Introduction to the Yang-Mills Equations

Final Project for Math 581
11 May 2012

Sébastien Picard
sebastien.picard@mail.mcgill.ca

Contents

1. Introduction

We discuss some topics on the Yang-Mills equations. We first develop some of the geometric background needed to understand the statement of the equation; namely, the notions of a connection, curvature, and the Hodge star operator. We then state the equation and briefly discuss gauge transformations. As an aside, which is interesting from a physics standpoint, we derive the Maxwell's equations from the Yang-Mills framework. Since Yang-Mills theory is a generalization of Hodge theory, we spend some time carefully proving the Hodge theorem. We conclude by investigating the Yang-Mills equations over a compact four-manifold and discuss selfdual and antiselfdual instantons.

The majority of these notes were written while following Jost's *Riemannian Geometry and Geometric Analysis* [1], with some inspiration from more physics-oriented texts such as [2]. The last section was written while following [4]. The prerequisites are graduate level courses in smooth manifolds and partial differential equations. The Einstein summation convention is always assumed.

2. Connections and Curvature

In this section, we explore connections on vector bundles. Let $\pi : E \to M$ to be a vector bundle of rank n over a smooth manifold M. We use the notation $\Gamma(E)$ to denote the set of smooth sections $\sigma : M \to E$ of the vector bundle E. In the case $E = TM$, we have that $\sigma \in \Gamma(TM)$ is a smooth vector field on M. A connection on a vector bundle in a generalization of the directional derivative in $\mathbb{R}^d$. The precise definition is the following:

Definition 2.1. *Let M be a smooth manifold, and $\pi : E \to M$ a vector bundle. A connection on E is a map*

$$D : \Gamma(TM) \times \Gamma(E) \to \Gamma(E),$$
$$(X, \sigma) \mapsto D_X\sigma$$

such that for any $\tau, \sigma \in \Gamma(E)$, $X, Y \in \Gamma(TM)$, and $f \in C^\infty(M)$, the following four properties hold:

(1) $D_{X+Y}\sigma = D_X\sigma + D_Y\sigma,$

(2) $D_{fX}\sigma = fD_X\sigma,$

(3) $D_X(\sigma + \tau) = D_X\sigma + D_X\tau,$

(4) $D_X(f\sigma) = X(f)\sigma + fD_X\sigma.$

Rule (4) is called the Leibniz rule. We start by investigating the local properties of D. Given any point $q \in M$, by the definition of a vector bundle, there exists a neighbourhood U containing q and a diffeomorphism $\Phi : \pi^{-1}(U) \to U \times \mathbb{R}^n$ that acts linearly on the fibers. (Φ is called a *local trivialization.*) If $\{e_1, \dots, e_n\}$ is the standard basis of $\mathbb{R}^n$, we obtain local sections $\varepsilon_i : M \to E$, $\varepsilon_i(p) = \Phi^{-1}(p, e_i)$ such that at each $p \in U$, $\{\varepsilon_i(p)\}$ forms a basis of the fiber E_p. We can thus write any section $\sigma : M \to E$ in the form

$$\sigma(p) = \sigma^i(p)\varepsilon_i(p),$$

where the σ^i are functions on U. We define the trivial connection D^0 as follows:

$$D^0_X(\sigma) := X(\sigma^i)\varepsilon_i.$$

It is easy to check that D^0 indeed defines a connection. Now, take any given connection D. Locally, we can define the difference

$$A(X) := D^0_X - D_X,$$
$$A : \Gamma(TM) \to C^\infty(M, \mathrm{End}(\mathbb{R}^n)),$$
$$X \mapsto A(X).$$

Hence for each smooth vector field X, we obtain $A(X)$, which is a matrix-valued function on M. Said in another way, given $p \in M$, $(A(X))(p)$ acts on a vector $\sigma^i(p)\varepsilon_i(p)$. We can check that A is C^∞-linear. Indeed, we have $A(X)(\sigma^i\varepsilon_i) = \sigma^i A(X)\varepsilon_i$, since the derivative terms in the Leibniz rule cancel:

$$A(X)(\sigma^i\varepsilon_i) = X(\sigma^i)\varepsilon_i - X(\sigma^i)\varepsilon_i - \sigma^i D_X\varepsilon_i = \sigma^i(D^0_X - D_X)\varepsilon_i.$$

Conversely, any map $A : \Gamma(TM) \to C^\infty(M, \mathrm{End}(\mathbb{R}^n))$ that is C^∞-linear defines a connection $D_X = D^0_X + A(X)$. Therefore, the maps A classify connections on a trivial bundle.

We thus obtain the following local decomposition for any connection D:

$$D_X = D^0_X + A(X) \tag{1}$$

.

Written out explicitly, under this decomposition we have

$$D_X\sigma = X(\sigma^i)\varepsilon_i + A_{ij}(X)\sigma^j\varepsilon_i$$

We now define the curvature operator corresponding to a connection D.

Definition 2.2. *Let D be a connection on the vector bundle $E \to M$. Given vector fields $V, W \in \Gamma(TM)$, we define the curvature $F(V, W) : \Gamma(E) \to \Gamma(E)$ in the following way:*

$$F(V, W)\sigma = D_V D_W\sigma - D_W D_V - D_{[V,W]}\sigma.$$

We investigate the form in F in local coordinates. The first claim is that if we have $V = V^i\partial_i$, $W = W^i\partial_i$, we can write $F(V, W) = V^jW^kF_{jk}$, where

$$F_{jk} := F(\partial_j, \partial_k).$$

Since partial derivatives commute, we have $[\partial_j, \partial_k] = 0$, and hence

$$F_{jk} = D_j D_k - D_k D_j, \tag{2}$$

where for convenience, we denote $D_i := D_{\partial_i}$.

Before launching into the computation, we note how to write the Lie bracket in coordinates:

$$[V, W] = (V^i \partial_i W^k - W^i \partial_i V^k)\partial_k.$$

We now compute

$$\begin{aligned}
F(V^i\partial_i, W^k\partial_k)\sigma &= D_{V^i\partial_i} D_{W^k\partial_k}\sigma - D_{W^k\partial_k} D_{V^i\partial_i}\sigma - D_{(V^i\partial_i W^k - W^i\partial_i V^k)\partial_k}\sigma \\
&= D_{V^i\partial_i} W^k D_k\sigma - D_{W^k\partial_k} V^i D_i\sigma - (V^i\partial_i W^k)D_k\sigma - (W^i\partial_i V^k)D_k\sigma \\
&= V^i\partial_i W^k D_k\sigma + V^i W^k D_i D_k\sigma - W^k\partial_k V^i D_i\sigma - W^k V^i D_k D_i\sigma \\
&\quad - V^i\partial_i W^k D_k\sigma - W^i\partial_i V^k D_k\sigma \\
&= V^i W^k (D_i D_k\sigma - D_k D_i\sigma) \\
&= V^i W^k F(\partial_i, \partial_k)\sigma.
\end{aligned}$$

This shows that $F(V, W) = V^j W^k F_{jk}$.

The next step is to write F_{jk} in terms of A. As usual, locally we write $\sigma = \sigma^i \varepsilon_i$. Using the decomposition from (1), and the notation $A_i = A(\partial_i)$, we compute

$$\begin{aligned}
F_{jk}\sigma &= (D_j D_k - D_k D_j)\sigma^i\varepsilon_i \\
&= (D_j^0 + A_j)(D_k^0 + A_k)\sigma^i\varepsilon_i - (D_k^0 + A_k)(D_j^0 + A_j)\sigma^i\varepsilon_i \\
&= (D_j^0 + A_j)(\partial_k\sigma^i\varepsilon_i + A_k\sigma^i\varepsilon_i) - (D_k^0 + A_k)(\partial_j\sigma^i\varepsilon_i + A_j\sigma^i\varepsilon_i) \\
&= \partial_j\partial_k\sigma^i\varepsilon_i + \partial_j(A_k\sigma^i)\varepsilon_i + A_j\partial_k\sigma^i\varepsilon_i + A_jA_k\sigma^i\varepsilon_i - \partial_k\partial_j\sigma^i\varepsilon_i - \partial_k(A_j\sigma^i)\varepsilon_i - A_k\partial_j\sigma^i\varepsilon_i - A_kA_j\sigma^i\varepsilon_i \\
&= (\partial_j A_k)\sigma^i\varepsilon_i + A_k\partial_j\sigma^i\varepsilon_i + A_j\partial_k\sigma^i\varepsilon_i + A_jA_k\sigma^i\varepsilon_i - (\partial_k A_j)\sigma^i\varepsilon_i - A_j\partial_k\sigma^i\varepsilon_i - A_k\partial_j\sigma^i\varepsilon_i - A_kA_j\sigma^i\varepsilon_i \\
&= (\partial_j A_k - \partial_k A_j + A_jA_k - A_kA_j)\sigma^i\varepsilon_i.
\end{aligned}$$

We have thus derived the following useful local expression:

$$F_{jk}\sigma = (\partial_j A_k - \partial_k A_j + [A_j, A_k])\sigma^i\varepsilon_i. \tag{3}$$

Written this way, it is easy to see the skew-symmetry of the curvature operator:

$$\begin{aligned} F_{jk} &= \partial_j A_k - \partial_k A_j + [A_j, A_k] \\ &= -(\partial_k A_j - \partial_j A_k - [A_j, A_k]) \\ &= -(\partial_k A_j - \partial_j A_k + [A_k, A_j]) \\ &= -F_{kj}. \end{aligned}$$

There is a different way to look at the curvature operator F. The ultimate goal is to generalize the definition of the connection D in a way akin to the exterior derivative d. The rest of this section introduces new notation that will be used throughout the document. We start by rewriting the curvature operator, and then move on to generalize the connection D.

We rewrite the curvature operator in the following way:

$$F = F_{jk} \otimes dx^j \wedge dx^k. \tag{4}$$

We use the so-called "Alt convention" for the wedge product; for covectors $\omega^i \in \Gamma(TM^*)$ and vector fields $V_i \in \Gamma(TM)$, we have

$$\omega^1 \wedge \cdots \wedge \omega^k (V_i, \dots, V_k) = \frac{1}{k!} \det(\omega^i(V_j)).$$

To justify (4), we compute for vector fields V, W

$$\begin{aligned} F(V, W) &= F_{jk} \otimes dx^j \wedge dx^k \ (V^i \partial_i \otimes W^l \partial_l) \\ &= \frac{F_{jk}}{2}(V^j W^k - W^j V^k) \\ &= \frac{1}{2}(V^j W^k F_{jk} - V^k W^j F_{jk}) \\ &= \frac{1}{2}(V^j W^k F_{jk} + V^k W^j F_{kj}) \\ &= V^j W^k F_{jk}. \end{aligned}$$

Now that we have tweaked the definition of the curvature F, we revisit the definition of a connection D.

Given a vector bundle $E \to M$ of rank n, we can construct the bundle $\mathrm{End}(E)$. If E_p is the fiber over $p \in M$, we can look at the space of linear transformations on the vector space E_p. The fibers of $\mathrm{End}(E)$ are given by this space of linear transformations, and allows us to the define the vector bundle $\mathrm{End}(E) \to M$ of rank n^2.

