Proc. R. Soc. Lond. A. **362**, 425–461 (1978)
Printed in Great Britain

Self-duality in four-dimensional Riemannian geometry

BY M. F. ATIYAH,† F.R.S., N. J. HITCHIN†
AND I. M. SINGER‡

† *Mathematical Institute, University of Oxford,*
24–29 *St Giles', Oxford, U.K.*
‡ *Department of Mathematics, University of California,*
Berkeley, California 94720, *U.S.A.*

(*Received* 23 *December* 1977)

We present a self-contained account of the ideas of R. Penrose connecting four-dimensional Riemannian geometry with three-dimensional complex analysis. In particular we apply this to the self-dual Yang–Mills equations in Euclidean 4-space and compute the number of moduli for any compact gauge group. Results previously announced are treated with full detail and extended in a number of directions.

INTRODUCTION

This paper is essentially an amplification of our previous note (Atiyah, Hitchin & Singer 1977) concerning the deformation theory of self-dual solutions of the Yang–Mills equations in Euclidean 4-space. Besides providing full details for all the results announced in Atiyah *et al.* (1977), we broaden its scope by setting it in a natural differential-geometric context, and we refine the results in a number of directions. We also include an account of the relation between self-duality and holomorphic structures. The basic ideas here are those of R. Penrose, in his twistor programme, but there are some advantages in having a presentation of Penrose's ideas in the framework of real Riemannian geometry. The application of twistor theory to Yang–Mills fields is due to R. S. Ward and has been further developed by Atiyah & Ward (1977). Our presentation overlaps slightly with their work but has a more differential-geometric viewpoint. We do not pursue the complex analytic implications themselves.

We begin in § 1 with a review of four-dimensional Riemannian geometry, and the particular rôle of the duality $*$-operator. The special feature of four-dimensions is that the rotation group $SO(4)$ is not simple but is locally isomorphic to $SU(2) \times SU(2)$. For this reason the Riemann curvature tensor has an extra decomposition: the conformally invariant part W (the Weyl tensor) decomposes under $*$ as $W = W_+ \oplus W_-$. A self-dual metric (or conformal structure) is one for which $W_- = 0$, and such metrics will be our main concern. On the geometric side each of the $SU(2)$ factors of $SO(4)$ defines a spin bundle with fibre $\mathbb{C}^2$ and hence a projective spin bundle with fibre the complex projective line $P_1(\mathbb{C})$. This fibration will play an important rôle. In § 1 and elsewhere we employ freely the concepts and techniques of modern differential geometry and we give, as far as possible, a coordinate-free treatment.

In § 2 we consider the $*$-operator in the context of connections and curvature on fibre bundles. We introduce the notion of a self-dual connection, i.e. one whose curvature Ω satisfies $*\Omega = \Omega$, which for the physicists gives an absolute minimum of the Yang–Mills functional. The material here is by now standard except for the observation that self-duality of the spin bundles coincides with the Einstein property of the metric.

We shall show that the self-duality property of metric and connections can be interpreted as the integrability conditions for complex structures. In order to have a unified treatment of both cases which is also conformally invariant, we give in § 3 a general integrability result associated to first order differential equations. This is applied first in § 4 to the twistor equation to show that, if X is a self-dual 4-manifold, the projective bundle $P(V_-)$ of anti-self-dual spinors inherits the structure of a complex analytic 3-manifold. The conformal invariance of the twistor equation implies that the complex structure of $P(V_-)$ depends only on the conformal structure of X. The most noteworthy example arises when X is the 4-sphere in which case $P(V_-)$ can be identified with complex projective 3-space.

In § 5 we apply the integrability result of § 3 to prove that a hermitian vector bundle, over a complex manifold, with a connection whose curvature is of type (1,1) has a natural holomorphic structure. This is then combined with the results of § 4 to prove that a self-dual bundle over a self-dual 4-manifold lifts to give a holomorphic bundle over $P(V_-)$.

Section 6 is devoted to studying the moduli space of irreducible self-dual connections (over a self-dual manifold of positive scalar curvature). As outlined in Atiyah *et al.* (1977) we derive a general formula for the dimension of each component of the moduli space. We also prove that the moduli space is globally a (Hausdorff) manifold.

In § 7 we specialize to $SU(2)$-bundles over S^4 and give a differential-geometric treatment of the 't Hooft solutions, and we then proceed to consider other simple Lie groups G in § 8. We prove that if $k \geqslant k(G)$ then irreducible self-dual G-bundles over S^4 with Pontrjagin index k exist and we compute the dimension of the moduli space. The lowest value of $k(G)$ is found for all G. These results, which refine the information given in Atiyah *et al.* (1977), are very similar to recent results of Bernard, Christ, Guth & Weinberg (1977).

Finally in § 9 we give a differential-geometric proof that the moduli space for self-dual $SU(2)$-bundles over S^4 with $k = 1$ is the hyperbolic 5-space corresponding to the known solutions described in § 7. Another proof of this fact by using algebraic geometric methods is given in Hartshorne (1977).

1. Four-dimensional Riemannian geometry

Let X be an oriented Riemannian manifold of even dimension $2l$, and let Λ^p denote the bundle of exterior p-forms with $A^p = \Gamma(\Lambda^p)$ its space of smooth sections. The Hodge star operator $*: \Lambda^p \to \Lambda^{2l-p}$ is defined by

$$\alpha \wedge *\beta = (\alpha, \beta)\, \omega \in \Lambda^{2l}$$

where $\alpha, \beta \in \Lambda^p$, (α, β) is the induced inner product on p-forms and ω is the volume form.

Of particular interest is the star operator on forms in the middle dimension $p = l$, where $* : \Lambda^l \to \Lambda^l$ satisfies $*^2 = (-1)^l$. On l-forms $*$ is conformally invariant, for if we change the metric by multiplying by a scalar λ, the inner product on tangent vectors is multiplied by λ and on l-forms by λ^{-l}. On the other hand the volume form is multiplied by λ^l and $(\alpha, \beta)\,\omega = \alpha \wedge *\beta$ remains the same.

If $l = 1$, then $*^2 = -1$ and defines the complex structure on a Riemann surface. We are interested in the case $l = 2$, i.e. when X is a four-dimensional manifold. In this instance $*^2 = +1$ and the bundle Λ^2 splits into a direct sum,

$$\Lambda^2 = \Lambda^2_+ \oplus \Lambda^2_-,$$

where $\Lambda^2_\pm$ are the ± 1 eigenspaces of $*$. We call them the bundles of *self-dual* and *anti-self-dual* 2-forms respectively.

The 2-forms are important in Riemannian geometry because of their relation with the curvature tensor, and this decomposition has a profound influence on the underlying geometry of four dimensions.

The Riemann curvature tensor defines in general a self-adjoint transformation $\mathscr{R}: \Lambda^2 \to \Lambda^2$ given by

$$\mathscr{R}(e_i \wedge e_j) = \tfrac{1}{2} \sum_{k,l} R_{ijkl}\, e_k \wedge e_l,$$

where $\{e_i\}$ is a local orthonormal basis of 1-forms. In four dimensions, we can write $\mathscr{R}$ as a block matrix relative to the decomposition $\Lambda^2 = \Lambda^2_+ \oplus \Lambda^2_-$:

$$\mathscr{R} = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix}, \tag{1.1}$$

where $B \in \operatorname{Hom}(\Lambda^2_-, \Lambda^2_+)$ and $A \in \operatorname{End} \Lambda^2_+$, $C \in \operatorname{End} \Lambda^2_-$ are self-adjoint.

This representation of $\mathscr{R}$ gives us a complete decomposition of the curvature tensor into irreducible components (Singer & Thorpe 1969):

$$\mathscr{R} \to (\operatorname{tr} A, B, A - \tfrac{1}{3} \operatorname{tr} A, C - \tfrac{1}{3} \operatorname{tr} C),$$

where $\operatorname{tr} A = \operatorname{tr} C = \frac{1}{4}$ scalar curvature, B is the traceless Ricci tensor and the last two components, which we denote W_+ and W_-, together give the conformally invariant Weyl tensor, $W = W_+ + W_-$. Note that the metric is Einstein iff $B \equiv 0$, conformally flat iff $W \equiv 0$. Both of these special forms of metric occur in higher dimensions, but there is one specialization which is only valid in four dimensions and with which we shall be primarily concerned.

15-2

Definition. An oriented Riemannian 4-manifold is *self-dual* if its Weyl tensor $W = W_+$, i.e. if $W_- \equiv 0$.

Since the Weyl tensor and the star operator are conformal invariants, it is clear that this is a property of the underlying conformal structure, and the choice of orientation.

EXAMPLES

1. If X is conformally flat, then $W_+ = W_- = 0$ and X is clearly self-dual. Hence for example the 4-sphere S^4, $S^1 \times S^3$ and the 4-torus T^4 with natural metrics are all self-dual.

2. The complex projective plane $P_2(\mathbb{C})$ with its standard metric and orientation is self-dual.

3. Any 2-complex dimensional Kähler manifold with vanishing Ricci tensor is *anti-self-dual* with respect to its canonical orientation. The recent proof of the Calabi conjecture by Yau (1977) thus yields an anti-self-dual metric on any $K3$ surface.

There are topological restrictions on manifolds which carry a self-dual conformal structure. In particular the *signature* of X must be non-negative, for the first Pontrjagin class p_1 may be represented by the integral

$$\begin{aligned} p_1(X) &= \frac{1}{8\pi^2}\int_X \Sigma\mathscr{R}(e_i \wedge e_j) \wedge \mathscr{R}(e_i \wedge e_j) \\ &= \frac{1}{4\pi^2}\int_X (|W_+|^2 - |W_-|^2)\,\omega \\ &= \frac{1}{4\pi^2}\int_X |W_+|^2\omega \quad \text{if} \quad X \text{ is self-dual.} \end{aligned}$$

Since the signature τ is equal to $\frac{1}{3}p_1$ we have $\tau \geqslant 0$ with equality iff $W_+ = W_- = 0$, i.e. iff X is conformally flat.

In particular, since the 4-sphere S^4 has zero signature and up to diffeomorphism a unique conformally flat structure, we deduce that there is on S^4 a *unique* self-dual conformal structure.

The 2-forms can be identified, by using the metric, with skew-adjoint transformations of Λ^1, and then the decomposition $\Lambda^2 = \Lambda^2_+ \oplus \Lambda^2_-$ corresponds to the isomorphism of Lie algebras $so(4) \cong so(3) \oplus so(3)$, $\Lambda^2_\pm$ being bundles of three-dimensional Lie algebras. It is convenient to use this decomposition on the Lie group level (Spin $(4) \cong SU(2) \times SU(2)$) and introduce, at least locally, the two complex spinor bundles V_+ and V_-: the bundles of *self-dual* and *anti-self-dual* spinors. In the usual way, the complex endomorphism bundle of the total spin bundle $V = V_+ \oplus V_-$ is isomorphic to the complexified Clifford algebra bundle of Λ^1, which is isomorphic as a graded vector bundle to $\Lambda^*_c = \bigoplus_p \Lambda^p_c$, where $\Lambda^p_c = \Lambda^p \otimes \mathbb{C}$.

(The Clifford algebra is the algebra generated by Λ^1 subject to the relations

$$e_i \,.\, e_j + e_j \,.\, e_i = -2\delta_{ij},$$

where $\{e_i\}$ is a local orthonormal basis.)

Under this isomorphism, p-forms (for $p = 1, 2$) act on spinors in the following way:

$$\Lambda^1_c \cong \mathrm{Hom}(V_+, V_-) \cong \mathrm{Hom}(V_-, V_+).$$

$\Lambda^2_{+c} \subset \mathrm{End}\, V_+$ consists of the traceless endomorphisms of V_+, and the real bundle Λ^2_+ the traceless skew hermitian endomorphisms. Similarly $\Lambda^2_{-c} \subset \mathrm{End}\, V_-$, and self-dual forms act trivially on anti-self-dual spinors. Since $V_+ \cong V_+^*$, symplectically, we also have $\Lambda^2_{+c} \cong S^2V_+$. (Here S^mV_+ denotes the mth symmetric power bundle of the two-dimensional complex bundle V_+.)

We can likewise express the decomposition of the curvature tensor (1.1) in spinor form. The bundle of self-dual Weyl tensors is identified with the bundle of self-adjoint traceless endomorphisms of Λ^2_+. The bundle of all endomorphisms of Λ^2_+ is $\Lambda^2_+ \otimes \Lambda^2_+ \cong S^2V_+ \otimes S^2V_+ \cong S^4V_+ \oplus S^2V_+ \oplus S^0V_+$. S^2V_+ is the bundle of skew-adjoint transformations, the one-dimensional bundle S^0V_+ consists of the scalar transformations, hence the bundle of self-dual Weyl tensors is S^4V_+. Similarly the bundle of anti-self-dual Weyl tensors is S^4V_-, and of traceless Ricci tensors $S^2V_+ \otimes S^2V_-$.

The real bundle $\Lambda^1 \subset \Lambda^1_c \cong \mathrm{Hom}(V_+, V_-)$ deserves some consideration in connection with the almost complex structures on X. If we fix a non-zero spinor $\phi \in (V_+)_x$ at a point x, then this gives a real isomorphism $\Lambda^1_x \cong (V_-)_x$ defined by Clifford multiplication $\alpha \mapsto \alpha . \phi$ and so identifies Λ^1_x with a complex vector space and furnishes the tangent space of X at x with a complex structure compatible with the metric and orientation. Clearly multiplying ϕ by a scalar $\lambda \in \mathbb{C}^*$ defines the same complex structure, so the projective space $P(V_+)_x$ parametrizes a set of compatible complex structures.

The subgroup of $\mathrm{Spin}(4) \cong SU(2) \times SU(2)$ which leaves fixed ϕ up to a scalar multiple is $S^1 \times SU(2)$, the double covering of $U(2) \subset SO(4)$. Hence the projective space $P(V_+)_x$ is naturally isomorphic to $SO(4)/U(2)$, the space of *all* complex structures compatible with the metric and orientation.

There is a dual way of looking at this, where we take not the Clifford multiplication map $\Lambda^1 \otimes V_+ \to V_-$, but its adjoint

$$V_- \to V_+ \otimes \Lambda^1,$$
$$\psi \mapsto \sum_i e_i . \psi \otimes e_i,$$

where $g = \Sigma e_i \otimes e_i$ is the metric tensor. Now if we are given $\phi \in V_+$, we get a map of V_- into Λ^1_c:

$$\psi \mapsto \sum_i (e_i . \psi, \phi)\, e_i \tag{1.2}$$

by means of the symplectic form on V_+, and the image of V_- in Λ^1_c is the subspace $\Lambda^{1,0}$ of $(1, 0)$ forms which equivalently defines the complex structure.

The symplectic and hermitian structure on V_+ defines an isomorphism $V_+ \cong V_+^* \cong \overline{V}_+$ which is antilinear. Under this isomorphism, $\phi \otimes \overline{\phi}$ defines an element of $(S^2V_+)_x$ which is invariant under $U(2)$: the hermitian form in $(\Lambda^2_+)_x$. The space $(\Lambda^2_-)_x$ then consists of $(1, 1)$ forms orthogonal to this form.

Using the metric, we see that this 2-form defines a skew-adjoint transformation I of the tangent space T_x such that $I^2 = -1$ and $\det I = +1$. This is the usual definition of an almost complex structure.

It is clear now that we may consider two general classes of objects – self-dual and anti-self-dual – on a 4-manifold. The natural category with which we shall be concerned is that of self-dual objects on self-dual spaces. The objects of primary importance, which come to us from physics, are the self-dual Yang–Mills fields or *self-dual connections* which we consider next.

2. Self-dual gauge fields

Connections can always be viewed in two ways: as defined on principal bundles, or vector bundles. We shall use both methods, so we begin by reviewing the relation between them.

On a *principal* G-bundle P over X, a connection is defined by a 1-form ω with values in the Lie algebra $\mathfrak{g}$ of G, and its curvature Ω is the $\mathfrak{g}$-valued 2-form

$$\mathrm{d}\omega + \tfrac{1}{2}[\omega, \omega],$$

which descends to X as a section of $\mathfrak{g} \otimes \Lambda^2$ where $\mathfrak{g}$ now denotes the vector bundle associated to P by the adjoint representation.

