PRINCIPAL BUNDLE AND ITS APPLICATIONS

BOWEN LIU

Abstract.

Contents

0. Preface

0.1. **About this lecture.**

0.2. **Some notations.**

1. M is used to denote a smooth manifold, and $x \in M$ denotes its point.
2. TM and Ω^k_M denote tangent bundle and bundle of k-forms over M respectively.
3. v is used to denote vector in tangent space.
4. X is used to denote a vector field on M, then X_x denote the value of X at point $x \in M$, similarly for a k-form ω.
5. For a vector bundle E over M, $C^\infty(M, E)$ denotes its (smooth) sections.
6. G is used to denote a Lie group, with Lie algebra $\mathfrak{g}$.
7. $\pi : P \to M$ is used to denote a principal G-bundle over M, and $p \in P$ denotes its point.
8. $\widetilde{X}$ is used to denote vector field on principal bundle P, so do $\widetilde{\omega}$ and $\widetilde{v}$.

Part 1. Principal bundle and its geometry

1. Principal bundle

1.1. A glimpse of fiber bundle.
Fix topological spaces E, B, F.

Definition 1.1.1 (fiber bundle)**.** A fiber bundle with fiber F over B is a surjective map $\pi : E \to B$ such that for any $p \in B$, there exists an open neighborhood $U \ni p$ and a homeomorphism φ such that the following diagram commutes

$$\begin{tikzcd} \pi^{-1}(U) \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & U \times F \arrow[dl, "\pi_1"] \\ & U & \end{tikzcd}$$

We always use $F \to E \xrightarrow{\pi} B$ to denote this fiber bundle and

1. B is called base space;
2. $E_x = \pi^{-1}(x)$ is called the fiber of E at x;
3. (U, φ) is called a local trivialization at point p, and use $E|_U$ to denote $\pi^{-1}(U)$.

Example 1.1.1 (trivial bundle)**.** Consider $E = B \times F$ and $\pi : E \to B$ is just the projection onto the first summand.

Example 1.1.2. Consider $E = S^n$ and $B = \mathbb{RP}^n$, then natural map $\pi : E \to B$ is a fiber bundle with $\mathbb{Z}/2\mathbb{Z}$. It's clear that this fiber bundle is not trivial, since S^n is connected.

Example 1.1.3 (Hopf fibration)**.** Recall that

$$\mathbb{CP}^n = \{\text{the set of all complex lines through origin in } \mathbb{C}^{n+1}\}$$

Consider the canonical open covering $\{U_i\}$ of $\mathbb{CP}^n$, that is

$$U_i = \{[z_0 : \cdots : z_n] \mid z_i \neq 0\}$$

Now view $S^{2n+1} \subset \mathbb{R}^{2n+2} = \mathbb{C}^{n+1}$ as the set of all $(z_0, \ldots, z_n) \in \mathbb{C}^{n+1}$ with $|z_0|^2 + \cdots + |z_n|^2 = 1$. Then the projection map $\pi : \mathbb{C}^{n+1} - \{0\} \to \mathbb{CP}^n$ restricts to a surjective smooth map

$$\pi : S^{2n+1} \to \mathbb{CP}^n$$

We claim that it's a fiber bundle with fiber S^1. Indeed, by definition we have

$$\pi^{-1}(U_i) = \{(z_0, \ldots, z_n) \in S^{2n+1} \mid z_i \neq 0\}$$

and local trivialization map can be taken as

$$\begin{aligned} \varphi_i : \pi^{-1}(U_i) &\to U_i \times S^1 \\ z &\mapsto ([z_0 : \cdots : z_n], \frac{z_i}{|z_i|}) \end{aligned}$$

It's also not trivial which can be seen by considering their fundamental groups.

Example 1.1.4. The covering space is a fiber bundle with discrete set as fiber.

1.2. **Principal bundle.**

1.2.1. *Definitions.* Briefly speaking, given a Lie group G and a smooth manifold M, a principal G-bundle P is a fiber bundle with fiber G equipped with a suitable smooth right G-action on it. For a smooth right G-action we mean a smooth map

$$\begin{aligned} P \times G &\to P \\ (p, g) &\mapsto pg \end{aligned}$$

Definition 1.2.1 (principal G-bundle)**.** A principal G-bundle is a surjective smooth map $\pi : P \to M$ between smooth manifolds such that:

1. There is a smooth right G-action on P;
2. For all $x \in M$, $\pi^{-1}(x)$ is a G-orbit;
3. For all $x \in M$, there exists an open subset U_α and a diffeomorphism φ_α such that the following diagram commutes

$$\begin{tikzcd} \pi^{-1}(U_\alpha) \arrow[rr, "\varphi_\alpha"] \arrow[dr, "\pi"'] & & U_\alpha \times G \arrow[dl, "\pi_1"] \\ & U_\alpha & \end{tikzcd}$$

   If we write $\varphi_\alpha(p) = (\pi(p), g_\alpha(p))$, then we require $g_\alpha(pg) = g_\alpha(p)g$ for any $g \in G$.

Remark 1.2.1. Note that G acts on P freely and transitively, which can be seen from local trivialization.

Notation 1.2.1. $\mathcal{P}_G M$ is used to denoted the set of all principal G-bundles over M up to isomorphism.

Example 1.2.1. $S^n \to \mathbb{RP}^n$ is a $\mathbb{Z}/2\mathbb{Z}$-principal bundle, where $\mathbb{Z}/2\mathbb{Z}$ acts on S^n as $x \mapsto -x$.

Example 1.2.2. $S^{2n+1} \to \mathbb{CP}^n$ is a $U(1)$-principal bundle, where $U(1)$ acts on S^{2n+1} as $(z_0, z_1, \ldots, z_n) \mapsto (z_0 e^{i\theta}, z_1 e^{i\theta}, \ldots, z_n e^{i\theta})$.

Definition 1.2.2 (isomorphism between principal G-bundle)**.** For two principal G-bundles $(P, M, \pi), (P', M, \pi')$, if there exists a G-equivariant smooth map $\widetilde{f} : P' \to P$ making the following diagram commute

$$\begin{tikzcd} P \arrow[rr, "\widetilde{f}"] \arrow[dr, "\pi"'] & & P' \arrow[dl, "\pi'"] \\ & M & \end{tikzcd}$$

Then P and P' are called isomorphic principal G-bundle.

Remark 1.2.2. Although here we put no restrictions on injectivity or surjectivity of $\widetilde{f}$, these information are encoded in the equivariance of $\widetilde{f}$ and properties of principal G-bundle:

1. $\widetilde{f}$ is injective: For any $p_1, p_2 \in P$, if $\widetilde{f}(p_1) = \widetilde{f}(p_2)$, then p_1, p_2 lie in same fiber, since above diagram commutes. If $p_1 = p_2 g$ for $g \in G$, then $\widetilde{f}(p_1) = \widetilde{f}(p_2)g$, which impiles $g = e$, since G acts on P' freely, that is $p_1 = p_2$;
2. $\widetilde{f}$ is surjective: For any $p' \in P'$, if $\pi'(p') = x$, then $p' \in P'_x$. So choose an arbitrary element $p \in P_x$, there must be some $g \in G$ such that $\widetilde{f}(pg) = p'$, since P'_x is a G-orbit and $\widetilde{f}$ is G-equivariant.

Definition 1.2.3 (trivial principal bundle)**.** A principal G-bundle P is called trivial, if there exists a principal G-bundle isomorphism $\varphi : P \to M \times G$.

Lemma 1.2.1. If $\widetilde{f} : M \times G \to M \times G$ is an isomorphism between trivial principal G-bundles, then there exists $\varphi : M \to G$ such that

$$\widetilde{f}(x, g) = (x, \varphi(x)g)$$

Proof. Define $\varphi(x)$ via $\widetilde{f}(x, 1) = (x, \varphi(x))$. □

1.2.2. *Transition functions.* By (3) of Definition 1.2.1, there exists an open covering $\{U_\alpha\}$ together with G-equivariant diffeomorphism φ_α. If $U_\alpha \cap U_\beta \neq \varnothing$, then

$$\begin{aligned} \varphi_{\alpha\beta} := \varphi_\alpha \circ \varphi_\beta^{-1} : (U_\alpha \cap U_\beta) \times G &\to (U_\alpha \cap U_\beta) \times G \\ (x, g) &\mapsto (x, g_\alpha(\varphi_\beta^{-1}(x, g))) \end{aligned}$$

If we denote

$$g_{\alpha\beta}(x) = g_\alpha(\varphi_\beta^{-1}(x, g))$$

we obtain G-equivariant map

$$g_{\alpha\beta} : U_\alpha \cap U_\beta \to \operatorname{Diff} G = \{f : G \to G \mid f \text{ is diffeomorphism}\}$$

But you can check that a diffeomorphism $f : G \to G$ which is G-equivariant must take the form $x \mapsto gx$, which impiles in fact we have transition functions of principal G-bundle take the form

$$g_{\alpha\beta} : U_\alpha \cap U_\beta \to G$$

Conversely, it's clear you can recover a principal G-bundle from its transition functions.

1.2.3. *Section.*

Definition 1.2.4 (global section)**.** A global section is a smooth map $s : M \to P$ such that $\pi \circ s = \mathrm{id}$.

Proposition 1.2.1. A principal bundle admits a section if and only if it is trivial.[1]

Proof. If $s : M \to P$ is a smooth section, we define

$$\begin{aligned}\varphi : P &\to M \times G \\ p &\mapsto (\pi(p), g(p))\end{aligned}$$

where $g(p) \in G$ such that $p = s(\pi(p))g(p)$, it always exsits since the right action of G is transitive on each fiber and it is unique since the action is free on each fiber. Clearly, it's G-equivariant, since

$$\varphi(ph) = (\pi(ph), g(ph)) = (\pi(p), g(p)h)$$

and the last equality holds since

$$ph = s(\pi(ph))g(ph) = s(\pi(p))g(ph) = pg^{-1}(p)g(ph) \implies h = g^{-1}(p)g(ph)$$

And it's easy to see φ is a bijection, with inverse map

$$\begin{aligned}\varphi^{-1} : M \times G &\to P \\ (p, g) &\mapsto s(p)g\end{aligned}$$

The smoothness of the section and of the G-action on P imply smoothness. □

1.3. **Associated fiber bundle.** Given a principal G-bundle P and a smooth manifold F admitting a smooth left G-action on it. Then we can construct a fiber bundle $P \times_G F$ with fiber F with base space M as follows

$$P \times_G F := P \times F / \sim$$

where $(p, f) \sim (p', f')$ if and only if $p' = pg, f' = g^{-1}f$. Let's check $P \times_G F$ is a fiber bundle.

Proof. Consider the map taking an equivalence class $[p, f]$ to $\pi(p)$. To see the local structure, since we already have the local structure of principal bundle P, i.e. for any $x \in M$, there exists open $U_\alpha \ni x$ and $\varphi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times G$. Now we define the local trivialization of $P \times_G V$ as

$$\begin{aligned}\varphi_\alpha^V : (P \times_G V)|_{U_\alpha} &\to U_\alpha \times V \\ (p, v) &\mapsto (\pi(p), g_\alpha(p)v)\end{aligned}$$

First note that this is well-defined, since

$$(pg, g^{-1}v) \mapsto (\pi(pg), g_\alpha(pg)g^{-1}v) = (\pi(p), g_\alpha(p)gg^{-1}v) = (\pi(p), g_\alpha(p)v)$$

And this map is one to one, and invertible, its inverse sends $(x, v) \in U_\alpha \times V$ to the equivalence class of $(\varphi_\alpha^{-1}(x, e), v)$. Directly check as follows

$$\begin{aligned}\varphi_\alpha^V(\varphi_\alpha^{-1}(x, e), v) &= (x, ev) \\ &= (x, v)\end{aligned}$$

since $\pi(\varphi_\alpha^{-1}(x, e)) = x$ and $g_\alpha(\varphi_\alpha^{-1}(x, e)) = e$. □

[1]This is in sharp contrast with vector bundles, which always admit sections.

Remark 1.3.1 (transition function of associated bundle). Though we've found the local trivialization of $P \times_G V$, it's also necessary to see what does the transition functions look like.

Let U_α, U_β be open sets with non-empty intersection $U_{\alpha\beta}$, and $\varphi_\alpha, \varphi_\beta$ be local trivializations of principal bundles, with transition functions

$$\begin{aligned} \varphi_\alpha \circ \varphi_\beta^{-1} : U_{\alpha\beta} \times G &\to U_{\alpha\beta} \times G \\ (x, g) &\mapsto (x, g_{\alpha\beta}(x) g) \end{aligned}$$

then we can compute the transition functions of associated vector bundles as follows

$$\begin{aligned} \varphi_\alpha^V \circ (\varphi_\beta^V)^{-1} : U_{\alpha\beta} \times V &\to U_{\alpha\beta} \times V \\ (x, v) &\mapsto (\varphi_\beta^{-1}(x, e), v) \mapsto (x, g_{\alpha\beta}(x) v) \end{aligned}$$

Example 1.3.1 (associated vector bundle)**.** Now let's consider a special case, that is associated vector bundles. Given a representation of G, that is a group homomorphism $\rho : G \to \mathrm{GL}(V)$, thus you can construct a vector bundle $P \times_G V$. However, there is a more simple way to construct in transition functions viewpoint: By Remark 1.3.1, we can see the transition function of this associated vector bundle is $\{\rho \circ g_{\alpha\beta}\}$, where $\{g_{\alpha\beta}\}$ is transition function of P.

Remark 1.3.2 (Relations between vector bundle and principal bundle). If we consider real vector bundles, we have the following one to one correspondence

$$\phi : \mathcal{P}_{\mathrm{GL}(n,\mathbb{R})} M \to \mathrm{Vect}_n^{\mathbb{R}} M$$

given by $P \mapsto P \times_\rho \mathbb{R}^n$, where $\rho : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$ is trivial representation. The inverse ψ is given by considering frame bundle of vector bundle. Furthermore, if we endow vector bundle a Riemannian metric, then it can be regarded as a $\mathrm{O}(n)$-principal bundle, and one can show it's independent of the choice of Riemannian metric, thus in fact we have the following one to one correspondence

$$\mathcal{P}_{\mathrm{O}(n)} M \Longleftrightarrow \mathrm{Vect}_n^{\mathbb{R}} M$$

Similarly we also have

$$\mathcal{P}_{\mathrm{U}(n)} M \Longleftrightarrow \mathrm{Vect}_n^{\mathbb{C}} M$$

Example 1.3.2. There are two important examples of associated bundles that we will use later.

1. The associated bundle obtained from G acts on G by conjugation, denoted by $\mathrm{Conj}\, P$;
2. The associated vector bundle obtained from G acts on $\mathfrak{g}$ by adjoint representation, denoted by $\mathrm{Ad}\, P$.

Remark 1.3.3. A philosophy of geometry is that we can study the objects lying over this geometric objects to study this geometric object itself, and that's why we study the vector bundle over a smooth manifold. Note that for a principal G-bundle, you can obtain a vector bundle from a representation of G. However, there are too many representations of G, so special representations may correspond to special vector bundles.

Proposition 1.3.1. There is a one to one correspondence

$$C^\infty(M, P\times_G F) \xleftrightarrow{1-1} \{f : P\to F \mid f \text{ is smooth and } f(xg) = g^{-1}f(x)\}$$

Proof. For smooth function $f : P \to F$ which is G-equivariant, we define $s_f \in \Gamma(M, P\times_G F)$ as

$$s_f(x) = \{(p, f(p)) \mid \pi(p) = x\}, \quad x\in M$$

Here we need to check our definition is independent of the choice of p. Indeed, if we choose pg, then $s_f(x) = (pg, f(pg)) = (pg, g^{-1}f(p)) \sim (p, f(p)) \in P\times_G F$.

Conversely, given $s\in C^\infty(M, P\times_G F)$, then for any $p\in P$, we consider $\pi(p) = x\in M$ and write $s(x) = [(p, v)]$, then we define $f(p) = v$. It's clear $f(pg) = g^{-1}f(p)$, since $[(p, v)] = [(pg, g^{-1}v)]$. □

In fact, this proposition is not a coincidence, and it's a quite important motivation which explains why we need principal bundles. If $\pi : P\to M$ is a principal G bundle, and E is a vector bundle over M such that E is an associated vector bundle of P, then if we use π to pull E back to P, we claim that the vector bundle π^*E is the trivial bundle $P\times V$ over P. Indeed, we define the following bundle map

$$\begin{aligned} \psi : P\times V &\to P\times_G V \\ (p, v) &\mapsto [p, v] \end{aligned}$$

and consider the following diagram

$$\begin{tikzcd} P\times V \arrow[r] \arrow[d, "\psi"] & P \arrow[d, "\pi"] \\ E = P\times_G V \arrow[r] & M \end{tikzcd}$$

Clearly $P\times V$ satisfies the universal property of pullback, thus by uniqueness we obtain $\pi^*E\cong P\times V$.

It's clear sections of trivial bundle $P\times V$ can be regard as smooth functions $f : P\to V$, and by relation between sections of bundle and its pullback bundle, there is no surprise you have one to one correspondence in Proposition 1.3.1.

Remark 1.3.4. More generally, we can use π to pull $(P\times_G V)\otimes E'$ back to P, and prove it's $(P\times V)\otimes\pi^*E'$ by the same method. The cases we will encounter are $E' = T^*M$ or $E' = \bigwedge^k T^*M$. We use $\Omega^k_M(P\times_G V)$ to denote $(P\times_G V)\otimes\bigwedge^k T^*M$, the generalization tells that we have the one to one correspondence between sections of $\Omega^k_M(P\times_G V)$ and sections of $(P\times V)\otimes\pi^*\bigwedge^k T^*M$ with equivariant conditions, we will call such forms basic forms, a conception we will define later.

1.4. **Reduction of principal bundle.** Given a principal G-bundle $\pi : P \to M$ and a H-principal bundle $\pi' : P' \to M$. Furthermore, there is a Lie group homomorphism $\alpha : H \to G$.

Definition 1.4.1 (reduction)**.** If there exists a smooth map $\varphi : P' \to P$ such that the following diagram commutes

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

and φ is H-equivariant, that is for any $p \in F, h \in H$

$$\varphi(ph) = \varphi(p)\alpha(h)$$

Then P is called an extension of P' from H to G and P' is called an reduction of P from G to H.

Remark 1.4.1. Here are two cases we're concern about:

1. $H < G$ is a subgroup, α is an inclusion.
2. $\alpha : H \to G$ is surjective, for example, H is universal covering of G.

Extension of principal bundle always exists, and it's unique, according to the following proposition.

Proposition 1.4.1. Given a Lie group homomorphism $\alpha : H \to G$ and a H-principal bundle P', there exists a unique extension of P' from H to G.

Proof. Existence: Note that $\alpha : H \to G$ gives a smooth left H-action on G, then consider associated fiber bundle $P' \times_H G$, it's a principal G-bundle, and if we define

$$\begin{aligned} \varphi : P' &\to P' \times_H G \\ p' &\mapsto [p', 1] \end{aligned}$$

Then φ is desired equivariant map which makes diagram commutes.

Uniqueness: If there is another extension $\varphi' : P' \to P$, in order to make the following diagram commutes

$$\begin{tikzcd} & P' \times_H G \arrow[d, dashed, "\psi"] \\ P' \arrow[ur, "\varphi"] \arrow[r, "\varphi'"'] & P \end{tikzcd}$$

we define ψ by $\psi([p, 1]) = \varphi'(p)$. Thus principal G-bundles $P' \times_H G$ and P are isomorphic to each other. $\square$

However, reduction may not exist.

