NIGEL J. HITCHIN

The Yang-Mills equations and the topology of 4-manifolds

Astérisque, tome 105-106 (1983), Séminaire Bourbaki,
exp. nº 606, p. 167-178
<http://www.numdam.org/item?id=SB_1982-1983__25__167_0>

Séminaire BOURBAKI
35e année, 1982/83, n° 606 Février 1983

THE YANG-MILLS EQUATIONS
AND THE TOPOLOGY OF 4-MANIFOLDS
[after Simon K. Donaldson]

by Nigel J. HITCHIN

§ 1. The result

(1.1) THEOREM (S.K. Donaldson [8]).— *Let* X *be a compact, smooth, simply connected, oriented* 4*-manifold such that the intersection form* Q *on* $H^2(X,\mathbb{Z})$ *is positive definite. Then there exists an integral basis for* $H^2(X,\mathbb{Z})$ *such that* $Q(u,u) = u_1^2 + u_2^2 + \dots + u_r^2$.

This theorem should be contrasted with

(1.2) THEOREM (M.H. Freedman [9]).— *Let* Q *be any unimodular quadratic form over* $\mathbb{Z}$. *Then there exists a compact, simply connected, topological* 4*-manifold* X *such that* Q *is equivalent to the intersection form on* $H^2(X,\mathbb{Z})$.

There are sufficient examples of definite unimodular forms (see [17]) to see that Donaldson's theorem imposes strong restrictions on smooth 4-manifolds.

(1.3) Proof of Theorem (1.1)

Let $r = \text{rank } H^2(X,\mathbb{Z})$ and $2n = \# \{u \in H^2(X,\mathbb{Z}) \mid Q(u,u) = 1\}$. The proof consists of constructing (as in § 3 - § 7) an oriented cobordism between X and n copies of $\mathbb{C}P^2$. Let p of these have the canonical orientation of the complex structure and $q = n - p$ the opposite orientation. Then

(i) By the cobordism invariance of signature,

$$r = \text{Sign } X = (p-q) \text{ Sign } \mathbb{C}P^2 = p - q \leq n .$$

(ii) Let $\{\pm x_1, \pm x_2, \dots, \pm x_n\} = \{u \in H^2(X,\mathbb{Z}) \mid Q(u,u) = 1\}$, then $Q(x_i,x_j) \in \mathbb{Z}$ but by the Cauchy-Schwarz inequality $|Q(x_i,x_j)| < 1$ if $i \neq j$. Hence $\{x_1,\dots,x_n\}$ is orthonormal and $n \leq r$.

(iii) From (i) and (ii) $n = r$ and $\{x_1,\dots,x_n\}$ is an orthonormal basis for $H^2(X,\mathbb{R})$. Thus for $u \in H^2(X,\mathbb{Z})$, $u = \sum_{i=1}^n Q(u,x_i)x_i = \sum_{i=1}^n u_i x_i$ with $u_i \in \mathbb{Z}$ and $\{x_1,\dots,x_n\}$ is a basis for $H^2(X,\mathbb{Z})$. Hence $Q(u,u) = \sum_{i=1}^n u_i^2$.

§ 2. Background

(2.1) Let X be an oriented riemannian 4-manifold. A 2-form $\alpha \in \Omega^2$ is said to be *self-dual* (resp. *anti-self-dual*) if $*\alpha = \alpha$ (resp. $*\alpha = -\alpha$) where $* : \Omega^2 \longrightarrow \Omega^2$ denotes the Hodge star operator.

Let G be a compact Lie group and P a principal G-bundle over X. A connection A on P has curvature $F(A) \in \Omega^2(\mathfrak{g})$ where $\mathfrak{g}$ denotes the vector bundle associated to P by the adjoint representation. For any bundle V associated to P a connection A defines a differential operator $d_A : \Omega^p(V) \longrightarrow \Omega^{p+1}(V)$. The metric on X defines the formal adjoint $d_A^* : \Omega^{p+1}(V) \longrightarrow \Omega^p(V)$. The *Bianchi identity*, satisfied by all connections, is $d_A F(A) = 0$. The *Yang-Mills equations* are $d_A^* F(A) = 0$.

A connection A on P is said to be *self-dual* if $F(A) = *F(A)$. In this case $d_A^* F(A) = *d_A * F(A) = *d_A F(A) = 0$ by the Bianchi identity, so a self-dual connection automatically satisfies the Yang-Mills equations.

The Yang-Mills equations describe the critical points for the Yang-Mills functional (or action).

$$\|F(A)\|^2_{L^2} = \int_X |F(A)|^2 d\mu .$$

The self-dual connections give the absolute minimum for compact X which, if $G = SU(2)$, may be expressed via the Chern-Weil theorem as $-8\pi^2 c_2(P)$ where $c_2(P)$ is the 2nd Chern class of the associated rank 2 vector bundle.

