

# THE TOPOLOGY OF THE SPACE OF STABLE BUNDLES ON A COMPACT RIEMANN SURFACE

GEORGIOS D. DASKALOPOULOS

## Abstract

We develop a singular Morse theory for the Yang-Mills functional on the space of holomorphic structures on a bundle over a compact Riemann surface. We also examine the relation with the algebraic methods.

## 1. Introduction

The cohomology of the moduli space of semistable bundles over a curve was computed by algebraic methods by Harder and Narasimhan [13] and by transcendental methods by Atiyah and Bott [1]. The main idea behind the second approach was to employ techniques from differential and symplectic geometry. Roughly speaking, the approach of Atiyah and Bott is based on the observation that for many moduli problems, including that of holomorphic bundles, the algebraic geometric notion of stability is related to the Morse theory of certain associated functionals.

F. Kirwan pursued this observation further, by examining a very broad class of group actions on nonsingular projective varieties and symplectic manifolds, and obtained inductive formulas for computing the equivariant cohomology of the set of semistable points of these actions ([16], [17]). She also indicated explicitly how Morse theory is related to the ideas of geometric invariant theory and geometric quantization ([22], [11]). Results in the same direction were also obtained independently by L. Ness [24].

The situation studied by Atiyah and Bott differs from Kirwan's mainly because they are dealing with an infinite-dimensional problem. The space of holomorphic structures on a fixed smooth bundle  $E$  over a Riemann surface  $M$  is an infinite-dimensional manifold  $\mathfrak{B}$ . The group  $\mathfrak{g}^{\mathbb{C}}$  of complex gauge transformations is an infinite-dimensional Lie group acting on  $\mathfrak{B}$  and  $[\mathfrak{B}] = \mathfrak{B}/\mathfrak{g}^{\mathbb{C}}$  parametrizes the space of holomorphic bundles over  $M$  of fixed topological type. Atiyah and Bott define a stratification  $\{\mathfrak{B}_{\mu}\}$

of  $\mathfrak{B}$  by  $\mathfrak{g}^{\mathbb{C}}$ -invariant subspaces corresponding to the type of filtration of the associated holomorphic bundle, and they proceed algebraically to show that these spaces form locally closed submanifolds. They use the stratification directly to compute cohomology, leaving the connection with Morse theory only at a conjectural level.

The main goal of this paper is to justify the Atiyah-Bott optimism about the existence of a Morse Theory for the Yang-Mills functional and, furthermore, to examine its relationship with the algebraic methods. Donaldson took the first step in this project by reproving the theorem of Narasimhan and Seshadri using analytic methods [5]. More specifically, Donaldson showed that every stable complex orbit contains a unique up to real gauge equivalence Hermitian-Einstein connection. However, as it leaves out, there is nothing special about stable orbits. We extend the analysis showing the real convergence of all complex orbits. In particular, we show that the gradient flow of the Yang-Mills functional converges at infinity. This allows us to define a Morse theory for the Yang-Mills functional and obtain a stratification of the space  $\mathfrak{B}/\mathfrak{g}$  of holomorphic structures modulo real gauge equivalence into certain subspaces having only orbifold singularities. Moreover, we show that the stratification above coincides with that of Harder and Narasimhan. In other words, our analytic cells coincide with the algebraic ones.

Now, a few words about the Morse theory that we are using: it is not difficult to verify that the Yang-Mills functional satisfies an equivariant Palais-Smale Condition  $C$  on  $\mathfrak{B}$  (cf. (4.1)). This provides us with all the necessary compactness for the flow to converge upon passing to the quotient. We exploit the flow directly to build up our space  $\mathfrak{B}/\mathfrak{g}$ . This is in some sense closer to the spirit of Smale [31], and avoids the handle attaching techniques of Palais and Bott, which are better suited to less singular situations ([26], [2]).

More explicitly, the material in the paper is organized as follows: §2 contains some preparatory definitions and lemmas needed in later sections. The main result of the section is the following decomposition theorem of the complex gauge group corresponding to the standard Iwasawa decomposition of the general linear group.

**Theorem A.** *With the notation as in §2, we have the homeomorphism*

$$\mathfrak{g}^{\mathbb{C}} \simeq \mathfrak{g}_{\star}^{\mathbb{C}} \times_{\mathfrak{g}_{\text{DIAG}}} \mathfrak{g}.$$

This theorem will be essential in §§3 and 5 when we define retractions of our strata onto the Yang-Mills connections.

In §3, we present a different proof of the following theorem due to Atiyah and Bott.

**Theorem B.** *The space of holomorphic structures of type  $\mu$  is a locally closed submanifold of the space of holomorphic structures. Furthermore, with the notation as in §2, we have the following homeomorphisms:*

$$\mathfrak{B}_\mu \simeq \mathfrak{g}^{\mathbb{C}} \times_{\mathfrak{g}^{\mathbb{C}}} \mathfrak{B}_*^{ss}, \quad \mathfrak{B}_\mu \simeq \mathfrak{g} \times_{\mathfrak{g}_{\text{DIAG}}} \mathfrak{B}_*^{ss}.$$

The formulas of Atiyah and Bott about equivariant cohomology follow easily from the above theorem and its corollaries.

In §4, we establish the link with Morse theory. As explained in [1] the Yang-Mills functional  $f: \mathfrak{B} \rightarrow \mathbb{R}$  defines a gradient flow which preserves the stratification  $\{\mathfrak{B}_\mu\}$ . We verify Condition *C* for  $f$  and define a stable-unstable manifold stratification in the sense of Morse theory via  $f$ .

In §5, we are concerned with some technical questions about the flow of  $f$ . Our main theorem is the following conjecture due to Atiyah and Bott.

**Theorem C.** *For each  $\mu$ , the gradient flow of the Yang-Mills functional defines a continuous deformation retract of the stratum  $\mathfrak{B}_\mu$  onto the critical set.*

As an immediate consequence, we obtain the following generalization of the uniqueness result of Donaldson in [5].

**Corollary.** *The closure of each complex gauge orbit contains exactly one minimizing real gauge orbit. In particular, the flow converges at infinite time.*

In §6 we combine the results of §§4 and 5 to give a proof of the following conjecture due to Atiyah and Bott.

**Theorem D.** *The Morse stratification of the Yang-Mills functional coincides with the stratification of Harder and Narasimhan.*

Finally, in §7 we indicate how one can use the techniques developed in the previous sections to produce new information about the moduli space of stable bundles. We specialize for the sake of concreteness to the case of bundles of rank 2 and take the Chern class to be 0, since the Chern class 1 case has already been treated by Atiyah and Bott. The key result of the section is the following “cell decomposition theorem” of the semistable stratum.

**Theorem E.** *The complement of the stable stratum  $\mathfrak{B}_s$  in the semistable stratum  $\mathfrak{B}_{ss}$  can be stratified by locally closed submanifolds of codimension at least  $2g - 2$ .*

Theorem E, combined with Theorem B and the Riemann Roch Theorem, implies that the inclusion of the stable stratum in  $\mathfrak{B}$  is a homotopy equivalence up to dimension  $2g - 3$ . This, combined with known results about the topology of the gauge group, allows us to compute certain homotopy and cohomology groups of the moduli space of stable bundles

$\mathcal{M} = \mathfrak{B}_s/\mathfrak{g}^{\mathbb{C}}$ . For example we can show, provided the genus  $g$  of the base Riemann surface  $M$  is at least 3, that:

$$\pi_1(\mathcal{M}) \simeq H_1(M, \mathbb{Z}), \quad \pi_2(\mathcal{M}) \simeq \mathbb{Z} \oplus \mathbb{Z}_2.$$

Generalizations of the above isomorphisms for arbitrary rank as well as an interpretation of the torsion class in  $\pi_2(\mathcal{M})$  will be given in a subsequent joint paper with K. Uhlenbeck [4].

We now discuss briefly the arguments involved in the proofs of our main Theorems B and C. Both proofs essentially amount to proving a slice theorem in a neighborhood of the critical set. More precisely, in Theorem B we start with a critical holomorphic structure  $\bar{\partial}$  of type  $\mu$  and define a projection  $q$  of the tangent space  $\mathfrak{B}$  at  $\bar{\partial}$  onto the subspace of  $H^1(M, \text{End } E)$  perpendicular to the space of 1-forms preserving the Harder-Harasimhan filtration associated to  $\bar{\partial}$ . We proceed to show that the stratum  $\mathfrak{B}_\mu$  is locally cut out by the equation  $q = 0$ . Indeed, if  $D = \bar{\partial} + A$ , with  $A$  sufficiently small and  $q(A) \neq 0$ , then we are able to construct a one-parameter subgroup of complex gauge transformations  $\{g_t\}_{t \geq 0}$  such that the energy of  $g_t(D)$  is less than the minimum allowable energy in the stratum  $\mathfrak{B}_\mu$ . This is of course a contradiction, since the stratum  $\mathfrak{B}_\mu$  is complex gauge invariant. The above establishes the manifold structure of  $\mathfrak{B}_\mu$ . The explicit homeomorphism of Theorem B is constructed quite naturally via the analysis in §2.

The proof of Theorem C can be outlined as follows: Our main step is to show that the gradient flow of the Yang-Mills functional converges to a unique critical point modulo real gauge equivalence. By Theorem B, it is enough to prove the convergence of the flow for the semistable stratum, the stable case having been established by Donaldson. By approximating semistable holomorphic structures by stable ones, we reduce the problem to a local analysis near the critical set. The final step is established by a slice theorem at the flat bundles, similar in spirit to the one used in Theorem B.

Along the same lines one could ask if the methods of this paper could be extended to other infinite-dimensional moment map problems, e.g., the one encountered by Hitchin [14]. It seems to us that being in the right range for the Sobolev theorems to work, the main analytical tools can be carried through, and it becomes plausible that a similar approach could be pursued. The study of the topology of the representation space of  $\pi_1 M$  into  $SL(n, \mathbb{C})$ , which is related to the Hitchin problem, is a very interesting problem.

Duistermaat recently established the convergence of the flow of the norm square of the moment map associated to the group actions studied by Kirwan. His methods depend upon the analyticity properties of the flow, and they are not directly applicable to the infinite-dimensional setting; one would hope though to understand his techniques in relation to ours.

Finally, we would like to comment that there are very close links between our Morse theory approach and the infinite-dimensional geometric invariant theory as developed in [6], [7], [30], [34]. Our one-parameter subgroups used in the proof of Proposition (3.5) correspond to the optimal one-parameter subgroups of Kempf [15], related to a Hilbert-Mumford numerical function. Several of our results mentioned above could also be seen in this setting. Due to the length of this paper these questions will be treated elsewhere.

### 2. Definitions and basic lemmas

Throughout the paper we fix a  $C^\infty$  vector bundle  $E$  over a compact Riemann surface  $M$  of genus  $g \geq 2$  with rank  $n$  and first Chern class  $k$ . For most parts of the paper we also fix a Hermitian metric  $K$  on  $E$ .

Let  $F$  be a subbundle of  $E$ ; we define the *normalized Chern class* of  $F$  by

$$\mu(F) = \frac{c_1(F)}{\text{rk}(F)},$$

where  $c_1(F)$  denotes the first Chern class of  $F$  and  $\text{rk}(F)$  the rank of  $F$ .

(2.1) **Definition.** A *Harder-Narasimhan filtration* of  $E$  is a finite sequence of  $C^\infty$  subbundles

$$(*) \quad 0 = E_0 \subset E_1 \subset \dots \subset E_r = E$$

such that  $\mu(F_1) > \dots > \mu(F_r)$ . Here  $F_i$  denotes the quotient

$$F_i = E_i/E_{i-1}, \quad i = 1, \dots, r,$$

which is also  $C^\infty$  isomorphic to the orthogonal complement of  $E_{i-1}$  in  $E_i$ . Consider the vector  $\mu = (\mu_1, \dots, \mu_n)$ , whose  $n_1 := \text{rk}(E_1)$  first components are equal to  $\mu(F_1)$ , the next  $n_2 := \text{rk}(E_2)$  are equal to  $\mu(F_2)$ , etc.;  $\mu$  is called the *type of the filtration*.

(2.2) **Definition.** Let  $GL(E)$  be the set of bundle isomorphisms of  $E$ , and let  $U(E)$  be the set of isomorphisms which preserve the metric. Let  $\text{End}(E)$  be the set of bundle endomorphisms, and  $u(E)$  be the set of

skew adjoint ones. Also, given a filtration  $(*)$ , let  $UT(E, *)$  be the set of bundle endomorphisms preserving  $(*)$  and let  $IUT(E, *)$  be the set of bundle isomorphisms that preserve  $(*)$ . Clearly  $\text{End}(E)$ ,  $u(E)$ , and  $UT(E, *)$  are vector bundles, whereas  $GL(E)$ ,  $U(E)$ , and  $IUT(E, *)$  are only fiber bundles.

We recall that the space  $\mathfrak{A}$  of smooth connections on  $E$ , compatible with the given metric, is an affine space corresponding to the vector space  $C^\infty(T^*M \otimes u(E))$ . Similarly the space  $\mathfrak{B}$  of smooth holomorphic structures on  $E$  is an affine space with corresponding vector space  $C^\infty(T^*M'' \otimes \text{End } E)$ , where  $T^*M''$  denotes the space of  $(0, 1)$  forms on  $M$ . For the purpose of this paper it is technically more correct to work with Sobolev completions of the above spaces. More specifically, we introduce the spaces of  $L^2_1$  connections and  $L^2_1$  holomorphic structures as in [1, §14]. They are both affine spaces corresponding to the vector spaces  $L^2_1(T^*M \otimes u(E))$  and  $L^2_1(T^*M'' \otimes \text{End } E)$  of  $L^2_1$  sections of the bundles  $T^*M \otimes u(E)$  and  $T^*M'' \otimes \text{End } E$  respectively.

By associating to a given holomorphic structure, and the given metric, the unique compatible connection, we obtain [1, §5] an identification  $j: \mathfrak{B} \rightarrow \mathfrak{A}$ . This extends naturally to an isomorphism

$$(2.3) \quad j: \mathfrak{B}_1^2 \rightarrow \mathfrak{A}_1^2.$$

(2.4) **Definition.** We define the gauge groups  $\mathfrak{g}^2$ ,  $\mathfrak{g}^{\mathbb{C}^2}$ ,  $\mathfrak{g}_*^{\mathbb{C}^2}$ ,  $\mathfrak{g}_{\text{DIAG}}^2$ , and  $\mathfrak{g}_{\text{DIAG}}^{\mathbb{C}^2}$  to be the groups of  $L^2_2$  sections of the fiber bundles  $U(E)$ ,  $GL(E)$ ,  $IUT(E, *)$ ,  $\prod_{i=1}^r U(F_i)$ , and  $\prod_{i=1}^r GL(F_i)$  respectively.

According to [27, §5], the above groups are smooth Hilbert Lie groups and the inclusions

$$\begin{aligned} \prod_{i=1}^r U(F_i) &\hookrightarrow \prod_{i=1}^r GL(F_i) \hookrightarrow IUT(E, *) \hookrightarrow GL(E), \\ \prod_{i=1}^r U(F_i) &\hookrightarrow U(E) \hookrightarrow GL(E) \end{aligned}$$

induce at the sections level inclusions of closed embedded subgroups. We shall also deal with the corresponding groups  $\mathfrak{g}$ ,  $\mathfrak{g}^{\mathbb{C}}$ ,  $\mathfrak{g}_*^{\mathbb{C}}$ ,  $\mathfrak{g}_{\text{DIAG}}^{\mathbb{C}}$ , and  $\mathfrak{g}_{\text{DIAG}}^{\mathbb{C}}$  of smooth sections of the bundles given in (2.4).

One could define the more general Banach manifold  $\mathfrak{A}_1^p, \mathfrak{g}^p, \dots$ , provided  $p > 1$ . These spaces can also be used for most parts of the paper. The gauge groups act on the spaces  $\mathfrak{B}$  and  $\mathfrak{B}_1^2$  by push-forward. More specifically, we define

$$\beta: \mathfrak{g}^{\mathbb{C}^2} \times \mathfrak{B}_1^2 \rightarrow \mathfrak{B}_1^2$$

by  $\beta(g, D'') = g(D'') = g \circ D'' \circ g^{-1}$ . The above action is smooth and restricts to smooth actions of the other groups as well. We can also define an action of  $\mathfrak{g}^{C^2}$  on  $\mathfrak{A}_1^2$ ,

$$\alpha: \mathfrak{g}^{C^2} \times \mathfrak{A}_1^2 \rightarrow \mathfrak{A}_1^2,$$

as follows: Let  $D \in \mathfrak{A}_1^2$ , where  $D = j(D'')$ , and let  $g \in \mathfrak{g}^{C^2}$ . We set  $\alpha(g, D) = g(D) := g \circ D'' \circ g^{-1} + (g \circ D'' \circ g^{-1})^*$ , where  $*$  denotes the fiberwise adjoint with respect to the Hermitian metric on  $E$ . Again  $\alpha$  is a smooth map and restricts to smooth actions of the other groups as well. For more details we refer to [6]. These actions generalize the actions of the unitary groups imposed by geometrical considerations. We note that the map (2.3) is an equivariant map for the above actions.

(2.5) **Definition.** A  $C^\infty$  holomorphic structure  $D''$  on  $E$  is called compatible with the filtration  $(*)$ , or preserving  $(*)$ , if for all  $s$  in the sheaf of smooth sections of  $E$  that lie in  $E_i$ ,  $D''s$  defines a section of  $E_i \otimes T^*M''$ . Clearly,  $D''$  induces holomorphic structures on  $E_i$  and the inclusions  $E_{i-1} \subset E_i$ ,  $i = 1, \dots, r$ , become holomorphic inclusions.

It is easy to see that if  $D''_0$  and  $D''_1$  are holomorphic structures preserving  $(*)$ , then  $D''_0 - D''_1 \in C^\infty(T^*M'' \otimes UT(E, *))$ . Therefore, the set of  $(*)$  preserving smooth holomorphic structures, is the affine space

$$\mathfrak{B}_* = D''_0 + C^\infty(T^*M'' \otimes UT(E, *)).$$

Clearly  $\mathfrak{B}_*$  is preserved by  $\mathfrak{g}_*^C$ . As usual, we work with the  $L^2_1$  completion:

$$\begin{aligned} \mathfrak{B}_{*1}^2 &:= D''_0 + L^2_1(T^*M'' \otimes UT(E, *)) \\ (2.6) \quad &= D''_0 + \prod_{i \geq j} L^2_1(T^*M'' \otimes \text{Hom}(F_i, F_j)). \end{aligned}$$

By an obvious approximation argument, it can be shown that  $\mathfrak{B}_{*1}^2$  is preserved by  $\mathfrak{g}_*^{C^2}$ .

Since any  $\mathfrak{g}_*^{C^2}$  orbit in  $\mathfrak{B}_{*1}^2$  contains a  $C^\infty$  holomorphic structure [1, §14], we have the following:

(2.7) **Lemma.** *With the notation as above,  $\mathfrak{B}_{*1}^2 = \mathfrak{g}_*^{C^2} \cdot \mathfrak{B}_*$ .*

(2.8) **Definition.** A holomorphic structure on  $E$  is said to be *semi-stable (stable)* if for any proper nontrivial holomorphic subbundle  $F \subset E$ ,

$$\mu(F) \leq \mu(E) \quad (\mu(F) < \mu(E)).$$

Elementary arguments [13, Proposition 1.3.9] show that every holomorphic structure on  $E$  has a unique *holomorphic filtration*

$$(*) \quad 0 = E_0 \subset E_1 \subset \dots \subset E_r = E,$$

where  $F_i := E_i/E_{i-1}$  ( $i = 1, \dots, r$ ) is semistable and  $\mu(F_i) > \mu(F_{i+1})$  ( $i = 1, \dots, r-1$ ). The type  $\mu$  of  $(*)$  is called the *type of the holomorphic bundle  $E$* . For each  $\mu$ , let  $\mathfrak{B}_\mu \subset \mathfrak{B}$  be the set of holomorphic structures  $D''$  preserving some filtration in the sense of Harder and Narasimhan. We denote by  $\mathfrak{B}_{ss}$  the semistable stratum. The existence and uniqueness of the filtration implies that  $\{\mathfrak{B}_\mu\}$  forms a partition of  $\mathfrak{B}$ , whereas the naturality of the filtration implies that the sets  $\mathfrak{B}_\mu$  are preserved by  $\mathfrak{g}^C$ . Define

$$\mathfrak{B}_{\mu 1}^2 := \mathfrak{g}^{C2} \cdot \mathfrak{B}_\mu.$$

Clearly,  $\mathfrak{B}_{\mu 1}^2$  is preserved by  $\mathfrak{g}^{C2}$ , and, since any  $\mathfrak{g}^{C2}$  orbit contains a smooth holomorphic structure, we have:

(2.9) **Lemma.**  $\{\mathfrak{B}_{\mu 1}^2\}$  is a partition of  $\mathfrak{B}_1^2$ .

