J. DIFFERENTIAL GEOMETRY
86 (2010) 163-188

GEOMETRY OF MINIMAL ENERGY YANG–MILLS CONNECTIONS

Mark Stern

Abstract

We prove that energy minimizing Yang–Mills connections on compact homogeneous 4-manifolds are either instantons or split into a sum of instantons on passage to the adjoint bundle. We prove related results for Calabi–Yau 3-folds and for $3-$dimensional manifolds of nonnegative Ricci curvature.

1. Introduction

Let G be a compact Lie group and E a principal $G-$bundle on a complete oriented Riemannian manifold, M. Let A denote a connection on E and ∇_A the associated covariant derivative on the adjoint bundle, $ad(E)$. The Yang–Mills energy of A is

$$YM(A) := \|F_A\|^2,$$

where F_A denotes the curvature of A, and $\|\cdot\|$ denotes the L_2 norm. A connection is called a *Yang–Mills connection* if it is a critical point of the Yang–Mills functional. A Yang–Mills connection, A, is called *stable* if the second variation of the Yang–Mills functional is nonnegative at A. It is called a local minimum of the Yang–Mills functional if it minimizes $YM(A)$ among all nearby connections. We call a connection *abelian* if its curvature takes values in an abelian subalgebra of the adjoint bundle.

In four dimensions, F_A decomposes into its self-dual and anti-self-dual components,

$$F_A = F_A^+ + F_A^-,$$

where $F_A^{\pm}$ denotes the projection onto the ± 1 eigenspace of the Hodge star operator. A connection is called self-dual (respectively anti-self-dual) if $F_A = F_A^+$ (respectively $F_A = F_A^-$). A connection is called an *instanton* if it is either self-dual or anti-self-dual. On compact oriented 4-manifolds, an instanton is always an absolute minimizer of the Yang–Mills energy. Not all Yang–Mills connections are instantons. See [**SSU**] and [**SS**] for (unstable) examples of $SU(2)$ Yang–Mills connections on S^4 which are neither self-dual nor anti-self-dual. Moreover,

The author was partially supported by NSF grant DMS-0504890.
Received 03/27/2009.

not all stable abelian Yang–Mills connections are instantons; otherwise every harmonic 2-form representing an integral cohomology class would be self-dual or anti-self-dual, which is clearly false.

This leads to the

Converse question: In four dimensions are local minima for the Yang–Mills energy necessarily direct sums of instantons and abelian connections?

Partial positive results for low rank G were obtained by Bourguignon, Lawson, and Simons in [**BLS**] and [**BL**], where they use a variational argument to show that if $G = SU(2)$ or $SU(3)$, and M is a compact oriented 4-dimensional homogeneous space, then the connection is either an instanton or abelian. In this note, we settle this question on the (possibly reduced) adjoint bundle of E, for any compact group, G, when M is a complete nonnegatively curved homogeneous manifold. We also give related results for oriented $3-$manifolds and Calabi–Yau geometries in three complex dimensions.

Our main result is the following theorem.

Theorem 1.1. *Let E be a principal $G-$bundle on a complete oriented nonnegatively curved homogeneous Riemannian 4-manifold, M. Let A be a finite energy, locally minimizing Yang–Mills connection on E. The adjoint bundle, $ad(E)$, contains two ∇_A-stable subbundles, k^+ and k^-, satisfying $F_A^\pm$ is a section of $\Lambda^2T^*M \otimes k^\pm$,*

$$[k^+, k^-] = 0,$$

and the curvature of k^+ is self-dual and that of k^- is anti-self-dual.

Our proof of Theorem 1.1 extends the variational argument of Bourguignon, Lawson, and Simons. Let A_t be a smooth family of connections on E with $A_0 = A$. The assumption that A is a local minimum of the Yang–Mills energy implies the variational inequality

$$\frac{d^2}{dt^2} YM(A_t)_{|t=0} \geq 0. \tag{1.2}$$

The proof of the theorem relies on choosing useful families of test connections with the difference, $A_t - A$, constructed from F_A. In [**BL**], the test connection $A_t = A + ti_XF_A^+$ was used, where i_X denotes interior multiplication by the vector field X, and X runs over a basis of Killing vector fields. Our results rely on recognizing this variation as only the first term in an infinite family of related variations.

The curvature is the only natural object from which to construct test variations, but we need a map from $2-$forms to $1-$forms in order to create test variations from the curvature $2-$form. In homogeneous geometries, interior product with Killing vector fields provides such a map. More generally, this suggests we seek new information on the geometry of locally minimizing Yang–Mills connections in special geometries where there exist natural maps

$$\Phi : \Lambda^2 T^* M \otimes ad(E) \to \Lambda^1 T^* M \otimes ad(E).$$

When such Φ exist, we can consider variations with $\frac{dA}{dt}(0) = \Phi(F_A)$ and seek additional results. Covariant constant 3-forms induce natural maps from 2-forms to 1-forms. Hence, one expects new results for G_2 manifolds, Calabi-Yau 3 folds, and oriented 3-dimensional manifolds. We note that for manifolds of dimension greater than 4, interesting stable Yang–Mills connections need not always exist. For example, Simons (see [**BL**]) proved the nonexistence of nonflat stable Yang–Mills connections on S^n, $n > 4$. This nonexistence result has subsequently been generalized in many directions, for example [**KOT**], [**OP**], [**P**],[**Sh**], and [**X**]. Our results below include no new *existence* results for minimizing connections in higher dimensions.

On a Kahler $m-$fold with Kahler form ω, the curvature decomposes as

$$F_A = F_A^{2,0} + F_{A0}^{1,1} + \frac{1}{m}(\Lambda F_A)\omega + F_A^{0,2},$$

where Λ denotes the adjoint of exterior multiplication by ω, and $F_{A0}^{1,1} = F_A^{1,1} - \frac{1}{m}(\Lambda F_A)\omega$. Define the energy $E'(A) = \|F_A^{0,2}\|^2$.

Theorem 1.3. *Let A be a $G-$connection on a bundle, E, on a complete Calabi Yau threefold, M. Assume that A is a stable critical point of E'. If M is noncompact, assume also that $F_A^{0,2} \in L^4$. Then $F_A^{0,2}$ is covariant constant and takes values in an abelian subbundle of $ad(E) \otimes \mathbb{C}$. If $Hol(M) = SU(3)$, then E is holomorphic.*

Some of our results for 3-manifolds are presumably already known (see for example [**JT**, Chapter II, Corollary 2.3] for the case of $\mathbb{R}^3$), but we include them here as they fall in the same family of techniques as the preceding results. We say an $n-$manifold has *local flat factors* if it is locally isometric to a product of an open subset in $\mathbb{R}^k$, $k \geq 1$, and an $(n-k)-$manifold.

Theorem 1.4. *Let E be a bundle on a complete 3-dimensional manifold, M, with nonnegative Ricci curvature. Let A be a stable finite energy Yang–Mills connection on E. If M is noncompact, assume also that $F_A \in L^4$. Then*

$$\nabla_A F_A = 0.$$

Moreover, F_A takes values in a flat abelian subbundle of $ad(E)$, and $F_A = 0$ unless M has local flat factors.

The additional assumption here that $F_A \in L_4$ if M is noncompact is easily established under additional geometric hypotheses of bounded geometry. (See Remark 2.18.) We remark that applying the preceding theorem to S^3 gives an analytic proof (see Section 4) of the triviality of $\pi_2(G)$ for all compact Lie groups G.

Acknowledgements. We wish to thank Savdeep Sethi for stimulating conversations and for his interest in this work. We thank Benoit Charbonneau for helpful comments. We also thank the referee for his extensive comments, especially his suggested extension (Proposition 2.17) of our theorems to complete noncompact manifolds.

2. Preliminaries

Let M be a complete Riemannian manifold and E a principal G bundle over M, with G a compact Lie group. Let $ad(E)$ denote the adjoint bundle of E, endowed with a $G-$invariant inner product. Let $A^p(M, ad(E))$ denote the smooth $p-$forms with values in $ad(E)$. Given a connection A on E, we denote by ∇_A the corresponding covariant derivative on $A^*(M, ad(E))$ induced by A and the Levi-Civita connection of M. Let d_A denote the exterior derivative associated to ∇_A.