The Yang-Mills functional and Yang-Mills equations are invariant under (i) conformal changes of the metric on X (ii) automorphisms of the principal bundle P ("*gauge transformations*").

(2.2) The initial mathematical development of the study of self-dual connections, motivated by the interest of mathematical physicists, concentrated on the case $X = S^4$ and an explicit description of all solutions was possible [2] using the twistor approach of R. Penrose and R.S. Ward [6] which converted the problem into one of holomorphic bundles on $\mathbb{C}P^3$.

More recently the self-duality equations have been studied on more general 4-manifolds. There are three major lines of thought which have spurred this progress :

(2.3) If X is a Kähler manifold, the space of anti-self-dual 2-forms $\Omega^2_- = \Omega^{1,1}_0$, the space of primitive 2-forms of type (1,1). A vector bundle with an anti-self-dual connection is then automatically endowed with a holomorphic structure (see [3]) and is moreover *stable* in the sense of Mumford and Takemoto (see [8], [11]). Converse results have been conjectured and in some cases proved ([13], [8]).

(2.4) The analysis of self-dual connections has been pushed forward by the fundamental results of K.K. Uhlenbeck ([20], [21]). Amongst these is the following *removable singularity theorem* : 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 ; then there is a bundle automorphism $g : B^4\setminus\{0\} \longrightarrow SU(2)$ such that $g(A)$ extends smoothly over B^4.

(2.5) The existence of self-dual connections is assured under very general circumstances by a theorem of C.H. Taubes [19] : Let X be a compact, oriented, riemannian 4-manifold with positive definite intersection form Q, and let P be a principal $SU(2)$ bundle over X with $c_2(P) \leq 0$. Then P admits an irreducible self-dual connection. Taubes'construction makes use of an implicit function theorem which involves L^p estimates on curvature. It should be noted that anti-self-dual harmonic 2-forms may certainly obstruct the existence of self-dual connections, as can be seen by considering $\mathbb{C}P^2$ with opposite orientation. There are no stable rank 2 bundles on $\mathbb{C}P^2$ with $c_2(P) = 1$ [16] and hence by (2.3) no anti-self-dual connections. Taubes'hypotheses and result are the starting point for Donaldson's theorem.

(2.6) As an example of a self-dual connection, take $X = \mathbb{R}^4$ and $G = SU(2)$. Then in terms of a quaternionic coordinate $x \in \mathbb{H} \cong \mathbb{R}^4$ and using the isomorphism $SU(2) \cong \mathrm{Im}\,\mathbb{H}$ the 1-*instanton* [7] solution of the self-duality equations is given by

$$A_\lambda = \mathrm{Im}\left(\frac{x d\bar{x}}{\lambda^2 + |x|^2}\right) \quad \text{with} \quad F(A_\lambda) = \frac{\lambda^2 dx \wedge d\bar{x}}{(\lambda^2 + |x|^2)^2}$$

and action $8\pi^2$.

(2.7) PROPOSITION.— *Let* A *be a self-dual* $SU(2)$ *connection on* $\mathbb{R}^4$ *with action* $8\pi^2$. *Then up to a gauge transformation and a translation of* $\mathbb{R}^4$, A *is equal to* A_λ *for some* $\lambda \in \mathbb{R}$.

Proof.— By conformal invariance and stereographic projection A is defined on $S^4\setminus\{x\}$, and by the removable singularity theorem is defined on a bundle $P \longrightarrow S^4$. Now use [3] § 9 or [2] or [6].

§ 3. The moduli space

(3.1) The cobordism in the proof of (1.1) is modelled on a moduli space of self-dual connections whose general structure is described next.

Let X be as in Theorem (1.1), and given a riemannian metric. Let P be a principal $SU(2)$ bundle over X with $c_2(P) = -1$. Using the covariant derivative of a fixed smooth connection A_0 on P, one may define Sobolev spaces $L^p_q(V)$ of sections of any associated vector bundle V.

Let $\mathcal{A}$ denote the affine space of connections on P differing from A_0 by an element of $L^2_3(\Omega^1(g))$, and let $\mathcal{G}$ denote the group of L^2_4 sections of $PX_{Ad}G$ ($\subset \mathrm{End}\ V$ for some faithful representation). Then $\mathcal{G}$ is a Banach Lie group of gauge transformations acting smoothly on $\mathcal{A}$ by $g(A) = A - (d_A g)g^{-1}$. Let $\mathcal{B}$ denote the quotient space with projection $p : \mathcal{A} \to \mathcal{B}$, and $p(A) = [A]$.

