J. DIFFERENTIAL GEOMETRY
18 (1983) 279–315

AN APPLICATION OF GAUGE THEORY TO FOUR DIMENSIONAL TOPOLOGY

S. K. DONALDSON

I.1 Introduction

This paper contains a detailed account of a result in 4-manifold topology, announced in [7], which is proved by analytical and geometrical methods.

It is a consequence of a theorem of J. H. C. Whitehead [22] that the homotopy type of a closed simply connected 4-manifold is entirely determined by the cup-square

$$Q\colon H^2(X;\mathbf{Z}) \to H^4(X;\mathbf{Z}) \cong \mathbf{Z}.$$

If we fix an orientation this becomes a quadratic form on the free Abelian group $H^2(X;\mathbf{Z})$ with determinant ± 1, realised dually on homology as the "intersection form". Many writers have discussed the problem of finding which forms may arise from 4-manifolds of various kinds [11], [12].

Very recently M. H. Freedman has shown [8] that any form of determinant ± 1 may be realised by a simply connected topological 4-manifold. Moreover he proves that the form, together with one extra piece of data (the Kirby-Siebenmann obstruction in $\mathbf{Z}/2$, always zero for smooth manifolds or for even forms) determines the manifold up to homeomorphism.

It has been known for 30 years that some forms cannot be realised by a *smooth* simply connected 4-manifold. We recall that one may divide the forms on the one hand into the even and odd forms (i.e., whether the form takes only even values) and on the other hand into definite and indefinite forms. Then *Rohlin's theorem* asserts that any even form coming from a smooth simply connected 4-manifold has signature divisible by 16. In particular the even definite form E_8 of rank 8 cannot occur in this way. The theorem which we prove here bears instead on the definite forms which we can take without loss of generality to be positive.

Received December 31, 1982.

Theorem 1. *Let X be a compact smooth simply connected oriented 4-manifold with the property that the associated form Q is positive definite. Then that form is equivalent, over the integers, to the standard diagonal form, so in some base*:

$$Q(u_1, u_2, \cdots, u_r) = u_1^2 + u_2^2 + \cdots + u_r^2.$$

There are many nontrivial definite forms [16, Chapter 5] so many of the manifolds constructed by Freedman cannot be given a differentiable structure. The most obvious problem remaining in this area is to discover for which values of m, n the manifold constructed by Freedman corresponding to the form

$$2mE_8 + n\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

can be smoothed. In particular when $m = 1$: the form for $n = 3$ is realised by, for example, a smooth quartic surface in $\mathbf{CP}^3$, while the form for $n = 0$ cannot, according to the theorem, come from a smooth simply connected 4-manifold. It is not yet clear whether the methods which we use here can be extended to the intermediate cases.

In another direction, the fundamental group does not seem to play a very important role in these methods, and one could hope to extend the result a long way here. It will appear that the present proof works under the weaker, but rather strange, hypothesis that there be no nontrivial representations of $\pi_1(X)$ in $SU(2)$.

Acknowledgement. The author wishes to thank Professor M. F. Atiyah for his advice, help and encouragement during the course of this work, and is also very grateful to M. J. Hopkins and C. H. Taubes for many useful conversations.

I.2. Method of proof

The bare structure of the proof of Theorem 1 is much simpler than the more technical material which makes up the bulk of this paper, so we give it now.

We will find associated to X a space $\mathfrak{M}^\sigma(X)$ which can be compactified to an orientable 5-manifold with boundary X and a certain number of point singularities, one singularity for each pair $\pm\alpha$ of solutions of

$$Q(\alpha) = 1, \quad \alpha \in H^2(X; \mathbf{Z}).$$

Call this number $n(Q)$. A neighborhood of one of these singular points in $\mathfrak{M}^\sigma(X)$ will have the form of a cone on $\mathbf{CP}^2$, so if we remove these neighborhoods, we get an orientable manifold with boundary the disjoint union of X and $n(Q)$ copies of $\mathbf{CP}^2$. We do not know, at this stage of the argument, how the orientations of the copies of $\mathbf{CP}^2$ compare.

The only arithmetical fact about quadratic forms which we use in this observation:

Lemma 2. *If Q is any positive definite quadratic form over $\mathbf{Z}$, then*

$$n(Q) \leqslant \operatorname{rank}(Q)$$

with equality if and only if Q is equivalent to the standard form.

Proof. By induction on $r = \operatorname{rank}(Q)$. If α is any solution of $Q(\alpha) = 1$, we may split:

$$\mathbf{Z}^r = \mathbf{Z}.\alpha \oplus \alpha^{\perp},$$
$$\beta \to (\beta.\alpha)\alpha \oplus (\alpha - (\beta.\alpha)\alpha),$$

and, because Q is definite, $n(Q) = n(Q|_{\alpha^{\perp}}) + 1$, $\operatorname{rk} Q = \operatorname{rk}(Q|_{\alpha^{\perp}}) + 1$.

On the other hand, the *signature* of the manifold X is $\sigma(Q) = \operatorname{rank}(Q)$ since Q is definite, and signature is an invariant of oriented cobordism, so

$$\operatorname{rank}(Q) \leqslant n(Q).\sigma(\mathbf{CP}^2) = n(Q).$$

Thus by our observation we must have $\sigma(Q) = n(Q)$, and Q is the standard form.

I.3. Connections and self-duality

The space $\mathfrak{M}^{\sigma} = \mathfrak{M}^{\sigma}(X)$ will be defined by using the ideas and methods developed for the study of the "gauge theories" of mathematical physics. There are now several thorough expositions ([5], [13] for example) of these ideas, so we will only say enough here to fix notation and the basic facts which we shall use.

If G is a compact Lie group, and $\begin{smallmatrix} P \\ \downarrow \\ X \end{smallmatrix}$ a principal G bundle over the 4-manifold X, one may define the space $\mathcal{A}$ of all *connections* on P. Any two connections A, B differ by an element

$$A - B \in \Omega^1(\mathfrak{g})$$

of the vector space of 1-forms with values in the bundle of Lie algebras associated to P, so $\mathcal{A}$ is an affine space. For any vector bundle V associated to P a connection A induces a differential operator

$$d_A \colon \Omega^p(V) \to \Omega^{p+1}(V),$$

and if X has a Riemannian metric, we get adjoints $d_A^* \colon \Omega^{p+1}(V) \to \Omega^p(V)$, and covariant derivatives $\nabla_A \colon \Omega^p(V) \to \Omega^p(V) \otimes \Omega^1$. The curvature $F(A)$ of any connection is an element of $\Omega^2(\mathfrak{g})$, and

$$F(A+a) = F(A) + d_A a + \tfrac{1}{2}[a, a], \qquad a \in \Omega^1(\mathfrak{g}).$$

The group $\mathcal{G}$ of automorphisms of P acts as a symmetry group on all this structure, and connections are *equivalent* if they are in the same orbit of $\mathcal{G}$ on $\mathcal{Q}$. Considering $\mathcal{G}$ as $\Gamma(P \times_{\mathrm{Ad}} G)$ this action is given by

$$g(A) = A - (d_A g)g^{-1}.$$

We will be concerned with the cases $G = SU(2)$, $U(1)$ and only very briefly with the second. The topological classification of such bundles is by

$$G = SU(2);\ P \text{ classified by } c_2(P) \in H^4(X; \mathbf{Z}) \cong \mathbf{Z},$$
$$G = U(1);\ P \text{ classified by } c_1(P) \in H^2(X; \mathbf{Z}).$$

In the second case we summarise all we need in

Proposition 3. *If P is a $U(1)$ bundle over X, and A any connection on P, then the image of $c_1(P)$ in $H^2(X; \mathbf{R})$ is represented in de Rham cohomology by the closed 2-form $-\frac{1}{2\pi i}F(A)$ ($\mathfrak{g}$ is in this case a trivial bundle); all 2-forms within this cohomology class occur as the curvature of some connection, and on the simply connected manifold X any two connections with the same curvature are isomorphic. In particular, if X has a Riemannian metric, there is a unique connection on P, up to equivalence, with harmonic curvature*: $\Delta F = 0$.

The corresponding characteristic class representation for an $SU(2)$ bundle is:

$$c_2(P) = \frac{1}{8\pi^2}\int_X Tr(F(A) \wedge F(A)) \in \mathbf{Z},$$

which also has a version for manifolds with boundary. If Y is a compact oriented 3-manifold, and A a connection on an $SU(2)$ bundle over Y, there is an invariant (see [5] Appendix 3)

$$Tc_2(A) \in \mathbf{R}/\mathbf{Z}$$

with the property that any time $Y = \partial Z$ and A extends to a connection $\tilde{A}$ over the oriented 4-manifold Z:

$$Tc_2(A) = \frac{1}{8\pi^2}\int_Z Tr(F(\tilde{A}) \wedge F(\tilde{A})), \quad \mathrm{mod}\, \mathbf{Z}.$$

(The analogous invariant for $U(1)$ connections over circles is the holonomy around the circle.)

Self-duality. On an oriented Riemannian 4-manifold the 2-forms decompose:

$$\Omega^2 = \Omega^2_+ \oplus \Omega^2_-$$

into the spaces of "self-dual" and "anti self-dual" forms, defined by the ± 1 eigenspaces of the operator $*: \Omega^2 \to \Omega^2$, and thus by the property:

$$\alpha \in \Omega^2_\pm \text{ if } \alpha \wedge \alpha = \pm(\alpha \wedge *\alpha) = \pm|\alpha|^2.\text{vol}. \tag{1}$$

There is a similar decomposition of bundle valued forms and so of the curvature $F(A)$ of a connection on P:

$$F(A) = F_+(A) + F_-(A) \in \Omega^2_+(\mathfrak{g}) \oplus \Omega^2_-(\mathfrak{g}).$$

The tool which we shall use to prove Theorem 1 is the notion of a *self-dual connection* [2], that is to say, a connection with self-dual curvature, $F_-(A) = 0$. These were defined and studied for the following reason: for an $SU(2)$ connection A the "*Yang-Mills action*"

$$\|F(A)\|^2_{L^2} = \int_X |F(A)|^2\, d\mu = \int_X |F_+|^2 + |F_-|^2\, d\mu$$

is bounded below by the absolute value of

$$-8\pi^2 c_2(P) = \int_X -Tr(F \wedge F) = \int_X |F_+|^2 - |F_-|^2\, d\mu$$

by (1) above, and for $c_2 \leqslant 0$ there is equality if and only if A is self-dual. Conversely if the bundle P admits a self-dual connection, then $c_2 \leqslant 0$; and if $c_2 = 0$, so P is topologically trivial, any self-dual connection is *flat*, i.e., $F = 0$. In general a flat connection over a manifold is (up to isomorphism) equivalent to a representation of π_1 in the structure group, so for the simply connected manifold X of the theorem any flat connection is isomorphic to the standard product connection on $SU(2) \times X$; we call this connection $\boldsymbol{\theta}$.

This decomposition into self-dual and anti self-dual parts gives a differential operator acting on 1-forms:

$$d^-: \Omega^1 \to \Omega^2_-, \text{ the composite } \Omega^1 \xrightarrow{d} \Omega^2 \xrightarrow{\pi} \Omega^2_-.$$

Similarly if we have any connection A, there is

$$d_A^-: \Omega^1(V) \to \Omega^2_-(V),$$

with Laplacians Δ_A on $\Omega^0(V)$, $\Omega^1(V)$, $\Omega^2_-(V)$ given respectively by $d_A^* d_A$, $d_A d_A^* + d_A^* d_A^-$, $d_A^- d_A^*$; and associated harmonic spaces H_A^0, H_A^1, H_A^2; *if A is self-dual*, these are the harmonic spaces associated to the elliptic complex:

$$\Omega^0(V) \xrightarrow{d_A} \Omega^1(V) \xrightarrow{d_A^-} \Omega^2_-(V).$$

For the rest of this paper we suppose that X is given some Riemannian metric and that $\begin{array}{c} P \\ \downarrow \\ X \end{array}$ is an $SU(2)$ bundle with $c_2(P) = -1$.

I.4. Instantons and conformal invariance

All the self-dual connections on the bundle P $(c_2(P) = -1)$ over the standard Riemannian S^4 are explicitly known, and this example provides the basic model for the general theory.

Notice first that the self-duality condition depends only upon the conformal class of the Riemannian metric. Moreover the action

$$\int_U |F|^2 \, d\mu$$

contained in any open set U is likewise independent of conformal changes. This means on the one hand that the conformal group $SO(5,1)^+$ of S^4 acts on the set of self-dual connections on P; on the other hand that any self-dual connections on S^4 can be interpreted by the conformal equivalence $\mathbf{R}^4 \to S^4 \setminus \{pt\}$ as a self-dual connection or "*instanton*" on $\mathbf{R}^4$, and that the Chern class can be recovered from

$$\int_{\mathbf{R}^4} |F|^2 \, d\mu. = -8\pi^2 c_2. \tag{2}$$

A theorem of Uhlenbeck implies that the converse is also true. For later use we state a form of this result as

Proposition 4 [20, *Theorem* 4.1]. *If A is an $SU(2)$ connection (on the trivial bundle) over the punctured ball $B^4 \setminus \{0\}$, self-dual with respect to some smooth Riemannian metric on B^4 and with finite action*:

$$\int_{B^4 \setminus \{0\}} |F(A)|^2 \, d\mu < \infty,$$

then there is a bundle automorphism g: $B^4 \setminus \{0\} \to SU(2)$ such that $g(A)$ extends smoothly over B^4.