We are interested in stable minima of the Yang–Mills energy

$$YM(A) = \|F_A\|^2,$$

where F_A denotes the curvature of A. Critical points of this energy satisfy the Yang–Mills equation

$$d_A^* F_A = 0, \tag{2.1}$$

where d_A^* denotes the (L_2-)adjoint of d_A. In addition, all connections satisfy the Bianchi identity

$$d_A F_A = 0. \tag{2.2}$$

If A_t is a smooth one parameter family of connections, then

$$\frac{d}{dt} F_{A_t} = d_{A_t}(\frac{dA}{dt}). \tag{2.3}$$

More generally, if $\psi \in A^1(M, ad(E))$ then

$$F_{A+\psi} = F_A + d_A\psi + \psi \wedge \psi. \tag{2.4}$$

Here we note that our convention on exterior products of $ad(E)$ valued forms is normalized by

$$(dx^I \otimes v_I) \wedge (dx^J \otimes v_J) = \frac{1}{2}(dx^I \wedge dx^J) \otimes [v_I, v_J].$$

As a notational convenience, we will often denote by $e(w)$ exterior multiplication on the left by a form w (possibly with $ad(E)$ coefficients). Its adjoint is denoted $e^*(w)$. Thus

$$e(w)h := w \wedge h, \text{ and } \langle f, e(w)h\rangle = \langle e^*(w)f, h\rangle.$$

If A minimizes the Yang–Mills energy, then of course it satisfies the inequality

$$\|F_A\|^2 \leq \|F_{A+\psi}\|^2, \tag{2.5}$$

for all smooth compactly supported ψ. Replacing ψ by $t\psi$ in (2.5), using (2.4), and taking the limit as $t \to 0$ leads to the second variation inequality

$$0 \leq \|d_A\psi\|^2 + 2\langle F_A, \psi \wedge \psi\rangle. \tag{2.6}$$

When considering noncompact manifolds, we will consider variations with ψ *not* compactly supported. Let $\{y_j\}_{j=1}^{\infty}$ be a sequence of functions, $0 \leq y_j \leq 1$, with $lim_{j\to\infty} y_j = 1$ pointwise and $|dy_j|$ uniformly bounded. If we assume merely that $\psi \in C^1 \cap L^2 \cap L^4$, then replacing ψ by $y_j\psi$ in (2.6) yields $0 \leq \|d_A\psi\|^2 + 2\langle F_A, \psi \wedge \psi\rangle$ upon passing to the limit. Hence we may apply this variational inequality to $\psi \in C^1 \cap L^2 \cap L^4$.

The ψ we will most often consider in (2.6) will be constructed from F_A and its derivatives. Hence we need to establish some standard estimates for $\nabla_A^k F_A$ to ensure these test variations are in $L^2 \cap L^4$. Let M be a complete noncompact homogeneous 4−manifold. Let A be a finite energy Yang–Mills connection on a G−bundle E over M. Let $B_r(p)$ denote the ball of radius r about p. The following lemma is a specialization of [**N**, Lemma 3.1] to our context.

Lemma 2.7. *There exist constants* $\epsilon = \epsilon(M, G)$ *and* $C = C(M, G)$ *such that if*

$$r^{4-n} \int_{B_r(p)} |F_A|^2 dv < \epsilon, \text{ then } \sup_{B_{r/4}(p)} |F_A|^2 \leq C r^{-n} \int_{B_r(p)} |F_A|^2 dv,$$

for $r < \frac{1}{2}\min\{1, \textit{injectivity radius of } M\}$, $p \in M$.

Cover M with a collection of balls, $\{B_s(p_i)\}_{i=1}^{\infty}$, of radius $s = \frac{1}{8}\min\{1,$ injectivity radius of $M\}$, and with a uniform bound on the number of balls containing any point of M.

Corollary 2.8. *For every* $\delta > 0$, $\sup_{x \in B_{2s}(p_i)} |F_A|(x)^2 < \delta$, *for* i *sufficiently large.*

Proof. As A is finite energy, for each $\delta > 0$, there exists N_δ so that we may remove a finite number of balls so that $Cs^{-n} \int_{B_{8s}(p_i)} |F_A|^2 dv < \delta$, for $i > N_\delta$. Now Lemma 2.7 yields the desired bound. q.e.d.

We specialize the existence and properties of the Uhlenbeck gauge to our situation.

Lemma 2.9. *(Uhlenbeck Gauge)*[**U1**]. *(See also* [**W**, p.91 and p.145]*).* *Let* A *be a connection on a* G−*bundle. Fix* $q > \frac{n}{2}$. *There exist constants*

$\beta(q)=\beta(M,G,q)$, $b(q)=b(M,G,q)$, and $C_{k,q}$ such that if

$$\int_{B_{2s}(p_i)} |F_A|^q dv \leq \beta(q), \tag{2.10}$$

then there exists a trivialization of E over $B_{2s}(p_i)$ in which the connection form $\hat{A}$ satisfies

$$d^*\hat{A}=0, \quad \text{and}$$

$$\|\hat{A}\|_{W^{1,q}(B_s(p_i))} \leq b(q)\|F_A\|_{L^q}.$$

If we assume also that A is a critical point of the Yang–Mills functional, then $\forall k$

$$\|\hat{A}\|_{W^{k+1,q}(B_s(p_i))} \leq C_{k,q}(1+\|\hat{A}\|_{W^{k,q}(B_s(p_i))}+\|\hat{A}\|^3_{W^{k,q}(B_s(p_i))}). \tag{2.11}$$

Corollary 2.12. *$|\nabla^k_A F_A|$ is pointwise bounded on M for all k.*

Proof. By Corollary 2.8, $|F_A|$ is pointwise bounded, and the hypothesis (2.10) holds for i sufficiently large. The result now follows from an induction argument using (2.11) and the Sobolev embedding theorem. q.e.d.

Lemma 2.13. *For A a critical point of the Yang–Mills functional with finite energy on a homogeneous manifold, $\nabla^k_A F_A \in L_2$ for all k.*

Proof. The Yang–Mills equation and Bianchi identity imply that F_A satisfies the equation

$$0=\nabla^*_A\nabla_A F_A+(\hat{F}_A+\hat{R})(F_A), \tag{2.14}$$

where, in a local orthonormal frame,

$$\hat{F}_A(F_A)=-[F_{ij},F_{jk}]dx^i\wedge dx^k,$$

and

$$\hat{R}(F_A)=-R_{ijmk}F_{jm}dx^i\wedge dx^k.$$

Here R denotes the Riemann curvature tensor. Assume now that the sequence of cutoff functions, $\{y_m\}_{m=1}^{\infty}$, is C^k bounded for all k. Then

$$0=\langle\nabla^*_A\nabla_A F_A+(\hat{F}_A+\hat{R})(F_A),y_m^2F_A\rangle_{L_2}$$

(2.15)

$$\Rightarrow 0=\|\nabla_A y_m F_A\|^2_{L_2}-\|[\nabla_A,y_m]F_A\|^2_{L_2}+\langle(\hat{F}_A+\hat{R})(F_A),y_m^2F_A\rangle_{L_2}.$$

From Corollary 2.12, we see that $|F_A|$ is bounded in *sup* norm. Taking $m\to\infty$, equation (2.15) now implies that

$$\nabla_A F_A \in L_2.$$

Differentiating equation (2.14), we obtain

$$0=\nabla^*_A\nabla_A\nabla_A F_A+(\hat{F}_A+\hat{R})(\nabla_A F_A)+[\nabla_A,\nabla^*_A\nabla_A+(\hat{F}_A+\hat{R})]F_A. \tag{2.16}$$

The quantity $[\nabla_A,\nabla^*_A\nabla_A+(\hat{F}_A+\hat{R})]F_A$ is a sum of terms involving at most one derivative of F and bounded coefficients. Hence taking the L_2-inner product of (2.16) with $\nabla_A F_A$ and integrating by parts, we obtain

$$\|\nabla_A^2 F_A\|^2 \leq c_1 \|\nabla_A F_A\|^2 + c_2 \|F_A\|^2.$$

We now induct. Suppose that $\nabla_A^b F_A \in L_2$, for $b \leq k$. Differentiating (2.14) k times, gives $\nabla_A^k(\nabla_A^* \nabla_A + (\hat{F}_A + \hat{R}))F_A = 0$, and

$$\begin{aligned} 0 &= \langle \nabla_A^k(\nabla_A^* \nabla_A + (\hat{F}_A + \hat{R}))F_A, y_m^2 \nabla_A^k F_A \rangle \\ &\geq \|\nabla_A y_m \nabla_A^k F_A\|^2 - \|dy_m \otimes \nabla_A^k F_A\|^2 - \sum_{0 \leq j \leq k} C_j \|\nabla_A^j F_A\|^2, \end{aligned}$$

for some constants C_j determined by the sup norms of $|\nabla_A^b F_A|$, $b \leq k$ and the other coefficients of $[\nabla_A^k, \nabla_A^*]$. Taking the limit as $m \to \infty$ gives $\nabla_A^{k+1} F_A \in L_2$. q.e.d.

These bounds now allow us to construct test variations from covariant derivatives of F_A.

Proposition 2.17. *Let M be a complete noncompact homogeneous $4-$manifold. Let A be a smooth finite energy Yang–Mills connection on a $G-$bundle E over M. Then $\nabla_A^k F_A \in L_2 \cap L_4 \cap C^\infty$, for all k.*

Proof. We are left to show $\nabla_A^k F_A \in L_4$. By Kato's inequality, $\nabla_A^{k+1} F_A \in L_2 \Rightarrow d|\nabla_A^k F_A| \in L_2$. Hence the scalar function $|\nabla_A^k F_A| \in W^{1,2}$. The Sobolev embedding theorem holds (see [**He**]) on manifolds of bounded geometry. Hence $|\nabla_A^k F_A| \in L_4$. q.e.d.

Remark 2.18. The assumption of homogeneous geometry is stronger than necessary in this section and can be replaced by weaker assumptions of bounded geometry. The constants in Lemmas 2.7 and 2.9 are then replaced by constants depending on bounds on the injectivity radius, the Riemannian curvature, and its derivatives. These geometric bounds would also then be used to bound the coefficients, C_j, in the proof of Lemma 2.13.