(3.2) Recall that a connection on P is *reducible* if its holonomy group is a proper subgroup of SU(2) . Since X is simply-connected and P is topologically non-trivial, the only possible reduction is to $U(1) \subset SU(2)$. Let $\Gamma_A \subset \mathcal{G}$ denote the subgroup of covariant constant sections with respect to the connection A . Then A is reducible iff $\Gamma_A \cong U(1)$. The equivalence classes of irreducible connections form an open subset $\mathcal{B}^* \subset \mathcal{B}$.

(3.3) PROPOSITION.— (i) *$\mathcal{B}$ is a Hausdorff space in the quotient topology.*

(ii) *$\mathcal{B}^*$ is a Banach manifold with charts constructed from the slices* $T_{A,\varepsilon} = \{A + a \mid d_A^* a = 0 , \|a\|_{L^2_3} < \varepsilon\}$ *of the action of $\mathcal{G}$.*

(iii) $p : p^{-1}(\mathcal{B}^*) \to \mathcal{B}^*$ *is a principal $\mathcal{G}/\pm 1$ bundle with a connection defined by the slices.*

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

Proof.— Standard methods (see [3], [12], [14]) using Banach space inverse and implicit function theorems.

(3.4) Let $\mathcal{M} \subset \mathcal{B}$ denote the subspace of equivalence classes of *self-dual* connections on P . $\mathcal{M}$ is the *moduli space*. If $A \in \mathcal{A}$ is reduced to a connection on a principal U(1) bundle $Q \subset P$, then (since $\pi_1(X) = 0$) its equivalence class is determined by its curvature $F(A) \in \Omega^2$. If A is self-dual, F(A) is a self-dual closed 2-form, hence harmonic. By Hodge theory F(A) is determined by its cohomology class $2\pi i c_1(Q)$. The reduction to U(1) is well-defined modulo the Weyl group, so $[A] \in \mathcal{M}$ is determined by $\pm c_1(Q)$. Since $c_2(P) = -c_1(Q)^2 = -1$ there are n *distinguished points in* $\mathcal{M}$ representing the reducible self-dual connections, where $2n = \# \{u \in H^2(X,\mathbb{Z}) \mid Q(u,u) = 1\}$. From (2.5) there are also irreducible connections.

(3.5) If A is a self-dual connection on P , then there exists an elliptic complex [3]

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

where d_A^- is the projection of d_A onto the anti-self-dual 2-forms. Let H_A^p ($0 \le p \le 2$) denote the associated harmonic spaces, then by the Atiyah-Singer index theorem (see [3])

$$-\sum_{p=0}^{2}(-1)^p \dim H_A^p = 8|c_2(P)| - \frac{3}{2}(\chi(X) - \mathrm{Sign}(X)) = 5 .$$

(3.6) PROPOSITION.— *Let* A *be a self-dual connection on* P .

Then there exists a neighbourhood U *of* $0 \in H^1_A$ *and a smooth map* $\phi : U \longrightarrow H^2_A$ *such that :*

(i) *if* A *is irreducible, a neighbourhood of* $[A] \in \mathcal{M}$ *is diffeomorphic to* $\phi^{-1}(0) \subseteq H^1_A$.

(ii) *if* A *is reducible, a neighbourhood of* $[A] \in \mathcal{M}$ *is diffeomorphic to* $\phi^{-1}(0)/\Gamma_A$.

Proof.— The connection $A+a$ is self-dual iff

$$\Phi(A+a) = F_-(A+a) = d_A^- a + \frac{1}{2}[a,a] = 0 \in L^2_2(\Omega^2_-(g)) .$$

Restricted to a slice $T_{A,\varepsilon}$ the derivative $D\Phi_A$ of Φ at A is the Fredholm operator $d_A^- : \text{Ker } d_A^*(\subseteq L^2_3(\Omega^1(g))) \longrightarrow L^2_2(\Omega^2_-(g))$, and so Φ is a Fredholm map ([1], [18]). After a local diffeomorphism Φ may be represented as $\Phi(x) = (D\Phi_A)x + \phi(x)$. The argument is analogous to the methods applied to moduli of complex structures [10].

(3.7) As a consequence of (3.5) and (3.6), if A is irreducible and $H^2_A = 0$, then $\mathcal{M}$ is a smooth 5-manifold in a neighbourhood of [A] . A particular case when this holds for all irreducible A is when the underlying metric on X is self-dual with positive scalar curvature (see [3]). Note that if A is reducible, Γ_A acts on H^1_A by complex multiplication ($b_1(X) = 0$) so that if $H^2_A = 0$, $H^1_A/\Gamma_A \cong \mathbb{C}^3/S^1$ from the index theorem and $\dim H^0_A = \dim \Gamma_A = 1$.