Given a filtration  $(*)$  as in (2.1), let  $\mathfrak{B}_*^{ss}$  denote the set of holomorphic structures  $D'' \in \mathfrak{B}_*$  such that  $F_i$  is semistable for any  $i = 1, \dots, r$ . We set

$$\mathfrak{B}_{*1}^{ss2} := \mathfrak{g}_*^{C2} \cdot \mathfrak{B}_*^{ss}.$$

(2.10) **Lemma.** If  $(*)$  is a filtration of type  $\mu$ , then:

- (i)  $\mathfrak{B}_\mu = \mathfrak{g}^C \cdot \mathfrak{B}_*^{ss}$ ,
- (ii)  $\mathfrak{B}_{\mu 1}^2 = \mathfrak{g}^{C2} \mathfrak{B}_*^{ss} = \mathfrak{g}^{C2} \mathfrak{B}_{*1}^{ss2}$ .

*Proof.* (i) Clearly  $\mathfrak{g}^C \cdot \mathfrak{B}_*^{ss} \subset \mathfrak{B}_\mu$ . For the other direction let  $D'' \in \mathfrak{B}_\mu$ . The holomorphic structure  $D''$  preserves some filtration isomorphic to  $(*)$  by some element of  $\mathfrak{g}^C$ , hence  $D'' \in \mathfrak{g}^C \cdot \mathfrak{B}_*^{ss}$ .

(ii) We have

$$\begin{aligned} \mathfrak{B}_{\mu 1}^2 &= \mathfrak{g}^{C2} \cdot \mathfrak{B}_\mu = \mathfrak{g}^{C2} \cdot \mathfrak{g}^{C2} \cdot \mathfrak{B}_*^{ss} = \mathfrak{g}^{C2} \cdot \mathfrak{B}_*^{ss}, \\ \mathfrak{g}^{C2} \cdot \mathfrak{B}_{*1}^{ss2} &= \mathfrak{g}^{C2} \cdot \mathfrak{g}_*^{C2} \cdot \mathfrak{B}_*^{ss} = \mathfrak{g}^{C2} \cdot \mathfrak{B}_*^{ss}. \quad \text{q.e.d.} \end{aligned}$$

We now need to introduce the Yang-Mills equations. Let  $f$  be any smooth function on the Lie algebra  $u(n)$  of the unitary group  $U(n)$  which is invariant under the adjoint action and is convex [1, §8]. Given such an  $f$ , we define a functional  $f$  on the space of metric connections in the obvious way:

$$f(D) = \int_M f(*F(D)) * 1,$$

where  $F(D)$  denotes the curvature of the connection  $D$ . We simply normalize so that the volume of  $M$  is 1. In this case  $f(x) = \text{tr } x^* x$ , the above functional is called the *Yang-Mills functional*. It is quite easy to verify [8, p. 198] that the Yang-Mills functional defines a smooth map  $f: \mathfrak{A}_1^2 \rightarrow \mathbb{R}$ .

The critical points of  $f$  are called the *Yang-Mills connections* and they are described by the equation

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

Here  $D^*$  is the  $L^2$ -adjoint of  $D$  and, just as in usual Hodge theory, it is given by  $\pm * D*$ , where  $*$  is the standard Hodge  $*$ -operator on  $M$  [9, Chapter 0, §6]. We denote by  $\mathfrak{N}$  the set of Yang-Mills connections. An elementary argument [1, §5] shows that, if  $D \in \mathfrak{N}$ , then the eigenvalues of  $\frac{\sqrt{-1}}{2\pi} * F(D)$  are constant, say  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ . Let  $\mu := (\lambda_1, \dots, \lambda_n)$  and let  $\mathfrak{N}_\mu$  be the set of critical points  $D$  such that  $\frac{\sqrt{-1}}{2\pi} * F(D)$  has eigenvalues  $\lambda_1, \dots, \lambda_n$ . Note that  $\mathfrak{N} = \bigcup_\mu \mathfrak{N}_\mu$  and  $j^{-1}(\mathfrak{N}_\mu) \subset \mathfrak{B}_\mu$ . Furthermore, let  $f$  be a convex, invariant function on  $u(n)$ , as before. Assume also that the second derivative  $f''$  of  $f$  is everywhere positive. Then, it is quite easy to show that the critical points of the associated functional are exactly the Yang-Mills connections [1, (8.7)]. Moreover, if  $D \in \mathfrak{N}_\mu$ ,  $\mu = (\lambda_1, \dots, \lambda_n)$ , let  $\Lambda_\mu$  be the diagonal matrix whose nonzero entries are equal to  $-2\pi\sqrt{-1}\lambda_j$ ,  $j = 1, \dots, n$ . We define  $f(\mu) := f(\Lambda_\mu)$ . It follows immediately from the definition that  $f(D) = f(\mu)$ . The usefulness of the above notation will be clarified in Proposition (2.11).

We recall that on the set of all possible  $\mu = (\mu_1, \dots, \mu_n)$  appearing as types of Harder-Narasimhan filtrations as in (2.1), we define a partial order  $\leq$  as follows:

$$\lambda \leq \mu \quad \text{if} \quad \sum_{j \leq i} \lambda_j \leq \sum_{j \leq i} \mu_j, \quad i = 1, \dots, n - 1.$$

For more details we refer to [1, §§1, 7].

We now come to the following ‘‘Sobolev’’ version of the result in [1, §8]:

(2.11) **Proposition.** *Let  $\mathfrak{A}_{\mu_1}^2 = j(\mathfrak{B}_{\mu_1}^2) \subset \mathfrak{A}_1^2$ . Then the following hold.*

- (i) *For any  $D \in \mathfrak{A}_{\mu_1}^2$  and any convex invariant function  $f$  as before,  $f(D) \geq f(\mu)$ .*
- (ii)  *$D \in \mathfrak{A}_{\mu_1}^2$  iff  $\inf_{g \in \mathfrak{g}^{\mathbb{C}^2}} f(g(D)) = f(\mu)$ .*

*Proof.* The result is known if we replace  $\mathfrak{A}_{\mu_1}^2$ ,  $\mathfrak{B}_{\mu_1}^2$  by  $\mathfrak{A}_\mu$ ,  $\mathfrak{B}_\mu$ , and  $\mathfrak{g}^{\mathbb{C}^2}$  by  $\mathfrak{g}^{\mathbb{C}}$ . The rest follows by an obvious approximation argument by smooth elements. q.e.d.

As a corollary we obtain the following ‘‘Sobolev version’’ of the result of Shatz. For more details see [1, (7.8)] and [29].

(2.12) **Proposition.** *Under the partial ordering on the stratification defined in [1], we obtain*

$$\overline{\mathfrak{B}_{\lambda 1}^2} \subset \bigcup_{\mu \geq \lambda} \mathfrak{B}_{\mu 1}^2.$$

In particular,  $\mathfrak{B}_{ss1}^2$  is open.

The proof is the same as in [1] and is omitted.

By writing elements of  $\mathfrak{B}_{*1}^{ss2}$  as appropriate extensions of semistable structures of lower rank, using induction and (2.12) it can be shown that  $\mathfrak{B}_{*1}^{ss2}$  is an open subset of  $\mathfrak{B}_{*1}^2$ .

In the following spaces we divide out by the obvious diagonal group actions. The same argument as in [8, Theorem 3.2] shows that they can be given the structure of a smooth manifold. Moreover the natural quotient maps are smooth submersions:

- (1)  $\mathfrak{g}^2 \times_{\mathfrak{g}_{\text{DIAG}}^2} \mathfrak{g}_*^{\mathbb{C}2}$ .
- (2)  $\mathfrak{g}^{\mathbb{C}2} \times_{\mathfrak{g}_*^{\mathbb{C}2}} \mathfrak{B}_{*1}^2$  and  $\mathfrak{g}^2 \times_{\mathfrak{g}_{\text{DIAG}}^2} \mathfrak{B}_{*1}^2$ .
- (3)  $\mathfrak{g}^{\mathbb{C}2} \times_{\mathfrak{g}_*^{\mathbb{C}2}} \mathfrak{B}_{*1}^{ss2}$  and  $\mathfrak{g}^2 \times_{\mathfrak{g}_{\text{DIAG}}^2} \mathfrak{B}_{*1}^{ss2}$ .
- (4) Since  $\mathfrak{B}_{*1}^{ss2} \subset \mathfrak{B}_{*1}^2$  is open, it follows that  $\mathfrak{g}^{\mathbb{C}2} \times_{\mathfrak{g}_*^{\mathbb{C}2}} \mathfrak{B}_{*1}^{ss2}$  is an open submanifold of  $\mathfrak{g}^{\mathbb{C}2} \times_{\mathfrak{g}_*^{\mathbb{C}2}} \mathfrak{B}_{*1}^2$ . Similarly for  $\mathfrak{g}^2 \times_{\mathfrak{g}_{\text{DIAG}}^2} \mathfrak{B}_{*1}^{ss2}$  in  $\mathfrak{g}^2 \times_{\mathfrak{g}_{\text{DIAG}}^2} \mathfrak{B}_{*1}^2$ .

We now return to our gauge groups. Let  $A$  belong to  $GL(n)$ , the group of invertible  $n \times n$  matrices. Then we can write uniquely  $A = B \cdot U$ , where  $B$  is upper triangular with positive diagonal elements and  $U$  belongs to  $U(n)$ , the group of unitary  $n \times n$  matrices. This follows directly from Iwasawa decomposition. More generally, let

$$(*) \quad 0 \subset \mathbb{C}^{n_1} \subset \mathbb{C}^{n_1+n_2} \subset \dots \subset \mathbb{C}^{n_1+\dots+n_r} = \mathbb{C}^n$$

be a filtration of subspaces of  $\mathbb{C}^n$ . It follows from the above result that any  $A \in GL(n)$  can be written as  $A = B \cdot U$ , where  $B$  belongs to  $IUT(\mathbb{C}^n, *)$ , the set of  $(*)$ -preserving  $n \times n$  matrices and  $U \in U(n)$ . Moreover, it is not hard to see that such a decomposition is unique up to  $(*)$ -diagonal matrices, i.e., matrices that preserve the splitting  $\mathbb{C}^{n_1} \oplus \dots \oplus \mathbb{C}^{n_r}$  of  $\mathbb{C}^n$ .

We define a map

$$\varphi : IUT(\mathbb{C}^n, *) \times U(n) \rightarrow GL(n) \times U(n_1) \times \dots \times U(n_r)$$

as follows: Let  $(B, U) \in IUT(\mathbb{C}^n, *) \times U(n)$ , where

$$B = \begin{pmatrix} A_1 & \dots & * \\ & \ddots & \vdots \\ 0 & & A_r \end{pmatrix},$$

with  $A_i \in GL(n_i)$ ,  $i = 1, \dots, r$ . Write  $A_i = B_i U_i$ , where  $B_i$  is upper triangular with positive diagonal elements and  $U_i \in U(n_i)$ , and let

$$D = \begin{pmatrix} U_1 & & 0 \\ & \ddots & \\ 0 & & U_r \end{pmatrix} \in U(n_1) \times \dots \times U(n_r).$$

We set  $\varphi(B, U) = (B \cdot U, D)$ .

(2.13) **Lemma.**  $\varphi$  is an isomorphism.

*Proof.* Assume  $\varphi(B, U) = \varphi(\tilde{B}, \tilde{U})$ . Then  $B U = \tilde{B} \tilde{U}$  and  $D = \tilde{D}$ , where  $B, \tilde{B}, U, \tilde{U}, D$ , and  $\tilde{D}$  are the obvious matrices. Thus  $D := \tilde{B}^{-1} B = \tilde{U} U^{-1}$  is  $(*)$ -upper triangular and unitary, hence  $(*)$ -diagonal. Therefore

$$C = \begin{pmatrix} \tilde{A}_1^{-1} A_1 & & 0 \\ & \ddots & \\ 0 & & \tilde{A}_r^{-1} A_r \end{pmatrix} \in \prod_{i=1}^r U(n_i).$$

Since  $\tilde{A}_i^{-1} A_i = U_i^{-1} \tilde{B}_i^{-1} B_i U_i \in U(n_i)$ , we have  $C_i := \tilde{B}_i^{-1} B_i \in U(n_i)$ . By the choice of  $B_i$  and  $\tilde{B}_i$ ,  $C_i$  is also upper triangular with positive diagonal, hence  $C_i = Id_{n_i}$ . Hence  $\tilde{B}_i = B_i$ ,  $\tilde{U}_i = U_i$ , and  $\tilde{A}_i = A_i$ . Thus  $C = Id$ , and consequently  $B = \tilde{B}$  and  $U = \tilde{U}$ , showing that  $\varphi$  is injective.

To show that  $\varphi$  is surjective, let  $(A, D) \in GL(n) \times U(n_1) \times \dots \times U(n_r)$ . Write  $A = \tilde{B} \tilde{U}$ , where  $\tilde{U} \in U(n)$ ,

$$\tilde{B} = \begin{pmatrix} \tilde{A}_i & & * \\ & \ddots & \\ 0 & & \tilde{A}_r \end{pmatrix} \in IUT(\mathbb{C}^n, *), \quad D = \begin{pmatrix} U_1 & & 0 \\ & \ddots & \\ 0 & & U_n \end{pmatrix},$$

and  $\tilde{A}_i = \tilde{B}_i \tilde{U}_i$  as before. Set

$$B = \tilde{B} \begin{pmatrix} \tilde{U}_1^{-1} U_1 & & 0 \\ & \ddots & \\ 0 & & \tilde{U}_r^{-1} U_r \end{pmatrix}, \quad U = \begin{pmatrix} U_1^{-1} \tilde{U}_1 & & 0 \\ & \ddots & \\ 0 & & U_r^{-1} \tilde{U}_r \end{pmatrix} \tilde{U}.$$

Then clearly  $\varphi(B, U) = (A, D)$ . q.e.d.

We now come to the bundle-version of the above lemma. Let  $(*)$  be as in (2.1). Define the map

$$f: IUT(E, *) \times U(E) \rightarrow GL(E)$$

fiberwise by  $f(B_x, U_x) = B_x \cdot U_x$ .

(2.14) **Lemma.** *f is a fibered morphism of bundles, i.e., f is a fiber bundle map over M and  $f: IUT(E, *) \times U(E) \rightarrow GL(E)$  is a fiber bundle.*

*Proof.* Clearly, f is a smooth fiber bundle map over M. Also f is onto. It remains to check local triviality. Let V be a coordinate chart in M such that all bundles appearing in Lemma (2.14) are trivial over V. Let  $W = GL(E)|_V \simeq V \times GL(n)$ . Define

$$\begin{aligned} \sigma: f^{-1}(W) &\simeq V \times IUT(\mathbb{C}^n, *) \times U(n) \\ &\rightarrow V \times GL(n) \times (U(n_1) \times \dots \times U(n_r)) \\ &\simeq W \times U(n_1) \times \dots \times U(n_r) \end{aligned}$$

by  $\sigma(x, B_x, U_x) = (x, \varphi(B_x, U_x))$ . The map  $\varphi$  on the right-hand side is the one defined in Lemma (2.13);  $\sigma$  is fiberwise an isomorphism and clearly smooth in x. This completes the proof.

(2.15) **Corollary.** *The map*

$$L_2^2(f): \mathfrak{g}_*^{\mathbb{C}^2} \times \mathfrak{g}^2 \rightarrow \mathfrak{g}^{\mathbb{C}^2}$$

*is a smooth fiber bundle with fiber isomorphic to  $\mathfrak{g}_{\text{DIAG}}^2$ .*

*Proof.* By [27, (14.16)], the map

$$L_2^2(f): \mathfrak{g}_*^{\mathbb{C}^2} \times \mathfrak{g}^2 = L_2^2(IUT(E, *) \times U(E)) \rightarrow L_2^2(GL(E)) = \mathfrak{g}^{\mathbb{C}^2}$$

is a fiber bundle with fiber isomorphic to  $L_2^2(\zeta^*(IUT(E, *) \times U(E)))$ , where  $\zeta$  is a  $C^\infty$  section of  $GL(E)$ . But  $\zeta^*(IUT(E, *) \times U(E))$  is isomorphic to  $U(F_1) \times \dots \times U(F_r)$ , hence  $L_2^2(\zeta^*(IUT(E, *) \times U(E))) \cong \mathfrak{g}_{\text{DIAG}}^2$ .

(2.16) **Theorem** (Decomposition theorem of the complex gauge group).

*We have the homeomorphism*

$$\mathfrak{g}^{\mathbb{C}^2} \simeq \mathfrak{g}_*^{\mathbb{C}^2} \times_{\mathfrak{g}_{\text{DIAG}}^2} \mathfrak{g}^2.$$

*Proof.* If p denotes the map  $L_2^2(f)$ , and  $\pi$  the natural quotient map, then there is a map g such that the following diagram commutes:

$$\begin{array}{ccc} \mathfrak{g}_*^{\mathbb{C}^2} \times \mathfrak{g}^2 & \xrightarrow{p} & \mathfrak{g}^{\mathbb{C}^2} \\ & \searrow \pi & \uparrow g \\ & & \mathfrak{g}_*^{\mathbb{C}^2} \times \mathfrak{g}^2 / \mathfrak{g}_{\text{DIAG}}^2 \end{array}$$

The map g is clearly a homeomorphism since both p and  $\pi$  are quotient maps. q.e.d.

We finish this section by discussing the space of equivalence classes of connections modulo the real gauge group, known as the *moduli space of*

connections. More precisely, let  $[\mathfrak{A}_1^2] := \mathfrak{A}_1^2/\mathfrak{g}^2$  and  $[\mathfrak{A}] = \mathfrak{A}/\mathfrak{g}$  denote the obvious quotient spaces endowed with the quotient topology. Unfortunately, the above spaces are not manifolds due to the presence of split connections [8, §3]. However, it is shown in the above cited reference that  $[\mathfrak{A}_1^2]$  and  $[\mathfrak{A}]$  are Hausdorff topological spaces. We denote by  $\pi$  the natural projection, and we set  $[\mathfrak{A}_{\mu 1}^2] := \pi(\mathfrak{A}_{\mu 1}^2)$ ,  $[\mathfrak{R}_\mu] := \pi(\mathfrak{R}_\mu)$ ,  $[\mathfrak{R}] := \pi(\mathfrak{R})$ , etc. In general we use the bracket notation to denote an element of the quotient space. For example, if  $x \in \mathfrak{A}_1^2$  we set  $[x] := \pi(x) \in [\mathfrak{A}_1^2]$ .

Due to the invariance of the Yang-Mills functional by the real gauge group, it defines a well-defined map:

$$f: [\mathfrak{A}_1^2] \rightarrow \mathbb{R}.$$

Therefore, the Yang-Mills equations can be defined in  $[\mathfrak{A}_1^2]$ . Some caution is needed due to the fact that  $[\mathfrak{A}_1^2]$  is not a manifold. However, it is only in the moduli space that the Yang-Mills equations are well behaved elliptic equations, and in this paper we shall mainly work in  $[\mathfrak{A}_1^2]$ . For example, we will obtain the convergence of the gradient flow of  $f$  in  $[\mathfrak{A}_1^2]$ , leaving the equivariant convergence of the flow in  $\mathfrak{A}_1^2$  to an upcoming paper.

(2.17) **Definition-Notation.** As a final remark we make the following notational conventions: Let  $\mathfrak{A}_s := j(\mathfrak{B}_s)$ ,  $\mathfrak{A}_{ss} := j(\mathfrak{B}_{ss})$ ,  $\mathfrak{A}_\mu := j(\mathfrak{B}_\mu)$ , etc. However, since as  $\mathfrak{g}$ -spaces  $\mathfrak{A}$  and  $\mathfrak{B}$  are isomorphic via  $j$ , we often will not distinguish between them. For example, notation like  $F(D'')$  or  $f(D'')$  will appear throughout the paper.

### 3. The manifold structure of $\mathfrak{B}_\mu$

(3.1) **Lemma.** *The natural map*

$$\varphi: \mathfrak{g}^{\mathbb{C}^2} \times_{\mathfrak{g}_*^{\mathbb{C}^2}} \mathfrak{B}_{*1}^2 \rightarrow \mathfrak{B}_1^2$$

defined by  $\varphi[g, D''] = g(D'')$  is a smooth immersion.

*Proof.* It is quite trivial to check that the map  $\varphi$  is well defined. The immersion statement is verified as follows: First observe that, by translating by the obvious group action, it suffices to check that  $\varphi$  is an immersion at the point  $(id, D'')$ . Consider the maps:

$$\mathfrak{g}_*^{\mathbb{C}^2} \xrightarrow{i_{D''}} \mathfrak{g}^{\mathbb{C}^2} \times \mathfrak{B}_{*1}^2 \xrightarrow{\pi} \mathfrak{g}^{\mathbb{C}^2} \times_{\mathfrak{g}_*^{\mathbb{C}^2}} \mathfrak{B}_{*1}^2,$$

where  $i_{D''} := (g^{-1}, g(D''))$  and  $\pi(g, D'') := [g, D'']$ . By an elementary computation, we have  $(di_{D''})_{id}(s) = (-s, D''s)$ , and, since  $\pi \circ i_{D''} = \text{constant}$ , we obtain  $d(\pi \circ i_{D''})_{id} = 0$  and thus  $\text{Im}(di_{D''})_{id} \subset \ker(d\pi)_{(id, D'')}$ .

