GEOMETRY AND GAUGE-THEORETIC PHYSICS: AN INSTANTON CONNECTION

Arjun Kudinoor

Senior Thesis in Mathematics
Advisor: Professor Mike Miller Eismeier

COLUMBIA UNIVERSITY
March 2023

Abstract

Broadly, the subjects of this thesis are differential geometry and mathematical gauge theory as they relate to physics. Motivated by the formulation of Maxwell's equations using differential forms, we build and use tools like connections, curvature, and characteristic classes in differential geometry to study spaces of self dual solutions to the nonabelian analogue of Maxwell's equations, namely the Yang–Mills equations. We also examine spaces of these self dual solutions, called instantons, modulo gauge transformations. These spaces are called moduli spaces. Finally, we introduce concepts like Clifford bundles, Dirac operators, and the Atiyah–Singer index theorem to prove the Atiyah–Hitchin–Singer theorem regarding the dimensionality of moduli spaces of $SU(2)$-instantons on the 4-sphere.

Acknowledgements

I would like to thank my thesis advisor, Professor Mike Miller Eismeier for his incredible support throughout the thesis-writing process. This thesis would not have been possible without his guidance. His lectures, office hours, advising sessions, and unimaginably quick and thorough responses to emails have been invaluably helpful. I am especially grateful that Professor Miller encouraged me to learn the material in this thesis by teaching it to him during our advising meetings — this approach to learning not only helped me understand the material better, but also helped me develop my skills as an educator in mathematics. In addition to advising me on matters related to my undergraduate education, I am also grateful for Professor Miller's support as I applied to graduate schools and for his advice as it relates to my career goals. I owe Professor Miller a huge debt of gratitude and am incredibly fortunate to know him as my teacher, advisor, and mentor.

I also want to thank Professor Brian Cole and Professor Sebastian Will in Columbia's Physics department, as well as Professor Robert Camley at the University of Colorado at Colorado Springs for encouraging and enabling my pursuit of a research career in physics and mathematics. I'm also grateful to the faculty in the Physics and Mathematics departments who have taught and mentored me. I also thank my high school mathematics teacher Mr. Jeremy Beckman for inspiring me to pursue an education in mathematics.

I thank Columbia University, the Science Research Fellows program, and the Barry Goldwater Scholarship and Excellence in Education Foundation for supporting me financially during my research experiences at Columbia. Specifically, I would like to thank my Science Research Fellows advisor, Professor Vesna Gasperov, for her encouragement of my research efforts throughout the last four years.

I would also like to thank my friends, Aswath Suryanarayanan, Param Gujral, Shloka Janapaty, Avishi Poddar, Baha Topbas, Jonathan Marty, Sebastian Rossi, and Vivian Cribb. They have been incredibly supportive and their feedback has been invaluable.

Finally, none of this would have been possible without the unwavering support of my family. Words cannot express my gratitude to my parents, Srinivasan Kudinoor and Sharmila Ayyengar, and my sister, Shalini Kudinoor, for their love and encouragement of all my pursuits. They have provide me with every opportunities I need to succeed, for which I am forever grateful.

This thesis is dedicated in loving memory to my grandfather,
Sri. K. V. Seshadri.

Contents

Introduction **6**

1 Connections, Curvature, and Gauge Transformations **11**

1.1 Connections and Curvature on Vector Bundles . . . 11
1.1.1 Fiber and Vector Bundles . . . 11
1.1.2 Connections and Curvature . . . 12
1.1.3 Connections and Curvature as Differential Forms . . . 13
1.1.4 Complex vector bundles . . . 15
1.1.5 Characteristic Classes . . . 17
1.2 Connections, and Curvature on Principal G-Bundles . . . 21
1.2.1 Principal G-Bundles . . . 21
1.2.2 G-Equivariance . . . 21
1.2.3 Horizontal and Vertical Tangent Bundles . . . 23
1.2.4 Connections and Curvature on Principal G-Bundles . . . 24
1.2.5 The Adjoint Bundle . . . 26
1.3 Gauge Transformations . . . 27
1.3.1 Gauge Transformations . . . 27
1.3.2 Electromagnetism From Principal $U(1)$-Bundles . . . 28

2 Yang–Mills Gauge Theory **30**

2.1 Yang–Mills Instantons . . . 30
2.1.1 Self-Dual Connections . . . 30
2.1.2 The Yang–Mills Equations . . . 31
2.1.3 Instantons . . . 33
2.1.4 Moduli Spaces of Instantons . . . 33
2.2 An Introduction to Index Theory . . . 35
2.2.1 Fredholm Operators and The Index . . . 35
2.2.2 Clifford Bundles and Dirac Operators . . . 36
2.2.3 Elliptic Operators . . . 39
2.3 The Atiyah–Hitchin–Singer Theorem . . . 41
2.3.1 Step 1.1: Tangent Moduli Spaces . . . 42
2.3.2 Step 1.2: A Dirac Operator and the Index Theorem . . . 43
2.3.3 Steps 2 and 3: Local and Global Moduli Spaces . . . 45
2.3.4 Restricting to the Atiyah–Hitchin–Singer theorem . . . 45

Introduction

Differential geometry in physics

Differential geometry is the the framework we use to describe and study geometry of spaces with smooth structure, like differentiable manifolds. In physics, differential geometry is a powerful tool to describe the geometry of spacetime, the structures of gauge theories, and the mathematical underpinnings of symmetries and conservation laws. The most popular application of differential geometry to physics in in general relativity, where colloquially, people say "spacetime curvature is equivalent to gravity". This equivalence is made precise in the Einstein field equations,

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu},$$

where $G_{\mu\nu}$ is the Einstein tensor, Λ is the cosmological constant, $g_{\mu\nu}$ is the metric tensor, κ is the Einstein gravitational constant, and $T_{\mu\nu}$ is the stress-energy tensor. Although it is not important for the purpose of this thesis to know exactly what these terms mean, it is important to note the following interpretation of the Einstein equations. The left hand side of the above equation is geometric — it uses ideas from differential geometry including the spacetime metric and its curvature. The right hand side of the equation is physical — it describes the physics of the gravitational field using the stress-energy tensor. Thus, Einstein's equations explicate an equivalence between differential geometric ideas like spacetime curvature and physical ideas like the strength of the gravitational field.

This equivalence between curvature and field strength is not unique to gravity. In fact, this equivalence is present in the equations of motion that describe the other three fundamental forces, namely the weak, strong, and electromagnetic forces. The Yang–Mills equations describe the free-field weak and strong interactions. The electromagnetic interaction is described by Maxwell's equations, which are an Abelian instance of the Yang–Mills equations in the free-field case. Let us briefly examine the use of differential geometry in the construction of Maxwell's equations,

$$\begin{aligned} \vec{\nabla} \cdot \vec{B} &= 0, \qquad & \partial_t \vec{B} + \vec{\nabla} \times \vec{E} &= 0, \\ \vec{\nabla} \cdot \vec{E} &= \rho, \qquad & \vec{\nabla} \times \vec{B} - \partial_t \vec{E} &= \vec{J}, \end{aligned}$$

where $\vec{E}$ is the electric field, $\vec{B}$ is the magnetic field, ρ is the electric charge density, and $\vec{J}$ is the electric current density.

A geometric construction of Maxwell's equations

We usually think of the electric and magnetic fields as vectors in $\mathbb{R}^3$. However, to tease out the geometric aspects of Maxwell's equations, we will express the electric and magnetic fields in terms of differential forms.

In the static case, the top two Maxwell's equations above take the form $\vec{\nabla} \cdot \vec{B} = 0$ and $\vec{\nabla} \times \vec{E} = 0$. In the language of differential forms, the divergence becomes an exterior derivative on 2-forms on $\mathbb{R}^3$ and the curl becomes an exterior derivative on 1-forms on $\mathbb{R}^3$. So, we will treat the magnetic field as a 2-form on $\mathbb{R}^3$ and the electric field as a 1-form on $\mathbb{R}^3$. We write

$$\begin{aligned} B &= B_x dy \wedge dz + B_y dz \wedge dx + B_z dx \wedge dy, \\ E &= E_x dx + E_y dy + E_z dz. \end{aligned}$$

Then, the top two Maxwell's equations become

$$dB = 0 \quad \text{and} \quad dE = 0,$$

where d is the exterior derivative on p-forms in $\mathbb{R}^3$.

Generalizing to the time-dependent case, our electric and magnetic fields lie on 4-dimensional Minkowski spacetime $\mathbb{R}^{3+1}$ with metric signature $(+,-,-,-)$ and usual coordinates $(x^0, x^1, x^2, x^3) = (t, x, y, z)$. Then, we similarly define the electric field 1-form E and the magnetic field 2-form B on $\mathbb{R}^4$ as

$$\begin{aligned} B &= B_x dy \wedge dz + B_y dz \wedge dx + B_z dx \wedge dy, \\ E &= E_x dx + E_y dy + E_z dz. \end{aligned}$$

We then construct a 2-form $F = E \wedge dt + B$, called the *electromagnetic field strength*. Motivated by the Bianchi identity, let us set $dF = 0$. Then,

$$0 = dF = dB + d(E \wedge dt) = dB + dE \wedge dt,$$

where in the last equality we have used the fact that $d^2 t = 0$. For a differential form $\omega = a_I dx^I$ on $\mathbb{R}^3$, we may split $d\omega$ into the spatial and temporal derivatives by

$$d\omega = \partial_\mu \omega_I dx^\mu \wedge dx^I = d_s \omega + \partial_t a_I dt \wedge dx^I,$$

where $d_s \omega = \partial_i a_i dx^I \wedge dx^i$. Thus, $dF = dB + dE \wedge dt = 0$ yields

$$d_s B + dt \wedge \partial_t B + (d_s E + dt \wedge \partial_t E) \wedge dt = d_s B + (\partial_t B + d_s E) \wedge dt = 0.$$

Thus,

$$dF = 0 \implies d_s B = 0 \text{ and } \partial_t B + d_s E = 0,$$

which correspond to $\nabla \cdot \vec{B} = 0$ and $\partial_t \vec{B} + \nabla \times \vec{E} = 0$, respectively. These are precisely the top two Maxwell's equations. The other two Maxwell's equations, namely $\nabla \cdot \vec{E} = \rho$ and $\nabla \times \vec{B} - \partial_t \vec{E} = \vec{J}$, are given by

$$*d * F = J,$$

where $*$ is the Hodge star operator and the current density J is a 1-form on $\mathbb{R}^4$, given by$J = \rho dt - (J_x dx + J_y dy + J_z dz)$. Thus, Maxwell's equations take the form

$$dF = 0 \text{ and } *d*F = J.$$

Thus, the use of differential forms yields a compact expression of the fundamental equations of electromagnetism. To see the connection to geometry, observe that since $dF = 0$, we may write $F = dA$ for a complex-valued 1-form A. This is possible because $\mathbb{R}^{3+1}$ is contractible and, by the Poincaré lemma (see Ch. 4.3 in [1]), all closed forms are exact on contractible spaces. A is called an *electromagnetic gauge potential*, *gauge field*, or *vector potential*. In language that will be introduced later, the vector potential A is a *connection* 1-form on a $U(1)$-principal bundle over $\mathbb{R}^{3+1}$ and $F = dA$ is the *curvature* 2-form of this connection.

We can motivate this terminology of connection 1-forms and curvature 2-forms using a heuristic picture of derivatives in calculus. Suppose we have a function $f : \mathbb{R} \to \mathbb{R}$. Then, it's derivative $f'(x)$ is the slope of f at the point $x \in \mathbb{R}$. So, $f'(x)$ defines a direction, a vector in the plane on which the graph of f, 1-dimensional manifold, lies. The span of this tangent vector defines the tangent space to the graph at $(x, f(x))$. This tangent space is isomorphic to the cotangent space of differential 1-forms at $(x, f(x))$. This is analogous to viewing the vector potential 1-form A as a connection, or *covariant derivative*, at a point in the manifold $\mathbb{R}^4$. Differentiating $f'(x)$, we get $f''(x)$ — the curvature of f. We call it the curvature because the sign of $f''(x)$ tells us whether a function is concave up, concave down, or an inflection point. Similarly, the exterior derivative of the connection is its curvature.

Thus, using differential equations to write Maxwell's equations as $dF = 0$ and $*d*F = J$ explicates an equivalence between the geometry of electromagnetic fields and the underlying physics of electric currents.

A primer on gauge theory

To understand the role of geometry in particle physics beyond electromagnetism, we introduce another mathematical tool that describes the symmetries enjoyed by physical systems of particle fields: gauge theory.

Note that the electromagnetic field strength F is invariant under the transformation $A \mapsto A + d\phi$, where ϕ is a smooth, purely imaginary function on $\mathbb{R}^4$. So, any physical information about the field strength we derive from A can also be derived from $A + d\phi$.

We call this property of F *gauge invariance* and the transformation $A \mapsto A + d\phi$ a *gauge transformation*. The study of quantities and properties left invariant under gauge transformations is the primary goal of gauge theory.

There is another way to write this gauge transformation that makes its useful group-theoretic structure more apparent. With some foresight that A is a $U(1)$-*gauge field*, suppose we transform A under the Abelian unitary group $U(1)$ as follows: For each $\lambda \in U(1)$, λ acts on A as

$$\lambda^{-1} A = \lambda^{-1} A \lambda + \lambda^{-1} d\lambda.$$

Since $\lambda \in U(1)$, we can write $\lambda = e^{\phi}$, where $\phi : \mathbb{R}^4 \to i\mathbb{R}$ is a purely imaginary smooth function. Then,

$$\lambda^{-1} A = e^{-\phi} A e^{\phi} + e^{-\phi} d\phi e^{\phi} = A + d\phi,$$

where the last equality follows from $U(1)$ being Abelian. This is precisely the gauge transformation we started with. Hence, we say electromagnetism is a $U(1)$-gauge theory.

The implications of gauge theory are widespread not only in electromagnetism, but in particle physics as a whole. In particular, Yang–Mills gauge theories generalize the Abelian $U(1)$-gauge group in electromagnetism to nonabelian gauge groups. The electroweak interaction is described by an $SU(2) \times U(1)$ Yang–Mills gauge theory, and the strong interaction is described by an $SU(3)$ Yang–Mills gauge theory. Hence, the standard model of particle physics is described by an $SU(3) \times SU(2) \times U(1)$ Yang–Mills gauge theory.