§ 4. A key result

(4.1) An important tool in understanding the global structure of the moduli space is the following : (see also [15]).

(4.2) PROPOSITION.— *Let* $\tilde{A}_i \in \mathcal{A}$ *be a sequence of self-dual connections on* P . *Then there is a subsequence such that either :*

(i) *each* $\tilde{A}_i$ *is gauge equivalent to* $A_i \in \mathcal{A}$ *converging in* C^∞ *to a self-dual connection* A_∞ *on* P , *and hence* $[\tilde{A}_i] \longrightarrow [A_\infty] \in \mathcal{M}$.

or

(ii) *there is a point* $x \in X$ *and trivializations* ρ_i *of* $P|_K$ *on the complement* K *of any geodesic ball about* x *such that* $\rho_i^* \tilde{A}_i \longrightarrow \vartheta$ *(the trivial flat connection) in* $C^\infty(K)$.

Proof.— The proof uses two lemmas :

(4.3) *Lemma.— Given* $L, C > 0$ *let* $\{f_i\}$ *be a sequence of integrable functions on* X *with* $f_i \geq 0$ *and* $\int_X f_i d\mu \leq L$. *Then there exists a subsequence, a finite set* $\{x_1,\ldots,x_\ell\} \subset X$ *and a countable collection* $\{B_\alpha\}$ *of geodesic balls in* X *such that the half-sized balls cover* $X\setminus\{x_1,\ldots,x_\ell\}$ *and for each* α , $\limsup \int_{B_\alpha} f_i d\mu < C$.

Proof.— Elementary : the x_i's are characterized by the property that each lies in no ball with $\limsup \int_B f_i d\mu \le \frac{1}{2} C$.

(4.4) *Lemma.— Let* h_i *be a sequence of metrics on* B^4 *, sufficiently close to the Euclidean metric, and converging in* $C^\infty(\bar{B}^4)$ *to* h_∞ *. Let* $\tilde{A}_i$ *be a sequence of connections on the trivial bundle over* B^4 *with* $\tilde{A}_i$ *self-dual with respect to* h_i *. Then there is a constant* C *(independent of* h_i *and* $\tilde{A}_i$*) such that if* $\int_{B^4} |F(\tilde{A}_i)|^2 d\mu \le C$ *, there is a subsequence such that* A_i *(gauge equivalent to* $\tilde{A}_i$*) converges in* $C^\infty(\frac{1}{2}\bar{B}^4)$ *to* A_∞ *, a connection which is self-dual with respect to* h_∞ *.*

Proof.— Consequence of ([21] Theorem (1.3)).

(4.5) To obtain (4.2) first consider a geodesic coordinate system χ on a geodesic ball $B \subset X$ of radius r . Thus χ defines a diffeomorphism $\chi : B_r^4 \longrightarrow B$ from the euclidean ball of radius r to B . Pulling back the metric h , and putting it on the Euclidean unit ball by dilation gives a metric

$$h_r = \chi^* h(rx) = r^2(\delta_{ij} + r^2 0(|y|^2))dy_i dy_j \ .$$

Choose r small enough that the metric $r^{-2}h_r$ on B^4 satisfies the condition for (4.4). By conformal invariance each $\tilde{A}_i$ is self-dual with respect to h_r .

Now in Lemma (4.3) take the constant C from (4.4), $f_i = |F(\tilde{A}_i)|^2$ and $L = 8\pi^2$. Thus from (4.4) on each ball $\frac{1}{2}\bar{B}_\alpha$ some subsequence converges (after gauge transformations) to $A_\infty(\alpha)$. By a diagonal argument the convergence may be achieved simultaneously for all α .

The gauge transformations introduced in the above process give rise to connection matrices $A_i(\alpha) \longrightarrow A_\infty(\alpha)$ in $C^\infty(\frac{1}{2}B_\alpha)$ and transition functions $g_i(\alpha,\beta) : \frac{1}{2}B_\alpha \cap \frac{1}{2}B_\beta \longrightarrow SU(2)$ satisfying :

$$(4.6) \qquad A_i(\alpha) = -dg_i(\alpha,\beta)g_i(\alpha,\beta)^{-1} + g_i(\alpha,\beta)A_i(\beta)g_i(\alpha,\beta)^{-1} \ .$$

The compactness of SU(2) gives a uniform bound to dg_i in (4.6) and so one can find a uniformly convergent subsequence. Repeatedly applying (4.6) gives convergence in C^∞ , and using a diagonal argument one obtains a subsequence

$$(A_i(\alpha), g_i(\alpha,\beta)) \longrightarrow (A_\infty(\alpha), g_\infty(\alpha,\beta))$$

for all (α,β) simultaneously. This represents a self-dual connection on a bundle Q over $X\backslash\{x_1,\ldots,x_\ell\}$. Furthermore, if $K \subset X\backslash\{x_1,\ldots,x_\ell\}$ is compact then by induction on the number of balls $\frac{1}{2}B_\alpha$ covering K (see [21] Sect. 3) one obtains isomorphisms $\rho_i : Q|_K \longrightarrow P|_K$ such that $\rho_i^* : A_i \longrightarrow A_\infty$ in $C^\infty(K)$.

