

LECTURE 10: THE ATIYAH-GUILLEMIN-STERNBERG CONVEXITY THEOREM

CONTENTS

1. The Atiyah-Guillemin-Sternberg Convexity Theorem	1
2. Proof of the Atiyah-Guillemin-Sternberg Convexity theorem	3
3. Morse theory	6
4. Student presentation	9

1. THE ATIYAH-GUILLEMIN-STERNBERG CONVEXITY THEOREM

¶ The statement of AGS convexity theorem.

Let (M, ω) be a compact connected symplectic manifold. In this and next lecture, we will study the special case where $G = T$ is a compact connected abelian Lie group, acting in Hamiltonian fashion on M .

Since T is abelian, from standard Lie theory we know that $T \simeq \mathbb{T}^k$ for some k . It follows that $\mathfrak{t} \simeq \mathbb{R}^k$, and thus $\mathfrak{t}^* \simeq \mathbb{R}^k$. We have seen that in this case the moment map $\mu : M \rightarrow \mathfrak{t}^*$ is a T -invariant map.

The main theorem we want to prove is

Theorem 1.1 (Atiyah-Guillemin-Sternberg Convexity Theorem). *Let (M, g, T, μ) be a compact connected Hamiltonian \mathbb{T} -manifold. Then the image of μ is a convex polyhedron in \mathfrak{t}^* whose vertices are the image of the T -fixed points.*

Remark. In the general case of a Hamiltonian action of a compact Lie group G , the image of μ might be much more complicated. However, one can prove that the intersection of $\mu(M)$ with each *Weyl chamber* is a convex polyhedron.

¶ Examples.

Example. The moment map for the standard rotational S^1 -action on S^2 is

$$\mu(x, y, z) = z.$$

So the image of μ is the interval $[-1, 1]$. Observe that the pre-images of the points ± 1 are north/south poles of S^2 , which are exactly the fixed points of the S^1 -action.

Example. Consider the standard \mathbb{T}^2 on $\mathbb{C}\mathbb{P}^2$ via

$$(e^{it_1}, e^{it_2}) \cdot [z_0 : z_1 : z_2] = [z_0 : e^{it_1} z_1 : e^{it_2} z_2]$$

The moment map is

$$\mu([z_0 : z_1 : z_2]) = -\frac{1}{2} \left(\frac{|z_1|^2}{|z|^2}, \frac{|z_2|^2}{|z|^2} \right)$$

The image is a triangle in \mathbb{R}^2 with vertices $(0, 0)$, $(-\frac{1}{2}, 0)$, $(0, -\frac{1}{2})$. Observe that these vertices are the image of $[1 : 0 : 0]$, $[0 : 1 : 0]$ and $[0 : 0 : 1]$ of $\mathbb{C}\mathbb{P}^2$, which are exactly the fixed points of the T^2 -action. One can easily extend this example to arbitrary dimension, in which case the image of the moment map is a simplex.

Example. Hirzebruch surfaces.

¶ Application: Schur-Horn theorem.

Let A be an Hermitian matrix whose eigenvalues are

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

Arrange the diagonal entries of A in increasing order as

$$a_1 \leq a_2 \leq \cdots \leq a_n.$$

It was shown by Schur that for each $1 \leq k \leq n$,

$$\sum_{i=1}^k a_i \geq \sum_{i=1}^k \lambda_i.$$

Horn prove the converse: For any sequences λ_k and a_k satisfying all the above inequalities, there exists an Hermitian matrix A whose diagonal entries are a_k 's and whose eigenvalues are λ_k 's.