3. Conservative Decompositions

Suppose that $\Lambda^2 T_M^* \otimes ad(E)$ decomposes into two orthogonal subbundles

$$\Lambda^2 T_M^* \otimes ad(E) = \Lambda^+(E) \oplus \Lambda^-(E), \tag{3.1}$$

such that ∇_A preserves this decomposition. Let $P^\pm$ denote the projection onto these summands. We call such a decomposition *conservative* if there exists $a, b \in \mathbb{R}$, not both zero, so that

$$a\|P^+ F_A\|^2 + b\|P^- F_A\|^2 \ \text{ is independent of } A. \tag{3.2}$$

The following elementary lemma clarifies the importance of conservative decompositions.

Lemma 3.3. *Given a conservative decomposition, a connection minimizes $YM(A)$ if and only if it minimizes $\|P^- F_A\|^2$ (equivalently, if and only if it minimizes $\|P^+ F_A\|^2$).*

Proof. By (3.1) we have $\|F_A\|^2 = \|P^+ F_A\|^2 + \|P^- F_A\|^2$. The lemma then follows readily from (3.2). q.e.d.

Consequently, for energy minimizing connections and conservative decompositions we have the additional critical point equation :

$$0 = d_A^* P^- F_A \tag{3.4}$$

and the refined variational inequalities:

$$\|P^- F_A\|^2 \leq \|P^- F_{A+\psi}\|^2, \tag{3.5}$$

and

$$0 \leq \|P^- d_A \psi\|^2 + 2\langle P^- F_A, \psi \wedge \psi\rangle. \tag{3.6}$$

The following elementary lemma shows how to begin to extract information about the curvature from the variational inequalities (3.5) and (3.6).

Lemma 3.7. *Let $\psi \in A^1(M, ad(E))$ satisfy $\psi \in L_2 \cap L_4$,*

$$0 = P^- d_A \psi, \ \ \textit{and} \ \ 0 = \langle P^- F_A, \psi \wedge \psi\rangle.$$

Then

$$e^*(\psi) P^- F_A = 0.$$

Proof. Consider the variation $A_t = A + t\psi + t^{\frac{3}{2}} w$, for $w \in A^1(M, ad(E))$ arbitrary. Then expanding (3.5) we have

$$\|P^- F_A\|^2 \leq \|P^- F_A\|^2 + 2\langle P^- F_A, d_A(t\psi + t^{\frac{3}{2}} w)$$
$$+ t^2 \psi \wedge \psi + 2t^{\frac{5}{2}} \psi \wedge w\rangle + t^2 \|P^- d_A \psi\|^2 + O(t^3).$$

Invoking (3.4) and our hypotheses on ψ, this reduces to

$$0 \leq 2\langle P^- F_A, 2t^{\frac{5}{2}} \psi \wedge w\rangle + O(t^3).$$

Replacing w by $-w$, we see that

$$0 = \langle e^*(\psi) P_A^-, w\rangle$$

for all w, and the lemma follows. q.e.d.

In the following sections we will consider $1-$forms ψ constructed from F_A that satisfy the hypotheses of Lemma 3.7 and use them to uncover information about F_A and A.

4. Dimension 4: (Anti-)Self-Duality and Homogeneous Spaces

In this section we assume that M is a 4-dimensional oriented Riemannian homogeneous space with nonnegative sectional curvature. Denote the group of isometries of M by K and its Lie algebra by $\mathfrak{k}$. Identify $\mathfrak{k}$ with the Lie algebra of Killing vector fields on M. Fixing a base point $o \in M$ and a metric on $\mathfrak{k}$ induces a decomposition $\mathfrak{k} = \mathfrak{p} \oplus \mathfrak{u}$, where $\mathfrak{u}$ is the Lie algebra of the isotropy group of o and therefore also the kernel of the evaluation map $\mathfrak{k} \to T_oM$. Because K is the product of an abelian and a compact group, we may choose the metric on $\mathfrak{k}$ to be invariant under the adjoint action of K and so that for every x the evaluation map $\mathfrak{k} \to T_xM$ is an isometry when restricted to the orthogonal complement of its kernel.

Let $\{X_j\}_{j=1}^D$ be a basis of $\mathfrak{k}$. Let $\phi_{j,t}\colon M \to M$, $j = 1, \dots, D$, $t \in \mathbb{R}$, be the associated one parameter families of isometries. We define a pullback map

$$\phi_{j,t}^* : (\Lambda^2 T_M^* \otimes ad(E))_{\phi_{j,t}(x)} \to (\Lambda^2 T_M^* \otimes ad(E))_x$$

by defining the action of $\phi_{j,t}^*$ on the $ad(E)$ factor to be parallel transport along the curve $t \to \phi_{j,t}(x)$. Fix j. Away from a fixed point of $\phi_{j,t}$, we may choose a local frame that is parallel on the integral curves of X_j through all points in a neighborhood of x. In such a frame the connection form, which we also denote A, satisfies

$$i_j A = 0, \text{ and } i_j F_A = i_j d_A A, \tag{4.1}$$

where $i_j = i_{X_j}$ denotes interior multiplication by X_j.

Given a local frame $\{s_a\}_a$ for $ad(E)$, we write an $ad(E)$ valued $p-$form f as $\sum_a f^a \otimes s_a$. Then we have

$$\begin{aligned} \phi_{j,t}^* d_A f &= \sum_a \phi_{j,t}^*(df^a) \otimes \phi_{j,t}^* s_a + \sum_{a,b} (-1)^p \phi_{j,t}^* f^a \otimes \phi_{j,t}^*(A_a^b s_b) \\ &= d_A \phi_{j,t}^* f + 2(\phi_{j,t}^* A - A) \wedge \phi_{j,t}^* f. \end{aligned} \tag{4.2}$$

In four dimensions (oriented) we have the decomposition of $\Lambda^2 T^*M \otimes ad(E)$ given by the decomposition into self-dual and anti-self-dual summands:

$$\Lambda^2 T^*M \otimes ad(E) = (\Lambda_+^2 T^*M \otimes ad(E)) \oplus (\Lambda_-^2 T^*M \otimes ad(E)).$$

Thus the projections onto the summands are given by

$$P^{\pm} = \frac{1}{2}(1 \pm *),$$

where $*$ denotes the Hodge star operator. Let $p_1(A, E)$ denote the first pontrjagin form of E determined by the connection A. Recall that

$$\|F_A^+\|^2 - \|F_A^-\|^2 = c \int_M p_1(A, E), \tag{4.3}$$

where $c > 0$ is determined by normalization of the inner product on $ad(E)$. Chern–Weil theory says that the right hand side is independent of A on compact manifolds. On noncompact manifolds it is constant under variations of A that decay suitably at ∞. Hence $P^{\pm}$ define a conservative decomposition. Because $d_A^* = - * d_A *$, we have for this decomposition

$$d_A P^{\pm} F_A = 0.$$

Having lifted the action of $\phi_{j,t}^*$ to $ad(E)$, we obtain an extension of the Lie derivative $L_j = L_{X_j}$ to $ad(E)$ valued forms. It satisfies the usual relation

$$L_j = d_A i_j + i_j d_A,$$

and of course

$$L_j f = \frac{d}{du}_{|u=0} \phi_{j,u}^* f.$$

Because $\phi_{j,t}$ is an isometry we have

$$\phi_{j,t}^* P^- = P^- \phi_{j,t}^*,$$

and hence, infinitesimally,

$$[L_j, P^-] = 0. \tag{4.4}$$

4.1. Variations. Set

$$F_A^{\pm} = P^{\pm} F_A.$$

The stability results in [**BLS**] and [**BL**] followed in large part from consideration of the variations $A + t i_j F_A^+$. Our theorems in 4-dimensions rely on recognizing these variations as an approximation to the variations $A + i_j F_j^+(t)$, where

$$F_j^+(t) := \int_0^t \phi_{j,s}^* F_A^+(x) ds.$$

Heuristically, this variation may be thought of as an attempt to test whether the isometry invariance of the Yang–Mills energy extends to isometry invariance when only a self-dual component of the connection is shifted by the isometry.

Next we use (4.2) to expand $F_{A+i_jF_j^+(t)}$.

$$\begin{aligned}F_{A+i_jF_j^+(t)} &= F_A + d_A i_j F_j^+(t) + i_j F_j^+(t) \wedge i_j F_j^+(t) \\ &= F_A + L_j F_j^+(t) - 2i_j \int_0^t (A - \phi_{j,s}^* A) \wedge \phi_{j,s}^* F_A^+(x) ds \\ &\quad + i_j F_j^+(t) \wedge i_j F_j^+(t) \\ &= F_A + \int_0^t \frac{d}{ds} \phi_{j,s}^* F_A^+(x) ds \\ &\quad + 2i_j \int_0^t \int_0^s \frac{\partial}{\partial u} \phi_{j,u}^* A \wedge \phi_{j,s}^* F_A^+(x) duds \\ &\quad + i_j \int_0^t \int_0^t \phi_{j,u}^* F_A^+(x) \wedge i_j \phi_{j,s}^* F_A^+(x) duds.\end{aligned}$$

Using (4.1), we have

$$\frac{\partial}{\partial u} \phi_{j,u}^* A = \phi_{j,u}^* i_j d_A A = \phi_{j,u}^* i_j F_A.$$