Outline and expectations of this thesis

Although much of the motivation in this introduction is based in physics, this thesis is an exposition on the mathematical foundation for gauge-theoretic physics. In particular, this thesis is divided into two chapters dedicated to differential geometry and Yang–Mills gauge theory.

The first chapter covers topics in differential geometry like connections, curvature, and characteristic classes. In particular, we examine connections and curvature on two kinds of fiber bundles — vector bundles and principal bundles — both of which will be important when describing solutions to the Yang–Mills equations, i.e. the nonabelian analogue of Maxwell's equations. In chapter 1, we also introduce gauge transformations and return to electromagnetism as an example of a $U(1)$-gauge theory.

In the second chapter, we apply the geometric tools developed in Chapter 1 to study Yang–Mills gauge theory. We examine spaces of self dual connections, called instantons, that minimize the Euclidean Yang–Mills action functional. Instantons are widely studied in quantum field theory, condensed matter physics, geometry, and gauge theory. In this thesis, we focus on a geometric and gauge-theoretic examination of instantons by studying spaces of instantons modulo gauge transformations. These spaces are called moduli spaces. The final theorem we prove in this thesis is the Atiyah–Hitchin–Singer theorem.

Theorem (Atiyah–Hitchin–Singer Theorem)**.** The dimension of the moduli space of instantons on a principal $SU(2)$-bundle over S^4, with Pontryagin index $k > 0$ is $8k - 3$.

Prior to the statement and proof of this theorem, the known self dual solutions (instantons) of the $SU(2)$ Yang–Mills equations on S^4, called the t'Hooft solutions, were constructed using $5k + 4$ parameters. However, the Atiyah–Hitchin–Singer theorem implies the existence of new instanton solutions, which turn out to be deformations of the t'Hooft solutions themselves. To prove this theorem, we develop additional mathematical framework on index theory, Clifford bundles, and Dirac operators in Chapter 2.

Since this thesis extensively employs ideas from differential geometry, it requires knowledge of differential geometry as taught at the undergraduate level. Familiarity with electromagnetism and quantum field theory would be helpful, but is not strictly necessary. One goal of this thesis is to elucidate the potency of differential geometry in our understanding of Yang–Mills theories in particle physics. Conversely, another goal is to demonstrate how Yang–Mills gauge theories inspire the study of geometry, especially as it relates to moduli spaces of instantons.

Chapter 1

Connections, Curvature, and Gauge Transformations

1.1 Connections and Curvature on Vector Bundles

We embark on our quest to study Yang–Mills gauge theory with an introduction to some fundamental concepts in differential geometry: (vector and principal) bundles, connections, and curvature. We will observe that the electromagnetic gauge field may be thought of as a *principal connection* on the *principal* $U(1)$*-bundle* with field strength as its *curvature.*

1.1.1 Fiber and Vector Bundles

Definition 1.1.1. A **smooth fiber bundle** is a structure (E, X, F, π), where the **total space** E is a smooth manifold equipped with a smooth projection map $\pi \in C^\infty(E, X)$ to a **base** manifold X, with a **fiber** manifold F, that satisfies the following conditions:

- *Surjectivity:* π is surjective
- *Local triviality:* For any $x \in X$, there is an open neighbourhood U of x for which there is a diffeomorphism $T_U : \pi^{-1}(U) \to U \times F$, and $\text{proj}_1 \circ T_U = \pi$ on $\pi^{-1}(U)$. Here, $\text{proj}_1 : U \times F \to U$ is projection onto the first coordinate.

We call U a **trivializing open subset** of the fiber bundle. We will sometimes abuse notation and refer to the total space E or the map $\pi : E \to X$ as a fiber bundle, with an implicit understanding of the other components of the fiber bundle structure (E, X, F, π).

Since we will only work with smooth bundles, maps, and forms in this document, we we will suppress the word smooth.

Example 1.1.1 (Trivial Bundle)**.** The total space $E = X \times F$ equipped with $\pi_1 = \text{proj}_1 : E \to X$ is trivially a fiber bundle over X. So, we call it a **trivial bundle**.

Definition 1.1.2. A **local section** of a fiber bundle E on an open subset $U \subseteq X$ is a map $s : U \to E$ such that

$$\pi(s(x)) = x$$

If $U = X$, then s is a **global section** (or simply a section). We denote the space of smooth sections of a fiber bundle E as $C^\infty(E)$.

Definition 1.1.3. If the fibers of a fiber bundle E are vector spaces, then we call E a **vector bundle**.

Example 1.1.2. The tangent bundle TX (and cotangent bundle T^*X) of a manifold X are vector bundles over X with tangent (and cotangent) spaces as fibers.

1.1.2 Connections and Curvature

Let X be a smooth manifold, $\mathcal{F}$ be the space of all smooth functions on X, and E be a vector bundle on X. Denote the space of smooth sections of E by $C^\infty(E)$. For tangent bundle TX on X, $C^\infty(TX)$ are the **vector fields** on X.

Definition 1.1.4. A **covariant derivative operator** ∇ on E is a linear map

$$\nabla : C^\infty(TX) \times C^\infty(E) \to C^\infty(E)$$

assigning to each vector field X and section s of E a new section $\nabla_X s$ such that for all $f \in \mathcal{F}$:

- $\nabla_{fX} s = f \nabla_X s$ ($\mathcal{F}$-linearity on X)
- $\nabla_X(fs) = f\nabla_X s + (df)(X)s$

Remark 1.1.1. We call the second condition above the Leibniz rule. Suppressing X, we may rewrite it as $\nabla(fs) = f\nabla s + df \cdot s$. This allows us to think of a covariant derivative operator ∇ as a map $C^\infty(E) \to C^\infty(T^*X \otimes E)$ such that

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

for all $f \in C^\infty(X)$ and $s \in C^\infty(E)$. Thus, a covariant derivative operator may be defined by a **a connection** d_A (or more simply A), a first-order differential operator

$$d_A : C^\infty(E \otimes \Lambda^0) \to C^\infty(E \otimes \Lambda^1),$$

where Λ^p denotes the bundle of exterior p-forms on X.

Remark 1.1.2. We will often refer to the connection d_A and the covariant derivative operator interchangeably as A.

Definition 1.1.5. Let ∇ be a covariant derivative on a vector bundle E. The **curvature** of ∇ is a multilinear map $R : C^\infty(TX) \times C^\infty(TX) \times C^\infty(E) \to C^\infty(E)$ such that

$$R(X, Y)s = \nabla_X \nabla_Y s - \nabla_Y \nabla_X s - \nabla_{[X,Y]} s$$

Since R is $\mathcal{F}$-linear in all three arguments, it is defined pointwise, i.e. at a point $x \in X$, $R(X, Y)s$ depends only on the values of X, Y and s at m, and not on their values at nearby points. Thus, R is induced by a vector bundle map $TX \otimes TX \to \text{End}(E)$, where $\text{End}(E)$ is the set of endomorphisms of E. We call this map the **curvature tensor** of ∇.

Furthermore, since $R(X, Y)$ is antisymmetric and bilinear in X and Y, we may regard it as an $\text{End}(E)$-valued 2-form on X. This 2-form may be computed locally, as we will see in the next section.

Before we do that, it is worth noting an equivalent definition of curvature using the connection in remark 1.1.1.

Definition 1.1.6. There is a natural way to extend the connection $d_A : C^\infty(E \otimes \Lambda^0) \to C^\infty(E \otimes \Lambda^1)$ to a map

$$d_A : C^\infty(E \otimes \Lambda^p) \to C^\infty(E \otimes \Lambda^{p+1})$$
$$d_A(s \otimes \alpha) = (d_A s) \wedge \alpha + s \otimes d\alpha.$$

Then, there is an $\text{End}(E)$-valued 2-form F_A, also called the **curvature** of the connection A, such that for any

$$F_A = d_A^2.$$

One should read this as follows: For any $\omega \in C^\infty(E \otimes \Lambda^p)$, $d_A^2(\omega) = F_A \wedge \omega$, where the right hand side is a combination of the wedge product and the contraction $\text{End}(E) \otimes E \to E$.

1.1.3 Connections and Curvature as Differential Forms

Understanding the local behavior of connections and curvature necessitates a local conceptualization of sections of vector bundles. We do this using *frames*.

Definition 1.1.7. A **frame** for a vector bundle E with fiber dimension r over an open set U is a collection of sections $e_1, ..., e_r$ of E over U such that at each point $p \in U$, the elements $e_1(p), ..., e_r(p)$ form a basis for the fiber F at p.

Suppose $U \subseteq X$ is a trivializing open set for vector bundle E over a manifold X. Let $e_1, ..., e_r$ be a frame for E over U. Let $X \in C^\infty(TU)$ be a C^∞ vector field on U. As a section of E over U, $\nabla_X e_j$ is a linear combination of the e_i's with coefficients ω^i_j depending on X:

$$\nabla_X e_j = \omega^i_j(X) e_i$$

where we implicitly sum over repeated indices. On U, any section $s \in C^\infty(U, E)$ is a linear combination $s = a^j e_j$. If desired, the section $\nabla_X s$ can be computed from $\nabla_X e_j$ by linearity and the Leibniz rule.

Definition 1.1.8. The $\mathcal{F}$-linearity of $\nabla_X e_j$ in X implies that ω^i_j is $\mathcal{F}$-linear in X. So, ω^i_j is a 1-form on U (see Corollary 7.27 in [2]). The 1-forms ω^i_j are called the **connection forms** of the connection ∇ relative to the frame $e_1, ..., e_r$ on U.

A similar computation establishes that curvature may be represented as a 2-form.

Definition 1.1.9. For vector fields $X, Y \in C^\infty(TU)$, the section $R(X,Y)e_j$ is a linear combination

$$R(X,Y)e_j = \Omega^i_j(X,Y)e_i$$

with alternating and $\mathcal{F}$-bilinear coefficients Ω^i_j. The Ω^i_j's are 2-forms, called the **curvature forms** of the connection ∇ relative to the frame $e_1, ..., e_r$ on U.

Theorem 1.1.1 (Second Structural Equation)**.** Let ∇ be a connection on a vector bundle E over a manifold X with fiber dimension r. Relative to a frame $e_1, ..., e_r$ over a trivializing open subset $U \subseteq X$, the connection forms and curvature forms are related by

$$\Omega^k_j = d\omega^k_j + \omega^k_i \wedge \omega^i_j.$$

Proof. Let $X, Y \in C^\infty(TU)$ be arbitrary vector fields. Then,

$$\begin{aligned}\nabla_X \nabla_Y e_j &= \nabla_X(\omega^k_j(Y)e_k) \\ &= X\omega^k_j(Y)e_k + \omega^k_j(Y)\nabla_X e_k \text{ (Leibniz rule)} \\ &= X\omega^k_j(Y)e_k + \omega^k_j(Y)\omega^i_k(X)e_i \\ &= X\omega^k_j(Y)e_k + \omega^k_i(X)\omega^i_j(Y)e_k \text{ (relabeling indices)}\end{aligned}$$

$$\nabla_Y \nabla_X e_j = Y\omega^k_j(X)e_k + \omega^k_i(Y)\omega^i_j(X)e_k$$

$$\nabla_{[X,Y]}e_j = \omega^k_j([X,Y])e_k$$

Thus,

$$\begin{aligned}R(X,Y)e_j &= \nabla_X\nabla_Y e_j - \nabla_Y\nabla_X e_j - \nabla_{[X,Y]}e_j \\ &= (X\omega^k_j(Y) - Y\omega^k_j(X) - \omega^k_j([X,Y]))e_k + (\omega^k_i(X)\omega^i_j(Y) - \omega^k_i(Y)\omega^i_j(X))e_k \\ &= (d\omega^k_j + \omega^k_i \wedge \omega^i_j)(X,Y)e_k\end{aligned}$$

Since $R(X,Y)e_j = \Omega^k_j e_k$,

$$\Omega^k_j = d\omega^k_j + \omega^k_i \wedge \omega^i_j. \quad \square$$

Suppressing the indices, we write the curvature 2-form of the connection ω as

$$\Omega = d\omega + \omega \wedge \omega.$$

Thinking of ω as an $\mathrm{End}(E)$-valued 1-form, it will instead be useful to write Ω as

$$\Omega = d\omega + \frac{1}{2}[\omega, \omega],$$

where $[\cdot, \cdot]$ is the Lie bracket.

We note that the form of Ω above looks similar to the form of the electromagnetic field strength F. In particular, $F = dA$, where A is the electromagnetic gauge potential 1-form, and $\Omega = d\omega + \frac{1}{2}[\omega, \omega]$, where ω is a matrix of connection 1-forms. In fact, in the next section, we will think of the gauge potential as a connection on a *principal G-bundle*, a fiber bundle with additional group structure. Doing so will allow us to identify the field strength as the curvature of the gauge potential.

1.1.4 Complex vector bundles

At this point, it is important for us to distinguish between real and complex vector bundles. This is because, in the subsection that follows, we will define some useful topological notions on vector bundles called characteristic classes. The definitions and properties of these characteristic classes differ between real and complex vector bundles.

Definition 1.1.10. Suppose E is a vector bundle over a manifold X, with fiber dimension k. We call E a **real vector bundle** if its fibers are real vector spaces. So, the projection map $\pi : E \to X$ is locally of the form $U \times \mathbb{R}^k \to U$. On the other hand, we call E a **complex vector bundle** if its fibers are complex vector spaces. So, the projection map $\pi : E \to X$ is locally of the form $U \times \mathbb{C}^k \to U$.

Complex vector bundles admit a *complex structure* that allows us to promote real vector bundles to complex ones.

Definition 1.1.11. Suppose E is a real vector bundle on a manifold X. We call an automorphism $J : E \to E$ a **complex structure on** E if at each fiber E_x at $x \in X$, $J_x : E_x \to E_x$ has the property that

$$J_x^2 = -\text{id}.$$

If E is a real vector bundle with complex structure J, then we can promote E to a complex vector bundle by defining complex multiplication on vectors in each fiber E_x by

$$(a + ib)v = av + J(bv),$$

for all $a, b \in \mathbb{R}$ and all $v \in E_x$. On the other hand, if $E \to X$ is a complex vector bundle, then we can define a complex structure $J : E \to E$ by identifying $J_x : E_x \to E_x$, at each fiber E_x, with scalar multiplication by i.