Lemma 1.4.1. Let $\alpha : H \to G$ be a Lie group homomorphism, P is a principal G-bundle with transition functions $\psi_{\alpha\beta} : U_{\alpha\beta} \to G$. The following statements are equivalent:

1. There exists reduction of P from G to H;
2. There exists $\varphi_{\alpha\beta} : U_{\alpha\beta} \to H$ such that $\alpha \circ \varphi_{\alpha\beta} = \psi_{\alpha\beta}$.

Corollary 1.4.1. Let P be a principal G-bundle and H is a Lie subgroup of G, then there exists a reduction of P from G to H if and only if there exists transition functions of P valued in H.

Example 1.4.1. If $E \to M$ is a complex vector bundle with a hermitian inner product, then a local trivialization

$$\varphi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{C}^n$$

gives a hermitian inner product on $\mathbb{C}^n$. Thus a transition function must preserve the inner product, thus

$$\begin{tikzcd}
U_\alpha \cap U_\beta \arrow[r] \arrow[dr, dashed] & \mathrm{GL}_n(\mathbb{C}) \\
& \mathrm{U}(n) \arrow[u, hook]
\end{tikzcd}$$

This gives a reduction of $\mathrm{GL}_n(\mathbb{C})$-principal bundle to a $\mathrm{U}(n)$-principal bundle.

Example 1.4.2. If $E \to M$ is a real vector bundle, by the same argument we can always reduce its frame bundle P, that is a $\mathrm{GL}_n(\mathbb{R})$-principal bundle, to a $\mathrm{O}(n)$-principal bundle. Furthermore,

1. P can be reduced to a $\mathrm{SO}(n)$-principal bundle if and only if E is orientable;
2. P can be reduced to a $\{e\}$-principal bundle if and only if E is trivial.

Example 1.4.3. Let M be an oriented Riemannian manifold, then TM is a $\mathrm{SO}(n)$-principal bundle. Consider universal covering $\mathrm{Spin}(n) \xrightarrow{2:1} \mathrm{SO}(n)$. If there exists a reduction from $\mathrm{SO}(n)$ to $\mathrm{Spin}(n)$, then we say M admits a spin structure.

2. Connection of principal bundle

2.1. **Forms valued in vector space.** In this section, let M be a smooth manifold, V a vector space with basis $\{e_\alpha\}$ and G a Lie group with Lie algebra $\mathfrak{g}$. A k-form valued in vector space V can be written as

$$\omega = \omega^\alpha e_\alpha$$

where ω^α are k-forms. We use $\Omega^k_M(V)$ to denote the bundle of k-forms valued in V. It's an easy generalization of differential forms, just by replacing $\mathbb{R}$ with a general vector space, and properties of k-forms also hold for k-forms value in V.

Definition 2.1.1 (exterior derivative)**.** Let $\omega = \omega^\alpha e_\alpha$ be a k-form valued in V, then its exterior derivative is

$$\mathrm{d}\omega = \mathrm{d}\omega^\alpha e_\alpha$$

Proposition 2.1.1 (Cartan's formula)**.** Let $\omega = \omega^\alpha e_\alpha$ be a k-form valued in V, then

$$\begin{aligned}(\mathrm{d}\omega)(X_1, \dots, X_{k+1}) = &\sum_{i=1}^{k+1}(-1)^{i+1}X_i\omega(X_1, \dots, \widehat{X_i}, \dots, X_{k+1})\\ &+\sum_{i<j}(-1)^{i+j}\omega([X_i, X_j], X_1, \dots, \widehat{X_i}, \dots, \widehat{X_j}, \dots, X_{k+1})\end{aligned}$$

where X_i are vector fields.

Definition 2.1.2 (wedge product)**.** Let ω_1, ω_2 are forms valued in V with degree k and l respectively, then

$$(\omega_1 \wedge \omega_2)(X_1, \dots, X_{k+l}) := \frac{1}{k! \times l!}\sum_{\sigma \in S_{k+l}}(-1)^{|\sigma|}\omega_1(X_{\sigma(1)}, \dots, X_{\sigma(k)}) \otimes \omega_2(X_{\sigma(k+1)}, \dots, X_{\sigma(k+l)})$$

where X_i are vector fields.

Proposition 2.1.2. Let $\omega_i, i = 1, 2, 3$ be forms valued in V, then

1. $(\omega_1 \wedge \omega_2) \wedge \omega_3 = \omega_1 \wedge (\omega_2 \wedge \omega_3)$, where $\omega_i, i = 1, 2, 3$ are forms valued in V;
2. $\mathrm{d}(\omega_1 \wedge \omega_2) = \mathrm{d}\omega_1 \wedge \omega_2 + (-1)^{\deg \omega_1}\omega_1 \wedge \mathrm{d}\omega_2$,

Definition 2.1.3. Let $T : V \to W$ be a linear map between vector spaces, and ω is a k-form valued in V, then $T\omega$ is a k-form valued in W, which is defined as

$$T\omega(X_1, \dots, X_k) := T(\omega(X_1, \dots, X_k))$$

where X_i are vector fields.

Example 2.1.1. Let ω_1, ω_2 be forms with degree k and l respectively, then by our definition one has $\omega_1 \wedge \omega_2 \in \Omega^{k+l}_M(\mathfrak{g} \otimes \mathfrak{g})$. It's a little bit different from standard definition of wedge product, since $\omega_1\wedge\omega_2$ should be a $(k+l)$-form, not a $(k+l)$-form valued in $\mathbb{R}\otimes\mathbb{R}$. If we consider

$$\begin{aligned} T:\mathbb{R}\otimes\mathbb{R}&\to\mathbb{R}\\ a\otimes b&\mapsto ab \end{aligned}$$

Then $T(\omega_1\wedge\omega_2)$ is a $(k+l)$-form, coincides with standard definition, so we just denote $T(\omega_1\wedge\omega_2)$ by $\omega_1\wedge\omega_2$ for convenience.

Example 2.1.2. Let ω_1 be a k-form valued in $\mathfrak{g}$, and ω_2 a l-form valued in V. Given a representation $\rho: G\to \mathrm{GL}(V)$, it induces a representation of Lie algebra, that is $\rho_*:\mathfrak{g}\to\mathfrak{gl}(V)$. If we consider

$$\begin{aligned} T:\mathfrak{g}\otimes V&\to V\\ \xi\otimes v&\mapsto \rho_*(\xi)v \end{aligned}$$

Then we have $T(\omega_1\wedge\omega_2)$ is a $(k+l)$-form valued in V, we just denote it by $\omega_1\wedge\omega_2$ for convenience.

Example 2.1.3. Let ω_1,ω_2 be forms valued in $\mathfrak{g}$ with degree k and l respectively, by our definition $\omega_1\wedge\omega_2$ is a $(k+l)$-form valued in $\mathfrak{g}$. If we consider

$$\begin{aligned} T:\mathfrak{g}\otimes\mathfrak{g}&\to\mathfrak{g}\\ \xi\otimes\eta&\mapsto[\xi,\eta] \end{aligned}$$

Then we have $T(\omega_1\wedge\omega_2)$ is a $(k+l)$-form valued in $\mathfrak{g}$, we just denote it by $\omega_1\wedge\omega_2$ for convenience.

Remark 2.1.1. If Lie group $G=\mathrm{GL}(n,\mathbb{R})$, then $\mathfrak{g}=\mathfrak{gl}(n,\mathbb{R})$ consists of matrix. Thus in this case for any $\xi,\eta\in\mathfrak{g}$, we can define T as multiplying them together to obtain an element in $\mathfrak{gl}(n,\mathbb{R})$. However, these two definitions may cause some misunderstandings.

Example 2.1.4. Let ω be a 1-form valued in $\mathfrak{g}$, then for vector fields X,Y, one has

$$\begin{aligned} \omega\wedge\omega(X,Y)&=T((\omega_1\wedge\omega_2)(X_1,X_2))\\ &=T(\frac{1}{1!\times 1!}(\omega(X)\otimes\omega(Y)-\omega(Y)\otimes\omega(X)))\\ &=[\omega(X),\omega(Y)]-[\omega(Y),\omega(X)]\\ &=2[\omega(X),\omega(Y)] \end{aligned}$$

Remark 2.1.2. If T is choose as in Remark 2.1.1, then in this case we have

$$\omega\wedge\omega(X,Y)=[\omega(X),\omega(Y)]$$

That's where misunderstanding lies. Different authors may use different notations, so be careful!

Proposition 2.1.3. Let ω be a 1-form valued in $\mathfrak{g}$, then

$$(\omega\wedge\omega)\wedge\omega=\omega\wedge(\omega\wedge\omega)=0$$

Proof. For arbitrary vector fields X, Y and Z, one has

$$\begin{aligned}(\omega\wedge\omega)\wedge\omega(X,Y,Z) &= \frac{1}{2!\times 1!}\{[\omega\wedge\omega(X,Y),\omega(Z)]+[\omega\wedge\omega(Y,Z),\omega(X)]+[\omega\wedge\omega(Z,X),\omega(Y)]\\ &\quad -[\omega\wedge\omega(Y,X),\omega(Z)]+[\omega\wedge\omega(Z,Y),\omega(X)]+[\omega\wedge\omega(X,Z),\omega(Y)]\}\\ &= \frac{2}{2!\times 1!}\{[[\omega(X),\omega(Y)],\omega(Z)]+[[\omega(Y),\omega(Z)],\omega(X)]+[[\omega(Z),\omega(X)],\omega(Y)]\\ &\quad -[[\omega(Y),\omega(X)],\omega(Z)]+[[\omega(Z),\omega(Y)],\omega(X)]+[[\omega(X),\omega(Z)],\omega(Y)]\}\end{aligned}$$

This equals to zero according to Jacobi identity of Lie bracket. □

Proposition 2.1.4. Let ω_1, ω_2 be forms valued in $\mathfrak{g}$ with degree k and l respectively, then

$$\omega_1\wedge\omega_2 = (-1)^{kl+1}\omega_2\wedge\omega_1$$

Proof. Note that for a k-form ω_1 and a l-form ω_2, we have

$$\omega_1\wedge\omega_2 = (-1)^{kl}\omega_2\wedge\omega_1$$

But in this case, there is one more -1 coming from Lie bracket. □

2.2. Maurer-Cartan form.

Example 2.2.1 (Maurer-Cartan form)**.** The Maurer-Cartan form θ, which is defined by

$$\theta_g := (L_{g^{-1}})_*$$

is a $\mathfrak{g}$-valued 1-form on G. Indeed, since tangent bundle of Lie group is trivial, so we may assume vector field X is left-invariant, then

$$\theta_g(X_g) = (L_{g^{-1}})_*(L_g)_*X_e = X_e \in \mathfrak{g}$$

where X_g means value of X at point $g\in G$.

Remark 2.2.1. If G is a matrix group, we also use $g^{-1}\mathrm{d}g$ to denote its Maurer-Cartan form, which is easy to compute. For example,

Example 2.2.2. Consider $G = \mathrm{SO}(2)\subset \mathrm{GL}(2,\mathbb{R})$. We may parametrize $\mathrm{SO}(2)$ by

$$g(\theta) = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

where $\theta\in\mathbb{R}$. Then directly compute we have

$$\begin{aligned}\omega &= g^{-1}\mathrm{d}g\\ &= \begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}-\sin\theta\mathrm{d}\theta & -\cos\theta\mathrm{d}\theta\\ \cos\theta\mathrm{d}\theta & -\sin\theta\mathrm{d}\theta\end{pmatrix}\\ &= \begin{pmatrix}0 & -\mathrm{d}\theta\\ \mathrm{d}\theta & 0\end{pmatrix}\end{aligned}$$

2.3. **Motivation.** In fact, here we use principal G-bundle as a tool to study geometry of vector bundle E, that is to give a connection on E, if E is an associated vector bundle of P. Recall a connection on E is defined as the following linear operator

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

satisfying Leibniz rule.

Suppose vector bundle E is associated to principal G-bundle $\pi : P \to M$, and written as $P \times_G V$, then from Proposition 1.3.1, we have a one to one correspondence between sections of E with G-equivariant maps from P to V. Given a section s of E, if we use s^P to denote the G-equivariant map obtained from one to one correspondence, it's easy to take derivatives of s^P to obtain a 1-form on P valued in V, that is a G-equivariant fiber-wise linear map from TP to V.

However, this 1-form does not by itself define a covariant derivative of s. As what we've defined, $\nabla s \in C^\infty(M, \Omega^1_M(E))$, so by Remark 1.3.4, a covariant derivative appears upstairs on P is supposed to be a G-equivariant section over $(P \times V) \otimes \pi^* T^* M$, that is a G-equivariant fiber-wise linear map from $\pi^* TM$ to V.

To see what is missing, it is important to keep in mind that TP has some properties that arise from the fact that P is a principal bundle over M. In fact, we have the following exact sequence

$$0 \to \ker \pi_* \to TP \to \pi^* TM \to 0 \tag{2.1}$$

This exact sequence is quite important, let's make following remarks:

Remark 2.3.1. The map from $\ker \pi_*$ is clearly an inclusion. And the map from TP to $\pi^* TM$ is characterized as follows

$$\begin{aligned} TP &\to \pi^* TM \subset P \times TM \\ v &\mapsto (p, \pi_* v) \end{aligned}$$

where $v \in T_p P$.

Remark 2.3.2. $\ker \pi_*$ is isomorphic to trivial bundle $P \times \mathfrak{g}$. Indeed, we have the following bundle isomorphism

$$\begin{aligned} \psi : P \times \mathfrak{g} &\to \ker \pi_* \\ (p, X) &\mapsto \sigma(X)_p := \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} pe^{tX} \end{aligned}$$

where $\sigma(X)_p$ means the value of $\sigma(X)$ at p. It's clear $\sigma(X) \in \ker \pi_*$, since for each $p \in P$,

$$\begin{aligned} \pi_*(\sigma(X)_p) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \pi(pe^{tX}) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \pi(p) \\ &= 0 \end{aligned}$$

Remark 2.3.3 (G-equivariance of exact sequence)*.* The action of G on P can be lifted to the exact sequence (2.1), as follows:

Let $R_g : P \to P$ denote the action of $g \in G$ on P, given by $p \mapsto pg$.

1. The G action on TP is given by $(R_g)_* : TP \to TP$, and it descends to $\ker \pi_*$ since if $v \in \ker \pi_*$, then

$$\begin{aligned}\pi_*((R_g)_* v) &= (\pi \circ R_g)_*(v)\\ &= \pi_*(v)\\ &= 0\end{aligned}$$

2. The G action on π^*TM is given by sending defined by sending a pair $(p, v) \in P \times TM$ to the pair (pg, v). $(pg, v) \in \pi^*TM$ since $\pi(pg) = \pi(p) = \pi(v)$.

Furthermore, we claim the exact sequence (2.1) is equivariant with respect to the lifts.

1. It automatically holds for inclusion map from $\ker \pi_*$ to TP, since G action on $\ker \pi_*$ is obtain from descending the one on TP;
2. It holds for the map from TP to π^*TM, since for $v \in TP$ we have $(R_g)_* v$ is sent to $(pg, \pi_*(R_g)_* v)$, that is exactly $(pg, \pi_* v)$, since $\pi \circ R_g = \pi$.

If we want to identify $\ker \pi_*$ as $P \times \mathfrak{g}$, we need to choose a G-action on $\mathfrak{g}$ properly such that the isomorphism ψ is G-equivariant. It turns out to be adjoint representation. Indeed, we compute as follows

$$\begin{aligned}(R_g)_* \psi(p, X) &= (R_g)_* \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} p \exp(tX) \right)\\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} p \exp(tX) g\\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (pg) \left(g^{-1} \exp(tX) g\right)\\ &= \psi(pg, \mathrm{Ad}(g^{-1})X)\end{aligned}$$

2.4. **Connection on principal bundle.** So if we want to obtain a fiber-wise linear map $\pi^*TM \to V$ from a fiber-wise linear map $TP \to V$, we need exact sequence (2.1) splitting. In other words, we desire there exists a G-equivariant $\omega : TP \to P \times \mathfrak{g}$, such that $\omega|_{P \times \mathfrak{g}}$ is identity. Such ω is called a connection on principal G-bundle P.

Definition 2.4.1 (connection on principal bundle)**.** Let $\pi : P \to M$ be a principal G-bundle. If $\omega \in C^\infty(P, \Omega^1_P(\mathfrak{g}))$ satisfies

1. For any $X \in \mathfrak{g}$, $\omega(\sigma(X)) = X$;
2. For any $g \in G$, $R_g^* \omega = \mathrm{Ad}(g^{-1}) \circ \omega$

Then ω is called a connection on P.

Notation 2.4.1. We use $\mathscr{A}(P)$ to denote the set of all connections on P.

Remark 2.4.1 (horizontal distribution viewpoint). If we define $H = \ker\omega$, then

$$TP = H \oplus (P \times \mathfrak{g})$$

such that $(R_g)_* H_p = H_{pg}$. H is called a horizontal distribution and $P \times \mathfrak{g}$ is called vertical distribution. Conversely, give a horizontal distribution, one can also construct a connection, they're the same things.

Example 2.4.1 (connection on trivial principal bundle)**.** Consider trivial principal G-bundle $P = M \times G$. Recall we have a Maurer-Cartan form θ, which is a 1-form valued in $\mathfrak{g}$. Then we can use $\pi_2 : M \times G \to G$ to pull it back to P to obtain a 1-form on P valued in $\mathfrak{g}$, which is called Maurer-Cartan form on trivial principal G-bundle, and it's denoted ω_{mc}. Now let's check ω_{mc} gives a connection on trivial principal bundle.

1. For any $X \in \mathfrak{g}$, we have

$$\begin{aligned}
\omega_{mc}(\sigma(X)) &= \pi_2^*\theta(\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}(x, g)e^{tX}) \\
&= \theta(\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} ge^{tX}) \\
&= (L_{g^{-1}})_*(\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} ge^{tX}) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} e^{tX} \\
&= X
\end{aligned}$$

2. It suffices to check $R_g^*\theta = \mathrm{Ad}(g^{-1}) \circ \theta$ holds for $g \in G$. For any left-invariant vector field X, recall that $\theta(X) = X_e$, thus

$$R_g^*\theta(X) = \theta((R_g)_*X) = ((R_g)_*X)_e = (L_{g^{-1}})_*(R_g)_*X_e$$

that's exactly $\mathrm{Ad}(g^{-1}) \circ \theta(X)$.

Remark 2.4.2. It's clear to see $\ker\omega_{mc} = \pi^*TM$, since ω_{mc} is pullback from a 1-form on G, thus in this case

$$TP = \pi^*TM \oplus \pi_2^*TG$$

that's exactly canonical splitting of TP.

2.5. **Gauge group.**

Definition 2.5.1 (gauge group)**.** For a principal G-bundle P, the gauge group $\mathscr{G}(P)$ is the group of G-automorphism of P, that is G-equivariant diffeomorphism $\Phi : P \to P$ such that $\pi = \pi \circ \Phi$.

Definition 2.5.2 (gauge transformation)**.** An element in $\mathscr{G}(P)$ is called gauge transformation.

Remark 2.5.1 (local expression of gauge transformation). For a gauge transformation Φ, if we consider its action on local trivialization $\varphi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha$, we have $\varphi_\alpha(\Phi(p)) = (\pi(p), g_\alpha(\Phi(p)))$, which induces a map $\bar{\phi}_\alpha : \pi^{-1}(U_\alpha) \to G$ by

$$\widetilde{\phi}_\alpha(p) = g_\alpha(\Phi(p))g_\alpha(p)^{-1}$$

By the equivariance of g_α and Φ we have $\widetilde{\phi}$ is G-invariant, which impiles $\widetilde{\phi}_\alpha(p) = \phi_\alpha(\pi(p))$ for some $\phi_\alpha : U_\alpha \to G$.