This and additional manipulations give

$$F_{A+i_jF_j^+(t)} = F_A^- + \phi_{j,t}^* F_A^+ - 2\Phi_j(t), \tag{4.5}$$

where

$$\begin{aligned}\Phi_j(t) &= \int_0^t \int_0^s \phi_{j,u}^* i_j F_A \wedge i_j \phi_{j,s}^* F_A^+(x) duds \\ &\quad - \frac{1}{2} \int_0^t \int_0^t i_j \phi_{j,u}^* F_A^+(x) \wedge i_j \phi_{j,s}^* F_A^+(x) duds \\ &= \int_0^t \int_0^s \phi_{j,u}^* i_j F_A \wedge i_j \phi_{j,s}^* F_A^+(x) duds \\ &\quad - \frac{1}{2} \int_0^t \int_0^s \phi_{j,u}^* i_j F_A^+ \wedge i_j \phi_{j,s}^* F_A^+(x) duds \\ &\quad - \frac{1}{2} \int_0^t \int_s^t i_j \phi_{j,u}^* F_A^+(x) \wedge i_j \phi_{j,s}^* F_A^+(x) duds.\end{aligned}$$

Changing the order of integration in the last term and cancelling reduces the preceding to

$$\Phi_j(t) = \int_0^t \int_0^s i_j \phi_{j,u}^* F_A^-(x) \wedge i_j \phi_{j,s}^* F_A^+(x) duds. \tag{4.6}$$

For later application it is useful to Taylor expand Φ_j. We have for all integers B,

$$\Phi_j(t) = \sum_{a,b\geq 0}^{a+b=B} \frac{t^{a+b+2} i_j L_j^a F_A^-(x) \wedge i_j L_j^b F_A^+(x)}{(a+1)!b!(a+b+2)} + O(t^{B+3}). \tag{4.7}$$

With this notation, (3.5) becomes

$$\|F_A^-\|^2 \leq \|F_A^- - 2P^-\Phi_j(t)\|^2, \tag{4.8}$$

or in a more useful form:

$$\langle F_A^-, \Phi_j(t)\rangle \leq \|P^-\Phi_j(t)\|^2, \tag{4.9}$$

Equivalently

$$\|F_A^+\|^2 = \|\phi_{j,t}^* F_A^+\|^2 \leq \|\phi_{j,t}^* F_A^+ - 2P^+\Phi_j(t)\|^2, \tag{4.10}$$

The right hand side of (4.9) is evidently $O(t^4)$, implying the nonpositivity of the $O(t^2)$ terms in the left hand side. The Taylor expansion (4.7) gives for each j

$$0 \geq \langle F_A^-, i_j F_A^- \wedge i_j F_A^+\rangle \ \text{(no } j \text{ sum)}. \tag{4.11}$$

Switching the roles of P^- and P^+, we similarly deduce

$$0 \geq \langle F_A^+, i_j F_A^+ \wedge i_j F_A^-\rangle \ \text{(no } j \text{ sum)}. \tag{4.12}$$

Lemma 4.13. *Let f^+ be a self-dual 2-form and f^- an anti-self-dual 2-form. Let $\{e_1, e_2, e_3, e_4\}$ be a local orthonormal frame for TM. Then $\sum_a i_{e_a} f^+ \wedge i_{e_a} f^+$ is self-dual, and $\sum_a i_{e_a} f^- \wedge i_{e_a} f^-$ is anti-self-dual. If ϕ_1, ϕ_2, and ϕ_3 are $ad(E)$ valued 2-forms, then*

$$\sum_a \langle \phi_1, i_{e_a}\phi_2 \wedge i_{e_a}\phi_3\rangle = \sum_a \langle \phi_3, i_{e_a}\phi_1 \wedge i_{e_a}\phi_2\rangle.$$

Proof. This is an elementary computation. q.e.d.

As proved in [**BLS**],[**BL**], we now obtain our first commutation result:

Proposition 4.14.

$$0 = [F_{st}^+, F_{ij}^-],$$

for all indices s, t, i, j.

Proof. Summing (4.11), we obtain

$$0 \geq \sum_j \langle F_A^-, i_j F_A^- \wedge i_j F_A^+\rangle.$$

Because the evaluation map $\mathfrak{k} \to T_m M$ is an isometry on the orthogonal complement to its kernel, the pointwise inner product,

$$\sum_j \langle F_A^-, i_j F_A^- \wedge i_j F_A^+\rangle(m) = \sum_a \langle F_A^-, i_{e_a} F_A^- \wedge i_{e_a} F_A^+\rangle(m),$$

for a local orthonormal frame $\{e_a\}_{a=1}^4$. Applying Lemma 4.13, we see that

$$\sum_a \langle F_A^-, i_{e_a} F_A^- \wedge i_{e_a} F_A^+\rangle(m) = \sum_a \langle i_{e_a} F_A^- \wedge i_{e_a} F_A^-, F_A^+\rangle(m) = 0.$$

Hence the inequality (4.11) is actually an equality for each j:

$$0 = \langle F_A^-, i_j F_A^- \wedge i_j F_A^+\rangle = \langle F_A^+, i_j F_A^- \wedge i_j F_A^-\rangle \text{ (no } j \text{ sum)}. \tag{4.15}$$

Symmetrically we obtain

$$0 = \langle F_A^-, i_j F_A^+ \wedge i_j F_A^+\rangle \text{ (no } j \text{ sum)}. \tag{4.16}$$

On the other hand, we have

$$P^- d_A i_j F_A^+ = -P^- i_j d_A F_A^+ + L_j P^- F_A^+ = 0.$$

Hence $\psi = i_j F_A^+$ satisfies the hypotheses of Lemma 3.7, implying

$$e^*(i_j F_A^+) F_A^- = 0.$$

Expanding this equality in components, using the duality relations, and allowing $X_j(m)$ to run over a basis of $T_m M$, we obtain the claimed commutation result:

$$0 = [F_{st}^+, F_{ij}^-],$$

for all indices s, t, i, j. q.e.d.

4.2. Inductive hypothesis. In order to move beyond the commutation of the self-dual with the anti-self-dual components of the curvature to the construction of ∇_A stable subbundles k^+ and k^- of $ad(E)$ with self-dual (respectively anti-self-dual) curvature, we wish to prove the following proposition.

Proposition 4.17. *$[\nabla_A^i F_A^+, \nabla_A^j F_A^-] = 0$ for all i and j.*

In this and the next subsection, we prove this proposition by induction on $i + j$.

Denote by A_N the inductive hypothesis:

$$A_N : \qquad [\nabla_A^i F_A^+, \nabla_A^j F_A^-] = 0, \text{ for } i + j < N. \tag{4.18}$$

We have established A_1 in Proposition 4.14. We will show A_N implies A_{N+1}. Assume A_N holds, for some $N \geq 1$. In the inductive hypothesis, powers of covariant derivatives can be replaced by powers of Lie derivatives, since they differ by lower order terms.

Observe that

$$A_N \Rightarrow \Phi_j(t) = O(t^{N+2}). \tag{4.19}$$

Hence A_N and equation (4.9) imply

$$\langle F_A^-, \Phi_j(t)\rangle \leq O(t^{2N+4}). \tag{4.20}$$

Set

$$S_j(t) = \langle F_A^-, \Phi_j(t)\rangle.$$

Then $S_j(0) = 0$, and

$$S_j'(t) = \langle F_A^-, \int_0^t i_j \phi_{j,u}^* F_A^- \wedge i_j \phi_{j,t}^* F_A^+ du\rangle \tag{4.21}$$

$$= \langle F_A^-, \phi_{j,t}^* \int_{-t}^0 i_j \phi_{j,u}^* F_A^- \wedge i_j F_A^+ du \rangle.$$

Hence $S_j'(0) = S_j''(0) = 0$.

In order to use the variational inequality (4.20) we need to estimate $S_j(t)$. Taylor expanding S_j' gives

$$S_j'(t) = \langle F_A^-, \sum_{a,b\geq 0}^{a+b=m} \frac{t^a}{a!} L_j^a \int_{-t}^0 i_j \frac{u^b}{b!} L_j^b F_A^- \wedge i_j F_A^+ du \rangle + O(t^{m+2})$$

Hence

$$S_j(t) = \langle F_A^-, \sum_{a,b\geq 0}^{a+b=m} \frac{(-1)^b t^{a+b+2}}{a!(b+1)!(a+b+2)} L_j^a [i_j L_j^b F_A^- \wedge i_j F_A^+] \rangle + O(t^{m+3})$$

$$= \langle F_A^-, \sum_{a,b\geq N}^{a+b=m} \frac{(-1)^b t^{a+b+2}}{a!(b+1)!(a+b+2)} L_j^a [i_j L_j^b F_A^- \wedge i_j F_A^+] \rangle + O(t^{m+3})$$

Hence, using A_N to eliminate lower order terms, (4.20) implies

$$\langle F_A^-, (-1)^N L_j^N [i_j L_j^N F_A^- \wedge i_j F_A^+] \rangle \leq 0. \tag{4.22}$$