(4.7) Let $B_j^!$ be a small punctured ball centred on x_j $(1 \le j \le \ell)$. Since $\int_{B_j^!} |F(\tilde{A}_i)|^2 d\mu \le 8\pi^2$, by Fatou's lemma $\int_{B_j^!} |F(A_\infty)|^2 d\mu \le 8\pi^2$. Hence by the removable singularity theorem (2.4) the connection A_∞ and bundle Q extend over X . By the definition of x_j , $\lim \int_{B_j^!} |F(\tilde{A}_i)|^2 d\mu > \frac{1}{2} C$ for all balls B_j hence for a sufficiently small ball

$$\int_{B'_j} |F(A_\infty)|^2 d\mu < \lim \int_{B'_j} |F(\tilde{A}_i)|^2 d\mu \ .$$

(4.8) On the other hand, since all connections are self-dual these integrands are Chern forms. They may therefore be evaluated mod. $8\pi^2\mathbb{Z}$ by boundary integrals (Chern-Simons invariants). Hence by uniform convergence on the boundary ∂B_j ,

$$\int_{B'_j} |F(A_\infty)|^2 d\mu = \lim \int_{B'_j} |F(\tilde{A}_i)|^2 d\mu \quad \text{mod. } 8\pi^2\mathbb{Z} \ .$$

(4.9) However, since $\int_{B'_j} |F(A_\infty)|^2 d\mu \geq 0$ and $\int_{B'_j} |F(\tilde{A}_i)|^2 d\mu \leq 8\pi^2$ the only possibilities from (4.7) and (4.8) are :

(i) $\ell = 0$ or

(ii) $\lim \int_{B'_j} |F(\tilde{A}_i)|^2 d\mu = 8\pi^2$ and $\int_X |F(A_\infty)|^2 d\mu < 8\pi^2$ and hence Q is trivial and A_∞ flat. Thus Proposition (4.2) follows.

(4.10) The proposition shows that a self-dual connection on P can only degenerate by having its curvature concentrate in the neighbourhood of a point. An example is the instanton A_λ in (2.6) as $\lambda \to 0$.

§ 5. The boundary of $\mathcal{M}$

(5.1) Let $\beta : \mathbb{R} \to \mathbb{R}$ be a bump function approximating and dominated by $\chi_{[-1,1]}$ and set $R_A(x,s) = \int_X \beta(d(x,y)/s)|F(A)|^2 d\mu_y$, where $d(x,y)$ is the geodesic distance in X . Then define

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

where K is chosen so that $\lambda(A_1) = 1$ for the instanton A_1 . Donaldson introduces this convenient but ad hoc function as a measure of the concentration of curvature : if β is replaced by $\chi_{[-1,1]}$ then $\lambda(A)$ becomes the radius of the smallest ball containing half the action. In any case a ball of radius $\lambda(A)$ contains more than half the action and hence any sequence $[A_i] \in \mathcal{M}$ without convergent subsequences has $\lambda(A_i) \to 0$ from (4.2). It is thus a measure of the distance from the boundary.

(5.3) PROPOSITION.— *There exists* $\lambda_0 > 0$ *such that if* A *is a self-dual connection on* P *with* $\lambda(A) < \lambda_0$ *, then the minimum in* (5.2) *is attained at a unique point* $x(A) \in X$.

Proof.— Take a small geodesic ball of radius r centred on a minimum x for A , and pull back the metric and connection as in (4.5) to the Euclidean ball of radius $r/\lambda(A)$. For each sequence of connections with $\lambda(A_i) \to 0$, the pulled-back connections $\hat{A}_i$ satisfy $\lambda(\hat{A}_i) = 1$ by construction and applying (4.4) and (4.2) there is a subsequence converging to a self-dual connection on $\mathbb{R}^4$. From the classification (2.7) and normalization this is the instanton A_1 . Since $\lambda(\hat{A}_i) = 1$, from

(4.2) every subsequence converges and since the limit is unique, $\hat{A}_i \to A_1$ as $\lambda(A_i) \to 0$. Now the function R_{A_1} has a unique non-degenerate minimum so for sufficiently small $\lambda(A)$, so will $R_{\hat{A}}$. Any two minima for A must however be separated by a distance of at most $2\lambda(A)$, since the ball of radius $\lambda(A)$ about each contains more than half the action, thus a unique minimum for $R_{\hat{A}}$ implies a unique one for R_A .

