#a + a\#a\#a$. Thus we see that a satisfies

$$\frac{d}{dt}a + \bar{\Delta}a = f + a\#F_{\mathrm{bg}} + \nabla_{\mathrm{bg}}a\#a + a\#a\#a\,,$$

where $f \in L^{2,2}$ and is smooth for $t > 0$.

Moreover, since

$$\bar{\Delta} = \Delta_1 + \nabla_1\#\bar{A} + \bar{A}\#\bar{A}\,,$$

where $\bar{A} = A_{\mathrm{bg}} + a$, we obtain

$$\begin{aligned}\frac{d}{dt}a + \Delta_1 a &= f + F_{\mathrm{bg}}\#a + A_{\mathrm{bg}}\#\nabla_1 a + \nabla_1 A_{\mathrm{bg}}\#a \\ &\quad + A_{\mathrm{bg}}\#A_{\mathrm{bg}}\#a + a\#\nabla_1 a + A_{\mathrm{bg}}\#a\#a + a\#a\#a\,.\end{aligned}$$

Local existence of a unique solution $a \in V_T = L^2(H^{2,2}) \cap H^{1,2}(L^2)$ to (21) now follows from Lemma 3.2.

In fact, since $a(0) = 0$, from (9) we have

$$\|a\|_{L^\infty(H^{1,2})} \leq C\,\|a\|_V \leq C\left\|\left(\frac{d}{dt} + \Delta_1\right)a\right\|_{L^{2,2}}\,,$$

on any interval $[0, T]$, where $0 < T \leq T_1$, $C = C(\eta)$. The eight terms I, …, VIII occurring on the right may be estimated as follows:

$$\begin{aligned}\mathrm{II} = \|F_{\mathrm{bg}}\#a\|_{L^{2,2}} &\leq \|F_{\mathrm{bg}}\|_{L^{2,4}}\,\|a\|_{L^{\infty,4}} \leq C\,\|F_{\mathrm{bg}}\|_{L^2(H^{1,2})}\,\|a\|_{L^\infty(H^{1,2})} \\ &\leq C\,\|A_{\mathrm{bg}}\|_{L^2(H^{2,2})}\,\|a\|_{L^\infty(H^{1,2})} \leq \epsilon\,\|a\|_{L^\infty(H^{1,2})} \leq \epsilon\,\|a\|_V\,,\end{aligned}$$

if $0 < T \leq T(\epsilon, D_1, A_0)$;

$$\begin{aligned}\mathrm{III} = \|A_{\mathrm{bg}}\#\nabla_1 a\|_{L^{2,2}} &\leq \|A_{\mathrm{bg}}\|_{L^{\infty,4}}\,\|\nabla_1 a\|_{L^{2,4}} \\ &\leq C\,\|A_0\|_{H^{1,2}}\,\|a\|_{L^2(H^{2,2})} \leq \epsilon\,\|a\|_{L^2(H^{2,2})} \leq \epsilon\,\|a\|_V\,,\end{aligned}$$

if $\|A_0\|_{H^{1,2}} < \frac{\epsilon}{C}$, with $C = C(\eta)$.

The estimate for $I = \|f\|_{L^{2,2}}$ is trivial; IV and V may be estimated like II. Moreover, we have

$$\mathrm{VI} = \|a\#\nabla_1 a\|_{L^{2,2}} \leq \|a\|_{L^{\infty,4}}\,\|\nabla_1 a\|_{L^{2,4}} \leq C\,\|a\|_{L^\infty(H^{1,2})}\,\|a\|_{L^2(H^{2,2})} \leq \epsilon\,\|a\|_V\,,$$

if $\|a\|_{L^\infty(H^{1,2})} \leq \frac{\epsilon}{C}$. Finally,

$$\begin{aligned}\mathrm{VII} &= \|A_{\mathrm{bg}}\#a\#a\|_{L^{2,2}} \leq \|A_{\mathrm{bg}}\|_{L^{\infty,4}}\,\|a\#a\|_{L^{2,4}} \leq C\,\|A_{\mathrm{bg}}\|_{L^\infty(H^{1,2})}\,\|a\#a\|_{L^2(H^{1,2})} \\ &\leq C\,\|A_0\|_{H^{1,2}}\left(\|\nabla_{\mathrm{ref}}a\#a\|_{L^{2,2}} + \|a\|^2_{L^{4,4}}\right) \\ &\leq C\,\|A_0\|_{H^{1,2}}\,\|a\|_{L^{\infty,4}}\left(\|\nabla_{\mathrm{ref}}a\|_{L^{2,4}} + \|a\|_{L^{2,4}}\right) \\ &\leq C\,\|A_0\|_{H^{1,2}}\,\|a\|_{L^\infty(H^{1,2})}\,\|a\|_{L^2(H^{2,2})} \leq \epsilon\,\|a\|_V\end{aligned}$$

and similarly

$$\mathrm{III} = \|a\#a\#a\|_{L^{2,2}} \leq C\,\|a\|^2_{L^\infty(H^{1,2})}\,\|a\|_V \leq \epsilon\,\|a\|_V$$

if $\|a\|_{L^\infty(H^{1,2})} \leq \frac{\epsilon}{C}$.

Hence it suffices to first choose D_1 such that $\|A_0\|_{H^{1,2}} < \epsilon$ for some convenient $\epsilon > 0$ depending only on the bundle η and then choose $T = T(\epsilon, D_1, A_0) > 0$ to obtain a-priori bounds and hence existence of $a \in L^2(H^{2,2}) \cap C^0(H^{1,2}) \cap H^{1,2}(L^2)$ on $[0, T]$ by the contracting map principle. Moreover, a is smooth for $t > 0$ by the theory of linear parabolic equations; see for instance Ladyžhenskaya-Solonnikov-Ural'ceva [11].

4.4. Local existence for the Yang-Mills flow

In order to obtain the desired local solution D of (3), (4) from the solution $\bar{D}$ constructed above, we need to solve the evolution equation

$$\frac{d}{dt}S = S \circ s \tag{22}$$

for the gauge transformaton relating D and $\bar{D}$. Recall that

$$s = -\bar{D}^* a \in L^2(H^{1,2})$$

and is smooth for $t > 0$.

