YANG-MILLS ON RIEMANN SURFACES

JEFFREY JIANG

These are notes I've made while working through the paper *The Yang-Mills Equations over Riemann Surfaces* [1] by Atiyah and Bott.

Contents

1. Preliminary Setup

To discuss the Yang-Mills functional, we must first fix some data. The setup will consist of the following ingredients

(1) A compact manifold M.
(2) A compact connected Lie group G.
(3) A principal G-bundle $P \to M$.

With this data, we have two associated bundles

$$\begin{aligned} \operatorname{Ad} P &:= P \times_G G \\ \mathfrak{g}_P &:= P \times_G \mathfrak{g} \end{aligned}$$

Where the action of G on G is by conjugation, and the action of G on $\mathfrak{g}$ is the adjoint action. We note that these bundles both contain additional structure – $\operatorname{Ad} P$ is a bundle of groups (not a principal bundle), and $\mathfrak{g}_P$ is a bundle of Lie algebras. The space of sections $\Gamma(M, \operatorname{Ad} P)$ has a natural group structure given by pointwise multiplication, and is called the ***gauge group*** $\mathscr{G}(P)$. Likewise, the space of sections $\Gamma(M, \mathfrak{g}_P)$ has a natural Lie algebra structure given by the pointwise Lie bracket, and can be naturally identified with the Lie algebra of $\mathscr{G}(P)$, as we shall see. An alternate characterization of these spaces of sections comes from a general characterization of sections of associated bundles

Proposition 1.1. *We have natural correspondences*

$$\Gamma(M, \mathrm{Ad}\, P) \longleftrightarrow \left\{ f : P \to G \ : \ f(p \cdot g) = g^{-1} f(p) g \right\}$$
$$\Gamma(M, \mathfrak{g}_P) \longleftrightarrow \left\{ f : P \to \mathfrak{g} \ : \ f(p \cdot g) = \mathrm{Ad}_{g^{-1}} f(p) \right\}$$

From the above correspondence we a group isomorphism $\mathscr{G}(P) \to \mathrm{Aut}(P)$, where $\mathrm{Aut}(P)$ denotes the group of G-equivariant diffeomorphisms $\varphi : P \to P$ such that the following diagram commutes :

$$\begin{tikzcd} P \arrow[rr, "\varphi"] \arrow[dr] & & P \arrow[dl] \\ & M & \end{tikzcd}$$

The isomorphism is given by mapping a G-equivariant map $f : P \to g$ to the automorphism $\varphi_f : P \to P$ defined by $\varphi_f(p) = p \cdot f(p)$. We get a similar identification for $\Gamma(M, \mathfrak{g}_P)$. Regarding a section $\phi \in \Gamma(M, \mathfrak{g}_P)$ as a G-equivariant map $P \to \mathfrak{g}$, we get a bundle automorphism $\exp(\phi)$ defined by $p \mapsto p \cdot \exp(\phi(p))$. From this perspective, we see that $\Gamma(M, \mathfrak{g}_P)$ can be identified with the G-invariant vertical vector fields on P, which are the infinitesimal generators of the action of $\mathscr{G}(P)$ on P.

In the the case of $\Gamma(M, \mathfrak{g}_P)$, we can extend the correspondence to the spaces of $\mathfrak{g}_P$ valued forms. The kernel of the differential of the projection $P \to M$ gives a subbundle of TP, which has a natural identification with the trivial bundle $\underline{\mathfrak{g}} = P \times \mathfrak{g}$. Then we can identify the space of sections of $\Lambda^k T^* M \otimes \mathfrak{g}_P$, (i.e. the space $\Omega^k_M(\mathfrak{g}_p)$ of $\mathfrak{g}_p$ valued k-forms with a subspace of the space $\Omega^k_P(\mathfrak{g})$ of $\mathfrak{g}$-valued k-forms ω on P satisfying :

1. $R_g^* \omega = \mathrm{Ad}_{g^{-1}} \omega$, where $R_g : P \to P$ denotes the right action of $g \in G$.
2. $\iota_\xi \omega = 0$ for any $\xi \in \mathfrak{g}$, where ι denotes interior multiplication, and we identify ξ with the constant vector field ξ under the identification of the vertical space with $\mathfrak{g}$.

We have maps

$$\Omega^p_M(\mathfrak{g}_p) \otimes \Omega^q_M(\mathfrak{g}_P) \to \Omega^{p+q}_M(\mathfrak{g}_p \otimes \mathfrak{g}_p)$$
$$(\omega_1 \otimes \xi_1) \otimes (\omega_2 \otimes \xi_2) \mapsto (\omega_1 \wedge \omega_2) \otimes (\xi_1 \otimes \xi_2)$$

From now on, we will usually omit the tensor symbol for $\mathfrak{g}_P$-valued forms in favor of juxtaposition, i.e. we write $\omega\xi$ instead of $\omega \otimes \xi$. Using the Lie bracket, we then get

$$\Omega^p_M(\mathfrak{g}_p) \otimes \Omega^q_M(\mathfrak{g}_P) \to \Omega^{p+q}_M(\mathfrak{g}_p)$$
$$\omega \otimes \eta \mapsto [\omega, \eta]$$

We note that this is *not* skew-symmetric, instead, given $\omega \in \Omega^p_M(\mathfrak{g}_P)$ and $\eta \in \Omega^q_M(\mathfrak{g}_P)$, we have

$$[\omega, \eta] = (-1)^{pq+1} [\eta, \omega]$$

For any semisimple Lie group G (in particular, for any compact Lie group G), we have an inner product $\langle \cdot, \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ that is invariant under the adjoint action (e.g. the Killing form). Invariance under the adjoint action gives us

$$\begin{aligned}\langle[\xi_1,\xi_2],\xi_3\rangle &= \langle[-\xi_2,\xi_1],\xi_3\rangle \\ &= \frac{d}{dt}\bigg|_{t=0}\langle \mathrm{Ad}_{\exp(-t\xi_2)}\,\xi_1,\xi_3 \\ &= \frac{d}{dt}\bigg|_{t=}\langle \mathrm{Ad}_{\exp(t\xi_2)}\,\mathrm{Ad}_{\exp(-t\xi_2)}\,\xi_1, \mathrm{Ad}_{\exp(t\xi_2)}\,\xi_3\rangle \\ &= \langle\xi_1,[\xi_2,\xi_3]\rangle\end{aligned}$$

Fixing one such inner product induces a fiber product on the trivial bundle $P\times\mathfrak{g}$, and invariance guarantees that this descends to a fiber product on $\mathfrak{g}_P$. This give us pairings

$$\begin{aligned}\Omega^p_M(\mathfrak{g}_P)\otimes\Omega^q_M(\mathfrak{g}_P) &\to \Omega^{p+q}_M \\ \omega\otimes\eta &\mapsto \langle\omega,\eta\rangle\end{aligned}$$

which also satisfies the identity

$$\langle[\omega,\eta],\xi\rangle = \langle\omega,[\eta,\xi]\rangle$$

We again note that this is not symmetric or skew symmetric, and instead behaves like the wedge product, i.e. for $\omega\in\Omega^p_M(\mathfrak{g}_P)$ and $\eta\in\Omega^q_M(\mathfrak{g}_P)$, we have

$$\langle\omega,\eta\rangle = (-1)^{pq}\langle\eta,\omega\rangle$$

which can be seen by writing $\omega=\omega^i\xi_i$ and $\eta=\eta^i\xi_i$ for an orthonormal basis $\{\xi_i\}$ for $\mathfrak{g}$. We then fix an orientation and Riemannian metric on M, which gives us a Hodge star operator $\star:\Omega^p_M\to\Omega^{n-p}_M$ and a Riemannian volume form dV_g. The Hodge star extends to $\mathfrak{g}_P$-valued k-forms, where given $\omega\in\Omega^p_M$ and $\xi\in\Gamma(M,\mathfrak{g}_P)$, we define $\star(\omega\xi)=(\star\omega)\xi$. Then given $\omega_1\xi_1,\omega_2\xi_2\in\Omega^p_M(\mathfrak{g}_P)$, we have

$$\langle\omega_1\xi_1,\star\omega_2\xi_2\rangle = \langle\omega_1,\omega_2\rangle_g\langle\xi_1,\xi_2\rangle$$

where $\langle\cdot,\cdot\rangle_g$ denotes the fiber metric on Λ^pT^*M induced by g. This gives us an inner product on each $\Omega^p_M(\mathfrak{g}_P)$ defined by

$$(\theta,\varphi) = \int_M\langle\theta,\star\varphi\rangle$$

Which gives us the L^2 norm on $\Omega^p_M(\mathfrak{g}_P)$ with $\|F\|^2_{L^2}=(F,F)$.