If we consider on the overlaps $x \in U_{\alpha\beta}$ with $p = \pi^{-1}(x)$. Then

$$\begin{aligned}
\phi_\alpha(x) &= g_\alpha(\Phi(p))g_\alpha(p)^{-1} \\
&= g_\alpha(\Phi(p))g_\beta(\Phi(p))^{-1}g_\beta(\Phi(p))g_\beta(p)^{-1}g_\beta(p)g_\alpha(p)^{-1} \\
&= g_{\alpha\beta}(x)\phi_\beta(x)g_{\alpha\beta}(x)^{-1}
\end{aligned}$$

This shows $\{\phi_\alpha\}$ defines a global section of associated bundle obtained from G acts on G by conjugation, that is $\operatorname{Conj} P$ defined in Example 1.3.2. In fact, we have the following one to one correspondence.

Proposition 2.5.1. There is one to one correspondence between the group $\mathscr{G}(P)$ and $C^\infty(M, \operatorname{Conj} P)$.

Proof. We have already seen that a gauge transformation can give an element in $C^\infty(M, \operatorname{Conj} P)$. Conversely, by Proposition 1.3.1, there is a one to one correspondence between $C^\infty(M, \operatorname{Conj} P)$ and smooth functions $f : P \to G$ which is G-equivariant. For such f, consider $\Phi_f : P \to P$ given by $\Phi_f(p) = pf(p)$.

1. $\pi \circ \Phi_f = \pi$, since $\pi \circ \Phi_f(p) = \pi(pf(p)) = \pi(p)$
2. It's G-equivariant since

$$\begin{aligned}
\Phi_f(pg) &= pgf(pg) \\
&= pgg^{-1}f(p)g \\
&= pf(p)g \\
&= \Phi_f(p)g
\end{aligned}$$

The two maps we constructed are clearly inverse to each other, giving the desired correspondence. □

Now we're going to show $\mathscr{G}(P)$ acts on $\mathscr{A}(P)$.

Lemma 2.5.1. For any $X \in \mathfrak{g}$ and $\Phi \in \mathscr{G}(P)$, then

$$\Phi_*(\sigma(X)) = \sigma(X)$$

Proof. Direct computation shows

$$\begin{aligned}\Phi_*\sigma(X) &= \Phi_*(\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} pe^{tX}) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \Phi(pe^{tX}) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \Phi(p)e^{tX} \\ &= \sigma(X)\end{aligned}$$

□

Proposition 2.5.2. $\mathscr{G}(P)$ acts on $\mathscr{A}(P)$ via pullback.

Proof. For $\omega \in \mathscr{A}(P)$ and $\Phi \in \mathscr{G}(P)$, let's check $\Phi^*\omega \in \mathscr{A}(P)$.

1. For any $X \in \mathfrak{g}$, we have
$$\begin{aligned}\Phi^*\omega(\sigma(X)) &= \omega(\Phi_*\sigma(X)) \\ &= \omega(\sigma(X)) \\ &= X\end{aligned}$$
2. Note that $R_g^*\Phi^* = (R_g \circ \Phi)^* = (\Phi \circ R_g)^*$, thus
$$\begin{aligned}R_g^*(\Phi^*\omega) &= \Phi^*(R_g^*\omega) \\ &= \Phi^*(\mathrm{Ad}(g^{-1}) \circ \omega) \\ &= \mathrm{Ad}(g^{-1}) \circ \Phi^*\omega\end{aligned}$$

□

Remark 2.5.2. Gauge theory concerns about "space" of orbit of $\mathscr{G}(P)$, that is $\mathscr{A}(P)/\mathscr{G}(P)$.

2.6. **Local expression of connection.** Instead of considering connection 1-form living on P, we want to convert it into the one living on base manifold M, since we want to use it to study connection of vector bundle over M. To do this, we divide the process into three steps:

1. Given a connection on trivial principal G-bundle, correspond it to a 1-form on M;
2. Figure out how does this correspondence transform under gauge transformation;
3. Since a G-principal is locally trivial, and you can regard transition functions as gauge transformation, then together above two step to conclude.

2.6.1. *Baby case.* Fix a trivial principal G-bundle $P = M \times G$ and the following notations:

1. $\pi : P \to M$ is natural projection, given by $p = (x, g) \mapsto x \in M$;
2. $i : M \to P$ is natural inclusion, given by $x \mapsto (x, e) \in P$.

Lemma 2.6.1. For any $A \in C^\infty(M, \Omega^1_M(\mathfrak{g}))$, there exists a unique $\widetilde{A} \in C^\infty(P, \Omega^1_P(\mathfrak{g}))$ such that

1. $i^*\widetilde{A} = A$;
2. $\widetilde{A}(\sigma(X)) = 0$, where $X \in \mathfrak{g}$;
3. $R_g^*\widetilde{A} = \mathrm{Ad}(g^{-1}) \circ \widetilde{A}$.

Proof. Let's construct $\widetilde{A}$ pointwise.

(a) For $p = (x, e) \in M \times G$, we have

$$T_pP = T_xM \oplus \mathfrak{g}$$

Then $\widetilde{A}$ is uniquely determined at this point according to (1) and (2).

(b) At point $p' = (x, g)$, it's clear $p' = pg$ and $(R_g)_* : T_pP \to T_{p'}P$ is an isomorphism, then for arbitrary $v \in T_{p'}P$, we may assume $v = (R_g)_*w$ for some $w \in T_pP$, then

$$\begin{aligned} \widetilde{A}_{p'}(v) &= \widetilde{A}_{pg}((R_g)_*w) \\ &= (R_g^*\widetilde{A})_p(w) \\ &= \mathrm{Ad}(g^{-1}) \circ \widetilde{A}(w) \end{aligned}$$

□

Proposition 2.6.1. $i^* : \mathscr{A}(P) \to C^\infty(M, \Omega^1_M(\mathfrak{g}))$ is bijective, that is

$$\begin{tikzcd} C^\infty(P, \Omega^1_P(\mathfrak{g})) \arrow[r, "i^*"] & C^\infty(M, \Omega^1_M(\mathfrak{g})) \\ \mathscr{A}(P) \arrow[u, hook] \arrow[ur, "1-1"] & \end{tikzcd}$$

Proof. For any $A \in C^\infty(M, \Omega^1_M(\mathfrak{g}))$, by Lemma 2.6.1 we have $\omega_{mc} + \widetilde{A}$ is also a connection on P. Thus consider

$$\begin{aligned} \tau : \Omega^1_M(\mathfrak{g}) &\to \mathscr{A}(P) \\ A &\mapsto \omega_{mc} + \widetilde{A} \end{aligned}$$

It's clear τ is surjective, since for any $\omega \in \mathscr{A}(P)$, we have

$$\tau(i^*(\omega - \omega_{mc})) = \omega_{mc} + \omega - \omega_{mc} = \omega$$

Now it suffices to show $i^*\tau = \mathrm{id}$, which impiles τ is injective thus bijective. Indeed, for $A \in \Omega^1_M(\mathfrak{g})$,

$$i^*\tau(A) = i^*(\omega_{mc} + \widetilde{A}) = i^*\widetilde{A} = A$$

since $i^*\omega_{mc} = 0$. □

2.6.2. *How to glue.* For gauge transformation Φ we can write it as

$$\Phi(x,g)=(x,\varphi(x)g)$$

where $\varphi: M\to G$ is smooth. So for any $\omega\in\mathscr{A}(P)$, if we write it as $\omega=\omega_{mc}+\widetilde{A}$. Then

$$\begin{aligned}i^*\Phi^*\omega&=i^*\Phi^*(\omega_{mc}+\widetilde{A})\\&=i^*\Phi^*\pi_2^*\theta+i^*\Phi^*\widetilde{A}\\&=\varphi^*\theta+i^*\Phi^*\widetilde{A}\end{aligned}$$

since $\pi_2\circ\Phi\circ i(x)=\pi_2\circ\Phi(x,e)=\pi_2(x,\varphi(x))=\varphi(x)$ for $x\in M$. So it suffices to compute $i^*\Phi^*\widetilde{A}$. For any vector field X, we have

$$\begin{aligned}(i^*\Phi^*\widetilde{A})(X)&=\widetilde{A}(\Phi_*i_*(X))\\&=\widetilde{A}(\varphi_*X)\\&=\mathrm{Ad}(\varphi^{-1})\circ\widetilde{A}(X)\end{aligned}$$

Thus we have

$$i^*(\Phi^*\omega)=\varphi^*\theta+\mathrm{Ad}(\varphi^{-1})\circ\widetilde{A}$$

2.6.3. *General case.* Let $\pi:P\to M$ be a principal G-bundle with local trivialization $\{(U_\alpha,\varphi_\alpha)\}$, we use $i_\alpha:U_\alpha\to U_\alpha\times G$ to denote natural inclusion. For a connection $\omega\in\mathscr{A}(P)$, then we can write it locally on

$$i_\alpha^*(\varphi_\alpha^{-1})^*\omega|_{\pi^{-1}(U_\alpha)}=A_\alpha\in\Omega^1_{U_\alpha}(\mathfrak{g})$$

Furthermore,

$$A_\alpha=\mathrm{Ad}(g_{\alpha\beta}^{-1})\circ A_\beta+g_{\alpha\beta}^*\theta$$

where $g_{\alpha\beta}:U_\alpha\cap U_\beta\to G$ are transition functions. Thus we have the following one to one correspondence.

Proposition 2.6.2.

$$\mathscr{A}(P)\overset{1-1}{\Longleftrightarrow}\{(A_\alpha)\in\prod_\alpha\Omega^1_{U_\alpha}(\mathfrak{g})\mid A_\alpha=\mathrm{Ad}(g_{\alpha\beta}^{-1})\circ A_\beta+g_{\alpha\beta}^*\theta\}$$

Remark 2.6.1. From this viewpoint, for two connection ω_1 and ω_2, it's clear $\omega_1-\omega_2$ gives a global section of associated vector bundle $\mathrm{Ad}\,P$, thus $\mathscr{A}(P)$ is an affine space modelled on $C^\infty(M,\mathrm{Ad}\,P)$.

3. Curvature of principal bundle

3.1. Definition.

Definition 3.1.1 (curvature)**.** Let P be a principal G-bundle and $\omega \in \mathscr{A}(P)$. Curvature of ω is defined as

$$\Omega := \mathrm{d}\omega + \frac{1}{2}\omega \wedge \omega \in C^{\infty}(P, \Omega^2_P(\mathfrak{g}))$$

Example 3.1.1. If $P = M \times G$, and $\omega = \omega_{mc}$, then $\Omega = 0$.

Proof. It suffices to check Maurer-Cartan form $\theta \in \Omega^1_G(\mathfrak{g})$ satisfying

$$\mathrm{d}\theta + \frac{1}{2}\theta \wedge \theta = 0$$

which is called Maurer-Cartan equation. Let X, Y are two left-invariant vector fields on G, then

$$\theta(X) = (L_{g^{-1}})_* X_g = (L_{g^{-1}})_*(L_g)_* X_e = X_e$$

is constant. Thus

$$\mathrm{d}\theta(X, Y) = -\theta([X, Y]) = -\frac{1}{2}\theta \wedge \theta(X, Y)$$

since $X(\theta(Y)) = Y(\theta(X)) = 0$. □

Theorem 3.1.1 (Bianchi identity)**.**

$$\mathrm{d}\Omega + \omega \wedge \Omega = 0$$

Proof.

$$\begin{aligned}
\mathrm{d}\Omega &= \mathrm{d}(\mathrm{d}\omega + \frac{1}{2}\omega \wedge \omega) \\
&= \frac{1}{2}\mathrm{d}\omega \wedge \omega - \frac{1}{2}\omega \wedge \mathrm{d}\omega \\
&= -\omega \wedge \mathrm{d}\omega \\
&= -\omega \wedge (\Omega - \frac{1}{2}\omega \wedge \omega) \\
&= -\omega \wedge \Omega
\end{aligned}$$

□

Definition 3.1.2 (horizontal form)**.** Let ω be a 2-form on P valued in vector space V, it's called horizontal, if $\omega(\sigma(X), \text{-}) = 0$ for arbitrary $X \in \mathfrak{g}$.

Proposition 3.1.1. Ω is a horizontal 2-form.

Proof. Divide computations into two parts:

1. If $X, Y \in \mathfrak{g}$ and write $X' = \sigma(X), Y' = \sigma(Y)$, then

$$\begin{aligned}
\mathrm{d}\omega(X', Y') &= X'(\omega(Y')) - Y'\omega(X') - \omega([X', Y']) \\
&= X'(Y) - Y'(X) - [X, Y] \\
&= -[X, Y] \\
&= -\frac{1}{2}\omega \wedge \omega(X', Y')
\end{aligned}$$

2. If $X \in \mathfrak{g}$ and Y is a horizontal vector field, note that
$$\frac{1}{2}\omega \wedge \omega(\sigma(X), Y) = 0$$
since $\omega(Y) = 0$. So it suffices to compute
$$\begin{aligned} \mathrm{d}\omega(\sigma(X), Y) &= \sigma(X)(\omega(Y)) - Y\omega(\sigma(X)) - \omega([\sigma(X), Y]) \\ &= -\omega([\sigma(X), Y]) \\ &= -\omega(\mathscr{L}_{\sigma(X)}Y) \end{aligned}$$
However, we have
$$\mathscr{L}_{\sigma(X)}Y = \lim_{t\to 0} \frac{Y \circ \phi_t - Y}{t}$$
where ϕ_t is the flow generated by $\sigma(X)$, thus it's clear $\omega(\mathscr{L}_{\sigma(X)}Y) = 0$.

□

Remark 3.1.1. Now let's give another explaination about horizontal: Given a horizontal distribution $H \subset TP$, we define the horizontal projection $h : TP \to TP$ to be the projection onto the horizontal distribution along the vertical distribution. Since both H and V are invariant under the action of G, so is h.

Then $\Omega = h^*\mathrm{d}\omega$. Indeed, it suffices to show for vector fields X, Y, one has
$$\mathrm{d}\omega(hX, hY) = \mathrm{d}\omega(X, Y) + \frac{1}{2}\omega \wedge \omega(X, Y)$$
Consider the following cases:

1. Let X, Y be horizontal. In this case there is nothing to prove, since $\omega(X) = \omega(Y) = 0$ and $hX = X, hY = Y$;
2. If one of X, Y are vertical, then it's clear both sides are zero, since both Ω and $h^*\mathrm{d}\omega$ are horizontal.

That is, $\Omega(X, Y) = 0$ if and only if $[hX, hY]$ is horizontal. In other words, the curvature of the connection measures the failure of integrability of the horizontal distribution $H \subset TP$.

3.2. **Local expression of curvature.** Let $\pi : P \to M$ be a principal G-bundle with local trivialization $\{(U_\alpha, \varphi_\alpha)\}$, we use $i_\alpha : U_\alpha \to U_\alpha \times G$ to denote natural inclusion. For connection $\omega \in \mathscr{A}(P)$, its curvature is defined as
$$\Omega = \mathrm{d}\omega + \frac{1}{2}\omega \wedge \omega$$
If we define $\Omega_\alpha = (\varphi_\alpha^{-1})^*\Omega$, which is a 2-form on $U_\alpha \times G$, and
$$F_\alpha := i_\alpha^*\Omega_\alpha \in C^\infty(U_\alpha, \Omega^1_{U_\alpha}(\mathfrak{g}))$$
By definition one has
$$F_\alpha = \mathrm{d}A_\alpha + \frac{1}{2}A_\alpha \wedge A_\alpha$$

Now we're going to show on $U_{\alpha\beta}$, one has

$$F_\beta = \mathrm{Ad}(g_{\alpha\beta}^{-1}) \circ F_\alpha$$

where $g_{\alpha\beta} : U_{\alpha\beta} \to G$ are transition function. Note that

$$\begin{aligned} F_\alpha &= \mathrm{d}A_\alpha + \frac{1}{2}A_\alpha \wedge A_\alpha \\ &= \mathrm{d}(\mathrm{Ad}(g_{\alpha\beta}^{-1}) \circ A_\beta + g_{\alpha\beta}^*\theta) + \frac{1}{2}(\mathrm{Ad}(g_{\alpha\beta}^{-1}) \circ A_\beta + g_{\alpha\beta}^*\theta) \wedge (\mathrm{Ad}(g_{\alpha\beta}^{-1}) \circ A_\beta + g_{\alpha\beta}^*\theta) \end{aligned}$$

Since θ satisfies Maurer-Cartan equation, one has

$$g_{\alpha\beta}^*(\mathrm{d}\theta + \frac{1}{2}\theta \wedge \theta) = 0$$

In order to give a neat computation of Ad, we here assume G is a matrix group[2]. Then

$$\begin{aligned} \mathrm{d}(\mathrm{Ad}(g_{\alpha\beta}^{-1}) \circ A_\beta) &= \mathrm{d}(g_{\alpha\beta}^{-1}A_\beta g_{\alpha\beta}) \\ &= \mathrm{d}g_{\alpha\beta}^{-1}A_\beta g_{\alpha\beta} + g_{\alpha\beta}^{-1}\mathrm{d}A_\beta g_{\alpha\beta} + g_{\alpha\beta}^{-1}A_\beta \mathrm{d}g_{\alpha\beta} \\ &= \mathrm{d}g_{\alpha\beta}^{-1}A_\beta g_{\alpha\beta} + g_{\alpha\beta}^{-1}A_\beta \mathrm{d}g_{\alpha\beta} + \mathrm{Ad}(g_{\alpha\beta}^{-1}) \circ \mathrm{d}A_\beta \end{aligned}$$

And

$$\begin{aligned} \mathrm{d}g_{\alpha\beta}^{-1}A_\beta g_{\alpha\beta} + g_{\alpha\beta}^{-1}A_\beta \mathrm{d}g_{\alpha\beta} &= -g_{\alpha\beta}^{-1}\mathrm{d}g_{\alpha\beta} \wedge g_{\alpha\beta}^{-1}A_\beta g_{\alpha\beta} + g_{\alpha\beta}^{-1}A_\beta g_{\alpha\beta}g_{\alpha\beta}^{-1}\mathrm{d}g_{\alpha\beta} \\ &= -g_{\alpha\beta}^*\theta \wedge \mathrm{Ad}(g_{\alpha\beta}^{-1})A_\beta + \mathrm{Ad}(g_{\alpha\beta}^{-1})A_\beta \wedge g_{\alpha\beta}^*\theta \end{aligned}$$

Remark 3.2.1. In order words, $\{F_\alpha\}$ gives a global section of $\Omega_M^2(\mathrm{Ad}\,P)$, which is denoted by F_ω.

3.3. **Basic forms.**

Recall our curvature form $\Omega \in C^\infty(P, \Omega_P^2(\mathfrak{g}))$ has the following properties:

1. Ω is horizontal;
2. It's Ad-equivariant, that is

$$R_g^*\Omega = \mathrm{Ad}(g^{-1}) \circ \Omega$$

Definition 3.3.1 (basic form)**.** Let $\rho : G \to \mathrm{GL}(V)$ be a representation of G, a k-form on P valued in V is called basic, if it satisfies

1. Ω is horizontal;
2. It's ρ-equivariant, that is

$$R_g^*\Omega = \rho(g^{-1}) \circ \Omega$$

The set of all basic k-forms is denoted by $C^\infty(P, \Omega_P^k(V))^{\mathrm{basic}}$.

Example 3.3.1. The curvature form is a basic 2-form.

