J. DIFFERENTIAL GEOMETRY
29 (1989) 499-544

THE GEOMETRY OF THE YANG-MILLS MODULI SPACE FOR DEFINITE MANIFOLDS

DAVID GROISSER & THOMAS H. PARKER

0. Introduction

The moduli space $\mathscr{M}$ of self-dual connections on a compact Riemannian 4-manifold carries a natural L^2 Riemannian metric. In [9] the authors explicitly computed this metric on the moduli space $\mathscr{M}_1(S^4)$ of self-dual $k = 1$ SU(2) connections on the standard 4-sphere. The result was a complete description of $\mathscr{M}_1(S^4)$ as a concrete Riemannian 5-manifold.[1] Its geometry turns out to be that of a slightly distorted hemisphere of S^5; in particular it has finite diameter and volume and its boundary is isometric to S^4 (up to a constant conformal factor $4\pi^2$). In this paper we examine the Riemannian geometry of the moduli space $\mathscr{M}$ of $k = 1$ self-dual SU(2)-connections on a general class of 4-manifolds: compact oriented simply-connected 4-manifolds M with positive-definite intersection form. For such manifolds (M, g) the moduli space $\mathscr{M}$ is (possibly after perturbing the metric g) a smooth 5-manifold except at the finite set of points $\{p_1, \cdots, p_n\} \in \mathscr{M}$ corresponding to the reducible self-dual connections [8]. The well-known result of Donaldson [6] asserts that there is a compact set $K \subset \mathscr{M}$ such that $\mathscr{M} - K$ is a disjoint union of $N + 1$ ends (Figure 1). One end—the collar of $\mathscr{M}$—is diffeomorphic to $(0, 1) \times M$. Each of the others is diffeomorphic to a cone on $\mathbb{C}P^2$ with vertex at a reducible connection $p_i \in \mathscr{M}$. The basic question of whether $\mathscr{M}$ has finite diameter and volume depends on the geometry of the ends.

The L^2 Riemannian metric on $\mathscr{M}$ is obtained from metrics on the infinite-dimensional spaces used in constructing $\mathscr{M}$. This construction is standard, and goes as follows (see [2], [8], [10] for details). Given a compact Lie group G and a principal G-bundle $P \to M$ we consider the affine space $\mathscr{A}$ of all smooth connections on P and the gauge group $\mathscr{G}$ of all automorphisms of P covering the identity. A connection $A \in \mathscr{A}$ is called *self-dual* if its curvature

Received November 20, 1987. The authors were partially supported by National Science Foundation grants DMS-8405661 (Groisser) and DMS-8603461 (Parker).

[1]This has been done independently by Doi, Matsumoto and Matumoto [5].

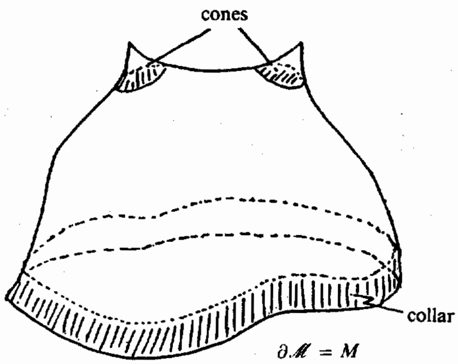

FIGURE 1

F_A is a self-dual 2-form ($F_A = *F_A$). The *moduli space* $\mathscr{M} = \mathscr{M}(P, g)$ *of self-dual connections* is the space $\mathscr{SD}/\Gamma$, where $\mathscr{SD} \subset \mathscr{A}$ is the set of self-dual connections.

When G is equipped with a bi-invariant metric h, the metrics g and h determine inner products on the spaces $\Omega^k(\operatorname{Ad} P)$ of k-forms with values in the vector bundle $\operatorname{Ad} P = P \times_{\operatorname{Ad}} \mathfrak{g}$ (a bundle of Lie algebras; $\mathfrak{g}$ is the Lie algebra of G). We can then define Riemannian metrics on the spaces $\mathscr{A}$, $\mathscr{A}/\mathscr{G}$ and $\mathscr{M}$, as follows. First, at each $A \in \mathscr{A}$ the canonical identification between the tangent space $T_A\mathscr{A}$ and $\Omega^1(\operatorname{Ad} P)$ gives an L^2 inner product on $\mathscr{A}$ (obtained by integrating the pointwise inner product against the Riemannian volume form) which is invariant under the action of $\mathscr{G}$ on $\mathscr{A}$. This $\mathscr{G}$-action is locally free on the open dense set $\mathscr{A}^*$ of irreducible connections and, by completing $\mathscr{A}$ and $\mathscr{A}/\mathscr{G}$ in appropriate Sobolev norms in the usual way, we can give $\mathscr{B}^* = \mathscr{A}^*/\mathscr{G}$ the structure of a Hilbert manifold. The L^2 metric on $\mathscr{A}$ then descends to a (weak) Riemannian metric on $\mathscr{B}^*$ by declaring $\mathscr{A}^* \to \mathscr{B}^*$ to be a Riemannian submersion (see [9, §2]). Finally, $\mathscr{M}^* = \mathscr{M} \cap \mathscr{B}^*$ is a finite-dimensional manifold (with singularities), and hence inherits a Riemannian metric by restriction. We denote this metric by $\mathscr{g}$.

This L^2 metric can also be described in terms of harmonic forms. For any self-dual connection A one has the "fundamental elliptic complex"

$$0 \to \Omega^0(\operatorname{Ad} P) \xrightarrow{d_A} \Omega^1(\operatorname{Ad} P) \xrightarrow{d_A^-} \Omega^2_-(\operatorname{Ad} P) \to 0, \tag{0.1}$$

where d_A is the exterior covariant derivative and d_A^- is d_A followed by the orthogonal projection p_- onto the space of anti-self-dual 2-forms. The principal stratum $\mathscr{M}'^*$ of $\mathscr{M}^*$ consists of the gauge orbits $[A]$ of those A for which $d_A : \Omega^0 \to \Omega^1$ is injective and $d_A^- : \Omega^1 \to \Omega^2_-$ is surjective. For such A, the tangent space $T_{[A]}\mathscr{M}'^*$ can be identified with the harmonic space

$\mathscr{H}_A = \{\omega \in \Omega^1(\operatorname{Ad} P) | d_A^* \omega = 0,\ d_A^- \omega = 0\}$. The metric on $\mathscr{M}'^*$ is simply the restriction of the L^2 metric to $\mathscr{H}_A$ (this is well-defined since the assignment $A \mapsto \mathscr{H}_A$ is $\mathscr{G}$-equivariant and $\mathscr{G}$ acts isometrically).

The analysis underlying these descriptions of the L^2 metric was described in detail in a previous paper [9, §§1,2]. We will generally adhere to the notation introduced in that paper, and will assume that the reader is familiar with the background presented there.

The L^2 metric on $\mathscr{M}$ is analogous to the Weil-Petersson metric on Teichmüller space, which had been studied extensively. In both cases it is difficult to make explicit statements about the Riemannian geometry because any such statement necessarily involves *global* analytic quantities on the original manifold. In fact, the metric on $\mathscr{M}$ is more complicated because it inherently depends on the metric on M, whereas the Weil-Petersson metric depends only on the topology of the underlying Riemann surface.

In this paper we examine the geometry of the moduli space near the ends depicted in Figure 1. The paper is divided into two parts, corresponding to the two types of ends—the cones and the collar.

The first two sections are devoted to studying the geometry of the cones. Our approach is to resolve the singularities of $\mathscr{M}$ using the "based moduli space" $\tilde{\mathscr{M}}$. Specifically, we fix a basepoint $x_0 \in M$ and consider the based gauge group $\mathscr{G}_0 = \{g \in \mathscr{G} | g(x_0) = \operatorname{Id}\}$ and its orbit space $\tilde{\mathscr{B}} = \mathscr{A}/\mathscr{G}_0$. In §1 we prove that $\tilde{\mathscr{B}}$ is a smooth Riemannian Hilbert manifold with an isometric SO(3)-action and that $\tilde{\mathscr{B}}^* \to \mathscr{B}^*$ is a Riemannian submersion. Restricting to self-dual connections gives a smooth Riemannian manifold $\tilde{\mathscr{M}}$ with an SO(3)-action whose orbit space is $\mathscr{M}$. The map $\tilde{\mathscr{M}} \to \mathscr{M}$ desingularizes the cones.

This viewpoint leads to a much more concrete picture of the geometry of the cones in $\mathscr{M}$. Questions which a priori involve global analysis (perturbation theory for Green operators, for example) are reduced to rather straightforward questions about *finite-dimensional* Riemannian geometry. We analyze this in §2, and obtain the following description of the metric and sectional curvatures of the cones.

Theorem I. *Let (M, g) be a compact oriented 1-connected 4-manifold with positive-definite intersection form, and let $(\mathscr{M}_k, \mathscr{g})$ be the moduli space of self-dual* SU(2) *connections on the bundle P over M with instanton number $k = \frac{1}{4}p_1(\operatorname{Ad} P) \geq 1$ (with its L^2 metric). Fix a reducible connection $[A] \in \mathscr{M}_k$. Let g_0 be the standard metric on $\mathbb{P} = \mathbb{CP}^{4k-2}$ (see Definition 2.7). Then there are a number r_0, a neighborhood U of $[A]$ in $\mathscr{M}_k$, and a diffeomorphism*

$$F\colon (0, r_0) \times \mathbb{P} \to U - \{[A]\},$$

which extends to a homeomorphism from the open cone $[0, r_0) \times \mathbb{P}$ *to* U. *In this coordinate system:*

(a) *The metric satisfies*

$$F^* \mathscr{g} = dr^2 \oplus r^2(g_0 + O(r^2)). \tag{0.2}$$

(b) *As* $r \to 0$, *the sectional curvatures* σ *of* $F^* \mathscr{g}$ *satisfy*

$$\begin{aligned} \sigma\left(\frac{\partial}{\partial r}, X\right) &= O(1), \\ \sigma(X, Y) &= \frac{3}{r^2} g_0(JX, Y) + O(1), \end{aligned} \tag{0.3}$$

where $X, Y \in T\mathbb{P}$ *and* J *is the complex structure on* $\mathbb{P}$.

This theorem shows that both the metric and the sectional curvatures of the cones in $\mathscr{M}$ are, to leading order, those of the standard cone on $\mathbb{P}$. The higher-order terms in the expansions (0.2) and (0.3) can be expressed in terms of the Green operators of certain Laplacians constructed from the connection A (see §2 for details).

The expansion (0.2) of the metric shows that U has finite volume and that the radial rays to $[A]$ in $\mathscr{M}_k$ have finite length. Thus the geometry of the cones is as depicted in Figure 1.

The second part of this paper is an analysis of the geometry of the collar when (M, g) is as in Theorem I and $k = 1$. In essence, our approach is to compare the local geometry of $\mathscr{M}$ in the collar with the corresponding geometry of the moduli space $\mathscr{M}_1(S^4)$, which was described in [9].

The collar consists of a self-dual connections ("instantons") whose energy densities $|F_A|^2$ are sharply concentrated bump-functions. Each such instanton A has unique center point $p(A) \in M$ and scale $\lambda(A) \in \mathbb{R}^+$ (cf. §4). These define a map

$$\Psi\colon \text{Collar of } \mathscr{M} \to 0(0, \lambda_0) \times M \tag{0.4}$$

which Donaldson [6, §III] has shown to be a diffeomorphism. The inverse map Ψ^{-1} provides a convenient coordinate chart which we use to describe the metric in the collar.

The most direct approach to computing the metric on $\mathscr{M}$ is to identify $T_{[A]}\mathscr{M}$ with the harmonic space $\mathscr{H}_A$ and evaluate the L^2-norms of these $\omega \in \mathscr{H}_A$. Of course, for a general manifold (M, g) we cannot explicitly solve for these ω. Instead, we define in §3 a set of 1-forms $\{\tilde{\omega}_A \in \Omega^1(\operatorname{Ad} P)\}$ which are approximately harmonic then $[A]$ is in the collar. The span of these $\{\tilde{\omega}_A\}$

defines a space $\hat{T}_A$ which approximates the tangent space $T_{[A]}\mathcal{M}$. We then compute the L^2 norms of the $\{\tilde{\omega}_A\}$ and obtain estimates—with explicit λ-dependence—on how good the approximation is.

Our choice of the approximation $\hat{T}_A$ is motivated by analogy with the case $M = S^4$. In that special case, T_A is *precisely* $\{\omega = i_Z F_A\}$, where Z is a conformal vector field on S^4 obtained by projecting a constant vector field on $\mathbb{R}^5$ onto TS^4 [9, Proposition 4.3]. As $\lambda = \lambda(A) \to 0$ the forms ω become concentrated around the center point of A. One expects a similar localization of the harmonic forms to occur on a general manifold (which has no conformal vector fields). Thus given a connection A with center $p \in M$ and scale size $\lambda \ll 1$, we use normal coordinates at p to define vector fields (obtained from the four coordinate vector fields and the radial vector field $r\partial/\partial r$ on $\mathbb{R}^4$) which are nearly conformal in a neighborhood of p. The approximately harmonic forms $\tilde{\omega}_A$ are then defined by contracting F_A with these vector fields. These forms have support near p, and hence the estimates on their L^2 norms (done in §3) are essentially local calculations.

§4 is devoted to the calculation of the differential of the coordinate chart Ψ^{-1}; we write it as an explicit bundle map $T((0, \lambda_0) \times M) \to T\mathcal{M}$ plus an error term, and keep track of the λ-dependence of the error term. These calculations parallel those of Donaldson [6, §III], but our purposes require considerably more detail (Ψ_*^{-1} has an expansion in powers of λ which we must compute to an additional order in λ). The improvements require combining the gauge theory with a certain amount of Riemannian geometry on M.

The results of §§3 and 4 are tied together in §5, where we construct an approximate inverse to Ψ_* and again estimate how good our approximation is. This enables us to conclude that the L^2 metric $\mathcal{g}$ on the collar is asymptotic to a product metric. Specifically, we prove:

Theorem II. *Let $\mathcal{h}$ denote the product metric $4\pi^2(2d\lambda^2 \oplus g)$ on $\mathbb{R} \times M$, and let Ψ be the collar map* (0.4). *Then $\mathcal{g} \sim \Psi^*\mathcal{h}$ as $\lambda \to 0$. More precisely, given $\varepsilon > 0$ there exists $\lambda_0 > 0$ such that for any $[A] \in \mathcal{M}$ with $\lambda(A) < \lambda_0$ and any $W \in T_{[A]}\mathcal{M}$,*

$$(0.5) \qquad (1-\varepsilon)\Psi^*\mathcal{h}(W,W) \leq \mathcal{g}(W,W) \leq (1+\varepsilon)\Psi^*\mathcal{h}(W,W).$$

Theorem II allows us to attach a geometric boundary to $\mathcal{M}$. To do this, we let $\overline{\mathcal{M}}$ be the completion of $(\mathcal{M}, \mathcal{g})$ as a metric space. In §5 we prove that the metric on $\mathcal{M}$ extends to a C^0 metric on $\overline{\mathcal{M}}$. Equation (0.5) then implies that the scale size λ is (asymptotically) proportional to the distance to $\partial\overline{\mathcal{M}}$. In particular, *the distance to the boundary is finite.* This observation immediately leads to several important conclusions about the metric space structure of the moduli space $(\mathcal{M}, \mathcal{g})$.

Theorem III. (a) $\overline{\mathscr{M}}$ *is compact, and hence has finite diameter and volume.*

(b) $\mathscr{M}$ *is incomplete.*

(c) *The distance function* $\rho([A]) = \operatorname{dist}([A], \partial\mathscr{M})$ *is asymptotic to* $(8\pi^2)^{1/2}\lambda$ *as* $\lambda \to 0$ (*i.e., the ratio of these functions approaches* 1, *uniformly in* $[A]$).

Part (c) above shows that $\rho([A])$, which is a function depending on the geometry of the moduli space, is essentially equivalent to the scale λ, which is a characteristic of the individual instantons on M. In particular, each instanton has a natural scale size $\rho([A])/\sqrt{8\pi^2}$ which is independent of the arbitrary choices (of cut-off function, etc.) involved in the local definition of λ given in §4.

Another implication of Theorem II is that the function λ on the collar extends smoothly to $\overline{\mathscr{M}}$, and that the boundary $\partial\overline{\mathscr{M}}$ is precisely the set $\{\lambda = 0\}$. Thus, formally, $\partial\overline{\mathscr{M}}$ consists of "instantons of scale zero". Such instantons are completely characterized by their center point $p \in M$. This suggests that $\partial\overline{\mathscr{M}}$ and M might be equal *as Riemannian manifolds.* The next theorem asserts that this is true, except for a factor of $4\pi^2$. It generalizes Corollary C of [9], which dealt with the case $M = S^4$.

Theorem IV. *The metric completion* $\overline{\mathscr{M}^*}$ *of the moduli space* $(\mathscr{M}^*, g)$ *is a compact singular manifold-with-boundary. Its singularities are the* (*isolated*) *cone points described in Theorem* I, *and its boundary* $\partial\overline{\mathscr{M}^*} = \partial\overline{\mathscr{M}} = \{\lambda = 0\}$ *is a smooth submanifold of* $\overline{\mathscr{M}}$ *isometric to* $(M, 4\pi^2 g)$.

If we truncate the cones of $\overline{\mathscr{M}}$ by removing the open neighborhoods U_i of the singular points we obtain a smooth, compact, Riemannian manifold-with-boundary which is a cobordism from M to a disjoint union of $\mathbb{C}P^2$'s. From this one can easily show that the intersection form of M is standard. This is almost exactly Donaldson's original proof of this fact [6]. However, the above theorems give a much sharper picture of the moduli space; in particular, they show that Donaldson's topological compactification of the moduli space is naturally implemented by the L^2 metric.

Finally, we briefly consider the moduli spaces of instantons on more general 4-manifolds and with instanton number $k > 1$. These moduli spaces are stratified manifolds. The various strata consist formally of multi-instantons, some of whose scales are zero. Thus they are similar to the boundary of the collar in Figure 1, except that they usually have high codimension and a rather complicated topological structure (as described in part in the work of Taubes and Donaldson). Nevertheless, much of the analysis of the second part of this paper should carry over to these general strata. One therefore expects Theorem III to be valid for general moduli spaces.

PART I. GEOMETRY OF THE CONES

1. The moduli space of the based gauge group

Reducible connections prevent the action of the gauge group $\mathscr{G}$ on $\mathscr{A}$ from being locally free, and consequently the orbit space $\mathscr{B}$ and the moduli space $\mathscr{M} \subset \mathscr{B}$ are usually not manifolds. In the next section we will examine the geometry of $\mathscr{M}$ in neighborhoods of the singular points—the cones of Figure 1. This section provides the analytical foundation on which that discussion is based. This analysis is of independent interest. It requires no assumptions on M, P, and G, save that the 4-manifold M be compact and oriented, and that the Lie group G be compact and semisimple.

From a topological perspective the singularities of $\mathscr{B}$ are best understood by fixing a basepoint $x_0 \in M$ and considering the normal subgroup $\mathscr{G}^0 \subset \mathscr{G}$ consisting of those automorphisms of P which restrict to the identity on P_{x_0}, the fiber over x_0. This "based" gauge group acts freely on $\mathscr{A}$, so $\mathscr{A} \to \tilde{\mathscr{B}} = \mathscr{A}/\mathscr{G}^0$ is a principal $\mathscr{G}^0$-bundle. On the other hand, the full gauge group $\mathscr{G}$ acts on $\mathscr{A}$ and the stabilizer at each *irreducible* connection is the center of $\mathscr{G}$. This center is the group $\mathscr{Z} = \Gamma(P \times_{\mathrm{Ad}} Z)$, where Z is the center of G ($\mathscr{Z} \cong Z$ is finite since G is semisimple). Thus the action of $\mathscr{G}$ on $\mathscr{A}$ induces an action of $\mathscr{G}/(\mathscr{Z} \times \mathscr{G}^0) \cong G/Z$ on $\tilde{\mathscr{B}}$ which is *free* on the open dense subset $\tilde{\mathscr{B}}^*$ of irreducible connections, and whose orbit space is $\mathscr{B} = \mathscr{A}/\mathscr{G}$. Restricting to the self-dual connections $\mathscr{S}\mathscr{D} \subset \mathscr{A}$, we get a moduli space $\tilde{\mathscr{M}} \subset \tilde{\mathscr{B}}$ with a G/Z-action whose orbit space is the usual moduli space $\mathscr{M} = \mathscr{S}\mathscr{D}/\mathscr{G} \subset \mathscr{B}$.

$$(1.1) \qquad \begin{array}{ccc} \tilde{\mathscr{M}} & \hookrightarrow & \tilde{\mathscr{B}} \\ \downarrow & & \downarrow \\ \mathscr{M} & \hookrightarrow & \mathscr{B} \end{array}$$

The singularities of $\mathscr{M}$ and $\mathscr{B}$ can then be described by studying this G/Z-action.