Hence we may recover a self-dual connection on P (up to isomorphism) from an instanton on $\mathbf{R}^4$ with total action $8\pi^2$.

According to the classification, established in [2] for example, the equivalence classes of self-dual connection on $\begin{smallmatrix} P \\ \downarrow \\ S^4 \end{smallmatrix}$ form a single orbit under the conformal group. There is a single $SO(5)$ invariant class (coming from the natural connection on the fibration $\begin{smallmatrix} S^7 \\ \downarrow \\ S^4 \end{smallmatrix}$), hence the set of these equivalence classes is parametrised by a *moduli space* $\mathfrak{M}(S^4) = SO(5,1)^+/SO(5) = B^5$, the open 5-ball. Interpreted as instantons on $\mathbf{R}^4$ these connections can be specified by a "*centre*" in $\mathbf{R}^4$, about which they have $SO(4)$ symmetry, and a *scale* which we can measure, for example, by the radius of the ball about the centre containing action $4\pi^2$ (i.e., one half of the total). Then the conformal group acts by translations on the centres and dilations $x \to \alpha . x$ on the radii. Using quaternionic notation: $\mathbf{R}^4 \cong \mathbf{H}$, $su(2) \cong Im\mathbf{H}$, the instanton $I_{0,\lambda}$ with centre 0 and radius λ can be given by the explicit connection matrix:

$$I_{0,\lambda}(x) = \mathrm{Im}\left(\frac{x d\bar{x}}{\lambda^2 + |x|^2} \right)$$

with curvature

$$F(I_{0,\lambda}) = \lambda^2 \frac{dx\, d\bar{x}}{\left(\lambda^2 + |x|\right)^2}.$$

(We shall not make any use of these equations.)

Thus one sees that the compactification $\bar{B}^5 = \mathfrak{M}(S^4) \cup S^4$ has an intrinsic interpretation in terms of the connections; a point x of S^4 represents the limit point of a sequence of connections whose curvature becomes concentrated in diminishing balls about x. The whole object of §§II and III of this paper is to show that an analogous moduli space $\mathfrak{M}(X)$ exists for the manifold X of the theorem, and that it may be compactified in the same way.

I.5. Taubes' theorem on the existence of self-dual connections

According to the Hodge Theory the second real cohomology group of any compact Riemannian manifold is represented by the space of harmonic 2-forms. The * operator commutes with the Laplacian Δ; so on a 4-manifold these harmonic forms decompose into self-dual and anti self-dual parts; $\mathcal{H}^2 = \mathcal{H}^2_+ \oplus \mathcal{H}^2_-$. By the defining property (1) the dimensions of these two spaces are just the numbers of positive and negative eigenvalues in a diagonalisation of the quadratic form:

$$Q_{\mathbf{R}}(\alpha) = \int_X \alpha \wedge \alpha, \quad \alpha \in H^2(X; \mathbf{R}).$$

(This is very close to some of the earliest applications of the theory of harmonic forms [9, §52.2]); so an equivalent form of the hypothesis in Theorem 1 that the intersection form is positive definite is the statement that the space $\mathcal{H}^2_-(X)$ of anti-self dual harmonic forms vanishes.

In the recent paper [18] C. H. Taubes constructs self-dual connections on the bundle P under precisely this hypothesis [18, Theorem 1.1]. They are constructed there by means of an "implicit function theorem" which we state here for later use:

Proposition 5 [18, *Theorem* 2.2]. *Let* A_0 *be a connection on* $\begin{matrix} P \\ \downarrow \\ X \end{matrix}$. *There are constants* $C, \varepsilon_0 > 0$, *depending only on the Riemannian structure of* X, *the*

L^3-norm $\|F_-(A_0)\|_{L^3}$ and the 1st eigenvalue $\mu(A_0)$ of the operator $\Delta_{A_0} = d^-_{A_0}d^*_{A_0}$ on $\Omega^2_-(\mathfrak{g})$ such that if

$$\delta(A_0) = \|F_-(A_0)\|_{L^2} + \|F(A_0)\|_{L^4}\|F_-(A_0)\|_{L^{4/3}} < \varepsilon_0,$$

then there is a self-dual connection $A_0 + d^*_{A_0}u = A_0 + a = q(A_0)$ with $u \in \Omega^2_-(\mathfrak{g})$ and

(i) $\|\nabla_{A_0}a\|_{L^2} \leqslant C.\delta(A_0)$,

(ii) $\|a\|_{L^2} \leqslant C.\|F_-(A_0)\|_{L^{4/3}}$.

The construction is $\mathcal{G}$-invariant, and $q(A_0)$ varies smoothly with A_0.

Given any point of X and $\lambda > 0$, Taubes applies this to a connection A_0 constructed to have most of its curvature in a ball of radius λ around that point. The hypothesis $\mathcal{H}^2_-(X) = 0$ means that the 1st eigenvalue of Δ on Ω^2_- is positive and ensures that $\mu(A_0)$ is bounded away from zero as $\lambda \to 0$.

A much easier observation, which we need to complete these preliminaries, is that for the 4-manifold X with positive definite form each $U(1)$ bundle has a unique self-dual connection up to isomorphism. This follows directly from Proposition 3 and the fact that the harmonic curvature form is automatically self-dual.

II.1. Sobolev spaces, reducible connections and properties of the space $\mathcal{B}$

We shall need to refer to Sobolev spaces $L^p_k(p \geqslant 1, k \in N)$ of sections of bundles associated to P; that is, of sections locally represented by functions with their first k derivatives in L^p. A smooth connection A gives definite norms $\|\ \|_{L^p_k(A)}$ on these spaces, for example:

$$\|s\|^2_{L^2_1(A)} = \|\nabla_A s\|^2_{L^2} + \|s\|^2_{L^2} = \int_X \left(|\nabla_A s|^2 + |s|^2\right) d\mu.$$

Moreover it is convenient in this §II to widen the definition of $\mathcal{C}$ to allow A differing from a smooth connection by an element of $L^2_3(\Omega^1(\mathfrak{g}))$. This differentiability is high enough not to make any essential differences in the properties discussed in §I (and also the norms $\|\ \|_{L^2_k(A)}$ are defined for $k \leqslant 3$) but allows us to work in Banach spaces.

We define the *moduli space* $\mathfrak{M}(X)$ to be the set of equivalence classes of self-dual connections on P, that is to say a subset of the quotient space:

$$\mathcal{B} = \mathcal{C}/\mathfrak{g},$$

and use the notation $p\colon \mathcal{C} \to \mathcal{B}$, $A \to [A]$. The techniques which we use here to study the moduli space $\mathfrak{M}$ are essentially those of [2], but we emphasise the role of the ambient space $\mathcal{B}$ rather more.

We discuss first the *reducible connections* on P. These may be understood from various different points of view. If we pick a base point in X, a connection gives, by parallel transport, a representation of the space of loops in X into $SU(2)$. The connection is reducible if this representation takes values in some proper subgroup of $SU(2)$. Since P is topologically nontrivial, the only possibility is that this subgroup be a copy of S^1. Alternatively a connection is reducible if there is a decomposition

$$E = L \oplus L^{-1}$$

of the $\mathbf{C}^2$ vector bundle E associated to P compatible, in the obvious sense, with the connection. Topologically the complex lie bundles are classified by the first Chern class in $H^2(X; \mathbf{Z})$, and since by the Cartan formula

$$c_2(L \oplus L^{-1}) = -c_1(L)^2$$

the number of such topological reductions is the number

$$n(Q) = \tfrac{1}{2}\left|\left\{\alpha \in H^2(X; \mathbf{Z}) \mid Q(\alpha) = 1\right\}\right|$$

of §I.2. (The factor $\frac{1}{2}$ appearing since we have a choice of L, L^{-1}.) Hence by Proposition 3 and the remark at the end of §I.5, there is a unique self-dual connection on P, up to equivalence, for each of these $n(Q)$ reductions.

If A is a reducible connection on P corresponding to a splitting $E = L \oplus L^{-1}$, then for any $e^{i\theta} \in S^1$ the element γ of the group of automorphisms $\mathcal{G}$ corresponding to a constant rotation of L by $e^{i\theta}$ and L^{-1} by $e^{-i\theta}$ is a covariant constant; $d_A\gamma = 0$. Hence γ fixes A in $\mathcal{Q}$, and moreover the action of γ on bundle-valued forms commutes with all the differential operators associated with A. In particular the stabiliser $\Gamma_A \cong S^1$ of A in $\mathcal{Q}$ acts on the harmonic sub-spaces H_A^1, H_A^2 of the $\mathfrak{g}$ valued forms. The bundle $\mathfrak{g}$ splits naturally into the direct sum of a trivial line bundle and the complex line bundle L^2. It will become important to note that since

$$H^1(X; \mathbf{R}) = \mathcal{H}_-^2(X) = 0,$$

the harmonic spaces H_A^1, H_A^2 are contained entirely within the complex part $\Omega^p(L^2)$, acted on by Γ_A in the obvious way.

Conversely, if a connection A has a nontrivial stabiliser Γ_A in $\mathcal{G}$, then it is reducible. Similarly a connection is reducible if and only if the harmonic space H_A^0 of covariant constant sections of $\mathfrak{g}$ is not zero, and this is then one dimensional.

Theorem 6. (i) *$\mathcal{B}$ is a Hausdorff space, and the open subset $\mathcal{B}^*$ of equivalence classes of irreducible connections forms a Banach manifold, with charts constructed from "slices":*

$$T_{A,\varepsilon} = \{A + a \mid d_A^* a = 0, \|a\|_{L_3^2} < \varepsilon\}. \ (\varepsilon \textit{ sufficiently small}).$$

Moreover p: $p^{-1}(\mathcal{B}^*) \to \mathcal{B}^*$ is a principle $\mathcal{G}/\pm 1$ bundle.

(ii) *If A is reducible, then Γ_A acts on $T_{A,\varepsilon}$ and the map $T_{A,\varepsilon/\Gamma_A} \to \mathcal{B}$ is a homeomorphism to a neighborhood of $[A]$ in $\mathcal{B}$, smooth off the fixed point set of Γ_A.*

Proof. Results of this type are now standard [2], [13], so we shall be brief.

First, the groups $\mathcal{G}$, $\mathcal{G}/\pm 1$ are Lie groups in the L_4^2 topology and have Lie algebra $L_4^2(\Omega^0(\mathfrak{g}))$; moreover they act smoothly on $\mathcal{A}$. Second, we may always reduce the problem to a local one in the following sense: if $A \in \mathcal{A}$, and ε is sufficiently small, then any $g \in \mathcal{G}$ for which there is an $A + a$ with $\|a\|_{L_3^2}$, $\|g(A + a) - A\|_{L_3^2}$ both less than ε may be factored as:

$$g = \gamma\tilde{g} \text{ with } \gamma \in \Gamma_A, \ \|\tilde{g} - 1\|_{L_4^2} \leqslant \text{const } \varepsilon.$$

To see this, consider $\mathcal{G}$ as a subset of $L_4^2(\Omega^0(\mathfrak{g} \otimes \mathbf{C}))$ and write $g = g_1 + g_2$, $g_1 \in H_A^0$, $g_2 \in H_A^{0\perp}$ ($\perp$ with respect to L^2 inner product). Then $d_A g_1 = 0$ and

$$\|d_A g_2\|_{L_k^2} \geqslant \text{const.}\|g_2\|_{L_{k+1}^2}, \qquad k = 0, 1, 2, 3,$$

so that

$$b = g(A + a) - A = {}^{-}d_A g g^{-1} + g a g^{-1},$$
$$d_A g_2 = d_A g = bg - ga.$$

Thus from $\|b\|_{L_3^2}, \|a\|_{L_3^2} < \varepsilon$ we deduce in turn that $\|g_2\|_{L_k^2} < \text{const.}\varepsilon$ for $k = 1, 2, 3, 4$, so that the constant component g_1 differs from an element $\gamma \in \Gamma_A$ by $0(\varepsilon)$. Then write

$$g = g_1 + g_2 = \gamma + (g_1 - \gamma + g_2)$$
$$= \gamma\big(1 + \gamma^{-1}\{g_1 - \gamma + g_2\}\big) = \gamma\tilde{g}.$$

By the same argument, if A, $B \in \mathcal{A}$, then for small enough ε any $g \in \mathcal{G}$ moving an element within ε of A to an element within ε of B can be factored into an automorphism sending A to B and an automorphism close to 1, so $\mathcal{B}$ is Hausdorff in the quotient topology.

Now the proof is standard calculus.

(i) The set $\mathcal{B}^* \subset \mathcal{B}$ representing irreducible connections is open since the condition $H_A^0 = 0$ is an open condition ("semi-continuity of cohomology"), and $\mathcal{G}/\pm 1$ acts freely on $p^{-1}(B^*)$. The smooth map

$$S\colon T_{A,\varepsilon} \times \mathcal{G}/\pm 1 \to \mathcal{A}, \quad S(A + a, g) = g(A + a)$$

has derivative at $a = 0$, $g = 1$:

$$DS\colon \operatorname{Ker} d_A^* \times \Omega^0(\mathfrak{g}) \to \Omega^1(\mathfrak{g}), \quad (a, v) \to a + d_A v,$$

which is an isomorphism since $\operatorname{Ker} d_A = H_A^0 = 0$ and $\Omega^1(\mathfrak{g}) = \operatorname{Im} d_A \oplus \operatorname{Ker} d_A^*$. Hence S is locally a diffeomorphism and so, appealing to the discussion above, a diffeomorphism to its image. The cited properties of $p^{-1}(\mathcal{B}^*) \to \mathcal{B}^*$ follow easily from this.