On a *vector* bundle E over X, a connection is defined by its covariant derivative ∇, which is a first order linear differential operator

$$\nabla : A^{\circ}(E) \rightarrow A^1(E),$$

where $A^p(E) = \Gamma(E \otimes \Lambda^p)$ is the space of smooth sections of $E \otimes \Lambda^p$. The covariant derivative has a natural extension

$$\mathrm{D}_1 : A^1(E) \rightarrow A^2(E), \tag{2.1}$$

defined by $$\mathrm{D}_1(e \otimes \alpha) = \nabla e \wedge \alpha + e \otimes \mathrm{d}\alpha,$$

where $e \in A^0(E)$ and $\alpha \in A^1$. The curvature Ω is then defined as the composition $\mathrm{D}_1 \nabla \in A^2(\mathrm{End}\, E)$. The relation is easy to describe. A representation of G on a vector space E defines an associated vector bundle $P \times_G E$ and a local section of P a distinguished local basis $\{e_i\}$ of E. Pulling back ω via the section and applying the representation we get a matrix of 1-forms ω_{ij} and define $\nabla e_i = \sum_j \omega_{ij} \otimes e_j$. Conversely if E has a G-structure preserved by ∇, then this defines ω on the principal bundle of G-frames.

For the physicist the curvature $\Omega \in A^2(g)$ is called the *gauge field*, and the connection form ω the *gauge potential*. The concept of equivalence of two connections which is appropriate here is that of gauge equivalence.

Definition. A *gauge transformation* on a principal G-bundle P is a diffeomorphism $f: P \rightarrow P$ such that (1) $f(gp) = gf(p)$ $g \in G$, $p \in P$, (2) f preserves each fibre, i.e. acts trivially on the base space X.

The infinite dimensional group $\mathscr{G}$ of all gauge transformations consists of sections of the bundle of groups $P \times_G G$ where G acts on itself by conjugation. Locally, $f \in \mathscr{G}$ may be represented as a G-valued function on X and a connection as a $\mathfrak{g}$-valued 1-form ω. The action of f on a connection is then locally given by

$$f^{-1}\omega = f^{-1}\,\mathrm{d}f + (\mathrm{Ad}\,f^{-1})\,(\omega).$$

Under a representation, $\mathscr{G}$ is mapped into the group $\Gamma(\mathrm{Aut}\,E)$ of automorphisms of the vector bundle E and its action on the covariant derivative is that of conjugation:

$$f^{-1}(\nabla) = f^{-1}\nabla f.$$

Connections have many invariants under gauge transformations. For example if p is any invariant polynomial on the Lie algebra $\mathfrak{g}$, then the Chern–Weil construction defines a differential form $p(\Omega)$ by applying p to the curvature. This form is gauge-invariant. There is another property of a connection which is gauge-invariant, and this is the notion of a self-dual connection, self-dual Yang–Mills field, self-dual gauge field, or instanton.

Definition. On a 4-manifold X, a connection is said to be *self-dual* if its curvature Ω is in $A^2_+(\mathfrak{g})$ (i.e. $\Omega = *\Omega$) and *anti-self-dual* if $\Omega \in A^2_-(\mathfrak{g})$ ($\Omega = -*\Omega$).

In fact since the star operator is conformally invariant on 2-forms, the property of self-duality of a connection is invariant under the larger group of transformations of a principal bundle consisting of those which act by conformal transformations on the base space X.

EXAMPLES

1. Take G to be $U(1)$. Since G is abelian, the curvature Ω is a closed 2-form such that $\Omega/2\pi i$ defines an integral class, the first Chern class, in $H^2(X, \mathbb{R})$. Given such a 2-form Ω, we can always find a $U(1)$ connection with Ω as its curvature. If the connection is *self-dual*, then $\mathrm{d}^*\Omega = *\mathrm{d}*\Omega = *\mathrm{d}\Omega = 0$, so $\Omega/2\pi i$ is in the harmonic space $H^2_+ \cap H^2(X, \mathbb{Z})$/torsion. This set may or may not be zero, depending on the conformal structure of X. For example, it can be shown that a flat torus T^4 has non-flat self-dual $U(1)$ connections iff it is an abelian variety.

2. Fix a Riemannian metric on X and consider the $SO(3)$ bundle Λ^2_+ with the induced Riemannian connection. The adjoint bundle $\mathfrak{g}$ is in this instance Λ^2_+ itself and the curvature of the induced connection is that part of the Riemann curvature tensor which lies in $\Lambda^2_+ \otimes \Lambda^2$, i.e.

$$\Omega = A + B^* \in A^2(\Lambda^2_+)$$

in the decomposition of § 1. Since $B^* \in A^2_-(\Lambda^2_+)$, this connection is self-dual iff $B \equiv 0$, in other words iff the metric is Einstein. In that case, Λ^2_- with the induced connection is anti-self-dual, and if X is a spin manifold, the spinor bundle V_+ is self-dual and V_- is anti-self-dual (as bundles with $SU(2)$-connections). We thus have the following proposition.

Proposition 2.2. *Let X be a 4-manifold with an Einstein metric. Then the induced connections on the bundle of self-dual spinors V_+ and the bundle of self-dual 2-forms Λ^2_+ are self-dual. The induced connections on the corresponding anti-self-dual bundles are anti-self-dual. Conversely, if the induced connections on V_+ and Λ^2_+ are self-dual, the metric is an Einstein metric.*

In this way we immediately get non-trivial self-dual $SU(2)$ connections on S^4. On $P_2(\mathbb{C})$, the $SO(3)$ connection on Λ^2_+ reduces to $SO(2) = U(1)$, but we still have a non-trivial anti-self-dual connection on Λ^2_-.

A bundle with a self-dual connection must satisfy some topological conditions. We see this in the above example of $U(1)$ connections where the first Chern class c_1 must be positive in some sense. The same is true of the first Pontrjagin class:

If E is a hermitian vector bundle with connection, then the 4-form

$$[1/(2\pi \mathrm{i})^2]\operatorname{tr}(\Omega^2)$$

represents the characteristic class

$$p_1(E) = (c_1^2 - 2c_2)(E) \in H^4(X, \mathbb{Z})$$

which, evaluated on the fundamental cycle of the compact manifold X, gives an integer.

If the connection is self-dual, this integer is *positive*, since if $\alpha \in \Lambda^2_+$, $\alpha^2 = |\alpha|^2\omega$, so

$$p_1(E) = -\frac{1}{4\pi^2}\int \operatorname{tr}\Omega^2 = \frac{1}{4\pi^2}\int |\Omega|^2\,\omega \geqslant 0,$$

and if $p_1(E) = 0$, the connection is flat. Similarly, if the connection is anti-self-dual, $p_1(E) \leqslant 0$.

In this context, the self-dual connections give absolute minima for the Yang–Mills functional

$$\frac{1}{8\pi^2}\int_X |\Omega|^2\omega = \frac{1}{8\pi^2}\int_X (|\Omega_+|^2 + |\Omega_-|^2)\,\omega,$$

for this is always greater than or equal to the topological invariant (essentially $p_1(E)$)

$$\frac{1}{8\pi^2}\int_X (|\Omega_+|^2 - |\Omega_-|^2)\,\omega,$$

with equality iff $\Omega_- = 0$, i.e. iff the connection is self-dual. These, then, are our basic objects of study: self-dual connections modulo the notion of gauge equivalence. We are going to link the structure of a self-dual connection with that of a self-dual base space, and our next task is to provide ourselves with a useful tool for this purpose.

3. An integrability theorem

It is well known that the vanishing of the curvature tensor of a connection is the condition for the integrability of the horizontal subspaces in the principal bundle or, in the vector bundle approach, for the existence of a local basis of solutions of the differential equation $\nabla s = 0$. Our self-dual condition involves the vanishing of *part* of the curvature, and we shall relate this to the solutions of a more general differential equation $\bar{\mathrm{D}}s = 0$, so in this section we shall consider first order differential operators from a particular point of view.

Let E be a vector bundle, which for the moment we suppose real. A section $s \in \Gamma(E)$ defines by duality a function $s^{\vee}$ on the total space of the dual bundle E^*:

$$s^{\vee}(\epsilon_x) = \langle s(x), \epsilon_x \rangle$$

and its derivative $\mathrm{d}s^{\vee}$ is a 1-form on E^*. In local terms, choose a basis $(e_1, \ldots, e_k)$ for E and take the dual basis $(\epsilon_1, \ldots, \epsilon_n)$ for E^*, then we can parametrize E^* locally by

$$(\lambda_1, \ldots, \lambda_n, x_1, \ldots, x_n) \mapsto \sum_i \lambda_i \epsilon_i(x) \in E^*.$$

Now if $s = \sum_i f_i e_i$, then

$$s^{\vee}(\lambda_1, \ldots, \lambda_n, x_1, \ldots, x_n) = \Sigma \lambda_i f_i(x)$$

and

$$\mathrm{d}s^{\vee} = \Sigma \mathrm{d}\lambda_i f_i + \Sigma \lambda_i \mathrm{d}f_i.$$

Now let $\bar{\mathrm{D}} : \Gamma(E) \to \Gamma(F)$ be a linear first order differential operator and let $I_x \subset F_x$ denote the subspace consisting of all $\bar{\mathrm{D}}(s)_x$ where s runs over all local sections of E at x. We shall say that $\bar{\mathrm{D}}$ is of *constant rank* if dim I_x is independent of x. When this is the case we shall associate with $\bar{\mathrm{D}}$ a vector bundle $V(\bar{\mathrm{D}})$ on $E^*\backslash 0$ (the complement of the zero section). $V(\bar{\mathrm{D}})$ is the sub-bundle of the cotangent bundle $T^*(E^*\backslash 0)$ whose fibre at a point $\epsilon_x \in E_x^*$ consists of all 1-forms $(\mathrm{d}s^{\vee})_{\epsilon_x}$ where s is a local section of E satisfying $\bar{\mathrm{D}}(s)_x = 0$. The constant rank assumption ensures that the spaces $V(\bar{\mathrm{D}})_{\epsilon_x}$ have constant dimension and that $V(\bar{\mathrm{D}})$ is indeed a vector bundle. By using the language of jets (Taylor expansions) this construction can be reformulated as follows: The derivation $\mathrm{d}s^{\vee}$ of sections of E factors through the universal derivative, the 1-jet $j_1(s) \in \Gamma(J_1(E))$ and we get a homomorphism of vector bundles over E^*:

$$V : p^*J_1(E) \to T^*E^*,$$

where $p : E^* \to X$ is the projection. V is surjective off the zero section and is characterized by the property

$$V(p^*j_1(s)) = \mathrm{d}s^{\vee}.$$

It is easy to see that V is well defined by this property and that when it is restricted to $p^*(E \otimes \Lambda^1) \subset p^*J_1(E)$,

$$V((e_x \otimes \alpha_x)_{\epsilon_x}) = -\langle e_x, \epsilon_x \rangle p^*\alpha_x \in (T^*E^*)_{\epsilon_x}.$$

A linear first-order differential operator $\bar{\mathrm{D}}$ is defined as a homomorphism from $J_1(E)$ to F, whose kernel R we assume to be a vector bundle. Then the vector bundle $V(\bar{\mathrm{D}})$ defined above is just $V(p^*R)$.

Example

If $\bar{\mathrm{D}} = \nabla$, the covariant derivative of a connection on E, then $R \subset J_1(E)$ is isomorphic to E under the natural homomorphism $J_1(E) \to E$ and $V(\nabla) \subset T^*(E^*\backslash 0)$

is the annihilator of the bundle of horizontal subspaces on E^*. Locally, relative to the basis $(e_1, \ldots, e_k)$ we obtain the connection matrix ω_{ij} where

$$\nabla_{e_i} = \sum_j \omega_{ij} \otimes e_j$$

and then $V(\nabla)$ is spanned by the 1-forms

$$\theta_i = \mathrm{d}\lambda_i - \sum_j \omega_{ij} \lambda_j \quad 1 \leqslant i \leqslant k.$$

Consider now a general differential operator $\overline{\mathrm{D}}$ of the form $\sigma\nabla$ where $\sigma : E \otimes \Lambda^1 \to F$ is the *symbol* of $\overline{\mathrm{D}}$, and ∇ is a connection on E. Let $S_1 = R \cap E \otimes \Lambda^1$ be the kernel of the symbol homomorphism and $S_2 \subset E \otimes \Lambda^2$ the image of $S_1 \otimes \Lambda^1$ under exterior multiplication.

PROPOSITION 3.1. *$V(\overline{\mathrm{D}}) \subset T^*(E^* \backslash 0)$ is involutive iff* (1) $\mathrm{D}_1 \Gamma(S_1) \subset \Gamma(S_2)$, (2) $\Omega\Gamma(E) \subset \Gamma(S_2)$, *where $\Omega : A^0(E) \to A^2(E)$ is the curvature of ∇ and D_1: $A^1(E) \to A^2(E)$ the extended covariant derivative* (2.1). (Note that the first condition is a 'torsion' condition on the connection – the vanishing of certain components of the connection matrix – and the second condition is a condition on the curvature.)

Proof. Recall that $V \subset T^*M$ is involutive if for any section $v \in \Gamma(V)$, $\mathrm{d}v = \Sigma v_i \wedge \alpha_i$ for 1-forms $v_i \in \Gamma(V)$, in other words iff $\mathrm{d}\Gamma(V) \subset \Gamma(V_2)$ where $V_2 \subset \Lambda^2$ is the image of $V \otimes \Lambda^1$ under exterior multiplication.

If $\overline{\mathrm{D}} = \sigma\nabla$, then $V(\overline{\mathrm{D}}) = V(p^*(S_1 \oplus E))$ and is thus spanned by the 1-forms

$$\theta_i = \mathrm{d}\lambda_i - \Sigma\omega_{ij} \lambda_j \, (1 \leqslant i \leqslant k) \text{ in } V(p^*E),$$

and

$$\sigma_i^\vee = \Sigma s_{ijk} \lambda_j \, \mathrm{d}x_k \, (1 \leqslant i \leqslant m) \text{ in } V(p^*S_1),$$

where $\sigma_i = \Sigma s_{ijk} e_j \otimes \mathrm{d}x_k$ locally span S_1. Now $T^*(E^* \backslash 0)$ has a local basis $\mathrm{d}\lambda_1, \ldots, \mathrm{d}\lambda_k$, $\mathrm{d}x_1. \ldots, \mathrm{d}x_n$, or equivalently $\theta_1, \ldots, \theta_k$, $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$, hence V_2 is spanned by forms $\theta_i \wedge \theta_j$, $\theta_i \wedge \mathrm{d}x_j$, $\sigma_i^\vee \wedge \mathrm{d}x_j$. The forms $\sigma_i^\vee \wedge \mathrm{d}x_j$ span $V(S_2)$. Taking the exterior derivative of θ_i and $\sigma_i^\vee$ we find

$$\begin{aligned} \mathrm{d}\theta_i &= -\Sigma\lambda_j \, \mathrm{d}\omega_{ij} - \Sigma \mathrm{d}\lambda_j \wedge \omega_{ij} \\ &= -\Sigma\lambda_j \, \mathrm{d}\omega_{ij} - \Sigma\omega_{jk} \wedge \omega_{ij} \lambda_k - \Sigma\theta_j \wedge \omega_{ij} \\ &= -\Sigma\lambda_k \, \Omega_{ik} - \Sigma\theta_j \wedge \omega_{ij}, \end{aligned}$$

which is a section of V_2 iff $\Sigma\lambda_k \, \Omega_{ik}$ is a section of $V(S_2)$ i.e. iff $\Omega\Gamma(E) \subset \Gamma(S_2)$.

$$\begin{aligned} \mathrm{d}\sigma_i^\vee &= \Sigma s_{ijk} \, \mathrm{d}\lambda_j \wedge \mathrm{d}x_k + \Sigma\lambda_j \, \mathrm{d}s_{ijk} \wedge \mathrm{d}x_k \\ &= \Sigma s_{ijk} \, \theta_j \wedge \mathrm{d}x_k + \Sigma s_{ijk} \, \omega_{mj} \, \lambda_m \wedge \mathrm{d}x_k + \Sigma\lambda_j \, \mathrm{d}s_{ijk} \, \mathrm{d}x_k \\ &= \Sigma s_{ijk} \, \theta_j \wedge \mathrm{d}x_k + (\mathrm{D}_1 \sigma_i)^\vee \end{aligned}$$

which is a section of V_2 iff $(\mathrm{D}_1 \sigma_i)^\vee$ is a section of $V(S_2)$, i.e. iff $\mathrm{D}_1 \Gamma(S_1) \subset \Gamma(S_2)$. Hence $V(\overline{\mathrm{D}})$ is involutive iff the conditions of the theorem hold.

