INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONE: 2240-1
CABLE: CENTRATOM - TELEX 460392 - I

SMR.304/15

COLLEGE

ON

GLOBAL GEOMETRIC AND TOPOLOGICAL METHODS IN ANALYSIS

(21 November - 16 December 1988)

THE YANG-MILLS EQUATIONS.

J.D.S. Jones
Mathematics Institute
University of Warwick
Coventry CV4 7AL
U.K.

These are preliminary lecture notes, intended only for distribution to participants.

THE YANG-MILLS EQUATIONS

by

John D.S. JONES

Mathematics Institute
University of Warwick
Coventry, CV4 7AL
ENGLAND

December 1988

INTRODUCTION

These notes are the notes for the course on the Yang-Mills equations. The purpose of the course is to introduce the Yang-Mills equations and to describe their "instanton", or point-like, solutions. There are three basic references which cover everything I have to say and quite a lot more:

M.F. Atiyah, Geometry of Yang-Mills Fields, Lezione Fermiane, Scuola Normale Superiore, Pisa, 1979.

D.S. Freed and K.K. Uhlenbeck, Instantons and Four-Manifolds, MSRI Publications, Springer-Verlag, 1984

H. Blaine Lawson, Jr., The Theory of Gauge fields in Four Dimensions, Regional Conference Series in Mathematics 58, AMS 1985.

1 Yang-Mills theory on $\mathbb{R}^4$

To begin with I will describe Yang-Mills theory on $\mathbb{R}^4$; for most of the time we will think of $\mathbb{R}^4$ with its usual Euclidean metric but it will sometimes be convenient to vary the metric. On $\mathbb{R}^4$ it is possible to give the key points in the theory and to describe the point-like solutions to the Yang-Mills equations with the minimum technical apparatus. The rest of the notes will be devoted to developing the full global versions of these constructions.

In Yang-Mills theory there is a ***gauge group*** or ***structure group*** G. For definiteness I will take this group to be either the unitary group U(n) or the special unitary group SU(n). Recall that these groups are defined as follows:

$$U(n) = \left\{ A \in GL_n(\mathbb{C}) \mid A^*A = AA^* = I \right\}$$
$$SU(n) = \left\{ A \in U_n \mid \det A = 1 \right\}.$$

I am using standard notation, in particular $GL_n(\mathbb{C})$ is the group of invertible, n×n matrices over $\mathbb{C}$, A^* is the conjugate transpose of A, I is the identity matrix and det is the determinant. Let $\mathfrak{g}$ be the Lie algebra of G. In our special cases these Lie algebras are

$$\mathfrak{u}(n) = \left\{ A \in M_n(\mathbb{C}) \mid A^* = -A \right\}$$
$$\mathfrak{su}(n) = \left\{ A \in \mathfrak{u}(n) \mid \mathrm{Trace}\, A = 0 \right\}.$$

A ***G-connection*** or ***G-potential*** on $\mathbb{R}^4$ is a smooth $\mathfrak{g}$-valued 1-form A on $\mathbb{R}^4$, i.e.

$$A = \sum_{\mu=1}^{4} A_\mu dx_\mu, \qquad A_\mu : \mathbb{R}^4 \to \mathfrak{g}.$$

where the A_μ are smooth functions. A ***gauge transformation*** is a smooth function

$$g : \mathbb{R}^4 \to G.$$

These gauge transformations act on connections by the formula

$$g^*(A) = g^{-1}Ag + g^{-1}dg.$$

The simplest way to make sense of this formula is to remember that both g and A are matrix valued functions. Two connections A_1 and A_2 are ***gauge equivalent*** if there is a gauge transformation g such that $g^*A_1 = A_2$. The ***curvature*** or ***gauge field*** of this connection is defined by

$$F_A = dA + A\wedge A$$

where $A\wedge A$ is defined by using the combination of matrix product and the exterior product of forms. It is not difficult to compute F_A explicitly in terms of coordinates

$$F = \sum_{\mu<\nu} F_{\mu\nu}\, dx_\mu \wedge dx_\nu$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu].$$

Here I am using the notation ∂_μ for the partial derivative $\partial/\partial x_\mu$ and $[\,,\,]$ is the usual commutator, or Lie bracket, of matrices $[X,Y] = XY - YX$. In particular note that this formula shows that F is a $\mathfrak{g}$-valued 2-form. It is easy to work out how the curvature changes under a gauge transformation of the connection,

$$F_{g^*A} = g^{-1}F_A g\,.$$

Next we need the ***Hodge star operator***. In general this is an operator

$$* : \Omega^p(M) \to \Omega^{n-p}(M)$$

where M is an n-dimensional oriented manifold equipped with a metric. This operator is characterised by two facts:

The Yang-Mills equations

(1) It is linear over functions

(2) $\alpha\wedge(*\beta) = \langle\alpha, \beta\rangle\, \mathrm{vol}$ $\qquad \alpha, \beta \in \Omega^p(M)$.

Here the inner product is the inner product on forms induced by the metric and vol is the volume form determined by the metric and the orientation.

For example, on $\mathbb{R}^4$ with its usual metric and orientation the star operator on 2-forms is given by the following formulas:

$$*(dx_1\wedge dx_2) = dx_3\wedge dx_4$$
$$*(dx_1\wedge dx_3) = -\,dx_2\wedge dx_4$$
$$*(dx_1\wedge dx_4) = dx_2\wedge dx_3$$
$$*(dx_2\wedge dx_3) = dx_1\wedge dx_4$$
$$*(dx_2\wedge dx_4) = -\,dx_1\wedge dx_3$$
$$*(dx_3\wedge dx_4) = dx_1\wedge dx_2$$

By way of contrast, and to illustrate how the star operator depends on the metric, one can check that if we compute the star operator using the Lorentz metric $dx_1^2 + dx_2^2 + dx_3^2 - dx_4^2$ then it is given by the following formulas

$$*(dx_1\wedge dx_2) = dx_3\wedge dx_4$$
$$*(dx_1\wedge dx_3) = -\,dx_2\wedge dx_4$$
$$*(dx_1\wedge dx_4) = -\,dx_2\wedge dx_3$$
$$*(dx_2\wedge dx_3) = dx_1\wedge dx_4$$
$$*(dx_2\wedge dx_4) = dx_1\wedge dx_3$$
$$*(dx_3\wedge dx_4) = -\,dx_1\wedge dx_2.$$