[2]In fact, most interesting cases we're concern about are matrix group

Proposition 3.3.1. Let $\rho: G \to \mathrm{GL}(V)$ be a representation of G, if we use E to denote associated vector bundle $P \times_\rho V$, then there is an one to one correspondence

$$C^\infty(P, \Omega^k_P(V))^{\text{basic}} \overset{1-1}{\Longleftrightarrow} C^\infty(M, \Omega^k_M(E))$$

Proof. Given $\widetilde{\omega} \in C^\infty(P, \Omega^k_P(V))^{\text{basic}}$, now we're going to construct a $\omega \in C^\infty(M, \Omega^k_M(E))$ pointwise, that is for arbitrary $x \in M$ and $v_1, \dots, v_k \in T_xM$, we give an assignment:

$$\omega_x(v_1, \dots, v_k) \in E_x$$

Recall that E_x is an equivalent class $[p, v]$, where $p \in P, v \in V$. Choose arbitrary $p \in \pi^{-1}(x) \in P$, since $\pi_*: T_pP \to T_xM$ is surjective, we can choose $\widetilde{v}_i \in T_pP$ such that $\pi_*(\widetilde{v}_i) = v_i, i = 1, \dots, k$.

Now we define

$$\omega_x(v_1, \dots, v_k) := [(p, \widetilde{\omega}_x(\widetilde{v}_1, \dots, \widetilde{v}_k))]$$

It's well-defined, that is it is independent of the choice of p and $\widetilde{v}_1, \dots, \widetilde{v}_k$. Indeed, choose $p' = pg \in \pi^{-1}(x)$ and $\widetilde{v}'_1, \dots, \widetilde{v}'_k \in T_{p'}P$ with $\pi_*(\widetilde{v}'_i) = X_i, i = 1, \dots, k$. Note that for each i, one has $(R_g)_*\widetilde{v}_i - \widetilde{v}'_i$ is vertical, since $\pi_*((R_g)_*\widetilde{v}_i - \widetilde{v}'_i) = 0$. Thus

$$\begin{aligned}
\widetilde{\omega}_{p'}(\widetilde{v}'_1, \dots, \widetilde{v}'_k) &\overset{(1)}{=} \widetilde{\omega}_{p'}((R_g)_*\widetilde{v}_1, \dots, (R_g)_*\widetilde{v}_k) \\
&= (R_g^*\omega)_p(\widetilde{v}_1, \dots, \widetilde{v}_k) \\
&\overset{(2)}{=} \rho(g^{-1}) \circ \omega_p(\widetilde{v}_1, \dots, \widetilde{v}_k)
\end{aligned}$$

where

(1) holds from $\widetilde{\omega}$ is horizontal;
(2) holds from $\widetilde{\omega}$ is G-equivariant.

This shows ω is well-defined, since $[(p, \widetilde{\omega}_x(\widetilde{v}_1, \dots, \widetilde{v}_k))] = [(p', \rho(g^{-1}) \circ \omega_p(\widetilde{v}_1, \dots, \widetilde{v}_k))]$ in E. Conversely, from above construction, there is a formula

$$\omega_x(X_1, \dots, X_k) = [(p, \widetilde{\omega}_p(\widetilde{X}_1, \dots, \widetilde{X}_k))] \tag{3.1}$$

So it's clear how to construct $\widetilde{\omega}$ when you have $\omega \in C^\infty(M, \Omega^k_M(E))$. □

Remark 3.3.1. Above proposition is the key tool to study the geometry of vector bundle E via principal bundle, if E can be constructed as an associated vector bundle $P \times_\rho V$. Furthermore, it gives a explicit proof of motivation we said in Remark 1.3.4. In particular, if we consider the case $k = 0$, then we have

$$\{f: P \to V \mid f(xg) = \rho(g^{-1})f(x)\} \overset{1-1}{\Longleftrightarrow} C^\infty(M, E)$$

since the former is exactly $\Omega^0_P(V)^{\text{basic}}$. This shows Proposition 1.3.1 again.

3.4. **Relations between connections on principal bundle and its associated bundle.** Now we're going to define connection on vector bundle $E = P \times_\rho V$ using connection ω on principal G-bundle P. Thanks to Proposition 3.3.1, it suffices to construct

$$\mathrm{d}_\omega : C^\infty(P, \Omega_P^0(V))^{\mathrm{basic}} \to C^\infty(P, \Omega_P^1(V))^{\mathrm{basic}}$$

Note that there is a natural exterior derivative

$$\mathrm{d} : C^\infty(P, \Omega_P^0(V)) \to C^\infty(P, \Omega_P^1(V))$$

However, it may not descend down to basic forms. Here we define

$$\begin{aligned} \mathrm{d}_\omega : C^\infty(P, \Omega_P^0(V))^{\mathrm{basic}} &\to C^\infty(P, \Omega_P^1(V))^{\mathrm{basic}} \\ s &\mapsto \mathrm{d}s + \rho_*(\omega)s \end{aligned}$$

where $\rho_* : \mathfrak{g} \to \mathfrak{gl}(V)$ is representation of Lie algebra induced by ρ. Let's show d_ω is well-defined.

1. It's G-equivariant:

$$\begin{aligned} (R_g)^*(\mathrm{d}_\omega s) &= R_g^*(\mathrm{d}s + \rho_*(\omega)s) \\ &= \mathrm{d}R_g^* s + R_g^*(\rho_*(\omega)s) \\ &= \mathrm{d}R_g^* s + \rho_*(R_g^*\omega)R_g^* s \\ &= \mathrm{d}R_g^* s + \rho_*(\mathrm{Ad}(g^{-1})\omega)R_g^* s \\ &= \mathrm{d}R_g^* s + \mathrm{Ad}(\rho(g^{-1}))(\rho_*\omega)\rho_g^* s \\ &= \mathrm{d}\rho(g^{-1})s + \mathrm{Ad}(\rho(g^{-1}))(\rho_*\omega)\rho(g^{-1})s \\ &= \mathrm{d}(\rho(g^{-1})s) + \rho(g^{-1})(\rho_*(\omega)s) \\ &= \rho(g^{-1})\mathrm{d}_\omega s \end{aligned}$$

2. It's horizontal: For arbitrary vertical $\sigma(X)$

$$\begin{aligned} \mathrm{d}_\omega(s) &= \mathrm{d}s(\sigma(X)) + \rho_*(\omega(\sigma(X)))s \\ &= \mathrm{d}s(\sigma(X)) + \rho_*(X)s \end{aligned}$$

So it suffices to check

$$\sigma(X)(s) = -\rho_*(X)s$$

Indeed,

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} se^{tX} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho(e^{-tX})s$$

More generally, we can define

$$\begin{aligned} \mathrm{d}_\omega : C^\infty(P, \Omega_P^k(V))^{\mathrm{basic}} &\to C^\infty(P, \Omega_P^{k+1}(V))^{\mathrm{basic}} \\ s &\mapsto \mathrm{d}s + \rho_*(\omega) \wedge s \end{aligned}$$

And one can check it's well-defined by the same method as above.

Remark 3.4.1. The case we're most interested in is $V = \mathfrak{g}$ and ρ is adjoint representation. Since in this case $\rho_*(X)$ acts on Y is exactly $[X, Y]$, where $X, Y \in \mathfrak{g}$. In particular,

$$\mathrm{d}_\omega(s) = \mathrm{d}s + \omega \wedge s$$

where $s \in C^\infty(P, \Omega^k_P(\mathfrak{g}))^{\text{basic}}$ and above wedge is wedge of forms valued in $\mathfrak{g}$.

Now we're going to back to base manifold M to give description of connection ∇ on E. Here are two methods:

1. Take $E = \Omega^1_M(P \times_{\mathrm{Ad}} \mathfrak{g})$ as an example, since we need this example later. If U_α is a local trivialization and $\pi_\alpha : U_\alpha \times G \to U_\alpha$ is the projection to the first factor, section s of E on U_α is a section of $\Omega^1_{U_\alpha}(\mathfrak{g})$, denoted by s_α, then $\widetilde{s}_\alpha := \pi^*_\alpha s_\alpha$ is a basic 1-form valued in $\mathfrak{g}$, and $\mathrm{d}_\omega \widetilde{s}_\alpha = \mathrm{d}\widetilde{s}_\alpha + \omega \wedge \widetilde{s}_\alpha$, then by using $i_\alpha : U_\alpha \to U_\alpha \times G$ to pullback, one has
$$\mathrm{d}s_\alpha + A_\alpha \wedge s_\alpha \tag{3.2}$$
where $A_\alpha \in C^\infty(U_\alpha, \Omega^1_{U_\alpha}(\mathfrak{g}))$ is given by ω.
2. Let X be a vector field on M, s a section of E, we can give an explicit formula of $\nabla_X s$ via formula (3.1). For $x \in M$, choose $p \in \pi^{-1}(x)$, $\widetilde{X}$ is horizontal such that $\pi_*(\widetilde{X}) = X$ and $\widetilde{s} : P \to V$ is G-equivariant map which corresponds to s, then
$$\begin{aligned}(\nabla_X s)_x &= [(p, \widetilde{\nabla}\widetilde{s}(\widetilde{X}_p)) \\ &= [(p, (\mathrm{d}\widetilde{s} + \rho_*(\omega)\widetilde{s})(\widetilde{X}_p))] \\ &\overset{(1)}{=} [(p, \mathrm{d}\widetilde{s}(\widetilde{X}_p))] \\ &= [(p, \widetilde{X}_p(\widetilde{s}))]\end{aligned}$$
where (1) holds from $\widetilde{X}$ is horizontal.

From the second method, it's easy to see write

$$([\nabla_X, \nabla_Y]s - \nabla_{[X,Y]}s)_x = [(p, -\rho_*(\omega([\widetilde{X}, \widetilde{Y}]_p))\widetilde{s})]$$

here X, Y are vector fields on M. A natural question is what's the relation of curvature of ω and ∇.

Recall curvature of ω, denoted by Ω, is basic 2-form, and it can be regarded as a section of $\Omega^2_M(\mathrm{Ad}\, P)$, we use $\Omega \mapsto \Theta$ to denote this correspondence. Now let's compute Θ via formula (3.1): For $x \in M, v, w \in T_x M$, choose $\widetilde{v}, \widetilde{w}$ such that $\pi_*(\widetilde{v}) = v$ and $\pi_*(\widetilde{w}) = w$. Without lose of generality, we may assume $\widetilde{v}, \widetilde{w}$ are horizontal, then

$$\Theta_x(v, w) = [(p, \Omega_p(\widetilde{v}, \widetilde{w}))] \in (\mathrm{Ad}\, P)_x$$

Note that

$$\begin{aligned}\Omega_p(\widetilde{v},\widetilde{w}) &= \mathrm{d}\omega(\widetilde{v},\widetilde{w}) + \frac{1}{2}\omega\wedge\omega(\widetilde{v},\widetilde{w}) \\ &\overset{(1)}{=} \mathrm{d}\omega(\widetilde{v},\widetilde{w}) \\ &\overset{(2)}{=} -\omega([\widetilde{v},\widetilde{w}])\end{aligned}$$

where

(1) holds from $\widetilde{v},\widetilde{w}$ are horizontal;
(2) holds from Cartan's formula.

Note that $\operatorname{Ad} P$ can act on $P\times_\rho V$ as follows

$$[(p,X)]\times[(p,v)]\mapsto[(p,\rho_*(X)v)]$$

So $\Theta_x(v,w)$ can act on E_x, that is $\Theta\in C^\infty(M,\Omega^2_M(\operatorname{End} E))$. Thus we have the following theorem.

Theorem 3.4.1. Let $\pi: P\to M$ be a principal G-bundle, $E=P\times_\rho V$ an associated vector bundle of P. For vector fields X,Y over M and section s of E, then

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

4. Flat connection and holonomy

4.1. **Lifting of curves.** Let $\pi : P \to M$ be a principal G-bundle equipped with connection ω, consider smooth curve $\gamma : [0,1] \to M$ and a point $p \in \pi^{-1}(\gamma(0))$, we claim there exists a unique smooth map $\widetilde{\gamma} : [0,1] \to P$ such that

1. The following diagram commutes:

$$\begin{tikzcd} & P \arrow[d, "\pi"] \\ {[0,1]} \arrow[ur, dashed, "\widetilde{\gamma}"] \arrow[r, "\gamma"] & M \end{tikzcd}$$

2. $\widetilde{\gamma}'(t)$ is horizontal;
3. $\widetilde{\gamma}(0) = p$.

Proof. For convenience we assume G is a matrix group, and without lose of generality, we may assume P is trivial principal G-bundle $M \times G$, since it's a local problem. In this case we write $\widetilde{\gamma} = (\gamma(t), g(t))$, it's clear $\pi \circ \widetilde{\gamma} = \gamma$.

For conditions (2) and (3), it's an ODE with initial value in fact: Note that we can write connection $\omega = \omega_{mc} + \widetilde{A}$, so $\widetilde{\gamma}'(t)$ is horizontal if and only if

$$\begin{aligned}(\omega_{mc} + \widetilde{A})(\widetilde{\gamma}'(t)) &= (\omega_{mc} + \widetilde{A})((\gamma'(t), g'(t)) \\ &= g^{-1}(t)g'(t) + \widetilde{A}((\gamma'(t), g'(t))) \\ &= g^{-1}(t)g'(t) + \mathrm{Ad}(g^{-1}(t)) \circ A_{\gamma(t)}(\gamma'(t)) \\ &= g^{-1}(t)g'(t) + g^{-1}(t)A_{\gamma(t)}(\gamma'(t))g(t) \\ &= 0\end{aligned}$$

This completes the proof. □

4.2. **Flat connection.**

Definition 4.2.1 (flat connection)**.** Let P be a principal G-bundle, a connection $\omega \in \mathscr{A}(P)$ is called flat, if its curvature form $\Omega = 0$.

Theorem 4.2.1. The following are equivalent:

1. ω is flat;
2. For any $p \in M$, there exists $p \in U \subset M$ and local trivialization $\varphi : \pi^{-1}(U) \to U \times G$ such that $\omega|_U = \varphi^* \omega_{mc}$.

Proof. Hallmark of proof is to see curvature vanishes if and only if horizontal distribution is integrable. □

Remark 4.2.1. From this theorem, we can see a flat connection is just a topology information.

Corollary 4.2.1. The following are equivalent:

1. There is a flat connection on P;

2. There is a local trivialization $\varphi_\alpha : P|_{U_\alpha} \to U_\alpha \times G$ such that transition functions $\{g_{\alpha\beta} : U_\alpha \cap U_\beta \to G\}$ are locally constant.

Proof. From (2) to (1): Note that a connection on P is given by the following data:

$$A_\beta = \mathrm{Ad}(g_{\alpha\beta}^{-1})A_\alpha + g_{\alpha\beta}^*\theta$$

If $g_{\alpha\beta}$ are locally constant, then $g_{\alpha\beta}^*\omega = 0$, so just take all $A_\alpha = 0$ to obtain a flat connection.

From (1) to (2): If ω is a flat connection, then we can choose a local trivialization $\{U_\alpha, \varphi_\alpha\}$ such that $\omega|_{\pi^{-1}(U_\alpha)}$ are $\varphi_\alpha^*\omega_{mc}$. In this local trivialization, we have all $A_\alpha = 0$, thus $g_{\alpha\beta}^*\theta = 0$, which impiles $g_{\alpha\beta}$ is locally constant. $\square$

4.3. **Holonomy.** Now we give a smooth closed curve $\gamma : [0,1] \to M$ and $p \in \pi^{-1}(\gamma(0))$. Consider its lifting $\widetilde{\gamma} : [0,1] \to P$, note that

$$\widetilde{\gamma}(1) \in \pi^{-1}(\gamma(1)) = \pi^{-1}(\gamma(0))$$

So there exists $g \in G$ such that $\widetilde{\gamma}(1) = \widetilde{\gamma}(0)g$, since P_p is an orbit of G. Such element g is called holonomy, and it's denoted by $\mathrm{Hol}(\gamma, p)$, since it only depends on γ and p.

Proposition 4.3.1. For holonomy, the following properties hold.

1. If we change base point p to pg, then

$$\mathrm{Hol}(\gamma, pg) = g^{-1}\,\mathrm{Hol}(\gamma, p)g$$

2. For two smooth closed curves γ_1, γ_2, we have

$$\mathrm{Hol}(\gamma_1\gamma_2, p) = \mathrm{Hol}(\gamma_1, p)\,\mathrm{Hol}(\gamma_2, p)$$

Proof. Clear. $\square$

From (2) of above proposition, Hol can be regarded as a group homomorphism to some extend, so if we want to give a homomorphism

$$\mathrm{Hol} : \pi_1(M) \to G$$

It suffices to check when $\mathrm{Hol}(\gamma, p)$ is independent of homotopy class. Consider the following homotopy

$$\gamma_s : (-\varepsilon, \varepsilon) \times [0,1] \to M$$

such that $\gamma_0 = \gamma$. If we write its lifting on local trivialization as $\widetilde{\gamma}_s(t) = (\gamma_s(t), g_s(t))$, then the following equation holds

$$\frac{\partial g_s}{\partial t}(t) + A_{\gamma(t)}(\frac{\partial g_s}{\partial t}(t))g_s(t) = 0$$

So if ω is a flat connection, then it reduces to for arbitrary $s \in (-\varepsilon, \varepsilon)$, one has $\frac{\partial g_s}{\partial t}(t) = 0$. This shows it's independent of s.

Theorem 4.3.1 (Riemann-Hilbert correspondence)**.** There is a one to one correspondence

$$\{\text{flat connections on } P\}/\text{isomorphism} \overset{1-1}{\Longleftrightarrow} \mathrm{Hom}(\pi_1(M), G)/\text{conjugate}$$

Part 2. Chern-Weil theory

5. Chern-Weil theory

5.1. **G-invariant polynomial.** Let G be a Lie group with Lie algebra $\mathfrak{g}$, note that G can act on $\mathfrak{g}$ via adjoint representation, then G can also act on dual space of $\mathfrak{g}$, that is $\mathfrak{g}^*$, and thus on $\operatorname{Sym}^k \mathfrak{g}^*$. To be explicit, for $p \in \operatorname{Sym}^k \mathfrak{g}^*$ and $g \in G$, one has

$$gp(x_1, \dots, x_k) := p(\operatorname{Ad}(g)x_1, \dots, \operatorname{Ad}(g)x_k)$$

Definition 5.1.1 (G-invariant polynomial)**.** The set of G-invariant polynomials of degree k is

$$I^k(\mathfrak{g}) := \{p \in \operatorname{Sym}^k \mathfrak{g}^* \mid gp = p, \forall g \in G\}$$

and

$$I(\mathfrak{g}) := \bigoplus_{k \geq 0} I^k(\mathfrak{g})$$

5.2. **Chern-Weil homomorphism.** Let $\pi : P \to M$ be a principal G-bundle, and ω is a connection on P with curvature Ω.

Proposition 5.2.1. For $p \in I^k(\mathfrak{g})$,

$$p(\Omega) := p \circ (\underbrace{\Omega \wedge \cdots \wedge \Omega}_{k \text{ times}})$$

is a $2k$-form defined on P. Then

1. $p(\Omega)$ is horizontal, G-invariant, closed;
2. There exists a unique $2k$-form $p(\Theta)$ on M such that $\pi^*(p(\Theta)) = p(\Omega)$ and $\mathrm{d}p(\Theta) = 0$;
3. $[p(\Theta)] \in H^{2k}(M)$ is independent of the choice of connection ω.

Proof. For (1). It's clear $p(\Omega)$ is horizontal, since Ω is horizontal; To see it's G-invariant, note that

$$\begin{aligned} R_g^*(p(\Omega)) &= p(R_g^*\Omega) \\ &\overset{(a)}{=} p(\operatorname{Ad}(g^{-1}) \circ \Omega) \\ &\overset{(b)}{=} p(\Omega) \end{aligned}$$

where

(a) holds from Ω is G-equivariant;
(b) holds from p is G-invariant.