(ii) This time DS is surjective but has Kernel H_A^0. Thus the restriction

$$S\colon T_{A,\varepsilon} \times \exp\left(H_A^{0\perp}\right) \to \mathcal{A}$$

($H_A^{0\perp}$ the complement in the L^2 inner product) is a local diffeomorphism. Similarly, since multiplication $\Gamma_A \times \exp(H_A^{0\perp}) \to \mathcal{G}$ has derivative 1 at the identity there is a unique splitting of $g \in \mathcal{G}$ close to 1 as $g = \gamma\tilde{g}$, $\gamma \in \Gamma_A$, $\tilde{g} \in \exp(H_A^{0\perp})$. So in this case a neighborhood of $[A]$ in $\mathcal{B}$ is homeomorphic to $T_{A,\varepsilon/\Gamma_A}$, and $p\colon T_{A,\varepsilon} \to \mathcal{B}^*$ is clearly smooth off the fixed point set of Γ_A.

We see then that in the local model of a neighborhood of a reducible connection in $\mathcal{B}$:

$$T_{A,\varepsilon}/\Gamma_A \subset \Omega^1(L^2)/S^1 \times \Omega^1_X,$$

the component $\operatorname{Ker} d^* \subset \Omega^1_X$ corresponds to the deformations of A within the space of equivalence classes of reducible connections, and that transverse to this set $\mathcal{B}$ has a quotient singularity of the form Complex Hilbert space$/S^1$.

II.2. Local properties of the moduli space

If $A \in \mathcal{A}$ is self-dual the self-dual connections within the slice $T_{A,\varepsilon}$ are given by the set $Z(\Phi)$ of zeros of the map

$$T_{A,\varepsilon} = \left\{A + a \,|\, d_A^* a = 0, \|a\|_{L_3^2} < \varepsilon\right\} \xrightarrow{\Phi} L_2^2\left(\Omega_-^2(\mathfrak{g})\right),$$

$$\Phi(A + a) = F_-(A + a) = d_A^- a + \tfrac{1}{2}[a, a].$$

The essential fact which we use is that the derivative of Φ:

$$(D\Phi)_A\colon \left(\operatorname{Ker} d_A^* \subset L_3^2(\Omega^1(\mathfrak{g})\right) \xrightarrow{d_A^-} L_2^2\left(\Omega_-^2(\mathfrak{g})\right)$$

is a *Fredholm* operator (recall that a bounded linear map is Fredholm if it has finite dimensional kernel and cokernel and a closed range). This is an immediate consequence of the ellipticity of the differential operator: $d_A^* + d_A^-$ since it is well known that elliptic operators over compact manifolds are

Fredholm. Thus first, it has an index

$$\begin{aligned}\operatorname{Index}(D\Phi)_A &= \dim \operatorname{Ker}(D\Phi)_A - \dim \operatorname{coker}(D\Phi)_A \\ &= \operatorname{index}\big(d_A^* + d_A^-\colon \Omega^1(\mathfrak{g}) \to \Omega^0(\mathfrak{g}) \oplus \Omega_-^2(\mathfrak{g})\big) + \dim H_A^0.\end{aligned}$$

The index of $d_A^* + d_A^-$ is independent of $A \in \mathcal{A}$ and can be computed by the Atiyah-Singer index theorem in terms of the original data X, P to be [2]:

$$\operatorname{index}(d_A^* + d_A^-) = 8|c_2(P)| - \tfrac{3}{2}(\chi(X) - \tau(X)) = 5,$$

since by our assumptions on X, $\chi = 2 + b_2$, $\tau = b_2$. Thus the index of $(D\Phi)_A$ is 5 or 6 as A is irreducible or not, corresponding as we shall see to the fact that if A is reducible, we have to divide $Z(\Phi)$ by Γ_A to get a true neighborhood of $[A]$ in $\mathcal{B}$.

Second, it is a general property of nonlinear maps of this type that we may reduce them locally to a linear part and a finite dimensional nonlinear part. This was done by Kuranishi in [10] and by Smale in [16]; we shall use the following simple lemma, in which all maps are to be intepreted as germs about the origin.

Lemma 7. *Let $E \xrightarrow{f} F$, $f(0) = 0$, be a smooth map between Banach spaces with the property that $L = (Df)_0$ is a Fredholm operator of index $k > 0$. If we choose complements*

$$E = \ker L \oplus E', \quad F = \operatorname{Im} L \oplus F',$$

so that $\dim \operatorname{Ker} L = \dim F' + k$, *then there are a diffeomorphism ψ of E and a smooth map $\phi\colon E \to F'$ such that*

$$f \circ \psi^{-1}(x) = Lx + \phi(x) \in \operatorname{Im} L \oplus F' = F.$$

Proof. By the open mapping theorem, $L|_{E'}\colon E' \to \operatorname{Im} L$ has a bounded inverse $\tilde{L}$ which we can extend to F in the obvious way, so that $L\tilde{L} = 1 - \pi$; $\pi\colon F \to F'$ the projection.

Then $\psi\colon E \to E$, $\psi(x') = x' + \tilde{L}(f(x') - Lx')$, has derivative 1_E at the origin and is thus a local diffeomorphism with

$$\begin{aligned}L\psi(x') &= Lx' + L\tilde{L}(f(x') - Lx') \\ &= Lx' + (1 - \pi)(f(x') - L(x')) \\ &= f(x') - \pi(f(x') - Lx').\end{aligned}$$

So if $x' = \psi^{-1}(x)$, then

$$f(\psi^{-1}(x)) = Lx + \pi\big(f(\psi^{-1}(x)) - L\psi^{-1}(x)\big) = Lx + \phi(x),$$

as required.

Note also that, with the obvious definition (c.f. [16]), a point $y = y_1 + y_2 \in \operatorname{Im} L \oplus F' = F$ is a *regular value* for f if and only if y_2 is a regular value for the finite dimensional map $\phi|_{L^{-1}(y_1)}$, and for any such regular value the preimage $f^{-1}(y)$ is a smooth submanifold of E of dimension K.

Applying this first to the self-duality equations in $T_{A,\varepsilon}$, and choosing the harmonic space H_A^2 as the complement of $\operatorname{Im} d_A^-$, we have

Proposition 8. *For any self-dual A, ε sufficiently small there is a map ϕ from a neighborhood of the origin in the harmonic space H_A^1 to the harmonic space H_A^2 such that if A is irreducible, a neighborhood of $[A]$ in $\mathfrak{M}$ is carried by a diffeomorphism onto*

$$Z(\phi) = \phi^{-1}(0) \subset H_A^1 \left(\textit{and} \dim H_A^1 - \dim H_A^2 = 5\right)$$

and if A is reducible, so H_A^1, H_A^2 are complex spaces, ϕ is Γ_A equivariant and a neighborhood of $[A]$ is modelled on

$$Z(\phi)/\Gamma_A \cong Z(\phi)/S^1 \quad \left(\textit{and} \dim_{\mathbf{R}} H_A^1 - \dim_{\mathbf{R}} H_A^2 = 6\right).$$

This position follows straight from the lemma: the second part because the symmetry group Γ_A acts on the whole procedure.

Thus just as H_A^0 detects the reducibility of a connection, the vanishing of the space H_A^2 signals whether or not 0 is a regular value for Φ, and so whether or not the moduli space has generic behaviour about A: either a smooth 5-manifold or the quotient of $H_A^1 \cong \mathbf{C}^3$ by $\Gamma_A \cong S^1$. The 1st eigenvalue μ of $\Delta_{A_0} = d_{A_0}^- d_{A_0}^*$ on $\Omega_-^2(\mathfrak{g})$ in Taubes implicit function theorem (Proposition 5) appears for essentially the same reason. Anticipating here Proposition 18 we can say that the subset $K \subset \mathfrak{M}$ of classes where $H_A^2 \neq 0$ is *compact*.

II.3. Perturbation of the moduli space

For the proof of the main theorem used in §I.2 we have to establish the existence of a smooth 5-dimensional space with only finitely many explicitly known singularities. If the harmonic space H_A^2 vanishes for all self-dual A, then the moduli space $\mathfrak{M}$ would have this property, but in general we know very little about these spaces, so we argue that one may find a perturbation σ of the self duality equations such that the corresponding zero set $\mathfrak{M}^\sigma$ has generic form.

This is the formal setting for the perturbation: the group $\mathcal{G}/\pm 1$ acts on the Banach spaces

$$L_3^2(\Omega_-^2(\mathfrak{g})) \hookrightarrow L_2^2(\Omega_-^2(\mathfrak{g})).$$

So we get a pair of bundles, $\mathcal{E}^3 \hookrightarrow \mathcal{E}^2$ say, over the manifold $\mathfrak{B}^*$, associated to the principal bundle $p^{-1}(\mathfrak{B}^*)$; and the anti-self dual component of the curvature induces a canonical section Φ of $\mathcal{E}^2$. The allowable perturbations σ are to be smooth sections of $\mathcal{E}^3$: the "uniform norm" $\|\sigma(A)\|_{L^2_3(A)}$ in the fibre over $[A] \in \mathfrak{B}^*$ is well defined as is the "covariant derivative" $(\nabla\sigma)_A$, a linear map

$$(\nabla\sigma)_{[A]}\colon \operatorname{Ker} d_A^* \cong (T\mathfrak{B}^*)_{[A]} \to \mathcal{E}_A^3 \cong L^2_3(\Omega^2_-(\mathfrak{g})),$$

(in effect the slices $T_{A,\varepsilon}$ define a connection on $\mathcal{E}^k$). If we interpret σ as a section of $\mathcal{E}^2$, $\nabla(\Phi + \sigma)$ is Fredholm and of the same index as $\nabla\Phi$, since $\nabla\sigma$ factors through L^2_3, and $L^2_3 \hookrightarrow L^2_2$ is compact.

All of this makes sense over the whole of $\mathfrak{B}$ provided that σ is defined in a $T_{A,\varepsilon}$ with the appropriate Γ_A symmetry. We give such σ the topology of uniform convergence of $\|\sigma\|_{L^2_3(A)}$ and convergence of $\nabla\sigma$ in operator norm, uniformly on compact sets.

It is convenient to treat the reducible connections first. Suppose that A is a reducible self-dual connection with $H_A^2 \cong \mathbf{C}^p$, $p > 0$ so that we have an S^1 equivariant map

$$H_A^1 \cong \mathbf{C}^{3+p} \xrightarrow{\phi} H_A^2 \cong \mathbf{C}^p$$

with $(D\phi)_0 = 0$ by construction. Pick a complex linear epimorphism $\alpha\colon H_A^1 \to H_A^2$, and for sufficiently small $\eta > 0$ define $\tilde{\phi}$, also equivariant, by

$$\tilde{\phi}(z) = \phi(z) + \eta.\beta(|z|/\eta).\alpha(z),$$

where $\alpha(z) \in H_A^2$, and β is a cut off function. Then in a small enough neighborhood of 0, $\tilde{\phi}^{-1}(0)$ is a smooth 6-dimensional manifold, and $\tilde{\phi}^{-1}(0)/S^1 \cong \mathbf{C}^3/S^1$ is a cone on $\mathbf{CP}^2$. In the notation of Lemma 7 we replace

$$\Phi\colon T_{A,\varepsilon} \to L^2_2(\Omega^2_-(\mathfrak{g})), \quad \Phi(\psi^{-1}X) = Lx + \phi(x)$$

by $\tilde{\Phi}$

$$\tilde{\Phi}(\psi^{-1}X) = Lx + \tilde{\phi}(x),$$

equal to Φ outside a small neighborhood of $[A]$.

Doing this for each of the $n(Q)$ reducible self-dual connections we can pass without loss of generality to the case when the compact set K is contained in $\mathfrak{B}^*$ since all our subsequent perturbations will be supported in $\mathfrak{B}^*$. It is convenient then to prove

Lemma 9. *If $V \subset\subset U \subset \mathfrak{B}^*$ are open sets with U covered by finitely many slices $T_{A,\varepsilon}$, then the set G of "good perturbations" σ, supported in U and such that the zero set $\overline{V} \cap Z(\Phi + \sigma)$ is cut out transversally (i.e., $(\nabla(\Phi + \sigma))$ is onto for all $[A] \in \overline{V} \cap Z(\Phi + \sigma)$) is open and dense.*

Given Lemma 9 we can find three open sets $K \subset U_1 \subset\subset U_2 \subset\subset U_3$, each covered by finitely many slices $T_{A,\varepsilon}$, since K is compact. The set of good perturbations over U_1 supported in U_2 is open and dense, and by applying

Lemma 9 to $(U_2\backslash \bar{U}_1) \subset U_3\backslash \bar{U}_1$ and the fact that Φ is already good outside U_1, we can find the required small perturbation σ such that $\mathcal{M}^\sigma = Z(\Phi + \sigma)$ is a smooth 5-manifold with $n(Q)$ singular points, equal to $\mathcal{M}$ outside a compact set.

