LECTURE 30: CHERN-WEIL THEORY

1. Invariant Polynomials

We start with some necessary backgrounds on invariant polynomials. Let V be a vector space. Recall that a k-tensor $T \in \otimes^k V^*$ is called *symmetric* if

$$T(v_{\sigma(1)}, \cdots, v_{\sigma(k)}) = T(v_1, \cdots, v_k), \qquad \forall \sigma \in S_k.$$

We will denote the space of all symmetric k-tensors on V by $\mathsf{S}^k V^*$. Like the wedge product, we can define a *symmetric product* $\circ : \mathsf{S}^k V^* \times \mathsf{S}^l V^* \to \mathsf{S}^{k+l} V^*$ via

$$T_1 \circ T_2(v_1, \cdots, v_{k+l}) = \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} T_1(v_{\sigma(1)}, \cdots, v_{\sigma(k)}) T_2(v_{\sigma(k+1)}, \cdots, v_{\sigma(k+l)}).$$

As usual we write $\mathsf{S}^0 V^* = \mathbb{R}$.

Let $T \in \mathsf{S}^k V^*$ be any symmetric k-tensor on V. Then T induces a map $P_T : V \to \mathbb{R}$ by

$$P_T(v) := T(v, \cdots, v).$$

The map P_T is called a "degree k homogeneous polynomial on V" since it satisfies

$$P_T(tv) = t^k P_T(v), \qquad \forall t \in \mathbb{R}.$$

Conversely, given any degree k homogeneous polynomial P_T on V, one can recover $T \in \mathsf{S}^k V^*$ by the standard *polarization formula*

$$T(v_1, \cdots, v_k) := \frac{1}{k!} \frac{\partial^k}{\partial t_1 \cdots \partial t_k} P_T(t_1 v_1 + \cdots + t_k v_k).$$

(Note: $P_T(t_1 v_1 + \cdots + t_k v_k)$ is a degree k homogeneous polynomial in $t_1, \cdots, t_k$.)

Exercise

1. Check that the correspondence $T \leftrightsquigarrow P_T$ is bijective.
2. Prove: For any symmetric tensors T_1, T_2 on V, one has $P_{T_1 \circ T_2} = P_{T_1} P_{T_2}$.

In applications we will take $V = \mathfrak{g}$ be the Lie algebra of a Lie group G. For simplicity we assume $G \subset \mathrm{GL}(r, \mathbb{R})$ be a linear Lie group. Then the adjoint action of G on $\mathfrak{g}$ induces a G-action on $\mathsf{S}^k(\mathfrak{g}^*)$ by

$$(g \cdot T)(X_1, \cdots, X_k) = T(gX_1 g^{-1}, \cdots, gX_k g^{-1}), \qquad \forall g \in G, X_i \in \mathfrak{g}.$$

Definition 1.1. A symmetric k-tensor $T \in \mathsf{S}^k(\mathfrak{g}^*)$ is called *G-invariant* if

$$g \cdot T = T, \qquad \forall g \in G.$$

The set of all G-invariant elements in $\mathsf{S}^k(\mathfrak{g}^*)$ is denoted by $I^k(G)$.

By definition it is easy to see $T \in \mathsf{S}^k(\mathfrak{g}^*)$ is *G-invariant* if and only if P_T is G-invariant.

Example. Consider $G = \mathrm{GL}(r,\mathbb{R})$. For any positive integer k, we let p_k denote the degree k homogeneous polynomial in the expansion

$$\det\left(\lambda I - \frac{1}{2\pi}A\right) = \sum_{k=0}^{r} p_k(A)\lambda^{r-k}, \qquad \forall A \in \mathfrak{gl}(r,\mathbb{R}).$$

More explicitly,

$$p_0(A) = 1, \quad p_1(A) = -\frac{\mathrm{Tr}A}{2\pi}, \quad p_2(A) = \frac{(\mathrm{Tr}A)^2 - \mathrm{Tr}(A^2)}{2(2\pi)^2}, \quad \cdots, \quad p_r(A) = (-1)^r\frac{\det A}{(2\pi)^r}.$$

Obviously each p_k is G-invariant. So by the correspondence above, we get $T_{p_k} \in I^k(\mathrm{GL}(r,\mathbb{R}))$. They are called *Pontrjagin polynomials*.