Let (t_k) be a sequence of numbers $0 < t_k \le T$, $t_k \searrow 0$ as $k \to 0$. Solve (22) with initial data $S_k = id$ at $t = t_k$ to obtain a sequence $S_k = S_k(t) \in \mathcal{G}$ of gauge transformations depending smoothly on t for $0 < t \le T$. Clearly, $S_k = S_l^{-1}(t_k) \circ S_l$. Let

$$D_k = (S_k^{-1})^* \bar{D} = S_l(t_k)^* D_l$$

be the corresponding connections. For each k, $D_k = D_k(t)$ is of class C^∞ for $0 < t \le T$ and is a classical solution of (3). Also remark that

$$D_k(t_k) = \bar{D}(t_k) \to D_0 \quad \text{in} \quad H^{1,2} \tag{23}$$

as $k \to \infty$.

Moreover, by the energy inequality (12) and invariance of the energy under gauge transformations we have uniform bounds

$$\begin{aligned} \left\| \frac{d}{dt} D_k \right\|_{L^{2,2}}^2 &= \|D_k^* F(D_k)\|_{L^{2,2}}^2 \le \text{YM}\,(D_0)\,, \\ &\sup_t \text{YM}\,\big(D_k(t)\big) \le \text{YM}\,(D_0) \end{aligned} \tag{24}$$

for any k. Thus, for each k the limit

$$D_k(0) = \lim_{t \searrow 0} D_k(t)$$

exists in L^2 and in view of (23), moreover,

$$D_k(0) \to D_0 \quad \text{in} \quad L^2$$

as $k \to \infty$.

Similarly, by (22) we have

$$\left\| \frac{d}{dt} S_l \right\|_{L^{2,4}}^2 \le \|s\|_{L^{2,4}}^2 \le C\, \|s\|_{L^2(H^{1,2})}^2\,,$$

whence

$$S_l(0) = \lim_{t \searrow 0} S_l(t) \in L^4\,,$$

exists for any l.

Fix some $l = \hat{l}$ and let $\hat{S} = S_{\hat{l}}$, $\hat{D} = D_{\hat{l}}$, $\hat{S}_0 = \hat{S}(0)$, $\hat{D}_0 = \hat{D}(0)$, $\hat{S}_k = \hat{S}(t_k) \in \mathcal{G}$. Then, by the above

$$\hat{S}_k \to \hat{S}_0 \quad \text{in} \quad L^4 ,$$
$$D_k(0) = \hat{S}_k^*(\hat{D}_0) \to D_0 \quad \text{in} \quad L^2 .$$

Moreover, if we let $\hat{D}_0 = D_1 + \hat{A}_0$, $\hat{A}_0 \in L^2$, we find

$$\begin{aligned} \hat{S}_k^*(\hat{D}_0) - D_0 &= \hat{S}_k^*(D_1) - D_1 + \hat{S}_k^{-1}\hat{A}_0\hat{S}_k - A_0 \\ &= \hat{S}_k^{-1} \circ (D_1\hat{S}_k) + \hat{S}_k^{-1}\hat{A}_0\hat{S}_k - A_0 \to 0 \quad \text{in} \quad L^2 . \end{aligned}$$

Thus, also

(25) $$\lim_{k\to\infty} D_1\hat{S}_k = \lim_{k\to\infty} (\hat{A}_0\hat{S}_k - \hat{S}_k A_0) \in L^2$$

exists and necessarily coincides with the distributional limit $D_1\hat{S}_0$; that is,

$$\hat{S}_k \to \hat{S}_0 \quad \text{in} \quad H^{1,2} .$$

But this implies that $D_k = \hat{S}_k^*(\hat{D})$ converges uniformly to some

$$D = \hat{S}_0(\hat{D}) \in C^0(L^2)$$

with $D(0) = D_0$ and $\frac{d}{dt}D \in L^{2,2}$ by (24).

Similarly,

$$F(D_k) = S_k^*\big(F(\hat{D})\big)$$

converges in L^2, locally uniformly for $t > 0$. Since $D_k \to D$ in $C^0(L^2)$, we moreover have convergence

$$F(D_k) \rightharpoonup F(D)$$

in the sense of distributions. Together, these results imply

$$F(D_k) \to F(D) \quad \text{in} \quad C^0(]0,T]; L^2) .$$

In the same way, from (24) and since $D \in C^0([0,T]; L^2)$, we also obtain that

$$F(D(t)) \rightharpoonup F(D_0)$$

weakly in L^2 as $t \to 0$. Finally, since by (24) also

$$\limsup_{t\to 0} \|F(D(t))\|_{L^2}^2 \le \|F(D_0)\|_{L^2}^2 ,$$

we obtain strong convergence $F(D(t)) \to F(D_0)$ in L^2 as $t \to 0$; that is, $F(D) \in C^0([0,T]; L^2)$.

Hence D in fact is a weak solution to the Yang-Mills evolution problem (3), (4). Moreover, D satisfies (12) with equality.

5. Gauge normalization

For the proof of uniqueness, we also consider the gauge-equivalent version (16) of (3). However, we need to specify a gauge condition.

Before we go into the details of the proof, observe that uniqueness of s and hence of the evolving gauge transformations S determined by (18) can only hold if the operator $D\colon \Omega^0(\text{ad}\,\eta) \to \Omega^1(\text{ad}\,\eta)$ is invertible. For smooth connections this condition is equivalent to an algebraic-topological condition on the connection.

5.1. Irreducible connections

Given a connection $D \in \mathscr{D}$ on η, we denote

$$\Gamma = \Gamma(D) = \{S \in \mathscr{G}; S^*(D) = D\}$$

the isotropy subgroup of D. Γ is a Lie group with Lie algebra

$$\gamma = \{s \in \Omega^0(\text{ad }\eta); Ds = 0\} .$$

A connection D is called *irreducible* iff $\Gamma(D)$ consists only of sections with values in the center of G.

In the special case $G = SO(n)$, $n \geq 3$, or $SU(n)$, $n \geq 2$ the center of G is trivial, and hence we obtain:

$$\ker(D) \cap \Omega^0(\text{ad }\eta) = \{0\}$$

iff D is irreducible.

Remark that, if $G = SU(2)$ and D is not irreducible, then either η is trivial and $D = d$ is the trivial connection, or η splits into a sum of line bundles and D restricts to a connection on each factor. (See Donaldson-Kronheimer [6], p. 131 ff., and Freed-Uhlenbeck [7], p. 47.) Thus, at least in many cases of interest, it does not seem to constitute a loss of generality to assume that all connections D we encounter in the evolution (3) are irreducible.

For connections D of class $H^{1,2}$ we require irreducibility in the sense that

$$\|s\|_{H^{1,2}} \leq C \, \|Ds\|_{L^2} \tag{26}$$

for all $s \in H^{1,2}(\Omega^0(\text{ad }\eta))$ with constants $C = C(D)$. This constant can be chosen locally uniformly, as follows.