Proof of Lemma 9. If σ is any perturbation bounded in $\| \ \|_{L^2_3}$ and $0 \leqslant R < \infty$, then the set of $[A]$ in $\bar{V}$ such that $\|(\Phi + \sigma)(A)\|_{L^2_3(A)} \leqslant R$ is compact (for in each of the slices $T_{A,\varepsilon}$ covering[1] U an L^2_3 bound on $\Phi + \sigma$ gives an L^2_3 bound on $(d_A^- + d_A^*)a$ and so an L^2_4 bound on a by ellipticity, and $L^2_4 \hookrightarrow L^2_3$ is compact).

That G is open, in the topology we defined, follows almost immediately from this and the fact that the set of surjective operators on a Banach space is open in operator norm.

Similarly, using this compactness, to prove any given σ is in $\bar{G}$ we may reduce to the case when U is contained in some $T_{A,\varepsilon}$ and the local representation

$$f = \Phi + \sigma \colon E = \operatorname{Ker} d_A^* \cap L^2_3 \to L^2_2 = F$$

can be decomposed in the manner of Lemma 7 over U (since a finite intersection of open dense sets is open and dense). We may then suppose that the finite dimensional complement F' lies in L^2_3, and so choose a regular value of f in F' arbitrarily small in L^2_3 norm by applying the usual Sard lemma to ϕ: $\operatorname{Ker} L \to F'$. Then we extend this to get an element of G by using a bump function $\beta(\|a\|)$.

II.4. Orientability of $\mathcal{M}^\sigma$

To prove that the 5-manifold $\mathcal{M}^\sigma \cap \mathcal{B}^*$ is orientable, for any good perturbation σ, we again use a general fact about Fredholm operators [1], [3]. Suppose V, W are bundles over some (compact) manifold M and that T is a compact *parameter space* with bundles $\tilde{V}$, $\tilde{W}$ over $M \times T$ such that $\tilde{V}_t = \tilde{V}|_{M\times\{t\}} \cong V$, $\tilde{W}_t \cong W$. If we are given a family $\{L_t\}$ of Fredholm operators:

$$L_t \colon \Gamma(\tilde{V}_t) \to \Gamma(\tilde{W}_t)$$

varying continuously with $t \in T$, there is an element

$$\operatorname{ind}\{L_t\} \in KO(T),$$

[1] We ought here to suppose that the slices $T_{A,\varepsilon}$ are centred on smooth connections A—this is possible since the smooth connections are dense in $\mathcal{Q}$.

which has the property that if $\dim \operatorname{Ker} L_t$ is constant for t in some closed subset $S \subset T$ — so that $\operatorname{Ker}\{L_t\}$, $\operatorname{coKer}\{L_t\}$ form vector bundles over S —

$$(*) \qquad \operatorname{ind}\{L_t\}\,|_S = \operatorname{Ker}\{L_t\} - \operatorname{CoKer}\{L_t\} \in KO(S).$$

(The numerical index used previously is then the image of this KO index under the augmentation: $KO(T) \to \mathbf{Z}$.) Since the reducible connections are a nuisance here, we choose a base point $x_0 \in X$ and form the restricted gauge group $\mathcal{G}_0 \subset \mathcal{G}$ of automorphisms equal to the identity on P_{x_0}. Then $\mathcal{G}_0$ acts freely on $\mathcal{C}$, so we get bundles

$$\tilde{V} = \widetilde{(T^*X) \otimes \mathfrak{g}}, \quad \tilde{W} = \widetilde{\mathfrak{g} \oplus \Lambda^2(T^*X) \otimes \mathfrak{g}},$$

over the product $\mathcal{C}/\mathcal{G}_0 \times X = \hat{\mathcal{B}} \times X$, and a family of operators:

$$L_{[A]} = d_A^* + d_A^- + (\nabla\sigma)_A\colon \Gamma\big(\tilde{V}_{[A]}\big) \to \Gamma\big(\tilde{W}_{[A]}\big).$$

Note first that the natural projection

$$\hat{\mathcal{B}} \xrightarrow{\pi} \mathcal{B}$$

becomes a fibration when restricted to the dense open subset $\mathcal{B}^* \subset \mathcal{B}$ of irreducible connections, with fibre $SO(3)$; hence the tangent bundle to $\mu^\sigma \cap \mathcal{B}^*$ will be orientable if and only if its lift to $\pi^{-1}(\mu^\sigma \cap \mathcal{B}^*)$ is.

Second we recall that the characteristic class w_1 is defined on KO, and since by the property $(*)$ above:

$$\pi^*(T\mu^\sigma) = \operatorname{Ind}\{L_{[A]}\} \quad \text{in } KO\big(\pi^{-1}(\mu^\sigma \cap \mathcal{B}^*)\big),$$

it suffices to prove that $w_1(\operatorname{Ind} L_{[A]}) = 0 \in H^1(\hat{\mathcal{B}}; \mathbf{Z}/2)$.

Thirdly observe that for all $t \in [0, 1]$ the operator

$$d_A^* + d_A^- + t(\nabla\sigma)_A\colon L_3^2 \to L_2^2$$

is Fredholm, so that by forming the corresponding family of operators over a cylinder, we have

$$\operatorname{Ind}\{L_{[A]}\} = \operatorname{Ind}\{d_A^* + d_A^-\} \in KO(\hat{\mathcal{B}}).$$

Now the next lemma, together with these preliminary remarks, proves that $\mu^\sigma \cap \mathcal{B}^*$ is orientable.

Lemma 10. *Let M be a simply-connected 4-manifold, P_k an $SU(n)$ bundle over M with Chern class $c_2 = k$. If we form the space:*

$$\hat{\mathcal{B}}(M, n, k)$$

of equivalence classes of connections on P_k, and the corresponding family of operators $d^* + d^-$ on Lie algebra valued forms, then we have

$$w_1(\operatorname{Ind}(d^* + d^-)) = 0 \in H^1\big(\hat{\mathcal{B}}(M, n, k); \mathbf{Z}/2\big).$$

Proof. The lemma is stated in more generality than we need since the proof will use the stable range $n \geqslant 3$.

The adjoint representation of $SU(3)$ on $\mathfrak{su}(3)$, when restricted to $SU(2)$, decomposes into the direct sum

$$\mathfrak{su}(2) \oplus \mathbf{R} \oplus V,$$

where V is a complex vector space. This means that if we regard $\hat{\mathfrak{B}}(M, 2, k)$ as being included as a subset of reducible connections in $\hat{\mathfrak{B}}(M, 3, k)$, in the obvious way, the stable real vector bundle

$$\operatorname{ind}\{d^* + d^-\}(M, 3, k),$$

when restricted to $\hat{\mathfrak{B}}(M, 2, k)$, differs from

$$\operatorname{ind}\{d^* + d^-\}(M, 2, k)$$

by the sum of $(\pm)$ a trivial real bundle and a complex stable bundle. Hence

$$w_1(\operatorname{ind}\{d^* + d^-\}(M, 2, k))$$

is the restriction of $w_1(\operatorname{ind}\{d^* + d^-\}(M, 3, k)) \in H^1(\hat{\mathfrak{B}}(M, 3, k), \mathbf{Z}/2)$ to $H^1(\hat{\mathfrak{B}}(M, 2, k); \mathbf{Z}/2)$.

Then we claim that $\hat{\mathfrak{B}}(M, n, k)$ is simply connected for $n \geqslant 3$; this will obviously complete the proof of the lemma. $\mathcal{G}_0 = \mathcal{G}_0(M, n, k)$ acts freely on the contractible affine space of connections so $\pi_1(\hat{\mathfrak{B}}(M, n, k))$ is isomorphic to the set of connected components of $\mathcal{G}_0(M, n, k)$.

The 4-manifold M is simply connected so it is, homotopically, a wedge of 2-spheres with a 4-cell attached [12]. The bundle P_k is trivial over the 2-skeleton and, since $\pi_2(SU(n)) = 0$, any element of $\mathcal{G}_0$ can be deformed to be the identity on a neighborhood of the 2-skeleton. Thus collapsing the 2-skeleton, we may reduce to the case when $M = S^4$. Similarly P_k is trivial on the complement of a point, so we may reduce to the case when $k = 0$ and the bundle is trivial. Then the components of $\mathcal{G}_0(S^4, n, 0)$ are the homotopy groups

$$\pi_4(SU(n)),$$

which vanish for $n \geqslant 3$.

Notes. (i) $\operatorname{ind}\{d^* + d^-\}$ is not strictly defined on the noncompact space $\hat{\mathfrak{B}}$. However, this plainly does not affect the argument, which only involves compact subsets.

(ii) For the case at hand, when $k = 1$, $n = 2$, one may alternatively argue that the full quotient $\mathfrak{B} = \hat{\mathfrak{B}}/SO(3)$ is simply connected.

(iii) The point to this proof is that we may regard the group $SU(2)$ as being the first member of either the family of symplectic groups $Sp(n)$ or of the special unitary groups $SU(n)$. For the manifold S^4, or more generally for any spin manifold, there is an essential loop in the spaces corresponding to the symplectic groups, and it is detected by the index of the Dirac operator [3], whereas this loop does not persist in the other family, as we have seen above.

III

In this section we prove

Theorem 11. *There is an open subset $\mathfrak{M}_{\lambda_0}$ of the moduli space $\mathfrak{M} = \mathfrak{M}(X)$ of self-dual connections on P which is a smooth 5-manifold diffeomorphic to $X \times (0, \lambda_0)$, $\lambda_0 > 0$, and with the property that the complement $\mathfrak{M} \setminus \mathfrak{M}_{\lambda_0}$ is compact.*

Once this has been done the proof of Theorem 1 will be complete—since the perturbed space $\mathfrak{M}^\sigma$ of §II is identical to $\mathfrak{M}$ outside a compact set and so may be compactified to have boundary X by using the collar of Theorem 11.

The proof which we give may be regarded as an essay on the fundamental results of C. H. Taubes and K. K. Uhlenbeck [18], [20], [21], and is divided into four parts:

III.1. Convergence of a sequence of connections over the 4-ball.

III.2. Convergence of a sequence of connections over X.

III.3. The definition of $\mathfrak{M}_{\lambda_0}$ and of a covering map $\mathfrak{M}_{\lambda_0} \xrightarrow{p} X \times (0, \lambda_0)$.

III.4. Proof that p is 1-sheeted, hence a diffeomorphism.

III.1.

Proposition 12. *There is a constant $C > 0$ such that if $\{m_i\}$ is a sequence of metrics on the 4-ball B^4, each sufficiently close to the Euclidean metric and converging in $C^\infty(\bar{B}^4)$ to a limiting m_∞; and if $\{\tilde{A}_i\}$ is a sequence of connections (on the trivial bundle) over B^4 with $\tilde{A}_i$ self-dual with respect to metric m_i and satisfying*

$$\int_{B^4} \left|F(\tilde{A}_i)\right|^2 d\mu \leqslant C$$

Then there is a subsequence $\{i'\} \subset \{i\}$, and A_i equivalent to $\tilde{A}_i$ such that $A_{i'}$ converge in C^∞ on the half sized ball $\frac{1}{2}B^4 = \{X \in \mathbf{R}^4 \,||X|< \frac{1}{2}\}$ to a limiting connection A_∞; self-dual with respect to the metric m_∞.

This proposition follows from Theorem 1.3 of [21]. Identifying in the usual way a connection on the trivial bundle with a matrix of 1-forms, we may suppose by that result that the A_i are chosen to satisfy:

(i) $d^*A_i = 0$,

(ii) $\|A_i\|_{L_1^p} \leqslant \text{const.}\|F(\tilde{A}_i)\|_{L^p}$, for any $p \geqslant 2$.

Just as in §II the constraint (i), together with the self-duality condition,

$$d^{-}A_i + \tfrac{1}{2}[A_i, A_i]_{-} = 0,$$

gives an elliptic system of equations, and for C sufficiently small the standard convergence argument applies (by the a priori inequality, Theorem 3.5 of [20] we can take $\|F(\tilde{A}_i)\|_{L^p}$, and so $\|A_i\|_{L^p_1}$, bounded for any given p).

III.2

The local result of §III.1 above allows us to prove the following theorem for a sequence of connections defined over X.

Theorem 13. *Let $\tilde{A}_i \in \mathcal{C}$ be a sequence of self-dual connections on the bundle P over X. Then there is a subsequence $\{i'\}$ such that either one of the following holds.*

(i) *Each $\tilde{A}_{i'}$ is equivalent to an $A_{i'} \in \mathcal{C}$, converging in C^∞ to a self-dual connection A_∞ on P; hence*

$$[\tilde{A}_i] \to [A_\infty] \quad \textit{in } \mathfrak{M} \subset \mathfrak{B}.$$

(ii) *There are a point $x \in X$ and on the complement K of any geodesic ball about x bundle isomorphisms*

$$\rho_{i'}\colon K \times SU(2) \to P|_K$$

such that $\rho_{i'}^*(\tilde{A}_{i'}) \to \theta$ (the product connection) in $C^\infty(K)$.

There is considerable overlap in this theorem with the results of [15], one may compare also [14]. We prove a combinatorial lemma first.