We can give a symplectic geometric proof of Schur-Horn's theorem. For any $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$, let \mathcal{H}_λ be the set of all $n \times n$ Hermitian matrices whose eigenvalues are precisely $\lambda_1, \dots, \lambda_n$. As was explained in Yuguo's presentation, \mathcal{H}_λ can be identified with a $U(n)$ -coadjoint orbit and thus is a symplectic manifold. The coadjoint action of $U(n)$ on \mathcal{H}_λ is a Hamiltonian action. This action restricts to a Hamiltonian action of the maximal torus $\mathbb{T}^n \simeq T \subset U(n)$ on \mathcal{H}_λ , whose moment map is

$$\mu : \mathcal{H}_\lambda \rightarrow \mathbb{R}^n, \quad A = (a_{ij})_{n \times n} \mapsto (a_{11}, a_{22}, \dots, a_{nn}) =: \vec{a}.$$

The fixed points of the T action are the diagonal matrices, whose diagonal entries has to be $\lambda_\sigma = (\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(n)})$, where σ is a permutation of $(1, 2, \dots, n)$. So according to the Atiyah-Guillemin-Sternberg convexity theorem, a vector \vec{a} is the diagonal of a Hermitian matrix in \mathcal{H}_λ if and only if \vec{a} lies in the convex hull of the points λ_σ . This is equivalent to the Schur-Horn inequalities.

Remark. Kostant extends Schur-Horn's theorem to more general coadjoint orbits.

¶ **Application: number of T -fixed points.**

As an application of the AGS convexity theorem, we can prove

Proposition 1.2. *Let (M, g) be a compact connected symplectic manifold admitting a Hamiltonian T -action. Suppose there exists $m \in M$ such that T acts locally free at m . Then there must be at least $k + 1$ fixed points, where $k = \dim T$.*

Proof. Since the action is locally free at m , the stabilizer T_m is a finite subgroup of T , and thus $\mathfrak{t}_m = \{0\}$. It follows that

$$\text{Im}(d\mu_m) = \mathfrak{t}_m^0 = \mathfrak{t}^*.$$

So $d\mu_m$ is surjective, i.e. μ is a submersion near m . It follows that $\mu : M \rightarrow \mathfrak{t}^*$ is an open map near m . In particular, $\mu(p)$ is an interior point of $\mu(M)$. So $\mu(M)$ is a non-degenerate convex polytope in $\mathfrak{t}^* \simeq \mathbb{R}^k$, which has at least $k + 1$ vertices. Since each vertex is the image of a T -fixed point, the T action has at least $k + 1$ fixed points. \square

2. PROOF OF THE ATIYAH-GUILLEMIN-STERNBERG CONVEXITY THEOREM

We will follow the Guillemin-Sternberg approach to prove the convexity theorem. The Atiyah approach is sketched in Ana Canas de Silver's book.

¶ **The equivariant Darboux theorems.**

Suppose G is a compact connected Lie group acting smoothly on M . Suppose m_0 is a fixed point of the G -action. We can endow with M a G -invariant Riemannian metric. Then sufficiently small geodesic balls around m_0 is a contractible invariant domain, and using the ordinary Poincaré lemma it is easy to prove the following *invariant Poincaré lemma*: Any invariant closed k -form in a neighborhood of m_0 is the differential of an invariant $k - 1$ -form. As a consequence, we can prove

Theorem 2.1 (Equivariant Darboux theorem). *Let (M, ω_i) , $i = 1, 2$ be symplectic G -spaces. Let m be a fixed point of G so that $\omega_1(m) = \omega_2(m)$. Then there is an invariant neighborhood U of m and an equivariant diffeomorphism f of U into M so that $f(m) = m$ and $f^*\omega_2 = \omega_1$.*

Proof. Apply Moser's trick as before. Details left as an exercise. \square

Now suppose (M, ω) be a symplectic G -manifold, and let m be a G -fixed point. Then the isotropy action of G on $T_m M$,

$$g \cdot v = dg_m(v),$$

is a linear G -action, i.e. a representation, of G on $T_m M$. Moreover, if we fix an G -invariant Riemannian metric on M , then the exponential map $\exp : T_m M \rightarrow M$ is equivariant. It follows from the previous theorem that