Example. Consider $G = \mathrm{SO}(r)$, where $r = 2p$ is even. For any $A = (a^i_j) \in \mathfrak{so}(r)$ we let

$$\mathrm{Pf}(A) = \frac{1}{(4\pi)^{r/2}(r/2)!}\sum_{\sigma\in S_r}(-1)^\sigma a^{\sigma(1)}_{\sigma(2)}a^{\sigma(3)}_{\sigma(4)}\cdots a^{\sigma(r-1)}_{\sigma(r)}.$$

It is homogeneous of degree $r/2$. One can check that for any $A = (a^i_j) \in \mathfrak{so}(r)$,

$$[\mathrm{Pf}(A)]^2 = \frac{\det A}{(2\pi)^r}.$$

In particular, Pf is an $\mathrm{SO}(r)$-invariant. It is called the *Pfaff polynomial.*

Here is another description of the Pfaffian: Given any skew-symmetric matrix $A \in \mathfrak{so}(r)$, where $r = 2p$, one can construct a linear 2-form $\omega = \sum_{i<j} a_{ij}e^i \wedge e^j$, where $e^1, e^2, \cdots, e^{2p}$ is the standard basis of $(\mathbb{R}^{2p})^*$. Then $\mathrm{Pf}(A)$ is the number such that

$$\frac{1}{p!}\frac{\omega^p}{(2\pi)^p} = \frac{\omega\wedge\cdots\wedge\omega}{p!(2\pi)^p} = \mathrm{Pf}(A)e^1\wedge e^2\wedge\cdots\wedge e^{2p}.$$

G-invariant symmetric tensors admits the following nice property:

Proposition 1.2. *For any $T \in I^k(G)$ and any $X, X_1, \cdots, X_k \in \mathfrak{g}$, we have*

$$T([X,X_1],X_2,\cdots,X_k) + \cdots + T(X_1,\cdots,X_{k-1},[X,X_k]) = 0.$$

Proof. By definition, T is a k-tensor, i.e. it is multi-linear. So the conclusion follows from

$$\begin{aligned} 0 &= \frac{d}{dt}\Big|_{t=0} T(e^{tX}X_1e^{-tX},\cdots,e^{tX}X_ke^{-tX}) \\ &= T(\frac{d}{dt}\Big|_{t=0} e^{tX}X_1e^{-tX}, X_2,\cdots,X_k) + \cdots + T(X_1,\cdots,X_{k-1},\frac{d}{dt}\Big|_{t=0} e^{tX}X_ke^{-tX}) \\ &= T([X,X_1],X_2,\cdots,X_k) + \cdots + T(X_1,\cdots,X_{k-1},[X,X_k]). \end{aligned}$$

□

Now suppose $T \in I^k(G)$ is a symmetric k-tensor on $\mathfrak{g}$. For $\mathfrak{g}$-valued differential forms $\eta_1, \cdots, \eta_k \in \Omega^*(U) \otimes \mathfrak{g}$ one can define $T(\eta_1, \cdots, \eta_k) \in \Omega^*(U)$ by extending linearly the relation

$$T(\omega_1 \otimes X_1, \cdots, \omega_k \otimes X_k) := (\omega_1 \wedge \cdots \wedge \omega_k) T(X_1, \cdots, X_k).$$

Note that if ω is a 2-form on M, or more generally is any even-form on M, then for any differential form $\tilde{\omega}$ on M, one has $\omega \wedge \tilde{\omega} = \tilde{\omega} \wedge \omega$. As a consequence we see

Corollary 1.3. *If $\eta_1, \cdots, \eta_k \in \Omega^{\text{even}}(U) \otimes \mathfrak{g}$, then for any $T \in I^k(G)$ and $\eta \in \Omega^*(U) \otimes \mathfrak{g}$ one has*

$$T([\eta, \eta_1], \eta_2, \cdots, \eta_k) + \cdots + T(\eta_1, \cdots, \eta_{k-1}, [\eta, \eta_k]) = 0.$$

We remark that in the above expression,

$$[\eta_1, \eta_2] := \eta_1 \wedge eta_2 - (-1)^{k_1 k_2} \eta_2 \wedge \eta_1, \quad \forall \eta_i \in \Omega^{k_i}(U) \otimes \mathfrak{g},$$

where the wedge product $\wedge$ between two $\mathfrak{g}$-valued differential forms were defined last time:

$$(\omega_1 \otimes X_1) \wedge (\omega_2 \otimes X_2) := (\omega_1 \wedge \omega_2) \otimes (X_1 X_2).$$

(As a consequence, if $\omega_i \in \Omega^{k_i}(U)$ and $A_i \in \mathfrak{g}$, then

$$[\omega_1 \otimes X_1,\ \omega_2 \otimes X_2] = (\omega_1 \wedge \omega_2) \otimes [X_1, X_2].$$

This can be viewed as an alternative definition.)