Lemma 5.1. *Suppose D_0 satisfies (26) with $C_0 = C(D_0)$, $D_0 \in H^{1,2}$. There exists an $H^{1,2}$-neighborhood $\mathscr{V}$ of D_0 and a constant $C > 0$ such that any $D \in \mathscr{V}$ is irreducible and there holds*

$$\|s\|_{H^{1,2}} \leq C \, \|Ds\|_{L^2}$$

uniformly for $s \in H^{1,2}(\Omega^0(ad\ \eta))$.

Proof. Assume by contradiction that for sequences $A_k \in H^{1,2}(\Omega^1(\text{ad }\eta))$, $s_k \in H^{1,2}(\Omega^0(\text{ad }\eta))$ with $A_k \to 0$ in $H^{1,2}$ as $k \to \infty$, $\|s_k\|_{H^{1,2}} = 1$ for all k, we have

$$\|D_k s_k\|_{L^2} = \|(D_0 + A_k)s_k\|_{L^2} \to 0$$

as $k \to \infty$. Then from

$$\begin{aligned} C_0^{-1} = C_0^{-1} \|s_k\|_{H^{1,2}} &\leq \|D_0 s_k\|_{L^2} \leq \|D_k s_k\|_{L^2} + \|[A_k \# s_k]\|_{L^2} \\ &\leq o(1) + \|A_k\|_{L^4} \, \|s_k\|_{L^4} \leq o(1) , \end{aligned}$$

where $o(1) \to 0$ as $k \to \infty$, we obtain a contradiction. □

5.2. *Gauge fixing*

We need a global analogue of Uhlenbeck's [22] theorem on the existence of local Coulomb gauges, depending smoothly on the connection.

Let D_0 be a connection of class $H^{1,2}$ and satisfying (26), and let $D_{\text{bg}}(t) = D_1 + A_{\text{bg}}$, $0 \le t \le T$, be a family of background connections such that $D_{\text{bg}}(0) = D_0$, $A_{\text{bg}} \in C^\infty$ for $t > 0$ and $A_{\text{bg}} \in L^2(H^{2,2}) \cap H^{1,2}(L^2)$, as determined by (19). Let $H^{1,2}$-clos $(\mathcal{G})$ denote the $H^{1,2}$-closure of $\mathcal{G}$.

Proposition 5.2. *Let D be a weak solution of (3), (4) on $[0, T[$ as in Theorem 2.3 (i). There exist $T_0 > 0$ and a family of gauge transformations*

$$S = S(t) \in C^0\big([0, T_0]; H^{1,2}\text{-}clos\ (\mathcal{G})\big)$$

with

$$s = S^{-1} \circ \frac{d}{dt} S \in L^2(H^{1,2})$$

such that

$$\bar{D} = S^*(D) = D_{bg} + \bar{a}$$

satisfies $\bar{a} \in L^\infty(H^{1,2}) \cap H^{1,2}(L^2)$, $\bar{a}(t) \to 0$ in $H^{1,2}$ as $t \to 0$, and $\bar{D}^* \bar{a} = 0$.

Proof. (i) Consider first the case that D_0, D_{bg} and D are smooth. It is convenient to express

$$D = D_{\text{bg}} + a =: D_a\,, \quad \bar{D} = D_{\text{bg}} + \bar{a} = D_{\bar{a}}$$

and to denote the corresponding curvatures by

$$F(D_a) = F_a\,, \quad F(D_{\bar{a}}) = F_{\bar{a}}\,.$$

We apply the implicit function theorem.

For $2 < p < 4$ and fixed $t \ge 0$ introduce the map

$$L: H^{1,p}\big(\Omega^1(\text{ad}\ \eta)\big) \times H^{2,p}(\mathcal{G}) \to L^p\big(\Omega^0(\text{ad}\ \eta)\big)$$

$$L(a, S) = D^*_{\bar{a}} \bar{a}\,,$$

where by abuse of notation we let

$$\bar{a} = \bar{a}(a, S) = S^*(D_a) - D_{\text{bg}}\,.$$

Recall that for $p > 2$ the class $H^{2,p}(\mathcal{G})$ is a manifold; see for instance Uhlenbeck [22]. Moreover, L is of class C^1 with

$$l_{\bar{a}}(s) := \left|\frac{\partial L}{\partial S}\right|_{(a,S)}(s) = D^*_{\bar{a}} D_{\bar{a}} s + D_{\bar{a}} s \# \bar{a}\,.$$

Multiplying by s, we find

$$\big(l_{\bar{a}}(s), s\big) = \|D_{\bar{a}} s\|^2_{L^2} + (D_{\bar{a}} s \# \bar{a}, s) \ge \frac{1}{2} \|D_{\bar{a}} s\|^2_{L^2} - C\, \|\bar{a}\|_{L^4}\, \|s\|_{H^{1,2}}\,.$$

Hence by (26) and Lemma 5.1 the operator $l_{\bar{a}}$ is invertible if $t \ge 0$ and $\|\bar{a}\|_{L^4}$ are sufficiently small.

It follows that there exists $T_0 > 0$ and $S = S(t) \in C^1([0,T_0]; H^{2,p}\mathscr{G})$ such that $S(0) = id$ and
$$L\big(a(t), S(t)\big) = 0;$$
that is, $S^*(D_a) = D_{\bar{a}}$ is of class $C^1(H^{1,p})$ and satisfies $D_{\bar{a}}^*\bar{a} = 0$.