Note how the connectedness of the moduli space for $\mathbb{R}^4$ is essential for this argument.

(5.4) Let $\mathcal{M}_{\lambda_0} = \{[A] \in \mathcal{M} \mid \lambda(A) < \lambda_0\}$, and define $p : \mathcal{M}_{\lambda_0} \longrightarrow X \times (0,\lambda_0)$ by $p(A) = (x(A),\lambda(A))$.

(5.5) PROPOSITION.— (i) $\mathcal{M}\backslash\mathcal{M}_{\lambda_0}$ *is compact.*

(ii) $\mathcal{M}_{\lambda_0}$ *is a smooth manifold.*

(iii) p *is a smooth covering map.*

Proof.— (i) Immediate from Proposition (4.2).

(ii) As $\lambda(A) \to 0$, $[A] \to \vartheta$ in $C^\infty(X\backslash B(x(A),r))$ from (4.2). Then using an argument of Taubes [19], $H^2_A = 0$. The result follows from (3.6).

(iii) p is smooth because the minimum of R_A is non-degenerate, and proper by (4.2). Thus one only needs to check that the derivative of p is an isomorphism. Taubes' implicit function theorem provides an inverse.

(5.6) PROPOSITION.— p *is a diffeomorphism.*

Proof.— This is the most technical part of Donaldson's proof, and involves delicate curvature estimates. The idea is to show that any two self-dual connections A , B with $x(A) = x(B)$ and $\lambda(A) = \lambda(B)$ sufficiently small may be joined by a short path in $\mathcal{M}$ (see [8]).

§ 6. Perturbation of $\mathcal{M}$

(6.1) If $H^2_A = 0$ for all self-dual connections then $\mathcal{M}$ is a smooth manifold except at the $n(Q)$ points corresponding to the reducible connections. This may not be true in general and there may be a subset $K \subset \mathcal{M}$ (compact from (5.5)) for which $H^2_A \neq 0$. A perturbation of $\mathcal{M}$ is then necessary to obtain a manifold.

(6.2) Perturbation around the reducible connections is dealt with in a straightforward manner : the finite-dimensional map $\phi(x)$ in the decomposition $\Phi(x) = (D\Phi_A)x + \phi(x)$ is modified by a nearby map with surjective derivative. Then, as in (3.6) a neighbourhood of $[A]$ is diffeomorphic to $\mathbb{C}^3/S^1$ - a cone on $\mathbb{C}P^2$. One may assume, then, that $K \subset \mathcal{M} \cap \mathcal{B}^*$.

(6.3) The group $\mathcal{G}/\pm 1$ acts on the Banach spaces $L^2_3(\Omega^2_-(\mathfrak{g}))$ and $L^2_2(\Omega^2_-(\mathfrak{g}))$ and associated to the principal $\mathcal{G}/\pm 1$ bundle $p^{-1}(\mathcal{B}^*)$ over $\mathcal{B}^*$ one obtains vector bundles $\mathcal{E}^3 \subset \mathcal{E}^2$ with norms and connections. There is a canonical section $\Phi = F_-(A)$ of $\mathcal{E}^2$ and one seeks perturbations $\sigma \in C^\infty(\mathcal{B}^*,\mathcal{E}^3)$, such that $\Phi + \sigma$ vanishes non-degenerately.

(6.4) PROPOSITION.— *There exists* $\sigma \in C^\infty(\mathcal{B}^*,\mathcal{E}^3)$ *, supported in a neighbourhood of* K *, such that* $(\Phi+\sigma)^{-1}(0)$ *is a smooth* 5*-manifold.*

Proof.— Covering K with a finite number of slices $T_{A,\varepsilon}$ and shrinking, take open sets U_1 , U_2 with $K \subset U_1$ and $\overline{U}_1 \subset U_2$, and let σ be a bounded section of $\mathcal{E}^3$ supported in U_2 . Then $\hat{K} = \{[A] \in \overline{U}_1 \mid \|(\Phi+\sigma)(A)\|_{L^2_3} \le R\}$ is compact. This follows from the fact that U_2 is covered by a finite number of slices and on each one $\Phi(A) = d_A^- a + \frac{1}{2}[a,a] + \sigma(A)$ with $d_A^* a = 0$ and $\|a\|_{L^2_3} < \varepsilon$. Thus L^2_3 bounds on $\sigma(A)$, a and $(\Phi+\sigma)(A)$ give an L^2_3 bound on $(d_A^- + d_A^*)a$ and so by ellipticity an L^2_4 bound on a . Since $L^2_4 \subset L^2_3$ is compact the statement follows. Thus if $\Phi + \sigma$ vanishes non-degenerately in $\overline{U}_1$, so do nearby sections $\Phi + \sigma'$ in the topology of uniform convergence of σ and its derivative on compact sets.