Passing to the Lie algebras, we get a complex:

$$L\mathfrak{g}_*^{\mathbb{C}^2} \xrightarrow{di_{D''}} L\mathfrak{g}^{\mathbb{C}^2} \times T_{D''}\mathfrak{B}_{*1}^2 \xrightarrow{d\pi} T_{[id, D'']}\mathfrak{g}^{\mathbb{C}^2} \times_{y_*^{\mathbb{C}^2}} \mathfrak{B}_{*1}^2.$$

Consider  $d\pi(s, A) \in T_{[id, D]}(\mathfrak{g}^{\mathbb{C}^2} \times_{g_*^{\mathbb{C}^2}} \mathfrak{B}_{*1}^2)$  and let  $\bar{\varphi}$  be the obvious lift of  $\varphi$  according to the following diagram:

$$\begin{array}{ccc} \mathfrak{g}^{\mathbb{C}^2} \times \mathfrak{B}_{*1}^2 & & \mathfrak{B}_1^2 \\ \downarrow \pi & \searrow \bar{\varphi} & \\ \mathfrak{g}^{\mathbb{C}^2} \times_{g_*^{\mathbb{C}^2}} \mathfrak{B}_{1*}^2 & \nearrow \varphi & \end{array}$$

Since  $d\varphi \circ d\pi(s, A) = d\bar{\varphi}(s, A)$ , it follows that  $d\varphi(d\pi(s, A)) = 0$  if and only if  $0 = d\bar{\varphi}(s, A) = -D''s + A$  or equivalently  $A = D''s$ . Hence

$$d\pi(s, A) = d\pi(s, D''s) = d\pi(di_{D''}(-s)) = d(\pi \circ i_{D''})_{id}(-s) = 0,$$

proving that  $d\varphi_{[id, D'']}$  is injective. To show that the image splits note that, since  $\pi$  is a submersion, we have:

$$\text{Im } d\varphi_{[id, D'']} = \text{Im } d\bar{\varphi}_{[id, D'']} = D''(L\mathfrak{g}^{\mathbb{C}^2}) + L_1^2(T^*M'' \otimes UT(E, *)).$$

This, together with the ellipticity of  $D''$ , implies that the image is a closed subspace of a Hilbert space, and therefore splits. q.e.d.

Following Uhlenbeck and Yau [34], it is often useful to view holomorphic subbundles of  $E$  as orthogonal projections of  $E$  satisfying the additional property  $(1 - \pi)D''(\pi) = 0$ . The next lemma is quite general and does not make use of the fact that our base manifold is a Riemann surface. It will mainly be used in subsequent sections to obtain regularity of weak subbundles of  $E$ , i.e.,  $L_1^2$  projections  $\pi$  of  $E$  satisfying  $(1 - \pi)D''(\pi) = 0$ . The operator  $\Lambda$  is the usual contraction operator with the volume form on  $M$  [9, p. 111]. We will denote by  $\Delta, \Delta', \Delta''$  respectively the  $d$ -,  $\partial$ -, and  $\bar{\partial}$ -Laplacians on  $M$  with coefficients in  $\text{End } E$  [9, p. 152].

(3.2) **Lemma.** *Let  $\pi: E \rightarrow E$  be an automorphism of  $E$  such that  $\pi = \pi^*$ ,  $\pi^2 = \pi$ , and  $(1 - \pi)D''(\pi) = 0$ . Then*

$$\Delta(\pi) + 2\sqrt{-1}\Lambda[D''(\pi), D'(\pi)] + \sqrt{-1}(2\pi - 1)[\Lambda F, \pi] = 0.$$

*Proof.* The equation  $(1 - \pi)D''(\pi) = 0$  implies  $\pi D''(\pi) = D''(\pi)$ . By taking adjoints, we clearly have  $\pi(D'(\pi)) = 0$ .

By applying  $D''$  to the last equation, we obtain

$$D''(\pi) \cdot D'(\pi) + \pi D''(D'(\pi)) = 0,$$

which implies

$$\begin{aligned} &\sqrt{-1}\Lambda D''(\pi) \cdot D'(\pi) + \pi \Delta'(\pi) \\ &= \sqrt{-1}D''(\pi) \cdot D'(\pi) + \sqrt{-1}\Lambda D''D'(\pi) = 0. \end{aligned}$$

From the last equation we obtain

$$\frac{1}{2}\pi\Delta(\pi) + \sqrt{-1}(\frac{1}{2}\pi[\Lambda F, \pi] + \Lambda D''(\pi) \cdot D'(\pi)) = 0.$$

By applying  $D'$  to the equation  $(1 - \pi)D''(\pi) = 0$ , we get

$$-D'(\pi) \cdot D''(\pi) + (1 - \pi)D'D''(\pi) = 0,$$

which implies

$$-\sqrt{-1}\Lambda D'(\pi) \cdot D''(\pi) + (1 - \pi)\sqrt{-1}\Lambda D'D''(\pi) = 0.$$

The last equation, rewritten in terms of the Laplacian, gives

$$-\sqrt{-1}\Lambda D'(\pi) \cdot D''(\pi) + \frac{1}{2}(1 - \pi)(-\Delta(\pi) + \sqrt{-1}[\Lambda F, \pi]) = 0,$$

which yields

$$\frac{1}{2}\pi\Delta(\pi) - \frac{1}{2}\Delta(\pi) - \sqrt{-1}(\Lambda D'(\pi) \cdot D''(\pi) - \frac{1}{2}(1 - \pi)[\Lambda F, \pi]) = 0.$$

Combining this with the formula derived before, we complete the proof.

(3.3) **Lemma.** *Assume  $D''_1, D''_2 \in \mathfrak{B}_{*1}^{ss2}$  and  $g(D''_2) = D''_1$ . Then  $g \in \mathfrak{g}_*^{C^2}$ .*

*Proof.* Choose first  $h, h_1 \in \mathfrak{g}_*^{C^2}$  such that  $\tilde{D}'' := h(D'')$  and  $\tilde{D}''_1 := h_1(D''_1)$  belong to  $\mathfrak{B}_*^{ss}$ . We set also  $k := h_1 g h^{-1}$  and observe that  $k(\tilde{D}'') = D''_1$ . Since  $D''$  preserves the filtration  $(*)$ ,  $k(D'')$  preserves the filtration  $k(*)$ , obtained by taking the images under  $k$  of the subbundles in  $(*)$ . Let  $\pi$  denote a projection corresponding to a weak subbundle entering into the filtration  $k(\cdot)$ . Then  $\pi$  is clearly  $L_2^2$ . By Lemma (3.2) and the obvious elliptic regularity, we can immediately deduce that for any  $p \geq 2$ ,  $\pi \in L_2^p$ . By bootstrapping, we obtain  $\pi \in C^\infty$  and therefore the filtration  $k(*)$  is smooth. Now, the uniqueness of the Harder-Narasimhan filtration implies that  $k(*) = (*)$ . Hence, by the choice of  $h$  and  $h_1$ , we obtain  $g \in \mathfrak{g}_*^{C^2}$  and the proof is complete.

(3.4) **Corollary.** *The map  $\varphi$  (3.1) is a smooth injective immersion.*

*Proof.* The restriction of a smooth immersion to an open subset is a smooth immersion. It remains to verify the injectivity assertion. Assume  $\varphi[g, D''] = \varphi[g_1, D''_1]$ . Then  $D''_1 = g_1^{-1}g(D'')$  and, by (3.3),  $g_1^{-1}g \in \mathfrak{g}_*^C$ . Thus  $[g_1, D''_1] = [g_1, g^{-1}g_1(D'')] = [g, D'']$ . q.e.d.

We now proceed to show that the  $\mathfrak{B}_{\mu 1}^2$  are locally closed submanifolds of  $\mathfrak{B}_1^2$ . Our proof is essentially a slice theorem in a neighborhood of the

critical connections. Our slice is demonstrated by the decomposition of  $H^1(M, \text{End } E)$  into  $H^1(M, UT(E, *))$  and its perpendicular subspace. Our proof depends quite heavily on the Implicit Function Theorem, and it should be viewed as the transcendental analogue of the algebraic arguments of Atiyah and Bott [1, §15]. More specifically, we show:

(3.5) **Proposition.** *The set  $\mathfrak{B}_{\mu_1}^2$  is a locally closed embedded submanifold of  $\mathfrak{B}_1^2$ .*

*Proof.* Fix a filtration  $(*)$  of type  $\mu$  as in Definition (2.1), and let  $D_0'' \in \mathfrak{G}_{\mu_1}^2$  preserving  $(*)$ . We start by assuming that  $D_0''$  defines a critical connection. Let  $\mathcal{P} := \ker D_0''$  and let  $S$  be the  $L^2$ -orthogonal complement of  $\mathcal{P}$  in  $L_2^2(\text{End } E)$ . We endow  $E$  with the holomorphic structure  $D_0''$  and identify  $H^1(M, \text{End } E)$  with the appropriate subspace of  $L_2^2(T^*M'' \otimes \text{End } E)$  via Hodge theory. We define the map

$$f: S \times H^1(M, \text{End } E) \rightarrow \mathfrak{B}_1^2$$

by  $f(u, A) = e^u(D_0'' + A)$ . Clearly  $f(0, 0) = D_0''$ , and for any tangent vector  $(\delta u, \delta A)$  of  $S \times H^1(M, \text{End } E)$  at  $(0, 0)$  we have

$$(\delta f)_{(0,0)}(\delta u, \delta A) = D_0''(\delta u) + \delta A.$$

Thus, the map  $(\delta f)_{(0,0)}$  is an isomorphism of the corresponding tangent spaces, and  $f$  defines an isomorphism between neighborhoods  $U_0$  of  $(0, 0)$  in  $S \times H^1(M, \text{End } E)$  and  $\mathcal{O}_{D_0''}$  of  $D_0''$  in  $\mathfrak{B}_1^2$ . Let  $H^1(M, UT(E, *))$  be considered as a subspace of  $H^1(M, \text{End } E)$ , and let  $p: H^1(M, \text{End } E) \rightarrow H^1(M, \text{End } E)$  be the orthogonal projection onto  $H^1(M, UT(E, *))$  with respect to the  $L^2$ -inner product on  $L_2^2(\text{End } E)$ . Of course on the finite-dimensional space  $H^1(M, \text{End } E)$  all the norms are equivalent, and it makes no difference which one we are using. Let  $q = 1 - p$ , where 1 is the identity operator on  $H^1(M, \text{End } E)$ . We claim that our submanifold  $\mathfrak{B}_{\mu_1}^2$  is cut out by the equation  $q = 0$  in  $\mathfrak{B}_1^2$ . More specifically, if  $u \in S$  and  $A \in H^1(M, UT(E, *))$ , then clearly  $e^u(D_0'' + A) \in \mathfrak{B}_{\mu_1}^2$ . Conversely, let  $D'' \in \mathfrak{B}_{\mu_1}^2 \cap \mathcal{O}_{D_0''}$ . Then,  $D''$  can be written as  $D'' = e^u(D_0'' + B + \sigma)$ , where  $B$  and  $\sigma$  belong to  $H^1(M, \text{End } E)$ ,  $p(B) = B$ , and  $q(\sigma) = \sigma$ . We are going to show that  $\sigma = 0$ . Assume that  $\sigma \neq 0$  and express  $\sigma = (\sigma_{ij})$  with respect to the decomposition

$$(3.6) \quad H^1(M, \text{End } E) = \bigoplus_{i,j=1}^r H^1(M, \text{Hom}(F_i, F_j))$$

defined by the filtration  $(*)$ . Our contradiction will be obtained in several steps. In fact, we will show that if  $\sigma \neq 0$ , then we can find a one-parameter subgroup  $\{g_t: t \geq 0\}$  in the complex gauge group such that the energy of  $g_t(D'')$  is less than the minimum energy of the stratum  $\mathfrak{B}_{\mu_1}^2$ . Of course, this is a contradiction since the strata  $\mathfrak{B}_{\mu_1}^2$  are complex gauge invariant. Now, we present the full details of the proof.

We first claim that, with respect to the decomposition (3.6),  $\sigma$  is strictly lower triangular. Indeed set  $A = p(\Sigma)$ . By the orthogonality of  $A$  and  $\sigma$ , we obtain

$$0 = (A, \sigma)_{L^2} = \sum_{i \leq j} |\sigma_{ij}|_{L^2}^2,$$

hence  $\sigma_{ij} = 0$  for  $i \leq j$ , proving our claim.

The following claim is crucial: We can find  $\varepsilon, \eta > 0$  such that for

$$\sigma = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ \sigma_{21} & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \sigma_{r1} & \sigma_{r2} & \cdots & \sigma_{rr-1} & 0 \end{pmatrix},$$

with  $\sigma_{ij} \in H^1(M \text{Hom}(F_j, F_i))$  and  $|\sigma_{ij}|_{L^2} = \varepsilon$ , the following inequality holds:

$$|F(D''_0 + \sigma)|_{L^2}^2 < |F(D''_0)|_{L^2}^2 - \eta.$$

The proof is by computation:  $F(D''_0 + \sigma) = F(D''_0) - \sigma \wedge \sigma^* - \sigma^* \wedge \sigma$ , where  $\sigma^* \in H^{1,0}(M, \text{End } E)$  denotes the obvious metric adjoint of  $\sigma$ . Since  $D_0$  is critical of type  $\mu$ , we can find positive constants  $c_1, c_2$ , and  $c_3$  such that

$$\begin{aligned} |F(D''_0 + \sigma)|_{L^2}^2 &\leq |F(D''_0)|_{L^2}^2 + c_1 \sum |\sigma_{ij}|_{L^2}^4 + c_2 \sum |\sigma_{\mu\nu}^* \sigma_{kl}|_{L^2}^2 \\ &\quad + c_3 \sum_{\lambda > \rho} |\sigma_{\lambda\rho}|_{L^2}^2 (\mu_\lambda - \mu_\rho) \\ &\leq |F(D''_0)|_{L^2}^2 + c_1 \sum |\sigma_{ij}|_{L^2}^4 + c_2 \sum |\sigma_{\mu\nu}|_{L^2}^2 |\sigma_{kl}|_{L^2}^2 \\ &\quad + c_3 \sum_{\lambda > \rho} |\sigma_{\lambda\rho}|_{L^2}^2 (\mu_\lambda - \mu_\rho); \end{aligned}$$

the summing indices  $i, j, \dots, \lambda, \rho$  run between 1 and  $r$ . But on the finite-dimensional space  $H^1(M, \text{End } E)$  all norms are equivalent, hence

after changing  $c_1, c_2,$  and  $c_3,$  if necessary, we get

$$|F(D_0'' + \sigma)|_{L^2}^2 \leq |F(D_0'')|_{L^2}^2 + c_1 \sum |\sigma_{ij}|_{L^2_1}^4 + c_2 \sum |\sigma_{\mu\nu}|_{L^2_1}^2 |\sigma_{k\lambda}|_{L^2_1}^2 + c_3 \sum_{\lambda > \rho} |\sigma_{\lambda\rho}|_{L^2_1}^2 (\mu_\lambda - \mu_\rho).$$

Since  $\mu_\lambda < \mu_\rho$  if  $\lambda > \rho$  and  $|\sigma_{\alpha\beta}|_{L^2_1} = \varepsilon$  for  $\alpha, \beta = 1, \dots, r,$  we obtain

$$|F(D_0'' + \sigma)|_{L^2}^2 \leq |F(D_0'')|_{L^2}^2 + \varepsilon^2 (A\varepsilon^2 + B),$$

where  $A > 0$  and  $B < 0.$  Thus, we can find  $\varepsilon, \eta > 0$  such that

$$|F(D_0'' + \sigma)|_{L^2}^2 < |F(D_0'')|_{L^2}^2 - \eta \text{ for } |\sigma|_{L^2} = \varepsilon,$$

as claimed. Of course by changing  $\varepsilon, \eta,$  if necessary, we also obtain the inequality

$$|F(D_0'' + \sigma)|_{L^2} < |F(D_0'')|_{L^2} - \eta.$$

After these preparatory remarks we return to our task of obtaining a contradiction to our assumption  $\sigma \neq 0.$  We first note that we may cut our neighborhoods  $U_0$  and  $\mathcal{O}_{D_0''}$  more, if necessary, so that  $U_0$  is a product neighborhood of  $(0, 0)$  in  $S \times \text{Im } p \times \text{Im } q.$  We also may assume, without loss of generality, that the projection of  $U_0$  onto  $\text{Im } p$  is contained in a  $(\delta - L_1^2)$ -neighborhood of  $0 \in \text{Im } p,$  where  $\delta$  is chosen so that, if  $|\sigma|_{L^2_1} = \varepsilon,$  then

$$|F(D_0'' + B + \sigma) - F(D_0'' + \sigma)|_{L^2} < \eta/2.$$

This can be achieved in view of the inequality

$$|F(D_0'' + B + \sigma) - F(D_0'' + \sigma)|_{L^2} \leq |D_0(B)|_{L^2} + |[\sigma, B]|_{L^2} + \frac{1}{2}|[B, B]|_{L^2}$$

and the obvious function space inclusions. We may also assume that the projection of  $U_0$  onto  $\text{Im } q$  is contained in an  $(\varepsilon - L_1^2)$ -neighborhood of  $0 \in \text{Im } q.$

Now we are ready to define our one-parameter subgroup. Let  $\mu = (\mu_1, \dots, \mu_n)$  be as usual the type of the filtration  $(*)$ . We view  $\mu$  as a diagonal infinitesimal gauge transformation and set  $g_t = \exp(-t\mu)$  for  $t \geq 0.$  Since  $D_0''$  preserves the splitting,

$$g_t(D_0'' + \sigma) = D_0'' + \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ e^{t(\mu(F_1) - \mu(F_2))} \sigma_{21} & 0 & \dots & 0 & 0 \\ e^{t(\mu(F_1) - \mu(F_3))} \sigma_{31} & e^{t(\mu(F_2) - \mu(F_3))} \sigma_{32} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ e^{t(\mu(F_1) - \mu(F_r))} \sigma_{r1} & e^{t(\mu(F_2) - \mu(F_r))} \sigma_{r2} & \dots & e^{t(\mu(F_{r-1}) - \mu(F_r))} \sigma_{rr-1} & 0 \end{pmatrix},$$

$|g_t(D''_0 + \sigma) - D''_0|_{L^2_1} = |\sigma|_{L^2_1} < \varepsilon$ . Since  $\mu(F_i) < \mu(F_j)$  for  $i < j$ , it follows that  $|g_t(D''_0 + \sigma) - D''_0|_{L^2_1}$  increases to infinity with  $t$ . Hence, by continuity, we can find  $t > 0$  such that

$$|g_t(D''_0 + \sigma) - D''_0|_{L^2_1} = \varepsilon.$$

Also

$$g_t(D''_0 + B) = D''_0 + \begin{pmatrix} B_{11} & e^{t(\mu(F_2) - \mu(F_1))} B_{12} & \dots & e^{t(\mu(F_r) - \mu(F_1))} B_{1r} \\ 0 & B_{22} & \dots & e^{t(\mu(F_r) - \mu(F_2))} B_{2r} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & B_{rr} \end{pmatrix},$$

hence

$$\begin{aligned} |g_t(D''_0 + B) - D''_0|_{L^2_1} &= \sum_i |B_{ii}|_{L^2_1}^2 + \sum_{i < j} e^{2t(\mu(F_j) - \mu(F_i))} |B_{ij}|_{L^2_1}^2 \\ &\leq \sum_{i < j} |B_{ij}|_{L^2_1}^2 = |B|_{L^2_1}^2 < \delta. \end{aligned}$$

Therefore, by previous computations, we have

$$\begin{aligned} |F(g_t(D''_0 + \sigma))|_{L^2} &< |F(D''_0)|_{L^2} - \eta, \\ |F(g_t(D''_0 + B + \sigma)) - F(g_t(D''_0 + \sigma))|_{L^2} &< \eta/2. \end{aligned}$$

Thus, if  $f$  denotes the Yang-Mills functional, the above inequalities imply

$$f(g_t(D''_0 + B + \sigma)) < f(D''_0) = f(\mu),$$

contradicting (2.11). Hence  $\sigma = 0$  and the manifold structure of  $\mathfrak{B}^2_{\mu_1}$  is established at the critical connections. If  $D''_0$  is a general point in  $\mathfrak{B}^2_{\mu_1}$ , then, according to (2.11), we can find a gauge transformation  $g$  such that  $g(D''_0)$  is in a neighborhood of a critical point. Since gauging with  $g$  is a homeomorphism in  $\mathfrak{B}^2_1$ , we can translate the manifold structure to  $D''_0$  via  $g^{-1}$ . This defines a manifold atlas of  $\mathfrak{B}^2_{\mu_1}$  compatible with the one of  $\mathfrak{B}^2_1$  and completes the proof.