Note that in the case $\overline{\mathrm{D}} = \nabla$, then $S_1 = 0$ and the condition is precisely the vanishing of the curvature. In the real situation, we could now apply the Frobenius integrability theorem and obtain a foliation on $E^*\backslash 0$. Suppose instead that E is a complex vector bundle, then $s^\vee$ is a complex valued function on E^*, $\mathrm{d}s^\vee$ a section of the complexified bundle and $V(\bar{\mathrm{D}})$ a sub-bundle of $T^*_{\mathrm{c}}(E^*\backslash 0)$. If now ∇ and $\bar{\mathrm{D}}$ commute with the complex structure on E, then proposition (3.1) still holds, where the λ_i's are to be treated as complex numbers.

There is one particularly important class of subbundles of the complexified cotangent bundle: those $V \subset T^*_{\mathrm{c}}$ such that $V \cap \bar{V} = 0$ and $V + \bar{V} = T^*_{\mathrm{c}}$, for such a bundle defines an almost complex structure. If it is *involutive*, then the Newlander–Nirenberg theorem (1957) implies that the structure is integrable.

4. Self-dual spaces and complex manifolds

As an application of proposition 3.1, we shall prove the following:

THEOREM 4.1. *Let X be an oriented 4-manifold. Then a conformal structure on X defines in a natural way an almost complex structure on $P(V_-)$, which is integrable iff $W_- = 0$, i.e. iff X is self-dual.* (Note that the spin representations are well defined projective representations of $SO(4)$, so we need not assume that X is a spin manifold.)

Proof. The question is local, so we may consider the vector bundle V_-. We also choose a metric within the conformal structure. There are two natural differential operators defined on V_-: the *Dirac operator*,

$$\mathrm{D}: \Gamma(V_-) \xrightarrow{\nabla} \Gamma(V_- \otimes \Lambda^1) \xrightarrow{\sigma} \Gamma(V_+),$$

whose symbol is Clifford multiplication, and the so-called *twistor operator* (Penrose 1975)

$$\bar{\mathrm{D}}: \Gamma(V_-) \xrightarrow{\nabla} \Gamma(V_- \otimes \Lambda^1) \xrightarrow{\bar{\sigma}} \Gamma(V_+^\perp),$$

where $\bar{\sigma}$ is orthogonal projection $1 - \sigma^*\sigma$ onto the kernel of σ in $\Gamma(V_- \otimes \Lambda^1)$. Locally, $\bar{\mathrm{D}}\psi = \nabla\psi + \frac{1}{4}\Sigma e_i . \mathrm{D}\psi \otimes e_i \in \Gamma(V_+^\perp)$, where $g = \Sigma e_i \otimes e_i$ is the metric tensor. We are going to apply proposition 3.1 to the operator $\bar{\mathrm{D}}$, so first consider $S_1 = \ker \bar{\sigma} \subset V_- \otimes \Lambda^1$. This is just V_+ embedded as $\psi \mapsto \Sigma e_i . \psi \otimes e_i$. Now at $\phi \in V_-^*$, $V(S_1)$ is $\Sigma\langle e_i . \psi, \phi\rangle e_i$, i.e. the $\Lambda^{1,0}$ subspace of Λ^1_c parametrized by the symplectic dual spinor to ϕ (see 1.2)). Hence $V(\bar{\mathrm{D}}) \cong V(p^*S_1 \oplus p^*V_-)$ is a four-dimensional complex sub-bundle of $T^*_{\mathrm{c}}(V_-^*\backslash 0)$ and such that $V(\bar{\mathrm{D}}) \cap \overline{V(\bar{\mathrm{D}})} = 0$, in other words an almost complex structure. To check its integrability we just have to verify the conditions of (3.1).

1. If $\phi = \Sigma e_i . \psi \otimes e_i$, then since the Riemannian connection preserves the metric tensor g,

$$\nabla\phi = \Sigma e_i . \nabla\psi \otimes e_i.$$

Furthermore, since ∇ is torsion free, $\mathrm{D}_1\phi = A(\nabla\phi) = \Sigma e_i . \nabla_j \psi \otimes e_j \wedge e_i \in \Gamma(S_2)$ where A denotes alternation.

2. $\Gamma(S_2)$ consists of sections which are in the image of the map $V_+ \otimes \Lambda^1 \to V_- \otimes \Lambda^2$ defined by $\psi \otimes \alpha \mapsto \Sigma e_i . \psi \otimes e_i \wedge \alpha$. If $\Sigma \psi_j \otimes e_j$ is in the kernel of this map,

$$\Sigma e_i . \psi_j \otimes e_i \wedge e_j = 0,$$

and hence $e_i \psi_j = e_j \psi_i$ for $i \neq j$ from which it follows that $e_i \psi_i = -e_j \psi_j = 0$, and so

$$V_+ \otimes \Lambda^1 \subset V_- \otimes \Lambda^2.$$

If we decompose these bundles into irreducible components,

$$V_+ \otimes \Lambda^1 \cong (S^2 V_+ \otimes V_-) \oplus V_-,$$

and
$$V_- \otimes \Lambda^2 \cong (S^2 V_+ \otimes V_-) \oplus V_- \oplus S^3 V_-.$$

Hence $\Omega \Gamma(V_-) \subset \Gamma(S_2)$ iff the composition

$$V_- \xrightarrow{\Omega} V_- \otimes \Lambda^2 \longrightarrow S^3 V_-$$

is zero. But this is precisely the $S^4 V_-$ component of the Riemann curvature tensor, i.e. W_-.

So the sub-bundle is involutive, and the complex structure integrable, iff X is self-dual, i.e. $W_- \equiv 0$. The structure is clearly invariant under scalar multiplication by $\lambda \in \mathbb{C}^*$ on $V_-^* \backslash 0$ and hence the quotient space $P(V_-^*) = P(V_-)$ has a complex structure. The twistor equation $\bar{\mathrm{D}}\psi = 0$ can be made conformally invariant by giving the bundle V_- conformal weight $-\frac{1}{2}$ (Fegan 1976). Since the sub-bundle $\mathrm{V}(\bar{\mathrm{D}})$ is defined by the operator $\bar{\mathrm{D}}$ it follows that the complex structure on $P(V_-)$ is defined in terms of the conformal structure on X. In particular any orientation preserving conformal transformation of X induces a holomorphic transformation of $P(V_-)$.

Remarks

1. If we fix a Riemannian metric in the conformal class, the almost complex structure on $P(V_-)$ can be seen more geometrically as follows. First, using the Riemannian connection we can split the tangent bundle of $P(V_-)$ into vertical and horizontal parts. On the vertical part we have the complex structure of the fibres (complex projective lines). On the horizontal part at a point $\phi \in P(V_-)_x$ over $x \in X$ we put the complex structure on Λ^1_x parametrized by ϕ as explained in §1. Note that this complex structure defines the opposite orientation on Λ^1_x since $\phi \in P(V_-)_x$. One can now proceed to verify directly that the vanishing of W_- is precisely the integrability condition for the almost complex structure. Finally, one can check that conformally equivalent metrics on X give rise to the same almost complex structure on $P(V_-)$. In our approach the conformal invariance follows from the known conformal invariance of the twistor equation.

2. Since the complex structure on $P(V_-)$ descends from one on $V_-^* \backslash 0$, we have locally a holomorphic line bundle H over $P(V_-)$ whose principal bundle is $V_-^* \backslash 0$. Globally, the spin bundle V_- and the line bundle H exist only if $w_2(X) = 0$, but in

all cases H^2 is a well defined bundle. H^2 is in fact canonically associated to the complex structure on $P(V_-)$ and is not an extra piece of data: we have $H^{-4} \cong K$, the canonical line bundle of holomorphic 3-forms on $P(V_-)$. This can be seen as follows: the bundle R associated to the twistor equation $\bar{\mathrm{D}}\phi = 0$ is an extension

$$V_+ \to R \to V_-$$

of the spin bundle V_- of conformal weight $-\frac{1}{2}$ by V_+ of weight $1-\frac{1}{2} = \frac{1}{2}$ and hence R has a canonical $SL(4, \mathbb{C})$ structure. R is mapped into $T^*(V_-^*\backslash 0)$ as the holomorphic cotangent bundle and so $V_-^*\backslash 0$ has a non-vanishing section ω of $\Lambda^4 T^*$. Locally $\omega = \sigma_1^\vee \wedge \sigma_2^\vee \wedge \theta_1 \wedge \theta_2$ in the notation of (3.1) where $\sigma_\alpha^\vee = \Sigma \lambda_j \langle e_i . \psi_\alpha, \phi_j \rangle e_i$ and $\theta_i = \mathrm{d}\lambda_i - \Sigma \omega_{ij} \lambda_j$ and one may show that $\mathrm{d}\omega = 0$, i.e. ω is holomorphic, using the expressions for $\mathrm{d}\theta_i$ and $\mathrm{d}\sigma_\alpha^\vee$ in (3.1) and the fact that $W_- = 0$. On the other hand ω is homogeneous of degree 4 in λ and hence trivializes $H^4 K$ on $P(V_-)$. Hence $K \cong H^{-4}$.

The fibres of $P(V_-) \to X$ are complex submanifolds, projective lines, of $P(V_-)$. The normal bundle of each fibre is trivial as a real bundle, but not as a holomorphic one as we shall see next. Take a fibre $(V_-^*)_x$, with λ_1, λ_2 linear coordinates. Then from proposition 3.1, the conormal bundle N^* of this fibre is spanned by $\sigma_1^\vee$ and $\sigma_2^\vee$ where

$$\sigma_\alpha^\vee = \Sigma \langle e_i . \psi_\alpha, \phi \rangle e_i,$$

where $\phi \in (V_-^*)_x$ and $\psi_\alpha \in (V_+)_x$. Now $\sigma_1^\vee$ and $\sigma_2^\vee$ are *holomorphic* sections of N^* on $(V_-^*)_x$ since $\mathrm{d}\sigma_\alpha^\vee$ in (3.1) contains no $\mathrm{d}\bar{\lambda}_i$ terms. Hence they trivialize N^* on $(V_-^*)_x$. But they are linear in ϕ, so on $P(V_-^*)_x$ they trivialize HN^* where H is the hyperplane bundle, and so $N \cong H \oplus H$, or more invariantly $N \cong (V_+^*)_x \otimes H$. Next note that the holomorphic sections $H^0(P_1, \mathcal{O}(H))$ are parametrized naturally by $(V_-)_x$ and so the space of holomorphic sections of N,

$$H^0(P_1, \mathcal{O}(N)) \cong (V_+^* \otimes V_-)_x \cong \mathrm{Hom}\,(V_+, V_-)_x \cong (\Lambda_c^1)_x$$

Consider a decomposable element $\phi \otimes \psi \in (V_+^* \otimes V_-)_x$. In $H^0(P_1, \mathcal{O}(N))$ this defines a section ϕl_ψ where l_ψ is the linear form corresponding to ψ. This is a section of N which vanishes somewhere, namely at θ where $l_\psi(\theta) = \langle \psi, \theta \rangle = 0$. Conversely, the set of sections in $H^0(P_1, \mathcal{O}(N))$ which vanish somewhere corresponds to the set of decomposable elements in $(V_+^* \otimes V_-)_x$, or equivalently elements of rank less than or equal to 1 in $\mathrm{Hom}\,(V_+, V_-)_x$. If $\alpha \in (\Lambda_c^1)_x$ annihilates a spinor ψ by Clifford multiplication, then $0 = \alpha^2 . \psi = -(\alpha, \alpha)\psi$ and so α is of length zero, and conversely if $(\alpha, \alpha) = 0$, α is of rank less than or equal to 1. It follows then that there is a 1–1 correspondence between holomorphic sections of N which vanish somewhere, and the null-cone in $(\Lambda_c^1)_x$ defining the (complex) conformal structure.

This is a principle of fundamental importance in the Penrose twistor programme; the conformal differential geometry over the complex numbers has been coded into the holomorphic structure of $P(V_-)$ and the holomorphic lines thereon.

In fact, given a complex 3-manifold Z and a line $P_1(\mathbb{C}) \subset Z$ with normal bundle $H \oplus H$, then $H^1(P_1, \mathcal{O}(N)) = 0$ and it follows from a theorem of Kodaira (1962) that the line belongs to a 4-parameter complex analytic family whose tangent space at C is naturally isomorphic to $H^0(C, \mathcal{O}(N))$. The set of sections which vanish somewhere defines naturally a complex conformal structure which, as Penrose (1976) shows, is self-dual. In our case, we need a positive definite real conformal structure and hence an extra piece of information. This information is a *real structure* on the complex manifold $P(V_-)$, that is, an anti-holomorphic involution $\tau: P(V_-) \to P(V_-)$, defined by the quaternionic structure $J: V_- \to V_-$ in each fibre. On each fibre $P_1(\mathbb{C})$, J is anti-holomorphic and, regarding $P_1(\mathbb{C})$ as S^2, is the antipodal map and so takes the complex structure I on Λ^1_x to $-I$. On the direct sum of the complex structures, τ is thus anti-holomorphic.

τ has no fixed points on $P(V_-)$ but does leave the fibres invariant. This means that the base space X, which parametrizes the fibres, is mapped naturally into the fixed point set of τ with respect to its natural action on the complex four-dimensional family of lines described above, and the action on the tangent space of this fixed point set reduces the complex conformal structure to a real one. As a consequence of this, any self-dual space X has a real analytic structure, relative to which the conformal structure is real analytic.

To sum up, theorem 4.1 translates a self-dual space into a complex 3-manifold with no real points, fibred by a real family of lines having normal bundle $H \oplus H$.

Examples

1. Take X to be $\mathbb{R}^4$ with its flat conformal structure, then $P(V_-)$ is just $S^2 \times \mathbb{R}^4$ given the complex structure of the total space of the holomorphic bundle $H \oplus H$ over $P_1(\mathbb{C})$. The holomorphic sections give the 4-parameter family $\mathbb{C}^4$ of lines. If X is replaced by a flat torus $\mathbb{R}^4/\Gamma$, we get the non-Kähler manifolds described by Blanchard (1956) and Calabi.

2. If $X = S^4$, then $P(V_-) = P_3(\mathbb{C})$. The lines are just the ordinary lines in $P_3(\mathbb{C})$, parametrized by the Klein quadric Q_4. The real structure on $P_3(\mathbb{C})$ is given by a quaternionic structure $J: \mathbb{C}^4 \to \mathbb{C}^4$ and the real lines are defined by quaternionic planes in $\mathbb{C}^4$, i.e. $P_1(\mathbb{H}) \cong S^4$, also to be thought of as the real quadric $S^4 \subset Q_4$.

3. If $X = P_2(\mathbb{C})$, then $P(V_-)$ is the flag manifold F_3. A point of F_3 is a pair (x, l) where $x \in P_2(\mathbb{C})$ and $l \subset P_2(\mathbb{C})$ is a line containing x. The standard metric on $P_2(\mathbb{C})$ is induced from a hermitian form on $\mathbb{C}^3$ which defines an antilinear map from $\mathbb{C}^3$ to its dual space and thus an antiholomorphic map from $P_2(\mathbb{C})$ to $P_2^*(\mathbb{C})$. Under this map a pair (x, l) goes to a pair $(\bar{x}, \bar{l})$ where $\bar{l} \in \bar{x}$. This is the real structure on F_3.