The space of such non-degenerate perturbations is also dense : at each point take a slice on which there is a decomposition $\Phi + \sigma = L + \phi$ where L is linear and ϕ finite dimensional. By compactness, take a finite subcovering and modify $\Phi + \sigma$ by substracting a regular value of ϕ , extended by a bump function. By Sard's theorem such perturbations can be made arbitrarily close in L^2_3 norm to $\Phi + \sigma$.

The section Φ itself vanishes non-degenerately outside $\overline{U}_1$. By the density argument choose a perturbation σ sufficiently small that $\Phi + \sigma$ (by the openness argument on $U_2\backslash\overline{U}_1$) vanishes non-degenerately on $\overline{U_2\backslash\overline{U}_1}$. Then $\Phi + \sigma$ is non-degenerate everywhere. Let $\mathcal{M}^\sigma = (\Phi+\sigma)^{-1}(0)$, a 5-manifold with n quotient singularities $\mathbb{C}^3/S^1$ and boundary X .

§ 7. Orientability of $\mathcal{M}^\sigma$

(7.1) On the manifold $\mathcal{M}^\sigma \cap \mathcal{B}^*$ one must consider the Stiefel-Whitney class $w_1(\mathrm{Ker}\,\nabla(\Phi+\sigma))$. The singular points can be avoided by using the gauge transformations $\mathcal{G}_0 \subset \mathcal{G}$ which are the identity at a fixed point $x_0 \in X$. These then act freely on $\mathcal{A}$ to give quotient $\hat{\mathcal{B}} \xrightarrow{\pi} \mathcal{B}$. Over $\mathcal{B}^*$, π gives a principal $SO(3)$ bundle, so $T\mathcal{M}^\sigma \cap \mathcal{B}^*$ is orientable iff its pull back to $\pi^{-1}(\mathcal{M}^\sigma \cap \mathcal{B}^*)$ is.

(7.2) The vector bundle $\mathrm{Ker}\,\nabla(\Phi+\sigma)$ restricted to any compact subset $Y \subset \pi^{-1}(\mathcal{M}^\sigma \cap \mathcal{B}^*)$ defines an element of $KO(Y)$. This is the *index class* [5] of the family of Fredholm operators $d_A^* + d_A^- + (\nabla\sigma)A$, which by considering the deformation $d_A^* + d_A^- + t(\nabla\sigma)A$, $0 \le t \le 1$, is independent of σ . Since w_1 factors through KO , the orientability can be decided by considering $\mathrm{ind}(d^*_A + d^-_A) \in KO(Y)$ where Y is a loop. Since this is now defined for all equivalence classes of connections, the loop may be deformed in $\hat{\mathcal{B}}$.

(7.3) If $SU(2)$ is embedded in $SU(3)$ in the standard way, the Lie algebra bundle $\tilde{g}$ of the associated $SU(3)$ connection $\tilde{A}$ splits as $\tilde{g} = g \oplus \mathbb{R} \oplus V$ where V is a *complex* rank 2 bundle and $\mathbb{R}$ a *trivial* bundle, all preserved by the connection. Thus $w_1(\mathrm{ind}(d^*_{\tilde{A}} + d^-_{\tilde{A}})) = w_1(\mathrm{ind}(d^*_A + d^-_A))$ and so the loop may be deformed in the space $\hat{\mathcal{B}}_3$ of equivalence classes of $SU(3)$ connections.

(7.4) PROPOSITION.— $\pi_1(\hat{\mathcal{B}}_3) = 0$.
Proof.— Since the group $\mathcal{G}_0$ of $SU(3)$ gauge transformations preserving x_0 acts freely, $\pi_1(\hat{\mathcal{B}}_3) \cong \pi_0(\mathcal{G}_0)$. The principal bundle P is trivial on the complement of a point and in particular on the 2-skeleton of X . Since $\pi_2(SU(3)) = 0$ any element of $\mathcal{G}_0$ can be deformed to one which is the identity on the 2-skeleton. Collapsing the 2-skeleton of X gives a sphere S^4 . The homotopy type of $\mathcal{G}_0$ on S^4 is independent of $c_2(P)$ (see [4]) , so the question reduces to the trivial bundle. But $\pi_4(SU(3)) = 0$, so $\mathcal{G}_0$ is connected.

Thus $\mathcal{M}^\sigma \cap \mathcal{B}^*$ is an oriented 5-manifold which, putting in the boundaries of the quotient singularities provides the cobordism of Theorem (1.1).