Lemma 4.23. *If A_N holds then*

$$\langle F_A^-, L_j^N [i_j L_j^N F_A^- \wedge i_j F_A^+] \rangle = 0.$$

Proof. Set $S(X_j, N) = \langle F_A^-, L_j^N [i_j L_j^N F_A^- \wedge i_j F_A^+] \rangle$. The inequality $(-1)^N S(X_j, N) \leq 0$ holds when X_j is replaced by any Killing vector. We will show that the average of $S(X_j, N)$ over the unit sphere of $\mathfrak{k}$ is zero. Hence $S(X_j, N)$ is zero for each j (and each choice of basis of $\mathfrak{k}$). To see this we consider $S(\sum_k y^k X_k, N)$ and integrate the resulting degree $2N+2$ homogeneous polynomial in y over the unit sphere. Integration of homogeneous degree $2N+2$ polynomials over the sphere projects onto the span of the radial function, $(r^2)^{N+1}$. Expanding in a multi-index notation where $L_J = L_{j_1} \cdots L_{j_{|J|}}$, we write

$$S(\sum_k y^k X_k, N) = \sum_{|I|=N, |J|=N} \sum_{m,p=1}^{dim \mathfrak{k}} y^I y^J y^m y^p \langle F_A^-, L_I [i_m L_J F_A^- \wedge i_p F_A^+] \rangle.$$

That integration over the unit sphere projects onto radial functions implies that upon integration of S, we are left with a linear combination of coefficients, $\langle F_A^-, L_I [i_m L_J F_A^- \wedge i_p F_A^+] \rangle$, of S where the indices are contracted pairwise. We will see that all such contractions vanish. The condition A_N allows us to replace the Lie derivatives by covariant derivatives, as the difference vanishes in the inner product. We can also drop commutators of derivatives by A_N, as all such commutations drop the degree of the differentiation and thus lead to terms which vanish by A_N. We can then use the Yang–Mills equation and A_N to equate to zero all inner products containing $\sum_k i_k L_k F_A^{\pm}$ terms or $\sum_k L_k^2 F_A^{\pm}$ terms. We also invoke Lemma 4.13 to remove terms with paired indices $m = p$ on the interior products. Thus m and p must both pair with elements of I, for if p pairs with an L_{j_r}, we can use the Leibniz formula, A_N and the Yang–Mills equation to eliminate the corresponding term. This leaves two L_j terms which must pair with each other, but since

$$0 = -(d_A d_A^* +_A^* d_A) F_A^{\pm}$$

$$= \sum_j L_j^2 F_A^{\pm} \quad \text{(modulo terms vanishing in the inner product by } A_N\text{)},$$

we find the average vanishes. Hence

$$\langle F_A^-, L_j^N [i_j L_j^N F_A^- \wedge i_j F_A^+] \rangle = 0,$$

as claimed. q.e.d.

Now we apply a variant of Lemma 3.7 to obtain a commutation result.

Lemma 4.24. *If A_N holds then*

$$e^*(i_j L_j^N F_A^+) F_A^- = 0.$$

Proof. Once again we let ψ be a smooth compactly supported $ad(E)$ valued $1-$form. Then

$$\|F_A^-\|^2 \le \|F^-_{A + i_j F_j^+(t) + t^p \psi}\|^2$$

$$= \|F_A^- - 2P^- \Phi_j(t) + t^p d_A \psi + t^{2p} \psi \wedge \psi + 2t^p \psi \wedge i_j F_j^+(t)\|^2$$

$$= \|F_A^-\|^2 - 4S_j(t) + 4t^p \langle F_A^-, \psi \wedge i_j F_j^+(t) \rangle + O(t^{2N+4} + t^{2p}).$$

Lemma 4.23 implies $S_j(t) = O(t^{2N+3})$. Hence Taylor expanding again, we get

$$0 \le \sum_{b=0}^{N} 4t^p \langle F_A^-, \psi \wedge \frac{t^{b+1}}{(b+1)!} i_j L_j^b F_A^+ \rangle + O(t^{2N+3} + t^{2p})$$

$$= 4t^p \langle F_A^-, \psi \wedge \frac{t^{N+1}}{(N+1)!} i_j L_j^N F_A^+ \rangle + O(t^{2N+3} + t^{2p} + t^{N+p+2}).$$

Choosing $p = N + \frac{3}{2}$ gives

$$0 \le \langle F_A^-, \psi \wedge i_j L_j^N F_A^+ \rangle.$$

Replacing ψ with $-\psi$ makes the inequality an equality, and we conclude

$$0 = e^*(i_j L_j^N F_A^+) F_A^-$$

as desired. q.e.d.

4.3. An Algebraic Reduction. We now complete our induction argument by proving the following proposition.

Proposition 4.25. *The assumption A_N implies A_{N+1}.*

Proof. Assume A_N holds. By Lemma 4.24, $e^*(i_j L_j^N F_A^+)F_A^- = 0$ for all j. The Lie derivative L_j differs from the covariant derivative ∇_j by zero order terms which commute with every element of $ad(E)$. Hence we have

$$0 = e^*(i_j \nabla_j^N F_A^+)F_A^-. \tag{4.26}$$

Fix a point x and a basis for the infinitesimal isometries so that $X_j(x) = e_j$, $j = 1, 2, 3, 4$ is an orthonormal basis of T_xM. Then expanding (4.26) in this frame, yields for all k and t,

$$\sum_s [\nabla_k^N F_{ks}^+, F_{st}^-] = 0 \qquad \text{(no k sum)}. \tag{4.27}$$

In fact, replacing e_k by $u^1 e_1 + \cdots u^4 e_4$, for any $(u_1, u_2, u_3, u_4) \in \mathbb{R}^4$, we have for all t,

$$\sum_{a,s} [(\sum_j u^j \nabla_j)^N u^a F_{as}^+, F_{st}^-] = 0. \tag{4.28}$$

Set

$$p_{ik}(u) = [(\sum_j u^j \nabla_j)^N F_{1i}^+, F_{1k}^-],$$

and let $p_{ik,a} := \frac{\partial p_{ik}}{\partial u^a}$. In our local orthonormal frame, duality implies that the components of F_A^+ and F_A^- (and their covariant derivatives) satisfy $F_{12}^\pm = \pm F_{34}^\pm$, $F_{13}^\pm = \pm F_{42}^\pm$, and $F_{14}^\pm = \pm F_{23}^\pm$, and consequently each term $\big[\big(\sum_j u^j \nabla_j\big)^N u^a F_{as}^+, F_{st}^-\big]$ in (4.28) is equal to $p_{ik}(u)$ for some i and k. Thus we may expand (4.28) as

$$\begin{aligned} 0 &= u^1(p_{22} + p_{33} + p_{44}) - u^2(p_{34} - p_{43}) \\ &\quad + u^3(p_{24} - p_{42}) - u^4(p_{23} - p_{32}). \end{aligned} \tag{4.29}$$

$$\begin{aligned} 0 &= u^1(p_{34} - p_{43}) + u^2(-p_{22} + p_{44} + p_{33}) \\ &\quad - u^3(p_{32} + p_{23}) - u^4(p_{42} + p_{24}). \end{aligned} \tag{4.30}$$

$$\begin{aligned} 0 &= u^1(-p_{24} + p_{42}) - u^2(p_{23} + p_{32}) \\ &\quad + u^3(-p_{33} + p_{44} + p_{22}) - u^4(p_{43} + p_{34}). \end{aligned} \tag{4.31}$$

$$\begin{aligned} 0 &= u^1(p_{23} - p_{32}) - u^2(p_{24} + p_{42}) - u^3(p_{34} + p_{43}) \\ &\quad + u^4(-p_{44} + p_{33} + p_{22}). \end{aligned} \tag{4.32}$$

In addition to these equations, we have the Yang–Mills equations and the Bianchi identities. These are best encoded as relations between the derivatives of the p_{ij} as follows.

$$\sum_{i=2}^{4} p_{ik,i} = N[(u^j\nabla_j)^{N-1}(d_A^* F^+)_1, F^-_{1k}] = 0,$$

and

$$\begin{aligned} p_{ik,a}(u) &= N[(u^j\nabla_j)^{N-1}F^+_{1i,a}, F^-_{1k}] \\ &= N[(u^j\nabla_j)^{N-1}F^+_{1a,i}, F^-_{1k}] + N[(u^j\nabla_j)^{N-1}F^+_{ai,1}, F^-_{1k}] \\ &= p_{ak,i}(u) + N[(u^j\nabla_j)^{N-1}F^+_{ai,1}, F^-_{1k}]. \end{aligned}$$

This reduces to the following 12 equations.

$$p_{2k,2} + p_{3k,3} + p_{4k,4} = 0. \tag{4.33}$$

$$p_{2k,1} - p_{4k,3} + p_{3k,4} = 0. \tag{4.34}$$

$$p_{2k,4} - p_{4k,2} - p_{3k,1} = 0. \tag{4.35}$$

$$p_{2k,3} + p_{4k,1} - p_{3k,2} = 0. \tag{4.36}$$

We may also use A_N to shift the derivative to the F^-_A term, yielding the following relations among the derivatives.

$$\sum_{k=2}^{4} p_{ik,k} = -N[(u^j\nabla_j)^{N-1}F^+_{1i}, (d_A^* F^-_A)_1] = 0,$$

and

$$\begin{aligned} p_{ik,a} = N[(u^j\nabla_j)^{N-1}F^+_{1i,a}, F^-_{1k}] &= -N[(u^j\nabla_j)^{N-1}F^+_{1i}, F^-_{1k,a}] \\ = -N[(u^j\nabla_j)^{N-1}F^+_{1i}, F^-_{ak,1}] &- N[(u^j\nabla_j)^{N-1}F^+_{1i}, F^-_{1a,k}] \\ = N[(u^j\nabla_j)^{N-1}&F^+_{1i,1}, F^-_{ak}] + p_{ia,k}. \end{aligned}$$

This yields 12 additional equations.

$$p_{i2,2} + p_{i3,3} + p_{i4,4} = 0. \tag{4.37}$$

$$p_{i2,1} + p_{i4,3} - p_{i3,4} = 0. \tag{4.38}$$

$$p_{i2,4} + p_{i3,1} - p_{i4,2} = 0. \tag{4.39}$$

$$p_{i2,3} - p_{i4,1} - p_{i3,2} = 0. \tag{4.40}$$

Homogeneity allows us to reconstruct the p_{ik} readily from their derivatives:

$$p_{ik} = \frac{1}{N} u^a p_{ik,a}. \tag{4.41}$$

Hence, equations (4.33) - (4.40) allow in (4.29) - (4.32) the replacement of $p_{22}, p_{23}, p_{24}, p_{32}$, and p_{42} by linear combinations of p_{33}, p_{44}, p_{34}, and p_{43}. It is more convenient, however, to differentiate (4.29) - (4.32) and make the replacement at the level of derivatives.