We want to define extend our definition of a connection D to allow it to act on sections of $\mathrm{End}(E)$. The motivation for the definition is to have a Leibniz rule; if $T \in \Gamma(\mathrm{End}(E))$, $\sigma \in \Gamma(E)$, and V is a smooth vector field, then we should have

$$D_V(T\sigma) = (D_V T)\sigma + T(D_V \sigma).$$

With this motivation in mind, we define

$$(D_V T)\sigma := D_V(T\sigma) - T(D_V\sigma). \tag{5}$$

Locally, if we write D in its decomposition $D = D^0 + A$, we obtain

$$(D_j T)\sigma = (D_j^0 + A_j)(T\sigma) - T((D_j^0 + A_j)\sigma) = (\partial_j T)\sigma + T\partial_j\sigma + A_j T\sigma - T\partial_j\sigma - TA_j\sigma.$$

After cancellation, we see that

$$D_j T = \partial_j T + [A_j, T]. \tag{6}$$

x

In particular, since $F_{jk} \in \Gamma(\mathrm{End}(E))$, we have

$$D_i F_{jk} = \partial_i F_{jk} + [A_i, F_{jk}]. \tag{7}$$

We denote by $\Omega^p(M)$ the space of smooth p-forms on M. On this space, we have the notion of the exterior derivative $d : \Omega^p(M) \to \Omega^{p+1}(M)$. Inspired by the exterior derivative, we would like to generalize the connection operator D in a similar fashion.

Let us look at our current definition of D from a different angle. Originally, we defined D as a map $D : \Gamma(TM) \times \Gamma(E) \to \Gamma(E)$, $(X, \sigma) \mapsto D_X\sigma$. However, we can also view this as a map $D : \Gamma(E) \to \Gamma(E) \otimes \Gamma(T^*M)$. Since T^*M is the space of 1-forms on M, this can be rewritten as $D : \Gamma(E) \to \Gamma(E) \otimes \Omega^1(M)$. When viewed in this way, in local coordinates, we can see the action of D by using the shorthand $D_i \otimes dx^i$. For given a section $\sigma : M \to E$ and any smooth vector field V written out in coordinates $V^k\partial_k$, we have the action

$$(D_i\sigma \otimes dx^i)(V^k\partial_k) = V^i D_i\sigma,$$

which agrees with the previous definition. We now introduce some more notation. Let

$$\Omega^p(E) := \Gamma(E) \otimes \Omega^p(M).$$

Locally, for $\sigma \in \Gamma(E)$ and $\omega \in \Omega^p(M)$, we have $\sigma = \sigma^i \varepsilon_i$ and $\omega = \omega_{(i_1,\ldots,i_p)} dx^{i_1} \wedge \cdots \wedge dx^{i_p} = \omega_I dx^I$, where $I = (i_1, \ldots, i_p)$ is a multi-index. Using a property of tensors, elements of $\Omega^p(E)$ are locally of the form

$$\sigma \otimes \omega = \sigma^i \varepsilon_i \otimes \omega_I dx^I = \sigma^i \omega_I \varepsilon_i \otimes dx^I := \sigma_I \otimes dx^I,$$

where σ_I is defined as a useful short-hand notation.

As the main motivating example for this definition, we see from (4) that $F \in \Omega^2(\mathrm{End}(E))$.

On $\Omega^p(E)$, we define the following wedge product; for $\sigma \in \Gamma(E)$, $\omega_1 \in \Omega^1(M)$, $\omega_2 \in \Omega^p(M)$, define

$$(\sigma \otimes \omega_1) \wedge \omega_2 := \sigma \otimes (\omega_1 \wedge \omega_2).$$

In other words, we simply let the wedge act on the p-form and leave the vector bundle section untouched.

Using this notation, we extend the definition of the connection in the following way: for $0 \leq p \leq d$, where d is the dimension of the manifold M, $\sigma \in \Gamma(E)$, $\omega \in \Omega^p(M)$,

$$D : \Omega^p(E) \to \Omega^{p+1}(E),$$

$$D(\sigma \otimes \omega) = D\sigma \wedge \omega + \sigma \otimes d\omega.$$

In coordinates, if we write an element of $\Omega^p(E)$ as $\sigma_I \otimes dx^I$, we obtain

$$D(\sigma_I \otimes dx^I) = (D_k \sigma_I \otimes dx^k) \wedge dx^I = D_k \sigma_I \otimes dx^k \wedge dx^I.$$

In short-hand, when acting on elements of the local form $\sigma_I \otimes dx^I$, it may be useful to intuitively think of the connection as being of the form $D = D_i \otimes dx^i \wedge$.

For example, we compute $D \circ D$:

$$\begin{aligned} D \circ D &= D_i D_j \otimes dx^i \wedge dx^j \\ &= \frac{1}{2}(D_i D_j \otimes dx^i \wedge dx^j + D_j D_i \otimes dx^j \wedge dx^i) \\ &= \frac{1}{2}(D_i D_j - D_j D_i) \otimes dx^i \wedge dx^j \\ &= \frac{1}{2} F_{ij} \otimes dx^i \wedge dx^j. \end{aligned}$$

Hence

$$D \circ D = \frac{1}{2} F. \tag{8}$$

Proposition 2.1. (Bianchi Identity) *The curvature F of a connection D satisfies $DF = 0$.*

Proof. Recall that locally we have $F = F_{ij} \otimes dx^i \wedge dx^j$. We use (7) to expand $D_k F_{ij}$, then use (3) to write F_{ij} in terms of A.

$$\begin{aligned} DF &= D_k F_{ij} \otimes dx^k \wedge dx^i \wedge dx^j \\ &= (\partial_k F_{ij} + [A_k, F_{ij}]) \otimes dx^k \wedge dx^i \wedge dx^j \\ &= (\partial_k(\partial_i A_j - \partial_j A_i + [A_i, A_j]) + [A_k, \partial_i A_j - \partial_j A_i + [A_i, A_j]]) \otimes dx^k \wedge dx^i \wedge dx^j \\ &= (\partial_k [A_i, A_j] + [A_k, \partial_i A_j] - [A_k, \partial_j A_i] + [A_k, [A_i, A_j]]) \otimes dx^k \wedge dx^i \wedge dx^j \\ &= ((\partial_k A_i) A_j + A_i \partial_k A_j - (\partial_k A_j) A_i - A_j \partial_k A_i + A_k \partial_i A_j \\ &\quad - (\partial_i A_j) A_k + (\partial_j A_i) A_k - A_k \partial_j A_i + [A_k, [A_i, A_j]]) \otimes dx^k \wedge dx^i \wedge dx^j \\ &= A_k [A_i, A_j] \otimes dx^k \wedge dx^i \wedge dx^j. \end{aligned}$$

The rest follows from manipulating indices and the skew-symmetry of the wedge product:

$$\begin{aligned} DF &= (A_k(A_i A_j - A_j A_i) - (A_i A_j - A_j A_i) A_k) \otimes dx^k \wedge dx^i \wedge dx^j \\ &= A_k A_i A_j \otimes dx^k \wedge dx^i \wedge dx^j + A_j A_i A_k \otimes dx^k \wedge dx^i \wedge dx^j \\ &\quad - (A_k A_j A_i \otimes dx^k \wedge dx^i \wedge dx^j + A_i A_j A_k \otimes dx^k \wedge dx^i \wedge dx^j) \\ &= A_k A_i A_j \otimes dx^k \wedge dx^i \wedge dx^j - A_j A_i A_k \otimes dx^j \wedge dx^i \wedge dx^k \\ &\quad - (A_k A_j A_i \otimes dx^k \wedge dx^i \wedge dx^j - A_i A_j A_k \otimes dx^i \wedge dx^k \wedge dx^j) \\ &= 0. \end{aligned}$$

□

3. The Hodge Star Operator

Let V be a vector space of dimension d with a positive definite inner product $(\cdot,\cdot)$. Recall that an *orthonormal basis* of V is a basis $\{e_1,\dots,e_n\}$ of V such that

$$(e_i, e_j) = \delta^{ij}.$$

An *orientation* on V is a choice of an orthonormal basis $\{e_1,\dots,e_d\}$ which is declared to be positive. Any other basis $\{v_1,\dots,v_d\}$ that is related to $\{e_1,\dots,e_d\}$ by a change of basis matrix with positive determinant is said to be *positively oriented*. If the change of basis matrix has negative determinant, the basis $\{v_1,\dots,v_d\}$ is said to be *negatively oriented*.

Recall that if $\{e_1,\dots,e_d\}$ is a basis for V, then $e_{i_1}\wedge\cdots\wedge e_{i_p}$, $\{i_1 < i_2 < \cdots < i_p\}$ is a basis for $\Lambda^p(V)$.

We define the Hodge star operator to be a linear map $*:\Lambda^p(V)\to\Lambda^{d-p}(V)$. By linearity, it suffices select a positive orthonormal basis $\{e_1,\dots,e_d\}$ for V and to define the star operator on a basis element of $\Lambda^p(V)$ of the form $e_{i_1}\wedge\cdots\wedge e_{i_p}$. We define

$$*(e_{i_1}\wedge\cdots\wedge e_{i_p}) = e_{j_1}\wedge\cdots\wedge e_{j_{d-p}},$$

where $\{e_{j_1},\dots,e_{j_{d-p}}\}$ is chosen such that $\{e_{i_1},\dots,e_{i_p},e_{j_1},\dots,e_{j_{d-p}}\}$ is a positive orthonormal basis for V. If another grouping $\{e_{k_1},\dots,e_{k_{d-p}}\}$ is chosen such that $\{e_{i_1},\dots,e_{i_p},e_{k_1},\dots,e_{k_{d-p}}\}$ is a positive orthonormal basis for V, then $\{e_{k_1},\dots,e_{k_{d-p}}\}$ differs from $\{e_{j_1},\dots,e_{j_{d-p}}\}$ by an even number of permutations, and by skew-symmetry $e_{j_1}\wedge\cdots\wedge e_{k_{d-p}} = e_{j_1}\wedge\cdots\wedge e_{k_{d-p}}$.

We must show that the Hodge star operator does not depend on the choice of positive orthonormal basis $\{e_1,\dots,e_d\}$. Take $\{f_1,\dots,f_d\}$ to be another positive orthonormal basis of V, with $e_i = A_{ij}f_j$. Since both these bases are positively oriented, we have $\det A > 0$, and hence by orthonormality we must have $\det A = 1$. Therefore,

$$*(e_1\wedge\cdots\wedge e_p) = *(Af_1\wedge\cdots\wedge Af_p) = *((\det A)\ f_1\wedge\cdots\wedge f_p).$$

A useful property of the Hodge star operator is the following:

Proposition 3.1. *For all $\alpha\in\Lambda^p(V)$, the Hodge star operator satisfies*

$$**\,\alpha = (-1)^{p(d-p)}\alpha.$$

Proof. It suffices to show the identity for an orthonormal basis element $e_{i_1}\wedge\cdots\wedge e_{i_p}$. Suppose $\{e_{i_1},\dots,e_{i_p},e_{j_1},\dots,e_{j_{d-p}}\}$ is a positive orthonormal basis for V. First, we notice that

$$e_{i_1} \wedge \cdots \wedge e_{i_p} \wedge e_{j_1} \wedge \cdots \wedge e_{j_{d-p}} = (-1)^{p(d-p)} e_{j_1} \wedge \cdots \wedge e_{j_{d-p}} \wedge e_{i_1} \wedge \cdots \wedge e_{i_p},$$

since moving each e_j across p elements e_i changes the sign by $(-1)^p$, and this occurs for $(d-p)$ elements e_j. In other words, $(-1)^{p(d-p)}$ is the determinant of the matrix changing the basis $\{e_{i_1}, \ldots, e_{i_p}, e_{j_1}, \ldots, e_{j_{d-p}}\}$ to $\{e_{j_1}, \ldots, e_{j_{d-p}}, e_{i_1}, \ldots, e_{i_p}\}$. We can now compute

$$\begin{aligned} **(e_{i_1} \wedge \cdots \wedge e_{i_p}) &= *(e_{j_1} \wedge \cdots \wedge e_{j_{d-p}}) \\ &= (-1)^{p(d-p)}\ e_{i_1} \wedge \cdots \wedge e_{i_p}. \end{aligned}$$

□

We now consider the star operator on an orientable Riemannian manifold (M, g) of dimension d. Recall that a smooth manifold is said to be *orientable* if there exists a covering by smooth charts $\{(U_\alpha, \varphi_\alpha)\}$ such all that the transition maps $\varphi_\alpha \circ \varphi_\beta^{-1}$ have positive Jacobian determinant on $\varphi_\beta(U_\alpha \cap U_\beta)$. Intuitively, we have a choice of orientation for each tangent space T_pM that fits together nicely with the tangent space of the other points on the manifold.