(ii) In order to obtain the analogous result for weak solutions D as in Theorem 2.3, we apply the above reasoning to the approximating sequence of smooth connections
$$D_k = \hat{S}_k^*(\hat{D}) = D_{\text{bg}} + a_k ,$$
on $[t_k, T[$ whose existence is guaranteed by Theorem 2.3 (i), with initial condition
$$D_k(t_k) = D_{k0} \to D_0 \quad \text{in } H^{1,2}$$
for some sequence $t_k \to 0$. Note that
$$\left|\frac{d}{dt}D_k\right| = \left|\frac{d}{dt}D\right| \in L^{2,2}.$$
Moreover, we choose corresponding smooth background connections $D_{\text{bg},k} = D_1 + A_{\text{bg},k}$, where $A_{\text{bg},k}(t)$ solves (19) with initial data
$$A_{\text{bg},k}(t_k) = A_{\text{bg}}(t_k) + a_k(t_k).$$
Observe that

(27) $$D_k(t_k) = D_{\text{bg},k}(t_k).$$

Also note that by Lemma 5.1 the data D_{k0} satisfy condition (26) with a uniform constant C. Moreover, $A_{\text{bg},k} \in V_T$, and, given $\epsilon > 0$, by suitable choice of D_1 and choosing a smaller time $T > 0$, if necessary, we can achieve that
$$\begin{aligned}\epsilon_k(T) \quad := \quad & \left\|\frac{d}{dt}D_k\right\|^2_{L^{2,2}} + \left\|\frac{d}{dt}A_{\text{bg},k}\right\|^2_{L^{2,2}} + \|A_{\text{bg},k}\|^2_{L^\infty(H^{1,2})} + \|A_{\text{bg},k}\|^2_{L^2(H^{2,2})} \\ & + \|F_{\text{bg},k}\|^2_{L^2(H^{1,2})} < \epsilon ,\end{aligned}$$
uniformly in k, $k \geq k_0(\epsilon)$.

By (27), the reasoning in part (i) is applicable for D_k, yielding a C^1-family of smooth gauge transformations $S_k = S_k(t)$ on some interval $[t_k, T_k]$, $t_k < T_k \leq T$, such that
$$\bar{D}_k = S_k^*(D_k) = D_{\text{bg},k} + \bar{a}_k$$
satisfies

(28) $$\bar{D}_k^*\bar{a}_k = 0.$$

The following estimates establish a uniform lower bound on T_k and suitable a-priori bounds on $\bar{a}_k, S_k$ that will allow us to pass to the limit $k \to \infty$ in these relations.

Let $s_k = S_k^{-1}\frac{d}{dt}S_k$ and extend $s_k(t) = 0$, $\bar{a}_k(t) = 0$, that is, $S_k(t) = id$, $D_{\text{bg},k}(t) = D_k(t) = D_{k_0}$ for $0 \leq t \leq t_k$.

Lemma 5.3. *There exist constants $C = C(D_0)$, $T_0 = T(D_0)$ such that for $0 < T_k \leq T_0$ and sufficiently large k there holds*
$$\left\|\frac{d}{dt}\bar{a}_k\right\|^2_{L^{2,2}} + \|\bar{a}_k\|_{L^\infty(H^{1,2})} + \|s_k\|^2_{L^2(H^{1,2})} \leq C\epsilon_k(T_k).$$

Proof. The proof requires several steps. We drop the index k for simplicity.

Claim 1. There exist constants $C = C(D_0) > 0$, $T_0 = T(D_0) > 0$, $\epsilon > 0$ such that for sufficiently large k we have

$$\left\| \frac{d}{dt}\bar{a} \right\|^2_{L^{2,2}} + \|s\|_{L^2(H^{1,2})} \leq C\epsilon(T)$$

for $0 < T \leq T_0$, provided $\|\bar{a}\|_{L^{\infty,4}} \leq \epsilon$.

Proof. Differentiating the equation (28), we obtain

$$\begin{aligned} 0 = \frac{d}{dt}\left(D^*_{\bar{a}}\bar{a}\right) &= D^*_{\bar{a}}D_{\bar{a}}s + D_{\bar{a}}s\#\bar{a} + D^*_{\bar{a}}\left(S^*\left(\frac{d}{dt}D\right) - \frac{d}{dt}D_{\rm bg}\right) \\ &\quad + \left(S^*\left(\frac{d}{dt}D\right) - \frac{d}{dt}D_{\rm bg}\right)\#\bar{a}\,, \end{aligned}$$

where we used (1). Multiplying by s and integrating, we find

$$\begin{aligned} \|D_{\bar{a}}s\|^2_{L^{2,2}} &= (D^*_{\bar{a}}D_{\bar{a}}s, s) \\ &\leq \|D_{\bar{a}}s\|_{L^{2,2}}\left(\|\bar{a}\|_{L^{\infty,4}}\|s\|_{L^{2,4}} + \|f\|_{L^{2,2}}\right) + \|f\|_{L^{2,2}}\|\bar{a}\|_{L^{\infty,4}}\|s\|_{L^{2,4}}\,, \end{aligned}$$

where we denote

$$f = S^*\left(\frac{d}{dt}D\right) - \frac{d}{dt}D_{\rm bg}$$

for brevity.

Since

$$\|f\|_{L^{2,2}} \leq \left\|\frac{d}{dt}D\right\|_{L^{2,2}} + \left\|\frac{d}{dt}D_{\rm bg}\right\|_{L^{2,2}}$$

and since by (26), moreover,

$$\|s\|_{L^{2,4}} \leq C\,\|s\|_{L^2(H^{1,2})} \leq C\,\|D_{\bar{a}}s\|_{L^{2,2}}\,,$$

we obtain the desired estimate for s. Finally, we note that

$$\frac{d}{dt}\bar{a} = \frac{d}{dt}(S^*(D_a) - D_{\rm bg}) = D_{\bar{a}}s + f\,,$$

and the claim follows. □

Let $\delta > 0$ be as in Lemma 3.3 and let $R_0 > 0$ be chosen such that (15) is satisfied, which is possible because $F(D) \in C^0(L^2)$. By (13), it is in fact possible to choose $R_0 = R_0(D_0)$ on a time interval of length CR_0^2.

Claim 2. There exist constants $C = C(D_0)$, $T_0 = T(D_0) > 0$, $\epsilon = \epsilon(D_0) > 0$ such that

$$\|\bar{a}\|^2_{L^{\infty,4}} \leq C\left(\|F_{\bar{a}} - F_{\rm bg}\|^2_{L^{\infty,2}} + \epsilon(T)\right)$$

for $0 < T \leq T_0$, provided $\|\bar{a}\|_{L^{\infty,4}} \leq \epsilon$.