2. Chern-Weil Theory

Let E be a rank r vector bundle over M, and ∇ a connection on M. Let

$$R_\nabla \in \Omega^2(M; \mathrm{End}(E))$$

be the curvature 2-form of ∇. As we have seen last time, locally after choosing a local frame of E, one can represent ∇ by the matrix of connection 1-form

$$\Gamma \in \Omega^1(U) \otimes \mathfrak{gl}(r, \mathbb{R}),$$

and

$$R_\nabla = d\Gamma + \Gamma \wedge \Gamma \in \Omega^2(U) \otimes \mathfrak{gl}(r, \mathbb{R}).$$

The Bianchi identity reads

$$dR_\nabla = R_\nabla \wedge \Gamma - \Gamma \wedge R_\nabla = [R_\nabla, \Gamma].$$

Now suppose $T \in I^k(G)$, where $G = \mathrm{GL}(r, \mathbb{R})$. We can define

$$P_T(R_\nabla) := T(R_\nabla, \cdots, R_\nabla) \in \Omega^{2k}(U).$$

Since T is G-invariant and since $\widetilde{R}_\nabla = g^{-1} R_\nabla g$ in different local frame (where $g \in \mathrm{GL}(r, \mathbb{R})$ is the "transferring matrix"), we see that $P_T(R_\nabla)$ in different trivialization charts can be glued together to a globally-defined $2k$-form

$$P_T(R_\nabla) \in \Omega^{2k}(M).$$

(Note R_∇ is not an element in $\Omega^2(M)$: it sits in $\Omega^2(M; \mathrm{End}(E))$.)

Now we state the main theorem in Chern-Weil theory:

Theorem 2.1 (Chern-Weil)**.** *Let E be a vector bundle over M. Then*

1. *For any $T \in I^k(G)$ and any linear connection ∇ on E, $P_T(R_\nabla)$ is a closed $2k$-form.*
2. *The de Rham cohomology class $[P_T(R_\nabla)] \in H^{2k}_{dR}(M)$ is independent of the choices of ∇.*
3. *The Chern-Weil map*
$$\mathcal{CW} : (I^*(G), \circ) \to (H^*_{dR}(M), \wedge), \quad T \mapsto [P_T(R_\nabla)]$$
is a ring homomorphism.

Proof. (1) According to the Bianchi identity $dR_\nabla = [R_\nabla, \Gamma]$ and Corollary 1.3,
$$dP_T(R_\nabla) = dT(R_\nabla, \cdots, R_\nabla) = T(dR_\nabla, \cdots, R_\nabla) + \cdots + T(R_\nabla, \cdots, dR_\nabla) = 0.$$
So $P_T(R_\nabla)$ is closed. (The second equality follows from the fact that T is multi-linear.)

(2) The idea to prove that the de Rham cohomology class $[P_T(R_\nabla)]$ is independent of the choices of the linear connection ∇ is to construct a chain homotopy, as we did in Lecture 24. We let ∇^0 and ∇^1 be two connections and let Γ^0, Γ^1 be the (local) matrices of connection 1-forms. By definition it is easy to check that the collection of matrices of 1-forms
$$\widetilde{\Gamma} = (1-s)\Gamma^0 + s\Gamma^1 \in \Omega^1(U \times \mathbb{R}) \otimes \mathfrak{gl}(r, \mathbb{R})$$
defines a connection $\widetilde{\nabla}$ on a new vector bundle $E \times \mathbb{R}$ over $M \times \mathbb{R}$ in the obvious way. (The connection $\widetilde{\nabla}$ can be constructed globally as follows: Consider the canonical projection map $\pi : M \times \mathbb{R} \to M$. Then the pull-back π^*E is a rank r vector bundle over $M \times \mathbb{R}$. Moreover, the pull-backs $\pi^*\nabla^0$ and $\pi^*\nabla^1$ are two connections on π^*E. It follows that $\widetilde{\nabla} = (1-s)\nabla^0 + s\nabla^1$ is a linear connection on π^*E.) We need