(3.7) **Theorem.** *The set  $\mathfrak{B}^2_{\mu_1}$  of holomorphic structures of type  $\mu$  is a locally closed submanifold of  $\mathfrak{B}^2_1$ . Furthermore, if  $(*)$  is a filtration of type  $\mu$ , the following homeomorphisms hold:*

$$\mathfrak{B}^2_{\mu_1} \simeq \mathfrak{g}^{C2} \times_{\mathfrak{g}^{C2}} \mathfrak{B}^{ss2}_{*1}, \quad \mathfrak{B}^2_{\mu_1} \simeq \mathfrak{g}^2 \times_{\mathfrak{g}^2_{\text{DIAG}}} \mathfrak{B}^{ss2}_{*1}.$$

*Proof.* The first statement was proven in (3.5). The proof of the second statement runs as follows: According to (3.2), the map

$$\varphi: \mathfrak{g}^{C2} \times_{\mathfrak{g}^{C2}} \mathfrak{B}^{ss2}_{*1} \rightarrow \mathfrak{B}^2_1$$

defined by  $\varphi(g, D'') := g(D'')$  is a smooth injective immersion onto its image. Since the image of  $\varphi$  in  $\mathfrak{B}_1^2$  is  $\mathfrak{B}_{\mu_1}^2$  and  $\mathfrak{B}_{\mu_1}^2$  is an embedded submanifold of  $\mathfrak{B}_1^2$ , we have that the image of  $d\varphi_{[id, D'']}$  lies in  $T_{D_0''} \mathfrak{B}_{\mu_1}^2$ . But as shown in the course of the proof of (3.1),

$$\text{Im } d\varphi_{[id, D_0'']} = D_0''(L\mathfrak{g}^{\text{C}2}) + L_1^2(T^*M'' \otimes UT(E, *)).$$

By taking perpendicular subspaces with respect to the  $L^2$  norm in  $T_{D_0''} \mathfrak{B}_1^2$ , we obtain that  $\text{Im } d\varphi_{[id, D_0'']}^\perp = \text{Im } q$ , where  $q$  is the projection used in the proof of Proposition (3.5). Thus  $\text{Im } d\varphi_{[id, D_0'']}$  and  $T_{D_0''} \mathfrak{B}_{\mu_1}^2$  have the same codimension in  $T_{D_0''} \mathfrak{B}_1^2$ . Hence,  $d\varphi_{[id, D_0'']}$  is onto  $T_{D_0''} \mathfrak{B}_{\mu_1}^2$ , proving that  $d\varphi_{[id, D_0'']}$  is an isomorphism. By translating and using the inverse function theorem, we show that  $\varphi$  is a diffeomorphism, as claimed. Finally, the homeomorphism  $\mathfrak{g}^{\text{C}2} \simeq \mathfrak{g}^2 \times_{\mathfrak{g}_{\text{DIAG}}^2} \mathfrak{g}_*^{\text{C}2}$  of Theorem (2.16) completes the proof. q.e.d.

The following lemma and corollary are due to Atiyah and Bott. They are the main tools in their inductive computation of cohomology. We include them for the sake of completeness. Corollary (3.9) will be used in §5.

(3.8) **Lemma.** *There is a  $\mathfrak{g}_{\text{DIAG}}^2$ -equivalent deformation retract of  $\mathfrak{B}_{*1}^{\text{SS}2}$  onto  $\prod_{i=1}^r \mathfrak{B}_{\text{SS}1}^2(F_i)$ , which restricts to a  $\mathfrak{g}_{\text{DIAG}}$ -equivalent retraction of  $\mathfrak{B}_*^{\text{SS}}$  onto  $\prod_{i=1}^r \mathfrak{B}_{\text{SS}}(F_i)$ .*

*Proof.* Let  $D'' = D_0'' + A'' \in \mathfrak{B}_{\text{SS}1}^2$ , with  $D_0''$  preserving the splitting, and let  $g_t$  be the one-parameter subgroup introduced in the proof of (3.5). If we write

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ & \ddots & & \\ 0 & 0 & \cdots & A_{rr} \end{pmatrix},$$

then it is easy to see that  $\lim_{t \rightarrow \infty} g_t(D_0'' + A'') = D_\infty''$ , where

$$D_\infty'' = D_0'' + \begin{pmatrix} A_{11} & & 0 \\ & \ddots & \\ 0 & & A_{rr} \end{pmatrix}.$$

Define  $H: \mathfrak{B}_{*1}^{\text{SS}2} \times [0, \infty] \rightarrow \mathfrak{B}_{*1}^{\text{SS}2}$  by

$$H(D'', t) = \begin{cases} g_t(D'') & \text{if } t < \infty, \\ D_\infty'' & \text{if } t = \infty. \end{cases}$$

The map  $H$  is clearly continuous,  $H(D'', 0) = D''$ ,  $H(D'', \infty) \in \prod_{i=1}^r \mathfrak{B}_{ss1}^2(F_i)$ , and, since  $g_i(D'') = D''$  holds for  $D'' \in \prod_{i=1}^r \mathfrak{B}_{ss1}^2(F_i)$  and  $t \in [0, \infty]$ ,  $H$  is the identity map on  $\prod_{i=1}^r \mathfrak{B}_{ss1}^2(F_i)$ . In order to prove equivariance, we let  $g \in \mathfrak{g}_{\text{DIAG}}^{\text{C2}}$ ,  $t \neq \infty$ , and observe  $H(g(D''), t) = g_t g(D'') = g g_t(D'') = g \cdot H(D'', t)$ . Since the last equation persists at the limits, the proof is complete.

(3.9) **Corollary.** *There is a deformation retract of  $[\mathfrak{B}_{\mu 1}^2]$  onto  $\prod_{i=1}^r [\mathfrak{B}_{ss1}^2(F_i)]$ , which restricts to a retraction of  $[\mathfrak{B}_{\mu}]$  onto  $\prod_{i=1}^r [\mathfrak{B}_{ss}(F_i)]$ .*

*Proof.* Define  $[H]: [\mathfrak{B}_{\mu 1}^2] \times [0, \infty] \rightarrow \mathfrak{B}_{\mu 1}^2$  by  $[H][D'', t] = [H(D'', t)]$ . The map  $[H]$  is well defined in view of the equality  $\mathfrak{g}_{*}^{\text{C2}} \cap \mathfrak{g}^2 = \mathfrak{g}_{\text{DIAG}}^2$ ; so it remains to show continuity. Let  $D''_i \rightarrow D''$ ,  $t_i \rightarrow t \in [0, \infty]$ ; write  $D''_i = [g_i, \tilde{D}''_i]$  and  $D'' = [g, \tilde{D}'']$ , where  $\tilde{D}''_i, \tilde{D}'' \in \mathfrak{B}_{*1}^{ss2}$  and  $g_i, g \in \mathfrak{g}^2$  with  $\tilde{D}''_i \rightarrow \tilde{D}''$ . By (3.7),  $\lim_i H(\tilde{D}''_i, t_i) = H(\tilde{D}'', t)$ , hence also  $\lim_i [H(\tilde{D}''_i, t_i)] = [H(\tilde{D}'', t)]$ , proving the continuity of  $[H]$ .

#### 4. Connection with Morse theory

In the previous sections we studied the space  $\mathfrak{B}$  of holomorphic structures without relying very heavily on the analytical aspects of Yang-Mills theory. In this section we make the connection with Morse theory. More precisely, by examining the flow of the gradient defined in §2, we obtain a Morse-theoretical stratification of our configuration space  $\mathfrak{A} \simeq \mathfrak{B}$  along the lines conjectured by Atiyah and Bott. Were the Yang-Mills functional sufficiently nondegenerate, the above stratification would be nothing but the stable-unstable manifold decomposition obtained via a Morse function. It will not be until §6 that we prove the manifold structure of the stratification obtained in this section. However, in this section we will set up the machinery for §6.

Let  $\mathfrak{A}_1^2$  be the space of  $L_1^2$ -connections on  $E$ . We endow the tangent spaces of  $\mathfrak{A}_1^2$  with the  $L^2$ -metric. Thus,  $\mathfrak{A}_1^2$  becomes (an incomplete) Riemannian manifold. The Yang-Mills functional  $f: \mathfrak{A}_1^2 \rightarrow \mathbb{R}$  is smooth and with respect to the Riemannian structure in  $\mathfrak{A}_1^2$  defines a gradient. With the notation explained in §2, it is given by the following formula [1, §4]:

$$\nabla f(D) = D^* F(D).$$

Similarly we have an associated Hessian, which is given by

$$H_D(\eta, \eta) = \langle D^* D\eta + *[ *F(D), \eta], \eta \rangle_{L^2},$$

where  $\eta \in T_D \mathfrak{A}_1^2$  [1, §4]. More generally, if  $f$  is a convex, invariant functional as in §2, we have similar formulas for  $\nabla f$  and  $H$  involving also  $f'$  and  $f''$  [1, §8].

Our goal in this section will be to study the gradient flow of the Yang-Mills functional  $f$ . We first show that  $f$  satisfies an equivariant Palais-Smale Condition  $C$  on  $\mathfrak{A}_1^2$  and thus also a Condition  $C$  on  $[\mathfrak{A}_1^2]$ .

(4.1) **Proposition.** *Let  $\{D_i\} \in \mathfrak{A}_1^2$  with  $\sup_i \{f(D_i)\} < \infty$  and  $|D_i^* F(D_i)|_{L^2} \rightarrow 0$  as  $i \rightarrow \infty$ . Then, we can find a sequence  $\{g_i\}$  in  $\mathfrak{g}$  and a subsequence  $\{D_j\}$  of  $\{D_i\}$  so that  $g_j(D_j) \xrightarrow{L_1^2} D_\infty$ , where  $D_\infty$  is a (smooth) critical connection. In particular,  $[D_j] \rightarrow [D_\infty]$  in  $[\mathfrak{A}_1^2]$ .*

*Proof.* Our proposition is a consequence of Uhlenbeck’s weak compactness theorem [32]. In fact, [32] assures us that we may choose  $\{g_i\}$  and subsequence  $\{D_j\}$  of  $\{D_i\}$  so that  $g_j(D_j) \xrightarrow{L_1^2} D_\infty$  (weakly). Thus, we only have to improve the convergence. The method is quite standard and parallels the verification of Condition C for the geodesics problem [26]. For notational convenience, we denote  $g_j(D_j)$  by  $D_j := D_0 + A_j$ ,  $F(D_j)$  by  $F_j$ , and set  $D_\infty := D_0 + A_\infty$ ; note that  $|D_j^* F_j|_{L^2} \rightarrow 0$  persists after this change. First notice that by the compactness  $L_1^2 CL^4$ ,  $A_j \xrightarrow{L^4} A_\infty$ . According to the  $L^4$ -slice theorem (cf. [34]) we may also assume  $D_\infty * (A_j - A_\infty) = 0$ . We start by showing that  $D_\infty$  is a critical point. Since

$$D_j * F_j - D_\infty * F_\infty = D_\infty (*F_j - *F_\infty) + [A_j - A_\infty, *F_j],$$

the above expression converges to 0 weakly in  $L_{-1}^2$ . But  $|D_j * F_j|_{L^2} \rightarrow 0$ , hence  $D_\infty * F_\infty = 0$ , and therefore  $D_\infty$  is critical. By setting  $\varepsilon_j := C|D_j^* F_j|$  we have

$$(4.2) \quad \begin{aligned} \langle D_\infty^* F_\infty - D_j^* F_j, A_\infty - A_j \rangle_{L^2} &= -\langle D_j^* F_j, A_\infty - A_j \rangle \\ &\leq C|D_j^* F_j|_{L^2} |A_\infty - A_j|_{L^2} \leq \varepsilon_j |A_\infty - A_j|_{L_1^2}. \end{aligned}$$

Also, since the  $L^2$  norms of  $F_j$  and  $D_\infty * F_j$  are bounded and  $A_j \xrightarrow{L_1^2} A_\infty$  (weakly), we obtain the following equality in view of the appropriate Sobolev embedding theorems:

$$\begin{aligned} &\langle D_\infty^* F_\infty - D_j^* F_j, A_\infty - A_j \rangle_{L^2} \\ &= \langle F_\infty, D_\infty(A_\infty - A_j) \rangle_{L^2} - \langle F_j, D_j(A_\infty - A_j) \rangle_{L^2} \\ &= |F_\infty - F_j|_{L^2}^2 + \text{terms converging to zero.} \end{aligned}$$

Moreover, as  $D_\infty * (A_j - A_\infty) = 0$ , we obtain

$$C|A_\infty - A_j|_{L^2} \leq |A_\infty - A_j|_{L^2} + |F_\infty - F_j|_{L^2} + \text{terms converging to zero.}$$

Hence

$$|F_\infty - F_j|_{L^2} \geq C|A_\infty - A_j|_{L^2} + \eta_j,$$

where  $\lim_{j \rightarrow \infty} \eta_j = 0$ . Therefore we also have

$$|F_\infty - F_j|_{L^2}^2 \geq C|A_\infty - A_j|_{L^2}^2 + 2\eta_j|A_\infty - A_j|_{L^2} + \eta_j^2$$

with  $\lim_{j \rightarrow \infty} \eta_j = 0$ . Hence,

$$(4.3) \quad \langle D_\infty^* F_\infty - D_j^* F_j, A_\infty - A_j \rangle_{L^2} \geq C|A_\infty - A_j|_{L^2}^2 + \delta_j|A_\infty - A_j|_{L^2} + \delta_j,$$

where  $\lim_{j \rightarrow \infty} \delta_j = 0$ . From (4.2) and (4.3) we obtain

$$\varepsilon_j|A_\infty - A_j|_{L^2} \geq C|A_\infty - A_j|_{L^2}^2 + \delta_j|A_\infty - A_j|_{L^2} + \delta_j,$$

which, together with  $\lim_{j \rightarrow \infty} \varepsilon_j = \lim_{j \rightarrow \infty} \delta_j = 0$ , proves (4.1). q.e.d.

Proposition (4.1) will provide us with all the necessary compactness needed to proceed. In the sequel we will study the equation

$$(4.4) \quad \begin{cases} \partial D / \partial t = -D^* F(D), \\ D(0) = D_0 \end{cases}$$

on the space  $\mathfrak{A}_1^2$  of connections on  $E$  compatible with the fixed metric  $K$ . As explained in [6], there is an equivalent way of viewing (4.4): We fix the holomorphic structure  $\bar{\partial} = D_0''$  and  $E$  and consider the equation

$$(4.5) \quad \begin{cases} \partial h / \partial t = 2\sqrt{-1}h(*F_{Kh} - \mu I) \\ \quad = -\{\Delta_0 h + \sqrt{-1}(*F_0 h + hF_0 - 2\mu h) \\ \quad \quad + 2\sqrt{-1}*(D_0'' h h^{-1} D_0' h)\} \\ h(0) = I \end{cases}$$

for positive endomorphisms  $h$ . (Here  $D_0'$ ,  $D_0''$ ,  $\Delta_0$ , and  $F_0$  are the operators corresponding to  $\bar{\partial}$  and  $K$ .) The correspondence between equations (4.4) and (4.5) is given as follows: Let  $\{D(t)\}$  be a solution of (4.4), where  $D(t) = g(t)D_0$  for some complex gauge transformation  $g(t)$ . Then  $h(t) = g(t)^* g(t)$  satisfies (4.5). Conversely, assume that  $h(t)$  satisfies (4.5). If we set  $g_t = h(t)^{1/2}$ , then some connection real gauge equivalent to  $g_t(D_0)$  satisfies (4.4). The approaches of fixing the holomorphic structure (and varying the metric) or fixing the metric (and varying the

holomorphic structure) are geometrically equivalent. They are related by the “moving frame”

$$g_t : (E, \bar{\partial}, Kh_t) \rightarrow (E, g_t(\bar{\partial}), K),$$

$$g_t * g_t = h_t,$$

which is an isometry and a holomorphic isomorphism. Hence the two approaches are completely equivalent.

Equation (4.5) is nonlinear parabolic and standard techniques can be applied to show (cf. [6]):

- (i) Short time existence for the solutions.
- (ii) Uniqueness.
- (iii) Extension of solutions for all time.

Moreover, by following [12, Proof of Theorem on p. 122], we also get smooth dependence upon the initial conditions. In particular:

- (iv) For any finite  $T > 0$ , there is a continuous flow

$$[\Phi] : [\mathfrak{A}_1^2] \times [0, T] \rightarrow [\mathfrak{A}_1^2]$$

defined by  $[\Phi][D, t] = [D_t]$ .

(4.6) **Lemma.** *Let  $D = D(t)$  be a solution of (4.4). Then*

$$\sup_{x,t} |F_{x,t}| < +\infty.$$

*Proof.* The equation  $\partial D/\partial t = -D^*F$  gives  $\partial F/\partial t = -D * D * F$ , hence

$$\frac{\partial *F}{\partial t} = -\Delta(*F) = -\nabla^* \nabla(*F).$$

Thus, by taking inner products pointwise, we obtain

$$\left\langle \frac{\partial *F}{\partial t}, *F \right\rangle_x = -\langle \nabla^* \nabla(*F), *F \rangle_x = -\frac{1}{2} \Delta_M |*F|_x^2 - |\nabla *F|_x^2,$$

hence

$$\frac{\partial}{\partial t} |*F|_x^2 + \Delta_M |*F|_x^2 = -2|\nabla *F|_x^2 \leq 0.$$

From the classical maximum principle, it follows that

$$\sup_{x,t} |*F_{x,t}| \leq \sup_x |*F_{x,0}| < \infty.$$

(4.7) **Lemma.** *If  $D = D(t)$  is the solution of  $\partial D/\partial t = -D^*F(D)$ , then  $|D^*F(D)|_{L^2} \rightarrow 0$  as  $t \rightarrow \infty$ .*

*Proof.* Since  $|\partial D/\partial t|_{L^2}^2 = -\frac{1}{2}(\partial/\partial t)|F|_{L^2}^2$ , we obtain for all  $T > 0$

$$\int_0^T \left| \frac{\partial D}{\partial t} \right|_{L^2}^2 dt = \frac{1}{2} (|F_0|_{L^2}^2 - |F_T|_{L^2}^2) \leq \frac{1}{2} |F_0|_{L^2}^2$$

and consequently

$$\int_0^\infty \left| \frac{\partial D}{\partial t} \right|_{L^2}^2 dt = \int_0^\infty |D^* F|_{L^2}^2 dt < \infty.$$

Hence, we can find a sequence of real numbers  $t_i \rightarrow +\infty$ , such that  $|D^*(F(t_i))|_{L^2} \rightarrow 0$ . By differentiating our original equation once more, we obtain

$$\frac{\partial^2 D}{\partial t^2} = -D^* \frac{\partial F}{\partial t} - * \left[ \frac{\partial D}{\partial t}, *F \right],$$

hence

$$\frac{\partial}{\partial t} \left| \frac{\partial D}{\partial t} \right|_{L^2}^2 = -2 \operatorname{Re} \left\langle \frac{\partial F}{\partial t}, \frac{\partial F}{\partial t} \right\rangle - 2 \operatorname{Re} \left\langle \frac{\partial D}{\partial t}, * \left[ \frac{\partial D}{\partial t}, F \right] \right\rangle.$$

According to (4.6) there is a constant  $B > 0$  such that

$$(4.8) \quad \frac{\partial}{\partial t} \left| \frac{\partial D}{\partial t} \right|_{L^2}^2 \leq -2 \left| \frac{\partial F}{\partial t} \right|_{L^2}^2 + B \left| \frac{\partial D}{\partial t} \right|_{L^2}^2.$$

By integrating, we obtain

$$\left| \frac{\partial D}{\partial t}(T) \right|_{L^2}^2 + 2 \int_0^T \left| \frac{\partial F}{\partial t} \right|_{L^2}^2 dt \leq \left| \frac{\partial D}{\partial t}(0) \right|_{L^2}^2 + B \int_0^T \left| \frac{\partial D}{\partial t} \right|_{L^2}^2 dt.$$

But since

$$\int_0^\infty \left| \frac{\partial D}{\partial t} \right|_{L^2}^2 dt < \infty,$$

we have

$$\int_0^\infty \left| \frac{\partial F}{\partial t} \right|_{L^2}^2 dt < \infty.$$

Let  $T_i$  be any sequence of real numbers converging to  $\infty$ , and let  $t_i$  be our sequence of real numbers with the property  $|(\partial D/\partial t)(t_i)|_{L^2} \rightarrow 0$ . Set  $t_{\lambda_i} := \max\{t_j : t_j \leq T_i\}$ . Rename  $t_{\lambda_i} = t_i$  and note that  $t_i \rightarrow \infty$ . By integrating (4.8) from  $t_i$  to  $T_i$ , we get

$$\begin{aligned} 0 &\leq \left| \frac{\partial D}{\partial t}(T_i) \right|_{L^2}^2 + 2 \int_{t_i}^{T_i} \left| \frac{\partial F}{\partial t} \right|_{L^2}^2 dt \\ &\leq \left| \frac{\partial D}{\partial t}(t_i) \right|_{L^2}^2 + B \int_{t_i}^{T_i} \left| \frac{\partial D}{\partial t} \right|_{L^2}^2 dt, \end{aligned}$$

which implies that

$$\left| \frac{\partial D}{\partial t}(T_i) \right|_{L^2} = |D^* F(T_i)|_{L^2} \rightarrow 0.$$

Since  $T_i$  was arbitrary, this completes the proof. q.e.d.