Proof. By Lemma 3.3 for any t we have

$$\|\bar{a}\|_{L^4}^2 \leq C\|D_{\bar{a}}\bar{a}\|_{L^2}^2 + CR_0^{-2}\|\bar{a}\|_{L^2}^2 ,$$

while by (2) there holds

$$D_{\bar{a}}\bar{a} = F_{\bar{a}} - F_{\mathrm{bg}} + \bar{a}\#\bar{a} .$$

Finally, from Claim 1 and since $\bar{a}(0) = 0$ we obtain

$$\|\bar{a}\|_{L^{\infty,2}}^2 \leq CT \left\| \frac{d}{dt}\bar{a} \right\|_{L^{2,2}}^2 \leq CT\epsilon(T) .$$

Together this yields

$$\|\bar{a}\|_{L^{\infty,4}}^2 \leq C\|F_{\bar{a}} - F_{\mathrm{bg}}\|_{L^{\infty,2}}^2 + C\|\bar{a}\|_{L^{\infty,4}}^4 + CTR_0^{-2}\epsilon(T) ,$$

and the claim follows. □

Claim 3. Under the assumptions of Claim 2 there holds

$$\|\bar{a}\|_{L^\infty(H^{1,2})}^2 \leq C\left(\|F_{\bar{a}} - F_{\mathrm{bg}}\|_{L^{\infty,2}}^2 + \epsilon(T)\right)$$

Proof. It suffices to estimate (t fixed)

$$\|\nabla_{\mathrm{ref}}\,\bar{a}\|_{L^2} \leq \left\|\bar{\nabla}\bar{a}\right\|_{L^2} + \|(A_1 + A_{\mathrm{bg}} + \bar{a})\#\bar{a}\|_{L^2}^2 ,$$

where $\bar{\nabla}$ is the covariant derivative corresponding to $\bar{D} = \bar{D}_{\bar{a}}$. Now, by Lemma 3.3, (15), (28) and (2), we have

$$\begin{aligned} \left\|\bar{\nabla}\bar{a}\right\|_{L^2} &\leq C\left\|\bar{D}\bar{a}\right\|_{L^2}^2 + CR_0^{-2}\|\bar{a}\|_{L^2}^2 \\ &\leq C\|F_{\bar{a}} - F_{\mathrm{bg}}\|_{L^2}^2 + C\|\bar{a}\|_{L^4}^4 + CTR_0^{-2}\epsilon(T) , \end{aligned}$$

and in view of Claim 2 the assertion follows. □

Claim 4. Under the assumptions of Claim 2 above there holds

$$\|F_{\bar{a}} - F_{\mathrm{bg}}\|_{L^{\infty,2}}^2 \leq C\epsilon(T) .$$

Proof. By (16) and the second Bianchi identity there holds

$$\frac{d}{dt}F_{\bar{a}} + \Delta_{\bar{a}}F_{\bar{a}} = D_{\bar{a}}D_{\bar{a}}s$$

and hence also

$$\frac{d}{dt}(F_{\bar{a}} - F_{\mathrm{bg}}) + \Delta_{\bar{a}}(F_{\bar{a}} - F_{\mathrm{bg}}) = D_{\bar{a}}D_{\bar{a}}s - \frac{d}{dt}F_{\mathrm{bg}} - \Delta_{\bar{a}}F_{\mathrm{bg}} .$$

Multiplying by $F_{\bar{a}} - F_{\mathrm{bg}}$ and integrating, we obtain

$$\begin{aligned} \mathrm{I} &:= \frac{1}{2}\|F_{\bar{a}} - F_{\mathrm{bg}}\|_{L^{\infty,2}} + \|D_{\bar{a}}(F_{\bar{a}} - F_{\mathrm{bg}})\|_{L^{2,2}}^2 + \|D_{\bar{a}}^*(F_{\bar{a}} - F_{\mathrm{bg}})\|_{L^{2,2}}^2 \\ &\leq \mathrm{II} + \mathrm{III} + \mathrm{IV} . \end{aligned}$$

By the second Bianchi identity, Claim 1 and since by assumption $\|\bar{a}\|_{L^{\infty,4}} \leq \epsilon$, we obtain

$$\begin{aligned} \text{II} &= (D_{\bar{a}}D_{\bar{a}}s, F_{\bar{a}}) - (D_{\bar{a}}D_{\bar{a}}s, F_{\text{bg}}) = -(D_{\bar{a}}s, D_{\bar{a}}^* F_{\text{bg}}) \\ &\leq \|D_{\bar{a}}s\|_{L^{2,2}} \left(\|F_{\text{bg}}\|_{L^2(H^{1,2})} + \|\bar{a}\|_{L^{\infty,4}} \|F_{\text{bg}}\|_{L^{2,4}} \right) \leq C\epsilon(T). \end{aligned}$$

Moreover

$$\begin{aligned} \text{III} &= \left(\frac{d}{dt} F_{\text{bg}}, F_{\bar{a}} - F_{\text{bg}} \right) = \left(D_{\text{bg}} \left(\frac{d}{dt} D_{\text{bg}} \right), F_{\bar{a}} - F_{\text{bg}} \right) \\ &= \left(\frac{d}{dt} D_{\text{bg}}, D_{\text{bg}}^* \left(F_{\bar{a}} - F_{\text{bg}} \right) \right) = \left(\frac{d}{dt} D_{\text{bg}}, D_{\bar{a}}^*(F_{\bar{a}} - F_{\text{bg}}) + \bar{a}\#(F_{\bar{a}} - F_{\text{bg}}) \right) \\ &\leq C\epsilon(T) + \frac{1}{4} \|D_{\bar{a}}^*(F_{\bar{a}} - F_{\text{bg}})\|_{L^{2,2}}^2 + C \|\bar{a}\|_{L^{\infty,4}}^2 \|F_{\bar{a}} - F_{\text{bg}}\|_{L^{2,4}}^2 \end{aligned}$$

Now by Lemma 3.3,

$$\begin{aligned} \|F_{\bar{a}} - F_{\text{bg}}\|_{L^{2,4}}^2 &\leq C \left(\|D_{\bar{a}}^*(F_{\bar{a}} - F_{\text{bg}})\|_{L^{2,2}}^2 + \|D_{\bar{a}}(F_{\bar{a}} - F_{\text{bg}})\|_{L^{2,2}}^2 \right) \\ &\quad + CTR_0^{-2} \|F_{\bar{a}} - F_{\text{bg}}\|_{L^{\infty,2}}, \end{aligned}$$

whence for small $\epsilon > 0$, $T > 0$ we obtain

$$\text{III} \leq C\epsilon(T) + \frac{1}{3} \text{I}.$$

Finally, we may estimate

$$\begin{aligned} \text{IV} &= (-\Delta_{\bar{a}} F_{\text{bg}}, F_{\bar{a}} - F_{\text{bg}}) \\ &= -\big(D_{\bar{a}} F_{\text{bg}}, D_{\bar{a}}(F_{\bar{a}} - F_{\text{bg}})\big) - \big(D_{\bar{a}}^* F_{\text{bg}}, D_{\bar{a}}^*(F_{\bar{a}} - F_{\text{bg}})\big) \\ &\leq C \left(\|D_{\bar{a}} F_{\text{bg}}\|_{L^{2,2}}^2 + \|D_{\bar{a}}^* F_{\text{bg}}\|_{L^{2,2}}^2 \right) + \frac{1}{3} \text{I} \\ &\leq C \left(\|F_{\text{bg}}\|_{L^2(H^{1,2})}^2 + \|\bar{a}\|_{L^{\infty,4}}^2 \|F_{\text{bg}}\|_{L^{2,4}}^2 \right) + \frac{1}{3} \text{I} \\ &\leq C\epsilon(T) + \frac{1}{3} \text{I}. \end{aligned}$$