Associated to the connection is its ***covariant derivative*** operator

$$D_A = d + A.$$

This operator acts on functions, or more generally forms which take their values in a representation of G and produces forms with values in the same representation. For example if we use the standard action of U(n) or SU(n) on $\mathbb{C}^n$ and $f : \mathbb{R}^4 \to \mathbb{C}^n$ is a smooth function then $D_A f$ is the $\mathbb{C}^n$-valued 1-form

$$D_A f = df + Af ;$$

where

$$Af = \sum_{\mu=1}^{4} (A_\mu f) dx_\mu$$

As another example suppose that F is a $\mathfrak{g}$-valued 2-form, then

$$D_A F = dA + [A \wedge F].$$

where the symbol $[A \wedge F]$ means that we use the combination of the Lie bracket [,] of matrices and the exterior product $\wedge$ of forms. I can now state the ***Yang-Mills equations***:

$$D_A F_A = 0$$
$$D_A(*F_A) = 0.$$

The first of these equations is an identity, the ***Bianchi identity*** which is always satisfied when F_A is the curvature of the connection A. I will use the term ***Yang-Mills connection*** for a solution of these equations. The Yang-Mills connections are the Euler-Lagrange equations of the ***Yang-Mills functional***

$$Y(A) = \int_{\mathbb{R}^4} F_A \wedge (*F_A) = \int_{\mathbb{R}^4} |F_A|^2 \text{ vol.}$$

I will prove this in the more general global context later in these notes.

There is a simple analogy which might help to explain some of the significance of the Yang-Mills equation. If we replace D_A by the ordinary exterior derivative operator on forms, and replace F_A by a 2-form f we get the equations

$$df = 0, \quad d(*f) = 0$$

and a form which satisfies these equations is a ***harmonic form***. Furthermore these equations are the Euler-Lagrange equations of the functional

$$\int f \wedge (*f) = \int |f|^2 \text{ vol.}$$

So a Yang-Mills connection is one whose curvature is "harmonic" with respect to its own covariant derivative and this gives them a natural appeal.

These equations are a system of second order non-linear partial differential equations equations. There is a simpler, and for our present purposes, more important, system of equations whose solutions are Yang-Mills connections. These are the ***self dual Yang-Mills equations*** and the ***anti self dual Yang-Mills equations***:

$$*F_A = F_A \qquad \text{(self dual)}$$
$$*F_A = -F_A \qquad \text{(anti-self dual).}$$

In view of the Bianchi identity it is clear that a solution of these equations is automatically a Yang-Mills connection. A connection which satisfies the self dual equations will be called a ***self dual connection*** and one which satisfies the anti self dual equations will be called an ***anti self dual connection***.

Of course the Yang-Mills functional and the full Yang-Mills equations make sense in any dimension. However the self dual equations arespecial to dimension 4 since it is only in dimension 4 that the $*$ operator takes 2-forms to 2-forms and so it only makes sense to look for connections with self dual curvature in 4-dimensions.

There is another special feature of 4-dimensions; Yang-Mills theory is conformally invariant in dimension 4. This means that if we change the metric h on $\mathbb{R}^4$ to a new metric h' which is conformally equivalent to h, i.e. $h' = \lambda^2 h$ where λ is a real valued function, then the Yang-Mills functional Y, which looks as though it ought to depend on the metric, is in fact unchanged. To see this it is not difficult to check that the inner product on 2-forms given by the metric h and metric h' is $\lambda^{-4}\langle\,,\rangle$ where $\langle\,,\rangle$ is the inner product on 2-forms given by the metric h and the volume form in the new metric is given by λ^4 vol. So the star operator remains unchanged and also the Yang-Mills functional Y is unchanged. This conformal invariance is a very important feature of Yang-Mills theory in dimension 4.

An interesting special case to look at is the case where G is the circle group U(1) and we work over $\mathbb{R}^4$ with the Lorentz metric. Since the Lie algebra of the circle is $\mathbb{R}$ we can identify a connection with an ordinary 1-form on $\mathbb{R}^4$

$$A = A_1dx_1 + A_2dx_2 + A_3dx_3 + A_4dx_4.$$

Since the Lie algebra of the circle is commutative, that is the bracket [,] is identically zero, the components of the curvature become

$$F_{ij} = \partial_i A_j - \partial_j A_i$$

and the Bianchi identity comes out as

$$\partial_i F_{jk} + \partial_j F_{ki} + \partial_k F_{ij} = 0.$$

A neat way to write out the Yang-Mills equation is as follows:

$$\partial_1 F_{i1} + \partial_2 F_{i2} + \partial_3 F_{i3} - \partial_4 F_{i4} = 0.$$

If we now make the following substitutions

$$H_1 = \partial_2 A_3 - \partial_3 A_2 = F_{23}$$
$$H_2 = -\partial_3 A_1 + \partial_1 A_3 = -F_{13}$$
$$H_3 = \partial_1 A_2 - \partial_2 A_1 = F_{12}$$
$$E_1 = \partial_1 A_4 - \partial_4 A_1 = F_{14}$$
$$E_2 = \partial_2 A_4 - \partial_4 A_2 = F_{24}$$
$$E_3 = \partial_3 A_4 - \partial_4 A_3 = F_{34}$$
$$x_1 = x,\ x_2 = y,\ x_3 = z,\ x_4 = t$$

then we get Maxwell's equations (in vacuo)

$$\text{curl } E = -\partial H/\partial t, \quad \text{div } H = 0$$
$$\text{curl } H = -\partial E/\partial t, \quad \text{div } E = 0.$$

2 *The basic instanton on $\mathbb{R}^4$*

Now I will describe the basic SU(2)-instanton, i.e self dual solution to the Yang-Mills equations on $\mathbb{R}^4$ with structure group SU(2). This is easiest to describe using the formalism of quaternions and so I will begin by explaining some of the basic properties of quaternions. The algebra of ***quaternions*** $\mathbb{H}$ is the algebra over $\mathbb{R}$, with unit 1, generated by i, j, k subject to the relations

$$i^2 = j^2 = k^2 = -1$$
$$ij = -ji = k,\ jk = -kj = i,\ ki = -ik = j.$$

The general quaternion is of the form

$$x = x_1 + ix_2 + jx_3 + kx_4 \qquad x_i \in \mathbb{R}$$

and the product is defined using the above formulae. The ***conjugate*** of x is given by

$$\bar{x} = x_1 - ix_2 - jx_3 - kx_4$$

and it is easy to check that

$$|x|^2 = x\bar{x} = \bar{x}x = \sum_{\mu=1}^{4} x_\mu^2$$

where $|x|$ is the length of the vector x. It follows that provided $x \neq 0$

$$x \frac{\bar{x}}{|x|^2} = \frac{\bar{x}}{|x|^2} x = 1$$

so $\mathbb{H}$ is a skew-field, i.e. the non-commutative analogue of a field.