To see it's closed,

$$\mathrm{d}p(\Omega) = p(\mathrm{d}\Omega \wedge \Omega \wedge \cdots \wedge \Omega + \Omega \wedge \mathrm{d}\Omega \wedge \cdots \wedge \Omega + \dots)$$

Bianchi identity impiles

$$\mathrm{d}\Omega + \omega \wedge \Omega = 0$$

If we substitute $\mathrm{d}\Omega$ by $-\omega\wedge\Omega$ in above, then it suffices to show $\mathrm{d}p(\Omega)$ is horizontal. To see this, given a vertical vector field X, then $\mathscr{L}_X p(\Omega)=0$, since $p(\Omega)$ is horizontal, then by Cartan formula

$$\begin{aligned}0&=\mathscr{L}_X p(\Omega)\\&=\mathrm{d}\iota_X p(\Omega)+\iota_X\mathrm{d}p(\Omega)\\&=\iota_X\mathrm{d}p(\Omega)\end{aligned}$$

For (2). Note that $\operatorname{im}\pi^*=\{\tau\in C^\infty(P,\Omega_P^{2k})\mid\tau \text{ is horizontal and } G\text{-invariant}\}$ and π^* is injective impiles uniqueness. It's closed, since

$$\pi^*(\mathrm{d}p(\Theta))=\mathrm{d}\pi^*(p(\Theta))=\mathrm{d}p(\Omega)=0$$

For (3). Let ω' be another connection on P, consider principal G-bundle $P\times\mathbb{R}$ over $M\times\mathbb{R}$, and connection $\widetilde{\omega}=(1-t)\omega+t\omega'$ on it, with curvature $\widetilde{\Omega}$. If we use i_0,i_1 to denote inclusion from M to $M\times\{0\}$ and $M\times\{1\}$ respectively, then it's clear

$$\begin{aligned}p(\Theta)&=i_0^*p(\widetilde{\Omega})\\p(\Theta')&=i_1^*p(\widetilde{\Omega})\end{aligned}$$

Furthermore, the homotopy invariance of de Rham cohomology impiles $i_0^*,i_1^*:H^{2k}(M\times\mathbb{R})\to H^{2k}(M)$ are the same map. □

Theorem 5.2.1 (Chern-Weil homomorphism)**.** There is a ring homomorphism

$$\begin{aligned}W(P,\text{-}):I(\mathfrak{g})&\to H^*(M)\\p&\mapsto[p(\Theta)]\end{aligned}$$

Proof. It suffices to show

$$p\odot q(\Theta)=p(\Theta)\wedge q(\Theta)$$

Note that π^* is injective, thus it suffices to check

$$p\odot q(\Omega)=p(\Omega)\wedge q(\Omega)$$

and that's clear. □

5.3. **Transgression.** In this section we will show for a given principal G-bundle P and a connection ω on it with curvature Ω, $[p(\Omega)]=0\in H^{2k}(P)$, where $p\in I^k(\mathfrak{g}),k\geq1$.

To see this, let's introduce the funtorial Chern-Weil homomorphism. Given the following homomorphism between principal G-bundles

$$\begin{tikzcd} P' \arrow[r, "\widetilde{f}"] \arrow[d, "\pi'"] & P \arrow[d, "\pi"] \\ M' \arrow[r, "f"] & M \end{tikzcd}$$

where $P'=f^*P$.

Proposition 5.3.1 (funtorial)**.** For all $p \in I(\mathfrak{g})$, we have

$$W(f^*P, p) = f^*W(P, p)$$

Proof. Given a connection $\omega \in \mathscr{A}(P)$ with curvature Ω, and use ω' to denote the pullback connection $\widetilde{f}^*\omega \in \mathscr{A}(P')$ with curvature Ω'. For any $p \in I(\mathfrak{g})$, it's clear

$$p(\Omega') = \widetilde{f}^*p(\Omega)$$

Then

$$\begin{aligned}(\pi')^*(p(\Theta')) &= \widetilde{f}^*\pi^*p(\Theta) \\ &= (\pi')^*f^*p(\Theta')\end{aligned}$$

which impiles $p(\Theta') = f^*p(\Theta)$, since $(\pi')^*$ is injective. □

Example 5.3.1. Let $P = M \times G$ be trivial bundle, then we can regard it as

$$\begin{tikzcd} M \times G \arrow[r, "\widetilde{f}"] \arrow[d, "\pi'"] & G \arrow[d, "\pi"] \\ M \arrow[r, "f"] & \{\mathrm{pt}\} \end{tikzcd}$$

So for any $p \in I^k(\mathfrak{g}), k \geq 1$, we have

$$W(P, p) = f^*W(G, p) = 0$$

since $W(G, p) \in H^{2k}(\{\mathrm{pt}\}) = 0$ if $k \geq 1$.

Remark 5.3.1. This example shows if P is a trivial principal G-bundle, then the Chern-Weil homomorphism $W(P, \text{-})$ is trivial on $I(\mathfrak{g})$.

Now let's consider the following case

$$\begin{tikzcd} f^*P \arrow[r, "\widetilde{f}"] \arrow[d, "\pi'"] & P \arrow[d, "\pi"] \\ P \arrow[r, "f"] & M \end{tikzcd}$$

where $f = \pi$. In fact we can write f^*P down as

$$\begin{aligned}f^*P &= \{(x', x) \in P \times P \mid f(x') = \pi(x)\} \\ &= \{(x', x) \in P \times P \mid \pi(x') = \pi(x)\}\end{aligned}$$

It's clear it has global section, given by

$$\begin{aligned}s : P &\to f^*P \\ x &\mapsto (x, x)\end{aligned}$$

so f^*P is trivial principal bundle. Thus for any $p \in I^k(\mathfrak{g}), k \geq 1$, we have

$$W(f^*P, p) = 0 \in H^{2k}(P)$$

However, funtorial impiles

$$\begin{aligned} W(f^*P, p) &= f^*W(P, p) \\ &= f^*[p(\Theta)] \\ &= \pi^*[p(\Theta)] \\ &= p(\Omega) \end{aligned}$$

This shows $[p(\Omega)] = 0$ in $H^{2k}(P)$.

6. Characteristic class

6.1. Chern class.

Proposition 6.1.1. Let $G = \mathrm{U}(n)$ with Lie algebra $\mathfrak{g} = \mathfrak{u}(n)$. For any $X \in \mathfrak{g}$, consider

$$\det(I - \frac{t}{2\pi i}X) = \sum_{k=0}^{n} c_k(X)t^k$$

Then

1. For each $1 \leq k \leq n$, $c_k \in I(\mathfrak{g})$;
2. $I(\mathfrak{g})$ is generated by $c_1, \ldots, c_n$

Proof. For (1). For arbitrary $g \in G$, note that

$$\begin{aligned}\det(I - \frac{t}{2\pi i}\operatorname{Ad}(g)X) &= \det(I - \frac{t}{2\pi i}gXg^{-1}) \\ &= \det(g^{-1}g - \frac{t}{2\pi i}gXg^{-1}) \\ &= \det(I - \frac{t}{2\pi i}X)\end{aligned}$$

For (2). Note that any $X \in \mathfrak{g}$ is diagonalizable, so without lose of generality we may assume $X = \operatorname{diag}\{\lambda_1, \ldots, \lambda_n\}$. Then $I(\mathfrak{g})$ consists of symmetric polynomial of $\lambda_1, \ldots, \lambda_n$. Then the proof follows since any symmetric function can be expressed in terms of elementary symmetric functions and

$$\begin{aligned}c_1 &= -\frac{1}{2\pi}\lambda_1 + \cdots + \lambda_n \\ &\vdots \\ c_n &= (\frac{1}{2\pi})^n \lambda_1 \ldots \lambda_n\end{aligned}$$

□

Let E be a complex vector bundle of rank n over M equipped with a hermitian metric, then consider its frame bundle we obtain a $\mathrm{U}(n)$-principal bundle $\pi : P \to M$, then choose an arbitrary connection ω on it with curvature Ω, then by Chern-Weil theory there exists a unique $2k$-form $c_k(\Theta)$ on M such that $\pi^*(c_k(\Theta)) = c_k(\Omega)$.

Definition 6.1.1 (chern class)**.** The k-th Chern class of E is defined as

$$c_k := [c_k(\Theta)] \in H^{2k}(M, \mathbb{C})$$

Definition 6.1.2 (chern polynomial)**.** The Chern polynomial is defined as

$$c(t) = \det(I - \frac{t}{2\pi i}\Theta) = \sum_{k=0}^{n} c_k t^k$$

Proposition 6.1.2.

$$c_k \in H^{2k}(M, \mathbb{R})$$

Proof. Note that $\mathfrak{u}(n)$ consists of skew-symmetric matrices, then for arbitrary $X \in \mathfrak{u}(n)$, one has

$$\begin{aligned}\det(I - \frac{t}{2\pi i}X) &= \det(I + \frac{t}{2\pi i}\overline{X}^t) \\ &= \overline{\det(I - \frac{t}{2\pi i}X)} \\ &= \sum_{k=0}^{n} \overline{c}_k t^k\end{aligned}$$

which impiles $c_k = \overline{c}_k$. □

Proposition 6.1.3. Let E, F are two complex vector bundles, then

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

Proof. If ∇_E, ∇_F are connections on E, F respectively, then $\nabla_E \oplus \nabla_F$ gives a connection on $E \oplus F$, with curvature $\begin{pmatrix} \Theta_E & 0 \\ 0 & \Theta_F \end{pmatrix}$, and thus

$$c(E \oplus F) = \det \begin{pmatrix} I - \frac{1}{2\pi i}\Theta_E & 0 \\ 0 & I - \frac{1}{2\pi i}\Theta_F \end{pmatrix} = c(E)c(F)$$

□

6.2. **Pontrjagin class.** Now let E be a (real) vector bundle of rank n over M equipped with a Riemannian metric, then its frame bundle is a $\mathrm{O}(n)$-principal bundle P. For any $X \in \mathfrak{o}(n)$, consider

$$\det(I - \frac{t}{2\pi}X) = \sum_{k=0}^{n} q_k(X)t^k$$

By the same argument as above one can show $q_k \in I(\mathfrak{g})$, thus we pick arbitrary connection ω of P with curvature Ω, then it gives rise to a closed $2k$-form $q_k(\Theta)$ on M for each k. Note that $X + X^t = 0$, then

$$\det(I + \frac{t}{2\pi}X) = \det(I - \frac{-t}{2\pi}X)$$

which impiles

$$q_k(X) = q_k(-X) = (-1)^k q_k(X)$$

Thus we can conclude $q_k = 0$ for odd k.

Definition 6.2.1 (Pontrjagin class)**.** $[p_k(\Theta)] := [q_{2k}(\Theta)] \in H^{4k}(M, \mathbb{R})$ is called k-th Pontrjagin class of E.

Proposition 6.2.1. Let E be a vector bundle with its complexification $E^c = E \otimes \mathbb{C}$, which is a complex vector bundle, then

$$p_k(E) = (-1)^k c_{2k}(E^c)$$

Proof. □

If we consider oriented vector bundle E, then its frame bundle is a $\mathrm{SO}(n)$-principal bundle. Then

Lemma 6.2.1. Let E be a oriented vector bundle of rank n, then
1. If $n = 2m + 1$, then $I(\mathfrak{so}(n))$ is generated by $q_2, \dots, q_{2m}$;
2. If $n = 2m$, then $I(\mathfrak{so}(n))$ is generated by $q_2, \dots, q_{2m}, e$, where

$$e(\mathrm{diag}\{\begin{pmatrix} 0 & \lambda_1 \\ -\lambda_1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & \lambda_m \\ -\lambda_m & 0 \end{pmatrix}\}) = \lambda_1 \dots \lambda_m$$

Definition 6.2.2 (Euler class)**.** Let E be an oriented vector bundle of rank $2m$, then $[\frac{1}{(2\pi)^m} e(\Theta)] \in H^{2m}(M, \mathbb{R})$ is called the Euler class of E, denoted by $e(E)$.

Remark 6.2.1. For an oriented $2m$-dimensional manifold M, $e(TM)$ is the Euler number of M. See [JM74].

7. The classifying space

In last section, we have defined characteristic classes via a geometrical method, that is we use connections. However, they're topological invariants. In this section, we will give another explaination about characteristic class, and explain why it computes the right thing.

7.1. The universal G-bundle.

In this section, we work on category of topological spaces (In particular, CW-complexes) instead of smooth manifolds.

Definition 7.1.1 (weakly homotopy)**.** Let X, Y be topological spaces, X is weakly homotopy to Y, if there exists a continuous map $f : X \to Y$ such that f induces isomorphisms between homotopy groups of X and Y.

Definition 7.1.2 (weakly contractible)**.** A topological space X is called weakly contractible, if it's weakly homotopy to a point.

Remark 7.1.1. A contractible space is weakly contractible, and by Whitehead's theorem, a CW-complex is weakly contractible if and only if it's contractible.

Definition 7.1.3 (classifying space)**.** For a principal G-bundle $EG \to BG$, where EG, BG are topological spaces. If EG is weakly contractible, then

1. BG is called a classifying space for G;
2. EG is called a universal G-bundle.

Remark 7.1.2. Note that in definition the classifying space for G is just a topological space, in fact, we can choose it as a CW-complex. Indeed, since for any topological space, there exists a CW-complex which is weakly homotopic to it. Then for a classifying space BG, there exists a CW-complex BG' and a weakly homotopy $g : BG' \to BG$, then $g^*EG \to BG'$ is also a universal G-bundle.

Theorem 7.1.1. Let $EG \to BG$ be a universal G-bundle, then for all CW-complexes X, then the following map is bijective.

$$\begin{aligned} \phi : [X, BG] &\to \mathcal{P}_G X \\ f &\mapsto f^*P \end{aligned}$$

where $[X, BG]$ denotes the set of all continuous maps up to homotopy.

Proof. See [Mit01]. □

Remark 7.1.3. This theorem impiles why BG is called classifying space, since it can be used to classify principal G-bundles over a given CW-complex.

However, untill now we still don't know whether classifying space exists or not. The following theorem is due to [Mil56].

Theorem 7.1.2. Let G be any topological group, then there exists a classifying space for G.

Now let's see some examples of classifying space for special Lie group G.

Proposition 7.1.1. Let G be a discrete group, then $PK(G,1) \to K(G,1)$ is a universal G-bundle, and hence $K(G,1)$ is a classifying space for G.

Proof. It's clear path space $PK(G,1)$ is contractible. □

Remark 7.1.4. In [Liu22] we have already computed $K(G,1)$ for groups, for example, $K(\mathbb{Z},1) = \mathbb{S}^1$, $K(\mathbb{Z}_2,1) = \mathbb{RP}^\infty$ and so on.

Proposition 7.1.2. $V_n(\mathbb{R}^\infty) \to Gr_n(\mathbb{R}^\infty)$ is a universal $\mathrm{GL}(n,\mathbb{R})$-bundle, and hence $Gr_n(\mathbb{R}^\infty)$ is a classifying space for $\mathrm{GL}(n,\mathbb{R})$.

Proof. It suffices to show $V_n(\mathbb{R}^\infty)$ is contractible. Since we have already computed low dimensional homotopy groups of $V_n(\mathbb{R}^N)$ in [Liu22], and then telescope construction completes the proof. □

Corollary 7.1.1. For all CW-complexes X, $[X, Gr_n(\mathbb{R}^\infty)] \to \mathrm{Vect}_n^{\mathbb{R}} X$.

Proof. See Remark 1.3.2. □

Remark 7.1.5. The analogous result with $\mathbb{R}$ replaced by $\mathbb{C}$ also holds.

7.2. **Homotopical properties of classifying spaces.** In this section we collect some Homotopical properties of classifying spaces.

Theorem 7.2.1. Let G be any topological group, then G is weakly equivalent to the loop space ΩBG.

Corollary 7.2.1. For $n \geq 1$, $\pi_n(BG) = \pi_{n-1}(G)$.

Theorem 7.2.2. Let G be a topological space and H a subgroup, then the homotopy fiber of $BH \to BG$ is G/H, up to weakly equivalent.

Theorem 7.2.3. Let G be a topological space and H a subgroup, then there is a fiberation $BH \to BG \to B(G/H)$.

Example 7.2.1. The exact sequences $1 \to \mathrm{SO}(n) \to \mathrm{O}(n) \to \mathbb{Z}_2 \to 1$ and $1 \to \mathrm{SU}(n) \to \mathrm{U}(n) \to S^1 \to 1$ give rise to fiberation

$$B\,\mathrm{SO}(n) \to B\,\mathrm{O}(n) \to \mathbb{RP}^\infty$$

and

$$B\,\mathrm{SU}(n) \to B\,\mathrm{U}(n) \to \mathbb{CP}^\infty$$

7.3. **Another viewpoint to characteristice class.**

Proposition 7.1. The cohomology ring of $B\,\mathrm{U}(n)$ with integer coefficients is $\mathbb{Z}[c_1,\ldots,c_n]$.

Proof. If we consider $\mathrm{U}(n-1)$ as a subgroup of $\mathrm{U}(n-1)$, then we have the following filteration

$$\begin{tikzcd} S^{2n-1} \cong \mathrm{U}(n)/\mathrm{U}(n-1) \arrow[r] & B\,\mathrm{U}(n) \arrow[d] \\ & B\,\mathrm{U}(n-1) \end{tikzcd}$$

Apply Leray spectral sequence this this fiberation and use the fact that the cohomology ring of $\mathbb{CP}^\infty$ is $\mathbb{Z}[c_1]$ to conclude. $\square$

Definition 7.1 (universal Chern class)**.** The generators $c_1, \cdots, c_n$ of $H^*(B\,\mathrm{U}(n), \mathbb{Z})$ are called the universal Chern classes of $\mathrm{U}(n)$-bundles.

Definition 7.2 (Chern class)**.** The k-th Chern class of the $\mathrm{U}(n)$-bundle $\pi : E \to M$ with classifying map $f_\pi : M \to B\,\mathrm{U}(n)$ is defined as

$$c_k(E) := f_\pi^*(c_k) \in H^{2k}(M, \mathbb{Z})$$

Proposition 7.2. The cohomology ring of $B\,\mathrm{O}(n)$ with $\mathbb{Z}_2$ coefficients is $\mathbb{Z}_2[w_1, \ldots, w_n]$.

Proof. The same as above, just note that cohomology ring of $\mathbb{RP}^\infty$ with $\mathbb{Z}_2$ coefficient is $\mathbb{Z}_2[w_1]$. $\square$

Definition 7.3 (universal Steifel-Whitney class)**.** The generators $w_1, \cdots, w_n$ of $H^*(B\,\mathrm{O}(n), \mathbb{Z}_2)$ are called the universal Steifel-Whitney classes of $\mathrm{O}(n)$-bundles.

Definition 7.4 (Steifel-Whitney class)**.** The k-th Steifel-Whitney class of the $\mathrm{O}(n)$-bundle $\pi : E \to M$ with classifying map $f_\pi : M \to B\,\mathrm{O}(n)$ is defined as

$$w_k(E) := f_\pi^*(w_k) \in H^{2k}(M, \mathbb{Z}_2)$$

Part 3. The Yang-Mills equations on Riemann surface

8. The Yang-Mills equations

In this section we assume G is a compact Lie group, since we desire Killing form of G is non-degenerate, and (M, g) is an oriented compact Riemannian manifold, since we need to consider integration.

8.1. The Yang-Mills functional.

Let P be a principal G-bundle, V is a vector space and $\rho : G \to \mathrm{GL}(V)$ is a representation of G. If we want to construct an inner product on $\Omega^k_M(P \times_\rho V)$, firstly on each local trivialization U_α, view such forms as forms with values in V, so all we need is an inner product on V, since we already have a Riemannian metric g on M, which induces an inner product on forms.