From an analytical perspective the situation is more complicated. There are standard "slice theorems" which show that $\mathscr{M}^*$ and $\mathscr{B}^*$ are Hilbert manifolds (cf. [2, §6] or [8, §3]), and it is frequently asserted that the same arguments show that $\tilde{\mathscr{M}}$ and $\tilde{\mathscr{B}}$ are manifolds. A closer examination reveals that the usual method for obtaining a slice of the $\mathscr{G}$-action—which makes use of the L^2 inner product on $T\mathscr{A}$—will not work for $\mathscr{G}^0$, essentially because $\mathscr{G}^0$ is not a closed subgroup in the L^2 topology. However, $\mathscr{G}^0$ *is* closed after completing in an appropriate Sobolev topology, and if we use this Sobolev metric to define slices, we do get a slice theorem. This yields smooth structures on $\tilde{\mathscr{M}}$ and $\tilde{\mathscr{B}}$, and the vertical arrows in (1.1) become Riemannian submersions with respect to the Sobolev Riemannian metrics on $\tilde{\mathscr{M}}$ and $\tilde{\mathscr{B}}$. Furthermore, with our specific choice of Sobolev norm, *the Sobolev metric and the L^2 metric on $\mathscr{M}$ are equal.* This key fact will enable us, in §2, to use the Sobolev metric on $\tilde{\mathscr{M}}$ to obtain information about the L^2 geometry of $\mathscr{M}$ near the singular points. In the remainder of this section we will carry out the construction just described, giving the details of the Hilbert space structure and the Sobolev metrics on the spaces in (1.1).

We begin by recalling the fundamental elliptic sequence (0.1) (which is a complex if and only if A is self-dual). The Laplacians $\Box_A^0 = d_A^* d_A + \mathrm{Id}$ and $\Box_A^1 = d_A d_A^* + 2(d_A^-)^* d_A^- + \mathrm{Id}$ are invertible for each $A_0 \in \mathscr{A}$. Fix a smooth connection A_0 and write $E = T^*M \otimes \mathrm{Ad}\, P$. For each integer $s \geq 0$ we define the Sobolev s-norm on $\Gamma(E)$ by

$$\langle \phi, \eta \rangle_{s;A_0} = \langle \phi, (\Box_{A_0}^1)^s \eta \rangle_{L^2}. \tag{1.2}$$

Remark. This definition extends to all real s as follows. By the generalized Hodge theorem [1] there is a complete orthonormal basis $\{\phi_k\}$ of $L^2(E)$ where each ϕ_k satisfies $\Box_{A_0}^1 \phi_k = \lambda_k \phi_k$ for some real positive eigenvalue λ_k. Hence we can expand any $\phi \in \Gamma(E)$ as $\sum a_k \phi_k$ and define

$$\|\phi\|_{s;A_0}^2 = \sum \lambda_k^s a_k^2. \tag{1.3}$$

This agrees with (1.2) for integral $s \geq 0$ (note that for any $s \in \mathbb{R}$ we can choose $k \in \mathbb{N}$ with $s \leq k$ and then $\|\phi\|_s^2 \leq \|\phi\|_k^2 = \langle \phi, \Box^k \phi \rangle$ is finite, so the sum in (1.3) converges).

Let $L_s^2(E)$ denote the completion of $\Gamma(E)$ in the Sobolev s-norm. By identifying $\mathscr{A} = T_{A_0}\mathscr{A} = \Gamma(E)$, we obtain a complete space of connections $\mathscr{A}_s$. One can similarly complete the gauge group to a group $\mathscr{G}_{s+1}$ (see [9, §1] for details). The results of Uhlenbeck [16, §1] imply that, for $s > 1$, $\mathscr{G}_{s+1}$ is a smooth Lie group acting smoothly on $\mathscr{A}_s$, and that the topologies on these spaces are independent of the choice of the connection A_0 used to define them. Furthermore, each $g \in \mathscr{G}_{s+1}$ is continuous (as a section of $P \times_{\mathrm{Ad}} G$) and $\mathscr{G}_{s+1}^0 = \{g \in \mathscr{G}_{s+1} | g(x_0) = \mathrm{Id}\}$ is a closed Lie subgroup of $\mathscr{G}_{s+1}$. The Lie algebra of $\mathscr{G}_{s+1}^0$ is $\mathfrak{g}_{s+1}^0 = \{X \in L_{s+1}^2(\mathrm{Ad}\, P) | X(x_0) = 0\}$. (We define the L_{s+1}^2 metric on $\Gamma(\mathrm{Ad}\, P)$ analogously to (1.2), with $(\Box_A^0)^{s+1}$ replacing $(\Box_A^1)^s$.)

We can eliminate the special role of the connection A_0 by introducing the natural L_s^2 Riemannian metric on $\mathscr{A}_s$. It is defined at $A \in \mathscr{A}$ by replacing A_0 in (1.2) or (1.3) by A. The Sobolev inequalities imply that for $s \geq 1$ the norms $\|\cdot\|_{s;A}$ on $\Gamma(E)$ are all locally uniformly equivalent and that $\|\cdot\|_s$ is a smooth Riemannian metric on $\mathscr{A}_s$.

The space $\mathscr{A}_s$ thus carries both a strong L_s^2 metric and a weak L^2 metric. Each defines a slice for the action of $\mathscr{G}_{s+1}$. Since the infinitesimal action of

$\mathscr{G}_{s+1}$ at $A \in \mathscr{A}$ is the map $d_A \colon L^2_{s+1}(\operatorname{Ad} P) \to L^2_s(E)$, the L^2 slice is defined as the L^2-orthogonal complement of $\operatorname{Im}(d_A)$. Thus this slice is

$$H_A = \{\eta \in L^2_s(E) | d^*_A \eta = 0\}. \tag{1.4}$$

Similarly, the L^2_s slice H^s_A is defined by

$$0 = \langle d_A X, \eta \rangle_s = \langle X, d^*_A (\Box^1_A)^s \eta \rangle_{L^2} \quad \forall X. \tag{1.5}$$

For general A this does not lead to an expression as simple as (1.4). However, when $F^-_A = 0$ we have $d^*_A \Box^1_A = \Box^0_A d^*_A$, so that (1.5) is equivalent to $\langle X, (\Box^0_A)^s d^*_A \eta \rangle = 0$; i.e. $H^s_A = \ker((\Box^0_A)^s \circ d^*_A)$. Because $\Box^0_A$ is invertible, this simplifies to

$$H^s_A = \ker(d^*_A).$$

Thus *the slices H_A and H^s_A coincide (independently of s) when A is self-dual.*

It is trickier to describe the corresponding slices for the action of the based gauge group $\mathscr{G}^0$. First consider the L^2_s slice $\tilde{H}^s_A$ for the action of $\mathscr{G}^0_{s+1}$ on $\mathscr{A}_s$. As above, $\eta \in \tilde{H}^s_A$ if and only if η satisfies (1.5) for all $X \in \mathfrak{g}^0_{s+1}$. Thus, for such η, there is some v in the fiber $(\operatorname{Ad} P)_{x_0}$ for which η solves the distributional equation

$$d^*_A (\Box^1_A)^s \eta = \delta_v, \tag{1.6}$$

where the delta function δ_v is defined by $\langle \delta_v, Y \rangle = (v, Y(x_0))$ (here $(\ , \ )$ is the inner product on $(\operatorname{Ad} P)_{x_0}$). To obtain a more useful and direct form of (1.6) we separately consider the cases where A is reducible and irreducible.

First suppose that A is irreducible, so the Laplacian $\Delta^0_A = d^*_A d_A$ on $\Gamma(\operatorname{Ad} P)$ is invertible. Let $G_A(x, y)$ denote the Green function of Δ^0_A; for distinct x, y this is a linear map from $(\operatorname{Ad} P)_x$ to $(\operatorname{Ad} P)_y$. Fixing $x = x_0$, each $v \in (\operatorname{Ad} P)_{x_0}$ thus determines a section $G^v_A(y)$ of $\operatorname{Ad} P$ which is smooth for $y \neq x_0$, is singular at $y = x_0$, and satisfies $d^*_A d_A G^v_A = \delta_v$. Hence (1.6) can be written as $d^*_A((\Box^1_A)^s \eta - d_A G^v_A) = 0$.

When A is reducible Δ^0_A is not invertible. In this case we can solve (1.6) only for certain $v \in (\operatorname{Ad} P)_{x_0}$. Indeed, if v satisfies (1.6) and $\Phi \in \ker(\Delta^0_A) = \ker(d_A)$, then $(v, \Phi(x_0)) = \langle (\Box^1_A)^s \eta, d_A \Phi \rangle = 0$, so v lies in the orthogonal complement to the subspace $K_{x_0} = \operatorname{span}\{\Phi(x_0) | \Phi \in \ker(d_A)\} \subset (\operatorname{Ad} P)_{x_0}$. Conversely, when $v \perp K_{x_0}$ we can solve (1.6) by modifying the argument in the preceding paragraph, as follows. Even when A is reducible Δ^0_A is invertible on the L^2-orthogonal complement of its kernel, and this inverse is given by convolution with the Green function defined by

$$G_A(x, y) = \sum_{\lambda_i > 0} \lambda_i^{-1} (\phi_i(x), \cdot) \phi_i(y), \tag{1.7}$$

where $\{\phi_i \in \Gamma(\operatorname{Ad} P)\}$ is an L^2-orthonormal basis of eigenfunctions with eigenvalues $\{\lambda_i\}$. Fixing $x = x_0$, we obtain sections $G^v_A(y)$ of $\operatorname{Ad} P$ as above; these are again singular at $y = x_0$ and smooth elsewhere. They satisfy $\Delta^0_A G^v_A = \delta_v - \sum(\Phi_j(x_0), v)\Phi_j$, where $\{\Phi_j\}$ is an L^2-orthonormal basis of $\ker(d_A)$. (The construction of such Green functions is standard; one obtains them by an integral transform of the heat kernel (as in [13, §2]), or by directly proving the convergence of (1.7) [3].) In particular, when $v \perp K_{x_0}$ we again have $\Delta^0_A G^v_A = \delta_v$. Thus for all connections A, the equation (1.6) defining the slice $\tilde{H}_A$ is equivalent to

$$d^*_A((\Box^1_A)^s \eta - d_A G^v_A) = 0 \tag{1.8}$$

for some $v \perp K_{x_0}$, where $K_{x_0} = \{0\}$ if A is irreducible and where G_A is defined by (1.7). Henceforth whenever we write G^v_A we assume $v \perp K_{x_0}$.

Now suppose that A is self-dual. Then $\Box^0_A d^*_A = d^*_A \Box^1_A$, so multiplying (1.8) by $(\Box^0_A)^{-s}$ gives

$$d^*_A(\eta - d_A(\Box^0_A)^{-s} G^v_A) = 0. \tag{1.9}$$

Hence *when A is self-dual,*

$$\tilde{H}^s_A = \operatorname{span}\{H_A, d_A(\Box^0_A)^{-s} G^v_A\} = H_A \oplus \operatorname{span}\{d_A(\Box^0_A)^{-s} G^v_A\} \tag{1.10}$$

and this splitting is L^2_s-orthogonal. Thus for self-dual A the L^2_s-horizontal slice for $\mathscr{A} \to \mathscr{B}$ is spanned by the Δ^0_A-harmonic forms (which are smooth) and the sections $d_A(\Box^0_A)^{-s} G^v_A$.

Remarks. (1) One can check that $d_A(\Box^0_A)^{-s} G^v_A \in L^2_p$ for $p < 2s - 1$ so, since $s > 1$, this space lies in the tangent space to $\mathscr{A}_s$, on which the L^2 metric is well defined. This is where we would run into trouble were we only to use the L^2 metric throughout. Formally, we would find $\tilde{H}_A = \operatorname{span}\{H_A, d_A G^v_A\}$, but $d_A G^v_A \notin L^2$ so the L^2 metric on $\tilde{H}_A$ would not be defined.

(2) If we let A vary through self-dual connections, then, as we pass through a reducible connection A_0 some of the sections G^v_A vary discontinuously; furthermore, $\dim(H_A \cap \ker d^-_A)$ jumps up by $\dim(\ker(d_{A_0}))$ and

$$\dim(\operatorname{span}\{d_A(\Box^0_A)^{-s} G^v_A\})$$

jumps down. Nevertheless, one can show that $\tilde{\mathscr{H}}^s_A = \tilde{H}^s_A \cap \ker(d^-_A)$ varies smoothly.

(3) When A is irreducible, $W^{s+1}_A = \operatorname{span}\{(\Box^0_A)^{-s} G^v_A\}$ is the L^2_{s+1}-orthogonal complement of $\mathfrak{g}^0_{s+1}$ in $\mathfrak{g}_{s+1}$. For if $\phi \in \mathfrak{g}^0_{s+1}$ we have

$$\langle \phi, (\Box^0_A)^{-s} G^v_A \rangle_{s+1} = \langle \phi, \Box^0_A G^v_A \rangle_{L^2} = \langle \phi, \delta_v \rangle = 0,$$

so W_A^{s+1} is orthogonal to $\mathfrak{g}_{s+1}^0$, and since $\operatorname{codim}(\mathfrak{g}_{s+1}^0) = \dim(\operatorname{Ad} P)_{x_0} = \dim(W_A^{s+1})$, W_A^{s+1} is a complement. When A is reducible W_A^{s+1} is still orthogonal to $\mathfrak{g}_{s+1}^0$, but is too small to be a complement. In this case $W_A^{s+1} \oplus \ker(d_A)$ is a (nonorthogonal) complement.

Theorem 1.1. *For each $s > 1$, $\tilde{\mathscr{B}}_s = \mathscr{A}_s/\mathscr{G}_{s+1}^0$ is a smooth Hausdorff manifold and $\mathscr{A}_s \to \tilde{\mathscr{B}}_s$ is a smooth fibration of Hilbert manifolds. There is a natural (strong) L_s^2 metric on $\tilde{\mathscr{B}}_s$ which makes this fibration a Riemannian submersion.*

Proof. We begin by showing that (1.8) defines a slice for the action of $\mathscr{G}_{s+1}^0$. Fix $A \in \mathscr{A}_s$ and consider the affine subspace $S_A = A + \tilde{H}_A^s \subset \mathscr{A}_s$. The group action gives a map

$$\Phi\colon \mathscr{G}_{s+1}^0 \times S_A \to \mathscr{A}_s,$$

which is smooth for $s > 1$ [16, Lemma 1.2]. Identifying $T_A S_A$ with $\tilde{H}_A^s$, the differential of Φ at (Id, A) is

$$D\Phi\colon \mathfrak{g}_{s+1}^0 \times \tilde{H}_A^s \to T_A\mathscr{A}_s,$$
$$(X, \eta) \mapsto d_A X + \eta.$$

Suppose that $d_A X + \eta = 0$. Then $\langle d_A X, d_A X\rangle_s = -\langle d_A X, \eta\rangle_s = 0$, implying $d_A X = 0$. In particular $|X| = \text{const} = 0$ since $X(x_0) = 0$. Hence $X = 0$, $\eta = 0$, and $D\Phi$ is injective.

To show that it is surjective, note that for every $A \in \mathscr{A}_s$, $s > 1$, there is a Poincaré inequality for $\mathfrak{g}_{s+1}^0$: there is a constant c such that $\|X\|_{L^2} \le c\|d_A X\|_{L_s^2}$ for every $X \in \mathfrak{g}_{s+1}^0$ (the proof is similar to that of [11, Theorem 3.6.5], using the facts that the embedding $L_{s+1}^2 \hookrightarrow C^0$ is compact and that $d_A X = 0$ implies $X = 0$ as above). This immediately extends to the inequality $\|X\|_{L_{s+1}^2} \le c'\|d_A X\|_{L_s^2}$, which implies that $d_A\mathfrak{g}_{s+1}^0 \subset L_s^2(E)$ is closed. Hence $L_s^2(E)$ is the orthogonal direct sum $d_A\mathfrak{g}_{s+1}^0 \oplus \tilde{H}_A^s$, so every element of $T_A\mathscr{A}_s = L_s^2(E)$ lies in the image of $D\Phi$. Therefore $D\Phi$ is an isomorphism, so by the inverse function theorem Φ is a local diffeomorphism of neighborhoods $\mathscr{O}_1 \times \mathscr{O}_2 \to \mathscr{O}_s$.

The argument for $\mathscr{O}_2$ injects into the quotient (i.e., that $g \cdot \mathscr{O}_2 \cap \mathscr{O}_2 = \varnothing$ for every $g \ne 1$ in $\mathscr{G}_{s+1}^0$) and that the quotient is Hausdorff proceeds exactly as in the slice theorem for $\mathscr{A} \to \mathscr{B}$ (see [2, §6], [8, §3], [10, §II.10], or [12, §4]).

Finally, given $[A] \in \tilde{\mathscr{B}}_s$, use the isomorphism $T_{[A]}\tilde{\mathscr{B}}_s \cong \tilde{H}_A^s$ (for any representative A of $[A]$) to pull back the L_s^2 metric from $\tilde{H}_A^s$ to $T_{[A]}\tilde{B}_s$. Our L_s^2 metric on $\tilde{\mathscr{B}}_s$ is gauge-invariant, so the result is independent of the choice of A and is well defined. By construction, $\mathscr{A}_s \to \tilde{\mathscr{B}}_s$ is then a Riemannian submersion. q.e.d.

As noted in the proof above, the L_s^2 metric on $\mathscr{A}_s$ is gauge-invariant, so the residual action of $\mathscr{G}/\mathscr{G}_0 \cong G/Z$ on $\tilde{\mathscr{B}}$ is isometric. This action is free on $\tilde{\mathscr{B}}^* \subset \tilde{\mathscr{B}}$. Thus we have a commutative diagram of Riemannian submersions,

$$(1.11) \qquad \begin{array}{ccc} \mathscr{A}_s^* & \xrightarrow{\ \pi_0'\ } & \tilde{\mathscr{B}}_s^* \\ {\scriptstyle \pi'}\big\downarrow & \swarrow{\scriptstyle \pi_1} & \\ \mathscr{B}_s^* & & \end{array}$$

where all spaces have their L_s^2 metrics; moreover π_1' is a principal G/Z-bundle.

We can now restrict this diagram to the self-dual connections. Let $\mathscr{A}' = \{A \in \mathscr{A} \,|\, \ker(d_A^-)^* = \{0\}\}$. An application of the implicit function theorem shows that $\mathscr{SD}_s' = \mathscr{SD}_s \cap \mathscr{A}_s'$ is a smooth Hilbert manifold for $s > 1$ (cf. [8, §3]). Writing $\mathscr{M}_s' = \mathscr{SD}_s'/\mathscr{G}_{s+1}$ and $\tilde{\mathscr{M}}_s = \mathscr{SD}_s'/\mathscr{G}_{s+1}^0$, we have a commutative diagram

$$(1.12) \qquad \begin{array}{ccc} \mathscr{SD}_s' & \xrightarrow{\ \pi_0\ } & \tilde{\mathscr{M}}_s' \\ {\scriptstyle \pi}\big\downarrow & \swarrow{\scriptstyle \pi_1} & \\ \mathscr{M}' & & \end{array}$$

where π_0 is a Riemannian submersion, and π, π_1 are submersions over the subspace $\mathscr{M}'^* = \mathscr{M}' \cap \mathscr{B}^*$.

Since $T_A(\mathscr{SD}') = \ker(d_A^-)$, we obtain slices for π and π_0 by intersecting H_A^s and $\tilde{H}_A^s$ with $\ker(d_A^-)$. Denoting these restricted slices by $\mathscr{H}_A^s$ and $\tilde{\mathscr{H}}_A^s$, we thus have identifications

$$(1.13) \qquad \begin{aligned} T_{[A]}\mathscr{M}_s'^* &= \mathscr{H}_A^s \stackrel{\text{def}}{=} H_A^s \cap \ker(d_A^-), \\ T_{[A]}\tilde{\mathscr{M}}' &= \tilde{\mathscr{H}}_A^s \stackrel{\text{def}}{=} \tilde{H}_A^s \cap \ker(d_A^-). \end{aligned}$$

These will be used frequently in later sections.

We now state the main theorem of this section. It describes the natural Riemannian metrics on the moduli space $\tilde{\mathscr{M}}'$.

Theorem 1.2. (a) *The differentiable structure on $\tilde{\mathscr{M}}_s'$ is independent of $s > 1$. This based moduli space $\tilde{\mathscr{M}}' = \tilde{\mathscr{M}}_s'$ is thus a well-defined smooth finite-dimensional manifold.*

(b) *$\tilde{\mathscr{M}}'$ carries an analytic family of smooth metrics $\{\mathscr{g}_s | s > 1\}$ induced by the L_s^2 norms. The action of G/Z on $\tilde{\mathscr{M}}'$ is isometric with respect to each $\mathscr{g}_s$.*

(c) *For each $s > 1$ there is a commutative diagram (1.12) where $\tilde{\mathscr{M}}_s' = \tilde{\mathscr{M}}'$. The map π_0 is a Riemannian submersion with respect to the L_s^2-induced* *metrics and, over* $\mathcal{M}'^*$, π *is a Riemannian submersion with respect to both the* L^2*- and* L^2_s*-induced metrics. For each* $s > 1$, π_1 *is a Riemannian submersion from* $(\tilde{\mathcal{M}}'^*, g_s)$ *to* $(\mathcal{M}'^*, g)$.