The complex 4-parameter family of lines on F_3 is given by taking pairs (y, m) where $y \in P_2(\mathbb{C})$ and $m \subset P_2(\mathbb{C})$ is a line *not* passing through y. For each line l through y, we associate the point $(l \cap m, l) \in F_3$ and so get the $P_1(\mathbb{C})$ of lines through y mapped into F_3. The real lines are those where the 1-dimensional vector space defined by y, and the 2-dimensional space defined by m in $\mathbb{C}^3$, are orthogonal relative to the hermitian form, and such pairs are simply parametrized by $y \in P_2(\mathbb{C})$.

Note that the projection from F_3 to $P_2(\mathbb{C})$ is not the usual holomorphic one: instead of associating to (x, l) the point x, we associate the point $x^{\perp}$ orthogonal to

x in l relative to the hermitian structure of $\mathbb{C}^3$. Note further that the fibre $P(V_-)_x$ parametrizes complex structures on Λ^1_x with *opposite* orientation to the integrable complex structure on $P_2(\mathbb{C})$. That complex structure defines a section of $P(V_+)$, the reduction of Λ^2_+ to $SO(2)$ mentioned in § 2.

4. If X is any conformally flat manifold, then $P(V_-)$ has a flat holomorphic projective connection. If $\pi_1(X) \to SO(5,1)$ is the holonomy representation on X, its composite with the projective spin representation $SO(5,1) \to SO(6, \mathbb{C}) \to PSL(4, \mathbb{C})$ is the holonomy representation on $P(V_-)$. This gives a way of constructing many examples of compact projectively flat complex manifolds.

5. Self-dual connections and holomorphic bundles

A second application of Proposition (3.1) is the following (see also Griffiths 1966).

THEOREM 5.1. *Let X be a complex manifold, E a C^∞ hermitian vector bundle with connection ∇ whose curvature Ω is of type* (1, 1), *i.e.* $\Omega \in A^{1,1}(\operatorname{End} E)$. *Then E has a natural holomorphic structure and ∇ is the unique* (1, 0) *hermitian connection.*

Proof. We apply (3.1) to the following differential operator:

$$\bar{\mathrm{D}} : \Gamma(E) \xrightarrow{\nabla} \Gamma(E \otimes \Lambda^1_{\mathrm{c}}) \xrightarrow{\sigma} \Gamma(E \otimes \Lambda^{0,1}).$$

$\bar{\mathrm{D}}$ is formally like the Dolbeault $\bar\partial$ operator. We are going to show that there is a complex structure on E for which $\bar{\mathrm{D}} = \bar\partial$.

Here $S_1 = \ker \sigma = E \otimes \Lambda^{1,0}$ and so clearly $V(\bar{\mathrm{D}}) \cap \overline{V(\bar{\mathrm{D}})} = 0$. Moreover $\dim V(\bar{\mathrm{D}}) = \dim X + \dim E$, so again we have an almost complex structure on $E^*\backslash 0$. It remains to verify the integrability conditions of (3.1).

1. If $e \otimes \mathrm{d}z \in \Gamma(S_1)$, then

$$D_1(e \otimes \mathrm{d}z) = \Sigma \nabla_i e \otimes \mathrm{d}z_i \wedge \mathrm{d}z + \Sigma \nabla_{\bar{i}} e \otimes \mathrm{d}\bar{z}_i \wedge \mathrm{d}z$$

which is in $A^{2,0}(E) \oplus A^{1,1}(E) \subset \Gamma(S_2)$.

2. If $\Omega \in A^{1,1}(\operatorname{End} E)$, then

$$\Omega\Gamma(E) \subset A^{1,1}(E) \subset \Gamma(S_2).$$

(In fact we see here that it is only necessary to assume that the curvature has no component of type (0, 2). For a hermitian connection this of course means that the (2, 0) component vanishes as well.)

Hence (3.1) implies that the almost complex structure is integrable, i.e. $E^*\backslash 0$ is a complex manifold. In fact the complex structure extends across the zero section since $V(\bar{\mathrm{D}})$ is generated by the forms

$$\mathrm{d}\lambda_i - \Sigma \omega_{ij} \lambda_j \quad \text{and} \quad \lambda_i \, \mathrm{d}z_j,$$

or equivalently
$$\mathrm{d}\lambda_i - \Sigma \omega_{ij} \lambda_j \quad \text{and} \quad \mathrm{d}z_j,$$

and this extends. What is more, the zero section is a complex submanifold relative to this complex structure, so its normal bundle is holomorphic. But the normal bundle is naturally isomorphic to E^*, so E^* and E are holomorphic vector bundles. The isomorphism between the conormal bundle N^* and E may be described as follows: to each section $s \in \Gamma(E)$ associate the 1-form $\mathrm{d}s^{\vee}$ of § 3 restricted to the zero section. In local coordinates this is $e_i \mapsto \mathrm{d}\lambda_i|_0$ which is a local section of N^*, since the zero section is defined by $\lambda_1 = \ldots = \lambda_n = 0$. But the zero section is also *holomorphic* and so is locally given by $w_1 = \ldots = w_n = 0$ where w_i are holomorphic functions. Being holomorphic they must satisfy

$$\mathrm{d}w_i = \Sigma A_{ij}(\mathrm{d}\lambda_j - \omega_{jk}\,\lambda_k) + \Sigma B_{ij}\,\mathrm{d}z_j. \tag{5.2}$$

Now $\mathrm{d}w_i|_0 = \Sigma A_{ij}\,\mathrm{d}\lambda_j|_0$ is a local holomorphic section of N^*. Using the isomorphism between N^* and E, we can apply the covariant derivative $\nabla_{\bar{l}}$ in the direction $\partial/\partial\bar{z}_l$ and find

$$\nabla_{\bar{l}}(\mathrm{d}w_i|_0) = \Sigma(\partial A_{ij}/\partial\bar{z}_l)\,\mathrm{d}\lambda_j|_0 + \Sigma A_{ij}\,\omega^{\bar{l}}_{jk}\,\mathrm{d}\lambda_k|_0.$$

From (5.2) we have $A_{ij} = \partial w_i/\partial\lambda_j$ and $\partial w_i/\partial\bar{z}_l = -\Sigma A_{ij}\,\omega^{\bar{l}}_{jk}\,\lambda_k$, hence at $\lambda_1 = \ldots = \lambda_n = 0$, $\partial A_{ij}/\partial\bar{z}_l = \partial^2 w_i/\partial\bar{z}_l\,\partial\lambda_j = -\Sigma A_{ik}\,\omega^{\bar{l}}_{kj}$, and so $\nabla_{\bar{l}}(\mathrm{d}w_i|_0) = 0$ and ∇ is a $(1,0)$ connection. In other words $\bar{\mathrm{D}} = \bar{\partial}$.

Remarks

1. Although the theorem is expressed in terms of connections on vector bundles, it is equally true for principal bundles with compact structure group, for if E has a G-structure for $G \subset U(k)$ which is preserved by the connection, then the holomorphic bundle $P \times_{U(k)} GL(k, \mathbb{C})/G^{\mathrm{c}}$ has a covariant constant section where P is the bundle of unitary frames of E. Since ∇ is a $(1,0)$ connection, every covariant constant section is holomorphic and hence the holomorphic frame bundle reduces to G^{c}. In the framework of principal G^{c}-bundles the holomorphic structure can be described geometrically as follows. Using the connection we split the tangent bundle to the principal bundle into horizontal and vertical parts. Using the complex structure of base (X) and fibre (G^{c}) we then get an almost complex structure on the principal bundle. One can then check that the integrability for this almost complex structure is the vanishing of the $(0,2)$-component of the curvature tensor. The complex structure is clearly invariant under the action of G^{c} and hence defines a holomorphic bundle.

2. The condition $\Omega \in A^{1,1}(\mathrm{End}\, E)$ is unchanged by a gauge transformation. For two gauge-equivalent connections, the two complex structures on E are not the same, but are *equivalent* under the gauge transformation.

3. If $\dim_{\mathbb{C}} X = 1$, then $A^2 = A^{1,1}$ so every hermitian connection on a C^∞ vector bundle over a Riemann surface defines a complex structure.

4. If $\dim_{\mathbb{C}} X = 2$ and X has a hermitian metric, then a bundle E on the underlying real 4-manifold is anti-self-dual iff $\Omega \in A^2_-(\mathrm{End}\ E) = A^{1,1}_0(\mathrm{End}\ E)$ where $A^{1,1}_0$ is the space of $(1,1)$ forms orthogonal to the hermitian form (see § 1). It follows from the theorem, then, that each such bundle is holomorphic. Moreover, because the curvature is orthogonal to the hermitian 2-form any holomorphic section of E is covariant constant by the vanishing theorem of Kobayashi & Wu (1970).

As an example we saw in § 2 that the $SO(3)$ bundle on Λ^2_- on $P_2(\mathbb{C})$ was anti-self-dual. The theorem gives its complexification the holomorphic structure of the sub-bundle of End T of traceless endomorphisms of the holomorphic tangent bundle. The manifold on which we really want to use theorem 5.1 is the complex 3-manifold $P(V_-)$ of theorem 4.1. The main result, following the ideas of Penrose and Ward, is that the concept of a self-dual connection on a bundle over a self-dual space transforms into that of a holomorphic bundle on a complex manifold. More precisely:

THEOREM 5.2. *Let E be a hermitian vector bundle with self-dual connection over a self-dual 4-manifold X and let $F = p^*E$ be the pulled back bundle on $P(V_-)$. Then* (1) *F is holomorphic on $P(V_-)$;* (2) *F is holomorphically trivial on each fibre;* (3) *there is a holomorphic isomorphism $\sigma: \tau^*\bar{F} \to F^*$, where $\tau: P(V_-) \to P(V_-)$ is the real structure, such that σ induces a* positive definite *hermitian structure on the space of holomorphic sections of F on each fibre. Conversely, every such bundle on $P(V_-)$ is the pull-back of a bundle E with self-dual connection on X.*

Proof. 1. From remark 2 of theorem 4.1, the complex structure on the horizontal space Λ^1_x at a point $\phi_x \in P(V_-)_x$ has the *opposite* orientation to that of X. Hence by remark 4 above,

$$(\Lambda^2_+)_x \cong (\Lambda^{1,1}_0)_x$$

and so $p^*\Lambda^2_+$ is a sub-bundle of $\Lambda^{1,1}$ of $P(V_-)$. It follows then, from theorem 5.1 that if E has a self-dual connection, the curvature of $F = p^*E$ is in $\Lambda^{1,1}(\text{End } F)$ and so F is holomorphic.

2. Along each fibre P_1, the pulled-back connection is clearly trivial, since any basis $(e_1, \dots, e_k)$ of E at x is covariant constant on the fibre at x. But since ∇ is a $(1,0)$ connection every covariant constant section is holomorphic and so along each fibre, F is holomorphically trivial.

3. On E^* take a local unitary basis $(\epsilon_1, \dots, \epsilon_k)$ then in theorem 5.1 the complex structure on F^* is defined by the forms

$$\mathrm{d}\lambda_i - \Sigma\omega_{ij}\lambda_j \quad \text{and} \quad \mathrm{d}z_j.$$

The first type of form is invariant under τ, since they are pulled back from the base space X. Hence applying τ, we get a complex structure on $\bar{F}^*$ defined by the forms

$$\mathrm{d}\bar{\lambda}_i - \Sigma\bar{\omega}_{ij}\bar{\lambda}_j \quad \text{and} \quad \mathrm{d}z_j.$$

But since the connection is hermitian, $\bar{\omega}_{ji} = -\omega_{ij}$ and we get a holomorphic structure on $\bar{F}^*$ defined by

$$\mathrm{d}\bar{\lambda}_i + \Sigma\omega_{ji}\bar{\lambda}_j \quad \text{and} \quad \mathrm{d}z_j,$$

and this under the hermitian isomorphism $\bar{F}^* \cong F$ gives the induced holomorphic structure on F, i.e.

$$\tau^*\bar{F} \cong F^*.$$

Conversely, suppose F is a bundle on $P(V_-)$ satisfying 1, 2 and 3. If F is holomorphically trivial on each fibre, then there is a natural isomorphism

$$H^0(P(V_-)_{p(z)}, \mathcal{O}(F)) \xrightarrow{\cong} F_z, \tag{5.3}$$

obtained by evaluation of a holomorphic section at each point z in the fibre. Since the holomorphic sections are solutions of an elliptic equation, the spaces $H^0(P(V_-)_x, \mathcal{O}(F))$ over X form a vector bundle E and so $F \simeq p^*E$.

The isomorphism $\sigma: \tau^*\overline{F} \simeq F^*$ induces a hermitian structure on the holomorphic sections on each fibre, which by hypothesis is positive definite, and thus gives a hermitian structure on E and F. Take the unique $(1,0)$ connection ∇ on F preserving this hermitian structure.

We claim that ∇ is the pull-back of a self-dual connection on E. First, by the uniqueness of $(1,0)$ connections, ∇ on each fibre† is the flat connection defined by the trivialization (5.3), hence any section† of E is covariant constant along the fibre when pulled back and so ∇ defines a connection on E over X. Secondly, since ∇ is of type $(1,0)$, its curvature $\Omega \in A^{1,1}(P(V_-), \operatorname{End} E)$ but since ∇ is also pulled back from X, $\Omega \in p^*A^2(X, \operatorname{End} E)$. Now, as is easily verified,

$$A^{1,1}(P(V_-)) \cap p^*A^2(X) = p^*A^2_+(X).$$

Hence ∇ on E is self-dual.

Remarks

1. A consequence of this theorem is that a self-dual connection on a self-dual space is gauge-equivalent to a real analytic one. A particular consequence of this is that the restricted, local and infinitesimal holonomy groups are all the same (Kobayashi & Nomizu 1963).

2. Since the idea behind this theorem is to encode all the information on X into the holomorphic structure of objects on $P(V_-)$, there remains the question about whether the isomorphism $\sigma: \tau^*\overline{F} \to F^*$ is an extra piece of information. In fact this is not so: if the connection on E is *irreducible* (i.e. E has no sub-bundles preserved by ∇), then σ is *unique*, modulo a scalar multiple. This again is a consequence of the theorem of Kobayashi & Wu (1970), for any two isomorphisms differ by an automorphism of the bundle F. We can introduce a hermitian metric on $P(V_-)$ (by pulling back the metric of X to the horizontal subspaces and adding the metric of the fibre) such that Ω is of type $(1,1)$ and orthogonal to the hermitian form. It follows that any holomorphic section of End F is covariant constant and, if ∇ is irreducible, must be a scalar.

3. Theorem 5.2 is again valid for a principal G-bundle, where G is a compact real form of the complex group G^c, by taking a faithful representation of G. A principal G-bundle P on X with self-dual connection is then characterized as a holomorphic principal G^c bundle P^c on $P(V_-)$, trivial on the fibres of $P(V_-)$ and an antiholomorphic map $f: P^c \to P^c$ such that f acts on the base space $P(V_-)$ by τ and $f(g.p) = \bar{g}f(p)$ where $g \mapsto \bar{g}$ denotes the real structure on G^c. If the connection on P is irreducible (i.e. G is the closure of the holonomy group), then f is uniquely defined modulo the centre of G.

This reinterpretation of self-dual connections on self-dual spaces makes the

† See footnote on p. 461.

construction of such objects in principle easier, for we have at our disposal the techniques and apparatus of complex analysis, sheaf theory and algebraic geometry. This approach is treated elsewhere (Atiyah & Ward 1977; Hartshorne 1977). The aim of this paper is to use a differential geometric approach on the manifold X, and so we leave the complex analysis, though noting that it does provide justification for linking together the twin concepts of self-dual manifold and self-dual connection.

6. Moduli of self-dual connections

We have seen in § 2 that the space of self-dual connections on a principal G-bundle P is acted upon by the group $\mathscr{G}$ of gauge transformations of P, so that given one such connection we get infinitely many. The situation is analogous to that of complex structures on a manifold: given one, we get many by applying diffeomorphisms. It makes sense, however, to talk of complex structures modulo this notion of equivalence, and the space of such structures, when it exists, is called the space of moduli. We use the same notation here: the space of all self-dual connections on P modulo gauge equivalence will be called the *space of moduli* of self-dual connections on P.