Given  $[D] \in [\mathfrak{A}_1^2]$ , we define the *limit set*  $\omega[D]$  of  $[D]$  to be the set of all  $[D_\infty] \in [\mathfrak{A}_1^2]$ , such that any neighborhood of  $[D_\infty]$  contains  $[D_t]$  with  $t$  arbitrarily large. By combining now Condition C with Lemma (4.7) we obtain the following:

(4.9) **Corollary.** *For any  $D$ ,  $\omega[D]$  is a nonempty subset of the critical set  $[\mathfrak{N}]$ .*

Moreover, we show:

(4.10) **Lemma.**  *$\omega[D]$  is connected.*

*Proof.* Assume  $\omega[D] = (\omega[D] \cap U) \cup (\omega[D] \cap V)$ , where  $U$  and  $V$  are open subsets of  $[\mathfrak{A}]$ ,  $U \cap V = \emptyset$ ,  $\omega[D] \cap U \neq \emptyset$ , and  $\omega[D] \cap V \neq \emptyset$ . We claim that we can find  $T \geq 0$  such that  $\{[D_t]\}_{t \geq T} \subset U \cup V$ . For, otherwise we can take  $t_i \rightarrow \infty$ , such that  $[D_{t_i}] \notin U \cup V$ . By Condition C we can choose subsequence  $[D_{t_i}] \rightarrow D_\infty \in \omega[D] \subset U \cup V$  hence  $[D_{t_i}] \in U \cup V$ , a contradiction. Consequently,  $\{[D_t]\}_{t \geq T} \subset U \cup V$ . Since  $\{[D_t]\}_{t \geq T}$  is connected, it lies in one of them, say  $U$ . But then the closure of  $\{[D_t]\}_{t \geq T}$  lies in  $[\mathfrak{A}] - V$ , hence also  $\omega[D]$ , contradicting  $\omega[D] \cap V \neq \emptyset$ . q.e.d.

Before we proceed, we need to observe that the critical sets  $[\mathfrak{N}_\mu]$  are open and closed in  $[\mathfrak{N}]$ . This is a consequence of the convexity arguments explained in [1, §12], as follows: Write  $[\mathfrak{N}] = \bigcup_i [\mathfrak{N}^i]$  as the union of its connected components. If  $[\mathfrak{N}^i] \cap [\mathfrak{N}_\mu] \neq \emptyset$ , then we claim that  $[\mathfrak{N}^i] \subset [\mathfrak{N}_\mu]$ . For, if  $[\mathfrak{N}^i] \cap [\mathfrak{N}_\lambda] \neq \emptyset$ , then we have  $f(\mu) = f(\mathfrak{N}_\mu) = f(\mathfrak{N}_\lambda) = f(\lambda)$  for all convex functionals  $f$ , hence  $\lambda = \mu$ . Therefore  $[\mathfrak{N}_\mu]$ , being the union of its connected components, is open and closed. Now we can define

$$[\mathfrak{C}_\mu] := \{[D] \in [\mathfrak{A}_1^2] : \omega[D] \subset \mathfrak{N}_\mu\}, \quad \mathfrak{C}_\mu := \pi^{-1}[\mathfrak{C}_\mu].$$

We call  $\mathfrak{C}_\mu$  and  $[\mathfrak{C}_\mu]$  the *Morse strata* of  $f$ .

(4.11) **Lemma.** (i)  $\{[\mathfrak{C}_\mu]\}_\mu$  forms a partition of  $[\mathfrak{A}_1^2]$ .

(ii)  $\{\mathfrak{C}_\mu\}_\mu$  forms a partition of  $\mathfrak{A}_1^2$ .

*Proof.* (i) Let  $[D] \in [\mathfrak{A}]$ . Since  $\omega[D] \neq \emptyset$ , then  $[D] \in [\mathfrak{C}_\mu]$  for some  $\mu$ ; hence the  $[\mathfrak{C}_\mu]$ 's cover  $[\mathfrak{A}]$ . Now, the connectedness of  $\omega[D]$  proves that they form a partition.

(ii) Clear from (i).

(4.12) **Proposition.** *For each  $\mu$  there exists a neighborhood  $[V_\mu]$  of  $[\mathfrak{N}_\mu]$  such that  $[V_\mu] \cap [\mathfrak{A}_{\mu 1}^2] \subset [\mathfrak{C}_\mu]$ .*

*Proof.* Let  $[D] \in [\mathfrak{A}_{\mu_1}^2] \cap [\mathfrak{C}_\lambda]$ . Since the flow preserves the gauge orbits, we can choose complex gauge transformations  $g_i$  such that  $g_i(D) \xrightarrow{L_1^2} D_\infty \in \mathfrak{N}_\lambda$ . By (2.11), for any convex functional  $f$ , we have  $f(g_i(D)) \geq f(\mu)$ ; thus also  $f(\lambda) = f(D_\infty) \geq f(\mu)$ , hence  $\lambda \geq \mu$ . By restricting now to a neighborhood  $[V_\mu]$  of  $[\mathfrak{N}_\mu]$ , we will exclude the possibility  $\lambda > \mu$ . The set  $V_\mu$  is constructed as followed: Let  $f$  be the Yang-Mills functional; there are only finitely many  $\mu_1 > \dots > \mu_i = \mu > \dots > \mu_n$  such that  $f(\mu_i) = f(\mu)$ . Let  $\varepsilon > 0$  such that  $\mathfrak{N}_{\mu_1}, \dots, \mathfrak{N}_{\mu_n}$  are the only critical sets intersecting  $f^{-1}(f(\mu) - 2\varepsilon, f(\mu) + 2\varepsilon) =: U_{f(\mu)}$ . Pick convex functionals  $f_i$  such that

$$f_i(\mu_i) > f_i(\mu_{i+1}) \quad (i = 1, \dots, n - 1)$$

and define

$$V_{\mu_i} := \bigcap_{j=1}^{i-1} f_j^{-1}[f_j(\mu_i) - \eta_j, f_j(\mu_i) + \eta_j] \cap U_{f(\mu_i)},$$

where  $\eta_i = (f_i + \mu_i) - f_i(\mu_{i+1})/2$ . We claim that, if  $[D] \in [V_\mu] \cap [\mathfrak{A}_{\mu_1}^2] \cap [\mathfrak{C}_\lambda]$  and  $\lambda \geq \mu$ , then  $\lambda = \mu$ . Let us assume  $\lambda > \mu$ . Since  $D \in U_{f(\mu)}$ , then  $f(D) < f(\mu) + \varepsilon$  and  $f(D_\infty) < f(\mu) + \varepsilon$ . Since  $\{\lambda: f(\mu) \leq f(\lambda) < f(\mu) + \varepsilon\} \subset \{\mu_1, \dots, \mu_n\}$ , we have  $\lambda \in \{\mu_1, \dots, \mu_n\}$ ; say  $\lambda = \mu_k$  for some  $k < i$ . But then

$$\begin{aligned} f_k(\mu_k) &= f_k(D_\infty) \leq f_k(D) < f_k(\mu_i) + \eta_k \\ &= f_k(\mu_i) + \frac{1}{2}[f_k(\mu_k) - f_k(\mu_{k+1})], \end{aligned}$$

which combined with the  $k + 1 \leq i$ ,  $\mu_{k+1} \geq \mu_i$ , and  $f_k(\mu_{k+1}) \geq f_k(\mu_i)$ , shows

$$f_k(\mu_k) < f_k(\mu_i) + \frac{1}{2}[f_k(\mu_k) - f_k(\mu_i)] < f_k(\mu_k),$$

a contradiction. This proves  $\lambda = \mu$  and completes the proof. q.e.d.

The reverse of the inclusions proved in Proposition (4.12) will be shown in §6 after we establish the convergence of the gradient flow in §5.

### 5. Retraction along the flow

In this section we show that the gradient lines converge rather than spiraling around the critical set. Furthermore, the flow defines a retraction of the strata considered in §3 onto the Yang-Mills connections. We also show that the retraction above coincides with the retraction defined by one-parameter subgroups via algebraic considerations.

Let  $D$  be a stable connection. According to [5], there is a Yang-Mills connection  $D_\infty$  on the complex gauge orbit of  $D$ , unique up to real gauge equivalence. Therefore, if  $\mathfrak{A}_s$  denotes the set of stable holomorphic structures, we have a map  $r: [\mathfrak{A}_s] \rightarrow [\mathfrak{A}_0]$  defined by  $r([D]) = [D_\infty]$ . We want to extend the map  $r$  to the whole semistable stratum:

$$(5.1) \quad r: [\mathfrak{A}_{ss}] \rightarrow [\mathfrak{A}_0].$$

For this, we need the following result of Seshadri [28]:

(5.2) **Proposition.** *Let  $(E, D'')$  be a semistable bundle. Then, there is a filtration of holomorphic subbundles*

$$(*) \quad 0 = E_0 \subset E_1 \subset \dots \subset E_r = E$$

such that  $E_i/E_{i-1}$  ( $i = 1, \dots, r$ ) is stable and  $\mu(E_i/E_{i-1}) = \mu(E_{i+1}/E_i)$  ( $i = 1, 2, \dots, r - 1$ ). Moreover, the isomorphism class of  $E_1 \oplus E_2/E_1 \oplus \dots \oplus E_r/E_{r-1}$  depends only upon the isomorphism class of  $E$ .

We shall call the filtration of (5.2) a *Seshadri filtration* of  $E$ . Now we define  $r$  as follows: Let  $D''$  be a semistable holomorphic structure and let  $(*)$  be a Seshadri filtration. Let  $F_i$  be the orthogonal complement of  $E_{i-1}$  in  $E_i$  and let  $D''_0$  be any holomorphic structure on  $E$  that splits as  $D''_0 = D''_0^{(1)} \oplus \dots \oplus D''_0^{(r)}$  with respect to  $E \simeq_{C^\infty} F_1 \oplus \dots \oplus F_r$ . Clearly, with respect to the filtration  $(*)$ ,  $D''$  has the form:

$$D'' = D''_0 + \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1r} \\ 0 & A_{22} & \dots & A_{2r} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & A_{rr} \end{pmatrix}.$$

We set

$$r(D'') = r(D''_0^{(1)} + A_{11}) \oplus \dots \oplus r(D''_0^{(r)} + A_{rr}).$$

The next two lemmas prove that  $r$  is well defined and complex gauge invariant.

(5.3) **Lemma.** *Assume that  $D''_1$  and  $D''_2$  are Yang-Mills, semistable, and complex gauge equivalent. Then  $D''_1$  is real gauge equivalent to  $D''_2$ .*

*Proof.* Let  $D''_2 = g(D''_1)$ , where  $g \in \mathfrak{g}^C$ . Without loss of generality, we may assume  $g = g^*$ ,  $g$  positive. By [5, pp. 276–277], we have  $D_2(g^2) = 0$  and therefore  $D_2$  splits according to the eigenvalues of  $g$ , say  $E = F_1 \oplus \dots \oplus F_s$  with  $g|_{F_i} = c_i Id_{F_i}$ . Consider the maps  $E_i \rightarrow E \xrightarrow{Id} E \rightarrow F_j$ . Clearly, for each  $i$  there is  $j = j(i)$  such that  $\lambda_i: E_i \rightarrow E \rightarrow F_{j(i)}$  is nonzero. The map  $\lambda_i$  is a monomorphism or generically an epimorphism.

In the second case, if we denote by  $S_i$  the sheaf  $\text{Im } \lambda_i \subset F_{j(i)}$ , we have the inequalities

$$\mu(E) = \mu(E_i) \leq \mu(S_i) \leq \mu(F_{j(i)}) = \mu(E).$$

Hence,  $\mu(S_i) = \mu(F_{j(i)})$ , or, since  $\text{rk } S_i = \text{rk } F_{j(i)}$ , we obtain  $c_1(S_i) = c_1(F_{j(i)})$  and thus  $c_1(F_{j(i)}/S_i) = 0$ . But, since  $\lambda_i$  is generically an epimorphism,  $F_{j(i)}/S_i$  is a torsion sheaf, hence  $F_{j(i)}/S_i = 0$ , implying that  $S_i = F_{j(i)}$ . Therefore,  $\lambda_i$  is onto, and, by the stability of  $E_i$ ,  $\lambda_i$  is an isomorphism. In any case  $\lambda_i$  is a monomorphism, so  $E_i$  is a subbundle of  $F_{j(i)}$  and  $g|_{E_i} = c_i Id_{E_i}$ . Hence, with respect to  $E \simeq E_1 \oplus \dots \oplus E_r$ , we can write

$$g = \begin{pmatrix} c_1 Id_1 & & 0 \\ & \ddots & \\ 0 & & c_r Id_r \end{pmatrix}, \quad D''_1 = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix}.$$

Thus,  $D''_2 = g(D''_1) = D''_1$ .

(5.4) **Lemma.** *If  $D'' \in \mathfrak{A}_{ss}$  and  $g \in \mathfrak{g}^C$ , then  $r(g(D'')) = r(D'')$ .*

*Proof.* According to (5.2),  $r(g(D'')) = \tilde{g}r(D'')$ , where  $\tilde{g} \in \mathfrak{g}^C$ . The rest follows by (5.3).

(5.5) **Proposition.** *If  $D_0$  is Yang-Mills and semistable, then we can find  $\varepsilon > 0$  and  $K > 0$  such that for all  $B \in L^2_1(T^*M \otimes u(E))$  and  $|B|_{L^2} < \varepsilon$  the following hold:*

- (i) *There exists a  $u$   $L^2$ -perpendicular to  $\mathcal{P} := \ker D''_0$  such that  $e^u(D_0 + B) = D_0 + A$ , where  $|u|_{L^2_2} \leq K\varepsilon$  and  $D''_0 A' = 0$ .*
- (ii) *Let  $D_0 + \tilde{B}$  be complex gauge equivalent to  $D_0 + B$  with  $\tilde{B} \in L^2_1(T^*M \otimes u(E))$  and  $|\tilde{B}|_{L^2_1} < \varepsilon$ , and let  $D_0 + \tilde{A}$  be the connection obtained from  $D_0 + \tilde{B}$  via (i); then  $D_0 + \tilde{A} = p(D_0 + A)$ , where  $p \in \mathcal{P}$ .*
- (iii) *If  $D''_0 + B''$  preserves the same filtration with  $D''_0$ , then so does  $A''$ .*

*Proof.* (i) Let  $S$  be the space of sections of  $\text{End}(E)$  which are  $L^2$ -perpendicular to  $\mathcal{P}$ , and define

$$\mathcal{D}: S \times L^2_1(T^*M' \otimes \text{End } E) \rightarrow S$$

by

$$\mathcal{D}(u, B') = (\sqrt{-1}\Lambda D''_0(e^{-u^*} B' e^{u^*} + e^{-u^*} D'_0 e^{u^*}))^*,$$

where  $*$  denotes the fiberwise adjoint with respect to the metric on  $E$ . We observe that

$$\left. \frac{\partial \mathcal{D}}{\partial U} \right|_{(0,0)} (\delta u) = (\sqrt{-1}\Lambda D''_0 D'_0 (\delta u)^*)^* = \Delta''_0 (\delta u),$$

which is invertible as a map from  $S$  to itself. Hence, by the implicit function theorem, we can solve the equation  $\mathcal{D}(u, B') = 0$  for  $B'$   $\varepsilon$ -close to  $0 \in L^2_1(T^*M' \otimes \text{End } E)$  and obtain the estimate  $|u|_{L^2} \leq K|B'|_{L^2}$  for some positive constant  $K$  depending upon  $\mathcal{D}$  and  $\varepsilon$ . By setting

$$(D_0 + A) = e^u(D_0 + B)$$

property (i) is clearly satisfied.

(ii) Let us assume now that  $D_0 + \tilde{B} = g(D_0 + \tilde{A})$  for some  $g \in \mathfrak{g}^C$  with  $|g|_{L^2} = 1$  and  $|\tilde{B}|_{L^2_1} < \varepsilon$ . Let  $D_0 + \tilde{A} = e^{\tilde{u}}(D_0 + \tilde{B})$  be obtained as in (i) and let  $p = e^{\tilde{u}}ge^{-\tilde{u}}$ . Of course,  $D_0 + \tilde{A} = p(D_0 + A)$  and  $|p|_{L^2} \leq e^{2K\varepsilon}$ . We claim that  $p \in \mathcal{P}$ . Write

$$D'_0 p^* = p^* \tilde{A}' - A' p^*, \quad D''_0 D'_0 p^* = D''_0(p^* \tilde{A}' - A' p^*).$$

Let  $p = \hat{p} + q$ , where  $\hat{p} \in \mathcal{P}$  and  $q \in \mathcal{P}^\perp$ . Then, since  $D''_0$  is Yang-Mills,  $D''_0 A' = D''_0 \tilde{A}' = 0$ , we obtain  $D''_0 D'_0 q = D''_0(q \tilde{A}' - A' q)$ . Since  $|A|_{C^0} \leq C\varepsilon$ , integrating against  $q$

$$\begin{aligned} \gamma|q|_{L^2} &\leq |D'_0 q|_{L^2} \leq D'_0 q|_{L^2} (|\tilde{A}' q|_{L^2} + |A' q|_{L^2}) \\ &\leq 2C\varepsilon|q|_{L^2}|q|_{L^2}, \end{aligned}$$

where  $C$  and  $\gamma$  are Sobolev constants. But also  $|q|_{L^2} \leq |p|_{L^2} \leq e^{2K\varepsilon}$ . Hence by making  $\varepsilon$  smaller, if necessary, we obtain a contradiction unless  $q = 0$ .

(iii) Let  $(*)$  be a filtration preserved by  $D''_0$  and let  $S'$  be the subset of  $S$  consisting of the endomorphisms of  $E$  that preserve  $(*)$ . The map  $\mathcal{D}$  restricts to a smooth map

$$\mathcal{D}: S' \times L^2_1(T^*M'' \otimes UT(*)) \rightarrow S'.$$

The rest follows as in (i).

(5.6) **Lemma.** *If  $\{D_i\}$  is a sequence of semistable connections on  $E$  converging in  $L^2_1$  to a critical semistable connection  $D$ , and  $\{\pi_i\}$  is a sequence of endomorphisms of  $E$  satisfying  $\pi_i = \pi_i^*$ ,  $\pi_i^2 = \pi_i$ ,  $(1 - \pi_i)D''_i(\pi_i) = 0$ , and  $\mu(\pi_i) = \mu(E)$ , then there is a subsequence  $\{\pi_j\}$  of  $\{\pi_i\}$  converging in  $L^2_1$  to  $\pi$ , where  $D''(\pi) = 0$ .<sup>1</sup>*

*Proof.* By taking subsequences, if necessary, we may assume that the  $\pi_i$ 's have the same topological type. Recall [34] that the Chern class of  $\pi_i$

<sup>1</sup> The convergence can be easily improved by using (3.2) and a standard linearization argument; see, for example, [8, Proposition 8.3].

is defined by

$$c_1(\pi_i) = \frac{\sqrt{-1}}{2\pi} \int_M \text{tr } \pi_i F(D_i) + |D''_i \pi_i|_{L^2}^2.$$

Since  $F(D_i) \xrightarrow{L^2} F(D) = \mu(E) \cdot d$  and  $|\pi_i|_{L^2}$  is bounded, we have that, after passing to a subsequence,  $\pi_i \xrightarrow{L^2} \pi$  (weakly). Hence  $\pi_i \xrightarrow{L^2} \pi$ , and  $|D''_i \pi_i|_{L^2} \rightarrow 0$  since  $\mu(\pi) = \mu(\pi_i) = \mu(E)$ . Therefore  $|D'' \pi_i|_{L^2} \rightarrow 0$ , which combined with the  $\pi_i \xrightarrow{L^2} \pi$  proves the lemma. q.e.d.