But if we desire such inner product $\langle -, - \rangle$ can be glued well on overlaps, we need to require that it is G-invariant, that is, for all $g \in G, v, w \in V$,

$$\langle \rho(g)w, \rho(g)w \rangle = \langle v, w \rangle$$

since if $\omega \in C^\infty(M, \Omega^k_M(P \times_\rho V))$ is represented locally by $\omega_\alpha \in C^\infty(U_\alpha, \Omega^k_{U_\alpha}(V))$, then on a non-empty overlap $U_{\alpha\beta}$, we have $\omega_\alpha = \rho(g_{\alpha\beta})\omega_\beta$.

The case we're most interested in is $V = \mathfrak{g}$, since curvature of a connection is a section of $\Omega^2_M(\mathrm{Ad}\,\mathfrak{g})$. So we what we need is an inner product on Lie algebra $\mathfrak{g}$ which is invariant under the adjoint action. Since G is compact, its Killing form is a non-degenerate inner product, that's what we're looking for!

Thus we have an pointwise inner product on the bundle $\Omega^k_M(\mathrm{Ad}\,\mathfrak{g})$, and denote it by $\langle -, - \rangle$, and define a global inner product on $\Omega^k_M(\mathrm{Ad}\,\mathfrak{g})$ as

$$(\alpha, \beta) := \int_M \langle \alpha, \beta \rangle \,\mathrm{vol}$$

where $\alpha, \beta \in C^\infty(M, \Omega^k_M(\mathrm{Ad}\,\mathfrak{g}))$.

Definition 8.1.1 (Hodge star operator)**.** There exists an operator

$$* : C^\infty(M, \Omega^k_M(\mathrm{Ad}\,\mathfrak{g})) \to C^\infty(M, \Omega^{n-k}_M(\mathrm{Ad}\,\mathfrak{g}))$$

For $\beta \in C^\infty(M, \Omega^k_M(\mathrm{Ad}\,\mathfrak{g}))$, $*\beta$ is given by

$$\alpha \wedge *\beta = \langle \alpha, \beta \rangle \,\mathrm{vol}, \quad \forall \alpha \in C^\infty(M, \Omega^k_M(\mathrm{Ad}\,\mathfrak{g}))$$

With some of the preliminary results established, we arrive at the Yang-Mills functional.

Definition 8.1.2 (Yang-Mills functional)**.** The Yang-Mills functional is the map $YM : \mathscr{A}(P) \to \mathbb{R}$ given by

$$YM(\omega) := \|F_\omega\|^2 = \int_M \langle F_\omega, F_\omega \rangle \,\mathrm{vol}$$

where F_ω is curvature of connection ω, which is a section of $\Omega^2_M(\mathrm{Ad}\,\mathfrak{g})$.

Remark 8.1.1. By using Hodge star operator, we may rewrite Yang-Mills functional as follows
$$YM(\omega)=\int_M F_\omega\wedge *F_\omega$$
The advantages of writing Yang-Mills functional in this way is that we can use some properties of Hodge operator to simplify our computations

Proposition 8.1.1. Yang-Mills functional YM is gauge invariant, that is for any gauge transformation $\Phi\in\mathscr{G}(P)$, one has $YM(\Phi^*\omega)=YM(\omega)$ holds for connection ω.

Proof. On each local trivialization U_α, the curvature of $\Phi^*\omega$ is given by $\mathrm{Ad}(\phi^{-1})\circ F_\alpha$, where ϕ is given by $\Phi|_{U_\alpha}(x,g)=(x,\phi(x)g)$, thus Yang-Mills functional is gauge invariant follows from inner product $\langle\text{-},\text{-}\rangle$ is adjoint invariant. $\square$

Definition 8.1.3 (Yang-Mills connection)**.** A Yang-Mills connection is a connection $A\in\mathscr{A}(P)$ which is a local extremum of Yang-Mills functional.

Notation 8.1.1. $\mathscr{A}_{YM}(P)$, or briefly $\mathscr{A}_{YM}$ denotes the set of all Yang-Mills connections.

8.2. **The variational problem.** Let's see how to use a second-order partial differential equation to characterize Yang-Mills connection. Recall that $\mathscr{A}(P)$ is an affine space modelled on $\Omega^1_M(\mathrm{Ad}\,\mathfrak{g})$. This means the tangent space to $\mathscr{A}(P)$ at any point is isomorphic to $\Omega^1_M(\mathrm{Ad}\,\mathfrak{g})$.

Given $\omega\in\mathscr{A}(P)$ and $\tau\in C^\infty(M,\Omega^1_M(\mathrm{Ad}\,\mathfrak{g}))$. The directional derivative of Yang-Mills functional at ω in the direction τ is given by
$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} YM(\omega+t\tau)$$
And Yang-Mills condition states that this vanishes for all τ. In order to see what this means, firstly we need the following lemma.

Lemma 8.2.1. Given $\omega\in\mathscr{A}(P)$ and $\tau\in C^\infty(M,\Omega^1_M(\mathrm{Ad}\,\mathfrak{g}))$, then
$$F_{\omega+\tau}=F_\omega+\mathrm{d}_\omega\tau+\frac{1}{2}\tau\wedge\tau$$
where d_ω is connection induced by ω on $\Omega^1_M(\mathrm{Ad}\,\mathfrak{g})$.

Proof. On local trivialization U_α one has
$$\begin{aligned}(F_{\omega+\tau})_\alpha&=\mathrm{d}(A_\alpha+\tau_\alpha)+\frac{1}{2}(A_\alpha+\tau_\alpha)\wedge(A_\alpha+\tau_\alpha)\\&=(F_\omega)_\alpha+\mathrm{d}\tau_\alpha+\frac{1}{2}(A_\alpha\wedge\tau_\alpha+\tau_\alpha\wedge A_\alpha)+\frac{1}{2}\tau_\alpha\wedge\tau_\alpha\\&\overset{(1)}{=}(F_\omega)_\alpha+\mathrm{d}\tau_\alpha+A_\alpha\wedge\tau_\alpha+\frac{1}{2}\tau_\alpha\wedge\tau_\alpha\\&\overset{(2)}{=}(F_\omega)_\alpha+\mathrm{d}_\omega\tau_\alpha+\frac{1}{2}\tau_\alpha\wedge\tau_\alpha\end{aligned}$$
where

(1) holds from both A_α, τ_α are 1-form valued in $\mathfrak{g}$;
(2) holds from (3.2).

□

Proposition 8.2.1 (first variation formula)**.** Let ω be a Yang-Mills connection, then we have

$$\mathrm{d}_\omega^* F_\omega = 0$$

Proof. Direct computation shows

$$\begin{aligned} YM(\omega + t\tau) &= \int_M \langle F_{\omega+t\tau}, F_{\omega+t\tau}\rangle \operatorname{vol} \\ &= \int_M \langle F_\omega + \frac{t^2}{2}(\tau\wedge\tau) + t\mathrm{d}_\omega\tau, F_\omega + \frac{t^2}{2}(\tau\wedge\tau) + t\mathrm{d}_\omega\tau\rangle \operatorname{vol} \end{aligned}$$

The coefficient of linear term is

$$\int_M \langle F_\omega, \mathrm{d}_\omega\tau\rangle + \langle \mathrm{d}_\omega\tau, F_\omega\rangle \operatorname{vol} = 2\int_M \langle \mathrm{d}_\omega\tau, F_\omega\rangle \operatorname{vol}$$

Let $\mathrm{d}_\omega^* = (-1)^{2n+1} * \mathrm{d}_\omega *$ denote the formal adjoint to d_ω. Then we have

$$\int_M \langle \mathrm{d}_\omega\tau, F_\omega\rangle \operatorname{vol} = \int_M \langle \tau, \mathrm{d}_\omega^* F_\omega\rangle \operatorname{vol}$$

this shows

$$\mathrm{d}_\omega^* F_\omega = 0$$

□

Definition 8.2.1 (Yang-Mills equations)**.** A connection $\omega \in \mathscr{A}(P)$ is called satisfying Yang-Mills equations, if

$$\begin{cases} \mathrm{d}_\omega F_\omega = 0 \\ \mathrm{d}_\omega^* F_\omega = 0 \end{cases}$$

Remark 8.2.1*.* The first equation is also called Bianchi identity.

Example 8.2.1. In the case that $G = U(1)$, we have that the curvature of a connection A can be identified as a section of Ω_M^2. Indeed, the curvature form takes value in the bundle $\operatorname{Ad}\mathfrak{g}$, but here $G = U(1)$ is abelian, thus the adjoint action on $\mathfrak{u}(1)$ is trivial, so

$$\operatorname{Ad}\mathfrak{g} = M \times \mathfrak{u}(1) = M \times \mathbb{R}$$

is trivial bundle. Furthermore, ω is a Yang-Mills connection if and only if F_ω is a harmonic 2-form, that is $\Delta F_\omega = 0$, where $\Delta = \mathrm{d}\mathrm{d}^* + \mathrm{d}^*\mathrm{d}$. Indeed, thanks to $U(1)$ is abelian again, d_ω can be reduced to d, since for arbitrary form β, we have $\omega \wedge \beta = 0$. This follows from in the definition of wedge product of forms valued in Lie algebra we used Lie bracket, and abelian Lie algebra has trivial Lie bracket. Note that F_ω is harmonic if and only if

$$\begin{cases} \mathrm{d}^* F_\omega = 0 \\ \mathrm{d} F_\omega = 0 \end{cases}$$

It's a standard result in differential geometry, which can be seen from

$$\begin{aligned}0 &= \int_M \langle \Delta F_\omega, F_\omega \rangle \operatorname{vol} \\ &= \int_M \langle \mathrm{d}\mathrm{d}^* F_\omega, F_\omega \rangle + \langle \mathrm{d}^*\mathrm{d} F_\omega, F_\omega \rangle \operatorname{vol} \\ &= \int_M \|\mathrm{d}^* F_\omega\|^2 + \|\mathrm{d} F_\omega\|^2 \operatorname{vol}\end{aligned}$$

9. GIT quotient and sympletic quotient

Note that the Yang-Mills functional is guage invariant, so if a connection ω solves the Yang-Mills equations, so does any gauge transformed $\Phi^*\omega$. In other words, the gauge group acts on $\mathscr{A}_{YM}$. The quotient $\mathscr{A}_{YM}/\mathscr{G}$ is the space of classical solutions. In general it is infinite dimensional, and the topology of this space may be quite bad. For example it may be neither Hausdorff or a smooth manifold. But adding some restrictions, we do have a good correspondence, and that's main theorem for this section.

9.1. A Fairy Tale.

To get a picture of the action of gauge group on $\mathscr{A}(P)$, let's study a finite-dimensional analogue: Let V be a complex vector space with a hermitian inner product $\|\cdot\|$, and S^1 acts on it by unitary matrices, that is there is a group homomorphism $S^1 \to \mathrm{U}(V)$, if we consider the complexification of this action, we obtain $\mathbb{C}^* \to \mathrm{GL}(V)$. The goal is to understand the quotient space $V/\mathbb{C}^*$, but this space can be quite unpleasant. Let's see an example:

Example 9.1.1. Let $\lambda \in \mathbb{C}^*$ acting on $\mathbb{C}^2$ by $(x,y) \mapsto (\lambda^{-1}x, \lambda y)$. The orbits are

1. the conics $xy = c, c \neq 0 \in \mathbb{C}$;
2. the axes $y = 0, x \neq 0$ or $x = 0, y \neq 0$;
3. the origin.

It's clear that the topology on the orbit space is not Hausdorff, since origin lies in the closure of axes. But note that $\mathbb{C}^2 \setminus \{\text{axes}\}/\mathbb{C}^*$ is Hausdorff. Indeed, it's homeomorphic to $\mathbb{C}$.

The problems arise from that axes are not closed orbits. More generally, if we want to form Hausdorff quotients, we just need to consider only closed orbits which are closed sets.

Definition 9.1.1 (stable)**.** A point $v \in V$ is stable if its orbit under $\mathbb{C}^*$ is closed.

Now let's see a criterion for whether an orbit is closed or not.

Theorem 9.1.1. A point $v \in V$ is stable if and only if the function $p_v : \mathbb{C}^* \to \mathbb{R}$, defined by $p_v(g) := \|g(v)\|^2, g \in \mathbb{C}^*$, attains its minimum.

Remark 9.1.1. Note that since the norm is $\mathrm{U}(V)$-invariant, then the function p_v is S^1-invariant and descend to a function on $\mathbb{C}^*/S^1$, given by

$$p_v(x) := \|e^x(v)\|^2$$

where $x \in \mathbb{C}^*/S^1$. In Example 9.1.1, we have $e^x(v_1, v_2) = (e^{-x}v_1, e^x v_2)$, so we have

$$p_v(x) = \|v_1\|^2 e^{-2x} + \|v_2\|^2 e^{2x}$$

Take its derivative and let it equal zero

$$\frac{\mathrm{d}p_v(x)}{\mathrm{d}x} = -2\|v_1\|^2 e^{-2x} + 2\|v_2\|^2 e^{2x} = 0$$

we have this function take its minimum at

$$\frac{1}{2}(\log(\|v_1\|) - \log(\|v_2\|))$$

if both v_1 and v_2 are not zero, and at 0 if $v = 0$. Furthermore, the minimum is not attained along the two punctured axes. In fact, Example 9.1.1 is quite representative.

Proof. Note that the S^1-action is reducible, so V splits into an orthogonal direct sum $V_1 \oplus \cdots \oplus V_n$ of 1-dimensional representations where S^1 acts on V_m as $v_m \mapsto \lambda^{j_m} v_m$ for some weight j_m. Therefore we have

$$p_v(x) = \sum_m \|v_m\|^2 e^{2j_m x} = \sum_{k=-\infty}^{\infty} a_k e^{kx}$$

where only finitely many $a_k \neq 0$. We divide our analysis into three cases:

1. $a_k = 0$ for all $k \neq 0$. In this case the minimum is obviously attained and the orbit is obviously closed since $j_m = 0$ so the action fixes v.
2. $a_k = 0$ for all $k < 0$ (resp. $k > 0$) and $a_k \neq 0$ for some $k > 0$ (resp. $k < 0$). In this case the minimum is obviously not attained and the orbit is obviously not closed since $e^x(v)$ tends to an orbit of the first type as $x \to -\infty$ (resp. ∞).
3. There is a $k > 0$ and a $k' < 0$ such that $a_k \neq 0$ and $a_{k'} \neq 0$. In this case the minimum is obviously attained (just do the calculus). We will now show that this implies v is stable.

Conversely, if v is not stable, then p_v does not attain its minimum. Indeed, if v is not stable then its orbit is not closed so there exists $w \in V$ such that $w \in \overline{\mathbb{C}^*(v)}$ but $w \notin \mathbb{C}^*(v)$, so either $w = \lim_{x\to\pm\infty} e^x(v)$. The corresponding limit $\lim_{x\to\pm\infty} p_v(x) = p_v(w)$ is finite and hence the j_m are either all nonpositive or all nonnegative. Since $w \neq v$ there must be one j_m which is nonzero. It's now easy to see that the function $p_v(x)$ is of type II and hence does not attain its minimum. □

So the above criterion says if we want to understand the space of stable points, it's neccessary to understand the critical point of p_v. Take derivative we have

$$\frac{\mathrm{d}p_v}{\mathrm{d}x} = 2\sum_{m=1}^{n} j_m \left\|v_m\right\|^2 e^{2j_m x}$$

Suppose v is stable and without lose of generality we assume its minimum occurs at $x = 0$. Therefore the orbit of a stable vector contains a zero of the function

$$\mu = \sum_{m=1}^{n} j_m \|v_m\|^2 : V \to \mathbb{R}$$

So we can restate above criterion as follows, and it's a fairy tale version of Kempf-Ness.

Theorem 9.1.2 (Kempf-Ness)**.** Let V^s denote the space of stable vectors under the action of $\mathbb{C}^*$. Then

$$V^s/\mathbb{C} = \mu^{-1}(0)/S^1$$

Let's define more conceptions which we will see a more abstract version later. Let V be a vector space and $Q : V \otimes V \to \mathbb{R}$ a non-degenerate bilinear form. Then we can make a 1-form ω into a vector field X by defining

$$\omega(Y) = Q(X, Y), \quad \forall Y \in TV$$

In particular, if $f : V \to \mathbb{R}$ is a function and consider its derivative 1-form $\mathrm{d}f$. Then it corresponds to a vector field $\mathrm{Qgrad}(f)$ by defining

$$\mathrm{d}f(Y) = Q(\mathrm{Qgrad}(f), Y)$$

We call f the Hamiltonian generating $\mathrm{Qgrad}(f)$, and μ the moment map for the circle action.

Example 9.1.2. Take $f(x, y) = x^2 + y^2$ and $Q = \mathrm{d}x \wedge \mathrm{d}y$. Then

$$\mathrm{Qgrad}\, f = -y\partial_x + x\partial_y$$

9.2. **Kempf-Ness Theorem.** In this section, we give an abstract version of Kempf-Ness theorem. Let X be a Kähler manifold, $K \subset G$ denote the maximal compact subgroup, which has the property that its complexification is isomorphic to G.

Suppose that the action of K on X is sympletic, i.e. the action of any $k \in K$ preserves the Kähler metric on X. Let $\mathfrak{k}$ denote the Lie algebra of K. Then the infinitesimal action of K is given by the Lie algebra homomorphism $\mathfrak{k} \to \mathfrak{X}(X)$ defined by $\xi \mapsto X_\xi$, where

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

Definition 9.2.1 (Hamiltonian)**.** A symplectic action of K on X is Hamiltonian if for each $\xi \in \mathfrak{k}$, there exists a function $H_\xi : X \to \mathbb{R}$ such that for all $p \in X$ and $v \in T_pX$ we have

$$\omega_p((X_\xi)_p, v) = (\mathrm{d}H_\xi)_p(v)$$

and the mapping $\xi \mapsto H_\xi$ is K-equivariant with respect to the right action of K on $\mathfrak{k}$ by the adjoint action and precomposition with right translation R_k on $C^\infty(X)$. The functions H_ξ are called Hamiltonian functions.

Definition 9.2.2 (moment map)**.** Suppose we have a Hamiltonian action of K on X. A moment map for the action is a K-equivariant map $\mu : X \to \mathfrak{k}^*$(where the action on $\mathfrak{k}^*$ is the coadjoint action) such that for any $p \in X, v \in T_pX$ and $\xi \in \mathfrak{k}$, we have

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

Remark 9.2.1. Let's make coadjoint action more clear:

Let G be a Lie group and $\mathfrak{g}$ be its Lie algebra. Let $\mathrm{ad} : G \to \mathrm{Aut}(\mathfrak{g})$ denote the adjoint representation of G. Then we can define its coadjoint representation $\mathrm{ad}^* : G \to \mathrm{Aut}(\mathfrak{g}^*)$ as

$$\langle \mathrm{ad}_g^* \mu, Y \rangle = \langle \mu, \mathrm{ad}_{g^{-1}} Y \rangle$$

for $g \in G, Y \in \mathfrak{g}, \mu \in \mathfrak{g}^*$.

Remark 9.2.2. One thing to note is that the Hamiltonian functions can be recovered by the moment maps. If a Hamiltonian action admits a moment map, then

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

Let $\langle \cdot, \cdot \rangle$ be an inner product on $\mathfrak{k}^*$ that is invariant under the coadjoint action, and $\| \cdot \|$ be the induced norm. Since X is compact, then the map $\|\mu\|^2 : X \to \mathbb{R}$ attains its minimum, and without lose of generality we assume that the minimum value is 0.

Definition 9.2.3 (symplectic quotient)**.** The symplectic quotient of X by K is the quotient space

$$\mu^{-1}(0)/K$$

Remark 9.2.3. The symplectic quotient can also be referred to as the symplectic reduction. It should be noted that the symplectic quotient depends on our choice of moment map.