If $H \subset G$ is a subgroup, then any self-dual H connection defines a self-dual G-connection, so that the space of moduli of G-connections contains H-connections for all subgroups $H \subset G$. (Actually the map from the moduli space of H to that of G-connections is not quite an inclusion because H may have outer automorphisms induced by inner automorphisms of G.) It is more natural, then to consider the space of all *irreducible* self-dual G-connections, i.e. those for which the connection does not reduce to any proper closed subgroup $H \subset G$. Since irreducibility is an open condition, this will form an open set in the space of all connections, though as we shall see later it may be empty.

We know nothing *a priori* about these moduli spaces, the next theorem shows that under fairly general conditions, they are finite dimensional manifolds:

Theorem 6.1. *Let X be a compact self-dual Riemannian 4-manifold with positive scalar curvature. Let P be a principal G-bundle over X where G is a compact semi-simple Lie group.*

Then, the space of moduli of irreducible self-dual connections on P is either empty or a manifold of dimension

$$p_1(\mathfrak{g}) - \tfrac{1}{2}\dim G(\chi - \tau),$$

where $p_1(\mathfrak{g})$ is the first Pontrjagin class of the bundle associated to P by the adjoint representation – the 'adjoint bundle', χ is the Euler characteristic of X, and τ is the signature of X.

Proof. The proof is in three parts: infinitesimal, local, and global.

1. Compute the dimension of the space of infinitesimal deformations of a self-dual connection using the Atiyah–Singer index theorem and a vanishing theorem.

2. Use the method of Kuranishi to apply Banach space inverse and implicit

function theorems to integrate the infinitesimal deformations and obtain a local moduli space.

3. Show that the local moduli spaces give local coordinates on the global moduli space and that this global space is a Hausdorff manifold.

1. First assume the space is non-empty, and that we have at least one self-dual connection ω. If ω' is another connection, the two differ by an element $\tau \in A^1(\mathfrak{g})$ and the relation between the two curvatures is

$$\Omega' - \Omega = \mathrm{D}_1 \tau + \tfrac{1}{2}[\tau, \tau],$$

where $\mathrm{D}_1 : A^1(\mathfrak{g}) \to A^2(\mathfrak{g})$ is the extended covariant derivative (2.1). Thus if ω_t is a 1-parameter family of self-dual connections,

$$\Omega_t = \Omega + \mathrm{D}_1 \tau_t + \tfrac{1}{2}[\tau_t, \tau_t],$$

and

$$p_-(\mathrm{D}_1 \tau_t + \tfrac{1}{2}[\tau_t, \tau_t]) = 0 \in A^2_-(\mathfrak{g}), \tag{6.2}$$

where p_- is the projection onto Λ^2_-, i.e.

$$p_- \alpha = \tfrac{1}{2}(\alpha - *\alpha).$$

Differentiating with respect to t and putting $t = 0$, we get $p_-(\mathrm{D}_1 \dot{\tau}) = 0 \in A^2_-(\mathfrak{g})$ where $\dot{\tau} = (\partial \tau_t / \partial t)|_{t=0}$. If the family were obtained by a 1-parameter family of gauge transformations, i.e. $\omega_t = f_t^{-1} . \omega$ where $f_t \in \mathcal{G}$, then from § 2 we would have $\dot{\tau} = \nabla \dot{f}$ where $\dot{f} \in \Gamma(\mathfrak{g}) = A^0(\mathfrak{g})$. Thus a 1-parameter family of self-dual connections defines an element in $\ker p_- \mathrm{D}_1 / \mathrm{Im}\, \nabla$. Now $p_- \mathrm{D}_1 \nabla = p_-(\mathrm{D}_1 \nabla) = p_-(\Omega) = 0$ since the connection is self-dual, so we have defined an element in the first cohomology group $H^1(\mathfrak{g})$ of the following complex:

$$0 \to A^0(\mathfrak{g}) \xrightarrow{\mathrm{d}_0} A^1(\mathfrak{g}) \xrightarrow{\mathrm{d}_1} A^2_-(\mathfrak{g}) \to 0,$$

where $\mathrm{d}_0 = \nabla$ and $\mathrm{d}_1 = p_- \mathrm{D}_1$. It is easy to see that this complex is elliptic and hence the cohomology groups are finite dimensional. The aim is to compute $h^1 = \dim H^1(\mathfrak{g})$ by calculating the alternating sum $h^0 - h^1 + h^2$ by the index formula and then using a vanishing theorem to show that $h^0 = h^2 = 0$.

In fact $h^0 = \dim H^0(\mathfrak{g})$ and these are the covariant constant sections of $\mathfrak{g}$ which correspond to the Lie algebra of the centralizer of the holonomy group. Since by hypothesis the connection is irreducible, this is the centre of G, which, since G is semi-simple, is zero-dimensional. Hence $h^0 = 0$.

To proceed with the vanishing of h^2, it is useful to replace, in the standard way, the above elliptic complex by a single elliptic operator

$$\mathrm{d} + \mathrm{d}^* : A^1(\mathfrak{g}) \to A^0(\mathfrak{g}) \oplus A^2(\mathfrak{g}),$$

where d^* denotes the formal adjoint of $\mathrm{d}_0 = \nabla$ by means of the Riemannian metric on X.

We can write this in terms of the Dirac operator D associated to the metric:

$$\mathrm{D} : \Gamma(V_+ \otimes V_- \otimes \mathfrak{g}) \to \Gamma(V_- \otimes V_- \otimes \mathfrak{g}),$$

since these two elliptic operators have the same symbol and factor through the same connection; in other words they can both be written in the form $\mathrm{D}\phi = \Sigma\sigma_i \nabla_i \phi$ where $\nabla_i \phi$ is the covariant derivative of ϕ relative to the connection induced on $\Lambda^1 \otimes \mathfrak{g} \simeq V_+ \otimes V_- \otimes \mathfrak{g}$ by the Riemannian connection and the given self-dual connection.

Let us now consider in general, the Dirac operator

$$\mathrm{D}: \Gamma(V \otimes E) \to \Gamma(V \otimes E)$$

on spinors with values in some auxiliary bundle E with connection. We can consider D as the composition

$$\Gamma(V \otimes E) \xrightarrow{\nabla} \Gamma(V \otimes E \otimes \Lambda^1) \xrightarrow{C} \Gamma(V \otimes E),$$

where ∇ is the covariant derivative of the connection induced on $V \otimes E$ by the Riemannian connection on V and the given connection on E, and C is Clifford multiplication by Λ^1 on V. We can now write $\mathrm{D}^2 = (C\nabla)^2 = C^2\nabla^2$, where C^2: $V \otimes \Lambda^1 \otimes \Lambda^1 \to V$ is Clifford multiplication $\psi \otimes \alpha \otimes \beta \mapsto \beta \,.\, \alpha \,.\, \psi$. Now Clifford multiplication is defined by the property $\alpha \,.\, \beta + \beta \,.\, \alpha = -2(\alpha, \beta)$, hence C^2 on the symmetric part of $\nabla^2\psi$ is just $-\mathrm{Tr}\ \nabla^2\psi = \nabla^*\nabla\psi$ where ∇^* is the formal adjoint of ∇, and the skew part of ∇^2 is the curvature K of the bundle $V \otimes E$, considered as a section of End $(V \otimes E) \otimes \Lambda^2$ and hence of Hom $(V \otimes E,\ V \otimes E \otimes \Lambda^2)$. We thus get the Weitzenböck decomposition of D^2:

$$\mathrm{D}^2\psi = \nabla^*\nabla\psi + C(K)\,\psi,$$

where $C(K)$ is the self-adjoint composition

$$V \otimes E \xrightarrow{K} V \otimes E \otimes \Lambda^2 \xrightarrow{C} V \otimes E,$$
$$\psi \otimes e \otimes \alpha \mapsto \alpha \,.\, \psi \otimes e.$$

The general vanishing theorem is now clear: if $\mathrm{D}\psi = 0$, then $\mathrm{D}^2\psi = 0$, so

$$0 = \int (\nabla\psi, \nabla\psi) + (C(K)\,\psi, \psi),$$

and if the endomorphism $C(K)$ is positive definite, then this expression is positive unless $\psi \equiv 0$. We apply this to the case $E = V_- \otimes \mathfrak{g}$, and consider D on $V_- \otimes V_- \otimes \mathfrak{g}$. Then the curvature of the bundle $V_- \otimes V_- \otimes g$ is

$$K = K(V_- \otimes V_-) \otimes 1 + 1 \otimes K(\mathfrak{g}).$$

The connection on P is self-dual, so $K(\mathfrak{g}) \in \Lambda^2_+(\mathfrak{g})$. But Λ^2_+ acts trivially on V_- by Clifford multiplication (see §1), so $C(1 \otimes K(\mathfrak{g})) = 0$. Now $K(V_- \otimes V_-) \in \Gamma(\Lambda^2 \otimes \Lambda^2)$ but only the components in $\Lambda^2_- \otimes \Lambda^2_-$ act non-trivially on V_-, and this part of the Riemann curvature tensor corresponds to C in the decomposition (1.1). If the space X is self-dual, then $W_- \equiv 0$ and C is a scalar, essentially the scalar curvature. We then get

$$C(K) = \tfrac{1}{3}R \quad \text{on} \quad \Lambda^2_- \otimes \mathfrak{g},$$

and so if $R > 0$, $C(K) > 0$ and if $\mathrm{D}\psi = 0$, then $\psi = 0$. Hence $h^2 = 0$ in the complex.

We now have the vanishing result and so

$$\operatorname{index}(\mathrm{d}+\mathrm{d}^*) = -h^0+h^1-h^2 = h^1$$

and we may use the index theorem (Atiyah & Singer 1968) to evaluate the alternating sum. For the Dirac operator

$$\mathrm{D}: \Gamma(V_+ \otimes E) \to \Gamma(V_- \otimes E), \quad \operatorname{index}\mathrm{D} = \operatorname{ch}(E)\hat{\mathscr{A}}(X)[X],$$

where $\operatorname{ch}(E)$ = Chern character of E, and $\hat{\mathscr{A}}(X) = \hat{A}$ polynomial in the Pontrjagin classes of X. In four dimensions $\hat{\mathscr{A}}(X) = 1-\frac{1}{24}p_1(X)$. If $E = V_- \otimes \mathfrak{g}$, then

$$\begin{aligned}\operatorname{index}\mathrm{D} &= \operatorname{ch}(\mathfrak{g})\operatorname{ch}(V_-)\hat{\mathscr{A}}(X)[X]\\ &= p_1(\mathfrak{g})+\dim G\,(\operatorname{index}\mathrm{D}: \Gamma(V_+\otimes V_-)\to\Gamma(V_-\otimes V_-))\\ &= p_1(\mathfrak{g})-\tfrac{1}{2}\dim G(\chi-\tau),\end{aligned}$$

and so this is the dimension of $H^1(\mathfrak{g})$, which is going to be the tangent space at ω to the space of moduli. Recall that $p_1(\mathfrak{g}) \geqslant 0$ for a self-dual connection.

2. We have completed the infinitesimal computation. Next we have to integrate these infinitesimal deformations and show that every element in $H^1(\mathfrak{g})$ is defined by a 1-parameter family. For this we follow closely the argument of Kuranishi (1965). From (6.2), a self-dual connection is given by a solution of the non-linear equation

$$p_-(\mathrm{D}_1\tau+\tfrac{1}{2}[\tau,\tau]) = 0 \quad \tau\in A^1(\mathfrak{g}).$$

For ease of notation, let us write this as

$$\mathrm{d}\tau+\{\tau,\tau\} = 0,$$

where d is the differential in the complex

$$A^0(\mathfrak{g}) \xrightarrow{\mathrm{d}} A^1(\mathfrak{g}) \xrightarrow{\mathrm{d}} A^2_-(\mathfrak{g}).$$

Consider the set Φ of self-dual connections

$$\Phi = \{\tau\in A^1(\mathfrak{g}): \mathrm{d}\tau+\{\tau,\tau\} = 0 \quad\text{and}\quad \mathrm{d}^*\tau = 0\}.$$

We shall show that a neighbourhood of the origin of Φ is a local space of moduli, but first we show that Φ is an h^1 dimensional manifold with tangent space $H^1(\mathfrak{g})$ at the origin.

Take the Green operator G of the Laplacian $\mathrm{dd}^*+\mathrm{d}^*\mathrm{d} = \Delta$ of the complex. Then

$$G\Delta = 1-H,$$

where H is the orthogonal projection onto the harmonic subspace. Define a map $F: A^1(\mathfrak{g})\to A^1(\mathfrak{g})$ by

$$F(\tau) = \tau+G\mathrm{d}^*\{\tau,\tau\};$$

then

$$\begin{aligned}\mathrm{d}F(\tau) &= \mathrm{d}\tau+\mathrm{d}G\mathrm{d}^*\{\tau,\tau\}\\ &= \mathrm{d}\tau+G\mathrm{dd}^*\{\tau,\tau\}\\ &= \mathrm{d}\tau+\{\tau,\tau\} \quad \text{since } h^2 = 0,\end{aligned}$$

and
$$\begin{aligned} \mathrm{d}^*F(\tau) &= \mathrm{d}^*\tau + \mathrm{d}^*G\mathrm{d}^*\{\tau,\tau\} \\ &= \mathrm{d}^*\tau + G\mathrm{d}^*\mathrm{d}^*\{\tau,\tau\} \\ &= \mathrm{d}^*\tau \quad \text{since} \quad d^2 = 0. \end{aligned}$$

Hence F maps Φ into the harmonic subspace $H^1 \subset A^1(\mathfrak{g})$ of dimension h^1. If we give $A^1(\mathfrak{g})$ the Sobolev $\|_k$ norm topology for sufficiently large k, then by the Sobolev inequalities F extends to the Banach space completion of $A^1(\mathfrak{g})$ as a holomorphic map whose differential at $\tau = 0$ is the identity. The inverse function theorem and the regularity theorems for elliptic operators imply that F is invertible on C^∞ sections and F^{-1} restricted to H^1 gives local coordinates for Φ.

We now have an h^1-dimensional family Φ to realize all the infinitesimal deformations at ω. We want to show that Φ is (a) locally *complete*, i.e. any self-dual connection sufficiently close to ω is gauge equivalent to one in Φ; (b) locally *universal*, i.e. any C^∞ *family* of nearby self-dual connections is gauge equivalent to a C^∞ family in Φ; (c) locally *effective*, i.e. no two self-dual connections in Φ sufficiently close to ω are gauge equivalent by a small gauge transformation. For this we use the exponential map $\exp A^0(\mathfrak{g}) \to \mathscr{G}$ and show the following

LEMMA. *There exists a neighbourhood U of* 0 *in* $A^1(\mathfrak{g})$ *and* W *of* 0 *in* $A^0(\mathfrak{g})$ *such that for any $\tau \in U$, there is a unique $X \in W$ with* $\mathrm{d}^*(\omega - \exp(X)(\omega+\tau)) = 0$.

Proof.
$$\begin{aligned} \exp(-X)(\omega+\tau) - \omega &= \mathrm{e}^{-X}(\nabla \mathrm{e}^{X}) + \mathrm{Ad}(\exp(-X))\tau \\ &= \nabla X + \tau + R(X,\tau), \end{aligned}$$

where $R(tX, t\tau) = t^2 R(X,\tau,t)$ and $R(X,\tau,t)$ is C^∞ in t for small t and in τ and the 1-jet of X.