For each point $p \in M$, we have an induced inner product space T_p^*M with inner product $g^{ij}(p) = g_{ij}(p)^{-1}$. The orientation on T_pM also induces an orientation on T_p^*M: if $\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_d}\}$ is an orientation on T_pM, then the dual basis $\{dx_1, \ldots, dx_d\}$ is the induced orientation on T_p^*M. Therefore, we can define the star operator $* : \Lambda^p(T_p^*M) \to \Lambda^{d-p}(T_p^*M)$. This yields a base point preserving operator on the space of p-forms:

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

Understanding the Hodge star operator is best done by illustration with an example.

Example 3.1. *Take $M = \mathbb{R}^4$ with oriented cotangent basis $\{dx^1, dx^2, dx^3, dx^4\}$ and Euclidean inner product $g^{ij} = \delta^{ij}$. We compute*

$*(dx^1 \wedge dx^2) = dx^3 \wedge dx^4$
$*(dx^1 \wedge dx^3) = -dx^2 \wedge dx^4$
$*(dx^1 \wedge dx^4) = dx^2 \wedge dx^3$
$*(dx^2 \wedge dx^3) = dx^1 \wedge dx^4$
$*(dx^2 \wedge dx^4) = -dx^1 \wedge dx^3$
$*(dx^3 \wedge dx^4) = dx^1 \wedge dx^2$

The next task is to define the Hodge star operator on the curvature $F \in \Omega^2(\mathrm{End}(E))$. However, instead of defining the star operator on all of $\Omega^p(\mathrm{End}(E))$, we shall restrict ourselves to those elements of $\Omega^p(\mathrm{End}(E))$ whose endomorphisms on each fiber are skew-symmetric. Denote this space by

$$\Omega^p(\mathrm{Ad}\ E).$$

Before showing $F \in \Omega^2(\mathrm{Ad}\ E)$, we must make some more definitions. Given a vector bundle $\pi : E \to M$, a family of positive definite inner products on each of the fibers E_p, varying smoothly with $p \in M$, is called a *bundle metric*.

Definition 3.1. *Let E be a vector bundle over a smooth manifold M with bundle metric $\langle \cdot, \cdot \rangle$. A metric connection D on E is a connection on E such that for any smooth vector field $X \in \Gamma(TM)$ and sections $\sigma, \mu \in \Gamma(E)$, we have*

$$X\langle \sigma, \mu \rangle = \langle D_X \sigma, \mu \rangle + \langle \sigma, D_X \mu \rangle.$$

Proposition 3.2. *Let E be a vector bundle over a smooth manifold M with bundle metric $\langle \cdot, \cdot \rangle$ and metric connection D. Locally, if we write $D_X = D_X^0 + A(X)$, then $A(X_p)$ is a skew-symmetric matrix for every $p \in M$ and $X \in \Gamma(TM)$.*

Proof. Let $q \in M$. It can be shown that there exists an open set U containing q and local sections $\{\varepsilon_1, \dots, \varepsilon_n\}$ such that for all $p \in U$, $\{\varepsilon_1(p), \dots, \varepsilon_n(p)\}$ is an orthonormal basis for E_p. The proof uses Gram-Schmidt orthonormalization and is omitted. Given this set of local sections, select two (not necessarily distinct) members $\varepsilon_i, \varepsilon_j$. Notice that for all vector fields X,

$$D_X^0(\varepsilon_i) = X(1)\varepsilon_i = 0.$$

Also, since $\langle \varepsilon_i, \varepsilon_j \rangle = \delta_{ij}$ does not vary with $p \in U$, it is a locally constant function and we have $X\langle \varepsilon_i, \varepsilon_j \rangle = 0$. By definition of a metric connection,

$$0 = X\langle \varepsilon_i, \varepsilon_j \rangle = \langle (D_X^0 + A(X))\varepsilon_i, \varepsilon_j \rangle + \langle \varepsilon_i, (D_X^0 + A(X))\varepsilon_j \rangle = \langle A_{ki}(X)\varepsilon_k, \varepsilon_j \rangle + \langle \varepsilon_i, A_{lj}(X)\varepsilon_l \rangle.$$

Hence

$$A_{ki}(X)\langle \varepsilon_k, \varepsilon_j \rangle = -A_{lj}(X)\langle \varepsilon_i, \varepsilon_l \rangle,$$

and therefore $A_{ji}(X) = -A_{ij}(X)$ as required.

□

Proposition 3.3. *The curvature operator F of a metric connection D is an element of* $\Omega^2(\text{Ad } E)$.

Proof. For fixed j, k, recall that

$$F_{jk} = \partial_j A_k - \partial_k A_j + [A_j, A_k].$$

From the previous proposition, we know that A_j, A_k are $n \times n$ skew-symmetric matrices (where n is the rank of the bundle E). The space of skew-symmetric matrices (usually denoted $\mathfrak{o}(n)$) is a Lie algebra, hence it is closed under linear combinations and taking the Lie bracket. Therefore, F_{jk} is a skew-symmetric endomorphism on the fibers of E, hence $F \in \Omega^2(\text{Ad } E)$.

□

Note that the relation $F_{ij} = -F_{ji}$ is a completely different skew-symmetry than the one shown above! In the previous proposition we fixed i, j and showed that each $F_{ij} \in \Omega^2(\text{Ad } E)$, hence showing that $F \in \Omega^2(\text{Ad } E)$. The skew-symmetry $F_{ij} = -F_{ji}$ arises from varying i, j.

We now define the Hodge star operator on $\Omega^p(\text{Ad } E)$ by letting it act solely on the p-form. Let $\sigma \in \Gamma(\text{Ad } E), \omega \in \Omega^p(M)$,

$$\begin{gathered} * : \Omega^p(\text{Ad } E) \to \Omega^{d-p}(\text{Ad } E), \\ *(\sigma \otimes \omega) := \sigma \otimes *\omega. \end{gathered}$$

At last, we have all the necessary ingredients to define the Yang-Mills equations.

4. Statement of the Yang-Mills Equation

Definition 4.1. *Let E be a vector bundle over an orientable Riemannian smooth manifold M. A metric connection D is called a Yang-Mills connection if*

$$*D * F = 0.$$

Since the Hodge star operator has an inverse, we see that this equation is equivalent to $D * F = 0$.

We briefly discuss gauge transformations. Denote by $\mathrm{Aut}(E)$ the bundle over M whose fiber over $p \in M$ is the group of orthogonal linear transformations of E_p. A section g of $\mathrm{Aut}(E)$ is called a *gauge transformation*. The group of all such gauge transformations is called the *gauge group* G. We let g act on D by conjugation: if $\sigma \in \Gamma(E)$, then

$$(g \cdot D)\sigma := g^{-1} D(g\sigma).$$

How does A transform under this action by g? So far, we have written the local decomposition $D_X = D_X^0 + A(X)$. However, using the usual exterior derivative, we can write this without having to feed the vector field X into D. From now onwards, we will sometimes simply write $D = d + A$. Using this expression, we see

$$\begin{aligned} g \cdot (d + A)\sigma &= g^{-1}(d + A)(g\sigma) \\ &= g^{-1} d(g\sigma) + g^{-1} A(g\sigma) \\ &= g^{-1}(dg)\sigma + g^{-1} g d\sigma + g^{-1} A(g\sigma) \\ &= (d + g^{-1} dg + g^{-1} A g)\sigma. \end{aligned}$$

Therefore, we have

$$g \cdot A = g^{-1} dg + g^{-1} A g \tag{9}$$

We let g act on $\Omega^p(\mathrm{End}(E))$ in the following way

$$g \cdot (F_I \otimes dx^I) := (g^{-1} F_I g) \otimes dx^I.$$

We will show that if D is a Yang-Mills connection on E, then the connection $\tilde{D} = g \cdot D$ is also a Yang-Mills connection. First, we compute the curvature $\tilde{F}$ of $\tilde{D}$. By (2), we have locally

$$\begin{aligned}
\tilde{F}_{jk} &= \tilde{D}_j \tilde{D}_k - \tilde{D}_k \tilde{D}_j \\
&= g^{-1} D_j g g^{-1} D_k g - g^{-1} D_k g g^{-1} D_j g \\
&= g^{-1}(D_j D_k - D_k D_j) g \\
&= g^{-1} F_{jk} g.
\end{aligned}$$

Therefore, the curvature of D transforms as $\tilde{F} = g^{-1} F g$. Computing the Yang-Mills equations in coordinates using (5), for any $\sigma \in \Gamma(E)$, we obtain

$$\begin{aligned}
(*\tilde{D} * \tilde{F})\sigma &= (\tilde{D}_i \tilde{F}_{jk})\sigma \otimes *(dx^i \wedge *(dx^j \wedge dx^k)) \\
&= (\tilde{D}_i(\tilde{F}_{jk}\sigma) - \tilde{F}_{jk}(\tilde{D}_i \sigma)) \otimes *(dx^i \wedge *(dx^j \wedge dx^k)) \\
&= (g^{-1} D_i(F_{jk}(g\sigma)) - g^{-1} F_{jk}(D_i(g\sigma))) \otimes *(dx^i \wedge *(dx^j \wedge dx^k)) \\
&= g^{-1}((D_i F_{jk})(g\sigma)) \otimes *(dx^i \wedge *(dx^j \wedge dx^k)) \\
&= g \cdot (*D * F)\sigma.
\end{aligned}$$

It follows that if D is a Yang-Mills connection on E, then $g \cdot D$ is also a Yang-Mills connection. Thus the space of Yang-Mill connections on a vector bundle E of rank $n > 1$ is either empty or infinite. To fix this problem, we can seek solutions of the Yang-Mills equations modulo gauge transformations by a gauge group G.

5. Short Excursion into Hyperbolic Yang-Mills: The Maxwell Equations

So far, we have only considered Riemannian manifolds. For many applications to physics, we need a slight generalization and allow so-called *pseudo-Riemannian* manifolds. The difference is that we now only require the inner product on the tangent space to be nondegenerate; the positive-definitive condition is dropped. Only in this section will we consider pseudo-Riemannian manifolds; for all other sections we will return to working with a Riemannian manifold. The goal of this section is simply to do some computations with the Yang-Mills equations and see how the Maxwell's equations can be derived from them, and it is not logically required for the subsequent sections.