The imaginary part of x is

$$\operatorname{Im} x = ix_2 + jx_3 + kx_4$$

and we will denote the space of purely imaginary quaternions by $\operatorname{Im}\mathbb{H}$.

We can identify the algebra of quaternions with a subalgebra of the algebra of 2×2 complex matrices as follows:

$$1 \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad i \to \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \quad j \to \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad k \to \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$$

Under this identification $SU(2)$ gets identified with the space of unit quaternions, i.e. those quaternions x with $|x| = 1$ and in particular this shows that topologically $SU(2)$ is the 3-sphere S^3. The Lie algebra $\mathfrak{su}(2)$ gets identified with the space of purely imaginary

quaternions, that is all quaternions of the form $xi + yj + zk$. In particular, when we identify $\mathbb{R}^4$ with $\mathbb{H}$, an $SU(2)$-connection is given by

$$A = \sum_{\mu=1}^{4} A_\mu(x) dx_\mu$$

where now the $A_\mu : \mathbb{R}^4 \to \operatorname{Im}\mathbb{H}$ are smooth functions.

Just as in the theory of complex variables where one uses the complex differentials

$$dz = dx + idy, \qquad d\bar{z} = dx - idy,$$

we shall consider the quaternion differential

$$dx = dx_1 + idx_2 + jdx_3 + kdx_4$$

and its conjugate

$$d\bar{x} = dx_1 - idx_2 - jdx_3 - kdx_4$$

The use of the quaternion differentials is particularly well-suited to the study of self duality since a routine computation shows that

$$dx \wedge d\bar{x} = -2\left\{ i(dx_1 \wedge dx_2 + dx_3 \wedge dx_4) + j(dx_1 \wedge dx_3 + dx_4 \wedge dx_2) + k(dx_1 \wedge dx_4 + dx_2 \wedge dx_3) \right\}$$

and we see that the coefficients of i,j,k are a basis of the self dual 2-forms. Therefore, this form $dx \wedge d\bar{x}$ is a self dual $su(2)$ valued 2-form. Similarily the coefficients of $d\bar{x} \wedge dx$ give a basis for the anti self dual 2-forms and $d\bar{x} \wedge dx$ is an anti self dual $su(2)$ valued 2-form.

If $f : \mathbb{H} \to \mathbb{H}$ is any smooth function then the expression

$$A(x) = \mathrm{Im}\,(f(x)dx) = \frac{1}{2}\left\{f(x)dx - d\bar{x}\,\overline{f(x)}\right\}$$

is an SU(2)-connection. Here these expressions are computed formally starting from the formulas .

$$f(x) = f_1(x)dx_1 + if_2(x)dx_2 + jf_3(x)dx_3 + kf_4(x)dx_4$$
$$dx = dx_1 + idx_2 + jdx_3 + kdx_4$$

and the usual rules for multiplying quaternions and differential forms. Now we compute the curvature $F = dA + A\wedge A$ of this connection; it is given by

$$F = \mathrm{Im}\left\{ df\wedge dx + fdx\wedge fdx \right\}$$

We are now in a position to describe the basic instanton on $\mathbb{R}^4$. In fact it is slightly easier to begin by constructing the basic anti self dual connection on $\mathbb{R}^4$. Consider the SU(2) connection A defined by the formula

$$A(x) = \mathrm{Im}\left\{\frac{\bar{x}\,dx}{1+|x|^2}\right\} = \frac{1}{2}\left\{\frac{\bar{x}\,dx - d\bar{x}\,x}{1+|x|^2}\right\}$$

It is not difficult to read off explicit formulas for the components A_μ from this quaternion notation; for example

$$A_1(x) = \frac{-ix_2 - jx_3 - kx_4}{1+|x|^2} \quad , \quad A_2(x) = \frac{ix_1 - jx_4 + kx_3}{1+|x|^2}$$

We now compute the curvature of this potential as above,

$$F = \mathrm{Im}\left\{ \frac{d\bar{x}\wedge dx}{1+|x|^2} + \bar{x}\,d((1+|x|^2)^{-1} + \frac{\bar{x}\,dx\wedge\bar{x}\,dx}{(1+|x|^2)^2} \right\}$$

Now if we write $|x|^2 = x\bar{x}$, the middle term in this expression becomes

$$-\frac{\bar{x}\,dx\wedge\bar{x}\,dx}{(1+|x|^2)^2} - \frac{\bar{x}x\,d\bar{x}\wedge dx}{(1+|x|^2)^2}$$

and when we substitute this in the formula for F we get

$$F = \frac{d\bar{x}\wedge dx}{(1+|x|^2)^2}$$

Therefore from the above remarks relating quaternion differentials and self duality we see that F is anti self dual, that is $*F = -F$ and therefore the original connection A is an anti self dual connection. This is the ***basic anti-instanton*** i.e. anti self dual connection on $\mathbb{R}^4$.

Clearly if we interchange the roles of x and $\bar{x}$ we will get a self dual connection A with curvature F given by the following formulas

$$A = \mathrm{Im}\left\{\frac{x\,d\bar{x}}{1+|x|^2}\right\} \quad , \quad F = \frac{dx\wedge d\bar{x}}{(1+|x|^2)^2}$$

This is the ***basic instanton*** on $\mathbb{R}^4$.