Now
$$\mathrm{d}^*(\exp(-X)(\omega+\tau) - \omega) = 0 \text{ iff}$$
$$\mathrm{d}^*\mathrm{d}X + \mathrm{d}^*\tau + \mathrm{d}^*R(X,\tau) = 0. \tag{6.3}$$

Applying G, we have
$$X + G\mathrm{d}^*\tau + G\mathrm{d}^*R(X,\tau) = 0,$$

since $H^0(\mathfrak{g}) = 0$. Take neighbourhoods $U_1 \subset A^1(\mathfrak{g})$, $W_1 \subset A^0(\mathfrak{g})$ of the origin and let F be the mapping

$$F: U_1 \times W_1 \to A^0(\mathfrak{g}), \quad (\tau, X) \mapsto X + G\mathrm{d}^*\tau + G\mathrm{d}^*R(X,\tau).$$

R is continuous in the $\|_k$ topology on $A^1(\mathfrak{g}) \times A^0(\mathfrak{g})$ to the $\|_{k-1}$ topology on $A^0(\mathfrak{g})$, so by the Sobolev inequalities F is continuous in the $\|_k$ topology for small enough U_1, W_1. Hence F can be extended to the completion and by (6.3) and the implicit function theorem in Banach spaces, for sufficiently small U, any $\tau \in U$ admits a unique solution $X(\tau)$ in the completion of $A^0(\mathfrak{g})$, which is small. By elliptic regularity, since X satisfies the equation

$$\Delta X + \mathrm{d}^*R(X,\tau) + \mathrm{d}^*\tau = 0,$$

X is C^∞ if τ is C^∞. Hence if ω' is self-dual and sufficiently close to ω, there exists a gauge transformation $\exp(X)$ for small X such that $\exp(X)\omega' \in \Phi$, and we have local *completeness*.

If $\omega(s)$, $s \in \mathbb{R}^k$ is a C^∞ family of self-dual connections, then for all $s \in \mathbb{R}^k$ such that $\omega(s) \in U$,

$$\mathrm{d}^*(\omega - \exp(-X(s))\,\omega(s)) = 0$$

where $X(s)$ is C^∞ in s, since it is obtained from a C^∞ mapping of Banach spaces. The map $s \mapsto \exp(-X(s))\,\omega(s)$ then defines a gauge-equivalent C^∞-family in Φ, so we have local *universality*.

The uniqueness in the lemma implies local *effectiveness*, where $\mathscr{G}$ is given the topology induced from the $\|_k$ topology on $A^0(\mathfrak{g})$ by the exponential map.

We now have our local moduli space: a neighbourhood of the origin in Φ.

3. The space of all connections $\mathscr{A}$ on P is an affine space isomorphic to $A^1(\mathfrak{g})$. If $\mathscr{A}^+$ denotes the self-dual irreducible connections, then the *global space of moduli* is just $\mathscr{A}^+/\mathscr{G}$, the quotient space under the action of the gauge group. We give $\mathscr{M} = \mathscr{A}^+/\mathscr{G}$ the quotient topology from the $\|_k$ topology on $A^1(\mathfrak{g})$. This clearly has a countable basis of open sets.

We want to show that $\mathscr{M}$ is a manifold. First we show that $\mathscr{M}$ is a *Hausdorff* space, in fact a metric space with the $\|_0$ topology from $\mathscr{A}$. Suppose $\omega, \omega+\tau \in \mathscr{A}$ are connections in different orbits under $\mathscr{G}$, then since the $\|_0$ norm is invariant under $\mathscr{G}$, we can define the distance between two orbits as

$$\inf_{f\in\mathscr{G}} |\omega+\tau-f^{-1}\omega|_0 = \inf_{f\in\mathscr{G}} |\tau - f^{-1}\nabla f|_0,$$

so long as this is *non-zero*. We now map $\mathscr{G}$ into the sections of the bundle of groups $\mathrm{Ad}\,G \subset \mathrm{End}\,\mathfrak{g}$ and regard f as a section of the vector bundle $\mathrm{End}\,\mathfrak{g}$. Then

$$|\tau - f^{-1}\nabla f|_0^2 = |f\tau - \nabla f|_0^2 = (\Delta f, f)_0 \geqslant \lambda |f_1|_0^2,$$

where $\Delta = \mathrm{D}^*\mathrm{D}$, $\mathrm{D}f = \nabla f - f\tau$, λ is the smallest non-zero eigenvalue of Δ, and f_1 is the projection of f onto the orthogonal complement of the zero eigenspace H. Consider the function F on the finite dimensional space H defined by

$$F(s) = \int \inf_{f\in(\mathrm{Ad}\,G)_x} |s(x) - f|^2.$$

Since by hypothesis there are no sections of $\mathrm{Ad}\,G$ satisfying $\tau = f^{-1}\nabla f$, $F(s) > 0$ but as G is compact, $F(s) > \mu > 0$ and so if $f \in \mathscr{G}$, $|f_1|_0^2 > \mu$. Hence

$$|\tau - f^{-1}\nabla f|_0^2 > \lambda\mu,$$

and the distance between two orbits is *positive*.

We next want to show that the local moduli spaces Φ give local coordinates on $\mathscr{M}$: we need to prove that a sufficiently small neighbourhood maps *injectively* into $\mathscr{M}$, in other words is globally effective. To see this, suppose ω and $\omega+\tau$ are in Φ and are gauge equivalent by an arbitrary gauge transformation f, then

$$f^{-1}\nabla f = \tau.$$

Again map $\mathscr{G}$ into $\Gamma(\mathrm{End}\,\mathfrak{g})$ by the adjoint representation (the kernel of this map is just the finite group of central gauge transformations), and decompose

End $\mathfrak{g} = E_0 \oplus E_1$ where E_0 is the trivial bundle corresponding to the endomorphisms of the Lie algebra invariant under G. Now Ad $\mathscr{G} \cap A^0(E_0)$ is the identity section so if we decompose $f \in \mathrm{Ad}\,\mathscr{G}$ as $f = f_0 + f_1$, then $|f_1|$ measures the distance of f from the identity. Furthermore, pointwise

$$|\tau|^2 = |f^{-1}\nabla f|^2 = |\nabla f|^2 = |\nabla f_0|^2 + |\nabla f_1|^2, \tag{6.4}$$

and integrating
$$|\tau|_0^2 \geqslant \lambda |f_1|_0^2, \tag{6.5}$$

where λ is the smallest eigenvalue of $\nabla^*\nabla$ on E_1, non-zero since the connection is irreducible. Hence if τ is small in the $\|_0$ norm, f is close to the identity in the $\|_0$ norm in End $\mathfrak{g}$.

Unfortunately this is not the topology for local completeness: we have to lift the section through the exponential map. However the pointwise formula (6.4) shows that if τ is C^0 close to zero, then the length $|f_1|$ is close to a constant, which by (6.5) is close to zero, hence f can be lifted back to $X \in A^0(\mathfrak{g})$ as a smooth section via the exponential map. By taking derivatives, if τ is C^k close to zero, X is C^k and hence $\|_k$ close to zero, and eventually in the neighbourhood W of the lemma and hence equal to zero. On a sufficiently small neighbourhood of Φ, the C^k and $\|_k$ topologies coincide and the above estimates can be made uniform in τ, so that finally we have some open set $U \subset \Phi$ which maps injectively into $\mathscr{A}^+/\mathscr{G}$.

This map is continuous, and also open since its inverse image in $\mathscr{A}^+$ is $\bigcup_{f \in \mathscr{G}} fU$ which is open by the lemma, so we get a homeomorphism of an open set of Euclidean space to a neighbourhood of $\mathscr{A}^+/\mathscr{G}$, and a topological manifold structure on it. The differentiability follows from the local *universality* of the local moduli spaces.

Hence, finally, $\mathscr{M} = \mathscr{A}^+/\mathscr{G}$ is a differentiable manifold of dimension $p_1(\mathfrak{g}) - \frac{1}{2}\dim G\,(\chi - \tau)$.

What we have not done as yet is to show that there exist self-dual connections and moduli spaces. We shall do this next, and look at the case which is of most physical interest: the 4-sphere S^4.

7. The 't Hooft solutions

From a global point of view, the first thing to notice about the 4-sphere S^4 is that it is simply-connected, and so the holonomy group of any connection is *connected*. Secondly, $H^2(S^4, \mathbb{Z}) = 0$, so there are no non-trivial *abelian* self-dual connections and, more generally, the holonomy group of any self-dual G-connection with compact G is *semi-simple*. Furthermore, since the second cohomology group with arbitrary coefficients is zero, there is a unique lifting to a connection on the universal covering group bundle, which splits into a product of bundles with simply-connected *simple* Lie groups as structure group. The general problem can thus be treated by considering an irreducible self-dual connection with holonomy group G where G is a compact simply-connected simple Lie group.

The first such group is $SU(2)$, and we shall describe next a range of self-dual examples on S^4 due to 't Hooft; see also Jackiw, Nohl & Rebbi (1977).

A principal G-bundle on S^4 is classified topologically by an element in $\pi_3(G)$, which since $SU(2) \cong S^3$ is just the integers $\mathbb{Z}$, and the integer invariant which classifies the bundle is the second Chern class $c_2(E)$ where E is the associated two-dimensional complex vector bundle. Note that for a self-dual connection $p_1(E) = -2c_2(E)$ is *positive*, so only bundles with negative c_2 may have such connections. The 't Hooft examples realize all negative values of c_2.

To construct them, we go back to the twistor equation in (4.1):

$$\overline{\mathrm{D}}\phi = \nabla\phi + \tfrac{1}{4}\Sigma e_i\,.\,\mathrm{D}\phi \otimes e_i = 0$$

On $\mathbb{R}^4$, this has, even locally, a four-dimensional vector space T of solutions which consist of affine linear spinors of the form $x\,.\,\phi+\psi$ where $\phi \in \Gamma(V_+)$, $\psi \in \Gamma(V_-)$ are covariant constant spinors. By interpreting these spinors to have conformal weight $-\frac{1}{2}$, they extend to the whole of S^4 under stereographic projection.

We have an *evaluation* map

$$p: S^4 \times T \to V_-,$$

$$p(x, \phi) = \phi(x),$$

since the vectors in T are sections of V_-.

In $\mathbb{R}^4$, there is a distinguished two-dimensional subspace of T of constant spinors. By choosing a scale m, i.e. by fixing a flat metric on $\mathbb{R}^4$, we have a hermitian structure on this subspace (the conformal structure gives it a quaternionic structure) and its product with S^4 is a flat hermitian sub-bundle L of $S^4 \times T$. The projection p from L to V_- is surjective except at the point at infinity on S^4.

Now $\mathbb{R}^4$ was obtained by stereographic projection from $a_0 = \infty$. In general take $(k+1)$ distinct points $a_0, \ldots, a_k$ and look at the flat, non-intersecting bundles $L_0, \ldots, L_k \subset S^4 \times T$. We take the direct sum $L = L_0 \oplus \ldots \oplus L_k$ and consider the projection map onto the direct sum $\overset{k}{\underset{1}{\oplus}} V_-$ of k copies of V_- defined by

$$(\phi_0, \ldots, \phi_k) \mapsto (p(\phi_0 - \phi_1), \ldots, p(\phi_0 - \phi_k)).$$

This is surjective if the points are distinct and thus has a two-dimensional bundle E as kernel. E has an induced hermitian structure from L, and $c_2(E) = -kc_2(V_-) = -k$.

We put a connection on E in the obvious way: restrict the flat connection of L to E and use the hermitian structure to project back to E:

$$\nabla_E: A^0(E) \xrightarrow{\nabla_L} A^1(L) \to A^1(E).$$

We claim this connection is *self-dual*.

To see this, we need only consider the local situation. Outside the points $a_1, \ldots, a_k$, projection onto L_0 is an isomorphism, hence a local section of L_0 defines a local section of E as follows: to a section ϕ of L_0 we must associate a section $(\phi, \phi_1, \ldots, \phi_k)$ where ϕ_i is a linear spinor vanishing at a_i, i.e. of the form $(y-a_i)\,.\,\psi_i$ and whose value at x, $(x-a_i).\psi_i(x)$, is $\phi(x)$. Choosing ϕ to be a constant section of L_0, we have $\psi_i(x) = (x-a_i)^{-1}\phi$.

First of all, this local parametrization does not preserve norms, for if $|\phi|^2 = 1$ in L_0, then in E

$$|\phi|_E^2 = |\phi|^2 + \Sigma|x_i^{-1}.\phi|^2 m_i^2 = 1 + \Sigma m_i^2/r_i^2 = \rho,$$

where $x_i = x - a_i$ and m_i^2 is the scale which determines the hermitian structure on L_i.

Secondly, in the *flat* connection on $L_0 \oplus \dots \oplus L_k$,

$$\nabla\phi = (0, x_1.\nabla(x_1^{-1}).\phi, \dots, x_k.\nabla(x_k^{-1}).\phi)$$

and projecting back onto E

$$\nabla_E\phi = \sum_{i,\alpha}(x_i^{-1}.\phi_\alpha, \nabla(x_i^{-1}).\phi)\,\phi_\alpha/|\phi_\alpha|_E^2,$$

where $\{\phi_\alpha\}$ is a unitary basis of constant sections of L_0. Hence the connection matrix relative to this basis is

$$\omega = -\sum_{i-1}^{k}(x_i^{-1}/\rho)\,\nabla(x_i^{-1}).$$

However using the fact that $x^{-1} = -x/r^2$ in Clifford multiplication we see that

$$-x^{-1}.\nabla(x^{-1}).\phi = \Sigma e_j.(x/r^4)\otimes e_j$$

and since $\mathrm{d}(1/r^2) = -2x/r^4$, then

$$\omega = -\tfrac{1}{2}\Sigma e_j.\mathrm{d}(\log\rho)\otimes e_j, \tag{7.1}$$

where $\rho = 1 + \Sigma m_i^2/r_i^2$.

ω is thus locally given as the image of a 1-form $\alpha = -\frac{1}{2}\mathrm{d}(\log\rho)$ under the natural map

$$\Lambda^1 \to \Lambda_-^2 \otimes \Lambda^1,$$

$$\alpha \mapsto \Sigma(e_j.\alpha)_- \otimes e_j,$$

where Λ_-^2 is identified with the Lie algebra of $SU(2)$. The *curvature* $\Omega = \mathrm{d}\omega + \frac{1}{2}[\omega, \omega]$ is then in the image of natural maps

$$\Lambda^1 \otimes \Lambda^1 \to \Lambda_-^2 \otimes \Lambda^2$$

both from the derivative term $\nabla\alpha$ and the quadratic term $\alpha\otimes\alpha$. For E to be *self-dual*, we require the component in $\Lambda_-^2 \otimes \Lambda_-^2 \cong 1 \oplus \Lambda_-^2 \oplus S^4V_-$ to be zero. But there are only two such natural maps: the inner product $\Lambda^1\otimes\Lambda^1 \to 1$, and the skew part $\Lambda^1\otimes\Lambda^1\to\Lambda_-^2$. For the derivative term, since $\mathrm{d}\alpha = 0$ the skew-part vanishes and since $\alpha\otimes\alpha$ is symmetric, it vanishes on this too. The remaining scalar condition is $\mathrm{d}^*\alpha + 2(\alpha,\alpha) = 0$ which for $\alpha = -\frac{1}{2}\mathrm{d}\ (\log\rho)$ reduces to $\Delta\rho = 0$ which is satisfied for $\rho = 1 + \Sigma m_i^2/r_i^2$.

Hence these connections are all *self-dual*.

Remarks

1. If we take the spinor bundle V_+ on S^4, then $c_2(V_+) = -1$ and the 't Hooft construction gives a self-dual connection on V_+, given a pair of points a_0, a_1 and a scale. On the other hand, we know that a metric of constant curvature also gives a self-dual connection on V_+. There is a relation: for every pair of points in S^4, there is a constant curvature metric for which the points are antipodal and this metric is fixed by choosing a scale. The two connections produced in different ways are then gauge-equivalent.

2. The four-dimensional space T of solutions to the equation $\bar{\mathrm{D}}\phi = 0$ is in fact naturally the dual of the vector space whose projective space is $P(V_-)$ over S^4 (cf. theorem 4.1). The isomorphism between this space and $L_0 \oplus L_1$ gives it a symplectic structure since L_0 and L_1 are $SU(2) = \mathrm{Sp}(1)$ spaces. The holomorphic bundle F on $P_3(\mathbb{C})$ corresponding via theorem 5.2 to E on S^4 is then $F = H^\perp/H$ where $H \subset P_3(\mathbb{C}) \times \mathbb{C}^4$ is the Hopf bundle and $H^\perp$ its orthogonal space relative to the symplectic structure.