Definition 1.2. A connection on a principal bundle $\pi:P\to M$ is a choice of G-invariant splitting of the exact sequence of vector bundles over P

$$0\longrightarrow\underline{\mathfrak{g}}\longrightarrow TP\longrightarrow\pi^*TM\longrightarrow 0$$

i.e. a distribution $H\subset TP$ such that

(1) $(R_g)_*H_p=H_{p\cdot g}$
(2) $H\oplus\underline{\mathfrak{g}}=TP$

Equivalently, it is the data of a $\mathfrak{g}$-valued 1-form $A\in\Omega^1_P(\mathfrak{g})$ satisfying

(1) $R_g^*A=\mathrm{Ad}_{g^{-1}}A$
(2) $\iota_\xi A=\xi$ for all $\xi\in\mathfrak{g}$.

Note in particular that by a dimension count, we have that $\pi_*|_H : H \to TM$ is an isomorphism. This implies that given a tangent vector v at x and a point $p \in P$ in the fiber over x, we get a unique horizontal lift $\tilde{v} \in H_p$. For a fixed principal G-bundle $\pi : P \to M$, we let $\mathscr{A}(P)$ denote the space of all connections on P, which is an affine space over $\Omega^1_M(\mathfrak{g}_P)$. The connection form A on P induces an exterior covariant derivative on any associated vector bundle $E = P \times_G V$ arising from a linear representation $\rho : G \to \mathrm{GL}(V)$. Let $\dot{\rho} : \mathfrak{g} \to \mathrm{End}(V)$ be the derivative of ρ at the identity. Then the exterior covariant derivative is given by

$$\begin{aligned} d_A : \Omega^p_M(E) &\to \Omega^{p+1}_M(E) \\ \psi &\mapsto d\psi + \dot{\rho}(A) \wedge \psi \end{aligned}$$

In particular, we get an exterior covariant derivative on $\mathfrak{g}_P$, which is given by

$$d_A\psi = d\psi + [A, \psi]$$

Proposition 1.3. *Let $\phi \in \Omega^p_M(\mathfrak{g}_P)$ and $\psi \in \Omega^q_M(\mathfrak{g}_P)$. Then*

$$d\langle \phi, \psi \rangle = \langle d_A\phi, \psi \rangle + (-1)^p \langle \phi, d_A\psi \rangle$$

Proof. We compute

$$\begin{aligned} \langle d_A\phi, \psi \rangle + (-1)^p \langle \phi, d_A\psi \rangle &= \langle d\phi, \psi \rangle + \langle [A, \phi], \psi \rangle + (-1)^p (\langle \phi, d\psi \rangle + \langle \phi, [A, \psi] \rangle) \\ &= \langle d\phi, \psi \rangle + \langle [A, \phi], \psi \rangle + (-1)^p (\langle \phi, d\psi \rangle + \langle [\phi, A], \psi \rangle) \\ &= \langle d\phi, \psi \rangle + \langle [A, \phi], \psi \rangle + (-1)^{2p+1} \langle [A, \phi], \psi + (-1)^p \langle \phi, d\psi \rangle \\ &= \langle d\phi, \psi \rangle + (-1)^p \langle \phi, d\psi \rangle \end{aligned}$$

Writing $\phi = \phi^i \xi_i$ and $\psi = \psi^i \xi_i$ in an orthonormal basis $\{\xi_i\}$ for $\mathfrak{g}$, this becomes

$$\begin{aligned} \langle d_A\phi, \psi \rangle + (-1)^p \langle \phi, d_A\psi \rangle &= \sum_i d\phi^i \wedge \psi^i + (-1)^p \phi^i \wedge d\psi^i \\ &= d\langle \phi, \psi \rangle \end{aligned}$$

■

Given any distribution $E \subset TP$, we get a Frobenius tensor $\phi_E : E \otimes E \to TP/E$ given by $X \otimes Y \to [X, Y] \mod E$ where we extend X and Y to local vector fields. The Frobenius tensor should be thought of as the obstruction to the existence of an integral submanifold for the distribution E. In the case of a connection H on a principal bundle $P \to M$, we can extend to all of TP by first projecting onto H, and have an identification of $TP/H \cong \underline{\mathfrak{g}}$, and the Frobenius tensor is given by $X \otimes Y \mapsto A([X, Y])$, where A is the connection 1-form, and is called the ***curvature form*** of the connection, and is denoted F_A. In terms of differential forms, we have that for horizontal vectors ξ_1, ξ_2 on TP,

$$dA(\xi_1, \xi_2) = \xi_1 A(\xi_2) - \xi_2 A(\xi_1) - A([\xi_1, \xi_2])$$

The fact that ξ_1 and ξ_2 are horizontal implies that they are in the kernel of A, which gives us $dA(\xi_1, \xi_2) = -F_A(\xi_1, \xi_2)$. We also know that F_A vanishes on vertical vectors, and since

$A(X) = X$ for $X \in \mathfrak{g}$, we get that

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

It can be shown that F_A transforms by the adjoint action under pullback, and vanishes on vertical vectors, so it descends to a $\mathfrak{g}_P$-valued 2-form on the base manifold M.

Another thing to note is that there is a natural right action of the gauge group $\mathscr{G}(P)$ on the space of connections $\mathscr{A}(P)$. Interpreting the elements of $\mathscr{G}(P)$ as bundle automorphisms $\varphi : P \to P$ and elements of $\mathscr{A}(P)$ as $\mathfrak{g}$-valued 1-forms A on P, the action is simply pullback, $(\varphi, A) \mapsto \varphi^* A$. To show that this defines an action, we must check that $\varphi^* A$ satisfies the conditions

(1) $R_g^* \varphi^* A = \mathrm{Ad}_{g^{-1}} \varphi^* A$
(2) $\iota_\xi \varphi^* A = \xi$ for all $\xi \in \mathfrak{g}$.

Which are all simple consequences of the G-equivariance of φ and the transformation law for A. For a specific formula, let $\varphi : P \to P$ be an element of the gauge group, and let $g_\varphi : P \to G$ be its associated G-equivariant map. Then

$$\varphi^* A = \mathrm{Ad}_{g_\varphi^{-1}} A + g_\varphi^* \theta$$

where $\theta \in \Omega^1_G(\mathfrak{g})$ denotes the ***Maurer-Cartan form***

$$\theta_g(v) = (dL_{g^{-1}})_g(v)$$

which satisfies the ***Maurer-Cartan equation***

$$d\theta + \frac{1}{2}[\theta, \theta] = 0$$

Proposition 1.4. *Let $A \in \mathscr{A}(P)$ be a connection and $\varphi : P \to P$ an element of $\mathscr{G}(P)$ with associated G-equivariant map $g_\varphi : P \to G$. Then*

$$F_{\varphi^* A} = \mathrm{Ad}_{g_\varphi^{-1}} F_A$$

Proof. Using the transformation law for $\varphi^* A$ we compute

$$\begin{aligned} F_{\varphi^* A} &= d(\mathrm{Ad}_{g_\varphi^{-1}} A + g_\varphi^* \theta) + \frac{1}{2}[\mathrm{Ad}_{g_\varphi^{-1}} A + g_\varphi^* \theta, \mathrm{Ad}_{g_\varphi^{-1}} A + g_\varphi^* \theta] \\ &= \mathrm{Ad}_{g_\varphi^{-1}} dA + g_\varphi^* d\theta + \frac{1}{2}\Big([\mathrm{Ad}_{g_\varphi^{-1}} A, \mathrm{Ad}_{g_\varphi^{-1}} A] + [\mathrm{Ad}_{g_\varphi^{-1}} A, g_\varphi^* \theta] + [g_\varphi^* \theta, \mathrm{Ad}_{g_\varphi^{-1}}] + [g_\varphi^* \theta, g_\varphi^* \theta]\Big) \\ &= \mathrm{Ad}_{g_\varphi^{-1}} dA + \frac{1}{2}[\mathrm{Ad}_{g_\varphi^{-1}} A, \mathrm{Ad}_{g_\varphi^{-1}} A]] \end{aligned}$$

Where we use skew-symmetry and the Maurer-Cartan equation. ∎

We similarly compute the infinitesimal action of the Lie algebra $\Gamma(M, \mathfrak{g}_P)$.