Lemma 14. *Given $L, C > 0$ and a sequence of functions $f_i \geqslant 0$ on X with*

$$\int_X f_i \, d\mu \leqslant L$$

one may find a subsequence $\{i'\} \subset \{i\}$, a finite set $\{x_1, x_2, \cdots, x_l\} \subset X$ and a countable collection $\{B_\alpha\}$ of small geodesic balls in X such that the half sized balls cover $X \setminus \{x_1, \cdots, x_l\}$, and for each α

$$I(\alpha, i') = \int_{B_\alpha} f_{i'} \, d\mu$$

is eventually less than C.

Proof of Lemma 14. X has a countable base of neighborhoods $\{B_q\}_{q \in N}$ made up of small geodesic balls, and for each q we have

$$I(q, i) = \int_{B_q} f_i \, d\mu \leqslant L.$$

Thus we may find, for each fixed q, a subsequence i' such that $I(q, i')$ converges, and by a "diagonal argument" we can arrange that this happens for all q simultaneously: $I(q, i') \to I(q)$ say. That is just the standard fact that some subsequence of the f_i converge in the sense of distributions.

Suppose that $x_1, \cdots, x_l$ are points of X each of which lies in no ball B_q with $I(q) \leqslant c/2$ (i.e., the limiting distribution "contains δ-functions" at the points x_j). Then we may choose disjoint balls B_j containing x_j and satisfying

$$\int_{B_j} f_{i'}\, d\mu > c/2 \quad \text{for all large enough } i';$$

so that

$$L \geqslant \int_X f_{i'}\, d\mu \geqslant \sum_{j=1}^{l} \int_{B_j} f_{i'}\, d\mu \geqslant \tfrac{1}{2} lc.$$

Thus l is at most $2L/c$, and in particular there are only finitely many points of this type. Then we can select the required cover of $X \setminus \{x_1, x_2, \cdots, x_l\}$ from among the balls B_q.

Proof of Theorem 13. Suppose that $B \subset X$ is any geodesic ball of radius r. We can find a geodesic coordinate system x^i in B such that

$$ds^2 = \left(\delta_{ij} + 0|x|^2\right) dx^i\, dx^j.$$

If we expand this by a factor r^{-1} to give a map:

$$\chi_B \colon B^4 \to B, y^i \to ry^i = x^i,$$

the induced metric on B^4 is $r^2(\delta_{ij} + r^2 0|y|^2) dy^i\, dy^j$, which is conformally equivalent to a metric differing, with all its derivatives, by $0(r^2)$ from the Euclidean metric.

Thus we may apply Lemma 14 to the sequence of action densities:

$$f_i = \left|F\left(\tilde{A}_i\right)\right|^2, \quad \int_X f_i\, d\mu = 8\pi^2 \quad \text{by §I.3},$$

with the constant C of Proposition 12, and choose the balls B_α so small that the hypotheses of Proposition 12 apply to the metrics on B^4 induced by the "conformal charts"

$$\chi_\alpha \colon B^4 \to B_\alpha.$$

Because of the conformal invariance of the Yang-Mills action we may apply Proposition 12 to deduce that for each α some subsequence of the $\chi_\alpha^*(A_{i'})$ converge after suitable bundle automorphism on $\frac{1}{2}B^4$ (all ensuing subsequences will be suppressed in the notation), and by a diagonal argument we may suppose this is true for all α simultaneously.

Thus in terms of local representations for the connections $\tilde{A}_{i'}$ over the cover $\frac{1}{2}B$ we have connection matrices

$$A_{i'}(\alpha) \to A_{\infty}(\alpha) \quad \text{in } C^{\infty}\left(\tfrac{1}{2}B_{\alpha}\right)$$

with transition functions $g_{i'}(\alpha, \beta)$: $\frac{1}{2}B_{\alpha} \cap \frac{1}{2}B_{\beta} \to SU(2)$ satisfying the compatability condition

$$(1) \qquad A_{i'}(\alpha) = -\, dg_{i'}(\alpha, \beta) g_{i'}(\alpha, \beta)^{-1} + g_{i'}(\alpha, \beta) A_{i'}(\beta) g_{i'}(\alpha, \beta)^{-1}.$$

The $g_{i'}(\alpha, \beta)$ are bounded since $SU(2)$ is compact, and the $A_{i'}(\alpha)$ converge in C^{∞}; so $dg_{i'}(\alpha, \beta)$ is uniformly bounded, and we may thus find a subsequence converging uniformly. Then (1) implies that this convergence is in C^{∞}. Once again we may arrange the subsequence so that this happens for all (α, β) simultaneously.

Thus the data $(A_{\infty}(\alpha), g_{\infty}(\alpha, \beta))$ represents a self-dual connection on a bundle Q over $X \backslash \{x_1, \cdots, x_l\}$; and if B_j is some small ball centred on $x_j (j = 1, \cdots, l)$, then

$$(2) \qquad \int_{B_j \backslash \{x_j\}} |F(A_{\infty})|^2 \, d\mu \leqslant \lim \int_{B_j \backslash \{x_j\}} |F(A_{i'})|^2 \, d\mu \leqslant 8\pi^2.$$

The bundle $Q|_{B_j \backslash \{x_j\}}$ is topologically trivial, so by Uhlenbeck's Removability of Singularities Theorem (Proposition 4) the connection A_{∞} and the bundle Q extend over all of X. Thus there is strict inequality in (3) above (by the definition of x_j in Lemma 14). On the other hand, all connections are self-dual, so

$$(3) \qquad \frac{1}{8\pi^2} \int_{B_j} |F(\tilde{A}_i)|^2 \, d\mu = \frac{1}{8\pi^2} \int_{B_j} Tr(F \wedge F) = -Tc_2\left(\tilde{A}_i|_{\partial B_j}\right), \quad \operatorname{mod} \mathbf{Z}.$$

Finally given any compact $K \subset X \backslash \{\kappa_1, \cdots, \kappa_l\}$ we may construct bundle isomorphisms

$$\rho_{i'}: Q|_K \to P|_K$$

such that $\rho_{i'}^{*}(\tilde{A}_{i'}) \to A_{\infty}$ in $C^{\infty}(K)$, by induction on the number of balls $\frac{1}{2}B_{\alpha}$ covering K just as in [21, §3]. In particular,

$$Tc_2\left(A_{\infty}|_{\partial B_j}\right) = \lim_{i'} Tc_2\left(\tilde{A}_{i'}|_{\partial B_j}\right),$$

so from (2), (3) $l = 0$ or 1 and we have *either* one of the following:

(i) When $l = 0$, $Q \cong P$; so we find the convergent sequence $A_{i'}$ in $\mathcal{C}$.

(ii) When $l = 1$, Q is trivial; so $A_{\infty} \cong \theta$ by the discussion of §I.3, and we find the required bundle maps over each K.

III.3

Theorem 13 shows that one needs to consider connections with curvature concentrated in small balls in order to compactify the moduli space $\mathfrak{M}$ (just as for the case $X = S^4$ in §I.4). Taubes constructs self-dual connections of this type depending upon a point in X and a scale $\lambda_0 > \lambda > 0$; thus he constructs a map:

$$X \times (0, \lambda_0) \to \mathfrak{M},$$

and one can prove directly [19] that this map is a diffeomorphism to its image. However, we find it easier here to define a map p from an open subset of $\mathfrak{M}$ representing these concentrated solutions to $X \times (0, \lambda_0)$—thus approximately inverse to Taubes construction. The definition of any such map is bound to be rather arbitrary; the one we choose is slightly complicated but suits our later needs.

We could define the "radius" and position of a self-dual connection by taking the smallest ball which contains one half of the total action, just as for the instantons. However, we will "smooth" this definition in the following way.

Let β be a smooth even bump function approximating, and dominated by $\chi_{[-1,1]}$, and for any pair x, y of points in a Riemannian manifold and $s > 0$ set

$$\beta_s(x, y) = \beta\left(\frac{d(x, y)}{s}\right).$$

Then for the basic instanton action density

$$f = |F(I_{0,1})|^2 \quad \text{on } \mathbf{R}^4,$$

the function R_I: $\mathbf{R}^4 \times \mathbf{R} \to \mathbf{R}$ given by

$$R_I(x, s) = \int_{\mathbf{R}^4} \beta_s(x, y) f(y)\, d\mu_y$$

has, as one can very easily see, the following properties:

(i) $\partial R_I/\partial s \neq 0$, so by the implicit function theorem and the fact that R_I is monotone there is a smooth function $s(x)$ on $\mathbf{R}^4$ such that $R_I(x, s(x)) = 4\pi^2$.

(ii) $s(x)$ (which is rotationally symmetric) has a unique nondegenerate minimum at 0 with minimum value K say.

Each of these is an open property, so will be shared by functions sufficiently close to R_I.

Definition 15. *For any connection A over a Riemannian manifold define*

(a) $R_A(x, s) = \int \beta_s(x, y) |F(A)_y|^2 \, d\mu_y$,

(b) $\lambda(A) = K^{-1} \min\{s \mid \exists x,\ R_A(x, s) = 4\pi^2\}$.

Thus $\lambda(A)$ is essentially the radius of the smallest ball containing action $4\pi^2$, but this definition has been smoothed by β. Clearly, by Theorem 13, any sequence $[A_i] \in \mu$ without convergent subsequences has $\lambda(A_i) \to 0$.

Suppose that A is a self-dual connection on P, $\lambda(A) = \lambda$, and we pick a point $x \in X$ where the minimum in Definition 15(b) is attained. Then for some fixed small $r > 0$ we can take a "conformal chart" in the manner of Theorem 13, but this time scaling by a factor λ:

$$\chi_\lambda : (r/\lambda)B^4 \subset \mathbf{R}^4 \to B(x, r) \subset X.$$

Define $\hat{A} = \chi_\lambda^*(A)$, and let the induced metric on $(r/\lambda)B^4$ be $\lambda^2 \hat{m}$; so as in Theorem 13, $\hat{m}$ tends to the Euclidean metric as $\lambda \to 0$, uniformly with all its derivatives on bounded regions in $\mathbf{R}^4$.

Theorem 16. (i) *On each bounded region of* $\mathbf{R}^4$, $\hat{A} \to I_{0,1}$, *uniformly in* C^∞ *for* $[A] \in \mathfrak{M}$, *as* $\lambda(A) \to 0$.

(ii) *There is a function* $\delta(r)$ *tending to zero as* $r \to 0$ *such that the curvature densities satisfy a bound*:

$$|F(\hat{A}(y))| \leqslant \text{const.}|y|^{-4+\delta}, \qquad y \in \mathbf{R}^4,\ |y| \leqslant \tfrac{1}{2} r\lambda^{-1}.$$

Note. We use the convention of writing const. for a general constant depending at bottom only on the Riemannian metric of X and always independent of λ.

Proof. (i) We apply Proposition 12 again. If A_i is a sequence of self-dual connections on P with $\lambda(A_i) \to 0$, then the corresponding $\hat{A}_i$ are self-dual in metric $\hat{m}_i$ and satisfy

$$\int_{\mathbf{R}^4} |F(\hat{A})|^2 \, d\mu \leqslant 8\pi^2.$$

The metrics $\hat{m}_i$ converge to the Euclidean metric on $\mathbf{R}^4$; thus we may run the argument of Theorem 13 again to deduce that the $\hat{A}_i$ have a convergent subsequence with limit a self-dual connection on $\mathbf{R}^4$. This time we do not have curvature gathering over points because of the normalisation chosen: $\lambda(\hat{A}_i) = 1$. The convergence is uniform on bounded regions, and again by the choice of normalisation the limit must be the instanton $I_{0,1}$ (by the classification discussed in §I.4 and the property (ii) of R_I above).

Since the limit is the same for every possibly subsequence the rate of convergence depends only upon $\lambda(A)$.

(ii) This bound will become rather important in our development of §III.4. To preserve continuity we give the proof in the Appendix.

Corollary 17. *There is a $\lambda_0 > 0$ such that if A is a self-dual connection on P with $\lambda(A) < \lambda_0$, then the minimum in Definition* 15(*b*) *is attained at a unique point $x(A)$ in X and the map*

$$p\colon \mathfrak{M}_{\lambda_0} = \{[A] \in \mathfrak{M} \mid \lambda(A) < \lambda_0\} \to X \times (0, \lambda_0),$$
$$p(A) = (x(A), \lambda(A)),$$

extends to a smooth map on some neighborhood of $\mathfrak{M}_{\lambda_0}$ in $\mathfrak{B}$. If $[A_t]$ is any smooth path in $\mathfrak{B}$, $A_0 \in \mathfrak{M}_{\lambda_0}$, there is a bound:

$$\left|\frac{d}{dt}p(A_t)\right| \leqslant \text{const.}\lambda(A_0)\left\|\frac{d}{dt}F(A_t)\right\|_{L^2},$$

both derivatives being evaluated at $t = 0$.

Proof. Any two minima must be separated by a distance of at most 2λ, since the ball of radius λ about each contains more than $\frac{1}{2}$ of the total action. Since the question is also conformally invariant, we may change scale and pass to $\hat{A}$.

The curvature densities $|F(\hat{A})|^2_{\hat{m}}$ converge uniformly on bounded regions of $\mathbf{R}^4$ to the instanton action density $f = |F(I_{0,1})|^2$ as $\lambda(A) \to 0$. So by the openness of the two properties of R_I noted before Definition 15 there will be a unique minimum corresponding to $\hat{A}$ for sufficiently small λ, and the position of a nondegenerate minimum varies smoothly with parameters, so the map p defined above is smooth.