The number of parameters in the construction of the 't Hooft solutions is $5k+4$ in general: $(k+1)$ points on S^4 and the scales $m_0, \ldots, m_k$ modulo a scalar multiple. For $k \leqslant 2$ this set of data has conformal symmetries and the number of *effective* parameters is less, but for $k > 2$ the $5k+4$ parameters are effective (Jackiw *et al.* 1977). We may check, then, from theorem 6.1 whether we have conceivably realized all moduli.

For S^4, $\chi = 2$ and $\tau = 0$. The three-dimensional adjoint bundle of an $SU(2)$ connection may be identified, after complexification, with S^2E where E is the two-dimensional vector bundle with $c_2(E) = -k$, so since

$$p_1(\mathfrak{g}) = p_1(S^2E) = (c_1^2 - 2c_2)(S^2E),$$

we see from the Chern character formula

$$\begin{aligned} \mathrm{ch}(S^2E) &= \mathrm{ch}(E)^2 - \mathrm{ch}(1) \\ &= (2+kx)^2 - 1 \\ &= 3 + 4kx \end{aligned}$$

that $p_1(\mathfrak{g}) = 8k$ and so the dimension of the space of moduli is $8k-3$ (see Atiyah *et al.* 1977; Schwarz 1977). Since $8k-3 > 5k+4$ for $k > 2$ we see the power of theorem 6.1, for it asserts the existence of new self-dual $SU(2)$ connections by deformation of the 't Hooft solutions. We shall see next that, using the theorem repeatedly, we can prove the existence of irreducible self-dual G-connections on S^4 for an arbitrary simple Lie group G.

8. Solutions for simple G

For an arbitrary compact simple Lie group we again have $\pi_3(G) \cong \mathbb{Z}$ and so a principal G-bundle over S^4 is classified by an integer k, called by the physicists the Pontrjagin index, topological charge or instanton number. For our purposes we need the generator of $\pi_3(G)$ explicitly and for this we go to the Morse-theoretic proof of $\pi_3(G) \cong \mathbb{Z}$ given by Bott (1956).

Let G be a simply-connected compact simple Lie group with maximal torus T. If t is the Lie algebra of T and $X \in t$, then the straight line $0X$ from the origin to X defines a geodesic in G whose index is twice the number of root planes crossed. Furthermore, every such line in the fundamental chamber crosses the root plane $\theta(t) = 1$ where θ is the highest root. Hence if X is in this root plane, any non-minimal geodesic from the identity $e \in G$ to $\exp X$ must have index not less than 4. The space Ω^{d} of minimal geodesics from e to $\exp X$ is $C(\exp X)/T$ where $C(g)$ is the centralizer of g and this is just a 2-sphere S^2. Since this is a manifold, it follows (see Milnor 1963, §22) that the induced homomorphism $\pi_2(\Omega^{\mathrm{d}}) \to \pi_2(\Omega G)$ is an isomorphism, where ΩG is the loop space of G. Hence

$$\pi_3(G) \cong \pi_2(\Omega G) \cong \pi_2(S^2) \cong \mathbb{Z}.$$

The Lie algebra of $C(\exp X)$ is $t \oplus e_\theta$ where e_θ is the two-dimensional root space of θ. This root space generates a three-dimensional subalgebra which is the Lie algebra of a subgroup K of $C(\exp X)$, and such that $K/K \cap T = C(\exp X)/T$. Hence if $SU(2)$ is the universal covering group of K, we have a homomorphism $\rho: SU(2) \to G$ which induces an isomorphism $\rho_*: \pi_3(SU(2)) \to \pi_3(G)$. (In fact, since G is 2-connected, the Hurewicz homomorphism $\pi_3(G) \to H_3(G)$ is an isomorphism and ρ induces an isomorphism $\rho_*: H_3(SU(2)) \to H_3(G)$. From this it follows that K itself is a copy of $SU(2)$, since the covering homomorphism $H_3(SU(2)) \to H_3(SO(3))$ is of degree 2.) In topological terms it means that any principal G-bundle of index k reduces to a principal $SU(2)$-bundle of index k.

With this information, we can calculate the Pontrjagin number $p_1(\mathfrak{g})$ in terms of k. The first Pontrjagin class $p_1(E)$ of a vector bundle E is defined under the Chern–Weil homomorphism by the invariant polynomial $\operatorname{tr} A^2$ on End E. On $\mathfrak{g}$, this is the invariant polynomial $\operatorname{tr}(\operatorname{ad} X)^2$: the Killing form. Since the bundle reduces to $SU(2)$, we can compute $p_1(\mathfrak{g})$ by restricting the Killing form of G to $SU(2)$ and so

$$p_1(\mathfrak{g}) = \frac{B_G}{B_{SU(2)}} p_1(su(2)),$$

where B_G is the restricted Killing form and $B_{SU(2)}$ the Killing form of $SU(2)$.

Now
$$\frac{B_G}{B_{SU(2)}} = \frac{B^\vee_{SU(2)}}{B^\vee_G} = \frac{B^\vee_{SU(2)}(\theta,\theta)}{B^\vee_G(\theta,\theta)},$$

where $B^\vee$ is the induced form on $\mathfrak{g}^*$ and θ the highest root. The square of the length of the highest root can be readily computed (see Bourbaki 1968) to give the following table

G	A_l	B_l	C_l	D_l	G_2	F_4	E_6	E_7	E_8
$B^\vee(\theta,\theta)$	$\frac{1}{l+1}$	$\frac{1}{2l-1}$	$\frac{1}{l+1}$	$\frac{1}{2l-2}$	$\frac{1}{4}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{30}$
		$l \geqslant 3$		$l \geqslant 4$					

where the simply-connected groups corresponding to A_l, B_l, C_l, D_l are of course $SU(l+1)$, Spin$(2l+1)$, Sp(l) and Spin$(2l)$. This yields the following table for the dimension of the space of moduli of irreducible self-dual G-connections of index k (if it exists):

TABLE 8.1.

G	$\dim \mu = p_1(\mathfrak{g}) - \dim G$	
$SU(n)$	$4nk - n^2 + 1$	
Spin (n)	$4(n-2)k - \frac{1}{2}n(n-1)$	$(n \geqslant 7)$
Sp (n)	$4(n+1)k - n(2n+1)$	
G_2	$16k - 14$	
F_4	$36k - 52$	
E_6	$48k - 78$	
E_7	$72k - 133$	
E_8	$120k - 248$	

(These numbers have been independently computed by Bernard *et al.* (1977). Note that since $p_1(\mathfrak{g}) \geqslant 0$ for a self-dual connection, $k \geqslant 0$.

If dim G is large compared with k, then some of these numbers become negative and consequently there are no irreducible self-dual connections with such an index k. We are going to determine next the precise range of k for which there exist irreducible self-dual G-connections. Roughly speaking we shall start with a G-connection and an embedding of G in H, and by a parameter count show that there are more H-connections than reductions to G. Beginning with $SU(2)$ and the 't Hooft solutions, we inductively get solutions for H. There are a number of difficulties, however.

The first is keeping track of the holonomy group: if we start with an irreducible G-connection on a principal bundle P and if by deformation we get a connection on the associated H-bundle Q, what do we know about the holonomy group of the deformed connection? In fact, the holonomy group has a semi-continuity property. If ω_t is a smooth family of connections with holonomy group Φ_t, then for sufficiently small t,

$$\dim \Phi_t \geqslant \dim \Phi_0. \tag{8.2}$$

In our case, it is easiest to see this by using the real analyticity (remark 2 of theorem 5.2), for here the Lie algebra of the holonomy group is isomorphic to that of the infinitesimal holonomy group and this, at a point $p \in P$ is spanned by the values at p of the $\mathfrak{g}$-valued functions $\tilde{V}_k \dots \tilde{V}_1 \Omega(\tilde{X}, \tilde{Y})$ where $\tilde{X}$, $\tilde{Y}$, ... are the horizontal lifts of vector fields X, Y,... on S^4, and Ω is the curvature of the connection (see Kobayashi & Nomizu 1963).

If $\Omega_1, \dots, \Omega_n$ form a basis for this Lie algebra at $t = 0$, then they remain linearly independent under a small deformation, and hence the inequality (8.2).

In the case of equality, then by the rigidity of semi-simple subalgebras of semi-simple Lie algebras (Richardson 1967), there is a gauge transformation in Q which transforms the deformed connection to a G-connection on P. Hence, if $G \subset H$ is a semi-simple subgroup of *maximal* dimension, a deformed G-connection is either gauge-equivalent to a connection on P, or is an *irreducible* H-connection, i.e. one whose holonomy group is the whole of H.

Now every simply-connected compact simple Lie group has a distinguished connected subgroup of highest dimension, as can be seen from Dynkin's (1957) paper. We list them:

TABLE 8.3

$$
\begin{aligned}
S(U(n)\times U(1)) &\subset SU(n+1)\\
\mathrm{Spin}\,(n) &\subset \mathrm{Spin}\,(n+1)\\
\mathrm{Sp}(n)\times \mathrm{Sp}(1) &\subset \mathrm{Sp}(n+1)\\
SU(3) &\subset G_2\\
\mathrm{Spin}\,(9) &\subset F_4\\
F_4 &\subset E_6\\
E_6\times U(1) &\subset E_7\\
E_7\times SU(2) &\subset E_8
\end{aligned}
$$

Furthermore, each of the simple factors is, in Dynkin's terminology, of index 1 in the group on the right hand side. The index of a simple subgroup G in a simple Lie group H is defined as the ratio of the invariant inner product on $\mathfrak{h}$ to the invariant inner product on $\mathfrak{g}$ where the inner products are normalized to make the length of the highest root 2. By comparison with the discussion at the beginning of this section, this is just the degree of the induced map on π_3 (which incidentally shows immediately that it is an integer). Hence the inclusions of the simple factors preserve the index k of a bundle on S^4.

The second difficulty is in the definition of the space of moduli, as we have really only considered moduli of irreducible connections in the large. However, for the parameter count we only need the local moduli spaces. We note that in the proof of theorem 6.1, $h^2 = 0$ whatever the holonomy group and this gives us an h^1-dimensional manifold of self-dual connections which is locally complete. The condition $h^0 = 0$ ensures local effectiveness.

Now if ω is an irreducible self-dual G-connection, and $G \subset H$ is a subgroup of maximal dimension, such that the associated H-connection has $H^0(\mathfrak{h}) = 0$, then we can apply the following argument: if the dimension of the local moduli space of the H-connections centred on ω is larger than the dimension for G-connections, there must exist an irreducible self-dual H-connection. Suppose not, then each sufficiently close H-connection has holonomy group of dimension dim G, and by the rigidity argument above is thus gauge-equivalent to a G-connection. But this contradicts the local effectiveness of the space of G-connections. To see whether $H^0(\mathfrak{h}) = 0$ is a mere group theoretic condition, for $H^0(\mathfrak{h})$, the Lie algebra of covariant constant sections of $\mathfrak{h}$, is naturally isomorphic to the Lie algebra of the centralizer $C(G)$ of G in H.

We use this argument first to show that there exist irreducible $\mathrm{Sp}(n)$ connections for $k \geqslant n$. We use induction: suppose it is true for $n-1$. Then if $k \geqslant n$, $k-1 \geqslant n-1$ and we have an irreducible $\mathrm{Sp}(n-1)\times\mathrm{Sp}(1)$ connection by taking the product of an $\mathrm{Sp}(n-1)$ connection of index $(k-1)$ and an $\mathrm{Sp}(1)$ connection of index 1. Since each inclusion is an isomorphism on π_3, this defines an $\mathrm{Sp}(n)$ connection of index k. Now count the parameters by using (8.1): for $\mathrm{Sp}(n-1)\times \mathrm{Sp}(1)$ connections we get $4n(k-1)-(n-1)(2n-1)+5$ which is less that the dimension of $\mathrm{Sp}(n)$ connections $4(n+1)k-n(2n+1)$ if $k>1$. To start the induction, the 't Hooft examples give $\mathrm{Sp}(1)\cong SU(2)$ connections for $k\geqslant 1$.

Similarly irreducible $\mathrm{Spin}(n)$ connections exist for $k>\frac{1}{4}(n-1)$, except for $n=4$ which is not simple anyway ($\mathrm{Spin}(4)\cong SU(2)\times SU(2)$). Here for $n<7$ we use the isomorphisms $\mathrm{Spin}(3)\cong SU(2)$, $\mathrm{Spin}(5)\cong \mathrm{Sp}(2)$, $\mathrm{Spin}(6)\cong SU(4)$ to obtain the formula from (8.1).

From the inclusion $\mathrm{Spin}(9)\subset F_4$ we get irreducible F_4 connections for $k\geqslant 3$ and from $F_4\subset E_6$ we get E_6 connections for $k\geqslant 3$.

We cannot use this argument however, for $SU(n)\subset SU(n+1)$ or $E_6\subset E_7$ since we have a 1-dimensional *centralizer* $U(1)$ and so $H^0(\mathfrak{h})\neq 0$ and we no longer have effectiveness. Nevertheless, the Kuranishi space still exists as a manifold, whose dimension by the index theorem is $p_1(\mathfrak{h})-\dim H+\dim C(G)$.

We use here a different argument: since $H^0(\mathfrak{h})$ is the space of solutions of an elliptic equation, either there is an open neighbourhood of ω in the Kuranishi space of H-connections on which $H^0(\mathfrak{h})$ is of constant rank and forms a vector bundle, or in every neighbourhood $H^0(\mathfrak{h})$ changes dimension. If it changes dimension, the holonomy group changes and by semi-continuity must be H. If $H^0(\mathfrak{h})$ is of constant rank, we can find a smooth one parameter family of connections ∇_t with $\nabla_0=\omega$ and $\nabla_t X_t=0$.

Differentiating with respect to t, and putting $t=0$, we have

$$\dot{\nabla}X_0+\nabla\dot{X}_0=0,$$

i.e.

$$\nabla\dot{X}_0+[\dot{\tau},X_0]=0.$$

In other words $[\dot{\tau},X_0]\in A^1(\mathfrak{h})$ defines the zero element in $H^1(\mathfrak{h})$ of the complex in theorem 6.1. Since we can find a family by integration such that $\dot{\tau}$ is arbitrary, we deduce that $H^0(\mathfrak{h})$ acts trivially on $H^1(\mathfrak{h})$ for the connection ω. Let us consider $SU(n)\subset SU(n+1)$ in the light of this comment.

The orthogonal complement $su(n)^{\perp}$ of the Lie algebra $su(n)$ in $su(n+1)$ decomposes into a trivial one dimensional factor, the Lie algebra of the centralizer $u(1)$ and an irreducible *complex* representation of $SU(n)$ on which $U(1)$ acts as scalar multiplication. Now $H^1(su(n+1))=H^1(su(n))\oplus H^1(su(n)^{\perp})$ for an $SU(n)$ connection ω and so if $\dim H^1(su(n)^{\perp})>1$, then $U(1)$ acts as scalar multiplication on a non-trivial subspace of $H^1(su(n+1))$, hence $\dim H^0(su(n+1))$ cannot be constant, and by our argument there must be an irreducible $SU(n+1)$ connection arbitrarily close to ω. Now from (8.1), $\dim H^1(su(n)^{\perp})>1$ if $k>\frac{1}{4}(2n+1)$, so irreducible $SU(n)$ connections exist if $k>\frac{1}{4}(2n-1)$.

In particular, irreducible $SU(3)$ connections exist for $k\geqslant 2$ and with the inclusion $SU(3)\subset G_2$, and the first argument G_2 connections exist for $k\geqslant 2$. A similar argument to $SU(n)$ gives irreducible E_7 connections for $k\geqslant 3$.

For $E_7 \times SU(2) \subset E_8$ we can take the product of an index $k\,E_7$ connection and an index 1 $SU(2)$ connection to obtain irreducible index $(k+1)$ E_8 connections for $k \geqslant 3$. However, if we consider the inclusion of the index 3 E_7 connection in E_8 we can apply the second argument for centralizer $SU(2)$ and deduce that a nearby connection has a larger holonomy group. From Dynkin, this must be $E_7 \times SU(2)$ or E_8, and if the former then we would obtain an irreducible E_7 connection of index less than 3. But this would give a value of $\dim H^1(e_8)$ under inclusion of

$$120k - 248 + 3 < 0 \quad \text{if} \quad k < 3.$$

Hence there are no irreducible E_7 connections for $k < 3$ and there exist irreducible E_8 connections for $k \geqslant 3$.