Proposition 1.5. *The vector field corresponding to $\phi \in \Gamma(M, \mathfrak{g}_P)$ is $A \mapsto d_A \phi \in \Omega^1_M(\mathfrak{g}_P)$*

[1] Our convention for the sign of the curvature is opposite from many other conventions, which usually sets $F_A = dA + \frac{1}{2}[A, A]$

Proof. We compute the vector field at a connection $A \in \mathscr{A}(P)$ to be

$$\begin{aligned}
&\frac{d}{dt}\bigg|_{t=0} \operatorname{Ad}_{\exp(t\phi)^{-1}} A + \exp(t\phi)^*\theta = -[\phi, A] + \frac{d}{dt}\bigg|_{t=0} (dL_{\exp(-t\phi)} d(\exp(t\phi))) \\
&= [A,\phi] + \left(\frac{d}{dt}\bigg|_{t=0} dL_{\exp(-t\phi)}\right) d(\exp(0)) + dL_{\exp(0)} \left(\frac{d}{dt}\bigg|_{t=0} d(\exp(t\phi))\right) \\
&= [A,\phi] + d\phi \\
&= d_A\phi
\end{aligned}$$

where for the third equality we use the product rule, and in the fourth equality we use the fact that $\exp(0) = \mathrm{id}$ and that the derivative of $\exp(t\phi)$ as $t \to 0$ is ϕ. ∎

For any other connection $A + \eta$ with $\eta \in \Omega^1_M(\mathfrak{g}_P)$, a quick computation yields

$$F_{A+\eta} = F_A + \frac{1}{2}[\eta, \eta] + d_A\eta$$

From this description, we can relate the curvature F_A with the covariant derivative. Note that for the line of connections $A + t\eta$, we have that

$$\frac{d}{dt}\bigg|_{t=0} F_{a+t\eta} = \frac{d}{dt}\bigg|_{t=0} F_A + \frac{t^2}{2}[\eta,\eta] + t d_A\eta = d_A\eta$$

So $d_A\eta$ measures the infinitesimal change of the curvature F_A in the direction η.

2. The Yang-Mills Functional

With the setup done, we have the ingredients necessary to define the Yang-Mills functional.

Definition 2.1. The ***Yang-Mills functional*** is the map $L : \mathscr{A}(P) \to \mathbb{R}$ given by

$$L(A) = \|F_A\|^2_{L^2} = \int_M \langle F_A, \star F_A \rangle$$

We immediately see that the Yang-Mills functional is invariant under $\mathscr{G}(P)$ in the following sense – for any gauge transformation φ we have that $L(\varphi^* A) = L(A)$, which follows immediately from the invariance of $\langle \cdot, \cdot \rangle$ and the transformation law for curvature.

Our goal now will be to find the Euler-Lagrange equations for the Yang-Mills functional by computing the first and second variations. Using the Hodge star operator, we construct the formal adjoint with respect to the inner product $d_A^* : \Omega^p_M(\mathfrak{g}_P) \to \Omega^{p-1}_M(\mathfrak{g}_P)$ in the same manner as for classical Hodge theory on a Riemannian manifold. Explicitly, the formula on p-forms is given by

$$d_A^* = (-1)^{n(p+1)+1} \star d_A \star$$

where $n = \dim M$. We then compute the first variation of L.

Proposition 2.2 (***The First Variation***)**.** *For a local extremum $A \in \mathscr{A}(P)$ of the Yang-Mills functional, we have*

$$d_A \star F_A = 0$$

The local extremum connection A is then called a ***Yang-Mills connection***, *and the space of Yang-Mills connections is denoted $\mathscr{A}_{\mathrm{YM}}(P)$.*

Proof. Consider a variation $A + t\eta$ with $t \in \mathbb{R}$ and $\eta \in \Omega^1_M(\mathfrak{g}_P)$. We have that the curvature is given by

$$F_{A+t\eta} = F_A + \frac{t^2}{2}[\eta, \eta] + t d_A \eta$$

This then gives us

$$\begin{aligned} \|F_{A+t\eta}\|_{L^2} &= \int_M \langle F_{A+t\eta}, F_{A+t\eta} \rangle \\ &= \int_M \langle F_A + \frac{t^2}{2}[\eta, \eta] + t d_A \eta, \star(F_A + \frac{t^2}{2}[\eta, \eta] + t d_A \eta) \rangle \end{aligned}$$

Expanding this out, we get that the term that is linear in t is

$$\int_M \langle F_A, \star d_A \eta \rangle + \langle d_A \eta, \star F_A \rangle = 2(F_A, d_A \eta)$$

where we use symmetry of $(\cdot, \cdot)$. Since A is extremal, we have that this term must vanish, giving us that $(F_A, d_A \eta) = (d_A^* F, \eta) = 0$ for every η. Then since we have (up to sign) $d_A^* = \star d_A \star$, and $\star$ is an isomorphism, this implies $d_A \star F_A = 0$. ∎

Proposition 2.3 (***The Second Variation***)**.** *At a Yang-Mills connection $A \in \mathscr{A}(P)$, we have*

$$d_A^* d_A \eta + \star[\eta, \star F_A] = 0$$

Proof. We differentiate the first variational equation with respect to t, i.e. we compute

$$\left.\frac{d}{dt}\right|_{t=0} d^*_{A+t\eta} F_{A+t\eta}$$

We expand out

$$\begin{aligned} d^*_{A+t\eta} F_{A+t\eta} &= \pm \star d_{A+t\eta} \star F_{A+t\eta} \\ &= \pm \left(\star d_A \star \left(F_A + t d_A \eta + \frac{t^2}{2}[\eta, \eta] \right) + t \star \left[\eta, \star \left(F_A + t d_A \eta + \frac{t^2}{2}[\eta, \eta] \right) \right] \right) \end{aligned}$$

Taking the term linear in t yields

$$\pm \left(\star d_A \star d_A \eta + \star[\eta, \star F_A] \right)$$

Giving us that at an extremal connection A, we have

$$d_A^* d_A \eta + \star[\eta, \star F_A] = 0$$

∎

The second variation should be thought of as the Hessian to the Yang-Mills functional, which will allow us to apply Morse theory techniques to the space of connections.

3. The $U(1)$ Case

We first restrict to the special case $G = U(1)$. In this case, the Lie algebra is abelian, so the adjoint action of $U(1)$ on $\mathfrak{u}(1)$ is trivial, giving us that $\mathfrak{g}_P$ is a trivial bundle. Identifying $\mathfrak{u}(1)$ with $\mathbb{R}$, we can then identify $\mathfrak{u}(1)$ valued forms on P with ordinary differential forms. Likewise, using triviality of $\mathfrak{g}_P$, we can identify $\mathfrak{g}_P$-valued forms with ordinary differential forms on M. The vertical bundle in this case is a trivial line bundle over P, and there is a unique $U(1)$-invariant vertical vector field on P, which on each fiber restricts to the vector field dual to the Maurer-Cartan form θ. Then given a connection A on P, we have that $dA = \pi^* F_A$, since $[A, A] = 0$. This immediately tells us that F_A is closed, since d commutes with pullback. Furthermore, for any other connection $A + \eta$, we have that

$$F_{A+\eta} = F_A + \frac{1}{2}[\eta, \eta] + d_A \eta$$

Then since $d_A = d$ and $[\eta, \eta] = 0$, this gives us that $F_{A+\eta} = F_A + d\eta$, which tells us that the cohomology class of F_A is independent of our choice of A. Using our sign convention, this is equal to $-2\pi i c_1(P)$. In addition, it tells us that every representative of the curvature class can be realized as the curvature of some connection. Furthermore, in this situation, the Yang-Mills functional reduces to the standard Hodge theory picture, since the L^2 norm will coincide with the L^2 norm on differential forms. Therefore, a Yang-Mills connection on P is equivalent to finding the unique connection whose curvature minimizes the L^2 norm in the cohomology class $2\pi i c_1(P)$. By standard Hodge theory, there exists a unique harmonic representative of the curvature class, and the Yang-Mills connections for P are a torsor over the space Z^1_M of closed 1-forms. In total, this gives us the fibration

$$\begin{tikzcd} Z^1_M \arrow[r] & \mathscr{A}_{\mathrm{YM}}(P) \arrow[d] \\ & 2\pi i c_1(P) \end{tikzcd}$$

In the flat case $c_1(P) = 0$, the Yang-Mills connections are just flat connections, which are parameterized by conjugacy classes of homomorphisms $\pi_1(M) \to U(1)$.