An important example of an nondegenerate inner product that is not positive-definite is the following:

Example 5.1. (Lorentz Inner Product) *Let V be a four dimensional vector space. If $x = (x^0, x^1, x^2, x^3)$ and $y = (y^0, y^1, y^2, y^3)$, we define*

$$(x, y) = -x^0y^0 + x^1y^1 + x^2y^2 + x^3y^3.$$

The definition of orthonormality is tweaked in the following way: an *orthonormal basis* of V is a basis $\{e_1, \ldots, e_n\}$ of V such that

$$(e_i, e_j) = \pm\delta^{ij}.$$

The definition of the Hodge star operator is also altered slightly. Select a positive orthonormal basis $\{e_1, \ldots, e_d\}$ for V and to define the star operator on a basis element as follows:

$$*(e_{i_1} \wedge \cdots \wedge e_{i_p}) = (e_{i_1}, e_{i_1}) \ldots (e_{i_p}, e_{i_p})\ e_{j_1} \wedge \cdots \wedge e_{j_{d-p}},$$

where $\{e_{j_1}, \ldots, e_{j_{d-p}}\}$ is chosen such that $\{e_{i_1}, \ldots, e_{i_p}, e_{j_1}, \ldots, e_{j_{d-p}}\}$ is a positive orthonormal basis for V. Following the same argument as before, it can be shown that the Hodge operator is well-defined.

Example 5.2. *Take $M = \mathbb{R}^4$ with oriented cotangent basis $\{dx^0, dx^1, dx^2, dx^3\}$ and Lorentz inner product $g^{ij} = \mathrm{diag}(-1, 1, 1, 1)$. We compute*

$*(dx^0 \wedge dx^1) = -dx^2 \wedge dx^3$
$*(dx^0 \wedge dx^2) = dx^1 \wedge dx^3$

$*(dx^0 \wedge dx^3) = -dx^1 \wedge dx^2$
$*(dx^1 \wedge dx^2) = dx^0 \wedge dx^3$
$*(dx^1 \wedge dx^3) = -dx^0 \wedge dx^2$
$*(dx^2 \wedge dx^3) = dx^0 \wedge dx^1$

With these generalizations, we can look at the Yang-Mills equations over a pseudo-Riemannian manifold. Let us compute the Yang-Mills equations explicitly for the case $M = \mathbb{R}^4$ with the Lorentz metric $g^{ij} = \mathrm{diag}(-1, 1, 1, 1)$. We take the positive orientation for the cotangent bundle to be $\{dx^0, dx^1, dx^2, dx^3\}$. The Yang-Mills equation reads

$$D_i F_{jk} \otimes dx^i \wedge *(dx^j \wedge dx^k) = 0.$$

Since we must equate the coefficient of each of the wedges $dx^\alpha \wedge dx^\beta \wedge dx^\gamma$ to zero, we have $\binom{4}{3} = 4$ distinct equations. The first thing to notice is that if i, j, k are all distinct, then the right-hand side is trivially zero because of a repeated index in the wedge product. We compute explicitly the coefficient in front of $dx^1 \wedge dx^2 \wedge dx^3$, which turns out to be

$$2(D_1 F_{10} + D_2 F_{20} + D_3 F_{30}).$$

The factor of 2 comes from the skew-symmetry $F_{ij} = -F_{ji}$ and the skew-symmetry of the wedge product. Substituting the expression (7) for $D_i F_{jk}$ and equating the whole thing to zero yields the first equation

$$\partial_1 F_{10} + [A_1, F_{10}] + \partial_2 F_{20} + [A_2, F_{20}] + \partial_3 F_{30} + [A_3, F_{30}] = 0.$$

We now compute the coefficient in front of the $dx^0 \wedge dx^a \wedge dx^b$ term, where $a, b \in \{1, 2, 3\}$ and $a < b$. Let $c \in \{1, 2, 3\} \backslash \{a, b\}$. Then the coefficient in front of the $dx^0 \wedge dx^a \wedge dx^b$ is

$$2(-D_0 F_{0c} + D_a F_{ac} + D_b F_{bc}).$$

Thus we have three equations, one for each $c \in \{1, 2, 3\}$:

$$-\partial_0 F_{0c} - [A_0, F_{0c}] + \sum_{i=1}^{3} \partial_i F_{ic} + [A_i, F_{ic}] = 0.$$

Adding the first equation, the Yang-Mills equations for $d = 4$ with Lorentz metric become the following set of four equations:

$$-\partial_0 F_{0\beta} - [A_0, F_{0\beta}] + \sum_{i=1}^{3} \partial_i F_{i\beta} + [A_i, F_{i\beta}] = 0, \tag{10}$$

where $\beta \in \{0, 1, 2, 3\}$.

Using this formulation, we can derive Maxwell's equations for 3 dimensional space in vacuum from the Yang-Mills equations. The Maxwell's equations for 3 dimensional space in vacuum are:

$$\partial_t E = \nabla \times B, \qquad \partial_t B = -\nabla \times E,$$

and

$$\nabla \cdot E = 0, \qquad \nabla \cdot B = 0,$$

where $E, B : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$.

Consider the Yang-Mills equation in $d = 4$ Minkowski space over a rank $n = 2$ vector bundle. As discussed in the chapter on Hodge theory, since skew-symmetric 2×2 matrices commute, the space $\mathrm{Ad}(E)$ reduces to the trivial bundle and the endomorphisms A_i, F_{jk} are real (we can generalize this to complex) numbers. Since all commutators involving A vanish, the Yang-Mills equations (10) then read

$$-\partial_0 F_{0\beta} + \sum_{i=1}^{3} \partial_i F_{i\beta} = 0, \tag{11}$$

for $\beta \in \{0, 1, 2, 3\}$. Since F is a curvature operator, it must satisfy the Bianchi identiy:

$$DF = \partial_k F_{ij} \otimes dx^k \wedge dx^i \wedge dx^j = 0.$$

Therefore, after setting the coefficient of each $dx^\alpha \wedge dx^\beta \wedge dx^\gamma$ terms to zero, we see that F must also satisfy

$$\partial_\alpha F_{\beta\gamma} + \partial_\gamma F_{\alpha\beta} + \partial_\beta F_{\gamma\alpha} = 0, \tag{12}$$

for any $\alpha, \beta, \gamma \in \{0, 1, 2, 3\}$.

The last step is to write the matrix with entries F_{ij} in the following way

$$F_{ij} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix}$$

We identify $\partial_0 = \partial_t$, $\partial_1 = \partial_x$, $\partial_2 = \partial_y$, $\partial_3 = \partial_z$, and read off (11) for $\beta = 0$:

$$\partial_x E_x + \partial_y E_y + \partial_z E_z = 0.$$

The cases $\beta = 1, 2, 3$ yield respectively

$$\begin{aligned}
\partial_t E_x &= \partial_y B_z - \partial_z B_y, \\
\partial_t E_y &= \partial_z B_x - \partial_x B_z, \\
\partial_t E_z &= \partial_x B_y - \partial_y B_x.
\end{aligned}$$

Hence we have recovered two of the Maxwell equations: $\nabla \cdot E = 0$ and $\partial_t E = \nabla \times B$.

We now read off (12) for $(\alpha, \beta, \gamma) = (123)$:

$$\partial_x B_x + \partial_y B_y + \partial_z B_z = 0.$$

For the cases $(\alpha, \beta, \gamma) = (012), (013), (023)$ we obtain respectively

$$\begin{aligned}
\partial_t B_z &= \partial_y E_x - \partial_x E_y, \\
\partial_t B_y &= \partial_x E_z - \partial_z E_x, \\
\partial_t B_x &= \partial_z E_y - \partial_y E_z.
\end{aligned}$$

We have thus obtained the last two Maxwell equations: $\nabla \cdot B = 0$ and $\partial_t B = -\nabla \times E$.

6. Hodge Theory

We shall see that for two dimensional vector bundles, Yang-Mills theory reduces to Hodge theory. The goal of this section is to prove Hodge's theorem. The proofs in this section follow [1].

Throughout this section, we let M be a compact, oriented, Riemannian manifold of dimension d. Using the Hodge star operator $* : \Omega^p(M) \to \Omega^{d-p}(M)$, we define an L^2 inner product on the space $\Omega^p(M)$:

$$(\alpha, \beta) = \int_M \alpha \wedge *\beta,$$

for any $\alpha, \beta \in \Omega^p(M)$. As usual, the norm is given by $||\alpha||_{L^2}^2 = (\alpha, \alpha)$.

Recall the exterior derivative $d : \Omega^p(M) \to \Omega^{p+1}(M)$. We define d^* as the adjoint of d with respect to the inner product $(.,.)$.

$$d^* : \Omega^{p+1}(M) \to \Omega^p(M),$$
$$(d\alpha, \beta) = (\alpha, d^*\beta),$$

for any $\alpha \in \Omega^p(M)$, $\beta \in \Omega^{p+1}(M)$.

We use the definitions of d, d^* to introduce the notion of a harmonic p-form.

Definition 6.1. *The Laplace-Beltrami operator* $\Delta : \Omega^p(M) \to \Omega^p(M)$ *is defined as*

$$\Delta\alpha = dd^*\alpha + d^*d\alpha.$$

We say that an element $\alpha \in \Omega^p(M)$ is *harmonic* if $\Delta\alpha = 0$. It is immediate from the definition that the Laplace-Beltrami operator is self-adjoint:

$$(\Delta\alpha, \beta) = (\alpha, \Delta\beta),$$

for any $\alpha, \beta \in \Omega^p(M)$.

An equivalent definition of harmonicity of a p-form α is that $d\alpha = 0$ and $d^*\alpha = 0$. Indeed, first suppose $\Delta\alpha = 0$. Then

$$0 = (\Delta\alpha, \alpha) = (dd^*\alpha, \alpha) + (d^*d\alpha, \alpha) = (d^*\alpha, d^*\alpha) + (d\alpha, d\alpha).$$

Since $||d^*\alpha||_{L^2}^2 + ||d\alpha||_{L^2}^2 = 0$, we must have $d\alpha = 0$ and $d^*\alpha = 0$. On the other hand, if

$d\alpha = 0$ and $d^*\alpha = 0$, then $dd^*\alpha + d^*d\alpha = 0$.

For computations involving d^*, it is usually easier way to think of the operator in the following way.

Proposition 6.1. *The operator* $d^* : \Omega^p(M) \to \Omega^{p-1}(M)$ *satisfies, for any* $\beta \in \Omega^p(M)$,

$$d^*\beta = (-1)^{d(p-1)+1} * d * \beta.$$

Proof. Let $\alpha \in \Omega^{p-1}(M)$. Since $\beta \in \Omega^p(M)$, we have $d * \beta \in \Omega^{d-p+1}(M)$. By Proposition (3.1), $** d * \beta = (-1)^{(d-p+1)(p-1)} d * \beta$. We compute using the Leibniz rule for exterior derivatives:

$$d(\alpha \wedge *\beta) = d\alpha \wedge *\beta + (-1)^{p-1}\alpha \wedge d * \beta = d\alpha \wedge *\beta + (-1)^{(p-1)} \, (-1)^{(d-p+1)(p-1)}\alpha \wedge (**) d * \beta.$$

Since even powers of (-1) are the identity, we compute

$$(-1)^{(p-1)+(d-p+1)(p-1)} = (-1)^{2p-2+p(p-1)+d(p-1)} = (-1)^{d(p-1)}.$$

Therefore, after integrating we obtain

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

By Stokes' theorem, the left hand side is zero, and hence by definition of the L^2 inner product on forms we obtain

$$(d\alpha, \beta) = (\alpha, (-1)^{d(p-1)+1} * d * \beta).$$

□

We now recall the definition of DeRham cohomology. For any p-form α, if $d\alpha = 0$, we say α is *closed*. If there exists a $p-1$ form η such that $\alpha = d\eta$, we say that α is *exact*. Since $dd\eta = 0$ for any $p-1$ form η, exact forms are closed. It therefore makes sense to define the quotient vector space

$$H^p_{dR}(M) := \frac{\{\text{Closed } p \text{ forms}\}}{\{\text{Exact } p \text{ forms}\}}.$$

In other words, we identify two elements $\alpha_1, \alpha_2 \in \Omega^p(M)$ to be the same element of $H^p_{dR}(M)$ (we also say they are *cohomologous*) if $\alpha_1 - \alpha_2 = d\eta$ for some $\eta \in \Omega^{p-1}(M)$. This partitions the vector space of closed p-forms into *cohomology classes*.