Now we will use the earlier remark that the self dual and anti self dual equations are conformally invariant to construct many more solutions from these two basic instantons. First note that the basic instantons are invariant under the transformations

$$x \to axb$$

where a and b are unit quaternions. These transformations generate the group SO(4) acting on $\mathbb{R}^4$ in the usual way so the basic instanton is invariant under SO(4). The other simple natural transformations to use are $T_{\lambda,a}$

$$x \to \lambda(x-a);$$

since these are conformal transformations they transform instantons to instantons. We will use the notation $A^\lambda(a)$ for the instanton $T_{\lambda,a}{}^*A$, and $F^\lambda(a)$ for its curvature, explicitly

$$A^\lambda(a) = \mathrm{Im}\left\{\frac{\lambda^2(x-a)\,d\bar{x}}{1+\lambda^2|x-a|^2}\right\}, \quad F^\lambda(a) = \frac{\lambda^2\,dx\wedge d\bar{x}}{(1+\lambda^2|x-a|^2)^2}\ .$$

Note that the pointwise norm of the curvature of attains its maximum value $\lambda^2|dx\wedge d\bar{x}|$ when $x = a$ and, since the pointwise norm of the curvature is a gauge invariant quantity, it follows that for distnct pairs (a,λ) the self dual connections $A^\lambda(a)$ cannot be gauge equivalent.

Now we turn to the main task of these notes, the global formulation of this Yang-Mills theory and the project of taking these solutions $A^\lambda(a)$ and glueing them into a general 4-manifold, or more generally superimposing k such solutions at distinct points in the manifold.

3 Covariant derivatives and connections and curvature

I will begin the formulation of the global Yang-Mills theory with the gobal theory of connections. I will take the analytic approach to the theory of connections, so a connection will be defined by its covariant derivative operator. Let E be a smooth vector bundle over a smooth manifold M. Then a smooth section of E is a smooth function $s : M \to E$ such that $\pi\circ s = 1$, where π is the projection of the bundle E; so $s(x) \in E_x$ where E_x is the fibre of E at the point $x \in M$. We treat sections of E as generalized functions, in the sense that s(x) is a

function of x but it takes its values in the variable vector space E_x, and we try to differentiate these sections.

I will use the notation

$$C^\infty(E) = C^\infty(M;E)$$

for the vector space of smooth sections of E and I will write

$$\Omega^p(E) = \Omega^p(M;E)$$

for the space $C^\infty(E\otimes\Lambda^pT^*)$ where T^* is the cotangent bundle of M; this is the space of differential forms on M with values in the bundle E. At a point $x \in M$ an element $\omega \in \Omega^p(E)$ is an alternating multi-linear map

$$\omega_x : T_x \times \ldots \times T_x \to E_x$$

where T_x is the tangent space to M at x and there are p factors. Locally every element of $\Omega^p(E)$ can be written as a linear combination of elements of the form $s\otimes\alpha$ where s is a section of E and α is a p-form.

A connection will be defined to be the analogue, for sections of E, of the total derivative operator $d : C^\infty(M) \to \Omega^1(M)$; recall that locally

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\,dx_i\ .$$

A ***connection*** on E is a linear map

$$\nabla : C^\infty(E) \to \Omega^1(E)$$

which satisfies the following form of Leibnitz rule

$$\nabla(fs) = f(\nabla s) + s \otimes df$$

where $s \in C^\infty(E)$ is a smooth section of E and $f \in C^\infty(M)$ is a function on M. We often refer to ∇ as the ***covariant derivative***; from this point of view there is no distinction between the connection and the covariant derivative and the two terms will be used interchangeably.

A choice of a local trivialisation of the bundle E will be called a ***local frame***, or ***local gauge*** for E. Such a local frame is given by $k = \dim E$ pointwise linearly independent sections $e_1, \dots, e_k$ of E defined on an open set U. Using a local gauge for E we can express the connection as a matrix of 1-forms on U as follows. Let s be a section of E, defined over U, then $s = s_1 e_1 + \dots + s_k e_k$ where the s_i are functions. From Liebnitz rule we get the following formula for ∇s:

$$\nabla s = \sum_{i=1}^{k} s_i \nabla e_i + e_i \otimes ds_i .$$

Now write

$$\nabla e_i = \sum_{j=1}^{k} e_j \otimes a_{ji}$$

where the a_{ji} are 1-forms; then we get the formula

$$\nabla s = ds + As$$

where A is the matrix of one forms $A = (a_{ij})$ defined on U. This matrix A is the ***connection matrix*** determined by the gauge $e = (e_1, \dots, e_k)$. Sometimes, when it is necessary to emphasize the dependence of A on e, we will write $A = A(e)$

We must work out how the connection matrix depends on the frame e. Let $f = (f_1, \dots, f_m)$ be a new frame and let g be the transition function, or the gauge transformation from the frame e to the frame f. So g is a matrix valued function and;

$$e_i = \sum_{j=1}^{m} g_{ji} f_j .$$

Then it is straightforward to check that

$$A(e) = g^{-1} A(f) g + g^{-1} dg$$

So a connection on E is given locally by a matrix of 1-forms A which transforms by the above formula under the transition functions of the bundle E and the corresponding covariant derivative is the operator $d + A$. As these notes progress we will encounter many examples of connections.

There are several constructions which are suggested by the analogy between the covariant derivative of a section of E and the total derivative of a function. Suppose that V is a vector field on M; then we define an operator $\nabla_V : C^\infty(E) \to C^\infty(E)$ by setting

$$(\nabla_V s)(x) = ((\nabla s)(x))(V(x)).$$

We will refer to the operator ∇_V as ***covariant differentiation in the direction V***. In terms of these operators ∇_V Liebnitz rule can be formulated as follows:

$$\nabla_V (fs) = (Vf)s + f(\nabla_V s).$$

where Vf is the operation of vector field V on the function f. We can get a local formula for ∇ in terms of the ∇_V

$$\nabla s = \sum_{i=1}^{n} (\nabla_i s)\otimes dx_i$$

where we have chosen local coordinates $x = (x_1, \dots, x_n)$ for M and ∇_i is the covariant derivative in the direction of the i-th coordinate vector field. These operators ∇_i which are locally defined operators on sections of E are analogous to the partial derivatives $\partial_i = \partial/\partial x_i$, which are locally defined operators on functions on M.

The connection allows us to define parallel transport along a smooth curve $\gamma : [0,1] \to M$. Suppose that s is a section of E along γ, (i.e $s(t) \in E_{\gamma(t)}$ for all $t \in [0,1]$), then s is ***parallel*** along γ if

$$\nabla_{\gamma'} s = 0$$

where γ' is the velocity vector of γ. There is the usual uniqueness theorem; if we are given a curve γ in M and a vector $e \in E_{\gamma(0)}$ then there is a unique section s of E along γ which is parallel and has the property that $s(0) = e$. The ***parallel transport*** operator

$$P_\gamma : E_{\gamma(0)} \to E_{\gamma(1)}$$

is defined by $P_\gamma(e) = s(1)$ where s is the unique parallel section of E along γ with the property that $s(0) = e$.