With this information, we can construct the Yang-Mills moduli space $\mathscr{A}_{\mathrm{YM}}(P)/\mathscr{G}(P)$ in this simple case. Since the conjugation action is trivial, the bundle $\mathrm{Ad}(P)$ is a trivial bundle, so the gauge group is just $\mathscr{G}(P) = \mathrm{Map}(M, U(1))$. Given $f : M \to U(1)$, the action of f on a connection A is given by

$$A \mapsto A + \pi^* f^* \theta$$

The gauge group acts on Z^1_M in the same way, so if we fix some reference Yang-Mills connection A_0 to identify $\mathscr{A}_{\mathrm{YM}}(P)$ with Z^1_M, the group actions are identified. Therefore, it suffices to compute the quotient of $Z^1_M/\mathscr{G}(P)$. To compute this quotient, we do it in two steps, first quotienting by the identity component $\mathscr{G}_0(P)$, and then quotienting by the component group $\pi_0\mathscr{G}(P)$. The components of $\mathscr{G}(P)$ are simply the homotopy classes of maps $M \to S^1$, so $\mathscr{G}_0(P)$ consists of all nullhomotopic maps $M \to S^1$. Let dx denote the standard form on $\mathbb{R}$. Then any nullhomotopic map $f \in \mathscr{G}_0(P)$ lifts to a function $\tilde{f} : M \to \mathbb{R}$ that exponentiates to f. Since the Maurer-Cartan form on S^1 pulls back to dx, we find that

$f^*\theta = d\tilde{f}$. Therefore, the action of $\mathscr{G}_0(P)$ on Z^1_M is given by the addition of exact 1-forms, giving us that the quotient $Z^1_M/\mathscr{G}_0(P)$ is $H^1(M,\mathbb{R})$. Then since S^1 is a $K(\mathbb{Z},1)$, we know that homotopy classes of maps $M \to S^1$ are in bijection with $H^1(M,\mathbb{Z})$. Therefore, by quotieting by $\mathscr{G}_0(P)$ and putting everything together, we get isomorphism

$$\mathscr{A}_{\mathrm{YM}}(P)/\mathscr{G}(P) \cong Z^1_M/\mathscr{G}(P) \cong \mathrm{Jac}(M) := H^1(X,\mathbb{R})/H^1(X,\mathbb{Z})$$

which is isomorphic to a torus $\mathbb{T}^{b_1(M)}$. However, we note that in general, this isomorphism is highly non-canonical – to make the identification of $\mathscr{A}_{\mathrm{YM}}(P)$ with Z^1_M, we need to fix a reference connection. In general, there is no canonical choice of reference except in the case where P is a trivial bundle, in which case the trivial connection defines a canonical reference connection. As we'll see soon, this reflects the fact that $\mathscr{A}_{YM}(P)$ is a $\mathrm{Jac}(M)$-torsor when P is a $U(1)$-bundle.

4. Yang-Mills Over a Riemann Surface

We now restrict our attention to when M is a orientable surface, with genus $g > 0$. Let $Q \to M$ be a principal $U(1)$ bundle with $c_1(Q) = 1$, i.e.

$$\frac{1}{2\pi i}\int_M c_1(Q) = 1$$

Then fix a Riemannian metric on M with volume form ω such that $\int_M \omega = 1$ and Yang-Mills connection A on Q. Since $c_1(Q) = 1$, we have that $[c_1(Q)] = [\omega]$. Furthermore, since $\star\omega = 1$ and $-2\pi i c_1(Q) = [F_A]$, we get that the curvature of A must be equal to $-2\pi i\omega$ to minimize the Yang-Mills functional. Similarly, for any other $U(1)$-bundle P, the curvature of a Yang-Mills connection must be $-2\pi i c_1(P)\omega$. Then let $\widetilde{M} \to M$ be the universal cover of M. Since the genus of M is at least 1, $\widetilde{M}$ is contractible, so the pullback of Q along the covering projection gives us a trivial $U(1)$ bundle over $\widetilde{M}$

$$\begin{array}{ccc} \widetilde{M}\times U(1) & \longrightarrow & Q \\ \downarrow & & \downarrow \\ \widetilde{M} & \longrightarrow & M \end{array}$$

Then we have a covering map $\widetilde{M}\times\mathbb{R} \to U(1)$, using the usual covering $\mathbb{R} \to U(1)$, giving a principal $\mathbb{R}$-bundle over $\widetilde{M}$. Then if we consider the composite map $\widetilde{M}\times\mathbb{R} \to \widetilde{M} \to M$, this is a fiber bundle over M. Furthermore, since the action of $\mathbb{R}$ on on $\widetilde{M}\times\mathbb{R}$ commutes with the $\pi_1(M)$ action on $\widetilde{M}$, we know that this is a principal bundle with structure group $\Gamma_{\mathbb{R}}$, where $\Gamma_{\mathbb{R}}$ is a central extension of $\pi_1(M)$ by $\mathbb{R}$. Let J denote the element of $\mathbb{R} \subset \Gamma_{\mathbb{R}}$ corresponding to $1 \in \mathbb{R}$. Then consider M as the quotient of the $2g$-gon. The holonomy about the path traversed by the boundary is exactly the product $\prod_i [a_i, b_i]$ of the commutators of representatives of generators of $\pi_1(M)$, and has holonomy equal to 2π, which follows from the fact that the holonomy about any loop bounding a disk is equal to the integral of the curvature, and the fact that $c_1(Q)$ is represented by the curvature class of any connection, divided by $2\pi i$. Therefore, if we consider the pullback connection on $\widetilde{M}\times U(1)$, the holonomy about the lifts of the boundary path to $\widetilde{M}$ will also be 2π, which then lifts to translation by 1 in the bundle $\widetilde{M}\times\mathbb{R}$. This gives the relation that $\prod_i [a_i, b_i] = J$, which gives us a presentation of the group $\Gamma_{\mathbb{R}}$. We let $\Gamma_{\mathbb{Z}}$ denote the central extension obtained in a similar manner using $\mathbb{Z}$ instead of $\mathbb{R}$. Since $\pi_1(M)$ is discrete, $\widetilde{M}$ is a flat bundle, so the pullback connection on $\widetilde{M} \times U(1)$ still has curvature $-2\pi i\omega$, and the curvature also remains unchanged after lifting to $\widetilde{M} \times \mathbb{R}$.

Suppose we have a homomorphism $\rho : \Gamma_{\mathbb{R}} \to G$ to a compact group G. This then gives us an associated bundle $P = (\widetilde{M} \times \mathbb{R}) \times_{\Gamma_{\mathbb{R}}} G$, which is a principal G bundle In addition to ρ, we get a Lie algebra homomorphism $\dot{\rho} : \mathbb{R} \to \mathfrak{g}$. Using this, $\dot{\rho}(A) \in \Omega^1_P(\mathfrak{g})$ determines a connection on P, which has curvature $\dot{\rho}(F_A)$. Furthermore, we have that

$$\dot{\rho}(d_A \star F_A) = d_{\dot{\rho}(A)} \star \dot{\rho}(F_A)$$

which tells us that $\dot{\rho}(A)$ is a Yang-Mills connection on P. The main theorem is that every Yang-Mills connection on every principal bundle arises in this way.

Theorem 4.1. *The above construction gives a bijective correspondence*

$$\mathsf{Hom}(\Gamma_{\mathbb{R}}, G)/G \longleftrightarrow \{G\text{-bundles } P \to M \text{ equipped with a Yang-Mills connection}\} \ / \sim$$

Where the action of G on $\mathsf{Hom}(\Gamma_{\mathbb{R}}, G)$ is by conjugation, and the equivalence relation $\sim$ is given by $(P_1, A_1) \sim (P_2, A_2)$ if there exists an isomorphism $\varphi : P_1 \to P_2$ of principal bundles such that $\varphi^* A_2 = A_1$.

Proof. We first give an outline of the proof strategy, since the proof is rather long and involved. Before that, we note that we have not fixed a principal bundle, so a homomorphism $\Gamma_{\mathbb{R}} \to G$ must provided us with the data of a principal G-bundle $P \to M$, as well as a Yang-Mills connection on P. Our strategy will be as follows:

(1) For any principal G-bundle P with a Yang-Mills connection, use $\star F_A$ to identify an orbit of the adjoint action on $\mathfrak{g}$. Then reduce to the case where $\star F_A$ takes the constant value $X \in \mathfrak{g}$, where X is fixed by the adjoint action.

(2) Pass the case where we can replace G with a quotient group $\overline{G}$ that is the product of a torus and a semisimple group, which are easy cases that are easily established.

Since M is 2-dimensional, we have that $\star F_A \in \Omega^0_M(\mathfrak{g}_P)$, so we may regard it as a G-equivariant map $P \to \mathfrak{g}$, i.e. $\star F_A(p \cdot g) = \mathrm{Ad}_{g^{-1}} \star F_A(p)$. This then tells us that the image of $\star F_A$ is exactly one orbit of $\mathfrak{g}$ under the adjoint action. Fix a nonzero element $X \in \mathfrak{g}$ lying in the image of $\star F$, and then consider the preimage $P_X := (\star F_A)^{-1}(X)$. Let $G_X \subset G$ be the stabilizer of X under the adjoint action. Then G_X acts on P_X, since given any $p \in P_X$, and $g \in G_X$, we have that $\star F_A(p \cdot g) = \mathrm{Ad}_{g^{-1}} F_A(p) = X$. This action is clearly transitive and free, so P_X defines a reduction of structure group from G to G_X, giving us a bundle isomorphism $P_X \times_{G_X} G \cong P$. Furthermore, since $d_A \star F = 0$, we have that $\star F$ is constant in the horizontal directions, so the differential of $\star F$ vanishes in the horizontal directions, so the horizontal distribution is contained in the tangent bundle of P_X. Therefore, the connection A on P restricts to a Yang-Mills connection on P_X (which we also denote A). This restricted connection has the property that $\star F_A$ is the constant map with value $X \in \mathfrak{g}$, so F_A is just the 2-form $X \otimes \omega \in \Omega^2_M(\mathfrak{g}_{P_X})$. This tells us that every Yang-Mills connection on any bundle P arises from such a connection on the reduced bundle P_X for some $X \in \mathfrak{g}$. Then suppose we have a homomorphism $\rho : \Gamma_{\mathbb{R}} \to G$ with derivative $\dot{\rho} : \mathbb{R} \to \mathfrak{g}$. The image of $1 \in \mathbb{R}$ under $\dot{\rho}$ determines an element $X_\rho \in \mathfrak{g}$. Centrality of $\mathbb{R}$ in $\Gamma_{\mathbb{R}}$ then implies that the image of $\Gamma_{\mathbb{R}}$ preserves X_ρ under the adjoint action, so we may regard ρ as a homomorphism $\Gamma_{\mathbb{R}} \to G_{X_\rho}$. Combining this observation with the previous one, we can then reduce to the case where X is preserved by all of G (i.e. $G = G_X$, which is equivalent to X lying in the center of $\mathfrak{g}$).