Proof. (a) For any $1 \le s \le s'$ the inclusions $\mathcal{A}_{s'} \hookrightarrow \mathcal{A}_s$ and $\mathcal{G}^0_{s'+1} \hookrightarrow G^0_{s+1}$ induce a smooth inclusion $\tilde{\mathcal{B}}_{s'} \to \tilde{\mathcal{B}}_s$ which restricts to a map $\iota\colon \tilde{\mathcal{M}}'_{s'} \to \tilde{\mathcal{M}}'_s$. Since every self-dual connection is gauge-equivalent to a smooth one, ι is a bijection and, given $[A] \in \tilde{\mathcal{M}}'_{s'}$, we can choose a smooth connection A representing both $[A]$ and $\iota([A])$. The slice theorem obtained in the proof of Theorem 1.1, together with the regular value theorem (applied to the function $A \mapsto F^-_A$), gives us smooth maps $\Phi_{s'}, \Phi_s$ from neighborhoods U' of $[A]$ in $\mathcal{SD}'_{s'}$ (respectively, U of $\iota([A])$ in $\mathcal{SD}'_s$) to $\tilde{\mathcal{H}}^{s'}_A$ (respectively $\tilde{\mathcal{H}}^s_A$). These define smooth local charts for $\tilde{\mathcal{M}}'_{s'}$ and $\tilde{\mathcal{M}}'_s$ around $[A]$ and $\iota([A])$. To prove that the smooth structures on $\tilde{\mathcal{M}}'_{s'}$ and $\tilde{\mathcal{M}}'_s$ are equivalent, it suffices to show that the overlap function $\tilde{\mathcal{H}}^{s'}_A \to \tilde{\mathcal{H}}^s_A$ is locally a diffeomorphism. To do this, let $\Psi_{s'}$ denote the restriction of $\Phi_{s'}$ to the slice $A + \tilde{\mathcal{H}}^{s'}_A$. Then $\Psi^{-1}_{s'}(\eta) = A + \eta$, so $\Psi^{-1}_{s'}$ is an affine map from the finite-dimensional vector space $\tilde{\mathcal{H}}^{s'}_A$ to $\mathcal{A}'_{s'} \subset \mathcal{A}'_s$. It is therefore bounded, and hence smooth, as a map from $\tilde{\mathcal{H}}^{s'}_A$ to $\tilde{\mathcal{H}}^s_A$. Since $\Phi_s\colon \mathcal{A}'_s \to \tilde{\mathcal{H}}^s_A$ is smooth it follows that the composition $\Phi_s \circ \Psi^{-1}_{s'}\colon \tilde{\mathcal{H}}^{s'}_A \to \tilde{\mathcal{H}}^s_A$ (i.e. the overlap map) is smooth. Since $\tilde{\mathcal{H}}^{s'}_A$ intersects the vertical space $d_A\mathfrak{g}^0_{s+1}$ trivially, the differential of this map at the origin is invertible. We conclude that the overlap map is locally a diffeomorphism.

(b) The L^2_s metric on $\tilde{\mathcal{M}}'_s$ (or $\mathcal{M}'^*_s$) is simply the restriction of the L^2_s metric on $\mathcal{A}_s$ to $\tilde{\mathcal{H}}^s_A$ (or $\mathcal{H}^s_A$). The action of $\mathcal{G}_{s+1}$ on $\mathcal{A}_s$ is L^2_s-isometric, so preserves the distribution $\tilde{\mathcal{H}}^s_A$ and induces a g_s-isometric action of $\mathcal{G}_{s+1}/\mathcal{G}^0_{s+1} = G/Z$ on $\mathcal{M}$.

Observe that by (1.10) and the paragraph following (1.5) we have $H^s_A = H_A$ (so $\mathcal{H}^s_A = \mathcal{H}_A$ is independent of s), and $\tilde{H}^s_A = H_A \oplus V^s_A$, where $V^s = \operatorname{span}\{d_A(\Box^0_A)^{-s}G^v_A\}$ and this splitting is L^2_s-orthogonal. Since $F^-_A = 0$, $V^s_A \subset \ker(d^-_A)$, so intersecting with $\ker(d^-_A)$ gives an L^2_s-orthogonal splitting $\tilde{\mathcal{H}}^s_A = \mathcal{H}^s_A \oplus V^s_A = \mathcal{H}_A \oplus V^s_A$. Furthermore, by the definition of $\Box^1_A$ and the L^2_s metric, $\langle \eta, \eta\rangle_s = \langle \eta, \eta\rangle_{L^2}$ for $\eta \in \mathcal{H}^s_A = \mathcal{H}_A$. Thus $g_s|_{\mathcal{H}_A}$ is independent of s. This means that the s-dependence of g_s is completely determined by the restriction to the fibers of $\tilde{\mathcal{M}}' \to \mathcal{M}'$ (in particular, $\pi_1\colon (\tilde{\mathcal{M}}'^*, g_s) \to (\mathcal{M}'^*, g)$ is a Riemannian submersion for any $s > 1$). These fibers are the orbits of G/Z, along which g_s is invariant. We claim that if we allow s to assume complex values, $\operatorname{Re}(s) > 1$, then $\{g_s\}$ is an analytic family, i.e. $g_s \in \operatorname{Sym}^2(T^*_{[A]}\tilde{\mathcal{M}}')\otimes\mathbb{C}$ depends analytically on s for each $[A] \in \tilde{\mathcal{M}}'$. To verify this, recall that at each irreducible $[A] \in \tilde{\mathcal{M}}'$, the g_s-vertical tangent space is spanned by

$\{d_A(\square_A^0)^{-s}G_A^v | v \in (\operatorname{Ad} P)_{x_0}\}$. This is also true for reducible $[A]$ provided we restrict $\square_A^0$ and G_A to the orthogonal complement of $\ker \square_A^0$, as in (1.7). In either case we have that, for $v, w \in (\operatorname{Ad} P)_{x_0}$,

$$\begin{aligned}\langle d_A(\square_A^0)^{-s}G_A^v, d_A(\square_A^0)^{-s}G_A^w\rangle_{L_s^2} &= \langle(\square_A^0)^{-s}G_A^v, d_A^*(\square_A^1)^s d_A(\square_A^0)^{-s}G_A^w\rangle \\ &= \langle(\square_A^0)^{-s}G_A^v, d_A^* d_A G_A^w\rangle \\ &= \langle(\square_A^0)^{-s}G_A^v, \delta_w\rangle = ((\square_A^0)^{-s}G_A^w(x_0), w)\end{aligned}$$

depends analytically on s (cf. [14, §8]).

(c) We have already shown that π_0 and π are Riemannian submersions of the L_s^2 metrics. The map π is an L^2-Riemannian submersion by the definition of the metric on $\mathscr{M}$, and we have just verified the statement about π_1.

2. Geometry near the reducible connections

In this section we will describe the L^2 metric on the moduli space in a neighborhood of a reducible connection. Our approach is to use the fibration $\tilde{\mathscr{M}} \to \mathscr{M}$ of Theorem 1.2 to reduce the problem to a calculation in finite-dimensional Riemannian geometry.

We will work on a principal $G = \mathrm{SU}(2)$ bundle over a 4-manifold M satisfying $b_1(M) = b_2^-(M) = 0$, but will allow any instanton number $k \geq 1$. Fix a Sobolev norm $s > 1$. By a theorem of Uhlenbeck [8, §3] we have, after perturbing the metric on M if necessary, that $\mathscr{M}'^* = \mathscr{M}^*$ (i.e. that $\ker d_A^- = \{0\}$ for all self-dual A). The moduli spaces $\tilde{\mathscr{M}}$ and $\mathscr{M}^*$ are then smooth, and $\mathscr{M} - \mathscr{M}^*$ consists of the gauge-equivalence classes of reducible self-dual connections whose holonomy reduces the bundle P to an S^1-bundle. The set of such reducible connections is in 1-1 correspondence with $\{u \in H^2(M; \mathbb{Z}) | u \cup u = 1\}$, and hence consists of a finite number of points in $\mathscr{M}$ (cf. [10, §4.3]). Our goal is to describe the geometry of $\mathscr{M}$ near these points.

The reducible connections can also be characterized in terms of the isometric action of $\mathscr{G}/(\mathscr{G}_0 \times \mathscr{Z}) \cong \mathrm{SO}(3)$ on $\tilde{\mathscr{M}}$ described in the previous section. The stabilizer at an irreducible point is trivial, so $\tilde{\mathscr{M}}^* \to \tilde{\mathscr{M}}$ is a principal SO(3)-bundle and $\dim \tilde{\mathscr{M}} = \dim \mathscr{M} + 3 = 8k$. At a reducible point $[A_0] \in \tilde{\mathscr{M}}$ the stabilizer is a circle $S^1 \subset \mathrm{SO}(3)$ which can be described as follows. Since A_0 is reducible there is a section $\Phi = \Phi^{A_0}$ of $\operatorname{Ad} P$ satisfying $\nabla^{A_0}\Phi = 0$. This Φ has constant length $(d|\Phi|^2 = 2(\Phi, \nabla\Phi) = 0)$ which we normalize to be 1. For convenience we use, in this section only, the metric on $\operatorname{Ad} P$ induced by $-\frac{1}{2}$ the Killing form of su(2), so the identity $[u,[v,w]] = (u,w)v - (u,v)w$ holds. For each unit vector $u \in \mathrm{su}(2)$ we then have $\exp(2\pi u) = -1$, and $\exp(tu) = \pm 1 \Leftrightarrow t$ is a multiple of 2π. (Although our metric here is *twice* that used in §§3–5 (where, as is customary, our metric is minus the trace form of the standard representation on $\mathbf{C}^2$), all the results of the present section—in particular Theorem 2.8 and Corollary 2.9—are completely independent of the choice of normalization.) For any $t \in \mathbf{R}$, $\exp(t\Phi)$ is a gauge transformation fixing A_0 and $\exp(t\Phi) \in \mathscr{Z} \Leftrightarrow t$ is a multiple of 2π. Hence $\exp(t\Phi)$, $0 \le t \le 2\pi$, projects to a circle in $\mathscr{G}/(\mathscr{G}_0 \times \mathscr{Z}) = \mathrm{SO}(3)$ which acts isometrically on $\tilde{\mathscr{M}}$ fixing the point $p = [A_0]$. We now make four observations.

(i) Since $g_t = \exp(t\Phi) \in \mathscr{G}$ takes ∇ to $g_t \circ \nabla \circ g_t^{-1} = \nabla - t\nabla\Phi + O(t^2)$, the circle action generated by Φ on $\mathscr{A}$ is given infinitesimally by the Killing vector field

$$\hat{\xi}(A) = -\nabla^A \Phi. \tag{2.1}$$

The corresponding circle action on the quotient $\tilde{\mathscr{M}}$ has an infinitesimal generator the Killing field

$$\xi([A]) = \pi_* \hat{\xi}(A).$$

(ii) The differential of this S^1 action at $p = [A_0]$ is the *isotropy representation* of S^1 on $T_p\tilde{\mathscr{M}}$. To compute it, first note that the differential of the S^1 action on $\mathscr{A}^s$ at A_0 is given by

$$B \mapsto \frac{d}{ds} g_t(\nabla^{A_0} + sB) g_t^{-1}|_{s=0} = (\operatorname{Ad} g_t)B \quad \text{for } B \in T_{A_0}\mathscr{A}^s.$$

This differential preserves the vertical subspace $\{d_{A_0}X | X \in L^2_{s+1}(\operatorname{Ad} P)$ satisfies $X(x_0) = 0\}$ since $(\operatorname{Ad} g_t) d_A X = d_A((\operatorname{Ad} g_t)X)$ with $((\operatorname{Ad} g_t)X)(x_0) = 0$. It must therefore preserve the L^2_s orthogonal complement (since S^1 acts isometrically). Thus if we identify $T_p\tilde{\mathscr{M}}$ with $\tilde{\mathscr{H}}_A$ as in (1.13), then the isotropy representation is

$$\operatorname{Ad} g_t \colon \tilde{\mathscr{H}}_A \to \tilde{\mathscr{H}}_A. \tag{2.2}$$

(iii) Writing $\operatorname{Ad} g_t = \exp(tJ)$, where $J = \operatorname{ad}\Phi$, we see that the infinitesimal generator of the isotropy representation is

$$J = [\Phi, \cdot] \colon \tilde{\mathscr{H}}_A \to \tilde{\mathscr{H}}_A. \tag{2.3}$$

(iv) The infinitesimal isotropy representation can also be described in terms of the local Riemannian geometry of $\tilde{\mathscr{M}}$ by linearizing the Killing vector field ξ at p as follows. Let $\{\Psi_s\}$ be the flow of ξ, choose $B \in T_p\tilde{\mathscr{M}}$, and extend B arbitrarily to a neighborhood of p. Then since $\xi(p) = 0$,

$$\begin{aligned} \frac{d}{dt}[(\Psi_s)_* B]_p &= (\Psi_s)_*(-L_\xi B)|_p = (\Psi_s)_*(\nabla_B \xi - \nabla_\xi B)|_p \\ &= (\Psi_s)_{*p}(JB), \end{aligned} \tag{2.4}$$

when J is defined by $J(B) = (\nabla_B^{\tilde{\mathscr{M}}}\xi)(p)$ (here $\nabla^{\tilde{\mathscr{M}}}$ is the Levi-Civita connection on $\tilde{\mathscr{M}}$). It then follows from the uniqueness theorem for ordinary differential equations that $(\Psi_s)_{*p} = \exp(sJ)$. Thus (2.3) is alternatively given by

$$J = \nabla^{\tilde{\mathscr{M}}}\xi \in \operatorname{End}(T_p\tilde{\mathscr{M}}). \tag{2.5}$$

(Indeed, one can check directly that (2.3) follows from (2.1) and (2.5).) Note that since the isotropy representation is orthogonal, J is skew-symmetric.

These observations reduce the original gauge theory problem to the following problem in pure Riemannian geometry. We are given an $8k$-dimensional Riemannian manifold $\tilde{W}$ ($\tilde{\mathscr{M}}$ in our application) on which SO(3) acts isometrically. We assume that the action is free except along a finite number of exceptional orbits $\{\mathscr{O}_i \subset \tilde{W}\}$ at each point of which the isotropy subgroup is S^1 (so each $\mathscr{O}_i$ is diffeomorphic to $\mathrm{SO}(3)/S_1 = S_2$). Let $\pi\colon \tilde{W} \to W$ denote the projection onto the orbit space and let $x_i = \pi(\mathscr{O}_i)$. Then there is a Riemannian metric on $W^* = W - \{x_i\}$ such that $\pi\colon \tilde{W}^* = \tilde{W} - \{\mathscr{O}_i\} \to W^*$ is a Riemannian submersion. We seek a description of the geometry of W near each x_i.

This geometry problem is solved by Theorem 2.8 below. Our answer will make use of the following geometric quantities. The infinitesimal SO(3)-action is a linear map L which associates to each $v \in \mathrm{SO}(3)$ a Killing vector field $L(v)$ on W; we will often write ξ_v for $L(v)$. The pointwise adjoint of L is an $\mathfrak{so}(3)$-valued 1-form L^* on W. For each $x \in W$, the operator $L^*L \in \operatorname{End}(\mathfrak{so}(3))$ is selfadjoint and nonnegative; it is strictly positive off $\bigcup \mathscr{O}_i$ and has a (simple) zero eigenvalue at each point $x \in \mathscr{O}_i$. (In our gauge theory problem L is the restriction of d_A to the Lie algebra $\mathfrak{g} = \mathfrak{so}(3) = \{(\Box_A^0)^{-s}G_A^v\}$, and L^*L is the restriction to $\mathfrak{g}$ of the Laplacian $d_A^*d_A$ of the fundamental elliptic complex (0.1).) The fact that the smallest eigenvalue of L^*L approaches zero as x approaches $\mathscr{O}_i$ will be central to the discussion below.

Fix a point $p \in \tilde{W}$ on an exceptional orbit $\mathscr{O} = \pi^{-1}(x_0)$. Let $v_p \in \mathfrak{so}(3)$ be a unit-length generator of its isotropy subgroup $S_p^1 \subset \mathrm{SO}(3)$. Then the infinitesimal action of S_p^1 is the Killing vector field $\xi = L(v_p)$ and, as above, the isotropy representation is generated by $J = \nabla\xi$. The local topology around $\mathscr{O}$ is described by the following well-known theorem.

Differentiable Slice Theorem. *Let G be a compact Lie group acting isometrically on a finite-dimensional Riemannian manifold $\tilde{W}$. Let G_p be the isotropy subgroup of a point $p \in \tilde{W}$ with orbit $\mathscr{O}_p$ (so G_p acts on the normal bundle $N_p\mathscr{O}$ by the isotropy representation). Then the exponential map is an equivariant diffeomorphism from $N_p\mathscr{O}$ to a neighborhood of $\mathscr{O}_p$ in $\tilde{W}$.*

This theorem is a simple consequence of the uniqueness theorem for solutions to ordinary differential equations.

To apply this we must first identify the isotropy representation of our SO(3) action.

Lemma 2.1. $J^2 = -\mathrm{Id}$. *Hence* $\dim \tilde{W}$ *is even and* $T_p\tilde{W} = T_p\mathscr{O} \oplus N_p\mathscr{O} \cong \mathbb{C} \oplus \mathbb{C}^{4k-1}$ *as* S^1_p*-representations.*

Proof. Since J is skew-symmetric we can choose a basis $\{e_l\}$ of $T_p\tilde{W}$ so that J is given by a matrix of the form

$$J = \begin{pmatrix} 0 & a_1 & & & & & & \\ -a_1 & 0 & & & & & & \\ & & \ddots & & & & & \\ & & & 0 & a_k & & & \\ & & & -a_k & 0 & & & \\ & & & & & 0 & & \\ & & & & & & \ddots & \\ & & & & & & & 0 \end{pmatrix}, \qquad a_i > 0.$$

The isotropy subgroup S^1_p acts linearly on $T_p\tilde{W}$ by $\exp(tJ)$; it follows that the a_i are integers and that the stabilizer of e_{2l-1} is $\mathbb{Z}/a_l\mathbb{Z} \subset S^1$ for $l \le k$, and the stabilizer of e_l, $l > 2k$, is S^1_p. Applying the Differentiable Slice Theorem with $G = S^1_p$, we see that for small $\varepsilon > 0$ the point $z = \exp_p(\varepsilon e_l) \in \tilde{W}$ has the same stabilizer as does e_l. These points z near p are of two types.

(i) If $z \notin \mathscr{O}_p$, then by hypothesis the action of SO(3)—and hence that of S^1_p—is free so the stabilizer of z is $1 \in S^1$.

(ii) If $z \in \mathscr{O}_p$, then $z = g \cdot p$ for some $g \in \mathrm{SO}(3)$ (not unique). Hence the stabilizer of z in SO(3) is $gS_p^{-1}g^{-1}$ and its stabilizer in S^1_p is $S^1_p \cap gS^1_p g^{-1} = 1$.

Thus the stabilizer of $z \ne p$ is always trivial. We conclude that J has no kernel and that each a_i is 1, so $J^2 = -\mathrm{Id}$. The lemma follows since $T_p\mathscr{O} = \{L_p(v) | v \in \mathfrak{so}(3)\}$ is an S^1_p-invariant subspace of $T_p\tilde{W}$, and hence so is $N_p\mathscr{O}$.

Remark. In our gauge theory application, Lemma 2.1 can also be proved by entirely analytic methods. This is done in the appendix.

We now describe the local structure of the fibration $\tilde{W} \to W$ near $\mathscr{O}$. For this, we use polar coordinates to identify the normal space $N_p\mathscr{O} - \{p\}$ with $(0, \infty) \times \Sigma_p$, where $\Sigma_p \cong S^{8k-3}$ is the unit sphere. The Differentiable Slice Theorem then implies that the exponential map

$$\begin{aligned} \tilde{F}_p = \exp_p \colon (0, \varepsilon) \times \Sigma_p &\to \tilde{U}^* \\ &= \{\exp_p X | X \in N_p\mathscr{O},\ 0 < |X| < \varepsilon\} \subset \tilde{W}^* \end{aligned} \tag{2.6}$$

is an S^1_p-equivariant diffeomorphism for ε sufficiently small. Since $N_p\mathcal{O} = \mathbb{C}^{4k-1}$ by Lemma 2.1, the quotient of Σ_p by S^1_p is diffeomorphic to $\mathbb{P} = \mathbb{C}P^{4k-2}$, and $\tilde{F}$ induces a diffeomorphism

$$F_p\colon (0,\varepsilon) \times \mathbb{P} \to U^* \subset W^*, \tag{2.7}$$

where $U^* = \pi\tilde{U}^*$ is open in W. (F_p gives the analog of a polar normal coordinate system on W^*.) Hence $U = U^* \cup \{x_0\}$ is a neighborhood of $x_0 = \pi(p)$ homeomorphic to a cone on $\mathbb{P}$ (diffeomorphic off the vertex).