Now let us assume that  $D_0$  is Yang-Mills semistable preserving a Seshadri filtration  $0 \subset E_1 \subset E$ . We also assume without loss of generality that  $E_1$  is not isomorphic to  $E/E_1 := E_2$ . Consider for  $i, j \in \{1, 2\}$  the operator

$$C^\infty(\text{Hom}(E_i, E_j)) \xrightarrow{D''_0} C^\infty(T^*M'' \otimes \text{Hom}(E_i, E_j)).$$

By Riemann-Roch and the assumption that our base curve has genus  $g \geq 2$  we can pick  $\beta$  and  $\gamma$  nonzero elements of the cokernel of  $D''_0$ , where  $\beta \in C^\infty(T^*M'' \otimes \text{Hom}(E_2, E_1))$  and  $\gamma \in C^\infty(T^*M'' \otimes \text{Hom}(E_1, E_2))$ . We set

$$\Sigma = \begin{pmatrix} 0 & \beta \\ \gamma & 0 \end{pmatrix};$$

clearly  $D'_0 \Sigma = 0$ .

(5.7) **Lemma.** *If  $\Sigma$  is as above, then there exists an  $\varepsilon_0 > 0$  such that for any  $0 < \varepsilon \leq \varepsilon_0$  the holomorphic structure  $D''_\varepsilon := D''_0 + \varepsilon \Sigma$  is stable.*

*Proof.* Assume that there is a sequence of positive numbers  $\varepsilon_i \rightarrow 0$ , such that  $D''_{\varepsilon_i}$  is not stable. By (5.6) we may assume that  $D''_{\varepsilon_i}$  preserves a Seshadri filtration  $0 \subset F_i \subset E$  with  $F_i$  topologically equivalent to  $E_1$  or  $E_2$ , say, e.g.,  $E_1$ . Choose  $g_i$  to be a complex gauge transformation such that  $g_i(F_i) = E_1$ . Therefore,  $g_i(D''_{\varepsilon_i})$  preserves the filtration  $0 \subset E_1 \subset E$  with stable quotients. We also may assume, without loss of generality, that  $g_i(D_{\varepsilon_i}) \xrightarrow{L^2} D$ , where  $D$  is Yang-Mills preserving the same Seshadri filtration. By a standard argument [8, p. 59], the automorphisms  $g_i/|g_i|_{L^2}$  and  $g_i^{-1}/|g_i^{-1}|_{L^2}$  converge to nontrivial endomorphisms  $h_1$  and  $h_2$ . Clearly,  $h_1: (E, D_0) \rightarrow (E, D)$  and  $h_2: (E, D) \rightarrow (E, D_0)$  are holomorphic maps. By combining  $h_1$  and  $h_2$  we show that  $D_0$  and  $D$  are gauge equivalent, say  $D = g(D_0)$ . By the stability of  $D_0$  and  $D$  restricted to  $E_1$  and  $E_2$ ,

we have that  $g$  preserves the splitting  $E_1 \oplus E_2$ , hence  $g^{-1}g_i(D_{\varepsilon_i})$  preserves a Seshadri filtration  $0 \subset E_j \subset E$ , where  $j \in \{1, 2\}$ . Observe also that  $g^{-1}g_i(D_{\varepsilon_i}) \xrightarrow{L_1^2} D_0$ . The rest follows from (5.5)(ii) and (iii). *q.e.d.*

Now let  $D''$  be a semistable holomorphic structure, preserving a non-trivial Seshadri filtration  $0 \subset E_1 \subset E$ . We claim that  $D''$  can be approximated in the  $L_1^2$  norm by stable holomorphic structures. First choose a one-parameter family of gauge transformations  $\{g_t\}_{t \geq 0}$  such that  $\lim_{t \rightarrow \infty} g_t(D) =: D_0$  is Yang-Mills, preserving the same filtration as  $D$ . Then, choose (by (5.7))  $\varepsilon_0 > 0$  such that for any  $\Sigma$  as in (5.7) with  $|\beta|_{L_1^2} < \varepsilon_0, |\gamma|_{L_1^2} < \varepsilon_0$ , the holomorphic structure  $D''_0 + \Sigma$  is stable. Choose  $t$  and adjust  $\varepsilon_0$ , if necessary, so that  $g_t(D)$  lies in an  $\varepsilon_0$ -neighborhood of  $D_0$  such that (5.5) holds in the  $\varepsilon_0$ -neighborhood. We may assume without loss of generality that  $g_t(D'')|_{E_1} \oplus g_t(D'')|_{E/E_1} = D''_0$ . Then, by the same argument as in (5.5)(i) and (iii) we obtain  $D''_0 + A$  gauge equivalent to  $D$ , where

$$A = \begin{pmatrix} 0 & \gamma \\ 0 & 0 \end{pmatrix}, \quad \gamma \in H^1(M, \text{Hom}(E/E_1, E_1)), |\gamma|_{L_1^2} < \varepsilon_0.$$

Choose  $\delta_i \in H^1(M, \text{Hom}(E_1, E/E_1)), |\delta_i|_{L_1^2} \rightarrow 0$ . Then  $D''_0 + \begin{pmatrix} 0 & \gamma \\ \delta_i & 0 \end{pmatrix}$  are stable, converging to  $D''_0 + A$ . Hence,  $D''_0 + A$  can be approximated by stable holomorphic structures and therefore the same holds for  $D$ . Now we are ready to prove the following:

(5.8) **Proposition.** *The set  $\mathfrak{A}_s$  of stable connections is dense in  $\mathfrak{A}_{ss}$ .*

*Proof.* Let  $(E, D'')$  be a semistable holomorphic structure. We have already shown that (5.8) holds in the case where  $D''$  satisfies a two-step Seshadri filtration. The general case follows by induction on the length  $r$  of the Seshadri filtration. Let us assume that  $r \geq 2$  and Proposition (5.8) holds for all semistable holomorphic structures with Seshadri filtrations of length  $< r$ . Let

$$0 = E_0 \subset E_1 \subset \dots \subset E_r = E$$

be a Seshadri filtration of  $D''$ . By our inductive hypothesis, we can approximate  $D''|_{E_{r-1}}$  by stable holomorphic structures. Thus we can approximate  $D''$  by holomorphic structures that preserve a two-step Seshadri filtration. Our result follows by our inductive hypothesis.

(5.9) **Lemma.** *Suppose  $D$  and  $\tilde{D}$  are Yang-Mills,  $D_i$  Yang-Mills stable,  $D_i \xrightarrow{L_1^2} D$ ,  $\tilde{D}_i$  is complex gauge equivalent to  $D_i$ , and  $\tilde{D}_i \xrightarrow{L_1^2} \tilde{D}$ . Then  $D$  is a real gauge equivalent to  $\tilde{D}$ .*

*Proof.* Assume that  $D$  is not gauge equivalent to  $\tilde{D}$ . First observe that we can find  $c \geq 1$  with the following property: There is  $\tau > 0$  such that if  $B \in L^2_1(T^*M \otimes u(e))$  with  $|D^*B|_{L^2} \leq \tau$  and  $|DB + B \wedge B|_{L^2} \leq \tau$ , then

$$(5.10) \quad |B|_{L^2} \leq c(2\tau + |B|_{L^4}).$$

This can be achieved as follows. Consider the operator

$$\mathcal{D} : L^2_1(T^*M \otimes u(E)) \rightarrow L^2_{-1}(T^*M \otimes u(E))$$

defined by the formula  $\mathcal{D}(B) = DD^*B + D^*(DB + B \wedge B) + B$ . By our assumption,  $\mathcal{D}(B) - B|_{L^2_{-1}} \leq c_1\{|D^*B|_{L^2} + |DB + B \wedge B|_{L^2}\} \leq c_1\tau$ . But, provided  $\tau$  is sufficiently small, we obtain

$$\begin{aligned} |B|_{L^2} &\leq c_2|\mathcal{D}(B)|_{L^2_{-1}} \\ &\leq c_2\{|DD^*B|_{L^2_{-1}} + |D^*(DB + B \wedge B)|_{L^2_{-1}} + |B|_{L^2_{-1}}\} \\ &\leq c\{2\tau + |B|_{L^2}\} \end{aligned}$$

for some  $c > 0$ , proving (5.10).

We pick  $\varepsilon$  as in (5.5) and let  $0 < \eta < \frac{1}{2} \min\{\varepsilon, \delta\}$ , where  $\delta > 0$  is such that the real slice theorem [8, Theorem (3.2)] holds in a  $\delta$ -neighborhood of  $D$ . We flow  $\tilde{D}_i$  along the heat equation until we obtain the connection  $D_i + \tilde{B}_i$ , where  $\tilde{B}_i \in L^2_1(T^*M \otimes u(E))$ ,  $D_i^*(\tilde{B}_i) = 0$ , and  $|\tilde{B}_i|_{L^4} = \eta/2c$ . This is possible by our assumption that  $D$  and  $\tilde{D}$  are not gauge equivalent for some  $\eta > 0$  as before. Since  $\tilde{D}_i$  is a minimizing sequence, the same is true for  $D_i + \tilde{B}_i$ , hence  $|D_i\tilde{B}_i + \tilde{B}_i \wedge \tilde{B}_i|_{L^2} \rightarrow 0$ . Since  $D_i \xrightarrow{L^2_1} D$ , it follows that  $|D\tilde{B}_i + \tilde{B}_i \wedge \tilde{B}_i|_{L^2} \rightarrow 0$  and  $|D^*\tilde{B}_i|_{L^2} \rightarrow 0$ . Therefore, according to (5.10), we obtain for  $i$  sufficiently large  $|\tilde{B}_i|_{L^2_1} \leq \eta < \varepsilon/2$ . By passing to a subsequence, if necessary, we may also assume that  $\tilde{B}_i \xrightarrow{L^2_1} \tilde{B}$  (weakly), where  $D + \tilde{B}$  is Yang-Mills and  $|\tilde{B}|_{L^4} = \eta/c$ . Summing up, so far we have connections  $D_i = D + B_i$  and  $\tilde{D}_i = D + \tilde{B}_i$ , where  $D$  and  $D_i$  are Yang-Mills stable,  $\tilde{D}_i$  is complex gauge equivalent to  $D_i$ ,  $B_i \xrightarrow{L^2_1} 0$ ,  $\tilde{B}_i \xrightarrow{L^2_1} \tilde{B}$  (weakly),  $|B_i|_{L^2_1} < \varepsilon$ ,  $|\tilde{B}|_{L^4} = \eta/c$ , and  $D^*\tilde{B} = 0$ . Now we are ready to apply (5.5) to obtain  $u_i$  and  $\tilde{u}_i$   $L^2$ -perpendicular to  $\mathcal{D} := \ker D$ , with

$$|u_i|_{L^2} \leq K|B_i|_{L^2_1}, \quad |\tilde{u}_i|_{L^2} \leq K|\tilde{B}_i|_{L^2_1}.$$

We set as in (5.5)

$$e^{u_i}(D + B_i) = D + A_i, \quad e^{\tilde{u}_i}(D + \tilde{B}_i) = D + \tilde{A}_i,$$

where there exists  $p_i \in \mathcal{P}$  such that

$$D + \tilde{A}_i = p_i(D + A_i).$$

Now we are almost done with the proof of (5.9). By choosing subsequences, we may assume that  $\tilde{u}_i \xrightarrow{L^2_2} u$  (weakly),  $\tilde{A}_i \xrightarrow{L^2_1} \tilde{A}$  (weakly),  $A_i \xrightarrow{L^2_1} 0$  (weakly), and  $D + \tilde{B} = e^u(D + \tilde{A})$ . We also may assume, by taking further subsequences, if necessary, that we can write  $p_i = \sum_{\alpha=1}^k \lambda_i^\alpha \pi_i^\alpha$ , where  $\lambda_i^1 > \dots > \lambda_i^k$ ,  $\pi_i^\alpha \in C^\infty(\text{End } E)$ ,  $\pi_i^\alpha = \pi_i^{\alpha*} = \pi_i^{\alpha^2}$ ,  $D''\pi_i^\alpha = 0$ , and  $\pi_i^\alpha \xrightarrow{L^2_2} \pi_\infty^\alpha$  with  $\text{rk}(\pi_i^\alpha) = \text{rk}(\pi_\infty^\alpha) = r(a) > 0$  ( $i = 1, 2, \dots$ ,  $\alpha = 1, \dots, k$ ). Since  $A_i \xrightarrow{L^2_1} 0$  and  $\tilde{A}_i \xrightarrow{L^2_1} \tilde{A}$  (weakly),  $\tilde{A}''$  has to be strictly upper triangular with respect to the filtration

$$0 \subset \pi_\infty^1(E) \subset \pi_\infty^1(E) \oplus \pi_\infty^2(E) \subset \dots \subset \pi_\infty^1(E) \oplus \dots \oplus \pi_\infty^k(E) = E.$$

But, since  $D + \tilde{B} = e^u(D + \tilde{A})$ , we have  $r(D + \tilde{B}) = r(D + \tilde{A}) = r(D) = [D]$ , hence  $D + \tilde{B}$  is gauge equivalent to  $D$ . But, since  $|\tilde{B}| < \delta$  and  $D^*\tilde{B} = 0$ , it follows that  $\tilde{B}$  has to be 0, contradicting  $|\tilde{B}|_{L^4} = \eta/c$ . This completes the proof.

(5.11) **Corollary.** *The map  $r: [\mathfrak{A}_s] \rightarrow [\mathfrak{R}_0]$  is continuous.*

Before we proceed further we need to establish the continuity of  $r$  on  $[\mathfrak{A}_{ss}^{\text{spl}}]$ , the set of holomorphic structures that split as a direct sum of stable ones. More precisely,  $D \in \mathfrak{A}_{ss}^{\text{spl}}$  if there is a complex gauge transformation  $g$  such that  $g(D)$  splits as a direct sum of stable connections (for the rank-2 case cf. [8, p. 54]). We start by proving a simple lemma which is of independent interest.

(5.12) **Lemma.** *Let  $\pi_i \in L_k^p(\text{End } E)$ ,  $\text{rk } \pi_i = s$ ,  $\pi_i^2 = \pi_i$ , and  $\pi_i \xrightarrow{L_k^p} \pi_\infty$  ( $pk > 2$ ). Then we can write  $\pi_i = u_i \tilde{\pi}_i u_i^{-1}$ , with  $u_i, \tilde{\pi}_i \in L_k^p(\text{End } E)$ ,  $u_i$  preserving the metric,  $u_i \xrightarrow{L_k^p} id$ , and  $\tilde{\pi}_i(E) = \pi_\infty(E)$ .*

*Proof.* It is easy to check  $\pi_\infty^2 = \pi_\infty$  and  $\text{rk } \pi_\infty = s$ . We may assume that the metric on  $E$  is chosen so that  $\pi_\infty$  is an orthogonal projection. The problem is local, and we may assume that the bundles  $E$ ,  $F := \pi_\infty(E)$ , and  $F^\perp := (1 - \pi_\infty)(E)$  are all product bundles of dimensions  $n$ ,  $s$ , and  $n - s$ , respectively. Let  $\{e_\infty^1, \dots, e_\infty^n\} = \vec{e}_\infty$  be a  $C^\infty$  global orthonormal frame of  $E$  such that  $\{e_\infty^1, \dots, e_\infty^s\}$  and  $\{e_\infty^{s+1}, \dots, e_\infty^n\}$  form corresponding frames for  $F$  and  $F^\perp$ . By the inclusion  $L_k^p \subset C^0$ ,

the frames  $\{\pi_i(e_\infty^1) =: f_\infty^1, \dots, \pi_i(e_\infty^s) =: f_\infty^s\}$  and  $\{(1 - \pi_i)(e_\infty^{s+1}) =: f_\infty^{s+1}, \dots, (1 - \pi_i)(e_\infty^n) =: f_\infty^n\}$  are global  $L_k^p$  frames for  $F_i := \pi_i(E)$  and  $\tilde{F}_i := (1 - \pi_i)E_i$ . Let  $\vec{e}_i = \{e_i^1, \dots, e_i^n\}$  be the frame obtained by Gram-Schmidt orthonormalization of  $\vec{f}_i = \{f_i^1, \dots, f_i^n\}$ . Since  $\vec{f}_i$  is fiberwise very close to the orthonormal frame  $\vec{e}_\infty$ , the Gram-Schmidt procedure is  $C^\infty$ . Hence, by the composition lemma  $C^\infty \times L_k^p \rightarrow L_k^p$ , the assignment

$$u_i: \{e_\infty^1, \dots, e_\infty^n\} \rightarrow \{e_i^1, \dots, e_i^n\}$$

is  $L_k^p$  and  $u_i \xrightarrow{L_k^p} id$ . We set  $\tilde{\pi}_i = u_i^{-1}\pi_i u_i$  and observe that  $\tilde{\pi}_i(e_\infty^j) = u_i^{-1}\pi_i e_i^j = u_i^{-1}e_i^j = e_\infty^j$  for  $j \leq s$ . Thus  $\text{Im } \tilde{\pi}_i \supset \text{Im } \pi_\infty$ , which, together with the equalities  $s = \text{rk } \pi_i = \text{rk } \tilde{\pi}_i = \text{rk } \pi_\infty$ , proves the lemma.

(5.13) **Corollary.** *The map  $r: [\mathfrak{A}_{ss}^{\text{spl}}] \rightarrow [\mathfrak{N}_0]$  is continuous.*

*Proof.* We have to show that if  $D_i'', D_\infty'' \in \mathfrak{A}_{ss}^{\text{spl}}$  and  $D_i'' \xrightarrow{L_1^2} D_\infty''$ , then  $r(D_i'') \rightarrow r(D_\infty'')$ . If the  $D_i''$ 's are stable, then this follows from (5.9). The general case will follow by induction on the rank of  $E$ . If  $\text{rk } E = 1$ , our statement follows again from (5.9). Assume the result to be true for all bundles with  $\text{rank} < n$  and let  $E$  be a bundle of  $\text{rank} = n$ . Suppose that all the  $D_i''$ 's are split and choose  $\pi_i: E \rightarrow E$  such that  $\pi_i^2 = \pi_i = \pi_i^*$  and  $D_i''(\pi_i) = 0$ . By standard elliptic estimates,  $\pi_i \xrightarrow{L_2^2} \pi_\infty$ , and, by (5.12),  $\pi_i = u_i \tilde{\pi}_i u_i^{-1}$  with  $u_i \xrightarrow{L_2^2} id$  and  $\tilde{\pi}_i(E) = \pi_\infty(E)$ . Since  $D_i'' \xrightarrow{L_1^2} D_\infty''$ , it follows that  $u_i^{-1}(D_i'') \xrightarrow{L_1^2} D_\infty''$  and  $u_i^{-1}(D_i'')$  preserves the bundles  $\pi_\infty(E)$  and  $(1 - \pi_\infty)(E)$ . But then  $u_i^{-1}(D_i'')|_{\pi_\infty(E)} \xrightarrow{L_1^2} D_\infty''|_{\pi_\infty(E)}$  and  $u_i^{-1}(D_i'')|_{(1-\pi_\infty)(E)} \xrightarrow{L_1^2} D_\infty''|_{(1-\pi_\infty)(E)}$ ; hence, by our inductive hypothesis

$$r(D_i'') = r(u_i^{-1}(D_i'')) \rightarrow r(D_\infty''). \quad \text{q.e.d.}$$

As in the unstable case, it is useful to view  $r(D'')$  as limit points of one-parameter subgroups. Given  $D''$  semistable, there is a one-parameter subgroup  $\{e^{t\alpha}\}_{t \in \mathbb{R}}$  such that

$$\lim_{t \rightarrow \infty} e^{t\alpha}(D'') = D''_{sp},$$

with  $D''_{sp} \in \mathfrak{A}_{ss}^{\text{spl}}$  and  $r(D''_{sp}) = r(D'')$ . More specifically, if  $D''$  preserves the Seshadri filtration  $(*)$  of (5.2), then we can choose  $\alpha$  to be an

infinitesimal complex gauge transformation by setting

$$\alpha = r \cdot Id_{F_1} \oplus (r - 1)Id_{F_2} \oplus \cdots \oplus Id_{F_r},$$

where  $Id_{F_i}$  denotes the identity endomorphism on  $F_i$ ,  $i = 1, \dots, r$ . We will proceed to show that  $r$  is continuous on the whole stratum  $[\mathfrak{A}_{ss}]$ .

(5.14) **Lemma.** *Let  $D_i \in \mathfrak{A}_s$  and  $D_i \xrightarrow{L_1^2} D$ . Then  $r(D_i) \rightarrow r(D)$ .*

*Proof.* If  $[D] \in [\mathfrak{A}_s]$ , we are done by (5.13). If  $[D] \in [\mathfrak{A}_{ss}]$  choose a one-parameter subgroup as before such that

$$e^{t\alpha}(D) \xrightarrow{L_1^2} D_{sp} \quad (t \rightarrow \infty).$$

By definition,  $r(D) = r(D_{sp})$ . Let  $V$  be an open neighborhood of  $r(D)$ . According to (5.13) there is a neighborhood  $U$  of  $D_{sp}$  in  $[\mathfrak{A}]$  such that  $r(\tilde{U}) \subset V$ , where  $\tilde{U} = U \cap [\mathfrak{A}_{ss}^{spl}]$ . Choose  $T$  such that  $e^{T\alpha}(D) \in U$ . Since  $D_i \xrightarrow{L_1^2} D$ , we have that  $e^{T\alpha}(D_i) \xrightarrow{L_1^2} e^{T\alpha}(D)$  ( $i \rightarrow \infty$ ), hence we can find  $i_0$  such that, for all  $i \geq i_0$ ,  $e^{T\alpha}(D_i) \in U$ . Since  $D_i \in \mathfrak{A}_s$ , it follows that  $e^{T\alpha}(D_i) \in \mathfrak{A}_s$ ; hence  $e^{T\alpha}(D_i) \in \tilde{U}$ , which implies that, for  $i \geq i_0$ ,  $r(D_i) = r(e^{T\alpha}(D_i)) \in V$ . Since  $V$  was arbitrary,  $r(D_i) \rightarrow r(D)$ .