We then want to reduce to the cases where G is either a torus or a semisimple group. This follows from the fact that any compact group G arises as $H \times_D S$, where H is a maximal torus and $S = [G, G]$ is the maximal connected semisimple subgroup, and $D = H \cap S$ is a finite subgroup of the center of S. Quotienting by D, we get a finite covering $G \to \overline{G} = \overline{H} \times \overline{S}$, where $\overline{H} = H/D$ and $\overline{S} = S/D$. We then claim that we can reduce to the case where the structure group is $\overline{G}$. To see this, we note that if we quotient P by the action of D to obtain $\overline{P}$, we get a finite sheeted covering $P \to \overline{P}$. Since this covering is a local diffeomorphism, we get an identification $TP \cong \pi^* T\overline{P}$ where $\pi : P \to \overline{P}$ is the covering projection. Therefore, we get that the horizontal distribution on P induces a horizontal distribution on $\overline{P}$, and conversely, we get that a connection on $\overline{P}$ lifts to a horizontal distribution on P, so we may reduce to the case $P = \overline{P}$. Then since $\overline{P}$ is a principal bundle where the structure group is a product group, $\overline{P}$ is isomorphic to a principal $\overline{H} \times \overline{K}$-bundle obtained by taking the fiber product of an $\overline{H}$-bundle with a $\overline{S}$-bundle, and the connection on $\overline{P}$ is equivalent to the data of a connection on each of these two bundles. Then since we assumed that the Lie algebra element $X \in \mathfrak{h} \oplus \mathfrak{s}$ is central, and the center of S is finite, we have that $X \in \mathfrak{h}$. Therefore, we see that a Yang-Mills connection for $\overline{P}$ is equivalent to a flat connection for a principal $\overline{S}$-bundle and a Yang-Mills connection for a principal $\overline{H}$-bundle. Furthermore, any map $\Gamma_{\mathbb{R}} \to G$ gives a map $\overline{\rho} : \Gamma_{\mathbb{R}} \to \overline{G}$ by composition with the quotient map. Then since we assume that $\mathring{\rho}(1) = X$ is central, we have that the image of $\mathbb{R} \subset \Gamma_{\mathbb{R}}$ is a 1-parameter subgroup of G lying in the center. Furthermore, we have that $\mathbb{Z}$ is contained in the commutator subgroup $[\Gamma_{\mathbb{R}}, \Gamma_{\mathbb{R}}]$, so its image under ρ is central in G and lies in the maximal semisimple subgroup $S = [G, G]$, so $\rho(\mathbb{Z}) \subset D$. Therefore, the map $\overline{\rho}$ descends to a map $\Gamma_{\mathbb{R}}/\mathbb{Z} \to \overline{G}$. We note $\Gamma_{\mathbb{R}}/\mathbb{Z} \cong U(1) \times \pi_1(M)$. Then since $\overline{G}$ is the product group $\overline{H} \times \overline{S}$, we have that $\overline{\rho}$ is given by a pair of homomorphisms $\alpha : U(1) \times \pi_1(M) \to \overline{H}$ and $\beta : U(1) \times \pi_1(M) \to \overline{S}$. Furthermore, since $\overline{S}$ is semisimple, it has a finite center, so centrality of X implies that β is trivial on the $U(1)$ factor, so it is actually a map $\beta : \pi_1(M) \to \overline{S}$, which is exactly the data of a principal $\overline{S}$-bundle with flat connection, which in particular, is a $\overline{S}$-bundle with Yang-Mills connection.

We are left to consider the homomorphism α, which amounts to understanding Yang-Mills connections when structure group is a torus $U(1) \times \cdots \times U(1)$. By passing to associated bundles, this is equivalent to passing to a direct sum of Hermitian line bundles, each equipped with a Yang-Mills connection. Therefore, it suffices to verify this for the case where the torus is just $U(1)$. In this case, the map ρ necessarily factors through $\pi_1(M)$, since $U(1)$ is abelian, and the derivative $\mathring{\rho}$ evaluated at the identity gives the first Chern class of the line bundle. Using our fixed reference bundle, this allows us to construct a Yang-Mills connection by tensoring our reference bundle tensored with itself $\mathring{\rho}(1)$ many times with the flat line bundle determined by the induced map $\pi_1(M) \to U(1)$. ■

While we have shown that the data of a principal G-bundle $P \to M$ with Yang-Mills connection is equivalent to a homomorphism $\Gamma_{\mathbb{R}} \to G$, we have not yet shown how to recover the isomorphism class of the principal bundle associated to such a homomorphism. Moreover, we have not yet shown that an arbitrary principal G-bundle admits a single Yang-Mills connection.

Theorem 4.2. *Every principal G-bundle admits a Yang-Mills connection.*

Proof. We use the fact that for a group $\overline{G}$, we have that principal $\overline{G}$-bundles over the surface M are classified by $H^2(M, \pi_1(G)) = \pi_1(G)$. In our case, by the reductions we made in the proof of the previous theorem, $\overline{G} = \overline{H} \times \overline{S}$, so we have

$$\pi_1(G) = \pi_1(\overline{H}) \oplus \pi_1(\overline{S})$$

By restricting to a copy of $U(1)$, the map α determines a class in $\pi_1(\overline{H}) \cong \mathbb{Z}^n$, which can be thought of as choosing the first Chern class for each direct summand of a vector bundle. For β, we need to do some additional work. The homomorphism $\beta : \pi_1(M) \to \overline{S}$ defines a group action of $\pi_1(M)$ on $\overline{S}$, which can be lifted to an action of $\Gamma_{\mathbb{Z}}$ on the universal cover of $\overline{S}$, giving a homomorphism of $\Gamma_{\mathbb{Z}}$ to the universal covering group. The image of the central element $J \in \Gamma_{\mathbb{Z}}$ in the universal cover is an element of the center (since we assumed that $G = G_X$), which is equivalent to an element of $\pi_1(\overline{S})$, since $\overline{S}$ is the quotient of the universal cover by a subgroup of the center. We then note that since $\overline{S}$ is compact and semisimple, every element of $\overline{S}$ is a commutator. Therefore, we can always find a homomorphism $\pi_1(M) \to \overline{S}$ that maps J to any element of the center $Z(\overline{S})$. Therefore, we can always find a map $\pi_1(M) \to \overline{S}$ whose lift to Γ_Z maps the central element J to any element of $\pi_1(\overline{S})$, considered as a subgroup of the center of the universal cover. In addition, it is clear that any element of $\pi_1(\overline{H})$ is realized as the restriction of a map $U(1) \times \pi_1(M) \to \overline{H}$. Therefore, every element of $\pi_1(\overline{G})$ can be realized via the maps α and β coming from the homomorphism $\overline{\rho} : \Gamma_{\mathbb{R}} \to \overline{G}$, so this shows that every isomorphism class of principal G-bundle can be obtained from a homomorphism $\Gamma_{\mathbb{R}} \to G$, and also shows that every principal G-bundle admits a Yang-Mills connection. ■

We now specialize to the case where $G = \mathrm{U}(n)$, which will be important for our study of holomorphic vector bundles over M. Given a principal $\mathrm{U}(n)$-bundle $P \to M$, we obtain a complex vector bundle $E \to M$ by taking the associated bundle corresponding to the defining representation $\mathrm{U}(n) \hookrightarrow \mathrm{GL}_n\mathbb{C}$. In the other direction, given a complex vector bundle $E \to M$, we obtain a principal $\mathrm{U}(n)$-bundle by fixing a Hermitian form on E and taking the unitary frame bundle of E. Let $P \to M$ be a principal $\mathrm{U}(n)$ bundle with Yang-Mills connection A, and let $X \in \mathfrak{u}_n$ be the Lie algebra element determined by the curvature of A. Writing $X = -2\pi i\Lambda$ for a Hermitian matrix Λ, the Yang-Mills condition implies that the trace of Λ is integral equal to the Chern class of P. Furthermore, using the eigenspace decomposition of X, we find that the stabilizer group G_X is a product of $U(n_i)$, where n_i is the multiplicity of the i^{th} eigenvalue λ_i. It can then be shown (though we don't show it here) that we must have that $n_i\lambda_i \in \mathbb{Z}$.