Specializing to the case $\tilde{W} = \tilde{\mathcal{M}}$, we have arrived at the fact, first observed by Donaldson, that each irreducible connection $[A_0] \in \mathcal{M}$ has a neighborhood in $\mathcal{M}$ diffeomorphic to a cone on $\mathbb{C}P^{4k-2}$.

The next step is to write the metric g on W in the coordinate system (2.7), expressing it in terms of the metric $\tilde{g}$ on $\tilde{W}$ and the operator L. This is a very natural geometric problem—essentially a Gauss lemma on the orbit space—which does not seem to be in the literature.

The calculations are best done by introducing appropriate Jacobi fields. We will first describe a geodesic $\tilde{\gamma}$ in $\tilde{W}$ and its projection γ in W. Then, given a vector Y tangent to W at a point in γ, we will lift it to the corresponding point of $\tilde{\gamma}$, extend it to a Jacobi field $\tilde{K}$ along $\tilde{\gamma}$, and examine the projected vector field $K = \pi_*\tilde{K}$ along γ. In doing this we will let π denote both the projection $\tilde{W} \to W$ and the corresponding projection $(0,\varepsilon) \times \Sigma_p \to (0,\varepsilon) \times \mathbb{P}$ obtained from (2.6) and (2.7). We will also simply write $\tilde{F}$ for $\tilde{F}_p$, and F for F_p.

Given $(r,\tau) \in (0,\varepsilon) \times \mathbb{P}$ choose a vector $\tilde{\tau} \in \Sigma_p \subset N_p\mathcal{O}$ with $\pi\tilde{\tau} = \tau$ and consider the geodesic $\tilde{\gamma}(t) = \exp_p(t\tilde{\tau})$ in $\tilde{W}$. Its tangent vector field, which we denote by $\tilde{T}(t)$, satisfies $\tilde{T}(0) = \tau$. Observe that $\tilde{\gamma}$ is everywhere perpendicular to the Killing vector fields $\xi = L(v)$, $v \in \mathfrak{so}(3)$. (This is a more stringent condition for $t \neq 0$ than for $t = 0$, since for $t \neq 0$ the ξ's span a 3-dimensional space, while at $t = 0$ their span is only 2-dimensional.) This follows since (a) $\tilde{g}_p(\tilde{T},\xi) = 0$ because $\tilde{T} = \tau \in N_p\mathcal{O}$ at p, and (b) $\frac{d}{dt}\tilde{g}(\tilde{T},\xi) = \tilde{g}(\tilde{T}, \nabla_{\tilde{T}}\xi) = 0$ using the geodesic equation $\nabla_{\tilde{T}}\tilde{T} = 0$ and Killing's equation $\tilde{g}(X, \nabla_Y \xi) = -\tilde{g}(Y, \nabla_X \xi)$. Thus $\tilde{\gamma}$ is horizontal with respect to the Riemannian submersion $\tilde{W}^* \to W^*$, so projects to an arclength-parametrized geodesic $\gamma = \pi\tilde{\gamma}$ in W. (The curve γ is independent of the choice of $\tilde{\tau}$.) In particular, $F_*\frac{\partial}{\partial r} = \pi_*\tilde{\tau}$ has unit length in TW.

Now given $Y \in T_\tau\mathbb{P}$, the horizontal lift $\tilde{Y}$ of $\tilde{\tau}$ is the unique vector in $N_p = N_p\mathcal{O}$ satisfying

(i) $\pi_*\tilde{Y} = Y$,

(ii) $\tilde{Y} \in T_{\tilde{\tau}}\Sigma_p$, and

(iii) $\tilde{Y}$ is orthogonal to the infinitesimal action of S^1_p at $\tilde{\tau}$.
(In (ii) and (iii) we are identifying $T_{\tilde{\tau}}N_p$ with N_p.) Since the tangent space to the S^1_p-orbit through $\tilde{\tau} \in N_p$ is precisely $\mathbb{R} \cdot J\tilde{\tau} \subset N_p = T_{\tilde{\tau}}N_p$, conditions (i) and (iii) are equivalent to $(\tilde{Y}, \tilde{\tau}) = (\tilde{Y}, J\tilde{\tau}) = 0$.

Lemma 2.2. *With the above notation, the differential of F at $(r, \tau) \in (0, \varepsilon) \times P$ is given by*

$$F_{*(r,\tau)}(a, Y) = \pi_*(a\tilde{T} + (\exp_p)_{*r\tilde{\tau}}(r\tilde{Y})). \tag{2.8}$$

Proof. Let $\alpha(t)$ be a curve in $\mathbb{P}$ with $\alpha(0) = \tau$ and $\alpha'(0) = Y$. It has a unique lift to a horizontal curve $\tilde{\alpha}(t)$ in Σ_p with $\pi\tilde{\alpha} = \alpha$ and $\tilde{\alpha}(0) = \tilde{\tau}$. Differentiating the equation $F(r, \alpha(t)) = \pi \circ \exp_p(r\tilde{\alpha}(t))$ at $t = 0$ we obtain

$$F_{*(r,\tau)}(0, Y) = \pi_* \circ \exp_{p^*(r\tilde{\tau})}(r\tilde{Y}).$$

Similarly, one checks that $F_{*(r,\tau)}(1, 0) = \pi_*(\exp_p)_{(r\tilde{\tau})}(\tilde{\tau})$; we write this as simply $\pi_*\tilde{T}$. Taking a linear combination of these two formulas then yields (2.8). q.e.d.

To evaluate the last term in (2.8), we consider the family of geodesics

$$\tilde{\gamma}(s, t) = \tilde{F}(t, \tilde{\tau} + s\tilde{Y}) = \exp_p(t(\tilde{\tau} + s\tilde{Y})),$$

which are variations of the geodesic $\tilde{\gamma}(t)$ defined above. Then $\tilde{K}_Y(t) = \tilde{\gamma}_* \frac{\partial}{\partial s}|_{s=0}$ is the Jacobi field along $\tilde{\gamma}$ with initial conditions $\tilde{K}_Y(0) = 0$ and $(\nabla_{\tilde{T}} K_Y)(0) = \tilde{Y}$. It therefore satisfies the Jacobi equation

$$\nabla_{\tilde{T}}\nabla_{\tilde{T}}\tilde{K}_Y = \tilde{R}(\tilde{T}, \tilde{K}_Y)\tilde{T}, \tag{2.9}$$

where $\tilde{R}$ is the Riemannian curvature of $\tilde{W}$ (cf. [4, §1.4]). Its value at $t = r$ is

$$\tilde{K}_y(r) = (\exp_p)_{*(r\tilde{\tau})}(r\tilde{Y}). \tag{2.10}$$

Proposition 2.3. *The metric $\mathscr{g}$ of W satisfies*

$$F_p^*\mathscr{g} = dr^2 \oplus g_r, \tag{2.11}$$

where g_r is the metric on $\mathbb{P}$ given at $(r, \tau) \in (0, \varepsilon) \times \mathbb{P}$ by

$$g_r(Y, Y) = [\tilde{g}(\tilde{K}_Y, \tilde{K}_Y) - \tilde{g}(L^*\tilde{K}_Y, (L^*L)^{-1}L^*\tilde{K}_Y)]|_{\tilde{F}(r,\tau)}. \tag{2.12}$$

Proof. Since π is a Riemannian submersion, (2.8) and (2.10) imply

$$(F_p^*\mathscr{g})\left(a\frac{\partial}{\partial r} + Y, a\frac{\partial}{\partial r} + Y\right) = \tilde{g}(\mathrm{hor}(a\tilde{\tau} + \tilde{K}_Y), \mathrm{hor}(a\tilde{\tau} + \tilde{K}_Y)), \tag{2.13}$$

where hor denotes the component perpendicular to the SO(3)-orbits in $\tilde{W}$. We have already observed that $\tilde{T}$ is horizontal and of unit length. A basic fact about Jacobi fields is that the conditions $\tilde{K}(0) = 0$ and $(\nabla_{\tilde{T}}\tilde{K})(0) \perp \tilde{T}$

imply that $\tilde{K} \perp \tilde{T}$ for all t; hence $\tilde{K}_Y \perp \tilde{T}$. Since the vertical component of $\tilde{K}_Y$ is automatically perpendicular to the horizontal vector $\tilde{T}$, it follows that $\operatorname{hor}(\tilde{K}_Y) \perp \tilde{T}$. Thus the right-hand side of (2.13) is simply $a^2 + \tilde{g}(\operatorname{hor}(\tilde{K}_Y), \operatorname{hor}(\tilde{K}_Y))$. Finally, since the tangent space to the SO(3)-orbit at $\tilde{\gamma}(t)$ is image$(L) \subset T_{\tilde{\gamma}(t)}\tilde{W}$, the horizontal projection (for $t \neq 0$) is $\operatorname{hor} = I - L(L^*L)^{-1}L^*$. The proposition follows.

Now consider the term $f(t) = \tilde{g}(\tilde{K}_Y(t), \tilde{K}_Y(t)) = |\tilde{K}(t)|^2$ in (2.12). Its Taylor series near $t = 0$ is easily calculated by successively differentiating, using (2.9), and evaluating at $t = 0$ (see [4, §1.4] for details). We obtain $f(0) = f'(0) = 0$, $f''(0) = 2|\tilde{Y}|^2$, $f'''(0) = 0$, $f''''(0) = 8(\tilde{Y}, R(T, \tilde{Y})T)$, and hence

$$(2.14) \qquad |\tilde{K}(t)|^2 = t^2|\tilde{Y}|^2 - \tfrac{1}{3}t^4(\tilde{R}(\tilde{T},\tilde{Y})\tilde{Y},\tilde{T}) + O(t^5|\tilde{Y}|^2).$$

Unfortunately this procedure cannot be directly applied to the second term in (2.12) because it involves evaluating at $t = 0$, where $(L^*L)^{-1}$ is not defined. Therefore we next examine $L^*\tilde{K}$ as $t \to 0$.

Definition 2.4. At $p \in \mathscr{O} \subset \tilde{W}$, $L_p^*L_p \in \operatorname{End}(\mathfrak{so}(3))$ has a 1-dimensional kernel. Fix $v_p \in \ker L_p^*L_p$ with $|v_p| = 1$. For $q \in \tilde{W}$ near p let $\lambda_0(q) \leq \lambda_1(q) \leq \lambda_2(q)$ be the eigenvalues of $L_q^*L_q$ and let $v(q) \in \mathfrak{so}(3)$ be the unique λ_0-eigenvector with $|v(q)| = 1$ and $v(p) = v_p$. (Note that the functions $\lambda_0(q)$ and $v(q)$ are smooth.)

Lemma 2.5. *Let $v = v(\tilde{\gamma}(t))$, $\lambda_i(t) = \lambda_i(\tilde{\gamma}(t))$, and $L_t = L_{\tilde{\gamma}(t)}$. Then*

(a) $\lambda_1(0) = \lambda_2(0)$,

(b) $v_t = v_p + O(t^2)$ *as* $t \to 0$,

(c) $\lambda_0(t) = t^2 + O(t^3)$ *as* $t \to 0$.

Proof. (a) It is straightforward to check that $L_p^*L_p$ commutes with the adjoint action of the isotropy subgroup S_p^1 of $\mathfrak{so}(3)$. But S_p^1 acts irreducibly on the orthogonal complement of $\{\operatorname{span}(v_p)\}$ in $\mathfrak{so}(3)$, so Schur's lemma implies that $L_p^*L_p$ is a multiple of the identity on this subspace.

(b) For each $w \in \mathfrak{so}(3)$ consider the function $\phi(t) = (w, L_t^*L_tv_t) = \lambda_0(t)(w, v_t)$. Since $\lambda_0(t)$ is a smooth nonnegative function vanishing at 0 we have $\lambda_0(0) = \dot{\lambda}(0) = 0$ and hence $\dot{\phi}(0) = \dot{\lambda}_0(w, v) + \lambda_0(w, v)|_{t=0} = 0$, where a dot denotes d/dt. On the other hand, $\phi(t) = (L_tw, L_tv_t) = (\xi_w, \xi_{v_t})$, so

$$\dot{\phi}(0) = (\nabla_{\tilde{T}}\xi_w, \xi_{v_t}) + (\xi_w, \nabla_{\tilde{T}}\xi_{v_t}) + (\xi_w, \xi_{\dot{v}_t})|_{t=0}.$$

But $\xi_v(0) = 0$ and $\nabla_{\tilde{T}}\xi_v(0) = J(\tilde{T})$ by (2.5). Hence at $t = 0$

$$0 = (\xi_w, J(\tilde{T})) + (w, L^*L\dot{v}).$$

However, at $t = 0$, $\xi_w \in T_p\mathscr{O}$, $T \in N_p\mathscr{O}$ and J preserves $N_p\mathscr{O}$ (cf. Lemma 2.1). Thus $0 = (w, L^*L\dot{v})$ $\forall w$, so $\dot{v} \in \ker L = \operatorname{span} v_p$. By differentiating $|v_t|^2 = 1$ we also see that $(\dot{v}, v_p) = 0$, so $\dot{v}(0) = 0$ and (b) follows.

(c) Since $|v_t|^2 = 1$ we have $\lambda_0(t) = (v_t, L_t^* L_t v_t) = |\xi_{v_t}|^2$ and hence

$$\dot{\lambda}_0(t) = 2(\xi_{v_t}, \nabla_{\tilde{T}}\xi_{v_t} + \xi_{\dot{v}_t}).$$

Differentiating again and evaluating at $t = 0$ where $\xi_v = 0$ and $\dot{v}_t = 0$ yields

$$\ddot{\lambda}_0(0) = 2|\nabla_{\tilde{T}}\xi_v(0)|^2 = 2|J(\tilde{T})|^2 = 2|\tilde{T}|^2 = 2$$

(using Lemma 2.1). Then (c) follows by Taylor's theorem.

Lemma 2.6. *As $t \to 0$*

$$(L^*\tilde{K}, (L^*L)^{-1}L^*\tilde{K})(\tilde{\gamma}(t)) = \lambda_1^{-1}t^4|(\nabla_{\tilde{T}}L^*)(\tilde{Y})|^2 + O(t^5), \tag{2.15}$$

where λ_1 is the nonzero eigenvalue of L^*L at $t = 0$.

Proof. We first fix $w \in \mathrm{so}(3)$ and compute the Taylor series of the function

$$f_w(t) = (L^*\tilde{K}(\gamma(t)), w) = (\tilde{K}, \xi_w)(\gamma(t)).$$

For this we differentiate f_w three times and evaluate at $t = 0$, noting that

(a) $\tilde{K}$ satisfies $\tilde{K}(0) = 0$, $(\nabla_{\tilde{T}}\tilde{K})(0) = \tilde{Y}$ and equation (2.9), so $\nabla^2_{\tilde{T}}\tilde{K}(0) = 0$ and $\nabla^3_{\tilde{T}}\tilde{K}(0) = \tilde{R}(\tilde{T}, \tilde{Y})\tilde{T}$.

(b) A Killing vector field is a Jacobi field along every geodesic, so $\nabla^2_T\xi_w = R(T, \xi_w)T$ and $\nabla^3_T\xi_w = (\nabla_T R)(T, \xi_w)T + R(T, \nabla_T\xi_w)T$. The result is

$$f_w(t) = t^2(\tilde{Y}, \nabla_{\tilde{T}}\xi_w) + \tfrac{2}{3}t^3(R(\tilde{T}, \tilde{Y})\tilde{T}, \xi_w) + O(t^4). \tag{2.16}$$

For each t, let $\{w_i(t)\}_{i=0}^2$ be an orthonormal basis of $\mathfrak{so}(3)$ with

$$(L^*L)_{\gamma(t)}w_i(t) = \lambda_i(\gamma(t))w_i(t)$$

and $w_0(t) = v(\gamma(t))$ (see Definition 2.4). Then, writing $\lambda_i(\gamma(t)) = \lambda_i(t)$,

$$\begin{aligned}(L^*\tilde{K}, (L^*L)^{-1}L^*\tilde{K})|_{\gamma(t)} &= \sum_{i=0}^{2}\lambda_i(t)^{-1}(L^*\tilde{K}(\gamma(t)), w_i(t))^2 \\ &= \lambda_0(t)^{-1}f^2_{w_0(t)}(t) + \sum_{i=1}^{2}\lambda_i(t)^{-1}f^2_{w_i(t)}(t).\end{aligned} \tag{2.17}$$

According to Lemma 2.5(b) we have $w_0(t) = v_p + O(t^2) \in \mathfrak{so}(3)$, and hence $\xi_{w_0(t)} = \xi_{v_p} + o(t^2)$. Substituting this into (2.16) shows that $f_{w_0(t)}(t) = f_{v_p}(t) + O(t^4)$. But $f_{v_p}(t)$ is also $O(t^4)$ since $(\tilde{Y}, \nabla_{\tilde{T}}\xi_{v_p}) = (\tilde{Y}, J\tilde{Y}) = 0$ and $\xi_{v_p}(p) = 0$. Hence $f^2_{w_0(t)}(t) = O(t^8)$, and by Lemma 2.5(c) the first term of (2.17) is therefore $O(t^6)$.

Similarly, for $i = 1, 2$, (2.16) gives

$$f^2_{w_i(t)}(t) = (\tilde{Y}, \nabla_{\tilde{T}}\xi_{w_i(t)})^2 t^4 + O(t^5)$$
$$= (\tilde{Y}, \nabla_{\tilde{T}}\xi_{w_i(0)})^2 + O(t^5).$$

Furthermore, Lemma 2.5(a) shows that as $t \to 0$, $\lambda_1(t)$ and $\lambda_2(t)$ have the same nonzero limit, which we call λ_1. Thus if we set $v_i = w_i(0)$ the last term in (2.17) is

$$t^4 \lambda_1^{-1} \sum_{i=1}^{2} (\tilde{Y}, \nabla_{\tilde{T}}\xi_{v_i})^2 + O(t^5). \tag{2.18}$$

In fact, since $(\tilde{Y}, \nabla_{\tilde{T}}\xi_{v_0}) = (\tilde{Y}, J\tilde{T}) = 0$, we can include $i = 0$ in the sum without changing the value.

Now observe that, in general,

$$((\nabla_V L^*)(W), v) = (\nabla_V(L^*W) - L^*(\nabla_V W), v)$$
$$= V(W, \xi_v) - (\nabla_V W, \xi_v) = (W, \nabla_V \xi_v).$$

Hence

$$\sum_{i=0}^{2} (\tilde{Y}, \nabla_{\tilde{T}}\xi_{v_i})^2 = \sum_{i=0}^{2} ((\nabla_{\tilde{T}} L^*)(\tilde{Y}), v_i)^2 = |(\nabla_{\tilde{T}} L^*)(\tilde{Y})|^2. \tag{2.19}$$

The lemma follows from (2.17), (2.18), and (2.19).

We can now write down an expansion for the metric on W near a singular point. The leading term in this expansion involves the homogeneous metric on $\mathbb{P}$, which we normalize as follows.

Definition 2.7. Let g_0 be the metric induced on $\mathbb{P} = \mathbb{CP}^{4k-2}$ by the Riemannian submersion (Hopf fibration) $\Sigma \to \mathbb{P}$, where Σ is the unit sphere in $\mathbb{C}^{4k-1}$. The sectional curvatures σ_0 of this metric are given in terms of the complex structure J of $\mathbb{P}$ by

$$\sigma_0(X, Y) = 1 + 3(JX, Y)^2 \tag{2.20}$$

for orthogonal unit vectors $X, Y \in T\mathbb{P}$.

Theorem 2.8. *Let $F\colon (0, \varepsilon) \times \mathbb{P} \to U^*$ be the diffeomorphism (2.7) onto a punctured neighborhood of a singular point $x_0 \in W$. Let r be the distance in W to x_0 and let g_0, σ_0 be as above.*

(a) *The metric g of W satisfies*

$$F^* g = dr^2 \oplus r^2(g_0 + r^2 Q + O(r^3)), \tag{2.21}$$

where Q is the quadratic form on $T\mathbb{P}$ defined at $\tau \in \mathbb{P}$ as the limiting value of the curvature of W as we approach x_0 along the ray $F(r, \tau)$:

$$Q(X, Y) = -\frac{1}{3} \lim_{r \to 0} \left((F^* R)_{(r,\tau)} \left(\frac{\partial}{\partial r}, X \right) Y, \frac{\partial}{\partial r} \right). \tag{2.22}$$

(b) *The sectional curvatures σ of W satisfy*

$$(F^*\sigma)(X,Y) = \frac{\sigma_0(X,Y)-1}{r^2} + l(X,Y) + O(r), \tag{2.23}$$

where l is a function homogeneous of degree zero in X, Y and depending linearly on Q and its second covariant derivatives.