Denote the quantities on the righthand side of equations (4.29) - (4.32) by I_1, I_2, I_3, and I_4, respectively. Differentiating I_j and replacing the derivatives of $p_{22}, p_{23}, p_{24}, p_{32}$, and p_{42} by linear combinations of derivatives of p_{33}, p_{44}, p_{34}, and p_{43} gives

$$\begin{aligned} 0 = I_{1,1} = & (p_{22} + p_{33} + p_{44}) \\ & + u^1(p_{43,2} - p_{34,2} + 2p_{33,1} + 2p_{44,1}) - u^2(p_{34,1} - p_{43,1}) \\ & + u^3(2p_{44,3} - p_{34,4} - p_{43,4}) - u^4(p_{43,3} + p_{34,3} - 2p_{33,4}) \end{aligned} \tag{4.42}$$

$$\begin{aligned} 0 = I_{2,2} = & u^1(p_{34,2} - p_{43,2}) - (p_{22} - p_{44} - p_{33}) \\ & + u^2(p_{34,1} - p_{43,1} + 2p_{44,2} + 2p_{33,2}) \\ & + u^3(2p_{33,3} + p_{34,4} + p_{43,4}) + u^4(p_{43,3} + p_{34,3} + 2p_{44,4}) \end{aligned} \tag{4.43}$$

$$\begin{aligned} 0 = I_{3,3} = & u^1(2p_{44,1} + p_{43,2} - p_{34,2}) \\ & - u^2(p_{34,1} - p_{43,1} + 2p_{33,2}) - (p_{33} - p_{44} - p_{22}) \\ & + u^3(-2p_{33,3} + 2p_{44,3} - p_{43,4} - p_{34,4}) - u^4(p_{43,3} + p_{34,3}) \end{aligned} \tag{4.44}$$

$$\begin{aligned} 0 = I_{4,4} = & u^1(p_{43,2} - p_{34,2} + 2p_{33,1}) \\ & - u^2(p_{34,1} - p_{43,1} + 2p_{44,2}) - u^3(p_{34,4} + p_{43,4}) \\ & + u^4(2p_{33,4} - 2p_{44,4} - p_{34,3} - p_{43,3}) + (-p_{44} + p_{33} + p_{22}). \end{aligned} \tag{4.45}$$

Combining (4.42) and (4.43) and using (4.41) gives

$$\begin{aligned} 0 = I_{1,1} + I_{2,2} = & 2p_{33} + 2p_{44} + 2u^1(p_{33,1} + p_{44,1}) + 2u^2(p_{33,2} + p_{44,2}) \\ & + 2u^3(p_{33,3} + p_{44,3}) + 2u^4(p_{33,4} + p_{44,4}) \\ = & 2(N+1)(p_{33} + p_{44}). \end{aligned}$$

Hence $p_{33} = -p_{44}$. Similarly $0 = I_{1,1} - I_{3,3} = (N+1)2p_{33}$, and we conclude that

$$0 = p_{33} = p_{44}. \tag{4.46}$$

Take the u^2 derivative of I_1 and use the relations (4.33) - (4.40) and (4.46) to replace all $p_{ij,k}$ in the resulting equation by derivatives of $p_{34} - p_{43}$. This yields

$$\begin{aligned} 0 = I_{1,2} = & - u^1(p_{34,1} - p_{43,1}) - (p_{34} - p_{43}) - u^2(p_{34,2} - p_{43,2}) \\ & - u^3(p_{34,3} - p_{43,3}) - u^4(p_{34,4} - p_{43,4}) \\ = & - (N+1)(p_{34} - p_{43}). \end{aligned}$$

Hence $p_{34} = p_{43}$. Similarly $0 = I_{3,4} = -(N+1)(p_{43} + p_{34})$, and we conclude

$$0 = p_{34} = p_{43}. \tag{4.47}$$

Using homogeneity, (4.46), (4.47), and (4.33) - (4.40), we deduce that all remaining p_{ik} vanish. We conclude: for all i, k,

$$0 = [(u^j \nabla_j)^N F^+_{1i}, F^-_{1k}].$$

Hence for all i, k, s, t

$$0 = [\nabla^N F^+_{si}, F^-_{tk}].$$

Finally we have, via application of A_N, that for $0 \leq a \leq N$,

$$0 = [\nabla^a F^+_{si}, \nabla^{N-a} F^-_{tk}].$$

Hence A_{N+1} holds. This completes our proof of Proposition 4.25, and, by induction, our proof of Proposition 4.17. q.e.d.

4.4. Splitting. Define subsheaves of the sheaf of sections of $ad(E)$ by setting $k^+_x \subset ad(E)_x$ (respectively k^-_x) to be the subspace generated by the components of $\nabla^a F^+(x)$, (respectively $\nabla^a F^-(x)$) as a runs through all nonnegative integers. By definition, the subsheaves $\mathcal{K}^\pm$ of sections of $ad(E)$ which take values in $k^\pm$ are preserved by the connection. This gives a reduction of the adjoint bundle (which need not be proper - for example when A is an instanton).

Fix a point o on M. Choose a gauge in a neighborhood of o in which the connection form A satisfies $d^*A = 0$. (See, for example, [**U1**] or [**W**] for the existence of such gauges). In this gauge the Yang–Mills equations become a nonlinear elliptic system for A. The homogeneous manifolds we are considering are all real analytic, and this system has analytic coefficients. Hence A and F_A are real analytic in this gauge (See [**M**]. See also [**JT**, Chapter V, Theorem 1.1]).

Fix a point o in M and analytic coordinates in a neighborhood of o. Let $X_1(x), \cdots, X_d(x)$ be linear combinations of components (in the analytic coordinate system) of $\nabla^a_A F^+_A(x)$, any a, such that $\{X_1(o), \cdots, X_d(o)\}$ is a basis of k^+_o. Then $X_1(x), \cdots, X_d(x)$ are linearly independent in a neighborhood of o. Suppose that these vectors do not span k^+_x, x near o. Then there exists an analytic local section, X_{d+1}, constructed from the components of the covariant derivatives of F^+_A which, at x, is linearly independent of $X_1(x), \cdots, X_d(x)$. Then $X_1 \wedge \cdots \wedge X_d \wedge X_{d+1}$ vanishes to infinite order at o. By analyticity, this implies $X_1 \wedge \cdots \wedge X_d \wedge X_{d+1}$ is identically zero in a neighborhood of o, contradicting our assumption. Thus we see that when they are nonzero, k^+ and k^- define subbundles of $ad(E)$ (although one of these is the zero bundle when A is an instanton). This gives the following theorem.

Theorem 4.48. *Let E be a principal $G-$bundle on a complete, oriented, nonnegatively curved homogeneous Riemannian manifold, M. Let A be a locally minimizing Yang–Mills connection on E. The adjoint bundle, $ad(E)$, contains two ∇_A-stable subbundles, k^+ and k^-, satisfying $F^\pm_A$ is a section of $\Lambda^2 T^*M \otimes k^\pm$,*

$$[k^+, k^-] = 0,$$

and the curvature of k^+ is self-dual and that of k^- is anti-self-dual.

Proof. The connection ∇_A preserves $k^\pm$ by construction. Hence the curvature of each subbundle is simply the restriction of the adjoint action of F_A to that subbundle. By Proposition 4.17, $[k^+, k^-] = 0$; in particular, $F_A^\pm$ acts trivially on $k^\mp$, and the curvature operator on $k^\pm$ is $F_A^\pm$, yielding the asserted self-duality and anti-self-duality. q.e.d.

Remark 4.49. Observe that in passing to the adjoint bundle and therefore the adjoint representation of the curvature, we may lose some information about the curvature. In the extreme case of a line bundle, the adjoint bundle is trivial and its induced curvature vanishes. More generally, we lose information about any summand of the curvature in the kernel of the adjoint representation.