From the perspective of vector bundles, these constraints have a more natural interpretation. The reduction of structure group to G_X corresponds to a direct sum decomposition of the vector bundle, while the constraint on the eigenvalues corresponds to the Chern classes of the factors coinciding with the Chern class of the total bundle, which follows from the formula for the determinant line bundle of a direct sum of vector bundles.

5. The Holomorphic Viewpoint

Fix a smooth complex vector bundle $E \to M$ of rank n and first Chern class k. The data of a holomorphic structure on E is equivalent to the data of a first order differential operator $\overline{\partial}_E : \Omega^0_M(E) \to \Omega^1_M(E)$ satisfying $\overline{\partial}_E^2 = 0$, and the holomorphic sections of E are obtained by taking the kernel of $\overline{\partial}_E$. Much like connections, the any such operator may be written in a smooth local trivialization of E as

$$\overline{\partial}_E = \overline{\partial} + B$$

where $\overline{\partial}$ is the usual operator on $\mathbb{C}^n$ and B is $\mathrm{GL}_n\mathbb{C}$-valued $(0,1)$-form. This corresponds to the space of holomorphic structures being an affine space over $\Omega^{0,1}_M(\mathrm{End}\, E)$. The group $\mathrm{Aut}(E)$ of smooth bundle automorphisms of E acts on the space of holomorphic structures by conjugation, i.e. under an automorphism $g \in \mathrm{Aut}(E)$, the action is given by the mapping $\overline{\partial}_E \mapsto g\overline{\partial}_E g^{-1}$. Furthermore, the orbits of this group action are exactly the isomorphism classes of holomorphic structures on E, so the quotient by this action would be the moduli space of holomorphic structures on E. However, the quotient by this group action is not very nice, so we must restrict ourselves to structures that are stable in the sense of GIT.

Definition 5.1. The ***slope*** of a complex vector bundle $E \to M$ is defined to be

$$\mu(E) := \frac{c_1(E)}{\mathrm{rank}(E)}$$

where we use the orientation on M to determine the isomorphism $H^2(M,\mathbb{Z}) \cong \mathbb{Z}$.

Definition 5.2. A holomorphic vector bundle $E \to M$ is

(1) ***Stable*** if for all holomorphic subbundles $F \subset E$, we have the strict inequality $\mu(F) < \mu(E)$.
(2) ***Semistable*** if for all holomorphic subbundles $F \subset E$, we have inequality $\mu(F) \leq \mu(E)$.

If we let $\mathscr{C}(E)$ denote the space of holomorphic structures on E and let $\mathscr{C}_s(E)$ and $\mathscr{C}_{ss}(E)$ denote the subspace of stable and semistable holomorphic structures respectively, then we can construct the moduli space of of holomorphic vector bundles of rank n and degree k as

$$\mathcal{N}(n,k) := \mathscr{C}_{ss}(E)/\,\mathrm{Aut}(E)$$

In the case that n and k are coprime, we have that stability and semistability coincide, and this will be the main situation of interest.

Understanding the moduli space of semistable bundles will also give information regarding unstable holomorphic bundles. One reason for this is the existence of a canonical filtration of an arbitrary holomorphic vector bundle.

Theorem 5.3. *Let $E \to M$ be a holomorphic vector bundle. Then there exists a filtration*

$$0 = E_0 \subset E_1 \subset \cdots \subset E_m = E$$

of E such that E_{i+1}/E_i is semistable and we have

$$\mu(E_1/E_0) > \mu(E_2/E_1) > \cdots > \mu(E_n/E_{n-1})$$

6. The Symplectic Viewpoint

Even though the space $\mathscr{A}(P)$ is infinite dimensional, it has enough structure to be viewed as a symplectic "manifold," but we will gloss over the formal details. Because $\mathcal{A}(P)$ is an affine space over $\Omega^1_M(\mathfrak{g}_P)$, we may work with it as a manifold, where the tangent space at any point is $\Omega^1_M(\mathfrak{g}_P)$. We will be relatively cavalier with the details, though all we are doing can be made formal by passing to Sobolev completions of spaces of sections.

In the case that M is a Riemann surface, then the Hodge star $\star : \Omega^1_M(\mathfrak{g}_P) \to \Omega^1_M(\mathfrak{g}_P)$ can be viewed as a complex structure on $\mathscr{A}(P)$. In addition, after fixing a Riemannian metric on M, we get a trivialization $\Omega^2_M \cong \mathbb{R}$ using the natural orientation induced by the complex structure. This allows us to view the pairing

$$(\omega, \eta) \mapsto \int_M \langle \omega, \eta \rangle$$

as a symplectic form on $\mathscr{A}(P)$. In addition, these structures are visibly compatible, which gives $\mathscr{A}(P)$ a Kähler structure.

Recall that if we have a symplectic left action of a group G on a symplectic manifold (M, ω) we get an induced map $\mathfrak{g} \to \mathfrak{X}(M)$ mapping ξ to the vector field X_ξ defined by

$$(X_\xi)_p = \left.\frac{d}{dt}\right|_{t=0} \exp(t\xi) \cdot p$$

The action is ***Hamiltonian*** if for all $\xi \in \mathfrak{g}$ there exists a function $H_\xi : M \to \mathbb{R}$ called a ***Hamiltonian function*** such that the vector field X_ξ satisfies the identity

$$\omega_p((X_\xi)_p, v) = (dH_\xi)_p v$$

for all points $p \in M$ and tangent vectors $v \in T_pM$, and the mapping $\xi \mapsto H_\xi$ is G-equivariant with respect to the right actions $\xi \cdot g = \mathrm{Ad}_{g^{-1}} \xi$ and $f \cdot g = f \circ L_g$, where $L_g : M \to M$ is the symplectomorphism determined by left multiplication by g. Given a Hamiltonian action of G on M, a ***moment map*** for the action is a G-equivariant map $\mu : M \to \mathfrak{g}^*$ such that for any $p \in M$ and $\xi \in \mathfrak{g}$, we have

$$d\mu_p(v)(\xi) = \omega_p((X_\xi)_p, v)$$

in which case the Hamiltonian function is recovered by the formula

$$H_\xi(p) = \mu(p)(\xi)$$

We claim that the action of $\mathscr{G}(P)$ on $\mathscr{A}(P)$ is Hamiltonian. We first note that the action is symplectic, since $\langle \cdot, \cdot \rangle$ is Ad-invariant, and the action of a gauge transformation φ on a tangent vector $\eta \in \Omega^1_M(\mathfrak{g}_P)$ is by $\varphi \cdot \eta = \mathrm{Ad}_{g_\varphi^{-1}} \eta$ where $g_\varphi : P \to G$ is the associated G-equivariant map. To show that the action is Hamiltonian, we note that each $\phi \in \mathrm{Lie}(\mathscr{G}(P)) = \Omega^0_M(\mathfrak{g}_P)$ determines a map $H_\phi : \mathscr{A}(P) \to \mathbb{R}$ given by

$$H_\phi(A) = \int_M \langle F_A, \phi \rangle$$

and the mapping $\phi \mapsto H_\phi$ is clearly $\mathscr{G}(P)$-equivariant. We then compute for $A \in \mathscr{A}(P)$ and $\psi \in T_A\mathscr{A}(P) = \Omega^1_M(\mathfrak{g}_P)$

$$\begin{aligned} d(H_\phi)_A(\psi) &= \frac{d}{dt}\bigg|_{t=0} \int_M \langle F_{A+t\psi}, \phi\rangle \\ &= \int_M \langle d_A\psi, \phi\rangle \\ &= \int_M d\langle \psi, \phi\rangle - \int_M \langle \psi, d_A\phi\rangle \\ &= \int_M \langle d_A\phi, \psi\rangle \end{aligned}$$

Then noting that $d_A\phi$ is the vector field determined by ϕ, this shows that the action is Hamiltonian, with the functions H_ϕ defining the Hamiltonian functions. We then claim that mapping $\mu(A) = F_A$ defines a moment map. We see this, we first note that μ is $\mathscr{G}(P)$-equivariant by our formulas regarding how a connection and its curvature transform under gauge transformation. For $\psi \in \Omega^1_M(\mathfrak{g}_P)$, we have that

$$d\mu_A(\psi)(\phi) = \int_M \langle d_A\psi, \phi\rangle$$

which from our above computation is equal to $\int_M \langle d_A\phi, \psi\rangle$, so μ defines a moment map for this Hamiltonian action.