§ 8. Examples

(8.1) Let $X = S^4$, with the canonical metric. Then any self-dual connection on P is gauge equivalent to f^*A where $f : S^4 \longrightarrow \mathbb{H}P^1$ is a *conformal* map and A is the canonical connection on the quaternionic Hopf bundle. Since isometries of $\mathbb{H}P^1$ preserve A , the moduli space is $SO(5,1)/SO(5) \cong$ hyperbolic 5-space. This is the ball B^5 with boundary S^4 . There are many ways of proving this ([2], [3], [6]).

(8.2) Let $X = \mathbb{C}P^2$ with its canonical metric. In the non-compact component of $\mathcal{M}$, any connection is gauge equivalent to f^*A where $f : \mathbb{C}P^2 \longrightarrow \mathbb{H}P^2$ is equivalent under the action of $SU(3)$ on $\mathbb{C}P^2$ to a map of the form

$$(z,w) \longmapsto \left(z \, , \frac{ajz + w}{\sqrt{1 - a^2}}\right) \qquad a \in [0,1)$$

in affine coordinates. When $a = 0$ this is the standard embedding $\mathbb{C}P^2 \subset \mathbb{H}P^2$ and gives the reducible connection. The moduli space is a cone on $\mathbb{C}P^2$ where a is essentially the distance from the vertex. This was proved by Donaldson (unpublished) using the algebraic geometry of the flag manifold F_3 , and the Penrose/Ward approach.

§ 9. References

[1] ABRAHAM R., ROBBIN J. - Transversal mappings and flows, Benjamin, New York (1967).

[2] ATIYAH M.F., HITCHIN N.J., DRINFELD V.G., MANIN, YU I. - *Construction of instantons*, Phys. Lett. 65A(1978), 185.

[3] ATIYAH M.F., HITCHIN N.J., SINGER I.M. - *Self-duality in four-dimensional Riemannian geometry*, Proc. R. Soc. (Lond.) A362(1978), 425-461.

[4] ATIYAH M.F., JONES J.D.S. - *Topological aspects of Yang-Mills theory*, Commun. Math. Phys. 61(1978), 97-118.

[5] ATIYAH M.F., SINGER I.M. - *The index of elliptic operators* IV , Ann. of Math. 93(1971), 119-138.

[6] ATIYAH M.F., WARD R.S. - *Instantons and algebraic geometry*, Commun. Math. Phys. 55(1977), 117-124.

[7] BELAVIN A.A., POLYAKOV A.M., SCHWARTZ A.S., TYUPKIN Y.S. - *Pseudoparticle solutions of the Yang-Mills equations*, Phys. Lett. 59B(1975), 85.

[8] DONALDSON S.K. - D. Phil. thesis (Oxford) (1982).

[9] FREEDMAN M.H. - *The topology of four-dimensional manifolds*, J. Differential Geometry 17(1982), 357-453.

[10] KURANISHI M. - *A new proof for the existence of locally complete families of complex structures*, Proc. Conf. on Complex Analysis, Minneapolis, Springer-Verlag, New York, (1964), 142-156.

[11] LÜBKE M. - *Stability of Einstein-Hermitian Vector bundles*, Preprint (Bayreuth).

[12] MITTER P.K., VIALLET C.M. - *On the bundle of connections and the gauge arbit manifold in Yang-Mills theory*, Commun. Math. Phys. 79(1981), 457-472.

[13] NARASIMHAN M.S., SESHADRI C.S. - *Stable and unitary vector bundles on a compact Riemann surface*, Ann. of Math. 82(1965), 540-567.

[14] PARKER T.H. - *Gauge theories on four-dimensional riemannian manifolds*, Commun. Math. Phys. 85(1982), 563-602.

[15] SEDLACEK S. - *A direct method for minimizing the Yang-Mills functional over 4-manifolds*, Commun. Math. Phys. 86(1982), 515-527.

[16] SCHWARZENBERGER R.L.E. - *Vector bundles on the projective plane*, Proc. Lond. Math. Soc. 11(1961), 623-640.

[17] SERRE J.-P. - A course in arithmetic, Springer-Verlag, New York, (1973).

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

[19] TAUBES C.H. - *Self-dual Yang-Mills connections on non-self-dual 4-manifolds*, J. Differential Geometry 17(1982), 139-170.

[20] UHLENBECK K.K. - *Removable singularities in Yang-Mills fields*, Commun. Math. Phys. 83(1982), 11-30.

[21] UHLENBECK K.K. - *Connections with* L^p *bounds on curvature*, Commun. Math. Phys. 83(1982), 31-42.

Nigel J. HITCHIN

St. Catherine's College
GB - OXFORD OX1 3UJ