Thus we obtain $\text{I} \leq C\epsilon(T)$, as desired. □

In view of Claims 1,2, and 4 above there exist $T_0 > 0$, C such that

$$\|\bar{a}_k\|_{L^{\infty,4}}^2 \leq C\epsilon(T) < \epsilon$$

on $[0, T]$ for any $T \leq T_0$ and sufficiently large $k \geq k_0(\epsilon)$, where $\epsilon > 0$ is the constant in Claims 1–4.

The assertion of the lemma follows. □

Proof of Proposition 5.2 (continued)

By Lemma 5.3 we can choose $0 < T_0 \leq T$ such that for large k the linearized operators $l_{\bar{a}_k}$ corresponding to the gauge equation (28) are uniformly invertible on $[0, T_0]$. Hence $T_k \geq T_0$ for large k.

Moreover, Lemma 5.3 implies that the sequence $\bar{a}_k$ is uniformly equi-continouos in L^2, and – by Rellich's theorem – pointwise relatively compact in L^2. Arzéla-Ascoli's theorem thus yields uniform covergence of a sub-sequence

$$\bar{a}_k \to \bar{a} \quad \text{in } C^0(L^2).$$

Hence also

$$\bar{D}_k = D_{\text{bg},k} + \bar{a}_k \to \bar{D} = D_{\text{bg}} + \bar{a}$$

uniformly in L^2 as $k \to \infty$. Passing to the limit in (28), moreover, we obtain

$$D^*_{\bar{a}} \bar{a} = 0 .$$

Finally, $\frac{d}{dt}\bar{a} \in L^{2,2}$ and $\bar{a} \in L^\infty(H^{1,2})$ by lower semi-continuity, and

$$\|\bar{a}(t)\|_{H^{1,2}} \leq C\epsilon(t) \to 0$$

as $t \to 0$.

Similarly, $S_k \to S$ uniformly in L^2 with $s = S^{-1}\frac{d}{dt}S \in L^2(H^{1,2})$. In fact, by an argument similar to (25) we find $S_k \to S$ uniformly in $H^{1,2}$, and

$$D_{\bar{a}} = S^*(D_a) .$$

This ends the proof. □

6. Uniqueness

Given $D_0 \in H^{1,2}$, a family of background connections D_{bg} as in Section 4.2, let $D_a = D_{\text{bg}} + a$ be a local weak solution to (3) and $D_{\bar{a}} = D_{\text{bg}} + \bar{a}$ the corresponding family of normalized connections according to Proposition 5.2.

By (16), $D_{\bar{a}}$ weakly solves the initial value problem

$$\frac{d}{dt} D_{\bar{a}} = -D^*_{\bar{a}} F_{\bar{a}} + D_{\bar{a}} s , \tag{29}$$

$$D^*_{\bar{a}}(\bar{a}) = 0 , \tag{30}$$

$$\bar{a}(0) = 0 , \tag{31}$$

where $F_{\bar{a}} = F(D_{\bar{a}})$, and

$$\begin{gathered} \bar{a} \in L^\infty\Big([0,T]; H^{1,2}\big(\Omega^1(\text{ad}\ \eta)\big)\Big) \cap H^{1,2}\Big([0,T]; L^2\big(\Omega^1(\text{ad}\ \eta)\big)\Big) , \\ F_{\bar{a}} \in C^0\Big([0,T]; L^2\big(\Omega^2(\text{ad}\ \eta)\big)\Big) , \\ s \in L^2\Big([0,T]; H^{1,2}\big(\Omega^0(\text{ad}\ \eta)\big)\Big) \end{gathered} \tag{32}$$

on some interval $[0, T]$. $\bar{a}$ attains its initial data in the $H^{1,2}$-sense. The following result shows that — provided D_0 is irreducible — the solution $D_{\bar{a}}$ above is unique.

Proposition 6.1. *For any $D_0 \subset H^{1,2}$ satisfying* (26) *there exists $T > 0$ and a unique solution $(\bar{a}, s)$ of* (29)–(31) *on* $[0, T]$ *satisfying* (32).

In addition, $\bar{a} \in L^2(H^{2,2})$, and $\bar{a}$ and s are smooth for $t > 0$. Finally, if D_0 is smooth, $\bar{a}$ and s are smooth up to $t = 0$.

Proof. (i) The existence of $(\bar{a}, s)$ was shown above.

(ii) To see higher regularity and uniqueness, in a first step we establish suitable a-priori estimates for solutions in the above class.

Estimate for s: From the second Bianchi identity and (30) we obtain

$$D_{\bar{a}}^* D_{\bar{a}} s = D_{\bar{a}}^* \left(\frac{d}{dt} D_{\bar{a}} \right) = D_{\bar{a}}^* \left(\frac{d}{dt} A_{\text{bg}} \right) + \bar{a} \# \frac{d}{dt} \bar{a} .$$

Multiplying by s, we find

$$\begin{aligned} \| D_{\bar{a}} s \|_{L^{2,2}}^2 &= (D_{\bar{a}}^* D_{\bar{a}} s, s) = \left(\frac{d}{dt} A_{\text{bg}}, D_{\bar{a}} s \right) + \left(\bar{a} \# \frac{d}{dt} \bar{a}, s \right) \\ &\leq C \left\| \frac{d}{dt} A_{\text{bg}} \right\|_{L^{2,2}} \| D_{\bar{a}} s \|_{L^{2,2}} + C \, \| \bar{a} \|_{L^\infty(H^{1,2})} \left\| \frac{d}{dt} \bar{a} \right\|_{L^{2,2}} \| s \|_{L^2(H^{1,2})} , \end{aligned}$$

whence

$$\| s \|_{L^2(H^{1,2})} \leq C \, \| D_{\bar{a}} s \|_{L^{2,2}}^2 \leq C \left\| \frac{d}{dt} A_{\text{bg}} \right\|_{L^{2,2}}^2 + C \| \bar{a} \|_{L^\infty(H^{1,2})}^2 \left\| \frac{d}{dt} \bar{a} \right\|_{L^{2,2}}^2 .$$