Theorem 2.2 (Equivariant Darboux theorem, version 2). *Let (M, ω) be a symplectic G -manifold and m a fixed point of the G -action. Then there is an invariant neighborhood U of m in M and an equivariant diffeomorphism φ from (U, ω) into $(T_m M, \omega_m)$ so that $\varphi(m) = 0$ and $\varphi^* \omega_m = \omega$.*

One can identify the symplectic vector space $(T_m M, \omega_m)$ with the complex space \mathbb{C}^n . So

Theorem 2.3 (Equivariant Darboux theorem, version 3). *Let (M, ω) be a symplectic G -manifold and m a fixed point of the G -action. Then there is an invariant neighborhood U of m in M and local complex coordinates z_1, \dots, z_n so that on U , the symplectic form can be written as*

$$\omega = \frac{1}{2i} \sum_k dz_k \wedge d\bar{z}_k,$$

and the G -action becomes a linear symplectic G -action on \mathbb{C}^n .

¶ Darboux theorem for the moment map.

Now let's turn back to the case $G = T$ is a torus of dimension n . Let $\alpha_1, \dots, \alpha_n \in \mathfrak{t}^*$ be the weights of the isotropy representation of T on $T_m M$. In other words,

$$e^{it} \cdot (z_1, \dots, z_n) = (e^{i\alpha_1(t)} z_1, \dots, e^{i\alpha_n(t)} z_n).$$

Theorem 2.4. *Let $p \in U$ be a point whose coordinate is z . Then*

$$\mu(p) = \mu(m) + \sum_k \frac{|z_k|^2}{2} \alpha_k.$$

Proof. Exercise. □

Remark. Locally a symplectic action is always Hamiltonian since the first cohomology vanishes.

¶ The local convexity.

Now let's go back to the theorem.

Let U be the invariant neighborhood given by the equivariant Darboux theorem above. Then the image of U under the moment map μ , near the point $\mu(m)$, is

$$\left\{ \mu(m) + \sum_{k=1}^n s_k \alpha_k \mid 0 \leq s_k \leq \varepsilon \right\}.$$

In other words, we proved

Proposition 2.5. *Let (M, ω, T, μ) be a Hamiltonian T -space and m a T -fixed point. Then there exists a neighborhood U of m so that $\mu(U)$ is a cone with vertex $\mu(m)$.*

This local convexity theorem has a relative version: Let $T_1 \subset T$ be a subgroup, and $m_1 \in M$ is a fixed point under the induced T_1 action. As we have seen in lecture 8, the moment map μ_1 of this T_1 -action is $\mu_1 = d\iota^T \circ \mu$. Note that the map $d\iota^T : \mathfrak{t}^* \rightarrow \mathfrak{t}_1^*$ is nothing else but the projection $\pi : \mathfrak{t}^* \rightarrow \mathfrak{t}_1^*$ if we identify \mathfrak{t}_1^* as a subspace of \mathfrak{t} . Applying the previous arguments, we can find a neighborhood U_1 of m_1 so that $\mu_1(U_1)$ is the cone with vertex $\mu_1(m_1)$ described as above. It follows that its preimage is

Proposition 2.6. *Let (M, ω, T, μ) be a Hamiltonian T -space, T_1 a subgroup of T , and m_1 a T_1 -fixed point. Let $\alpha_1^1, \dots, \alpha_{n_1}^1$ be the weights of the isotropy representation of T_1 on $T_{m_1}M$. Then there exists a neighborhood U_1 of m_1 so that $\mu(U_1)$, near $\mu(m_1)$, is*

$$\left\{ \mu(m) + \pi^{-1} \left(\sum_{k=1}^{n_1} s_k \alpha_k \right) \mid 0 \leq s_k \leq \varepsilon \right\}.$$

¶ The global convexity.