Next we analyse the space of all connections $\mathcal{A}(E)$ on E. Let End(E) be the vector bundle whose fibre at the point $x \in M$ is the space $End(E_x)$, i.e. the space of all linear maps from E_x to itself. Note that a section of End(E) defines a linear map of vector bundles $E \to E$. Therefore, if A is an element of $\Omega^1(End(E))$ it defines an operator

$$A : C^\infty(E) \to \Omega^1(E).$$

This operator is linear over functions (i.e $A(fs) = fA(s)$) and it is a standard exercise in the

theory of vector bundles to check that any linear map $C^\infty(E) \to \Omega^1(E)$ which is linear over functions arises from some $A \in \Omega^1(End(E))$.

Now suppose we are given a connection ∇ and an element $A \in \Omega^1(End(E))$; then, since A is linear over functions, the operator

$$\nabla + A : C^\infty(E) \to \Omega^1(E)$$

satisfies Liebnitz rule and is a connection. Conversely given two connections ∇_0 and ∇_1 on E the difference

$$\nabla_1 - \nabla_0 : C^\infty(E) \to \Omega^1(E)$$

is linear over functions and it follows that there exists a unique $A \in \Omega^1(End(E))$ such that $\nabla_1 = \nabla_0 + A$. A sophisticated way of saying the same thing is as follows.

Lemma The space $\mathcal{A}(E)$ of all connections on E is an affine space associated to the vector space $\Omega^1(End(E))$.

This means that if we pick a connection ∇ on E then we may identify $\Omega^1(End(E))$ with $\mathcal{A}(E)$ using the mapping $A \to \nabla + A$. But there is no preferred connection on E, that is the space $\mathcal{A}(E)$ has no natural origin.

If the bundle has extra structure then this extra structure often determines a canonical connection, an origin for the space of connections, but the connection will depend on the extra structure. For example if the bundle E is trivial and we are given an isomorphism t of E with the product bundle then we can define a connection on E by the formula $\nabla_t = t^{-1} \circ d \circ t$. This connection does depend on the choice of t. As another example a Riemannian structure on X determines a canonical connection on TX, see the notes for the course on Differential Geometry, but this connection depends on the Riemannian structure.

It is not difficult to show that every vector bundle admits a connection. We will give two ways to construct a connection on E. The first way is based on the observation that if f is a smooth function on M and ∇_0, ∇_1 are connections on E then $f\nabla_0 + (1-f)\nabla_1$ is also a connection. The trivial bundle admits a connection, for example the total derivative d, and using local trivialisations of E we get locally defined operators on sections of E. The above observation allows us to piece these locally defined operators together, using a partition of unity, to obtain a global operator. I will now give this construction precisely.

Pick a covering of M by open sets U_i and local trivialisations t_i of the bundle $E|_{U_i}$. Pick a partition of unity φ_i subordinate to the covering U_i. Let

$$\nabla_i : C^\infty(E) \to C^\infty(E|_{U_i})$$

be the operator given by

$$\nabla_i = t_i^{-1} \circ d \circ t_i \circ r_i$$

where $r_i : C^\infty(E) \to C^\infty(E|_{U_i})$ is given by restricting sections of E to the open set U_i, $t_i : C^\infty(E|_{U_i}) \to C^\infty(U_i; \mathbb{R}^k)$ is the isomorphism given by the local trivialisations and d is the total derivative operator acting on vector valued functions. We now get an operator

$$\varphi_i\nabla_i : C^\infty(E) \to C^\infty(E)$$

by extending $\varphi_i\nabla_i(s)$ to be zero outside U_i. It is easy to check that

$$\nabla = \sum_i \varphi_i\nabla_i$$

is a connection on E.

The second general method of constructing connections is to use an embedding i of

E in a product bundle $M\times\mathbb{C}^N$ (or $M\times\mathbb{R}^N$ if the bundle is real). It is a standard fact in the theory of vector bundles that any bundle over a compact space can be embedded as a sub-bundle of a product bundle. Let $p : M\times\mathbb{C}^N \to E$ be the orthogonal projection onto E. We now define the operator

$$\nabla : C^\infty(E) \to C^\infty(E\otimes T^*) = \Omega^1(E)$$

to be the composition

$$C^\infty(E) \xrightarrow{i} C^\infty(M; \mathbb{C}^N) \xrightarrow{d} C^\infty(\mathbb{C}^N\otimes T^*) \xrightarrow{p\otimes 1} C^\infty(E\otimes T^*)$$

It is straight forward to check that ∇ is a connection on E.

We can represent this connection by a matrix valued 1-form on M without using a gauge for the bundle E, but instead by working in the ambient space $\mathbb{C}^N$ and producing an $N\times N$ matrix rather than a $k\times k$ matrix where $k = \dim E$. Let $q = 1-p$ be the complementary projection and take the $M_N(\mathbb{C})$-valued 1-form

$$B = qdq.$$

A function $f : M \to \mathbb{C}^N$ is actually section of E, if and only if $pf = f$, or $qf = 0$. If we now compute the covariant derivative operator $\nabla_B = d + B$ acting on sections of E, using the fact that $0 = d(qf) = (dq)f + qdf$, we see that

$$\nabla_B f = df + (qdq)f = df - q^2 f = pdf = \nabla f$$

where ∇ is the covariant derivative operator acting on sections of E defined above. Therefore ∇_B extends ∇ to all $\mathbb{C}^N$ valued functions on M.

Now if we differentiate the equation $q^2 = q$ we get $(dq)q + qdq = dq$ and so

$$B^2 = qdq \wedge qdq = q(dq - qdq) \wedge dq = 0.$$

Hence the field $F_B = dB + B \wedge B$ is just

$$F_B = dq \wedge dq.$$

If we now restrict this field to the bundle E, we get the curvature F of ∇ is given by

$$F = p\, dp \wedge dp.$$

This is a particularly useful way of constructing connections.

As a particular example of this construction note that if X is a submanifold of $\mathbb{R}^N$ with the induced Riemannian metric then we get a natural embedding of TX in the trivial bundle $X \times \mathbb{R}^N$ which preserves the metric. The above construction gives the standard covariant derivative operator on vector fields and we recover the usual definition of the Levi-Civita connection on TX.