Corollary 4.50. *Let M be a compact homogeneous 4-manifold. Let E be a principal $G-$bundle over M with locally minimizing Yang–Mills connection A. Suppose the first Pontrjagin number of E is greater than or equal to zero (respectively less than or equal to zero). Then if the Yang–Mills energy of A is greater than the topological lower bound determined by the first Pontrjagin number of E, $ad(E)$ has a nontrivial subbundle with anti-self-dual curvature (respectively self-dual curvature).*

Proof. From equation (4.3), we have for some $c > 0$, $\|F_A^+\|^2 - \|F_A^-\|^2 = c\int_M p_1(A, E)$. The topological lower bound is then $\|F_A^+\|^2 + \|F_A^-\|^2 \geq |c\int_M p_1(A, E)|$. The hypotheses of the corollary imply

$$\|F_A^+\|^2 + \|F_A^-\|^2 > |\|F_A^+\|^2 - \|F_A^-\|^2|,$$

which is possible only if both $\|F_A^+\|$ and $\|F_A^-\|$ are nonzero. Hence both k^+ and k^- have positive dimension. q.e.d.

5. Dimension 3

We now consider applications of the variational inequality to three dimensions.

Theorem 5.1. *Let Y be a complete 3-dimensional manifold with nonnegative Ricci curvature. Let A be a Yang–Mills minimizing connection on a bundle E on Y. If M is noncompact, assume further that $F_A \in L^4$. Then*

$$\nabla_A F_A = 0,$$

and

$$Ric(*F_A, *F_A) = 0.$$

Moreover, F_A takes values in a flat abelian subbundle of $ad(E)$, and $F_A = 0$ unless M has local flat factors.

Proof. The variational inequality (2.5) gives

$$\|F_A\|^2 \leq \|F_{A+t*F_A}\|^2 = \|F_A + t^2(*F_A) \wedge (*F_A)\|^2.$$

Hence,

$$0 \leq \langle F_A, (*F_A) \wedge (*F_A) \rangle. \tag{5.2}$$

Let R denotes the Riemann curvature. Combining the Yang–Mills equation and the Bochner formula, we have

$$0 = \|d_A * F_A\|^2 + \|d_A^* * F_A\|^2$$

$$= \|\nabla_A * F_A\|^2 + \int_Y Ric(*F_A, *F_A)dv + 2\langle F_A, (*F_A) \wedge (*F_A) \rangle.$$

Thus if the Ricci curvature of Y is nonnegative, we conclude from this Bochner formula and inequality (5.2) that

$$0 = \langle F_A, (*F_A) \wedge (*F_A) \rangle,$$

and

$$\nabla_A F_A = 0. \tag{5.3}$$

If the Ricci curvature is strictly positive at some point, then $F_A = 0$. Equation (5.3) implies that the subbundle H of $ad(E)$ generated by the components of F_A is stable under ∇_A.

Applying Lemma 3.7 with $P^- = 1$ and $\psi = *F_A$, we deduce

$$0 = [F_{ij}, F_{st}],$$

for all i, j, s, t. Hence H is an abelian flat subbundle of $ad(E)$. Now $\nabla_A F_A = 0$ implies the Riemannian curvature acts trivially on the subbundle of T^*M determined by F_A. Thus M has local flat factors unless $F_A = 0$.

q.e.d.

Up to diffeomorphism, the only simply connected 3-manifold with strictly positive Ricci curvature is S^3 (see [**Ha**]). Uhlenbeck's compactness theorem [**U1**] implies that every $G-$bundle on a compact 3-manifold has a Yang–Mills minimizing $G-$connection. Theorem 5.1 then implies the minimizing connection is flat on S^3 and its smooth finite quotients. On S^3 it is therefore trivial as a $G-$bundle. As G-bundles on S^3 are classified by $\pi_2(G)$, this gives an analytic proof of the well known fact that

$$\pi_2(G) = 0$$

for all compact connected Lie groups. (See [**Bo**, Section 18]). We similarly deduce that all $G-$bundles on T^3 admit $G-$connections with covariant constant curvature.

6. Calabi–Yau 3-folds

Let M be a compact Calabi–Yau 3-fold, with Kahler form ω and nonzero covariant constant $(3,0)$ form Ω. Let $L := e(\omega)$. Let A be a connection on a $G-$bundle, E, over M.

Decompose the curvature, F_A as

$$F_A = F_A^{2,0} + F_{A0}^{1,1} + \phi_A\omega + F_A^{0,2},$$

where

$$\phi_A := \frac{1}{3}L^* F_A.$$

The Kahler identity (see for example [**We**, Theorem 3.16, p.187])

$$LF_A = *(F_A^{2,0} + 2\phi_A\omega - F_{A0}^{1,1} + F_A^{0,2}) \tag{6.1}$$

implies, after wedging with F_A, taking the trace, and integrating, that (6.2)

$$4\|F_A^{0,2}\|^2 + 9\|\phi_A\|^2 - \|F_A\|^2 = -\int tr(F\wedge F)\wedge\omega = c\int_M p_1(A,E)\wedge\omega$$

and is therefore independent of the connection.

Define two new energies,

$$E'(A) := 4\|F_A^{0,2}\|^2, \text{ and } E''(A) = 9\|\phi_A\|^2.$$

Minimizing the Yang–Mills energy is therefore equivalent to minimizing $E'(A) + E''(A)$:

$$YM(A) = E'(A) + E''(A) + \text{topological constant.} \tag{6.3}$$

The energy functional E'' plays an important role in the study of Hermitian–Einstein connections. (See [**K**], [**D**], [**UY**], and [**LT**]). Recall that a connection on a holomorphic vector bundle on a Kahler manifold is called *Hermitian–Einstein* if $\phi_A = \alpha I_E$, for some constant α. Here I_E denotes the identity endomorphism. Hermitian–Einstein connections are critical points of E''.

We now study stable critical points of $E'(A)$. Critical points of E' satisfy

$$\bar\partial_A^* F_A^{0,2} = 0. \tag{6.4}$$

Define an $ad(E)$ valued $(0,1)$ form, ψ_A, so that

$$e^*(\psi_A)\bar\Omega = F_A^{0,2}. \tag{6.5}$$

More explicitly, in a local special unitary frame,

$$\psi_A = F_{23}^{2,0}d\bar z^1 + F_{31}^{2,0}d\bar z^2 + F_{12}^{2,0}d\bar z^3.$$

Applying $\bar\partial_A^*$ to each side of (6.5) gives

$$e^*(\bar\partial_A\psi_A)\bar\Omega = 0,$$

and therefore

$$\bar{\partial}_A \psi_A = 0. \tag{6.6}$$

The Bianchi identity implies $\partial_A F_A^{2,0} = 0$, which is equivalent to

$$\bar{\partial}_A^* \psi_A = 0. \tag{6.7}$$

Assume now that A is a stable critical point of E'. By definition, the second variation of $E'(A)$ is then positive:

$$0 \le \|\bar{\partial}_A \eta^{0,1}\|^2 + 2\mathrm{Re}\langle F_A^{0,2}, \eta^{0,1} \wedge \eta^{0,1}\rangle,$$

for all $ad(E)$ valued $(0,1)$ forms $\eta^{0,1}$. Taking $\eta = \lambda\psi_A + \bar{\lambda}\bar{\psi}_A$, $\lambda \in \mathbb{C}$, as our test variation and applying (6.6) gives

$$0 \le \|\bar{\partial}_A \lambda\psi_A\|^2 + 2\mathrm{Re}\langle F_A^{0,2}, \lambda\psi_A \wedge \lambda\psi_A\rangle = 2\mathrm{Re}\bar{\lambda}^2\langle F_A^{0,2}, \psi_A \wedge \psi_A\rangle.$$

Choosing $\lambda^2 = -\langle F_A^{0,2}, \psi_A \wedge \psi_A\rangle$, we see that $\langle F_A^{0,2}, \psi_A \wedge \psi_A\rangle = 0$. We may now argue as in Lemma 3.7 to deduce

$$0 = e^*(\psi_A) F_A^{0,2}.$$

In components, this is equivalent to

$$0 = [F_{su}^{0,2}, F_{ab}^{0,2}], \tag{6.8}$$

all s, u, a, b. The components of $F_A^{0,2}$ thus generate an abelian subalgebra of $ad(E) \otimes \mathbb{C}$. In particular,

$$\psi_A \wedge \psi_A = 0, \tag{6.9}$$

and

$$F^{0,2}_{A+t(\lambda\psi_A + \bar{\lambda}\bar{\psi}_A)} = F_A^{0,2}, \tag{6.10}$$

for all t. We can also show that

$$\phi_{A+t(\lambda\psi_A + \bar{\lambda}\bar{\psi}_A)} = \phi_A, \tag{6.11}$$

for all t. To see this, first observe that for a general smooth one parameter family of connections, A_t,

$$\frac{d}{dt}\phi_{A_t} = \frac{d}{dt}\frac{1}{3}L^* F_{A_t} = \frac{1}{3}L^* d_{A_t}\frac{d}{dt}A_t = \frac{1}{3}(i\bar{\partial}_{A_t}^* - i\partial_{A_t}^*)\frac{d}{dt}A_t.$$

Here we have used the Hodge identities:

$$[L^*, \bar{\partial}_A] = -i\partial_A^*, \text{ and } [L^*, \partial_A] = i\bar{\partial}_A^*. \tag{6.12}$$

Choosing $A_t = A + t(\lambda\psi_A + \bar{\lambda}\bar{\psi}_A)$, $\lambda \in \mathbb{C}$, we have

$$\frac{1}{3}(i\bar{\partial}_{A_t}^* - i\partial_{A_t}^*)\frac{d}{dt}A_t = \frac{1}{3}(i\lambda\bar{\partial}_{A_t}^*\psi_{A_t} - i\bar{\lambda}\partial_{A_t}^*\bar{\psi}_{A_t}) = 0.$$

Note that here we have used $\psi_A = \psi_{A_t}$ for this variation (see (6.10)) to extend (6.7) from $\bar{\partial}_A^*\psi_A = 0$ to $\bar{\partial}_{A_t}^*\psi_A = 0$. We have the following lemma.