Hodge theory looks at harmonic forms in $H^p_{dR}(M)$. The main theorem of this section states that every cohomology class has a unique harmonic representative. To prove this theorem, we use Sobolev spaces and standard elliptic theory. We state the results from standard elliptic theory that will be used before proceeding to prove the theorem.

First we recall the definition of Sobolev space. For $u \in L^1_{\text{loc}}(U)$, we say that $\partial^\alpha u = v$ is the αth weak derivative of u if for all $\varphi \in C^\infty_c(U)$ we have

$$\int_U u \partial^\alpha \varphi = (-1)^{|\alpha|} \int_U v\varphi.$$

We can define Sobolev spaces using this notion of weak derivatives. Let $k \in \mathbb{N}$, and define

$$H^{k,2}(\Omega) = \{f \in L^2(\Omega) : \int_\Omega |\partial^\alpha f|^2 < \infty, \text{ for all } |\alpha| \leq k\}$$

$$||f||_{H^{k,2}(\Omega)} = \left(\sum_{|\alpha| \leq k} \int_\Omega |\partial^\alpha f|^2 \right)^{1/2}.$$

The space $H^{k,2}$ is a Banach space. We denote the Banach space $H^{k,2}_0(\Omega)$ as the closure of $C^\infty_c(\Omega)$ with respect to the Sobolev norm given above.

Theorem 6.1. Sobolev Lemma *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Let k, s be positive integers such that $s > n/2 + k$. Then $H^{s,2}_0(\Omega)$ is continuously embedded into $C^k(\overline{\Omega})$.*

Theorem 6.2. Rellich-Kondrachov Theorem *Let Ω be a bounded domain. Then $H^{1,2}_0(\Omega)$ is compactly embedded into $L^2(\Omega)$.*

Theorem 6.3. Elliptic Regularity *Let $u \in H^{1,2}(\Omega)$ be a weak solution of the equation $\Delta u = f$:*

$$\int_\Omega u \Delta \varphi = \int_\Omega f \varphi,$$

for all $\varphi \in H^{1,2}_0(\Omega)$. If $f \in H^{s,2}(\Omega)$, then $u \in H^{s+2,2}(\tilde{\Omega})$ for every $\tilde{\Omega}$ whose closure is a compact subset of Ω.

We adapt the definition of the Sobolev spaces such that the theory can be applied to $\Omega^p(M)$. We define the norm

$$||\alpha||^2_{H^{1,2}(M)} := (d\alpha, d\alpha) + (d^*\alpha, d^*\alpha) + (\alpha, \alpha).$$

We call $H^{1,2}_p(M)$ the completion of the space $\Omega^p(M)$ with respect to the norm $||\cdot||^2_{H^{1,2}(M)}$. It turns out that, locally, this norm is equivalent to the Euclidean Sobolev norm defined above. The proof uses results from Riemannian geometry involving the existence of normal coordinates, and will be omitted. The important conclusion is that we can apply all theorems developed for Sobolev spaces in a Euclidean setting to our compact, oriented Riemannian manifold.

Lemma 6.1. *Let $\beta \in \Omega^p(M)$ be a closed form that is orthogonal to the kernel of d^*. Then there exists a constant $C > 0$ such that*

$$(\beta, \beta) \leq C(d^*\beta, d^*\beta).$$

Proof. Proceed by contradiction. Suppose there exists a sequence $\{\beta_n\}$ of closed forms $\beta_n \in \Omega^p(M)$ such that each β_n is orthogonal to the kernel of d^*, and

$$(\beta_n, \beta_n) \geq n(d^*\beta_n, d^*\beta_n).$$

We define the following scaling factor

$$\lambda_n := \frac{1}{(\beta_n, \beta_n)^{1/2}}.$$

We then have

$$\begin{aligned} 1 &= (\lambda_n\beta_n, \lambda_n\beta_n) \\ &= \lambda_n^2(\beta_n, \beta_n) \\ &\geq \lambda_n^2 n(d^*\beta_n, d^*\beta_n) \\ &= n(d^*\lambda_n\beta_n, d^*\lambda_n\beta_n). \end{aligned}$$

Using the above bound and the fact that β_n are closed, we can bound the sequence $\{\lambda_n\beta_n\}$ in the Sobolev norm:

$$||\lambda_n\beta_n||^2_{H^{1,2}(M)} = (d\lambda_n\beta_n, d\lambda_n\beta_n) + (d^*\lambda_n\beta_n, d^*\lambda_n\beta_n) + (\lambda_n\beta_n, \lambda_n\beta_n) \leq \frac{1}{n} + 1.$$

Since M is a compact manifold, we can use the Rellich-Kondrachov theorem on the bounded sequence $\{\lambda_n\beta_n\} \subset H^{1,2}(M)$. We then use sequential compactness to extract a subsequence $\{\lambda_{n_k}\beta_{n_k}\}$ that converges in the L^2 norm to a p-form $\psi \in H^{1,2}(M)$. For convenience, we replace the original sequence with this new subsequence; i.e. assume that $\{\lambda_n\beta_n\}$ converges to ψ in the L^2 norm.

Since $(d^*\lambda_n\beta_n, d^*\lambda_n\beta_n) \leq 1/n$, we see that $d^*\lambda_n\beta_n \to 0$ in L^2. Then for arbitrary $\varphi \in \Omega^p(M)$, we have

$$\begin{aligned}(d^*\psi, \varphi) &= (\psi, d\varphi) \\ &= \lim(\lambda_n\beta_n, d\varphi) \\ &= \lim(d^*\lambda_n\beta_n, \varphi) = 0.\end{aligned}$$

Since this holds for all $\varphi \in \Omega^p(M)$, we have that $d^*\psi = 0$. But then since the β_n are orthogonal to the kernel of d^*, we know that $(\lambda_n\beta_n, \psi) = 0$. This leads to the following contradiction: on one hand, $||\psi||_{L^2} = 1$ since ψ is the L^2 limit of the sequence $\{\lambda_n\beta_n\}$ where $||\lambda_n\beta_n||_{L^2} = 1$ for all positive integers n. On the other hand,

$$||\psi||_{L^2} = (\lim \lambda_n\beta_n, \psi) = \lim(\lambda_n\beta_n, \psi) = 0.$$

□

The lemma will be used to prove the Hodge theorem, which is the main theorem of this section.

Theorem 6.4. Hodge Theorem *Let M be a compact Riemannian manifold of dimension d. Let $0 \leq p \leq d$. In each de Rham cohomology class of $H^p_{dR}(M)$, there exists a unique harmonic representative.*

Proof. We start by showing uniqueness. If $p = 0$ is this is trivial. If $p > 0$, suppose there exists two elements $\omega_1, \omega_2 \in \Omega^p(M)$, both harmonic, such that $\omega_1 - \omega_2 = d\eta$ for some $\eta \in \Omega^{p-1}(M)$. Then since $d^*\omega_1 = d^*\omega_2 = 0$ by harmonicity, we have

$$||\omega_1 - \omega_2||^2_{L^2} = (\omega_1 - \omega_2, d\eta) = (d^*(\omega_1 - \omega_2), \eta) = 0.$$

Therefore $\omega_1 = \omega_2$.

The proof of existence is more involved. Fix a closed form $\omega_0 \in \Omega^p(M)$. We want to find a harmonic element $\omega \in \Omega^p(M)$ such that $\omega - \omega_0 = d\eta$ for some $\eta \in \Omega^{p-1}(M)$. The strategy is to minimize the functional $||\cdot||_{L^2}$ over a certain family of p-forms in order to obtain the candidate cohomologous harmonic form. Let $(L^2)^p(M)$ denote the space of sections of $\Lambda^p(M)$ such that $||\alpha||_{L^2} < \infty$. Define

$\mathcal{F} = \{\omega \in (L^2)^p(M) : \text{there exists } \alpha \in (L^2)^{p-1}(M) \text{ such that for all } \varphi \in \Omega^p(M), (\omega - \omega_0, \varphi) = (\alpha, d^*\varphi)\}.$

We denote the infimum of the L^2 norm over this family by b:

$$b = \inf_{\omega \in \mathcal{F}} ||\omega||_{L^2}^2.$$

Take a minimizing sequence $\{\omega_n\} \subset \mathcal{F}$ such that $||\omega_n||_{L^2}^2 \to b$. Without loss of generality, we can assume that $||\omega_n||_{L^2}^2 \leq 2b$. By the Banach-Alaoglu theorem, every norm bounded sequence in L^2 has a weakly convergent subsequence. In other words, there exists a subsequence $\{\omega_{n_k}\}$ and $\omega \in (L^2)^p(M)$ such that for all $\varphi \in \Omega^p(M)$,

$$(\omega_{n_k}, \varphi) \to (\omega, \varphi).$$

After relabelling and removing terms from the original sequence, we may assume that $\{\omega_n\}$ converges weakly to ω. The goal is now to show that ω is harmonic. We will show that ω is weakly harmonic, and then use standard elliptic theory to complete the proof. The first step is to show that $\omega \in \mathcal{F}$.

Define $\eta = \omega - \omega_0$. Then for all $\varphi \in \Omega^p(M)$ such that $d^*\varphi = 0$, we have $(\eta, \varphi) = 0$:

$$(\eta, \varphi) = \lim(\omega_n - \omega_0, \varphi) = \lim(\alpha_n, d^*\varphi) = 0.$$

Using this fact, we construct a linear functional f on the space $d^*(\Omega^p(M))$. We define f as

$$f(d^*\varphi) = (\eta, \varphi),$$

for all $\varphi \in \Omega^p(M)$. We must check that f is well-defined. Indeed, if $d^*\varphi_1 = d^*\varphi_2$, then

$$f(d^*\varphi_1) - f(d^*\varphi_2) = (\eta, \varphi_1) - (\eta, \varphi_2) = (\eta, \varphi_1 - \varphi_2) = 0.$$

We show that f is bounded. Define π to be the orthogonal projection onto the kernel of d^*. Then

$$f(d^*\varphi) = f(d^*(\varphi - \pi(\varphi))) = (\eta, \varphi - \pi(\varphi)).$$

By Lemma 6.1, since $\varphi - \pi(\varphi)$ is in the orthogonal complement of the kernel of d^*, we have

$$||\varphi - \pi(\varphi)||_{L^2} \leq C \; ||d^*(\varphi - \pi(\varphi))||_{L^2} = C||d^*(\varphi)||_{L^2}.$$

Using this estimate and the Cauchy-Schwarz-Bunyakovsky inequality, we can show that f is a bounded linear functional:

$$|f(d^*\varphi)| = |(\eta, \varphi - \pi(\varphi))| \leq ||\eta||_{L^2}\ ||\varphi - \pi(\varphi)||_{L^2} \leq C||\eta||_{L^2}\ ||d^*\varphi||_{L^2}.$$

By density, we can extend f to the L^2 closure of $d^*(\Omega^p(M))$. This allows us to use the Riesz representation theorem to obtain an $\alpha \in (L^2)^p(M)$ such that $f(d^*\varphi) = (\alpha, d^*\varphi)$. By the definition of f, we have shown that for all $\varphi \in \Omega^p(M)$,

$$(\eta, \varphi) = (\alpha, d^*\varphi).$$

It follows that $\omega \in \mathcal{F}$. Also, for any $\varphi \in \Omega^{p+1}(M)$,

$$(\omega, d^*\varphi) = (\omega_0, d^*\varphi) + (\eta, d^*\varphi) = (d\omega_0, \varphi) + (\alpha, d^*d^*\varphi) = 0.$$

(If for any $\varphi \in \Omega^{p+1}(M)$, we have $(\omega, d^*\varphi) = 0$, we say that $d\omega = 0$ weakly.)