Lemma 2.2. *Let $\iota_0, \iota_1 : M \to M \times \mathbb{R}$ be the inclusions*
$$\iota_0(x) = (x, 0), \qquad \iota_1(x) = (x, 1).$$
Then there exists a collection of linear operators $Q : \Omega^k(M \times \mathbb{R}) \to \Omega^{k-1}(M)$ so that
$$\iota_0^*\omega - \iota_1^*\omega = dQ(\omega) - Q(d\omega), \qquad \forall \omega \in \Omega^k(M \times \mathbb{R}).$$

Proof. According to Lemma 2.7 in Lecture 24 (see line -9 in page 6 there), there exists a linear map $\widetilde{Q} : \Omega^k(M \times \mathbb{R}) \to \Omega^{k-1}(M \times \mathbb{R})$ so that
$$\omega - \phi_1^*\omega = d\widetilde{Q}\omega + \widetilde{Q}d\omega,$$
where $\phi_1 : M \times \mathbb{R} \to M \times \mathbb{R}$ is the map $\phi_1(p, a) = (p, a+1)$. It follows that
$$\iota_0^*\omega - \iota_1^*\omega = \iota_0^*\omega - \iota_0^*\phi_1^*\omega = \iota_0^*d\widetilde{Q}\omega + \iota_0^*\widetilde{Q}d\omega = d(\iota_0^*\widetilde{Q})\omega + (\iota_0^*\widetilde{Q})d\omega.$$
So the conclusion holds for $Q = \iota_0^*\widetilde{Q} : \Omega^k(M \times \mathbb{R}) \to \Omega^{k-1}(M)$. □

Back to the proof of the main theorem. By definition
$$\iota_0^*\widetilde{\Gamma} = \Gamma^0 \quad \text{and} \quad \iota_1^*\widetilde{\Gamma} = \Gamma^1.$$
As a result, we see
$$\iota_0^*R_{\widetilde{\nabla}} = R_{\nabla^0} \qquad \text{and} \qquad \iota_1^*R_{\widetilde{\nabla}} = R_{\nabla^1}.$$

So according to the above lemma,

$$P_T(R_{\nabla^0}) - P_T(R_{\nabla^1}) = \iota_0^* P_T(R_{\tilde{\nabla}}) - \iota_1^* P_T(R_{\tilde{\nabla}}) = dQ(P_T(R_{\tilde{\nabla}})) + Qd(P_T(R_{\tilde{\nabla}})).$$

But we just proved that $P_T(R_{\tilde{\nabla}})$ is closed. So $[P_T(R_{\nabla^0})] = [P_T(R_{\nabla^1})]$.
(3) It is straightforward to show that for $T \in I^k(G)$ and $S \in I^l(G)$,

$$P_{T\circ S}(R_\nabla) = P_T(R_\nabla) \wedge P_S(R_\nabla).$$

Details left as an exercise. (Note: the wedge products between 2-forms commute!) □

As a consequence, for any $T \in I^k(G)$ (where $G = \mathrm{GL}(r, \mathbb{R})$) and any rank r vector bundle E over M, one gets a de Rham cohomology class

$$T(E) := [P_T(R_\nabla)] \in \Omega^{2k}(M).$$

We have

Proposition 2.3. *Let $\varphi : N \to M$ be any smooth map, and let E be any vector bundle of rank r over M. Then for any $T \in I^*(G)$, one has*

$$\varphi^* T(E) = T(\varphi^* E).$$

The proof is straightforward (you have to play with the conceptions of pull-back line bundles, pull-back connections etc) and is left as an exercise.

Definition 2.4. For any $T \in I^*(G)$, the cohomology class $[P_T(R_\nabla)]]$ is called the *characteristic class* of the vector bundle E corresponding to T.

Obviously one has

- If two vector bundles E_1 and E_2 over M are isomorphic, then the characteristic classes $T(E_1) = T(E_2)$.
- If E is the trivial bundle, then $T(E) = [P_T(0)]$. (since one can take the trivial connection, so that $R_\nabla = 0$.)

Example. Let E be any vector bundle over M. Let p_k be the Pontrjagin polynomial that we alluded to above. Then

$$P_k(E) := [p_{2k}(R_\nabla)] \in H^{4k}_{dR}(M)$$

is called the kth *Pontrjagin characteristic class* of E. The Pontrjagin class $P_k(M)$ of M is defined to be the Pontrjagin class of TM. They are important topological invariants to study manifolds. For example, using the Pontrjagin classes one can compute (via the Hirzebruch signature theorem) the *signature* of a manifold. They are also used to compute the $\widehat{A}$-*genus* of a manifold which can be used to study the existence of positive scalar curvature metric.

(By choosing a Riemannian metric on E, one can always reduce the structural group of E from $\mathrm{GL}(r, \mathbb{R})$ to $\mathrm{O}(r)$ by taking orthonormal frames only. Using the new structure group, one can prove that $[p_{odd}(R_\nabla)]$ are all zero. See the computation below for Chern classes.)

Example. Now suppose E is an oriented vector bundle over M of rank r. Then the structural group of E can be reduced to $\mathrm{SO}(r)$. Thus we get a characteristic class

$$\chi(E) := [\mathrm{Pf}(R_\nabla)] \in H^r_{dR}(M).$$

This is nothing else but the *Euler characteristic class* of E that we mentioned in Lecture 28. As we have seen in Lecture 28, the Euler class of E is closely related to the problem "whether E admits a non-vanishing global section".

Example. All discussions above are over $\mathbb{R}$, but they can be generalized to objects over $\mathbb{C}$. For example, one can talk about *complex vector bundle* E over smooth manifold M: They are vector bundles over M whose fibers are $\mathbb{C}^r$ and whose structural groups are $\mathrm{GL}(r,\mathbb{C})$. For any $A \in \mathfrak{gl}(r,\mathbb{C})$ (= the set of all $r \times r$ complex matrices), one can define c_k to be such that

$$\det\left(\lambda I_r - \frac{1}{2\pi i}A\right) = \sum c_k(A)\lambda^{n-k}.$$

Again c_k is a homogeneous polynomial of degree k and is $\mathrm{GL}(r,\mathbb{C})$-invariant. As a result, for any complex vector bundle E of complex rank r, one gets a complex-valued de Rham cohomology class

$$c_k(E) := c_k(R_\nabla) \in H^{2k}_{dR}(M;\mathbb{C})$$

which is called the k^{th} *Chern class* of E. If one fix an Hermitian metric on E (i.e. fix a Hermitian inner product on each fiber of E which vary smoothly with respect to base points – this is always possible by using partition of unity), then one can reduce the structural group to $U(n)$. But for $A \in \mathfrak{u}(n)$ (= the set of all $r \times r$ complex matrices with $A + \overline{A}^T = 0$), one has

$$\det\left(\lambda I_r - \frac{1}{2\pi i}A\right) = \overline{\det\left(\bar{\lambda} I_r - \frac{1}{2\pi i}A\right)},$$

i.e.

$$\sum c_k(A)\lambda^{n-k} = \sum \overline{c_k(A)}\lambda^{n-k}.$$

It follows that each c_k is a real-valued polynomial and thus each $c_k(E) \in H^{2k}_{dR}(M)$, i.e. $c_k(E)$ is a (real-valued) de Rham cohomology class. They are of fundamental importance in algebraic topology, differential geometry, algebraic geometry and mathematical physics. For example, they appear in the famous Atiyah-Singer index theorem.

Remark. More generally, one can define characteristic classes for principle bundles.

Remark. Historically, characteristic classes first appeared in algebraic topology. In general, characteristic classes are very hard to compute. The Chern-weil theory gives a differential geometry description of various characteristic classes so that they are explicitly computable.

However, there are still many characteristic classes that don't have Chern-Weil type description. The first characteristic classes people found are the the Stiefel-Whitney classes $w_k(E) \in H^k(M;\mathbb{Z}_2)$. These classes are very useful in topology. For example, $w_k(TM)$ can be used to characterize whether a smooth n-dimensional manifold (without boundary) M can be realized as the boundary of a smooth $(n+1)$-dimensional manifold. One can also Stiefel-Whitney classes w_k to justify whether a smooth n-dimensional manifold M can be embedded into $\mathbb{R}^{n+k}$.

Remark. Finally we mention another important family of characteristic classes: the Wu classes. The Wu classes are different from the previously mentioned characteristic classes: they are not characteristic classes associated to vector bundles, but associated to manifolds M themselves. Using Wu classes one can

- Prove the Stiefel-Whitney class of a manifold (i.e. of the tangent bundle) is a homotopy invariant.
- Compute the Stiefel-Whitney classes of M via the Steenrod square operation. (The Wu formula)

There was another Wu class in USTC: the 1960-1965 class in mathematics. This class is the main origin of geometry and topology in USTC.

Dedicated to WU Wen-tsun (1919-2017)!

We are all "descendants" of Wu's class.

—The End—