Likewise for the second part; the statement is again scale invariant, so we may pass to the corresponding path $\hat{A}_t, f_t = |F(\hat{A}_t)|^2$. Thus

$$\begin{aligned}\left.\frac{d\lambda(\hat{A})}{dt}\right|_{t=0} &= k^{-1}\int_{\mathbf{R}^4}\beta_k(0, y)\dot{f}_t(y)\,|_{t=0}\,d\mu_y \\ &\leqslant \text{const.}\int_{\mathbf{R}^4}|\dot{f}_t(y)|\,d\mu_y \\ &\leqslant \text{const.}\|F(\hat{A})\|_{L^2}\left\|\frac{d}{dt}F(\hat{A}_t)\right\|_{L^2}.\end{aligned}$$

Similarly

$$\left\|\frac{d}{dt}x(\hat{A})\right\|_{t=0} = H^{-1}\left\{\int_{\mathbf{R}^4}[\operatorname{grad}_y\beta_k(0, y)]\,\dot{f}_t(y)\,|_{t=0}\,d\mu_y\right\},$$

where H is the Hessian corresponding to the minimum at $t = 0$.

To conserve notation we will allow ourselves to rechoose the positive number λ_0 throughout the rest of this section, to give the space $\mathfrak{M}_{\lambda_0}$ the desired properties. In any case $\mathfrak{M}\setminus\mathfrak{M}_{\lambda_0}$ is compact by Theorem 13.

Proposition 18. *Let* $[A] \in \mathcal{M}_{\lambda_0}$.

(i) *Given* $r > 0$, $[A] \to \theta$ *in* $C^\infty(X \setminus B(x(A), r))$ *as* $\lambda(A) \to 0$.

(ii) *The first eigenvalue* $\mu(A)$ *of* Δ_A *on* $\Omega^2_-(\mathfrak{g})$ *tends to the first eigenvalue* μ_1 *of* Δ *on* Ω^2_- *as* $\lambda(A) \to 0$. *Since* $\mu_1 > 0$, $\mathcal{M}_{\lambda_0}$ *is a smooth 5-manifold for* λ_0 *sufficiently small by the discussion of* §II.

(iii) $\|F(A)\|_{L^4} \leq \text{const.}\lambda(A)^{-1}$.

Proof. (i) is just a re-statement of Theorem 13, alternative (ii), with uniformity in λ just as in Theorem 16.

(ii) If ω is the first (normalised) eigenfunction of Δ_A, then $\|d_A^*\omega\|^2_{L^2} = \mu(A)$, and we compare ω with $\omega' = (1 - \beta_r)\omega$ where β_r is supported in $B(x(A), r)$, equal to 1 in the half-sized ball. Since $\|d_\theta^*\omega'\|^2_{L^2} \geq \mu_1\|\omega'\|^2_{L^2}$, and $A \to \theta$ on supp ω' by (i) above, we deduce the result by taking r sufficiently small. This goes precisely as [18, Proposition 8.8], so we omit the details.

(iii) The L^4-norm on 2-forms transforms by a factor λ^{-1} under a scale change of λ, so it plainly suffices to show $\|F(\hat{A})\|_{L^4}$ bounded; but this follows from Theorem 16 (ii) once r, and so $\delta(r)$, is taken small enough for

$$\left(1/|y|^{[4-\delta(r)]}\right)$$

to be integrable at infinity.

Theorem 19. *For sufficiently small* λ_0,

$$p\colon \mathcal{M}_{\lambda_0} \to X \times (0, \lambda_0)$$

is a covering map.

Proof. First note that Theorem 13 implies that p is a proper map, and it is easy to see that a proper local homeomorphism is a (finite) covering, so it suffices to prove that

$$dp\colon (T\mathcal{M})_{[A]} \to (TX)_{x(A)} \times \mathbf{R}$$

is an isomorphism for $[A] \in \mathcal{M}$, $\lambda(A)$ sufficiently small. By the application of the index theorem described in §II.2 we know each space is 5-dimensional, so it suffices again to exhibit a linear map:

$$\alpha\colon (TX)_{x(A)} \times \mathbf{R} \to (T\mathcal{M})_{[A]}$$

such that $dp \circ \alpha$ is invertible, which will certainly be true if

$$(1) \qquad |(dp \circ \alpha)\xi - \xi| < |\xi| \quad \text{for all } \xi \in TX \times \mathbf{R}.$$

To define α we show that one can displace a solution A by an approximately conformal transformation of X. Choose r so small that $\delta(r) < 1$ say, and let f_t be the flow on X given in local geodesic coordinates x^i by the vector field

$$V = \beta_{r/2}(v^i + \phi x^i)\frac{\partial}{\partial x^i},$$

where $\xi = (v^i, \phi)$. Thus $f_t = 1$ outside $B(x(A), r/2)$. Lifting to the bundle f_t defines a flow on $\mathcal{B}$; set $[A_t] = q(f_t(A)) \in \mathfrak{M}$ where q is Taubes' map as in Proposition 5, so $[A_t]$ is defined for sufficiently small t (depending on λ).

Then put $\alpha(\xi) = \frac{d}{dt}[A_t]|_{t=0} \in (T\mathfrak{M})_{[A]}$. α is plainly linear since the whole process varies smoothly with ξ. To check property (1): observe first that if m_t is a path of Euclidean structures on $\mathbf{R}^4$ with m_0 the standard one, then for $\omega \in \Lambda^2\mathbf{R}^4$,

$$\frac{d}{dt}\left|\omega - *_{m_t}\omega\right| \leqslant \text{const.}\,|\omega|\left|\pi\left(\frac{dm}{dt}\right)\right|,$$

where π is the projection onto the "trace free" part. Hence on X,

$$\frac{d}{dt}|F_-(f_t(A)| \leqslant \text{const.}|F(A)|\,|\pi \mathfrak{L}_V m|.$$

(All derivatives $\frac{d}{dt}$ in this proof are to be understood as being evaluated at $t = 0$; $\mathfrak{L}_V$ is the Lie derivative.) Furthermore, one easily sees that, because x^i are geodesic and so f_t approximately conformal,

$$|\pi \mathfrak{L}_V m| \leqslant \text{const.}\,\rho|\xi|, \quad 0 \leqslant \rho \leqslant r/2$$

where the left-hand side is evaluated at $x \in X$, and $\rho = d(x, x(A))$. Scaling down the bounds of Theorem 16 with $\delta = 1$ we have

$$\begin{aligned} |F(A)| &\leqslant \text{const}\,\lambda^{-2}|\xi| \quad 0 \leqslant \rho \leqslant \lambda,\\ &\leqslant \text{const}\,\lambda\rho^{-3}|\xi| \quad \lambda \leqslant \rho \leqslant r/2. \end{aligned}$$

So

$$\frac{d}{dt}\|F_-(f_t(A))\|_{L^p} \leqslant \text{const}\left\{\int_0^\lambda (\rho\lambda^{-2})^p\rho^3 d\rho + \int_\lambda^{\frac{1}{2}r}(\lambda\rho^{-3})^p\rho^3 d\rho\right\}^{1/p}|\xi|,$$

which gives

$$\frac{d}{dt}\|F_-(f_t(A))\|_{L^2}, \quad \frac{d}{dt}\|F_-(f_t(A))\|_{L^{4/3}},$$

each bounded by const $\lambda.|\xi|$. (We would improve the bound on the $L^{4/3}$ norm by taking δ smaller, but this is not necessary here.) So using the bounds on $q(A_0) - A_0$ in Proposition 5 and recalling that $A_t = q(f_t(A))$, we obtain

$$\begin{aligned} \left\|\frac{d}{dt}(A_t - f_t(A))\right\|_{L_1^2(A)} &\leqslant \text{const.}\left\{\frac{c}{dt}\|F_-(f_t(A))\|_{L^2}\right.\\ &\qquad\left. + \lambda^{-1}\frac{d}{dt}\|F_-(f_t(A))\|_{L^{4/3}}\right\}\\ &\leqslant \text{const.}|\xi|, \end{aligned}$$

since by Proposition 18 $\|F(A)\|_{L^4} \leqslant \text{const.}\lambda^{-1}$, and the eigenvalue $\mu(A)$ is not small. But this implies $\|\frac{d}{dt}(F(A_t) - F(f_t(A)))\|_{L^2} \leqslant \text{const.}\,|\xi|$. So by Corollary 17, applied to the path $A_t - f_t(A)$,

$$\left|\frac{d}{dt}p(A_t) - \frac{d}{dt}pf_t(A)\right| \leqslant \text{const.}\lambda|\xi|,$$

i.e.,

$$\left|(dp \circ \alpha)\xi - \frac{d}{dt}pf_t(A)\right| \leqslant \text{const.}\lambda|\xi|.$$

If the metric on X were flat in the ball of radius r about $x(A)$, then by the construction $\frac{d}{dt}p(f_t(A))$ would equal ξ exactly. In general one easily sees that the difference is bounded by $\text{const.}\lambda.|\xi|$, so

$$|(dp \circ \alpha)\xi - \xi| \leqslant \text{const.}\lambda.|\xi|,$$

and, property (1) is satisfied once $\lambda = \lambda(A)$ is sufficiently small.

III.4. The proof that p is a diffeomorphism

To prove that $p\colon \mathfrak{M}_{\lambda_0} \to X \times (0, \lambda_0)$ is a diffeomorphism, it suffices now to show that if we pick some point x_0 in X, then for λ_0 sufficiently small any pair $[A], [B] \in \mathfrak{M}_{\lambda_0}$ of self-dual connections with

$$x(A) = x(B) = x_0, \quad \lambda(A) = \lambda(B) = \lambda$$

can be joined by a short path in $\mathfrak{M}$ (in a sense made precise below). The metric on X has always been at our choice, and there is a small saving of labour here if we suppose it to be *flat* in some ball $B(x_0, r)$—although this does not, of course, affect the truth of the result as one easily sees by deformation.

First we note a small lemma.

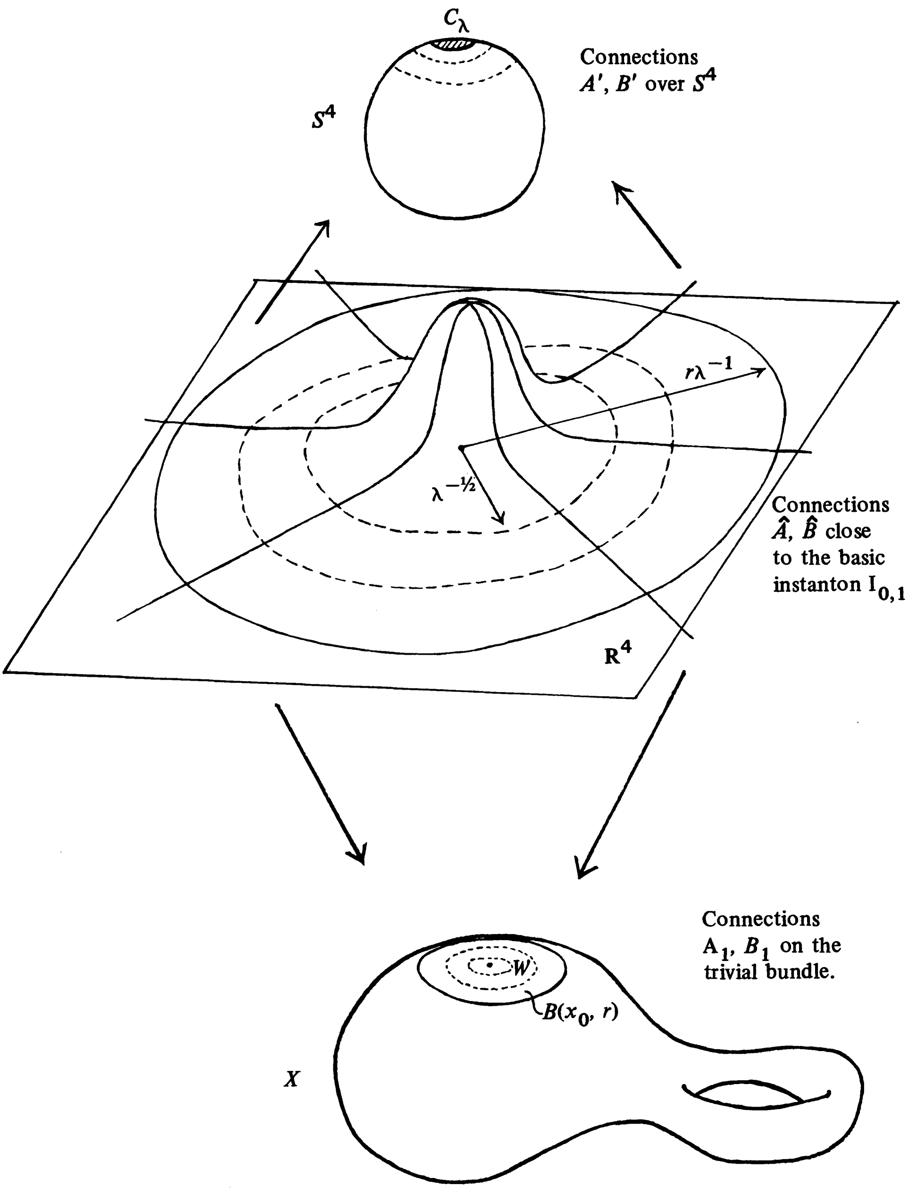
Lemma 20. *There are constants $K, C > 0$, independent of R, such that if A is a connection defined on the trivial bundle over the annulus*

$$\{x \in \mathbf{R}^4 \mid R/2 < |x| < 2R\}$$

with $\|F(A)\|_{C^0} \leqslant K/R^2$, we may find a connection $\tilde{A}$ over the ball $2RB^4$, equal to A in $R < |x| < 2R$, and having

$$\|F(\tilde{A})\|_{C^0} \leqslant C.\|F(A)\|_{C^0}.$$

Proof. Under conformal re-scaling by factor R^{-1}, the norms $\|F(A)\|_{C^0}$, $\|F(\tilde{A})\|_{C^0}$ each transform by a factor R^2, so it suffices to do this when $R = 1$.