Since $\omega \in \mathcal{F}$, we know that $b \leq ||\omega||_{L^2}^2$. On the other hand, by weak lower semicontinuity of the L^p norm, if g_i converges to g weakly in L^p, then $||g||_p \leq \liminf ||g_i||_p$. It follows that

$$||\omega||_{L^2}^2 \leq \liminf ||\omega_n||_{L^2}^2 = b.$$

Hence $||\omega||_{L^2}^2 = b$ and ω is the minimizer of the functional $||\cdot||_{L^2}^2$ in $\mathcal{F}$.

We now show that ω is weakly harmonic. It was already shown that that $d\omega = 0$ weakly, so it only remains to show that $d^*\omega = 0$ weakly. (If for any $\varphi \in \Omega^{p-1}(M)$, we have $(\omega, d\varphi) = 0$, we say that $d^*\omega = 0$ weakly.) By calculus of variation, since ω is the minimizer of the functional $||\cdot||_{L^2}^2$ in $\mathcal{F}$, we must have

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} ||\omega + \varepsilon d\varphi||_{L^2}^2 = 0$$

for all $\varphi \in \Omega^{p-1}(M)$. (Indeed, $\omega + \varepsilon d\varphi \in \mathcal{F}$ since $(\omega - \omega_0 + \varepsilon d\varphi, \psi) = (\alpha + \varepsilon\varphi, d^*\psi)$ for all $\psi \in \Omega^p(M)$.) Passing the derivative under the integral is not a problem since the derivative with respect to ε of the integrand is absolutely integrable. (One can use the dominated convergence theorem.) We thus compute

$$\begin{aligned}
0 &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \int_M (\omega + \varepsilon d\varphi) \wedge *(\omega + \varepsilon d\varphi) \\
&= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \int_M \omega \wedge *\omega + \omega \wedge *\varepsilon d\varphi + \varepsilon d\varphi \wedge *\omega + \varepsilon^2 d\varphi \wedge *d\varphi \\
&= \int_M \omega \wedge *d\varphi + d\varphi \wedge *\omega \\
&= 2(\omega, d\varphi).
\end{aligned}$$

Therefore, $d^*\omega = 0$ weakly. It follows that ω is weakly harmonic:

$$(\omega, \Delta\varphi) = 0,$$

for all $\varphi \in \Omega^p(M)$. A bootstrap argument finishes the proof. By elliptic regularity, $\omega \in H^{k,2}(M)$ for all positive integers k. By the Sobolev lemma, $\omega \in C^k$ for all k, hence $\omega \in \Omega^p(M)$.

□

Now that we have established some Hodge theory, we return to the Yang-Mills equations and show that it reduces to Hodge theory in the case of a vector bundle $E \to M$ of rank $n = 2$.

Recall that $\mathrm{Ad}(E)$ is the vector bundle over M whose fiber at $p \in M$ is the set of skew-symmetric endomorphisms of the fiber E_p. For the case $n = 2$, a skew symmetric endomorphism is of the form

$$\begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix}$$

where $x \in \mathbb{R}$. Furthermore, these matrices commute:

$$\begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix}\begin{pmatrix} 0 & y \\ -y & 0 \end{pmatrix} = \begin{pmatrix} 0 & -xy \\ -xy & 0 \end{pmatrix} = \begin{pmatrix} 0 & y \\ -y & 0 \end{pmatrix}\begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix}.$$

In this case, a skew-symmetric endomorphisms of the fiber E_p is determined by selecting a single real number, and furthermore, all Lie brackets vanish. We can thus view $\mathrm{Ad}(E)$ as the trivial bundle $M \times \mathbb{R}$.

Locally, the curvature is of the form $F = F_{jk} \otimes dx^j \wedge dx^k$ where F_{jk} is now a real number. Since all commutators vanish, we have $D_i F_{jk} = \partial_i F_{jk}$ (7). Hence the Bianchi identity (2.1) becomes

$$dF = 0,$$

where d is the usual exterior derivative. Therefore, the curvature F is a closed 2-form. The Yang-Mills equations becomes $*D*F = *d*F = 0$, which by Proposition 6.1 means that

$$d^*F = 0. \tag{13}$$

Hence a connection D satisfies the Yang-Mills equations if and only if it is harmonic. Thus, in the case $n = 2$, Yang-Mills theory reduces to Hodge theory.

7. Variational Formulation

The Yang-Mills equations arise from minimizing a functional called the Yang-Mills functional. In this section we will define the Yang-Mills functional, and recover the Yang-Mills equations from it. Recall that M is always assumed to be an oriented, compact Riemannian manifold of dimension d. The first task is to define an L^2 product on $\Omega^p(\mathrm{Ad}E)$.

Let $A_1, A_2 \in \mathfrak{o}(n)$. We define

$$A_1 \cdot A_2 := -\mathrm{tr}(A_1 A_2). \tag{14}$$

This defines a positive definite inner product on the space of skew-symmetric $n \times n$ matrices $\mathfrak{o}(n)$.

For p-forms $\omega_1, \omega_2 \in \Lambda^p T_x^* M$, we define the following inner product:

$$\langle \omega_1, \omega_2 \rangle := *(\omega_1 \wedge *\omega_2).$$

Putting these two together, for $\mu_1, \mu_2 \in \mathrm{Ad}E_x$, $\omega_1, \omega_2 \in \Lambda^p T_x^* M$, we obtain

$$\langle \mu_1 \otimes \omega_1, \mu_2 \otimes \omega_2 \rangle := \mu_1 \cdot \mu_2 \ \langle \omega_1, \omega_2 \rangle. \tag{15}$$

Extending this linearly, we obtain an inner product on $\Omega^p(\mathrm{Ad}E)$. For $\mu \in \Omega^p(\mathrm{Ad}E)$, we will denote

$$|\mu|^2 := \langle \mu, \mu \rangle.$$

We now define an L^2 inner product on $\Omega^p(\mathrm{Ad}E)$, and denote it by round brackets:

$$(\mu_1 \otimes \omega_1, \mu_2 \otimes \omega_2) := \int_M \langle \mu_1 \otimes \omega_1, \mu_2 \otimes \omega_2 \rangle * (1).$$

Here, we use $*(1)$ to denote the volume form of M. A computation using the definition of the Hodge star operator (which we omit) shows that

$$*(1) = \sqrt{\det g_{ij}} dx^1 \wedge \cdots \wedge dx^d.$$

For $\nu \in \Omega^p(\mathrm{Ad}E)$, we will denote

$$||\nu||_{L^2}^2 = (\nu, \nu).$$

We can thus define the operator adjoint to D with respect to this L^2 inner product:

$$D^* : \Omega^p(\mathrm{Ad}E) \to \Omega^{p-1}(\mathrm{Ad}E),$$
$$(D^*\nu, \mu) = (\nu, D\mu),$$

for all $\mu \in \Omega^{p-1}(\mathrm{Ad}E), \nu \in \Omega^p(\mathrm{Ad}E)$.

From Proposition (6.1), we know that the adjoint $d^* : \Omega^p(M) \to \Omega^{p-1}(M)$ can be written of the form $d^* = (-1)^{d(p-1)+1} * d*$. Using this fact, it can be shown that

$$D^* = (-1)^{d(p-1)+1} * D * . \tag{16}$$

Definition 7.1. *Let M be a compact, oriented, Riemannian manifold. Let E be a vector bundle over M, with bundle metric. Let D be a metric connection on E, and let $F_D \in \Omega^2(\mathrm{Ad}E)$ be its curvature. We define the Yang-Mills functional as the following:*

$$YM(D) := ||F_D||^2_{L^2} = \int_M |F_D|^2 * (1) = \int_M \langle F_D, F_D\rangle * (1).$$

Suppose D is a critical point of the Yang-Mills functional. We consider variations of the form $D + \varepsilon B$, where $B \in \Omega^1(\mathrm{Ad}E)$. The curvature can be found using (8). If $\sigma \in \Gamma(E)$, then using (5), we obtain

$$\begin{aligned}
\frac{1}{2}F_{D+\varepsilon B}\sigma &= (D+\varepsilon B)(D+\varepsilon B)\sigma \\
&= D\circ D\sigma + \varepsilon D_i(B_j\sigma)\otimes dx^i\wedge dx^j + \varepsilon B_j D_i\sigma\otimes dx^j\wedge dx^i + \varepsilon^2 B_i B_j\sigma\otimes dx^i\wedge dx^j \\
&= D\circ D\sigma + \varepsilon(D_iB_j)\sigma\otimes dx^i\wedge dx^j + \varepsilon B_j D_i\sigma\otimes dx^i\wedge dx^j + \varepsilon B_j D_i\sigma\otimes dx^j\wedge dx^i \\
&\quad + \varepsilon^2 B_i B_j\sigma\otimes dx^i\wedge dx^j \\
&= (D\circ D + \varepsilon(DB) + \varepsilon^2 B\wedge B)\sigma
\end{aligned}$$

Therefore, the Euler-Lagrange equations yield

$$0 = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} YM(D+\varepsilon B) = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0}\int_M \langle F_{D+\varepsilon B}, F_{D+\varepsilon B}\rangle * (1)$$

By substituting the above expression and dividing out a constant, we obtain

$$0 = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \int_M \langle D \circ D, D \circ D\rangle + \varepsilon(\langle D \circ D, DB\rangle + \langle DB, D \circ D\rangle) + \mathcal{O}(\varepsilon^2).$$

By symmetry of the inner product $\langle \cdot, \cdot \rangle$, we have

$$0 = 2(DB, D \circ D).$$

Hence by (8), we obtain the condition

$$0 = (DB, F_D) = (B, D^* F_D).$$

Since this holds for all $B \in \Omega^1(\mathrm{Ad}E)$, we must have

$$D^* F_D = 0.$$

By (16), this condition is equivalent to requiring D to be a Yang-Mills connection. Therefore, critical points of the Yang-Mills functional obey the Yang-Mills equations.