Proof. From Proposition 2.3, equation (2.14), and Lemma 2.6 we have

$$(F^*g)(Y,Y) = r^2|\tilde{Y}|^2 + r^4 Q_1(Y,Y) + O(r^5),$$

where $Q_1(Y,Y) = -\frac{1}{3}(\tilde{R}(\tilde{T},\tilde{Y})\tilde{Y},\tilde{T}) - \lambda_1^{-1}|(\nabla_{\tilde{T}}L^*)(\tilde{Y})|^2$. Since $\tilde{Y}$ is horizontal, $|\tilde{Y}|^2 = g_0(Y,Y)$. This identifies the metric g_r of Proposition 2.3, so we have

$$F^*g = dr^2 \oplus r^2(g_0 + r^2 Q_1 + O(r^3)). \tag{2.24}$$

It is a straightforward calculation to compute the curvature of a metric of the form (2.24). One finds that

$$\left((F^*R)\left(\frac{\partial}{\partial r}, Y\right)Y, \frac{\partial}{\partial r}\right) = -3Q_1(Y,Y) + O(r).$$

Thus the limit in (2.22) exists, the bilinear forms Q and Q_1 are equal, and (2.21) follows from (2.24).

To prove (2.23) we write the metric (2.21) as $dr^2 \oplus g_r$, where g_r is the induced metric on the level-set $\mathbb{P}_r = \{y|\text{dist}(x_0,y) = r\}$. By the Gauss equation, σ is related to the sectional curvature σ_r of g_r by

$$F^*\sigma(X,Y) = \sigma_r(X,Y) - r^{-2} + l_0(X,Y) + O(r),$$

where l_0 is some linear function of Q. By simple rescaling, σ_r is r^{-2} times the sectional curvature of the metric $r^{-2}g_r$. If we then consider $r^{-2}g_r = g_0 + r^2 Q + O(r^3)$ as a 1-parameter family of metrics on $\mathbb{P}$, the standard formulas for the variation of the curvature with respect to the metric give

$$\sigma_r(X,Y) = r^{-2}[\sigma_0(X,Y) + r^2 l_1(X,Y) + O(r^3)],$$

where l_1 is a linear function of $\nabla\nabla Q$. Equation (2.23) follows from the last two equations. q.e.d.

At the beginning of this section we cast our original gauge theory problem in terms of the Riemannian geometry of $\tilde{W}$. Theorem 2.8 solves this geometry problem. Returning to the gauge theory, we immediately obtain a proof of Theorem I of the introduction.

Proof of Theorem I. Given a reducible connection $[A] \in \mathscr{M}_k$, simply apply Theorem 2.8 with $x_0 = [A]$, $\tilde{W} = \tilde{\mathscr{M}}_k$ and $W = \mathscr{M}_k$, noting that σ_0 is given by (2.20). q.e.d.

Actually, Theorem 2.8 enables us to compute the $O(r^2)$ term in the expansion (0.2) of the metric on $\mathscr{M}$. For this, we use the general formula, derived in [9], which expresses the curvature of $\mathscr{M}^*$ in terms of the Green operators $G_A^0 = (d_A^* d_A)^{-1}$ and $G_A^2 = (d_A^-(d_A^-)^*)^{-1}$ of the fundamental elliptic complex (0.1). Specifically, Theorem 2.2 of [9] asserts that if $\overline{X}, \overline{Y} \in H_A = \ker(d^*) \cap \ker(d^-)$ represent $X, Y \in T_{[A]}\mathscr{M}^*$ then the curvature of $\mathscr{M}^*$ at $[A]$ is

$$\text{(2.25)} \qquad \begin{aligned} \langle R(X,Y)Y,X\rangle = 3\langle P_{\overline{X}}^*\overline{Y}, G_A^0 P_{\overline{X}}^*\overline{Y}\rangle &+ \langle P_{\overline{X}}^-\overline{X}, G_A^2(P_{\overline{Y}}^-\overline{Y})\rangle \\ &- \langle P_{\overline{X}}^-\overline{Y}, G_A^2(P_{\overline{X}}^-\overline{Y})\rangle. \end{aligned}$$

(Here $P_X : \Omega^k(\operatorname{Ad} P) \to \Omega^{k+1}(\operatorname{Ad} P)$ is the linear map obtained by bracketing with the $\operatorname{Ad} P$ factor and wedging with the $\Omega^k(M)$ factor, P_X^* is its pointwise adjoint, and $P_X^- = p_- \circ P_X$.)

Corollary 2.9. *In the notation of Theorem* I, *we have*

$$F^*\mathscr{g} = dr^2 \oplus r^2(g_0 + r^2 Q + O(r^3)),$$

where Q is the quadratic form on $\mathbb{P}$ defined as follows. Fix a reducible connection A in the gauge class $[A]$ and identify $\mathbb{P}$ with Σ/S^1, where Σ is the unit sphere in the harmonic subspace $H_A = \mathbb{C}^{4k-1}$ of $\Omega^1(\operatorname{Ad} P)$. Let Φ be a nonzero section of $\operatorname{Ad} P$ *with $d_A\Phi = 0$ and with $J = [\Phi, \cdot]$ as in* (2.2). *Now, given $X \in T_\tau\mathbb{P}$, choose a harmonic form $\tilde{T} \in \Sigma \subset H_A$ representing τ, and a harmonic form $\tilde{X} \in T_{\tilde{T}}\Sigma \subset H_A$ which projects to X and satisfies $\langle \tilde{X}, [\Phi, \tilde{T}]\rangle = 0$. Then*

$$\text{(2.26)} \qquad \begin{aligned} Q_\tau(X,X) = &-\langle P_{\tilde{T}}^*\tilde{X}, G_A^0 P_{\tilde{T}}^*\tilde{X}\rangle - \tfrac{1}{3}\langle P_{\tilde{T}}^-\tilde{T}, G_A^2 P_{\tilde{X}}^-\tilde{X}\rangle \\ &+ \tfrac{1}{3}\langle P_{\tilde{T}}^-\tilde{X}, G_A^2 P_{\tilde{T}}^-\tilde{X}\rangle, \end{aligned}$$

where G_A^0, G_A^2 are the Green operators for the reducible connection A.

Proof. Let $\gamma(t) = [A_t] = [A + t\tilde{T} + O(t^2)]$ be the radial geodesic in $\mathscr{M}$ given by $\gamma(t) = F(t,\tau)$. We will evaluate $Q_\tau(X,X)$ by combining equations (2.22) and (2.25).

Let $\tilde{\gamma}(t)$ be the horizontal lift of γ to $\tilde{\mathscr{M}}$ with initial tangent vector represented by $\tilde{T} \in T_A$. Extend $\tilde{X}$ to a horizontal vector field among $\tilde{\gamma}$, and write $\tilde{X}(\gamma(t)) = \tilde{X}_t$, $\tilde{T}(\gamma(t)) = \tilde{T}_t$. Inserting $\tilde{T}_t, \tilde{X}_t$, and A_t into (2.25), we first consider the term

$$\text{(2.27)} \qquad \langle P_{\tilde{T}_t}^*\tilde{X}_t, G_{A_t}^0 P_{\tilde{T}_t}^*\tilde{X}_t\rangle,$$

whose limit we must evaluate as $t \to 0$. Care must be taken because the smallest eigenvalue λ_t of $d_{A_t}^* d_{A_t}$ approaches zero as $t \to 0$. Therefore we let Φ_t be the eigensection of $d_{A_t}^* d_{A_t}$ with eigenvalue λ_t, normalized so that $\|\Phi_t\|_{L^2} = \|\Phi\|_{L^2}$, and write

$$\text{(2.28)} \qquad P_{\tilde{T}_t}^*\tilde{X}_t = (\mathrm{I})_t + (\mathrm{II})_t,$$

where $(\mathrm{I})_t = \|\Phi\|^{-2}\langle P^*_{\tilde{T}_t}\tilde{X}_t, \Phi_t\rangle\Phi_t$ is the L^2-orthogonal projection of $P^*_{\tilde{T}_t}\tilde{X}_t$ onto the eigenspace of the smallest eigenvalue. Since (2.28) is an orthogonal eigenspace decomposition, the expression (2.27) is just the sum $\langle(\mathrm{I})_t, G^0_{A_t}(\mathrm{I})_t\rangle +$ $\langle(\mathrm{II})_t, G^0_{A_t}(\mathrm{II})_t\rangle$.

Now $\langle(\mathrm{I})_t, G^0_{A_t}(\mathrm{I})_t\rangle = \|\Phi\|^2\lambda_t^{-2}\langle P^*_{\tilde{T}_t}\tilde{X}_t, \Phi_t\rangle^2$, and we claim that this term is $O(t^2)$. To establish this, consider $f(t) = \langle P^*_{\tilde{T}_t}\tilde{X}_t, \Phi\rangle = \langle \tilde{X}_t, [\tilde{T}_t, \Phi]\rangle$. By the definition of $\tilde{X}$, this function vanishes at $t = 0$, and its first derivative there is $\langle \nabla_{\tilde{T}}\tilde{X}, [\tilde{T}, \Phi]\rangle = -\langle \nabla_{\tilde{T}}\tilde{X}, J\tilde{T}\rangle$. On the other hand, $\tilde{X}$ is everywhere horizontal, so $\langle \tilde{X}_t, \xi_\Phi(\gamma(t))\rangle \equiv 0$. If we differentiate this equation twice, use the fact that $\nabla_{\tilde{T}}\nabla_{\tilde{T}}\xi_\Phi = \tilde{R}(\tilde{T}, \xi_\Phi)\tilde{T}$ (cf. the proof of Lemma 2.6), and evaluate at $t = 0$ (where $\xi_\Phi = 0$), we obtain $\langle \nabla_{\tilde{T}}\tilde{X}, J\tilde{T}\rangle = 0$. Therefore $f(t) = O(t^2)$. But, as in Lemma 2.5, we have $\Phi_t = \Phi + O(t^2)$, so $\langle P^*_{\tilde{T}}\tilde{X}_t, \Phi_t\rangle =$ $f(t) + O(t^2) = O(t^2)$. Since $\lambda_t = O(t^2)$ (as in Lemma 2.5) we conclude that the contribution to (2.27) involving $(\mathrm{I})_t$ is indeed $O(t^2)$.

Therefore the limiting value of (2.27) is the limit of $\langle(\mathrm{II})_t, G^0_{A_t}(\mathrm{II})_t\rangle$. Now $(\mathrm{II})_t$ and the restriction of $G^0_{A_t}$ to the orthogonal complement of Φ_t are both continuous at $t = 0$ (all eigenvalues but λ_t are bounded away from zero). Hence we obtain the limit of (2.27) simply by substituting A for A_t and $\lim_{t\to 0}(\mathrm{II})_t$ for $P^*_{\tilde{T}_t}\tilde{X}_t$. But $\lim_{t\to 0}(\mathrm{II})_t = P^*_{\tilde{T}}\tilde{X}$, since $\langle P^*_{\tilde{T}}\tilde{X}, \Phi\rangle = 0$. Therefore the limit of (2.27), multiplied by $-1/3$ as in (2.22), is in the first term in (2.26).

Since $G^2_{A_t}$ is uniformly bounded as $t \to 0$ ($(d^-_A)^* d^-_A$ is continuously invertible for *all* $[A] \in \mathscr{M}$, and $\mathscr{M}$ is locally compact), the remaining terms in (2.25) approach the corresponding ones in (2.26). q.e.d.

We conclude this section by discussing several examples which help to unravel the geometric meaning of the formula (2.21) for the metric. Let $\overline{g}$ be the standard metric on the unit sphere S^n.

Example 1. The metric $dr^2 \oplus r^2(c^2\overline{g})$ on $(0,1)\times S^n$ (where c is a constant) defines a "linear" cone of one of three types.

(a) If $c = 1$ we get simply the flat metric on the punctured unit ball in $\mathbb{R}^{n+1}$, and the metric extends smoothly over the vertex.

(b) If $0 < c < 1$ we have a flat cone embedded in $\mathbb{R}^{n+1}$ with vertex angle $\arcsin(c)$.

(c) If $c > 1$ the cone cannot be isometrically embedded in Euclidean space, and is therefore harder to visualize.

Example 2. Let (X, g_1) be any compact Riemannian manifold not homeomorphic to a sphere and consider the metric $dr^2 \oplus r^2 g_1$ on the punctured cone $C^* = (0,1)\times X$. In this case the cone is not homeomorphic to a manifold, unlike the cones in Example 1. The geometry of the cone falls into two categories, distinguished by the asymptotic behavior of the curvature of C^* near the vertex.

(a) Unless all the sectional curvatures of g_1 equal 1, then, as in Theorem 2.8(b), one can show that the curvature of the cone blows up as one approaches the vertex.

(b) If g_1 has constant curvature 1, then X is a quotient of a sphere, and it is easy to see that the universal cover $\tilde{C}^*$ is $\mathbb{R}^{n+1} - \{0\}$ with the flat metric. Consequently, C^* is a flat cone on a rational homology sphere.

Example 3. More general cone metrics have an expansion like that in (2.21). The linearization—which is always one of the above types—determines the nature of the singularity at the vertex to leading order. For example, on a Riemannian manifold, a normal coordinate neighborhood of any point is a cone on a sphere; the linearization is a cone of type 1(a) above.

Example 4. For a global example which realizes the hypotheses of our SO(3)-action on $\tilde{W}$, take $\tilde{W} = S^2 \times S^2 \times S^2 \times S^2$, the product of four unit spheres, with SO(3) acting by rotation on each factor. The point $p = (u_1, u_2, u_3, u_4)$ has trivial stabilizer unless all the u_i, considered as vectors in $\mathbb{R}^3$, lie on a line. Thus the exceptional points are all of the form $(u_1, \pm u_1, \pm u_1, \pm u_1)$, and the stabilizer of such a point is SO(2) acting in the plane orthogonal to u_1. The quotient W therefore has eight singular points, each with a neighborhood homeomorphic to a cone on $\mathbb{C}P^2$. Near such a singular point, the metric behaves, to leading order, like Example 2(a). Had we used three copies of S^2 instead of four, we would have obtained cones on $\mathbb{C}P^1$, behaving as in Example 1(c) with $c = 2$ (on $\mathbb{C}P^1 = S^2$ the metric g_0 of Definition 2.7 equals $4\overline{g}$).

The cones which occur in the moduli space are of the same character as those in Example 4; in particular, their linearizations are of type 2(a).

PART II. GEOMETRY OF THE COLLAR

3. The approximate tangent space

We now turn our attention to the collar of the moduli space $\mathscr{M}$ and study the asymptotic behavior of the metric $\mathscr{g}$ there, with the primary goal of proving Theorem II. Recall that each connection A has a scale $\lambda = \lambda([A])$ (the radius of the smallest ball containing half the "energy"), and that there is a constant $\lambda_0 > 0$ such that each $[A] \in \mathscr{M}$ with $\lambda \leq \lambda_0$ is irreducible and has a well-defined center $p = p([A]) \in M$. We then define the collar to be

$$\mathscr{M}_{\lambda_0} = \{[A] \in \mathscr{M} | \lambda([A]) \leq \lambda_0\}.$$

(The precise definitions of $\lambda(A)$ and $p(A)$ are given in §4; the details of those definitions are not important in this section.)

As noted in the introduction, given $[A] \in \mathscr{M}^*$ we can choose $A \in \mathscr{A}$ in the gauge class $[A]$ and identify (isometrically) $T_{[A]}\mathscr{M}$ with $H_A = \ker(d_A^*) \cap \ker(d_A^-) \subset \Omega^1(\operatorname{Ad} P)$. We will make this identification throughout the next three sections. Accordingly, we will replace the notation H_A by T_A, this being more suggestive of a tangent space. In this section we will construct, for $[A] \in \mathscr{M}_{\lambda_0}$, an approximation $\hat{T}_A$ to T_A and obtain estimates on $\|\mathrm{Id} - \pi_A\|$, where $\pi_A : \hat{T}_A \to T_A$ is the orthogonal projection. Our estimates will rely on three basic facts, the first two of which follow from Theorem 16 of [6] and the third from Theorem 21(i) of [6].

Fact A. Given $\varepsilon > 0, N > 0$, there exists $\lambda_0 = \lambda_0(\varepsilon, N)$ such that the curvature of each $[A] \in \mathscr{M}_{\lambda_0}$ satisfies

$$\sup_{B(p,N\lambda)} \left| |F_A|^2 - |F_\lambda|^2 \right| \leq \varepsilon \lambda^{-4}, \tag{3.1}$$

where $\lambda = \lambda([A])$, $B(p, N\lambda)$ is the ball of radius $N\lambda$ about the center p of $[A]$, and F_λ is the curvature of the standard instanton of scale λ on $\mathbb{R}^4$, pulled back to $B(p, N\lambda)$ by any choice of normal coordinates and local gauge about p. Here one can take $|F|^2$ to be defined either by the metric on M or by the Euclidean metric in the normal coordinate system defining F_λ; the assertion is true for either interpretation of the norm.

Fact B. There exists a constant C such that, given $\delta > 0$, there exist $r_0 > 0$ and $\lambda_0 = \lambda_0(\delta) > 0$ such that the curvature of each $[A] \in M_{\lambda_0}$ satisfies

$$|F_A|(q) \leq C\lambda^{2-\delta}/r^{4-\delta},$$

whenever $r = \operatorname{dist}(q, p([A])) \leq r_0$ (here $\lambda = \lambda([A])$).

Fact C. Let $\Omega \subset M$ be the complement of the ball $B(p, \sqrt{r_0 \lambda})$, where $p = p([A])$ and $\lambda = \lambda([A])$. Then there are constants λ_0 and c such that each $[A] \in \mathscr{M}_{\lambda_0}$ satisfies

$$\int_\Omega |F_A|^2 \leq c\lambda^2.$$

We will also need the following simple lemma.

Lemma 3.1. *Let $Z = \operatorname{grad}(\phi)$ be a gradient vector field and A a self-dual connection. Then*

(a) $d_A^*(i_Z F_A) = 0$,

(b) $d_A^-(i_Z F_A) = \sum_{j,k} (H^0\phi)_{jk} \theta^j \wedge (i_{e_k} F_A)$.

Here i_Z denotes contraction with Z, $\{e_j\}$ is any local orthonormal basis of TM, $\{\theta^j\}$ is the dual coframe, and $H^0\phi$ is the traceless part of the covariant Hessian $H\phi$ of ϕ.

Proof. (a) We have $*(i_Z F_A) = \pm d\phi \wedge *F_A$, so

$$d_A^*(i_Z F_A) = \pm * d_A(d\phi \wedge *F_A) = \pm * (d\phi \wedge d_A * F_A) = 0.$$

(b) Fixing $p \in M$, it suffices to verify this under the assumption that $(\nabla e_j)_p = 0$. Let ∇ denote both the Levi-Civita connection on M and its extension, by tensoring with ∇^A, to a connection on $\operatorname{Ad} P$-valued tensors. Then, at p,

$$\begin{aligned} d_A(i_Z F_A) &= \theta^j \wedge \nabla_{e_j}(i_Z F_A) = \theta^j \wedge (i_{\nabla_j Z} F_A + i_Z \nabla_k F_A) \\ &= (H\phi)_{jk}\theta^j \wedge (i_{e_k} F_A) - i_Z(\theta^j \wedge \nabla_j F_A) + \nabla_Z F_A. \end{aligned}$$

The middle term on the right is $-i_Z(d_A F_A) = 0$ by the Bianchi identity. Using $\nabla_Z(p_-) = 0$ we then have

$$d_A^-(i_Z F_A) = (H\phi)_{jk} p_-(\theta^j \wedge (i_{e_k} F_A)) + \nabla_Z(F_A^-),$$

and the last term vanishes since A is self-dual. Finally, a little algebra shows that the remaining term is precisely the expression in (b). q.e.d.

Notation. Choose $0 < \delta \ll q$ and let $r_0, \lambda_0 = \lambda_0(\delta)$ be as in Fact B with $\lambda_0 \le r_0 \le$ one-half the injectivity radius of M. Also, fix a function $b \in C_0^\infty(\mathbb{R})$ with $0 \le b(t) \le 1$, $b(t) = 1$ for $t \in [0,1]$, $b(t) = 0$ for $t \ge 2$, and $b'(t) \le 0$ for all t.

Given $[A] \in \mathscr{M}_{\lambda_0}$, we will consider the local geometry of M around the center point $p = p([A])$. Thus we let r denotes the distance to p (a function on M), $B = B(p, 2r_0)$ denote the ball of radius r_0 around p, and Ω denote the annulus $\{r_0 \le r \le 2r_0\}$; χ_B and χ_Ω will denote the corresponding characteristic functions. Let β be the cutoff function $\beta(r) = b(r/r_0)$; note that β is supported in B and its gradient is supported in Ω.

In this and subsequent sections we will use the letter c for a universal positive constant depending on the geometry of M and on r_0, but not on $[A] \in \mathscr{M}_{\lambda_0}$. Thus, for example, we will use the inequalities $|\chi_B dr| \le c\chi_B$ and $|\chi_B \Gamma^i_{jk}| \le cr\chi_B$, where Γ^i_{jk} are the Christoffel symbols in any normal coordinate system centered at p. The value of c will be constantly updated; for example when c is multiplied by 2 the result is immediately renamed c. Similarly, the value of λ_0 will be decreased as needed.

Definition 3.2. (a) Given $p \in M$ and $\mathbf{a} \in T_pM$, let $f_{\mathbf{a}}$ be the unique function on B such that

(i) $f_{\mathbf{a}}$ is linear in some—and hence any—normal coordinate system centered at p.

(ii) $f_{\mathbf{a}}(p) = 0$,

(iii) $(\operatorname{grad} f_{\mathbf{a}})(p) = \mathbf{a}$.