Thus once K is small enough we may find an "exponential gauge" in the manner of [20, §2], so that without loss of generality the connection matrix over the annulus satisfies

$$\|A\|_{C^0} \leqslant \text{const.}\|F(A)\|_{C^0}.$$

Then set $\tilde{A} = (1 - \beta)A$, whence

$$F(\tilde{A}) = d\tilde{A} + \tfrac{1}{2}[\tilde{A}, \tilde{A}] = (1 - \beta)F(A) - d\beta \wedge A + \tfrac{1}{2}(\beta^2 - 3\beta)[A, A],$$

so $\|F(\tilde{A})\|_{C^0} \leqslant \text{const.}\|F(A)\|_{C^0}$ as required.

This lemma is used to compare the connection A on the bundle P over X with the instanton and the flat connection. First we need some more notation: for the instanton $I_{y,v}$ over $\mathbf{R}^4$, with centre y and scale v, write $I'_{y,v}$ for the corresponding connection over S^4. Similarly, the connection $\hat{A}$ over the large ball $\frac{1}{2}r\lambda^{-1}B^4 \subset \mathbf{R}^4$ has curvature bounded by

$$|F(\hat{A})(y)| \leqslant \text{const}/|y|^4,$$

since in this case we may take $\delta = 0$ in Theorem 16. Thus we may conformally map this large ball to the complement in S^4 of a small cap C_λ over the north pole, of radius $0(\lambda)$, and the connection we get there has uniformly bounded curvature, independent of λ. By Lemma 20 we may extend this connection to a connection, A' say, defined over all of S^4, and maintain a uniform bound on the curvature of A'. Clearly A' has Chern class -1.

Theorem 21. (i) *If A_1 be the restriction of the connection A to $X \setminus B(x_0, \sqrt{\lambda})$, then after a suitable bundle automorphism*

$$\|A_1 - \theta\|_{L_1^p(\theta)} \leqslant \text{const.}\lambda^{2/p}, \quad p > 1.$$

(ii) *Similarly, after a suitable bundle automorphism,*

$$\|A' - I'_{0,1}\|_{L_1^p(I'_{0,1})} \leqslant \text{const.}\lambda^{4/p}, \quad p \geqslant 2.$$

Proof. (i) By Theorem 16 (ii), scaled down:

$$|F(A)_x| \leqslant \text{const.}\lambda^2/\rho^4, \qquad \rho = d(x, x_0).$$

So the curvature is uniformly bounded, independent of λ, on the annulus $\frac{1}{2}\sqrt{\lambda} \leqslant \rho \leqslant 2\sqrt{\lambda}$, and for λ sufficiently small we may extend A_1 over all of X (by Lemma 20) to a connection on the trivial bundle, preserving a uniform bound on the curvature. Thus

$$\|F_-(A_1)\|_{L^p} \leqslant \text{const.}\left[\text{vol.}B\left(x_1, \sqrt{\lambda}\right)\right]^{1/p} \leqslant \text{const.}\lambda^{2/p}.$$

Furthermore, just as in Proposition 18, the eigenvalue $\mu(A_1)$ is not small, so we may apply the implicit function theorem of Taubes (Proposition 5) to find a self-dual connection $A_1 + a = q(A_1)$, which is without loss of generality the product connection θ, with $\|a\|_{L_1^2(A)} \leqslant \text{const.}\lambda$. Moreover by the form of the construction, $a = d_{A_1}^* u$ so that

(a) $d_\theta^* a = d_{A_1}^* a + \{u, a\} = \{F_-(A_1), u\} + \{a, a\}$,

(b) $d_\theta^- a = \{a, a\} + F_-(A_1)$.

(We write $\{\ ,\ \}$ for algebraic bilinear expressions whose particular form is not important here.) By the ellipticity of $d^* + d^-$,

$$\|a\|_{L^p_1} \leqslant \text{const.}\left[\|(d_\theta^* + d_\theta^-)a\|_{L^p} + \|a\|_{L^p}\right],$$

and combined with standard estimates of the quadratic terms in (a), (b) this gives

$$\|a\|_{L^p_1} \leqslant \text{const.}\|F_-(A_1)\|_{L^p} \leqslant \text{const.}\lambda^{2/p},$$

once λ is sufficiently small. (The situation is not at all delicate here, compared with [18], because of the uniform bound on the curvature of A_1.)

(ii) We apply the same argument to the connection A' over S^4 with

$$\|F_-(A')\|_{L^p} \leqslant \text{const.}\,(\text{Vol}.C_\lambda)^{1/p} \leqslant \text{const.}\lambda^{4/p}.$$

One obtains a lower bound on the eigenvalues $\mu(A')$ by using the Weitzenböch formula for connections over S^4 [18, Proposition 2.2], [2, p. 145]:

$$\Delta_A\omega = \tfrac{1}{2}\nabla_A^*\nabla_A\omega + \tfrac{2}{3}s.\omega + \{F_-, \omega\}, \quad \omega \in \Omega_-^2(\mathfrak{g}),$$

where $s =$ Scalar curvature of $S^4 > 0$.

So this time we get $\|A' - I'_{y,\nu}\|_{L^p_1} \leqslant \text{const.}\lambda^{4/p}$ for some y, ν. Now in general, if A, $A + a$ are connections, then

$$F(A + a) = F(A) + d_A a + \tfrac{1}{2}[a, a]$$

implies

$$\|F(A + a) - F(A)\|_{L^2} \leqslant \text{const.}\left\{\|a\|_{L_1^2(A)} + \|a\|_{L^4}^2\right\},$$

and in dimension 4, $L_1^2(A) \hookrightarrow L^4$ with embedding constant independent of A [18, Lemma 5.2]. Thus in our case,

$$\|F(A') - F(I'_{y,\nu})\|_{L^2} \leqslant \text{const.}\lambda^2,$$

from which one easily sees, by the normalisation $\lambda(\hat{A}) = 1$, $x(\hat{A}) = 0$, that $|y| + |\nu - 1| \leqslant \text{const.}\lambda^2$, so

$$\|A' - I'_{0,1}\|_{L^p_1(A')} \leqslant \text{const.}\lambda^{4/p}, p \geqslant 2.$$

The bounds (i), (ii) of Theorem 21 will also, of course, be satisfied by the connection B. Thus to complete the proof of Theorem 11 we have to use these to show that A and B can be joined, after suitable bundle automorphism, by a path in $\mathcal{C}$ which is short enough to project onto the moduli space. This is slightly easier to write down if we use a small modification of Proposition 5.

Proposition 22. *Let A be any self-dual connection on P, there are constants C, ε, depending only upon the first eigenvalue $\mu(A)$ and the metric on X, such that if $b \in \Omega^1(\mathfrak{g})$ and*

$$\delta(A, b) = \|\nabla_A b\|_{L^2} + \|b\|_{L^2}\|F(A)\|_{L^4} < \varepsilon,$$

we may construct a self-dual connection $A + b + q(b, A) = A + b + a$ with

$$\|\nabla_A a\|_{L^2} \leq C\left\{\|F_-(A + b)\|_{L^2} + \|F(A)\|_{L^4}\|F_-(A + b)\|_{L^{4/3}}\right\},$$
$$\|a\|_{L^2} \leq C\|F_-(A + b)\|_{L^{4/3}}.$$

The construction is $\mathcal{G}$-invariant ($\mathcal{G}$ acting on A, a, b), and a varies smoothly with b.

This proposition follows by the method of proof of [18, Theorem 2.2], applied to the equation $F_-(A + b + d_A^* u) = 0$. In the equation corresponding to [18, (5.2)] we lose the term in $F_-(A_0)$, which means that we do not need to involve $\|F_-\|_{L^3}$ in [18, (5.12.b)].

Lemma 23. *Suppose that A, B are self-dual connections on P, $x(A) = x(B) = x_0$, $\lambda(A) = \lambda(B) = \lambda$. If λ is sufficiently small, then we may suppose that, after a suitable bundle automorphism; $B = A + b$ with*

$$\|\nabla_A b\|_{L^2} \leq \text{const.}\lambda^{1/2}, \quad \|b\|_{L^2} \leq \text{const.}\lambda^{3/2}.$$

Proof of Theorem 11, *given Lemma* 23. Suppose that, on the contrary, p: $\mathfrak{M}_{\lambda_0} \to X \times (0, \lambda_0)$, which is a covering by Theorem 19, had at least two sheets. Then for arbitrarily small λ we could find A, B as in Lemma 23 representing different sheets. By Proposition 18 the eigenvalue $\mu(A)$ is not small, and $\|F(A)\|_{L^4} \leq \text{const.}\lambda^{-1}$. Thus we can apply Proposition 22 to the path $A + tb$ from A to B, and find a path in $\mathfrak{M}$, $[A(t)]$ say: $[A(0)] = [A]$, $[A(1)] = [B]$. Moreover $\|F(A(t)) - F(A)\|_{L^2} \leq \text{const.}\lambda^{1/2}$, from which one easily sees, as in Corollary 17, that

$$|\lambda(A(t)) - \lambda(A)| + d\big(x(A(t)), x_0\big) \leq \text{const.}\lambda^{3/2}.$$

So, if λ is much less than λ_0, $A(t)$ stays in $\mathfrak{M}_{\lambda_0}$, and $x(A(t))$ is close to x_0. Hence $[A]$, $[B]$ are in the same local component of the covering and so equal.

On the other hand $\mathfrak{M}_{\lambda_0}$ is not empty. By Taubes construction there are self-dual connections on X with $\lambda(A)$ arbitrarily small, which one may see by estimating the curvature density of the solutions in [18], in the manner above.

Proof of Lemma 23. We proceed in three stages, corresponding to the three regions $X \setminus B(x_0, \sqrt{\lambda})$, $B(x_0, 2\sqrt{\lambda})$ and $W = B(x_0, 2\sqrt{\lambda}) \setminus B(x_0, \sqrt{\lambda})$.

(a) Using Theorem 21(i) and the Sobolev embedding theorems in dimension 4 we may suppose that A is given over $X \backslash B(x_0, \sqrt{\lambda})$ by a connection matrix A_1 with

$$\|A_1\|_{L^2} \leqslant \text{const.}\|A_1\|_{L_1^{4/3}} \leqslant \text{const.}\lambda^{3/2},$$

$$\|A_1\|_{C^0} \leqslant \text{const.}\|A_1\|_{L_1^8} \leqslant \text{const.}\lambda^{1/4},$$

$$\|A\|_{L_1^2} \leqslant \text{const.}\lambda.$$

Similarly B is represented by some B_1. Thus in this region $B = A + b_1$ (up to equivalence) with $\|b_1\|_{L_1^2} \leqslant \text{const.}\lambda^{1/2}$, $\|b_1\|_{L^2} \leqslant \text{const.}\lambda^{3/2}$.

(b) First we transform the bounds of Theorem 21(ii) on the connections A', B' over S^4 to the small ball $B(x_0, 2\sqrt{\lambda})$ by the obvious conformal transformation. Thus on S^4 we may suppose $B' = A' + b'$, $\|b'\|_{L_1^p(A')} \leqslant \text{const.}\lambda^{4/p}$, $p \geqslant 2$. So $\|b'\|_{C^0} \leqslant \text{const.}\lambda^{1/2}$ say, $\|b'\|_{L^2} \leqslant \text{Const.}\lambda^2$. Taking account of conformal factors gives, over $B(x_0, 2\sqrt{\lambda})$,

$B = A + b_2$ say, with

$$\|b_2\|_{L^2} \leqslant \text{const.}\lambda^2, \text{ and } \|b_2\|_{C^0(W)} \leqslant \text{const.}\lambda^{1/2}.$$

The norm on 1-forms $\|\nabla(\ )\|_{L^2}$ is independent of constant scale changes, but in going from S^4 to $B(x_0, 2\lambda^{1/2})$ we have a varying conformal factor which introduces a lower order term. One easily calculates

$$\begin{aligned}\|\nabla_{I_\lambda} b_2\|_{L^2} &\leqslant \|\nabla_{I_1}(\hat{A} - \hat{B})\|_{L^2} \\ &\leqslant \text{const.}\|b'\|_{L_1^2(I_1')} + \left[\int_{\mathbf{R}^4} \left(\frac{|\hat{A} - \hat{B}|}{1 + |y|}\right)^2 d\mu_y\right]^{1/2} \\ &\leqslant \text{const.}\lambda^{1/2},\end{aligned}$$

since $|(\hat{A} - \hat{B})(y)| \leqslant \text{const.}\|b'\|_{C^0}/|y|^2$, $y \in \mathbf{R}^4$. So also $\|b_2\|_{L_1^2(I_\lambda)} \leqslant \text{const.}\lambda^{1/2}$. Moreover, it does not matter whether we use $L_1^2(I_\lambda)$ or $L_1^2(A)$ here.