To investigate this idea of complex structure on complex vector bundles further, we restrict to examining *complexifications* of real vector bundles. Recall that the complexification of a real vector space V is the tensor product $V_{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C}$ between V and $\mathbb{C}$, where we view $\mathbb{C}$ as a two-dimensional real vector space. We make $V_{\mathbb{C}}$ a complex vector space by defining complex multiplication on $V_{\mathbb{C}}$ by

$$z(v \otimes w) = v \otimes zw,$$

for all $v \in V$, $z, w \in \mathbb{C}$.

Definition 1.1.12. The **complexification** of a real vector bundle E over X is the complex vector bundle

$$E_{\mathbb{C}} = E \otimes \mathbb{C}$$

over X, whose fibers over each $x \in X$ is $E_x \otimes_{\mathbb{R}} \mathbb{C}$.

$E_{\mathbb{C}}$ has a complex structure $J : E_C \to E_{\mathbb{C}}$, which sends the real subbundle $E \subset E_{\mathbb{C}}$ to JE. Since $J^2 = -\text{id}$, $E \cap JE$ is the submanifold of $E_{\mathbb{C}}$ consisting of all the zero vectors and $E + JE = E_{\mathbb{C}}$. So, $E \oplus JE$ is canonically isomorphic to $E_{\mathbb{C}}$.

Furthermore, since J is a bundle automorphism, the restriction $J|_E : E \to JE$ is a bundle isomorphism. Thus, $E \oplus E$ is isomorphic to $E_{\mathbb{C}}$ by the isomorphism

$$\phi : E \oplus E \to E_{\mathbb{C}}, \quad \phi(\alpha, \beta) = \alpha + J\beta.$$

Since $E \oplus E$ is isomorphic to $E_{\mathbb{C}}$, a complex vector bundle, one might ask what complex structure on $E \oplus E$ would make it a complex vector bundle — we give it the complex structure $\tilde{J} : E \oplus E \to E \oplus E$ on $E \oplus E$ such that on each fiber $(E \oplus E)_x$,

$$J_x : (E \oplus E)_x \to (E \oplus E)_x, \quad \tilde{J}_x(v, w) = (-w, v).$$

It is easy to see that $\tilde{J}_x^2 = -\text{id}$.

Definition 1.1.13. For a complex vector bundle E with complex structure $J : E \to E$, define the **conjugate bundle** $\overline{E}$ to be the bundle E with complex structure $-J$. In other words, $\overline{E}$ has the same base and fibers as E, but we define complex multiplication on each fiber E_x by setting

$$(a + ib)v = av - J(bv),$$

for all $a, b \in \mathbb{R}$ and $v \in E_x$. Hence, scalar multiplication by $\mathbb{C}$ on each fiber of $\overline{E}$ acts through complex conjugation.

Proposition 1.1.1. If E is a real vector bundle, then $E_{\mathbb{C}}$ is isomorphic to $\overline{E_{\mathbb{C}}}$.

Proof. $E_{\mathbb{C}}$ has complex structure $J : E_{\mathbb{C}} \to E_{\mathbb{C}}$. Define a map $\overline{\phi} : E \oplus E \to \overline{E_{\mathbb{C}}}$ by

$$\overline{\phi}(\alpha, \beta) = \alpha - J\beta.$$

$\overline{\phi}$ is an isomorphism in the same way as $\phi : E \oplus E \to E_{\mathbb{C}}$ is an isomorphism above. Since compositions preserve isomorphisms, $\overline{\phi} \circ \phi^{-1} : E_{\mathbb{C}} \to \overline{E_{\mathbb{C}}}$ is an isomorphism. Thus, $E_{\mathbb{C}}$ and $\overline{E_{\mathbb{C}}}$ are isomorphic. □

Definition 1.1.14. Suppose E is a complex vector bundle. Define the **underlying real vector bundle** $E_{\mathbb{R}}$ to be the vector bundle E with fibers treated as vector spaces over $\mathbb{R}$, as opposed to over $\mathbb{C}$.

Proposition 1.1.2. Suppose E is a complex vector bundle over X. Then, the complexification $(E_{\mathbb{R}})_{\mathbb{C}} = E_{\mathbb{R}} \otimes \mathbb{C}$ of its underlying real vector bundle $E_{\mathbb{R}}$ is isomorphic to $E \oplus \overline{E}$.

Proof. Suppose $J : E_{\mathbb{R}} \otimes \mathbb{C} \to E_{\mathbb{R}} \otimes \mathbb{C}$ is the complex structure on $E_{\mathbb{R}} \otimes \mathbb{C}$, that acts as multiplication by i on vectors within each fiber. Let $x \in X$ be arbitrary, and consider the fiberwise map

$$\phi : (E_{\mathbb{R}} \otimes \mathbb{C})_x \to E_x \oplus \overline{E}_x, \quad \phi(v \otimes 1 + w \otimes i) = (v + iw, v - iw),$$

where $v, w \in E_x$. We note that

$$\begin{aligned}\phi(i(v \otimes 1 + w \otimes i)) &= \phi(-w + v \otimes i) \\ &= (-w + iv, -w - iv) \\ &= i(v + iw, v - iw) \\ &= i\phi(v \otimes 1 + w \otimes i),\end{aligned}$$

where in the penultimate line, we use the fact that $\mathbb{C}$ acts on $\overline{E}_x$ through complex conjugation. Thus, ϕ is $\mathbb{C}$-linear and induces an isomorphism between $E_{\mathbb{R}} \otimes \mathbb{C}$ and $E \oplus \overline{E}$. □

The purpose of these definitions and propositions will become clear in the following subsection about characteristic classes of real and complex vector bundles.

1.1.5 Characteristic Classes

Recall that a primary goal of gauge theory is to study quantities that are invariant under "gauge transformations". In the spirit of understanding such invariants in future sections, we introduce the notion of characteristic classes. These are topological invariants that will prove useful when we examine the relationship between topological and analytic data of operators on manifolds in the next chapter.

Definition 1.1.15. Let X be an $k \times k$ matrix with entries X^i_j. A **polynomial** $P(X)$ on $\mathfrak{gl}_k(\mathbb{R}) = \mathbb{R}^{k\times k}$ is a polynomial in the entries of X. A polynomial $P(X)$ on $\mathfrak{gl}_k(\mathbb{R})$ is said to be **ad $\mathrm{GL}_k(\mathbb{R})$-invariant** or simply **invariant** if for all $A \in \mathrm{GL}_k(\mathbb{R})$,

$$P(A^{-1}XA) = P(X).$$

Example 1.1.3. $\det(X)$ and $\mathrm{tr}(X)$ are invariant polynomials on $\mathfrak{gl}_k(\mathbb{R})$.

Example 1.1.4. Let X be a $k \times k$ matrix of indeterminates and let λ be another indeterminate. The coefficients $f_l(X)$ of λ^{k-l} in

$$\det(\lambda I + X) = \lambda^k + f_1(X)\lambda^{k-1} + ... + f_k(X)$$

are polynomials on $\mathfrak{gl}_k(\mathbb{R})$. Each $f_l(X)$ is an invariant polynomial because for all $A \in GL_k(\mathbb{R})$ and any $k \times k$ matrix X of real numbers,

$$\det(\lambda I + A^{-1}XA) = \det(\lambda I + X) \implies f_k(A^{-1}XA) = f_k(X),$$

for all $A \in GL_k(\mathbb{R})$, $X \in \mathfrak{gl}_k(\mathbb{R})$.

Proposition 1.1.3. Suppose E is a real vector bundle over a manifold X, with fibers whose dimension is k. Let A be a connection on E with associated covariant derivative ∇ and curvature matrix Ω relative to any frame on E. If $P(X)$ is a homogeneous invariant polynomial of degree p on $\mathfrak{gl}_k(\mathbb{R})$, then $P(\Omega)$ defines a closed global $2p$-form on X, whose cohomology class $[P(\Omega)] \in H^{2p}(X)$ is independent of the connection.

Proof. This is proved as Theorem 23.3 in [2]. □

Definition 1.1.16. The cohomology class $[P(\Omega)]$ is called the **characteristic class** associated to P of the real vector bundle E.

Definition 1.1.17. With the polynomials $f_{2l}(X)$ defined as in Example 1.1.4, we define the l-th **Pontryagin class** $p_l(E)$ of a real vector bundle E over X with fiber dimension k to be

$$p_l(E) = \left[f_{2l}\Big(\frac{i}{2\pi}\Omega\Big)\right] \in H^{4l}(X).$$

For odd l, $p_l(E)$ are all 0 (see theorem 24.3 in [2]). So, we can expand $\det\left(\lambda I + \frac{i}{2\pi}\Omega\right)$ as

$$\det\left(\lambda I + \frac{i}{2\pi}\Omega\right) = \sum_{l=0}^{k} f_l\left(\frac{i}{2\pi}\Omega\right)\lambda^{k-l} = \sum_{i=0}^{\tilde{k}} p_i(E)\lambda^{k-2i},$$

where $\tilde{k} = \lfloor \frac{k}{2} \rfloor$.

Definition 1.1.18. Define the **total Pontryagin class** $p(E)$ of E to be the expression we get when we set $\lambda = 1$ in the equation above, i.e.

$$p(E) := \det\left(I + \frac{i}{2\pi}\Omega\right) = 1 + p_1 + ... + p_{\tilde{k}}.$$

Many times, it is useful to work with complex vector bundles. The analogue of Pontryagin classes on complex vector bundles are Chern classes. If $P(X)$ is a complex polynomial on $\mathfrak{gl}_k(\mathbb{C})$ that is invariant under conjugation by elements of $GL_k(\mathbb{C})$, then $P(\Omega)$ defines a closed global form on X, whose cohomology class $[P(\Omega)]$ is independent of the connection.

Definition 1.1.19. Setting $P(\Omega) = \det\left(I + \frac{i}{2\pi}\Omega\right)$, we obtain the **Chern classes** $c_i(E)$ of a complex vector bundle E from the **total Chern class**

$$c(E) := \det\left(I + \frac{i}{2\pi}\Omega\right) = 1 + c_1(E) + ... + c_k(E)$$

in the same way as we obtained the Pontryagin classes of real vector bundles above.

Both the Pontryagin and Chern classes satisfy the **Witney product formula**: Given complex vector bundles E and F,

$$c(E \oplus F) = c(E)c(F),$$

and if E and F are real vector bundles, then $p(E \oplus F) = p(E)p(F)$. See Theorem 24.6 in [2] for a proof. Furthermore, if E is a complex vector bundle, then

$$c_i(\overline{E}) = (-1)^i c_i(E).$$

See Lemma 1.4 in [3] for a proof. We will use this statement and the Whitney product formula to deduce the relationship between Chern classes of complex vector bundles and Pontryagin classes of their underlying real vector bundles.

Lemma 1.1.2. If E is a real vector bundle, then $c_i(E_{\mathbb{C}}) = c_i(E \otimes \mathbb{C}) = 0$ for all odd i.

Proof. Since $E_{\mathbb{C}}$ is isomorphic as a complex vector bundle to $\overline{E_{\mathbb{C}}}$,

$$c_i(E_{\mathbb{C}}) = c_i(\overline{E_{\mathbb{C}}}) = (-1)^i c_i(E_{\mathbb{C}}) = -c_i(E_{\mathbb{C}}),$$

for odd i. Thus, $2c_i(E_{\mathbb{C}}) = 0$. □

From this Lemma, we deduce that Pontryagin classes of a real vector bundle E over X arise as the even-indexed Chern classes of the complexification $E_{\mathbb{C}} = E \otimes \mathbb{C}$. That is, for a real vector bundle E,

$$p_i(E) = (-1)^i c_{2i}(E_{\mathbb{C}}).$$

We won't prove this fact here, but will use it to prove the following theorem.

Theorem 1.1.3. Let E be a complex vector bundle with underlying real vector bundle $E_{\mathbb{R}}$. Then,

$$1 - p_1(E_{\mathbb{R}}) + ... + (-1)^n p_n(E_{\mathbb{R}}) = (1 + c_1(E) + ... + c_n(E))(1 - c_1(E) + ... + (-1)^n c_n(E)).$$

Proof. Since $p_i(E_{\mathbb{R}}) = (-1)^i c_{2i}((E_{\mathbb{R}})_{\mathbb{C}}) = (-1)^i c_{2i}(E_{\mathbb{R}} \otimes \mathbb{C})$,

$$\begin{aligned} 1 - p_1(E_{\mathbb{R}}) + ... + (-1)^n p_n(E_{\mathbb{R}}) &= 1 + c_2(E_{\mathbb{R}} \otimes C) + ... + c_{2n}(E_{\mathbb{R}} \otimes C) \\ &= c(E_{\mathbb{R}} \otimes C), \end{aligned}$$

where in the last equality we have used the fact that $c_i(E_{\mathbb{R}} \otimes C) = 0$ for odd i (Lemma 1.1.2). By proposition 1.1.2 $E_{\mathbb{R}} \otimes \mathbb{C}$ is isomorphic to $E \oplus \overline{E}$. So,

$$\begin{aligned} c(E_{\mathbb{R}} \otimes C) &= c(E \oplus \overline{E}) \\ &= c(E)c(\overline{E}) \quad \text{(Whitney product formula)} \\ &= (1 + c_1(E) + ... + c_n(E))(1 + c_1(\overline{E}) + ... + c_n(\overline{E})) \\ &= (1 + c_1(E) + ... + c_n(E))(1 - c_1(E) + ... + (-1)^n c_n(E)), \end{aligned}$$

where in the last equality we have used the fact that $c_i(\overline{E}) = (-1)^i c_i(E)$. Thus,

$$1 - p_1(E_{\mathbb{R}}) + ... + (-1)^n p_n(E_{\mathbb{R}}) = (1 + c_1(E) + ... + c_n(E))(1 - c_1(E) + ... + (-1)^n c_n(E)).$$

□

Example 1.1.5. Suppose E is a complex vector bundle of complex rank 2 with underlying real vector bundle $E_{\mathbb{R}}$. Then, $(1+c_1(E)+c_2(E))(1-c_1(E)+c_2(E)) = 1-c_1(E)^2+2c_2(E)$. Thus, $p_1(E_{\mathbb{R}}) = c_1(E)^2 - 2c_2(E)$.

Remark 1.1.3. There is a rational polynomial in the Chern classes of vector bundles E with fiber dimension k called the **Chern character**, beginning with

$$\mathrm{ch}(E) = k + c_1(E) + \frac{1}{2}(c_1(E)^2 - 2c_2(E)) + \frac{1}{6}(c_1(E)^3 - 3c_1(E)c_2(E) + 3c_3(E)) + ...,$$

with the following properties

$$\mathrm{ch}(E \oplus F) = \mathrm{ch}(E) + \mathrm{ch}(F)$$
$$\mathrm{ch}(E \otimes F) = \mathrm{ch}(E)\mathrm{ch}(F)$$

for vector bundles E and F over X.

1.2 Connections, and Curvature on Principal G-Bundles

In this section, we will add some additional structure to fiber bundles. Connections on these fiber bundles that are, in some way, compatible with this structure become especially important. For example, general relativity concerns the study of tangent bundles over a spacetime endowed with a metric. The connections of import — called Riemannian connections — are those that are compatible with the metric (see Ch. 6.3 in [2]). In Yang–Mills theory, we examine fiber bundles with additional group structure — we call these *principal G-bundles*, where G is the structure group of the theory.

1.2.1 Principal G-Bundles

Definition 1.2.1. Let P equipped with a map $\pi : P \to X$ be a fiber bundle over a manifold X. Let $F = \pi^{-1}(x)$ be an arbitrary fiber of $x \in X$. Additionally, let G be a Lie group that acts on P by a smooth right action $R : P \times G \to P$. We call P a **principal G-bundle** (G-bundle for short) with **structure group** G if R is:

- *Fiber-preserving:* $R_g F \subseteq F$
- *Free:* If $g_1, g_2 \in G$ and $p \cdot g_1 = p \cdot g_2$ for some $p \in F$, then $g_1 = g_2$
- *Transitive:* For all $p, q \in F$, there exists a $g \in G$ such that $q = p \cdot g$.

With the definition above, we recognize a principal bundle P with structural group G over a smooth manifold X as a locally trivial fiber bundle whose fiber is G itself considered as a right G-space. Thus, G acts smoothly on P by right multiplication on each fiber and $X = P/G$.

Example 1.2.1 (Hopf bundle)**.** The group S^1 of unit complex numbers acts on the complex vector space $\mathbb{C}^{n+1}$ by left multiplication. This induces an action of S^1 on the unit sphere S^{2n+1} in $\mathbb{C}^{n+1}$. We may define the complex projective space $\mathbb{CP}^n$ as the orbit space of S^{2n+1} by S^1. The natural projection $S^{2n+1} \to \mathbb{CP}^n$ with fiber S^1 is a principal S^1-bundle. When $n = 1$, $S^3 \to \mathbb{CP}^1$ with fiber S^1 is called the **Hopf bundle**. Interestingly, the Hopf bundle appears in the mathematical study of quantum computing — the Hopf bundle explains how one can represent 2-level quantum states (*qubit* states) on a 2-sphere, called the *Bloch sphere* (see Ch. 8.2 in [4]).

Example 1.2.2 (Frame Bundle)**.** Let E be a vector bundle with fiber dimension k. Consider the space $\text{Fr}(E)$ whose fiber over a point $x \in E$ is the collection of all frames in the fiber E_x of E over x. Intuitively, one may act on an element in $\text{Fr}(E)$, a frame, by an invertible linear transformation (in $GL(k)$) to yield another frame in E_x. So, $\text{Fr}(E)$ is a principal $GL(k)$-bundle.

1.2.2 G-Equivariance

We aim to define a connection on principal G-bundles that are, in some way, compatible with the group structure. The compatiblity we are interested in is called *G-equivariance.*

Let us make this more precise.

Definition 1.2.2 (G-equivariance)**.** A manifold M equipped with a Lie group G that acts on M on the right (left) is called a **right (left) G-manifold**. We say that a map $f : M \to N$ between two right G-manifolds is **right G-equivariant** if for all $x \in M$ and $g \in G$,

$$f(x \cdot g) = f(x) \cdot g$$

We say that a map $f : M \to N$ between two left G-manifolds is **left G-equivariant** if for all $x \in M$ and $g \in G$,

$$f(g \cdot x) = g \cdot f(x)$$

If M is a right G-manifold and N is a left G-manifold, then a map $f : M \to N$ is **G-equivariant** if for all $x \in M$ and $g \in G$,

$$f(x \cdot g) = g^{-1} \cdot f(x)$$

Definition 1.2.3. Let $\text{conj}(g) : G \to G$ be a right group action on G given by conjugation

$$f \cdot g := \text{conj}(g)(f) = gfg^{-1}.$$

The differential of $\text{conj}(g)$ evaluated at the identity $e \in G$ gives the adjoint action

$$\text{ad}(g) = \text{conj}_*(g)(e) : T_eG \to T_eG.$$

We may identify $\mathfrak{g}$ with T_eG and invoke the chain rule to show that $\text{ad}(g_1) \circ \text{ad}(g_2) = \text{ad}(g_1 g_2)$. This gives us a homomorphism

$$\text{ad}(g) : G \to GL(\mathfrak{g})$$

called the **adjoint representation**.

Example 1.2.3. If G is a subgroup of $GL(k, \mathbb{R})$, then $\text{ad}(g)(\xi) = g\xi g^{-1}$.

Suppose G is a Lie group with associated Lie algebra $\mathfrak{g}$. Also suppose G acts smoothly on a manifold P on the right. To every element $A \in \mathfrak{g}$ we can associate a vector field $\underline{\text{A}}$ on P called the **fundamental vector field** on P associated to A as follows: For $x \in P$, define

$$\underline{\text{A}}_x = \frac{d}{dt}\bigg|_{t=0} x \cdot e^{tA} \in T_xP$$

$\underline{\text{A}}_x$ may be understood as the initial direction of the curve $c_x : t \mapsto x \cdot e^{tA}$ at a point $x \in P$:

$$\underline{\text{A}}_x = c_x'(0)$$

Proposition 1.2.1 (G-equivariance of Fundamental Vector Fields)**.** Suppose G is a Lie group with associated Lie algebra $\mathfrak{g}$. Suppose G acts smoothly on a manifold P by right-action $R_g : x \mapsto x \cdot g$. Then, for $A \in \mathfrak{g}$, the associated fundamental vector field $\underline{A}$ satisfies the following *G-equivariance* property:

$$(R_g)_*\underline{A} = \underline{\text{ad}(g^{-1})A}$$

Proof. We need to show that for all $x \in P$, $R_{g*}(\underline{A}_x) = \underline{\mathrm{ad}(g^{-1})A}_{xg}$.

For $x \in P$, define a map $j_x : G \to P$ by $j_x(g) = x \cdot g$. The differential of j_x using the curve $c(t) = e^{tA}$ is

$$j_{x*}(A) = \frac{d}{dt}\Big|_{t=0} j_x(e^{tA}) = \frac{d}{dt}\Big|_{t=0} (x \cdot e^{tA}) = \underline{A}_x$$

For $q \in G$,

$$(R_g \circ j_x)(q) = xqg = xgg^{-1}qg = j_{xg}(g^{-1}qg) = (j_{xg} \circ c_{g^{-1}})(q)$$

Thus,

$$R_{g*}(\underline{A}_x) = R_{g*}j_{x*}(A) = j_{xg*}c_{g^{-1}*}(A) = j_{xg*}(\mathrm{ad}(g^{-1})A) = \underline{\mathrm{ad}(g^{-1})A}_{xg} \quad \square$$

1.2.3 Horizontal and Vertical Tangent Bundles

Let P be a principal G-bundle. Let $\mathfrak{g}$ denote the Lie algebra associated with the Lie group G.

Definition 1.2.4. If we differentiate the G-action, we can associate each element u of $\mathfrak{g}$ with a G-invariant vector field X_u on P, called the **Killing field** corresponding to u. The Killing fields span a subbundle VP of TP, which is equal to the kernel of the map $\pi_* : TP \to TX$. Thus, each fiber of VP is canonically identified with $\mathfrak{g}$. We call VP the **subbundle of vertical tangent vectors** to P.

By the local triviality of a principal bundle, at every point $x \in P$, the differential $\pi_{*,x} : T_xP \to T_{\pi(x)}X$ of the projection is surjective. The vertical tangent subspace $V_xP \subseteq T_xP$ is $\ker(\pi_{*,x})$. This is precisely a short exact sequence of vector spaces

$$0 \to V_xP \to T_xP \xrightarrow{\pi_{*,x}} T_{\pi(x)}X \to 0$$

The differential $\pi_* : TP \to TX$ of $\pi : P \to X$ induces a bundle map $\tilde{\pi}_* : TP \to \pi^*TX$ over P, given by $\tilde{\pi}_{*,x}(X_x) = (x, \pi_{*,x}X_x)$. This is represented in the diagram

$$\begin{array}{ccc} TP & \xrightarrow{\tilde{\pi}_*} & \pi^*TX \\ & \searrow & \downarrow \\ & & TX \end{array}$$

$\tilde{\pi}_*$ is surjective because it maps the fiber T_xP onto the fiber $(\pi^*TX)_x \cong T_{\pi(x)}X$. Its kernel is VP by the short exact sequence above. Thus, there is a short exact sequence of vector bundles over P

$$0 \to VP \to TP \xrightarrow{\tilde{\pi}_*} \pi^*TX \to 0$$

If $v \in VP$, we say that v is **vertical**.

Proposition 1.2.2. For all $A \in \mathfrak{g}$, the fundamental vector field $\underline{A}$ is vertical at all points $x \in P$.

Proof. For $x \in P$, define $j_x : G \to P$ by $j_x(g) = x \cdot g$. Then,

$$(\pi \circ j_x)(g)) = \pi(x \cdot g) = \pi(x)$$

Since $\underline{A}_x = j_{x*}(A)$ (by proof of Proposition 1.2.1), and $\pi \circ j_x$ is a constant map,

$$\pi_{*,x}(\underline{A}_x) = (\pi_{*,x} \circ j_{x*})(A) = (\pi_x \circ j_x)_*(A) = 0$$

Thus, $\underline{A}_x \in \ker(\pi_{*,x})$ and $\underline{A}$ is vertical at all $x \in P$. □

Remark 1.2.1. Conversely, we can also identify every vertical tangent vector with a fundamental vector field of some element in $\mathfrak{g}$. More precisely, for $x \in P$, the differential at $e \in G$ of the map $j_x : G \to P$ is an isomorphism of $\mathfrak{g}$ onto the vertical tangent space: $(j_x)_{*,e} : \mathfrak{g} \xrightarrow{\sim} V_xP$.

Definition 1.2.5. Let P be a principal G-bundle with projection map $\pi : P \to X$ and vertical subbundle VP of the tangent bundle TP. We call a subbundle HP of TP a **bundle of horizontal tangent vectors** (or horizontal tangent subbundle) on P if $TP = VP \oplus FP$ as vector bundles, i.e. for all $x \in P$,

$$T_xP = V_xP + H_xP \text{ and } V_xP \cap H_xP = 0$$

1.2.4 Connections and Curvature on Principal G-Bundles

Suppose HP is a horizontal tangent subbundle of a principal G-bundle P equipped with projection $\pi : P \to X$. For $x \in P$, define a map $j_x : G \to P$ by $j_x(g) = x \cdot g$. This induces an isomorphism $j_{x*} : \mathfrak{g} \to V_xP$ (see Prop. 27.18 in [2]), allowing us to canonically identify the vertical tangent space V_xP with the Lie algebra $\mathfrak{g}$ associated with the Lie group G.

Let $\nu_x : T_xP = V_xP \oplus H_xP \to V_xP$ be the projection onto the vertical tangent space at the point $x \in P$. Note that this projection depends on the choice of horizontal tangent subbundle HP. Finally, define the map

$$\omega_x = j_{x*}^{-1} \circ \nu_x : T_xP \xrightarrow{\nu} V_xP \xrightarrow{j_{x*}^{-1}} \mathfrak{g}$$

With this definition, ω is a smooth $\mathfrak{g}$-valued 1-form on P.

Theorem 1.2.1. Let HP be a right-invariant horizontal tangent subbundle associated with a principal G-bundle P. Then, the associated $\mathfrak{g}$-valued 1-form as defined above satisfies the following conditions:

1. For all $A \in \mathfrak{g}$ and $x \in P$, we have $\omega_x(\underline{A}_x) = A$;
2. (G-equivariance) For all $g \in G$, $R_g^*\omega = \mathrm{ad}(g^{-1})\omega$

Proof. For 1, Since $\underline{A}_x$ is already vertical by Proposition 1.2.2, $\nu_x(\underline{A}_x) = \underline{A}_x$. So,

$$\omega_x(\underline{A}_x) = j_{x*}^{-1}(\nu_x(\underline{A}_x)) = j_{x*}^{-1}(\underline{A}_x) = A$$

For 2, let $x \in P$ and $X_x \in T_xP$ be arbitrary. Then, we need to show that

$$\mathrm{ad}(g^{-1})\omega_x(X_x) = R_g^*\omega_x(X_x) = \omega_{xg}(R_{g*}(X_x))$$

It suffices to show this for both the vertical and horizontal components of X_x since both sides of the above equation are $\mathbb{R}$-linear in X_x and X_x is the sum of its horizontal and vertical components. Let V_X be the vertical component of X_x and H_X be the horizontal component of X_x.

By remark 1.2.1, $V_X = \underline{A}_x$ for some $A \in \mathfrak{g}$. Then,

$$\begin{aligned}\omega_{xg}(R_{g*}(V_X)) = \omega_{xg}(R_{g*}(\underline{A}_x)) &= \omega_{xg}(\underline{\mathrm{ad}(g^{-1})A}_{xg}) \text{ (Proposition 1.2.1)} \\ &= \mathrm{ad}(g^{-1})A \text{ (by condition } \mathit{1}) \\ &= \mathrm{ad}(g^{-1})\omega_x(\underline{A}_x) \text{ (by condition } \mathit{1}) \\ &= \mathrm{ad}(g^{-1})\omega_x(V_X)\end{aligned}$$

Since the horizontal tangent subbundle HP is right-invariant, $R_{g*}H_X$ is horizontal. Thus, by definition of ω_x,

$$\omega_{xg}(R_{g*}(H_X)) = 0 = \mathrm{ad}(g^{-1})\omega_x(H_X)$$

□

Definition 1.2.6. A **connection** on a principal G-bundle is a $\mathfrak{g}$-valued smooth 1-form on P satisfying the conditions in Theorem 1.2.1.

A $\mathfrak{g}$-valued 1-form is a map $\alpha : TP \to \mathfrak{g}$. The Lie group G acts on TP on the right by the differentials of right translations and G acts on g on the left by the adjoint representation. By Definition 1.2.2, α is G-equivariant if and only if

$$\alpha(X_x \cdot g) = g^{-1} \cdot \alpha(X_x)$$

for all $x \in P$, $X_x \in T_xP$ and $g \in G$. Equivalently,

$$\alpha(R_{g*}X_x) = (\mathrm{ad}(g^{-1})) \cdot \alpha(X_x)$$

So, $\alpha : TP \to \mathfrak{g}$ is G-equivariant if and only if $R_g^*\alpha = \mathrm{ad}(g^{-1})\alpha$. Thus, condition *2* of a connection ω on a principal G-bundle says that $w : TP \to \mathfrak{g}$ is a G-equivariant map.

Definition 1.2.7. Let ω be a connection on a principal G-bundle P over X with values in $\mathfrak{g}$. Then, the **curvature** of ω is the $\mathfrak{g}$-valued 2-form

$$F_\omega = d\omega + \frac{1}{2}[\omega, \omega].$$

It turns out that the space of connections on a principal G-bundle is an affine space. To understand what this affine space is modeled on, we first introduce the concept of the adjoint bundle.

1.2.5 The Adjoint Bundle

Definition 1.2.8. Let P be a principal G-bundle over X. We can construct a vector bundle over X as follows: Let $\rho : G \to GL(F)$ be a representation of G on a vector space F. Then G operates on $P \times F$ by

$$(p, f) \cdot g = (pg^{-1}, \rho(g)f).$$

The quotient space $P \times_\rho F$ of $P \times F$ by the group action of G is a vector bundle over X with fibers isomorphic to F. $P \times_\rho F$ is the **vector bundle associated to P by the representation ρ.**

Example 1.2.4. $\mathrm{ad}(P) := P \times_{\mathrm{ad}} \mathfrak{g}$ is the vector bundle associated to P by the adjoint representation, simply called the **adjoint bundle** of P.

Theorem 1.2.2. For any principal G-bundle P over X, the space of all connections $\mathcal{A}(P)$ is an affine space modeled on $\Omega^1(X; \mathrm{ad}(P))$.

Although we will not prove this theorem here, its consequences are three-fold:

1. $\mathcal{A}(P)$ is nonempty
2. if $\nabla, \nabla' \in \mathcal{A}(P)$ are connections, then $\nabla - \nabla'$ is a 1-form on X with values in $\mathrm{ad}(P)$
3. If $\nabla \in \mathcal{A}(P)$ is a connection and $a \in \Omega^1(\mathrm{ad}(P))$, then $\nabla + a$ defined by

$$(\nabla + a)s = \nabla s + as$$

   is a connection.

Example 1.2.5. Let $P = \mathrm{Fr}(E)$ be the principal $GL(k)$-bundle of frames on a vector bundle E (see Example 1.2.2 for definition). Then $\mathrm{ad}(P) = \mathrm{End}(E)$ and the connections on P are $\mathfrak{gl}_k(\mathbb{R})$-valued 1-forms on E.

The adjoint bundle is also useful in understanding the curvature of principal connections. The curvature F_∇ of a connection ∇ on a principal G-bundle over X descends to X as a smooth section of $\mathrm{ad}(P) \otimes \Lambda^2$.

1.3 Gauge Transformations

1.3.1 Gauge Transformations

Definition 1.3.1. Let P be a principal G-bundle with projection $\pi : P \to X$. Denote the set of automorphisms of P by $\mathcal{G}(P)$:

$$\mathcal{G}(P) = \{\lambda : P \to P \mid \pi \circ \lambda = \pi \text{ and } \lambda(gp) = g\lambda(p) \text{ for all } p \in P, g \in G\}$$

$\mathcal{G}(P)$ is called the **gauge group** of P and its elements are called **gauge transformations**. When necessary, we will specify the association of the gauge group $\mathcal{G}$ to the structure group G by referring to gauge transformations in $\mathcal{G}$ as G**-gauge transformations**.

Proposition 1.3.1. Let G act on G by conjugation. Then, $\mathcal{G}(P)$ is in bijection with the set of G-equivariant maps $P \to G$.

Proof. Let $u : P \to G$ be a G-equivariant map, i.e. $u(pg^{-1}) = gu(p)g^{-1}$. This defines an automorphism $\lambda_u : P \to P$ by $\lambda_u(p) = pu(p)$. Note that

$$\lambda_u(pg^{-1}) = pg^{-1}u(pg^{-1}) = pg^{-1}gu(p)g^{-1} = pu(p)g^{-1} = \lambda_u(p)g^{-1}.$$

Now, let $\lambda : P \to P$ be an automorphism. Since λ preserves the base point of p, we may write $\lambda(p) = pu(p)$ for some $u : P \to G$. Since λ is an automorphism, $\lambda(pg^{-1}) = \lambda(p)g^{-1}$. So,

$$pg^{-1}u(pg^{-1}) = \lambda(pg^{-1}) = \lambda(p)g^{-1} = pu(p)g^{-1}.$$

Thus, $g^{-1}u(pg^{-1}) = u(p)g^{-1} \implies u(pg^{-1}) = gu(p)g^{-1}$ — so u is G-equivariant. □

Let $P \times_{\text{conj}} G$ be the fiber bundle associated to P where G acts on G by conjugation. Then for all $p \in P$ and $h \in G$, $g \in G$ acts on $(p, h) \in P \times G$ by

$$(p, h) \cdot g = (pg^{-1}, ghg^{-1}).$$

Sections of $P \times_{\text{conj}} G$ are precisely the G-equivariant maps $P \to G$. In other words, we can identify each section of $P \times_{\text{conj}} G$ with a map $u : P \to G$ such that

$$u(pg^{-1}) = gu(p)g^{-1}$$

(see Ch. 13.4 in [5]). Thus, Proposition 1.3.1 yields bijective descriptions of the gauge group:

$$\mathcal{G}(P) \cong C^\infty(P \times_{\text{conj}} G).$$

Example 1.3.1. The principal $G = GL_k(\mathbb{R})$-frame bundle $\text{Fr}(E)$ of a vector bundle E over X has gauge group

$$\mathcal{G}(\text{Fr}(E)) = \{\lambda \in C^\infty(\text{End}(E)) \mid \lambda(x) \in GL(E_x) \text{ for all } x \in X\}$$

So, we can identify $\mathcal{G}(\text{Fr}(E))$ with $GL(E)$.

Locally, a gauge transformation $\lambda \in \mathcal{G}(P)$ may be represented by a G-valued function on X and a connection A may be represented as a $\mathfrak{g}$-valued 1-form. Then, the action of λ on A is locally given by

$$\lambda^{-1}A = \lambda^{-1}d\lambda + \mathrm{ad}(\lambda^{-1})(A) = \lambda^{-1}d\lambda + \lambda^{-1}A\lambda.$$

Under a gauge transformation, the curvature F_A transforms to

$$F_{\lambda^{-1}A} = \lambda^{-1}F_A\lambda.$$

Definition 1.3.2. Two connections are called **gauge equivalent** if there is a gauge transformation that transforms one of the connections into the other one.

In addition to gauge equivalence, some quantities demonstrate *gauge invariance*, i.e. they are unchanged under gauge transformations. This idea is quite important in physics — physical observables are gauge invariant. For example, the electromagnetic field strength composed of the electric and magnetic fields is gauge invariant under $U(1)$-gauge transformations. We will see this in the following section.

1.3.2 Electromagnetism From Principal $U(1)$-Bundles

So far, we have constructed a general framework of connections and curvature as they relate to principal G-bundles. This framework can be applied to quickly demonstrate what we initially observed in the introduction of this thesis.

Let $E = X \times \mathbb{C}$ be a trivial complex line bundle over X. In other words, we assume $E = X \times \mathbb{C}$, so that the fiber E_x, over any point $x \in X$ equals $\mathbb{C}$. A connection ∇ on E can be identified with its connection matrix A, which is an $\mathrm{End}(E)$-valued 1-form. We call A a **vector potential**.

Since $\mathrm{End}(\mathbb{C})$ is canonically isomorphic to $\mathbb{C}$, A is a complex-valued 1-form. E becomes a $U(1)$-bundle if we think of its standard fiber, $\mathbb{C}$, as the fundamental representation of the group $U(1)$. In this case, the entries of A must live in $\mathfrak{u}(1) = \{ix \mid x \in \mathbb{R}\}$.

Now suppose we apply a gauge transformation λ to the vector potential A. Then, A will transform as

$$A' := \lambda^{-1}A = \lambda^{-1}A\lambda + \lambda^{-1}d\lambda = A + \lambda^{-1}d\lambda,$$

where the last equality follows from $U(1)$ being abelian. If we can write $\lambda = e^{\phi}$ for some imaginary function ϕ, then

$$\lambda^{-1}d\lambda = d\phi.$$

In this case, a gauge transformation amounts to adding an exact form $d\phi$ to A:

$$A' = A + d\phi.$$

This is precisely the gauge transform we observed in the introduction. Furthermore, the curvature F_A associated with this vector potential is invariant under $U(1)$-gauge transformations:

$$F_{A'} = \lambda^{-1} F_A \lambda = F_A.$$

This is also precisely what we observed in the introduction. This compact demonstration is our bridge into the geometric study of particle physics using gauge theory. Electromagnetism is the simplest gauge theory, a $U(1)$ abelian gauge theory. In the next chapter, we will examine nonabelian gauge theories that describe physics beyond electromagnetism. They will exhibit much richer properties of geometry and physics for us to study.

Chapter 2

Yang–Mills Gauge Theory

In this chapter, we will examine a spaces of solutions to the Yang–Mills equations, called *instantons*. We will concern ourselves with the spaces of such solutions modulo gauge equivalence, called *moduli spaces*. Finally, we will use the Atiyah–Singer theorem to prove the *Atiyah–Hitchin–Singer theorem*, a theorem about the dimension of moduli spaces of instantons in $SU(2)$-Yang–Mills theories.

2.1 Yang–Mills Instantons

2.1.1 Self-Dual Connections

Recall that on a principal G-bundle P over X, a connection A is defined by a $\mathfrak{g}$-valued 1-form, where $\mathfrak{g}$ is the Lie algebra associated with G. The curvature F_A of the connection is the $\mathfrak{g}$-valued 2-form

$$F_A = dA + \frac{1}{2}[A, A].$$

This descends to X as a smooth section of $\mathrm{ad}(P) \otimes \Lambda^2$, where $\mathrm{ad}(P) := P \times_{\mathrm{Ad}} \mathfrak{g}$ is the vector bundle associated to P by the adjoint representation.

Recall from Remark 1.1.1 that on a vector bundle E over X, a connection d_A (or just A) is a linear map

$$d_A : C^\infty(E \otimes \Lambda^0) \to C^\infty(E \otimes \Lambda^1).$$

This extends to a map

$$\begin{aligned} d_A : C^\infty(E \otimes \Lambda^1) &\to C^\infty(E \otimes \Lambda^2) \\ d_A(e \otimes \alpha) &= \nabla e \wedge \alpha + e \otimes d\alpha, \end{aligned}$$

where $e \in C^\infty(E \otimes \Lambda^0)$ and $\alpha \in \Lambda^1$. The curvature of the connection is

$$F_A = d_A^2.$$

Definition 2.1.1. Let X be an (even) $2n$-dimensional oriented manifold and let Λ^p denote the bundle of p-forms on X with inner product $\langle \cdot, \cdot \rangle$. The **Hodge star** operator $* : \Lambda^p \to \Lambda^{2n-p}$ maps p-forms β to $*\beta$ defined by

$$\alpha \wedge *\beta = \langle \alpha, \beta \rangle \mathrm{vol}$$

where $\alpha \in \Lambda^p$ and $\mathrm{vol} \in \Lambda^{2n}$ is the volume form on X.

It is worth noting that when $p = n$, $*^2 = (-1)^n$. So, for $n = 2$, i.e. X is a 4-dimensional manifold, $*^2 = 1$. In this case, the bundle Λ^2 is graded by $*$ into a direct sum

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

where $\Lambda_\pm$ are the ± 1-eigenspaces of $*$.

Definition 2.1.2. On a 4-manifold X, a connection A is said to be **self dual** if its curvature F_A is in $C^\infty(\mathrm{ad}(P) \otimes \Lambda^2_+)$ (i.e. $*F_A = F_A$) and **anti self dual** if F_A is in $C^\infty(\mathrm{ad}(P) \otimes \Lambda^2_-)$ (i.e. $*F_A = -F_A$).

Using the Hodge star operator we can decompose F_A into self dual and anti self dual components as

$$F_A = F_A^+ + F_A^-$$

where $*F_A^+ = F_A^+$ and $*F_A^- = -F_A^-$. So,

$$*F_A = F_A^+ - F_A^-.$$

Proposition 2.1.1. $F_A \wedge *F_A = F_A^+ \wedge *F_A^+ + F_A^- \wedge *F_A^-$.

Proof.

$$\begin{aligned} F_A \wedge *F_A &= (F_A^+ + F_A^-) \wedge (F_A^+ - F_A^-) \\ &= F_A^+ \wedge F_A^+ - F_A^+ \wedge F_A^- - F_A^- \wedge F_A^+ - F_A^- \wedge F_A^- \\ &= F_A^+ \wedge F_A^+ - F_A^- \wedge F_A^- \\ &= F_A^+ \wedge *F_A^+ + F_A^- \wedge *F_A^-. \end{aligned}$$

□

2.1.2 The Yang–Mills Equations

Let E be a Hermitian vector bundle over X with connection A whose curvature is F_A. We define the **Yang–Mills functional** on E by the integral

$$\mathcal{YM}(A) := \frac{1}{8\pi^2} \int_X F_A \wedge *F_A = \frac{1}{8\pi^2} \int_X \langle F_A, F_A \rangle \mathrm{vol}.$$

We are interested in studying Yang–Mills functionals because they are the *action* of massless bosons in physics. Yang–Mills functionals are crucial to our description of the standard model of particle physics — in particular the electromagnetic, electroweak, and strong interactions.

Motivated by the principle of least action in physics, we devote our attention to connections that extremize the Yang–Mills functional. The critical points of $\mathcal{YM}$ on the space of connections $\mathcal{A}(E)$ are called **Yang–Mills connections**.

Definition 2.1.3. Let A be a connection on a vector bundle E with associated exterior derivative $d_A : C^\infty(E \otimes \Lambda^p) \to C^\infty(E \otimes \Lambda^{p+1})$. We define the **formal adjoint** $d_A^* : C^\infty(E \otimes \Lambda^p) \to C^\infty(E \otimes \Lambda^{p-1})$ as the operator that satisfies

$$\langle d_A^* \alpha, \beta \rangle = \langle \alpha, d_A \beta \rangle.$$

One can write d_A^* explicitly as $d_A^* = \pm * d_A *$.

Proposition 2.1.2. Every Yang–Mills connection A satisfies the **Yang–Mills equations**

$$d_A * F_A = 0$$

Proof. Recall from Theorem 1.2.2 $\mathcal{A}(E)$ is an affine space modeled on the vector space $\Omega^1(\mathrm{End}(E))$ of $\mathfrak{g}$-valued 1-forms on $\mathrm{End}(E)$. Given a small perturbation $A + ta$ about A, where $t \in \mathbb{R}^+$ and $a \in \Omega^1(\mathrm{End}(E))$, the curvature is modified as

$$\begin{aligned}
F_{A+ta} &= d_A(A + ta) + \frac{1}{2}[(A + ta), (A + ta)] \\
&= d_A A + \frac{1}{2}[A, A] + t d_A a + \frac{1}{2} t^2 [a, a] \\
&= F_A + t d_A a + \frac{1}{2} t^2 [a, a].
\end{aligned}$$

We extremize $\mathcal{YM}$ by setting $\frac{d}{dt}\Big|_{t=0} \mathcal{YM}(A + ta) = 0$.

$$\begin{aligned}
\frac{d}{dt}\Big|_{t=0} \mathcal{YM}(A + ta) &= \frac{d}{dt}\Big|_{t=0} \int_X \langle F_A + t d_A a + \frac{1}{2} t^2 [a, a], F_A + t d_A a + \frac{1}{2} t^2 [a, a] \rangle \mathrm{vol} \\
&= \frac{d}{dt}\Big|_{t=0} \int_X (|F_A|^2 + t \langle F_A, d_A a \rangle) \mathrm{vol} \\
&\qquad \text{(higher order terms vanish at } t = 0) \\
&= \int_X \langle F_A, d_A a \rangle \mathrm{vol} \\
&= \int_X \langle d_A^* F_A, a \rangle \mathrm{vol} \\
&= 0
\end{aligned}$$

Since $a \neq 0$, $d_A^* F_A = \pm * d_A * F_A = 0$. Thus, $d_A * F_A = 0$. □

2.1.3 Instantons

We note that if A is a self dual connection, then $*F_A = F_A$ and $d_A * F_A = d_A F_A = 0$. The last equality is the Bianchi identity. Hence, the Yang–Mills equations are automatically satisfied for self dual connections. This warrants naming the self dual connections — we call them **Yang–Mills instantons** (or more simply just instantons).

Since an instanton A is self dual, $F_A^- = 0$. So,

$$F_A \wedge *F_A = F_A^+ \wedge *F_A^+ + F_A^- \wedge *F_A^- = F_A^+ \wedge *F_A^+.$$

This results in the following proposition.

Proposition 2.1.3. Instantons give absolute minima for the Yang–Mills functional on a vector bundle E over X.

Proof. To prove this proposition, we compare the Yang–Mills functional to the first Pontryagin class of E, a topological invariant. Given a connection A on E, the first Pontryagin class of E is

$$p_1(E) = -\frac{1}{4\pi^2}\int_X tr(F_A^2) = \frac{1}{4\pi^2}\int_X (F_A^+ \wedge *F_A^+ + F_A^- \wedge *F_A^-).$$

Comparing the Yang–Mills functional to $p_1(E)$, we get

$$\begin{aligned} \mathcal{YM}(A) &= \frac{1}{8\pi^2}\int_X |F_A|^2 = \frac{1}{8\pi^2}\int_X (F_A^+ \wedge *F_A^+ + F_A^- \wedge *F_A^-) \\ &\geq \frac{1}{8\pi^2}\int_X (F_A^+ \wedge *F_A^+ - F_A^- \wedge *F_A^-) = \frac{1}{2}p_1(E). \end{aligned}$$

Since $p_1(E)$ is a topological invariant, $\mathcal{YM}(A)$ is minimized if and only if $|F_A^-|^2 = 0$, i.e. when A is an instanton. When A is an instanton, $\mathcal{YM}(A) = \frac{1}{2}p_1(E)$. □

2.1.4 Moduli Spaces of Instantons

The Yang–Mills functional exhibits a particularly interesting property — it is invariant under gauge transformations, i.e.

$$\mathcal{YM}(\lambda \cdot A) = \frac{1}{8\pi^2}\int_X (\lambda^{-1}F_A\lambda) \wedge *(\lambda^{-1}F_A\lambda) = \frac{1}{8\pi^2}\int_X F_A \wedge *F_A = \mathcal{YM}(A),$$

where λ is a gauge transformation acting on connections. Gauge invariance of $\mathcal{YM}$ is incredibly important in physics because it implies that physical observables are the same under different gauges — physicists call this *gauge symmetry*. Gauge invariance of the Yang–Mills functional is interesting to mathematicians because it allows them to study spaces of instantons modulo gauge transformations, called *moduli spaces*.

Definition 2.1.4. Let P be a principal G-bundle over X. The **moduli space** of instantons on P is the space of instantons on P modulo gauge transformations.

If $H \subset G$ is a subgroup, then any self dual H-connection defines a self dual G-connection so that the moduli space of G-instantons contains H-instantons for all $H \subset G$. So, it makes more sense to consider the moduli space of G-instantons for which the instantons do not reduce to any proper closed subgroup $H \subset G$. We say that such instantons are **irreducible**. Denoting the space of irreducible instantons on P as $\mathcal{A}^+(P)$, the moduli space of irreducible instantons on P is

$$\mathcal{M}(P) := \mathcal{A}^+(P)/\mathcal{G}(P).$$

Remark 2.1.1. A rather surprising fact about moduli spaces of irreducible instantons is that, under fairly certain conditions, they are finite-dimensional (Hausdorff) manifolds. Without restricting to irreducible instantons, moduli spaces are not smooth manifolds.

A more surprising claim about moduli spaces of irreducible instantons is that when they are finite dimensional manifolds, we can calculate their dimensionality. The following theorem (called the *Atiyah–Hitchin–Singer theorem*) on moduli spaces of irreducible $SU(2)$-instantons on S^4 is an example. It will be the final focus of this document.

Theorem 2.1.1 (Atiyah–Hitchin–Singer Theorem)**.** The dimension of the moduli space of instantons on a principal $SU(2)$-bundle over S^4, with Pontryagin index $k > 0$ is $8k - 3$.

Before we prove this theorem, we first need to introduce a powerful tool called the *Atiyah–Singer index theorem*, which will help us prove the Atiyah–Hitchin–Singer theorem.

2.2 An Introduction to Index Theory

The Atiyah–Singer theorem is one of the most powerful theorems in mathematical physics. It has applications in the study of moduli spaces in mathematics and quantum field theories in physics. It equates (local) topological data about the characteristic classes of elliptic operators to their (global) analytic data about the number of linearly independent solutions to the operator's homogeneous differential equation. The data being equated here is something we call the *index* of the operator.

2.2.1 Fredholm Operators and The Index

Let H be a (separable) complex Hilbert space, and let $\mathcal{B}$ denote the Banach algebra of bounded linear operators $T : H \to H$ with finite operator norm. The norm of the operator T is defined by

$$||T|| = \sup\{||Tu|| \mid ||u|| \leq 1\} < \infty$$

Definition 2.2.1. An operator $T \in \mathcal{B}$ is called a **Fredholm operator** if it has finite dimensional kernel and cokernel. Recall that

$$\ker(T) = \{u \in H \mid Tu = 0\} \text{ and } \operatorname{coker}(T) = H/\operatorname{im}(T)$$

So, for a Fredholm operator T, $Tu = 0$ has finitely many linearly independent solutions. Additionally, to solve $Tu = v$, it is sufficient that v satisfy a finite number of linear conditions.

Definition 2.2.2. We define the **index** of a Fredholm operator T by

$$\operatorname{index}(T) = \dim(\ker(T)) - \dim(\operatorname{coker}(T))$$

Example 2.2.1. Let $L^2(\mathbb{Z}_+)$ be the (Hilbert) space of sequences $c = (c_0, c_1, ...)$ of complex numbers with square-summable absolute values, i.e.

$$\sum_{n=0}^{\infty} |c_n|^2 < \infty$$

Define the $\text{shift}^{\pm} : L^2(\mathbb{Z}_+) \to L^2(\mathbb{Z}_+)$ operators by

$$\text{shift}^+(c_0, c_1, c_2, ...) = (0, c_0, c_1, c_2, ...) \text{ and } \text{shift}^-(c_0, c_1, c_2, ...) = (c_1, c_2, ...)$$

It is easy to see that $\ker(\text{shift}^+) = 0$ and the cokernel is 1-dimensional. So, $\dim(\ker(\text{shift}^+)) = 0$ and $\dim(\operatorname{coker}(\text{shift}^+)) = 1$ Thus, shift^+ is Fredholm with

$$\operatorname{index}(\text{shift}^+) = 0 - 1 = -1$$

Similarly, one can show that shift^- is Fredholm with

$$\operatorname{index}(\text{shift}^-) = 1 - 0 = 1$$

Proposition 2.2.1. For finite dimensional vector spaces H, H', a linear transformation $T : H \to H'$ is Fredholm with $\text{index}(T) = \dim(H) - \dim(H')$.

Proof. Let H, H' be finite dimensional and $T : H \to H'$ be linear. Then, we have vector space isomorphisms

$$H/\ker(T) \to \text{im}(T) \to H'/\text{coker}(T)$$

So, $\dim(H) - \dim(\ker(T)) = \dim(H') - \dim(\text{coker}(T))$. Thus,

$$\text{index}(T) = \dim(\ker(T)) - \dim(\text{coker}(T)) = \dim(H) - \dim(H') \quad \square$$

Remark 2.2.1. The index of a Fredholm operator generalizes to the Euler characteristic $\chi(C)$ of a complex C:

$$\ldots \xrightarrow{T_{k+1}} V_k \xrightarrow{T_k} V_{k-1} \xrightarrow{T_{k-1}} V_{k-2} \xrightarrow{T_{k-2}} \ldots$$

of vector spaces and linear maps such that $T_k \circ T_{k+1} = 0$ with finite **Betti numbers**

$$b_k := \dim(\ker(T_k)/\text{im}(T_{k+1})).$$

We call $\ker(T_k)/\text{im}(T_{k+1})$ the k-th **cohomology space** $H^k(C)$. If all the Betti numbers of a complex are finite and the number of nonzero betti numbers are finite, then we define the **Euler characteristic** of the complex as

$$\chi(C) := \sum_i (-1)^k b_k.$$

The index of a Fredholm operator $T : H \to H$ is then given by the Euler characteristic of the complex

$$0 \to H \xrightarrow{T} H \to 0.$$

We will be concerned with studying the indices of Dirac operators, which act on sections of Clifford bundles. In the next section, we make these terms clearer.

2.2.2 Clifford Bundles and Dirac Operators

We being our discussion of Dirac operators with a look at Laplacians in some of the most important equations in physics, like the Schrödinger equation and Klein-Gordon equation. The Schrödinger equation is derived by "quantizing" the energy equation $E = T + V = \frac{p^2}{2m} + V$. Heuristically, this amounts to transforming functions (observables) into operators: $E \mapsto i\partial_t$, $p \mapsto -i\nabla$ (the spatial Laplacian), $V \mapsto V$ (whose operation is multiplication by V). Each of these operators act on a quantum state ψ. Then, we get the *Schrödinger equation*

$$E = \frac{p^2}{2m} + V \mapsto i\partial_t\psi = (-\frac{1}{2m}\nabla^2 + V)\psi$$

This equation, however, is inconsistent with special relativity. The energy equation from special relativity is $(-E^2+p^2)+m^2 = 0$. Quantizing in the same way, where operators act on "spinor fields" ψ, we get the *Klein-Gordon equation* that is satisfied by all components of free quantum fields.

$$(-E^2 + p^2) + m^2 = 0 \mapsto (\partial^2 + m^2)\psi = 0, \text{ where } \partial^2 = \partial_\mu \partial^\mu = \partial_t^2 - \nabla$$

We note that the Klein-Gordon equation is second-order in time, whereas the Schrödinger equation is of first-order in time. This gives the Schrödinger equation a particular advantage: time evolution. Knowledge of ψ at one given time allows you to determine ψ to any other point in time. This is not true of the Klein-Gordon equation, precisely due to the absence of first order derivatives in time. One might ask if there is a first-order linear differential equation describing quantum fields. There is one: it is called the *Dirac equation.*

$$(i\not{\partial} - m)\psi = 0$$

where $\not{\partial} = \gamma^\mu \partial_\mu$ is a *Dirac operator* for matrices γ^μ defined so that $(i\not{\partial} + m)(i\not{\partial} - m) = -(\partial^2 + m^2)$, i.e. multiplying the Dirac equation by the operator $(i\not{\partial}+m)$ yields the Klein-Gordon equation. Loosely, the Dirac equation is a factorization of the Klein-Gordon equation whose Dirac operator squares to the Laplacian. The matrices γ^μ are elements of a Clifford algebra, a concept made more precise below.

Definition 2.2.3. Let V be a vector space equipped with a symmetric bilinear form $\langle \cdot, \cdot \rangle$. A Clifford algebra for V is defined to be a unital algebra A which is equipped with a map $\phi : V \to A$ such that

1. $\phi(v)^2 = -\langle v, v \rangle 1$
2. If $\phi' : V \to A'$ is another map such that $\phi'(v)^2 = -\langle v, v \rangle 1$, then the following diagram commutes:

$$\begin{array}{ccc} V & \xrightarrow{\phi} & A \\ & \overset{\phi'}{\searrow} & \downarrow \\ & & A' \end{array}$$

Exercise 2.2.1. The second condition implies that the Clifford algebra for V (denoted $C\ell(V)$) is unique up to isomorphism.

Example 2.2.2. If V is equipped with a bilinear form that is identically 0, then the associated Clifford algebra is the exterior algebra $\Lambda^\bullet V$.

Definition 2.2.4. Let V be a real inner product space with orthonormal basis $e_1, ..., e_n$. Let S be a vector space which is also a left module over $C\ell(V)$ and let $C^\infty(V; S)$ denote the smooth S-valued functions on V. Each basis element e_i corresponds to a differential operator ∂_i on $C^\infty(V; S)$. Define an operator D on $C^\infty(V; S)$ by

$$Ds = \sum_i e_i(\partial_i s)$$

Then,

$$D^2 s = \sum_{i,j} e_i \partial_i (e_j \partial_j s) = \sum_{i,j} e_i e_j \partial_i \partial_j s = -\sum_i \partial_i^2 s$$

This is the Euclidean Laplacian. We call D the **Dirac operator** on $C^\infty(V; S)$.

We may extend this flat-space construction of the Dirac operator to a Riemannian manifold using Clifford bundles. If M is a Riemannian manifold, then TM is a bundle whose fibers are inner product spaces, so we may form the bundle of Clifford algebras $C\ell(TM)$. Let S be a bundle of Clifford modules, i.e. the fiber S_m at $m \in M$ is a left module over $C\ell(T_m M) \otimes \mathbb{C}$. The sections of S play the role of S-valued functions in the above flat-spaced construction. To differentiate these sections, we need to define a connection on S that is compatible with the metric on M. Here's how we define compatibility:

Definition 2.2.5. Let S be a bundle of Clifford modules over a Riemannian manifold M. S is a **Clifford bundle** if it is equipped with a Hermitian metric and compatible connection ∇ such that

1. The Clifford action of each $v \in T_m M$ on S_m is skew-adjoint, i.e. $\langle v \cdot s_1, s_2 \rangle + \langle s_1, v \cdot s_2 \rangle = 0$.
2. The connection on S is compatible with the Levi-Civita connection on M, i.e. $\nabla_X (Ys) = (\nabla_X Y)s + Y \nabla_X s$ for all vector fields X, Y and sections $s \in C^\infty(S)$.

Definition 2.2.6. The **Dirac operator** D of a Clifford bundle S is the first order differential operator on $C^\infty(S)$ defined by the following composition:

$$C^\infty(M; S) \xrightarrow{\nabla} C^\infty(T^*M \otimes S) \xrightarrow{g^\sharp} C^\infty(TM \otimes S) \xrightarrow{\cdot} C^\infty(M; S).$$

The first arrow is given by the connection. The second arrow is given by the metric, identifying TM with T^*M (in language more familiar to physicists, the metric maps covariant vectors to contravariant vectors). The third arrow is given by the Clifford action.

Thus, we may write D in terms of a local orthonormal basis e_i of sections of TM as

$$Ds = e_i g^{ij} \nabla_j s = e^i \nabla_i s,$$

where g^{ij} is the metric identifying TM and T^*M, $e^j = e_i g^{ij}$, and we implicitly sum over repeated indices. We may now calculate D^2. Choose an orthonormal frame e_i which is *synchronous* at some point $x \in M$, meaning at x, $\nabla_i e_j = 0$ and the Lie bracket of e_i and e_j vanishes at x. Thus at x,

$$\begin{aligned} D^2 s &= e^i \nabla_i e^j \nabla_j s \\ &= e^i e^j \nabla_i \nabla_j s && \text{(Leibniz rule)} \\ &= \nabla^i \nabla_i s + \sum_{j<i} e^i e^j (\nabla_i \nabla_j - \nabla_j \nabla_i) s \\ &= \nabla^i \nabla_i s + \sum_{j<i} e^i e^j K(e_i, e_j) s, \end{aligned}$$

where $K(e_i, e_j) = \nabla_i \nabla_j - \nabla_j \nabla_i$ is the curvature of the connection on S, and is an endomorphism of S. We call the term $\sum_{j<i} e^i e^j K(e_i, e_j)s$ the **Clifford contraction** $\mathcal{K}$ of the curvature applied to s. Additionally, we may write $\nabla^i \nabla_i$ as $\nabla^* \nabla$, where ∇^* is the formal adjoint of ∇ when we think of ∇ as a map $C^\infty(S) \to C^\infty(T^*M \otimes S)$. Here, the space of sections are equipped with the natural inner products associated with the metrics on their respective bundles. Thus,

$$D^2 s = \nabla^* \nabla s + \mathcal{K} s$$

This is called the **Weitzenbock formula**.

For the following discussion, it is worth describing additional structure called *grading* on Clifford bundles.

Definition 2.2.7. A module W over a Clifford algebra $C\ell(V)$ is **graded** if it is provided with a decomposition $W = W_+ \oplus W_-$ such that Clifford multiplication by any $v \in V$ interchanges the summands W_+ and W_-. A Clifford bundle S on Riemannian manifolds M is **graded** if it is provided with a decomposition $S = S_+ \oplus S_-$ that respects the metric and connection and makes each fiber S_m a graded Clifford module over $C\ell(T_m M)$.

The Dirac operator of a graded Clifford bundle anticommutes with the grading operator. So, it maps sections of $S_\pm$ to sections of $S_\mp$. We thus have maps

$$C^\infty(S_+) \xrightarrow{D_+} C^\infty(S_-) \xrightarrow{D_-} C^\infty(S_+)$$

where D_+ is the restriction of D to sections of S_+ and its adjoint $D_+^* = D_-$ is the restriction of D to sections of S_-.

Definition 2.2.8. The index of a graded Dirac operator D is

$$\text{index}(D_+) = \dim(\ker(D_+)) - \dim(\ker(D_-))$$

Example 2.2.3. Consider the de Rham operator $D = d + d^*$ with the grading operator defined by $\epsilon = (-1)^q$ on $\Omega^q(M)$. Then, Hodge theory tells us that index(D) is the Euler characteristic $\chi(M)$ of M in the topological sense.

2.2.3 Elliptic Operators

An important property of Dirac operators is that they are elliptic. In this section, we explain what that means and why it's important.

Definition 2.2.9. A linear differential operator $D : C^\infty(E) \to C^\infty(E)$ of order n on sections of vector bundles E over X is a map that can be written in local coordinates on X as

$$Ds(x) = \sum_{|\alpha| \leq n} c_\alpha(x) D^\alpha s$$

where $\alpha = (\alpha_1, ..., \alpha_n) \in (\mathbb{Z}_{\geq 0})^n$ is a multiindex with $|\alpha| = \sum_{i=1}^n \alpha_i$, for each α, $c_\alpha(x) : E \to E$ is a bundle homomorphism, and

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} ... \partial x_n^{\alpha_n}}.$$

Example 2.2.4. The Laplacian $\partial^2 = -\sum_{i=1}^n \partial_i^2$ is a differential operator.

Definition 2.2.10. We may construct a polynomial p_D by replacing each D^α with $\xi^\alpha := \xi_1^{\alpha_1} ... \xi_n^{\alpha_n}$, where $\xi = \xi_i dx^i \in T^*X$ is a covector field, as follows:

$$p_D(\xi) := \sum_{|\alpha| \leq n} c_\alpha(x) \xi^\alpha.$$

The highest homogeneous component of p_D,

$$\sigma_D(\xi) := \sum_{|\alpha| = n} c_\alpha(x) \xi^\alpha,$$

is called the **principal symbol** of D.

Example 2.2.5. The principal symbol of the Laplacian is $p_{\partial^2}(\xi) = -(\xi_1^2 + ... + \xi_n^2)$.

Definition 2.2.11. We say that an operator D is **elliptic** if for each nonzero $\xi \in T^*X$, the principal symbol $\sigma_D(\xi)$ is invertible. Equivalently, D is elliptic if the symbol is nonzero whenever ξ is nonzero.

Example 2.2.6. The Laplacian is elliptic. Dirac operators are also elliptic because they square to Laplacians.

We may extend the idea of ellipticity from differential operators to differential complexes using the concept of exact sequences.

Definition 2.2.12. A sequence

$$V_0 \xrightarrow{T_0} V_1 \xrightarrow{T_1} ... \xrightarrow{T_m} V_m$$

of vector spaces and linear maps is **exact** if $\ker(T_{i+1}) = \mathrm{im}(T_i)$.

Definition 2.2.13. Suppose we have a complex

$$C^\infty(E_0) \xrightarrow{D_1} C^\infty(E_1) \xrightarrow{D_2} ... \xrightarrow{D_m} C^\infty(E_m)$$

of differential operators on sections of vector bundles E_i such that $D_{i+1} \circ D_i = 0$. A differential complex with first order operators is **elliptic** if the sequences of symbols

$$0 \to \pi^* E_0 \xrightarrow{\sigma_{D_1}} \pi^* E_1 \xrightarrow{\sigma_{D_2}} ... \xrightarrow{\sigma_{D_m}} \pi^* E_m$$

is exact outside the zero section. Here, $\pi : T^*X \to X$ is the projection map and π^* is the pullback of a vector bundle.

Remark 2.2.2. Although we will not prove it here, it is worth noting that elliptic complexes have finite dimensional cohomology spaces. This is proved in Theorem 5.2 in [6].

2.3 The Atiyah–Hitchin–Singer Theorem

We are now ready to prove the Atiyah–Hitchin–Singer theorem.

Theorem 2.3.1 (Atiyah–Hitchin–Singer Theorem)**.** The dimension of the moduli space of instantons on a principal $SU(2)$-bundle over S^4, with Pontryagin index $k > 0$ is $8k - 3$.

Unexpectedly, we will find it easier to prove this theorem in greater generality first, in what I am calling the generalized Atiyah–Hitchin–Singer theorem.

Theorem 2.3.2 (Generalized Atiyah–Hitchin–Singer Theorem)**.** The dimension of the moduli space of instantons on a principal $SU(2)$-bundle P over S^4, with Pontryagin class $p_1(\mathrm{ad}(P)_{\mathbb{R}}) > 0$ of the underlying real vector bundle of the adjoint bundle of P, is $p_1(\mathrm{ad}(P)_{\mathbb{R}}) - 3$.

Following the proof of this theorem in [7] and theorem 6.1 in [8], one can prove this theorem in three steps:

1. Infinitesimal: Calculate the dimension of the space of infinitesimal deformations of an instanton.

2. Local: Integrate the local deformations to obtain a local moduli space.

3. Global: Show that the local moduli spaces give local coordinates on the global moduli space.

In our proof of the Atiyah–Hitchin–Singer theorem, we will need to use some results from applying the Atiyah–Singer index theorem to principal $SU(2)$-bundles on the 4-sphere.

Theorem 2.3.3 (Atiyah–Singer index theorem)**.** Let E be a vector bundle over X, a compact oriented even-dimensional manifold and let S be a canonically graded Clifford bundle over X with associated Dirac operator $D : C^\infty(S_+ \otimes E) \to C^\infty(S_- \otimes E)$. Then,

$$\mathrm{index}(D) = \mathrm{ch}(E)\hat{\mathcal{A}}(X)[X],$$

where $\mathrm{ch}(E)$ is the Chern character of E and $\hat{\mathcal{A}}(X)$ is what we call the $\hat{\mathcal{A}}$-*genus* of X.

Proof. For brevity, we omit the proof of this theorem. However, there are multiple ways to prove it and multiple textbooks dedicated to doing so. For example, Roe's *Elliptic Operators, Topology, and Asymptotic Methods* [9] proves the Atiyah–Singer theorem using the heat equation and an asymptotic expansion of the heat kernel. Bleecker and Booß-Bavnek's *Index Theory with Applications to Mathematics and Physics* [10] proves it using tools from topological K-theory. □

Example 2.3.1. Since we only need the Atiyah–Singer theorem to prove the Atiyah–Hitchin–Singer theorem, we will only be concerned with $X = S^4$ and $G = SU(2)$. On a 4-manifold, $\hat{\mathcal{A}}(X) = 1 - \frac{1}{24}p_1(X)$. On S^4, $p_1(S^4) = 0$. So,

$$\hat{\mathcal{A}}(S^4) = 1$$

.

We will use the Atiyah–Singer index theorem to achieve the first step in 2 sub-steps detailed below. Since our focus is on the application of the Atiyah–Singer index theorem to prove this theorem, we will forgo a rigorous proof of the second and third steps (which do not require the use of the Atiyah–Singer index theorem). I will instead give a brief explanation for how one might achieve them.

2.3.1 Step 1.1: Tangent Moduli Spaces

Let $\mathfrak{g} = \mathfrak{su}(2)$. To simplify notation, we denote the space of sections of $\text{ad}(P) \otimes \Lambda^p$ as $\Omega^p(\text{ad}(P))$, where $\text{ad}(P) = P \times_{\text{Ad}} \mathfrak{g}$ is the adjoint bundle of P. Assume there is at least one instanton $A \in \mathcal{A}^+(P)$, where $\mathcal{A}^+(P)$ is the set of self dual connections on P. Suppose A' is another connection. Since $A, A' : \Omega^0(\text{ad}(P)) \to \Omega^1(\text{ad}(P))$, by Theorem 1.2.2,

$$\tau := A' - A \in \Omega^1(\text{ad}(P)).$$

The difference in curvature is $F_{A'} - F_A = d_A\tau + \frac{1}{2}[\tau, \tau]$, where $d_A : \Omega^1(\text{ad}(P)) \to \Omega^2(\text{ad}(P))$ is the exterior derivative associated with A.

Now, let A_t be a family of instantons on P parametrized by t, with $A_0 = A$ and $\tau_t := A_t - A_0$. Then,

$$F_{A_t} - F_A = d_A\tau_t + \frac{1}{2}[\tau_t, \tau_t].$$

Since A and A' are self dual connections, $*F_A = F_A$ and $*F_{A_t} = F_{A_t}$. So,

$$*\left(d_A\tau_t + \frac{1}{2}[\tau_t, \tau_t]\right) = *(F_{A_t} - F_A) = F_{A_t} - F_A = d_A\tau_t + \frac{1}{2}[\tau_t, \tau_t].$$

So, given the projection map

$$p_- : \Lambda^2 \to \Lambda^2_-$$
$$p_-(\alpha) = \frac{1}{2}(\alpha - *\alpha)$$

onto anti self dual 2-forms,

$$p_-\left(d_A\tau_t + \frac{1}{2}[\tau_t, \tau_t]\right) = 0.$$

Differentiating this in t and setting $t = 0$ yields

$$p_-(d_A\dot{\tau}) = 0 \in \Omega^2_+,$$

where $\dot{\tau} = \frac{\partial \tau_t}{\partial t}\Big|_{t=0}$. The derivative of $\frac{1}{2}[\tau_t, \tau_t]$ vanishes when $t = 0$ because $[\tau_t, \tau_t]$ is quadratic in t. We see this by noticing that

$$\tau_0 = 0 \implies \tau_t = tg(t) \implies [\tau_t, \tau_t] = t^2[g(t), g(t)]$$

for some $g(t) \in \Omega^1(\text{ad}(P))$. If the family of instantons is generated by gauge transformations, i.e. $A_t = f_t^{-1} \cdot A$, then

$$\dot{\tau} = d_A f,$$

where $f \in \Omega^0(\mathrm{ad}(P))$. Thus, $\dot{\tau}$ defines an element

$$[\dot{\tau}] \in \ker(p_- d_A)/\mathrm{im}(d_A).$$

The space $\ker(p_- d_A)/\mathrm{im}(d_A)$ is well-defined because $\mathrm{im}(d_A) \in \ker(p_- d_A)$. This results from $p_- d_A(d_A) = p_-(d_A^2) = p_-(F_A) = 0$ (since A is self dual). Thus, $[\dot{\tau}]$ is an element of the first cohomology group $H_A^1(\mathrm{ad}(P))$ of the complex

$$0 \to \Omega^0(\mathrm{ad}(P)) \xrightarrow{\nabla} \Omega^1(\mathrm{ad}(P)) \xrightarrow{p_- d_A} \Omega^2_-(\mathrm{ad}(P)).$$

This complex is elliptic and so its cohomology groups are finite dimensional by Remark 2.2.2. We also note that since $\dot{\tau}$ is an infinitesimal deformation of A, it represents a vector in the space tangent to $\mathcal{M}(P)$. So,

$$T_{[A]}\mathcal{M}(P) \cong H_A^1(\mathrm{ad}(P))$$

for $[A] \in \mathcal{M}(P) = \mathcal{A}^+(\mathcal{P})/\mathcal{G}(P)$. Thus, our goal is to calculate the dimension of $H_A^1(\mathrm{ad}(P))$. We appeal to the index theorem to do this.

2.3.2 Step 1.2: A Dirac Operator and the Index Theorem

Before we apply the index theorem, we need to replace this complex by a Dirac operator, namely the elliptic operator

$$p_- d_A + d_A^* : \Omega^1(\mathrm{ad}(P)) \to \Omega^2_-(\mathrm{ad}(P)) \oplus \Omega^0(\mathrm{ad}(P)),$$

where d_A^* is the formal adjoint of d_A. We can do this because $\ker(p_- d_A + d_A^*) = \ker(p_- d_A)/\mathrm{im}(d_A)$. We see this by noting that $\ker(d_A^*) = \mathrm{im}(d_A)^\perp$. Additionally, $\mathrm{coker}(p_- d_A + d_A^*) = 0$. This is not so obvious, even though $\mathrm{coker}(p_- d_A) = 0$ (due to positive scalar curvature of S^4) and $\mathrm{coker}(d_A^*) = 0$ (due to irreducibility of A) individually. Thus,

$$\dim(\ker(p_- d_A)/\mathrm{im}(d_A)) = \dim(\ker(p_- d_A + d_A^*)) = \mathrm{index}(p_- d_A + d_A^*).$$

So,

$$\mathrm{index}(p_- d_A + d_A^*) = h^1.$$

In fact, $p_- d_A + d_A^*$ is a Dirac operator and can be written in terms of the Dirac operator associated to the Riemannian metric on X:

$$D : C^\infty(S_+ \otimes S_- \otimes \mathrm{ad}(P)) \to C^\infty(S_- \otimes S_- \otimes \mathrm{ad}(P)),$$

for associated graded Clifford bundles S. Here, $S_+ \otimes S_- = \Lambda^1$ and $S_- \otimes S_- = \Lambda^2_- \oplus \Lambda^0$.

Remark 2.3.1. It is worth noting that S_- is an $SU(2)$-bundle with fiber dimension 2. It has first Chern number $c_1(S_-) = 0$. In fact, if E is an $SU(n)$-bundle, $c_1(E) = 0$. This has to do with the fact that the determinant bundle $\det(E)$ is canonically trivialized for an $SU(n)$-bundle E and Chern classes obey $c_1(E) = c_1(\det(E))$.

Taking $E = S_- \otimes \mathrm{ad}(P)$, we have

$$D : C^\infty(S_+ \otimes E) \to C^\infty(S_- \otimes E).$$

Before we compute the index for D, we will first compute the index for

$$D_{S_+} : C^\infty(S_+) \to C^\infty(S_-),$$

the Dirac operator associated with $p_- d_A + d_A^* : \Omega^1(\mathrm{ad}(P)) \to \Omega^2_-(\mathrm{ad}(P)) \oplus \Omega^0(\mathrm{ad}(P))$ without twisting by E. This computation will be useful when we compute the index of D with twisting by E. Using Betti numbers, Hodge theory gives us

$$\mathrm{index}(D_{S_+}) = b_1 - b^- - 1,$$

where b_1 is the kernel dimension (harmonic 1-forms) and $b^- + 1$ is the cokernel dimension (anti self dual 2-forms and functions). From Remark 2.2.1, we know that the Euler characteristic is

$$\chi(S^4) = b_0 - b_1 + b_2 - b_3 + b_4 = 1 - b_1 + b_2 - b_1 + 1 = 2 - 2b_1 + b_2,$$

where we have used $b_i = b_{4-i}$ resulting from Poincaré duality (see Theorem 2.5 and Corollary 2.6 in [6]). Using $b_2 = b^+ + b^-$, we can write the signature τ of S^4 as

$$\tau(S^4) = b^+ - b^- = b_2 - 2b^-.$$

Thus,

$$\mathrm{index}(D_{S_+}) = b_1 - b^- - 1 = -\frac{1}{2}(\chi(S^4) - \tau(S^4)) = -\frac{1}{2}(2 - 0) = -1.$$

Now, we compute $\mathrm{index}(D)$. Using the Atiyah–Singer index theorem 2.3.3,

$$\begin{aligned}
\mathrm{index}(D) &= \mathrm{ch}(E)\hat{\mathcal{A}}(S^4)[S^4] && \text{(Atiyah–Singer index theorem)}\\
&= \mathrm{ch}(\mathrm{ad}(P))\mathrm{ch}(S_-)[S^4] && \text{(Remark 1.1.3 and Example 2.3.1)}\\
&= \Big(\dim(SU(2)) + \frac{1}{2}p_1(\mathrm{ad}(P)_\mathbb{R})x\Big)\Big(2 + \mathrm{index}(D_{S_+})x\Big)[S^4],
\end{aligned}$$

where in the last equality we used the Chern character formula from Remark 1.1.3

$$\mathrm{ch}(E) = \mathrm{rank}(E) + c_1(E) + \frac{1}{2}(c_1(E)^2 - 2c_2(E)),$$

and the facts that $c_1(\mathrm{ad}(P)) = 0$ (Remark 2.3.1) and $p_1(E_\mathbb{R}) = c_1(E)^2 - 2c_2(E)$ (Example 1.1.5). Recall that for a complex vector bundle E, $E_\mathbb{R}$ denotes the underlying real vector bundle of E. Additionally, x here is a generator of $H^4(S^4)$, for which $x[S^4] = 1$. Continuing our computation, we get

$$\begin{aligned}
\mathrm{index}(D) &= \Big(2\dim(SU(2)) + (p_1(\mathrm{ad}(P)_\mathbb{R}) + \dim(SU(2))\mathrm{index}(D_{S_+}))x\Big)[S^4]\\
&= p_1(\mathrm{ad}(P)_\mathbb{R}) + \dim(SU(2))\mathrm{index}(D_{S_+}) && \text{(since } x[S^4] = 1)\\
&= p_1(\mathrm{ad}(P)_\mathbb{R}) + (3)(-1)
\end{aligned}$$

Therefore,

$$\boxed{h^1 = \mathrm{index}(D) = p_1(\mathrm{ad}(P)_\mathbb{R}) - 3}.$$

This concludes step 1.

2.3.3 Steps 2 and 3: Local and Global Moduli Spaces

One can achieve step 2 using Banach space inverse and implicit function theorems to integrate the infinitesimal deformations of instantons and show that every element in $H^1_A(\mathrm{ad}(P))$ is defined by a 1-parameter family of instantons. So, we obtain local moduli spaces. In step 3, one must show that the local moduli spaces give local coordinates on the global moduli space. More detailed proofs for steps 2 and 3 can be found in section 6 of [8].

Thus, the dimension of the moduli space of instantons on a principal $SU(2)$-bundle P over S^4 is $p_1(\mathrm{ad}(P)_\mathbb{R}) - 3$.

Remark 2.3.2. The Atiyah–Hitchin–Singer theorem generalizes further to general compact semi-simple Lie groups G (see Theorem 6.1 in [8]). More precisely, the dimension of the moduli space of instantons on a principal G-bundle P over a compact self dual Riemannian 4-manifold X with positive scalar curvature is

$$p_1(\mathrm{ad}(P)_\mathbb{R}) - \frac{1}{2}\dim(G)(\chi - \tau),$$

where $p_1(\mathrm{ad}(P))$ is the first Pontryagin class of the adjoint bundle of P, χ is the Euler characteristic of X, and τ is the signature of X.

2.3.4 Restricting to the Atiyah–Hitchin–Singer theorem

In the special case where $G = SU(2)$ and $X = S^4$, the general Atiyah–Hitchin–Singer theorem restricts to the Atiyah–Hitchin–Singer theorem. In particular, the moduli space of instantons on a principal $SU(2)$-bundle P over S^4 has dimension $8k - 3$, where k is the *Pontryagin index* of P. Let's unpack what the Pontryagin index is and how this statement arises. Observe that

$$\mathrm{ch}(\mathrm{ad}(P)) = \dim(SU(2)) + \frac{1}{2}p_1(\mathrm{ad}(P)_\mathbb{R}) \implies p_1(\mathrm{ad}(P)_\mathbb{R}) = 2(\mathrm{ch}(\mathrm{ad}(P)) - 3).$$

In [8], Atiyah, Hitchin, and Singer identify $\mathrm{ad}(P)$ with the second symmetric power bundle S^2E of some rank-2 complex vector bundle E. The **Pontryagin index of** P is $k := -c_2(E)$ and $c_1(E) = 0$. They then state that $\mathrm{ch}(\mathrm{ad}(P)) = \mathrm{ch}(S^2E) = \mathrm{ch}(E)^2 - \mathrm{ch}(1)$. We know

$$\begin{aligned}\mathrm{ch}(E) &= \mathrm{rank}(E) + c_1(E) + \frac{1}{2}(c_1(E)^2 - 2c_2(E))\\ &= \mathrm{rank}(E) - c_2(E)\\ &= 2 + k\end{aligned}$$

Keeping terms to first order in k (really in the generator x of $H^4(S^4)$), $\mathrm{ch}(\mathrm{ad}(P)) = 4k + 4 - 1 = 4k + 3$. Thus,

$$p_1(\mathrm{ad}(P)_\mathbb{R}) = 2(4k + 3 - 3) = 8k.$$

Therefore, by Theorem 2.3.2, the moduli space of instantons on a principal $SU(2)$-bundle P over S^4 has dimension

$$p_1(\mathrm{ad}(P)_{\mathbb{R}}) - 3 = 8k - 3.$$

Note that we didn't show that such instantons and moduli spaces exist. Section 7 of [8] demonstrates their existence, called the t'Hooft solutions. However, the construction of the t'Hooft solutions depends on $5k+4$ parameters, which for $k > 2$ is less than the $8k-3$ parameters mentioned in the Atiyah–Hitchin–Singer theorem. Thus, the Atiyah–Hitchin–Singer theorem suggests the existence of new $SU(2)$-instantons via deformations of the t'Hooft solutions.

Bibliography

[1] Manfredo P. doCarmo. *Differential Forms and Applications*. Universitext. Springer Berlin, Heidelberg, Dec. 2012. ISBN: 9783642579516. DOI: https://doi.org/10.1007/978-3-642-57951-6.

[2] Loring W. Tu. *Differential Geometry. Connections, Curvature, and Characteristic Classes*. Graduate Texts in Mathematics. Springer Cham, June 2017. ISBN: 9783319550848. DOI: https://doi.org/10.1007/978-3-319-55084-8.

[3] Mark McLean. *Pontryagin Classes*. URL: https://www.math.stonybrook.edu/~markmclean/MAT566/lecture13.pdf.

[4] R. Mosseri. "Two-Qubit and Three-Qubit Geometry and Hopf Fibrations". In: *Topology in Condensed Matter*. Ed. by Michail Ilych Monastyrsky. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 187–203. ISBN: 978-3-540-31264-2. DOI: 10.1007/3-540-31264-1_9. URL: https://doi.org/10.1007/3-540-31264-1_9.

[5] Clifford Henry Taubes. "Flat connections and holonomy". In: *Differential Geometry: Bundles, Connections, Metrics and Curvature*. Oxford University Press, Oct. 2011. ISBN: 9780199605880. DOI: 10.1093/acprof:oso/9780199605880.003.0013. eprint: https://academic.oup.com/book/0/chapter/161658290/chapter-ag-pdf/44916550/book_12200_section_161658290.ag.pdf. URL: https://doi.org/10.1093/acprof:oso/9780199605880.003.0013.

[6] R. O. Wells. *Differential Analysis on Complex Manifolds*. 2nd ed. Graduate Texts in Mathematics. Springer New York, NY, Apr. 2013. ISBN: 9781475739466. DOI: https://doi.org/10.1007/978-1-4757-3946-6.

[7] M. F. Atiyah, Nigel J. Hitchin, and I. M. Singer. "Deformations of Instantons". In: *Proc. Nat. Acad. Sci.* 74 (1977), pp. 2662–2663. DOI: 10.1073/pnas.74.7.2662.

[8] M. F. Atiyah, N. J. Hitchin, and I. M. Singer. "Self-Duality in Four-Dimensional Riemannian Geometry". In: *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 362.1711 (1978), pp. 425–461. ISSN: 00804630. URL: http://www.jstor.org/stable/79638 (visited on 03/26/2023).

[9] John Roe. *Elliptic operators, topology and asymptotic methods*. Pitman research notes in mathematics series. Longman, 1998. ISBN: 0582325021.

[10] David Bleecker and Bernhelm Booß-Bavnek. *Index Theory with Applications to Mathematics and Physics*. International Press, 2013. ISBN: 9781571462640.