(b) Given $[A] \in \mathscr{M}_{\lambda_0}$ with center p and scale λ, let $X, Y_{\mathbf{a}}$ denote the vector fields

$$X = \operatorname{grad}(\beta(r) \cdot \tfrac{1}{2}r^2), \qquad Y_{\mathbf{a}} = \operatorname{grad}(\beta(r) f_{\mathbf{a}}). \tag{3.2}$$

These vector fields will appear frequently in the next several sections.

(c) Define

$$\hat{T}_A = \operatorname{span}\{i_X F_A, i_{Y_{\mathbf{a}}} F_A | \mathbf{a} \in T_pM\} \subset \Omega^1(\operatorname{Ad} P)\}.$$

The assignment $A \mapsto \hat{T}_A$ is $\mathscr{G}$-equivariant, so induces a vector bundle $\hat{T}\mathscr{M}$ over $\mathscr{M}_{\lambda_0}$; the fiber $\hat{T}_{[A]}$ is isomorphic to $\hat{T}_A$. We refer to both $\hat{T}_A$ and $\hat{T}_{[A]}$ as the *approximate tangent space.*

More Notation. Given $[A] \in \mathscr{M}_{\lambda_0}$, we let $\{x^i\}$ denote an arbitrary choice of normal coordinates centered at p. We use the $\{x^i\}$ for our computations, the results of which are independent of the choice of $\{x^i\}$. We use the normal coordinate system to identify T_pM with $\mathbb{R}^4$, so $\mathbf{a} = a_i\partial/\partial x^i$, $f_{\mathbf{a}} = \sum a_i x^i$, and $r^2 = \sum (x^i)^2$. Also we adopt the convention of implied summation over repeated indices.

We will show that, as $\lambda \to 0$, $\|\operatorname{Id} - \pi_A\| \to 0$ uniformly in $[A]$, justifying the term "approximate tangent space". The first step is to prove the following.

Proposition 3.3. *Let $[A] \in \mathscr{M}_{\lambda_0}$ and let X and $Y = Y_{\mathbf{a}}$ be as in Definition* 3.2. *Then*

(a) $d_A^*(i_X F_A) = d_A^*(i_Y F_A) = 0$,

(b) $\|d_A^-(i_X F_A)\|_2^2 \le c\lambda^{4-2\delta}$,

(c) $\|d_A^-(i_Y F_A)\|_2^2 \le c|\mathbf{a}|^2\lambda^2$.

Proof. X and Y have the form $Z = \operatorname{grad}(\beta f)$, where f is either $\frac{1}{2}r^2$ or $f_{\mathbf{a}}$. Hence Lemma 3.1(a) immediately gives statement (a). For (b) and (c), we apply Lemma 3.1(b) to $\phi = \beta f$. We have

$$H(\beta f) = \nabla d(\beta f) = (H\beta) f + (d\beta \otimes df + df \otimes d\beta) + \beta(Hf),$$

so Lemma 3.1(b) gives the pointwise bound

$$|d_A^-(i_Z F_A)| \le (|f|\,|H^0\beta| + 2|d\beta|\,|df| + \beta|H^0 f|)|F_A|. \tag{3.3}$$

Now

$$\begin{aligned} d\beta &= r_0^{-1} b'(r/r_0)\, dr, \\ \nabla d\beta &= r_0^{-2} b''(r/r_0)\, dr \otimes dr + r_0^{-1} b'(r/r_0) \nabla dr, \\ \nabla dr &= r^{-1}(x^i \nabla dx^i + dx^i \otimes dx^i - dr \otimes dr). \end{aligned}$$

Since $\nabla dx^i = -\Gamma^i_{jk}\, dx^k \otimes dx^k$ and $r^{-1} \le c$ on Ω, we have

$$|d\beta| \le c\chi_{\Omega}, \qquad |H^0\beta| \le |H\beta| \le c\chi_{\Omega}. \tag{3.4}$$

Substituting into (3.3) we find

$$|d_A^-(i_Z F_A)| \le c\chi_{\Omega}(|f| + |df|)|F_A| + c\chi_B |H^0 f|\,|F_A|,$$

whence

$$\|d_A^-(i_Z F_A)\|_2^2 \le c\int_\Omega (|f| + |df|)^2 |F_A|^2 + c\int_B |H^0 f|^2 |F_A|^2. \tag{3.5}$$

We apply this in two cases.

Case 1. Taking $f = \frac{1}{2}r^2$ (i.e. $Z = X$), we have $|f| \le cr^2$, $|df| \le cr$, so $\chi_{\Omega}(|f| + |df|) \le c\chi_{\Omega}$. We also have $Hf = x^i \nabla dx^i + dx^i \otimes dx^i$, and $\chi_B |g - dx^i \otimes dx^i| \le \chi_B \cdot cr^2$, so $\chi_B |H^0 f| \le \chi_B \cdot cr^2$.

Case 2. Taking $f = a_i x^i$ (i.e., $Z = Y_{\mathbf{a}}$) we have $|f| \le |\mathbf{a}| r$, $|df| \le c|\mathbf{a}|$, so $\chi_{\Omega}(|f| + |df|) \le \chi_{\Omega} \cdot c|\mathbf{a}|$, and $\chi_B |H^0 f| \le \chi_B |\nabla df| \le \chi_B \cdot c|\mathbf{a}| r$.

In each case we can substitute the appropriate bounds into (3.5) and estimate the integrals using Fact B. First, we have

$$\int_\Omega |F_A|^2 \le c\lambda^{4-2\delta} \int_\Omega r^{2\delta - 8} \le c\lambda^{4-2\delta}. \tag{3.6}$$

Next, we bound the integrals over B by writing $B = B_\lambda \cup (B - B_\lambda)$, where $B_\lambda = \{r \le \lambda\}$, and using Fact B on the annulus $B - B_\lambda$. Thus for $p \in \mathbb{Z}$,

$$\begin{aligned} \int_B r^p |F_A|^2 &\le c\lambda^p \int_{B_\lambda} |F_A|^2 + c\lambda^{4-2\delta} \int_{B - B_\lambda} r^{p+2\delta-8} \\ &\le c\lambda^p + c\lambda^{4-2\delta}\operatorname{sgn}(p + 2\delta - 4)(r^{p+2\delta-4}|_{\lambda}^{r_0}), \end{aligned} \tag{3.7}$$

since $\int_{B_\lambda} |F_A|^2 \le \|F_A\|_2^2 \le 8\pi^2$. In Case 1, (3.5), (3.6) and (3.7) yield

$$\|d_A^-(i_X F_A)\|_2^2 \le c\lambda^{4-2\delta} + c(\lambda^4 + \lambda^{4-2\delta}),$$

and statement (b) of the proposition follows. Similarly, in Case 2, (3.5), (3.6) and (3.7) plus the fact that $4 - 2\delta > 2$, give statement (c). q.e.d.

We wish next to consider the L^2 inner products of the $i_Z F_A$. For this we will need two lemmas.

Lemma 3.4. *For any two vector fields Z and Z', and any two self-dual bundle-valued 2-forms F and F',*

$$(i_Z F, i_{Z'} F') + (i_{Z'} F, i_Z F') = (Z, Z')(F, F')$$

pointwise. In particular,

$$(i_Z F, i_{Z'} F) = \tfrac{1}{2}(Z, Z')|F|^2. \tag{3.8}$$

Proof. Let e_Z denote exterior multiplication by the 1-form dual to Z. The well-known formula $e_Z i_{Z'} + i_{Z'} e_Z = (Z, Z') \cdot \mathrm{Id}$ holds for bundle-valued forms. Furthermore, e_Z is the pointwise adjoint of i_Z, and $e_Z = -*i_Z*$. Thus, since $*$ is an isometry and $F = *F, F' = *F'$,

$$\begin{aligned}(i_Z F, i_{Z'} F') + (i_{Z'} F, i_Z F') &= (F, e_Z i_{Z'} F') + (*i_{Z'} * F, *i_Z * F')\\ &= (F, e_Z i_{Z'} F') + (e_{Z'} F, e_Z F')\\ &= (F, (e_Z i_{Z'} + i_{Z'} e_Z) F) = (Z, Z')(F, F'). \quad \text{q.e.d.}\end{aligned}$$

In our estimates we will encounter integrals similar in form to $\int h x^i x^j |F_A|^2$, where h is a cutoff function and $\{x^i\}$ are normal coordinates at $p(A)$. The next lemma shows that as $\lambda \to 0$ these integrals approach the corresponding integrals on $\mathbb{R}^4$.

Lemma 3.5. *Let $h\colon [0,\infty) \to \mathbb{R}$ be piecewise continuous with compact support. For each multi-index $I = (i_1, \cdots, i_n)$ let x^I be the corresponding monomial on $\mathbb{R}^4$ and let $S \subset \mathbb{R}^4$ be the unit sphere. Define constants $K_I = \int_S x^I$ and $c_{m,n} = \int_0^\infty 48h(t)t^{n-m+3}(1+t^2)^{-4}\,dt$ (if $n-m+3 \le -1$ we assume* supp$(h) \subset (0,\infty)$). *Now let $\{x^i\}$ be normal coordinates on M at $p(A)$. Then as $\lambda = \lambda(A) \to 0$*

$$\lim_{\lambda \to 0} \left[\lambda^{m-n} \int_M h(r/\lambda) x^I r^{-m} |F_A|^2 \right] = c_{m,n} K_I,$$

uniformly in $[A] \in \mathscr{M}$. In particular, when $n = 1, 2$ this holds with $K_i = 0$ and $K_{ij} = \frac{1}{2}\pi^2 \delta_{ij}$.

Proof. Choose $\varepsilon > 0$ and suppose supp$(h) \subset [0, N]$. Let $\lambda_0 = \lambda_0(\varepsilon, N)$ be the constant supplied by Fact A, so (3.1) holds. In normal coordinates the metric and the volume form satisfy $g_{ij} = \delta_{ij} + O(r^2)$ and $dv_g = d^4x(1 + O(r^2))$, where d^4x is the Euclidean volume form on $\mathbb{R}^4$. Hence, writing $|F|_0^2 = \frac{1}{2}(F_{ij}, F_{ij})$, we have

$$\begin{aligned}&\left| \int h(r/\lambda) x^I r^{-m} |F_A|^2 \, dv_g - \int h(r/\lambda) x^I r^{-m} |F_\lambda|_0^2 \, d^4x \right|\\ &\qquad\qquad \le \varepsilon \left| \int h(r/\lambda) x^I r^{-m} \lambda^{-4} \, d^4x \right|.\end{aligned}$$

Now $|F_\lambda|_0^2 = 48\lambda^4(\lambda^2 + r^2)^{-4}$; see equation (3.4.6) of [10], for example. If we change variables to $u = x/\lambda$, set $t = |u|^2$, integrate in polar coordinates, and multiply through by λ^{m-n}, then we obtain

$$\left| \lambda^{m-n} \int h(r/\lambda) x^I r^{-m} |F_A|^2 - K_I c_{m,n} \right| \le c\varepsilon.$$

This proves the lemma.

Proposition 3.6. *For each $\varepsilon > 0$, $\exists\lambda_0 > 0$ such that for all A with $\lambda(A) = \lambda \le \lambda_0$,*

(a) $\|i_{Y_{\mathbf{a}}} F_A\|_2^2 = 4\pi^2 |\mathbf{a}|^2 (1 + O(\lambda^2))$,

(b) $\|i_X F_A\|_2^2 = 8\pi^2\lambda^2(1 + O(\varepsilon))$,

(c) $|\langle i_X F_A, i_{Y_{\mathbf{a}}} F_A\rangle| \le c\varepsilon \|i_X F_A\| \|i_{Y_{\mathbf{a}}} F_A\|$,

where, in (b), *we mean* $|\|i_X F_A\|_2^2 - 8\pi^2\lambda^2| \le \varepsilon\lambda^2$.

Proof. Let ϕ, ψ be functions equal to either $\frac{1}{2}r^2$ or $a_i x^i$, so $Z = \operatorname{grad}(\beta\phi)$ and $Z' = \operatorname{grad}(\beta\psi)$ are either X or $Y = Y_{\mathbf{a}}$. By the product rule and (3.4),

$$|(Z, Z') - \beta^2(d\phi, d\psi)| = |(d\beta, \phi\psi d\beta + \beta\phi d\psi + \beta\psi d\phi)| \\ \le c(|\psi\phi| + |d(\phi\psi)|)\chi_{\Omega}.$$

Now let $p = \deg(\phi\psi)$ (so $p = 2, 3$, or 4) and let $m =$ homogeneity of $|\mathbf{a}|$ in $\phi\psi$ (so $m = 0, 1$, or 2). Then $|\phi\psi| \le c|\mathbf{a}|^m r^p$ and $|d(\phi\psi)| \le c|\mathbf{a}|^m r^{p-1}$. Multiplying the inequality above by $\frac{1}{2}|F|^2$ and using (3.8) gives

$$|(i_Z F, i_{Z'} F) - \tfrac{1}{2}\beta^2(d\phi, d\psi)|F|^2| \le c|F|^2|\mathbf{a}|^m(r^p + r^{p-1})\chi_{\Omega} \\ \le c|F|^2|\mathbf{a}|^m\chi_{\Omega},$$

since r and r^{-1} are bounded on Ω. Integrating over M and using equation (3.6) yields

$$\langle i_Z F_A, i_{Z'} F_A\rangle = \tfrac{1}{2}\int \beta^2|F_A|^2(d\phi, d\psi) + O(|\mathbf{a}|^m\lambda^{4-2\delta}). \tag{3.9}$$

We will apply this in each of the three cases.

(a) Taking $\phi = \psi = a_i x^i$, we have $p = m = 2$, $d\phi = d\psi = a_i dx^i$, and hence $(d\phi, d\psi) = g^{ij}a_i a_j = |\mathbf{a}|^2(1 + O(r^2))$. Then (3.9) gives

$$\langle i_Y F, i_Y F\rangle = \tfrac{1}{2}|\mathbf{a}|^2\left(\int \beta^2|F|^2(1 + O(r^2)) + O(\lambda^{4-2\delta})\right) \\ = \tfrac{1}{2}|\mathbf{a}|^2\left(\int \beta^2|F|^2 + O(\lambda^2)\right),$$

where we have used (3.7) and the fact that $4 - 2\delta > 2$. Now write $\beta^2 = 1 - (1 - \beta^2)$. Note that $\|F\|_2^2 = 8\pi^2$, and $1 - \beta^2$ has support in $M - B(p, r_0)$. When λ is sufficiently small we can then apply Fact C to obtain

$$|\|i_Y F\|_2^2 - 4\pi^2|\mathbf{a}|^2| \le c|\mathbf{a}|^2\lambda^2.$$

(b) Next take $\phi = \psi = \frac{1}{2}r^2$, so $p = 4$, $m = 0$, $d\phi = d\psi = x^i dx^i$, and $(d\phi, d\psi) = g^{ij}x^i x^j = r^2$ (in normal coordinates, $g^{ij}x^j = x^i = g_{ij}x^j$). Choose $N \ge 1$ such that $N^{2\delta-2} < \varepsilon$ and let λ_0 be small enough that $N\lambda_0 < r_0$. Then

$\beta \equiv 1$ on $\{r \leq N\lambda\}$ for $\lambda \leq \lambda_0$, so (3.9) gives

$$\begin{aligned} \|i_X F\|_2^2 &= \tfrac{1}{2}\int \beta^2 r^2 |F|^2 + O(\lambda^{4-2\delta}) \\ &= \tfrac{1}{2}\int_{r\leq N\lambda} r^2|F|^2 + \tfrac{1}{2}\int_{r>N\lambda} \beta^2 r^2 |F|^2 + O(\lambda^{4-2\delta}). \end{aligned} \tag{3.10}$$

Now apply Lemma 3.5 in the case $h =$ characteristic function of $[0, n]$, $n = 0$, $m = -2$, to find

$$\begin{aligned} \lim_{\lambda\to 0}\left(\lambda^{-2}\int_{r\leq N\lambda} r^2|F|^2\right) &= \mathrm{vol}(S^3)\cdot\int_0^N 48t^5(1+t^2)^{-4}\,dt \\ &= 16\pi^2\left[1 - \frac{3N^4+3N^2+1}{(N^2+1)^3}\right], \end{aligned}$$

by a direct calculation. But $(3N^4+3N^2+1)(N^2+1)^{-3} \leq 7N^{-2} \leq 7\varepsilon$, so, by taking λ_0 sufficiently small,

$$\left|\tfrac{1}{2}\int_{r\leq N\lambda} r^2|F|^2 - 8\pi^2\lambda^2\right| \leq 8\varepsilon\lambda^2. \tag{3.11}$$

On the other hand, using Fact B as in (3.7) we obtain

$$\int_{r>N\lambda} \beta^2 r^2|F|^2 \leq c\lambda^{4-2\delta}(N\lambda)^{2\delta-2} \leq c\varepsilon\lambda^2. \tag{3.12}$$

The result now follows from (3.10) and (3.11).

(c) Take $\phi = \frac{1}{2}r^2$ and $\psi = a_i x^i$, so $p = 3$, $m = 1$, and $(d\phi, d\psi) = \psi$. (3.9) gives

$$\langle i_X F, i_Y F\rangle = \tfrac{1}{2}\int \beta^2\psi|F|^2 + O(|\mathbf{a}|\lambda^{4-2\delta}).$$

Choose $N \geq 1$ such that $N^{2\delta-3} < \varepsilon$, and choose λ_0 small enough that $N\lambda_0 \leq r_0$. Then, if $\lambda \leq \lambda_0$, we have

$$\int \beta^2\psi|F|^2 = \int_{r\leq N\lambda}\psi|F|^2 + \int_{r>N\lambda}\beta^2\psi|F|^2.$$

For the outer integral, we apply the argument used in (3.12), obtaining

$$\left|\int_{r>N\lambda}\beta^2\psi|F|^2\right| \leq c|\mathbf{a}|\lambda\varepsilon.$$

For the inner integral, Lemma 3.5 implies

$$\lim_{\lambda\to 0}\left(\lambda^{-1}\int_{r\leq N\lambda} x^i|F|^2\right) = 0,$$

so, for λ sufficiently small, we have $|\int\psi|F|^2| \leq \varepsilon|\mathbf{a}|\lambda$. Putting all this together,

$$|\langle i_X F, i_Y F\rangle| \leq c|\mathbf{a}|\lambda\varepsilon$$

for λ sufficiently small. By parts (a) and (b), $\|i_Y F\|$ and $\|i_X F\|$ are commensurate with $|\mathbf{a}|$ and λ, respectively. Hence statement (c) follows. q.e.d.

The next proposition shows that the projection $\pi\colon \hat{T}_A \to T_A$ satisfies $\pi_A = \mathrm{Id} + O(\lambda^{1-\delta})$.

Proposition 3.7. *There exist constants $c, \lambda_0 > 0$ such that if $Z = \operatorname{grad}(\beta(\frac{1}{2}a_0 r^2 + f_a))$ and $\lambda \leq \lambda_0$ then*

$$\|(1-\pi_A)i_Z F_A\| \leq c(|a_0|\lambda^{2-\delta} + |\mathbf{a}|\lambda), \tag{3.13}$$

$$\|\pi_A i_Z F_A\| = \|i_Z F_A\|(1 + O(\lambda^{1-\delta})). \tag{3.14}$$

Proof. For any $\omega \in \Omega^1(\operatorname{Ad} P)$, $(1-\pi_A)\omega$ is the projection of ω onto the L^2-orthogonal complement of the harmonic space. This can be expressed in terms of the Laplacians and Green operators of the fundamental elliptic complex, namely $(1-\pi_A)\omega = d_A G_A^0 d_A^* \omega + (d_A^-)^* G_A^2 d_A^- \omega$. Taking $\omega = i_Z F$ we have $d_A^* \omega = 0$ by Proposition 3.3(a). Hence

$$\begin{aligned}\|(1-\pi_A)\omega\|_2^2 &= \|(d_A^-)^* G_A^2 d_A^- \omega\|_2^2 \\ &= \langle G_A^2 d_A^- \omega, d_A^-(d_A^-)^* G_A^2 d_A^- \omega\rangle = \langle G_A^2 d_A^- \omega, d_A^- \omega\rangle.\end{aligned}$$

By Proposition 18(ii) of [6] there exist $\lambda_0, \mu > 0$ such that if $\lambda \leq \lambda_0$ then the first eigenvalue of $d_A^-(d_A^-)^*$ is $\geq \mu$. Therefore

$$\|(1-\pi_A)i_Z F\|_2^2 \leq \mu^{-1}\|d_A^- i_Z F\|_2^2,$$

and (3.13) follows by Proposition 3.3. Combining (3.13) and Proposition 3.6, we have $\|\pi_A i_Z F\| \leq \|i_Z F\| + \|(1-\pi_A)i_Z F\| \leq (1 + c\lambda^{1-\delta})\|i_Z F\|$; the reverse inequality is similar and we obtain (3.14).

4. The differential of the collar map

In this section we will derive some estimates on the differential of the collar map Ψ (0.4). We first recall Donaldson's precise definition of this collar map.