Estimate for $\bar{a}$: Using (30) we rewrite (29) as

$$\frac{d}{dt} \bar{a} + D_{\bar{a}}^* F_{\bar{a}} + D_{\bar{a}} D_{\bar{a}}^* \bar{a} = D_{\bar{a}} s - \frac{d}{dt} A_{\text{bg}} .$$

By (2) the left hand side equals

$$\left(\frac{d}{dt} + \Delta_{\bar{a}} \right) \bar{a} + D_{\bar{a}}^* F_{\text{bg}} + D_{\bar{a}}^* (\bar{a} \# \bar{a})$$

which is $\left(\frac{d}{dt} + \Delta_1 \right) \bar{a}$, the heat operator with respect to D_1 applied to $\bar{a}$, up to error terms

$$\nabla_1 \bar{a} \# \bar{a} + \bar{a} \# \bar{a} \# \bar{a} + A_{\text{bg}} \# \nabla_1 \bar{a} + \nabla_1 A_{\text{bg}} \# \bar{a} + A_{\text{bg}} \# \bar{a} \# \bar{a} + A_{\text{bg}} \# A_{\text{bg}} \# \bar{a} + D_{\bar{a}}^* F_{\text{bg}} .$$

Hence if $\| \bar{a} \|_{L^\infty(H^{1,2})}$ is sufficiently small, which we can achieve by choosing $T > 0$ sufficiently small, from Lemma 3.2 and routine estimates we obtain

$$\| \bar{a} \|_V^2 = \left\| \frac{d}{dt} \bar{a} \right\|_{L^{2,2}}^2 + \| \bar{a} \|_{L^2(H^{2,2})}^2 \leq C \left(\| D_{\bar{a}} s \|_{L^{2,2}}^2 + \| A_{\text{bg}} \|_V^2 \right) \leq C \, \| A_{\text{bg}} \|_V^2 .$$

Smoothness of $\bar{a}$ and s for $t > 0$ follows by standard methods; see for instance Ladyženskaya-Solonnikov-Ural'ceva [11].

(iii) Next we derive similar estimates for the difference (α, σ) of two solutions $(\bar{a}_1, s_1)$, $(\bar{a}_2, s_2)$ of (29), (30) with $\bar{a}_1(0) = \bar{a}_0 = \bar{a}_2(0)$. Let $D_1 = D_{\bar{a}_1}$, etc.

Note that (α, σ) satisfies

$$\frac{d}{dt}\alpha = -\left(D_1^* F_1 - D_2^* F_2\right) + D_1 s_1 - D_2 s_2 \,,$$

where

$$D_1 s_1 - D_2 s_2 = D_{\bar{a}}\sigma + \bar{a}\#\sigma + \alpha\# s \,,$$

denoting by $\bar{a}$ any convex linear combinations of $\bar{a}_1$ and $\bar{a}_2$, and similarly for s.

Moreover, for $i = 1, 2$, we have

$$D_i^* F_i = D_{\bar{a}}^* F_i + \alpha\# F_i \,,$$

and

$$F_1 - F_2 = D_{\bar{a}}\alpha + \alpha\#\bar{a} \,.$$

Also we note that

$$D_{\bar{a}}^*\alpha = D_{\bar{a}}^*\bar{a}_1 - D_{\bar{a}}^*\bar{a}_2 = \alpha\#\bar{a} \,.$$

Estimate for σ: By the second Bianchi identity we have

$$\begin{aligned} D_{\bar{a}}^* D_{\bar{a}}\sigma &= D_{\bar{a}}^*\left(\frac{d}{dt}\alpha\right) + D_1^* D_1^* F_1 - D_2^* D_2^* F_2 + \alpha\# D_{\bar{a}}^* F_{\bar{a}} + D_{\bar{a}}^*\left(\bar{a}\#\sigma + \alpha\# s\right) \\ &= \alpha\#\frac{d}{dt}\bar{a} + \frac{d}{dt}\alpha\#\bar{a} + \alpha\# D_{\bar{a}}^* F_{\bar{a}} + D_{\bar{a}}^*(\bar{a}\#\sigma + \alpha\# s) \,. \end{aligned}$$

Multiplying by σ and integrating by parts in the last term, we obtain

$$\begin{aligned} &\|\sigma\|_{L^2(H^{1,2})}^2 \le C\,\|D_{\bar{a}}^*\sigma\|_{L^2}^2 = C\left(D_{\bar{a}}^* D_{\bar{a}}\sigma, \sigma\right) \\ &\le C\left(\left\|\frac{d}{dt}\bar{a}\right\|_{L^{2,2}}\|\alpha\|_{L^{\infty,4}} + \|\bar{a}\|_{L^{\infty,4}}\left\|\frac{d}{dt}\alpha\right\|_{L^{2,2}} + \|D_{\bar{a}}^* F_{\bar{a}}\|_{L^{2,2}}\|\alpha\|_{L^{\infty,4}}\right)\|\sigma\|_{L^{2,4}} \\ &\quad + C(\|\bar{a}\|_{L^{\infty,4}}\|\sigma\|_{L^{2,4}} + \|\alpha\|_{L^{\infty,4}}\|s\|_{L^{2,4}})\,\|D_{\bar{a}}\sigma\|_{L^{2,2}} \\ &\le \left(\frac{1}{2} + C\,\|\bar{a}\|_{L^\infty(H^{1,2})}\right)\|\sigma\|_{L^2(H^{1,2})}^2 + C\,\|\bar{a}\|_{L^\infty(H^{1,2})}^2\left\|\frac{d}{dt}\alpha\right\|_{L^{2,2}}^2 \\ &\quad + C\left(\left\|\frac{d}{dt}\bar{a}\right\|_{L^{2,2}}^2 + \|D_{\bar{a}}^* F_{\bar{a}}\|_{L^{2,2}}^2 + \|s\|_{L^2(H^{1,2})}^2\right)\|\alpha\|_{L^\infty(H^{1,2})}^2 \,. \end{aligned}$$

That is, in view of our previous estimates for $(\bar{a}_i, s)$,

$$\|\sigma\|_{L^2(H^{1,2})} \le C\epsilon\left(\left\|\frac{d}{dt}\alpha\right\|_{L^{2,2}}^2 + \|\alpha\|_{L^\infty(H^{1,2})}^2\right) ,$$

where $\epsilon \to 0$ as $T \to 0$.