(5.15) **Lemma.** *The map  $r: [\mathfrak{A}_{ss}] \rightarrow [\mathfrak{A}_0]$  is continuous.*

*Proof.* Assume that we can find  $D^k \xrightarrow{L_1^2} D$  and  $r(D^k) \not\rightarrow r(D)$ . Since the critical set is compact, we can choose a subsequence, call it again  $D^k$ , so that

$$[D_\infty^k] := r(D^k) \rightarrow [D_\infty].$$

Choose  $D^{kl}$  stable such that  $D^{kl} \rightarrow D^k$  ( $l \rightarrow \infty$ ). According to (5.11),

$$r(D^{kl}) \rightarrow r(D^k) = [D_\infty^k] \quad (l \rightarrow \infty).$$

Let  $U$  and  $V$  be disjoint open neighborhoods of  $r(D)$  and  $[D_\infty]$ . For each  $k$ , choose  $l(k) \geq k$  such that  $r(D^{km}) \in V$  for all  $m \geq l(k)$ . For each  $k$  we can also choose  $m = m(k) \geq l(k)$  so that  $D^{km(k)}$  is stable and  $D^{km(k)} \xrightarrow{L_1^2} D$ . By (5.14),  $r(D^{km(k)}) \rightarrow r(D) \in U$  and  $r(D^{km(k)}) \in V$ . Hence for  $k$  large enough

$$r(D^{km(k)}) \in U \cap V = \emptyset,$$

a contradiction. This completes the proof. **q.e.d.**

We next extend  $r$  to a map  $r: [\mathfrak{A}_{ss1}^2] \rightarrow [\mathfrak{N}_0]$  as follows: Let  $D \in \mathfrak{A}_{ss1}^2$  and choose  $D_1$  complex gauge equivalent to  $D$  with  $D_1$  smooth. We set  $r(D) := r(D_1)$ . The following is a trivial corollary of Lemma (5.14).

(5.16) **Corollary.**  $r: [\mathfrak{A}_{ss1}^2] \rightarrow [\mathfrak{N}_0]$  is continuous.

We finally extend  $r$  to the whole moduli space  $r: [\mathfrak{A}_1^2] \rightarrow \bigcup_{\mu} [\mathfrak{N}_{\mu}]$  as follows: Let  $D = D_1 \oplus \dots \oplus D_r$ , where the  $D_i$ 's are semistable; we define

$$r(D) = r(D_1) \oplus \dots \oplus r(D_r).$$

Now assume  $[D] \in [\mathfrak{A}_{\mu 1}^2]$ . Let

$$[H_{\infty}] := [H(\cdot, \infty)]: [\mathfrak{A}_{\mu 1}^2] \rightarrow \prod_{i=1}^r [\mathfrak{A}_{ss1}^2(F_i)],$$

where  $[H]$  is the map of (3.9). We define  $r: [\mathfrak{A}_{\mu 1}^2] \rightarrow [\mathfrak{N}_{\mu}]$  by the formula

$$r := r|_{\prod_{i=1}^r [\mathfrak{A}_{ss1}^2(F_i)]} \circ [H_{\infty}].$$

(5.17) **Lemma.** If  $D'' \in \mathfrak{A}_{\mu 1}^2$  and  $g \in \mathfrak{g}^{C^2}$ , then  $r(gD'') = r(D'')$ .

Moreover,  $r$  is continuous on  $\mathfrak{A}_{\mu 1}^2$ .

*Proof.* Let  $(*)$  be a given filtration of type  $\mu$ ; we write  $E \simeq_{C^\infty} F_1 \oplus \dots \oplus F_r$  in the usual way. It is not difficult to verify that  $[H](gD'', \infty) = g_{sp}[H](D'', \infty)$ , where  $g_{sp}$  preserves  $E \cong F_1 \oplus \dots \oplus F_r$ . Hence,

$$r(g(D'')) = r(g_{sp}H(D'', \infty)) = r(H(D'', \infty)) = r(D'').$$

The continuity of  $r$  on  $[\mathfrak{A}_{\mu 1}^2]$  follows from (3.9) and (5.15). q.e.d.

After this long preparatory work we are ready to prove our main results.

(5.18) **Proposition.** If  $D_i'' \rightarrow D_{\infty}''$ , where  $D_i''$  belongs to the complex gauge orbit of  $D''$  and  $D_{\infty}''$  is Yang-Mills, then  $[D_{\infty}''] = r(D'')$ . Thus, there is exactly one minimizing limit point on each complex orbit.

*Proof.* Let  $D_i'' \rightarrow D_{\infty}''$  as above. Then  $r(D'') = r(D_i'') \rightarrow r(D_{\infty}'') = [D_{\infty}'']$ . Hence,  $r(D'') = [D_{\infty}'']$ . q.e.d.

As an immediate consequence we obtain the convergence of the heat flow at infinite time:

(5.19) **Corollary.** For any  $D''$  in  $\mathfrak{A}_1^2$ , the heat flow  $[\Phi(D'', t)]$  converges to  $r(D)$  as  $t \rightarrow \infty$ .

Finally we also have the following homotopy result.

(5.20) **Theorem.** For each  $\mu$  there is a deformation retract of the stratum  $[\mathfrak{A}_{\mu 1}^2]$  onto the critical set  $[\mathfrak{N}_{\mu}]$ .

*Proof.* According to Corollary (3.9) it is enough to prove our theorem for the semistable stratum. Let  $i$  denote the inclusion of  $[\mathfrak{N}_0]$  into  $[\mathfrak{A}_{ss}]$ .

We have to show that  $i \circ r$  is homotopic to the identity. The desired homotopy is given with the aid of the heat flow  $[\Phi(\cdot, \cdot)]$  as follows. Define

$$G: [\mathfrak{A}_{ss}] \times [0, \infty] \rightarrow [\mathfrak{A}_{ss}]$$

by  $G(D, t) = [\Phi(D, t)]$  if  $t \neq \infty$  and  $G(D, \infty) = r(D)$ . That  $G$  defines a homotopy between  $i \circ r$  and the identity map is a consequence of (5.17), (5.18), and remark (iv) following Proposition (4.1).

### 6. The equivalence of the two stratifications

In this section we compare the stratification  $\{\mathfrak{A}_\mu\}$  of  $\mathfrak{A}$  via stable bundles with the stratification of  $\mathfrak{A}$  obtained by the Morse function  $f$ . More precisely, show, as conjectured by Atiyah and Bott, that the two stratifications coincide. This follows quite easily from the results of §§4 and 5. Our original argument was making use of the distance decreasing property of the flow (4.5) restricted to a complex gauge orbit. However, such an argument is not necessary in view of the results proved in §5. The heat equation converges to a unique connection up to real gauge equivalence and this is certainly enough to imply the equality of the two stratifications.

We start with the following lemma.

(6.1) **Lemma.** *The Morse strata  $\mathfrak{C}_\mu$  are preserved by the complex gauge group.*

*Proof.* Let  $D_1 \in \mathfrak{C}_\mu$  and  $D_2 = g(D_1)$  for some complex gauge transformation  $g$ . By the convergence of the heat flow, we obtain, under the notation of §4,

$$\omega[D_1] = r(D_1) = r(D_2) = \omega[D_2].$$

Hence  $D_2 \in \mathfrak{C}_\mu$ , proving our claim.

(6.2) **Theorem.** *The Morse stratification of the Yang-Mills functional coincides with the stratification of Harder and Narasimhan. More precisely with the notation as in §§4 and 5 and for any  $\mu$ ,  $\mathfrak{A}_{\mu 1}^2 = \mathfrak{C}_\mu$ .*

*Proof.* Since  $\{\mathfrak{A}_{\mu 1}^2\}$  and  $\{\mathfrak{C}_\mu\}$  form partitions of  $\mathfrak{A}_1^2$ , it would be enough to show  $\mathfrak{A}_{\mu 1}^2 \subset \mathfrak{C}_\mu$ . Let  $D \in \mathfrak{A}_{\mu 1}^2$  and let  $[V_\mu]$  be the neighborhood of Proposition (4.12) with the property  $[V_\mu] \cap [\mathfrak{A}_{\mu 1}^2] \subset [\mathfrak{C}_\mu]$ . We flow  $D$  along the heat equation for time  $t$  until we obtain  $[D_t] \in [V_\mu]$ . Since also  $[D_t] \in [\mathfrak{A}_{\mu 1}^2]$ , from Proposition (4.12) it follows that  $[D_t] \in [\mathfrak{C}_\mu]$ . But  $D_t$  is complex gauge equivalent to  $D$ , hence, by Lemma (6.1),  $[D] \in [\mathfrak{C}_\mu]$ . Again Lemma (6.1) shows that  $D \in \mathfrak{C}_\mu$ . q.e.d.

Finally, by combining with the results of §5, we can state our main conclusion:

(6.3) **Theorem.** *Yang-Mills functional induces on the moduli space of metric connections a stable-unstable manifold stratification in the sense of Morse theory. Moreover, the gradient flow of the of the Yang-Mills functional defines a deformation retract of each stratum onto the critical set.*

### 7. Topological applications

In this section we will indicate how our methods can produce new information about the topology of the space of stable bundles. More precisely, we will specialize, as mentioned in the Introduction, to the case of bundles of rank 2 and Chern class 0 and we will compute some homotopy and cohomology groups of the moduli space of stable bundles. Although we are making use of the results in [1] on the topology of the gauge group, the spirit of our method differs from that of Atiyah and Bott. We are computing ordinary cohomology and homotopy instead of equivariant cohomology.

The main application of this section can be summarized in the following:

(7.1) **Theorem.** *Let  $\mathcal{M} = \mathfrak{A}_s/\mathfrak{g}^{\mathbb{C}}$  denote the moduli space of stable bundles of rank 2 and Chern class 0 over a Riemann surface of genus  $g \geq 3$ .<sup>2</sup> Let  $\mathcal{M}_0$  denote the subspace of  $\mathcal{M}$  consisting of bundles of trivial determinant. Then:*

- (i)  $\pi_1(\mathcal{M}) \simeq H_1(M, \mathbb{Z}), \pi_1(\mathcal{M}_0) \simeq 0,$
- (ii)  $\pi_2(\mathcal{M}) \simeq \pi_2(\mathcal{M}_0) \simeq \mathbb{Z} \oplus \mathbb{Z}_2,$
- (iii)  $\pi_i(\mathcal{M}) \simeq \pi_i(\mathcal{M}_0) \simeq \pi_{i-1}(\mathfrak{g}), 2 < i \leq 2(g - 1) - 2.$

Theorem (7.1) follows directly from the next proposition.

(7.2) **Proposition.** *Let  $\mathfrak{A}_s$  denote the set of stable holomorphic structures on a bundle of rank 2 and Chern class 0. Then the inclusion of  $\mathfrak{A}_s$  in  $\mathfrak{A}$  is a homotopy equivalence up to dimension  $2(g - 1) - 1$ . In particular,  $\pi_i(\mathfrak{A}_s) = 0$  if  $i \leq 2(g - 1) - 2$ .*

*Proof of Theorem (7.1) assuming (7.2).* Let  $\mathfrak{g}$  be the real gauge group as before, and let  $\bar{\mathfrak{g}} = \mathfrak{g}/u(1)$  be the quotient of  $\mathfrak{g}$  by its constant, central subgroup. Let  $\bar{\mathfrak{g}}^{\mathbb{C}}$  and  $\bar{\mathfrak{g}}^{\mathbb{C}} = \bar{\mathfrak{g}}^{\mathbb{C}}/\mathbb{C}^*$  denote the corresponding complexified groups. Since  $\bar{\mathfrak{g}}$  and  $\bar{\mathfrak{g}}^{\mathbb{C}}$  are homotopy equivalent and  $\bar{\mathfrak{g}}^{\mathbb{C}}$  acts freely on  $\mathfrak{A}_s$ , the moduli space  $\mathcal{M} = \mathfrak{A}_s/\bar{\mathfrak{g}}^{\mathbb{C}} = \mathfrak{A}_s/\bar{\mathfrak{g}}^{\mathbb{C}}$  is homotopy equivalent to  $\mathfrak{A}_s^{\bar{\mathfrak{g}}} := E\bar{\mathfrak{g}} \times_{\bar{\mathfrak{g}}} \mathfrak{A}_s$ . As usual  $E\bar{\mathfrak{g}} \rightarrow B\bar{\mathfrak{g}}$  denotes the unique up to homotopy equivalence universal bundle over the classifying space  $B\bar{\mathfrak{g}}$  of  $\bar{\mathfrak{g}}$ .

<sup>2</sup> The case  $g = 2$  is somewhat special. In fact, the moduli space is much more explicitly known in this case [23].

Let

$$(7.3) \quad \mathfrak{A}_s \rightarrow \mathfrak{A}_s^{\bar{g}} \rightarrow B\bar{g}$$

be the fibration associated to the universal one with fiber  $\mathfrak{A}_s$ . By applying Proposition (7.2) to the long exact sequence in homotopy associated to the fibration (7.3), we obtain

$$\pi_i(\mathfrak{A}_s^{\bar{g}}) \simeq \pi_i(B\bar{g}) \simeq \pi_{i-1}(\bar{g}).$$

Since  $\mathfrak{A}_s^{\bar{g}}$  is homotopy equivalent to the moduli space of stable bundles  $\mathcal{M}$ , we only have to compute  $\pi_i(\mathfrak{g})$ . Now the statements about  $\pi_i(\mathcal{M})$  follow directly by the long exact sequence in homotopy associated to the fibration  $U(1) \rightarrow \mathfrak{g} \rightarrow \bar{g}$  (cf. the discussion in [1] following (2.26)).

It is worth pointing out that the isomorphism  $\pi_i(\mathcal{M}) \simeq \pi_{i-1}(\mathfrak{g})$  for  $i \leq 2(n-1)(g-1)-2$  is induced via the boundary map in the long exact sequence in homotopy associated to the quotient fibration  $\mathfrak{A}_s \rightarrow \mathfrak{A}_s/\mathfrak{g}$  (recall that, by Theorem (5.20),  $\mathfrak{A}_s/\mathfrak{g}$  is homotopy equivalent to  $\mathcal{M}$ ). Thus, although for homotopy computations we do not need to introduce the space  $\mathfrak{A}_s^{\bar{g}}$ , we will use (7.3) to compute cohomology (cf. (7.10)).

In order to prove the statements about  $\mathcal{M}_0$ , consider the fibration

$$(7.4) \quad \mathcal{M} \xrightarrow{\det} J_0(M)$$

which associates to a given holomorphic bundle  $E$  its determinant line bundle  $\det(E) := \Lambda^n(E)$  in the Jacobian of  $M$  [1, §9]. Now if  $0$  is the trivial bundle in  $J_0(M)$ , the  $\mathcal{M}_0$  is nothing but the fiber  $\det^{-1}(0)$  of (7.4). Since  $J_0(M)$  is a torus and  $\det$  induces an isomorphism  $\det_* : \pi_1(\mathcal{M}) \rightarrow \pi_1(J_0)$  [1, §9], everything follows directly from the long exact sequence in homotopy associated to (7.4). q.e.d.

In order to prove (7.2) we will show that the complement of the stable stratum  $\mathfrak{A}_s$  in  $\mathfrak{A}$  can be stratified by locally closed submanifolds of  $\mathfrak{A}$  of real codimension at least  $2g-2$ . Let  $\mathfrak{A}_{ss}$  denote as before the semistable stratum. We have shown in §3 that the complement of  $\mathfrak{A}_{ss}$  in  $\mathfrak{A}$  can be stratified by locally closed submanifolds whose codimension, according to Riemann-Roch, is at least  $2g-2$ . We thus need only to stratify  $\mathfrak{A}_{ss} \setminus \mathfrak{A}_s$ .

In order to do so we recall from §5 that any semistable holomorphic bundle  $E$  preserves a holomorphic Seshadri filtration. Assuming that  $E$  is not stable the Seshadri filtration is nothing but a line subbundle

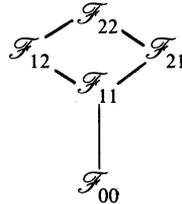
$$(*) \quad 0 \subset L \subset E.$$

As in the unstable case we will parametrize our stratification via the Seshadri filtration as follows:

(7.5) **Definition.** (i) Let  $D''$  be an *indecomposable* holomorphic structure. Assume  $D''$  is semistable and nonstable, and let  $L \subset E$  be the line subbundle of Chern class 0 preserved by  $D''$ . Let  $\mathcal{F}_{11}$  denote the set  $D''$  as before such that  $L$  is nonisomorphic to  $E/L$ , and let  $\mathcal{F}_{12}$  denote the set of  $D''$  such that  $L$  is isomorphic to  $E/L$ .

(ii) On the other extreme assume that  $D''$  splits as  $L_1 \oplus L_2$ . Let  $\mathcal{F}_{21}$  denote the set of  $D''$  such that  $L_1$  is nonisomorphic to  $L_2$ , and let  $\mathcal{F}_{22}$  denote the set of  $D''$  such that  $L_1$  is isomorphic to  $L_2$ .

(iii) We also set  $\mathcal{F}_{00}$  for the stable stratum. If on the set of indices  $ij$  as before ( $i, j = 0, 1, 2$ ) we put the lexicographic order, we obtain a partially ordered set:



The next lemma is the analogue of the result of Shatz [29] for the semistable case; cf. also Proposition (2.12).

(7.6) **Lemma.** *With the partial order on the set of indices  $\mu = ij$  defined as before,*

$$\overline{\mathcal{F}}_\mu \subset \bigcup_{\nu \geq \mu} \mathcal{F}_\nu.$$

*Proof.* Let  $\mu = ij$  and let  $D_i$  be a sequence of points in  $\mathcal{F}_{ij}$  converging to  $D$ . The openness of the stable strata (for various ranks) implies that  $D \in \mathcal{F}_\nu$ , where  $\nu = kl$  with  $k \geq i$ . If  $k > i$ , then we are done because  $\nu > \mu$ . If  $k = i$ , then  $l \geq j$  because the condition of two bundles being isomorphic to a closed condition. q.e.d.

Observe that from the last lemma it follows that  $\bigcup_{\nu \geq \mu} \mathcal{F}_\nu$  is a closed subspace of  $\mathcal{A}_{ss}$ . In fact we will make use of (7.6) in the proofs of (7.7) and (7.8) in this equivalent form. Now we are ready to state our “cell decomposition” theorem for  $\mathcal{A}_{ss}$ .

(7.7) **Theorem.** *The subspaces  $\mathcal{F}_\mu$ ,  $\mu = ij$  ( $i, j = 0, 1, 2$ ), are locally closed submanifolds of  $\mathcal{A}_{ss}$  of codimension at least  $2g - 2$ .*

*Proof of (7.2) assuming (7.7).* Let  $i \leq 2(g - 1) - 2$  and let  $f: S^i \rightarrow \mathcal{A}_s$  be a representative of a homotopy class in  $\pi_i(\mathcal{A}_s)$ . Since  $\pi_i(\mathcal{A}) = 0$ , by (7.2), there is a homotopy  $F$  of  $f$  to the constant loop in  $\mathcal{A}$ . However, since  $i + 1$  is less than the codimension of the submanifolds stratifying  $\mathcal{A} \setminus \mathcal{A}_s$ , we can deform  $F$  (cf. [10, Corollary on p. 73]) so that  $F$  misses

$\mathfrak{A}_{ss} \setminus \mathfrak{A}_s$ . In other words,  $f$  is homotopic to the constant loop in  $\mathfrak{A}_s$  and thus  $\pi_i(\mathfrak{A}_s) = 0$ . q.e.d.

In order to prove (7.7) we need to make some preparatory work. Let  $D''_0$  be Yang-Mills semistable and nonstable, and let  $L$  be a subbundle of  $E$  preserved by  $D''_0$ . We can use  $L$  to define a filtration

$$(*) \quad 0 \subset L \subset E$$

of  $C^\infty$  subbundles of  $E$ . Given  $\Sigma \in L^2_1(T^*M'' \otimes \text{End } E)$  we can decompose it with respect to  $(*)$  in the form

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

as in §3. The next lemma is the key to the proof of (7.7).