Now suppose the bundle E is a complex bundle and it has a hermitian metric $< , >$, in other words the structure group of the bundle is U(n). Then a connection ∇ on E is ***hermitian*** or a ***U(k) -connection*** if

$$d< s,t > = < \nabla s, t > + < s, \nabla t >.$$

If we pick an orthonormal local gauge for E, that is $k = \dim E$ pointwise linearly independent sections of E defined on some open set U $e_1, \ldots, e_k$ such that

$$< e_i, e_j > = \delta_{ij}$$

then the matrix of the connection in this gauge satisfies $A^* = -A$ so this is a U(k)-connection in the sense described in Section 1. If the structure group is SU(k) then in addition to the requirement that $d< s,t > = < \nabla s, t > + < s, \nabla t >$ we will also require that

$$\nabla(e_1 \wedge \ldots \wedge e_k) = 0$$

if $e_1, \ldots, e_k$ is an orthonormal local gauge. It is easy to see that this condition means that the matrix of ∇ determined by a suitable local gauge for E satisfies $A^* = -A$ and Trace $A = 0$ so this is an SU(k)-connection as described in Section 1.

If the bundle E is embedded as a sub-bundle of $M \times \mathbb{C}^N$ and the metric on the bundle is the same as the metric induced by the standard hermitian inner product on $\mathbb{C}^N$ and we now use the method described above to construct a connection on E then it is straightforward to check that we get a U(k)-connection on E.

Now we turn to the definition of the curvature of a connection. Unlike the partial derivatives ∂_i the operators ∇_i do not commute with each other and this is a reflection of the global twisting of the bundle. To define the curvature, first extend the covariant derivative operator ∇ to an operator

$$d^{\nabla} : \Omega^p(E) \to \Omega^{p+1}(E)$$

as follows: Locally d^{∇} is given by the formula

$$d^{\nabla}(s \otimes \omega) = \nabla(s) \wedge \omega + s \otimes d\omega.$$

and globally it is given by the natural extension of the formula for the exterior derivative of a differential form to sections of $E \otimes \Lambda^p T^*$ that is

$$d^{\nabla}\omega(V_0, \ldots, V_p) = \sum_{i=0}^{p} (-1)^i \nabla_{V_i}(V_0, \ldots, V_{i-1}, V_{i+1}, \ldots, V_p).$$

We now have the generalised de Rham sequence

$$\Omega^0(E) \to \Omega^1(E) \to \Omega^2(E) \to \dots$$

where the first operator is the covariant derivative and the others are the operators d^∇. Here we are identifying $C^\infty(E)$ with $\Omega^0(E)$. Now one very slick way of recording the fact that the partial derivative operators ∂_i commute is to simply say that $d^2 = 0$, where d is the exterior derivative. In the present case of the generalised de Rham sequence it is natural to form the operator $(d^\nabla)^2$ and to regard this as a measure of whether the operators ∇_i commute with each other.

We define the ***curvature*** $R = R(\nabla)$ of the connection ∇ to be the operator

$$R = d^\nabla \circ \nabla : \Omega^0(E) \to \Omega^2(E).$$

If we are given two vector fields V, W on M then the curvature operator gives an operator

$$R_{V,W} : \Omega^0(E) \to \Omega^0(E)$$

and from the definition it is easy to check that

$$R_{V,W} = \nabla_V \circ \nabla_W - \nabla_W \circ \nabla_V - \nabla_{[V,W]}$$

where $[V,W]$ is the Lie bracket of the two vector fields.

One of the key properties of the curvature is that it is linear over functions; $R(fs) = fR(s)$. This means that it is given by an element

$$R \in \Omega^2(End(E))$$

If we choose a local gauge e it is easy to see that the curvature is given by

$$F(e) = dA(e) + A(e) \wedge A(e)$$

where $A(e)$ is the matrix of the connection in the given gauge and that in a different gauge f the curvature is given by

$$F(f) = g^{-1} F(e) g$$

where g is the gauge transformation from the gauge e to the gauge f.

Now we will suppose that we are dealing with a $U(k)$ bundle E, i.e. a bundle equipped with a hermitian metric. Now introduce the group of ***gauge transformations*** $\mathcal{G}(E)$, this is just the group of all bundle isomorphisms of E, more precisely

$$\mathcal{G}(E) = \left\{ g \in C^\infty(End(E)) \mid g(x) \text{ is invertible for each } x \in X \right\}.$$

Under composition of maps $\mathcal{G}(E)$ becomes a bundle of groups, but beware, it is not a principle bundle. Analogously we may form

$$\mathfrak{g}(E) = C^\infty(End(E)).$$

There is a pointwise exponential map $\mathfrak{g}(E) \to \mathcal{G}(E)$. $\mathfrak{g}(E)$ is a Lie algebra using the pointwise Lie bracket $[g,h](x) = [g(x), h(x)]$ and we can regard this Lie algebra as the Lie algebra of the infinite dimensional Lie group $\mathcal{G}(E)$, provided we take due care over the necessary technicalities of infinite dimentions.

Now suppose that E is a G-bundle where G is $U(k)$ or $SU(k)$. Then we require that, ast each point $x \in M$, the gauge transformations $g \in \mathcal{G}(E)$ must be isometries and if we are dealing with $SU(n)$ we also require g to have determinant 1. Similarily an element $X \in \mathfrak{g}(E)$ is required to be skew adjoint at each point and if we are dealing with $SU(k)$ it is also required to have trace zero.

These gauge transformations g act on connections by forming the new covriant derivative operator $g^{-1}\circ\nabla\circ g$ and using a local gauge for E to identify the connection with a $\mathfrak{g}$ valued 1-form and to identify the gauge transformation g with a G-valued function we get the usual formula for the matrix of the connection $g^{-1}\circ\nabla\circ g$, i.e.

$$g^{-1}Ag + g^{-1}dg.$$

If E is a G bundle and ∇ is a G connection, where $G = U(k)$ or $SU(k)$ then a gauge transformation of ∇ is also a G-connection. Note also that the curvature of a G-connection is naturally an element of $\Omega^2(\mathfrak{g}(E))$.