This non-existence argument can be extended to the other cases: if an irreducible G connection defines by inclusion an H-connection, then we can split the bundle $\mathfrak{h}$ into a direct sum $\mathfrak{g} \oplus \mathfrak{g}^{\perp}$ and

$$\dim H^1(\mathfrak{g}^{\perp}) = p_1(\mathfrak{h}) - \dim H + \dim C(G) - p_1(\mathfrak{g}) + \dim G \geqslant 0.$$

So if $p_1(\mathfrak{h}) - \dim H + \dim C(G) < p_1(\mathfrak{g}) - \dim G$ we get *non-existence* of irreducible G-connections. For this we need inclusions with simple groups on the left hand side. We use the following inclusions, each one an isomorphism on π_3

$$\begin{gathered} SU(n) \subset SU(n+1), \\ \mathrm{Spin}(n) \subset \mathrm{Spin}(n+1), \\ \mathrm{Sp}(n) \subset \mathrm{Sp}(n+1), \\ G_2 \subset \mathrm{Spin}\,7, \\ F_4 \subset E_6, \\ E_6 \subset E_7, \\ E_7 \subset E_8. \end{gathered}$$

For E_8, $p_1(\mathfrak{g}) - \dim G = 120k - 248$ which is negative for $k \leqslant 2$ and this is sufficient. Using the values from (8.1) we obtain the following non-existence results: for $\mathrm{Sp}(n)$ if $k < n$; for $\mathrm{Spin}(n)$ if $k < \frac{1}{4}n$ $(n \geqslant 7)$; for $SU(n)\, k < \frac{1}{2}n$; for G_2 if $k = 1$ and F_4, E_6, E_7, E_8 if $k \leqslant 2$. Putting the existence and non-existence results together and taking account of the fact that k is an integer, we obtain the following:

THEOREM 8.4. *There exist irreducible self-dual G-connections on S^4 of index k iff for*

$$\mathrm{Sp}(n)\; k \geqslant n \quad SU(n)\; k \geqslant \tfrac{1}{2}n \quad \mathrm{Spin}(n)\; k \geqslant \tfrac{1}{4}n \quad (\text{where } n \geqslant 7)$$

$$G_2\; k \geqslant 2, \quad F_4, E_6, E_7, E_8\; k \geqslant 3.$$

(The above theorem corrects the inequality in our previous note (Atiyah *et al.* 1977).)

COROLLARY. *If $k \geqslant \frac{1}{2}n$, there exist simple holomorphic vector bundles of rank n on $P_3(\mathbb{C})$ with $c_1 = 0$, $c_2 = k$.*

To prove this, we just lift the bundle from S^4 to $P_3(\mathbb{C})$ and use theorem 5.2.

9. The case $k = 1$

An immediate corollary of theorem 8.4 is that every self-dual G-connection on S^4 of index 1 reduces to an $SU(2)$ connection. The dimension of the space of moduli for $k = 1$ is 5, and we have already found one five-dimensional space of self-dual $SU(2)$ connections on S^4 with $k = 1$, namely the Riemannian connections on V_+ relative to the constant positive curvature metrics within the fixed conformal structure. These are all equivalent under the conformal group $SO(5,1)$, each with isometry group $O(5)$ and hence are parametrized by the five-dimensional hyperbolic space $SO(5,1)/O(5)$. What we do not know is whether this family is complete. We shall show this next by differential geometric means:

THEOREM 9.1. *Let ω be any self-dual $SU(2)$ connection on S^4 of index 1. Then ω is gauge-equivalent to the Riemannian connection of a metric of constant sectional curvature within the conformal structure.*

Proof. The space $\mathscr{A}^+$ of self-dual connections on P is acted upon by the group of diffeomorphisms of P which commute with the group action of G and act on S^4 by orientation preserving conformal maps. This group $\mathscr{G}_1$ is an extension

$$\mathscr{G} \to \mathscr{G}_1 \to SO_0(5,1) \text{ where } \mathscr{G} \text{ is the gauge group}$$

and so $\mathscr{A}^+/\mathscr{G} = \mathscr{M}$ is acted upon by $SO_0(5,1)$, the component of the identity of $SO(5,1)$. Consider the point $\overline{\omega} \in \mathscr{M}$ defined by ω and look at the isotropy subgroup $H \subset SO_0(5,1)$ of $\overline{\omega}$. Since $\dim \mathscr{M} = 5$ and $\dim SO(5,1) = 15$, $\dim H \geqslant 10$. Now the Pontrjagin form of a self-dual connection defines a non-negative density $|\Omega|^2 \in A^4$, and this is furthermore gauge-invariant, so $|\Omega|^2$ must be preserved by H. However, since $|\Omega|^2$ is non-vanishing on some open set $U \subset S^4$, it defines a metric within the conformal structure on U, preserved by H. But the isometry group of a metric on an n-manifold always has dimension not more than $\frac{1}{2}n(n+1)$ with equality only for a space of constant curvature (Kobayashi & Nomizu 1963) and hence $\dim H = 10$, $|\Omega|^2$ defines a metric of constant curvature on U which since $|\Omega|^2$ is finite must be a metric of constant *positive* curvature on S^4.

Take the connection on $\mathfrak{g} \otimes \Lambda^p$ induced by ω on the adjoint bundle $\mathfrak{g}$ and the Riemannian connection on Λ^p of a constant curvature metric obtained above. Now by the Bianchi identity $D_2 \Omega = 0 \in A^3(\mathfrak{g})$ and since Ω is self-dual $*\Omega = \Omega$ and hence $D_1^* \Omega = 0 \in A^1(\mathfrak{g})$. We may write this as $D\Omega = 0$ where D is the Dirac operator on $\mathfrak{g} \otimes \Lambda_+^2 \cong V_+ \otimes V_+ \otimes \mathfrak{g}$ and as in § 6 use the Weitzenböck formula:

$$\begin{aligned} 0 &= \int (\nabla\Omega, \nabla\Omega) + \int (C(K) \,.\, \Omega, \Omega) \\ &= \int (\nabla\Omega, \nabla\Omega) + \int \tfrac{1}{3} R(\Omega, \Omega) + ([\Omega, \Omega], \Omega), \end{aligned} \tag{9.2}$$

where R is the scalar curvature, and $([\Omega, \Omega], \Omega)$ is essentially the determinant of $\Omega \in \Gamma(\mathrm{Hom}\,(\mathfrak{g}, \Lambda_+^2))$ where the three-dimensional bundles $\mathfrak{g}$ and Λ_+^2 are given the volume forms of their respective metrics.

Now the metric on S^4 is defined in terms of Ω, and it follows that the algebraic term in Ω is homogeneous in Ω. Instead of computing it explicitly we note that it must be a *positive* multiple of

$$(\operatorname{tr} \Omega * \Omega)^{\frac{3}{2}} + a \det \Omega,$$

for some $a \in \mathbb{R}$, since $R > 0$. But we know that the Riemannian connection of the constant curvature metric is a solution of the self-dual equations, so we can determine a. For the Riemannian connection $\Omega: \Lambda^2_+ \to \Lambda^2_+$ is a scalar and is covariant constant, hence both terms in (9.2) vanish and

$$\lambda^3 3^{\frac{3}{2}} + a\lambda^3 = 0,$$

and so $a = -3^{\frac{3}{2}}$.

However, in general if A is a 3×3 matrix,

$$(\det A^*A)^{\frac{1}{3}} \leqslant \tfrac{1}{3}(\operatorname{tr} A^*A),$$

hence $|\det A| \leqslant \frac{1}{3}(\operatorname{tr} A^*A)^{\frac{3}{2}}$ and

$$(\operatorname{tr} \Omega^*\Omega)^{\frac{3}{2}} - 3^{\frac{3}{2}} \det \Omega \geqslant 0,$$

with equality iff $\Omega^*\Omega$ is a scalar. This positivity condition yields the vanishing theorem in (9.2), i.e. the curvature Ω must satisfy

$$\nabla \Omega = 0 \quad \text{and} \quad \Omega^*\Omega = \text{scalar}.$$

Since $\nabla\Omega = 0$ and $\int |\Omega|^2 > 0$, $\Omega^*\Omega$ is some non-zero constant and so by an appropriate choice of scale for the constant curvature metric, the map $\Omega: \mathfrak{g} \to \Lambda^2_+$ preserves the inner product structures on $\mathfrak{g}$ and Λ^2_+ and since $\nabla\Omega = 0$, preserves the connections. In other words, the curvature Ω itself is a gauge transformation taking ω to a constant curvature connection.

Alternatively, instead of using the vanishing theorem we can use a homogeneity argument, for if the connection ω is taken into a gauge equivalent connection under the action of $SO(5)$, then there is a unique action of Spin(5) on the principal bundle P, preserving the connection. This follows from the fact that the connection has no non-trivial automorphisms. Since the connection is now homogeneous, an algebraic argument reduces it to the standard case.

For $k = 2$, there is quite a choice of irreducible connections: Spin(n) for $n \leqslant 8$, G_2 and $SU(3)$. All of these groups actually sit inside Spin(8) inducing an isomorphism on π_3 and so in some sense might be considered as degenerate forms of the twenty-dimensional space of Spin(8) connections.

10. The Dirac operator

The methods of the infinitesimal calculation in theorem 6.1, namely the index theorem plus a vanishing theorem, may also be used to calculate the dimension of the null-space of the Dirac operator

$$D: \Gamma(V_+ \otimes E) \to \Gamma(V_- \otimes E)$$

with a coefficient bundle E which is assumed to have a self-dual connection. In this case we must assume that the base space X is a spin manifold.

The vanishing theorem for harmonic sections of $V_- \otimes E$ reduces to the vanishing theorem of Lichnerowicz (1963) for harmonic spinors: if the scalar curvature $R > 0$, then there are no harmonic spinors. In contrast with the vanishing theorem of (6.1), this makes no additional self-dual assumption about the underlying conformal structure.

The index $h^0 - h^1$ is then just h^0, and may be calculated by the index theorem (Atiyah & Singer 1968);

$$\begin{aligned} \operatorname{index} D &= \operatorname{ch}(E)\hat{\mathscr{A}}(X)[X] \\ &= \tfrac{1}{2}p_1(E) - \tfrac{1}{8}\tau \dim E. \end{aligned}$$

For the 4-sphere $\tau = 0$ and we are left with index $D = \frac{1}{2}p_1(E)$.

In particular, if we have a self-dual connection on a principal G-bundle P and an irreducible representation $\rho: G \to \operatorname{Aut} E$, we can compute the dimension of the harmonic self-dual spinors with coefficients in the associated vector bundle E.

To compute $p_1(E)$ for simple G, we can see what multiple of the Killing form is the invariant form $\operatorname{tr}\rho(X)^2$ on $\mathfrak{g}$. But this is just

$$(\dim E/\dim G)\, C(E),$$

where $C(E)$ is the value of the Casimir operator on E. Hence

$$\operatorname{index} D = (\dim E/\dim G)\, \tfrac{1}{2}C(E)\, p_1(\mathfrak{g}).$$

All these terms can be computed algebraically by using the Weyl dimension formula and the formula for the Casimir operator in terms of the highest weight w of the representation ρ. It leads to the following formula:

$$h^0(D) = \frac{2\langle w + 2\delta, w\rangle k}{\dim G\langle \theta, \theta\rangle} \prod_i \frac{\langle w + \delta, \theta_i\rangle}{\langle \delta, \theta_i\rangle},$$

where δ is half the sum of the positive roots θ_i and θ is the highest root, all inner products taken relative to the Killing form.

Example

If $G = SU(2)$, with basic weight x, then the representation with highest weight lx has dimension $l+1$ and Casimir operator $\langle (l+2)x, lx\rangle = \frac{1}{8}l(l+2)$. Hence

$$h^0(D) = \left(\frac{l+1}{3}\right)\frac{l(l+2)}{16}8k = \frac{l(l+1)(l+2)}{6}k.$$

Jackiw & Rebbi have also obtained this formula.

References

Atiyah, M. F. & Singer, I. M. 1968 The index of elliptic operators. III. *Ann. Math. Princeton* **87**, 546–604.

Atiyah, M. F., Hitchin, N. J. & Singer, I. M. 1977 Deformations of instantons. *Proc. natn Acad. Sci. U.S.A.* **74**, 2662.

Atiyah, M. F. & Ward, R. S. 1977 Instantons and algebraic geometry. *Commun. Math. Phys.* **55**, 117–124.

Bernard, C. W., Christ, N. H., Guth, A. H. & Weinberg, E. J. 1977 Instanton parameters for arbitrary gauge groups. *Phys. Rev.* D **16**, 2967–2977.

Blanchard, A. 1956 Sur les variétés analytiques complexes. *Ann. sci. Éc. norm. sup. Paris* **73**, 157–202.

Bott, R. 1956 An application of the Morse theory to the topology of Lie groups. *Bull. Soc. math. Fr.* **84**, 251–281.

Bourbaki, N. 1968 *Groupes et algèbres de Lie*, chs 4, 5 and 6. Paris: Hermann.

Dynkin, E. B. 1957 Semisimple subalgebras of semisimple Lie algebras. *Am. math. Soc. Transl. Ser.* 2, **6**, 111–245.

Fegan, H. D. 1976 Conformally invariant first order differential operators. *Q. J. Math.* **27**, 371–378.

Griffiths, P. A. 1966 The extension problem in complex analysis. II. Embeddings with positive normal bundle. *Am. J. Math.* **88**, 366–446.

Hartshorne, R. 1977 Stable vector bundles and instantons. *Commun. Math. Phys.* (in the press).

Jackiw, R., Nohl, C. & Rebbi, C. 1977 Conformal properties of pseudoparticle configurations. *Phys. Rev.* D **15**, 1642–1646.

Kobayashi, S. & Nomizu, K. 1963 *Foundations of differential geometry*, vol. 1. New York: Interscience.

Kobayashi, S. & Wu, H. H. 1970 On holomorphic sections of certain hermitian vector bundles. *Math. Ann.* **189**, 1–4.

Kodaira, K. 1962 A theorem of completeness of characteristic systems for analytic families of compact submanifolds of complex manifolds. *Ann. Math. Princeton* **75**, 146–162.

Kuranishi, M. 1965 New proof for the existence of locally complete families of complex structures. In *Proceedings of the Conference on Complex Analysis* (ed. A. Aeppli *et al.*), Minneapolis 1964, pp. 142–154. New York: Springer Verlag.

Lichnerowicz, A. 1963 Spineurs harmoniques. *C.r. hebd. Séanc. Acad. Sci., Paris.* A–B **257**, 7–9.

Milnor, J. 1963 Morse theory. *Ann. Math. Stud.* no. 51. Princeton University Press.

Newlander, A. & Nirenberg, L. 1957 Complex analytic coordinates in almost complex manifolds. *Ann. Math. Princeton* **65**, 391–404.

Penrose, R. 1975 Twistor theory, its aims and achievements. In *Quantum gravity, an Oxford symposium* (ed. C. J. Isham *et al.*). Oxford University Press.

Penrose, R. 1976 Nonlinear gravitons and curved twistor theory. *Gen. relativity gravitation* **7**, 31–52.

Richardson, R. W. Jr. 1967 A rigidity theorem for subalgebras of Lie and associative algebras. *Illinois J. Math.* **11**, 92–110.

Schwarz, A. S. 1977 On regular solutions of Euclidean Yang–Mills equations. *Phys. Lett.* **67**B, 172–174.

Singer, I. M. & Thorpe, J. A. 1969 The curvature of 4-dimensional Einstein spaces. In *Global Analysis, Papers in honor of K. Kodaira* (ed. D. C. Spencer & S. Iyanaga), pp. 355–365. Princeton University Press.

Yau, S. T. 1977 On Calabi's conjecture and some new results in algebraic geometry. *Bull. Am. math. Soc.* (in the press).

Footnote to page 461. *Fibre* should be interpreted here as including the first formal neighbourhood. The trivialization (5.3) holds also in this sense because the co-normal bundle of a fibre has vanishing H^0 and H^1; any *section* should then be a *basis of sections*.