7. Comparison of the Holomorphic and Symplectic Perspectives

Given a smooth complex vector bundle $E \to M$, a Hermitian metric on E gives rise to the unitary frame bundle $\mathcal{B}_U(E)$, which is a principal $U(n)$-bundle. Furthermore, the isomorphism class of the bundle $\mathcal{B}_U(E)$ is independent of the choice of Hermitian metric, and is determined by the first Chern class of E. For notational compactness, let $P \to M$ denote the principal $U(n)$-bundle of unitary frames of E with respect to a fixed Hermitian metric on E. This Hermitian structure allows us to relate holomorphic stuctures on E with connections on P – given a holomorphic structure $\overline{\partial}_E$ on E, there exists a unique connection on E called the ***Chern connection*** that is compatible with both the holomorphic and Hermitian structures. Furthermore, a connection on P induces a connection on E, and the $(0,1)$ component of this connection necessarily squares to 0, since $\Omega^{2,0}_M = 0$, so it defines a holomorphic structure on E. Therefore, we have a bijection $\mathscr{A}(P) \to \mathscr{C}(E)$.

Using this comparison, we can compare the groups $\mathscr{G}(P)$ and $\mathrm{Aut}(E)$. In a similar fashion to $\mathscr{G}(P)$, an element of $\mathrm{Aut}(E)$ can be thought of as a global section of the associated bundle $\mathcal{B}(E) \times_{\mathrm{GL}_n\mathbb{C}} \mathrm{GL}_n\mathbb{C}$ constructed with the conjugation action. Ignoring analytic difficulties, this identifies the Lie algebra of $\mathrm{Aut}(E)$ with the global sections of $\mathcal{B}(E) \times_{\mathrm{GL}_n\mathbb{C}} \mathfrak{gl}_n\mathbb{C}$. In this way, we see that $\mathrm{Aut}(E)$ is the complexification of $\mathscr{G}(P)$, which essentially boils down to the fact that every complex matrix A can be expressed as $A = B + iC$ with B Hermitian and C skew-Hermitian.

With this in mind, we want to appeal to the usual relationship between a GIT quotient by a complex group and the symplectic quotient by its maximal compact subgroup. The analogies can be summarized as follows:

$$\begin{aligned} \text{Complex reductive group } G &\longleftrightarrow \operatorname{Aut}(E) \\ \text{Maximal compact subgroup } K &\longleftrightarrow \mathscr{G}(P) \\ \text{Moment map } \mu &\longleftrightarrow \text{Curvature } F_A \text{ of a connection} \\ \text{Norm square of the moment map} &\longleftrightarrow \text{Yang-Mills functional} \end{aligned}$$

In the finite dimensional case, the Kempf-Ness theorem establishes a homeomorphism between the GIT quotient by G and the symplectic quotient by K, which is a consequence of every G-orbit containing a unique K-orbit minimizing the norm square of the moment map. In the Yang-Mills situation, we have an infinite dimensional example, which is the Narasimhan-Seshadri theorem. Reformulating the statement in a more differential geometric way, the statement of the theorem is:

Theorem 7.1 (***Narasimhan-Seshadri***)**.** *Let $\mathscr{A}_s(P) \subset \mathscr{A}(P)$ denote the subspace of connections of P that are absolute minima of the Yang-Mills and arise from irreducible representations $\Gamma_{\mathbb{R}} \to U(n)$. Then the identification of $\mathscr{A}(P)$ with $\mathscr{C}(E)$ induces a homeomorphism $\mathscr{A}_s(P)/\mathscr{G}(P) \to \mathscr{C}_s(E)/\operatorname{Aut}(E)$.*

The reformulation and differential geometric proof of Narasimhan-Seshadri is due to Donaldson [3], where he proved that a stability of a holomorphic vector bundle $E \to M$ is equivalent to the existence of a unitary connection A with central curvature satisfying the condition that $\star F_A = -2\pi i \mu(E)$. In the case that E is degree 0, this agrees with the classical statement of Narasimhan-Seshadri by passing to the correspondence between flat bundles and representations of the fundamental group.

To continue this analogy, in the traditional GIT setting, the norm square of the moment map serves as a Morse function, which then gives a Morse stratification into stable and unstable submanifolds by looking at the gradient flow lines, which flow towards or away from the critical sets of the norm square of the moment map. In our case, we have a good candidate for the Morse strata of the Yang-Mills functional, which comes from the existence of Harder-Narasimhan filtrations. As before, let $E \to M$ be a holomorphic vector bundle with a fixed Hermitian metric and let $P \to M$ be its principal $U(n)$-bundle of unitary frames. Recall that the curvature of a Yang-Mills connection on P is equivalent to the data of a skew-Hermitian matrix X, and writing at $X = -2\pi i \Lambda$ with Λ Hermitian, the eigenvalues of X determine a reduction of structure group from $U(n)$ to the stabilizer subgroup o X, which is a subgroup isomorphic to $U(n_1) \times \ldots U(n_m)$ depending on the multiplicities of the eigenvalues. Comparing this with the holomorphic perspective, this exactly corresponds with the Harder-Narasimhan filtration of the bundle $E \to M$. In a paper of Shatz [4], the Harder-Narasimhan filtration gives a stratification of the space $\mathscr{C}(E)$ of holomorphic structures of E, where a holomorphic structure on E is of type $\mu \in \mathbb{Q}^k$ (where k is the length of the filtration) if the slopes of the quotients in the Harder-Narasimhan filtration (arranged in decreasing order) are given by the entries of μ. We let $\mathscr{C}_\mu(E) \subset \mathscr{C}(E)$ denote the subspace of holomorphic structures on E of type μ. Using the type, one can construct the Harder-Narasimhan polygon using the slopes, which is a convex polygon whose vertices are determined by the numerators and denominators of the slopes specified by μ. These polygons determine a partial ordering on the slope vectors, where we say $\lambda \geq \mu$ if the polygon corresponding to λ lies above the polygon corresponding to μ. Then the subspaces $\mathscr{C}_\mu(E)$ give a stratification of $\mathscr{C}(E)$ where

$$\mathscr{C}_\mu(E) \subset \bigcup_{\lambda \geq \mu} \mathscr{C}_\lambda(E)$$

called the ***Harder-Narasimhan stratification***. Furthermore, since the Harder-Narasimhan filtration is canonical, it is preserved by the action of $\mathrm{Aut}(E)$, so the action of $\mathrm{Aut}(E)$ on $\mathscr{C}$ restricts to actions of $\mathrm{Aut}(E)$ on the strata $\mathscr{C}_\mu(E)$, which we will need to compute the equivariant cohomology of $\mathscr{C}(E)$.

Using the identification of $\mathscr{C}(E)$ with $\mathscr{A}(P)$ to transport the Harder-Narasimhan stratification to $\mathscr{A}(P)$, and we can then give a more differential geometric interpretation of the strata $\mathscr{C}_\mu(E)$. Let $\mathscr{A}_\mu(P) \subset \mathscr{A}(P)$ denote the stratum corresponding to $\mathscr{C}_\mu(E)$, which under the identification $\mathscr{C}(E) \cong \mathscr{A}(P)$ corresponds to Yang-Mills minima whose Lie algebra element has eigenvalues coinciding with the type vector μ. Our goal will be to draw analogies between these Harder-Narasimhan strata and the hypothetical Morse strata for the Yang-Mills functional, and use the strata to compute the cohomology of the moduli space $\mathcal{N}(n,k)$ when n and k are coprime. In the differential geometric perspective, we consider a larger class of functionals on $\mathscr{A}(P)$ arising from convex functions on $\mathfrak{u}(n)$ that are invariant under the adjoint action, which will have the same critical sets as the Yang-Mills functional. Using these functionals, one can show that the gradient flow of these functionals is tangential to the $\mathrm{Aut}(E)$ orbits, and applying some elliptic theory to the connection Laplacian $d_A^* d_A + d_A d_A^*$ allows one to show that the critical sets have finite Morse index. Using these convex invariant functions also gives an alternative proof of the stratification

$$\mathscr{A}_\mu(P) \subset \bigcup_{\lambda \geq \mu} \mathscr{A}_\lambda(P)$$

Which suggests that the $\mathscr{A}_\mu$ play the role of Morse strata for the Yang-Mills functional modulo analytic difficulties with the critical sets, and problems regarding convergence of the gradient flow of the Yang-Mills functional. It was later shown by Daskalopoulos [2] that these analytic difficulties can be resolved, and that the Yang-Mills functional is an equivariantly perfect Morse function on $\mathscr{A}(P)$. However, it is not strictly necessary to use this fact to compute the cohomology of $\mathcal{N}(n,k)$, provided that we can show that the Harder-Narasimhan stratification is something called ***equivariantly perfect***, which we will discuss later.

8. The Cohomology of $\mathcal{N}(n,k)$

We restrict to the case where n and k are coprime, so stable bundles coincide with semistable bundles, and most of the work will be done using the holomorphic perspective. The strategy will be to compute the $\mathrm{Aut}(E)$-equivariant cohomology of the strata $\mathscr{C}_\mu(E)$ and then establish the equivariant perfection of the stratification to use an inductive Mayer-Vietoris and Kunneth formula procedure to compute the equivariant cohomology of $\mathscr{C}(E)$. Then by quotienting $\mathrm{Aut}(E)$ by a certain subgroup, we obtain a free action on $\mathscr{C}_s(E)$, and the equivariant cohomology by this quotient group will be the ordinary cohomology of the quotient space $\mathcal{N}(n,k)$, since in the coprime case, $\mathscr{C}_{ss}(E)$ and $\mathscr{C}_s(E)$ coicide.