8. Manifolds of Dimension $d = 4$

Throughout this entire section, we always assume that M is an orientable, compact Riemannian manifold of dimension $d = 4$. The Hodge star operator acts in a nice way on the 6 dimensional vector space $\Lambda^2(T_p^*M)$ as we shall now see. Suppose $g^{ij}(p) = \delta^{ij}$. It can be proved using techniques from Riemannian geometry that this can be assumed without loss of generality; coordinates with this feature are called normal coordinates centered at p. Define the following basis for $\Lambda^2(T_p^*M)$:

$$\begin{aligned} e_1 &= dx^1 \wedge dx^2, \quad e_2 = dx^1 \wedge dx^3, \quad e_3 = dx^1 \wedge dx^4, \\ e_4 &= dx^2 \wedge dx^3, \quad e_5 = dx^2 \wedge dx^4, \quad e_6 = dx^3 \wedge dx^4. \end{aligned}$$

As computed in Exercise (3.1), we have

$$\begin{aligned} *e_1 &= e_6, \quad *e_2 = -e_5, \quad *e_3 = e_4, \\ *e_4 &= e_3, \quad *e_5 = -e_2, \quad *e_6 = e_1. \end{aligned}$$

With respect to this basis, the Hodge $*$ operator has the following matrix form:

$$* = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The operator $*$ thus has two eigenvalues: 1 and -1. We have the decomposition $\Lambda^2(T_p^*M) = \Lambda^+ \oplus \Lambda^-$, where Λ^+ is the space spanned by the eigenvectors with eigenvalue 1, while Λ^- is the space spanned by the eigenvectors with eigenvalue -1. The eigenvectors associated with the eigenvalue 1 are

$$e_1 + e_6, \quad e_2 - e_5, \quad e_3 + e_4.$$

The eigenvectors associated with the eigenvalue -1 are

$$e_1 - e_6, \quad e_2 + e_5, \quad e_3 - e_4.$$

Elements of Λ^+ are called selfdual, while elements of Λ^- are called antiselfdual. Explicitly, a basis for Λ^+ is given by

$$f_1^+ = dx^1 \wedge dx^2 + dx^3 \wedge dx^4, \quad f_2^+ = dx^1 \wedge dx^3 - dx^2 \wedge dx^4, \quad f_3^+ = dx^1 \wedge dx^4 + dx^2 \wedge dx^3.$$

A basis for Λ^- is given by

$$f_1^- = dx^1 \wedge dx^2 - dx^3 \wedge dx^4, \quad f_2^- = dx^1 \wedge dx^3 + dx^2 \wedge dx^4, \quad f_3^- = dx^1 \wedge dx^4 - dx^2 \wedge dx^3.$$

A computation left to the reader shows that

$$f_i^+ \wedge f_j^- = 0. \tag{17}$$

We can naturally generalize the notion of selfdual and antiselfdual to elements of $\Omega^2(\text{Ad}E)$. This allows us to decompose an arbitrary $\mu \in \Omega^2(\text{Ad}E)$ as $\mu = \mu^+ + \mu^-$, where $*\mu^+ = \mu^+$ and $*\mu^- = -\mu^-$. The important fact about this decomposition is that it is orthogonal with respect to the inner product $\langle \cdot, \cdot \rangle$ defined in (15). Indeed, we can write

$$\begin{aligned} \mu^+ &= \mu_1^+ \otimes f_1^+ + \mu_2^+ \otimes f_2^+ + \mu_3^+ \otimes f_3^+, \\ \mu^- &= \mu_1^- \otimes f_1^- + \mu_2^- \otimes f_2^- + \mu_3^- \otimes f_3^-. \end{aligned}$$

To show orthogonality, we notice

$$\langle \mu_i^+ \otimes f_i^+, \mu_j^- \otimes f_j^- \rangle = \mu_i^+ \cdot \mu_j^- \; * (f_i^+ \wedge *f_j^-) = 0,$$

since $f_i^+ \wedge *f_j^- = -f_i^+ \wedge f_j^- = 0$ by (17). This orthogonal decomposition into selfdual and antiselfdual components will be an important ingredient when minimizing the Yang-Mills functional over a 4 dimensional manifold.

A metric connection D is called *selfdual instanton* if its curvature $F = F_{jk} \otimes dx^j \wedge dx^k$ is a selfdual 2-form. An *antiselfdual instanton* is defined similarly. The importance of selfdual and antiselfdual instantons comes from the following theorem:

Theorem 8.1. *If a metric connection D is selfdual or antiselfdual, then D is a Yang-Mills connection.*

Proof. Suppose $*F = \pm F$. Then the Bianchi identity yields

$$D * F = \pm DF = 0.$$

Hence D is a Yang-Mills connection.

□

The next objective will be to show that the converse also holds; that is, if D minimizes the Yang-Mills functional, then D is a selfdual or antiselfdual instanton. This will thus reduce the problem of solving the Yang-Mills equations to finding selfdual and antiselfdual instantons.

Before proceeding, we introduce the concept of Chern classes. Let E be a complex vector bundle of rank n over M. We recall the elementary symmetric polynomials P^j, which satisfy

$$\prod_{k=1}^{n}(1+x_k\tau)=\sum_{j=0}^{n}P^j(x_1,\ldots,x_n)\tau^j.$$

We define

$$P^j:\mathrm{Mat}(n\times n,\mathbb{C})\to\mathbb{C},$$

such that $P^j(B)$ is the elementary symmetric polynomial homogeneous of degree j of the eigenvalues of B. For example, if λ_i are the eigenvalues of the matrix B, we have

$$P^1(B)=\lambda_1+\cdots+\lambda_n,$$
$$P^2(B)=\lambda_1\lambda_2+\cdots+\lambda_{n-1}\lambda_n.$$

We notice that for any $A\in Gl(n,\mathbb{C})$, we have that $P^j(A^{-1}BA)=P^j(B)$. This follows from the fact that the eigenvalues are invariant under change of basis, and hence $A^{-1}BA$ only permutes the eigenvalues of A. The claim follows since elementary symmetric polynomials are invariant under permutation of the variables. We say that P^j is *invariant*.

Let $F_D\in\Omega^p(\mathrm{Ad}E)$ be the curvature of a connection D. Let $P^j(F_D)\in\Omega^{2j}(M)$ be the differential form of degree $2j$ defined by

$$P^j(F_D):=P^j((F_D)_{ij}\otimes dx^i\wedge dx^j).$$

Since P^j is invariant, this expression is well-defined since any other trivialization of F_D is related by a change of basis of the form $\varphi^{-1}F_D\varphi$. The following seemingly magical theorem will allow us to define Chern classes as invariants of the vector bundle $E\to M$. We state it without proof, and refer to the interested reader to [1].

Theorem 8.2. *$P^j(F_D) \in \Omega^{2j}(M)$ is a closed $2j$-form: $dP^j(F_D) = 0$. This allows us to define the cohomology class $[P^j(F_D)] \in H^{2j}(M)$. Futhermore, $[P^j(F_D)]$ does not depend on the choice of connection D.*

Definition 8.1. *The Chern classes of E are the following:*

$$c_j(E) = \left[P^j(\frac{i}{2\pi}F) \right] \in H^{2j}(M).$$

It will be useful to compute the first two Chern classes c_1, c_2 for our 4 dimensional, compact, orientable manifold M. Denote the eigenvalues of $iF/2\pi \in \Omega^2(\mathrm{Ad}E)$ by the 2-forms $\lambda_\alpha \in \Omega^2(M)$.

$$\begin{aligned}\sum_{j=0}^{n} c_j(E)\tau^j &= \det\left(\frac{i}{2\pi}\tau F + I\right) \\ &= \prod_{\alpha=1}^{n}(1+\lambda_\alpha \tau) \\ &= 1 + (\lambda_1 + \dots \lambda_n)\tau + (\lambda_1\lambda_2 + \dots + \lambda_{n-1}\lambda_n)\tau^2 \\ &\quad + \dots + (\lambda_1 \cdots \lambda_n)\tau^n.\end{aligned}$$

Since F is skew-symmetric, its trace is zero. Hence

$$c_1(E) = \sum_{i=1}^{n} \lambda_i = \mathrm{Tr}\frac{iF}{2\pi} = 0.$$

The second Chern class is given by

$$c_2(E) = \lambda_1\lambda_2 + \dots + \lambda_{n-1}\lambda_n = \frac{1}{2}\left((\sum_{i=1}^{n}\lambda_i)^2 - \sum_{i=1}^{n}\lambda_i^2 \right).$$

Rewritten in another way, we obtain

$$c_2(E) = \frac{1}{2}\left(\frac{i}{2\pi}\right)^2 (\mathrm{Tr}F \wedge \mathrm{Tr}F - \mathrm{Tr}(F \wedge F)) = \frac{1}{8\pi^2}\mathrm{Tr}(F \wedge F).$$

We investigate the term $\mathrm{Tr}(F \wedge F)$. First decompose $F = F^+ + F^-$ into its selfdual and antiselfdual projections. From (17), we know that the cross-terms vanish and by linearity of the trace operator

$$\mathrm{Tr}(F \wedge F) == \mathrm{Tr}((F^+ + F^-) \wedge (F^+ + F^-)) = \mathrm{Tr}(F^+ \wedge F^+) + \mathrm{Tr}(F^- \wedge F^-).$$

Since $*F^+ = F^+$ and $*F^- = -F^-$, we have

$$\mathrm{Tr}(F \wedge F) = \mathrm{Tr}(F^+ \wedge *F^+) - \mathrm{Tr}(F^- \wedge *F^-).$$

By definition (14), we see

$$\mathrm{Tr}(F \wedge F) = -F^+ \cdot F^+ + F^- \cdot F^- = -|F^+|^2 + |F^-|^2.$$

Integrating the second Chern class over M, we obtain an invariant depending only on the vector bundle $E \to M$.

$$\int_M c_2(E) = \frac{-1}{8\pi^2} \int_M (|F^+|^2 - |F^-|^2) * (1). \tag{18}$$

Looking back at the Yang-Mills functional, by orthogonality we see

$$YM(D) = \int_M |F_D|^2 * (1) = \int_M (|F_D^+|^2 + |F_D^-|^2) * (1).$$

Hence we are minimizing $\int_M |F^+|^2 + \int_M |F^-|^2$ subject to the constraint $const = \int_M |F^+|^2 - \int_M |F^-|^2$. The solution is that we must have either $F^+ = 0$ or $F^- = 0$. In other words, F must be a selfdual soliton or an antiselfdual soliton, depending on the sign of $\int_M c_2(E)$. If $c_2(E)$ is positive, we require F to be an antiselfdual soliton. This completes the proof that solving the Yang-Mills equation on M is equivalent to finding selfdual or antiselfdual instantons.

9. ASD Connections and Gauge Fixing

As in the previous section, we always assume that M is an orientable, compact Riemannian manifold of dimension $d = 4$. In the previous section, we showed using Chern classes and the variational formulation of the Yang-Mills equations that solving for Yang-Mills connections on M is equivalent to solving for selfdual or antiselfdual connection on M, depending on the sign of $c_2(E)$. If $c_2(E)$ is positive, we seek ASD (antiselfdual) connections. In this section, we will briefly discuss the problem of solving the ASD equation over M. The theorems and proofs will closely follow [4].