(c) The region W corresponds under the conformal transformation of (b) to a small annulus of radius $0(\lambda^{1/2})$ about the north pole of S^4. Thus we may find connection matrices A_2, B_2 for A, B respectively over $B(x_0, 2\lambda^{1/2}) \backslash \{x_0\}$ such that in W

$$\text{(i)} \qquad \begin{cases} \|A_2\|_{L^2(W)} \leqslant \text{const.}\lambda^{3/2}, \\ \|A_2\|_{C^0(W)} \leqslant \text{const.}\lambda^{1/2}, \qquad \text{with similar bounds on } B_2, \\ \|\nabla A_2\|_{L^2(W)} \leqslant \text{const.}\lambda^{1/2}. \end{cases}$$

(Since we may find such a connection matrix for $I_{0,\lambda}$, corresponding to a stationary framing around the north pole of S^4, and then A is close to $I_{0,\lambda}$ by (b) above.)

On W we have transition functions:

$$g, h\colon W \to SU(2),$$

$$\text{(ii)} \qquad \begin{cases} A_2 = -dgg^{-1} + gA_1g^{-1}, \\ B_2 = -dhh^{-1} + hB_1h^{-1}. \end{cases}$$

If it so happened that $g = h$, then we would be done, since the pair $A_2 - B_2$, $A_1 - B_1$ would represent a difference element $b \in \Omega^1(\mathfrak{g})$, $B = A + b$ after automorphism and satisfying the required bounds by parts (a), (b). So the final step is to check that we can find another set of data, $(B_1, \tilde{B}_2, g|_{W'})$ say, for $[B]$, where W' is the smaller annulus $B(x_0, \frac{3}{2}\lambda^{1/2})\backslash B(x_0, \lambda^{1/2})$, with $\tilde{B}_2 = B_2$ outside of W', and keeping $A_2 - \tilde{B}_2$ small.

Since the connection θ is fixed by a constant rotation of the bundle, we may suppose that at some point of W $g = h = 1$ (note that this would not be true if we had an irreducible flat connection in place of θ). Then (i) and (ii) together with part (a) give

$$\|dg\|_{C^0(W)} \leqslant \text{const.}\lambda^{1/4},$$

hence by integration, $\|g - 1\|_{C^0(W)} \leqslant \text{const.}\lambda^{3/4}$. Thus if $u = gh^{-1} = \exp(s)$, $\|s\|_{C^0(W)} \leqslant \text{const.}\lambda^{3/4}$ and L^p_k bounds on u are equivalent to L^p_k bounds on s. Set $v = \exp(\gamma_{\lambda^{1/2}}s)$, where

$$\gamma_{\lambda^{1/2}} = 0 \text{ on } B\left(x_0, \lambda^{1/2}\right) = 1 \text{ outside } B\left(x_0, \tfrac{3}{2}\lambda^{1/2}\right),$$

so that $(B_1, \tilde{B}_2, g|_{W'})$ is another representation of $[B]$ with

$$\tilde{B}_2 = -dvv^{-1} + vB_2v^{-1}.$$

Then one may check that $\|\tilde{B}_2\|_{L^2}$, $\|\tilde{B}_2\|_{L^2_1}$ are appropriately small. For example,

$$\begin{aligned}
\|\nabla\tilde{B}_2\|_{L^2(W)} &\leqslant \text{const.}\left\{\|\nabla B\|_{L^2(W)} + \|B_2\|_{L^2(W)} + \|\nabla\nabla v\|_{L^2(W)}\right\} \\
&\qquad\qquad \left(\text{since } |\nabla v| \to 0 \text{ uniformly}\right) \\
&\leqslant \text{const.}\left(\lambda^{1/2} + \|(\nabla\nabla\gamma)s\|_{L^2(W)} + \|\nabla\gamma.\nabla s\|_{L^2(W)}\right\} \\
&\leqslant \text{const.}\left(\lambda^{1/2} + \lambda^{-1}\|s\|_{L^2(W)} + \lambda^{-1/2}\|\nabla s\|_{L^2(W)}\right\} \\
&\leqslant \text{const.}\left(\lambda^{1/2} + \lambda^{-1}\lambda^{3/4}(\text{Vol } W)^{1/2} + \lambda^{-1/2}\lambda^{3/2}\right) \\
&\leqslant \text{const.}(\lambda^{1/2} + \lambda^{3/4} + \lambda) \leqslant \text{const.}\lambda^{1/2}.
\end{aligned}$$

Appendix. Outline proof of Theorem 16 (ii)

Uhlenbeck's theorem on the removeability of singularities (Proposition 4), applied to the point at infinity, shows that any self-dual connections I on $\mathbf{R}^4$ with finite action satisfies [20, Cor. 4.2]:

$$|F(I)(y)| \leqslant \text{Const.}(I)/|y|^4, \quad y \in \mathbf{R}^4.$$

The result which we need here is a small variant of this—since our connection $\hat{A}$ is only defined over a large ball. However the method of proof of [20] applies, and the proof becomes simpler if one only considers self-dual, rather than general Yang-Mills, connections. For completeness we shall explain this briefly, and refer to [20] for the lengthier parts of the argument.

It is convenient to work with a self-dual connection A defined over the cylinder $S^3 \times [0, T]$ which we suppose for simplicity has its standard Riemannian metric. Thus for each $t \in [0, T]$ we have a connection $A_t = A|_{S^3 \times \{t\}}$ on the standard 3-sphere. We shall show that there are constants $C, \varepsilon > 0$ (independent of T, A) such that if

$$\|F(A)\|_{L^2}^2 \leqslant \varepsilon, \; 0 \leqslant Tc_2(A_T),$$

then the supremum $M(t)$ of $|F(A)|$ on $S^3 \times \{t\}$ is bounded by

$$Ce^{-2t} \text{ for, say, } 1 \leqslant t \leqslant T-1.$$

The "a priori estimate" [20, Theorem 3.5] shows that if ε is sufficiently small, the absolute value of the curvature of a self-dual connection A at a point (x, t) is bounded by a multiple of the L^2 norm of the curvature in a ball of radius 1 in $S^3 \times [0, T]$ about (x, t). Thus it suffices to show that

$$\int_{S^3 \times [t, T]} |F(A)|^2 \, d\mu \leqslant \text{const.} e^{-4t}.$$

But this integral depends, as explained in §I.3, only on the boundary values:

$$\int_{S^3 \times [\varepsilon, T]} |F(A)|^2 \, d\mu = \frac{1}{8\pi^2}\left(Tc_2(A_t) - Tc_2(A_T)\right) \leqslant \frac{1}{8\pi^2} Tc_2(A_t),$$

by the hypothesis on A. We shall compare $J(t) = \dfrac{1}{8\pi^2} Tc_2(A_t)$ with its derivative

$$\frac{dJ}{dt} = -\int_{S^3 \times \{t\}} |F(A)|^2 \, d\mu.$$

Note first that if we choose a distinguished subspace $\mathbf{R}^3$ of a Euclidean space $\mathbf{R}^4$ there are natural isomorphisms

$$\Lambda^2 \mathbf{R}^4 \cong \Lambda^2 \mathbf{R}^3 \oplus \Lambda^1 \mathbf{R}^3 \cong \mathbf{R}^3 \oplus \mathbf{R}^3$$

corresponding in "electro-magnetic" notation to $F \to (\underline{E}, \underline{B})$, and the self-duality condition becomes $\underline{E} = \underline{B}$. Applying this to the tangent spaces of $S^3 \times \mathbf{R}$ we get

$$\frac{dJ}{dt} = -2\int_{S^3 \times \{t\}} |F(A_t)|^2$$

Thus both J, dJ/dt depend only upon the boundary value A_t.

We use the fact that for any 1-form ω on S^3,

$$\int_{S^3} \omega \wedge d\omega \leqslant \frac{1}{2}\int_{S^3} |d\omega|^2 \, d\mu.$$

(If we consider S^3 as the unit sphere in $\mathbf{H}$, then the bound is attained by each of the three imaginary quaternion components of the 1-form $\bar{x}\,dx$.) This is applied in the following way; once ε, and so also $M(t)$, is sufficiently small we can choose, for each fixed t a connection matrix A_t over S^3 with

$$\|A_t\|_{C^0} \leqslant \text{const.}\, M(t).$$

So J, dJ/dt differ by small amounts from their second order approximations, respectively:

$$J = \frac{1}{8\pi^2} Tc_2(A_t) = \int_{S^3} Tr\Big(A_t \wedge \Big(dA_t + \frac{3}{2}[A_t, A_t] \Big)\Big)$$

$$= \int_{S^3} Tr(A_t \wedge dA_t) + \text{Higher order terms.}$$

$$\frac{dJ}{dt} = -2\|F(A_t)\|^2 = -2\int_{S^3} |dA_t|^2 \, d\mu + \text{Higher order terms.}$$

Thus

$$J \leqslant -4\frac{dJ}{dt} + \text{Higher order terms.}$$

The higher order terms may be accommodated by iteration of the argument, just as in [20, Theorem 4.8], and once ε is chosen sufficiently small we can integrate this differential inequality to get

$$J \leqslant \text{Const.}\, J(1) e^{-4t} \leqslant \text{const.}\, e^{-4t}.$$

Hence the L^2 norm of $F(A)$ in $S^3 \times (t, T)$ is less than const.e^{-2t}, and so also the absolute value of F, for $1 \leqslant t \leqslant T - 1$.

Return now to the connection $\hat{A}$ of Theorem 16(ii), defined on the large ball of radius $r\lambda^{-1}$ in $\mathbf{R}^4$. If the metric $\hat{m}$ is flat, then given R we can conformally map the annulus

$$\{y \in \mathbf{R}^4 \mid R \leqslant |y| \leqslant r\lambda^{-1}\}$$

to $S^3 \times \mathbf{R}$. Since we know, by Theorem 16(i), that the connections $\hat{A}$ converge to the instanton on bounded regions of $\mathbf{R}^4$ as $\lambda \to 0$, we can choose R so large that the L^2 norm of the curvature is bounded by ε, and the boundary condition is satisfied, since $\hat{A}$ comes from a connection over X. Thus the bound above transforms to

$$|F(\hat{A})(y)| \leqslant \text{const.}/|y|^4.$$

If, more generally, the metric $\hat{m}$ is curved, then we can work the same argument with a small error δ in the exponent.

References

[1] M. F. Atiyah, *K-Theory*, Benjamin, New York, 1967.

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

[3] M. F. Atiyah & J. D. S. Jones, *Topological aspects of Yang-Mills theory*, Comm. Math. Phys. **61** (1978) 97–118.

[4] M. F. Atiyah & I. M. Singer, *The index of elliptic operators*. IV, Ann. of Math. **93** (1971) 119–138.

[5] J. P. Bourguignon & H. B. Lawson, Jr., *Yang-Mills theory: physical origins and differential geometric aspects*, Annals of Math. Studies, No. 102, Princeton University Press, Princeton, 1982, 395–421.

[6] S. S. Chern, *Complex manifolds without potential theory*, Math. Studies No. 15, Van Nostrand, Princeton, 1967.

[7] S. K. Donaldson, *Self-dual connections and the topology of smooth 4-manifolds*, Bull. Amer. Math. Soc. **8** (1983), 81–83.

[8] M. H. Freedman, *The topology of four-dimensional manifold* J. Differential Geometry **17** (1982) 357–453.

[9] W. V. D. Hodge, *The theory and applications of harmonic integrals*, Cambridge University Press, Cambridge, 1952.

[10] M. Kuranishi, *A new proof for the existence of locally complete families of complex structures*, Proc. Complex Analysis (Minneapolis), Springer, Berlin, 1964, 142–156.

[11] R. Mandelbaum, *Four dimensional topology: an introduction*, Bull. Amer. Math. Soc. **2** (1980) 1–159.

[12] J. W. Milnor, *On simply connected 4-manifolds*, Proc. Internat. Sympos. on Algebraic Topology, University of Mexico, 1958, 122–128.

[13] T. Parker, *Gauge theories on 4-dimensional Riemannian manifolds*, Comm. Math. Phys. **85** (1982) 563–602.

[14] J. Saks & K. K. Uhlenbeck, *On the existence of minimal immersions of 2-spheres*, Ann. of Math. **113** (1982) 1–24.

[15] S. Sedlacek, *A direct method for minimising the Yang-Mills functional*, Comm. Math. Phys. **86** (1982) 515–528.

[16] J. P. Serre, *A course in arithmetic*, Springer, Berlin, 1973.

[17] S. Smale, *An infinite dimensional version of Sard's theorem*, Amer. J. Math. **87** (1965) 861–866.

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

[19] ______, Private letter.

[20] K. K. Uhlenbeck, *Removeable singularities in Yang-Mills fields*, Comm. Math. Phys. **83** (1982) 11–30.

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

[22] J. H. C. Whitehead, *On simply connected 4-dimensional polyhedra*, Comment. Math. Helv. **22** (1949) 48–92.

OXFORD UNIVERSITY