We first tackle the problem of computing the $\mathrm{Aut}(E)$-equivariant cohomology of the strata $\mathscr{C}_\mu(E)$. To do this, we will first make some reductions to replace $\mathrm{Aut}(E)$ and $\mathscr{C}_\mu(E)$ with a smaller group and space that are homotopy equivalent to $\mathrm{Aut}(E)$ and $\mathscr{C}_\mu(E)$ respectively. Let $\mathscr{F}_\mu(E)$ denote the space of smooth filtrations of E of type μ, which can be interpreted as a subspace of the space of smooth sections of a flag bundle of E. There is a natural surjection $\mathscr{C}_\mu(E) \to \mathscr{F}_\mu(E)$ given by forgetting the holomorphic structures of the subbundles in the filtration, and the fiber $\mathscr{B}(E_\mu)$ of this map over a fixed filtration E_μ is the space of holomorphic structures on the terms of the filtration. Let $\mathrm{Aut}(E_\mu)$ denote the group of smooth automorphisms of E preserving the filtration E_μ. Then we can identify $\mathscr{F}_\mu(E)$ with the quotient space $\mathrm{Aut}(E)/\mathrm{Aut}(E_\mu)$, which is analogous to the identification of the space of partial flags in $\mathbb{C}^n$ with $\mathrm{GL}_n\mathbb{C}/P$ for a parabolic subgroup P. This identification gives $\mathrm{Aut}(E)$ the structure of a principal $\mathrm{Aut}(E_\mu)$-bundle over $F_\mu(E)$. Furthermore, this gives an identification

$$\mathscr{C}_\mu(E) \cong \mathrm{Aut}(E) \times_{\mathrm{Aut}(E_\mu)} \mathscr{B}(E_\mu)$$

Then let $E\,\mathrm{Aut}(E) \to B\,\mathrm{Aut}(E)$ be a universal bundle for $\mathrm{Aut}(E)$. Then the $\mathrm{Aut}(E)$-equivariant cohomology of $\mathscr{C}_\mu(E)$ is defined to the the ordinary cohomology of the total space of the associated bundle $E\,\mathrm{Aut}(E) \times_{\mathrm{Aut}(E)} \mathscr{C}_\mu(E)$. By our above observation, this is the same as the space $E\,\mathrm{Aut}(E) \times_{\mathrm{Aut}(E)} (\mathrm{Aut}(E) \times_{\mathrm{Aut}(E_\mu)} \mathscr{B}(E_\mu))$. Noting that $E\,\mathrm{Aut}(E) \times_{\mathrm{Aut}(E)} \mathrm{Aut}(E) \cong E\,\mathrm{Aut}(E)$, we have that this is the same as the space $E\,\mathrm{Aut}(E) \times_{\mathrm{Aut}(E_\mu)} \mathscr{B}(E_\mu)$. We then note that $E\,\mathrm{Aut}(E)$ is a contractible space with an action of $\mathrm{Aut}(E_\mu)$ (coming from restriction), so it may serve as the total space for a universal bundle for $\mathrm{Aut}(E_\mu)$. Thus, we have have

$$H^\bullet_{\mathrm{Aut}(E)}(\mathscr{C}_\mu(E)) \cong H^\bullet_{\mathrm{Aut}(E_\mu)}(\mathscr{B}(E_\mu))$$

We then reduce further. Let $\{E_i\}$ denote the subbundles appearing in the filtration E_μ of E. We have exact sequences of smooth vector bundles

$$0 \longrightarrow E_i \longrightarrow E_{i+1} \longrightarrow E_{i+1}/E_i \longrightarrow 0$$

which split in the smooth category. By fixing splittings for each of these sequences, we get a direct sum decomposition

$$E = D_1 \oplus \cdots \oplus D_r$$

compatible with the filtration E_μ, which we denote by E_μ^0. We then let $\mathrm{Aut}(E_\mu^0) \subset \mathrm{Aut}(E_\mu)$ be the automorphisms of E that respect this direct sum decomposition, and let $\mathscr{B}(E_\mu^0)$ denote the holomorphic structures that give the D_i the structure of semistable holomorphic bundles. We then clearly have that $\mathrm{Aut}(E_\mu^0) \cong \prod_i \mathrm{Aut}(D_i)$, which corresponds to writing an automorphism in block form. Similarly, we have that $\mathscr{B}(E_\mu^0) = \prod_i \mathscr{C}_{ss}(D_i)$. We then make two observations:

(1) $\mathrm{Aut}(E_\mu)$ deformations retracts onto $\mathrm{Aut}(E_\mu^0)$, which roughly corresponds to a parabolic subgroup of $\mathrm{GL}_n\mathbb{C}$ deformation retracting onto a product of $\mathrm{GL}_{n_i}\mathbb{C}$ via a Gram-Schmidt like procedure.
(2) The forgetful map $\mathscr{B}(E_\mu) \to \mathscr{B}(E_\mu^0)$ is a homotopy equivalence, which roughly corresponds to the space of splittings of an exact sequence of vector spaces being affine.

These two observations give us the second reduction, which when combined with the first gives us

$$H^{\bullet}_{\mathrm{Aut}(E)}(\mathscr{C}_\mu(E)) \cong H^{\bullet}_{\mathrm{Aut}(E^0_\mu)}(\mathscr{B}(E^0_\mu))$$

Using Kunneth and the identification of $\mathrm{Aut}(E^0_\mu)$ and $\mathscr{B}(E^0_\mu)$ with products, this gives us

$$H^{\bullet}_{\mathrm{Aut}(E)}(\mathscr{C}_\mu(E), \mathbb{Q}) \cong \bigotimes_{i=1}^{r} H^{\bullet}_{\mathrm{Aut}(D_i)}(\mathscr{C}_{ss}(D_i), \mathbb{Q})$$

Which tells us that computing the equivariant cohomology for the semistable locus for lower rank bundles will give us the equivariant cohomology of the Harder-Narasimhan strata.

With this in hand, our next goal is to establish equivariant perfection of the Harder-Narasimhan stratification, which will allow us to conclude that the equivariant Poincaré polyomial for $\mathscr{C}_{ss}(E)$ is equal to the "Morse polynomial" coming from the stratification, which is the polynomial

$$\sum_{\lambda} t^{k_\mu} P_{t,\mathrm{Aut}(E_\mu)}(\mathscr{C}_\mu(E))$$

where k_μ is the codimension of $\mathscr{C}_\mu(E)$ in $\mathscr{C}(E)$ and $P_{t,\mathrm{Aut}(E_\mu}(\mathscr{C}_\mu(E))$ denotes the equivariant Poincaré polynomial of the stratum $\mathscr{C}_\mu(E)$. To do this, we will need to take some results on faith, which give conditions for equivariant perfection.

Proposition 8.1. *If the equivariant Euler classes to the normal bundles N_μ of the strata $\mathscr{C}_\mu(E)$ are not zero divisors in $H^{\bullet}_{\mathrm{Aut}(E_\mu)}(N_\mu, k)$ for a coefficient field k, then the stratification is equivariantly perfect with respect to k.*

Proposition 8.2. *Let X be a connected G space where a subtorus of G acts trivially, and let $N \to X$ be an equivariant vector bundle. Then if the action of the subtorus on each fiber is a primitive representation and $H^{\bullet}(G, \mathbb{Z})$ has no torsion, then multiplication by the equivariant Euler class of N on $H^{\bullet}_G(X, \mathbb{F}_p)$ is injective.*

With this in mind, we want to identify the normal bundles to the strata and compute their equivariant Euler classes.

References

[1] M. F. Atiyah and R. Bott. "The Yang-Mills Equations over Riemann Surfaces". In: *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 308.1505 (1983), pp. 523–615. ISSN: 00804614. URL: http://www.jstor.org/stable/37156.

[2] Georgios D. Daskalopoulos. "The topology of the space of stable bundles on a compact Riemann surface". In: *J. Differential Geom.* 36.3 (1992), pp. 699–746. DOI: 10.4310/jdg/1214453186. URL: https://doi.org/10.4310/jdg/1214453186.

[3] S. K. Donaldson. "A new proof of a theorem of Narasimhan and Seshadri". In: *J. Differential Geom.* 18.2 (1983), pp. 269–277. DOI: 10.4310/jdg/1214437664. URL: https://doi.org/10.4310/jdg/1214437664.

[4] Stephen S. Shatz. "Degeneration and specialization in algebraic families of vector bundles". In: *Bull. Amer. Math. Soc.* 82.4 (July 1976), pp. 560–562. URL: https://projecteuclid.org:443/euclid.bams/1183538125.