Let $b \in C_0^\infty([0,\infty))$ be the cut-off function used in §3. For any $s > 0$ and any two points x, y in any Riemannian manifold, we define the bump-function

$$\gamma_s(x,y) = b\left(\frac{\operatorname{dist}(x,y)}{s}\right).$$

Definition [6, Definition 15]. For a connection A on a given complete oriented Riemannian manifold with metric g and volume form dv_g, define

$$R_A(s,x) = \int_M \gamma_s(x,y)|F_A(y)|^2\, dv_g(y), \tag{4.1}$$

$$\lambda(A) = K^{-1}\inf\{s | R_A(s,x) = 4\pi^2 \text{ for some } x\}. \tag{4.2}$$

Here K is that constant which makes λ(standard instanton on $\mathbb{R}^4$) $= 1$. (K depends on b; were b replaced by the characteristic function of $[0,1]$, K would be 1.) Because K will occur frequently, it will often be convenient to work with $\overline{\lambda} = K\lambda$ instead of λ. As $\overline{\lambda}$ (or λ) $\to 0$ we approach the "boundary" of $\mathscr{M}$ (in a sense we will make precise in §5).

By applying the implicit function theorem to (4.1), Donaldson shows that each sufficiently concentrated self-dual connection A has a unique well-defined center $p = p(A) \in M$. The quantities $\lambda(A)$ and $p(A)$ are characterized by the equations

$$R_A(\overline{\lambda}(A), p(A)) = 4\pi^2, \tag{4.3}$$

$$\frac{\partial R_A}{\partial p^i}(\overline{\lambda}(A), p(A)) = 0. \tag{4.4}$$

Both λ and p are gauge-invariant, so for λ_0 sufficiently small there is a well-defined map

$$\Psi \colon \mathscr{M}_{\lambda_0} \to (0, \lambda_0) \times M \tag{4.5}$$

given by $\Psi([A]) = (\lambda(A), p(A))$ on the collar $\mathscr{M}_{\lambda_0} = \{[A] \in M | \lambda(A) \leq \lambda_0\}$. Donaldson proves that Ψ is a diffeomorphism [6].

We will estimate the differential of Ψ by writing $\Psi_* = (\lambda_*, p_*)$ and examining first λ_* and then p_*.

Notation. For fixed p and λ, write $r = r(y) = \operatorname{dist}(p, y)$, and set $\gamma(y) = b(r/\overline{\lambda})$. In a fixed normal coordinate system $\{x^i\}$ centered at p we write $y^i = x^i(y)$, $\gamma_i = \partial_i\gamma$, and $\gamma_{ij} = \partial_i\partial_j\gamma$. These derivatives are supported in the annulus $\{\overline{\lambda} \leq r \leq 2\overline{\lambda}\}$ and are given by

$$\gamma_i(y) = \overline{\lambda}^{-1} b'(r/\overline{\lambda})\partial_i r = \overline{\lambda}^{-1} b'(r/\overline{\lambda}) y^i/r, \tag{4.6}$$

$$\gamma_{ij}(y) = \overline{\lambda}^{-2}\{b''(r/\overline{\lambda}) y^i y^j/r^2 + \overline{\lambda} b'(r/\overline{\lambda})[\delta_{ij}/r - y^i y^j/r^3]\}. \tag{4.7}$$

Proposition 4.1. *Let $A_t = A + t\eta + O(t^2)$ be a path of self-dual connections. Then $[\eta] = \pi_*\eta \in T_{[A]}\mathscr{M}$ satisfies*

$$\overline{\lambda}_*[\eta] = -4\lambda \frac{\int(\eta, i_{\nabla\gamma}F_A)}{\int(\nabla\gamma, \nabla(r^2))|F_A|^2}. \tag{4.8}$$

Furthermore, there exist $\lambda_0 > 0$ and $c > 0$ such that $[A] \in \mathscr{M}_{\lambda_0}$ implies

$$c^{-1} \leq -\int(\nabla\gamma, \nabla(r^2))|F_A|^2 \leq c. \tag{4.9}$$

Remark. Donaldson essentially derives this in the course of proving Corollary 17 of [6], but he makes a slight mistake (his first equality should be an inequality) that is of no consequence to his result but is important to us, corresponding to the denominator in (4.8). Therefore we give the complete proof here.

Proof. Differentiating (4.3) with respect to t, and writing $\overline{\lambda} = \overline{\lambda}(A)$, $p = p(A)$, $\overline{\lambda}_t = \overline{\lambda}(A_t)$, and $p_t = p(A_t)$, we have

$$\begin{aligned} 0 &= \frac{d}{dt} R_{A_t}(\overline{\lambda}_t, p_t)|_{t=0} \\ &= \left[\frac{d}{dt} R_{A_t}(\overline{\lambda}, p) + \frac{\partial R_A}{\partial x^i}(\overline{\lambda}, p)\frac{dp^i}{dt} + \frac{\partial R_A}{\partial s}(\overline{\lambda}, p)\frac{d\overline{\lambda}_t}{dt}\right]_{t=0}. \end{aligned}$$

The middle term vanishes by (4.4), so

$$\overline{\lambda}_*[\eta] = \frac{d\overline{\lambda}_t}{dt}|_{t=0} = \frac{-(d/dt)R_{A_t}(\overline{\lambda}, p)|_{t=0}}{(\partial R_A/\partial s)(\overline{\lambda}, p)}. \tag{4.10}$$

Since the curvature of A_t is $F_t = F_A + t d_A \eta + O(t^2)$, the numerator on the right-hand side of (4.10) is, from (4.1),

$$\frac{d}{dt} R_{A_t}(\overline{\lambda}, p)|_{t=0} = 2\int \gamma(d_A\eta, F_A) = 2\int (\eta, d_A^*(\gamma F_A)),$$

with

$$d_A^*(\gamma F_A) = \gamma d_A^* F_A - i_{\nabla\gamma} F_A = -i_{\nabla\gamma} F_A. \tag{4.11}$$

Similarly, differentiating (4.1) with respect to s gives

$$\frac{\partial R_A}{\partial s}(\overline{\lambda}, p) = -\int r\overline{\lambda}^{-2} b'(r/\overline{\lambda})|F_A|^2 = -\frac{1}{2\overline{\lambda}}\int (\nabla\gamma, \nabla r^2)|F_A|^2.$$

Multiplying the numerator and the denominator in (4.10) by $\overline{\lambda}$, we obtain (4.8). Moreover, by Lemma 3.5 and the monotonicity of b,

$$\int (\nabla\gamma, \nabla r^2)|F_A|^2 = -\int \frac{r}{\overline{\lambda}} b'(r/\overline{\lambda})|F_A|^2$$

approaches a (strictly) negative constant, uniformly in A, as $\overline{\lambda} \to 0$. Hence (4.9) follows as well. q.e.d.

The next task is to derive formulas for p_* analogous to those just obtained for λ_*. Up to this point, we have fixed $p \in M$ and used normal coordinates centered at p; we must now vary the center point p. The next lemma deals with this geometric complication.

Fix a number r_0 with $2r_0 <$ injectivity radius of M. Let $\{x^i\}$ be normal coordinates on the ball $B_{r_0}(p)$ centered at p. Given a point $y \in B_{r_0}(p)$ with coordinates $y^i = x^i(y)$ let $\rho(\cdot) = \operatorname{dist}(y, \cdot)$ be the distance function from y and set $r = \rho(p)$. With this notation we have

Lemma 4.2. *There exists a constant C depending only on the geometry of M such that, for each $p \in M$ and $y \in B_{r_0}(p)$,*

(a) $\frac{\partial \rho}{\partial x^i}(p) = -y^i/r$,

(b) $\left|\frac{\partial^2 \rho}{\partial x^i \partial x^j}(p) - \frac{\partial_{ij}}{r} + \frac{y^i y^j}{r^3}\right| \le Cr$.

Proof. Let $\sigma(t)$ be the geodesic from $y = \sigma(0)$ to $p = \sigma(1)$ parametrized proportionally to arclength; i.e., if $T = \sigma'(t)$ is the tangent vector to σ, then $\|T\| = \rho(p) = r$. Let $v, u_i \in T_y M$ be such that $\exp_y(v) = p$ and $(\exp_y)_{*v}(u_i) = (\partial/\partial x^i)(p)$. For each i we may consider the one-parameter variation

$$\sigma_s^i(t) = \exp_y(t(v + su_i)). \tag{4.12}$$

Since σ is varied through geodesics, the variational vector field $U_i = (\partial/\partial s)\sigma_s^i$ is a Jacobi field along σ, and satisfies $U_i(0) = 0$ and $U_i(1) = (\partial/\partial x^i)(p)$. The formula for first variation of arclength [4, equation 1.3] gives

$$\left.\frac{\partial \rho}{\partial x^i}\right|_p = \rho(p)^{-1}(U_i(1), T(1)), \tag{4.13}$$

since σ is a geodesic. But $T = -y^i \partial/\partial x^i$ in normal coordinates and $g_{ij}(p) = \delta_{ij}$, so we obtain (a).

Let V_i denote the vector field $\partial/\partial x^i$ on $B_{r_0}(p)$. If we replace p in (4.13) by an arbitrary point $q \in B_{r_0}(p)$, we obtain

$$V_i(p)|_q = \rho(q)^{-1}(V_i, T)|_q. \tag{4.14}$$

Applying the vector field U_j to (4.14) gives

$$U_j V_i(\rho)|_q = \{-\rho^{-2} U_j(\rho)(V_i, T) + \rho^{-1}[(\nabla_{U_j} V_i, T) + (V_i, \nabla_{U_j} T)]\}|_q.$$

When $q = p$, we obtain $U_j = V_j = \partial/\partial x^j$, $\nabla_{U_j} V_i = \Gamma_{ji}^k \partial/\partial x^k|_p = 0$, and $\nabla_{U_j} T = \nabla_T U_j$ since U_j, T are σ_*^j of $\partial/\partial s, \partial/\partial t$. Therefore, using part (a), we have

$$\left.\frac{\partial^2 \rho}{\partial x^i \partial x^j}\right|_p = -r^{-3} y^i y^j + r^{-1}(V_i, \nabla_T U_j)|_p. \tag{4.15}$$

But, from (4.12) it follows that we may write $U_j = tW$, where $W = W^k \partial/\partial z^k$ is some vector field with *constant* coefficients with respect to the normal coordinates $\{z^i\}$ centered at y; in these coordinates, $W = U_j(1)$. Hence, in these coordinates, writing $T = p^i \partial/\partial z^i$, we have

$$\begin{aligned}\nabla_T U_j &= W + tp^i W^k \nabla_{\partial/\partial z^i} \frac{\partial}{\partial z^k} \\ &= U_j(1) + tp^i W^k \Gamma_{ki}^l \frac{\partial}{\partial z^l}.\end{aligned}$$

Now set $t = 1$. Since $\rho(p) = r$ and $|\Gamma^l_{ki}(p)| \le Cr$, it follows that

$$|\nabla_T U_j - U_j(1)| \le Cr^2|W| = Cr^2|U_j(1)| = Cr^2, \tag{4.16}$$

noting that $U_j(1) = (\partial/\partial x^j)(p)$. Finally, $(V_i(p), U_j(1)) = (V_i(p), V_j(p)) = \delta_{ij}$, so statement (b) follows from (4.15) and (4.16).

Corollary 4.3. *Let Z be any vector field on M and let $A_t = A + t\pi_A i_Z F_A + O(t^2)$ be a path of self-dual connections. Let p and λ denote the center and scale of A. There exists a constant c, independent of A and Z, such that if $\{x^i\}$ are normal coordinates centered at p, then*

$$\overline{\lambda}^2 \left| \frac{d}{dt}\left(\frac{\partial R_{A_t}}{\partial x^i}(\overline{\lambda}, p)\right)_{t=0} - \int (Z, \nabla\gamma_i)|F_A|^2 \right| \le c\|(1-\pi_A)i_Z F_A\|.$$

Proof. Write $F = F_A$. Using Lemma 4.2 and (4.6), we have

$$\frac{\partial}{\partial x^i} b(r/\overline{\lambda}) = -\overline{\lambda}^{-1} b'(r/\overline{\lambda}) y^i/r = -\gamma_i.$$

Since $F_t = F + td_A\pi_A i_Z F + O(t^2)$ we find

$$\frac{d}{dt}\left(\frac{\partial R_{A_t}}{\partial x^i}(\overline{\lambda}, p)\right)\Big|_{t=0} = -\frac{d}{dt}\int \gamma_i |F_t|^2 = -2\int \gamma_i (d_A\pi_A i_Z F, F).$$

Integrating by parts and using (4.11), the last expression becomes

$$\int 2(\pi_A i_Z F, i_{\nabla\gamma_i} F).$$

Writing $\pi_A i_Z F = i_Z F - (1-\pi_A)i_Z F$, Lemma 3.4 shows that the integrand above equals $(Z, \nabla\gamma_i)|F|^2 - 2((1-\pi_A)i_Z F, i_{\nabla\gamma_i} F)$. Hence

$$\overline{\gamma}^2 \left| \frac{d}{dt}\left(\frac{\partial R_{A_t}}{dx^i}\right) - \int (Z, \nabla\gamma_i)|F|^2 \right| \le 2\|(1-\pi_A)i_Z F\|_2 \|\overline{\lambda}^2 i_{\nabla\gamma_i} F\|_2.$$

Now γ_i has support on $\{\overline{\lambda} \le r \le 2\overline{\lambda}\}$, where

$$\overline{\lambda}^2|\nabla\gamma_i| = \left| \left(b''\left(\frac{r}{\overline{\lambda}}\right)\frac{y_i}{r} - \overline{\lambda} b'\left(\frac{r}{\overline{\lambda}}\right) y_i/r^2 \right)\nabla r + \overline{\lambda} b'\left(\frac{r}{\overline{\lambda}}\right)\frac{1}{r}\nabla y^i \right|$$

is bounded by a constant, so $\|\overline{\lambda}^2 i_{\nabla\gamma_i} F\|_2 \le C\|F\|_2 \le C$, and the result follows.

We will also need a statement concerning the Hessian of R_A.

Lemma 4.4. *Let H_{ij} denote the Hessian $(\partial^2 R_A/\partial x^i\partial x^j)(\lambda, p)$. For λ sufficiently small, $\lambda^2 H_{ij}$ is uniformly invertible; i.e., there exist $\lambda_0, c > 0$ such that $[A] \in \mathscr{M}_{\lambda_0}$ implies*

$$c^{-1}|\xi|^2 \le \lambda^2 H_{ij}\xi^i\xi^j \le c|\xi|^2 \quad \forall \xi \in T_pM. \tag{4.17}$$

Moreover,

$$\lambda^2 H_{ij} = \int \lambda^2 \gamma_{ij} |F_A|^2 + O(\lambda^2), \tag{4.18}$$

where γ_{ij} *is as in* (4.7).

Remark. In fact we will show that $\lambda^2 H_{ij} \to \text{const} \cdot \delta_{ij}$ as $\lambda \to 0$.

Proof. Write (4.1) as

$$R_A(s,p) = \int b\left(\frac{\rho_y(p)}{\overline{\lambda}}\right) |F_A|^2(y)\, dv_g(y),$$

where $\rho_y(\cdot) = \text{dist}(y, \cdot)$. Letting $\partial_i = \partial/\partial x^i$, and writing ρ for ρ_y, we differentiate twice to obtain

$$\overline{\lambda}^2 H_{ij} = \int \left\{ b''\left(\frac{\rho(p)}{\overline{\lambda}}\right) \partial_i\rho\partial_j\rho + \overline{\lambda} b'\left(\frac{\rho(p)}{\lambda}\right)\partial_i\partial_j\rho \right\} |F_A|^2.$$

By Lemma 4.2 and (4.7) we can write this integrand as $\{\overline{\lambda}^{-2}\gamma_{ij} + \overline{\lambda}\phi_{ij}\}|F|^2$, where $|\phi_{ij}| \le Cr$ and $\text{supp}(\phi_{ij}) \subset B_{2\overline{\lambda}}(p)$. Hence

$$\int \overline{\lambda}\phi_{ij}|F|^2 \le c\overline{\lambda}^2 \int |F|^2 \le c\overline{\lambda}^2,$$

and we obtain (4.17).

To prove (4.18) we substitute (4.7) into Lemma 3.5 and find that

$$\begin{aligned}
\lim_{\lambda\to 0}\left[\int \overline{\lambda}^{-2}\gamma_{ij}|F_A|^2\right] &= 24\pi^2\delta_{ij}\int_0^\infty (t^3 b''(t/K) + 3Kb'(t,K)t^2)(1+t^2)^{-4}\,dt \\
&= 24\pi^2\delta_{ij}\int_0^\infty (1+t^2)^{-4}\frac{d}{dt}(Kt^3 b'(t/K))\,dt \\
&= -24\pi^2\delta_{ij}\int_0^\infty Kt^3 b'(t/K)\frac{d}{dt}(1+t^2)^{-4}\,dt \\
&= c_1\delta_{ij}.
\end{aligned}$$

Since $b' \le 0$ and $b' < 0$ somewhere, c_1 is strictly positive. Hence (4.19) follows from (4.17). q.e.d.

We finally obtain a statement for p_* analogous to the one for λ_* given in Proposition 4.1.

Proposition 4.5. *There exist* $c, \lambda_0 > 0$ *such that if* $[A] \in \mathscr{M}_{\lambda_0}$, *and* $A_t = A + t\pi_A i_Z F_A + O(t^2)$ *is a path of self-dual connections, then for any vector field* Z *on* M,

$$\begin{aligned}
\left|(p_*)^j(\pi_A i_Z F_A) + (\overline{\lambda}^2 H)^{-1}_{ij}\left[\int \overline{\lambda}^2 (Z, \nabla\gamma_i)|F_A|^2 + m_i\overline{\lambda}_*(\pi_A i_Z F_A)\right]\right| \\
\le c\|(1-\pi_A)i_Z F_A\|,
\end{aligned} \tag{4.19}$$

where

$$m_i = \overline{\lambda}^{-1} \int b''(r/\overline{\lambda}) y^i |F_A|^2. \tag{4.20}$$

Here $\overline{\lambda} = \overline{\lambda}(A)$, γ_i is given by (4.6), $(p_*)^j$ are the components of the image of p_* in $T_{p(A)}M$ with respect to normal coordinates $\{x^i\}$ centered at $p(A)$, and H_{ij} is the Hessian $(\partial^2 R_A/\partial x^i \partial y^i)(\overline{\lambda}, p)$.

Proof. Write $\overline{\lambda}_t = \lambda(A_t)$, $p_t = p(A_t)$, $\overline{\lambda} = \overline{\lambda}(A)$, and $p = p(A)$. Differentiating (4.4) yields

$$\begin{aligned} 0 &= \frac{d}{dt}\left(\frac{\partial R_{A_t}}{\partial x^i}(\overline{\lambda}_t, p_t)\right)_{t=0} \\ &= \left[\frac{d}{dt}\left(\frac{\partial R_{A_t}}{\partial x^i}(\overline{\lambda}, p)\right) + \frac{\partial^2 R_A}{\partial x^i \partial x^j}(\overline{\lambda}, p)\frac{dp_t^j}{dt} + \frac{\partial^2 R_A}{\partial s \partial x^i}(\overline{\lambda}, p)\frac{d\overline{\lambda}_t}{dt}\right]_{t=0}. \end{aligned}$$

By Lemma 4.4, H_{ij} is invertible, so we can solve this to obtain

$$\begin{aligned} \left.\frac{dp_t^i}{dt}\right|_{t=0} = -(\overline{\lambda}^2 H)_{ij}^{-1} \Bigg\{ &\overline{\lambda}^{-2} \left.\left(\frac{d}{dt}\frac{\partial R_{A_t}}{\partial x^i}\right)\right|_{t=0} (\overline{\lambda}, p) \\ &+ \overline{\lambda}^{-2} \frac{\partial^2 R_A}{\partial s \partial x^i}(\overline{\lambda}, p) \overline{\lambda}_*(\pi_A i_Z F_A) \Bigg\}. \end{aligned} \tag{4.21}$$

Differentiating (4.1) and using Lemma 4.2, we have

$$\begin{aligned} \overline{\lambda}^{-2} \frac{\partial^2 R_A}{\partial s \partial x^i} &= \int (b'' r/\overline{\lambda}^2 + b' r^2/\overline{\lambda}) y^i |F_A|^2 \\ &= m_i + \int b'(r^2/\overline{\lambda}) y^i |F_A|^2. \end{aligned} \tag{4.22}$$

But this last term vanishes, since by (4.1), (4.4) and Lemma 4.2

$$0 = \int \frac{\partial}{\partial x^i}(\gamma_{\overline{\lambda}}(x, y))|F|^2 = -\overline{\lambda}^{-1} \int b'(r/\overline{\lambda}) y^j |F|^2. \tag{4.23}$$

Next, we may use Corollary 4.3 and (4.17) to replace $\frac{d}{dt}(\partial R_{A_t}/\partial x^i)$ in (4.21) by $\int (Z, \nabla \gamma_i)|F_A|^2$ at the cost of the term $c\|(1 - \pi_A) i_Z F_A\|$ appearing on the right-hand side of (4.19). Thus substituting (4.22) and (4.23) into (4.21) we are done.