In the next section we shall prove

Lemma 2.7 (Guillemin-Sternberg lemma). *For any $X \in \mathfrak{g}$, the function $\mu^X : M \rightarrow \mathbb{R}$ has a unique local minimum/maximum.*

Proof of the Atiyah-Guillemin-Sternberg convexity theorem. Since M is compact, the image $\mu(M)$ is compact in \mathfrak{g}^* . Let $\xi \in \mathfrak{g}^*$ be a point in the boundary of $\mu(M)$. Take a point $m \in M$ so that $\mu(m) = \xi$. Let $T_1 = T_m$ be the stabilizer of m and let $\alpha_1, \dots, \alpha_n \in \mathfrak{t}_1$ be the weights of the isotropy representation of T_1 on $T_m M$. Then there exists a neighborhood U of m in M so that

$$\mu(U) = \left\{ \xi + \pi^{-1} \left(\sum_{k=1}^{n_1} s_k \alpha_k \right) \mid 0 \leq s_k \leq \varepsilon \right\}.$$

We denote $S = \pi^{-1}(\{\sum s_k \alpha_k \mid s_k \geq 0\})$. Let S_1 be a boundary component of S . Since S_1 has codimension at least 1, one can choose $X \in \mathfrak{g}$ so that

$$\langle \eta, X \rangle = 0 \text{ for } \eta \in S_k$$

and

$$\langle \eta, X \rangle < 0 \text{ for } \eta \text{ in the interior of } S.$$

Now suppose $\langle \xi, X \rangle = a$. Then for any $m' \in U$,

$$\mu^X(m') = \langle \mu(m'), X \rangle = \langle \xi + \eta, X \rangle \leq a,$$

i.e. a is a local maximum of μ^X . According to the Guillemin-Sternberg lemma, it is an absolute maximum. So $\langle \mu(M), X \rangle = \mu^X(M) \leq a$.

Applying this argument to all boundary components S_k of S , we conclude that $\mu(M)$ sits in the cone $\xi + S$. It follows that $\mu(M)$ is a convex polyhedron.

Finally a point $\mu(m)$ is a vertex if and only if $n_1 = n$ for the point m , i.e. the stabilizer of m is T itself. So m is a fixed point of the T -action. \square

3. MORSE THEORY

It remains to prove the Guillemin-Sternberg lemma.

¶ Morse-Bott functions.

Let M be a compact manifold and $f \in C^\infty(M)$ a smooth function. Recall that a point $m \in M$ is called a *critical point* of f if $df_m = 0$. The set of all critical points of f is denoted by $\text{Crit}(f)$.

Now let $m \in \text{Crit}(f)$ be a critical point of f . Consider the *Hessian* map $\text{Hess}_m(f) : T_m M \times T_m M \rightarrow \mathbb{R}$ defined by

$$(1) \quad \text{Hess}_m(f)(X_m, Y_m) = X(Yf)(m),$$

where X and Y are any vector fields whose value at m are X_m and Y_m respectively.

Lemma 3.1. *For $m \in \text{Crit}(f)$, $\text{Hess}_m(f)$ is well-defined, symmetric and bilinear.*

Proof. Symmetry follows from the fact

$$X(Yf)(m) - Y(Xf)(m) = [X, Y]f(m) = \langle df_m, [X, Y]_m \rangle = 0,$$

Since $X(Yf)(m) = X_m(Yf)$, the right hand side of (1) is independent on the choice of X and is bilinear on X_m . By symmetry, $X(Yf)(m) = Y(Xf)(m)$, the right hand side of (1) is also independent of the choice of Y and is bilinear on Y_m . \square

Recall that a critical point $m \in \text{Crit}(f)$ is called *non-degenerate* if the bilinear form $\text{Hess}_m(f)$ is non-degenerate. A function $f \in C^\infty(M)$ is called a *Morse function* if $\text{Crit}(f)$ is discrete, and each $m \in \text{Crit}(f)$ is non-degenerate. Obviously any Morse function on a compact manifold has only finitely many critical points.

Definition 3.2. A function $f \in C^\infty(M)$ is called a *Morse-Bott function* if $\text{Crit}(f)$ is a manifold (with different components, C_i , which may have different dimensions), and for each $m \in M$ the Hessian $\text{Hess}_m(f)$ is non-degenerate in the direction transverse to $\text{Crit}(f)$ (equivalently, the kernel of $\text{Hess}_m(f)$ is $T_m \text{Crit}(f)$).

¶ The index.

Now suppose f is a Morse-Bott function. Let $n^\pm(m)$ be the number of positive/negative eigenvalues of $\text{Hess}_m(f)$ at $m \in \text{Crit}(f)$. They are of course constant along each connected component C_i of $\text{Crit}(f)$. We shall denote them by n_j^\pm and call them the index/coindex of C_j .

Now we equip M with a Riemannian metric. Note that the gradient vector field ∇f vanishes exactly on $\text{Crit}(f)$. Using the Riemannian metric g one can identify

Hess f with $\nabla^2 f$. Let ϕ_t be the negative gradient flow of f . Then for any $m \in C_j$, one has the decomposition

$$T_m M = T_m C_j \oplus E_m^+ \oplus E_m^-,$$

where E^\pm are spanned by the eigenspace of positive/negative eigenvalues of $\nabla^2 f$ respectively.

Observe that as $t \rightarrow \infty$, every point moves to point in $\text{Crit}(f)$ under the flow ϕ_t . For each j we consider the set of points that flow to C_j as $t \rightarrow \infty$,

$$W^s(C_j) = \{m \in M \mid \phi_t(m) \rightarrow C_j \text{ as } t \rightarrow +\infty\}.$$

Fact: $W^s(C_j)$ is a submanifold of M of dimension $n_j^+ + \dim(C_j)$. Moreover, at each $m \in C_j$,

$$T_m W^s(C_j) = T_m C_j \oplus E_m^+.$$

We will call $W^s(C_j)$ the *stable submanifold* of C_j .

Similarly by studying the negative gradient flow as $t \rightarrow -\infty$, one gets the *unstable submanifold* of C_j ,

$$W^u(C_j) = \{m \in M \mid \phi_t(m) \rightarrow C_j \text{ as } t \rightarrow -\infty\},$$

which is a submanifold of dimension $n_j^- + \dim(C_j)$, whose tangent space at $m \in C_j$ is

$$T_m W^u(C_j) = T_m C_j \oplus E_m^-.$$

We shall use the following topological observation: If N is a submanifold of M of codimension at least 2, then $M \setminus N$ is connected. We have seen

$$n_j^+ = \text{codim}(W^u(C_j)), \quad n_j^- = \text{codim}(W^s(C_j)).$$

Note that n_j^- means each point in C_j is a local minimum.

Proposition 3.3. *Let f be a Morse-Bott function. If none of these n_j^- 's equals to 1, then there exists a unique j such that $n_j^- = 0$. In other words, there is a unique C_j on which f is local minimum, and thus an absolute minimum.*

Proof. By the assumption, if $n_j^- \geq 0$, then $n_j^- \geq 2$. The union of all these stable submanifolds has codimension at least 2. It follows that their complement, M^c , is a nonempty connected subset in M . Since M is the union of all stable submanifolds, there exists at least one critical component C_j with index $n_j^- = 0$.

Note that $W^s(C_j)$ is an open subset of M . Moreover, since M^c is connected, it cannot be a union of more than one disjoint open subset. It follows that such C_j is unique. So f attains its absolute minimum on C_j . \square

Remark. By the same way one can show that if none of the n_j^+ 's equals to 1, then there exists a unique C_k on which f is local maximum, and thus an absolute maximum.

¶ Proof of the Guillemin-Sternberg lemma.

Obviously the Guillemin-Sternberg lemma follows from

Theorem 3.4. *Let (M, ω, T, μ) be a compact connected Hamiltonian T -space. Then for any $X \in \mathfrak{g}$, μ^X is a Morse-Bott function, and all of its indices n_j^\pm 's are even.*

Proof. Let $X \in \mathfrak{g}$ and $m \in \text{Crit}(\mu^X)$. Let T_1 be the closure of the 1-parameter subgroup $\{\exp(tX) \mid t \in \mathbb{R}\}$. Then m is a fixed point for the action of T_1 . Let $\alpha_1, \dots, \alpha_n$ be the weights for the isotropy representation of T_1 on $T_m M$. Then as we have seen, the moment map of this T_1 action is locally of the form

$$\mu_1(p) = \mu_1(m) + \sum \frac{|z_k|^2}{2} \alpha_k.$$

On the other hand, $\mu_1 = \pi \circ \mu$, where π is induced by the inclusion $T_1 \subset T$. So

$$\mu^X(p) = \mu_1^X(p) = \mu_1^X(m) + \sum \frac{|z_k|^2}{2} \alpha_k(X) = \mu_1^X(m) + \sum \frac{\alpha_k(X)}{2} (x_k^2 + y_k^2).$$

It follows that the critical points of μ^X near m are defined locally by the equations

$$z_j = z_{j+1} = \dots = z_d = 0,$$

where j is chosen so that $\alpha_j(X) = \dots = \alpha_n(X) = 0$ and $\alpha_1(X) \neq 0, \dots, \alpha_{j-1}(X) \neq 0$, and d is half the dimension of M . In other words, the critical set $\text{Crit}(\mu^X)$ is a smooth submanifolds near each point $m \in \text{Crit}(\mu^X)$. So μ^X is a Morse-Bott function.

Moreover, the Hessian of μ^X at m has eigenvalues

$$\alpha_1(X), \alpha_1(X), \dots, \alpha_n(X), \alpha_n(X), 0, \dots, 0.$$

In particular, all indices are even. □

¶ Connectedness of level sets.

One can say more from the Morse-Bott theory.

Lemma 3.5. *Let f be a Morse-Bott function such that none of the n_j^\pm 's equals 1. Then for any regular value r of f , $f^{-1}(r)$ is a connected submanifold of M .*

Proof. According to the previous proposition, we see that there exists a unique C_j on which f attains its absolute minimum, and there exists a unique C_k one which f attains its local maximum. Let $M_0 = (M \setminus \text{Crit}(f)) \cup C_j \cup C_k$ be the complement of the union of all other critical levels. It is an open dense connected subset of M .

Consider the negative gradient flow ϕ_t on M_0 . Since r is a regular value, $f^{-1}(r)$ is a smooth submanifold of M , and $\min f < r < \max f$. So each orbit intersects with $f^{-1}(r)$ at a unique point. So the map

$$f^{-1}(r) \times \mathbb{R} \rightarrow M_0, \quad (m, t) \mapsto \phi_t(m)$$

is a diffeomorphism. Since M_0 is connected, $f^{-1}(r)$ must be connected. □

Proposition 3.6. *Let f be a Morse-Bott function such that none of the n_j^\pm 's equals 1. Then for any r , $f^{-1}(r)$ is connected.*

Proof. We may assume r is neither the minimum nor the maximum, because the connectedness is already proved for that two cases. Let M_1 be set of points which tends to the minimum of f as $t \rightarrow \infty$, and tends to the maximum as $t \rightarrow -\infty$. Then $f^{-1}(c) \cap M_1$ is a connected submanifold. It remains to prove that $f^{-1}(c) \cap M_1$ is dense in $f^{-1}(c)$.

Let $m \in f^{-1}(c)$, and U be an arbitray small neighborhood of m . Then $U \cap M_1$ is connected because we only delete a submanifold of codimension at least 2. Since c is neighter a local minimum nor a local maximum, one can find points m_1, m_2 in $U \cap M_1$ such that $f(m_1) < c$ and $f(m_2) > c$. So one can find a point m_3 in $U \cap M_1$ with $f(m_3) = c$. This completes the proof. \square

Corollary 3.7. *Under the assumption of AGS, $\mu^X(r)$ is connected for any X and any r .*

4. STUDENT PRESENTATION