Finally note that the connection ∇ defines a connection on the bundle $End(E)$ by the following formual. If $L \in \Omega^0(End(E))$ then L defines a map $L : \Omega^0(E) \to \Omega^0(E)$ which is linear over functions. Then we define $\nabla(L) = \nabla\circ L - L\circ\nabla$, it is easy to check that this gives a well-defined element of $\Omega^1(End(E))$ and so goves a connection on $End(E)$. This extends to $\Omega^p(End(E))$ where the analogous formula is $d^\nabla(\alpha) = \nabla\circ\alpha - (-1)^p\alpha\circ\nabla$. It is a nice simple exercise to see that this formula agrees with the formula given in Section 2 for the operator D_A applied to a Lie algebra valued 2-form.

4. *Some remarks on characteristic classes*

I will now give a very brief summary of the Chern-Weil method of constructing charateristic classes in de Rham cohomology using connections and curvature. Let E be a complex vector bundle and pick a connection ∇ on E with curvature $R \in \Omega^2(End(E))$. We can form the following inhomogeneous differential form

$$c(R) = \det\left(1 + \frac{i}{2\pi} R\right) \in \Omega(M)$$

where

$$\Omega(M) = \bigoplus_{i=1}^{n} \Omega^i(M)$$

is the space of all complex valued forms on M ($n = \dim M$). The important properties of this form $c(R)$ are summarised in the following theorem.

Theorem (1) The form $c(R)$ is closed, i.e. $dc(R) = 0$.

(2) If ∇_1 and ∇_2 are connections on E with curvatures R_1 and R_2 then there is a form $B \in \Omega(M)$ such that $dB = c(R_1) - c(R_2)$.

For a proof of this result see Appendix C in Milnor and Stasheff.

If we write $c(R) = c_0(R) + c_1(R) + \ldots + c_i(R) + \ldots$ where $c_i(R) \in \Omega^{2i}(M)$ is homogeneous of degree $2i$, then, by part (1) of the theorem, each $c_i(R)$ is a closed form and by part (2) the cohomology class $[c_i(R)] \in H^{2i}{}_{dR}(M)$ is independent of the choice of connection. Using de Rahm's theorem identifying singular cohomology and de Rham cohomology

$$H^*(M;\mathbb{C}) \cong H^*{}_{dR}(M)$$

the classes $[c_i(R)]$ give cohomology classes in $H^{2i}(M;\mathbb{C})$. The second theorem identifies these classes with the Chern classes of E, $c_i(E) \in H^{2i}(M;\mathbb{C})$ i.e. their images in cohomology with complex coefficients.

Theorem In $H^{2i}(M;\mathbb{C})$, $[c_i(R)] = c_i(E)$

Again this is proved in Appendix C of Milnor and Stasheff.

Let us look at the concrete application of this to Yang-Mills theory in 4-dimensions. Let E be an $SU(2)$-bundle over a closed oriented 4-dimensional manifold M. Pick a connection on E with curvature R. Then since we are dealing with 2-dimensional bundles

$$c_2(R) = \frac{1}{8\pi^2} \text{Trace}(R \wedge R)$$

Now define an integer k by the formula

$$k(E) = -\int_M c_2(R) = -\frac{1}{8\pi^2}\int_M \text{Trace}(R \wedge R).$$

This number is an integer since $c_2(R)$ is an integral characteristic class and therefore its integral around any closed cycle is an integer. It depends only on the bundle E and not on the connection ∇ in view of the previous theorems and Stokes theorem. The sign is conventional, but it seems to be so firmly entrenched that I dare not try to alter it. I can now state an imporatant result.

Theorem Two SU(2)-bundles E_1 and E_2 over a smooth closed oriented 4-manifold are isomorphic if and only if $k(E_1) = k(E_2)$.

5. The Yang-Mills functional

Let E be a vector bundle with structure group G over a compact closed Riemannian manifold M. As usual we will assume that G is $U(n)$ or $SU(n)$. Let ∇ be a G-connection on E. Then the curvature of ∇

$$R = R(\nabla) \in \Omega^2(\mathfrak{g}(E))$$

is a section of the bundle $\mathfrak{g}(E)\otimes\Lambda^2T^*$. The bundle $\mathfrak{g}(E)$ has a natural metric; on each fibre $\mathfrak{g}(E_x)$, which is the space of skew adjoint endomorphisms of E_x, it is given by the Killing form

$$-\text{Trace } AB.$$

So the bundle $\mathfrak{g}(E)\otimes\Lambda^2T^*$ has a natural metric and we can form the function $|R|$, the pointwise norm of the curvature. From this function we can form the L^2-norm of the curvature

$$\| R \|^2 = \int_M |R|^2 \text{ vol}$$

where dvol is the volume form on the Riemannian manifold M. By definition the ***Yang-Mills action*** of ∇ is given by

$$Y(\nabla) = \| R(\nabla) \|^2.$$

Using the star operator we see that this may be written differently

$$Y(\nabla) = -\int_M \text{Trace}(R \wedge (*R))$$

Let us work out the formula for Y in terms of local coordinates for M and a local orthonormal gauge for E so that the connection ∇ is given by a $\mathfrak{g}$-valued 1-form A where $\mathfrak{g}$ is the Lie algebra of the group G and the curvature is then given by the $\mathfrak{g}$-valued 2-form $F = dA + A\wedge A$ with coordinates

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu].$$

Now using the metric to raise and lower indices the formula for Y is

$$-\int \text{Trace}(F_{\mu\nu}F^{\mu\nu}) \text{ vol}.$$

The Yang-Mills action makes sense in all dimensions but the following lemma shows that in 4-dimensions it has a special feature.

Lemma If M is a 4-dimensional Riemannian manifold then the star operator on 2-forms and the Yang-Mills action is invariant under a conformal rescaling of the metric.

Proof This is just the same the proof given in Section 1, replacing the metric g by $\lambda^2 g$ multiplies the metric on 2-forms by λ^{-4} and the volume form by λ^4 so that both the star operator and the Yang-Mills action are unchanged.

A connection ∇ is called a ***Yang-Mills connection*** if it is a stationary point of Y. We will now work out the variational equations associated to this functional Y. Recall that the star operator $* : \Omega^p(M) \to \Omega^{n-p}(M)$ is linear over functions. This means that it is given by a bundle map $\Lambda^p T^* \to \Lambda^{n-p} T^*$; therefore extends as $*\otimes 1$ to a map $E\otimes\Lambda^p T^* \to E\otimes\Lambda^{n-p} T^*$ and so gives a star operator $\Omega^p(E) \to \Omega^{n-p}(E)$. The following lemma can be checked by a direct computation.