5. The metric on the collar

We now combine the results of the previous two sections to show that the metric on the collar is asymptotic, in a C^0 sense, to a product metric (Theorem II). This leads to a description of the metric completion of $\mathscr{M}$ (Theorems III and IV).

Definition 5.1. Given $[A] \in \mathscr{M}_{\lambda_0}$ with center p and scale λ, and given $(a_0, \mathbf{a}) \in \mathbb{R} \times T_pM$, we set

$$Z_{(a_0,\mathbf{a})} = \nabla(\beta(\tfrac{1}{2}\lambda^{-1}a_0r^2 + f_{\mathbf{a}})),$$

and define $\alpha_A \colon \mathbb{R} \times T_pM \to \hat{T}_A$ by

$$\alpha_A(a_0, \mathbf{a}) = -i_{Z_{(a_0,\mathbf{a})}}F_A,$$

where $\beta, r, f_{\mathbf{a}}$ and $\hat{T}_A$ are as defined in §3.

Observe that α_A depends on $\mathscr{G}$-equivariantly on A, so defines a bundle map α taking $T_{(\lambda,p)}(\mathbb{R} \times M)$ to $\hat{T}_{[A]}$ where $[A] = \Psi^{-1}(\lambda, p)$. Since the harmonic projection $\pi_A \colon \hat{T}_A \to T_A$ induces a bundle map $\pi \colon \hat{T}\mathscr{M} \to T\mathscr{M}_{\lambda_0}$, we can consider the composition

$$T_{(\lambda,p)}(\mathbb{R} \times M) \xrightarrow{\alpha} \hat{T}_{[A]} \xrightarrow{\pi} T\mathscr{M}.$$

Combining this with the differential of the collar diffeomorphism Ψ (cf. (4.5)) gives a diagram

$$\begin{array}{ccc} TM_{\lambda_0} & \xrightarrow{\Psi_*} & T((0,\lambda_0) \times M) \\ & {}_{\pi}\nwarrow \quad \swarrow_{\alpha} & \\ & \hat{T}\mathscr{M}_{\lambda_0} & \end{array} \tag{5.1}$$

We will use the results of §§3 and 4 to show that $\pi \circ \alpha$ is an approximate inverse to Ψ_* in the sense that (5.1) commutes up to terms which are $O(\lambda^{1-\delta})$.

Proposition 5.2. *There exist $c, \lambda_0 > 0$ such that $[A] \in \mathscr{M}_{\lambda_0}$ implies*

(a) $|\lambda_*[\pi_A\alpha_A(a_0, \mathbf{a})] - a_0| \le c(|a_0| + |\mathbf{a}|)\lambda^{1-\delta}$,

(b) $|p_*[\pi_A\alpha_A(a_0, \mathbf{a})] - a| \le c(|a_0| + |\mathbf{a}|)\lambda^{1-\delta}$.

In other words, $\Psi_* \circ \pi \circ \alpha = \mathrm{Id} \cdot (1 + O(\lambda^{1-\delta}))$.

Proof. (a) Write $F = F_A$ and take $\eta = \pi_A\alpha_A(a_0, \mathbf{a}) = -i_ZF + (1-\pi_A)i_ZF$ in Proposition 4.1. When λ is sufficiently small, $\beta \equiv 1$ on $\mathrm{supp}(\gamma)$, so using Lemma 3.4 we may replace $(i_ZF, i_{\nabla\gamma}F) = \frac{1}{2}(Z, \nabla\gamma)|F|^2$ by $\frac{1}{4}a_0(\nabla r^2, \nabla\gamma)|F|^2$ $+ \frac{1}{2}a_i(\nabla x^i, \nabla\gamma)|F|^2$. Hence the numerator of (4.8) is

$$\begin{aligned} \int (\eta, i_{\nabla\gamma}f) = &-\frac{1}{4}a_0 \int (\nabla r^2, \nabla\gamma)|F|^2 - \frac{1}{2}a_i \int (dx^i, d\gamma)|F|^2 \\ &+ \int ((1 - \pi_A)i_ZF, i_{\nabla\gamma}F). \end{aligned} \tag{5.2}$$

Now $(dx^i, d\gamma)(y) = \overline{\lambda}^{-1}b'(r/\overline{\lambda})g^{ij}y^j/r = \overline{\lambda}^{-1}b'(r/\overline{\lambda})y^j/r$, so by (4.23), the middle integral above vanishes. Therefore, multiplying (5.2) by $-4\overline{\lambda}$, dividing by the denominator of (4.8), and applying (4.9), we see that

$$\begin{aligned} |\lambda_*[\eta] - a_0| &\le c\lambda|\langle(1 - \pi_A)i_ZF, i_{\nabla\gamma}F\rangle_2| \\ &\le c\|(1 - \pi_A)i_ZF\|_2\|\lambda i_{\nabla\gamma}F\|_2. \end{aligned}$$

But $|\nabla\gamma| \le c\lambda^{-1}$ by (4.6), so $\|\lambda i_{\nabla\gamma} F\|_2 \le c\|F\|_2 \le c$. Part (a) now follows from Proposition 3.7.

(b) We apply Proposition 4.5. Again take λ small enough that $\beta \equiv 1$ on $\operatorname{supp}(\gamma)$. Then, on $\operatorname{supp}(\gamma)$,

$$\begin{aligned}(Z, \nabla\gamma_i) &= (d(\tfrac{1}{2}\lambda^{-1}a_0 r^2 + a_j y^j), d\gamma_i) \\ &= ((\lambda^{-1}a_0 y^j + a_j)\, dy^j, \gamma_{ik} dy^k) \\ &= (\lambda^{-1}a_0 y^j + a_j)(\gamma_{ij} + \gamma_{ik}(g^{jk} - \delta_{jk})).\end{aligned}$$

But $g^{jk} - \delta_{jk} = O(r^2)$, $|\gamma_{ik}| \le c\lambda^{-2}$, and $\overline{\lambda} \le r \le 2\overline{\lambda}$ on $\operatorname{supp}(\gamma)$, so

$$(Z, \nabla\gamma_i) = (\lambda^{-1}a_0 y^j + a_j)\gamma_{ij} + O(|a_0| + |\mathbf{a}|).$$

Thus, since $\int |F|^2 \le c$,

$$\int \overline{\lambda}^2 (Z, \nabla\gamma_i)|F|^2 = a_j \int \overline{\lambda}^2 \gamma_{ij}|F|^2 + \frac{a_0}{\lambda}\int y^j \overline{\lambda}^2 \gamma_{ij}|F|^2 + O((|a_0| + |\mathbf{a}|)\lambda^2).$$

By (4.18), the first term on the right is $\overline{\lambda}^2 H_{ij} a_j + O(|\mathbf{a}|\lambda^2)$. From (4.7) we compute $\overline{\lambda}^2 y^j \gamma_{ij} = b''(r/\overline{\lambda}) y^i$, so the second term on the right is $a_0 K m_i$, where m_i is as in (4.20), and K is as in (4.2). Therefore

$$\begin{aligned}&\int \overline{\lambda}^2 (Z, \nabla\gamma_i)|F|^2 + m_i \overline{\lambda}_*(\pi_A i_Z F_A) \\ &\qquad = \overline{\lambda}^2 H_{ij} a_j + K m_i(\lambda_*(\pi_A i_Z F_A) + a_0) + O((|a_0| + |\mathbf{a}|)\lambda^2).\end{aligned} \tag{5.3}$$

Now $|m_i|$ is bounded (in fact Lemma 3.5 shows that $m_i \to 0$ as $\lambda \to 0$), so we may use part (a) to bound the second term on the right-hand side of (5.3). We can then combine (4.19), (5.3), (4.17) and Proposition 3.7 to obtain

$$|(p_*)^j(\pi_A i_Z F_A) - a_j| \le c(|a_0|\lambda^{1-\delta} + |\mathbf{a}|\lambda),$$

and (b) then follows from the definition of α_A.

Proof of Theorem II. Since Proposition 3.7 shows that $\pi_A \circ \alpha_A$ is an isomorphism, it suffices to prove (0.5) for all W of the form $\pi_A i_Z F$, where $i_Z F = \alpha_A(a_0, \mathbf{a})$. But then

$$\Psi_* W = (a_0, \mathbf{a}) + O((|a_0| + |\mathbf{a}|)\lambda^{1-\delta})$$

by Proposition 5.2, and therefore

$$|(\Psi^*\mathscr{k})(W,W) - 4\pi^2(2a_0^2 + |\mathbf{a}|^2)| \le c\lambda^{1-\delta}(|a_0|^2 + |\mathbf{a}|^2) \le c\lambda^{1-\delta}(\Psi^*\mathscr{k})(W,W).$$

On the other hand, Propositions 3.6 and 3.7 imply that for λ_0 sufficiently small,

$$|\mathscr{g}(W,W) - 4\pi^2(2a_0^2 + |\mathbf{a}|^2)| \le \tfrac{1}{2}\varepsilon(\Psi^*\mathscr{k})(W,W).$$

Therefore taking λ_0 small enough, we obtain $|\mathscr{g}(W,W) - (\Psi^*\mathscr{h})(W,W)| \leq \varepsilon(\Psi^*\mathscr{h})(W,W)$ and the result follows.

Proof of Theorems III *and* IV. Let $[A] \in \mathscr{M}$ be a reducible connection and let $\overline{U}$ be the closure of a neighborhood of $[A]$ as in Theorem I. The form (0.2) of the metric shows that the completion of $\overline{U}^* = \overline{U} - \{[A]\}$ is obtained simply by putting the vertex $[A]$ back in; i.e., the completion of $\overline{U}^*$ is $\overline{U}$ itself. It follows that $\overline{\mathscr{M}^*}$ is identical to the completion $\overline{\mathscr{M}}$ of $(\mathscr{M}, \mathscr{g})$.

Now let $\{[A_i]\}$ be a Cauchy sequence which does not converge in $\mathscr{M}$. By Uhlenbeck's Compactness Theorem [8, Theorem 8.36] we have $\lambda_i = \lambda([A_i]) \to 0$, so the sequence eventually lies in the set $\mathscr{M}_{\lambda_0}$ with λ_0 as in Theorem II. The inequality (0.5) then implies that for i, j sufficiently large

$$\operatorname{dist}([A_i], [A_j])^2 \geq 2\pi^2(2|\lambda_i - \lambda_j|^2 + \operatorname{dist}(p_i, p_j)^2),$$

where $p_i = p([A_i]) \in M$. Hence $\{p_i\}$ is also Cauchy, converging to some $p_0 \in M$. It follows that the set of equivalence classes of Cauchy sequences not converging in $\mathscr{M}$ is in 1-1 correspondence with M, and that $\overline{\mathscr{M}}$ is homeomorphic to $\mathscr{M} \cup_\Psi ([0, \lambda_0) \times M)$ (which is compact by Uhlenbeck's theorem). Thus the metric topology on $\overline{\mathscr{M}}$ is independent of the choice of collar map Ψ (i.e., independent of the details of the cut-off function b used to define λ and p), so $\overline{\mathscr{M}}$ is a topological manifold-with-boundary in a natural way. The induced smooth structure on $\mathscr{M} \cup_\Psi ([0, \lambda_0) \times M)$ is also independent of the choice of Ψ, so $\overline{\mathscr{M}}$ inherits a natural smooth structure. The function λ on $\mathscr{M}$ extends smoothly to $\overline{\mathscr{M}}$ and we have $\partial M = \overline{\mathscr{M}} - \mathscr{M} = \{\lambda = 0\}$. In particular, $\mathscr{M}$ is incomplete, since instantons of scale size $\lambda = 0$ do not exist in $\mathscr{M}$.

We can define a Riemannian metric $\mathscr{g}$ on $\overline{\mathscr{M}}$ by declaring $\mathscr{g}$ to be $4\pi^2(2d\lambda^2 \oplus g)$ along $\partial\overline{\mathscr{M}}$ and to be the L^2-induced metric on the interior. Theorem II shows that $\mathscr{g}$ is continuous. The restriction to $\mathscr{g}$ to vectors tangent to $\partial\overline{\mathscr{M}}$ is $4\pi^2 g$, independent of the choice of Ψ again. The same theorem also shows that if σ is a curve lying in $\mathscr{M}_{\lambda_0}$ and $l_1(\sigma), l_2(\sigma)$ are the lengths of σ with respect to $\mathscr{g}$, $\Psi^*\mathscr{h}$ respectively, then $1 - \varepsilon \leq l_1(\sigma)/l_2(\sigma) \leq 1 + \varepsilon$. Since the distance to $\partial\overline{\mathscr{M}}$ with respect to $\Psi^*\mathscr{h}$ is exactly $\sqrt{8\pi^2}\,\lambda$, we conclude that the distance with respect to $\mathscr{g}$ is asymptotic to $\sqrt{8\pi^2}\,\lambda$. q.e.d.

Theorem II shows that the metric $\mathscr{g}$ is C^0-asymptotic to the product metric $\Psi^*\mathscr{h}$ as one approaches $\partial\overline{\mathscr{M}}$. It is natural to ask whether this is true for the derivatives of $\mathscr{g}$; i.e., is $\mathscr{g}$ C^l-asymptotic to a product metric? For $l = 1$ this would imply that $\partial\overline{\mathscr{M}}$ is a totally geodesic submanifold of $\overline{\mathscr{M}}$, and for $l = 2$ it would additionally imply that the curvature of $\mathscr{M}$ is asymptotic to the curvature of the cylinder $[0,1] \times M$. For the case $M = S^4$ the results of [9] give a complete answer.

Proposition 5.3. *Let $\mathscr{M}_1$ be the moduli space of $k = 1$ instantons on S^4 with its standard metric. Then*

(a) *$\partial\overline{\mathscr{M}}_1 \subset \overline{\mathscr{M}}_1$ is totally geodesic and is C^∞-isometric to the (round) 4-sphere of radius 2π (which has constant curvature $\frac{1}{4}/\pi^2$),*

(b) *along $\overline{\mathscr{M}}_1$, the sectional curvature of $\mathscr{M}_1$ in any two-plane spanned by one vector tangent and one vector normal to $\partial\overline{\mathscr{M}}_1$ is $\frac{3}{16}/\pi^2$.*

Proof. Part (a) is contained in Corollary C of [9]. For part (b) we use the fact that $\mathscr{M}_1$ is radially symmetric and conformally flat; i.e., its metric takes the form $\psi^2(\rho)\sum_{i=1}^5(dx^i)^2$ ([9, equation 6.5]; here $\rho = |x|$). From this we deduce that the scalar curvature of $\overline{\mathscr{M}}_1$ at the boundary is the limiting value s_∞ of $s = -4\psi^{-3}[2\psi'' + 8\rho^{-1}\psi' + \psi^{-1}(\psi')^2]$, and from the formulas for ψ in [9] one calculates $s_\infty = \frac{9}{2}/\pi^2$. Then part (a), together with radial symmetry, implies (b). (A completely different derivation from that in [9] of the metric on M_1, as well as a different proof of this proposition, appears in [5].) q.e.d.

Proposition 5.3 essentially shows that for $M = S^4$ the metric on the collar is asymptotic to a product in the C^1 topology, but not in the C^2 topology. At present the authors do not know if this is the situation for other M; in particular whether it is generally true that $\partial\overline{\mathscr{M}}$ is totally geodesic.

Appendix

Our description of the geometry of the cones $\mathscr{M}$ was obtained in §2 by reducing the problem to one of finite-dimensional Riemannian geometry. The results of that section can also be obtained directly (indeed, our original proof of Theorem 1 was completely analytic). Although the geometric approach is easier and conceptually clearer, it is important to bear in mind the fact that each statement about the geometry of $\tilde{\mathscr{M}}$ is equivalent to a statement about gauge fields on M. In this appendix we illustrate the analytic approach by giving a second proof of Lemma 2.1. (This proof generalizes that of Theorem 4.9 in [8].)

Lemma A.1. *Let $J = [\Phi, \cdot]$ be the infinitesimal isotropy representation (2.3) and assume $H^1(M;\mathbb{R}) = 0$. Then for any $\eta \in T_{[A]}\tilde{\mathscr{M}} = \tilde{\mathscr{H}}_A$ (defined by (1.13)) we have*

(a) $\langle \eta, \Phi df\rangle = 0\ \forall f \in C^\infty(M)$,

(b) *The 1-form $\omega = (\eta, \Phi) \in L^2_s(T^*M)$ vanishes pointwise,*

(c) *$J^2 = -\mathrm{Id}$, and $T_{[A]}\tilde{\mathscr{M}}$ decomposes as stated in Lemma 2.1.*

Proof. (a) The Laplacian $\Box_0 = d^*d + I$ on functions is an invertible elliptic operator, so given $f \in C^\infty(M)$, there is a unique $h \in C^\infty(M)$ with $\Box_0^s h = f$.

Write $h = \overline{h} + h(x_0)$, where $\overline{h} \in C^\infty(M)$ vanishes at x_0. Then $d\Box_0^s\overline{h} = d(f - h(x_0)) = df$, and hence

$$\begin{aligned}\langle \eta, \Phi df\rangle &= \langle \eta, \Phi d\Box_0^s \overline{h}\rangle = \langle \eta, d_A(\Box_A^0)^s(\Phi\overline{h})\rangle \\ &= \langle \eta, (\Box_A^1)^s d_A(\Phi\overline{h})\rangle.\end{aligned}$$

Since $\eta \in \tilde{\mathscr{H}}_A$ is a solution of the distributional equation (1.6), this becomes

$$\langle \eta, \Phi df\rangle = (v, \Phi\overline{h})(x_0)$$

and therefore vanishes since $\overline{h}(x_0) = 0$.

(b) For any $f \in C^\infty(M)$ and $\omega = (\eta, \Phi)$ we have, by (a),

$$\langle d^*\omega, f\rangle = \langle \omega, df\rangle = \langle \eta, \Phi df\rangle = 0,$$

which implies $d^*\omega \equiv 0$. Furthermore, $d^-\omega = (d_A^-\eta, \Phi) = 0$ because $d_A^-\eta = 0$ (η is tangent to the set of self-dual connections). Hence $d\omega = *d\omega$, so $d^*d\omega = - * dd\omega = 0$. Thus ω is a harmonic 1-form and, since $H^1(M;\mathbb{R}) = 0$, we conclude that $\omega \equiv 0$.

(c) The fact that $J^2 = -\mathrm{Id}$ now follows from (b), and the normalizations $|\Phi|^2 = 1$ and $[a,[b,c]] = (a,c)b - (a,b)c$ discussed at the beginning of §2. The decomposition of $T_{[A]}\tilde{\mathscr{M}}$ follows as in the proof of Lemma 2.1.

References

[1] M. F. Atiyah, R. Bott & V. K. Patodi, *On the heat equation and the index theorem*, Invent. Math. **19** (1973) 279–330.

[2] M. F. Atiyah, N. J. Hitchin & I. M. Singer, *Self-duality in four dimensional Riemannian geometry*, Proc. Roy. Soc. London Ser. A **362** (1978) 425–461.

[3] T. Aubin, *Non-linear analysis on manifolds: Monge-Ampére equations*, Springer, New York, 1982.

[4] J. Cheeger & D. Ebin, *Comparison theorems in Riemannian geometry*, North-Holland, Amsterdam, 1975.

[5] H. Doi, Y. Matsumoto & T. Matumoto, *An explicit formula of the metric on the moduli space of BPST-instantons over S^4*, A Fête of Topology, Academic Press, New York, 1988, 543–556.

[6] S. K. Donaldson, *An application of gauge theory to four dimensional topology*, J. Differential Geometry **18** (1983) 279–315.

[7] ——, *Connections, cohomology and the intersection forms of 4-manifolds*, J. Differential Geometry **24** (1986) 275–341.

[8] D. S. Freed & K. K. Uhlenbeck, *Instantons and four-manifolds*, Springer, New York, 1984.

[9] D. Groisser & T. H. Parker, *The Riemannian geometry of the Yang-Mills moduli space*, Comm. Math. Phys. **112** (1987) 663–689.

[10] H. B. Lawson, *The theory of gauge fields in four dimensions*, Amer. Math. Soc., Providence, RI, 1986.

[11] C. B. Morrey, *Multiple integrals in the calculus of variations*, Springer, Berlin, 1966.

[12] T. H. Parker, *Gauge theories on four dimensional Riemannian manifolds*, Comm. Math. Phys. **85** (1982) 1–40.

[13] T. H. Parker & S. Rosenberg, *Invariants of conformal Laplacians*, J. Differential Geometry **25** (1987) 199–222.

[14] R. T. Seeley, *Complex powers of an elliptic operator*, Proc. Sympos. Pure Math., Vol. **10**, Amer. Math. Soc., Providence, RI, 1967, 288–307.

[15] C. H. Taubes, *Self-dual connections on non-self-dual 4-manifolds*, J. Differential Geometry **17** (1982) 139–170.

[16] K. K. Uhlenbeck, *Connections with L^p bounds on curvature*, Comm. Math. Phys. **83** (1982) 31–42.

STATE UNIVERSITY OF NEW YORK, STONY BROOK
BRANDEIS UNIVERSITY