Lemma 6.13. *If A is an E' minimizing connection, then*

$$YM(A + t(\lambda\psi_A + \bar{\lambda}\bar{\psi}_A)) = YM(A), \ \ \forall t, \lambda. \tag{6.14}$$

Proof. This follows immediately from (6.3),(6.10), and (6.11). q.e.d.

The preceding lemma implies that the quartic polynomial in t, $YM(A + t(\lambda\psi_A + \bar{\lambda}\bar{\psi}_A))$, is in fact constant. Hence its t, t^2, t^3, and t^4 coefficients vanish. The four resulting equations, simplified by applying (6.9), are:

$$0 = \langle F_A, d_A(\lambda\psi_A + \bar{\lambda}\bar{\psi}_A)\rangle. \tag{6.15}$$

$$0 = \|d_A(\lambda\psi_A + \bar{\lambda}\bar{\psi}_A)\|^2 + 4|\lambda|^2\langle F_A, \psi_A \wedge \bar{\psi}_A\rangle \tag{6.16}$$

$$0 = \langle d_A(\lambda\psi_A + 2|\lambda|^2\bar{\lambda}\bar{\psi}_A), \psi_A \wedge \bar{\psi}_A\rangle \tag{6.17}$$

$$0 = |\lambda|^4\|\psi_A \wedge \bar{\psi}_A)\|^2. \tag{6.18}$$

The quartic relation immediately implies

$$0 = \psi_A \wedge \bar{\psi}_A, \tag{6.19}$$

and therefore the subalgebra of $ad(E) \otimes \mathbb{C}$ generated by the components of $F_A^{0,2}$ and $F_A^{2,0}$ is abelian. Simplifying the quadratic relation (6.16) with (6.19) yields

$$0 = d_A(\lambda\psi_A + \bar{\lambda}\bar{\psi}_A),$$

for all λ. The Bochner formula now gives

$$0 = \|d_A(\lambda\psi_A + \bar{\lambda}\bar{\psi}_A)\|^2 + \|d_A^*(\lambda\psi_A + \bar{\lambda}\bar{\psi}_A)\|^2 = \|\nabla_A(\lambda\psi_A + \bar{\lambda}\bar{\psi}_A)\|^2.$$

Here we have used the vanishing of the Ricci curvature on Calabi–Yau manifolds and (6.9) and (6.19) to eliminate the curvature terms from the Bochner formula. Hence

$$\nabla_A\psi_A = 0. \tag{6.20}$$

We conclude this discussion with the following theorem.

Theorem 6.21. *Let A be a G-connection on a bundle, E, on a complete Calabi–Yau 3-fold, M. Assume A is a stable critical point of E'. If M is noncompact, assume also that $F_A^{0,2} \in L_4$. Then $F_A - F_A^{1,1}$ is covariant constant and takes values in an abelian subbundle of $ad(E) \otimes \mathbb{C}$. If $Hol(M) = SU(3)$, then $(E, \bar{\partial}_A)$ is holomorphic.*

Proof. Let $R_{ij}dx^i \wedge dx^j$ denote the Riemann curvature tensor viewed as an $ad(T^*M)$ valued 2−form. The vanishing of $\nabla_A\psi$, (6.20) implies $0 = [\nabla_i, \nabla_j]\psi_A = (ad(F_{ij}) + R_{ij})\psi_A$ for all i, j. Because ψ_A takes values in an abelian subalgebra of $ad(E)$, $[F_{ij}, \psi_A] \perp R_{ij}\psi_A$. Hence $R_{ij}\psi_A = 0$, and the components of ψ_A are in the kernel of the Riemann curvature operator. This reduces the Riemannian holonomy group, unless $\psi_A = 0$, which implies $F_A^{0,2} = 0$. Recall that $(E, \bar{\partial}_A)$ determines a holomorphic structure if and only if $F_A^{0,2} = 0$. Thus we have the dichotomy: $\psi_A \neq 0$

implies a reduction of the holonomy of M, and $\psi_A = 0$ implies $(E, \bar{\partial}_A)$ is holomorphic. q.e.d.

References

[Bo] A. Borel, *Topology of Lie groups and characteristic classes*, Bull. Amer. Math. Soc **61**, (1955), 397–432, MR 0072426, Zbl 0066.02002.

[BL] J-P. Bourguignon & B. Lawson, *Stability and isolation phenomena for Yang-Mills fields*, Commun. Math. Phys. **79** (1981), no. 2, 189–230, MR 0612248, Zbl 0475.53060.

[BLS] J-P. Bourguignon, B. Lawson & J. Simons, *Stability and gap phenomena for Yang-Mills fields*, Proc. Nat. Acad. Sci. USA **76** (1979), no. 4, 1550–1553, MR 0526178, Zbl 0408.53023.

[D] S. Donaldson, *Anti self-dual Yang–Mills connections over complex algebraic surfaces and stable vector bundles*, Proc. London Math. Soc, **50** no. 3 (1985), 1–26, MR 0765366, Zbl 0529.53018.

[Ha] R. Hamilton, *Three-manifolds with positive Ricci curvature*, J. Diff. Geom **17**, no. 2, (1982), 255–306, MR 0664497, Zbl 0504.53034.

[He] E. Hebey, *Nonlinear Analysis on Manifolds: Sobolev Spaces and Inequalities*, Courant Lecture Notes in Mathematics **5**, American Mathematical Society, Providence, 2000, MR 1688256, Zbl 0981.58006.

[JT] A. Jaffe & C. Taubes, *Vortices and Monopoles*, Progress in Physics, 2. Birkhauser, Basel, 1980, MR 0614447, Zbl 0457.53034.

[K] S. Kobayashi, *Differential Geometry of Complex Vector Bundles*, Princeton University Press, Princeton, 1987, MR 0909698, Zbl 0708.53002.

[KOT] S. Kobayashi, Y. Ohnita & M. Takeuchi, *On Instability of Yang-Mills Connections*, Math. Zeit. **193** (1986), no. 2, 165–189, MR 0856147, Zbl 0634.53022.

[LT] M. Lubke & A. Teleman, *The Universal Kobayashi–Hitchin Correspondence on Hermitian Manifolds*, Memoirs of the AMS, **863**, American Mathematical Society, Providence, 2006, MR 2254074, Zbl 1103.53014.

[M] C.B. Morrey, *On the analyticity of the solutions of non-linear elliptic systems of partial differential equations: Part I. Analyticity in the interior*, American J. Math. **80**, (1958), no. 2, 198–218, MR 0106336, Zbl 0081.09402.

[N] H. Nakajima, *Compactness of the moduli space of Yang-Mills connections in higher dimensions*, J. Math. Soc. Japan **40** (1988), no. 3, 383–392, MR 0945342, Zbl 0647.53030.

[OP] Y. Ohnita & Y. Pan, *On weakly stable Yang-Mills fields over positively pinched manifolds and certain symmetric spaces*, Kodai Math. J. **13** (1990), no. 3, 317–332, MR 1078547, Zbl 0761.53018.

[P] T. Parker, *Conformal Fields and Stability*, Math. Zeit. **185** (1984), 305–319, MR 0731679, Zbl 0513.53007.

[SS] L. Sadun & J. Segert, *Non-self-dual Yang–Mills connections with nonzero Chern number*, Bulletin of the AMS **24**, (1991), no. 1, 163–170, MR 1067574, Zbl 0722.58014.

[Sh] C.L. Shen, *Weak Stability of Yang-Mills Fields over submanifolds of the sphere*, Arch. Math. **39** (1982), no. 1, 78–84, MR 0674536, Zbl 0479.53041.

[SSU] L. Sibner, R. Sibner & K. Uhlenbeck, *Solutions to Yang–Mills equations that are not self-dual*, Proc. Nat. Acad. Sci. **86**, (1989), 8610–8613, MR 1023811, Zbl 0731.53031.

[U1] K. Uhlenbeck, *Connections with L^p bounds on curvature*, Comm. Math. Phys. **83** (1982), no. 1, 31–42, MR 0648356, Zbl 0499.58019.

[UY] K. Uhlenbeck & S. T. Yau, *On the Existence of Hermitian-Yang-Mills Connections in Stable Vector Bundles*, Comm. Pure and Appl. Math. **39** (1986), 257–293, MR 0861491, Zbl 0615.58045.

[W] K. Wehreim, *Uhlenbeck Compactness*, European Mathematical Society, Zurich, 2004, MR 2030823, Zbl 1055.53027.

[We] R. Wells, *Differential Analysis on Complex Manifolds*, (Third ed.) Graduate Texts in Math. 65, Springer, New York, Berlin, 2008, MR 2359489, Zbl 1131.32001.

[X] Y. Xin, *Instability theorems of Yang-Mills fields*, Acta Math. Sci. **3** (1983), no. 1, 103–112, MR 0741362, Zbl 0543.58018.

Mathematics Department
Duke University, Box 90320
Durham, NC 27708-0320
E-mail address: stern@math.duke.edu