Lemma If $\omega \in \Omega^p(E)$, then $**\omega = (-1)^{p(n-p)}\omega$.

The operator $d^\nabla : \Omega^p(E) \to \Omega^{p+1}(E)$ has a formal adjoint $\delta^\nabla : \Omega^p(E) \to \Omega^{p-1}(E)$. This means that

$$\int_M < d^\nabla\alpha, \beta > \mathrm{vol} = \int_M < \alpha, \delta^\nabla\beta > \mathrm{vol}$$

for $\alpha \in \Omega^{p-1}(E)$ and $\beta \in \Omega^p(E)$. Here of course the inner products are induced by the Riemannian metric on M and the given inner product on E. The following standard lemma identifies the operator δ^∇ in terms of the $*$-operator.

Lemma Let M be an oriented n-dimensional Riemannian manifold and let E be a G-bundle

on M with a G-connection ∇; then

$$\delta^\nabla = (-1)^{n(p-1)-1} *d^\nabla* : \Omega^p(E) \to \Omega^{p-1}(E).$$

Proof It is sufficient to check the result in the case where E is the trivial line bundle. We must show that if $\alpha \in \Omega^{p-1}(M)$ and $\beta \in \Omega^p(M)$ then

$$\int_M < d\alpha, \beta > \mathrm{vol} = (-1)^{n(p-1)-1} \int_M < \alpha, *d(*\beta) > \mathrm{vol}.$$

From the definition of $*$ it follows that

$$\int_M < d\alpha, \beta > d\mathrm{vol} = \int_M d\alpha\wedge(*\beta);$$

since M is closed and $d(\alpha\wedge(*\beta)) = d\alpha\wedge(*\beta) + (-1)^{p-1}\alpha\wedge d(*\beta)$ Stokes theorem shows that

$$\int_M d\alpha\wedge(*\beta) = (-1)^p \int_M \alpha\wedge d(*\beta).$$

Now we use the formula for $**$ and the relation between $*$ and the inner product to get

$$\int_M \alpha\wedge d(*\beta) = (-1)^{(n-p+1)(p-1)} \int_M \alpha\wedge**d(*\beta) = \int_M < \alpha, *d(*\beta) > \mathrm{vol}.$$

Hence we see that

$$\int_M < d\alpha, \beta > \mathrm{vol} = (-1)^{(n-p+1)(p-1)+p} \int_M < \alpha, *d(*\beta) > \mathrm{vol}.$$

and simplifying the signs gives the required result.

Now I shall show how to prove that the ***Yang-Mills equations***

$$d^\nabla(R) = 0, \quad \delta^\nabla(R) = 0$$

are the variational equations of the Yang-Mills functional.

Lemma The following are equivalent

(1) ∇ is a Yang-Mills connection.

(2) $\delta^\nabla R = 0$.

(3) $\Delta R = 0$ where Δ is the Laplacian $d^\nabla\delta^\nabla + \delta^\nabla d^\nabla$.

Proof Fix A in $\Omega^1(X;\mathfrak{g}(E))$ and consider the curve of connections $\nabla^t = \nabla + tA$, $t \in \mathbb{R}$. It is straightforward to compute the curvature of $R^t = R(\nabla^t)$:

$$R^t = R + td^\nabla(A) + t^2[A, A].$$

where $[\,,\,]$ is defined by the combination of the Lie bracket of matrices and the exterior product of forms; therefore

$$\frac{d}{dt}\Big|_{t=0}(Y(\nabla^t)) = \int_M \langle d^\nabla A, R\rangle \text{ vol} = \int_M \langle A, \delta^\nabla R\rangle \text{ vol}.$$

and this shows that $\delta^\nabla R$ is orthogonal to every element of $\Omega^1(\mathfrak{g}(E))$ with respect to the inner product

$$\int_M \langle \alpha, \beta\rangle \text{ vol}.$$

However this inner product is non-degenerate so that

$$\frac{d}{dt}\Big|_{t=0}(Y(\nabla^t)) = 0$$

for all choices of A if and only if $\delta^\nabla R = 0$. This shows that ∇ is a stationary point of Y if and only if $\delta^\nabla R = 0$ and therefore shows that (1) is equivalent to (2). To see that (2) is equivalent to (3) we use the Bianch identity $d^\nabla R = 0$ and then argue as follows:

$$\int_M \langle \Delta R, R\rangle \text{ vol} = \int_M \langle (d^\nabla\delta^\nabla + \delta^\nabla d^\nabla)R, R\rangle \text{vol} = \int_M (|d^\nabla R|^2 + |\delta^\nabla R|^2) \text{ vol}$$

$$= \int_M |\delta^\nabla R|^2 \text{ vol}.$$

Now we will assume that M is 4-dimensional and that the structure group G of the bundle E is $SU(2)$. We will give a topological lower bound for the functional Y and show that the self dual connections are precisely the absolute minima of Y.

Lemma (1) $Y(\nabla) \geq 8\pi^2 k(E)$

(2) If $k(E) \geq 0$, then $Y(\nabla) = 8\pi^2 k(E)$ if and only if ∇ is self dual.

(3) If $k(E) < 0$, then there are no self dual connections on E.

Proof Let R be the curvature of ∇ and break R up into its self dual and anti self dual peices;

$$R = R^+ + R^-, \quad *R^+ = R^+, \quad *R^- = R^-.$$

It is easy to check directly that the self dual forms are orthogonal to the anti self dual forms so it follows that

$$Y(\nabla) = \int_M (|R^+|^2 + |R^-|^2) \text{ vol}$$

On the other hand one can check that $R^+ \wedge R^- = 0$ so now it follows that

$$\text{Trace } R\wedge R = \text{Trace } R^+\wedge R^+ + \text{Trace } R^-\wedge R^-$$

and therefore

$$8\pi^2 k(E) = -\int_M (|R^+|^2 - |R^-|^2) \text{ vol}$$

Therefore

$$Y(\nabla) = \int_M (|R^+|^2 + |R^-|^2) \text{ vol} \geq 8\pi^2 k(E)$$

and we have equality if and only if $R^- = 0$. Similarily it follows directly that if $k(E) < 0$ then there can be no self dual connections on E.