(7.8) **Lemma.** *Let  $D''_0$  be Yang-Mills semistable and nonstable, and let  $L$  be a subbundle of  $E$  preserving  $D''_0$ . Let  $(*)$  be as before the filtration defined by  $L$  and let  $\Sigma$  be an  $\text{End } E$ -valued  $(0, 1)$ -form satisfying:*

- (i)  $\Sigma$  is harmonic, i.e.,  $D'_0 \Sigma = 0$ ,
- (ii) with respect to the filtration  $(*)$

$$\Sigma = \begin{pmatrix} 0 & \beta \\ \gamma & 0 \end{pmatrix}$$

where both  $\beta$  and  $\gamma$  are nonzero.

Then, there exists  $\varepsilon = \varepsilon(D''_0) > 0$  such that for all  $\Sigma$  as above with  $|\Sigma|_{L^2_1} < \varepsilon$ :

Case (a)  $D''_0 + \Sigma \in \mathcal{F}_{00}$  if  $L$  is nonisomorphic to  $E/L$  with respect to  $D''_0$ .

Case (b)  $D''_0 + \Sigma \notin \mathcal{F}_{12} \cup \mathcal{F}_{22}$  if  $L$  is isomorphic to  $E/L$ .

*Proof.* We will deal with Case (a) first. Proof by contradiction: Suppose that there is a sequence

$$\Sigma_i = \begin{pmatrix} 0 & \beta_i \\ \gamma_i & 0 \end{pmatrix}$$

satisfying (i) and (ii) with  $|\beta_i|_{L^2_1} \rightarrow 0$  and  $|\gamma_i|_{L^2_1} \rightarrow 0$ , and  $D''_0 + \Sigma_i$  is not stable. Then, there are subbundles  $\pi_i$  of  $E$  such that  $\mu(\pi_i) = 0$ , and  $D''_0 + \Sigma_i$  preserves  $\pi_i$ . Thus, according to Lemma (5.6) and the footnote following it, we may assume after passing to a subsequence that  $\pi_i \rightarrow \pi$  in  $L^2_2$ , where  $D''_0(\pi) = 0$ . Since  $L$  is not isomorphic to  $E/L$ , the only bundles preserved by  $D''_0$  are  $L$  and  $E/L$ , and thus  $\pi(E) = L$  or  $(1 - \pi)(E) = L$ . Hence, by replacing  $\pi$  with  $(1 - \pi)$  if necessary, we may assume that  $\pi(E) = L$ . It is important to observe that the form of

$\Sigma_i$  does not alter under this change, since both  $\beta_i$  and  $\gamma_i$  were assumed to be nonzero.

Since  $\pi_i \rightarrow \pi$  in the  $L^2_2$ -topology, we may apply Proposition (5.12) to  $\pi_i$  to obtain holomorphic structures  $D''_i$  gauge equivalent to  $D''_0 + \Sigma_i$ ,  $D''_i \rightarrow D''_0$  in the  $L^2_1$ -topology, and  $D''_i$  preserves  $\pi(E) = L$ . Therefore, according to Proposition (5.5), we can choose  $p_i \in \ker D''_0$  such that  $p_i(D''_i) = D''_0 + p_i \Sigma_i p_i^{-1}$  preserves  $L$ . In other words if we express everything in terms of the filtration  $(*)$ , we obtain (after omitting the index  $i$ ) an equation of the form:

$$(7.9) \quad \begin{pmatrix} 0 & \beta \\ \gamma & 0 \end{pmatrix} = p^{-1} \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ 0 & \sigma_{22} \end{pmatrix} p, \quad p \in \ker D''_0.$$

But since  $L$  is nonisomorphic to  $E/L$ , the kernel of  $D''_0$  consists of gauge transformations of the form

$$p = \begin{pmatrix} p_1 & 0 \\ 0 & p_2 \end{pmatrix},$$

when  $p_1$  and  $p_2$  are constants. This contradicts (7.8) since  $\gamma \neq 0$ , and completes the proof of Case (a).

The proof of Case (b) is similar but with a few modifications to be pointed out. We first fix an identification of  $L$  with  $E/L$  such that the induced holomorphic structures from  $D''_0$  on  $L$  and  $E/L$  coincide under this identification. Again let  $\Sigma_i \rightarrow 0$  as before and assume that  $D''_0 + \Sigma_i$  belongs to  $\mathcal{F}_{12} \cup \mathcal{F}_{22}$ . Then there is a subbundle  $\pi_i$  preserving  $D''_0 + \Sigma_i$  as before and with the additional property that the induced holomorphic structures on the bundles  $\pi_i(E)$  and  $(1 - \pi_i)(E)$  are isomorphic. Again  $\pi_i \rightarrow \pi$  as before and we can gauge  $D''_0 + \Sigma_i$  to obtain a new holomorphic structure  $D''_0$  preserving  $\pi$  with  $D''_i \rightarrow D''_0$ . By gauging  $D''_i$  by an element of  $\ker D''_0$  we may assume  $\pi(E) = L$ . Finally, since we assumed that  $D''_0|_L = D''_0|_{E/L}$ , we may gauge  $D''_i$  so that also  $D''_i|_L = D''_i|_{E/L}$  and  $D''_i$  still converges to  $D''_0$ . Now we again apply (5.5) to obtain  $p_i \in \ker D''_0$  such that  $p_i(D''_0 + \Sigma_i) = D''_i$ . In other words, omitting the index  $i$  yields an equation of the form

$$(7.9') \quad \begin{pmatrix} 0 & \beta \\ \gamma & 0 \end{pmatrix} = p^{-1} \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ 0 & \sigma_{22} \end{pmatrix} p, \quad \sigma_{11} = \sigma_{22}.$$

By formally taking traces in (7.9') we get  $\sigma_{11} = \sigma_{22} = 0$ , and this clearly contradicts (7.9') since both  $\beta$  and  $\gamma$  are nonzero. q.e.d.

We now come to the proof of our main Theorem (7.7).

*Proof of (7.7).* We start with  $\mathcal{F}_{11}$  first. Let  $D'' \in \mathcal{F}_{11}$  and let

$$(*) \quad 0 \subset L \subset E$$

be the filtration preserved by  $D''$ . Let  $D''_0$  be the induced holomorphic structure on  $L \oplus E/L$ . Since the heat flow is continuous and converges to a unique real orbit (cf. Proposition (5.18) and Corollary (5.19)), we may flow down the complex orbit of  $D''$  so that  $D''_0$  is Yang-Mills and  $D''$  is arbitrarily close to  $D''_0$ . In other words, if we define a manifold structure of  $\mathcal{F}_{11}$  around  $D''$  close to the critical set, then we can use the flow to translate it everywhere.

Let  $\mathcal{P} = \ker D''_0 \subset L^2_2(\text{End } E)$ , let  $\mathcal{P}^\perp$  be the  $L^2$ -perpendicular subspace of  $\mathcal{P}$  in  $L^2_2(\text{End } E)$ , and let  $H^1(M, \text{End } E)$  denote the space of  $D''_0$ -harmonic 1-forms with coefficients in  $\text{End } E$ . Define

$$\varphi_0: \mathcal{P}^\perp \times H^1(M, \text{End } E) \rightarrow \mathfrak{A}$$

by  $\varphi_0(u, T) := e^u(D''_0 + T)$ . By the inverse function theorem we may assume that  $\varphi_0$  is a homeomorphism when restricted to an  $\varepsilon_0$ -neighborhood of  $0 \in \mathcal{P}^\perp \times H^1(M, \text{End } E)$ , where  $\varepsilon_0$  is less than the number  $\varepsilon(D''_0)$  of Lemma (7.8). Let  $U_0$  be the image of  $\varphi_0$  in  $\mathfrak{A}$ . As explained before we may assume that  $D'' = D''_0 + B \in U_0$ . With respect to the filtration (\*),

$$B = \begin{pmatrix} 0 & \beta \\ 0 & 0 \end{pmatrix}.$$

Clearly  $\beta \neq 0$ , since  $D''$  was assumed to be indecomposable, say,  $|\beta|_{L^2_1} = 2\varepsilon_1 > 0$ . Take  $U$  to be an  $\varepsilon_1$ -neighborhood of  $D''_0 + B = D''$  in  $U_0$  and define

$$\varphi: \mathcal{P}^\perp \times H^1(M, \text{End } E) \rightarrow \mathfrak{A}$$

by  $\varphi(u, T) := e^u(D'' + T)$ . By making  $\varepsilon_0$  and  $\varepsilon_1$  smaller, if necessary, we may assume by the inverse function theorem that  $\varphi$  is a homeomorphism into  $U$ , when restricted to an  $\varepsilon_0$ -neighborhood of  $0$  in  $\mathcal{P}^\perp \times H^1(M, \text{End } E)$ .

We now define relative to the filtration (\*) a projection

$$q: H^1(M, \text{End } E) \rightarrow H^1(M, \text{End } E)$$

by

$$q \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} = T_{21}.$$

We claim that the equation  $q = 0$  defines a chart of  $\mathcal{F}_{11}$  around  $D''$ . Indeed, let  $u \in \mathcal{P}^\perp$  and  $T \in H^1(M, \text{End } E)$  with  $|T|_{L^2_1} < \varepsilon_1$  and

$q(T) = 0$ . Then, by making  $\varepsilon_1$  smaller if necessary,  $D'' + T$  preserves  $L$  and induces nonisomorphic holomorphic structures on  $L$  and  $E/L$  (cf. Lemma (7.6)). Therefore,  $\varphi(u, T) = e^u(D'' + T) \in \mathcal{F}_{11}$ . Conversely, assume that  $D'' + T \in \mathcal{F}_{11}$  with  $|T|_{L^2_1} < \varepsilon_1$  and  $q(T) \neq 0$ . We will show that this is a contradiction. Indeed, write  $T = (T_{ij})$  as before. If  $T_{11} = T_{22} = 0$ , then our contradiction follows from (7.8) since  $T_{12} + \beta \neq 0$  and  $T_{21} = q(T) \neq 0$ . In general we can write

$$D'' + T = D''_0 + \begin{pmatrix} 0 & \beta + T_{12} \\ T_{21} & 0 \end{pmatrix} + \begin{pmatrix} T_{11} & 0 \\ 0 & T_{22} \end{pmatrix}$$

and, again by (7.8),  $D''_0 + \begin{pmatrix} 0 & \beta + T_{12} \\ T_{21} & 0 \end{pmatrix}$  is stable. Again, by shrinking  $U$  if necessary and recalling that the stable stratum is open, we obtain a contradiction.

The proof that  $\mathcal{F}_{12}$  is a submanifold is completely analogous. Again let  $D'' \in \mathcal{F}_{12}$  preserve a filtration  $(*)$ , let  $D''_0$  be the induced structure on  $L \oplus E/L$ , and assume as in Lemma (7.8) that  $D''_0|_L = D''_0|_{E/L}$  under a fixed identification of  $L$  with  $E/L$ . We proceed as before with the only difference being that in this case

$$q \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} = \begin{pmatrix} T_{11} - T_{22} & 0 \\ T_{21} & 0 \end{pmatrix}.$$

Indeed, we will have to show that the equation  $q = 0$  defines a chart of  $\mathcal{F}_{12}$  around  $D''$ . Clearly if  $q(T) = 0$ , then  $\varphi(u, T) = e^u(D'' + T) \in \mathcal{F}_{12}$ . Conversely assume  $q(T) \neq 0$ . If  $T_{21} = 0$ , then  $T_{11} \neq T_{22}$ , and  $D'' + T$  cannot belong to  $\mathcal{F}_{12}$  since it does not induce isomorphic structures to  $L$  and  $E/L$ . If  $T_{21} \neq 0$ , then everything follows as in the case of  $\mathcal{F}_{11}$  by Lemmas (7.6) and (7.8).

Finally, in order to show that  $\mathcal{F}_{21}$  and  $\mathcal{F}_{22}$  are submanifolds we start with  $D''$  in  $\mathcal{F}_{21}$  or  $\mathcal{F}_{22}$ . By proceeding as in the previous cases we can in fact assume that  $D'' = D''_0$  is Yang-Mills. Let  $\varphi_0$  be a coordinate chart of  $\mathcal{A}$  around  $D''_0 = D''$  as before. Then the map  $q$  defining the strata  $\mathcal{F}_{21}$  and  $\mathcal{F}_{22}$  is given respectively by

$$q \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} = \begin{pmatrix} 0 & T_{12} \\ T_{21} & 0 \end{pmatrix},$$

$$q \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} = \begin{pmatrix} T_{11} - T_{22} & T_{12} \\ T_{21} & 0 \end{pmatrix}.$$

The details are along the same lines described before and are left to the reader. This completes the proof of Theorem (7.7) and thus also the proofs of (7.1) and (7.2). q.e.d.

We would like to note that as an immediate consequence of our Proposition (7.2) we obtain Kirwan’s formula about the Poincaré series of the moduli spaces of stable bundles up to dimension  $2g - 3$  [18, (5.9) and (5.11)].<sup>3</sup> Indeed, since as we have seen in the proof of (7.2) the moduli space  $\mathcal{M}$  is homotopy equivalent to  $\mathfrak{A}_s^{\bar{g}}$  and the inclusion of  $\mathfrak{A}_s$  and  $\mathfrak{A}$  is a homotopy equivalence up to dimension  $2g - 3$ , we obtain the following equality on the Poincaré series:

$$\begin{aligned} P_t(\mathcal{M}) &= P_t(\mathfrak{A}^{\bar{g}}) \pmod{t^{2g-3}} \\ &= P_t(B\bar{g}) \pmod{t^{2g-3}}. \end{aligned}$$

Moreover, the fibration  $BU(1) \rightarrow Bg \rightarrow B\bar{g}$  is trivial over the rationals [1, §§2 and 9]; hence we obtain via the formula (2.9) in [1] for  $P_t(Bg)$

$$(7.10) \quad P_t(\mathcal{M}) = \frac{(1+t)^{2g}(1+t^3)^{2g}}{(1-t^2)(1-t^4)} \pmod{t^{2g-3}}.$$

We will finish by making a few remarks concerning the higher rank case. It is very natural to ask whether (7.1), (7.2), and (7.7) generalize for higher ranks. In this direction it is natural to conjecture the following:

**I.** Let  $\mathcal{M}$  denote the moduli space of stable bundles of rank  $n \geq 3$  and Chern class  $k$ . Let  $\mathcal{M}_0$  denote the subspace of  $\mathcal{M}$  consisting of bundles of fixed determinant. Then:

- (i)  $\pi_1(\mathcal{M}) \simeq H_1(M, \mathbb{Z}), \pi_1(\mathcal{M}_0) \simeq 0,$
- (ii)  $\pi_2(\mathcal{M}) \simeq \pi_2(\mathcal{M}_0) \simeq \mathbb{Z} \oplus \mathbb{Z}_p,$  where  $p = \text{g.c.d.}(n, k),$
- (iii)  $\pi_i(\mathcal{M}) \simeq \pi_i(\mathcal{M}_0) \simeq \pi_{i-1}(\mathfrak{g}), \quad 2 < i \leq 2(n-1)(g-1) - 2.$

**II.** The inclusion of the stable stratum  $\mathfrak{A}_s$  in  $\mathfrak{A}$  is a homotopy equivalence up to dimension  $2(n-1)(g-1) - 1$ . In particular,  $\pi_i(\mathfrak{A}_s) = 0,$  if  $i \leq 2(n-1)(g-1) - 2.$

**III.** The complement of  $\mathfrak{A}_s$  in  $\mathfrak{A}$  can be stratified by locally closed submanifolds of codimension at least  $2(n-1)(g-1).$

As in the rank-2 case III clearly implies II, which in turn implies I. Unfortunately, we cannot prove III except in the case  $n = 3$ . When  $n \geq 4,$  although all the analytical arguments go through, there does not seem to be a simple way of parametrizing the different strata. However, in an upcoming joint paper with K. Uhlenbeck [4] we will use alternative

---

<sup>3</sup> There is a slight error in Kirwan’s formulas [18, (5.9)]. There are no  $t^4$ -terms in the expression of  $q(t)$ . As very kindly explained to us by Frances Kirwan, a missing term from the formula for  $p(t)$  at the end of p. 259 cancels out the  $t^4$ -terms in the expression for  $A(t)$  and  $B(t)$  on p. 263 and hence there are no  $t^4$ -terms in the formula for  $q(t)$  [19].

techniques to bypass this difficulty and prove I and II directly without relying on III.

### Acknowledgments

The author is indebted to Karen Uhlenbeck for many useful discussions, as well as for her support and encouragement.

### References

- [1] M. F. Atiyah & R. Bott, *Yang-Mills equations on Riemann surfaces*, Philos. Trans. Roy. Soc. London A **308** (1982) 523–615.
- [2] R. Bott, *Nondegenerate critical manifolds*, Ann. of Math. (2) **60** (1954).
- [3] S. B. Bradlow, *Vortices on Kähler manifolds*, Ph.D. thesis, Chicago, 1988.
- [4] G. D. Daskalopoulos & K. K. Uhlenbeck, in preparation.
- [5] S. K. Donaldson, *A new proof of a theorem of Narasimhan and Seshadri*, J. Differential **18** (1983) 269–277.
- [6] —, *Antiselfdual connections over complex algebraic surfaces and stable vector bundles*, Proc. London Math. Soc. **50** (1985) 1–26.
- [7] —, *Infinite determinants, stable bundles and curvature*, Duke Math. J. **54** (1987) 231–247.
- [8] D. S. Freed & K. K. Uhlenbeck, *Instantons and four manifolds*, Math. Sci. Res. Inst. Publ. Vol. 1, Springer, Berlin, 1984.
- [9] P. Griffiths & J. Harris, *Principles of algebraic geometry*, Wiley, New York, 1978.
- [10] V. Guillemin & A. Pollack, *Differential topology*, Prentice-Hall, Englewood Cliffs, NJ, 1974.
- [11] V. Guillemin & S. Sternberg, *Geometric quantization and multiplicities of group representations*, Invent. Math. **67** (1982) 515–538.
- [12] R. S. Hamilton, *Harmonic maps of manifolds with boundary*, Lecture Notes in Math., Vol. 471, Springer, Berlin, 1975.
- [13] G. Harder & M. S. Narasimhan, *On the cohomology groups of moduli spaces of vector bundles on curves*, Math. Ann. **212** (1975) 215–248.
- [14] N. Hitchin, *The self duality equations on a Riemann surface*, Proc. London Math. Soc. (3) **55** (1982) 59–126.
- [15] G. Kempf, *Instability in invariant theory*, Ann. of Math. (2) **108** (1978) 299–317.
- [16] F. C. Kirwan, *Sur la cohomologie des espaces quotients*, C. R. Acad. Sci. Paris Sér. I **295** (1982) 261–264.
- [17] —, *Cohomology of quotients in symplectic and algebraic geometry*, Math. Notes, No. 31, Princeton University Press, Princeton, NJ, 1984.
- [18] —, *On the homology of compactifications of moduli spaces of vector bundles over a Riemann surface*, Proc. London Math. Soc. (3) **53** (1986) 237–266.
- [19] —, private communication.
- [20] M. Maruyama, *Stable vector bundles on an algebraic surface*, Nagoya Math. J. **58** (1975) 25–68.
- [21] J. Milnor, *Morse theory*, Annals of Math. Studies, No. 51, Princeton University Press, Princeton, NJ, 1975.
- [22] D. Mumford, *Geometric invariant theory*, Springer, Berlin and New York, 1965. (2nd ed., D. Mumford & J. Fogarty, Springer, Berlin, 1982.)

- [23] M. S. Narasimhan & R. Ramanan, *Moduli of vector bundles on a compact Riemann surface*, Ann. of Math. (2) **89** (1969) 14–51.
- [24] L. Ness, *A stratification of the null cone via the moment map*, Amer. J. Math. **106** (1984) 1281–1329.
- [25] C. Okonek, M. Schneider & H. Spindler, *Vector bundles on complex projective spaces*, Progr. in Math., Vol. 3, Birkhäuser, Boston, 1980.
- [26] R. S. Palais, *Morse theory on Hilbert manifolds*, Topology **2** (1963) 299–340.
- [27] —, *Foundations of global non-linear analysis*, Benjamin, New York, 1968.
- [28] C. S. Seshadri, *Space of unitary vector bundles on a compact Riemann surface*, Ann. of Math. (2) **85** (1967) 303–336.
- [29] S. S. Shatz, *The decomposition and specialization of algebraic families of vector bundles*, Compositio Math. **35** (1977) 163–187.
- [30] C. Simpson, *Systems of Hodge bundles and uniformization*, Ph.D. thesis, Harvard, 1987.
- [31] S. Smale, *Morse theory and a nonlinear generalization of the Dirichlet Problem*, Ann. of Math. (2) **80** (1964).
- [32] K. K. Uhlenbeck, *Connections with  $L^P$  bounds on curvature*, Comm. Math. Phys. **83** (1982) 31–42.
- [33] —, *Apriori estimates for Yang-Mills*, preprint.
- [34] K. K. Uhlenbeck & S. T. Yau, *On the existence of Hermitian-Yang-Mills connections in stable vectors bundles*, Comm. Pure Appl. Math. **39** (1986) S257–S293.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY