Equations of Gauge Theory

Karen Uhlenbeck

Notes by Laura Fredrickson

These notes are based on a series of lectures Professor Karen Uhlenbeck gave in 2012 at Temple University in Philadelphia. In a series of three lectures, Karen gave a history of the equations of gauge theory, from the Yang-Mills equations to the Kapustin-Witten equations, with a particular eye towards the relationship between the physics and mathematics communities. The notes are organized into three chapters, and are oriented towards the future of the Kapustin-Witten equations.

Chapter 1 discusses Yang-Mills equations, from their origins in the physics literature as a generalization of Maxwell's equations for electromagnetism to their current application in four-manifolds in Donaldson Theory and Floer Theory by providing new topological invariants. The first chapter discusses a number of real four-dimensional gauge theories, but focuses on the particularly interesting self-dual Yang-Mills equations. The Kapustin-Witten equations are a "complexification" (in some sense) of the self-dual Yang-Mills equations. As such, we are interested in which results about the self-dual Yang-Mills equations have complex counterparts that are true for the Kapustin-Witten equations.

Chapter 2 discusses Hitchin's equations, a set of equations over a two-dimensional Riemann surface, obtained by dimensional reduction of the self-dual Yang-Mills equations. While the self-dual Yang-Mills equations are a real gauge theory, featuring the real curvature of the real connection and real gauge group, Hitchin's equations are a complex gauge theory. They are the equations for a flat complex connection with an extra "zero moment map" condition. The Kapustin-Witten equations are also a complex gauge theory with a zero moment map condition, so the Kapustin-Witten equations are (in some sense) a particular higher-dimensional generalization of Hitchin's equations. In the second chapter, we discuss this mysterious zero moment map condition and the relation between Hitchin's equations and representation theory.

Finally, Chapter 3 is about the Kapustin-Witten equations. All the material in the previous two chapters comes together in this chapter. We prove results for the Kapustin-Witten equations that are the complex counterparts of results for the self-dual Yang-Mills equations or the higher-dimensional analog of results for Hitchin's equations. This includes some analytic estimates for solutions of Kapustin-Witten equations that would be the foundation for any topological results derived from the Kapustin-Witten equations.

More than a reference for the equations of gauge, these notes are a story about gauge theory: how obscure developments launched some of today's prominent areas of research, how strange and unnatural some of the developments in gauge theory must have appeared at the time, and why mathematicians trust the almost-magical intuition of the physicists. Karen peppered the lectures with stories of her own. To improve the readability, these storied parts are offset from the mathematical text in italics with wider margins.

These lectures were given at Temple University in Philadelphia, PA between February 7-9, 2012 as part of their Emil Grosswald lecture series.

Contents

Introduction

The story of the Yang-Mills equations exemplifies the changing relationship between mathematics and physics. Before the 1960s there was little collaboration between mathematicians and physicists. Now there is much collaboration and cross-fertilization between mathematics and physics, and more generally between math and the other sciences. The entire position of mathematics within the larger scientific community has changed.

In 1954 the Yang-Mills equations were published in a notable physics journal Physical Review, but the equations were not notable at the time either among physicists or mathematicians. Not much progress was made on the equations for over a decade because they were so difficult. Outside of the physics community, in 1963, mathematicians Atiyah and Singer published their Atiyah-Singer Index theorem. This result was spectacularly recognized. Using the Index Theorem, mathematicians were able to compute the dimension of the moduli space of solutions of these self-dual Yang-Mills equations arising from physics. This was one of the first major results in the study of the Yang-Mills equations.

In 1983, Donaldson used the Yang-Mills equations arising from physics in a purely mathematical context. Donaldson got new topological invariants for four-manifolds by studying the moduli space of solutions of the self-dual Yang-Mills equations over those four-manifolds. Donaldson's theorem and his method of proof opened up entirely new mathematical vistas.

> *So why are we mathematicians here? The physicists wrote down the Yang-Mills equations in 1954, and these equations have been employed in a lot of new mathematics, launching new fields like four-manifolds and particularly Donaldson Theory and Floer theory. The physicists in 1954 didn't even say that these equations were important!*
>
> *Now physicists such as Ed Witten and Anton Kapustin are writing down equations and they say they are important. So do we listen to the physicists or not? I'm giving this series of lectures with the thesis that we should.*

Timeline of Major Developments in Gauge Theory

- 1954 Yang-Mills equations appear in physics journal

- 1963 Atiyah-Singer Index Theorem

- 1977 Atiyah-Hitchin-Singer: deformations of instantons

- 1978 Atiyah-Drinfeld-Hitchin-Manin give a a description of solutions, construction of instantons gave a description of the solutions of Yang-Mills equations.

- 1983 Donaldson uses Yang-Mills in 4-manifold topology

- 1983 Atiyah-Bott: Yang-Mills over Riemann surface

- 1983 Hitchin: construction of monopole

- 1986 Connection with algebraic topology

- 1987 Hitchin self-dual equations on a Riemann surface

- 1989 Floer: "Witten's complex and ∞-dimensional Morse theory"

- 1994-1995 Taubes: Seiberg-Witten theory

- 2006 Preprint on Kapustin-Witten equations

- 2007 Kapustin-Witten equations published

Chapter 1

The Yang-Mills equations

1.1 Maxwell's Equations

The Yang-Mills equations are a glorification of Maxwell's equations, and consequently we turn our attention first to these. Maxwell's equations describe the time-evolution of electric and magnetic fields on $\mathbb{R}^3 \times \mathbb{R}$, where $\mathbb{R}^3$ is the space and $\mathbb{R}$ is the time. Electric charges and currents generate such electric and magnetic fields.

Let $\mathbf{E}$ be an electric field and $\mathbf{B}'$ be a magnetic field. In coordinates $\{dx^1, dx^2, dx^3\}$ on $\mathbb{R}^3$ these are given by

$$
\begin{aligned}
\mathbf{E} &= E_j \, dx^j \\
\mathbf{B}' &= B_j \, dx^j
\end{aligned}
$$

Here $\mathbf{E}, \mathbf{B}' \in \Omega^1(\mathbb{R}^3)$, but it is preferable to think of the magnetic field as a two-form $\mathbf{B} = \star\mathbf{B}' \in \Omega^2(\mathbb{R}^3)$. [1]

---

[1] The operator $\star$ indicates the point-wise linear *Hodge star operator*, dependent on the usual Riemannian metric on $\mathbb{R}^n$ that gives an isomorphism

$$
\star : \Omega^j \to \Omega^{n-j}.
$$

For any orthonormal basis $\{dx^1, \dots dx^n\}$, the Hodge star operator is defined by

$$
\star(dx^1 \wedge dx^2 \wedge \cdots \wedge dx^j) = dx^{j+1} \wedge \cdots \wedge dx^{n-1} \wedge dx^n
$$

Note a few properties of the Hodge star operator:

- $\star^2 = (-1)^{(j)(n-j)} : \Omega^j(\mathbb{R}^n) \to \Omega^j(\mathbb{R}^n)$.

- $\star$ is bijective

- The Hodge star defines an inner product on $\Omega^\bullet(M)$, the vector space of forms on a Riemannian manifold, M:

$$
\langle \alpha, \beta \rangle = \int_M \alpha \wedge \star\beta \tag{1.1}
$$

for $\alpha, \beta \in \Omega^\bullet(M, \mathbb{R})$. Note that

$$
\Omega^\bullet(M) = \oplus_{k=0}^{\dim M} \Omega^k(M)
$$

is a graded algebra, and $\langle \alpha, \beta \rangle = 0$ unless α and β are in the same grading.

There are two parts of **Maxwell's equations**: a homogeneous and non-homogeneous part. The homogeneous part is

$$
\begin{aligned}
d\mathbf{B} &= 0 & \in \Omega^3 \\
d\mathbf{E} &= \star\frac{\partial \mathbf{B}}{\partial t} & \in \Omega^2
\end{aligned}
\tag{1.2}
$$

and the inhomogeneous part

$$
\begin{aligned}
d^*\mathbf{E} &= \star d \star \mathbf{E} &= \frac{\rho}{\varepsilon_o} \in \Omega^0 \\
d^*\mathbf{B} &= \star d \star \mathbf{B} &= -\mu_o\varepsilon_o\frac{\partial \mathbf{E}}{\partial t} - \mu_0 J \in \Omega^1.
\end{aligned}
\tag{1.3}
$$

J and ρ are functions related to electric currents and charges, respectively; ε_o, μ_o are universal constants related to the speed of light by $c = \frac{1}{\sqrt{\mu_o\varepsilon_o}}$.

When $J = 0$ and $\rho = 0$, as is the case in a vacuum, Maxwell's equations simplify. The homogeneous part is

$$
\begin{aligned}
d\mathbf{B} &= 0 \\
d\mathbf{E} &= \star\frac{\partial \mathbf{B}}{\partial t}
\end{aligned}
$$

and the inhomogeneous part is

$$
\begin{aligned}
d^*\mathbf{E} &= 0 \\
d^*\mathbf{B} &= -\frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}.
\end{aligned}
$$

In this case, the electric and magnetic fields satisfy the wave equation

$$
\Delta\mathbf{E} = \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2}
$$

$$
\Delta\mathbf{B} = \frac{1}{c^2}\frac{\partial^2 \mathbf{B}}{\partial t^2}
$$

where $\Delta = dd^* + d^*d$. The coefficient $\frac{1}{c^2}$ appears in the above equations, because in a vacuum, waves propagate at the speed of light, c.

1.1.1   Invariance under Lorentz transformations

What are are the symmetries of Maxwell's equations? In the formulation of Maxwell's equations in Eqs. 1.2- 1.3, it is apparent that Maxwell's equations on $\mathbb{R}^3$ are invariant under the group of isomorphisms of $\mathbb{R}^3$, i.e. "Euclidean transformations." This group is generated by translations and rotations in space of $\mathbb{R}^3$. This section answers the question "Are Maxwell's

---

- The adjoint d^* of the usual exterior differential operator d is:

$$
d^* = \star \circ d \circ \star : \Omega^j \to \Omega^{j-1}.
$$

equations invariant under the larger group of Lorentz transformations?" which mix time and space. Since Lorentzian geometry is the geometry of general relativity, this question is synonymous with "Are Maxwell's equations compatible with general relativity?"

Maxwell's equations are indeed invariant under Lorentz transformations. However, it is not at all apparent from Maxwell's equations, as written in Eqs. 1.2-1.3. Here, we rewrite Maxwell's equations as equations on Minkowski space $\mathbb{R}^{3,1}$, the Lorentzian manifold with metric

$$g = -(cdt)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2.$$

The electric field $\mathbf{E} \in \Omega^1(\mathbb{R}^3)$ and the magnetic field $\mathbf{B} \in \Omega^2(\mathbb{R}^3)$ are combined into a two-form $\mathbf{F}$ known as the *electromagnetic field*:

$$\mathbf{F} = \mathbf{B} + \mathbf{E} \wedge dt \in \Omega^2(\mathbb{R}^{3,1}). \tag{1.4}$$

In coordinates $\{x^1, x^2, x^3, t = x^0\}$, the electromagnetic field is expressed

$$\mathbf{F} = \sum_{0 \leq j < k \leq 3} F_{jk} \, dx^j \wedge dx^k$$

where

$$F_{jk} = \begin{cases} E_k & \text{if } j = 0 \\ B_{jk} & \text{if } j \neq 0 \end{cases}$$

A simple computation gives **Maxwell's equations in 4-space**:

$$\begin{aligned} d\mathbf{F} &= 0 \quad \text{(homogeneous)} \\ d^*\mathbf{F} &= \mathcal{J} \quad \text{(inhomogeneous)} \end{aligned} \tag{1.5}$$

where $\mathcal{J} = (\mu_o J, \frac{\rho}{\varepsilon_o})$ is the electrostatic current. In this formulation, it is apparent that Maxwell's equations are invariant under Lorentz transformations because the Lorentzian transformations are precisely the isometries of Minkowski space $\mathbb{R}^{3,1}$.

1.1.2 Maxwell's equations as an $U(1)$ gauge theory

In the previous section, Maxwell's equations were equations for the electromagnetic field $\mathbf{F}$. Here we recast Maxwell's equations as equations for a quantity A, the electromagnetic potential. It is often more convenient to work with potentials than forces in physics because potentials are scalars and forces are vectors. As a scalar, the electromagnetic potential is more convenient here as well. The *electromagnetic potential* is defined to be a one-form $\tilde{A} \in \Omega^1(\mathbb{R}^{3,1})$ such that $\mathbf{F} = d\tilde{A}$. Locally, $\tilde{A}$ exists, though it is only unique up to cohomology class.

The "reality" of the electromagnetic potential $\tilde{A}$

An important philosophical question in physics is then "To what extent does the electromagnetic potential have reality?" From the discussion in the previous paragraph, certainly any physical meaning of $\tilde{A}$ can only depend on the cohomology class $[\tilde{A}]$. But is the electromagnetic potential just a convenient mathematical construct? Or are there physical consequences to the electromagnetic potential that do not arise from the electromagnetic field?

The Aharonov-Bohm effect (1959) demonstrates that electromagnetic potential *does* have physical meaning. The experimental set up is as pictured in Figure 1.1. An electric current
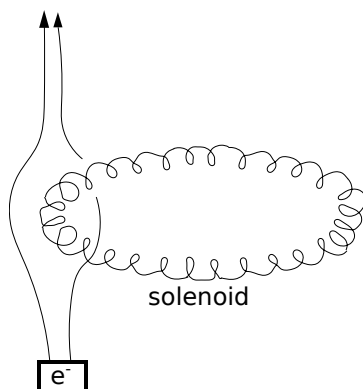

Figure 1.1: Aharonov-Bohm experiment

runs though a solenoid S producing a magnetic field outside in $\mathbb{R}^3 - S$. This electromagnetic field $\mathbf{F}$ vanishes on $\mathbb{R}^3 - S$. A beam of electrically-charged electrons is emitted, and them split into two topologically distinct beams through the magnetic field on $\mathbb{R}^3 - S$, as shown. One beam ψ_1 passes through the solenoid, and the other ψ_2 does not. The phases of the the two beams ψ_1 and ψ_2 are then compared via interference patterns.

And the phases are different! The phases of the beams are affected by the electromagnetic field, even though that field vanishes where the beams pass. Hence the electromagnetic potential is the truly fundamental quantity.

Gauge theoretic formulation

A theme throughout these lectures is that there were some developments in gauge theory that were "non-linear." Rather than proceeding forward in some logical way, there were developments that seemed to be going oddly sideways. In one such critical sideways development, Maxwell's equations were reinterpreted as an abelian $G = U(1)$ gauge theory on $\mathbb{R}^{3,1}$ in terms of the electromagnetic potential $\tilde{A}$. The electromagnetic potential is viewed as a connection on a principle $U(1)$-bundle over $\mathbb{R}^{3,1}$ and the electromagnetic field $\mathbf{F}$ is the curvature of that connection. This development will prove to be critical for the generalization of Yang-Mills that replaces the abelian group $G = U(1)$ with a non-abelian group.

Let P be a principal G-bundle over $\mathbb{R}^{3,1}$. When P is restricted to an open neighborhood $\mathcal{O}$ of $\mathbb{R}^{3,1}$, it is locally isomorphic to $\mathcal{O} \times G$. Locally, a connection $D_A = d + A$ on $P|_{\mathcal{O}}$ is

given by A, a $\mathfrak{g}$-valued one-form on $\mathbb{R}^{3,1}$. In this case, that means A is locally a $i\mathbb{R}$-valued one-form. This connection one-form A and the previously defined electromagnetic potential $\tilde{A}$ are related by

$$A = i\tilde{A},$$

hence the electromagnetic potential $\tilde{A}$ gives a connection D_A, defined in coordinates $\{t = x_0, x_1, x_2, x_3\}$ by,

$$D_A = \sum_{k=0}^{3} \left(\frac{\partial}{\partial x^k} + A_k \right) dx^k$$

The curvature of the connection is given by

$$F_A = (D_A)^2 = \left(\frac{\partial}{\partial x^j} A_k - \frac{\partial}{\partial x^k} A_j + \frac{1}{2}[A_j, A_k] \right) dx^j \wedge dx^k$$

but in this case, since $U(1)$ is abelian, the Lie bracket is trivial, and

$$i\mathbf{F} = F_A = dA$$

i.e., the curvature of the connection is precisely the electromagnetic field. Consequently, as a gauge theory, **Maxwell's equations** are simply:

$$
\begin{aligned}
D_A F_A &= d^2 A = & 0 \\
D_A^* F_A &= d^*dA = & \mathcal{J}.
\end{aligned}
\tag{1.6}
$$

Note that the homogeneous equation is satisfied by any connection since $d^2 = 0$. The differential operator d^*d looks similar to the Laplacian on forms, $\Delta = d^*d + dd^*$, the quintessential elliptic operator. We point this out here because in the different gauge theories in subsequent sections, this ellipticity will play a crucial role in the analysis.

1.2 Yang-Mills Equations

1.2.1 Origins

The innovation of Yang-Mills was to replace the abelian group $G = U(1)$ appearing in Maxwell's equations with a non-abelian group, consequently obtaining a non-abelian gauge theory, known as Yang-Mills gauge theory. In this sense, the Yang-Mills equations are a glorification of Maxwell's equations.

Yang-Mills replaced the Lie algebra associated to the abelian Lie group $U(1)$ with the Lie algebra associated to the compact non-abelian Lie group $SU(2)$.[2] The Yang-Mills equations on $\mathbb{R}^{3,1}$, like Maxwell's equations, are given in terms of a connection D_A. Locally, the connection $D_A = d + A$ is given by an $\mathfrak{su}(2)$-valued one-form

$$A = A_j dx^j \quad A_j \in \mathfrak{su}(2).$$

---

[2]Here it is not the particular choice of Lie group $SU(2)$ that matters, but rather the shift from an abelian to non-abelian Lie group. It is in this that all the difficulties immediately arise. One could instead take the non-abelian Lie group E_8 instead, for example.

The curvature F_A is locally a $\mathfrak{su}(2)$-valued two-form expressed by

$$F_A = (D_A)^2 = dA + \frac{1}{2}[A, A]^3.$$

The equations are

$$\begin{aligned} D_A F_A &= 0 \\ D_A^* F_A &= \mathcal{J}. \end{aligned}$$

(When $G = U(1)$, these equations are just Maxwell's equations. These same equations appeared in Eq. 1.6.) The homogeneous equation $D_A F_A = 0$ is known as *Bianchi identity*, and is true for all connections. The inhomogeneous part is

$$D_A^* F_A = \mathcal{J}$$

for $\mathcal{J} \in \Omega(\mathbb{R}^{3,1}, \mathfrak{su}(2))$. We will take $\mathcal{J} = 0$ for simplicity. The Yang-Mills equations are invariant under the infinite-dimensional group of gauge transformations

$$\mathcal{G} = \{g : M \to G\},$$

which acts on a connection $D_A = d + A$ by conjugation

$$g \cdot (d + A) = g(d + A)g^{-1} = -dgg^{-1} + gAg^{-1}.$$

Modulo these gauge transformations, the Yang-Mills equations are elliptic!

The **Yang-Mills equations**

$$\begin{aligned} D_A F_A &= 0 \\ D_A^* F_A &= 0. \end{aligned} \tag{1.8}$$

can be defined for a connection on a principle G-bundle, P, over *any* smooth Riemannian 4-manifold M.

How did the Yang-Mills equations go from being equations on the non-compact Lorentzian manifold $\mathbb{R}^{3,1}$ to equations on compact Riemannian manifolds? This is yet another development in gauge theory that was not "linear." The base space changed from the Lorentzian manifold $\mathbb{R}^{3,1}$ to the Riemannian manifold $\mathbb{R}^4$ via Wick rotation. However, there was little immediate advantage in this.

---

[3]Here $[\cdot, \cdot]$ refers to the usual extension of the Lie bracket on $\mathfrak{g} = \mathfrak{su}(2)$ to $\Omega^1(\mathcal{O}, \mathfrak{g})$ defined by taking the wedge of the basis 1-forms and the brackets of the $\mathfrak{g}$-valued coefficients. In coordinates,

$$[A_i dx^i, B_j dx^j] = [A_i, B_j] dx^i \wedge dx^j \tag{1.7}$$

1.2.2 Variational Formulation & self-dual Yang-Mills equations

The Yang-Mills equations are the Euler-Lagrange equations for the
Yang-Mills functional

$$\mathfrak{Ym}(D_A) = \int_M |F_A|^2 \mathrm{dvol} := \int_M \mathrm{tr}(F_A \wedge \star F_A^*) d\mathrm{vol} \tag{1.9}$$

The inner product appearing in the above equation is described in footnote [4]. The proof
that the Yang-Mills equations are the Euler-Lagrange equations is, similarly in footnote

There is a special class of solutions to the Yang-Mills equations that will be our focus later
in this chapter. These are the connections which satisfy the **(anti-)self-dual Yang-Mills
equations**

$$F_A = \pm \star F_A. \tag{1.10}$$

Any solution of the (anti-)self-dual Yang-Mills equations is also a solution of the Yang-Mills
equations. But more than that, solutions of the (anti-) self-dual Yang-Mills equations are
absolute minima of the Yang-Mills functional. The proof of this fact is delayed until Chapter
3 Proposition 3.1.1 because there is a similar statement that is true of the Kapustin-Witten
equations.

1.2.3 The analysis of the Yang-Mills equations

From the time the Yang-Mills equations appeared in a physics journal in 1954, a handful of
physicists were interested in studying single solutions of the Yang-Mills equations (known as
instantons) on the Lorentzian manifold $\mathbb{R}^{3,1}$, as well as the space of such solutions. However,
the equations were difficult to work with, and little progress was made for over ten years.

> *The next major development in the story of the so-far-only-physicist's Yang-Mills
> equations was one by mathematicians! The Atiyah-Singer Index Theorem said
> something about the dimension of the space of solutions to the Yang-Mills equa-
> tions. The mathematicians had something to say to the physicists, and it is right
> at this point that the relation between the mathematics and physics community
> began to change. It is here that the cooperation and cross-fertilization really be-
> gan.*
>
> *I don't know how to describe the tremendous difference this made. The entire
> position that mathematics held within the scientific community changed. The
> mathematicians actually had something to say!*

---

[4]The inner product on $\Omega^\bullet(M, P \times_G \mathfrak{g})$ comes from the invariant inner product on $\mathfrak{g}$ defined by

$$< A, B >_{\mathfrak{su}(n)} = tr(AB^*)$$

and the inner product on $\Omega^\bullet(M)$ defined via the Hodge star

$$< \alpha, \beta >_{\Omega^\bullet(M)} = \int_M \alpha \wedge \star \beta.$$

Of all the properties of the moduli space of solutions, the dimension of the moduli space is one of the easiest properties. However, determining even this easiest property of the moduli space of self-dual Yang-Mills equations requires the heavy machinery of the Atiyah-Singer index theorem and technical implicit-function-theorem-type arguments. After discussing the Atiyah-Singer Index Theorem in Section 1.2.3, we turn in Section 1.2.3 to the tools available for proving other properties about the moduli space of solutions of the self-dual Yang-Mills equations.

The Atiyah-Singer Index Theorem

The *Atiyah-Singer Index Theorem*(1963) is a remarkable theorem about elliptic differential operators on compact manifolds, relating the analytical index of the operator (which is turn is related to the dimension of the space of solutions) to the topological index (defined purely by topological data.)

We briefly rehearse the modifications of the Yang-Mills equations we described in the previous sections, since they are essential for the application of the Atiyah-Singer Index Theorem to the Yang-Mills equations. First, the Yang-Mills equations became equations on $\mathbb{R}^4$ rather than $\mathbb{R}^{3,1}$. (None of the analysis works right in $\mathbb{R}^{3,1}$!) Second, we replaced the non-compact base manifold $\mathbb{R}^4$ with the compact space $S^4 = \mathbb{R}^4 \cup \{\infty\}$, or, in fact, any compact 4-manifold M^4. The last, and perhaps strangest, modification is the switch from considering the full Yang-Mills equations to the self-dual Yang-Mills equations. This modification is strange because the full Yang-Mills equations have some nice properties: they are already elliptic, and, unlike the self-dual Yang-Mills equations, the full Yang-Mills equations are the Euler-Lagrange equations for a variational problem.

The Atiyah-Singer Index Theorem will tell us something about the hypothetical dimension of the moduli space of solutions $\mathcal{M}$ to the self-dual Yang-Mills equations

$$\mathcal{M} = \{D_A \big| \star F_A = F_A\}/\mathcal{G}.$$

on a principle bundle P over M^4. The hypothetical dimension is computed from the topological index, which is related to the topology of the principle bundle P.

For $G = SU(2)$, principle $SU(2)$-bundles on M^4 are classified by their second Chern class $c_2(P) \in H^4(X, \mathbb{Z}) = \mathbb{Z}$, an integer topological invariant. For the sphere S^4, we can explicitly state what the second Chern class $c_2(P)$ is measuring. The bundle over S^4 is constructed from trivial bundles over two patches: $S^4 - \{\infty\}$ and B_∞, a ball around ∞; and a gluing map, possibly twisting the bundle, on the intersection. This gluing map is

$$\phi : S^4 - \{\infty\} \cap B_\infty \to SU(2).$$

The domain $S^4 - \{\infty\} \cap B_\infty$ deformation retracts to S^3 and the Lie group $SU(2)$ is topologically S^3, so the gluing map ϕ gives a map

$$\phi_{S^3} : S^3 \to S^3,$$

and the degree of this map is the exactly second Chern class $c_2(P)$ of the bundle P.

The Atiyah-Singer Index Theorem says the following about the dimension of the moduli space:

Theorem 1.2.1 *Let P be a principle $SU(2)$-bundle over M^4 and let $c_2(P)$ be the above topological invariant. Then, the moduli space of self-dual instantons has hypothetical dimension*

$$\dim(\mathcal{M}_k) = 8c_2(P) - 3 \tag{1.11}$$

For S^4, this hypothetical dimension is realized.

Tools for Analysis of Yang-Mills

We now turn our attention to the analysis of solutions of self-dual Yang-Mills equations in $\mathbb{R}^4$. What tools are available? Here, there is not just one way to do it!

- **Twistor theory** In $\mathbb{R}^4$, you have twistor theory which transforms questions about partial differential equations into questions in complex geometry. Here, the anti-self-dual condition on connections over $\mathbb{R}^4$ is turned into a question about holomorphic bundles on the twistor space $\mathbb{CP}^3$.

- **Algebraic geometry**

- **Non-linear PDE**

  What's interesting about the theory is that non-linear PDE that has been the most successful and the richest. Of all these approaches–algebra, algebraic geometry,– the non-linear PDE has been richest and most useful. Take this with a grain of salt because non-linear PDE is what I do! But I still think if you look at the historical record, this is justified.

1.2.4 Donaldson Theory as an application of the self-dual Yang-Mills equations

We now turn towards an application of the self-dual Yang-Mills equations in mathematics. In the process, we will be able to say more about the moduli space of instantons. Donaldson proved a remarkable theorem regarding the topology of 4-manifolds. He proved the existence of non-smoothable topological 4-manifolds using the topology of the self-dual Yang-Mills equation. He exploited the interaction between the topology of the underlying manifold M^4 and the topology of the resulting moduli space of self-dual connections $\mathcal{M}$ on a particular principle bundle over M. He used remarkable proof techniques that mathematicians hope can be extended to use in the complex Kapustin-Witten equations discussed in Chapter 3.

Developments in 4-manifold topology prior to Donaldson's Theorem

Before we turn our attention to Donaldson's theorem, we briefly describe the developments in four-manifold topology before Donaldson's theorem. Before 1980, people knew very little about 4-manifold topology and had few examples of 4-manifolds.

I was at the Institute for Advanced Study in Princeton in 1979-1980. Shang came and gave a talk about four manifolds, and he discussed what people knew about 4-manifold topology. I came away from the talk with the realization that people didn't know much about four manifolds. They didn't know many examples. They didn't know how to work with them. They didn't know much at all.

Now, shortly after, in 1981, Mike Freedman proved a very important theorem in 4-manifold topology about the connection between 4-manifold topology and the classification of non-degenerate quadratic forms (equivalently, non-degenerate symmetric bilinear forms) over the integers. These non-degenerate quadratic forms are quite rare! From any 4-manifold M^4, one naturally obtains a quadratic form known as the *associated bilinear form* from the integral second cohomology class $H^2(M^4, \mathbb{Z})$, as follows:

$$
\begin{aligned}
\mathcal{Q}: \quad & H^2(M^4, \mathbb{Z}) \times H^2(M^4, \mathbb{Z}) \longrightarrow \mathbb{Z} \\
\mathcal{Q}: \quad & \qquad ([\alpha], [\beta]) \qquad \mapsto \int_{M^4} \alpha \wedge \beta.
\end{aligned} \tag{1.12}
$$

Example Here are some examples of simply-connected 4-manifolds and their associated bilinear forms:

$$
\begin{array}{lll}
S^4 & H^2(S^4) = 0 & Q = 0 \\
\mathbb{CP}^2 & H^2(\mathbb{CP}^2) = \mathbb{Z} & Q = \begin{pmatrix} 1 \end{pmatrix} \\
S^2 \times S^2 & H^2(S^2 \times S^2) = \mathbb{Z}^2 & Q = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
K3 & H^2(K3) & Q = 2E_8 + 3Q_0
\end{array}
$$

Freedman proved the converse: there is a 4-manifold associated to any non-degenerate quadratic form.

Theorem 1.2.2 *(Freedman) For any non-degenerate bilinear form Q over $\mathbb{Z}$ there exists a simply-connected topological (but not necessarily smooth!) manifold whose second homology realizes this form, i.e. the associated bilinear form to M is Q.*

Remark Freedman did not prove uniqueness. In fact, there is exactly one or two such manifolds. The extra piece of data to uniquely determine the manifold up to homeomorphism is the Kirby-Siebenmann obstruction in $\mathbb{Z}/2$.

This drastically expanded the number of examples of 4-manifolds mathematicians knew!

Donaldson's Theorem

Building on the work of Freedman, in 1983 Simon Donaldson, a graduate student of Atiyah at the time, proved a remarkable theorem.

> *What's the advantage of being a grad student? Well, if you don't know it's an impossible theorem, you might actually prove it! Donaldson was messing around with all this impossible stuff, and, lo and behold, he proved something!*

Theorem 1.2.3 *(Donaldson) Let M^4 is a smooth, compact, connected, simply-connected 4-manifold, with the property that the associated bilinear form Q is positive-definite (i.e. $b^- = 0$). Then Q is diagonalizable over the integers to the identity matrix. In particular, if Q is also negative-definite (i.e. $b^+ = 0$), then $M = S^4$.*

Donaldson's theorem implies that many of the of *topological* 4-manifolds that Freedman produced had *no differentiable structure*.

> *This not only a startling theorem, given that people didn't know anything about four-dimensional topology at this point; but the method of proof is extremely astonishing and opened a whole new world!*

Donaldson's proof is based on the topology of the space of self-dual Yang-Mills connections on a principle $SU(2)$ bundle over M^4, a smooth, compact, simply-connected 4-manifold. Here we sketch the method of proof:

Sketch of Proof Let M^4 satisfy the hypothesis of the theorem with associated bilinear form Q and fix a smooth Riemannian structure on M^4. There is (up to isomorphism) a unique principle $SU(2)$-bundle P over M^4 with second Chern class $c_2(P)[M^4] = -1$.

Consider the moduli space $\mathcal{M}$ of self-dual connections on P. What does $\mathcal{M}$ look like? By the Atiyah-Singer Index Theorem, since $c_2(P) = -1$, $\dim \mathcal{M} = 5$. From the work of Atiyah, $\mathcal{M}$ is orientable. $\mathcal{M}$ is non-compact, but the non-compact part looks like $M^4 \times (0,1)$. The moduli space also has singularities that look like cones on $\mathbb{CP}^2$. These singularities occur at those self-dual connections which are reducible to $U(1)$-connections.

Consequently, the moduli space is an oriented cobordism from M^4 to $\bigsqcup_{n_P} \mathbb{CP}^2$, where n_P is the number of reducible connections. The signature $\sigma = (b^+, b^-)$, coming from the bilinear form Q, is an oriented cobordism invariant, so

$$(b^+, 0) = \sigma(M) = \sigma \left(\bigsqcup_{n_P} \mathbb{CP}^2 \right) = (n_P, 0).$$

It turns out that the reducible connections correspond to $\alpha \in H^2(M; \mathbb{Z})$ such that $Q(\alpha, \alpha) = 1$, and the cobordism shows that the number of these reducible connections is equal to the full rank of the bilinear form Q. Hence, Q is diagonalizable. ∎
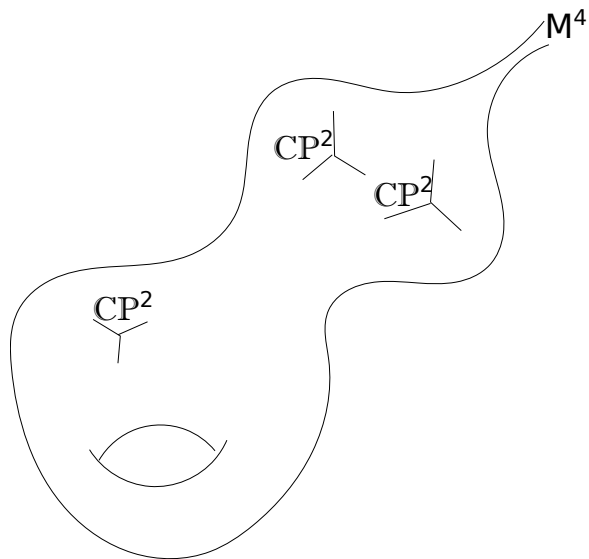
Figure 1.2: The moduli space of self-dual Yang-Mills connections

Ingredients in Donaldson's method of proof

Here, we list the ingredients in Donaldson's proof. *Why?* Later, we'll look at the Kapustin-Witten equations, a complex version of the real Yang-Mills equations, and an active area of research. It would be desirable to generalize the kind of mathematics that appears in Donaldson's proof to get results in this new complex setting.

Theory of principle bundles This is the basic underlying theory. At the time it wasn't very well known, but now it is a classical subject known by many.

Infinite dimensional topology : You need an infinite-dimensional version of Sard's theorem.

Finite dimensional topology : Need de Rham theory

Atiyah-Singer Index Theorem : This is necessary to compute dimension of $\mathcal{M}$. More than that, the Atiyah-Singer theory is crucial background because it introduced the crucial notion of Fredholm operator between function spaces, something Donaldson's method of proof relies on.

Non-compactness results : The manifold $\mathcal{M}$ is not compact. One needs to answer the question "Where does non-compactness come from?"

Orientation on $\mathcal{M}$ for cobordism.

non-linear PDE : As stated above, need the concept of a Fredholm operator. The Sobolev spaces are the right spaces to work in and we need some analytical results

- Need estimates
- Need genericity properties

- Need constructions using gluing procedures: Donaldson used *Taubes' Gluing Theorem*, a powerful analytic result, to construct solutions by taking a flat connection and gluing in singularities at a point. This is a non-linear generalization of how one can add two local solutions of a linear PDE to obtain a global solution. Here one patches together a solution to a non-linear equation in one region and a solution to a non-linear equation in another region into a global solution by perturbing the resulting solution using the Implicit Function theorem. This was a method of solving equations that was largely unexplored before gauge theory.

Besides the specific topics mentioned, the global analysis of the 1960s is at the basis of the entire framework. Given a non-linear elliptic operator on a underlying space, the solutions form manifold. In global analysis the the interaction between the topology of these solution manifolds and the topology of the underlying manifold was studied.

Donaldson Theory

Donaldson's theorem was the first result in an area now known as Donaldson Theory which gives topological constraints on 4-manifolds using the gauge theoretic self-dual Yang-Mills equations. Donaldson theory is difficult because the self-dual Yang-Mills are not easy to perturb in a gauge invariant way. In the next section, we introduce a set of equations that is easier to perturb and, similar to Donaldson Theory, gives topological constraints on 4-manifolds.

1.3 Seiberg-Witten Equations

In this section we briefly introduce the Seiberg-Witten equations, another set of equations that gives a lot of 4-dimensional topological invariants. The Seiberg-Witten equations are a technically easier set of equations to use than the self-dual Yang-Mills equations because they are an abelian gauge theory; and it is easier to make deformations in an abelian gauge theory!

1.3.1 Origins

In 1992, Ed Witten gave a talk at Harvard and Cliff Taubes was in the audience. After the talk Taubes walked up to talk with Witten. Witten told Taubes about some equations, the Seiberg-Witten equations, that he suspected that contained all the topological 4-manifold invariants that Donaldson theory was extracting from the self-dual Yang-Mills equations. Taubes went home and wrote his shortest paper–just over 15 pages!– on that.

So a physicist walks up to a you as a mathematician, and says "You should be doing this," then you go home and it works! And you have no idea why it works or why you should be doing it. We have 30 years of this pattern. Somehow the physicists know the interesting equations!

What are the Seiberg-Witten equations and why are they so magical? The magic of the Seiberg-Witten equations is that they are an *abelian* gauge theory with an exterior field ϕ vaguely referred to as a Higgs field. It is surprising that an abelian gauge theory could contain the information of the self-dual Yang-Mills equations, with non-abelian structure group $SU(2)$. However, one does not obtain such an interesting abelian gauge theory in the naive way of choosing an abelian gauge group. For example, the $G = U(1)$ Yang-Mills equations are just Maxwell's equations and one gets nothing except de Rham cohomology.

1.3.2 The Equations

The Seiberg-Witten equations are an abelian gauge theory with gauge group $G = U(1)$. The data consequently includes some principal $U(1)$-bundle with connection D_A with curvature F_A. The interest comes from the coupled exterior field ϕ, known as the Higgs field or spinor field. Coupling involves nothing more than writing an equation involving both the gauge field and the Higgs field.

Let M be a smooth oriented Riemannian 4-manifold with $\text{Spin}^{\mathbb{C}}$ structure and let P be a principle $\text{Spin}^{\mathbb{C}}(4)$ bundle. The data of the Seiberg-Witten equations is a connection A on a certain line bundle $\mathcal{L}$ and a section ϕ of a certain $\mathbb{C}^2$ bundle, known as the spinor field. Let ϕ be a spinor and A a $U(1)$-connection on $\mathcal{L}$. The **Seiberg-Witten equations** are:

$$
\begin{aligned}
F_A + \phi^T \phi &= 0 \\
\not{D}_A \phi &= 0
\end{aligned}
\tag{1.13}
$$

where $\not{D}_A$ is a Dirac operator and $\phi^T \phi$ is some particular sesquilinear map. A solution to the Seiberg-Witten equations is called a **monopole**. This is a general term for a solution of a first-order coupled equation. Consequently, solutions of the $\lambda = 0$ Bogomolny and Kapustin-Witten equations, which we will introduce in Chapter 2 and 3 respectively, are also called monopoles.

1.3.3 Useful Features of Equations

Because the Seiberg-Witten equations are an abelian gauge theory, we can globally fix the coordinates and take $d^* A = 0$ globally.

Moreover, the Seiberg-Witten equations are easy to perturb. The curvature F_A is Lie-algebra valued two-form, but when $G = U(1)$, because $\mathfrak{u}(1) = i\mathbb{R}$,

$$
F_A \in \Omega^2(M, \mathfrak{g}_P) \cong \sqrt{-1}\,\Omega^2(M, \mathbb{R}).
$$

Consequently F_A can be perturbed by adding any self-dual two-form ω. Adding a parameter λ, one obtains a family of perturbed equations:

$$
\begin{aligned}
F_A + \phi^T \phi + i\lambda\omega &= 0 \\
\not{D}_A \phi &= 0.
\end{aligned}
\tag{1.14}
$$

Letting $\lambda \to \infty$, one gets explicit type of solutions that concentrate in different ways.

1.3.4 Relation to Topology

Taubes exploited this method of perturbation in his papers. This process of using the Seiberg-Witten equations to get topological information was spectacularly successful. Still a great deal of 4-manifold topology is based on this.

> *How do topologists use these deep and beautiful analytical results? They first turn them into a combinatorial machine.*
>
> *Before I did gauge theory, I was involved in minimal surfaces. Back in the 1970s, minimal surfaces were used to prove results about 3-manifolds. Well, the first thing the topologists did was they replaced the beautiful theory of minimal surfaces with combinatorics. How did they do this? Well the three manifolds have geometric structure, so you concentrate this structure at the edges of your manifold.*
>
> *For 4-manifolds, I asked Peter Osratch, "Do you really use PDE anymore?" and he said, "No." All this PDE stuff has been replaced by combinatorial jazz.*

Though Witten suspected Seiberg-Witten theory had all the information of Donaldson theory, it turns out that Donaldson theory has more information. Both are still active areas of research. Generally, mathematicians are more interested in Seiberg-Witten theory because it is easier to work with. The physicists prefer the self-dual Yang-Mills equations to the Seiberg-Witten equations. Provided their intuition is correct–as it historically seems to be– there is still a lot of untapped topological information in the Yang-Mills equations.

Chapter 2

Hitchin's equations

In the previous chapter we saw that a lot of interesting equations have arisen from physics and string theory in particular. This, in turn, has generated a ton of mathematics. Our goal is to generate some more equations. We will do this is by dimensional reduction.

In this chapter we consider Hitchin's equations which Hitchin originally called "the self-dual equations on a Riemann surface" in his seminal 1987 paper. We say that Hitchin's equations are a "complex theory" because they are equations on a complex manifold. The Kapustin-Witten equations, which we will discuss in the third chapter, are also complex equations. Consequently, many aspects that we'll discuss for Hitchin's equations, like the zero moment map condition and the relation between solutions of Hitchin's equations and representation theory, will appear again in the third chapter.

2.1 Origin of Hitchin's equations

Hitchin's equations are the self-dual equations on a Riemann surface. They are obtained by dimensionally reducing the self-dual Yang-Mills equations $F_{\tilde{A}} = \star F_{\tilde{A}}$ on $\mathbb{R}^4$ to $\mathbb{R}^2$. After introducing complex coordinates and proving the conformal invariance of the resulting equation on $\mathbb{R}^2$, the Hitchin's equations can be defined on a Riemann surface. Here we dimensionally reduce in two steps: first, from $\mathbb{R}^4$ to $\mathbb{R}^3$ and obtain the Bogomolny equation; then, dimensionally reduce again to $\mathbb{R}^2$ to obtain Hitchin's equations on $\mathbb{R}^2$. We can then write Hitchin's equations for a Riemann surface. Note that this dimensional reduction is not the naive dimensional reduction where one simply throws away a dimension. This naive dimensional reduction to 2 or 3-dimensions doesn't lead to any interesting equations–only that $F_A = 0$. In contrast, in the Hitchin's equations, though we end up with a two-dimensional theory by dimensional reduction, we are still studying the self-dual Yang-Mills equations in four dimensions.

2.1.1 Bogomolny Equations from dimensional deduction to $\mathbb{R}^3$

Let $\tilde{A} \in \Omega^1(\mathbb{R}^4, \mathfrak{su}(2))$. In local coordinates,

$$\tilde{A} = \tilde{A}_1 dx^1 + \tilde{A}_2 dx^2 + \tilde{A}_3 dx^3 + \tilde{A}_4 dx^4$$

for $\tilde{A}_i \in \Gamma(\mathbb{R}^4, \mathfrak{su}(2))$. First, we assume that all fields $\tilde{A}_i$ are independent of x^4 i.e. $A_i = A_i(x^1, x^2, x^3)$. Then we have the following fields.

$$\frac{\partial}{\partial x^j} + A_j = \frac{\partial}{\partial x^j} + \tilde{A}_j \quad j = 1, 2, 3$$
$$\phi = \tilde{A}_4 \quad \text{Higgs field.}$$

The self-dual Yang-Mills equations on $\tilde{A} \in \Omega^1(\mathbb{R}^4, \mathfrak{su}(2))$ descend to equations on $A \in \Omega^1(\mathbb{R}^3, \mathfrak{su}(2))$ and we obtain:

$$F_A = \star D_A \phi \in \Omega^2(\mathbb{R}^3, \text{End}(\mathfrak{su}(2))) \tag{2.1}$$

The resulting equation $F_A = \star D_A \phi$ is known as the **Bogomolny equation**. It is derived from the self-dual Yang-Mills equation in footnote [1] to illustrate the process of dimensional reduction. However, this dimensional reduction is not the origin of the Bogomolny equation. The Bogomolny equation pre-dates the Yang-Mills equations and interestingly, though it can be obtained by dimensionally reducing the SDYM equations in quantum theory, its origin is actually cosmological.

Solutions (A, ϕ) of the Bogomolny equation (which also satisfy an appropriate boundary condition) are called *BPS monopoles* after Bogomolny-Prasad-Sommerfield.

The Bogomolny equations satisfy the Euler-Lagrange equations for a certain action integral $\mathcal{A}_\lambda$ with choice of a coupling constant λ:

$$\mathcal{A}_\lambda(D_A, \phi) := \int_{\mathbb{R}^3} |F_A|^2 + |D_A\phi|^2 + \lambda(1 - |\phi|^2)^2 dx^1 dx^2 dx^3. \tag{2.2}$$

The $\lambda = 0$ Euler-Lagrange equations are second-order equations

$$D_A^* F_A = [D_A\phi, \phi] \tag{2.3}$$
$$D_A^* D_A\phi = 0,$$

known as the the **BPST (Belavin-Polyakov-Schwarz-Tyupkin) equations**. The solutions of the Bogomolny equations are solutions of the $\lambda = 0$ BPST equations. Note that the Bogomolny equations are first-order equations, while the BPST equations are (harder) second-order equations. *What are the solutions to the Bogomolny equations and how do we find them?* As usual, we look for a solution exploiting symmetry (including gauge symmetry). In this way, Cliff Taubes reduced the Bogomolny PDE to an ODE and found a single monopole solution. The equations for the monopole are completely integrable, and as a result, there are many beautiful descriptions of the solution.

We briefly discuss two questions relating to the physics behind the above equations.

Why the particular form of the action $\mathcal{A}_\lambda$? The first two terms, as first derivative terms, are obvious to someone who has worked with the variational form of Laplace's equation. The more mysterious term is $(1 - |\phi|^2)^2$. However, the particular form of this equation is not related to any deep physics. The term $(1 - |\phi|^2)^2$ is a quartic in $|\phi|$, and as such, is the simplest non-linear term we could add. We could have instead taken a more complicated third term involving $|\phi|^6$.

How big is λ for cosmologists? The action $\mathcal{A}_\lambda(D_A, \phi)$ has some cosmological meaning, and so it makes sense to ask what λ is to cosmologists. Cosmologists are looking in universe for a very very very small but non-zero constant λ. In other words, the solutions of the Bogomolny equation (a dimensional reduction of the self-dual Yang-Mills equations) are solutions of the $\lambda = 0$, but these are not are not actually what the cosmologist are actually looking for in the universe, though they are very close to it.

2.1.2 Hitchin's equations from dimensional reduction to $\mathbb{R}^2$

In the previous section, we saw that dimensional reduction from $\mathbb{R}^4$ to $\mathbb{R}^3$ produced the interesting Bogomolny equations. Let's keep going and dimensionally reduce to $\mathbb{R}^2$!

Again, we started with $\tilde{A}$ our full connection on $\mathbb{R}^4$.

$$\tilde{A} = \sum_{i=1}^{4} \tilde{A}_i dx^i.$$

In the previous section, to dimensionally reduce to $\mathbb{R}^3$, we assumed that everything was invariant in the x^4 coordinate, Now, to dimensionally reduce to $\mathbb{R}^2$, we assume everything is invariant in both x^3 and x^4. Now, in addition to a connection A, we also get two Higgs fields, ϕ_1 and ϕ_2:

$$\begin{aligned}
\frac{\partial}{\partial x^j} + A_j &= \frac{\partial}{\partial x^j} + \tilde{A}_j \quad j = 1, 2 \\
\phi_1 &= A_3 \\
\phi_2 &= A_4
\end{aligned}$$

We package the two fields, ϕ_1 and ϕ_2, into a single Higgs field ϕ, and obtain the following objects:

$$\begin{aligned}
A &= A_1 dx^1 + A_2 dx^2 \quad \in \Omega^1(\mathbb{R}^2, \mathfrak{su}(2)) \\
\phi &= \phi_1 dx^1 + \phi_2 dx^2 \quad \in \Omega^1(\mathbb{R}^2, \mathfrak{su}(2))
\end{aligned}$$

2.2 Hitchin's Equations

The dimensional reduction of $F_{\tilde{A}} = \star F_{\tilde{A}}$ produces **Hitchin's equations** on a $\mathbb{R}^2$:

$$\begin{aligned}
F_A - \frac{1}{2}[\phi, \phi] &= 0 & (2.4) \\
D_A \phi &= 0 & (2.5) \\
D_A^* \phi &= 0 & (2.6)
\end{aligned}$$

More generally, these equations can be defined on a Riemann surface after introducing complex coordinates and proving conformal invariance. Because of their origin, Hitchin originally called these equations the "self-dual equations on a Riemann surface."

The moduli space of solutions of Hitchin's equations is

$$\mathcal{M}(\Sigma, G) = \left\{ (A, \phi) \middle|^{F_A - \frac{1}{2}[\phi,\phi]=0}_{\substack{D_A\phi=0 \\ D_A^*\phi=0}} \right\} / \mathcal{G} \tag{2.7}$$

where $\mathcal{G}$ is the group of real gauge transformations, i.e. automorphisms of the principal bundle P.

Tools for studying solutions of Hitchin's equations

Hitchin's equations are completely integrable! Consequently, they can be approached using twistor theory, which was developed exactly for this situation. Other approaches include algebraic geometry and non-linear PDE. Hitchin described the moduli space using algebraic geometric methods in his original paper. Of all these tools, non-linear PDE hasn't been used as much. However, non-linear PDE seems to be precisely the right tool to understand the noncompact part of the moduli space.

2.3   Moment maps in Hitchin's equations

One important interpretation of Hitchin's equations is that they are the equations for a complex flat connection with a zero moment map condition.

The equations for a complex flat connection are:

$$0 = \mathcal{F}_{A+i\phi} = F_A - \frac{1}{2}[\phi, \phi] + iD_A\phi$$

The real and imaginary parts of the equation $\mathcal{F}_{A+i\phi}$ are the following two of three equations in Hitchin's system of equations.

$$\begin{aligned} \mathrm{Re}(\mathcal{F}_{A+i\phi}) &= F_A - \tfrac{1}{2}[\phi, \phi] &= 0 \\ \mathrm{Im}(\mathcal{F}_{A+i\phi}) &= D_A\phi &= 0 \end{aligned}$$

The equation we're missing, $D_A^*\phi = 0$ is related to the zero moment map condition.

Consequently, one can see that given a solution of Hitchin's equations (A, ϕ), we get a complex flat connection $D_A + i\phi$. This map from solutions of Hitchin's equations to complex flat connections also works up to gauge equivalence. We have a map of moduli spaces

$$\left\{ (A, \phi) \middle|^{F_A - \frac{1}{2}[\phi,\phi]=0}_{\substack{D_A\phi=0 \\ D_A^*\phi=0}} \right\} / \mathcal{G} \rightarrow \left\{ A + i\phi \middle| \mathcal{F}_{A+i\phi} = 0 \text{ i.e. } {}^{F_A - \frac{1}{2}[\phi,\phi]=0}_{D_A\phi=0} \right\} / \mathcal{G}^{\mathbb{C}}.$$

Somewhat surprisingly, Hitchin proved that this map above is actually an isomorphism.

Theorem 2.3.1 *(Hitchin) We have the following correspondence:*

$$\left\{ (A, \phi) \middle|^{F_A - \frac{1}{2}[\phi,\phi]=0}_{\substack{D_A\phi=0 \\ D_A^*\phi=0}} \right\} / \mathcal{G} \cong \left\{ A + i\phi \middle| \mathcal{F}_{A+i\phi} = 0 \text{ i.e. } {}^{F_A - \frac{1}{2}[\phi,\phi]=0}_{D_A\phi=0} \right\} / \mathcal{G}^{\mathbb{C}}$$

On the right hand side of the correspondence, the larger complex gauge group compensates for the missing third equation $D_A^*\phi = 0$. The space on the right is interpreted as the moduli space of flat complex connections. In the next section (Section 2.3.1)we explain the connection between Hitchin's equations and representation theory. This comes from correspondence between flat complex connections over Σ and representations $\rho : \pi_1(\Sigma, G_\mathbb{C})$ of the fundamental group of Σ.

The space on the left, the Hitchin moduli space, is better for analysis. The group of complex gauge transformations do not fix the norms. And you can't have varying norms with PDE! Thus, real gauge theories are necessary for analysis. In Section 2.3.2, we discuss the zero moment condition and how the proof of the correspondence works. Particularly, we discuss how the equation $D_A^*\phi$ allows us to reduce from a complex gauge theory to a real gauge theory.

2.3.1 Representation Theory

For any group G' and any manifold M, there is a correspondence between flat G'-connections over M and representations $\pi_1(M) \to G'$. In the context of Hitchin's equations, we are looking at flat complex connections ($G' = G_\mathbb{C} = SL(2, \mathbb{C})$) over a Riemann surface Σ.

From flat connections to representations: Let $P_\mathbb{C} \to^p \Sigma$ be a principal $G_\mathbb{C}$-bundle with flat connection ∇. We get a representation as follows. Pick a base point $x_0 \in \Sigma$. We'll construct a representation

$$\rho : \pi_1(\Sigma, x_0) \to G_\mathbb{C}$$

For any $[\gamma] \in \pi_1(\Sigma, x_o)$, choose a point $\tilde\gamma(0) \in p^{-1}(x_0) = P_{\mathbb{C},x_0}$, the fiber of $P_\mathbb{C}$ over x_0. γ lifts to a unique flat section, $\gamma : [0,1] \to \Sigma$, of $P_\mathbb{C}$ starting at $\tilde\gamma(0)$. The representation
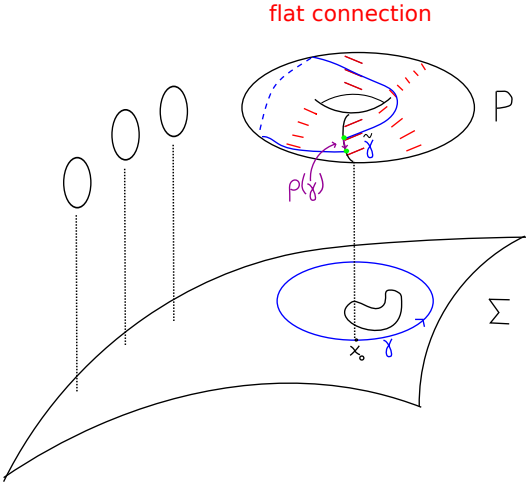

Figure 2.1: Getting a representation

$\rho(\gamma) \in \pi_1(\Sigma, x_0)$ is defined to be the difference of the endpoints of $\tilde\gamma$ i.e.

$$\tilde\gamma(1) = \rho(\gamma)\tilde\gamma(0)$$

This is known as the monodromy representation of the connection.

We made a few choices defining the representation

$$\rho : \pi_1(\Sigma, x_0) \to G_{\mathbb{C}}.$$

We chose a path γ in the class of $[\gamma] \in \pi_1(\Sigma, x_0)$. The resulting representation does not depend on the choice here precisely because the connection is *flat*. Second, we choose a starting point $\tilde{\gamma}(0)$. Since we are only interested in the difference of $\tilde{\gamma}(0)$ and $\tilde{\gamma}(1)$, the representation ρ does not depend on this choice either. Lastly, we chose a base point $x_0 \in \Sigma$. This *does* change the representation! Changing the base point changes the representation by conjugation. Consequently, we get a well-defined map from flat connections to conjugacy classes of representations

$$\{\rho : \pi_1(\Sigma, x_0) \to G_{\mathbb{C}}\} / G_{\mathbb{C}}.$$

From representations to flat connections:

Given $\rho : \pi_1(\Sigma, x_0) \to G_{\mathbb{C}}$, we can construct a bundle with flat connection as follows:

Let $\tilde{\Sigma}$ be the universal cover of Σ. Consider the trivial bundle

$$\tilde{\Sigma} \times G_{\mathbb{C}} \to \tilde{\Sigma}$$

with flat connection $\nabla = d$, where d is the ordinary exterior derivative on $\tilde{\Sigma}$ that doesn't see the $G_{\mathbb{C}}$-component.

To get a bundle on Σ, take this flat bundle upstairs, then twist it up and glue it together using the representation ρ.

$$\tilde{\Sigma} \times G_{\mathbb{C}} / (x, g) \sim (\gamma \cdot x, \rho(\gamma) g)$$

When the bundle is twisted up, the flat connection is twisted too.

In the same neighborhood of x_0 on Σ, the flat connection may look looks like d on Σ, but equally well like $g \circ d \circ g^{-1}$ for any $g = \rho(\gamma)$ since lifts $\gamma \in \pi_1(\Sigma)$ change the preimage of x_0 in $\tilde{\Sigma}$ that comes from changing the base point. In twisting d,

$$g \circ d \circ g^{-1} = d - dgg^{-1} = d + A + i\phi,$$

we get a new field, the "Higgs field" ϕ, from the complex part of the representation.

Given a complex flat connection $D_A + i\phi$, the pair (A, ϕ) satisfies two of the three equation in Hitchin's system of equations. However, the pair does not necessarily satisfy the third equation $D_A^* \phi = 0$.

2.3.2   The moment map

We've stated that Hitchin proved the following correspondence

Theorem 2.3.2 *(Hitchin)*

$$\left\{ (A, \phi) \Big|^{\substack{F_A - \frac{1}{2}[\phi,\phi]=0 \\ D_A\phi=0 \\ D_A^*\phi=0}} \right\} / \mathcal{G} \cong \left\{ A + i\phi \Big| \mathcal{F}_{A+i\phi} = 0 \ i.e. \ ^{\substack{F_A - \frac{1}{2}[\phi,\phi]=0 \\ D_A\phi=0}} \right\} / \mathcal{G}^{\mathbb{C}}$$

25

How exactly does this work? Hitchin realized that the complex gauge group acts symplectically on the space of flat connections. Consequently, get get a moment map

$$\mu : \mathcal{M} \rightarrow \mathcal{G}_{\mathbb{C}}^*$$
$$(A, \phi) \mapsto \left[\quad g \rightarrow D_A^{*[g]} \phi \quad \right].$$

We can make the complex $SL(2, \mathbb{C})$ gauge theory into a real $SU(2)$ gauge theory by taking the symplectic reduction of the space

$$\mu^{-1}(0)/\mathcal{G}.$$

Remark The moment map has nothing to do with the base manifold Σ being two-dimensional. Moreover, the symplectic action of the complex gauge group on the moduli space is not related to the symplectic structure–or lack-thereof–on the base manifold. There are moment maps even when the base manifold M has no symplectic structure, for example when $\dim_{\mathbb{R}} M = 3$.

There is another equivalent perspective on this equivalence between the two moduli spaces. Rather than using symplectic reduction, we can do something more concrete. We reduce from a complex to a real gauge group theory by showing that there is a distinguished real orbit within (almost) every complex gauge orbit. Let $h \in \mathcal{G}_{\mathbb{C}}$, a complex gauge transformation and look at it's action on $D_A + i\phi$:

$$h(D_A + i\phi)h^{-1} = D_{\tilde{A}_h} + i\tilde{\phi}_h$$

We seek to minimize the following action integral over h:

$$\mathcal{J}(A, \phi, h) = \int_M \mu(\tilde{A}_h, \tilde{\phi}_h)d\nu = \int_M |\tilde{\phi}_h|^2 d\nu. \tag{2.8}$$

From computing the first variation of the integral, we find that if h minimizes $\mathcal{J}$ then $D_{\tilde{A}_h}^* \tilde{\phi}_h = 0$. We have the following theorem of Hitchin which extends this observation:

Theorem 2.3.3 *(Hitchin) Suppose $D_A + i\phi$ is an irreducible flat complex connection on a compact manifold, then there exists a complex gauge transformation $h_0 \in \mathcal{G}_{\mathbb{C}}$ such that:*

- *The functional $F[h] := \int_M |\tilde{\phi}_h|^2 \mathrm{dvol}$ is minimized by h_0*

- $D_{\tilde{A}}^{*[h_0]} \tilde{\phi} = 0$

- *h_0 is unique up to real gauge transformations*

2.4 Extensions of Results and Applications

The previous result about the existence of such a h_0 has been extended to other situations. The above theorem proves the existence of h_0 given a flat complex connection on a Riemann surface. It would be great if such an existence theorem held for non-flat connections. For

example, the Kapustin-Witten equations, which we discuss in the next chapter, are equations for a not-necessarily-flat complex connection.

Hitchin's equations are at the center of a lot of mathematics. They appear in mirror symmetry and the geometric Langland's correspondence. Even the recent Field's medal in number theory uses Hitchin's equations.

Chapter 3

The Kapustin-Witten equations

The Kapustin-Witten equations are equations for a complex connection on a 4-manifold with a zero moment map condition. The Kapustin-Witten equations are a recent development in mathematics. Unlike the Yang-Mills equations and Hitchin equations which first appeared in 1954 and 1987 respectively, the Kapustin-Witten equations were introduced in 2007 by Kapustin and Witten in their paper "Electric-Magnetic Duality and the Geometric Langlands Program."

Though the Kapustin-Witten equations are equations on four-dimensional manifolds, the applications of the Kapustin-Witten equations are actually to three-manifolds. This is because there are no interesting solutions to the Kapustin-Witten equations on compact four-manifolds. The class of four-manifolds we consider are products $M^4 = X^3 \times \mathbb{R}$ of a three-manifold X^3 and $\mathbb{R}$. However, because $X^3 \times \mathbb{R}$ is a non-compact manifold, we must specify the behavior of solutions near $X^3 \times \{-\infty\}$ and $X^3 \times \{\infty\}$. There are many interesting boundary value problems arising from the Kapustin-Witten equations.

The Kapustin-Witten equations are, in some sense that we will discuss, both the complexification of the (real) self-dual Yang-Mills and the better 4-dimensional analog of Hitchin's equations on a Riemann surface. Consequently, in this section we will state results about the Kapustin-Witten equations that are similar to the results we stated about the Yang-Mills equations and the results we stated about Hitchin's equations. Because the Kapustin-Witten equations are so difficult, we will often use analogies with the easier equations we've discussed so far. To distinguish when we are talking about the Kapustin-Witten equations versus when we are talking or proving something in an analogous situation like the self-dual Yang-Mills equations or the Hitchin's equations, we will indent the text in all analogies.

The **Kapustin-Witten equations** are a $\theta \in S^1$ family of equations compactly expressed as

$$
\begin{aligned}
e^{i\theta}\mathcal{F}_{D_A+i\phi} &= \star\overline{e^{i\theta}\mathcal{F}_{D_A+i\phi}} && (3.1) \\
D_A^*\phi &= 0 && (3.2)
\end{aligned}
$$

with θ and $\theta + \pi$ giving the same equations, i.e. $\theta \in \mathbb{R}/\pi\mathbb{Z}$. The bar appearing in the above equation comes from the involution of complex structure on $G_\mathbb{C} \supset G$. Expanding the first

equation into real and imaginary parts one has:

$$\cos(\theta)(F_A - \frac{1}{2}[\phi, \phi])^+ = \sin(\theta)(D_A\phi)^+$$
$$\sin(\theta)(F_A - \frac{1}{2}[\phi, \phi])^- = -\cos(\theta)(D_A\phi)^-.$$

In the above equation, the Lie-algebra two-forms $F_A - \frac{1}{2}[\phi, \phi]$ and $D_A\phi$ above split[1] into self-dual and anti-self-dual parts.

Since we don't really know the applications of the Kapustin-Witten equations at this point, we don't really know if we gain anything from having a one-parameter θ-family or not. However, when we talk about "the" Kapustin-Witten equations, we usually mean the $\theta = \frac{\pi}{4}$-Kapustin-Witten equations given by

$$F_A - \frac{1}{2}[\phi, \phi] = \star D_A\phi \tag{3.3}$$
$$D_A^*\phi = 0.$$

> **The self-dual and anti-self-dual Yang-Mills equations** Recall that the anti-self-dual and self-dual Yang-Mills equations were

$$F_A = \star \pm F_A.$$

The Kapustin-Witten equations are a "complexification" of these equations. Note the following formal similarities (and differences) between the Kapustin-Witten equation and the first equation of the Kapustin-Witten equations (3.1):

$$e^{i\theta}\mathcal{F}_{D_A+i\phi} = \star \overline{e^{i\theta}\mathcal{F}_{D_A+i\phi}}.$$

- Both are equations on four-manifolds.

- The self-dual Yang-Mills equations were equations for a real connection D_A, while the Kapustin-Witten equations are equations for complex connection $D_A + i\phi$.

- In the (anti)self-dual Yang-Mills equations, the ordinary curvature F_A appears. In the Kapustin-Witten equations the complex curvature $\mathcal{F}_{D_A+i\phi}$ appears.

- In both, the Hodge star appears.

- The ± 1 coefficient appearing in the self-dual and anti-self-dual Yang-Mills equations is a real unit. The θ choice in the Kapustin-Witten equations is a choice of complex unit $e^{i\theta} \in \mathbb{C}^\times$.

◁

---

[1]Given a two-form ω, the self-dual part is $\omega^+ = \frac{1}{2}(\omega + \star\omega)$ and the anti-self-dual part is $\omega^- = \frac{1}{2}(\omega - \star\omega)$. More generally, for Lie-algebra-valued two-forms, B, the self-dual part is $B^+ = \frac{1}{2}(B + \star B^*)$ where $\star$ acts on the the two-form and $*$ acts on the Lie-algebra. The anti-self-dual part is similar.

3.1 The Kapustin-Witten equations are absolute minimizers of the complex Yang-Mills functional

In this section, we prove that the solutions of the Kapustin-Witten equations satisfy the Euler-Lagrange equations of the **complex Yang-Mills functional**

$$
\begin{aligned}
\mathfrak{Ym}_{\mathbb{C}}(D_A + i\phi) &= \int_M |\mathcal{F}_{A+i\phi}|^2 \mathrm{dvol} & (3.4)\\
&= \int_M \mathrm{tr}(\mathcal{F}_{A+i\ phi} \wedge \star\overline{\mathcal{F}_{A+i\phi}})\mathrm{dvol}.\\
&= \int_M (|F_A - \frac{1}{2}[\phi,\phi]|^2 + |D_A\phi|^2)\mathrm{dvol}
\end{aligned}
$$

but–more than that–they are absolute minima of this functional. Being an absolute minima is a much stronger statement than that they satisfy the Euler-Lagrange equations, known as the **complex Yang-Mills equations**

$$
\begin{aligned}
D_A^*\left(F_A - \frac{1}{2}[\phi \wedge \phi]\right) + [D_A\phi \wedge \phi] &= 0 & (3.5)\\
D_A^* D_A\phi + \left[\left(F_A - \frac{1}{2}[\phi \wedge \phi]\right) \wedge \phi\right] &= 0.
\end{aligned}
$$

which are obtained first order variation of the functional.

> ▷ **Self-dual Yang-Mills equations:** Recall that the Yang-Mills equations are the Euler-Lagrange equations of the (real) Yang-Mills functional
>
> $$
> \mathfrak{Ym}(D_A) = \int_M |F_A|^2 \mathrm{dvol}.
> $$
>
> Solutions of the (anti)self-dual Yang-Mills equations were the *absolute* minima of the Yang-Mills functional.
>
> The complex Yang-Mills functional is the "complexification" of the real Yang-Mills functional since the full complex curvature $\mathcal{F}_{D_A+i\phi}$ replaces the ordinary curvature F_A.
>
> We say that the Kapustin-Witten equations are a "complexification" of the self-dual Yang-Mills equations precisely because the Kapustin-Witten equations are the absolute minima of the complex Yang-Mills functional, in the same sense that the self-dual Yang-Mills equations are the absolute minima of the Yang-Mills functional. ◁

If you really feel like torturing yourself... actually, if you feel like torturing your student, give them the assignment of proving that the Kapustin-Witten equations satisfy the Euler-Lagrange equations of the complex Yang-Mills functional directly. This takes pages, and I've never actually done it myself. However there is a more elegant way to do it!

Rather than directly proving that the Kapustin-Witten equations satisfy the Euler-Lagrange equations, it is easier to prove the stronger statement that the Kapustin-Witten equations are absolute minimizer of the complex Yang-Mills functional. The proof of this fact is very similar to the proof that the (anti)self-dual Yang-Mills equations are absolute minimizers of the real Yang-Mills functional, so we prove both statements for comparison. This may seem a little tedious, but in future sections, we will simply state and prove statements about the self-dual Yang-Mills equations or Hitchin's equations on Riemann surfaces or 4-manifolds and then say "there is an analogous statement (with more complicated proof) for the Kapustin-Witten equations."

$\triangleright$ **Self-dual Yang-Mills equations:** For the *real* Yang-Mills functional, we have the following proposition:

Proposition 3.1.1 *Within a topological class $c_2(P) = k$ and with fixed boundary data $D_A|_{\partial M}$, solutions of the the (anti-) self-dual Yang-Mills are absolute minimizers of the Yang-Mills functional.*

Proof We decompose the Yang-Mills functional into three pieces: the first piece is minimized by the solutions of the (anti)self-dual Yang-Mills equations; the second piece is a topological term; the third piece is an integral over the boundary.

$$
\begin{aligned}
\mathfrak{Ym}(D_A) \quad &= \quad \int_M |F_A|^2 \mathrm{dvol} \\
&\overset{Lemma\,3.1.2}{=} \quad \int_M \frac{1}{2}|(F_A \pm \star F_A|^2 \mathrm{dvol} \mp \int_M \mathrm{tr}(F_A \wedge F_A) \\
&\overset{Fact\,3.1.3}{=} \quad \int_M |(F_A \mp \star F_A|^2 \mathrm{dvol} \pm 8\pi^2 c_2(P) \mp \int_{\partial M} CS(A)
\end{aligned}
$$

where $CS(A)$ is the Chern-Simons functional. We delay defining the Chern-Simons functional and stating and proving the referenced "Lemma 3.1.2" and "Fact 3.1.3" until immediately after the proof, for readability. Consequently, the functional is bounded below by a term depending only on the topology of the G-bundle P and the boundary data of the connection

$$
\mathfrak{Ym}(D_A) \geq 8\pi^2 c_2(P) + \int_{\partial M} CS(A)
$$

with equality if, and only if, D_A is self-dual or anti-self-dual. ∎

In the above proof, the main computational step is proving the decomposition of the Yang-Mills functional. Here, we prove the decomposition, though we will not do so for the Kapustin-Witten equations.

Lemma 3.1.2 *Given a Lie-algebra two-form F_A, here though of as the curvature of some connection, the following equation holds:*

$$
|F_A \pm \star F_A|^2 = 2|F_A|^2 \pm 2\mathrm{tr}(F_A \wedge F_A).
$$

Note that the $SU(n)$-invariant inner product on the forms valued in complexification $\mathfrak{sl}(n,\mathbb{C}) = \mathfrak{su}(n) \otimes_{\mathbb{R}} \mathbb{C}$ is

$$\langle A, B \rangle := \operatorname{tr}(A \wedge \star B^*)$$

for forms $A, B \in \Omega(M, \mathfrak{sl}(n,\mathbb{C}))$. When the forms are valued in the smaller space $A, B \in \Omega(M, \mathfrak{su}(n))$, as is the case here, the inner product simplifies because $B^* = -B$.

Proof

$$
\begin{aligned}
|F_A \pm \star F_A|^2 &= \langle F_A \pm \star F_A, F_A \pm \star F_A \rangle \\
&= |F_A|^2 + |\star F_A|^2 \pm \langle F_A, \star F_A \rangle \pm \langle \star F_A, F_A \rangle \\
&= 2|F_A|^2 \pm 2 \langle F_A, \star F_A \rangle \\
&= 2|F_A|^2 \pm 2\operatorname{tr}\left(F_A \wedge \star(\star F_A)^*\right) \\
&= 2|F_A|^2 \mp 2\operatorname{tr}\left(F_A \wedge F_A\right) \quad \blacksquare
\end{aligned}
$$

Now we consider the second term $\operatorname{tr}(F_A \wedge F_A)$ appearing in the previous lemma. This term is purely topological–rather than geometric–because the metric-dependent $\star$-operator does not appear. By integrating this term by parts we get a topological term plus and integral over the boundary.

Fact 3.1.3 *Given a G-bundle P over M^4 with connection D_A, the following equality holds:*

$$\int_M \operatorname{tr}(F_A \wedge F_A) = \text{topological term} + \int_{\partial M} CS(A). \tag{3.6}$$

When $\partial M = \emptyset$, the topological term is just the second Chern class $8\pi^2 c_2(P)$. In the boundary term, $CS(A)$ is the Chern-Simons three-form.

What is the Chern-Simons three-form? In a flat bundle, the formula for the Chern-Simons three-form is

$$\mathcal{CS}(A) := \operatorname{tr}(A \wedge F_A) - \frac{1}{6}\operatorname{tr}(A \wedge A \wedge A).$$

In a rough sense, most bundles with connection over a 3-manifold are topologically trivial, though they may not be flat. To give the formula for the Chern-Simons three-form on a topologically trivial but non-flat bundle one has to replace the traces appearing with the appropriate inner product in the twisted Lie algebra $\mathfrak{g}_P$. One important fact about the Chern-Simons functional is that:

Fact 3.1.4 *The Euler-Lagrange equations for the Chern-Simons functional are simply that the curvature is zero.*

?? The Chern-Simons functional can be defined "the functional whose Euler-Lagrange equations are simply that the curvature is zero." However, there is some ambiguity in this definition. ◁

We now extend the above argument to the more complicated complex situation of the Kapustin-Witten equations and prove that the Kapustin-Witten equations are absolute minima for the **complex Yang-Mills functional**

$$\mathfrak{Ym}_{\mathbb{C}}(D_A + i\phi) := \int_{M^n} |\mathcal{F}_{D_A + i\phi}|^2 \mathrm{dvol}. \tag{3.7}$$

The idea is that the Kapustin-Witten equations play the same role for the complex regime that the self-dual Yang-Mills equations play in the real regime. We will need to define the appropriate complex version of the Chern-Simons functional $\mathcal{CS}_{\mathbb{C}}[D_A + i\phi]$ is defined, by the complex analogue of Fact 3.1.4 to be the functional whose Euler-Lagrange equations are $\mathcal{F}_{D_A + i\phi} = 0$, i.e. the complex curvature vanishes. As in the real case, there is some ambiguity in this definition, by we will not discuss it. The complex Chern-Simons functional is the integral of the complex Chern-Simons three-form

$$\mathcal{CS}_{\mathbb{C}}(A + i\phi) := \int_{\partial M^4} CS_{\mathbb{C}}(A + i\phi).$$

For a flat bundle over $\partial M^4 = X^3$, the formula for the complex Chern-Simons 3-form is

$$CS_{\mathbb{C}}(A + i\phi) = \mathrm{tr}\left((A + i\phi) \wedge \mathcal{F}_{A+i\phi}\right) - \frac{1}{6}(A + i\phi) \wedge (A + i\phi) \wedge (A + i\phi).$$

This is the same formula we obtain by formally replacing the real connection A in the real Chern-Simons three-form formula with the complex connection $A + i\phi$

Proposition 3.1.5 *Solutions of the Kapustin-Witten equations are absolute minima for the complex Yang-Mills functional.*

Proof As in the real case, we decompose the complex Yang-Mills functional. We break the integral into three pieces: the first piece vanishes when the Kapustin-Witten equations are satisfied; the second is topological; the third is related the the complex Chern-Simons functional.

$$
\begin{aligned}
\int_{M^4} |\mathcal{F}_{\mathbb{C}}|^2 \mathrm{dvol} &= \int_{M^4} |e^{i\theta}\mathcal{F}_{\mathbb{C}}|^2 \mathrm{dvol} \\
&= \int_{M^4} \frac{1}{2}|e^{i\theta}\mathcal{F}_{\mathbb{C}} - \star\overline{e^{i\theta}\mathcal{F}_{\mathbb{C}}}|^2 + \mathfrak{Re}\left(e^{2i\theta} tr(\mathcal{F}_{\mathbb{C}} \wedge \mathcal{F}_{\mathbb{C}})\right) \mathrm{dvol} \\
&= \frac{1}{2}\int_{M^4} |e^{i\theta}\mathcal{F}_{\mathbb{C}} - \star\overline{e^{i\theta}\mathcal{F}_{\mathbb{C}}}|^2 \mathrm{dvol} \\
&\quad + \mathfrak{Re}\left(e^{i2\pi\theta}\text{top. term}\right) + \mathfrak{Re}\left(\int_{\partial M^4} e^{2i\theta}\mathrm{tr}\left(CS_{\mathbb{C}}(A + i\phi)\right)\right)
\end{aligned}
$$

(We omit the proof of the above decomposition, because the computations are formally similar to the self-dual Yang-Mills computations.) We conclude that

$$\int_{M^4} |\mathcal{F}_{\mathbb{C}}|^2 \mathrm{dvol} \geq \mathfrak{Re}\left(e^{i2\pi\theta}\text{top. term}\right) + \mathfrak{Re}\left(\int_{\partial M^4} e^{2i\theta}\mathrm{tr}\left(CS_{\mathbb{C}}(A + i\phi)\right)\right) \tag{3.8}$$

with equality, if and only if, $A + i\phi$ satisfies the θ-Kapustin-Witten equations. Solutions of the Kapustin-Witten equations are absolute minima of the complex Yang-Mills functional because at solutions of the Kapustin-Witten equations the functional is only dependent on the topology of the bundle and the boundary values. ∎

3.2   Zero Moment Map Condition

In Kapustin-Witten theory, we want to associate a moduli space of solutions of the Kapustin-Witten equations to a topological manifold M^4. However, we must *choose* a Riemannian structure on M^4 to define the Kapustin-Witten equations. The role the geometry of M^4 plays in Kapustin-Witten theory is subtle.

So far we've mostly discussed the first of the two the Kapustin-Witten equations. There is some of the geometry in

$$e^{i\theta} \mathcal{F}_{D_A + i\phi} = \star \overline{\left(e^{i\theta} \mathcal{F}_{D_A + i\phi} \right)} \tag{3.9}$$

because $\star$, the Hodge star operator, appears. However, this geometry is not very sharp because the $\star$-operator is conformally invariant so, for example, it doesn't see the difference between S^4, hyperbolic 4-manifolds, and $\mathbb{R}^4$.

We now consider the second of the Kapustin-Witten equations

$$D_A^* \phi = 0.$$

This equation is the zero moment map condition for the real gauge group. It is an important part of the Kapustin-Witten equations, because Eq. 3.9 is not elliptic–modulo the real gauge group, since we're doing real gauge theory–on its own. However, with this important trailer equation $D_A^* \phi = 0$,

Fact 3.2.1 *The $\left(\theta = \frac{\pi}{4} \right)$-Kapustin-Witten equations*

$$\begin{aligned} F_A - \frac{1}{2}[\phi, \phi] &= \star D_A \phi \tag{3.10} \\ D_A^* \phi &= 0 \tag{3.11} \end{aligned}$$

are elliptic.

There is a lot of geometry in this second equation!

3.3   Applications of the Kapustin-Witten equations

In Chapter 1, we saw that we got interesting topological invariants, the Donaldson invariants, from studying the solutions of the self-dual Yang-Mills equations on M^4. *What kind of invariants might we get from studying solutions of the Kapustin-Witten equations? Do we get new interesting invariants that we didn't get from studying solutions of the self-dual Yang-Mills equations?* In Section 3.3.1, towards answering this question, we discuss the relationship between the Kapustin-Witten equations and some of the other gauge theories we've discussed so far, particularly self-dual Yang-Mills and Hitchin's equations. In Section 3.3.2, we prove that studying the Kapustin-Witten equation on compact M^4 is generally no better than studying Hitchin's equations on M^4.

3.3.1 Relationship between the Kapustin-Witten Equations and other equations

The Kapustin-Witten equations are a generalization of both the self-dual Yang-Mills equations and Hitchin's equations.

In Chapter 2, we discussed Hitchin's equations on Riemann surfaces. However, Hitchin's equations can be defined on a Kähler manifold, M, of any dimension. **Hitchin's equations** are

$$\begin{aligned}
\mathcal{F}_{A+i\phi} &= 0 \\
D_A^* \phi &= 0.
\end{aligned}$$

These equations are overdetermined when $\dim_{\mathbb{R}} M > 2$. Certainly, any solution of Hitchin's equations on M^4 is also a solution of the Kapustin-Witten equations. However, if the Kapustin-Witten equations are worth anything, they must have other solutions–solutions that are not solutions of Hitchin's equations. Explicitly, for $\theta = 0$, the difference between Kapustin-Witten and Hitchin's equations is that solutions of the $(\theta = 0)$-Kapustin-Witten equations satisfy

$$\begin{aligned}
\left(F_A - \frac{1}{2}\phi \wedge \phi \right)^+ &= 0 \\
(D_A \phi)^- &= 0 \\
D_A^* \phi &= 0
\end{aligned}$$

while solutions of Hitchin's equations satisfy

$$\begin{aligned}
F_A - \frac{1}{2}\phi \wedge \phi &= 0 \\
D_A \phi &= 0 \\
D_A^* \phi &= 0,
\end{aligned}$$

i.e. *both* the self-dual and anti-self dual parts of both $F_A - \frac{1}{2}\phi \wedge \phi$ and $D_A\phi$ vanish for solutions of Hitchin's equations.

Similarly, if the Higgs field ϕ vanishes, the Kapustin-Witten equations become equations on a real connection D_A:

$$e^{i\theta} F_A = \star e^{-i\theta} F_A. \tag{3.12}$$

When $\theta = 0$, Eq. 3.12 is the self-dual Yang-Mills equations. When $\theta = \frac{\pi}{2}$, Eq. 3.12 is the anti-self-dual Yang-Mills equations. When $\theta \in [0,\pi) \backslash \{0, \frac{\pi}{2}\}$, any solution satisfies $F_A = 0$, i.e. it is a flat connection.

Figure 3.1 represents this general nesting of the solutions of the different gauge theories: The blue region in the above diagram contains the interesting new solutions we get when we study the Kapustin-Witten equations. In the next section we turn our attention to those solutions. We show that when M^4 is compact, we have no new solutions for most choices of parameter θ.
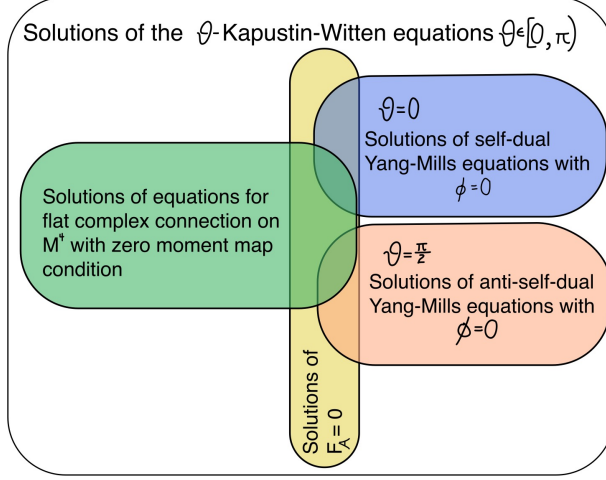
Figure 3.1: Relationship between Equations

3.3.2 Kapustin-Witten Equations on Compact Manifolds

The Kapustin-Witten equations, though equations on a 4-manifold, are generally used to study 3-manifolds instead. This is because for most choices of the parameter $\theta \in S^1$, there are no new interesting solutions of the Kapustin-Witten equations on a compact 4-manifold, as the next proposition shows.

Proposition 3.3.1 *Let M^4 be a compact 4-manifold. If $\theta \neq 0, \frac{\pi}{2} \in \mathbb{R}/\pi\mathbb{Z}$, then the only solutions of the Kapustin-Witten equations are those in which $\mathcal{F}_\mathbb{C} = 0$.*

This proposition says that when $\theta \neq 0, \frac{pi}{2} \in \mathbb{R}/\pi\mathbb{Z}$, every solution of the Kapustin-Witten equations is also a solution of Hitchin's equations. The proof of the above proposition is a corollary to the type of proof in the proof that Kapustin-Witten equations are the absolute minima.

Proof For any (A, ϕ), the curvature $\mathcal{F}_{\mathcal{A}=A+i\phi}$ satisfies

$$\int_{M^4} \left| e^{i\theta} \mathcal{F}_\mathcal{A} \right|^2 \mathrm{dvol} = \int_{M^4} \left\langle e^{i\theta} \mathcal{F}_\mathcal{A}, e^{i\theta} \mathcal{F}_\mathcal{A} - \star \overline{e^{i\theta} \mathcal{F}_\mathcal{A}} \right\rangle \mathrm{dvol} + e^{2i\theta} \int_{M^4} \mathrm{tr}(\mathcal{F}_\mathcal{A} \wedge \mathcal{F}_\mathcal{A}) \mathrm{dvol}.$$

If (A, ϕ) is a solution of the Kapustin-Witten equations, then the first term on the right-hand-side vanishes, and we have

$$\int_{M^4} \left| e^{i\theta} \mathcal{F}_{A+i\phi} \right|^2 \mathrm{dvol} = e^{2i\theta} \int_{M^4} \mathrm{tr}(\mathcal{F}_{A+i\phi} \wedge \mathcal{F}_{A+i\phi}) \mathrm{dvol}.$$

Notice left-hand-side is real. The right-hand-side is real only if $e^{2i\theta}$ is real or else everything vanishes. Consequently, for $\theta \neq 0, \frac{\pi}{2} \in \mathbb{R}/\pi\mathbb{Z}$, the right-hand-side is only real if $\mathcal{F}_{A+i\phi} = 0$, i.e. the complex connection is flat. ■

This proposition tells us that the topological applications of the Kapustin-Witten equations won't generally be to closed 4-manifolds. There are special values of the θ-parameter, e.g. $\theta = 0$ or $\theta = \frac{\pi}{2}$, where are new solutions of the Kapustin-Witten equations– solutions that are not solutions of Hitchin's equations. However, this will just be a modification of the real gauge theory in that case.

3.4 Boundary Value Problems in the Kapustin-Witten Equations

In the topological applications of the Kapustin-Witten equations, the manifold M^4 will usually be non-compact. Consequently, one must impose some conditions on the pair (A, ϕ) on the boundary $\partial \overline{M}$. There is not just one such choice! But not all choices are made equal.

We'll restrict our attention to non-compact four-manifolds $M^4 = X^3 \times \mathbb{R}$. Then, we are interested in solutions of Kapustin-Witten equations which satisfy certain boundary conditions on the $X_{-\infty} = X \times \{-\infty\}$ and $X_{\infty} = X \times \{+\infty\}$ ends. *Which boundary conditions should we impose?* One possible boundary condition–the simplest boundary condition– would be to impose that $(D_A + i\phi)|_{X_{\pm\infty}}$ is flat.

What topological information might this boundary value problem contain? We explain this by the following analogy:

> The Kapustin-Witten equations play the same role in complex Chern-Simons flow that the anti-self-dual Yang-Mills equations play in the (real) Chern-Simons flow.

We now digress, and briefly explain the role of the anti-self-dual Yang-Mills equation in Chern-Simons flow

> ▷ **Floer homology and solutions of the anti-self-dual Yang-Mills equation** Floer homology is a homology theory for three-manifolds. Floer homology is a particular infinite-dimensional version of (finite-dimensional) Morse homology. On the infinite dimensional space of connections on X^3, the (real) Chern-Simons functional on three-manifolds X plays the role of the Morse function. Floer homology, like Morse homology, is a homology theory constructed from the critical points of the functional and their respective domains of attraction under forward and backwards gradient flow. As stated in Fact **??**, the Euler-Lagrange equations for the Chern-Simons functional are the flat connections. Consequently, the critical points of the Chern-Simons functional are gauge equivalence classes of flat connections. Under Chern-Simons gradient flow, the orbits connecting these flat connections on X^3 (see Fig. 3.2) are, surprisingly, anti-self-dual instantons on the four-manifold $X^3 \times \mathbb{R}$! ◁

Figure 3.2: Critical points and orbits connecting them in Floer homology

Returning to the Kapustin-Witten equations, we had asserted that "The Kapustin-Witten equations play the same role in complex Chern-Simons flow that the anti-self-dual Yang-Mills equations play in the (real) Chern-Simons flow." The critical points of the the complex Chern-Simons functional are flat complex connections on X^3–the kinds of boundary connections we've specified in our simplest boundary value problem. The orbits connecting this flat connections are solutions of "the" $(\theta = \frac{\pi}{4})$-Kapustin-Witten equations on the four-manifold $X \times \mathbb{R}$. Consequently, the complex analog of Floer homology is related to the moduli space of solutions to the Kapustin-Witten equations with this boundary condition.

There are other more complicated possible boundary conditions. A natural questions is "What other kinds of invariants (and consequent applications) might we get from these more complicated boundary value problems?" People have all sorts of ideas about what the applications of the Kapustin-Witten theory will be, but before that can go forward, we need to do the analysis for the Kapustin-Witten BVP. The most interesting problem at the moment in Kapustin-Witten theory that is somehow tractable is to write down proper boundary value problems. However, writing down the appropriate boundary value problem for these is hard! For comparison, writing down the proper boundary value problem for the first-order Yang-Mills equations is already difficult, and not at all straightforward.

The Kapustin-Witten equations probably won't give new interesting four-manifold invariants, but they probably will give interesting knot invariants. Witten wrote a paper in 2011 titled "Khovanov Homology and Gauge Theory" discussing how one can compute the Khovanov homology of a knot–the "categorification','in some sense, of the celebrated Jones polynomial–using the Kapustin-Witten equations.

The knot K is embedded in the 3-manifold X^3, and one considers the solutions of the Kapustin-Witten equations on $X \times \mathbb{R}$. The embedding of the knot enters the boundary conditions, and one specifies some singular behavior of the complex connection near the knot.

3.5 Non-linear Analysis in the Kapustin-Witten Equations

In the previous sections, we've seen that non-linear analysis is a well-suited tool for understanding the non-compactness of the moduli space. In the real case, the space of solutions to Yang-Mills is almost compact. It fails to be compact by bubbles coming off and things concentrating at points. The non-compactness of the Hitchin moduli space is more complicated, and the non-compactness of Kapustin-Witten moduli space is every more complicated. We want to understand the non-compact part of the Kapustin-Witten moduli space. In this section, we discuss some aspects of this problem.

3.5.1 Bochner-Weitzenböck Formula

The Bochner-Weitzenböck Formula is an important tool in the analysis of solutions. The geometry of M^4 enters our estimates for solutions on M^4 through the Ricci curvature tensor featured in the Bochner-Weitzenböck formula. Note that the Ricci curvature tensor knows more about the geometry of M^4 than just its conformal geometry, and *can* distinguish the sphere, flat plane, and hyperbolic space.

The Bochner-Weitzenbck formula is a generalization of the better known Bochner formula

$$\nabla^* \nabla \alpha = (dd^* + d^*d)\,\alpha - \mathrm{Ric}(M)\alpha. \tag{3.13}$$

The Bochner formula relates the exterior differential d (on 1-forms) and full differential ∇ (on 1-tensors) via their relative Laplacians.

▷ We briefly digress to demonstrate how the Bochner formula may be used to extract topological information from geometric information.

Theorem 3.5.1 *(Bochner) Let M^n be a compact manifold with $Ric(M^n) > 0$. Then $H^1(M, \mathbb{R}) = 0$.*

Proof For any $\alpha \in \Omega^1(M, \mathbb{R})$,

$$\nabla^* \nabla \alpha = dd^* \alpha + d^* d\alpha - \mathrm{Ric}(M)\alpha.$$

Taking the inner product with α and integrating both sides we have

$$\int_{M^4} |\nabla \alpha|^2 = \int_{M^4} |d^* \alpha|^2 + \int_{M^4} |d\alpha|^2 - \int_{M^4} \mathrm{Ric}(M)|\alpha|^2. \qquad (3.14)$$

The Hodge-de Rham Theorem state that there is a harmonic representative $\alpha_0 \in [\alpha] \in H^1(M, \mathbb{R})$. This harmonic representative satisfies $d\alpha_0 = d^* \alpha_0 = 0$; consequently, from Eq. 3.14,

$$\int_{M^4} |\nabla \alpha_0|^2 = - \int_{M^4} \mathrm{Ric}(M)|\alpha_0|^2.$$

The left side is non-negative and the right side is non-positive, so both are zero. Consequently, $\alpha_0 \equiv 0$. Thus, there is only one homology class–the trivial one $[0]$– hence $H^1(M, \mathbb{R}) = 0$. ∎

Remark The non-abelian Hodge correspondence, which features Hitchin's equations, between flat $G_\mathbb{C}$-connections and $G_\mathbb{C}$-Higgs bundles really is a non-abelian $(G_\mathbb{C} \neq \mathbb{C}^\times)$ generalization of the well-known correspondence between the de Rham and Dolbeault cohomology classes. This correspondence is established through the existence of harmonic metrics on both the flat bundle and the Higgs bundle. Hitchin's equations are the equations for these harmonic bundles, which play the analogous role in nonabelian Hodge theory as the harmonic representatives in Hodge-de Rham theory. In Bochner's theorem above, we proved that the harmonic representative vanished. In Theorem 3.5.5 and 3.5.6 on the next page, we use the Bochner-Weitzenböck formula to prove a sort of non-abelian version of this vanishing theorem. ◁

The Bochner-Weitzenbck formula states:

Fact 3.5.2 *(Bochner-Weitzenbck) The connection D_A and related covariant derivative ∇_A are related by the following equation*

$$\nabla_A^* \nabla_A \phi = (D_A^* D_A + D_A D_A^*)\phi + \mathrm{Ricci}\,\phi + [F_A, \phi].$$

for all $\phi \in \Gamma(M, T^*M \otimes \mathfrak{g}_P) \equiv \Omega^1(M, \mathfrak{g}_P)$. The Ricci curvature appearing,

$$\mathrm{Ricci} : T^*M \to T^*M,$$

is obtained from the standard Ricci curvature tensor $\mathrm{Ric} : TM^{\otimes 2} \to \mathbb{R}$ *by the isomorphism of TM and T^*M through the metric on M^4.*

▷ **4D Hitchin's equations** Observe that if (A, ϕ) is a solution of Hitchin's equations, i.e. (A, ϕ) satisfies

$$
\begin{aligned}
F_A - \frac{1}{2}[\phi, \phi] &= 0 \\
D_A \phi &= 0 \quad \text{(This is topological)} \\
D_A^* \phi &= 0 \quad \text{(This is geometric)},
\end{aligned}
$$

then the Bochner-Weitzenbck formula simplifies to

$$
\nabla_A^* \nabla_A \phi - \text{Ricci } \phi - \frac{1}{2}[[\phi, \phi], \phi] = 0.
$$

◁

A similar statement holds for solutions of the complex Yang-Mills equations (Eq. 3.5), the Euler-Lagrange equations of the complex Yang-Mills functional.

Theorem 3.5.3 *If $D_A + i\phi$ is a solution of the complex Yang-Mills equations, then*

$$
\nabla_A^* \nabla_A \phi - \text{Ricci } \phi - \frac{1}{2}[[\phi, \phi], \phi] = 0. \tag{3.15}
$$

By taking the inner product of 3.15 with ϕ (thus only remembering the information in the same direction as ϕ), and integrating over M^4, we get the following corollary,

Corollary 3.5.4 *If $D_A + i\phi$ is a solution of the complex Yang-Mills equations, then*

$$
\frac{1}{2}\Delta|\phi|^2 - |\nabla_A \phi|^2 - |[\phi, \phi]|^2 - \text{Ricci}(\langle \phi \otimes \phi \rangle) \geq 0
$$

where $\langle \cdot, \cdot \rangle : \Gamma(M, T^*M \otimes \mathfrak{g}_P)^{\otimes 2} \to \Gamma(M, T^*M)^{\otimes 2}$.

▷ **4D Hitchin's equations** From this corollary we get a non-abelian relative of the Bochner Theorem (Theorem 3.5.3) In both, positive Ricci curvature gives some sort of vanishing of the harmonic representative.
Theorem 3.5.5 *Let M be a compact manifold with positive Ricci curvature. If (A, ϕ) is a solution of Hitchin's equations, i.e. $D_A + i\phi$ is a flat complex connection satisfying the zero moment map condition $D_A^* \phi$, then $\phi = 0$.*

Remark: We've already discussed that when M is compact and $\theta \neq 0, \frac{\pi}{2}$, then every solution of the Kapustin-Witten equations is also a solution of Hitchin's equations. From this theorem, we see that when Ricci > 0, we can say more: Every solution of the Kapustin-Witten equations is a flat (real) connection $F_A = 0$.

◁

A similar statement holds for complex Yang-Mills connections.

Theorem 3.5.6 *Let M is a compact manifold with positive Ricci curvature. If $D_A + i\phi$ is a solution of the complex Yang-Mills equations, then $\phi = 0$. The connection D_A is a solutions of the real Yang-Mills equations.*

Proof Since M is compact there is a positive constant $c > 0$ such that $\mathrm{Ric}(\xi, \xi) > c|\xi|^2$ for all $\xi \in T_x M$. Then,

$$\frac{1}{2}\Delta|\phi|^2 - |\nabla_A\phi|^2 - |[\phi, \phi]|^2 - c|\phi|^2 \geq 0 \tag{3.16}$$

Every term appearing on the left-hand side is non-positive. Consequently, to satisfy the inequality, every term must vanish. Therefore, $\phi = 0$.

3.5.2 Some Estimates

In my notes I have:

- For Ricci > 0, get beautiful estimates.

- For compact 4-manifold, also get estimates and $\partial M^4 \neq \phi$

3.5.3 Difficulties in the Getting Better Estimates

One of the main difficulties in the analysis of the Kapustin-Witten equations is that we see no way to obtain estimates on $\int_M |\phi|^2 \mathrm{dvol}$. In this subsection, we'll describe how having such local estimates on $\int_B |\phi|^2 \mathrm{dvol}$ would buy us a lot!

To get good estimates, we need to make a choice of gauge. The basic Uhlenbeck gauge fixing theorem fixes such a gauge, given estimates on the curvature $\int_{B'} |F_A|^2 \mathrm{dvol}$.

Theorem 3.5.7 *(Uhlenbeck) There exists a constant $K > 0$ such that if*

$$\int_{B'} |F_A|^2 \mathrm{dvol} < K$$

then there exists a choice of gauge in which

$$(a) \qquad d^*A \qquad = 0$$
$$(b) \qquad \left(\int_{B'} |A|^4 \mathrm{dvol}\right)^{\frac{1}{2}} \leq \int_{B'} |F_A|^2 \mathrm{dvol}.$$

However, the problem for the Kapustin-Witten equations is that we don't have the required estimates on the curvature to fix a good gauge! *Why not?* Fundamentally, it because we don't have a bound on the Higgs field $\int_B |\phi|^2 \mathrm{dvol}$. The following theorem, based on the Bochner-Weitzenbck formula, states that a bound of the Higgs field $\int_B |\phi|^2 \mathrm{dvol}$ would give the required, desired bounds on the curvature.

Theorem 3.5.8 *If $D_A + i\phi$ is a solution of the Kapustin-Witten equations for $\theta \in (0, \frac{\pi}{2})$ and $B' \subset B$, then*

$$\int_{B'} \left(|\nabla_A\phi|^2 + \frac{1}{2}|[\phi, \phi]|^2\right) \mathrm{dvol} \leq \int_B |\phi|^2 \mathrm{dvol}$$
$$\int_{B'} |F_A|^2 \mathrm{dvol} \leq C(\theta) \int_B |\phi|^2 \mathrm{dvol}.$$

(For proof, see the proof of Corollary 4.7 in Gagliardo-Uhlenbeck.)

For solutions of the Kapustin-Witten equations, we don't automatically get a bound on $|\phi|^2$. However, if we were to have the extra data of estimates on $|\phi|^2$, we would get estimates on $|F_A|^2$, and could fix a good gauge. The main theorem of this section, which is stated and proved in Gagliardo-Uhlenbeck Theorem 4.9, says that if we *were* to have estimates on $\int_B |\phi|^2$, the good gauge ensured by Uhlenbeck's theorem, would indeed be a particularly nice gauge.

Theorem 3.5.9 *Let (A, ϕ) be a solution of the θ-Kapustin-Witten equations. If $\int_B |\phi|^2 < K(\theta)$, then there is a gauge in which (A, ϕ) and all their derivatives are bounded in interior balls $B'' \subseteq B$.*

Unfortunately, we see no way of obtaining estimates on $\int_B |\phi|^2 \mathrm{dvol}$ for solutions of the Kapustin-Witten equations or more generally for solutions of the complex Yang-Mills equations. In contrast, for Hitchin's equations on Riemann surfaces, we could get beautiful estimates. We were aided in getting a bound on $|\phi|^2$ by the holomorphic quadratic differential, but no such suitable aid has been found for the Kapustin-Witten equations.

> ▷ **Hitchin's equations on Riemann surface** We show how the holomorphic quadratic differential gives estimates on $\int_B |\phi|^2$ for solutions (A, ϕ) of Hitchin's equations with $G = SU(2)$, for simplicity. The holomorphic quadratic differential is
> $$\varphi_2 = \mathrm{tr}(\phi^{(1,0)})^2.$$
> It is a fact that for any $\lambda > 0$, there exist conformal coordinates (x, y) on $M' \subset M$ such that $\varphi_2 = \lambda dz^2$ in the new coordinate $z = x + iy$. This subset M' is $M' = M \backslash Z$, where Z is the set of zeros of φ_2. In these (x, y)-coordinates, $\phi \in \Omega^1(M, \mathfrak{su}(2))$ can be written
> $$\phi = \phi_x dx + \phi_y dy.$$

Consequently,
$$\lambda = \mathrm{tr}\left(\frac{\phi_x - i\phi_y}{2}\right) = \frac{1}{4}\left(|\phi_x|^2 - |\phi_y|^2 - 2\langle\phi_x, \phi_y\rangle\right).$$

Applying the Bochner-Weitzenbck formula to (A, ϕ), a solution of Hitchin's equations, taking the inner product with $\phi_y dy$, and integrating, we get
$$\frac{1}{2}\Delta|\phi_y|^2 \geq \lambda|\phi_y|^2(\lambda + |\phi_y|^2).$$

By the application of the maximum principal, we see that $0 \leq |\phi_y|^2 \leq f_\lambda$ for $f_\lambda > 0$ a solution of $\left(\frac{1}{2}\Delta - \lambda\right) f_\lambda = 0$. Consequently, get bounds on $\int_B |\phi|^2$ (and also on $\int_{B'} |F_A|^2$) from bounds on $\int_B f_\lambda$ away from the zero locus, Z. This allows us to choose a good gauge. These bounds also allows us to say that the subsets $\{(A, \phi) : \int_M |\phi|^2 \leq K\}$ are compact for all K. For Kapustin-Witten we have no such results. ◁

3.6 Future Hopes for Kapustin-Witten Equations

- Knot Invariants

- Invariants coming from what the TFT attaches to 2-manifolds and 3-manifolds.

- Mostow Rigidity

- Applications to number theory through relationship between of Hitchin's equations and the Geometric Langlands program.