We first write out the equation explicitly. For F to be ASD, we require $F^+ = 0$ in the decomposition $F = F^+ + F^-$. If we write locally $F = F_{jk} \otimes dx^j \wedge dx^k$, from the definitions of f_i^+ and f_i^- given in the previous section (defined above (17)), a short computation yields

$$\begin{aligned}F &= F_{jk} \otimes dx^j \wedge dx^k \\ &= \frac{1}{2}(F_{12} + F_{34}) \otimes f_1^+ + \frac{1}{2}(F_{13} - F_{24}) \otimes f_2^+ + \frac{1}{2}(F_{14} + F_{23}) \otimes f_3^+ \\ &\quad + \frac{1}{2}(F_{12} - F_{34}) \otimes f_1^- + \frac{1}{2}(F_{13} + F_{24}) \otimes f_2^- + \frac{1}{2}(F_{14} - F_{23}) \otimes f_3^-.\end{aligned}$$

Therefore, for $F^+ = 0$, we require the coefficients in front of f_1^+, f_2^+, f_2^+ to be zero. Hence the ASD equations are

$$\begin{aligned}F_{12} + F_{34} &= 0,\\ F_{13} - F_{24} &= 0,\\ F_{14} + F_{23} &= 0,\end{aligned}$$

As it stands, the ASD equation is not elliptic. The problem is that the equation is invariant under gauge transformations. The trivial solution, where the connection is determined by $A = 0$, is gauge equivalent to the solution with $A = g^{-1}dg$ by (9). But the derivatives of $A = g^{-1}dg$ cannot be controlled by the L^2 norm. In order to obtain an elliptic equation, we need to specify an additional piece of information to remove the gauge invariance. This is called *gauge fixing*.

To illustrate the situation, we revisit the case $n = 2$. As discussed in the chapter on Hodge theory, since $\mathfrak{o}(2)$ is commutative, the Yang-Mills equation is linear in this case. By (3), since the commutators vanish, we obtain

$$F = (\partial_j A_k - \partial_k A_j) \otimes dx^j \wedge dx^k = \partial_j A_k \otimes dx^j \wedge dx^k + \partial_k A_j \otimes dx^k \wedge dx^j.$$

Therefore,

$$F = 2dA_i \otimes dx^i.$$

Since $A = A_i \otimes dx^i$, we have

$$F = 2dA.$$

The gauge group $O(2)$ also acts on A in a nice way in the linear case. Write $g \in O(2)$ as the exponential of a member of its Lie algebra: $g = e^X$ for $X \in \mathfrak{o}(2)$. Since everything commutes, the action of g on A according to (9) is given by

$$g \cdot A = A + dX.$$

Now, as shown in (13), the Yang-Mills equation in the $n = 2$ case is given by $d^*F = 0$. In terms of A, then Yang-Mills equation reads

$$d^*dA = 0.$$

To make this elliptic, we add the following gauge condition:

$$d^*A = 0.$$

Combining these two equations, we obtain the elliptic equation

$$\Delta A = (d^*d + dd^*)A = 0.$$

We will generalize this procedure to the nonlinear case. To indicate the dependence on A, we will sometimes denote a connection $d + A$ by D_A. We will denote the gauge group by G, which in our case will be the unitary group. Denote the gauge equivalence class of a connection D_A by

$$\mathcal{H} = \{g \cdot A : g \in G\}.$$

Given a connection D_{A_0}, we say that $B \in \mathcal{H}$ is in the *Coulomb gauge* relative to A_0 if

$$D^*_{A_0}(B - A_0) = 0.$$

We see that if $A_0 = 0$, then in the linear case this reduces to $d^*B = 0$, which is the condition discussed above.

Theorem 9.1. *There exists a constant $C(A)$ depending on A such that if B is another connection on E and $a = B - A$ satisfies*

$$||D_A D_A a||^2_{L^2} + ||a||^2_{L^2} < C(A),$$

then there exists $g \in G$ such that $g \cdot B$ is in the Coulomb gauge relative to A.

Before proving this theorem, we will state some theorems from partial differential equations that will be used in the proof. For $\sigma \in \Gamma(E)$, and D a metric connection, and k a positive integer, we have the Sobolev norm

$$||\sigma||^2_{H^{k,2}} = \sum_{i=0}^{k} \int_M |D^{(i)}\sigma|^2.$$

If we take the completion of the space of smooth sections $\Gamma(E)$ with respect to the $H^{k,2}$ norm, we obtain the Banach space denoted by $H^{k,2}$. This space is equivalent to the one defined using the regular definition from analysis in local coordinates and bundle trivializations. Therefore, we can use standard theorems from Sobolev space theory. We will need the following two theorems:

Theorem 9.2. (Implicit Function Theorem) *Let E be a product of Banach spaces E_1, E_2 and f be a smooth map. If the partial derivative $(D_2 f)$ at a point (ξ_1, ξ_2) is surjective and has a bounded right inverse, then for all η_1 close to ξ_1 there is a solution η_2 such that $f(\eta_1, \eta_2) = f(\xi_1, \xi_2)$.*

Theorem 9.3. (Fredholm Alternative) *Let E_1, E_2 be vector bundles with metrics over a compact manifold M. Let $L : \Gamma(E_1) \to \Gamma(E_2)$ be an elliptic operator. Then the formal adjoint L^* is also elliptic and a section σ of E_2 is in the image of L if and only if the L^2 inner product (σ, μ) is zero for all μ in the kernel of L^*.*

We now prove Theorem 9.1.

Proof. First, we compute using (5):

$$(g \cdot D_A)\sigma = g^{-1} D_A(g\sigma) = D\sigma - (D(g^{-1}g\sigma) - g^{-1}D(g\sigma)) = D\sigma - (D_A g^{-1})g\sigma.$$

Hence

$$g \cdot A = A - (D_A g^{-1})g.$$

Next, since $D_{A+a} = D_A + a$, we compute

$$\begin{aligned} g \cdot (A+a)\sigma &= (A+a)\sigma - (D_{A+a} g^{-1}) g\sigma \\ &= (A+a)\sigma - D_{A+a}(\sigma) + g^{-1} D_{A+a}(g\sigma) \\ &= (A+a)\sigma - D_A\sigma - a\sigma + g^{-1}a(g\sigma) + g^{-1}D_A(g\sigma) \\ &= A\sigma + g^{-1}a(g\sigma) - (D_A g^{-1})(g\sigma). \end{aligned}$$

Hence

$$g \cdot (A+a) = A + g^{-1}ag - (D_A g^{-1})g.$$

We are looking for a $g \in G$ such that $0 = D_A^*(g \cdot B - A)$. By the above computations, the equation that we need to solve is

$$D_A^*(g^{-1}ag - (D_A g^{-1})g) = 0.$$

Let $g = e^{-X}$, where $X \in \Gamma(\mathrm{End}(E))$ whose fiber endomorphisms are in the Lie algebra of G, and define

$$F(X, a) = D_A^*((D_A e^X)e^{-X} - e^X a e^{-X}).$$

We want to use Sobolev spaces, so we extend the domain of F to allow $X \in H^{3,2}$ and $a \in H^{2,2}$. We see that $\mathrm{Im}F \subset \overline{\mathrm{Im}D_A^*}$ where the closure is taken with respect to $H^{1,2}$. The total derivative of F at $(0,0)$ is given by

$$DF(\xi, b) = D_A^* D_A \xi - D_A^* b.$$

To apply the implicit function theorem, we need to show that the map $\xi \mapsto D_A^* D_A \xi$ is onto $\mathrm{Im}D_A^*$. We can then obtain a small solution X to the equation $F(X, a) = 0$.

Showing surjectivity will be done via the Fredholm alternative. Since $D_A D_A^*$ is elliptic, the equation $D_A^* D_A \xi = \eta$ has a solution if and only if η is L^2 orthogonal to the kernel of $D_A^* D_A$. If $D_A^* D_A \sigma = 0$, then

$$(D_A\sigma, D_A\sigma) = (\sigma, D_A^* D_A \sigma) = 0.$$

Hence $D_A\sigma = 0$. Therefore $(D_A^*\mu, \sigma) = (\mu, D_A\sigma) = 0$, and $\mathrm{Im}D_A^*$ is L^2 orthogonal to the kernel of $D_A^* D_A$. Since $||D_A D_A a||_{L^2}^2 + ||a||_{L^2}^2$ is an admissible norm on $H^{2,2}$, if $C(A)$ is small enough, the implicit value theorem provides a solution $g = e^{-X}$.

However, $X \in H^{3,2}$, and we are looking for a smooth $g = e^{-X}$. By a bootstrap argument, we can show that g is indeed smooth. Using that $D_A^*(gag^{-1} - (D_A g)g^{-1}) = 0$, one can compute

$$D_A^* D_A g = (D_A g g^{-1}, D_A g) + g D_A^* a g^{-1} + (D_A g, a) + (ga, g^{-1} D_A g),$$

where $(,)$ denotes the contraction on the one-form components. If g is continuous and $g \in H^{n,2}$, then the right hand side is in $H^{n-1,2}$. By elliptic regularity, since $D_A^* D_A$ is of order 2, we have $g \in H^{n+1,2}$. By the Sobolev embedding theorem, g is smooth.

□

We conclude this section by stating some theorems of Uhlenbeck without proof. The proofs are long and involved; they can be found in [4]. The first theorem states that for small curvature, the Coulomb gauge condition can be uniquely satisfied provided boundary conditions.

Theorem 9.4. *There are constants $\varepsilon_1, M > 0$ such that any connection defined by A on the trivial bundle over the closed unit ball $\overline{B}^4$ with $||F_A||_{L^2} < \varepsilon_1$ is gauge equivalent to a connection defined by $\tilde{A}$ over $\overline{B}^4$ such that*

$$d^* \tilde{A} = 0,$$
$$\lim_{|x| \to 1} \tilde{A}_r = 0,$$
$$||\tilde{A}||_{H^{1,2}} \leq M ||F_{\tilde{A}}||_{L^2}.$$

Here A_r denotes the radial component $\sum (x_i/r) A_i$ of the connection matrix, defined on $B^4 \backslash \{0\}$.

For suitable constants ε_1, M, the connection $\tilde{A}$ is uniquely determined by these properties, up to the transformation $\tilde{A} \to g_0 \cdot \tilde{A}$ for a constant $g_0 \in G$.

The next theorem states that for small curvature, the ASD equations in the Coulomb gauge has the nice property that the L^2 norm of the curvature controls the Sobolev norms of the connection.

Theorem 9.5. *There is a constant $\varepsilon_2 > 0$ such that if $\tilde{A}$ is any ASD connection on the trivial bundle over B^4 which satisfies the Coulomb gauge condition $d^* \tilde{A} = 0$ and $||\tilde{A}||_{L^4} \leq \varepsilon_2$, then for any interior domain D whose closure is compact and inside B^4 and any $l \geq 1$, we have*

$$||\tilde{A}||_{H^{l,2}(D)} \leq M_{l,D} ||F_A||_{L^2(B^4)},$$

for a constant $M_{l,D}$ *depending only on* l *and* D*.*

We have only scratched the surface of the mathematics involved in studying instantons on a 4-manifold. The study of the Yang-Mills equations on 4-manifolds has lead to deep insight on the classification of differentiable 4-manifolds, and instantons turn out to be a powerful tool. A detailed account can be found in [4].

References

[1] J. Jost, *Riemannian Geometry and Geometric Analysis*, Springer-Verlag, Berlin Heidelberg, 2005.

[2] M. Nielsen, *An introduction to Yang-Mills theories*, 2007, available at http://michaelnielsen.org/blog/introduction-to-yang-mills-theories.

[3] E. Meinrenken, *Riemannian Geometry*, Lecture Notes, University of Toronto, 2002, available at http://www.math.toronto.edu/mein/teaching/rieall.pdf.

[4] S. Donaldson and P. Kronheimer, *The Geometry of Four-Manifolds*, Oxford Mathematical Monographs, Oxford University Press, 1990.