Theorem 9.2.1. The symplectic quotient of X by K admits a unique Kähler structure such that the Kähler metric on $\mu^{-1}(0)/K$ is induced by the Kähler metric on X.

The relationship between the GIT quotient and the symplectic quotient is given by the Kempf-Ness theorem

Theorem 9.2.2 (Kempf-Ness)**.** Suppose a complex reductive group G acts on a Kähler manifold X such that the action of the maximal compact subgroup $K \subset G$ is Hamiltonian and admits a moment map $\mu : X \to \mathfrak{k}^*$. Then the G-orbit of any semistable point contains a unique K-orbit minimizing $\|\mu\|^2$. This establish a homeomorphism

$$X_{ss}/G \longleftrightarrow \mu^{-1}(0)/K$$

10. Holomorphic Vector Bundles and Hermitian Yang-Mills Connections

In the second part, we have already established the foundations of Yang-Mills equations in a general stage, or in other words, in the stage of Riemannian manifold (M, g).

As we have seen, when the dimension of underlying space is one, all curvature forms are trivial, so there is nothing interesting. Thus the first "non-trivial" theory arises when our underlying space is of dimension two.

This prototype theory merits a good deal of study due to the richness of structures naturally occurring on such manifold, such as a complex structure associated to the almost complex structure determined by the Hodge star operator $* : \Omega^p_M \to \Omega^{2-p}_M$. Furthermore, smooth Hermitian vector bundle E over Riemann surface have inherent holomorphic structures due to the vacuous integrability conditions on connections on E, in other words, this gives a correspondence between unitary connections and holomorphic structure $\overline{\partial}_E$ on E. Thus the study of Yang-Mills connections on Riemann surface can be put into a complex analytic framework.

Using such ideal, we give a description of Kempf-Ness theorem which relates symplectic quotient and GIT quotient. In this section, if the underlying space is a Riemann surface, we will see there is a parallel story for the action of gauge group $\mathscr{G}$ on the space of connections $\mathscr{A}(P)$.

We will complexify the action of $\mathscr{G}$ and state a theorem analogous to Kempf-Ness theorem, which is known as Narasimhan-Seshadri theorem.

Notation 10.0.1. For complex manifold X, we use Ω^k_X to denote the space of smooth complex-valued k-forms, and use $\Omega^{p,q}_X$ to denote the space of smooth (p, q)-forms.

10.1. Moment map in Yang-Mills theory.

When X is a Riemann surface, the space of connections has a natural symplectic form. As we already knonw, $\mathscr{A}(P)$ is affine modelled on $\Omega^1_X(\mathfrak{g}_P)$, then we consider the following non-degenerate symplectic form

$$Q(\alpha, \beta) = \int_X \alpha \wedge \beta, \quad \alpha, \beta \in \Omega^1_X(\mathfrak{g}_P)$$

where this integral do make senses since the real dimension of X is two.

Take $\phi \in \Omega^0_M(\mathfrak{g}_P)$, we can get a vector field on $\mathscr{A}(P)$ by the action of ∇ on ϕ, that is $V = \nabla\phi$.

Lemma 10.1.1. The function $f : \nabla \to -\int_X F_\nabla \wedge \phi$ is a Hamiltonian functions on $\mathscr{A}(P)$ generating V.

Proof. It suffices to check

$$Q(\nabla\phi, A) = \mathrm{d}f(A), \quad \forall A \in \Omega^1_X(\mathfrak{g}_P)$$

Integration by parts we have

$$\begin{aligned} Q(\nabla\phi, A) &= \int_X \nabla\phi \wedge A \\ &= -\int_X \phi \wedge \nabla A \\ &= -\int_X \nabla A \wedge \phi \end{aligned}$$

Note that $F_{\nabla+\varepsilon A} = F_\nabla + \varepsilon\nabla A + O(\varepsilon^2)$, then

$$\begin{aligned} \mathrm{d}f(A) &= \lim_{\varepsilon\to 0} \frac{-\int_X F_{\nabla+\varepsilon A}\wedge\phi + \int_X F_\nabla \wedge \phi}{\varepsilon} \\ &= -\int_X \nabla A \wedge \phi \end{aligned}$$

As desired. □

Remark 10.1.1. In our case the Lie algebra of gauge group is $\Omega^2_X(\mathfrak{g}_P)$ and the moment map is just

$$\nabla \mapsto -F_\nabla$$

The Yang-Mills functional is just the norm of the moment map.

Our ultimate goal is to relate moduli spaces of holomorphic vector bundles over X to Yang-Mills connections. Firstly, we want to consider $\mathscr{A}(P)$ as a space of holomorphic vector bundles.

11. Moduli space of semi-stable vector bundles

In thise section, the guiding problem is to classify holomorphic vector bundles on a Riemann surface with genus g, denoted by Σ_g. For the case $g = 0, 1$, there are complete classification results for holomorphic vector bundles on Σ_g, due to Grothendieck for the case of the Riemann sphere [Gro57], and due to Atiyah for the case of elliptic curves [Ati57]. So in the following discussion, we always assume $g \geq 2$.

11.1. **Stable bundle.**

Definition 11.1.1 (holomorphic vector bundle)**.** A holomorphic vector bundle is a complex vector bundle $\pi : E \to X$ such that the total space E is a complex manifold and π is holomorphic.

Definition 11.1.2 (degree)**.** Let $\pi : E \to X$ be a holomorphic vector bundle, its degree is defined as

$$\deg(E) := \int_X c_1(E)$$

where $c_1(E) \in H^2(X, \mathbb{Z})$ is the first Chern class of E.

Definition 11.1.3 (slope)**.** Let $\pi : E \to X$ be a holomorphic vector bundle, its slope is defined as

$$\mu(E) := \frac{\deg(E)}{\operatorname{rank}(E)}$$

Remark 11.1.1. One thing to note is that the slope of a holomorphic vector bundle is independent of the holomorphic structure, since both the degree and rank are topological invariants.

Definition 11.1.4 (slope stability)**.** Let $\pi : E \to X$ be a holomorphic vector bundle, it's

1. stable if for every non-trivial holomorphic subbundle F, $\mu(F) < \mu(E)$;
2. semi-stable if for every non-trivial holomorphic subbundle F, $\mu(F) \leq \mu(E)$;
3. unstable if it's not semi-stable.

Remark 11.1.2. For slope stability, we have the following remarks:

(a) It's clear that all holomorphic line bundles are stable, since they don't have non-trivial subbundles;
(b) A semi-stable vector bundle with coprime rank and degree is actually stable, since
(c) While the slope is a topological invariant, slope stability is not, since here we only consider holomorphic subbundles, which depends on the holomorphic structure.

Proposition 11.1.1. Let $E \to \Sigma_g$ be a holomorphic vector bundle, it's

1. stable if and only if for every non-trivial holomorphic subbundle F, $\mu(E/F) > \mu(E)$;

2. semi-stable if and only if for every non-trivial holomorphic subbundle F, $\mu(E/F) \geq \mu(E)$.

Proof. Denote r, r', r'' the ranks of $E, F, E/F$ respectively, and d, d', d'' their degrees respectively. From exact sequence

$$0 \to E \to E \to E/F \to 0$$

one has $r = r' + r''$ and $d = d' + d''$, thus

$$\frac{d'}{r'} < \frac{d' + d''}{r' + r''} \iff \frac{d'}{r'} < \frac{d''}{r''} \iff \frac{d' + d''}{r' + r''} < \frac{d''}{r''}$$

and likewise with the case semi-stable. □

A philosophy is that semi-stable bundles don't admit too many subbundles, since any subbundle they may have is of slope no greater than their own. This turns out to have many interesting consequences we're going to show, for example, the category of semi-stable bundles is abelian.

Lemma 11.1.1. If $\varphi : E \to E'$ is a non-zero homomorphism of holomorphic vector bundles over Σ_g, then

$$\mu(E/\ker\varphi) \leq \mu(\operatorname{im}\varphi)$$

Proposition 11.1.2. Let E, E' be two semi-stable bundles such that $\mu(E) > \mu(E')$, then any homomorphism $\varphi : E \to E'$ is zero.

Proof. If φ is non-zero, since E is semi-stable, then

$$\mu(\operatorname{im}\varphi) \overset{(1)}{\geq} \mu(E/\ker\varphi) \overset{(2)}{\geq} \mu(E) > \mu(E')$$

where

(1) holds from Lemma 11.1.1;
(2) holds from Proposition 11.1.1.

which contradicts to the semi-stablity of E'. □

Proposition 11.1.3. Let $\varphi : E \to E'$ be a non-zero homomorphism of semi-stable holomorphic of slope μ, then $\ker\varphi$ and $\operatorname{im}\varphi$ are semi-stable bundles of slope μ, and the natural map $E/\ker\varphi \to \operatorname{im}\varphi$ is an isomorphism.

Corollary 11.1.1. The category of semi-stable bundles of slope μ is abelian, and the simple object[3] in this category is the stable bundles of slope μ.

Proof. By Proposition 11.1.3 one has the category of semi-stable bundles of slope μ is abelian. A stable bundle E is simple in this category, since it admits no non-trivial subbundles with slope μ; Conversely, if a semi-stable bundle E is simple, then any non-trivial subbundle F satisfies $\mu(F) \leq \mu(E)$ since E is semi-stable and $\mu(F) \neq \mu(E)$ since E is simple, this shows E is stable. □

[3]Recall a simple object in an abelian category is an object with no non-trivial subobject.

Proposition 11.1.4. Let E, E' be two stable vector bundles over Σ_g with same slopes, and $\varphi : E \to E'$ be a non-zero homomorphism, then φ is an isomorphism.

Proof. Since $\varphi : E \to E'$ is a non-zero homomorphism between stable bundles with same slopes, then by Proposition 11.1.3 one has $\ker \varphi$ is either 0 or has slope $\mu(E)$, but E is actually stable, then $\ker \varphi$ must be 0, and since φ is strict, this shows φ is injective. Likewise, $\operatorname{im} \varphi \neq 0$ and has slope $\mu(E')$, then it must be E' since E' is stable. Then again by φ is strict, $\operatorname{im} \varphi = E'$ impiles φ is surjective. Therefore φ is an isomorphism. □

Proposition 11.1.5. If E is a stable bundle over Σ_g, then $\operatorname{End} E = \mathbb{C}$. In particular, $\operatorname{Aut} E = \mathbb{C}^*$.

Proof. Let φ be a non-zero endomorphism of E, by Proposition 11.1.4 one has φ is an automorphism, so $\operatorname{End} E$ is a field, which contains $\mathbb{C}$ as its subfield of scalar endomorphisms. For any $\varphi \in \operatorname{End} E$, by Cayley-Hamilton theorem one has φ is algebraic over $\mathbb{C}$, and since $\mathbb{C}$ is algebraic closed, this shows $\operatorname{End} E \cong \mathbb{C}$. □

Corollary 11.1.2. A stable bundle is indecomposable, that is it's not isomorphic to a direct sum of non-trivial subbundles.

Proof. The automorphism group of $E = E_1 \oplus E_2$ contains $\mathbb{C}^* \times \mathbb{C}^*$, so by Proposition 11.1.5 it can't be stable. □

Theorem 11.1.1 (Jordan-Hölder filteration)**.** Any semi-stable bundle of slope μ over Σ_g admits a filteration

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

by holomorphic subbundles such that for each $1 \leq i \leq k$, one has

1. E_i/E_{i-1} is stable;
2. $\mu(E_i/E_{i-1}) = \mu(E)$.

Proposition 11.1.6 (Seshadri)**.** Any two Jordan-Hölder filterations

$$S : 0 = E_0 \subset E_1 \subset \cdots \subset E_k = E$$

and

$$S' : 0 = E'_0 \subset E'_1 \subset \cdots \subset E'_l = E$$

of a semi-stable bundle E have same length, and the associated graded objects

$$\operatorname{gr}(S) : 0 = E_1/E_0 \oplus \cdots \oplus E_k/E_{k-1}$$

and

$$\operatorname{gr}(S') : 0 = E'_1/E'_0 \oplus \cdots \oplus E'_k/E'_{k-1}$$

satisfy $E_i/E_{i-1} \cong E'_i/E'_{i-1}$ for all $1 \leq i \leq k$.

Definition 11.1.5 (poly-stable bundle)**.** A holomorphic vector bundle E over Σ_g is called poly-stable if it is isomorphic to a direct sum $E_1 \oplus \cdots \oplus E_k$ of stable bundles of the same slope.

Example 11.1.1. A stable bundle is poly-stable.

Example 11.1.2. The graded object associated to any Jordan-Hölder filteration of a semi-stable bundle E is a poly-stable, and by Proposition 11.1.6, it's unique up to isomorphism, this isomorphic class is denoted by $\mathrm{gr}(E)$.

Definition 11.1.6 (S-equivalence class)**.** The graded isomorphism class $\mathrm{gr}(E)$ associated to a semi-stable bundle E is called the S-equivalence class of E. If $\mathrm{gr}(E) \cong \mathrm{gr}(E')$, E and E' are called S-equivalent, and denoted by $E \sim_S E'$.

Definition 11.1.7. The set $\mathcal{M}_{\Sigma_g}(r,d)$ of S-equivalence classes of semi-stable bundles of rank r and degree d over Σ_g is called its moduli set, it contains the set $\mathcal{N}_{\Sigma_g}(r,d)$ of isomorphism classes of stable bundles of rank r and degree d.

Theorem 11.1.2 (Mumford-Seshadri)**.** Let $g \geq 2, r \geq 1$ and $d \in \mathbb{Z}$.

1. The set $\mathcal{N}_{\Sigma_g}(r,d)$ admits a structure of smooth, complex quasi-projective variety of dimension $r^2(g-1)+1$;
2. The set $\mathcal{M}_{\Sigma_g}(r,d)$ admits a structure of complex projective variety of dimension $r^2(g-1)+1$;
3. $\mathcal{N}_{\Sigma_g}(r,d)$ is an open dense subvariety of $\mathcal{M}_{\Sigma_g}(r,d)$.

In particular, when r and d are coprime, $\mathcal{M}_{\Sigma_g}(r,d) = \mathcal{N}_{\Sigma_g}(r,d)$ is a smooth complex projective variety.

Proof. See [Mum62] and [Ses67]. □

11.2. **The Harder-Narasimhan filteration.**

Theorem 11.2.1 (Harder-Narasimhan)**.** Any holomorphic vector bundle E over Σ_g has a unique filteration

$$0 = E_0 \subset E_1 \subset \dots E_k = E$$

by holomorphic subbundles such that

1. for all $1 \leq i \leq k$, E_i/E_{i-1} is semi-stable;
2. the slope $\mu_i := \mu(E_i/E_{i-1})$ of successive quotients satisfies

$$\mu_1 > \mu_2 > \cdots > \mu_k$$

This filteration is called Harder-Narasimhan filteration.

Proof. See [HN75]. □

Remark 11.2.1. If we denote $r = \operatorname{rank} E, d = \deg E, r_i = \operatorname{rank}(E_i/E_{i-1})$ and $d_i = \deg(E_i/E_{i-1})$, one has

$$r_1 + \cdots + r_k = r, \quad d_1 + \cdots + d_k = d$$

The k-tuple

$$\vec{\mu} := (\underbrace{\mu_1, \dots, \mu_1}_{r_1 \text{ times}}, \dots, \underbrace{\mu_k, \dots, \mu_k}_{r_k \text{ times}})$$

is called the Harder-Narasimhan type of E. It's equivalent to the data of the k-tuple $(r_i, d_i)_{1\leq i\leq k}$. In the plane of coordinates (r, d), the polygonal line

$$P_{\vec{\mu}} := \{(0,0), (r_1, d_1), (r_1 + r_2, d_1 + d_2), \ldots, (r_1 + \cdots + r_k, d_1 + \cdots + d_k)\}$$

defines a convex polygon called the Harder-Narasimhan polygon of E. The slope of the line from $(0,0)$ to (r_1, d_1) is μ_1, that is the slope of E_1/E_0, and perhaps that's why it's called slope. It's indeed convex, since $\mu_1 > \cdots > \mu_k$. A vector bundle is semi-stable if and only if it is it own Harder-Narasimhan filteration, and if and only if its Harder-Narasimhan filteration is a single line from $(0,0)$ to (r, d).

Part 4. GIT quotient and sympletic quotient: the Kempf-Ness theorem

In this section, we mainly follows [Nov12] and [Bra12].

12. Geometric invariant theory

12.1. Introduction.
Many objects we want to take a quotient always have some sort of geometric structures, and we desire the quotients we obtain preserve geometric structure, for example:

Example 12.1.1. Suppose G is a Lie group and X is a smooth manifold, the quotient X/G will not always have the structure of a smooth manifold (For example, the presence of non-closed orbits, usually gives a non-Hausdorff quotient). However, if G acts properly and freely, then X/G has a smooth manifold structure, such that natural projection $\pi : X \to X/G$ is a smooth submanifold.

Geometric invariant theory (GIT) is the study of such question in the context of algebraic geometry, for example:

Example 12.1.2. Let $M_n(\mathbb{C})$ be the group of all $n \times n$ matrices over $\mathbb{C}$, then it can be given a geometric structure by regarding it as an affine variety. Consider the conjugate action of $\mathrm{GL}_n(\mathbb{C})$ on $M_n(\mathbb{C})$. Can we regard $M_n(\mathbb{C})/\operatorname{GL}_n(\mathbb{C})$ as a variety?

The answer of above question is yes, but good thing does not happen always, consider

Example 12.1.3. Let $\mathbb{C}^\times$ acts on $\mathbb{C}^2$ by $\lambda(x, y) := (\lambda x, \lambda y)$. The $\mathbb{C}^\times$-orbits are $\{(\lambda x, \lambda y) : \lambda \in \mathbb{C}^\times, (x, y) \neq (0, 0)\}$ as well as the origin $\{(0, 0)\}$. Now suppose that the set of orbits is a variety, then every point must be closed

So we need to be more careful when we constructing quotients in the category of varieties. As we have seen in smooth manifold, we can guess

1. only certain types of group (compared with Lie group) are allowed;
2. only certain types of group actions (compared with properly and freely) are allowed.

12.2. Good categorical quotient.

Definition 12.2.1 (G-invariant morphism)**.** A morphism $f : X \to Y$ is called G-invariant morphism, if it is constant on orbits.

Definition 12.2.2 (categorical quotient)**.** In any category, we call a G-invariant morphism $\pi : X \to Y$ is categorical quotient of X by G, when for any G-invariant morphism $f : X \to Z$, we have that f factors uniquely through π, that is

$$\begin{tikzcd} X \arrow[rr, "f"] \arrow[dr, "\pi"'] & & Z \\ & Y \arrow[ur, dotted, "\overline{f}"'] & \end{tikzcd}$$

Remark 12.2.1. Since categorical quotient is defined by its universal property, so it is unique when it exists.

However, for a quotient in the category of varieties, simple being a categorical quotient may not have a good geometric properties, so we need to define good categorical quotient. If G acts on a variety X, then we can get an action on the regular functions on X as follows. For $f \in \mathcal{O}(U), U \subset X$, we define

$$gf(x) = f(g^{-1}.x)$$

For the types of group action we are interested in, we require

$$gf \in \mathcal{O}(U), \quad \forall f \in \mathcal{O}(U)$$

Definition 12.2.3. A surjective G-invariant map of varieties $p : X \to Y$ is called a good categorical quotient of X by G, if the following three properties holds

1. For all open $U \subset Y$, $p^* : \mathcal{O}(U) \to \mathcal{O}(p^{-1}(U))^G$ is an isomorphism.
2. If $W \subseteq X$ is closed and G-invariant, then $p(W) \subset Y$ is closed.
3. If $V_1, V_2 \subseteq X$ are closed, G-invariants, and $V_1 \cap V_2 = \varnothing$, then $p(V_1) \cap p(V_2) = \varnothing$.

Remark 12.2.2. Note that the first requirement implies a good categorical quotient must be a categorical one: If $f : X \to Z$ is a G-invariant morphism, then $f^* : \mathcal{O}(Z) \to \mathcal{O}(X)$ must embed in $\mathcal{O}(X)^G$. If p is a good categorical quotient, then p^* is an isomorphism to $\mathcal{O}(X)^G$, so

$$\begin{tikzcd} \mathcal{O}(Z) \arrow[rr, "f^*"] \arrow[dr, dotted, "\overline{f^*}"'] & \arrow[d, phantom, "\circlearrowleft"] & \mathcal{O}(X)^G \arrow[r, hook] & \mathcal{O}(X) \\ & \mathcal{O}(Y) \arrow[ur, hook, two heads, "p^*"'] & & \end{tikzcd}$$

So f^* can factor through $\mathcal{O}(Y)$, and this factoring is unique since p^* is an isomorphism. By the anti-equivalence of category, the dual $f = \overline{f} \circ p$ is a unique factoring of f through p.

Remark 12.2.3. As we can see in the above Remark 12.2.2, the first requirement already implies categorical quotient, the more restrictions intend to avoid bad situation in geometry, such as Example 12.1.3

Notation 12.2.1. We denote by $X//G$ the good categorical quotient, or GIT quotient, of a variety X by a group G.

In the following, we will first construct GIT quotient in affine case, and this serves as a guide for projective case: we want to glue affine quotients to get projective one, since every projective variety admits an affine covering.

Unfortunately, we can not cover the whole of a projective variety, which leads to the concept of semistability.

It's natural to define $X//G = \operatorname{Spec}\mathcal{O}(X)^G$ in affine cases, since $X = \operatorname{Spec}\mathcal{O}(X)$, so G-invariant regular functions may representate the quotient we desire, but for this we require that $\mathcal{O}(X)^G$ is finitely generated.

Historically, whether the ring of invariants is finitely generated or not is knowns as Hilbert's 14-th problem. For general linear group over $\mathbb{C}$, Hilbert showed that the invariant rings are always finitely generated. However, Nagata gave an counterexample that $\mathcal{O}(X)^G$ is not finitely generated, and proved that for any reductive group, $\mathcal{O}^G$ is finitely generated, see Lemma 12.3.1.

12.3. **Reductive groups.** Now we focus on the reductive group which we can use to construct GIT quotient. We will define when a linear algebraic group is reductive and give some properties of it.

Definition 12.3.1 (algebraic group)**.** A (linear) algebraic group is a subgroup of $\mathrm{GL}_n(k)$ which is an affine variety, that is an irreducible algebraic set.

Example 12.3.1. The set of unitary matrices with determinant 1

$$\mathrm{SO}(2) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid ad - bc - 1 = 0 \right\}$$

is an algebraic group[4].

Example 12.3.2. $k^\times$ is also an algebraic group, by the embedding $\lambda \to \lambda I$.

Example 12.3.3. $\mathrm{GL}_n(k)$ is an algebraic group[5].

Definition 12.3.2. A linear algebraic group G over k is reductive if every representation $\rho : G \to \mathrm{GL}_n(k)$ has a decomposition as a direct sum of irreducible representations.

Proposition 12.3.1 (Maschke)**.** Let G be a finite group, then G is reductive.

Proposition 12.3.2. The multiplicative group $\mathbb{C}^\times$ is reductive.

Proof. Let $\rho : \mathbb{C}^\times \to \mathrm{GL}_n(\mathbb{C})$ be a representation of $\mathbb{C}^\times$, we will show ρ has a decomposition as a direct sum of irreducible representations. Assume ρ is not irreducible. Let $\langle -, - \rangle$ denote the standard inner product on $V = \mathbb{C}^n$, then define

$$\langle x, y \rangle := \int_0^{2\pi} \langle \rho(e^{i\theta})x, \rho(e^{i\theta})y \rangle \mathrm{d}\theta$$

[4]In general, special linear group $\mathrm{SL}(n)$ is always an algebraic group by considering the irreducible polynomial $\det -1$.

[5]We can check this by introducing a new variable T and consider irreducible polynomial $T \cdot \det -1$ with $n^2 + 1$ variables.

This form has the following property: $\langle\rho(g)x,\rho(g)y\rangle = \langle x; y\rangle$, where $x, y \in V, g = e^{i\psi} \in S^1 = \{z \in \mathbb{C}^\times : |z| = 1\}$. Indeed,

$$\begin{aligned}\langle\rho(e^{i\psi})x,\rho(e^{i\psi})y\rangle &= \int_0^{2\pi}\langle\rho(e^{i\theta}\rho(e^{i\psi}))x,\rho(e^{i\theta})\rho(e^{i\psi})y\rangle \mathrm{d}\theta\\ &= \int_0^{2\pi}\langle\rho(e^{i(\theta+\psi)})x,\rho(e^{i(\theta+\psi)})y\rangle \mathrm{d}\theta\\ &\overset{\phi=\theta+\psi}{=} \int_0^{2\pi}\langle\rho(e^{i\phi})x,\rho(e^{i\phi})y\rangle \mathrm{d}\phi\\ &= \langle x, y\rangle\end{aligned}$$

And also note that $\langle$-,-$\rangle$ is an inner product. If ρ is not irreducible, then there exists some $\mathbb{C}^\times$-invariant subspace U of V, let $W = U^\perp$ be the orthogonal complement of U with respect to $\langle$-,-$\rangle$. Then we can see W is S^1-invariant as follows

$$\begin{aligned}\langle u,\rho(g)w\rangle &= \langle\rho(g^{-1})u,\rho(g^{-1})\rho(g)w\rangle\\ &= \langle\rho(g^{-1})u, w\rangle\\ &= 0\end{aligned}$$

where $w \in W, u \in U, g \in S$. The last equality holds since U is S^1-invariant. What we need to do is to show W is $\mathbb{C}^\times$-invaraint.

Let N be the subset of $\mathbb{C}^\times$ which leaves W invariant, it contains S obviously. We will show that this set is closed in the Zariski topology. If we can do this, since all Zariski closed subset in $\mathbb{C}^\times$ are finite sets and whole space, so we can conclude $N = \mathbb{C}^\times$, as desired.

Let $W = \operatorname{span}\{e_1, \ldots, e_r\}$, and extends this basis to a basis $\{e_1, \ldots, e_n\}$ of V. Then we can regard W as solutions of equations

$$\langle v, e_i\rangle = 0, \quad i = r+1, \ldots, n$$

these define polynomials which take the coordinate of v as variables, which we call it f_i, so we can see W as a zero set of $\{f_{r+1}, \ldots, f_n\}$.

For each $i \in \{1, \ldots, r\}, j \in \{r+1, \ldots, n\}$, consider the set $\{T \in \mathrm{GL}(V) \mid f_j(Te_i) = 0\}$. If we fix i, j, this set is the zero set of a polynomial in the coordinates of T. So it's a closed set in $\mathrm{GL}(V)$, with respect to Zariski topology. Then we have $\{T \in \mathrm{GL}(V) \mid Te_i \in W\} = \bigcap_{j=r+1}^n \{T \in \mathrm{GL}(V) \mid f_j(Te_i) = 0\}$ is closed, so

$$\{T \in \mathrm{GL}(V) \mid Te_i \in W, \forall i \in \{1, \ldots, r\}\} = \bigcap_{i=1}^r \{T \in \mathrm{GL}(V) \mid Te_i \in W\}$$

is closed, thus we have

$$\begin{aligned}\{T \in \mathrm{GL}(V) \mid Tw \in W, \forall w \in W\} &= \{T \in \mathrm{GL}(V) : T(\lambda_1 e_1 + \ldots + \lambda_r e_r) \in U \text{ for all } \lambda_i \in \mathbb{C}\}\\ &= \{T \in \mathrm{GL}(V) : \lambda_1 (Te_1) + \ldots + \lambda_r (Te_r) \in U \text{ for all } \lambda_i \in \mathbb{C}\}\\ &= \{T \in \mathrm{GL}(V) : Te_i \in W \text{ for each } i \in \{1, 2, \ldots, r\}\}\end{aligned}$$

is closed with respect to Zariski topology, so $N = \rho^{-1}(\{T \in \mathrm{GL}(V) \mid Tw \in W, \forall w \in W\})$ is closed, as we desired. □

Remark 12.3.1. In fact, many classical groups such as $\mathrm{GL}_n(\mathbb{C}), \mathrm{SL}_n(\mathbb{C})$ are reductive, now we give a proof of $\mathbb{C}^\times$ is a reductive group.

Definition 12.3.3 (rationally)**.** For a reductive alegbraic group, we say that G acts rationally on a variety X if it acts by a morphism of varieties $G \times X \to X$.

But why we need reductive groups? and why this action? There are two key properties which might answer these questions.

Lemma 12.3.1. Let G be a reductive group acting rationally on an affine variety X, then $\mathcal{O}(X)^G$ is finitely generated.

Proof. See [CW04]. □

The following lemma is used in the construction of GIT quotient. It allows us to find a G-invariant function which separates disjoint G-invariant sets.

Lemma 12.3.2. Let G be a reductive group acting rationally on an affine variety $X \subset \mathbb{A}^n$. Let Z_1, Z_2 be two closed G-invariant subsets of X with $Z_1 \cap Z_2 = \varnothing$. Then there exists a G-invariant function $F \in \mathcal{O}(X)^G$ such that $F(Z_1) = 1, F(Z_2) = 0$.

Proof. See [Bra12]. □

12.4. **The affine quotient.** We now have enough tools to construct the quotient of an affine variety by a reductive group. For an affine variety X, the quotient of X by a reductive group G is just $\operatorname{Spec} \mathcal{O}(X)^G$. We will prove that this construction satisfies the required conditions being a good categorical quotient.

Theorem 12.4.1. Let X be an affine variety and G be a reductive group acting rationally on X. Let $p^* : \mathcal{O}(X)^G \to \mathcal{O}(X)$ be defined by the inclusion $\mathcal{O}(X)^G \subseteq \mathcal{O}(X)$. Then the dual of this map, $p : X \to Y := \operatorname{Spec} \mathcal{O}(X)^G$ is a good categorical quotient.

Now we give a concrete example to show how powerful the GIT construction is, and gives the answer to the Example 12.1.1 we mentioned at first.

Example 12.4.1. Consider the set X of 2×2 matrices over $\mathbb{C}$, embedded in $\mathbb{C}^4$ by

$$\begin{pmatrix} w & x \\ y & z \end{pmatrix} \mapsto (w, x, y, z)$$

It is an affine variety obviously, and consider the general linear group acts on it by conjugate action, then as the theorem above implies

$$X//G = \operatorname{Spec} k[w, z, y, z]^G$$

We know that there are two important invariants under conjugate action, that is, determinant and trace. In this case they are det $= wz - xy$ and tr $= w + z$, so we have an obvious inclusion

$$k[wz - xy, w + z] \subset k[w, x, y, z]^G$$

We will show that we in fact have equality.

Let $\lambda \in \mathbb{C}^\times$ be arbitrary and consider the matrix $A = \begin{pmatrix} 0 & 1 \\ \lambda & 0 \end{pmatrix}$. For all matrices $M = \begin{pmatrix} w & x \\ y & z \end{pmatrix}$, we can calculate as follows

$$\begin{aligned} A^{-1}MA &= \begin{pmatrix} 0 & -\frac{1}{\lambda} \\ -1 & 0 \end{pmatrix} \begin{pmatrix} w & x \\ y & z \end{pmatrix} \begin{pmatrix} 0 & 1 \\ \lambda & 0 \end{pmatrix} \\ &= \begin{pmatrix} z & \frac{y}{\lambda} \\ \lambda x & w \end{pmatrix} \end{aligned}$$

Let $f \in k[w, x, y, z]^G$, i.e. we require f satisfy that $f(w, x, y, z) = A.f(M) = f(A.M) = f(A^{-1}MA) = f(z, \frac{y}{\lambda}, \lambda x, w)$. That is

$$f(w, x, y, z) = f\left(z, \frac{y}{\lambda}, \lambda x, w\right)$$

From this equality, we can make the following observations

1. x must appear in the form xy to cancel λ in $A.f$.
2. z and w must appear in an symmetric way, i.e. must in the forms of $z+w$ or zw.

So we conclude $f \in k[xy, wz, z+w]$. Similarly consider matrix $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. And after the same calculation we can get

$$f(w, x, y, z) = f(w - x, w + x - y - z, y, y + z)$$

As we already have $f \in k[xy, wz, w+z]$, we can reformulate this requirement into

$$f(xy, wz, w + z) = f\left(wy + xy - y^2 - z, wy + wz - y^2 - yz, w + z\right)$$

We can see that this formular holds only when extra terms in $B.f$ must cancel with each other, which implies $f \in k[wz - xy, w + z]$, as desired. So we have the construction

$$\begin{aligned} X//G &= \operatorname{Spec} k[w, x, y, z]^G \\ &= \operatorname{Spec} k[wz - xy, w + z] \\ &= \operatorname{Spec} k[u, v] \\ &= \mathbb{C}^2 \end{aligned}$$

Remark 12.4.1. There is a high-dimensional analogous: if $\mathrm{GL}_n(\mathbb{C})$ acts on $M_n(\mathbb{C})$ by conjugate action, then

$$M_n(\mathbb{C}) // \mathrm{GL}_n(\mathbb{C}) = \mathbb{C}^n$$

See [Bri10] for more details.

12.5. **The projective quotient.** Now we construct projective quotient by gluing together affine quotients.

Let X be a projective variety, then X can be covered by some affine varieties X_{f_i}. In order to construct GIT quotient of X by G, it's natural for us to take quotient for every affine variety of G of the form $X_{f_i}//G = \mathrm{Spec}(\mathcal{O}(X_{f_i})^G)$, and cover the projective quotient by them. To do this, we need an action of G on the coordinates of X.

Our approach is to embed X in $\mathbb{P}^m$ for some m such that the action of G can be extended to a linear action on $\mathbb{A}^{m+1}$. This is called a linearisation of the action of G.

Definition 12.5.1. Let the group G act rationally on a projective variety X. Let $\varphi : X \hookrightarrow \mathbb{P}^m$ be an embedding of X that extends the group action, i.e. we have a rationally group action on $\mathbb{P}^m$ such that $\varphi(g.x) = g.\varphi(x)$. Let $\pi : \mathbb{A}^{m+1} \to \mathbb{P}^m$ be the natural projection. A linearisation of the action of G with respect to φ is a linear action of G on $\mathbb{A}^{m+1}$ that is compatible with the action of G on X in the following sense

1. For any $y \in \mathbb{A}^{m+1}, g \in G$
$$\pi(g.y) = g.(\pi(y))$$
2. For all $g \in G$, the map
$$\mathbb{A}^{m+1} \to \mathbb{A}^{m+1}, \quad y \mapsto g.y$$
is linear.

We write φ_G for a linearisation of the action of G with respect to φ.

Remark 12.5.1. Note that such action induces an action of G on $\mathcal{O}(X)$. we have $\mathcal{O}(X) \cong k[x_0, \dots, x_m]/I$ for some homogeneous ideal I, since X is isomorphic to the image $\varphi(X) \subseteq \mathbb{P}^m$. Using the fact that G acts on $k[x_0, \dots, k_m]$ by $g.f(x_0, \dots, x_m) := f(g^{-1}.(x_0, \dots, x_m))$, we can know that G also acts on $\mathcal{O}(X)$, and it's well-defined, since $g.f' \in I$ for $f' \in I$.

Example 12.5.1. Let $\mathbb{C}^\times$ act on $\mathbb{P}^1$ by $\lambda.(x_0, x_1) = (x_0 : \lambda x_1)$. A linearisation can be given by the obvious action on $\mathbb{A}^2$ with $\lambda.(x_0, x_1) = (x_0, \lambda x_1)$.

The above example illustrates a quite important issue when we are constructing projective quotient: good categorical quotient may not exist. The only possible G-invariant morphism sends all orbits to a point, since $(1, 0), (0, 1)$ are both in the closure of $(1, t)$. But this fails to seperate closed orbits, so is not a good categorical quotient.

The solution to such problem is to take an open G-invariant subset which has a good categorical quotient. We desire this subset to be covered by G-invariant open affine subsets so that we can cover the quotient by gluing together affine quotients. This leads us to the notion of semistability,

Definition 12.5.2. Let G be a reductive group acting on a projective variety X which has an embedding $\varphi : X \to \mathbb{P}^m$. A point $x \in X$ is called semistable (with respect to the linearisation φ_G) if there exists some G-invariant homogeneous polynomial f of degree greater than 0 in $\mathcal{O}(X)$, such that $f(x) \neq 0$ and X_f is affine.

Remark 12.5.2. Write $X^{\mathrm{as}}(\varphi_G)$ for the set of semistable points of X with respect to φ_G, or just X^{as} when it's not ambiguous.

For Example7.4.3, the set of semistable points of X with respect to φ_G is $X^{\mathrm{as}} = X_{x_0} = \mathbb{P}^1 \backslash \{(0:1)\}$. On this subset, the map to a point $p : X^{\mathrm{as}} \to \mathbb{P}^0$ is indeed a good categorical quotient.

Theorem 12.5.1. Let G be a reductive group acting rationally on a projective variety X embedded in $\mathbb{P}^m$ with a linearisation φ_G. Let R be the coordinate ring of X, then there is a good categorical quotient

$$p : X^{\mathrm{as}}(\varphi_G) \to X^{\mathrm{as}(\varphi_G)}//G \cong \operatorname{Proj} R^G$$

13. Symplectic quotient

14. The Kempf-Ness theorem

References

[Ati57] M. F. Atiyah. Vector bundles over an elliptic curve. *Proceedings of the London Mathematical Society*, s3-7(1):414–452, 1957.

[Bra12] Matthew Brassil. Geometric invariant theory, 2012.

[Bri10] Michel Brion. Introduction to actions of algebraic groups. *Les cours du CIRM*, 1(1):1–22, 2010.

[CW04] Harold Edward Alexander Eddy Campbell and David L Wehlau. *Invariant theory in all characteristics*, volume 35. American Mathematical Soc., 2004.

[Gro57] A. Grothendieck. Sur la classification des fibres holomorphes sur la sphere de riemann. *American Journal of Mathematics*, 79(1):121–138, 1957.

[HN75] G. Harder and M. S. Narasimhan. On the cohomology groups of moduli spaces of vector bundles on curves. *Mathematische Annalen*, 212(3):215–248, 1975.

[JM74] James D. Stasheff John Milnor. *Characteristic Classes*. Princeton University Press, 1974.

[Liu22] Bowen Liu. Spectral sequences and its application, 2022.

[Mil56] John W. Milnor. Construction of universal bundles, ii. *Annals of Mathematics*, 63:272, 1956.

[Mit01] Stephen A. Mitchell. Notes on principal bundles and classifying spaces. 2001.

[Mum62] David Mumford. Projective invariants of projective structures and applications. In *Proc. Internat. Congr. Mathematicians (Stockholm, 1962)*, pages 526–530, 1962.

[Nov12] Ramón Alejandro Urquijo Novella. Git quotients and symplectic reduction: the kempf-ness theorem, 4 2012.

[Ses67] C. S. Seshadri. Space of unitary vector bundles on a compact riemann surface. *Annals of Mathematics*, 85(2):303–336, 1967.

Yau Mathematical Sciences Center, Tsinghua University, Beijing, 100084, P.R. China,

Email address: `liubw22@mails.tsinghua.edu.cn`