Estimate for α: There holds

$$\frac{d}{dt}\alpha + \Delta_{\bar{a}}\alpha = \alpha\# F_{\bar{a}} + D_{\bar{a}}^*(\alpha\#\bar{a}) + D_{\bar{a}}(\alpha\#\bar{a}) + D_{\bar{a}}\sigma + \bar{a}\#\sigma + \alpha\# s \,,$$

whence for $0 < T \le T(\epsilon)$ and $\|A_0\|_{H^{1,2}} + \|\bar{a}\|_{L^\infty(H^{1,2})} < \epsilon$ similar to our estimates for (21) and in view of our estimate for σ we obtain

$$\left\| \frac{d}{dt}\alpha \right\|_{L^2} + \|\alpha\|_{L^\infty(H^{1,2})} + \|\alpha\|_{L^2(H^{2,2})} \leq C\epsilon \left(\left\| \frac{d}{dt}\alpha \right\|_{L^2} + \|\alpha\|_{L^\infty(H^{1,2})} + \|\alpha\|_{L^2(H^{2,2})} \right) .$$

Thus, for $\epsilon > 0$ sufficiently small, we have $\alpha = 0$ and $\sigma = 0$ as claimed. □

From Propositions 5.2 and 6.1 the uniqueness of the local solution $D = D_a$ to (3), (4) constructed in Sect. 3 follows.

Remark that by solving the initial value problem (29)–(31) we could give an alternative existence proof for (3).

7. Long-time existence and asymptotic behaviour

Proof of Theorem 2.3 (iii). Suppose $T < \infty$ is maximal such that (3), (4) possesses a weak solution D which is gauge-equivalent to a smooth solution $\hat{D} = (\hat{S}_0^{-1})^*(D)$ on $]0, T[$ and assume (by contradiction) that there exists $R > 0$ such that (15) holds. Then by Lemma 3.6

$$\lim_{t \to T} \hat{D}(t) = \hat{D}(T)$$

exists in $H^{1,2}$.

Thus for $t_0 < T$ sufficiently large the local solution $\hat{D}^{t_0}$ to the initial value problem (3) with initial data $\hat{D}(t_0)$ at time $t = t_0$ constructed in Section 4 extends to an interval $[t_0, t_1[$, $t_1 > T$. By uniqueness of weak solutions to (3) and equivariance of (3) under time-independent gauge transformations, necessarily $D(t) = (\hat{S}_0)^*(\hat{D}^{t_0}(t))$ on $[t_0, T[$. Hence $\hat{S}_0^*(\hat{D}^{t_0})$ extends the solution $D(t)$ to the interval $[t_0, t_1[$, contradicting the maximality of T. The finiteness assertion is proved as in the case of the evolution of harmonic maps, using the local form (13) of the energy inequality; see [18], p. 577. □

Blow-up analysis and asymptotics

The proof of Theorem 2.4 relies on Uhlenbeck's [21] and Sedlacek's [17] results for the blow-up analysis and on the local energy inequality (13); however, it promises to be rather technical, if carried out rigourously, and will be presented elsewhere. An indication of the proof of the asserted asymptotic behaviour has been given in some detail by Chen-Shen [5].

References

1. Atiyah, M., Bott, R.: The Yang Mills equations over Riemannian surfaces. Philos. Trans. R. Soc. London A **308**, 524–615 (1982)
2. Bethuel, F.: The approximation problem for Sobolev maps between manifolds. Acta Math. **167**, 153–206 (1991)
3. Bourguignon, J.P., Lawson, B.: Stability and isolation phenomena for Yang-Mills fields. Commun. Math. Phys. **79**, 189–230 (1981)
4. Chang, K.-C., Ding W.-Y., Ye, R.: Finite-time blow-up of the heat flow of harmonic maps from surfaces. (Preprint)

5. Chen, Y., Shen, C.-L.: Monotonicity formula and small action regularity for Yang-Mills flows in higher dimensions. Calc. Var. (to appear)
6. Donaldson, S.K., Kronheimer, P.: The topology of four-manifolds. Oxford, New York: Clarendon Press 1990
7. Freed, D., Uhlenbeck, K.: Instantons and four-manifolds. 2nd. edn. Berlin, Heidelberg, New York: Springer 1991
8. Hamilton, R.: Monotonicity formulas for parabolic flows on manifolds. Preprint
9. Helgason, S.: Differential geometry, Lie groups, and symmetric spaces. New York: Academic Press 1978
10. Jost, J.: Nonlinear methods in Riemannian and Kählerian geometry. (DMV Seminar, vol. 10) Basel: Birkhäuser 1991
11. Ladyžhenskaya, O.A., Solonnikov, V.A., Ural'ceva, N.N.: Linear and quasilinear equations of parabolic type. Providence, RI: AMS 1968
12. Lawson, B.: The theory of gauge filds in four dimensions. (CBMS Regional Conf. Series, vol. 58), Providence, RI: AMS 1987
13. Lions, J.L., Magenes, E.: Non-homogenous boundary value problems and applications, vol. I. (Grundlehren Math. Wiss., vol. 199) Berlin, Heidelberg, New York: Springer 1972
14. Naito, H.: Finite time blowing up for the Yang-Mills gradient flow in higher dimensions. Preprint, 1992
15. Nakajima, H.: Compactness of the moduli space of Yang-Mills connections in higher dimensions. J. Math. Soc. Jpn. **40**, 383–392 (1988)
16. Råde, J.: On the Yang-Mills heat equation in two and three dimensions. Preprint, 1991
17. Sedlacek, S.: The Yang-Mills functional over 4-manifolds. Commun. Math. Phys. **86**, 515–527 (1982)
18. Struwe, M.: The evolution of harmonic maps of Riemannian surfaces. Comment. Math. Helv. **60**, 558–581 (1985)
19. Struwe, M.: On the evolution of harmonic maps in higher dimensions. J. Diff. Geom. **28**, 485–502 (1988)
20. De Turck, D.M.: Deforming metrics in direction of their Ricci tensor. J. Diff. Geom. **18**, 157–162 (1983)
21. Uhlenbeck, K.: Removable singularities in Yang-Mills fields. Commun. Math. Phys. **83**, 11–30 (1982)
22. Uhlenbeck, K.: Connections with L^p-bounds on curvature. Commun. Math. Phys. **83**, 31–42 (1982)

This article was processed by the author using the LaTeX style file *pljour1* from Springer-Verlag.